GENOMIC STRUCTURE AND TRANSCRIPTIONAL REGULATION OF THE HUMAN UROKINASE RECEPTOR GENE

By
JINJUN DANG

A thesis submitted for the degree of Doctor of Philosophy
at the Australian National University

September 1996
STATEMENT

All the experimental work in this thesis was performed by the author, unless specifically stated otherwise in the text.

Jinjun Dang
ACKNOWLEDGMENTS

I would like to thank my supervisors, Dr. Yao Wang and Professor William F. Doe, for their supervision and encouragement for this research project. In particular, I would like to express my sincere gratitude to my chair supervisor, Dr. Yao Wang who provided invaluable direction and the support for this thesis.

I thank Dr. Donna Cohen and Professor Phil Board for their very helpful discussions and advice on the topics related to this thesis.

I am grateful to June Hornby, Caroline Jones, Jane Olsen and Sue Elsbury for their excellent technical assistance with various aspects of this work. I am also grateful to Dr. Peter Milburn and his colleagues for their assistance in DNA sequencing and oligonucleotide synthesis. My thanks also go to the staff at JCSMR photography section for their assistance in producing the figures used in this thesis.

I wish to express my warm thanks to June Hornby for all her excellent work in proofreading of this thesis.

The support of an Australian National University Postgraduate Scholarship is gratefully acknowledged.

Finally, I would like to thank my wife and my daughter for their encouragement and support during my PhD study.
ABSTRACT

Urokinase-type plasminogen activator (uPA) and its receptor (uPAR) have gained increasing medical importance in recent years as they may play important roles in a variety of physiological as well as pathological processes including cancer cell invasion and metastasis. This study describes the characterization of the genomic structure and the mechanisms of transcription regulation of the human uPAR gene.

To characterize the genomic structure of the human uPAR gene and its regulatory region, the entire uPAR gene including its 5′- and 3′-flanking regions was isolated and sequenced. The results showed that the human uPAR gene consists of 7 exons separated by 6 introns and spans over 22 kb. The seven exons of 101, 111, 144, 162, 135, 147 and 563 bp are separated by six introns of approximately 2.04, 2.62, 8.42, 0.906, 3.10 and 2.78 kb, respectively. Exons 1-7 encode 19, 37, 48, 54, 45 and 83 amino acid residues, respectively. Three cystine-rich domains of the uPAR protein are encoded by paired exons (2+3), (4+5) and (6+7), respectively. Sequence analysis of a 2.6 kb 5′-flanking region revealed a CpG-rich island and a number of DNA binding sites for the transcription factors Sp1, AP-1, AP-2, NF-κB, PEA3 and GATA. No potential TATA or CAAT boxes are located in the proximal region of the uPAR gene. Four Alu repetitive sequences were found in both orientations in the 5′-flanking region. Primer extension analysis demonstrated that the uPAR gene has multiple transcription initiation sites around 51 bp upstream of the first base of the translation start codon of the gene, which is designated +1.

To characterize cis-acting regulatory elements in the 5′-flanking region of the human uPAR gene, a series of deletions of 5′-flanking sequence was fused to the bacterial chloramphenicol acetyltransferase (CAT) gene and transiently transfected into human colon cancer cell line, HCT116, which constitutively expresses uPAR mRNA. A promoter region was identified between bases -141 and +47 which is necessary for the maximal promoter activity. This promoter region contains putative binding sites for the transcription factors Sp1, AP-1 and NF-κB. Either 5′ deletion to -60 or 3′ deletion to -10 in this region abolished the promoter activity. Co-transfection of uPAR promoter construct with c-fos and c-jun expression plasmids showed that AP-1 was able to
stimulate transcription from the uPAR promoter up to 4 to 5-fold. A distal silencer-like element was also found at the region between -1796 and -1458, which contains a full-length Alu sequence.

DNase I footprinting analysis revealed two protected regions, region I from -51 to -30 and region II from -118 to -79. Region I contains a NF-κB binding site and the region II includes two putative Sp1 binding sites. The specific binding of nuclear proteins from the HCT116 nuclear extracts to these NF-κB and Sp1 sites was confirmed using corresponding consensus sequences by electrophoretic mobility shift assay. Binding of NF-κB protein (p49) to the uPAR-NF-κB sequence was abolished by NF-κB consensus sequence, but not by mutant uPAR-NF-κB sequence. Binding of the nuclear protein to AP-1 site at -70 was examined using a uPAR-AP-1 oligonucleotide. These results indicate that multiple cis-acting elements may be involved in the transcription regulation of the uPAR gene.

The regulation of uPAR gene expression by sodium butyrate, a potential inhibitor of tumorigenesis, was also investigated in human colon cancer cells unstimulated or stimulated with tumor necrosis factor alpha (TNFα) or phorbol ester (PMA). Northern blot analysis demonstrated that while both uPA and uPAR mRNA levels were increased after stimulation by TNFα, PMA or cycloheximide, they were inhibited by butyrate at 2.5 to 25 mM. Nuclear run-on transcription assays indicated that uPA gene expression was modulated by butyrate at the transcriptional level. The uPAR gene expression, however, was regulated at both transcriptional and post-transcriptional levels in the presence or absence of TNFα. In the presence of PMA sodium butyrate may act at the post-transcriptional level on both genes.

The results presented in this thesis provide the structure of the uPAR gene and revealed a promoter, a silencer as well as several transcription factors involved in the transcriptional regulation of the uPAR gene. These data provide interesting avenues for future detailed molecular studies on uPAR gene regulation.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CAT</td>
<td>chloramphenicol acetyltransferase</td>
</tr>
<tr>
<td>DNase</td>
<td>deoxyribonuclease</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>ECM</td>
<td>extracellular matrix</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FCS</td>
<td>fetal calf serum</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>ONPG</td>
<td>o-nitrophenylβ-D-galactopyranoside</td>
</tr>
<tr>
<td>NaB</td>
<td>sodium butyrate</td>
</tr>
<tr>
<td>PAI-1</td>
<td>type-1 plasminogen activator inhibitor</td>
</tr>
<tr>
<td>PAI-2</td>
<td>type-2 plasminogen activator inhibitor</td>
</tr>
<tr>
<td>PAIs</td>
<td>plasminogen activator inhibitors</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PMA</td>
<td>phorbol 12-myristate 13-acetate</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SSC</td>
<td>sodium chloride/sodium citrate solution</td>
</tr>
<tr>
<td>SV40</td>
<td>Simian Virus 40</td>
</tr>
<tr>
<td>TNFα</td>
<td>tumor necrosis factor-alpha</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris (hydroxymethyl)-aminomethane</td>
</tr>
<tr>
<td>tPA</td>
<td>tissue-type plasminogen activator</td>
</tr>
<tr>
<td>uPA</td>
<td>urokinase-type plasminogen activator</td>
</tr>
<tr>
<td>uPAR</td>
<td>urokinase-type plasminogen activator receptor</td>
</tr>
<tr>
<td>VLL-AMC</td>
<td>H-D-Val-Leu-Lys 4-aminomethyl coumarin</td>
</tr>
</tbody>
</table>
The work presented in Chapter 2, and 5 has been published in:


The work presented in Chapter 3 has been published in abstract form in:

CONTENTS

STATEMENT ............................................................................................................. II
ACKNOWLEDGMENTS ............................................................................................. III
ABSTRACT .............................................................................................................. IV
ABBREVIATION ..................................................................................................... VI
PUBLICATIONS ....................................................................................................... VII

CHAPTER 1 INTRODUCTION
1.1 Structure of the Urokinase Receptor .................................................................. 2
1.1.1 Molecular structure ......................................................................................... 3
1.1.1.1 uPAR protein ............................................................................................... 3
1.1.1.2 Three domain structure ................................................................................. 3
1.1.1.3 Glycosyl-phosphatidylinositol (GPI)-anchored protein ......................... 4
1.1.1.4 Conservation of the uPAR gene ..................................................................... 4
1.1.1.5 Chromosome localization .............................................................................. 4
1.1.2 uPAR and Ly-6 family and snake venom α-neurotoxins .................................. 4
1.1.2.1 uPAR and family .......................................................................................... 4
1.1.2.2 uPAR and snake venom α-neurotoxins ....................................................... 5
1.2 uPA/uPAR System and Plasminogen Activation .................................................. 5
1.2.1 uPA/uPAR system ........................................................................................... 5
1.2.1.1 Urokinase .................................................................................................... 6
1.2.1.2 Plasminogen and plasmin ............................................................................ 6
1.2.1.3 Plasminogen activator inhibitors ................................................................. 7
1.2.1.4 α2-MR and gp330 ....................................................................................... 7
1.2.2 uPA/uPAR-mediated plasminogen activation ................................................... 7
1.2.2.1 Plasminogen activation ................................................................................. 7
1.2.2.2 Plasminogen activation occurs on cell surface .......................................... 8
1.2.2.3 Binding of uPA to its receptor accelerates plasminogen activation ........ 8
1.2.3 uPAR mediates internalization and degradation of uPA-serpin complexes ...... 9
1.2.3.1 The uPA/uPAR cycle .................................................................................. 9
1.2.3.2 uPAR-mediated internalization of uPA-PAIs complexes ............................ 9
1.3 uPAR and Tumor Metastasis .............................................................................. 10
1.3.1 Proteolysis is a key step of tumor metastasis .................................................. 10
1.3.1.1 Tumor metastasis ....................................................................................... 10
1.3.1.2 Requirement of proteolysis in tumor invasion .......................................... 10
1.3.1.3 Involvement of uPA-mediated proteolysis in tumor metastasis .............. 11
1.3.2 Correlation of uPAR with tumor invasion and metastasis ............................. 11
1.3.2.1 Elevated levels of uPAR is associated with human carcinomas ............. 11
1.3.2.2 uPAR focuses uPA activity at tumor invasive edge ................................... 12
1.3.2.3 uPAR and colon cancer .............................................................................. 12
1.3.2.4 uPAR is a potential prognostic factor ....................................................... 13
1.3.3 Blockade of uPAR in tumor invasion and metastasis .................................... 14
CHAPTER 2 GENOMIC ORGANIZATION OF THE HUMAN UROKINASE RECEPTOR GENE

2.1 Introduction ............................................................................................................. 32
2.2 Materials and Methods .......................................................................................... 34
2.2.1 Materials ............................................................................................................. 34
2.2.2 Southern blot hybridization of genomic DNAs ................................................. 34
2.2.2.1 Cell culture .................................................................................................... 34
2.2.2.2 Isolation of high molecular weight DNA from cultured cells ....................... 34
2.2.2.3 Restriction digestion of DNA and agarose gel electrophoresis ....................... 35
2.2.2.4 Blotting of DNA to reinforced nitrocellulose ............................................... 35
2.2.2.5 Probe labeling ............................................................................................... 36
2.2.2.6 Southern hybridization ................................................................................. 36
2.2.3 Isolation of genomic clones .............................................................................. 37
2.2.3.1 Plating bacteriophage λ .............................................................................. 37
2.2.3.2 Plaque transfer and hybridization ................................................................. 37
2.2.3.3 Large-scale preparation of phage DNA ......................................................... 38
2.2.4 Subcloning of the uPAR fragments .................................................................. 38
2.2.4.1 Isolation of uPAR gene fragments ................................................................. 38
2.2.4.2 Dephosphorylation of enzyme-digested vectors .......................................... 39
2.2.4.3 DNA ligation ............................................................................................... 39
2.2.5 Bacterial transformations ................................................................. 39
    2.2.5.1 Preparation of competent cells .................................................... 39
    2.2.5.2 Transformation ................................................................. 40
    2.2.5.3 Selection of plasmids of interest .................................................. 40
2.2.6 Preparation of plasmid DNA .......................................................... 40
    2.2.6.1 Minipreparation of plasmid DNA ................................................. 40
    2.2.6.2 Large scale preparation of plasmid DNA ....................................... 41
2.2.7 Nested deletion of plasmid DNA ...................................................... 41
2.2.8 DNA sequencing .............................................................................. 42
    2.2.8.1 T3 and T7 dye primer method ....................................................... 42
    2.2.8.2 Dye terminator method .............................................................. 42
2.2.9 Characterization of exon-intron boundaries and sequence assembly, analysis
     and alignment ..................................................................................... 43
2.2.10 Polymerase chain reactions (PCR) .................................................. 43
2.2.11 RNA preparation ............................................................................ 45
2.2.12 Primer extension analysis ............................................................... 45
2.3 Results ................................................................................................. 46
    2.3.1 Isolation and characterization of the human uPAR gene ...................... 46
        2.3.1.1 Southern blot analysis .............................................................. 46
        2.3.1.2 Isolation of uPAR genomic clones ............................................. 46
        2.3.1.3 Subcloning and sequencing strategy ......................................... 49
    2.3.2 Structure of the human uPAR gene ................................................ 49
        2.3.2.1 Exon/intron organization ......................................................... 49
        2.3.2.2 Domain structure and intron/exon boundaries ............................ 52
    2.3.3 Characterization of the 5'-flanking region ....................................... 52
    2.3.4 Identification of the transcription initiation sites .............................. 53
    2.3.5 Sequence analysis of the uPAR 5'-flanking region ........................... 56
        2.3.5.1 Consensus sequences for DNA binding proteins ....................... 56
        2.3.5.2 G+C-rich sequence and a dyad symmetry structure .................. 56
    2.3.6 Comparison of human and murine uPAR genomic structure .............. 57
2.4 Discussion .......................................................................................... 61

CHAPTER 3 CHARACTERIZATION OF PROMOTER AND SILENCER
IN THE HUMAN uPAR GENE
3.1 Introduction ...................................................................................... 66
3.2 Materials and Methods ....................................................................... 68
    3.2.1 Materials .................................................................................... 68
    3.2.2 Plasmid construction .................................................................... 68
        3.2.2.1 Plasmids .............................................................................. 68
        3.2.2.2 Deletion of 5'-flanking region ............................................... 69
        3.2.2.3 5' and 3' deletion of the uPAR promoter region .................... 69
        3.2.2.4 Internal deletion of 5'-flanking region .................................... 70
    3.2.3 Cell culture and transient transfections .......................................... 73
    3.2.4 Assays for β-galactosidase ........................................................... 74
    3.2.5 CAT assay ................................................................................... 74
        3.2.5.1 Thin layer chromatography (TLC) method ................................. 74
        3.2.5.2 Liquid scintillation counting (LSC) method .............................. 75
CHAPTER 4 MULTIPLE NUCLEAR FACTORS INTERACT WITH THE HUMAN uPAR PROMOTER SEQUENCES

4.1 Introduction ........................................................................................................... 88
4.2 Materials and Methods ....................................................................................... 90
4.2.1 Materials ........................................................................................................ 90
4.2.2 Preparation of nuclear extracts ....................................................................... 90
4.2.3 DNase I footprinting ....................................................................................... 91
   4.2.3.1 DNA probe .............................................................................................. 91
   4.2.3.2 Binding reaction and DNase I digestion .................................................. 91
   4.2.3.3 Preparation of G+A marker ..................................................................... 92
   4.2.3.4 Electrophoresis of digested DNA samples .............................................. 92
4.2.4 Electrophoretic mobility shift assay (EMSA) .................................................. 92
   4.2.4.1 DNA fragments and synthetic oligonucleotides ..................................... 92
   4.2.4.2 Binding reaction ...................................................................................... 93
   4.2.4.3 Electrophoresis of DNA-protein complexes ......................................... 93
4.3 Results .................................................................................................................. 95
   4.3.1 Identification of two protected regions in the uPAR promoter region .......... 95
   4.3.2 Binding of transcription factor NF-κB to the protected region I ................ 95
   4.3.3 Nuclear proteins interact with protected region II ....................................... 98
   4.3.4 Interactions of proteins with the 55 bp regulatory region (-8 to +47) .......... 100
   4.3.5 Binding of nuclear protein to the uPAR AP-1 element in the uPAR promoter region ................................................................. 104
4.4 Discussion ........................................................................................................... 104

CHAPTER 5 SODIUM BUTYRATE INHIBITS UROKINASE RECEPTOR EXPRESSION AT BOTH TRANSCRIPTION AND POST-TRANSCRIPTION LEVELS

5.1 Introduction .......................................................................................................... 110
5.2 Materials and Methods ....................................................................................... 112
   5.2.1 Materials .................................................................................................... 112
   5.2.2 Cell culture and stimulation ...................................................................... 112
   5.2.3 cDNA probes ............................................................................................ 113
   5.2.4 RNA preparation and Northern blot analysis ............................................. 113
      5.2.4.1 Preparation of formaldehyde-agarose denaturing gels ......................... 113
      5.2.4.2 Preparation of RNA samples for electrophoresis .................................. 113
      5.2.4.3 Electrophoresis of RNA ..................................................................... 113
      5.2.4.4 Blotting of RNA to reinforced nitrocellulose ....................................... 114
      5.2.4.5 RNA hybridization ............................................................................. 114
   5.2.5 uPAR binding assays .................................................................................... 114
   5.2.6 Nuclear run-on transcription analysis ........................................................ 115
<table>
<thead>
<tr>
<th>Chapter 5.2.6.1</th>
<th>Plasmid DNA blotting</th>
<th>115</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 5.2.6.2</td>
<td>Nuclei preparation</td>
<td>115</td>
</tr>
<tr>
<td>Chapter 5.2.6.3</td>
<td>Transcription elongation reaction</td>
<td>115</td>
</tr>
<tr>
<td>Chapter 5.2.6.4</td>
<td>Hybridization and autoradiography</td>
<td>116</td>
</tr>
<tr>
<td>5.3 Results</td>
<td></td>
<td>116</td>
</tr>
<tr>
<td>Chapter 5.3.1</td>
<td>Sodium butyrate (NaB) inhibits uPAR and uPA mRNAs expression</td>
<td>116</td>
</tr>
<tr>
<td>Chapter 5.3.2</td>
<td>Sodium butyrate inhibits TNFα or PMA-stimulated uPAR and uPA mRNAs expression</td>
<td>117</td>
</tr>
<tr>
<td>Chapter 5.3.3</td>
<td>Sodium butyrate inhibits uPAR protein levels on the cell surface</td>
<td>120</td>
</tr>
<tr>
<td>Chapter 5.3.4</td>
<td>Sodium butyrate inhibits cycloheximide-induced uPAR and uPA mRNAs expression</td>
<td>120</td>
</tr>
<tr>
<td>Chapter 5.3.5</td>
<td>Effects of sodium butyrate on uPAR and uPA transcription in isolated nuclei</td>
<td>120</td>
</tr>
<tr>
<td>5.4 Discussion</td>
<td></td>
<td>125</td>
</tr>
</tbody>
</table>

**CHAPTER 6 GENERAL DISCUSSION & FUTURE DIRECTION**

<table>
<thead>
<tr>
<th>6.1 General Discussion</th>
<th>128</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1.1 The aims and achievement of this thesis</td>
<td>128</td>
</tr>
<tr>
<td>6.1.2 Transcription regulation of uPAR gene is complex</td>
<td>129</td>
</tr>
<tr>
<td>6.1.3 Regulation of uPAR gene expression by sodium butyrate</td>
<td>134</td>
</tr>
<tr>
<td>6.2 Future Studies</td>
<td>136</td>
</tr>
<tr>
<td>6.2.1 To study signal transduction pathways in the regulation of uPAR expression</td>
<td>136</td>
</tr>
<tr>
<td>6.2.2 To study transcription factors involved in tumor metastasis</td>
<td>137</td>
</tr>
<tr>
<td>6.2.3 To test the functionality of uPAR regulatory elements in transgenic mice</td>
<td>137</td>
</tr>
</tbody>
</table>

**BIBLIOGRAPHY**
CHAPTER 1
INTRODUCTION

The transcriptional regulation of eukaryotic gene expression is central to cell differentiation and development. Gene transcription starts from upstream of the coding sequences and is controlled by the proximally located promoter, and further regulated by a distally located enhancer or silencer (Mitchell and Tjian, 1989; Tjian and Maniatis, 1994). In most cases, eukaryotic genes are regulated positively or negatively at the transcriptional level according to specific cell type or in response to a specific signal. Once this has occurred, all the other stages of gene expression such as RNA processing and translation follow and the appropriate protein is produced. Both the basal process of transcription itself and its regulation are involved in the interaction between specific short DNA sequences in the gene promoter/enhancer/silencer and transcription factors (Mitchell and Tjian, 1989; Herschbach and Johnson, 1993; Goodrich et al., 1996).

Cell migration is essential for both normal cells and malignant cells to carry out physiological and pathological functions including fertilization, embryogenesis, angiogenesis, and cancer metastasis (Lauffenburger and Horwitz, 1996). To cross tissue barriers, migrating cells have to selectively express several kinds of proteolytic enzymes, adhesion molecules and growth factors to degrade the basal lamina and extracellular matrix (ECM) and promote its movement (Liotta et al., 1991; Mignatti and Rifkin, 1993). One of the proteases, urokinase-type plasminogen activator (uPA), plays a central role in ECM degradation by catalyzing the formation of the broad spectrum protease plasmin from plasminogen directly on the cell surface (Danø et al., 1985; Blasi, 1988). uPA is expressed by both normal and tumor cells and interacts with its specific cell surface receptor (uPAR) thereby leading to localized and regulated proteolysis (for review see Blasi, 1993; Behrendt et al., 1995). uPA activity is inhibited by the plasminogen activator inhibitor type-1 (PAI-1), type-2 (PAI-2) and protease nexin-1 (PN-1) which trigger internalization of the uPA-uPAR complex (Conese and Blasi, 1995). The aberrant expression of these genes may lead to inappropriate tissue invasion and tissue injury (Fazioli and Blasi, 1994). The expression of uPA, uPAR and PAIs, therefore, is highly regulated. The uPA system provides an important model for the study of regulation of gene expression involved in cell migration and tumor metastasis.
The uPAR is increasingly considered to be an important factor in focusing uPA-mediated pericellular proteolysis. Binding of uPA to its receptor focuses its activity on the cell surface and promotes plasminogen activation (Blasi, 1993). Recent studies have shown that uPAR, a glycosylphosphatidylinositol (GPI)-anchored membrane protein, may have additional roles in signal transduction, adhesion and chemotaxis (Behrendt et al., 1995). Studying the regulation of this cellular receptor, therefore, may ultimately be valuable in understanding the mechanisms of cell migration and tumor metastasis. Both uPA and uPAR are expressed in many different cell types and regulated by various factors such as tumor promoters, cytokines and growth factors (Blasi, 1993; Conese and Blasi, 1995). While transcriptional regulation of uPA has been widely studied (Hansen et al., 1992; Nerlov et al., 1992; Novak et al., 1994; Lengyel et al., 1995), the mechanism of controlling uPAR transcription is poorly understood. Our laboratory has been interested in the structure of the uPAR gene and the factors that regulate its expression especially during the process of tumor invasion and metastasis. Previous studies in this laboratory have assigned the human uPAR gene to chromosome 19 (Webb et al., 1994) and demonstrated that tumor metastasis induced by human colon cancer cells is inhibited by blocking uPAR expression using antisense uPAR mRNA (Liang, PhD thesis, 1994). uPAR expression is up-regulated by phorbol ester and TNFα, and down-regulated by amiloride (Wang et al., 1994). Little was known, however, about the transcriptional regulation of the uPAR gene before the commencement of this project and data concerning the functional or mechanistic aspects of uPAR transcriptional regulation were lacking. Recently some work has been done in this rapidly advancing area of research field. Relevant literature published subsequently is addressed in the discussion sections of each chapter and in the final chapter. This chapter seeks to provide an overview of recent developments in my understanding of the uPAR gene including its structure, function and expression.

1.1 Structure of the Urokinase Receptor

The specific binding of urokinase to cell surfaces was first observed in human blood monocytes and the histiocytic lymphoma cell line U937 (Vassalli et al., 1985; Stoppelli et al., 1985). Subsequently, the membrane protein responsible for this binding was purified and characterized from PMA-stimulated U937 cells and designated as urokinase-type plasminogen activator receptor (Nielsen et al., 1988). A 1.4 kb of full
length human uPAR cDNA was isolated and sequenced in 1990 (Roldan et al., 1990). It is the first enzyme receptor on the cell surface to be identified and cloned.

1.1.1 Molecular structure

1.1.1.1 uPAR protein

The human uPAR is a single chain polypeptide with Mr of 55,000-60,000 which is highly glycosylated and decreases to Mr of 35,000 upon deglycosylation (Nielsen et al., 1988; Behrendt et al., 1990). The cDNA for uPAR encodes a 335-residue polypeptide of which the first 22 amino acids constitute the signal peptide (Roldan et al., 1990). After removal of its signal sequence, the protein is further truncated during the post-translational processing by removing a COOH-terminal signal peptide to allow the addition of a glycolipid membrane anchor, generating a mature uPAR protein composed of residues 1-283 (Ploug et al., 1991). Thus, the protein has no transmembrane or intracellular domains.

1.1.1.2 Three domain structure

An analysis of internal homology in the amino acid sequence revealed that the entire sequence of mature uPAR is composed of three cysteine-rich repeats of approximately 90 amino acids each (Behrendt et al., 1991). The consensus sequence derived from these repeats is primarily characterized by a conserved unique pattern of cysteine residues which constitutes 10% of the protein, although comparison of the individual repeats suggests that they are rather distantly related with 12-22% sequence identity (Behrendt et al., 1991). Each of these repeats corresponds to an autonomous structure domain and it has been proposed, therefore, that uPAR is composed of three homologous domains (Ploug et al., 1991). The first domain (residues 1-87) contains the ligand-binding region which is involved in high-affinity binding to the epidermal growth factor-like module of uPA with high species specificity. The remaining two COOH-terminal domain structures (residues 88-283) have no detectable uPA binding activity but the third domain is responsible for linking the protein to the plasma membrane (Ploug et al., 1991). The domains two and three, however, may be important for high affinity uPA-binding as the isolated first domain presents a 1500-fold lower ligand affinity than three-domain uPAR (Ploug et al., 1994). Recent studies have provided evidence that high ligand-affinity may be achieved by a combination of domain one with domain two or three (Riittinen et al., 1996).
1.1.1.3 Glycosyl-phosphatidylinositol (GPI)-anchored protein

The uPAR protein is linked to the cell membrane by a glycolipid anchor (Ploug et al., 1991). This is added during protein maturation in the endoplasmic reticulum by a combined cleavage/attachment process in which the GPI anchor replaces the COOH-terminal moiety (probably 30 residues). Two critical sequence elements required for successful glycolipid attachment, Ser282-Gly283-Ala284 and residues 306-313, have been identified (Møller et al., 1992). GPI-anchored proteins appear to play an important role in transmembrane signaling (Vidugiriene and Menon, 1994; Casey, 1995).

1.1.1.4 Conservation of the uPAR gene

After a 1.4 kb of full length human uPAR cDNA was cloned and analyzed (Roldan et al., 1990), work during the last few years has led to the isolation and characterization of complete uPAR cDNAs from other species including mouse (Kristensen et al., 1991), bovine (Reuning et al., 1993; Krätzschmar et al., 1993) and rat (Rabbani et al., 1994). The uPAR cDNAs from these different species have similar characteristics including three conserved cysteine-rich repeats. The interspecies conservation of the individual repeats is more than 60% (Ploug and Ellis, 1994).

1.1.1.5 Chromosome localization

The human uPAR gene has been assigned to the long arm of chromosome 19q13.2 (Børglum et al., 1992; Webb et al., 1994). However, the structure of the uPAR genomic DNA from any species had not been reported at the commencement of this project.

1.1.2 uPAR and Ly-6 family and snake venom α-neurotoxins

A protein database search using the consensus sequence of the putative domains revealed that internal repeats of uPAR appear to be related to a Ly-6 family as well as to snake venom α-neurotoxins (Ploug and Ellis, 1994).

1.1.2.1 uPAR and Ly-6 family

The similarities between the human uPAR in its three-domain structure and the Ly-6 family that contains a single cysteine-rich domain have been shown (Behrendt et al., 1991; Ploug and Ellis, 1994). Alignment of the uPAR protein sequence with murine Ly-6A revealed that there is a clear relationship between these two molecules (Palfree, 1991). The pattern of disulfide cross-links in the amino-terminal domain places uPAR within this Ly-6 superfamily of GPI-anchored proteins whose members include the
murine Ly-6 antigens, the human membrane inhibitor of reactive lysis (MIRL)/CD59 and the squid glycoprotein Spg (Behrendt et al., 1991; Ploug et al., 1993; Ploug and Ellis, 1994). Like uPAR, all members of this Ly-6 superfamily are glycolipid-anchored membrane glycoproteins, but uPAR is the only multi-domain member of the Ly-6 family.

1.1.2.2 uPAR and snake venom α-neurotoxins

Recent evidence has also highlighted similarities between the individual domains of uPAR and the large family of secreted, single domain snake venom α-neurotoxins (Ploug and Ellis, 1994). It has been found that these secreted proteins have similar size to the uPAR/Ly-6 domains. The individual uPAR domains may adopt the same overall structural topology as these toxins since the disulfide bond connection of the amino-terminal domain of uPAR has also been found to be homologous to that of the nonglycosylated snake venom α-neurotoxins (Ploug et al., 1993). ¹H-NMR assignments for MIR1/CD59 are consistent with a model in which this protein, homologous to the uPAR domains, has a gross folding topology similar to that of the snake α-neurotoxins (Fletcher et al., 1993). It has been suggested on the basis of their amino acid sequences that uPAR, Ly-6 superfamily and snake venom cytotoxin/neurotoxin family are evolutionally related (Ploug and Ellis, 1994).

1.2 uPA/uPAR System and Plasminogen Activation

The urokinase pathway of plasminogen activation is considered to be a central process in the regulation of pericellular proteolysis which occurs under both normal and pathological conditions, including cancer invasion and metastasis (Danzé et al., 1985; Blasi, 1993). Specific cell surface binding sites for urokinase promote plasminogen activation and focus this pericellular proteolysis on the cell surface (Blasi, 1993; Conese and Blasi, 1995). The receptor for uPA, therefore, has aroused great interest. This section will present some of the general concepts about uPAR-mediated proteolysis.

1.2.1 uPA/uPAR system

The uPA/uPAR system is composed of one enzyme (urokinase), one substrate (plasminogen), two receptors (urokinase receptor and the plasminogen receptor), and three high affinity plasminogen activator inhibitors (PAI-1, PAI-2 and protease nexin 1).
(for review see Vassalli et al., 1991; Blasi, 1993; Conese and Blasi, 1995). In addition, the endocytic receptors α₂-macroglobulin receptor (α₂-MR) and epithelial glycoprotein 330 (gp330) have been reported recently to be involved in internalization and degradation of uPA-serpin complexes. Each of these members will be described in turn as follows.

1.2.1.1 Urokinase

Urokinase is a serine protease which is produced and secreted as a 54-kDa single-chain glycosylated proenzyme (pro-uPA) and then converted by limited proteolysis into an active two-chain molecule which is linked by a disulfide bridge (Wun et al., 1982; Dano et al., 1985). The A chain of active uPA or its amino terminal fragment (ATF) can bind the urokinase receptor and the B chain contains the catalytic site of the enzyme and can bind specific inhibitors, PAI-1 and PAI-2 (Cubellis et al., 1990; Estreicher, et al., 1990). After secretion uPA is bound to its cell surface receptor with high specificity and affinity with a Kd of about 0.2 nM by both autocrine and paracrine mechanisms (Vassalli et al., 1985; Stoppelli et al., 1985; Blasi, 1993). uPAR recognizes both pro-uPA and active uPA and promotes conversion of pro-uPA into the active form (Vassalli et al., 1985; Appella et al., 1987). The active two-chain uPA is thought to catalyze plasmin formation. However, recent studies have suggested that single-chain uPA also efficiently initiates cell surface plasminogen activation (Manchanda and Schwartz, 1991; Schwartz, 1994).

1.2.1.2 Plasminogen and plasmin

Human plasminogen is a single-chain glycoprotein with a molecular weight of 92 kDa, containing 790 amino acids (Castellino and Powell, 1981). Plasminogen is an abundantly distributed plasma molecule which is synthesized as an inactive zymogen, primarily by the liver, and is present in most extracellular fluids (Raum et al., 1980). Inactive plasminogen can be activated by plasminogen activators. The resultant active plasmin consists of two polypeptide chains held together by disulfide bonds. The light chain (B-chain) has a molecular weight of 25 kDa and contains the active site. The full-length cDNA of human plasminogen has been isolated and sequenced (Forsgren et al., 1987). The plasminogen receptor, however, has not been isolated or characterized.

1.2.1.3 Plasminogen activator inhibitors

Three specific inhibitors of urokinase, PAI-1, PAI-2 and protease nexin 1 (PN-1), belong to the serpin superfamily (Ellis et al., 1990; Vassalli et al., 1991; Blasi, 1993).
These PAIs and uPA are covalently linked. Their interaction with uPA results in the inactivation of uPA and cleavage of the serpins. PAI-1 is the predominant form in plasma and is primarily produced by endothelial cells as well as by a number of other cell types. It can inhibit active uPA, but not pro-uPA (Cubellis et al., 1989; Andreasen et al., 1990). PAI-2 is produced mainly by the placenta and by activated monocytes/macrophages as well as some cells of neoplastic region. It is less efficient but more specific in uPA inhibition than PAI-1 (Vassalli et al., 1991). PN-1 not only inhibits uPA, but also inhibits plasmin, thrombin, and trypsin (Conese and Blasi, 1995).

1.2.1.4 \( \alpha_2 \)-MR and gp330

Both \( \alpha_2 \)-MR and gp330 belong to the low density lipoprotein receptor (LDL) family (Kounnas et al., 1992). These two trans-membrane receptors have been shown recently to bind uPA-PAI-1 complexes and mediate their clearance by endocytosis (Herz et al., 1992; Moestrup et al., 1994; Conese and Blasi, 1995).

1.2.2 uPA/uPAR-mediated plasminogen activation

1.2.2.1 Plasminogen activation

Plasminogen activation is a significant inducible extracellular proteolytic system involved in the regulation of cellular interactions (Blasi, 1993). There are two known specific plasminogen activators, tissue-type (tPA) and urokinase-type (uPA). tPA is the main PA in plasma where it plays a crucial role in intravascular fibrinolysis (Collen and Lijnen, 1991) and uPA is primarily involved in cell mediated proteolysis during cell migration, inflammation, fibrinolysis, tissue remodeling, tissue involution and trophoblastic invasion, wound healing, embryogenesis, invasion and metastasis (Danø et al., 1985; Conese and Blasi, 1995).

In urokinase-mediated plasminogen activation, plasminogen may bind to its specific cell surface receptors (Plow et al., 1986) and become activated by receptor bound uPA on the cell surface (Blasi, 1993). uPA converts the proenzyme plasminogen into the active, broad specificity, trypsin-like protease plasmin that is a key component in extracellular proteolysis, being capable of hydrolyzing proteins and peptides at either lysyl and arginyl bonds (Danø et al., 1985). In addition to cleaving the fibrin polymers of clots formed in blood, active plasmin either directly or indirectly degrades all components of the extracellular matrix including fibronectin, laminin, proteoglycan core proteins and vitronectin. It can also activate prometalloproteinases, thereby inducing the
action of other proteinases (Pellman et al., 1991; Mignatti and Rifkin, 1993). Therefore, the generation of plasmin can produce large-scale extracellular proteolytic activity.

1.2.2.2 **Plasminogen activation occurs on the cell surface**

The activation of plasminogen mediated by uPA mostly occurs at the cell surface since it requires the binding of both uPA and plasminogen to their cellular receptors (Blasi, 1993). Both uPA and plasminogen are able to bind to many cell types. Activation of plasminogen by receptor-bound uPA can be abolished by the plasminogen-binding antagonist tranexamic acid and by uPA inhibitors (Ellis et al., 1991; Blasi, 1993), showing that this concomitant binding is of critical importance. The uPA-mediated plasminogen activation is directed by uPAR. It has been found that production of active uPA occurs almost exclusively at the cell surface in cultured cells (Quax et al. 1991).

1.2.2.3 **Binding of uPA to its receptor accelerates plasminogen activation**

Binding of uPA to its cellular receptor can accelerate plasminogen activation by promoting conversion of pro-uPA to active uPA (Levi et al., 1993). The rate of activation of receptor-bound pro-uPA by plasmin is at least 20-fold higher than in solution (Quax et al. 1991). The receptor-bound uPA then accelerates plasminogen activation and allows uPA to function as a membrane-associated enzyme limiting proteolysis to the immediate pericellular environment (Kirchheimer et al., 1989a). In U937 cells, binding of uPA to its receptor causes a 40-fold higher efficiency of plasminogen activation with kinetic characteristics than those observed for uPA in solution phase (Ellis et al., 1989). Thus, the properties of uPA-uPAR complexes define a system in which plasminogen activation proceeds efficiently and is highly focused at the plasma membrane.

On the other hand, inhibition of uPA by PAI-1 and PAI-2 represents one of the mechanisms for regulating the cell surface plasminogen activation system (Blasi, 1993). In the uPA system, therefore, uPAR is a positive modulator for uPA-mediated plasminogen activation, whereas uPA inhibitors are negative modulators.

1.2.3 **uPAR mediates internalization and degradation of uPA-serpin complexes**

Studies of the interaction of uPA with uPAR have demonstrated that uPAR mediates internalization and degradation of the uPA-PAI-1 complex (Nykjær et al., 1992; Olson et al., 1992; Herz et al., 1992). The current view is that uPAR acts by
presenting the uPA-PAI-1 complex to α2-MR for internalization (Conese and Blasi, 1995).

1.2.3.1 The uPA/uPAR cycle

A scheme of the cellular cycle of uPA has been proposed by Blasi (1993). In this cycle, secreted pro-uPA binds to its receptor on the cell surface in an autocrine or paracrine fashion. Following uPA activation by plasmin, receptor-bound uPA can be inhibited by formation of the uPAR:uPA:PAIs complex. Receptor-bound active uPA is neither internalized nor degraded. However, uPAR efficiently mediates internalization and degradation of the uPA:PAI-1 complex. Once internalized, the uPA:PAI-1 complex is degraded in the lysosome but the internalized receptor is recycled to the cell surface (Blasi, 1993).

1.2.3.2 uPAR-mediated internalization of uPA-PAIs complexes

A binding role of uPAR appears to be important in the initial phase of these internalization processes. It has been shown that receptor binding does not shield uPA from the inhibition by PAI-1 and binding of the uPA-PAI-1 complex to uPAR results in internalization and degradation through an interaction with the α2-MR (Nykjær et al., 1992). A quaternary complex uPAR/uPA-PAI-1/α2-MR is formed on the cell surface before the internalization step (Conese and Blasi, 1995). Experiments carried out in the presence of uPAR antagonists have also shown that uPAR binding is required for internalization and degradation of uPA-PAI-2 complexes (Estreicher et al., 1990) and uPA-PN-1 complexes (Conese et al., 1994) through an uPAR/α2-MR-mediated endocytosis mechanism. Thus, binding of receptor-bound uPA to its specific inhibitors leads to internalization and degradation of the complex. However, a recent report showed that uPA-PAI-2 complexes are not internalized and degraded upon binding to uPAR in human myeloid cell line THP-1 cells (Ragno et al. 1995). The reason for this difference is not clear. In addition, the role of α2-MR was confirmed by using the α2-MR associated protein (α2MRAP) and antibodies against α2-MR to inhibit uPAR-mediated uPA-PAI-1 degradation (Nykjær et al., 1992). A requirement of α2-MR for internalization is also demonstrated in cell types that express uPAR but not α2-MR and lack the property of internalization and degradation (Conese and Blasi, 1995). These data show that both uPAR and α2-MR receptors are required for internalization and degradation of the uPA-PAI-1 complex.
Like \(\alpha_2\)-MR, the low density lipoprotein receptor-related protein (LRP) also mediates internalization and degradation of the \(uPA:PAI-1\) complex by co-operation with uPAR (Herz et al., 1992). After lysosomal degradation of the \(uPA:PAI-1\) complex, both LRP and uPAR return to the cell surface. Because \(uPA/uPAR\)-mediated plasminogen activation is involved with cell migration, the function of internalization of \(uPAR:uPA:serpin\) complexes may provide cells a mechanism for continuously changing the proteolytically active surface areas by re-positioning their \(uPA\) receptors (Conese and Blasi, 1995). In addition, \(uPAR\)-mediated internalization of its ligand may be involved in a signal transduction pathway.

### 1.3 \(uPAR\) and Tumor Metastasis

Metastasis is the most fearsome aspect of cancer since cancer patients usually do not die of their initial tumor, but die of metastasis. The malignant potential of solid tumors is related to their ability to invade adjacent tissue and to metastasize. There is growing evidence that the \(uPA/uPAR\) system is involved in tumor invasion and metastasis.

#### 1.3.1 Proteolysis is a key step of tumor metastasis

##### 1.3.1.1 Tumor metastasis

Tumor metastasis is an active process which consists of linked sequential steps including invasion, intravasation and extravasation from the circulatory system, then colonization and angiogenesis at a distant site (Liotta, 1991; Kohn, 1991; Fidler and Ellis, 1994). To produce metastases, tumor cells must complete all the steps in the process. First, tumor cells detach from the primary tumor mass and move into the blood vessels and lymphatics. Then they are carried in the general circulation through the body. Finally, they invade the surrounding tissues and begin to grow into secondary tumors in distant organs. The new tumor mass can induce the growth of new blood vessels, which may eventually become a departure point for more metastasizing cells.

##### 1.3.1.2 Requirement of proteolysis in tumor invasion

Tumor cell migration and invasion into the surrounding tissue are prerequisites for cancer metastasis. To complete metastasis steps, tumor cells have to degrade the basement membrane and surrounding extracellular matrix, which represents the first step in tumor spread (Dansø et al., 1985; Tryggvason et al, 1987). The acquisition of an
invasive phenotype by the tumor cells is thought to depend, in part, on their ability to hydrolyze these basement membrane components (Ossowski and Reich, 1983).

Different proteolysis systems are involved in tumor invasion and metastasis (Mignatti and Rifkin, 1993). Tumor cells can produce various specific degradative enzymes including serine-proteinase (uPA), metalloproteinases (collagenase IV), cysteine-proteinases (cathepsins B and L) and aspartic-proteinase (cathepsin D). All these proteases, especially urokinase, are considered to be intimately involved in the metastatic spread of tumor cells (Danø et al., 1985; Mignatti and Rifkin, 1993).

1.3.1.3 Involvement of uPA-mediated proteolysis in tumor metastasis

A link between uPA-mediated plasminogen activation and cancer is well-established (Blasi, 1993; Fazioli and Blasi, 1994). uPA has been shown to promote tumor invasion and metastasis by triggering a proteolysis cascade in which the activated plasmin can directly and indirectly degrade basement membrane and extracellular matrix (Ossowski et al., 1991; Blasi, 1993). The specific modulation of urokinase expression markedly affects the invasive and metastatic potential of tumor cells since tumor invasiveness and metastasis can be inhibited or stimulated in several model systems by directed increase or decrease of uPA (Ossowski, 1988; Ossowski et al., 1992; Reiter et al., 1993; Crowley et al., 1993). Furthermore, increased uPA expression has been found to be associated with many malignant tumors including carcinomas of the colon (De Bruin et al., 1987; Kohga et al., 1989), lung (Oka et al., 1991), prostate (Camiolo et al., 1984; Achbarou, et al., 1994), and breast (Duffy et al., 1990; Jänicke et al., 1990). In addition, it was noted that high levels of uPA are involved in the metastatic spread of cancer cells without requirement of uPAR (Cajot et al. 1989; Schultz and Zhang, 1992). These data have suggested that uPA-mediated proteolysis is crucial in tumor invasion and metastasis.

While high levels of uPA have been reported occasionally to enhance invasive and metastatic behavior in the absence of uPAR (Schultz and Zhang, 1992), there is convincing evidence that the binding of uPA to its specific receptor uPAR is central to the initiation and modulation of the focal proteolytic events on the cell surface that contribute to tumor invasion and metastasis (Stoppelli et al., 1986; Blasi, 1993).

1.3.2 Correlation of uPAR with tumor invasion and metastasis

1.3.2.1 Elevated levels of uPAR is associated with human carcinomas

Expression of uPAR has been reported in many tumor tissues and tumor cell lines and the ability of cancer cells to metastases is believed to depend on the receptor-bound uPA rather than total uPA activity (Hollas et al., 1991; Bruckner et al., 1992; Karikó et
al., 1993; Crowley et al., 1993). Increased uPAR expression has been detected in colon (Pyke et al., 1991; Ganesh et al., 1994a; Wang et al., 1994), breast (Del Vecchio et al., 1993; Jankun et al., 1993; Bianchi et al., 1994; Costantini et al., 1995), ovarian (Casslén et al., 1991), lung (Veale et al., 1990) and renal cancers (Wagner et al., 1995), compared to normal and benign tumor tissues. Breast carcinomas contain five times more uPAR and 19 times more uPA than benign breast lesions (Del Vecchio et al. 1993). The uPAR is overexpressed in invasive breast cancer but not in normal and benign breast tissues (Bianchi et al., 1994; Costantini et al., 1995). Human renal carcinomas cells modestly overexpress uPAR mRNA but have a significant decrease of uPA mRNA expression (Wagner et al., 1995). In addition, a strong correlation of uPAR levels in cells with invasive and metastatic potential was found in vitro and in vivo studies with human tumor cells lines from melanomas, breast, lung, and colon (Hollas et al., 1992; Pedersen et al., 1993; Wang et al., 1994; Stahl et al., 1994; Kobayashi et al., 1994; Liu et al., 1995). These findings suggest that an elevated level of uPAR is an accompaniment to the invasive phenotype of human cancers.

1.3.2.2 uPAR focuses uPA activity at the tumor invasive edge

The localization of uPA and uPAR in tumor specimens has been further analyzed by several investigators using immunocytochemical techniques. These studies have demonstrated that high numbers of uPAR are expressed at the invasive edge of cancer cells and the tumor mass (Pyke et al., 1991; Del Vecchio et al., 1993; Pyke et al., 1993a; Carriero et al., 1994; Wagner et al., 1995). In breast carcinoma, uPAR is expressed in invasive breast cancer but not in normal breast tissue (Bianchi et al., 1994). uPAR may have a role in providing tissue macrophages with a means of directing proteolysis at sites of breast cancer invasion as anti-uPAR antibodies mostly stained peritumoral macrophage-like cells (Pyke et al., 1993a). In renal cell carcinoma, tumor cells have also been found to modestly overexpress uPAR mRNA and to concentrate exogenous uPA at the tumor cell surface (Wagner et al., 1995). These data indicate that uPAR expression is required for focusing uPA activity at the invasive front. Thus movement of tumor cells through the extracellular matrix is likely to be facilitated by uPA/uPAR-mediated focal proteolysis.

1.3.2.3 uPAR and colon cancer

Human colorectal carcinogenesis has been found to be associated with an increased uPA expression (Blasi, 1993). This increased proteolytic activity may contribute to
colon cancer invasion and metastasis. Like uPA, elevated levels of uPAR have been reported in colorectal carcinomas (Pyke et al., 1991; Wang et al., 1994; Pyke et al., 1995). *In situ* hybridization studies on human colon adenocarcinomas have shown the occurrence of uPA mRNA exclusively in fibroblast-like stromal cells, whereas epithelial tumor cells expressed uPAR mRNA at invasive foci and appeared to be devoid of uPA mRNA (Grøndahl-hansen et al., 1991; Pyke et al., 1991). Therefore, increased uPA activity on the cancer cell surface may represent a mechanism by which uPAR produced by colorectal carcinoma cells recruits uPA to the invasive front. Further studies demonstrate that uPAR positive cells are most abundant at the invasive margin of colon cancer and are identified as various stromal-cell populations including macrophages, fibroblasts, neutrophilic and eosinophilic granulocytes, endothelial cells and cancer cells (Ohtani et al., 1995). The malignant tissues have been shown to have 30-400% more uPAR than the non-malignant adjacent tissues (Wang et al., 1994).

The expression of uPAR has been reported in a number of human colon cancer cell lines (Boyd et al., 1988; Hollas et al., 1991; Baker et al., 1992; Reiter et al., 1993; Wang et al., 1994). Hollas et al. (1991) demonstrated that the potential of ECM degradation by cultured colon cancer cells is dependent on the amount of cell surface receptor-bound uPA. For example, GEO colon cancer cells with low levels of both uPA and uPAR have low ability in ECM invasion, even in the presence of increased exogenous uPA. However, RKO cells with high levels of uPAR and low amounts of uPA can significantly degrade ECM. HCT116 cells with high levels of uPA and uPAR show high invasive potential. These data clearly demonstrate that uPAR in colon cancer cells can recruit exogenous uPA to the cell surface to facilitate extracellular matrix degradation. Therefore, the uPAR-directed uPA proteolysis system may play an essential role in the biology of colorectal carcinoma since uPA activity has been demonstrated to contribute to hepatic metastases of colorectal carcinoma (Sier et al., 1994).

**1.3.2.4 uPAR is a potential prognostic factor**

Measurement of uPA and PAI-1 has been shown to be of prognostic value and increased levels of uPA and PAI-1 may predict a high risk of metastasis and disease-free or overall survival of patients with breast cancer (Jänicke et al., 1994; Bouchet et al., 1994), and colorectal cancer (Ganesh et al., 1994b). Elevated levels of uPAR may also have a clinical implication since uPAR expression appears to be associated with
potential for tumor malignancy. Combined overexpression of uPA, uPAR and PAI-1 has been reported to be associated with breast cancer progression (Costantini et al., 1996). In the tumor progression of human melanoma, neither uPA nor uPAR was detectable in benign or early stages but appeared frequently in advanced primary melanoma and melanoma metastatic lesions (De Vries et al., 1994). In addition, high levels of uPAR expression in human tumors have been reported to correlate with a high risk of recurrence. In breast cancer, uPAR has also been suggested as a prognostic factor in patients with breast carcinoma since it plays a central role in the acquisition of an invasive phenotype (Del Vecchio et al., 1993). Recent studies by Duggan et al. (1995) have reported that patients with breast cancer containing high levels of uPAR had a worse prognosis than patients with low levels of the receptor. But as a prognostic marker in breast cancer, uPAR may be not as strong as uPA. In colorectal cancer, studies by Ganesh et al. (1994a) have also shown that a high uPAR concentration is an independent prognostic factor for 5-year overall survival. Similar data have also been obtained in squamous-cell lung cancer (Pedersen et al., 1994). Thus, uPAR is a likely potential prognostic factor.

1.3.3 Blockade of uPAR in tumor invasion and metastasis

Cancer invasion and metastasis can be inhibited by blockade of uPAR production on the tumor cell surface, providing direct evidence that uPAR is a key molecule in tumor cell invasion. While uPA inhibitors and antibody have been shown to suppress tumor cell metastasis effectively in a number of model systems (Blasi, 1993), data from both in vitro and animal experiments have shown that tumor invasion and metastasis can be inhibited by blocking the uPA receptor (Table 1.1). Different strategies have been used to inhibit or to abolish uPAR expression.

1.3.3.1 Blockade by inactive uPA

Blocking the binding of active uPA to its receptor by inactive mutant uPA (Crowley et al., 1993) or by the amino terminal fragment (ATF) of uPA (Kobayashi et al., 1993 and 1994; Lu et al., 1994) has been shown to decrease invasion by certain types of tumor cells significantly. Using the CAT assay as a sensitive method, it has been demonstrated that competitive displacement of uPA from its cellular binding site by catalytically inactive proteins can inhibit the metastatic capacity of a human prostate carcinoma cell line (PC3) in nude mice (Crowley et al., 1993). Using ATF as a
competitor, inhibition of tumor cell-induced invasion by breast adenocarcinoma cell line MDA-MB-231 and ovarian cancer cell line HOC-1 has also been observed (Lu et al., 1994; Kobayashi et al., 1993). Moreover, in an in vivo spontaneous metastasis model, a synthetic enzymatically inactive mouse peptide (amino acids 17-34) within the growth factor-like domain of uPA has been shown to decrease spontaneous lung metastasis by murine Lewis lung carcinoma (3LL) cells (Kobayashi et al., 1994). These results indicate that occupation of uPAR by inactive uPA or prevention of rebinding of uPA synthesized by tumor cells to the receptors specifically reduces tumor cell invasion and metastasis.

1.3.3.2 Blockade by antibody

Experiments in vitro have demonstrated that the specific antibodies for uPAR expression also decrease tumor cell invasion (Mohanam et al., 1993; Reiter et al., 1993). Polyclonal or monoclonal antibody against uPAR can efficiently block invasion by HT29TX66 human colon carcinoma cells (Reiter et al., 1993) and human glioblastoma cells (Mohanam et al., 1993) which is a similar result to that in competition experiments with a uPA peptide covering the growth-factor domain. These results suggest that blocking binding of uPA to uPAR by antibodies against uPAR is also an effective way to inhibit tumor invasion.

1.3.3.3 Blockade by recombinant soluble uPAR

Recombinant soluble uPAR has also been used as a scavenger for uPA to inhibit cancer cell invasion by blocking the uPA-uPAR interaction on the cancer cell surface. Purified soluble uPAR comprising amino acids 1-277 significantly reduces the proliferation and invasion of an invasive and tumorigenic human ovarian cancer cell line, OV-MZ-6. The soluble uPAR inhibits invasion of OV-MZ-6 cells by 75% using an in vitro matrigel invasion assay (Wilhelm et al., 1994).

1.3.3.4 Blockade by antisense RNA

Using antisense mRNA strategy to block uPAR expression is a new challenge to tumor metastasis. Recent studies by Kook et al. (1994) have shown that tumor metastasis caused by a highly malignant human epidermoid carcinoma cell line (HEp3) can be abolished by antisense mRNA which is 300 bases of the 5' end of the uPAR, including ATG codon. It was demonstrated that the antisense clones resulted in a reduced uPAR mRNA level and a 20-74% reduction in the uPAR sites on the cell surface. The diminished expression of surface uPAR leads to a reduction in invasiveness.
Table 1.1 Effects of uPAR blockade on cancer invasion & metastasis

<table>
<thead>
<tr>
<th>Tumor type</th>
<th>Cell line</th>
<th>Invasion in matrigel or CAM</th>
<th>Tumor metastasis in nude mice</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain</td>
<td>U251, UWR1-3</td>
<td>inhibited¹</td>
<td>ND</td>
<td>Mohanam et al., 1993</td>
</tr>
<tr>
<td>Colon</td>
<td>HT29</td>
<td>inhibited¹</td>
<td>ND</td>
<td>Reiter et al., 1993</td>
</tr>
<tr>
<td>Ovarian</td>
<td>HOC-1</td>
<td>inhibited²</td>
<td>ND</td>
<td>Kobayashi et al., 1993</td>
</tr>
<tr>
<td>Prostate</td>
<td>PC3</td>
<td>ND</td>
<td>inhibited²</td>
<td>Crowley et al., 1993</td>
</tr>
<tr>
<td>Breast</td>
<td>MDA-MB 231</td>
<td>inhibited²</td>
<td>ND</td>
<td>Lu et al., 1994</td>
</tr>
<tr>
<td>Epidermoid</td>
<td>HEp3</td>
<td>inhibited³</td>
<td>abolished³</td>
<td>Kook et al., 1994</td>
</tr>
<tr>
<td>Lung</td>
<td>VA-13</td>
<td>inhibited¹, ³</td>
<td>ND</td>
<td>Quattrone et al., 1995</td>
</tr>
<tr>
<td>Colon</td>
<td>HCT116</td>
<td>inhibited³</td>
<td>inhibited³</td>
<td>Liang et al., 1995</td>
</tr>
</tbody>
</table>

¹, uPAR antibody; 2, inactive uPA and ATF; 3, antisense mRNA; 4, antisense oligonucleotide.
of tumor cells *in vitro* and an increase in tumor latency *in vivo*. More recently, Quattrone et al. (1995) used anti-messenger RNA oligonucleotide to block invasion by SV40-transformed human fibroblasts, VA-13, which contain 10 times more uPAR than their normal counterpart and can efficiently invade the basement membrane matrigel. The data showed that the invasive properties of the transformed cells can be abolished by switching off uPAR expression using the anti-messenger oligodeoxynucleotides. In addition, studies in our laboratory have demonstrated that when transfected with antisense uPAR, the highly metastatic capacity of human colon cancer cells was significantly inhibited in the *in vitro* assay of ECM degradation and in the *in vivo* model of human colon cancer in immunodeficient mice (Liang, PhD thesis, 1994).

Direct interference with the uPA-uPAR interaction or blocking of uPAR expression by using inactive uPA, ATF, antibody or antisense mRNA not only provides more conclusive evidence on the role of uPAR in invasion and metastasis, but also provides the prospect of preventing uPA/uPAR-mediated tumor metastasis. The inhibition of uPAR expression is now a goal of anti-metastatic therapy (Fazioli and Blasi, 1994).

### 1.4 The Roles of uPAR in Cell Proliferation, Chemotaxis, Adhesion and Signal Transduction

As described above, uPAR plays an important role in enhancing and directing uPA-mediated cell surface proteolytic activity which has been shown to be required in normal and pathological conditions (Estreicher et al., 1990; Ellis et al., 1991; Blasi, 1993). Recent studies have suggested that uPAR has additional roles in cell proliferation (Kirchheimer et al., 1989b; Rabbani et al., 1990), chemotaxis (Gyetko et al., 1994; Busso et al., 1994), adhesion (Waltz et al., 1993; Waltz and Chapman, 1994), cell activation (Nykjær et al; 1994), and signal transduction (Dumler et al., 1993 and 1994; Anichini et al., 1994; Cao et al., 1995). These findings demonstrate that the GPI-anchored uPAR is a multifunctional protein.

#### 1.4.1 Effects of uPAR in cell proliferation

Involvement of uPAR in uPA-induced mitogenesis effects has been shown in several experiments (Kirchheimer et al., 1987; Kirchheimer et al., 1989b; Rabbani et al., 1990; De Petro et al., 1994). The binding of both catalytically active and inactive uPA to uPAR can stimulate cell proliferation in many different cell types, suggesting that
uPA-induced cell proliferation is independent of its catalytic activity. uPA has been indicated to act as a growth factor to induce DNA replication and cell division in serum-deprived cultured human skin fibroblasts (De Petro et al., 1994), human epidermal tumor cell line CCL20.2 (Kirchheimer et al., 1987), and in malignant and adjacent unaffected human renal cells (Kirchheimer et al., 1988a). After binding to uPAR, the ATF (amino acids 1-143, which contains a growth factor domain) of uPA has a growth factor-like activity in the osteosarcoma cell line SaOS-2 (Rabbani et al., 1990). Further studies have identified a fucosylated, but not defucosylated, growth factor domain (amino acids 4-43) that is responsible for the growth factor-like activity through binding to uPAR (Rabbani et al., 1990). uPA-induced DNA synthesis can be selectively blocked by uPA antibody and uPAR monoclonal antibody suggesting that interaction of uPA with uPAR is essential to uPA-induced mitogenic effects.

1.4.2 Effects of uPAR in chemotaxis

During cell migration, uPAR can function not only to focus plasmin-mediated ECM degradation on the front edge of migrating cells as described above, but also to mediate chemotaxis (Fibbi et al., 1988; Estreicher et al., 1990; Del Rosso et al., 1993; Anichini et al., 1994). uPA has been shown to act as a chemotaxin for neutrophils (Boyle et al., 1987) and to induce chemotaxis when binding to its specific receptor (Fibbi et al., 1988). Both active and inactive uPA or ATF have been shown to stimulate the migration of human neutrophils (Gudewicz and Gilboa, 1987), bovine adrenal capillary endothelial cells (Fibbi et al., 1988), human epidermal cells (Del Rosso et al., 1993), human fibroblasts and LB6 mouse fibroblasts (Anichini et al., 1994), suggesting that uPA-induced chemotaxis acts independently of its catalytic activity. In two human uPAR-transfected cell lines, a human epithelial cell line WISH and a mouse cell line LB6-uPAR, the exogenous pro-uPA produces up to a 3-fold enhancement of cell migration in matrigel and this receptor-bound uPA-induced migration is both time and concentration dependent, and independent of extracellular proteolysis (Busso et al., 1994). In addition, the rapid redistribution of the uPAR in response to a chemotactic signal has been demonstrated to achieve polarized expression of a powerful proteolytic cascade on the leading edge of migrating cells (Estreicher et al., 1990). Thus, the interaction of uPA and uPAR causes both plasminogen-dependent and plasminogen-independent effects during cell migration.
To be successfully recruited to inflammatory sites, monocytes must accomplish directional migration in a chemotactic gradient (Wilkinson and Haston, 1988). It has been found that human monocytes alter the plasma membrane distribution of uPAR on exposure to a chemotactic gradient, as uPAR becomes tightly localized to the leading edge of migration (Gyetko et al., 1994). Under ECM-free conditions, monocyte chemotaxis is unaffected by receptor-bound uPA activity. In polymorphonuclear leukocytes (PMNs), the interaction of uPAR with complement receptor type 3 (CR3) but not with uPA is required for chemotaxis as PMN chemotaxis is selectively inhibited by disrupting the association between uPAR and CR3 (Gyetko et al., 1995). Blocking uPAR expression with antisense oligonucleotides and monoclonal antibody completely ablates chemotaxis. uPAR was also found to dissociate from CR3 reversibly during neutrophil polarization which is required for cell migration (Kindzelskii et al., 1996). These results suggest that uPAR may play a crucial role in chemotaxis with or without uPA binding.

1.4.3 Effects of uPAR in adhesion

Binding of uPA to uPAR can trigger monocyte differentiation and adhesion independent of the uPA catalytic activity (Nusrat and Chapman, 1991). The interaction of uPA with uPAR induces a rapid cellular adhesion to soluble vitronectin and adsorbed vitronectin in leukemic cell lines exposed to differentiating cytokines (Waltz et al., 1993; Waltz and Chapman, 1994). Both catalytically inactive uPA (DFP-inactivated uPA) and ATF promote adherence of TGF-β1/D3-stimulated U937 cells to purified vitronectin adsorbed to tissue culture plastic (Waltz and Chapman, 1994). In monocytes, binding of uPA or ATF to uPAR promotes cell adhesion involving a cAMP-dependent pathway of signal transduction and requiring protein synthesis and gene expression (Li et al., 1995).

Recently, uPAR has been identified as an adhesion receptor for vitronectin, a adhesive protein in serum (Wei et al., 1994). uPAR binds vitronectin with high affinity in the absence of uPA and stable epithelial cell transfectants expressing membrane-anchored uPAR become strongly adhesive with altered morphology, but vitronectin binding to uPAR can be promoted by concurrent receptor binding to either uPA or ATF (Wei et al., 1994). As a receptor for vitronectin, therefore, uPAR induces adhesion by directly binding to vitronectin. In addition, the continued occupation of uPAR is required for adhesion since this cellular adhesion to vitronectin is reversible and
modulated by uPA/PAI-1 turnover (Waltz and Chapman, 1994). Together, these observations suggest that uPAR appears to modulate cellular trafficking by promoting attachment to vitronectin.

1.4.4 Effects of uPAR in cell activation

The ability of activated T lymphocytes to extravasate and reach inflammatory and malignant foci in the tissues is a basic function of cellular immunity. uPAR is expressed by peripheral blood monocytes, granulocytes and B-lymphocytes (Blasi 1988; Plesner et al., 1994) and its expression is increased during monocyte activation (Kirchheimer et al., 1988b). In vivo, uPAR expression is seen predominantly in inflammatory or malignant tissues. It has been found that uPAR is identical to the monocyte activation antigen previously designated as Mo3, now as CD87 (Min et al., 1992; Sitrin et al., 1994). Recent studies suggest that uPAR functions as a pan T cell activation antigen (Nykjaer et al., 1994) and its expression is highly regulated in cells of the immune system (Min et al., 1992; Nykjaer et al., 1994; Sitrin et al., 1994). By immunofluorescence flow cytometry, uPAR is barely detectable on freshly isolated monocytes, but is prominent on the surface of monocytes activated by culture in media containing soluble inflammatory factors, such as the bacterial products lipopolysaccharide (LPS) and muramyl dipeptide (Todd et al., 1985), and cytokines including TNFα, M-CSF, G-CSF, and IL-3 (Kirchheimer et al., 1988b; Todd et al., 1990). In response to PMA and TNFα, polymorphonuclear leukocytes translocate uPAR to the plasma membrane from internal stores (Plesner et al., 1994). In addition, uPAR expression is inducible in subsets of T cells by cytokines and all activated cultures show co-expression of uPAR and CD25 suggesting that uPAR acts as an activation antigen involved in extravasation and migration of activated T cells (Nykjaer et al., 1994). These studies suggest that uPAR expression in the immune cell plasma membrane may be related to the ability of the cell to respond to immune system challenge.

1.4.5 Effects of uPAR in signal transduction

The ability of cells to communicate with and respond to their external environment is critical for their continued existence. The specific membrane proteins present on the cell surface are functionally important in transmembrane signaling. Many of these cell-surface proteins are known to be anchored in the plasma membrane by covalent linkage
to GPI moiety which directs their membrane localization and function (Brown, 1993; Englund, 1993). GPI-anchored protein has been demonstrated to have roles in signal transduction (Morgan et al., 1993; Casey, 1995). So far, a number of intriguing observations support a role for uPAR in signal transduction acting as a signaling receptor. Following uPA binding to uPAR this GPI-anchored receptor can trigger an intracellular signal to induce several biological responses such as cell migration (Busso et al., 1994), differentiation (Nusrat and Chapman, 1991) and mitosis (Anichini et al., 1994)

1.4.5.1 uPAR triggers production of second messengers

uPA binding to its receptor has been demonstrated to cause production of some second messengers. In human fibroblasts and mouse fibroblasts transfected with the human uPAR gene, a second messenger diacylglycerol (DAG) is produced following chemotactic and mitogenic uPA-uPAR interaction (Anichini et al., 1994). This uPAR-dependent DAG production causes an insulin-like stimulation, but does not affect inositol lipid metabolism and intracellular $\text{Ca}^{2+}$ content. Both uPA and ATF are known to stimulate DAG formation and uPA is 3 to 4-fold more efficient than ATF. In addition, uPA-stimulated chemotactic migration and mitogenic activity in these cells can be impaired by down-regulation of protein kinase C (PKC) with PMA (Anichini et al., 1994). This uPA or ATF-stimulated DAG production was also observed in human epidermal cells (Del Rosso et al., 1993). These results suggest that the sole stimulation of uPAR is able to increase DAG formation directly and allows the cell to reach the DAG threshold level required to induce DNA synthesis or chemotaxis.

Moreover, recent studies by Cao et al. (1995) have demonstrated that uPAR is involved in regulation of calcium, a second messenger of uPA, which is responsible for production of superoxide anions by stimulated human neutrophils. In neutrophils from a patient with adhesion-deficient leukocytes, and 3T3 transfectants expressing CR3, uPAR or both, the calcium levels of neutrophils is increased in the presence of uPA (Cao et al., 1995). The uPA-dependent calcium rise is mediated in a CR3- and uPAR-dependent fashion, since this signal is inhibited by mAb directed against either CR3 or uPAR and requires intact uPA. Both mAb-mediated inhibition of signaling and gene transfection experiments suggest that CR3 and uPAR functionally interact to prime neutrophils for superoxide release.
1.4.5.2 uPAR-induced phosphorylation

uPAR appears to trigger phosphorylation of signal-transducing molecules in its signal transduction pathway. Recent studies by immunoprecipitation (Dumler et al., 1993) and immunocapture (Busso et al., 1994) have found that uPAR is associated with some phosphorylated protein, which provides further evidence to support a role of uPAR in signal transduction. uPAR binds to a 38 kDa phosphorylated protein after uPA stimulation in U937 cells (Dumler et al., 1993). The binding of uPAR to phosphorylated proteins has been found using immunocapture with antibodies against uPAR or pro-uPA (Busso et al., 1994). Following pro-uPA stimulation two phosphorylated proteins, cytokeratin 18 (CK18) and 8 (CK8), were shown to complex with uPAR in a human epithelial cell line, WISH. A similar pattern of phosphorylation was also observed with a lung carcinoma cell line, A459, but not HL60, LB6-uPAR or HEp3 cell lines, suggesting that these uPAR-complexes may be specific for simple epithelia (Busso et al., 1994). Further studies have suggested that the uPAR-associated kinase may belong to the PKC family since PKC-specific inhibitor, and not PKA or tyrosine kinase inhibitors, blocked the in vitro phosphorylation of CK8 and CK18 (Busso et al., 1994). As tyrosine phosphorylation is the hallmark of many signal transduction pathways that promote cell growth and differentiation, these results suggest that uPA might act as a para- or autocrine hormone and is mediated by uPAR in the uPA-stimulated signal transduction pathway.

The uPAR is a GPI-anchored protein that does not have a transmembrane domain or intracellular kinase domain (Roldan, et al., 1990). Therefore, a transmembrane protein partner has been implicated in uPAR-mediated signal transduction. Although the mechanisms of uPAR-mediated signaling are still unclear, identification of several uPAR-associated proteins will improve our understanding of the role of uPAR.

1.4.5.3 uPAR in regulation of gene expression

The regulation of specific gene expression by binding of uPA to the GPI-anchored uPAR has been reported (Dumler et al., 1994; Rao et al., 1995). The interaction of uPA with uPAR has been found to induce cathepsin B and gelatinase gene expression during macrophage differentiation (Rao et al., 1995). The expression and function of uPAR is believed to be important to the pattern of protease expression by differentiating phagocytes since antibodies to uPA or uPAR that prevent receptor binding can inhibit induction of both protease gene expression and protein production in PMA-stimulated
myeloid cells. This protease expression can also be blocked by mannosamine, an inhibitor of glycolipid anchor assembly, but not by anti-catalytic uPA antibodies and excess inactive uPA, suggesting that uPAR occupancy per se regulates protease expression (Rao et al., 1995).

Binding of uPA to uPAR has also been reported to induce a rapid and transient c-fos gene expression in OC-7 ovarian carcinoma cells (Dumler et al., 1993). uPAR may induce c-fos gene expression by a transmembrane signal pathway involved in protein tyrosine kinase (PTK) activation since the pretreatment of the cells with PTK inhibitors, but not inactive uPA, can abrogate the effect. In addition, a soluble uPAR fragment competing for binding of uPA to uPAR also inhibits c-fos induction. These results suggest that uPAR activation by uPA binding on the cell surface may initiate a transmembrane signal to activate nuclear transcription factors by PTK signal transduction pathway. Because c-fos is a component of transcription factor AP-1 (Angel and Karin, 1991), uPAR-induced c-fos gene expression may result in a pronounced biological response.

1.4.6 uPAR-knock-out mice

Carmeliet et al. (1994) have reported that transgenic mice lacking uPA develop and reproduce normally and exhibit remarkably limited phenotypic abnormalities in the absence of other challenging factors. Similarly, recent studies with uPAR-deficient mice has also shown that uPAR is not essential for mouse development or fertility because homozygous uPAR -/- mice were born and survived to adulthood with no overt phenotypic abnormalities (Bugge et al., 1995; Dewerchin et al., 1996). Although the uPAR-/- mice display normal endogenous thrombolytic capacity, the activated peritoneal macrophages fail to promote plasminogen activation in vitro suggesting an impaired uPA-mediated plasminogen-activating potential. In addition, the loss of the uPAR resulted in a redistribution of uPA in some tissues but had no impact on pro-uPA activation in the urogenital tract. Thus, in the absence of other challenging factors such as infection, injury, or other functional deficits, uPAR deficiency does not compromise fertility, development, or homeostasis. The availability of uPAR- and uPA-deficient mice will provide a valuable model to define the precise roles of uPAR in both normal and pathological conditions.
1.5 Regulation of uPAR Expression

Since the complete cDNA encoding for human uPAR from monocytes was isolated and sequenced (Roldan et al., 1990), regulation of uPAR expression has been investigated in many cell types.

1.5.1 uPAR expression

uPAR is expressed by many different cell types including peripheral blood monocytes (Miles and Plow, 1987), granulocytes (Nykjær et al., 1990) and B-lymphocytes (Plesner et al., 1994), activated T-cells (Nykjær et al., 1994), neutrophils (Gyetko et al., 1995), fibroblasts (Anichini et al., 1994), endothelial cells (Pepper et al., 1993; Langer et al., 1993) and epithelial cells (Limongi et al., 1995), as well as various tumor cell lines (Stoppeli et al., 1986; Boyd et al., 1988; Hollas et al., 1992; Mohanam et al., 1993).

In adult human tissues, however, the pattern of uPAR expression has not yet been investigated in sufficient detail. In circulating T lymphocytes uPAR expression is generally low, but higher expression is frequently found among patients suffering from infectious disease, e.g. virus infections (Nykjær et al., 1994). Migrating cells express higher uPAR which is localized to the leading front of migrating monocytes and invading tumor cells (Estreicher et al., 1990; Pyke et al., 1991; Pepper et al., 1993).

uPAR is expressed by tumor cells and by nonmalignant stromal cells infiltrating the tumors (Dano et al., 1994; Ohtani et al., 1995) and a significant higher expression of uPAR has been observed in malignant tumors than benign tissues (Pyke et al., 1991; Del Vecchio et al., 1993; Bianchi et al., 1994). A ligand-binding uPAR without glycolipid anchor has been found in the ascitic fluid and plasma of patients with ovarian cancer (Pedersen et al., 1993). An alternatively spliced mRNA variant of uPAR lacking the GPI anchor, which is generated by alternative splicing in the middle of the third domain, has also been characterized and is expressed in various human cell lines and tissues, but the corresponding protein has not been identified (Pyke et al., 1993b).

Cancer cells express higher uPAR numbers to promote cell invasion. The levels of urokinase receptor in several colon carcinoma cell lines have been determined by using a radioreceptor assay (Boyd et al., 1988). The aggressive and poorly differentiated colon cancer cell lines, RKO and HCT116, express higher numbers of the uPA receptor (>10^5).
sites/cell), while the more less invasive cell lines, CBS, GEO, and FET, possess fewer receptors (<1.5 X 10⁴ sites/cell). The relevance of uPAR expression to uPA-mediated laminin degradation in cultured colon cancer cells has also been observed (Schlechte et al., 1989). The invasive colon cancer cell lines possess 10-fold more uPA receptors than their non-invasive counterparts. This is due to a difference in transcription rates suggesting that transcriptional activation of the uPAR gene, but not gene amplification or enhanced mRNA stability, is the primary mechanism for the elevated expression of the receptor in invasive colon cancer (Wang et al., 1994). Altogether, uPAR appears to be expressed in various cell types, but the levels of expression may be regulated differently according to circumstance.

1.5.2 Regulation of uPAR expression

uPAR expression has been shown to be regulated by various factors including phorbol ester (Stoppelli et al., 1985; Lund et al., 1991a; Langer et al., 1993), cytokines (Kirchheimer et al., 1988b; Sitrin et al., 1994; Wang et al., 1994) and growth factors (Todd et al., 1990; Mandriota et al., 1995; Lund et al., 1995) in many different cell types.

1.5.2.1 Regulation by phorbol esters

Phorbol 12-myristate 13-acetate (PMA), a tumor promoter, has been shown to increase uPAR expression strongly in many cell types and tumor cell lines including monocytes (Picone et al., 1989; Lund et al., 1991a), neutrophils (Plesner et al., 1994), endothelial cells (Langer et al., 1993), HeLa cells (Estreicher et al., 1989), A549 lung carcinoma cells (Lund et al., 1991b and 1995) and colon cancer cells (Wang et al., 1994). PMA-induced expression of uPAR has been a typical model in studies of uPAR gene regulation since PMA treatment of cells leads to a profound increase in both mRNA and protein levels of uPAR (Stoppelli et al., 1985; Lund et al., 1991a). In human monocyte-like U937 cells, PMA-induced cell differentiation is accompanied by a strong and rapid increase of uPAR mRNA which is paralleled by a time-dependent increase in uPAR protein (Lund et al., 1991a). The uPAR mRNA level reaches a maximal 50-fold enhancement after 24 h of treatment. Nuclear run-on experiments have shown that uPAR expression is transcriptionally up-regulated by PMA in monocytes. A time-dependent increase in the uPAR gene transcription rate after exposure of U937 cells to PMA suggests that an increased gene transcription is responsible at least in part for the
accumulation of the uPAR mRNA and for the subsequent increase in uPA-binding capacity (Lund et al., 1991a). In A549 cells, PMA has been found to regulate uPAR expression both transcriptionally and post-transcriptionally since it can increase uPAR transcription and the stability of uPAR mRNA (Lund et al., 1995). PMA also induces a rapid increase of uPAR in isolated neutrophils (Plesner et al., 1994) and isolated T cells (Nykjør et al., 1994). PMA-induced increase of membrane uPAR is accompanied by enhanced receptor synthesis and elevated uPAR mRNA levels. This increased ATP-binding activity peaks within 36 h, whereas the nontumor-promoting phorbols PMA-Me and 4a-PDD have no effect (Nykjør et al., 1994).

PMA has been thought to act through the protein kinase C (PKC)-dependent signal transduction pathway since PKC is the only known receptor of PMA in cells (Nishizuka, 1986). It has been confirmed that PMA induces uPAR expression by the PKC pathway (Langer et al., 1993). In endothelial cells, PMA induces a time- and concentration-dependent increase of uPAR expression. This PMA-induced increase in uPAR mRNA is detectable within 30 minutes and a 6-fold increase was observed with PMA stimulation for 20 hours (Langer et al., 1993). H7, an inhibitor of PKC, can inhibit PMA-stimulated uPAR expression suggesting that PMA-stimulated uPAR expression is mediated by the PKC signal transduction pathway. This evidence indicates that uPAR expression is transcriptionally regulated by PMA through the PKC signal transduction pathway.

1.5.2.2 Regulation by cytokines

Studies into the regulation of uPAR expression by cytokines have suggested a more complicated mechanism. Data from several laboratories indicate that uPAR expression is modulated by a variety of cytokines including TNFα, IFNα, IFNβ, IL-2, IL-3, IL-4, CSF, GM-CSF and TGFβ. In monocytes, the macrophage-activators, IFNγ and TNFα, increase uPAR expression (Kirchheimer et al., 1988b). Treatment of cells with IFNγ increases the receptor numbers 5-fold, and with TNFα the increase is 2-fold. Co-stimulation with IFNγ and TNFα leads to a super-induction of uPAR expression (Kirchheimer et al., 1988b). Similarly, monocytes increase cell surface uPAR expression in response to CSF-1, GM-CSF and IL-3 (Kirchheimer et al., 1989a; Todd et al., 1990). Further studies have indicated that uPAR expression is regulated in a cytokine-specific fashion in U937 mononuclear phagocytes (Sitrin et al., 1994). IFNγ increases neosynthesis and surface expression of uPAR, whereas TNFα only induces a substantial
increase in uPAR mRNA with no effect on cell surface expression. In addition, IFNγ and TNFα produce strikingly different effects on the fate of newly synthesized uPAR. TNFα induces an increase in the level of uPAR protein released into conditioned medium resulting in a relatively constant level of uPAR and receptor-associated uPA on the cell surface, whereas IFNγ has no effect (Sitrin et al., 1994). This cytokine-specific modulation of uPAR may be important in regulating the function of mononuclear phagocytes in inflammation and tissue repair.

Regulation of uPAR expression in the immune cell plasma membrane may be intimately tied to the ability of the cell to respond to immune system challenge since its expression is inducible in subsets of T cells by a number of cytokines including IL-2, IL-4 and IL-7 (Wolf, 1994; Nykjaer et al., 1994). IL-2, IL-4 and IL-7 can increase uPAR presentation on 20-50% of the T cell population. Co-stimulation with IL-2 and IL-7 shows an additive effect. However, IL-4 can inhibit IL-2 or IL-7-induced increase of uPAR expression (Nykjaer et al., 1994). Regulation of uPAR expression by TGF-β1 appears to differ between cell types as TGF-β1 inhibits cytokine-induced uPAR expression in T cells (Nykjaer et al., 1994), but it increases uPAR expression in human A549 lung carcinoma cells (Lund et al., 1991b) and bovine vascular smooth muscle cells (Reuning et al., 1994). In A549 cells, an early increase in uPAR mRNA level was observed with a maximal 15-fold enhancement after 24 h of treatment. But nuclear run-on experiments demonstrated only a moderate 3-fold increase in transcription rate suggesting that TGF-β1 may regulate uPAR expression at both transcriptional and post-transcriptional levels (Lund et al., 1991b and 1995). Together, these data suggest that uPAR expression in different cell types is selectively regulated by different cytokines.

1.5.2.3 Regulation by cAMP

cAMP is an important metabolic effector molecule that controls a variety of cellular processes including the expression of specific genes such as uPA (Shinbo et al., 1995). Like uPA, uPAR expression is also regulated by intracellular cyclic AMP levels in some cell types. In cultured human umbilical vein endothelial cells (HUVEC), raising intracellular levels of cAMP by incubation with forskolin, a cAMP stimulator, induces a time- and concentration-dependent increase in expression of uPAR (Langer et al., 1993). Forskolin can induce an approximately 2-fold increase in the expression of uPAR mRNA within 2 hours. A similar increase is also observed with a cAMP analogue 8-
brcAMP. The cAMP-induced uPAR expression was confirmed by pretreatment of cells with deoxyadenosine monophosphate (DAM), an inhibitor of adenylyl cyclase, which leads to an inhibition of forskolin-stimulated uPAR expression (Langer et al., 1993). In A549 cells, however, cAMP analogues did not induce any change in the uPAR mRNA level (Lund et al., 1995). These data suggest that the cAMP-dependent pathway may be involved in controlling uPAR expression in a cell-specific manner.

1.5.2.4 Regulation by mitogens

It has also been reported that uPAR expression is increased by some mitogens such as thrombin (Reuning et al., 1994; Yang et al., 1995), retinoic acid (Tapiovaara et al., 1994), PHA and Con A (Nykjær et al., 1994), and vascular endothelial growth factor (VEGF) (Mandriota et al., 1995). For example, thrombin increases uPAR mRNA 4-fold and uPA binding to uPAR 2.5- to 5-fold by activation of the thrombin receptor in human retinal pigment epithelial (RPE) cells (Yang et al., 1995). In human promyelocytic NB4 cells, both uPA and uPAR mRNA levels were increased by retinoic acid which was accompanied by cell differentiation (Tapiovaara et al., 1993). VEGF, a potent angiogenic factor and endothelial cell-specific mitogen, has been demonstrated to induce angiogenesis which is accompanied by increased uPAR expression and uPA activity on the endothelial cell surface. uPAR mRNA was increased in a dose- and time-dependent manner in response to VEGF (Mandriota et al., 1995).

1.5.2.5 Regulation by uPA

Urokinase, a ligand for uPAR, also participates in regulation of uPAR expression (Kirchheimer et al., 1988b; Hollas et al., 1992). In human monocytes, uPA has been shown to inhibit INFγ-induced, but not TNFα-induced uPAR expression, suggesting that uPA, IFNγ and TNFα collectively regulate the number of uPAR on the cell surface (Kirchheimer et al., 1988b). Furthermore, overexpression of uPA has been shown to reduce uPAR expression on the cell surface via an autocrine stimulation (Hollas et al., 1992). The mechanisms of down-regulation of uPAR expression by its ligand is unclear.

1.5.2.6 Regulation by other factors

In addition to PMA, cytokines, cAMP and mitogens, uPAR expression is also up-regulated by many other agents including dexamethasone (Lund et al., 1991), bacterial products LPS (Todd et al., 1985; Alumus-Jacobs et al., 1995). Furthermore, uPAR expression is modulated by physical factors such as heat shock. Recent studies have shown that when cultured at 43°C, the level of uPAR mRNA in
human umbilical vein endothelial cells (HUVECs) increased up to 2.2-fold, which was followed by increased uPA binding and cellular fibrinolytic activity (Fukao et al., 1996).

1.5.3 Butyrate and gene regulation

Butyrate is one of the short chain fatty acids (SCFAs) produced in the lumen of the colon by bacterial fermentation of dietary fibre (Cummings et al., 1987). At concentrations present in the lumen it inhibits cell proliferation and stimulates cell differentiation (Clausen et al., 1991; Velcich and Augenlicht, 1993; Newmark, 1994). Butyrate also induces apoptosis in many cell types including colonic carcinoma cells (Hague et al., 1993; Calabresse et al., 1993; Filippovich et al., 1994). Butyrate has been reported to transcriptionally and post-transcriptionally modulate expression of many genes, such as c-fos and c-jun (Nishina et al., 1993; Rabizadeh et al., 1993; Souleimani and Asselin, 1993a), c-myc (Heruth et al., 1993; Souleimani and Asselin, 1993b), hsp70 (Filippovich et al., 1994), α-fetoprotein and albumin (Tsutsumi et al., 1994) in a positive or negative manner. For example, apoptosis induced by butyrate is accompanied by a marked increase in hsp70 mRNA (Filippovich et al., 1994). It rapidly increases c-fos and c-jun gene expression and reduces c-myc expression in many different cell types including colon carcinoma cells (Nishina et al., 1993; Heruth et al., 1993; Rabizadeh et al., 1993; Souleimani and Asselin, 1993a and 1993b). In colon cancer cells, butyrate does not affect c-myc expression at the level of transcriptional initiation or elongation, but at a post-transcriptional level. Cycloheximide blocks butyrate-dependent reduction of c-myc mRNA levels (Souleimani and Asselin, 1993b). Whereas butyrate induces c-fos gene expression at both transcriptional and post-transcriptional levels. Butyrate increases c-fos mRNA very rapidly at a post-transcriptional level and transcriptionally at later times. An ATF-CRE binding site located between -63 and -54 relative to the c-fos transcriptional start site has been identified to be responsible for butyrate-induced c-fos transcription (Souleimani and Asselin, 1993a). Butyrate-induced overexpression of c-jun and c-fos leads to a remarkable increase in AP-1 binding activity (Mollinedo et al., 1993). In addition, a butyrate-responsive element has been found in the 5' flanking regions of mouse calbindin-D28k gene (Gill and Christakos, 1993).

As well as effectively potentiating differentiation and apoptosis, butyrate has been suggested to have a protective role in tumorigenesis of the colon (Hague et al., 1993).
and 1995; Heerdt et al., 1994). Butyrate has been shown to modify the activity of various membrane bound enzymes (Wasserman et al., 1989). The effects of butyrate on the genes implicated in the cell surface proteases that may determine adenocarcinoma invasion and metastasis, however, are unknown. Given the pivotal role played by the uPA/uPAR system in colon cancer cell invasion and metastasis, it is important to study the effect of butyrate on their expression.

In summary, uPAR is expressed in many different cell types, which appears to be according to its function in physiological and pathological processes. Like uPA, uPAR expression is transcriptionally or post-transcriptionally regulated by various agents. These agents affecting intracellular signal transduction pathways participate in regulation of uPAR expression and thus may influence various uPAR-mediated processes. In many cultured cancer cells, the level of uPAR is high, but it can be further increased by these agents. However, while regulation of uPA transcription has been widely studied, the mechanism of controlling uPAR transcription remains poorly understood. Therefore, although the regulation of uPAR gene expression could be largely explained by the different transcription rates of the gene, elucidation of the mechanisms of uPAR gene activation will have to await the cloning and sequencing of the regulatory sequences of this gene. In addition, identification of regulatory elements and transcription factors that control uPAR gene transcription might be useful to develop a new strategy for prevention of uPAR-mediated tumor invasion and metastasis.

1.6 Aims of This Thesis

Based on the evidence presented in this introduction and knowledge of the uPAR gene, uPAR gene expression is critical to uPA-mediated pericellular proteolysis during cell migration and tumor metastasis. Many studies of uPAR have focused on its structure, function and expression as well as its relationship with tumor metastasis. A number of structural and functional features of the uPAR molecule have been elucidated. Regulation of uPAR expression by various agents in different cell types has been demonstrated. However, because its genomic organization and regulatory elements of its control region have not been isolated and characterized, the molecular mechanism of regulation of uPAR expression is not clear. Regulation of uPAR gene expression is
most likely due to the interaction of specific *cis*-acting DNA elements with specific trans-acting regulatory factors. Therefore, the major aims of this thesis were to investigate the molecular mechanisms of regulation of the uPAR gene expression. This will include cloning and characterizing the human uPAR gene, exploring the regulatory DNA elements and regulatory proteins that control uPAR gene expression and looking for new molecular evidence of uPAR regulation, with a view to a better understanding of the tumor metastasis process and the possibility of therapy prevention.

The specific aims were:

- to isolate and characterize the human uPAR gene and to analyze the putative regulatory elements located in the 5'-flanking region.
- to identify the regulatory elements of the human uPAR gene and to assess the contribution of these elements to uPAR transcription.
- to investigate the interaction of the promoter region with nuclear transcription factors which might control uPAR expression.
- to examine the effects of sodium butyrate on uPAR expression in human colon cancer cells.
CHAPTER 2
GENOMIC STRUCTURE OF THE HUMAN UROKINASE RECEPTOR GENE

2.1 Introduction

In eukaryotes, three different classes of genes are transcribed by distinct RNA polymerases. The first class is ribosomal RNA (rRNA) genes transcribed by RNA polymerase I; the second class is messenger RNA (mRNA)-encoding genes transcribed by RNA polymerase II; and the third class is small RNAs such as 5S RNA and transfer RNA (tRNA) transcribed by RNA polymerase III (Green, 1992). Most mRNA-encoding genes are interrupted by non-coding sequences and divided into exons and introns. A messenger RNA produced by RNA splicing consists only of the series of exons which are joined together in mRNA in the same order as their organization in DNA (Lewin, 1990).

Regulation of gene expression is controlled at many different levels such as activation of gene structure, initiation of transcription, transcript processing, transport to cytoplasm and translation of mRNA. The study of genomic DNA structure provides an important molecular basis for understanding transcription regulation of eukaryotic genes since most regulatory elements are located in the non-coding region of a gene. Enhancers or silencers, which positively or negatively control gene transcription, are distributed widely throughout the gene including 5'- and 3'-flanking regions and introns (Khoury and Gruss, 1983; Brand et al., 1985). Identification and characterization of these regulatory elements has become an essential step in investigating mechanisms of gene regulation. A study of the genomic structure of a gene, therefore, is necessary to understand not only gene evolution, but also regulation of gene expression.

uPAR gene expression has been implicated in many important biological processes including cell invasiveness and migration as described in chapter 1. The uPAR is a 55-60 kDa glycoprotein containing 313 amino acid residues and is organized as three repeats of similar cysteine-domains (Ploug et al., 1991). The 1.4 kb human uPAR cDNA has been cloned and characterized (Roldan et al., 1990). The uPAR protein is anchored on the cell surface by a GPI-anchor and specifically recognizes uPA. The receptor-bound
uPA catalyses the formation of plasmin on the cell surface to generate the proteolytic cascade that contributes to the breakdown of basement membranes and extracellular matrix (Blasi, 1993). uPAR is expressed on the surface of a number of cell types of normal and malignant origin. Recent evidence has also shown the soluble uPAR variants detected in ascitic fluid from patients with ovarian cancer (Pedersen et al., 1993) and in plasma from patients with paroxysmal nocturnal hemoglobinuria (Ploug et al., 1992), suggesting that mutation of the uPAR may correlate with malignancy or disease. In addition, it has been demonstrated that uPAR is expressed at the invasive front in human colon adenocarcinomas (Pyke et al., 1991) and in tumor-associated macrophages in human breast carcinoma (Pyke et al., 1993a). Blockade of the uPAR on human PC3 prostate carcinoma cells by inactive uPA markedly inhibited metastatic activity (Crowley et al., 1993). In vitro, the uPAR is inducible in T cells and monocytes by the tumor promoter, PMA, and some cytokines (Nielsen et al., 1988; Nykjaer et al., 1994). The regulation of the uPAR gene, therefore, may be important in the control of the malignant behavior of many types of cancer cell. Human uPAR has been characterized at the protein and the cDNA levels (Vassalli et al., 1985; Roldan et al., 1990); however, the genomic organization of the uPAR gene from any species has not previously been reported and the mechanisms that regulate transcription of the uPAR gene are largely unknown. To elucidate the mechanisms involved in the function and regulation of the uPAR gene, it is necessary to study the structural organization of this gene. The results of such study will also provide significant information for dissecting the cis-acting sequences necessary for the expression of the gene.

In this chapter, I describe the cloning and characterization of the human uPAR gene including its 5'-flanking region. In previous work performed in this laboratory part of the human uPAR gene containing exons 4-7 has been cloned and partially sequenced by screening a human genomic library using human uPAR cDNA as a probe. Subsequently, approximately 8 kb genomic DNA sequence covering the region from intron 3 to 6 was obtained from our collaborator Dr. Lorin Johnson, Salix Pharmaceuticals, Inc. Palo Alto, California, USA. Based on these results, genomic clones covering the full length uPAR gene were isolated by screening the same human genomic library and the genomic DNA including the 5'- and 3'-flanking regions of the gene was sequenced and characterized.
2.2 Materials and Methods

2.2.1 Materials

Radionucleotides [α-32P]dCTP (3000 Ci/mmol) and [γ-32P]ATP (3000 Ci/mmol) were obtained from Amersham International plc., UK. Restriction endonucleases were purchased from Bethesda Research Laboratories, Promega or Pharmacia. QuickPrimer™ kit was obtained from Pharmacia, Sweden. DNA ligation system kit was obtained from United States Biochemical Corp., USA. A human blood genomic (normal peripheral blood) library prepared in the Lambda GEM11 vector and the E. coli bacterial host KW251 used to propagate bacteriophage λ vectors were obtained from Promega, USA. pBluescript SK- vector was obtained from Stratagene, USA. The E. coli bacterial strain DH5α used for bacterial transformation was from Clontech, USA. Wizard™ minipreps DNA purification kit was obtained from Promega, USA. The 1.14 kb human uPAR cDNA (Roldan, et al., 1990; Webb et al., 1994), containing the entire coding region of uPAR, was a gift from Dr. E. K. O. Kruithof, Lausanne, Switzerland. The human histiocytic lymphoma cell line U937 (Fischer et al., 1980) and colon cancer cell line HCT116 (Brattain et al., 1981) were obtained from the American Type Tissue Collection. The human colon cancer cell line, LIM1215, was provided by Dr. R. T. Whitehead, Ludwig Institute for Cancer Research, Melbourne, Australia (Whitehead et al., 1985).

2.2.2 Southern blot hybridization of genomic DNAs

2.2.2.1 Cell culture

Human lymphoma cell line U937, human colon cancer cell lines HCT116 and LIM1215 were cultured in RPMI-1640 medium (Gibco-BRL) supplemented with 10% fetal calf serum (FCS), 2 mM L-glutamine, 50 μg/ml penicillin and 50 μg/ml gentamycin. All cell lines stained negative for Mycoplasma contamination using Hoechst stain 33258.

2.2.2.2 Isolation of high molecular weight DNA from cultured cells

High-molecular-weight genomic DNA was prepared from cultured cells according to the method described by Sambrook et al.(1989). After washing with PBS, ~1 x 10^8 cells were harvested by centrifugation in a 15 ml Falcon tube.
resuspended in 1 ml of extraction buffer (10 mM Tris-HCl, pH 8.0, 0.1 M EDTA, 20 µg/ml pancreatic RNase, 0.5% SDS) and protein digestion was carried out by adding proteinase K (0.1 mg/ml) at 56°C overnight. Following phenol extraction, the genomic DNA was precipitated with ethanol.

2.2.2.3 Restriction digestion of DNA and agarose gel electrophoresis

Digestion of genomic DNA with restriction enzymes was performed according to standard method (Sambrook et al., 1989). Generally, 0.5-1 µg of DNA was completely digested with 5 units of appropriate enzyme in 10 µl volume at 37°C for 2 h.

10 µg digested genomic DNA was separated by electrophoresis through an 1% agarose gel containing 0.5 µg/ml ethidium bromide in 1X TAE buffer (40 mM Tris-acetate, pH 8.0, 1 mM EDTA), at 100 volts for 3-4 h. The agarose gel was then photographed under UV light.

2.2.2.4 Blotting of DNA to reinforced nitrocellulose

After agarose gel electrophoresis, the DNA was alkaline-denatured with 500 ml of denaturing solution (0.5 M NaOH, 1.5 M NaCl) for 20 min and then neutralized with 500 ml of neutralization solution (0.5 M Tris-HCl, pH 7.5, 1.5 M NaCl) for 20 min. The denatured DNA was transferred onto reinforced nitrocellulose membranes (Schleicher & Schuell) as follows: the transfer apparatus was set up by placing a glass plate over a tray containing 20X SSC solution (0.3 M sodium citrate, pH 7.0, 3 M NaCl). Three sheets of pre-soaked Whatman 3MM paper were placed on the glass plate with their ends submerged in the 20X SSC solution. The gel was placed on top of the 3MM sheets and the pre-soaked reinforced nitrocellulose was placed on top of the gel. The transfer apparatus was then covered with plastic film to minimize evaporation and the section of plastic film covering the top piece of 3MM paper was cut out with a razor blade. A stack of paper towels was positioned over the top piece of 3MM paper in order to transfer the DNA from the gel to the reinforced nitrocellulose filter via capillary action. The transfer was left for ~20 h at room temperature with a ~500 g weight placed on top to ensure even transfer.

The reinforced-nitrocellulose filter was then removed and air dried on Whatman 3MM paper for 20 min. The filter was then baked between two sheets of Whatman paper in a vacuum oven at 80°C for 2 h.
2.2.2.5  Probe labeling

A 1.14 kb uPAR cDNA fragment was isolated from pBlue-uPAR by digestion with EcoRI and XbaI and used as a probe for both Southern blot and plaque hybridization. The random primer method was used for cDNA probe labeling according to the manufacturer's protocol (\textsuperscript{T}QuickPrimer™ Kit, Pharmacia). About 50 ng of cDNA fragment in 33 µl TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) was boiled for 5 min and then chilled on ice for 2 min. The labeling reaction was carried out by adding 10 µl of reagent mix, 5 µl of [\(\alpha\)-\textsuperscript{32}P]dCTP (3000 Ci/mmol, Amersham) and 1 µl of T7 DNA polymerase and incubating at 37°C for 10 min.

For characterization of the 5' flanking region, a 184 bp Smal-NruI fragment, which contains 47 bp exon 1 and 137 bp 5' flanking sequence, was isolated from genomic clone (\(\lambda\)11, see below) by digestion with Smal and NruI. About 50 ng of DNA fragment was labeled in 10 µl of reaction solution containing 3 µl of [\(\gamma\)-\textsuperscript{32}P]ATP and 1 µl of T4 polynucleotide kinase (Promega) at 37°C for 10 min.

After labeling, the DNA probe was purified by centrifugation through a Sephadex G-50 column (Nick™ column, Pharmacia) at 1,500 rpm for 5 min. The specific radioactivity of cDNA probes was measured by liquid scintillation counting. DNA probes were denatured by boiling 5 min then cooled on ice for 10 min before hybridization.

2.2.2.6  Southern hybridization

The baked reinforced nitrocellulose filter was pre-hybridized at 42°C for at least 4 h (up to overnight) in 30 ml of prehybridization solution (5X SSC, 50% formamide, 5X Denhardt's solution, 500 µg/ml of sonicated denatured salmon sperm DNA, 1 mM EDTA, 50 mM sodium phosphate, pH 6.7). The filter was then hybridized with \textsuperscript{32}P-labeled human uPAR cDNA probe in the same solution for 48 h. After hybridization, the filter was first washed in 500 ml of 2X SSC, 0.1% SDS at room temperature for 30 min and then washed in 500 ml of 0.1X SSC, 0.1% SDS at 56°C for 20 min. Autoradiographs were prepared by exposing the filters to Kodak XAR film at -70°C with intensifying screens.
2.2.3 Isolation of genomic clones

2.2.3.1 Plating bacteriophage λ

A human blood genomic library, constructed by cloning partial Sau3AI-digested human blood DNA (DNA fragments ranging over 9-23 kb) into the XhoI sites in the lambda GEM-11 vector (Promega), was used to isolate uPAR genomic clones. The bacteria cells KW251 were cultured at 37°C overnight by inoculating 50 ml of LB medium (1 liter containing 10 g bacto-tryptone, 5 g bacto-yeast extract and 10 g NaCl, pH 7.0) with a single bacterial colony. Cells were pelleted by centrifugation and then resuspended in approx. 20 ml of sterile 10 mM MgSO4 to an appropriate density. 2 X 10^4 bacteriophage particles were mixed with approx. 2 X 10^6 of freshly prepared KW251 bacteria cells and then incubated at 37°C for 20 min. After adding 10 ml of molten (50°C) top agarose, the infected bacteria mixture was immediately plated onto a 150-mm agar plate and incubated at 37°C for 14-18 h until the plaques reached a diameter of 1.5 mm. The plates were then chilled at 4°C for 1 h to allow the top agarose to harden.

2.2.3.2 Plaque transfer and hybridization

The formed plaques were transferred to reinforced nitrocellulose filters by carefully placing the filters on to the plate for 1 min. The orientation of filter was marked with three holes using an 18-gauge needle. The filters were denatured by placing them on the top of Whatman 3MM paper soaked with denaturing solution (see 2.2.2.4) for 5 min and neutralized on top of the 3MM paper soaked with neutralizing solution for 5 min, then rinsed in 2X SSC and air-dried on paper towels.

The plaque hybridization was performed with 32P-labeled uPAR cDNA probe as described in Southern blot hybridization (see above). After hybridization analysis, single positive clones were purified by three- or four-round screening. A total of 16 individual positive phage clones were obtained and purified after screening over 10^6 plaque-forming units.

To identify genomic clones containing the 5'-flanking region, all positive clones were re-hybridized with an 169 bp PCR product which covers 27 to 195 bp of uPAR cDNA (Roldan et al., 1990). Amplification of this fragment from the cDNA sequence was performed using primers derived from the uPAR cDNA (Exon 1+ and Exon 2- as forward and reverse primers, respectively, see Table 2.1). The PCR product was labeled with [γ-32P]ATP and T4 polynucleotide kinase.
2.2.3.3 Large-scale preparation of phage DNA

Positive plaques were purified and phage DNAs prepared by a liquid lysate procedure (Sambrook et al., 1989). The KW251 bacterial host was cultured in 10 ml LB medium at 37°C overnight. Then 500 ml of NZCYM medium (1 liter containing 10 g NZ amine, 5 g NaCl, 5 g bacto-yeast extract, 1 g casamino acids, 2 g MgSO\(_4\)\(\cdot\)H\(_2\)O, pH 7.0) was inoculated with 1 ml cell culture and incubated at 37°C with vigorous shaking until the OD\(_{600}\) of the cultures reached 0.5. Cells were infected with 10\(^{10}\) pfu of bacteriophage λ and were incubated 6 h until lysis occurred. Then, 10 ml of chloroform were added and the culture incubated for a further 10 min to complete cell lysis. The lysed culture was treated with pancreatic DNase I (1 µg/ml) and RNase (1 µg/ml) for 30 min at room temperature and then solid NaCl was added to a final concentration of 1 M. After removing debris by centrifugation at 10,000 rpm for 10 min at 4°C, solid polyethylene glycol (PEG 8000, Sigma) was added to a final concentration of 10% (w/v). The solution was kept on ice for 90 min to allow the bacteriophage particles to form a precipitate. After centrifugation, the pellet was resuspended in SM buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM MgSO\(_4\), 0.01% gelatin) and extracted twice with chloroform. Ultracentrifugation was performed at 36K rpm in SW55 rotor (Beckman) for 18 h. Following dialysis against Tris buffer (50 mM Tris-HCl, pH 8.0, 10 mM MgCl\(_2\), 10 mM NaCl), the purified bacteriophage was digested by adding proteinase K (100 µg/ml), 0.5% SDS and 20 mM EDTA at 56°C for 2h, and then extracted with phenol/chloroform and precipitated with ethanol.

2.2.4 Subcloning the uPAR fragments

2.2.4.1 Isolation of the uPAR gene fragment

The genomic inserts in prepared phage DNAs were characterized by digestion with SacI restriction enzyme, followed by Southern blot analysis using the uPAR cDNA as a probe. After complete digestion with the appropriate restriction enzyme, the digested DNA fragments were separated by 1% agarose electrophoresis with 1X TAE. The band of interest was then cut out using a razor blade and transferred into a dialysis bag with appropriate TE buffer. Electroelution was performed with 0.5X TAE at 50 volts for 2-3 h. DNA fragments were then purified by extraction with phenol/chloroform and precipitated with ethanol.
2.2.4.2 Dephosphorylation of enzyme-digested vectors

The vector used in subcloning was pBluescript SK- (Stratagene). For subcloning insert DNA with a single restriction endonuclease, the digested vector DNA was further treated with calf intestinal alkaline phosphatase (CIAP) before DNA ligation. Dephosphorylation of the vector DNA was carried out with 1 µl of CIAP (Boehringer, Germany) in a total volume of 100 µl containing 1X dephosphorylation buffer at 37°C for 1 h. The reaction was stopped by adding 2 µl of 0.5 M EDTA followed by phenol/chloroform extraction and ethanol precipitation.

2.2.4.3 DNA ligation

DNA ligation was performed by using a DNA ligation kit (Amersham) according to the manufacturer’s description. Generally, 10-50 ng of vector DNA was used and the ratio of vector to insert DNA was 1:1. The ligation reaction was carried out by adding 5 volumes of solution A and 1 volume of solution B relative to vector/insert mixture and incubating at 16°C for 4-16 h.

Of all overlapping genomic clones isolated, four were further characterized by restriction enzyme mapping with BamHI, HindIII and SacI. Seven contiguous restriction enzyme fragments including a 10 kb HindIII from lambda 11, an 8.0 kb SacI and a 3.9 kb HindIII from lambda 33, a 5.3 kb BamHI from lambda 36, as well as a 3.1 kb SacI, a 7.4 kb BamHI and a 2.4 kb HindIII from lambda 16 fragment were subcloned into pBluescript SK- for further analysis. Based on these constructs, further subcloning was carried out by using existing restriction sites.

2.2.5 Bacterial transformations

2.2.5.1 Preparation of competent cells

10 ml of LB medium was inoculated with a single bacterial colony (DH5α) and incubated at 37°C overnight with vigorous shaking. Then 2 ml of culture was transferred into 200 ml of LB medium in a 1 liter flask and continued to shake at 200 rpm at 37°C until the A₆₀₀ reached 0.45-0.55. The cells were chilled in ice water for 2 h and then collected by centrifugation for 5 min at 4°C. The cell pellet was resuspended in 200 ml of freshly prepared trituration buffer (100 mM CaCl₂, 70 mM MgCl₂, 40 mM sodium acetate, pH 5.5) and left on ice for 45 min. After centrifugation, the cells were gently resuspended in 20 ml of ice-cold trituration buffer and 80% pre-chilled glycerol was
added dropwise with gentle swirling to a final concentration of 15% (v/v). The competent cells were aliquoted in 0.2-1.0 ml quantities and stored at -70°C.

2.2.5.2 Transformation

100 µl of the competent cells were added to a pre-chilled 5 ml polystyrene tube (Falcon) containing 10 µl (5-20 ng) of the DNA to be transformed. The cells were cooled on ice for 30 min. Following heat shock at 42°C for 1 min and chilling on ice for 1 min, 1 ml of LB medium was added and the cells were incubated at 37°C with shaking at 225 rpm for 1 h. They were then pelleted by centrifugation for 2 min and resuspended in 50 µl of LB medium for plating on appropriate selective plates.

2.2.5.3 Selection of plasmids of interest

When using pBluescript vector, a blue/white selection method was used to screen recombinants (Sambrook et al., 1989). The positive clones were identified by minipreparation and restriction enzyme analysis.

2.2.6 Preparation of plasmid DNA

2.2.6.1 Minipreparation of plasmid DNA

Minipreparation of plasmid DNA was performed according to the alkali lysis method (Sambrook et al., 1989). A sterile 15 ml Falcon tube was used for a 1.5 ml miniculture of bacteria containing the plasmid of interest. The miniculture was grown up overnight in LB broth with ampicillin (100 µg/ml) in a 37°C shaker. Minicultures were transferred to 1.5 ml Eppendorf tubes for 1 min of centrifugation. The cell pellet was resuspended in 100 µl of Solution I (50 mM glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA) and then mixed with 200 µl of Solution II (1% SDS-0.2 N NaOH). After incubation on ice for 5 min, the solution was neutralized by adding 150 µl of Solution III (3.5 M potassium acetate, pH 5.2). Following centrifugation at top speed for 5 min, the supernatant was transferred into a new Eppendorf tube and plasmid DNA was precipitated with 2 volumes of ethanol. The pellet was washed with 75% ethanol, dried under vacuum, redissolved in 100 µl of TE buffer containing 100 µg/ml of bovine pancreatic RNase A (Sigma) and incubated at 37°C for 30 min.

For DNA sequencing, the plasmid DNA was further purified according to instructions with the Wizard™ Minipreps DNA Purification Kit (Promega). In brief,
plasmid DNA was mixed with 1 ml of resin, washed, eluted with 100 µl of pre-warmed TE buffer (50°C) and precipitated with ethanol.

2.2.6.2 Large scale preparation of plasmid DNA

Large scale preparation of the plasmid DNA was performed using a lysozyme-Triton method (Ausubel et al., 1987). One liter LB medium was inoculated with 5 ml of overnight LB culture and grown at 37°C to OD_{660} 0.4-0.6, then chloramphenicol (170 µg/ml) added. The culture was then shaken overnight (300 rpm) at 37°C. The bacteria were harvested by centrifuging for 10 min at 5K rpm, 4°C. The cells were washed with 20 ml PBS and resuspended in 12 ml cold 50 mM Tris-HCl, pH 8.0, 4 ml of 0.25 M EDTA and 10 mg lysozyme were added and the mixture placed on ice for 10 min. Then, an equal volume of Triton solution (50 mM Tris-HCl, pH 8.0, 50 mM EDTA, 0.1% Triton X-100) was added and the solution was gently mixed and placed on ice for 15 min. The viscous solution was transferred to tubes and centrifuged in the Beckman SS34 rotor for 60 min at 8K rpm. Then plasmid DNA in the supernatant was precipitated with 0.4 volumes of 10 M ammonium acetate and 2 volumes of isopropanol. The DNA pellet was dissolved in 7 ml of TE buffer and the suspension was centrifuged at 10 K rpm for 15 min. The volume of the supernatant was adjusted to 10 ml with TE buffer and 11g CsCl and 0.2 ml of ethidium bromide (10 mg/ml) were added. The solution was sealed in a Beckman Ti75 tube (Quick Seal Centrifuge Tube, 16 x 76 mm, Beckman) and centrifuged for 17 h at 60 K rpm. The plasmid band was carefully collected from the side of the tube with a syringe. Following extraction with butanol to remove ethidium bromide and dialysis against TE buffer (pH 7.5) overnight at 4°C with 3 changes of buffer, the isolated plasmid DNA was extracted twice with equal volumes of phenol and precipitated with ethanol.

2.2.7 Nested deletion of plasmid DNA

The deletion of specific regions of DNA insert was carried out by using the Double-stranded Nested Deletion Kit (Pharmacia) according to the manufacturer’s description. Double digestion at the poly linker site in pBluescript vector with appropriate enzymes generated a linearized double stranded DNA, which contained a blunt or 5'-overhanging nuclease-sensitive end adjacent to the target sequence and a nuclease-resistant end. The nested deletion was performed by adding 1 µl of exonuclease III (for 2 µg of DNA) in a
total volume of 20 µl at 30°C. 2 µl sample was removed from the reaction and mixed with 3 µl of S1 nuclease/buffer mix at each time point. Then S1 nuclease digestion was carried out at room temperature for 30 min and stopped by adding 1 µl of stop solution. After incubation at 65°C for 10 min, 3 µl of reaction mix was transferred to a new Eppendorf tube and nested deletions were re-ligated by adding 17 µl of ligation mix at room temperature for 2 h.

2.2.8 DNA sequencing

DNA sequencing was mainly performed by cycle sequencing using an Applied Biosystem Model 373A automatic DNA sequencer (Applied Biosystems) following the manufacturer’s instructions. Two sequencing methods were used for uPAR gene sequencing as follows.

2.2.8.1 T3 and T7 dye primer method

Both T3 and T7 dye primer kits (Applied Biosystems) were used for DNA sequencing. The sequencing reaction was carried out in a PCR machine in a total volume of 5 µl for A and C reactions (0.5 µg of template and 4 µl of reagent mix) and 10 µl for G and T reactions (1 µg of template and 8 µl of reagent mix) according to the manufacturer’s instructions. After heating at 95°C for 5 min, reaction in thermal cycler proceeded for 15 cycles of 95°C (30 sec), 55°C (30 sec) and 70°C (1 min), followed by 15 cycles of 95°C (30 sec), 55°C (1 sec) and 70°C (1 min). On completion of cycling, A, C, G, T samples were collected into an Eppendorf tube and precipitated with ethanol for sequencing.

2.2.8.2 Dye terminator method

For the dye terminator method, the sequencing reaction was performed in a total volume of 20 µl containing 1 µg of template, 10 pmol of appropriate primer and 9 µl of reagent mix (Dye terminator kit, Application Biosystems). After heating at 95°C for 5 min, reaction proceeded with 30 cycles of 95°C (30 sec), 50°C (15 sec) and 60°C (4 min). After reaction, the DNA sample was precipitated with ethanol and dried under vacuum for 2 min.

The particular oligonucleotide primers (Table 2.1), based on the uPAR cDNA sequence and uPAR sequence already obtained, were synthesized on an Applied
Biosystems 380B DNA synthesizer (Applied Biosystems Inc. USA) and provided by Biomolecular Resource Facility of The John Curtin School of Medical Research.

2.2.9 Characterization of exon-intron boundaries and sequence assembly, analysis and alignment

The intron/exon structure and intron sizes of the human uPAR gene were determined by restriction mapping and DNA sequence analysis. The sequence of the exon-intron boundaries was determined by nucleotide sequencing on both strands. Some primers were used to confirm the DNA sequence of the exon/intron junctions and the particular regions. The 5'-flanking sequence of the uPAR gene was determined by sequencing a 2.6 kb of Sphl-NruI DNA fragment isolated from the λ11 genomic clone on both strands by using nested deletion or particular primers. The sequences of the sense and antisense oligonucleotide primers used in uPAR DNA sequencing are shown in Table 2.1.

Analysis of DNA sequencing data and alignments of DNA sequences were performed using the MacVector program (version 4.0.1, dated 1992, International Biotechnologies, Inc.). The exon/intron junctions of the uPAR gene were identified with reference to previously published full-length uPAR cDNA sequence (Roldan et al., 1990) using the GT-AG rule (Mount, 1982).

2.2.10 Polymerase chain reactions (PCR)

PCR reactions were performed in a total volume of 100 µl containing 500 ng template DNA, 1.5 mM MgCl₂, 200 mM dNTPs (final concentration), 2.5 units of Taq polymerase, 10 pmol of each primer in 1X PCR buffer (Promega). Following denaturation at 94°C for 5 min, amplification proceeded in a DNA thermal cycler (PTC-100TM, MJ Research Inc. USA) for 30 cycles, with a denaturation step at 94°C for 1 min, an annealing step at 60°C for 1 min and an extension step at 72°C for 2 min. PCR products were purified by 1% agarose gel electrophoresis with 1X TAE followed by electroelution as described in section 2.2.4.1.
<table>
<thead>
<tr>
<th>Primer*</th>
<th>Sequence</th>
<th>Position**</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'-flanking region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-350(+)</td>
<td>5'—CCTGTTAGCCAAATATCAGCCACATTC-3'</td>
<td>-309</td>
</tr>
<tr>
<td>5'-350(-)</td>
<td>5'-GAATGTGCTGATATATCGCTTCAGG-3'</td>
<td>-309</td>
</tr>
<tr>
<td>5'-500(+)</td>
<td>5'-TACTCTGCTGCTGGGACCTGTCAGG-3'</td>
<td>-480</td>
</tr>
<tr>
<td>5'-650(-)</td>
<td>5'-GTGTGTTACTTAGAGGCTGAGGC-3'</td>
<td>-652</td>
</tr>
<tr>
<td>5'-950(+)</td>
<td>5'-TGCAGACATCACATACCTACACATGTCACCAC-3'</td>
<td>-916</td>
</tr>
<tr>
<td>5'-900(-)</td>
<td>5'-GGCTGTATCTTATCGGATTCCTG-3'</td>
<td>-845</td>
</tr>
<tr>
<td>5'-1300(-)</td>
<td>5'-TGGGAACCTGGAAGCTGACGTCAGG-3'</td>
<td>-1257</td>
</tr>
<tr>
<td>5'-1500(-)</td>
<td>5'-AACTGCAAGTTGTTGAAAAAGCTTATATAAG-3'</td>
<td>-1460</td>
</tr>
<tr>
<td>5'-1600(+)</td>
<td>5'-CCCAAGCCTGCAGTGCCGAGATTCAGGACC-3'</td>
<td>-1560</td>
</tr>
<tr>
<td>5'-1850(-)</td>
<td>5'-AATGCAAGGTGTCAGTGCTGCTGAGTCATG-3'</td>
<td>-1777</td>
</tr>
<tr>
<td>5'-1960(-)</td>
<td>5'-CCCCGTCATTTAGATGCTACACAG-3'</td>
<td>-1853</td>
</tr>
<tr>
<td>5'-2165(+)</td>
<td>5'-AATGCAAGTTTCTTATCCCTCTTGAGA-3'</td>
<td>-2085</td>
</tr>
<tr>
<td>5'-2300(-)</td>
<td>5'-TTGAGGTGGAAGCTGATGAG-3'</td>
<td>-2294</td>
</tr>
<tr>
<td>Exon</td>
<td>Exl(+), Exl(-), Ex2(+), Ex2(-), Ex3(+), Ex3(-), Ex4(-), Ex5(+), Ex6(-)</td>
<td></td>
</tr>
<tr>
<td>5'-ACAGGAGCTGCCCTCGCGACAT-3'</td>
<td>+27</td>
<td></td>
</tr>
<tr>
<td>5'-TGACCCATGTGCGAGGCGACCTCCTGT-3'</td>
<td>+27</td>
<td></td>
</tr>
<tr>
<td>5'-TGGGACCGCAACTGAGTGAGCAGCGACG-3'</td>
<td>+75</td>
<td></td>
</tr>
<tr>
<td>5'-CTGCGGATGTCACTGTAAGACC-3'</td>
<td>+2210</td>
<td></td>
</tr>
<tr>
<td>5'-ATCGTGTCGCTGACAGGAGCTGACG-3'</td>
<td>+2210</td>
<td></td>
</tr>
<tr>
<td>5'-TGCCGATGACAGCCAGCTTAC-3'</td>
<td>+4940</td>
<td></td>
</tr>
<tr>
<td>5'-GTAAGCTGAGCTGACAGGACG-3'</td>
<td>+4940</td>
<td></td>
</tr>
<tr>
<td>5'-GCTCATGTCTGATGAGCCACA-3'</td>
<td>+13476</td>
<td></td>
</tr>
<tr>
<td>5'-GCCAAGAGATGACGCCAAGC-3'</td>
<td>+14499</td>
<td></td>
</tr>
<tr>
<td>5'-GCTATAGCTGCGAGCTG-3'</td>
<td>+22873</td>
<td></td>
</tr>
<tr>
<td>Intron</td>
<td>Int1(-), Int2A(-), Int2B(-), Int3(-), Int4(-), Int6(+), Int6(-)</td>
<td></td>
</tr>
<tr>
<td>5'-TCATTAGAACCTGTTCAGCGAC-3'</td>
<td>+125</td>
<td></td>
</tr>
<tr>
<td>5'-TTTACTGCGTTCCAACTCTGCCG-3'</td>
<td>+2284</td>
<td></td>
</tr>
<tr>
<td>5'-TCAGGCTCAATCCACACTG-3'</td>
<td>+4414</td>
<td></td>
</tr>
<tr>
<td>5'-GATGATAGTATCTGAGG-3'</td>
<td>+5052</td>
<td></td>
</tr>
<tr>
<td>5'-TACCCATTTGCCAAACTGACTG-3'</td>
<td>+11155</td>
<td></td>
</tr>
<tr>
<td>5'-CAGGTCAGCTGAGGATGAGCAG-3'</td>
<td>+18430</td>
<td></td>
</tr>
<tr>
<td>5'-CAAGTAGCTGCGACTGAGGC-3'</td>
<td>+21654</td>
<td></td>
</tr>
</tbody>
</table>

* (+) sense; (-) antisense.
** relative to transcription start site.
2.2.11 RNA preparation

Total RNA from different cell lines was prepared according to the guanidinium/cesium chloride method (Sambrook et al., 1989). All water used in RNA work was pre-treated with diethyl pyrocarbonate (DEPC) as described by Sambrook et al. (1989). After washing with PBS, cultured cells (5-10 X 10^7) were lysed by adding 3 ml of guanidinium solution (4 M guanidinium isothiocyanate, 0.1 M Tris-HCl, pH 7.5, 1% β-mercaptoethanol). The viscous solution was transferred into a 15 ml Falcon tube, drawn up and down ten times through a 23-G needle to reduce viscosity and then transferred on to a CsCl cushion (1.7 ml of 5.7 M CsCl, 1 mM EDTA, pH 8.0) in a polyallomer centrifuge tube (13 x 51 mm, Beckman). Ultracentrifugation was performed using a SW55 rotor (Beckman) at 36K rpm for 20 h at 20°C. After carefully removing the supernatant using a Pasteur pipette and cutting off the bottom of the tube, the RNA pellet was resuspended in 750 µl of TES buffer (10 mM Tris-HCl, pH 7.4, 5 mM EDTA, 1% SDS) and extracted twice with phenol/chloroform/isoamyl alcohol (25:24:1). After precipitation with ethanol, the extraction procedure was repeated. Prepared RNA samples were precipitated with ethanol and stored at -70°C.

2.2.12 Primer extension analysis

Primer extension analysis was performed according to the standard method (Sambrook et al., 1989). A 28 mer oligonucleotide primer (Exon 1A-, from +27 to +54, Table 2.1) complementary to a region within exon 1 of the human uPAR gene was labeled at 5'-end with [γ-^32P]ATP and T4 polynucleotide kinase. The ^32P-labeled primer was purified by 5% polyacrylamide gel electrophoresis using 1X TBE. The labeled primer (5 X 10^5 cpm) was mixed with 50 µg of total RNA from human cell lines U937, HCT116, LIM1215 or with yeast tRNA (Sigma) and precipitated with ethanol, respectively. The pellet was dissolved in 10 µl of hybridization buffer (40 mM PIPES, pH 6.4, 1 mM EDTA, pH 8.0, 400 mM NaCl) and the mixture heated at 80°C for 2 min and maintained at 45°C for 3 h. After hybridization, the primer extension reaction was started by adding 20 µl of 5X reverse transcriptase buffer, 30 units of RNasin, 5 µl of 10 mM dNTP, 10 µl of 0.1 M DTT, and 100 units of AMV-RTase (Promega) and incubated at 42°C for 30 min. Elongated products were extracted with
phenol/chloroform and ethanol-precipitated, then fractionated through a 6% polyacrylamide gel containing 8 M urea, and detected by autoradiography at -70°C.

2.3 Results
2.3.1 Isolation and characterization of the human uPAR gene
2.3.1.1 Southern blot analysis

To characterize the genomic DNA fragments that encode the uPAR gene, high molecular-weight DNA from the human histiocytic cell line U937 and the colon cancer cell lines HCT116 and LIM1215 was analyzed by Southern blotting using the 1.14 kb human uPAR cDNA as a probe. This probe included the whole coding region of the uPAR cDNA (Roldan et al., 1990). Figure 2.1 shows that the human genomic DNA digested with the restriction enzymes SacI, EcoRI and HindIII generated multiple bands. There are five SacI bands (8.0, 6.2, 4.1, 3.1 and 0.8 kb), three EcoRI bands (11, 4.5 and 3.0 kb), and three HindIII bands (10, 3.9 and 2.4 kb). The hybridizing bands detected from genomic DNA of the three different human cell lines were identical to the fragments shown in the uPAR gene sequence.

2.3.1.2 Isolation of uPAR genomic clones

The human uPAR gene was isolated from a human blood genomic library in λGEM11 using a 1.14 kb human uPAR cDNA as a probe. A total of sixteen independent hybridization-positive clones from approximately 10^6 phage plaques were isolated and purified for further analysis. To confirm whether the cloned DNA fragments represented the whole human uPAR gene, genomic DNAs from these clones were digested with SacI and analyzed by Southern blotting using uPAR cDNA as a probe. As shown in Figure 2.2, all positive recombinant phages contain the uPAR gene fragments and revealed the same hybridizing bands detected with the genomic DNA from human histiocytic lymphoma and colon cancer cell lines. Five SacI bands (8.0, 6.2, 4.1, 3.1 and 0.8 kb) were shown in the phage clones, λ16 and λ33.
Figure 2.1 Southern blot analysis of various genomic DNA. Genomic DNAs from U937 (lanes 1, 2 and 3), LIM1215 (lanes 4, 5 and 6) and HCT116 (lane 7) cells were digested with restriction enzymes, HindIII (lanes 1, 4 and 7), EcoRI (lanes 2 and 5) and Sacl (lanes 3 and 6) and hybridized with the 1.14 kb human uPAR cDNA probe.
Figure 2.2  Southern blot hybridization of the uPAR genomic phage clones. Purified lambda DNA from uPAR genomic clones was digested with SacI and hybridized with a 1.14 kb human uPAR cDNA probe. Five SacI bands (8.0, 6.2, 4.1, 3.1 and 0.8 kb) were detected in two lambda genomic clones, λ16 (6.2, 3.1 and 0.8 kb) and λ33 (8.0 and 4.1 kb), respectively.
2.3.1.3 Subcloning and sequencing strategy

Of sixteen λ uPAR clones, four overlapping clones (λ11, λ33, λ36 and λ16) that covered the whole known uPAR cDNA were selected for further characterization. The relative positions of HindIII, BamHI and SacI sites were determined by digesting the λ genomic clones carrying the uPAR gene using combinations of restriction enzymes, DNA sequencing and Southern blot analysis. A restriction map of the human uPAR gene was generated as shown in Figure 2.3. uPAR genomic fragments, the 10 kb HindIII-HindIII fragments from λ11, the 8.0 kb SacI-SacI (S8.0) and the 3.9 kb HindIII-HindIII (H3.9) fragments from λ33, the 5.3 kb BamHI-BamHI (B5.3) from λ36 and the 3.1 kb SacI-SacI (S3.1), as well as the 7.4 kb BamHI-BamHI (B7.4) and the 2.4 kb HindIII-HindIII (H2.4) fragments from λ16, were isolated and subcloned into pBluescript SK- phagemid for sequencing. All subclones were sequenced and a 23 kb of continuous genomic DNA sequence including 2.6 kb at its 5’- and 250 bp at its 3’-end was determined according to the strategy outlined in Figure 2.3. Most of intron sequences were determined by single strand and the exon-intron borders as well as 5’- and 3’-flanking region were sequenced in both strands. There are at least three overlaps to show the consensus sequence in the introns.

2.3.2 Structure of the human uPAR gene

2.3.2.1 Exon/intron organization

The overall structure of the human uPAR gene was determined by Southern hybridization, restriction enzyme analysis and DNA sequencing. Alignment of genomic sequences with the full-length uPAR cDNA revealed that the gene is composed of seven exons separated by six introns and occupies about 23 kb. As shown in Figure 2.3, the uPAR gene was completely covered by three lambda uPAR clones (λ33, λ36 and λ16) and six overlapping subfragments (S8.0, H3.9, B5.3, S3.1, B7.4 and H2.4). No differences were found between the genomic exon sequences and the previously published sequence of uPAR cDNA (Roldan et al., 1990).
Figure 2.3 Sequence strategy and restriction map for the human uPAR gene. uPAR genomic DNA is shown with exons indicated by black boxes and numbered 1-7. Untranslated regions are shown by open boxes, and introns by horizontal lines. The uPAR gene was completely covered by four overlapping λGEM clones, λ 11, 33, 36 and 16. Fragment H10 was from λ11, fragments S8.0 and H3.9 were from λ33, fragment B5.3 was from λ36, and fragments S3.1, B7.4 and H2.4 were from λ16. Subfragments from λ clones were cloned into pBluescript SK- phagemid for sequencing. Horizontal arrows in or beneath DNA fragments indicate the direction of individual sequencing experiments. Restriction sites are shown on the map. B, BamHI; H, HindIII; S, SacI.

* λ16 genomic clone was isolated and partially sequenced by Dr. Yao Wang.
** PMN3500 and PMN4500 were isolated and sequenced by Dr. Lorin K. Johnson.
### Table 2.2  Intron/exon boundaries of the human uPAR gene

<table>
<thead>
<tr>
<th>Exon no.</th>
<th>Exon size (bp)</th>
<th>5'-splice donor</th>
<th>3'-splice acceptor</th>
<th>Intron size (bp)</th>
<th>Codon phase</th>
<th>Amino acid interrupted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>101+5'UTR</td>
<td>CCA G gtaggg.....ccccag CC TCT</td>
<td>2040</td>
<td>I</td>
<td>Ala (-4)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>111</td>
<td>GAA G gtgagc.....ttgcat AA GGA</td>
<td>2620</td>
<td>I</td>
<td>Glu (34)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>144</td>
<td>TCT G gtgagt.....aaacag GC CGG</td>
<td>8420</td>
<td>I</td>
<td>Gly (82)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>162</td>
<td>GAA G gtgagc.....ccatag GG CGT</td>
<td>906</td>
<td>I</td>
<td>Gly (136)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>135</td>
<td>CCA A gtaagg.....ccataag TC CTG</td>
<td>3100</td>
<td>I</td>
<td>Ile (181)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>147</td>
<td>CAC G gtaggg.....ccatag AA CCG</td>
<td>2780</td>
<td>I</td>
<td>Glu (230)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>563+3'UTR</td>
<td></td>
<td></td>
<td></td>
<td>Stop (313)</td>
<td></td>
</tr>
</tbody>
</table>

Note. Exon sequences are given in capital letters, and intron sequences in lowercase letters. Introns that do not split codon triplets are indicated by phase 0, interruption after the first nucleotide by phase I, and interruption after the second nucleotide by phase II. Amino acids encoded at the splice sites are indicated with numbers relative to the first amino acid of the mature protein.
All exons have been mapped on the gene and the structures of the exon/intron junctions were determined from the cDNA and genomic sequences. The lengths of the exons and introns, and the sequences at the splice sites, are shown in Table 2.2. The seven exons of the gene are 101, 111, 144, 162, 135, 147 and 563 bp long and encode 19, 37, 48, 45, 49 and 83 amino acid residues, respectively. The six introns of the uPAR gene are approximately 2.04, 2.62, 8.42, 0.906, 3.10 and 2.78 kb long, respectively. All introns begin with GT and end in AG which is consistent with consensus splice site sequences (Breathnach and Chambon, 1981; Mount, 1982). The resulting intron/exon junctions disrupt amino acid codons at all six splice donor/acceptor boundaries and all occur between the first and second bases of the codon triplets.

2.3.2.2 Domain structure and intron/exon boundaries

The uPAR mRNA encodes a protein containing 313 amino acids plus a 22 amino acid signal peptide (Roldan et al., 1990). Amino acid sequence analysis revealed that there are three cysteine-rich repeats containing approximately 90 amino acids each in this polypeptide (Ploug et al., 1991). Comparison of the uPAR exons with the three protein domains revealed that all of the introns interrupt the protein coding regions of the gene and the structural domains of this protein are almost perfectly divided among the 7 exons. While exon 1 provides the 5'UTR of the human uPAR mRNA and most of the signal peptide, exons (2+3), (4+5) and (6+7) encode protein domain I, II and III, respectively. Each domain is composed of a pair of exons and they are all interrupted by introns on either side in codon position 1 in the reading frame. Moreover, the hydrophobic sequence necessary for GPI-anchor attachment is encoded in exon 7. These results suggest a domain-specific manner in the exon/intron structure of the human uPAR gene.

2.3.3 Characterization of the 5'-flanking region

Two uPAR genomic clones, λ11 and λ33, were identified containing the 5' end of the gene by hybridization with the 169 bp of uPAR cDNA fragment (27 to 195) as a probe. To confirm that the cloned DNA fragments represented the 5'-flanking region, human genomic DNA from U937 cells was digested with PstI, EcoRI, XbaI and NruI and analyzed by Southern blotting using a end-labeled 184 bp SmaI-NruI DNA probe, which contains 137 bp of 5'-flanking region sequence and 47 bp of exon 1 and was isolated from 5' end of the uPAR gene (λ11). As shown in Figure 2.4B, Accl, Sacl or
PstI cleavage generated a single fragment of approximately 13 kb, 8.0 kb or 2.0 kb, respectively, while AccI plus NruI and SacI plus NruI double cleavage generated a single fragment of approximately 7.0 kb and 3.5 kb, respectively. Since there is a single NruI restriction site present in exon 1, this results showed a simple restriction map of the 5'-flanking region. The hybridizing bands detected with genomic DNA were identical to the restriction fragments shown in the uPAR genomic clone λ11, which indicates that the cloned DNA fragments represent the 5'-flanking region of the gene. Based on restriction analysis, DNA sequencing and Southern blot analysis, a restriction map of 5'-flanking region of the human uPAR gene is shown in Figure 2.4A.

2.3.4 Identification of the transcription initiation sites

Primer extension was used to identify the transcription initiation sites of the human uPAR gene. A synthetic oligonucleotide 28 mer DNA reverse complementary to nucleotides 27-54 (exon 1A-, Table 2.1) of the human uPAR cDNA (Roldan et al., 1990) was labeled with T4 polymerase kinase and [γ-32P]ATP, and hybridized with total RNA isolated from monocyte-like U937 cells, and the colon cancer cells HCT116 and LIM1215. The extended DNA products are shown in Figure 2.5. The results demonstrated that multiple alternative initiation sites clustered in the region around 51 bp upstream from the translation start (ATG) codon of the uPAR gene. In addition, another minor band was found at 77 bp suggesting that uPAR transcription starts at different nucleotides. An identical pattern of extended products was observed in these different cell types suggesting a similar transcription pattern among them. Because the intensity of 51 bp products is strongest, this site was designated as transcription initiation site (+1) which is one base pair different from recent published results in which the nucleotide at 52 bp relative to ATG start codon was designated as the transcription start site (Soravia et al., 1995). No primer-extended products were detected for negative control RNA (Yeast tRNA). The position of the major transcription start site is close to published +1 of cDNA (Roldan et al., 1990).
Figure 2.4 Southern blot and restriction map of 5'-flanking region of the human UPAR gene. A, a restriction map of 5'-flanking region of the human UPAR gene. The restriction sites for HindIII, Accl, SacI, PstI, XbaI, NruI and EcoRI are shown as H, A, S, P, X, N and R. B, Southern blot analysis of the 5'-flanking region. Genomic DNA from U937 cells was digested with PstI (lane 1), SacI (lane 2), SacI plus NruI (lane 3), Accl (lane 4) and Accl plus NruI (lane 5), respectively and hybridized with a 184 bp Smal-NruI (SN) probe.
Figure 2.5 Determination of the human uPAR gene transcription start sites by primer extension. $^{32}$P-Labeled primer Exon 1A- was hybridized to 50 μg of total RNA from U937 (lane 1), yeast tRNA (lane 2), LIM1215 (lane 3) or HCT116 (lane 4) and extended with reverse transcriptase. Lanes G, A, C, and T are reference dideoxy-sequencing reactions. Extended fragments at 51 and 77 bp relative to ATG start codon are indicated by arrows.
2.3.5 Sequence analysis of the uPAR 5'-flanking region

2.3.5.1 Consensus sequences for DNA binding proteins

To map the regulatory sequences in the 5'-flanking region of the uPAR gene precisely, a 2.6 kb of PstI-NruI fragment isolated from the uPAR genomic clone λ11 was subcloned into pBluescript SK- phagemid and sequenced. Figure 2.6 shows the nucleotide sequence of the 5'-flanking region of the human uPAR gene. Neither TATA-box nor CAAT box is found at the proximal region of the uPAR gene. But a search for consensus binding sites of transcriptional factors (Faisst and Meyer, 1992) indicated a number of putative sites for transcription factors that have been identified (Wang et al., 1995). There are two AP-1 sites located at positions -184 to -178 (TGAGTCA) and -70 to -64 (TGACTCG); one AP-2 binding site at position -170 to -163 (GCCAGCCC); two NF-κB binding sites at positions -606 to -598 (GGGTTTCAC) and -41 to -34 (AGGAGTCCC); two GATA binding sites at positions -2070 to -2065 (AGATAA) and -400 to -395 (AGATAA); two CFI consensus binding sites at positions -1703 to -1698 (ACATGG) and -549 to -544 (AAATGG); and three PEA3 binding sites at positions -2563 to -2558 (AGGAAG), -465 to -460 (AGGAAG) and -249 to -244 (AGGAAG) in the uPAR 5' upstream region. No consensus Sp1 sequence (GGGCGG) was found in the potential promoter region. However, three GGGAGG binding sites were found in the proximal region at positions -103 to -98, -94 to -89 and -45 to -40, respectively. The GGGAGG sequence is a variant of the canonical site and has been reported to have lower binding affinity for Sp1 (Letovsky and Dynan, 1989). An IL6 responsive element (IL6-RE, Kishimoto et al., 1994), was found at position -47 to -42 (CTGGGA), which overlaps with a Sp1 site and a NF-κB site. Moreover, a downstream promoter element (DPE), which has been found downstream of the promoter in many Drosophila genes (Burke and Kadonaga, 1996), is located at +11 to +17 (AGACGTG). In addition, four Alu sequences were also found within the 2.6 kb region in both orientations (Kariya et al., 1987).

2.3.5.2 G+C-rich sequence and a dyad symmetry structure

Sequence analysis revealed that the 5'-flanking sequence (between nucleotides -141 and +52) of the uPAR gene exhibits a high G+C content (65%) and characteristics of CpG island (Bird, 1986; Gardiner-Garden and Frommer, 1987). Examination of the sequence composition of the region around multiple transcription initiation sites also revealed an imperfect dyad symmetry structure (Figure 2.7). This stem and loop
structure starts at -32, ends at +27 and spans 60 bp. Multiple transcription start sites are located within this region. A comparison of the stem and loop structure of the uPAR gene with that of the human adenosine deaminase (ADA) gene (Valerio et al., 1985) and the human β-amyloid precursor protein gene (Donnelly et al., 1990) is shown in Figure 2.7. It has been shown that multiple transcription start sites, a dyad symmetry structure. TATA-less and G+C-rich sequence in the proximal region of the gene are common in TATA-less-G+C-rich promoters (Ackerman et al., 1993).

2.3.6 Comparison of human and murine uPAR genomic structure

Most recently, the murine uPAR gene has been isolated and characterized from a 129/SvJ murine genomic DNA library (Suh et al., 1994). A comparison of the structure of the human uPAR gene with that of the murine gene is shown in Figure 2.8. There is a high degree of structural similarity between the two species. Like human uPAR gene, the mouse gene also consists of seven exons and six introns. The size of the two genes, however, is quite different. While the human uPAR gene is 21.67 kb long, the murine gene is only 13.2 kb. Exons 1, 4 and 7 are larger in the human gene, whereas exons 2, 3, 5 and 6 are of identical size in both genes. Beyond that, the uPAR gene in both species has same exon/intron junctions. All 6 introns interrupt at the coding region and the splice junctions uniformly obey the GT-AG rule. It was found that the domain specific pattern is also perfectly conserved between the two species (Figure 2.8). In contrast to these findings, the exon/intron boundaries of both genes share low homologies. Finally, several different repetitive sequences were also observed in the introns and flanking regions of the mouse uPAR gene such as poly-pyrimidine. As in the human uPAR gene, intron 3 of the mouse gene contains a poly(TG), but only one Alu repeat was found. These data indicated that the organization of the uPAR gene is well conserved between human and murine species.
Figure 2.6 Nucleotide sequence of the human uPAR 5'-flanking region. The 2648 bp sequence of uPAR spans -2587 to +61 of the uPAR gene. The arrow indicates the major transcriptional start site (+1). The ATG initiator is in bold. The consensus sequences for transcription factors are underlined. Four Alu sequences are marked within the boxes.
Figure 2.7 Comparison of the potential stem-and-loop structure deduced from the proximal G+C-rich region (-32 to +27) of the human uPAR with that of human adenosine deaminase (ADA) and human β-amyloid precursor protein gene. A, human uPAR; B, human ADA (Valerio et al., 1985); C, human β-amyloid precursor protein (Donnelly et al., 1990). The major transcription sites within the structures are indicated by arrowheads.
Figure 2.8  Comparison of the human and mouse uPAR genes. Exons are indicated by boxes and numbered 1-7. Untranslated regions of the exons are shown as open areas. The sizes of each intron and exon of the uPAR gene are indicated below the boxes. The signal peptide encoded by exons and the putative protein domains encoded by pairs of exons are shown.
2.4 Discussion

The data presented in this chapter describe isolation, sequencing and characterization of the human genomic uPAR including its 5'- and 3'-flanking regions. The structure of the human uPAR gene was identified by analysis of four genomic clones isolated from a human blood genomic library. Characterization of the human uPAR genomic organization has revealed that this gene spans over 23 kb and consists of 7 exons separated by 6 introns. Exon sizes vary from 101 bp (exon 1) to 563 bp (exon 7) and introns from 903 bp (intron 4) to over 8 kb (intron 3). All boundary sequences of the exon/intron junctions conform to the GT-AG rule and all of the introns interrupt amino acid codons at these splice boundaries in the protein coding region.

The three domain model for uPAR protein structure was originally proposed according to a conserved pattern of cysteine residues of deduced amino acid sequence (Roldan et al., 1990; Behrendt et al., 1991). The data presented here provide further support for this view and suggest that the human uPAR gene has a domain-specific organization of all exons, since an alignment of the amino-acid sequences of the receptor protein with the exon/intron architecture strongly suggests a correlation between the morphological protein domain structure and genomic DNA organization (Wang et al., 1995). While the first exon contains the 5' untranslated region and most of the signal peptide, three domains are encoded by pairs of exons 2+3, 4+5 and 6+7, respectively. In addition, analysis of the exon/intron boundaries of the gene demonstrates that three cysteine-rich repeats of uPAR are encoded by symmetrical exon sets flanked by intron/exon boundaries of the same phase (type-1) suggesting that uPAR appears to have evolved by a combination of exon duplications and shuffling of individual domains (Patthy, 1987; Girard et al., 1991). The correlation of exon/intron organization and the protein domains has also been reported in many eukaryotic genes, such as human aggrecan (Valhmu et al., 1995), human lipoprotein-associated coagulation inhibitor gene (Girard et al., 1991) and human nidogene (Zimmermann et al., 1995). Like the human uPAR gene, these genes show domain-specific genomic organization.

Recently, Casey et al. (1995) have also isolated and characterized the human uPAR gene by screening a human placental genomic DNA library. The uPAR gene from the human placenta has the same genomic organization containing 7 exons separated by 6 introns. Interestingly, they have isolated a novel 1.7 kb uPAR cDNA from the HeLa cDNA library, in which exon 1 was 481 bp in length containing a unreported 380 bp
segment and exon 5 was absent, in comparison with uPAR cDNA (Roldan et al., 1990) and our data. It is unknown whether this long 5' untranslated segment of the uPAR gene is involved in tissue specific expression and uPAR function is affected by the absence of exon 5. Therefore, it will be very interesting to determine the biological significance of the uPAR mRNA variant.

The human uPAR gene contains several repetitive sequences, including 18 human Alu sequences, are located in the six introns and in the 5’ and the 3’ regions of the gene (Wang et al., 1995). Among them, four Alu repeat sequences were found within the 5’-flanking region (Figure 2.6). Although Alu sequences are widely dispersed (Kariya et al., 1987), the presence of multiple copies of the Alu sequence in asymmetric orientation may be responsible for most of the secondary structure of the uPAR gene. Multiple copies of the Alu sequence have also been found in other human genes, including CD59 (Tone et al., 1992), plasminogen activator inhibitor-1 (Bosma et al., 1988), hypoxanthine phospho-ribosyltransferase (Edwards et al., 1990) and retinoblastoma susceptibility protein (Toguchida et al., 1993). In the human CD59 gene, nine Alu repetitive sequences were distributed in each intron and 5’-flanking region in both orientations (Tone et al., 1992). Recently, several reports have indicated that the Alu sequence is involved in regulation of gene transcription in a positive or negative manner (Wu et al., 1990; Norris et al., 1995; Hewitt et al., 1995; Hanke et al., 1995).

Another notable sequence feature is a 23 TG tandem repeat which occurs in the third intron of the uPAR gene (Wang et al., 1995). TG tandem repeat sequences are abundant (5 x 10^5 copies/haploid human genome) and interspersed in eukaryotic genomes. It has been reported that the length polymorphism of poly(TG) may be a useful source of genetic markers for genome mapping (Gaillard and Strauss, 1994). It is noted that this poly(TG) found in the intron 3 is conserved in the human and murine uPAR gene. Recently, analysis of this poly(TG) repeat in colorectal cancer cell lines in our lab has revealed three different alleles that were not detected in a series of healthy control individuals, suggesting that this highly polymorphic microsatellite marker may be useful in linkage analysis of the role of uPAR in cancer (Kohonen-Corish et al., 1996). The poly(TG) tandem repeat has also been reported to have enhancer-like activity (Hamada et al., 1984), suggesting that the third intron may be involved in transcriptional regulation of the uPAR gene. Furthermore, sequences of alternating purines and pyrimidines, such as poly(dT,dA), which also occurs in the uPAR gene, have potential
to form a left-handed conformation (Z-DNA) under certain conditions (Arnott et al., 1980; Haniford and Pulleyblank, 1983). Whether or not these repeats have any functional significance is not yet clear.

The human uPAR gene has multiple transcription start sites in a G+C-rich region suggesting that transcription of the uPAR gene is initiated at different nucleotides. Of particular interest, multiple transcription initiation sites around a major site (at -51) are in a cluster, while a weaker start site is located at -77 relative to ATG codon, which is similar to the results obtained in HeLa cells (Soravia et al., 1995). Sequence analysis of the immediate 5'-upstream sequence reveals that the 5'-flanking region of the human uPAR gene lacks consensus TATA and CAAT boxes. Combined with a moderately high GC content, these results suggest that the potential uPAR promoter belongs to TATA-less G+C-rich promoters since its structure is highly similar to those presenting TATA-box-deficient G+C-rich promoter sequences with a sequence dyad symmetry around their major transcriptional start site (Ackerman et al., 1993). The absence of a TATA box together with high G+C content have also been found to be features of promoters of housekeeping genes such as hypoxanthine phosphoribosyltransferase (Patel et al., 1986) and adenosine deaminase (Valerio et al., 1985). Studies in mammalian molecular genetics have indicated that the G+C-rich promoters are very complex and prevalent. The use of multiple transcription start sites is common among these promoters (Ackerman et al., 1993). Therefore, existence of a TATA-less and G+C-rich promoter in the uPAR gene may represent a complicated mechanism of regulation of gene expression. Characterization of the uPAR gene promoter will provide further insight into this kind of promoter.

Many sequences governing transcriptional control and the nuclear factors with which they interact have been extensively characterized (Faisst and Meyer, 1992). Therefore, the homologies of such regulatory sequences in the proximal 2.6 kb region of the human uPAR gene were searched for since most cis regulatory elements are located in the 5'-flanking region of a gene, near the transcription start site (Mitchell and Tjian, 1989). On the basis of sequence similarity, several potential binding sites for transcription factors, such as AP-1, AP-2, GATA, NF-κB, PEA3 and Sp1 are shown in Figure 2.6.

Most recently, a 1.6 kb of the 5'-flanking sequence of the human uPAR gene has been reported by Soravia et al. (1995). Consistent with the data presented here, their
results showed that the uPAR gene has the same multiple transcription start sites located in the GC-rich region. A few differences, however, were noted between the uPAR sequences. For example, there is no Sp1 site at position -1242 and 24 T other than 26 T is located at position -801 suggesting a DNA polymorphism of the 5'-flanking region of the gene may exist among different cell types. In addition to these DNA binding sites for transcription factors, they also reported an Inr element located at position -12 to +1 (Soravia et al., 1995). In contrast to the 1.6 kb sequence, a 2.6 kb 5'-flanking sequence of the uPAR gene was sequenced and analyzed as shown in this chapter. Our expanded analysis of the 5'-flanking region of the human uPAR gene should facilitate future studies of its structure and function.

Murine, bovine and rat uPAR cDNAs have been reported to have similar characteristics to human uPAR cDNA (Krätzschmar et al., 1993; Ploug and Ellis, 1994). Recently, the murine uPAR gene has been isolated and characterized (Suh et al., 1995). A comparison of the human uPAR gene, as documented in this chapter, to the sequenced mouse uPAR gene reveals a high degree of conservation of sequence and structure organization between the two species (Figure 2.8). The two genes share the following characteristics: (1) both genes consist of 7 exons and 6 introns, in spite of their different size of introns and exons; (2) locations of splice junctions in the two genes are highly conserved with the introns occurring between the first and second nucleotides of codons; (3) both genes have a domain-specific exon structure in which three repeats are encoded by pairs of exons and the signal peptide and GPI-anchor attachment domain are encoded by exon 1 and exon 7, respectively; (4) the 5'-flanking regions of both genes show a high degree of sequence homology (Soravia et al., 1995). Comparison of the uPAR gene between the two species indicate that while the conservation of domain-specific exon/intron boundaries in structure suggests evolutionary conservation of uPAR protein function, high sequence homology in the promoter region suggests evolutionary conservation of transcriptional regulation.

In addition, the data present in this chapter provide further support for the conclusion, drawn on the basis of amino acid sequence, that uPAR is a member of the Ly-6 superfamily (Wang et al., 1995). It has been proposed that three cysteine-rich domains of uPAR arose as a result of internal triplication of an ancestral domain. The domains of Ly-6 family proteins are related to the internal repeats of uPAR (Behrendt et al., 1991). Further comparison of the uPAR gene structure with that of the CD59 and
Ly-6 family in terms of the organization of the exons, therefore, strongly suggests that human uPAR might have evolved from a common ancestral precursor resembling Ly-6 proteins. The conserved cysteine-rich domain encoded by pairs of exons in both uPAR and Ly-6 family genes may be important for the functions of these proteins.

The elucidation of the human uPAR exon structure addressed in this chapter provides insight into its domain-specific genomic organization and its relationship with the Ly-6 family and facilitates studies of genetic variations in the uPAR gene. The data described here will also provide important information to conduct gene targeting to investigate the relationship of structure and function in the human uPAR gene. Detailed analysis of 5'-flanking region will provide a basis toward our understanding of the transcriptional regulation of the uPAR gene.
CHAPTER 3
CHARACTERIZATION OF PROMOTER AND SILENCER
IN THE HUMAN uPAR GENE

3.1 Introduction

Eukaryotic transcription initiation for genes transcribed by RNA polymerase II (Pol II) is a primary control point in the regulation of gene expression. Transcription of each gene is controlled by a particular combination of positive and negative regulatory cis-acting elements (for review see Mitchell and Tjian, 1989; Roeder, 1991; Tjian and Maniatis, 1994). In most instances, such cis-acting elements are located in the 5'-flanking region of a gene. The control region in the immediate vicinity of a transcription start site is called the promoter. Regions that regulate a promoter from a distance in an orientation-independent fashion are called enhancers or silencers (Mitchell and Tjian, 1989; Johnson and McKnight, 1989). The interaction of the promoter, enhancer or silencer with specific regulatory proteins is necessary for accurate and regulated differential gene expression.

In eukaryotes, the promoter activation is complex due to the large number and different types of promoters. Promoters for genes transcribed by RNA polymerase II commonly contain two classes of DNA sequence elements, the TATA motif and the initiator (Inr) (Weis and Reinberg, 1992). Specific genes may contain either or both of these elements. The TATA box is the most common core promoter element located 25-30 bp upstream of transcription start site and binding of TATA binding protein (TBP), one of the components of the transcription machinery, to the TATA motif is the initial recognition step in transcription complex formation. The initiator has a loose consensus sequence and overlaps the transcription start site. It has been found in many TATA-less promoters and its interaction with Inr-binding proteins is critical in positioning RNA polymerase II (Weis and Reinberg, 1992). TFIID, a eukaryotic RNA polymerase II promoter recognition factor, has been shown to bind to both TATA box and Inr core promoter elements (Goodrich et al., 1996). According to GC content, the TATA-less promoters can be further divided into GC-rich and non-GC-rich promoters (Smale and Baltimore, 1989). The GC-rich promoters have been found primarily in housekeeping genes and usually contain multiple transcription start sites and several potential binding
sites for the Sp1 transcription factor. The non-GC-rich promoters are not constitutively active but rather are regulated during differentiation or development. Since one of the most important ways in which gene expression is regulated is through transcription initiation, defining promoter and potential regulatory elements represents an essential step toward our understanding of eukaryotic gene expression.

The expression of exogenous genes in cultured cells has proved to be invaluable in the analysis of transcriptional regulatory sequences. A plasmid that contains a reporter gene is frequently used to determine the regulatory region of a gene by transfecing the constructed plasmid DNA into cultured cells (for review see Alam and Cook, 1990). The most commonly used reporter gene is chloramphenicol acetyltransferase (CAT) gene (Gorman et al., 1982) which is cloned in a pSV2 vector containing a segment of pBR322 DNA carrying the plasmid origin of replication and the β-lactamase ampicillin resistance gene. This vector also contains the small t intron and a polyadenylation site from SV40 which are necessary for mature mRNA production in mammalian cells (Subramani and Southern, 1983). The cloning site for insertion of potential regulatory elements is located upstream of the CAT coding sequences. The chimeric genes generated with these constructs were assayed quantitatively for expression in recipient cells. Because there is no counterpart to the CAT gene in eukaryotes, the enzyme activity can be directly assayed in an extract of the cell. The CAT assay provides a simple, quantitative method and the level of CAT activity corresponds to the amount of enzyme produced after transfection, which in turn reveals the level of expression from the eukaryotic promoter.

The calcium phosphate-mediated transfection method, first described by Graham and Van der Eb (Graham and Van der Eb, 1973), has become one of the major methods for DNA transfer to mammalian cells since it is a highly efficient method for introducing DNA into a wide range of cells (Sambrook et al., 1989). This technique involves mixing purified DNA directly with CaCl₂ and phosphate buffer to form a fine calcium phosphate precipitate containing the DNA which is then placed on the cell monolayer. The precipitate binds to the plasma membrane and is taken into the cell by endocytosis. The calcium phosphate method can be used for the production of both transient and stable transformants.

The aim of this chapter was to investigate the mechanisms that regulate transcription of the uPAR gene by defining its promoter and potential regulatory
sequences in the 5'-flanking region, which are important in transcriptional regulation, using transient transfection assays. The initial approach used to delineate regulatory elements in this study was to compare expression of reporter constructs carrying 5' regions of the human uPAR gene progressively truncated from both 5' and 3' end. It was expected that this would broadly locate the regulatory elements of interest which could be further studied using DNA-protein binding analysis. Two human colon cancer cell lines, HCT116 and LIM1215, were selected and used as recipient cells in transient transfection assays in this study. They are highly invasive malignant cells and constitutively produce uPAR mRNA which is induced by PMA and cytokines (Schlechte et al., 1989; Wang et al., 1994). A CAT reporter gene system was utilized to study the relative transcriptional activity of the various 5'-flanking regions of the human uPAR gene. The calcium phosphate transfection procedure was used for uPAR-CAT chimeric DNA transfer.

3.2 Materials and Methods

3.2.1 Materials

n-butyryl coenzyme A and o-nitrophenol β-D-galactoside (ONPG) were obtained from Sigma Chemical Company, USA. D-threo-[dichloroacetyl-1-14C]chloramphenicol (57 mCi/mmol) was from Amersham Life Science, UK. Plasmids pCAT-Basic and pCAT-promoter were obtained from Promega Corporation, USA. Dulbecco's modified essential medium (DMEM) was from Gibco-BRL Life Technologies Inc., USA. TLC plates (TLC plastic sheets silica gel 60 F254) were purchased from Merker, Germany. c-fos and c-jun expression plasmids and pCMV-R which is the vector for c-fos and c-jun constructs were gifts from Dr. Donna Cohen of The John Curtin School of Medical Research.

3.2.2 Plasmid construction

3.2.2.1 Plasmids

The plasmid used for insertion of the 5'-flanking sequences was a promoterless and enhancerless pCAT-Basic which contains the coding region of the bacterial CAT gene and polyadenylation signal derived from simian virus 40 (SV40) gene. All CAT constructs were derived from this vector except those indicated. The pCAT-promoter
contains a SV40 promoter sequence and pCAT-RSV contains a SV40 promoter and a SV40 enhancer.

3.2.2.2 Deletion of 5'-flanking region

To generate human uPAR-CAT fusion constructs, a HindIII fragment of approximately 10 kb, containing 5'-flanking sequences, exon 1 of the uPAR gene and part of intron 1, was isolated from genomic clone λ11 (see chapter 2). The fragment was inserted in the sense orientation upstream of the CAT gene at the HindIII site of pCAT-Basic and designated as pCAT-H10 (Figure 3.1). Based on this construct, the first series of deletions was made by taking advantage of conveniently located restriction sites. First, pCAT-HN8.5 (HindIII-NruI) was generated by cutting pCAT-H10 at the NruI site in exon 1 and the Sall site of polylinker of pCAT-Basic to remove the ~1.5 kb NruI-HindIII fragment from 3' end, followed by converting the Sall site of the construct end with Klenow enzyme and religation as described in section 2.2.4.3. Then, the uPAR-CAT constructs, pCAT-AN7.0 (AccI-NruI), pCAT-SN2.6 (Sphi-NruI) and pCAT-PN1.3 (PstI-NruI), were generated by cutting pCAT-HN8.5 at AccI, Sphi and PstI restriction sites, respectively, within the 5'-flanking region and the corresponding sites in the polylinker of the pCAT-Basic and religated.

The pCAT-XN0.45 (XbaI-NruI) and pCAT-NX0.45 (NruI-XbaI) were generated by cutting a 456 bp of XbaI-XbaI fragment (450 bp of 5'-flanking sequence and 6 bp of polylinker sequence of pCAT-Basic) from pCAT-PN1.3 with XbaI and re-inserting into XbaI site of pCAT-Basic in both orientations. The pCAT-HX0.9 was generated by removing this 456 bp of XbaI-XbaI fragment from pCAT-PN1.3 by cutting with XbaI. Diagram of the CAT constructs is shown in Figure 3.2A.

3.2.2.3 5' and 3' deletion of the uPAR promoter region

To characterize the promoter region, a series of constructs was made by exonuclease III digestion. At first, the 456 bp of XbaI-XbaI fragment isolated from pCAT-XN0.45 was subcloned into a polylinker-modified pCAT-Basic vector in which Sall and AccI restriction sites of the polylinker were substituted by SpeI, BamHI and SmaI. This modification was made by replacing a polylinker region between PstI and XbaI in the pCAT-basic with a PstI-XbaI fragment from the polylinker of pBluescript. The generated construct was digested with PstI and BamHI and then digested with exonuclease III and religated using exonuclease III double-strands nested deletion kit (Pharmacia). The constructs pCAT-141 (-141/+47), pCAT-60 (-60/+47), pCAT-49 (-
49/+47), pCAT-29 (-29/+47) and pCAT-16 (-16/+47), were generated. pCAT-140 (-140/+47) was generated by cutting pCAT-XN0.45 with SmaI and religating.

To generate specific 5' deletion, the PCR product produced by using 5' primer D-100(+) with a PstI site and 3' primer Exon 1A- with NruI site (the PstI and NruI sites used for cloning sites are underlined, Table 3.1) was used to construct pCAT-100 (-100/+47). PCR reaction was performed as described in section 2.2.10. The generated PCR fragments were digested with PstI and NruI, respectively, and isolated by 1% agarose gel electrophoresis with 1X TAE and electroelution as described in section 2.2.4.1. The purified fragments were inserted into PstI and SmaI sites of pCAT-Basic.

Another series of 3' deletions of the uPAR promoter was generated by inserting three PCR fragments, -403 to -179, -403 to -46 and -403 to -10 (using 5' primer D-484(+): with PstI site and 3' primers D-179(-), D-46(-) and D-10(-) with XbaI site, respectively, Table 3.1) into XbaI site of pCAT-Basic. Three 3' deletion constructs, pCAT-403/179, pCAT-403/46 and pCAT-403/10, were generated.

3.2.2.4 Internal deletion of 5'-flanking region

To identify a potential negative element located at SphI/PstI segment (-2587 to -1268) upstream of the promoter region, a series of mutants containing internal deletions was made by joining different length fragments in this segment with the uPAR promoter region (pCAT-140). pCAT-S1 contains the SphI/PstI fragment. The constructs, pCAT-S2 to pCAT-S9 were generated by combining PCR fragments, -2104/-1268, -1796/-1268, -1268/-1796, -2104/-1458, -1796/-1458, -1718/-1458, -1613/-1458 and -1613/-1268 with the uPAR promoter region, respectively. PCR reactions were performed using 5' and 3' primers containing PstI site (D2104+, D1796+, D1718+, D1613+, D1458-, D1268-, Table 3.1). After PstI digestion and purification as described in section 2.2.4.1, these fragments were inserted into the PstI site of pCAT-140. In addition, the fragment -1796/-1458 was also inserted in front of the SV40 promoter in the pCAT-Promoter in both orientations and the generated constructs were named as pCAT-Pro/S and pCAT-Pro/AS, respectively.
Figure 3.1 The construct pCAT-Basic containing a uPAR sequence at the HindIII site. The pCAT-Basic vector lacks eukaryotic promoter and enhancer sequences. SV40 small T antigen region is located between 3064-3917 and multiple cloning sites (HindIII-XbaI) are located between 2242-2271. A 10 kb HindIII fragment of the human uPAR gene including the 5'-flanking region, exon I and part of intron 1, was inserted at HindIII site of pCAT-Basic. Some restriction sites used for deletions were marked. N, NruI; X, XbaI; P, PstI; S, SphI; A, AccI and H, HindIII. The transcription initiation site is indicated by an arrowhead.
Table 3.1 Primers used for constructing mutants with deletions in the 5'-flanking region of uPAR gene

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequences*</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-100(+)</td>
<td>5'-AACTGCAGGAGGGTTCGTTACGCGGAGGGGAGGAGG-3'</td>
</tr>
<tr>
<td>ExonIA(-)</td>
<td>5'-TGACCCATGTCGCGGAGGCACAGCTCCTGT-3'</td>
</tr>
<tr>
<td>D-484(+)</td>
<td>5'-AACTGCAGGAGCTGGGGGACAGCAGG-3'</td>
</tr>
<tr>
<td>D-10(-)</td>
<td>5'-GCTCTAGAGGAAGGAGCTTTTGTCG-3'</td>
</tr>
<tr>
<td>D-46(-)</td>
<td>5'-GCTCTAGACTCATGGAGTTGTGATCAC-3'</td>
</tr>
<tr>
<td>S-1613(+))</td>
<td>5'-AACTGCAGGGAGATTGCAATACCCCAG-3'</td>
</tr>
<tr>
<td>S-1718(+)</td>
<td>5'-AACTGCAGACCACCCTTGACAAAAATGTTG-3'</td>
</tr>
<tr>
<td>S-1796(+))</td>
<td>5'-AACTGCAGGTGCAGTGCTGCATTGCAGT-3'</td>
</tr>
<tr>
<td>S-2104(+)</td>
<td>5'-AACTGCAGTTTTCTTATTCCCCTCTCAGA-3'</td>
</tr>
<tr>
<td>S-1268(-)</td>
<td>5'-TGGAAGTGGAGGTGAGCTGAGCT-3'</td>
</tr>
<tr>
<td>S-1458(-)</td>
<td>5'-AACTGCAGTTTTGGAAAAAGGCTTATAATAG-3'</td>
</tr>
</tbody>
</table>

* restriction sites for PstI (CTGCAG), NruI (TCGCGA) or XbaI (TCTAGA) are underlined.
Plasmid DNA was prepared by the lysozyme-Triton procedure followed by two cycles of CsCl centrifugation and repeated phenol extraction as described in section 2.2.6.2. The orientations and boundaries of all pCAT-uPAR constructs were confirmed by restriction analysis and sequencing. Two synthetic oligonucleotide primers complementary to 5'‐ and 3'‐flanking sequences of polylinker (CAT-5', 5'-ATGCTTCGGCTCGTATGTTGT-3' and CAT-3', 5'-GGAAGCTTCCTTAGCTCC TGA-3') of pCAT-Basic, were used in DNA sequencing.

### 3.2.3 Cell culture and transient transfections

HCT116 or LIM1215 cells were maintained and subcultured in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal calf serum, 50 µg/ml penicillin and 50 µg/ml gentamycin. Cells were used between the tenth and fifteenth passage.

Transient transfection was performed according to the standard calcium phosphate method (Sambrook et al., 1989). Briefly, at 24 h before transfection, approximately 1.5 X 10⁶ cells were seeded in each 100 mm Petri dish in 10 ml of complete medium at 5% CO₂. After overnight culture, the medium was removed at 3 h before transfection and replaced with 9 ml of fresh complete medium. The calcium phosphate-DNA precipitates were prepared in HEPES-buffered saline. Each dish of cells was transfected with 10 µg of appropriate CAT construct and 2 µg of pRSV-β-galactosidase for 16 h. Following a wash with PBS, cultures were allowed to continue growth overnight in 10 ml of complete medium.

For transactivation experiments, 2.5 µg of the expression plasmids, c-fos and c-jun, were cotransfected with pCAT-uPAR constructs and pRSV-β-galactosidase. The pCMV-R vector was added to adjust the plasmid DNA to equal amounts. After 48 h transfection, cells were washed and harvested in PBS using a rubber policeman and transferred to an Eppendorf tube, then pelleted by centrifugation for 1 min. The supernatant was removed and replaced with 120 µl of 250 mM Tris-HCl, pH 7.8. The cells were lysed by three cycles of freezing on dry ice-ethanol for 5 min and thawing at 37°C for 5 min. Cell lysis was confirmed microscopically. The cell debris was pelleted by centrifugation and the supernatant assayed for β-galactosidase and chloramphenicol acyltransferase (CAT) activities.
In this study, the β-galactosidase activity was used to normalize the transfection efficiency for each dish. The β-galactosidase activity in a constant volume of extract was measured and the CAT assay was then carried out with amounts of extract that contain a defined amount of β-galactosidase activity as described by Sambrook et al. (1989).

3.2.4 Assays for β-galactosidase

Assays for β-galactosidase activity were performed according to the standard method (Sambrook et al., 1989). The β-galactosidase activity was measured in an assay that included 30 µl of cell lysate from each sample, 201 µl of 0.1 M sodium phosphate, pH 7.5, 3 µl of Mg²⁺ solution (0.1 MgCl₂, 4.5 M β-mercaptoethanol) and 66 µl of ONPG (4 mg/ml). After incubation at 37°C for 30 min, when a faint yellow color had developed, the reaction was terminated by the addition of 0.5 ml of 1 M Na₂CO₃. Absorption was then measured at 420 nm. The β-galactosidase activity was expressed as OD₄₂₀ X 100, extrapolated to 1 h of enzymic reaction as described previously (Herbomel et al., 1984). Under these conditions the β-galactosidase activity obtained ranged from 30 to 300 U per plate of HCT116 cells (~10⁵).

3.2.5 CAT assay

The CAT assay was performed according to the standard method (Sambrook et al., 1989). After β-galactosidase activity assay, the remaining supernatant was heated to 65°C for 10 min to inactivate endogenous acetylases. Additional debris was removed by centrifugation at 4°C for 15 min and the supernatant was transferred to a fresh Eppendorf tube. Appropriate standardized amounts (4-20 µl) of extract with the same β-galactosidase activity (2.5 units) was used for CAT activity assay. The reaction was carried out by adding a mixture of 80-100 µl of 1 M Tris, pH 7.8, 5 µl of n-butyryl-coenzyme A (5 µg/ml) and 1 µl of ¹⁴C-chloramphenicol (25 µCi/ml, 57 mCi/mmol, Amersham) to a final volume of 125µl and incubating at 37°C for 1h.

3.2.5.1 Thin layer chromatography (TLC) method

Following incubation, the CAT reaction was stopped by adding 1 ml of ethyl acetate and each sample was vortexed for 15 seconds. The aqueous and organic mix was then spun for 2 min in an Eppendorf centrifuge. The upper organic phase (900 µl)
was transferred to a new tube and the ethyl acetate was evaporated under vacuum. The reaction products were redissolved in 20 µl of ethyl acetate immediately before application to a silica gel thin layer plate (Merck). Ascending chromatography was conducted for 30 min with chloroform:methanol (95:5) as the running solvent. The plate was dried and exposed to Kodak XAR X-ray film for 16-24 h. The CAT activity was determined by quantitation using the PhosphorImager system (Molecular Diagnostics).

3.2.5.2 Liquid scintillation counting (LSC) method

The LSC method was also used in the CAT assay for a faster quantitation. Following CAT reaction, 300 µl of mixed xylene was added to each sample and the mixture vortexed for 30 seconds before centrifuging at room temperature for 5 min to separate the phases. The upper phase was transferred to a new tube and then back-extracted with 100 µl of 0.25 M Tris-HCl, pH 7.8. After vortexing for another 30 seconds and centrifuging, the upper xylene phase was back-extracted again. Finally, 200 µl of the xylene phase was removed and added to a scintillation Pony vial (Packard) containing 500 µl of Optiscint (LKB). The butyrylated 14C-chloramphenicol was measured in a scintillation counter. All assays were within the linear range of the assay.

3.3 Results

The screening of a human blood genomic DNA library with the 1.14 kb human uPAR cDNA probe resulted in the isolation of 16 independent clones. Two clones containing the 5'-flanking region of the uPAR gene, λ11 and λ33, have been identified and characterized (See chapter 2). The λ11 was used in transcription analysis of the 5'-flanking region of the human uPAR gene because it contains a longer upstream uPAR sequence.

3.3.1 Functional analysis of the human uPAR 5'-flanking region

To search for the DNA sequences required for uPAR gene expression, a series of constructs with different lengths of 5'-flanking sequence and 47 bp of the 5'-untranslated region was fused to the pCAT-Basic and transfected into HCT116 colon cancer cells. Each construct was cotransfected with the pRSV-β-galactosidase plasmid to normalize transfection efficiency. No detectable β-galactosidase activity was found in the HCT116 cells. CAT activities were measured from the cell lysates. As shown in
Figure 3.2, both pCAT-HN8.5 and pCAT-AN7.0 promoted CAT transcription and gave much higher levels of CAT activity compared to the negative control, pCAT-Basic. Deletion of the uPAR 5'-flanking sequence from the 5' end of pCAT-AN7.0 to the SphI site (pCAT-SN2.6) had a moderate decrease in CAT activity. However, deletion from the SphI to the PstI site (pCAT-PN1.3) showed significant increase in promoter activity suggesting that a silence-like activity may be located at the 1.3 kb region between SphI and PstI. Further deletion of the sequence from PstI to XbaI (pCAT-XN0.45) did not cause significant change in promoter activity. To determine whether this promoter is orientation-dependent, the 450-bp XbaI-NruI fragment (-403 to +47) with the highest promoter activity was subcloned into pCAT-Basic vector in anti-sense orientation (pCAT-NX0.45). As expected this construct did not show any CAT activity. Furthermore, no promoter activity was observed in pCAT-PX0.9 in which the XbaI-NruI segment was deleted. These results indicate that important sequences for uPAR promoter activity are presented between bases -403 and +47 and the region between SphI and PstI sites may contain a negative regulatory element.

3.3.2 Identification of the uPAR promoter region

To define the boundaries of the functional promoter and regulatory elements between -403 and +47, a further series of 5' deletion constructs were prepared. Based on the pCAT-XN0.45, constructs pCAT-141, pCAT-100, pCAT-60, pCAT-49, pCAT-29 and pCAT-16 were generated (Figure 3.3A). As shown in Figure 3.3B, deletion of sequences from -403 to -141 did not significantly affect promoter activity. Both pCAT-XN0.45 and pCAT-141 showed similar promoter activity. However, deletion to position -100 (pCAT-100) reduced the promoter activity by 60% and deletion to position -60 (pCAT-60) showed less than 5% of initial activity. Further deletion to positions -49, -29, and -16 (pCAT-49, pCAT-29 and pCAT-16) also reduced activity. These results indicate that the 81 bp segment located between -141 and -61 bp constitutes an important region that contributes to the uPAR promoter activity.
Figure 3.2  Functional analysis of the 5'-flanking region of the human uPAR gene. A, diagrammatic representation of CAT constructs containing 5'-flanking region of the uPAR gene. The constructs were generated using convenient restriction sites and fused into pCAT-Basic vector. Sequence in antisense orientation in pCAT-NX0.45 is indicated by an arrow. B, constructs were transiently transfected into HCT116 colon cancer cells and CAT activity was assayed after 48 h transfection. The mean values ± standard deviations of the relative activities were determined in at least three separate experiments except pCAT-NX0.45 and pCAT-PX0.9 (two separate experiments) are shown. The activity of pCAT-Basic was referred as 1.
A

(-58kb) (-7kb)

HindIII Accl Sphl PstI XbaI

pCAT-Basic

pCAT-HN8.5

pCAT-AN7.0

pCAT-SN2.6

pCAT-PN1.3

pCAT-XN0.45

pCAT-NX0.45

pCAT-PX0.9

pRSV-CAT

CAT

Relative CAT activity

0.67±0.27

11.11±3.89

12.01±1.49

9.56±5.69

16.06±7.29

16.33±1.13

14.1±0.76

0.10±0.03

61.58±6.45

B
Figure 3.3 5' deletion analysis of the human uPAR promoter. A, a diagram of CAT constructs with uPAR promoter sequence deleted from 5' end. The transcription initiation site is indicated by an arrow. B, constructs were transiently transfected into HCT116 colon cancer cells and CAT activity was assayed after 48 h transfection. The mean values ± standard deviations of the relative activities were determined in at least three separate experiments are shown. The activity of pCAT-Basic was referred as 1.
A

<table>
<thead>
<tr>
<th>Restriction Site</th>
<th>CAT Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>XbaI (-403)</td>
<td></td>
</tr>
<tr>
<td>Smal (-264)</td>
<td></td>
</tr>
<tr>
<td>Smal (-140)</td>
<td></td>
</tr>
<tr>
<td>NruI (+47)</td>
<td></td>
</tr>
</tbody>
</table>

- **pCAT-XN0.45**: 16.33±1.13
- **pCAT-141**: 17.13±0.72
- **pCAT-100**: 5.24±2.05
- **pCAT-60**: 1.56±0.25
- **pCAT-49**: 1.06±0.77
- **pCAT-29**: 1.91±1.51
- **pCAT-16**: 1.89±1.29
- **pCAT-Basic**: 0.67±0.27

B

[Image of gel electrophoresis results]
Another series of constructs prepared by deletion of pCAT-XN0.45 from the 3' end showed that CAT activity was significantly reduced in response to the deletion of three specific regions of the gene (Figure 3.4). There was a decrease in CAT activity of more than 90% when the region between bases -9 and +47 (pCAT-403/-10) was deleted, suggesting that the region between -9 and +47 is necessary for uPAR promoter activity. Further deletions to positions -46 (pCAT-403/-46) and -179 (pCAT-403/-179) completely abolished the CAT activity. Together, these results show that the region for maximal promoter activity of the human uPAR gene is located between -141 and +47. Two regions, -141 to -61 and -9 to +47, may be important for the uPAR gene expression. Similar results were also obtained when these CAT constructs were transfected into another human colon cancer cell line, LIM1215, (data not shown) suggesting that the uPAR promoter activity is not cell specific.

3.3.3 Deletion analysis of the silencer region

To identify the putative silencer-like activity of the region from -2587 to -1268, an internal deletion (pCAT-S1) was initially constructed by putting this 1.3 kb fragment in front of the uPAR promoter sequence in pCAT-140 and assayed for silencer activity in HCT116 cells. Surprisingly, this internal deletion significantly decreased the uPAR promoter activity compared to the native position, suggesting a distance-dependent inhibitory effect (Figure 3.5).

Because sequence analysis demonstrated that the 1.3 kb silencer-like region (Sphi-PstI) contains a conserved full-length Alu sequence (-1799 to -1518), we hypothesized that the Alu sequence may correlate with observed silencer activity. To test this, a series of internal deletions of the silencer-like region from the 5'- and 3'-ends was constructed by inserting various PCR fragments from this region in front of the uPAR promoter in pCAT-140 (Figure 3.5A). The deletion constructs were transfected into HCT116 cells, and assayed for silencer activity. As shown in Figure 3.5B, resulting constructs, pCAT-S2 (-2104 to -1268), pCAT-S3 (-1796 to -1268), pCAT-S4 (-1268 to -1796), pCAT-S5 (-2104 to -1458) and pCAT-S6 (-1796 to -1458), showed similar strong negative effects on the promoter activity. The full-length Alu repetitive sequence (pCAT-S6) retained the whole silencer activity, inhibiting more than 80% CAT activity. When the silencer fragment containing Alu sequences was placed in front of the uPAR promoter in antisense orientation (pCAT-S4), the same negative effect was also observed indicating
that this silencer element functions independently of its orientation. However, further deletion from 5' end of this Alu sequence to -1718 (pCAT-S7, -1718 to -1458) or -1613 (pCAT-S8, -1613 to -1458) showed less silencer activity than pCAT-S6. pCAT-S9 (-1613 to -1268) also showed less negative effect on the promoter activity. These results suggest that a full length Alu sequence may be necessary for maximal silencer activity.

To determine whether this Alu repeat sequence specifically affects the uPAR promoter activity, this fragment was also inserted upstream of the pCAT-Promoter vector, which contains a SV40 promoter, and transfected into HCT116 cells. As shown in Figure 3.5, this Alu silencer also decreased the SV40 promoter activity in both orientations (pCAT-Pro/S and pCAT-Pro/AS) compared to pCAT-Promoter, suggesting that the uPAR silencer was able to inhibit CAT activity in an orientation- and a promoter-independent fashion.

### 3.3.4 Transient transactivation with AP-1 transcription factor

To test whether the uPAR promoter could be transactivated by AP-1 transcription factor, either uPAR promoter construct pCAT-140 or pCAT-XN0.45, with c-fos, c-jun or c-fos plus c-jun expression plasmids, were transiently cotransfected into HCT116 cells, respectively. AP-1 was initially used since two AP-1 binding sequences at positions -184 and -70 were found within the 5' flanking region (see chapter 2) and uPAR expression in HCT116 cells is up-regulated by PMA or TNFα (Wang et al., 1994b), which have been demonstrated to increase uPA expression by activating AP-1 (Nerlov et al., 1992; Lengyel et al., 1994). Cotransfection with c-fos or c-jun expression plasmids only slightly induced transactivation of the uPAR promoter (Figure 3.6). However, the uPAR promoter activity was significantly increased up to 4- to 5-fold by cotransfection with c-fos plus c-jun expression plasmids. Both pCAT-XN0.45 and pCAT-140 showed similar enhanced promoter activity suggesting that the AP-1 binding sequence located at -70 may be required for the transactivation of AP-1. These results indicate that this AP-1 site may exert a significant influence on the uPAR promoter and provide a link between up-regulation of the uPAR gene expression and the stimuli that activate AP-1.
Figure 3.4 3' deletion analysis of the human uPAR promoter. A, a diagram of CAT constructs with uPAR promoter sequence deleted from 3' end. The transcription initiation site is indicated by an arrow. B, constructs were transiently transfected into HCT116 colon cancer cells and CAT activity was assayed after 48 h transfection. The mean values ± standard deviations of the relative activities were determined in at least three separate experiments are shown. The activity of pCAT-Basic was referred as 1.
<table>
<thead>
<tr>
<th>Vector</th>
<th>Relative CAT activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCAT-Basic</td>
<td>0.67 ± 0.27</td>
</tr>
<tr>
<td>pCAT-XN0.45(-403/+47)</td>
<td>16.33 ± 1.13</td>
</tr>
<tr>
<td>pCAT-403/-10</td>
<td>1.99 ± 1.26</td>
</tr>
<tr>
<td>pCAT-403/-46</td>
<td>0.60 ± 0.51</td>
</tr>
<tr>
<td>pCAT-403/-179</td>
<td>0.83 ± 0.54</td>
</tr>
</tbody>
</table>
Figure 3.5  Internal deletion analysis of 5'-flanking region of the human uPAR gene. A, a diagram of CAT constructs with internal deletion of the 5'-flanking region. The various PCR fragments amplified from the region between -2568 and -1268 were fused in front of the uPAR promoter (pCAT-140). Numbers of nucleotide relative to transcription initiation site are indicated. SV40 promoter in pCAT-Promoter vector was marked by black boxes. DNA sequences in antisense orientation are indicated by arrows. B, constructs were transiently transfected into HCT116 colon cancer cells and CAT activity was assayed after 48 h transfection. The mean values ± standard deviations of the relative activities were determined in at least three separate experiments are shown. The activity of pCAT-Basic was referred as 1.
A

<table>
<thead>
<tr>
<th>DNA Region</th>
<th>Relative CAT activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCAT-Basic</td>
<td>0.67±0.27</td>
</tr>
<tr>
<td>pCAT-140</td>
<td>17.30±5.16</td>
</tr>
<tr>
<td>pCAT-S1</td>
<td>6.51±2.01</td>
</tr>
<tr>
<td>pCAT-S2</td>
<td>4.30±1.19</td>
</tr>
<tr>
<td>pCAT-S3</td>
<td>2.79±0.72</td>
</tr>
<tr>
<td>pCAT-S4</td>
<td>2.75±0.69</td>
</tr>
<tr>
<td>pCAT-S5</td>
<td>2.61±0.89</td>
</tr>
<tr>
<td>pCAT-S6</td>
<td>2.52±0.57</td>
</tr>
<tr>
<td>pCAT-S7</td>
<td>6.02±2.80</td>
</tr>
<tr>
<td>pCAT-S8</td>
<td>6.88±4.33</td>
</tr>
<tr>
<td>pCAT-S9</td>
<td>6.31±2.32</td>
</tr>
<tr>
<td>pCAT-Promoter</td>
<td>2.60±1.88</td>
</tr>
<tr>
<td>pCAT-Pro/AS</td>
<td>1.23±0.57</td>
</tr>
<tr>
<td>pCAT-Pro/AS</td>
<td>0.54±0.24</td>
</tr>
</tbody>
</table>

B

[Image of gel electrophoresis results]
Figure 3.6 Cotransfection analysis of the 5'-flanking region of the human uPAR gene with c-fos and c-jun expression plasmids. 10 µg of pCAT-XN0.45 or pCAT-140 were cotransfected with 2.5 µg of c-fos, c-jun or c-fos plus c-jun expression plasmids into HCT116 cells. Same amount of plasmid DNA was adjusted by adding pCMV-R vector. CAT activity was assayed after 48 h transfection. The mean values ± standard deviations of the relative CAT activities determined in at least three separate experiments are shown.
3.4 Discussion

In this chapter, a promoter, a silencer and the transcriptional regulatory activities in the 5'-flanking region of the human uPAR gene were determined by transcription analysis in human colon cancer cells. This study elucidates the molecular mechanisms leading to the regulation of uPAR gene expression.

The results presented in this chapter clearly demonstrate that the human uPAR gene has a very strong promoter activity. The region required for maximal promoter activity resides in the 188 bp (-141 to +47) located immediately upstream of the major transcription initiation site. Any further deletions from both 5' and 3' ends result in significant loss of the promoter activity. The promoter activity was stable and reproducible in two human colon cancer cell lines tested, HCT116 and LIM1215. Two regions were found to be important for human uPAR promoter activity when deleting the promoter sequence from both 5' and 3' ends. Sequence analysis revealed that these two regions contain Sp1, Inr, as well as DPE motifs which play a primary role in the regulation of many gene promoters (Pugh and Tjian, 1990; Smale and Baltimore, 1989; Burke and Kadonaga, 1996).

The first region identified by 5' deletion analysis is 81 bp long, spanning from -141 to -61 and may contain multiple regulatory elements, since a gradual decline in promoter activity with progressive 5' deletion was observed. Two Sp1 binding sites (GGGAGG) were found at positions -103 and -94 within this region. The deletion of the sequence between -141 and -101, containing the first Sp1 site, caused a decrease of about 60% in expression activity compared to that of the intact promoter, whereas deletion of the second Sp1 site between -100 and -61 led to almost complete loss of promoter activity. These results demonstrated that both Sp1 binding sites might be required for maximal promoter activity. It has been demonstrated that Sp1 transcription factor is required to stimulate uPAR expression in Schneider cells, that lack endogenous Sp1, by cotransfection with Sp1 expression plasmid (Soravia et al., 1995). The data presented here suggest that the Sp1 binding sites presented between -141 and -61 might be important for the basal promoter activity of the uPAR gene.

The second region identified by 3' deletion analysis is 56 bp long, spanning from -9 to +47. A initiator element (-12 to +1) within this region has been indicated by Soravia et al. (1995). In addition, a DPE motif which has been recently reported to be responsible for the TFIID binding (Burke and Kadonaga, 1996) was also found at +11.
Since TFIID can recognize both Inr and DPE motifs, 3' deletion of the region between -9 and +47 in the uPAR promoter may inhibit uPAR promoter activity by abolishing formation of transcription machinery around the transcription start site. Downstream sequence of the promoter has been shown to be important for the promoter activity. For example, deletion of sequences from +9 to +38 on the glial fibrillary acidic protein gene promoter diminished transcription of this gene by abolishing the downstream promoter binding by TFIID (Nakatani et al., 1990). Moreover, it was also noted that deletion of the sequence from -9 to +47 abolished formation of an imperfect stem-loop structure around the transcription initiation site that may be important for the regulation of the TATA-less and G+C-rich promoters (Ackerman et al., 1993).

Comparison of the DNA sequence between the human and mouse uPAR promoters revealed significant homology (68% identity) only within a proximal region of about 135 nucleotides (Soravia et al., 1995). Both the human and mouse promoters lack TATA boxes and present a moderately high G+C content in the proximal region. The mouse uPAR sequence contains multiple canonical Sp1 binding sites, while the human sequence lacks the typical Sp1 consensus sequence but it contains five Sp1-like sequences in the proximal region (Soravia et al., 1995). The high homology between the human and mouse genes suggests that this region contains an important regulatory sequence and implicates a similar regulatory mechanism.

Internal deletion analysis demonstrated that a full-length Alu repetitive sequence is involved in uPAR silencer-like activity in an orientation-independent and position-dependent fashion. The 339 bp silencer region, spanning from -1796 to -1458, contains a highly conserved Alu sequence with 89% homology to human consensus Alu sequence (Kariya et al., 1987). Detailed silencer deletion experiments indicated that a full-length Alu sequence is required for maximal silencer activity because further deletion of this Alu repeat reduced the silencer effect. The Alu repeat sequence, which functions as either silencer or enhancer, has been implicated in transcription regulation in a number of genes, including keratin 18 (Neznanov and Oshima, 1993), e-globin (Wu et al., 1990), CD8α (Hanke et al., 1995), estrogen receptor (Norris et al., 1995) and Wilms' tumor (Hewitt et al., 1995) genes. The mechanisms involved in Alu sequence-modulated transcription are diverse. Because the Alu repeat contains a Poll III promoter it may block the passage of transcribing polymerases by formation of its own transcription complexes (Wu et al., 1990). It may also affect promoter activity by functioning as an
insulator (Neznanov et al., 1993) and forming secondary structure (Hanke et al., 1995). The interaction of Alu-binding proteins with the Alu repeat in a sequence-specific manner has been reported to contribute to its silencer activity (Hewitt et al., 1995). In addition, the introduction of base substitutions within the Alu repeat by evolution may lead to new transcription factor-binding sites, thus allowing Alu repeats to take on new roles as transcriptional enhancers or silencers (Brini et al., 1993). It is likely that the Alu repeat present within 5'-flanking region of the uPAR gene interferes with transcription of the uPAR mRNA by formation of a DNA-protein complex and secondary structure with other Alu repeat sequences. Sequence analysis showed that four Alu repeat sequences are located within the 2.6 kb of 5'-flanking sequences, suggesting the uPAR silencer may form secondary structures with downstream or upstream Alu repeats. Thus it is likely that the silencer element present within the 5'-flanking region of the uPAR gene interferes with transcription of the uPAR mRNA by formation of DNA-protein complexes and secondary structure with other Alu repeat sequences.

uPAR gene expression is induced by a number of growth factors, cytokines and chemical agents in a variety of cell types (Blasi, 1993). In HCT116 colon cancer cells, uPAR expression is up-regulated by phorbol esters, TNFα and down-regulated by amiloride (Wang et al., 1994). Transcription factor AP-1 has been reported to mediate cellular responses to many stimuli and to modulate gene transcription by its binding to the AP-1 motif TGAC/GTCA in the promoter/enhancer region of many different genes (Wang and Goldstein, 1993). Because two putative AP-1 binding sequences were found within the 5' region of the uPAR gene, it is interesting to determine whether the AP-1 transcription factor activates uPAR transcription. The data from cotransfection with c-fos and c-jun expression plasmids demonstrate that AP-1 strongly transactivates uPAR transcription. The AP-1 motif at -70, but not -184, may be responsible for the AP-1 mediated uPAR activation. This AP-1 motif is also conserved in the mouse uPAR promoter region (Soravia et al., 1995). Transcription factors integrate information from promoter sequences and signal transduction pathways to control the rate of gene expression. Thus, this finding establishes a potential molecular link between those signaling pathways through AP-1 and up-regulation of uPAR gene expression.
In summary, the results presented in this chapter demonstrate multiple regulatory elements located at the 5'-flanking region of the human uPAR gene suggesting that the transcription regulation of the uPAR gene is complex. Two regulatory regions, -141 to -60 and -10 to +47, appear to be involved in basal promoter activity, while the AP-1 binding site in the promoter may contribute to activation of the uPAR gene. In addition, an Alu repeat sequence may be involved in negative regulation of uPAR gene expression. Further study is required to characterize the specific DNA-binding proteins that regulate both the basal and inducible transcription of the human uPAR gene.
CHAPTER 4
MULTIPLE NUCLEAR FACTORS INTERACT WITH THE HUMAN uPAR PROMOTER SEQUENCES

4.1 Introduction

In eukaryotes, the regulation of gene transcription is emerging as a complicated process involving the interactions of a number of trans-acting factors with specific cis-acting DNA sequences (Mitchell and Tjian, 1989; Goodrich et al., 1996). Both the basal process of transcription itself and its regulation are controlled by specific DNA elements in the gene promoters, enhancers or silencers. These elements are recognized and bound by sequence-specific DNA-binding proteins that then influence the rate of gene transcription by either activating or repressing transcription initiation. Gene promoters utilize a large combinatorial array of DNA regulatory elements and transcription factors as receivers of physiological signals to direct spatial and temporal gene expression (Pugh and Tjian, 1990; Roeder, 1991). Thus, gene regulatory programs are governed by the activity of transcription regulatory proteins. Transcriptional regulatory proteins typically contain two functional domains: a DNA-binding domain that recognizes specific sequences and an activation domain that is required for transcription stimulation (Mitchell and Tjian, 1989). The study of these transcription factors has been a critical aspect of gene regulation.

A number of methods have been established to study the interaction between transcription factors and specific DNA sequence, such as DNA footprinting and electrophoretic mobility shift assay (EMSA). DNA footprinting is a high resolution technique which allows the determination of a short protein-binding site within a relatively large DNA fragment, and thus provides an essential step in the characterization of transcription factors (Ausubel et al., 1989). The general principle of in vitro DNA footprinting is based on a sequence-specific DNA-binding protein ‘protecting’ the nucleotides involved in DNA binding from enzyme or chemical cleavage. In the DNase I footprinting method, the DNA of interest is labeled at one end of one strand of the DNA duplex only, and then digested with DNase I enzyme. The limited digestion, such that each molecule of the DNA is cleaved randomly only once or a few times, results in a ladder of DNA fragments of varying size when the DNA is
subjected to denaturing polyacrylamide gel electrophoresis and detected by autoradiography. In the presence of DNA-binding proteins, the DNA is complexed with protein(s) and will be protected against cleavage. This results in a gap in the DNA ladder which is representative of a protein binding to a specific DNA sequence. By comparing the gap in the ladder to the Maxam and Gilbert sequencing reactions, the precise DNA sequence of the DNA-binding site of the protein can be determined.

The EMSA is a simple and sensitive method for determining interactions between protein and DNA (Ausubel et al., 1989). It has been widely used to identify and investigate DNA-binding proteins. The principle of the DNA mobility shift assay is based on the ability of a protein to bind to a radiolabeled DNA fragment in vitro, followed by electrophoretic separation of DNA-protein complexes from the unbound DNA on non-denaturing polyacrylamide gels. Thus, DNA-protein interaction can be determined by observing a DNA fragment with an altered mobility, while the specificity of the DNA-protein complex can be tested by competing for binding with non-radiolabeled DNA fragments. Although, unlike DNA footprinting, EMSA does not identify the DNA nucleotides recognized by the protein directly, this method provides a powerful tool for the identification of both known and novel factors binding to specific DNA sequences.

The expression of the uP AR gene is inducible in various cell types, including T cells, monocytes and many cancer cells, by phorbol ester, growth factors and cytokines (Blasi, 1993). Elevated expression of the uP AR gene has been implicated in tumor invasion and metastasis (Blasi, 1993; Wang et al., 1994). Many studies have shown that uP AR expression is regulated at the level of gene transcription, suggesting that the 5'-flanking region of the uP AR gene is important for regulated expression during cancer invasion. However, whether transcription factors are involved in up-regulation of uP AR transcription is unclear. The inappropriate regulation of transcription factors, for example NF-κB and their dependent genes, has been associated with various pathological conditions including toxic shock, radiation damage and cancer (Siebenlist et al., 1994; Singh and Aggarwal, 1995). Therefore, the study of transcription factors that control uP AR gene expression may provide a prospect for therapeutic intervention.

The isolation and characterization of the 5'-flanking region of the human uP AR gene provided us with an opportunity to study regulation of uP AR gene expression by
analyzing the cis-acting promoter elements. Deletion analysis revealed that two regulatory regions that contribute to uPAR promoter activity contain putative Sp1, Inr and DPE motifs. Cotransfection with AP-1 suggested that the AP-1 binding site located in the promoter region is involved in activation of the uPAR gene. Sp1 is a ubiquitously expressed transcription factor that plays a primary role in the regulation of a large number of gene promoters (Pugh and Tjian, 1990). Inr and DPE elements are important for the basal transcription in genes containing TATA-less promoters (Goodrich et al., 1996; Burke and Kadonaga, 1996). AP-1 has been shown to regulate transcription in TNFα- or PMA-inducible genes including uPA (Lee et al., 1993; Lengyel et al., 1995). This chapter examines the transcription factors that might be responsible for uPAR regulation and illustrates the interaction between nuclear factors and the uPAR promoter region by both DNA footprinting and EMSA analysis.

4.2 Materials and Methods

4.2.1 Materials

Phenylmethylsulfonyl fluoride (PMSF) and bovine serum albumin (BSA) were from Sigma. RQ1 RNase-Free DNase, Gel Shift Assay kit and NF-κB protein (p49) were obtained from Promega. Bio-Rad Protein Assay kit was from BioRad. Double-stranded poly(dI-dC) and yeast tRNA were obtained from Pharmacia-LKB Biotechnology. Oligonucleotides AP-1, Sp1, NF-κB, TFIID and GRE were purchased from Promega.

4.2.2 Preparation of nuclear extracts

The nuclear extract from HCT116 and LIM1215 human colon cancer cell lines was prepared as shown below (Dignam et al., 1983; Ausubel et al., 1989). Cells were grown in T175 Falcon tissue culture flasks. At confluency, cells (3-5 X 10^8) were washed with PBS, scraped using a policeman and collected in conical tubes. The cells were centrifuged for 10 min at 2,000 x g at 4°C. The pellet was resuspended in 5 packed cell volumes of cold Buffer A (10 mM HEPES, pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, 0.2 mM PMSF, 0.5 mM DTT) and immediately centrifuged for 5 min at 2,000 x g at 4°C. The pellet was resuspended in Buffer A to 3 packed cell volumes and allowed to swell on ice for 10 min. The cells were homogenized by 20 strokes in a Dounce homogenizer and nuclei were pelleted for 15 min at 3,500 x g at 4°C. The
packed nuclei volume was determined and the nuclei were resuspended in 0.5 packed nuclei volume Buffer D (20 mM HEPES, pH 7.9, 1.5 mM MgCl₂, 0.2 M KCl, 0.2 mM EDTA, 0.5 mM DTT, 25% glycerol, 0.2 mM PMSF). High salt buffer C (0.9 M KCl) was slowly added, with constant mixing, to a final concentration of 0.3 M KCl. The nuclei were extracted at 4°C for 30 min and then centrifuged 30 min at 4°C at 25,000 x g. The supernatant was dialyzed against 100 volumes of dialysis buffer (20 mM HEPES, pH 7.9, 100 mM KCl, 0.2 mM EDTA, 20% glycerol, 0.5 mM DTT, 0.2 mM PMSF) for 1 h and then centrifuged for 20 min at 4°C at 25,000 x g. Aliquots of nuclear extracts were stored at -70°C. Protein concentrations were measured by the Bio-Rad protein assay using bovine serum albumin as a standard.

4.2.3 DNase I footprinting

4.2.3.1 DNA probe

DNase I footprinting assays were performed as described below (Ausubel et al., 1989) with minor modifications. The 188 bp HindIII-XbaI fragment, which contains a Smal site in the 5' end and an AccI site in the 3' end of the -141/+47 uPAR promoter region, was first was isolated from pCAT-141 and dephosphorylated. The DNA fragment was labeled by T4 polynucleotide kinase and [γ-³²P]ATP, then cut with Smal or AccI to get upstrand or lower strand DNA probe labeled at one end of one strand. The labeled fragments were isolated on a 5% polyacrylamide gel with 0.5X TBE using a vertical electrophoresis apparatus. After electrophoresis, the gel was wrapped with plastic film and exposed to XAR X film for 5 min. The radiolabeled fragment of interest was then cut off with a razor blade, and isolated by electroelution and purified by Elutip D column.

4.2.3.2 Binding reaction and DNase I digestion

Approximately 1 ng (20,000 cpm) of DNA probe was added together with 2 µg of poly (dI-dC) to a total volume of 50 µl in binding buffer (25 mM Tris-HCl, pH 8.0, 50 mM KCl, 6.25 mM MgCl₂, 0.5 mM EDTA, 10% glycerol, 0.5 mM DTT), in the presence or absence of nuclear extract as indicated and kept on ice for 30 min. The binding reaction mixture was then digested at room temperature for 1 min by adding 50 µl of a solution containing 10 mM MgCl₂, 5 mM CaCl₂ and 0.1-0.5 units of RQ1 RNase-Free DNase which was diluted in cold 1 M Tris-HCl (pH 8.0) before use.
Reactions were stopped by addition of 90 µl of stop solution (200 mM NaCl, 30 mM EDTA, 1% SDS, 100 µg of yeast tRNA), and the DNA extracted with phenol-chloroform, ethanol precipitated, washed, dried, and resuspended in 4 µl of loading buffer (0.1 M NaOH:formamide[1:2 v/v], 0.1% xylene cyanol, 0.1% bromophenol blue).

4.2.3.3 Preparation of G+A marker
To define the position of the protected region, G+A sequence ladders were prepared according to the method described by Maxam and Gilbert (1980). The labeled target DNA (~5 ng) in 10 µl of TE buffer containing 1 µg of calf thymus DNA was incubated with 1 µl of 4% formic acid at 37°C for 30 min and the reaction stopped by adding 150 µl of the fresh diluted piperidine solution (1 M) and heating to 90°C for 30 min. The reaction was then placed on ice for 5 min and the DNA extracted with 1 ml of butanol with mixing by vortex. After centrifuging at high speed in an Eppendorf for 5 min, the DNA pellet was resuspended in 1% SDS and re-extracted with 1 ml of butanol. Finally, the DNA pellet was carefully rinsed twice with 0.5 ml of butanol, and dried under vacuum for 10 min and redissolved in 5-10 µl of DNA sequencing loading dye.

4.2.3.4 Electrophoresis of digested DNA samples
Before electrophoresis, the DNA samples were denatured by heating for 3 min at 80°C and cooling immediately on ice. Denatured DNA samples were analyzed by electrophoresis on 6% polyacrylamide-8M urea sequencing gels with 0.5X TBE at 1800 V for about 2h. The end-labeled probes subjected to the G+A sequencing reaction of Maxam and Gilbert were loaded on to the same gels for identification of protected sequences. The gels were then dried and autoradiographed.

4.2.4 Electrophoretic mobility shift assay (EMSA)
4.2.4.1 DNA fragments and synthetic oligonucleotides
DNA fragments were generated by restriction digestion with appropriate enzymes. A 72 bp DNA fragment (-137 to -66 bp) was excised from construct pCAT-141 using SmaI and HinfI. A 55 bp DNA fragment (-8 to +47) was obtained by digestion with Ddel and NruI. After purification by electrophoresis the DNA fragment was labeled with the Klenow fragment of DNA polymerase I and [α-32P]dCTP, or T4

92
polynucleotide kinase and [γ-32P]ATP, as described in section 2.2.2.5. The labeled DNA probes were purified by electrophoresis on a 5% polyacrylamide gel as described above.

The specific uPAR oligonucleotides used as DNA probes were synthesized according to the uPAR promoter sequence (Table 4.1). Complementary strands were annealed by combining equal amounts of each oligonucleotide in 25 mM Tris-HCl, pH 8.0, 5 mM MgCl₂ and 25 mM NaCl, heating to 90°C for 5 min and then cooling to room temperature. The oligonucleotides AP-1, NF-κB, Sp1, GRE and TFIID (Promega) were also used in these studies. The double-stranded oligonucleotide DNA was 5'-end-labeled using [γ-32P]ATP and T4 polynucleotide kinase or used as unlabeled competing DNA in competition experiments.

4.2.4.2 Binding reaction

EMSA with crude nuclear extracts were performed using the Gel Shift Assay System kit according to the Manufacturer's description (Promega). The nuclear extracts were preincubated for 5-10 min at room temperature in a 9 µl reaction mixture containing 10 mM Tris-HCl, pH 7.5, 1 mM MgCl₂, 50 mM NaCl, 0.5 mM EDTA, 0.5 mM DTT, 1 µg of poly(dI-dC), 4% glycerol and the indicated amount of nuclear extracts. Each reaction contained 10,000 cpm oligonucleotide (0.1-0.5 ng) or DNA probe (~1 ng) and the incubation was continued for 20 min. During competition experiments, specific or nonspecific competitor DNA or oligonucleotides as indicated were added to the samples before the addition of nuclear extract.

4.2.4.3 Electrophoresis of DNA-protein complexes

The incubation mixture was loaded on to a 4% or 5% non-reducing polyacrylamide gel in 0.5X TBE. Free DNA and DNA-protein complexes were resolved by electrophoresis at 100 volts at room temperature for about 2 h. After electrophoresis, gels were dried and then exposed to XAR X-ray film at -70°C with an intensifying screen.
### Table 4.1 Oligonucleotides used for electrophoretic mobility shift assay

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence*</th>
<th>Position**</th>
</tr>
</thead>
<tbody>
<tr>
<td>uNF-κB</td>
<td>5'-AACGTCTGGGAGGAGTCCCTGGGGCCACAA-3'</td>
<td>-23</td>
</tr>
<tr>
<td></td>
<td>3'-'TTGCAAGACCTCCCTACGGACCCCCGCTT-5'</td>
<td></td>
</tr>
<tr>
<td>mut-uNF-κB</td>
<td>5'-AACGTCTGGGCACCTCAACGGGGCCACAA-3'</td>
<td>-23</td>
</tr>
<tr>
<td></td>
<td>3'-'TTGCAAGACCTGGAAAGCCACCCCCGCTT-5'</td>
<td></td>
</tr>
<tr>
<td>uPAR-AP-1</td>
<td>5'-ATCTGGGGCTGACTCGCTCTTTCG-3'</td>
<td>-56</td>
</tr>
<tr>
<td></td>
<td>3'-'TAGACCCCGAATCGGAGAAAG-5'</td>
<td></td>
</tr>
<tr>
<td>72A</td>
<td>5'-CCCCGGGAAAGGAGTTGCGG-3'</td>
<td>-117</td>
</tr>
<tr>
<td></td>
<td>3'-'GGGCCCCTTCCCTTACTAACCGCC-5'</td>
<td></td>
</tr>
<tr>
<td>72B</td>
<td>5'-GGAGAGGAGTGCTACGGGAGGAGGGGGAC-3'</td>
<td>-89</td>
</tr>
<tr>
<td></td>
<td>3'-'CCTCCTCAAGAAGCCTCTCCCCCTCC-5'</td>
<td></td>
</tr>
<tr>
<td>72C</td>
<td>5'-GCAGCCACGGAATCCTGGGAGCTACG-3'</td>
<td>-64</td>
</tr>
<tr>
<td></td>
<td>3'-'CCGCGGCTCGTACGACCCCCGCCTGACG-5'</td>
<td></td>
</tr>
<tr>
<td>72D</td>
<td>5'-TTCTGCGGAGAGGAGTGCTACG-3'</td>
<td>-102</td>
</tr>
<tr>
<td></td>
<td>3'-'AAATGGAGGCTCCTCAAGATGCC-5'</td>
<td></td>
</tr>
<tr>
<td>72E</td>
<td>5'-GGAGAGGAGGGTTGACGGGAGGAGG-3'</td>
<td>-77</td>
</tr>
<tr>
<td></td>
<td>3'-'CTCTCCTCCTCCCTGGGAGGTAAG-5'</td>
<td></td>
</tr>
<tr>
<td>NF-κB</td>
<td>5'-AGTTGCAAGGACTTTCCCAGGC-3'</td>
<td>consensus</td>
</tr>
<tr>
<td></td>
<td>3'-'TCAGTCCTCGTAAAGGCTCG-5'</td>
<td></td>
</tr>
<tr>
<td>AP-1</td>
<td>5'-CGCTTGAGTGACTGGCCGCAA-3'</td>
<td>consensus</td>
</tr>
<tr>
<td></td>
<td>3'-'CGAGACTCTGAGGCTCG-5'</td>
<td></td>
</tr>
<tr>
<td>GRE</td>
<td>5'-TCGGACTGACAGGTGGTTAGCTACT-3'</td>
<td>consensus</td>
</tr>
<tr>
<td></td>
<td>3'-'AGGCTGACTGGCTTACAGAAGATCGTGA-5'</td>
<td></td>
</tr>
<tr>
<td>Spl</td>
<td>5'-ATTCGACTGAGGGGGCGCA-3'</td>
<td>consensus</td>
</tr>
<tr>
<td></td>
<td>3'-'TAAGCTAGCCCCGCCCTCG-5'</td>
<td></td>
</tr>
<tr>
<td>TFIID</td>
<td>5'-GCAGAAGCTATAAGGGAGGTAAG-3'</td>
<td>consensus</td>
</tr>
<tr>
<td></td>
<td>3'-'CGTCTCCTACTAACCATCT-5'</td>
<td></td>
</tr>
</tbody>
</table>

* DNA binding sites are underlined; mutant oligonucleotides are in bold.
** Relative to major transcription start site. Consensus oligonucleotides from Promega.
4.3 Results

4.3.1 Identification of two protected regions in the uPAR promoter region

To study the interaction of nuclear proteins with the uPAR promoter sequences, DNase I footprinting assays were performed with a $^{32}$P-labeled 188 bp of promoter fragment between -141 and +47 using nuclear extracts isolated from HCT116 cells. Figure 4.1 shows the footprinting assay results using increasing amounts of the crude nuclear extracts with coding (A) and non-coding strand (B) labeled DNA fragments. As shown in Figure 4.1A, using coding strand uPAR promoter DNA as a probe, two protected regions (I and II) marked with bars were found to interact with nuclear proteins in vitro. The first protected region was located between -51 and -30 and the second region, which is larger than region I, was located between -118 and -79 relative to the transcription start site. Sequence analysis revealed that region I contains an inverted NF-κB site (GAGGAGTCCC), a variant of the Sp1 binding site (GGGAGG) and an IL-6 responsive element (CTGGGA, Kishimoto, et al., 1994), whereas region II contains two variants of Sp1 binding sites (GGGAGG). Similarly, when the DNA fragment spanning -141 to +47 was labeled on the non-coding strand, the two regions were also protected from DNase I digestion (Fig. 4.1B).

4.3.2 Binding of transcription factor NF-κB to the protected region I

To verify that NF-κB transcription factor binds to the inverted NF-κB sequence located at the protected region I, a uPAR oligonucleotide (uNF-κB) representing the uPAR promoter region from -52 to -23 bp, which overlaps with an IL-6 responsive element and a variant site of Sp1, was synthesized. The labeled probe was used in EMSA with HCT116 nuclear extract or a purified NF-κB protein (p49). As shown in Figure 4.2A, incubation of labeled uNF-κB probe with HCT116 nuclear extracts produced two specific DNA-protein complexes, denoted as C1 and C2, and one non-specific shifted band denoted as NS (lane 2). Both C1 and C2 bands are specific because they can be abolished by increasing amounts of unlabeled oligonucleotide uNF-κB (lanes 3,4) and NF-κB consensus oligonucleotide (lanes 7,8), but not by mutant uNF-κB (lanes 5,6) in which the uNF-κB binding site was modified by replacing 7 nucleotides (Figure 4.2C).
Figure 4.1 DNase I footprinting analysis of the human uPAR promoter region. A 188 bp DNA fragment containing the human uPAR promoter (-141 to +47) was 5'-end-labeled using [γ-32P]ATP and T4 polynucleotide kinase, digested with Smal (coding) or Accl (non-coding). Increased amounts of HCT116 nuclear extracts, as indicated, were incubated with the 32P-end-labeled DNA fragment and subjected to DNase I footprint analysis. A, coding strand and B, non-coding strand. The regions of protection from DNase I digestion, indicated by bars, were identified by a dideoxy chain termination sequence ladder separated in parallel on a 6% polyacrylamide sequencing gel. The numbers of nucleotide relative to the transcription initiation site are indicated.
Figure 4.2 Identification of NF-κB binding to the uPAR promoter by EMSA.
A, EMSA performed with [γ-32P]ATP-labeled uPAR-NF-κB (uNF-κB) oligonucleotide (-52 to -23) and nuclear extracts from HCT116 cells. Lane 1, probe alone; lane 2, incubation with nuclear extracts; lanes 3 and 4, unlabeled uNF-κB oligonucleotide was added at 40- and 200-fold molar excess over probe; lanes 5 and 6, 40- and 200-fold molar excess of unlabeled mutant uNF-κB (mut-uNF-κB) competitor; lanes 7 and 8, 40- and 200-fold molar excess of unlabeled NF-κB consensus sequence. B, EMSA performed with [γ-32P]ATP-labeled uNF-κB fragment and NF-κB protein (p49). Lane 1, labeled probe alone; lane 2, incubation with p49; lanes 3-5, incubation with p49 in the presence of 200-fold molar excess of competitors, self, mut-uNF-κB and NF-κB consensus sequence, respectively. C, Sequences of uNF-κB and mut-uNF-κB. NF-κB binding site in the uPAR sequence is underlined. Substituted bases of mut-uNF-κB are indicated by asterisks.
### A

| Extract | - | + | + | + | + | + | + |
| Competitor | - | - | self | mut | NF-xB |
| Fold | 0 | 0 | 40 | 200 | 40 | 200 | 40 | 200 |

### B

| NF-xB (p49) | - | + | + | + |
| Competitor | - | - | self | mut | NF-xB |

### C

**uNF-κB:**

5' - AACGTCCTGGGAGGAGTCCCTGGGGCCACAA - 3'
3' - TTGCAAGACCTCTCTAGGACCCCCGTGGTT - 5'

**mut-uNF-κB:**

5' - AACGTCCTGGGCACTTCACCTGGGGCCACAA - 3'
3' - TTGCAAGACCTCTCTAGGACCCCCGTGGTT - 5'

---

| p49 | + |
| Free | 1 | 2 | 3 | 4 | 5 |

| Free | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
To confirm that NF-κB binds to the binding site located at uNF-κB, NF-κB protein (p49) was used in the EMSA assay. Figure 4.2B shows that incubation of the NF-κB protein with the uNF-κB probe formed one specific DNA-protein complex (lane 2). In the presence of 200-fold molar excess of the unlabeled uNF-κB (lane 3) and NF-κB consensus sequence (lane 5) this formation was abolished. The mutant uNF-κB had significantly reduced ability to compete with NF-κB binding activity (lane 4). These results indicated that NF-κB transcription factor may contribute to the regulation of uPAR expression by binding to its promoter region.

4.3.3 Nuclear proteins interact with protected region II

The 5'-deletion analyses (see chapter 3) demonstrated that the sequence between -141 and -61 is important for uPAR promoter activity in HCT116 cells and DNase I footprinting analysis also showed that several nuclear proteins may bind in the region between -118 and -79. To identify nuclear factors binding to this regulatory region, EMSA was performed using a labeled 72 bp DNA probe spanning -137 to -66. Incubation of the labeled 72 bp probe with the HCT116 nuclear extract produced multiple DNA-protein complexes, designated C1-C5 (Figure 4.3A, lane 2). These complexes represent a specific interaction between the 72 bp fragment and nuclear factors in the extracts since they can be abolished in the presence of competing unlabeled 72 bp fragment (Figure 4.3A, lanes 3-5).

To determine the DNA element responsible for these DNA-protein complexes in detail, a second EMSA analysis was performed for the farther mapping of the 72 bp regulatory element. In this experiment, a series of overlapping oligonucleotide competitor subfragments of the 72 bp element was synthesized (Figure 4.3C) and used in 200-fold molar excess as outlined in the legend of Figure 4.3. A distinct pattern of competition with nuclear protein using these unlabeled competitor fragments is shown in Figure 4.3B. Lane 1 shows the gel shift of the labeled 72 bp fragment and lane 2 exhibits the DNA-protein complexes formed with HCT116 nuclear extract without addition of competitor. Fragment 72A, which is located in the 5' end of the 72 bp fragment, can only compete with the protein in C5 complex (lane 3). Fragment 72B, which corresponds to the region from -118 to -89, competed with proteins in C4 and C5 (lane 4). Fragment 72C spanning from -90 to -64 failed to compete with proteins in
all five DNA-protein complexes (lane 5). Addition of fragment 72D, which spans from -126 to -102, showed partial competition with C1 and C5 DNA-protein complexes with the occurrence of an extra band on the top of C1. 72E, a 25 bp fragment from -101 to -77, was able to successfully compete with C1, C4 and C5 complexes and partially compete with C2 (lane 7). Like 72 D, a extra top band was also observed in the presence of the 72E. To verify that Sp1 transcription factor is indeed a component of the binding activity in nuclear extracts, the Sp1 consensus sequence was also used in this competition experiment. Like the 72E fragment, a similar competitive pattern was obtained after addition of Sp1 consensus competitor (lane 8). Sp1 oligonucleotide successfully competed with proteins in all major DNA-protein complexes except for C3. Unlike 72D and 72E, no extra shift band on the top was observed in the presence of the Sp1 consensus sequence. These results clearly suggest that Sp1 or Sp1-related nuclear factor(s) is a major nuclear protein responsible for interaction of nuclear factors with the 72 bp regulatory region. C2 may represent modified forms of Sp1 and/or Sp1-related proteins since Sp1 is known to be post-translationally modified (Jackson et al., 1990), while C3-C5 may represent the complexes formed between the DNA probe and the degraded Sp1 proteins (Chen et al., 1993; Zhang et al., 1994).

To determine the role of Sp1 transcription factor in the interaction of nuclear proteins with 72 bp region, a mutual competition experiment was performed with labeled oligonucleotides, Sp1 and 72E (Figure 4.4). Incubation of labeled 72E or Sp1 probe with nuclear extracts isolated from HCT116 showed a similar pattern of DNA-protein complexes to Figure 4.3 (Figure 4.4A and B). Using 72E as a probe, five DNA-protein complexes were observed (Figure 4.4A). Addition of 200-fold molar excess of the unlabeled 72E or Sp1 completely abolished the DNA-protein complexes C1-C4. In addition, a strong inhibition of the formation of complexes C1-C4 was observed in the presence of 72D and partial inhibition in the presence of 72B (Figure 4.4A), whereas 72A and 72C did not compete with the formation of these DNA-protein complexes. Using Sp1 consensus oligonucleotide as a probe, four DNA-complexes (C1-C4) were observed (Figure 4.4B). The top three DNA-protein complexes (C1-C3) can be completely competed away by both Sp1 self and 72E. In similarity to 72E, 72D also showed strong competition with DNA-protein complexes, C1 and C2. 72B also showed partial competition with this DNA binding. These data suggest that 72E, 72D and 72B have different ability to compete for Sp1 binding. Although both 72E and 72B
contain GGGAGG sequence, 72E showed much stronger competition ability than 72B. 72D may also contain a potential Sp1 binding site which is responsible for its competition ability. Taken altogether, these results clearly demonstrate that Sp1 or Sp1-related nuclear factor(s) is indeed involved in binding to the 72 bp fragments in the protected region II which is required for uPAR promoter activity.

4.3.4 Interaction of proteins with the 55 bp regulatory region (-8 to +47)

Since the 3'-deletion analysis to -10 in the promoter region reduced the CAT activity by >90%, it seemed that the sequences between -9 and +47 are important for the high level CAT expression in HCT116 cells. To identify nuclear factors binding to this region, EMSA was performed using the end-labeled 55 bp fragment (-8 to +47) as a probe. As shown in Figure 4.5, incubation of the 55 bp probe with the HCT116 nuclear extracts resulted in the formation of one prominent DNA-protein complex (C2 in Figure 4.5A, lane 2) and three less prominent complexes (C1, C3 and C4, lane 2). Specificity for the 55 bp sequence was demonstrated by competition experiments in the presence of either unlabeled 55 bp fragment or unlabeled oligonucleotides corresponding to the binding sites for two transcription factors, TFIID and AP-1. The results (Figure 4.5A) indicate that only the unlabeled 55 bp fragment was able to compete with radioactive C1-C3 DNA-protein complex formation (lanes 3-5), thereby demonstrating specificity of these complexes for the 55 bp fragment. In addition, C2 and C3 complexes were partially competed by TFIID (lane 6), but not by AP-1 (lane 7). C4 DNA-protein complex appears to be a non-specific DNA binding since it was not competed by unlabeled 55 bp fragment. In addition to an Inr element (Soravia et al., 1995), sequence analysis of this 55 bp region also revealed a downstream promoter element (DPE) located at position +11 to +17, which was recently found in Drosophila genes (Burke and Kadonaga, 1996). Since the unlabeled TFIID oligonucleotide was also able to partially compete with formation of C2 and C3 complexes, these results suggest that the TFIID or TFIID-related protein may be involved in the interaction of nuclear proteins with the 55 bp region.
Figure 4.3 EMSA analysis with ³²P-labeled 72 bp fragment (-137 to -66) of the human uPAR gene and nuclear extracts from HCT116 cells. A 72 bp fragment was end-labeled with [α-³²P]dCTP and Klenow fragment of DNA polymerase I, and incubated with 4 µg of nuclear extracts from HCT116 cells in the absence or presence of competitors. A, competition analysis with unlabeled 72 bp fragment. Lane 1, probe only; lane 2, incubation with nuclear extracts; lanes 3-5, incubation with nuclear extracts in the presence of 50-, 100- and 200-fold molar excess of unlabeled 72 bp fragment, respectively. B, competition analysis with oligonucleotides 72A-72E and Sp1. Lane 1, probe only; lane 2, incubation with nuclear extracts; lanes 3-7, incubation with nuclear extracts in the presence of 200-fold molar excess of 72A, 72B, 72C, 72D, and 72E, respectively; lane 8, incubation with nuclear extracts in the presence of 200-fold molar excess of Sp1 consensus oligonucleotide (5'-ATTCGATCGGGCGGGCGAGC-3'). C, sequences of the 72 bp fragment and oligonucleotides. The Sp1 binding sites in the 72 bp fragment are underlined.
### A

<table>
<thead>
<tr>
<th>Extract</th>
<th>+ + + +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fold</td>
<td>0 50 100 200</td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th>Extract</th>
<th>+ + + + + + +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Competitor</td>
<td>. . 72A 72B 72C 72D 72E Sp</td>
</tr>
</tbody>
</table>

### C

```plaintext
GAGAGGGGGGAGGGCCCCAGCCTC 72E
TTTGTGCGAGGAGTCTGACGG 72D
GGGCGCCACGCATCTGGGCTGACTCG 72C
GAGAGGCTCGACGGGAGGGAGG 72B
CCCGGGAGGGAAAGTTCTGGGGG 72A
```

<table>
<thead>
<tr>
<th></th>
<th>-137</th>
<th>-127</th>
<th>-117</th>
<th>-107</th>
<th>-97</th>
<th>-87</th>
<th>-77</th>
<th>-67</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CCCC</td>
<td>GAGG</td>
<td>CAGT</td>
<td>CAGG</td>
<td>AAGC</td>
<td>CAGT</td>
<td>CAGG</td>
<td>AAGC</td>
</tr>
</tbody>
</table>
Figure 4.4 Identification of Sp1 binding to the uPAR promoter by EMSA. A, EMSA with $^{32}$P-labeled 72E oligonucleotide. Lane 1, 72E probe alone; lane 2, incubation with 4 µg of nuclear extracts from HCT116 cells; lanes 3-8, incubation with nuclear extracts in the presence of 200-fold molar excess of the non-radioactive oligonucleotides, 72A, 72B, 72C, 72D, 72E and Sp1. B, EMSA with $^{32}$P-labeled Sp1 consensus oligonucleotide. Lane 1, Sp1 probe alone; lane 2, incubation with 4 µg of nuclear extracts; lanes 3-8, incubation with nuclear extracts in the presence of 200-fold molar excess of the non-radioactive oligonucleotides, Sp1, 72A, 72B, 72C, 72D and 72E.
A: 72E

| Extract |  - + + + + + + |
| Competitor |  - 72A 72B 72C 72D 72E Sp1 |

| C1 | C2 | C3 | C4 | C5 |
| Free | 1 2 3 4 5 6 7 8 |

B: Sp1

| Extract |  - + + + + + + |
| Competitor |  - Sp1 72A 72B 72C 72D 72E |

| C1 | C2 | C3 | C4 |
| Free | 1 2 3 4 5 6 7 8 |
Figure 4.5 EMSA analysis with $^{32}$P-labeled 55 bp fragment (-8 to +47) of the human uPAR gene and nuclear extracts from HCT116 cells. The labeled 55 bp fragment was incubated with nuclear extracts from HCT116 cells in the absence or presence of competitors. A, Competition analysis with unlabeled 55 bp fragment, TFIID or AP-1 oligonucleotides. Lane 1, probe only; lane 2, incubation with nuclear extracts; lanes 3-5, incubation with nuclear extracts in the presence of 50-, 100- and 200-fold molar excess of unlabeled 55 bp fragment, respectively; lanes 6 and 7, incubation with nuclear extracts in the presence of 200-fold molar excess of TFIID consensus sequence (5'-GCAGAGCATATAAGGTGAGGTAGGA-3') and AP-1 consensus sequence (5'-CGCTTGATGAGTCAGCCGGA-3'), respectively. B, Sequence of the 55 bp fragment. The DPE motif is underlined.
### A

| Extract | + | + | + | + | + | + | + |
| Competitor | - | - | self | TFIID AP-1 |
| Fold | 0 | 0 | 50 | 100 | 200 | 200 | 200 |

![ gel electrophoresis with lanes labeled C1, C2, C3, C4, Free with corresponding bands for each lane ]

### B

-8  
+1  
+11  
+21  
+31  
+41  

TGGGCCAGAGGAAGAGAAGAGTGCGAGGACCCGCGACAGGAGGCTGCCCTCGACTCCGGTCTTCGCTTCGTCCGGGAGCCCTCG

ACTCCGGTCATTCTCTTCTTCTGCACGTCCCTGGGGCGGCGTGTCCTCGACGGGAGC
4.3.5 Binding of nuclear protein to the uPAR AP-1 element in the uPAR promoter region

AP-1 transcription factor has been shown to be involved in the activation of many genes (Wang and Goldstein, 1993). Previous studies of cotransfection with c-jun/c-fos expression plasmids demonstrated that overexpression of AP-1 significantly increased the promoter activity of the human uPAR gene in HCT116 cells (see chapter 3). To test whether AP-1 binds to the uPAR promoter, the interaction of nuclear factors with the potential uPAR AP-1 binding sites was studied using EMSA. As shown in Figure 4.6, using a labeled uPAR AP-1 oligonucleotide (from -79 to -56 bp) as a probe, a major shifted band was observed after incubation with nuclear extracts from HCT116 cells. The specificity of this DNA-protein complex was examined by addition of 40 or 200-fold molar excess competitors. It can be competed away by itself and AP-1 consensus sequence, but not by unrelated GRE consensus sequence. It was noted that two extra shifted bands with lower and faster shift mobility occur in the presence of AP-1 consensus competitor. These results suggest that AP-1 may regulate uPAR transcription by interaction with the uPAR promoter.

4.4 Discussion

In this chapter, interactions of multiple nuclear factors with uPAR promoter sequences were examined by using DNA footprinting and EMSA analyses. Within the 188 bp SmaI-NruI fragment of the uPAR promoter, two protected regions have been localized by DNA footprinting analysis. The first region is from -51 to -30 that contains one NF-κB binding site and the second region is from -118 to -79 that contains two variants of Sp1 consensus sequence. Binding of nuclear proteins to these DNA sequences was further identified by EMSA suggesting that these DNA elements may be involved in the regulation of uPAR gene expression.

Binding activity of nuclear factors to the uPAR Sp1 binding sites within the region from -141 to -61 was demonstrated by DNase I footprinting and EMSA analysis suggesting that transcription factor Sp1 or Sp1-related protein is a good candidate for controlling uPAR expression. Firstly, repeated DNase I footprinting analysis using HCT116 cell nuclear extracts produced consistent protection patterns over putative Sp1 sites. Secondly, EMSA showed that multiple DNA-protein complexes formed with
Figure 4.6  EMSA analysis of uPAR-AP-1 binding site in the uPAR promoter region. A, EMSA performed with uPAR-AP-1 probe. Double-stranded uPAR-AP-1 oligonucleotide (-79 to -56) was labeled with T4 polynucleotide kinase and [γ-32P]ATP and incubated with nuclear extracts from HCT116 cells. Lane 1, probe only; lane 2, incubation with 4 µg of nuclear extract from HCT116 cells; lanes 3 and 4, unlabeled double stranded uPAR-AP-1 competitor was added at 40- or 200-fold molar excess over probe; lanes 5 and 6, 40- and 200-fold molar excess of unlabeled AP-1 consensus oligonucleotide; lanes 7 and 8, 40- and 200-fold molar excess of unlabeled GRE consensus oligonucleotide. B, EMSA analysis with the labeled AP-1 consensus oligonucleotide. Lane 1, probe only; lane 2, incubation with nuclear extract; lanes 3 and 4, incubation with nuclear extracts in the presence of 40- and 200-fold molar excess of unlabeled uPAR-AP-1 oligonucleotide. C, Sequences of uPAR-AP-1 and AP-1 consensus oligonucleotides. AP-1/GRE and AP-1 binding sites are underlined.
A: uPAR

| Extract | - | + | + | + | + | + | + |
| Competitor | - | - | self | AP-1 | GRE |
| Fold | 0 | 0 | 40 | 200 | 40 | 200 | 40 | 200 |

AP-1 ➔
Free

B: AP-1

| Extract | - | + | + |
| Competitor | - | - | uPAR |
| Fold | 0 | 0 | 40 | 200 |

AP-1 ➔
Free

C:

uPAR:

AP-1/GRE
5' -ATCTGGGGCTGACTCGCTCTTTCG-3' 3' -TAGACCCCGACTGAGCGAGAAAGC-5'
5' -CGCTTGATGAGTCAGCCGGAA-3' 3' -GCGAACTACTTCGAGGCCTTT-5'

AP-1:
the 72 bp fragment were abolished in the presence of homologous competitor and Sp1 consensus sequence, but not by the competitors without Sp1 site, suggesting that these DNA-protein complexes were specific. Combined with results of transcription analysis which showed significant reduction in promoter activity with progressive deletion of the region from -141 to -61 (see chapter 3), these data clearly demonstrate that the Sp1 or Sp1-related transcription factor plays an important role in uPAR gene expression.

Sp1, a zinc-finger-containing transcription factor, was originally identified in HeLa cells (Dynan & Tjian, 1983). It binds to the GGGCGG sequences (GC boxes) in the 21 bp repeat of the SV40 early promoter and activates transcription from the promoter in vitro. Subsequently, Sp1 binding motifs are also found in the TATA-less and GC-rich promoters as oligomeric forms to activate transcription (Briggs et al., 1986; Pugh and Tjian, 1990). It has been reported to function preferentially through an Inr element and Sp1-Inr interaction is important for transcription initiation (Colgan and Manley, 1995).

In a TATA-less promoter, Sp1 has been shown to be required for transcription reconstituted with fractionated general factors (Pugh and Tjian, 1990). Two Sp1 sites are located at the positions -103 and -95 within the regulatory region between -141 and -61. Because the Sp1 site (GGGAGG) is a variant of the canonical site (GGGCGG) and has been reported to have a lower binding affinity for Sp1 (Letovsky and Dynan, 1989), binding affinity of the uPAR Sp1 sites was examined by competition analysis using Sp1 consensus sequence as a probe (Figure 4.3B). Fragment 72B, which contains both Sp1 sites in which the second site is located at 3’ end of the fragment, showed a much weaker competition ability. However, fragment 72E, which contains only one Sp1 site at -95, had very strong competition ability which is similar to Sp1 consensus sequence. Fragment 72D also showed a strong competition ability for Sp1 binding, suggesting a potential Sp1 binding site located in this region. These results suggest that the Sp1 site within the 72E has much stronger binding affinity for Sp1 than those in 72B and 72D.

Recently, Soravia et al. (1995) have also demonstrated that Sp1 activates uPAR expression by cotransfecting the uPAR promoter construct and an Sp1 expression plasmid into Drosophila Schneider cells that lack endogenous Sp1. They also presented evidence for the Sp1 transcription factor binding to the Sp1 sites of the human uPAR promoter. Combined with our results that uPAR promoter activity was completely lost
after deletion of the Sp1 sites between -141 and -60, these data clearly indicate that interaction of Sp1 or Sp1-related transcription factors with this regulatory region plays a significant role in the transcriptional regulation of the human uPAR gene.

Nuclear factor κB (NF-κB) is a pleiotropic transcription factor which regulates expression of a large number of cellular and viral genes (Siebenlist et al., 1994; Thanos and Maniatis, 1995). The data presented here demonstrate that the transcription factor NF-κB can bind to the inverted NF-κB motif in the promoter region of the uPAR gene. This specific binding was examined by DNA footprinting and EMSA assays. Two specific complexes C1 and C2 were formed on the NF-κB motif in the uPAR promoter with nuclear extracts prepared from HCT116 cells. Binding of NF-κB to this inverted NF-κB site in the uPAR promoter region was further confirmed using purified NF-κB protein, p49.

It has been shown that NF-κB is activated by cytokines, phorbol esters, and viral trans-activators (Grimm and Baeuerle, 1993; Siebenlist et al., 1994). Like c-fos and c-jun, NF-κB is a immediate-early gene. Upon stimulation, NF-κB is rapidly activated and is able to dissociate immediately from 1κB and translocate into the nucleus. NF-κB is rapidly induced by TNFα in both proliferating and arrested cells (Duckett, et al., 1995). uPAR expression is induced by various cytokines, phorbol esters and growth factors, including TNFα and PMA (Langer et al., 1993; Wang et al., 1994; Lund et al., 1995). Binding of NF-κB to the uPAR promoter, therefore, may be involved in activation of the uPAR gene.

EMSA was used to identify proteins in HCT116 cells that could bind to the AP-1 motif. Using an oligonucleotide synthesized according to the uPAR promoter sequence as a probe (uPAR-AP-1), a major complex formed with nuclear extracts was detected. Because this potential AP-1 binding site in the uPAR promoter region overlaps a glucocorticoid receptor element (GRE), the specificity of the DNA-protein complex formed with nuclear extracts was examined. The major complex can specifically be abolished in competition with itself and the AP-1 consensus sequence but not by GRE consensus sequence suggesting that this DNA-protein complex is specific.

AP-1 is a family of dimers composed of the protein products from two gene families, jun and fos (Distel and Spiegelman, 1990; Vogt and Bos, 1990). The Jun/Fos heterodimer is more stable and has a higher affinity for AP-1 sites than does the Jun
homodimer. Mammalian Fos is unable to homodimerize or bind DNA. AP-1 was initially identified in HeLa cells as a factor that activates transcription in a site-specific manner (Lee et al., 1987). It mediates cellular responses to a diverse array of extracellular stimuli such as PMA, growth factors and oncogenes and has been implicated as part of the protein kinase C signal transduction pathway. The AP-1 transcription factor modulates gene transcription by its binding to a DNA sequence TGAC/GTCA in the promoter/enhancer region of many different genes (Angel and Karin, 1991). Previous studies have shown that cotransfection of uPAR promoter constructs with c-fos and c-jun expression plasmids increases the promoter activity up to 4 to 5-fold (see chapter 3). Therefore, AP-1 transcription factor binding to uPAR promoter may be involved in the upregulation of the uPAR gene expression in response to various stimuli that activate AP-1 activity.

A number of DNA-protein complexes was found after incubation of HCT116 nuclear extracts with the 55 bp regulatory region (-8 to +47) which is important for the uPAR promoter activity as shown in transcription analysis. Since this region contains the transcription initiation sites and most 5'UTR sequences of the uPAR gene, it was speculated that the DNA-binding activity observed in EMSA may represent the binding of a more general transcription factor to the uPAR promoter. It is interesting that the TFIID consensus sequence can partially compete with some of these DNA-protein complexes formed within this region. This preliminary experiment suggested that TFIID or TFIID-related proteins may be involved in the interaction between this region and transcription factors. TFIID is a general transcription factor which is necessary for RNA polymerase II-dependent transcription. Sequence analysis of the uPAR gene has revealed a consensus Inr element sequence of the dihydrofolate reductase (DHFR) gene located at -12 to +1 (Soravia et al., 1995). In TATA-less promoters, the Inr element plays an essential role in transcription complex formation and TFIID is required for Inr element-directed transcription (Weis and Reinberg, 1992). In addition, Inr element can direct TFIID binding to the downstream promoter sequences (Wang and Dyke, 1993). Most recently, a downstream promoter element (DPE, A/GGA/TCGTG) which is usually located about 30 bp downstream of the transcription initiation site in Drosophila genes has been reported (Burke and Kadonaga, 1996). The DPE sequence has been shown to mediate transcription of TATA-less promoters by acting in conjunction with the initiation site sequence to provide a binding site for TFIID in the
absence of a TATA box. Although the uPAR gene lacks a TATA box, a conserved
DPE sequence (+11 AGACGTG +17) located downstream of the initiation site and an
Inr element were found. It has been shown that TFIID can bind TATA box, Inr and
DPE elements (Goodrich et al., 1996; Burke and Kadonaga, 1996). If this is the case,
the interaction of TFIID with the Inr element in the uPAR promoter region will be
essential to uPAR transcription. It is of interest to examine further the interaction of
nuclear factors with this region.

The data presented in this chapter demonstrate that the presence of several
transcription factor binding sites in the human uPAR promoter offers a potential
mechanism by which trans-acting factors might cause activation of the promoter.
Combined with the data from transcription assays, these results indicate that the human
uPAR promoter is under complex regulation. While Sp1 or Sp1-like protein binds to
the promoter region for basal promoter activity, interaction of AP-1 and NF-κB with
the promoter region may be responsible for activation of the uPAR gene. In addition,
several nuclear proteins that bind to the region between -9 and +47 may also be
involved in uPAR transcription. Several studies have demonstrated that the expression
of uPAR can be regulated by cytokines, tumor promoter and growth factors (Blasi,
1993). The finding described in this chapter suggests that this regulation of human
uPAR transcription may be conferred by DNA sequences that lie within the promoter
region. Both AP-1 and NF-κB binding sites are thought to be the common elements
that mediate the induction of gene expression (Angel and Karin, 1991; Siebenlist, et al.,
1994). The effects of PMA and TNFα on gene expression are known to be mediated
by activation of the transcription factors AP-1 and NF-κB (Singh and Aggarwal, 1995;
Siebenlist, et al., 1994). Therefore, there is a possibility that these stimuli upregulate
uPAR expression through these DNA binding sites.

Gene regulation in eukaryotes is achieved by the combined action of multiple
regulatory proteins and transcription factors that interact with each other on complex
arrays of cis-regulatory elements in promoters and enhancers/silencers (Johnson and
McKnight, 1989). Binding of multiple transcription factors to the human uPAR
promoter region suggests that the regulation of uPAR transcription may depend in the
interplay of these diverse transcription factors. A cooperative interaction amongst these
transcription factors may be essential for regulated uPAR transcription in vivo.
CHAPTER 5
SODIUM BUTYRATE INHIBITS UROKINASE RECEPTOR
EXPRESSION AT BOTH TRANSCRIPTION AND POST-
TRANSCRIPTION LEVELS

5.1 Introduction

The aim of the project presented in this chapter was to investigate the effects of sodium butyrate on uPAR gene expression in human colon cancer cells stimulated and unstimulated with PMA or TNFα.

Butyrate is a short chain fatty acid which is produced in the lumen of the human colon by bacterial fermentation of dietary fibre (Clausen et al., 1991). At concentrations present in the lumen it inhibits cell proliferation and stimulates cell differentiation. In colorectal cancer, butyrate is of particular interest because it is representative of short-chain fatty acids that are natural effectors of colonic cell differentiation in \textit{in vivo} (Cummings et al., 1987) as well as \textit{in vitro} (Prasad, 1980). Butyrate also induces apoptosis of colonic carcinoma cells in culture, probably by a p53-independent pathway (Hague et al., 1993 and 1995; Filippovich et al, 1994). It has been reported to increase c-jun and decrease c-myc gene expression in colon cancer cells (Nishina et al., 1993; Souleimani et al., 1993a and 1993b; Heruth et al., 1993). Butyrate-responsive elements have been found in the 5' flanking regions of c-fos (Souleimani et al., 1993a) and mouse calbindin-D_{28k} genes (Gill et al., 1993). The effects of butyrate on the genes implicated in the cell surface proteases that may determine adenocarcinoma invasion and metastasis, however, are unknown.

As reviewed in chapter 1, proteolysis mediated by uPA and uPAR may be central to the processes of cell migration under normal and pathological conditions that include cancer invasion and metastasis. Receptor-bound uPA activates the proenzyme plasminogen to form plasmin which initiates a proteolytic cascade that contributes to the degradation of basement membranes and the extracellular matrix (Blasi, 1993). The activity of uPA on the cell surface is regulated by the level of uPAR and uPA inhibitors in the tumor cell microenvironment. Thus, the uPA/uPAR system controls cell migration and uPAR expression is a requirement for tumor invasion. In colon cancer, receptor-
bound uPA activity is focused at the invasive front (Pyke et al., 1991; Wang et al., 1994). High levels of uPA activity and uPAR gene expression in human cancers correlate with risk of recurrence and increasing evidence indicates that invasion and metastasis by adenocarcinoma cells can be prevented by inhibiting uPA and uPAR production (Blasi, 1993; Fazioli and Blasi, 1994). Thus, uPA/uPAR antagonists and inhibitors are viewed as potent antimetastatic agents. Given the pivotal role played by uPAR in colon cancer tumor metastasis, it is of importance to study the effect of dietary factors such as butyrate on uPAR expression.

Tumor necrosis factor alpha (TNFα) is a pro-inflammatory cytokine that has cytocidal activity against a number of cancer cell lines and induces haemorrhagic necrosis of cancers in mice (Komori et al., 1993). Phorbol 12-myristate 13-acetate (PMA) is one of the most potent tumor-promoting compounds and PMA-induced gene expression is among the events required for tumor promotion (Mayo et al., 1994). Both TNFα and PMA increase uPAR and uPA mRNAs in human colon cancer cells (Wang et al., 1994 and 1995), but whether butyrate inhibits the increased uPA and uPAR expression induced by TNFα or PMA is unclear. Similarly, whether butyrate regulates uPA and uPAR mRNA at the levels of transcription or post-transcription has received little attention.

To address the issue of transcriptional regulation of uPAR in tumor cells, two human colon cancer cell lines, HCT116 and LIM1215, were used as models. Both cell lines have been characterized in detail in this laboratory with respect to high invasiveness and overexpression of both uPA and uPAR on stimulation by PMA and TNFα (Wang et al., 1994). This chapter reports that sodium butyrate inhibits uPA and uPAR mRNA expression in a dose-dependent manner in colon cancer cell lines regardless of whether they have been stimulated with TNFα, PMA or cycloheximide. In the presence or absence of TNFα, butyrate acts, at least in part, at the transcriptional level on uPA gene expression, but acts at both the transcriptional and post-transcriptional levels on the expression of the uPAR gene. In the presence of PMA, however, butyrate mainly acts at the post-transcriptional level on the expression of both genes.
5.2 Materials and Methods

5.2.1 Materials

A recombinant human TNFα (activity: 1 unit ~ 0.0455 ng) was provided by the Asahi Chemical Company, Tokyo, Japan. Phorbol 12-myristate 13-acetate (PMA) and cycloheximide (CHX) were from Sigma Chem. Co., St. Louis, USA. Reinforced nitrocellular membrane was from Schleicher & Schuell. [α-32P]-dCTP and [α-32P]-UTP were from Amersham Ltd, UK. RQ1 RNase-Free DNase, Proteinase K and Random Priming kit were obtained from Promega, USA. The human uPA cDNA (pcUKLTR6) was provided by Dr. R. Miskin, the Weizmann Institute of Science, Israel (Miskin et al., 1989). Human 18S ribosomal DNA (18S rDNA) was provided by Dr. B. E. H. Maden, University of Liverpool, UK (McCalum and Maden, 1985). Human β-actin cDNA (Ponte et al., 1984) was a gift from Dr. H. D. Campbell, The Australian National University.

5.2.2 Cell culture and stimulation

The HCT116 and LIM1215 human colon cancer cells were maintained at 37°C, and 5% CO₂ in RPMI-1640 medium supplemented with 10% fetal calf serum, 2 mM glutamine, 50 µg/ml gentamycin and 50 µg/ml penicillin.

The following reagents were used in the experiments: sodium butyrate prepared in distilled water, PMA in DMSO, cycloheximide and rhTNFα in RPMI-1640. For Northern blot and nuclear run-on transcription analyses, cells were cultured in FCS-free medium for 4 h before stimulation. Then cells were treated with stimuli for 4 h in FCS-free medium. The viability of HCT116 and LIM1215 cells after stimulation was greater than 95% in all experiments.

For the uPAR binding assay, HCT116 colon cancer cells were cultured in 24-well plates in RPMI complete medium to ~90% confluency. The cell monolayers were rinsed with FCS-free RPMI-1640 medium once and cultured in this FCS-free medium for 4 h before stimulation. Cells were then treated with butyrate, PMA or PMA plus butyrate for 16 h.
5.2.3 cDNA probes

The cDNA probes were labeled with [α-32P]-dCTP using Quick Primer Kit and purified by centrifugation through a Nick column as described in section 2.2.2.5. The following cDNA probes were used: human uPAR cDNA, human uPA cDNA, β-actin cDNA and the human 18S ribosomal DNA (18S rDNA).

5.2.4 RNA preparation and Northern blot analysis

RNA was prepared from cells according to the method described in section 2.2.11. All water used in RNA work was pre-treated with DEPC as described in Sambrook et al. (1989).

5.2.4.1 Preparation of formaldehyde-agarose denaturing gels

Electrophoresis of RNA on formaldehyde-agarose gels was carried out in a Pharmacia GNA-200 electrophoresis apparatus (20 cm x 20 cm tray with a 350 ml gel volume). A 1.2% agarose gel was made by dissolving 4.2 g of agarose in 304.5 ml of DEPC-treated H₂O in a microwave oven, then cooling this solution to 60°C before adding 35 ml of 10X RNA running buffer (1X MOPS: 20 mM MOPS, pH 7.0, 5 mM sodium acetate, 10 mM EDTA) and 10.5 ml of 37% formaldehyde. The gel solution was then poured onto the gel mould and allowed to set at room temperature for 1 h.

5.2.4.2 Preparation of RNA samples for electrophoresis

10 µg of total RNA in 1X MOPS, 6.29% formaldehyde and 50% formamide was denatured by heating 15 min at 55°C. Samples were then placed on ice and 2 µl of RNA loading solution (50% glycerol, 1 mM EDTA, pH 8.0, 0.25% bromophenol blue, 0.25% xylene cyanol) was added just prior to loading on the formaldehyde-agarose gel.

5.2.4.3 Electrophoresis of RNA

The formaldehyde-containing agarose gel was set up in the appropriate electrophoresis apparatus (GNA-200, Pharmacia) and electrophoresis buffer was added to a level that did not permit buffer to enter the wells. RNA samples were loaded and electrophoresed into the gel at 150 V for 5 min before topping up the buffer level to cover the gel. RNA samples were electrophoresed at a constant voltage of 100 V for 2-3 h.
5.2.4.4 Blotting of RNA to reinforced nitrocellulose

After electrophoresis the gel was rinsed with distilled water and then submerged in 10X SSC for 30 min. Then RNA samples were transferred to reinforced nitrocellulose membranes with 20X SSC overnight. Then the membranes were air-dried on Whatman 3MM paper for ~20 min, sandwiched between two sheets of Whatman paper and baked in a vacuum oven at 80°C for 2 h.

5.2.4.5 RNA hybridization

RNA hybridization was performed according to Ausubel et al. (1989). Membranes were prehybridized in 20 ml of hybridization solution (5X SSC, 5X Denhardt's, 50 mM NaPO₄, pH 6.7 and 10 mg/ml Salmon sperm DNA and 50% formamide) in a cylinder (Hybaid) in a Hybaid oven at 42°C for 6-16 h and then hybridized with a ^32^P-labeled human uPAR cDNA or indicated cDNA probes at 42°C overnight. The membranes were first washed in 500 ml of 2X SSC, 0.1% SDS at room temperature for 20 min and then in 500 ml of 0.1X SSC, 0.1% SDS at 60°C for 30 min. Hybridization was analyzed by autoradiography. The membranes were exposed to Kodak XAR film and multiple film exposure times were used to ensure linearity of band intensities.

Each Northern blot was in turn rehybridized with the uPA cDNA and a human 18S rRNA probes. The intensities of mRNA bands in the autoradiographs were scanned and quantitated by a video densitometer (Model 620, BioRad). The mRNA signal intensity is expressed relative to the level of 18S rRNA in each sample to normalize for differences in RNA concentration.

5.2.5 uPAR binding assay

The ^125^I-uPA binding assay was performed according to Boyd and Brattain (1989). After stimulation, the cell monolayers were treated by adding 0.4 ml of acidic glycine buffer (50 mM glycine/HCl, pH 3.0, 0.1 M NaCl) at 4°C for 3 min and neutralized by adding 0.1 ml of HEPES buffer (0.5M HEPES pH 7.5, 0.1M NaCl) for 1 min. To determine the density of unoccupied uPARs on the cell surface, the acid elution step was omitted. After another two washes with binding buffer (RPMI-1640, 0.1% BSA, 20 mM HEPES, pH 7.4), the monolayer was incubated with ^125^I radiolabeled high molecular weight (HMW) human uPA (Actosolv) in 0.5 ml of binding buffer at room temperature for 45 min. To determine the extent of nonspecific binding, a 50-fold excess of cold uPA was added. Unbound uPA was removed by rinsing monolayers with
binding buffer three times and the cells were then completely lysed in 0.5 ml of 0.1 M NaOH at 37°C for 2 h. The radioactivity of the cell lysates was measured using a Packard Gamma Counter (SUTO-GAMMAR 5650, Packard Instrument Company Inc., IL., USA).

5.2.6 Nuclear run-on transcription analysis

Nuclear run-on transcription analysis was performed according to the method previously described (Linial et al., 1985; Mahajan and Thompson, 1987).

5.2.6.1 Plasmid DNA blotting

For nuclear run-on transcription analysis the following plasmid cDNAs and plasmid were used: the human uPAR cDNA, human uPA cDNA, human β-actin cDNA and plasmid pGEM. All plasmid DNAs (50 µg) were linearized by digestion with EcoRI restriction enzyme, denatured in 0.3 M NaOH at 65°C for 30 min followed by neutralization with an equal volume of 2 M ammonium acetate (pH 7.0) on ice for 15 min and then 5 ml of 6X SSC was added. The linearized plasmid DNAs were blotted onto reinforced nitrocellulose membranes by adding 5 µg of plasmid DNA into each slot using a Schleicher & Schuell Slot Blot apparatus according to the manufacturer’s description. Following washing with 6X SSC, the filter was air-dried and baked at 80°C for 2 h in a vacuum oven.

5.2.6.2 Nuclei preparation

After 4 h stimulation with TNFα, PMA or butyrate as described above, approx. 5 x 10^7 cells were washed twice with cold PBS and lysed in 4 ml of NP-40 lysis buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.5% NP-40). Nuclei were then pelleted by centrifugation at 600 x g for 3 min. After washing with 1 ml of cold nuclear storage buffer (50 mM Tris-HCl, pH 7.5, 40% glycerol, 5 mM MgCl₂, 0.1 mM EDTA), the nuclei was pelleted by centrifugation and then resuspended in 200 µl of the same buffer for use.

5.2.6.3 Transcription elongation reaction

For the transcription elongation reaction, each reaction was carried out in a total volume of 300 µl containing 210 µl of freshly prepared nuclei and 60 µl of 5X run-on buffer (25 mM Tris, pH 8.0, 12.5 mM MgCl₂, 750 mM KCl, 1.25 mM triphosphate of A, G, and C). 10 µl of α-32P-labeled uridine triphosphate (UTP, 100 µCi, 3,000
Ci/mM) was then added. The nuclei suspension was incubated at 30°C for 30 min. After reaction, 35 µl of RQ1 RNase-Free DNase and 10 mM CaCl₂ were added and incubated at 30°C for 15 min. The reaction was then stopped by adding 100 µl of 5X SET (5% SDS, 25 mM EDTA, 50 mM Tris, pH 7.4). Proteinase K was added to a concentration of 200 µg/ml and incubation continued at 37°C for 45 min. After extraction twice with phenol and chloroform, the RNA was precipitated by adding 10 M ammonium acetate to a final concentration of 2.3 M and an equal volume of isopropyl alcohol. The precipitate was centrifuged and the pellet was resuspended in 100 µl of TE buffer and purified by centrifuging through a spin column (Pharmacia). The eluate was treated with 0.2 M NaOH for 10 min on ice and then HEPES was added to a concentration of 0.24 M. The ³²P-labeled RNA was precipitated with ethanol and redissolved in hybridization buffer (10 mM Tris-HCl, pH 7.4, 0.2% SDS, 10 mM EDTA, 0.3 M NaCl, 1X Denhardt’s, and 250 µg/ml yeast tRNA) for use.

5.2.6.4 Hybridization and autoradiography
Following prehybridization at 65°C for 4 h, each filter was hybridized with the same amount of ³²P-labeled RNA (~1 X 10⁷ cpm) in 2 ml of hybridization solution at 65°C for 36 h. After hybridization, filters were washed for 30 min in 2X SSC-0.1% SDS at 60°C and then incubated at 37°C in 2X SSC with RNase A (10 µg/ml) for 30 min and were subsequently washed in 0.1X SSC-0.1% SDS at 60°C for 30 min. The filters were then exposed to XAR film at -70°C for 7-14 days using intensifying screens. The intensities of the blots were scanned and quantitated by a video densitometer (Model 620, BioRad). mRNA intensities are calculated relative to the intensity of the β-actin internal control.

5.3 Results
5.3.1 Sodium butyrate (NaB) inhibits uPAR and uPA mRNAs expression
To investigate the role of sodium butyrate in the regulation of uPAR mRNA, HCT116 cells were exposed to sodium butyrate concentrations varying from 0.25-25 mM for 4 h. Total cellular RNA was isolated from control cells and stimuli-treated cells and analyzed by Northern blotting. To determine the effects of butyrate on other members of the uPA/uPAR system, uPA mRNA expression was also determined in the same blot. As shown in Figure 5.1a, unstimulated HCT116 cells expressed low levels
of 1.4 kb uPAR and 2.5 kb uPA transcripts. Northern blot analyses show inhibition of uPAR and uPA mRNAs by sodium butyrate at 2.5 mM to 25 mM, in the range of concentrations found in the human colonic lumen. At 25 mM butyrate, uPAR mRNA levels were reduced by 60% and uPA mRNA expression was not detectable as determined by scanning densitometry.

5.3.2 Sodium butyrate inhibits TNFα- or PMA-stimulated uPAR and uPA mRNAs expression

To study whether expression of uPAR and uPA mRNAs is inhibited by butyrate in the presence of TNFα or PMA in human colon cancer cells, TNFα (440 units/ml) or PMA (30 ng/ml) plus various concentrations of butyrate (0.25-25 mM) were added to the HCT116 cells for 4h. As shown in Figure 5.1b and c, treatments of cells with PMA or TNFα induced a marked increased expression of uPAR and uPA. After PMA stimulation, uPAR and uPA mRNAs were increased 18-fold and 25-fold, respectively, while TNFα stimulation resulted in an increase of uPAR and uPA mRNAs 6-fold and 26-fold, respectively. In the presence of butyrate, however, a marked inhibition of uPAR and uPA mRNA overexpression in HCT116 cells was detected. At 25 mM, butyrate inhibited approximately 90% of the uPA mRNA induced by TNFα or PMA, whereas uPAR mRNA was inhibited by about 50% in the same experiment, suggesting that uPA mRNA was more sensitive to butyrate than that of uPAR (Figure 5.1b and c).

As in HCT116 cells, uPAR and uPA mRNAs were also strongly induced in LIM1215 colon cancer cells following treatment with PMA (Figure 5.2, lane 2). The effect of PMA induction was inhibited by sodium butyrate in a dose-dependent manner (Figure 5.2, lanes 3-5). These results suggest that butyrate may have similar inhibitory effects on uPAR and uPA mRNA expression in different colon cancer cell lines.

In our experiments, butyrate had no effect on the total RNA yield or on cell viability. When the uPAR probe was eluted from the filters and the same RNA blot then hybridized with 18S rDNA, no significant changes in 18S rRNA were detected, regardless of whether the cells had been exposed to butyrate, indicating that butyrate is selective in inhibiting gene expression.
Figure 5.1  Inhibition of uPA/uPAR mRNAs by sodium butyrate (NaB) in HCT116 cells. Cells were untreated or treated with PMA (30 ng/ml) or TNFα (440 units/ml) for 4 h in the presence or absence of NaB. The same Northern blot was hybridized to uPA cDNA, uPAR cDNA and 18S rDNA probes as indicated. (A) Dose-dependent inhibition of mRNA accumulation by NaB. Autoradiographic exposure times were 5 days (for uPA and uPAR cDNA as probes) and 1.5 h (for 18S rDNA as probe). Lane 1, untreated; Lanes 2-4, NaB at 0.25, 2.5 and 25 mM, respectively. (B) Dose-dependent inhibition of mRNA accumulation by PMA plus NaB. Autoradiographic exposure times were 48 h (for uPA), 24 h (for uPAR) and 1.5 h (for 18S rDNA). Lane 1, untreated; Lane 2, PMA; Lanes 3-5, PMA plus NaB. NaB at 0.25, 2.5 and 25 mM, respectively. (C) Dose-dependent inhibition of mRNA accumulation by TNFα plus NaB. Autoradiographic exposure times were 24 h (for uPA), 72 h (for uPAR) and 1.5 h (for 18S rDNA). Lane 1, untreated; Lane 2, TNFα; Lanes 3-5, TNFα plus NaB. NaB at 0.25, 2.5 and 25 mM, respectively. Each experiment was repeated at least three times and representative data shown in the figures.
Figure 5.2 Inhibition of uPA/uPAR mRNAs by sodium butyrate (NaB) in LIM1215 cells. Cells were treated with PMA (30 ng/ml) for 4 h in the presence or absence of NaB. The same Northern blot was hybridized to uPA cDNA, uPAR cDNA and 18S rDNA probes as indicated. Autoradiographic exposure times were 24 h (for uPA and uPAR cDNA as probes) and 1.5 h (for 18S rDNA as probe). Lane 1 untreated; Lane 2, PMA; Lanes 3-5, PMA plus NaB. NaB at 0.25, 1.0 and 10 mM, respectively.
1 2 3 4 5

uPA-

uPAR-

18S rRNA-
5.3.3 Sodium butyrate inhibits uPAR protein levels on the cell surface

To determine whether sodium butyrate affects uPAR protein expression, the amount of this receptor on the cell surface was indirectly examined by uPAR binding assay. HCT116 colon cancer cells were exposed to butyrate or PMA plus butyrate for 16 h and receptor-bound $^{125}$I-uPA was measured. The results showed that butyrate treatment alone, at 2.5 or 25 mM concentrations, resulted in a decrease of uPA binding to its receptor on the cell surface. PMA-induced uPAR protein production on the cell surface was inhibited by butyrate in a dose-dependent manner (Figure 5.3). The decrease in protein levels on cell surface caused by butyrate corresponded to the decrease of the uPAR mRNA. TNFα treatment (440 units/ml) induced lower levels of uPAR protein production on the cell surface than PMA (50 ng/ml) (Wang et al., 1994b).

5.3.4 Sodium butyrate inhibits cycloheximide-induced uPAR and uPA mRNAs expression

To investigate whether the inhibition of uPAR or uPA gene expression by butyrate depends on new protein synthesis, HCT116 cells were preincubated with CHX (20 µg/ml) for 30 min to ensure inhibition of protein synthesis at an early time point and then butyrate was added for 4h. Northern blot analysis showed that in HCT116 cells, CHX treatment alone induced uPAR and uPA mRNA expression suggesting that both genes are regulated by a labile repressor protein (Wang et al., 1994). uPAR and uPA mRNAs were also induced by CHX in HCT116 cells stimulated by TNFα or PMA (Figure 5.4), suggesting that stimulation of uPA and uPAR gene expression does not need de novo protein synthesis. uPAR or uPA mRNAs induced by CHX in the absence or presence of TNFα or PMA were inhibited by sodium butyrate (Figure 5.4) indicating that the effects of butyrate on uPA and uPAR gene expression do not require synthesis of new regulatory proteins in colon cancer cells. No difference was observed in the constitutively expressed 18S rRNA, which served as a control.

5.3.5 Effects of sodium butyrate on uPAR and uPA transcription in isolated nuclei

To determine whether effects of butyrate on the levels of uPAR or uPA mRNA act at the transcriptional level, nuclear run-on transcription assays were performed in
HCT116 cells untreated or treated for 4 h with 2.5 mM butyrate. As shown in Figure 5.5a, uPAR and uPA transcripts were expressed in nuclei from unstimulated HCT116 cells. While butyrate inhibited uPA mRNA expression, it had no effect on uPAR mRNA indicating that the inhibitory effect of butyrate alone on uPA expression was mediated, at least in part, at the transcriptional level. In contrast, the inhibition of uPAR mRNA by butyrate alone was not mediated at the transcriptional level, but may act via post-transcriptional mechanisms including the control points of mRNA processing, transport, translation or degradation. uPAR and uPA transcripts were increased in the cells exposed to TNFα (Figure 5.5c), but not PMA (Figure 5.5b). Both uPAR and uPA transcripts were also markedly reduced in the cells treated with butyrate plus TNFα (Figure 5.5c). These assays indicate that the inhibitory effect of butyrate in the presence of TNFα was mediated, at least in part, at the transcriptional level, whereas the inhibitory effect of butyrate in the presence of PMA may act via a post-transcriptional mechanism.
Figure 5.3 Effects of sodium butyrate (NaB) on uPAR protein production on the cell surface in PMA-stimulated or unstimulated HCT116 cells. Cells were deprived of serum for 4 h and then treated with NaB (0, 0.25, 2.5 and 25 mM) or PMA (30 ng/ml) plus NaB for 16 h. The results are shown as the mean ± SD of data from three separate experiments.
Figure 5.4 Inhibition of uPA/uPAR mRNAs by sodium butyrate (NaB) in the presence of cycloheximide (CHX) in HCT116 cells. Cells were untreated or treated with CHX (20 µg/ml) for 4 h in the presence or absence of NaB (2.5 mM), PMA (30 ng/ml) and TNFα (440 units/ml). The same Northern blot was hybridized to uPA cDNA, uPAR cDNA and 18S rDNA probes as indicated. Autoradiographic exposure times were 24 h (for uPA and uPAR as probes) and 4 h (for 18S rDNA as probe). The symbol '-' and '+' indicate untreated and treated cells respectively. Each experiment was repeated at least three times and representative data shown in the figures.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>uPA-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>uPAR-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18S rRNA-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHX</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PMA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TNFα</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Butyrate</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
Figure 5.5 Nuclear run-on transcription analysis of effects of sodium butyrate (NaB) on uPAR and uPA gene transcription in HCT116 cells. Nuclei were isolated from HCT116 cells treated with NaB (2.5 mM) (A), PMA (30 ng/ml) (B) or TNFα (2,000 unit/ml) plus NaB (2.5 mM) (C). Following in vitro transcriptions, RNA was purified from the nuclei and hybridized with uPAR cDNA, uPA cDNA, β-actin cDNA or vector DNA (pGEM) immobilized on nitrocellulose filters. Hybridized 32P-labeled nuclear RNA was detected by autoradiography. Autoradiographic exposure time was 7 days for experiment (A), 6 days for experiment (B) and 5 days for experiment (C). Each experiment was repeated at least twice and representative data shown in the figures.
A

control  NaB

uPAR-

uPA-

β-actin-

pGEM-

B

c控制  PMA

C

c控制  TNFα  TNFα+NaB
5.4 Discussion

This chapter reports that butyrate inhibits uPAR and uPA expression in the presence or absence of PMA and TNFα in colon cancer cells. While the effects of PMA on uPAR and uPA expression are mediated at the post-transcriptional level and TNFα at transcriptional level, butyrate acts at both transcription and post-transcription levels.

Previous reports showed that butyrate affects expression of genes including the c-fos and c-myc at both the transcription and post-transcription levels (Souleimani and Asselin, 1993a and 1993b). Butyrate induced the c-fos proto-oncogene expression very rapidly at the post-transcriptional level but up-regulated the c-fos gene at the transcriptional level at a later time in a colon carcinoma cell line, Caco-2 (Souleimani and Asselin, 1993a). c-myc gene expression, however, was down-regulated at the post-transcriptional level in the same cell line (Souleimani and Asselin, 1993b), indicating that complex mechanisms may be involved in the regulation of c-fos and c-myc gene expression by butyrate. In the case of the uPA system, butyrate may modulate gene expression at both the transcription (uPA) and post-transcription (uPAR) levels suggesting that different mechanisms may be involved in the regulation of uPA and uPAR gene expression.

PMA-induced gene expression has been considered to be one of the events required for tumor promotion (Mayo et al., 1994). The data presented in this chapter show that PMA-induced increases of uPAR and uPA mRNAs were markedly inhibited by sodium butyrate suggesting lumenal butyrate may also protect colon epithelial cells exposed to tumor promoters by down-regulating uPA system gene expression. Butyrate also inhibits PMA-induced uPAR protein production on the cell surface suggesting that butyrate inhibits uPAR expression at both mRNA and protein levels. In HCT116 colon cancer cells, the accumulation of uPAR and uPA mRNA induced by PMA may be mainly at a post-transcriptional level (Wang et al., 1994b). Using same approach, a 25-fold increase in uPAR transcription was observed in PMA-stimulated GEO colon cancer cells. However, no alteration was found in the invasive RKO cells (Wang et al., 1994a). PMA has been shown to increase mRNA stabilization in many genes such as mouse GM-CSF (Iwai et al., 1991) and ribonucleotide reductases (Amara et al., 1994) by regulatory elements at 3'-untranslated regions of these genes.

TNFα is produced by many cell types including activated monocytes, macrophages and lymphocytes (Beyaert and Fiers, 1994). Tumor infiltrating macrophages, but not
colon cancer cells, may also produce TNFα mRNA and protein in colon cancer tissue (Beissert et al., 1989). The inhibitory effect of butyrate on uPAR and uPA mRNAs stimulated by TNFα suggests that the tumor microenvironment may contribute to the regulation of uPAR and uPA gene expression in colonic cancers thereby influencing tumor invasiveness. Butyrate may also induce abnormal or transformed cell lines to convert to more normal phenotypes and function (Hague et al., 1993). These data combined with the inhibitory effects of butyrate on uPA and uPAR mRNA expression reported here, may offer the prospect of developing new therapeutic approaches to prevent invasion and metastasis by adenocarcinomas of the colon.

Cytokine-dependent activation of transcription factors such as NF-κB is one of the mechanisms by which signals are transmitted from the extracellular surface to the nucleus (Reddy et al., 1994). Human TNFα regulates gene expression in eukaryotic cells by binding to TNFα receptors (Yang et al., 1993) in endothelial cells. TNF receptor signal transduction pathways may involve cAMP, G-proteins, protein kinase C and other kinases (Niedbala and Stein-Picarella, 1993). The effects of TNFα may be mediated through the transcriptional factor NF-κB which is activated by the sphingomyelin pathway in HL-60 cells (Yang et al., 1993). The nuclear run-on experiment results reported here show that butyrate inhibits TNFα-induced uPAR mRNA at the transcriptional level suggesting that butyrate may affect uPAR gene expression by interacting with TNFα signaling pathways. Therefore, the inhibitory effects of butyrate on TNFα- and PMA-induced changes in uPAR and uPA mRNA levels may reflect inhibition of TNFα signaling pathways, inhibition of uPA gene transcription and degradation of both mRNAs at the level of post-transcription.

Butyrate is a naturally occurring short chain fatty acid that decreases cell growth, increases differentiation and induces apoptosis in colon cancer cell lines (Heruth et al., 1993), suggesting that locally produced butyrate may help protect colon epithelial cells from becoming neoplastic. Elevated uPAR and uPA expression has been reported for a number of human malignancies including colon, breast, bladder, and prostate cancer (Fazioli and Blasi, 1994). uPA-mediated proteolysis appears to be central to invasion by adenocarcinoma cells and inhibition of uPA activity has become a target of anti-metastasis therapy. The results presented in this paper show that in colon cancer cells, there is a strong correlation between the degree of inhibition of uPAR and uPA gene
expression and the concentration of sodium butyrate. Based on this and the evidence of up-regulation of uPA and uPAR gene expression by tumor promoters, it would seem that locally produced butyrate may have an important protective role in the colon.

In summary, the study presented in this chapter suggests that butyrate has strong effects on PMA- or TNFα-induced increase of uPAR as well as uPA expression in cultured human colon cancer cells. This inhibition occurs at both transcription and post-transcription levels. The regulation of uPA and uPAR gene expression by butyrate may represent a protective mechanism against increased uPAR-mediated tumor invasion and metastasis.
CHAPTER 6
GENERAL DISCUSSION AND FUTURE DIRECTION

6.1 General Discussion

6.1.1 The aims and achievement of this thesis

The aim of this thesis was to characterize the genomic structure and transcriptional regulation of the human uPAR gene. Towards this end, the organization of the 23 kb length of the human uPAR gene was determined and the mechanisms governing the transcriptional regulation of the uPAR gene were analyzed. Comparison of the uPAR genomic DNA with its cDNA showed that this gene contains 7 exons and 6 introns. Sequence analysis revealed that the uPAR promoter contains several putative cis-regulatory elements including AP-1, Sp1 and NF-κB, but no TATA or CAAT boxes. These results provided an opportunity to identify and characterize the regulatory elements and potential transacting factors involved in the regulation of the uPAR gene expression.

Recently, about 1.5 kb of the 5'-flanking region of the human uPAR gene has been described and characterization of the uPAR promoter showed that binding of Sp1 transcription factor to the promoter region is involved in the basal transcription of the uPAR gene (Soravia et al., 1995). The mechanisms of the regulation of uPAR gene expression, however, are still poorly understood. Little is known about the negative regulatory element involved in uPAR gene expression. In this thesis, several novel findings have been shown. (1) In human colon cancer cells, transient transfection analyses showed that a strong promoter is found in a 188 bp region between -141 and +47 relative to the transcription start site. Either 5' deletion to -60 or 3' deletion to -10 in this region abolished its promoter activity. These results suggest that both 5' and 3' regions of the uPAR promoter, bases -141 to -61 and -9 to +47, are necessary for the whole promoter activity. The first region (-141 to -61) contains the DNA binding sites for transcription factors Sp1 and AP-1, while the second region (-9 to +47) contains the putative Inr and a DPE element. (2) DNA footprinting analysis revealed two protected regions in the uPAR promoter; region I between -51 and -30 contains a NF-κB binding site and region II between -118 and -79 contains two potential Sp1 binding sites. (3) The transcription factor AP-1 can transactivate uPAR gene transcription in
human colon cancer cells. Co-transfection with the c-jun and c-fos expression plasmids stimulated the uPAR promoter activity 4 to 5-fold. (4) Mobility shift analysis showed that the NF-κB protein specifically binds to a putative reversed NF-κB motif around -38 bp. Mutation of the NF-κB sequence decreased this specific binding. (5) Extended sequencing revealed four Alu repeat sequences in both orientations in the 2.6 kb uPAR 5'-flanking region. Deletion analyses identified a negative regulatory region between -1796 and -1458. The silencer suppressed expression of the CAT reporter gene in a position-, orientation- and promoter-independent fashion. A full-length conserved Alu sequence may be necessary for maximal silencer activity. These results demonstrated that transcription of the human uPAR gene is positively or negatively regulated by these DNA elements.

The results presented in this thesis also indicated that sodium butyrate inhibits both uPAR and uPA gene expression in PMA- or TNFα-stimulated and unstimulated HCT116 cells at both transcription and post-transcription levels. This suggests that butyrate may have a protective role in tumor invasion and metastasis by colonic adenocarcinoma cells by inhibiting uPAR and uPA expression.

Characterization of the human uPAR gene and the cis-acting regulatory elements provides a basis for further studies on the regulation of uPAR gene expression. This knowledge will also contribute to the overall goal of understanding mechanisms in the control of cancer metastasis and may offer the prospect of elucidating specific molecular targets for future gene therapy research.

6.1.2 Transcription regulation of the uPAR gene is complex

The human uPAR gene promoter shows very strong activity compared to the SV40 promoter when transfected into human colon cancer cells. The ability of this uPAR promoter to direct CAT expression is orientation-dependent, as expected for a eukaryotic promoter. The uPAR promoter region has some features typical of promoters of housekeeping genes (Dynan, 1986; Pugh and Tjian, 1992): a high G+C content, TATA-less and multiple Sp1 binding sites. The Sp1 binding site appears to play an important role in directing transcription of other TATA-less promoters, such as mouse retinoblastoma (RBI) promoter (Zackshaus et al., 1993) and transcription enhancer factor-1 (TEF-1) promoter (Boam et al., 1995). A highly conserved promoter sequence has been reported between human and mouse uPAR genes (Soravia et al.,
The evolutionary conservation of DNA sequences usually indicates their functional importance. Like many other TATA-less and G+C-rich promoters, an imperfect three-dimensional stem structure has been speculated from the sequence around the multiple transcription start sites within the uPAR promoter region (Figure 2.7). This structure may improve uPAR transcription initiation as it may provide a 'landmark' to increase the speed and efficiency in the search for promoter by the transcription initiation complex (Ackerman et al., 1993).

As summarized in Figure 6.1, a number of regulatory elements were characterized in the 5'-flanking region of the uPAR gene. These include a strong promoter, a silencer and several DNA binding sites for transcription factors including Sp1, AP-1 and NF-kB. Sp1, a transcription factor that has been proved to play a primary role in the regulation of a large number of gene promoters (Pugh and Tjian, 1990), is required for uPAR promoter activity. This conclusion was drawn from the results in this study and others (Soravia et al., 1995). Data from transcription analysis and sequence analysis suggest that uPAR promoter activity may depend on the existence of Sp1 binding sites located at the region between -141 and -61. Interaction of nuclear factors with Sp1 binding sites located in this region was further confirmed by both DNase I footprinting analysis and the EMSA competition experiments. Sp1 or Sp1-related protein is the major nuclear protein binding to the 72 bp regulatory region between -137 and -66, since multiple DNA-protein complexes formed in this region were completely abolished in the presence of the Sp1 consensus sequence. Soravia et al. (1995) have also shown that Sp1 binds to the uPAR promoter at numerous sites and its expression is essential to the uPAR transcription. The data presented in this thesis further demonstrated that binding of Sp1 or Sp1-related protein to the 72 bp regulatory region plays an essential role in uPAR promoter activity.

Using EMSA, the AP-1 transcription factor, which is comprised of heterodimers of jun and fos family members or homodimers of jun or AP-1-related protein (Angel and Karin, 1991), was demonstrated to interact with the AP-1 motif of the uPAR promoter. Co-transfection experiments with the c-fos and c-jun expression plasmids also demonstrated that overexpression of AP-1 transcription factor markedly transactivated uPAR gene expression in HCT116 cells. Although two AP-1 binding sites are located within the 5'-flanking region of the uPAR gene at positions -184 and -70, the proximal site at -70 bp was responsible for this activation by AP-1. This AP-1 motif is conserved
Figure 6.1 Schematic representation of *cis*-acting regulatory sequences in the 5'-flanking region of the human uPAR gene. The regulatory elements that may govern uPAR gene transcription are indicated. The numbers indicate distance (in base pairs) from the major transcription start site (arrow). Four Alu repeat sequences are shown by stippled boxes and their orientation are indicated by open arrowheads.
between human and mouse uPAR promoters (Soravia et al., 1995). Binding of nuclear protein to this AP-1 site was confirmed by competition analysis using an AP-1 consensus sequence. It has been shown that binding of AP-1 to its cognate sequence appears to play an important role in controlling transcription regulation of many genes involved in various cell processes (Angel and Karin, 1991). Altered c-jun expression has been found to be an early event in human lung carcinogenesis suggesting that it may function as a mediator of growth factor signals (Szabo et al., 1996). Interaction of AP-1 and uPAR promoter, therefore, may provide a mechanism of activation of the uPAR gene in response to various stimuli that can activate AP-1 activity.

In addition to AP-1, NF-κB may be another potential transcription factor in the regulation of the uPAR gene. This transcription factor is involved in the transmission of signals from the cytoplasm to the nucleus and participates in the inducible expression of a variety of genes involved in the immune, inflammatory or acute phase responses (Schenk et al., 1994; Singh and Aggarwal, 1995; Pahl and Baueerle, 1995). Using a rel-related subunit of 49 kDa (p49), which has specific κB-binding activity and is capable of forming heterodimers with other rel proteins (Schmid et al., 1991), binding of NF-κB to the inverted uPAR NF-κB binding site at -41 bp was confirmed by EMSA analysis. This binding was also competed for by NF-κB consensus sequence. Binding of the p65/c-Rel heterodimeric complex to a uPA DNA element has been shown to be involved in PMA-induced expression of this gene (Hansen et al., 1992). Therefore, interaction of NF-κB with the uPAR promoter sequence suggests that NF-κB may contribute to the inducible expression of the uPAR gene.

Deletion of the 3' uPAR promoter region revealed a positive regulation sequence around the region from -9 to +47. This region may contribute to basal promoter activity because it contains part of the putative Inr sequence and a DPE motif. Overlapping with this region, several DNA-protein complexes were detected by the EMSA. TFIID, one of the general transcription factors that work in concert with RNA polymerase II to bring about promoter recognition and accurate transcription (Goodrich et al., 1996), was found to partially compete with some of this binding. In addition to binding to the TATA boxes, binding of TFIID to both Inr element and DPE motif has also been shown to be important for the basal transcription (Goodrich et al., 1996; Burke and Kadonaga, 1996). Study of the interaction between general
transcription factors and the promoter region has been essential to understanding the regulation of gene expression. Therefore, it is of interest to characterize further the DNA-protein interaction located at this region.

Negative regulation of gene transcription may involve many different mechanisms (Clark and Docherty, 1993; Herschbach and Johnson, 1993). A uPA transcriptional silencer which is able to block its transcription in a cell-specific and orientation independent manner has been identified in the 5'-flanking region of this gene (Cannio et al., 1991). In the uPAR gene, a 339 bp region (-1796 to -1458) far upstream containing a conserved full-length Alu repeat sequence was demonstrated to be required for maximal silencer activity from deletion experiments. This element showed a very strong negative effect when directly placed in front of the promoter region in both orientations, suggesting an orientation-independent silencer activity. It has been shown that the Alu repeat sequence can influence transcription of many genes in either a positive or negative manner (Wu et al., 1990; Norris et al., 1995; Hewitt, et al., 1955). Within the 2.6 kb upstream sequence, four Alu repeats were found in both orientations (Figure 6.1). The proximal two Alu repeats are in antisense orientation and the distal are in sense orientation. Thus, the proximal two Alu repeats may have the potential to base-pair with the distal repeats to form a cruciform structure to influence uPAR transcription in vivo. Alu repeats may also repress transcribing polymerases by transcription of the Pol III promoter and block the passage of transcribing polymerases by formation of transcription complexes (Wu et al., 1990; Englander et al., 1993). Because a total of 18 Alu repeat sequences were found through the human uPAR gene (Wang et al, 1995; this study), it will be of interest to search for possible roles of other Alu sequences in the regulation of uPAR gene expression.

Transcription regulation of a gene may involve interactions between activators bound to cis-acting elements or between these activators and factors bound to basal elements near the transcription start site (Li et al., 1994; Das et al., 1995). These protein-protein interactions also provide a link between signal transduction pathways (Kawana et al., 1995; Goodrich et al., 1996). For example, co-operative interaction of AP-1 and GATA-binding protein (GATA-2) can synergistically increase transcription activity of the endothelin-1 gene promoter in endothelial cells (Kawana et al., 1995). A functional cooperation between Sp1 and NF-κB/Rel has also been demonstrated in PMA-mediated HIV-1 long terminal repeat activation (Perkins et al., 1993; Li et al.,
In addition, interaction of Sp1 with Inr-binding protein YY1 (Lee et al., 1993) or with TFIID via the TBP-associated factor, TAF\textsubscript{II} 10 (Gill et al., 1994), has also been shown to be required for Sp1-mediated transcriptional activation. In the uPAR gene, presence of multiple cis-regulatory elements in the 5'-flanking region suggests a complex mechanism in controlling uPAR gene expression. In addition to the AP-1, Sp1, NF-κB, Inr and DPE elements, the 5'-flanking region of the human uPAR gene also contains consensus binding sites for transcription factors AP-2, PEA3, GATA and CF-1, which may play roles in governing transcription of the gene. Thus, AP-1 activation of the uPAR gene and contribution of Sp1 to promoter activity may also be involved in the interaction with other transcription factors. The integration of multiple DNA-protein and protein-protein interactions within the extended regulatory DNA sequences will allow uPAR expression under a fine regulation in various environments.

This work concentrated on the colon cancer cell line HCT116 but included Northern blot analysis and CAT assays on a second colon cancer cell line LIM1215. Had a uPAR-deficient cell line been available it would have been an important control to determine whether AP-1 or Sp1 transcription factors contribute to the elevated or basal expression of the uPAR gene.

### 6.1.3 Regulation of uPAR gene expression by sodium butyrate

A hallmark of the uPAR gene is its inducibility by a variety of agents including mitogens, cytokines and differentiation-specific agents. uPAR gene expression appears to be regulated at both transcriptional and post-transcriptional levels and does not require new protein synthesis (Wang et al., 1994b). The overexpression of the uPAR gene is often found at the invasive edge in many human cancers including colon cancer tissues (Pyke et al., 1991). The human colon cancer cell line HCT116, that displays a high capacity of invasiveness and shows inducibility of uPAR expression by PMA and TNFα, provides a useful model for studying the regulation of uPAR gene expression by various stimuli. Studies described in this thesis demonstrated that sodium butyrate can inhibit both uPAR and uPA gene expression in PMA- or TNFα-stimulated and unstimulated colon cancer cells suggesting that butyrate may have a protective role against uPA-mediated tumor invasion by repressing both uPAR and uPA expression.

Butyrate is a natural product of colonic bacterial fermentation and has varying effects on cell proliferation and differentiation by acting on chromosomal structure and gene expression (Tsutsumi et al., 1994). Recent studies have shown that it regulates gene expression at both transcription and post-transcriptional levels (Deng et al., 1992;
Souleimani and Asselin, 1993a and 1993b). Northern blot analysis in this thesis showed that butyrate inhibits uPAR and uPA mRNA expression in a dose-dependent manner in colon cancer cells regardless of whether they have been stimulated with TNFα, PMA or cycloheximide. In addition to mRNA levels, butyrate also inhibits uPAR expression on the cell surface. Butyrate inhibits uPAR gene expression superinduced by TNFα or PMA in the presence of a protein synthesis inhibitor, suggesting that inhibition of uPAR gene expression by butyrate does not require new protein synthesis. It seems likely that inhibition by butyrate must be triggered by the modification of pre-existing transcription factor(s).

Nuclear run-on transcription analysis has been used to examine the mechanisms of transcription regulation of uPAR and uPA genes by butyrate in colon cancer cells. Using this method, butyrate in the presence or absence of TNFα was demonstrated to act, at least in part, at the transcriptional level on uPA gene expression, but to act at both transcriptional and post-transcriptional levels on the expression of the uPAR gene. In the presence of PMA, however, butyrate acts mainly at the post-transcriptional level on the expression of uPAR and uPA genes. Therefore, both PMA- and TNFα-stimulated uPAR and uPA gene expression was inhibited by butyrate, but butyrate regulates each of these genes through different mechanisms in a gene-specific manner.

Recent study has indicated that butyrate modulates uPAR, uPA and PAI-1 gene expression through different signal transduction pathways. Inhibition of uPAR gene expression by butyrate is independent of both PKC and PKA signal transduction pathways as it can effectively suppress uPAR mRNA in the presence of Br-cAMP, PMA and okadaic acid (Antalis and Reeder, 1995). In the uPA gene, both AP-1 and PEA3 binding sites have been found in the enhancer region and proved necessary for stimulation of uPA gene expression by PMA and TNFα (Nerlov et al., 1992; Lengyel et al., 1995). A NF-κB binding site in the human uPA 5′-flanking region has also been reported to mediate in its overexpression by PMA (Hansen et al., 1992). Although these binding sites were also found in the 5′-flanking region of the uPAR gene and uPAR gene expression was transactivated by overexpression of AP-1 and up-regulated by PMA/TNFα, links between these cis-acting elements and the cellular response to stimuli remains to be elucidated.

Butyrate is a known inhibitor of histone deacetylase. The induced alterations in gene expression may be a result of changes in chromatin structure (Sealy and Chalkley, 1978). Recent studies have shown that butyrate is involved in the regulation of
promoter activity in many genes such as PLAP, c-fos, CD11c and calbindin-D28k. In contrast to its inhibitory role in PMA-stimulated uPAR expression in colon cancer cells, butyrate can act synergistically with PMA in inducing CD11c gene promoter activity. A butyrate responsive element has been identified in the 5'-flanking sequence of the PLAP (Deng et al., 1992), c-fos (Souleimani and Asselin, 1993a) and the calbindin-D28k genes (Gill and Christakos, 1993). However, no butyrate responsive consensus sequence has so far been reported. It has also been indicated that butyrate may regulate gene expression through several DNA binding sites. For example, removal of the consensus AP-1 sequence from the CD11c promoter region led to a reduced butyrate responsiveness (Rubio et al., 1995). Similarly, mutation of Sp1 sites within the human immunodeficiency virus-long terminal repeat (HIV-LTR) abolished its butyrate-inducible promoter activity (Bohan et al., 1989). Recent study has shown that regulation of uPA gene expression by butyrate may be through a PEA3 binding site in the 5' flanking region of this gene (Antalis and Reeder, 1995). If this is the case, a number of Sp1 binding sites, two PEA3 binding sites and two AP-1 binding sites located in the uPAR 5'-flanking region may be involved in butyrate inactivation. Identification of uPAR promoter elements that interact with several transcription factors should help to identify the signal transduction pathways that stimulate expression of the uPAR gene.

6.2 Future Studies

6.2.1 To study signal transduction pathways in the regulation of uPAR expression

uPAR gene expression has been demonstrated to be transcriptionally regulated by many factors including PMA, cytokines and growth factors in different cell types. Although a number of transfactors, such as Sp1, AP-1 and NF-κB have been shown to have a clear connection with the uPAR promoter, the links between these transcription factors and uPAR expression in response to the stimuli is still not clear. Therefore, future work on signal transduction pathways involved in the regulation of uPAR gene expression is required. These interesting studies may also include the questions of transcriptional regulation of the uPAR gene by sodium butyrate. Studies on signal transduction control of uPAR gene expression will represent an important future prospect for the regulation of uPAR gene transcription.
6.2.2 To study transcription factors involved in tumor metastasis

Clearly, one of the important future research priorities is to determine the transcription regulation of the uPAR gene in tumor metastasis. This study will not only allow us to identify transcription factors that are involved in controlling uPAR gene expression, but also provide an opportunity for interference of tumor metastasis by inhibiting uPAR expression. Transcription factors that are involved in human disease have become targets for therapy in recent years (Peterson and Baichwal, 1993). Recent study has provided a prospect for anti-metastasis therapy using antisense oligodeoxynucleotides, which reduce tumor cell invasiveness by inhibiting NF-κB-Rel A expression that is involved in PMA induction of the uPA gene (Reuning et al., 1995). Therefore, understanding the transcriptional regulation of the uPAR gene in tumor metastasis and interfering with its transcriptional regulation in cancer cells will be an important aspect in anti-metastasis study. Moreover, a striking result in this thesis is that butyrate inhibits both uPAR and uPA gene expression in PMA- and TNFα-stimulated and unstimulated human colon cancer cells, suggesting that butyrate may have a protective role in uPA-mediated tumor metastasis. A further study of transcription regulation of the uPAR gene by butyrate in tumor metastasis and the effects of butyrate on normal colon epithelial cells and cells from colonic adenomas and primary adenocarcinomas of the colon to test the clinical relevance would be interesting.

6.2.3 To test the functionality of uPAR regulatory elements in transgenic mice

Although the transcriptional unit and core promoter of the uPAR gene has been identified in this study, further studies will be required to confirm these findings in vivo. To test if the cis-acting sequences and trans-acting factors identified by regulation studies in vitro function in vivo, transgenic mice will be generated using constructs containing the 5'-flanking region of the uPAR gene linked to the CAT reporter gene. Detection of transgene expression in transgenic mice with systematic deletions of the 5'-flanking sequence will allow us to identify the specific regulatory elements required for uPAR expression in vivo.

A number of specific nuclear proteins have been shown in this thesis to bind the region around the transcription initiation site of the uPAR gene. However, the exact nature of these DNA-protein complexes is still unknown. The role of general transcriptional factors such as TFIID in transcriptional regulation of the eukaryotic
gene remains the challenging area of future investigation. Therefore, it would be very interesting to characterize the DNA-protein interaction in this region of the uPAR gene \textit{in vitro} and \textit{in vivo}. Transgenic mice will also be the best model for answering the question of whether there are additional genomic sequences regulating the uPAR gene expression. This study will provide a comprehensive understanding of transcriptional regulation of the uPAR gene \textit{in vivo}. I believe that the information obtained from this thesis will serve as a basis for further investigation into the mechanism of transcriptional regulation of the uPAR gene and uPAR-mediated biological processes.


Conese, M., Olson, D., and Blasi, F. (1994) Protease nexin-1-urokinase complexes are internalized and degraded through a mechanism that requires both urokinase receptor and α2-macroglobulin receptor. J. Biol. Chem. 269, 17886-17892.


Cubellis, M. V., Wun, T. C., and Blasi, F. (1990) Receptor-mediated internalization and degradation of urokinase is caused by its specific inhibitor PAI-1. EMBO J. 9, 1079-1085.


Lu, H., Yeh, P., Guitton, J-D., Mabilat, C., Desanlis, F., Maury, I., Legrand, Y., Soria, J., and Soria, C. (1994) Blockage of the urokinase receptor on the cell surface,


Olson, D., Pöllänen, J., Høyer-Hansen, G., Rønne, E., Sakaguchi, K., Wun, T-C., Appella, E., Danø, K., and Blasi, F. (1992) Internalization of the urokinase-
plasminogen activator inhibitor type-1 complex is mediated by the urokinase receptor. J. Biol. Chem. 267, 9129-9133.


phorbol ester, and sodium butyrate induce the CD11c intergrin gene promoter activity during myeloid cell differentiation. Blood 86, 3715-3724.


