REGULATION OF HUMAN ANTITHROMBIN GENE EXPRESSION.

1. MAPPING OF A DELETION INCLUDING IN PART THE PROMOTER REGION.
2. CHARACTERIZATION OF *CIS*-ACTING ELEMENTS AND TRANSCRIPTION FACTORS BINDING TO THIS REGION.

By
FRANCOISE FERNANDEZ-RACHUBINSKI.
Diplome d'Etat en Pharmacie,
C.E.S, Internat en Biologie Clinique.
Doctorat d'Universite

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
© Copyright by Francoise Fernandez-Rachubinski, March 1996

For Richard, Dorian and Ariane.

CHARACTERIZATION OF THE HUMAN ANTITHROMBIN GENE PROMOTER

DOCTOR OF PHILOSOPHY (1996)
(Medical Sciences)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE : REGULATION OF HUMAN ANTITHROMBIN GENE EXPRESSION
1. MAPPING OF A DELETION INCLUDING INCLUDING IN PART THE PROMOTER REGION.
2. CHARACTERIZATION OF *CIS*-ACTING ELEMENTS AND TRANSCRIPTION FACTORS BINDING TO THIS REGION.

AUTHOR : Francoise Fernandez-Rachubinski
Doctorat d'Universite (Toulouse)

SUPERVISOR : Dr. M.A. Blajchman, Professor, MD, FRCP(C)

NUMBER OF PAGES: xiv , 225

ABSTRACT

This thesis addresses regulatory aspects of the constitutive expression of human antithrombin, a serine-protease inhibitor involved in coagulation and thrombosis. An introductory part describes a 5'partial deletion, 480 nt upstream of the third exon of an antithrombin allele, in a kindred with thrombosis and hereditary deficiency in antithrombin. The lack of expression of the abnormal allele prompted the further characterization of elements regulating transcription in the 5' region of the normal antithrombin gene. Two *cis*-acting elements were found, both able to promote transcription in HepG2 and Cos I cells. The first promoter, at -150/+68 nt, encompassed the presumed transcriptional start site. The second element, in reverse orientation in regard to the first, was located at +895/+391 nt in the first intervening sequence. Following footprint analysis , it was shown that the 5'upstream promoter interacted with *trans*-acting factors at -92/-65 nt, -11/+37 nt and -124/-10 nt, in nuclear extracts from hepatic or non-hepatic origin. Several transcription-factors were subsequently identified, which interacted with these three elements either through direct binding or heterodimerization; they were the liver-enriched factors HNF4 and C/EBPα as well as the ubiquitous nuclear hormone receptors COUP-TF1, RXRα, PPARα and TRα. Furthermore, *in vivo* expression in HepG2, HeLa and BSC40 cells, of HNF4, C/EBPα and RXRα increased the efficiency of the 5'upstream promoter, while expression of COUP-TF1, TRα, HNF3α or β repressed the efficiency of the latter promoter and the activating effects of HNF4. In addition, activation by HNF4 was synergized by co-expression of RXRα in BSC40 and HeLa cells. These results suggest that the interplay of liver-enriched and ubiquitous factors modulates the constitutive expression of the antithrombin gene.

ACKNOWLEDGEMENTS

I would like to thank Dr Morris Blajchman for his supervision, and to extend my appreciation to the members of my supervisory commitee, Dr J. Capone, Dr J. Hassel, and Dr M. Hatton, for their helpful discussions and criticisms of this work.

I would like to acknowledge the Medical Research Council of Canada for financial support of this project.

I am indebted to Dr J. Wiener, who allowed completion of the work presented in this thesis. I am indebted to Dr R. Rachubinski and all the members of his laboratory, Sandra Marcus, Jennifer Smith, Baowei Zhang, Chris Winrow, Altaf Kassam, Gary and Christine Eitzen, Elaine Reklow, Pam Lagali, and Rachel Szilard, for their support, encouragements, use of their plasmids and antibodies, and equipment. Thanks also to L. Bardossi and B. Eng . My sincere thanks to Dr M. Pasdar for her support, and the use of her cell culture facilities. Thanks to Dr E. Shibuya, for the use of her scanner, and abuse of her computer and printer. Be ensured that R. Bradley's printing services were greatly appreciated.

Finally, I am more than grateful to D. Clayburn and M. Wyllie, who accepted to read this manuscript; thanks for their writing and editing suggestions.

LIST OF TABLES

CHAPTER I: INTRODUCTION

I.1 Tissue distribution of liver-enriched transcription factors.

I.2 Selected examples of liver specific promoters.

I.3 Classification of type I antithrombin deficiency.

CHAPTER II: MATERIALS AND EXPERIMENTAL PROTOCOLS.

II.1 List of EMSA oligonucleotides.

CHAPTER III: RESULTS

CHAPTER IV: DISCUSSION

IV.1 Homologies between the eukaryotic data bank and the antithrombin abnormal allele upstream of the breakpoint location.

IV.2 Signal sequences for eukaryotic promoter TATA elements in the antithrombin gene regions flanking exon 1.

LIST OF FIGURES

Chapter I: INTRODUCTION

I.1 Restriction map of the antithrombin gene.

Chapter II: MATERIALS AND EXPERIMENTAL PROTOCOLS

II.1 cDNA's and genomic probes.

II.2 Reporter plasmids and expression vectors.

Chapter III: RESULTS

III.1 Cloning and restriction map of genomic inserts for the 5' region of the antithrombin gene.

III.2 Nucleotide sequence of the antithrombin proximal region flanking exon 1.

III.3 Nucleotide sequence of the +391/+895 nt IVS 1 region.

III.4 Type I antithrombin deficient kindred: Pst1 Southern analysis and PCR confirmation of deletion in the 5'upstream region of the gene

III.5 Type I antithrombin deficient kindred: Hind III+Pst I Southern analysis

III.6 Type I antithrombin deficient kindred: Summary of Southern analysis and proposed restriction map of the deleted allele.

III.7 Type I antithrombin deficient kindred: PCR products generated during cloning of the breakpoint.

III.8 Type I antithrombin deficient kindred: Nucleotide sequence of the wild-type and mutant allele from exon 3 and upstream, and at the breakpoint location.

III.9 Mapping of 5'antithrombin cDNA ends by primer extension and S1 nuclease protection.

III.10 Mapping of the 5' antithrombin cDNA ends by RACE.

III.11 Deletional analysis of the antithrombin sequences flanking exon 1.

III.12 Deletional analysis of the -1100/+68 nt antithrombin region.

III.13 Activity of antithrombin -92/-65 nt and +1/+37 nt elements on heterologous promoters.

III.14 Strength of the 5' upstream promoter (-1100/+68 nt) in HepG2 and Cos1cells.

III.15 Effects of the antithrombin +895/+391 nt IVS1 element on SV40 promoter or enhancer activity.

III.16 Effects of the antithrombin -150/+68 nt upstream element on SV40 promoter or enhancer activity.

III.17 Footprint analysis of the upper-strand of the -150/+68 nt promoter.

III.18 Footprint of the lower-strand of the -150/+68 nt promoter.

III.19 Summary of footprinting studies.

III.20 EMSA with nuclear extracts and a -304/+68 nt or a -150/+68 nt element.

III.21 EMSA with nuclear extracts and elements A,B,and C.

III.22 Competition for binding of element A with oligonucleotides for COUP-TF1,HNF4, and HNF3.

III.23 Competition for binding of element A with single- and double-stranded oligonucleotides for PYBP.

III.24 EMSA with nuclear extracts from HepG2 cells and single-stranded PYBP oligonucleotides.

III.25 COUP-TF1 supershifts with element A and nuclear extracts.

III.26 HNF4, PPAR, RXR competition and supershifts with element A and nuclear extracts.

III.27 Competition assays with element A (wild-type and mutants) in nuclear extracts from HepG2 cells and binding profile of mutated elements A.

III.28 Competition assays and supershifts with element B at +1/+37 nt and nuclear extracts from HepG2 extracts and rat liver.

III.29 Competition assays and supershifts with with element C at -124/-101 nt and nuclear extracts from HepG2 cells.

III.30 SDS-PAGE of *in vitro* translated transcription factors.

III.31 Binding profile of element A (wild-type and mutants) by COUP-TF1.

III.32 Binding profile of element A (wild-type and mutants) by HNF4, PPAR, and RXR.

III.33 Binding profile of element A (wild-type and mutants) by TRα ; Interactions of nuclear hormone receptors on element A.

III.34 Interactions of HNF4, PPAR, and RXR, TR on element A.

III.35 Binding profile of element B by HNF3 and C/EBP.

III.36 Binding profile of element C by various transcription-factors.

III.37 Effect of expression of transcription-factors on the activity of the antithrombin 5'upstream promoter in HepG2 cells.

III.38 Effect of expression of transcription-factors on the activity of the antithrombin 5'upstream promoter in BSC 40 cells.

III.39 Effect of expression of transcription-factors on the activity of the antithrombin 5'upstream promoter in HeLa cells.

CHAPTER IV : DISCUSSION

IV.1 Hypothetical nuclear hormone receptors half-sites in direct-repeats and palindrome in elementsA andC.

CHAPTER V : CONCLUSION

V.1 Architecture of the antithrombin promoter region.

LIST OF ABBREVIATIONS

A	Ampere
AP 1,2,3,4	activator protein 1,2,3,4
nt	nucleotide
BSA	bovine serum albumin
CAT	chloramphenicol acetyl transferase
cDNA	DNA complementary to RNA
C/EBP	CCAAT enhancer binding protein
Ci	Curie
C1-inh	C1 inhibitor
COUP-TF	chicken ovalbumin upstream promoter-transcription factor
cpm	counts per minute
CTF/NF1	CCAAT transcription factor/nuclear factor 1
Da	dalton
DBP	site D albumin binding protein
ddNTP	2',3'-dideoxynucleoside 5'-triphosphate
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethylsulfoxide
dNTP	deoxynucleoside 5'-triphosphate
DTT	dithiothreitol
EDTA	ethylene diamine tetracetic acid
EMSA	electrophoretic mobility shift assays
g	gram
GTF	general transcription factor
HBS	HEPES buffered saline
HCII	heparin cofactor II
HEPES	n-(-2-hydroxyethyl)piperazine-N'- (2-ethanesulfonic acid)
HNF 1,2,3,4	hepatocyte nuclear factor 1,2,3,4
IgG	immunoglobulin G
Inr	initiator element
IPTG	isopropyl-1-thio-β-D-galactopyranoside
IVS	intervening sequence
k	kilo
LAP	liver activator protein
LFA1	liver factor A1
LUC	luciferase
μ	micro
m	milli

M	molar
min	minute
mol	mole
mRNA	messenger ribonucleic acid
n	nano
NFY	nuclear factor Y
ONPG	o-nitrophenylβ-D-galactopyranoside
p	pico
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethyl-sulfonyl fluoride
PYBP	pyrimidine binding protein
RACE	rapid amplification of cDNA ends
RFLP	restriction fragment length polymorphism
rpm	revolutions per minute
RAR	retinoic acid receptor
RXR	retinoid X receptor
SDS	sodium dodecyl sulfate
SSC	sodium chloride-sodium citrate buffer
TAF	TBP associated factors
TBE	tris-borate-EDTA buffer
TBP	TATA binding protein
TE	tris-EDTA buffer
TEMED	N,N,N',N'-tetramethylethylenediamine
TR	thyroid hormone receptor
Tris	Tris(hydroxymethyl)aminoethane
V	volts
Vit DR	vitamine d receptor
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

TABLE OF CONTENTS

ABSTRACT
ACKNOWLEDGEMENTS
LIST OF TABLES
LIST OF FIGURES
LIST OF ABBREVIATIONS

CHAPTER I: INTRODUCTION-REVIEW OF THE LITERATURE 1

INTRODUCTION 1

REVIEW OF THE LITERATURE 4

1. Antithrombin 4
 1.1 Introduction 4
 1.2 Antithrombin Expression 5
  1.2.1 Structural Features of the Gene, cDNA, and mRNA 5
  1.2.2 Regulation of Antithrombin Gene Expression 8

2. Transcriptional Machinery in Eukaryotes 11
 2.1 Introduction 11
 2.2 Core Promoter Elements 12
  2.2.1 Architecture of Regular Promoters 13
  2.2.2 Architecture of "TATA less" Promoters 14
 2.3 Enhancers, Activators, Coactivators, Repressors 15
  2.3.1 Cis-Acting Elements 15
  2.3.2 Trans-Acting Regulators 16
   (a) Coactivators 16
   (b) Activators 16
   (c) Repressors 18

3. Liver Specific Promoters 19
 3.1 Introduction 19
 3.2 Liver Enriched Factors 20
  3.2.1 HNF1 Family 20
  3.2.2 HNF3 Family 23
  3.2.3 HNF4 Family 24
  3.2.4 C/EBP Family 26
  3.2.5 Other Factors 29
 3.3 Single Stranded Binding Proteins and Liver Expression 30
 3.4 Specificity of Gene Expression in the Liver 32

4. General features of Regulatory Elements found in Proteinase Inhibitors and Clotting Factors with a Liver-Enriched Expression. 36
4.1 Serpins Genes other than Antithrombin, α_2Macroglobulin. 36
4.1.1 Cis-acting features for transcription initiation 37
4.1.2 Responsive Elements and Trans-acting factors 40
(a) in relation to constitutive expression in the liver 40
(b) in relation to inducible expression 40
4.2 Expression of Liver-enriched Clotting Factors. 42
4.2.1 Cis-acting features for transcription initiation 42
4.2.2 Responsive Elements and Trans-acting factors 42
(a) in relation to constitutive expression in the liver 42
(b) in relation to inducible expression 43

5. Pathology and Gene Expression Regulation of the Clotting Factors and Inhibitors of Coagulation 44
5.1 Antithrombin 44
5.2 Other Serpins and Clotting Factors 47
5.2.1 Factor IX 47
5.2.2 Protein C 48
5.2.3 Plasminogen Activator Inhibitor I 49
5.2.4 Fibrinogen 49

6. This Project 50

CHAPTER II : MATERIALS AND EXPERIMENTAL PROTOCOLS 52

1. MATERIALS 52

1.1 Chemicals 52
1.2 Radiochemicals 54
1.3 Restriction Enzymes, other Enzymes 54
1.4 Reagent Kits 54
1.5 cDNA's and Genomic Probes 55
1.5.1 cDNA probes 55
1.5.2 Genomic Probes 55
1.5.3 Probe Labelling 58
1.6 Cloning Vectors and Bacterial Strains 58
1.6.1 General Vectors 58
1.6.2 Reporter Plasmids 59
(a) Basic Plasmids 59
(b) Control Plasmids 59
1.6.3 In Vitro and In Vivo Expression Vectors for Transcription Factors 62
1.7 Antiseras, IgG fractions 63

2.EXPERIMENTAL PROTOCOLS 64

2.1 Restriction Mapping of the Deleted Allele in the Kindred with Type I Antithrombin Deficiency.
PCR Confirmation of Deletion in the 5' Upstream Region of the Gene 64
2.2 Cloning of the Breakpoint of the Deleted Allele 65
2.2.1 Preliminary Steps 65
2.2.2 Subcloning and Branch Capture PCR 65
2.2.3 Inverse PCR 65
2.2.4 Additional Steps 66
2.3 Cloning of Genomic Inserts For the 5' Region of the Antithrombin Gene 66
2.4 Mapping of the 5' Antithrombin cDNA Ends 67
2.4.1 S1 Nuclease Protection Assay 67
2.4.2 Primer Extension 68
2.4.3 Rapid Amplification of the 5' cDNA Ends 69
2.5 Plasmids and Antithrombin Constructs Used in Reporter Assays 69
2.6 Cell Culture and Maintenance, Transfection Protocols 71
2.7 Nuclear Extracts 73
2.8 Footprint Analysis 73
2.9 Gel Retardation Assays with Total Nuclear Extracts 75
2.10 *In Vitro* Transcription-Translation and Gel Retardation Assays with Transcription Factors 77
2.10.1 Transcription-Translation 77
2.10.2 Analysis of the Translated Products 78
2.10.3 Gel Retardation Assays 79

CHAPTER III : RESULTS 80

1. Genomic Clones Encompassing the 5' Upstream Region of the Antithrombin Gene and Restriction Mapping 80

2. Type I Antithrombin Deficient Kindred: Restriction Mapping of the Abnormal Allele. PCR Confirmation of Deletion in the 5' Upstream Region of the Gene 83

3. Type I Antithrombin deficient Kindred: Determination of the 5' Breakpoint of the Deleted Antithrombin Allele 91

4. Mapping of the Antihrombin cDNA 5' Ends 91
4.1 Primer Extension and S1 Nuclease Protection 91
4.2 5' cDNA RACE 95

5. Mapping of *Cis*-Elements Flanking Exon 1 in Reporter Assays 99
5.1 Deletional Analysis in the Plasmid pSVOA-LΔ5' 99
5.1.1 Elements Flanking Exon 1 100
(a) Mapping of the 5'upstream element 100
(b) Mapping of the IVS 1 element 102

5.1.2 Fine Mapping of the 5'Minimal Promoter 103
5.1.3 Strength of the 5'Upstream Promoter 106
5.2 Effects of the Antithrombin *Cis*-Elements on the SV40 Promoter and Enhancer 108

6. Footprint Analysis of the -150/+68 nt 5' Upstream Promoter 111

7. Gel Retardation Assays with Total Nuclear Extracts and the Elements -304/+68 nt and -150/+68nt 114

8. Gel Retardation Assays with Elements A, B, and C, and Total Nuclear Extracts 114

9. Identification of Factors Acting in *Trans* with element A at -92/-68nt 118
9.1 Competition Assays 118
9.1.1 Nuclear Hormone Receptors, HNF3 118
9.1.2 Pyrimidine Single Stranded Binding Protein 119
9.2 Supershifts 124
9.3 Effects of Mutant Elements A on Binding to Nuclear Extracts 127

10. Identification of Factors Acting in *Trans* with Element B at -14/+37 nt 129
10.1 Competition Assays and Supershifts 129
10.2 Effects of Mutant Elements B 131

11. Identification of Factors Acting in *Trans* with Element C at -124/-101 nt 131
11.1 Competition Assays 131
11.2 Supershifts Assays 133

12. In Vitro Transcription-Translation of Individual Factors 133

13. Interaction of Individual Nuclear Hormone Receptors with Element A 135
13.1 COUP-TF1 135
13.2 HNF4 137
13.3 Other Receptors 137
13.4 Receptor Interactions 139

14. Interaction of HNF3 and C/EBP with Element B at +1/+37nt 143

15. Interaction of Individual Transcription Factors with Element C at -124/-101 nt 145

16. Effect of *In Vivo* Expression of Transcription Factors on the Activity of the Antithrombin -150/+68 nt Promoter 145
16.1 Individual Effect 147
16.2 Combined Effect of Nuclear Hormone Receptors 148

CHAPTER IV: DISCUSSION 152

1. Type I Antithrombin Deficient Kindred 152

2. Transcriptional Start Site and General Control Sequences for Polymerase II Mediated Transcription in the Antithrombin Gene 154
2.1 Transcriptional Start Site 154
2.2 Eukaryotic Control Sequences for Promoter Elements 155

3. *Cis*-Acting Elements Mapped in Reporter Assays 159
3.1 Cell Lines Tested 159
3.2 5' Upstream Promoter 159
3.3 IVS 1 Element 162

4. Footprints and Deletional Analysis of the 5' Upstream Promoter 164

5. *In Vitro* Effects of Transcription-Factors Interacting with the 5' Upstream Promoter 168
5.1 Liver-Enriched Factors 169
5.1.1 HNF4 169
5.1.2 C/EBP 173
5.2 Ubiquitous Factors 174
5.2.1 Nuclear Hormone Receptors acting in Trans with Element A 174
(a) General Outline 174
(b) Receptors Identified 176
5.2.2 Single-Stranded DNA Binding Protein PYBP 180
5.2.3 Ubiquitous Factors Acting in Trans with Element B at +1/+37 nt 181

6. Effect of *In Vivo* Expression of Transcription-Factors to the Efficiency of the 5' Upstream Promoter 182
6.1 Factors with a Positive Effect 184
6.2 Factors with a Negative Effect 186
6.2.1 COUP-TFI 186
6.2.2 TRα 190
6.2.3 PPAR 190
6.2.4 HNF3 191

CHAPTER V : CONCLUSION AND FUTURE DIRECTIONS 192

1. 5' Upstream Promoter 192
1.1 Summary of the Results Presented in this Thesis 192
1.2 Prospective Studies on the Elements Acting in *Cis* with the 5'Upstream Promoter 196
1.3 Prospective Studies on the Factors Acting in *Trans* with the 5'Upstream Promoter 196
1.3.1 Liver-Enriched Factors 197
1.3.2 Ubiquitous Factors 200
1.4 Constitutive and Inducible Antithrombin Expression; Cross-Talk between Responsive Elements 203

2. IVS 1 Element 205

3. Impact of this Study for the Molecular Basis of Antithrombin Inherited Deficiency 206

REFERENCES 208

CHAPTER I
INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

The regulation of gene expression for the serine protases and their inhibitors (serpins) implicated in blood coagulation is not well understood. Transcription, translation and post-translational modifications are the three cellular events necessary for gene expression. These events precede secretion of these proteins. Transcription is the early process whereby DNA is copied into mRNA via RNA Polymerase II dependent mechanisms. Many experimental approaches can be chosen to investigate this critical step, including the use of minigene constructs, marked genes, enhancer traps, reporter genes, *in-vitro* binding assays, *in-vivo* footprinting, antisense strategies, somatic cell genetics or transgenes. The aim is to understand the conditions for constitutive and inducible gene expression mediated by RNA Polymerase II. This process is linked to the presence of critical regulatory *cis*-acting elements located on the gene itself. In addition, the interaction of these *cis*-acting elements with *trans*-acting factors is able to influence directly or to modulate transcription.

Other insights into the importance of the control mechanisms of gene expression have been provided as a result of the characterization of kindreds with inherited diseases of coagulation and thrombosis. In these cases, mutations or deletions in gene *cis*-acting regulatory sequences are responsible for a lack or an impairment of gene expression, with

an absence of mRNA or the presence of an abnormal message. This message may or may not be processed by the cell machinery. Most of these mutations involve the coding or splicing regions of the genes. In only very limited cases, mutations or deletions in the regulatory sequences of the serine protease and serpin encoding genes have been associated with defects in expression.

This thesis addresses aspects of the regulation of human antithrombin (AT) expression. Antithrombin is an inhibitory serpin largely responsible for the inactivation of thrombin and to a lesser degree of other activated serine proteases such as factors Xa, IXa, XIa and XIIa. The introductory part of this work describes the mapping of a natural deletion in a human AT gene. The presence of this deletion, located in part in the 5'region of the gene, was associated with the lack of expression of the deleted allele. This defect was found in a family who had an inherited deficiency in AT and who presented a thrombotic diathesis. The deleted allele was mapped by Southern analysis in various DNA samples from available members of the family. A combination of approaches was used to map and clone the abnormal allele at the breakpoint. Restriction analysis with partial probes synthesized by the polymerase chain reaction allowed us to walk along the gene. The abnormal allele was isolated with a biotinylated primer-streptavidin magnetic phase system. Subsequent amplification at the breakpoint location, subcloning and sequencing confirmed these findings.

During these studies, normal sequences in the 5' part of the gene were mapped extensively. The availability of this data and the fact that in the deleted allele, putative

regulatory sequences for AT gene expression were missing in the region encompassing exon 1 justified studies of this region. Therefore, the second and main part of this work describes some of the features of the regulatory sequences found. To map putative *cis*-acting elements, we selected a reporter gene assay approach. A preliminary step to this study was to confirm the start site of transcription in the AT gene by S1 nuclease assays, primer extension, and rapid amplification of 5'cDNA ends. A 6900 nt fragment from a normal genomic clone encompassing sequences upstream and downstream of the first AT exon, as well as deletions therefrom, were subcloned in a reporter luciferase vector. The cell lines selected for transfection were HepG2 cells, an hepatic transformed cell line of human origin, and three non hepatic cell lines, HeLa, Cos1, and BSC40. Transfection was mediated by calcium phosphate or liposomes. Two of the cell lines, HepG2 and Cos1, were able to sustain measurable and reproducible luciferase expression and allowed for the identification of regulatory elements for constitutive gene transcription. The first element was a classical 5'upstream promoter, located immediately upstream from the translational start site. The second element, also with a "promoter-like activity", was found in the first intervening sequence.

The upstream promoter was further characterized. Footprints and gel retardation assays with nuclear extracts prepared from the cell lines used for transfection allowed us to identify interactions between *cis*-acting elements and transcription factors. Competition assays, mutated elements and supershift assays with antibodies for binding candidates allowed us also to identify interactions with a number of ubiquitous and liver-enriched transcription factors. Direct binding of the factors identified was assessed with *in vitro*

translated material and gel retardation assays using a combination of wild-type and mutated binding elements. Finally, co-transfection experiments allowed for the deciphering of preliminary clues concerning the modulatory role of these factors for the constitutive AT expression mediated by the the 5'minimal promoter.

REVIEW OF THE LITERATURE

1 Antithrombin

1.1 Introduction

Antithrombin, a circulating 58 kDa glycosylated protein of 432 amino acids, is one of the first members of the serpin family identified (Rosenberg,1973). 50% of the AT pool is intravascular (2.5μM average concentration in plasma, or 70-130 U/dL), with a mean half life of 3 days (Pizzo,1994). AT is the major inhibitor of thrombin generated by the various pathways of the coagulation cascades. Thrombin mediated inactivation is catalysed in plasma by therapeutic agents such as unfractionated or low molecular weight heparins (Sheffield,1995). It is also activated at vascular sites by natural glycosaminoglycans such as heparan sulfates (Hatton,1994). AT inactivation occurs through cleavage at its reactive site. This step is followed by covalent bonding and esterification with cognate proteases (Sheffield,1995). The resulting complexes are believed to be uptaken by the liver *via* a receptor identical to that of other serpins such as α_1-antitrypsin (Pizzo,1989; Permulter,1990; Pizzo,1994). Whether this mechanism plays a role in the turnover and

expression of AT, such as has been suggested for other serpins-proteases complexes, is unknown. Crystallization of cleaved and native AT have both been reported (Mourey,1993; Schreuder,1993; Carrel, 1994).

1.2 Antithrombin Expression

1.2.1 Structural Features of the Gene, cDNA and mRNA

The AT gene is a single copy gene encompassing a 13500 nt region. It includes 7 exons and 6 intervening sequences, as shown in Fig.I.1. This figure was generated from published references (Prochownik,1985; Bock,1988) and as a result of extensive mapping of normal genomic clones (as is later explained in this thesis). The organization of the gene for animal species other than humans has not yet been reported. The gene locus is in the region 1q 22-25 of chromosome 1 (Bock,1985; NIH/CEPH collaborative mapping group,1992). Linkage analysis shows an association with the gene encoding the Duffy blood group antigen (Winter,1982). Various polymorphisms and repetitive sequences (mostly Alu repeats) are present in the exons and the non-coding parts of the gene (Lane,1994; Olds,1993). These polymorphisms, useful in linkage analysis, include an insertion-deletion polymorphism in the 5' upstream region (Bock,1983); the latter length polymorphism is due to the presence of a variable segment of 108 nt (called S) or 32 nt (called F) located 345 nt upstream of the AT mRNA initiation codon (Fig.I.1). It was also postulated that sequences at the 3'borders of both segments form intra-strand inverted-repeat structures with downstream sequences (Bock,1983); The significance of these

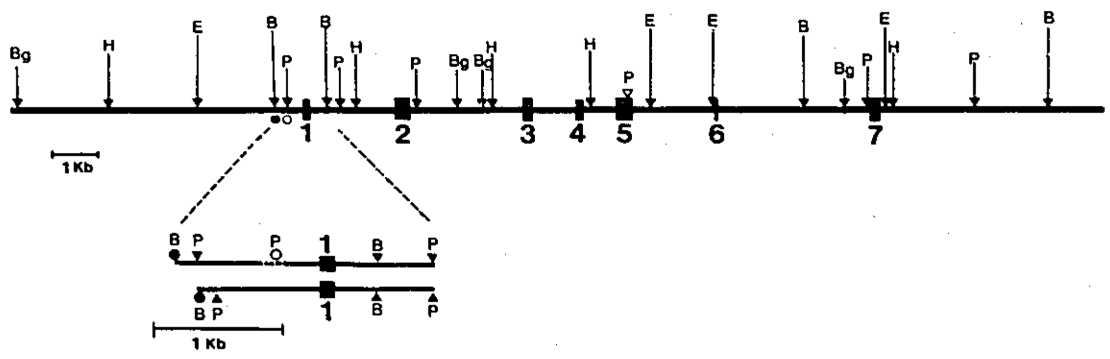


Figure I.1 :***Restriction map of the human antithrombin gene.*** The region flanking exon 1 has been expanded to show the length polymorphism. B, Bam HI; P, Pst I; E, Eco RI; Bg, Bgl II; H, Hind III. Each restricton site is indicated by an arrow. •, Bam H1 polymorphic site in the 5' region of the gene. ∘, Pst I polymorphic site in the 5' region of the gene. ▿, Intragenic Pst I polymorphic site in exon 5.

hypothetical structural arrangements for regulation of AT expression is unknown.

The human cDNA was cloned and expressed in 1982 (Bock,1982). Other groups later reported similar data (Prochownik,1983; Stackhouse,1983). AT cDNA's have also been cloned in various animals species (Stackhouse,1983; Wu,1992; Sheffield,1993; Amrani,1993; Niessen,1992) and expressed in various cell lines and cell-free systems (as reviewed by Tuddenham,1994; and by Sheffield,1995). The human cDNA contains 1392 nucleotides. The first 96 nucleotides encode a 32 amino-acid signal peptide. The next 1296 nucleotides encode the secreted protein. The termination codon is followed by 54 untranslated nucleotides, a polyadenylation signal AATAAA and a poly-A tail (87 nucleotides after the stop codon). The post-translational events preceding AT secretion are not known at the moment.

The major source of AT mRNA is the liver (Tremp,1995). Whether or not the kidney expresses AT constitutively is still debated (Prochownik,1985; d'Souza,1987; Sheffield,1993; Tremp,1995). Immunoreactive AT has also been detected in the vascular endothelium (Prochownik, 1985; Felsch,1994) and in the brain (Deschepper,1991). Interestingly, AT has been recently implicated in the pathogenesis of Alzheimer's disease (Kallaria,1993). In addition, the latest refinements in the polymerase chain reaction (PCR) have allowed for the detection of an ectopic AT expression in circulating white blood cells following amplification of total cellular RNA by reverse transcription (Berg,1992). Ectopic transcript analysis is of primary importance for the characterization of defects of AT gene expression in cases of inherited AT deficiencies. During the course of this work, the various tissues and cell lines utilized for transcriptional assays were tested for endogenous

AT expression; reverse transcription coupled to PCR amplification of polyA-mRNA (with primers specific to the human cDNA) allowed us to confirm the presence of an AT message in human liver mRNA as well as the human hepatoma cell-line HepG2 (Fig.III.10). In addition, the secretion of immunoreactive AT was also detected by ELISA (data not shown).

1.2.2 Regulation of Antithrombin Gene Expression

The expression of the AT gene, at least at the hepatic site, is believed to be constitutive. Unlike related serpins such as heparin-cofactor II or α_1-antitrypsin, AT is not known to behave like an acute-phase reactant. Nevertheless, developmental and hormonal factors are able to influence its expression (de Souza,1987; Amrani,1993). Effectively, AT levels are very low in foetuses and rise slowly in childhood to reach adult levels (Lao,1990). In addition, the expression of AT deficiencies is also age dependent as most cases are only clinically evident after the age of 15, suggesting hormonal and extracellular control mechanisms. For several decades, oral contraceptives have been known to decrease plasma AT levels (Laurell,1978). Conversely, androgens have been shown to increase these levels in plasma (Ford,1990) and, in fact, have been utilized for "boosting" AT expression in case of inherited AT deficiency (Winter,1984; Eyster,1985; Ford, 1990). Such an effect has also been reported for Protein-C and Protein-S (Ford,1990). During the course of this work, we observed that a synthetic derivative of ethisterone, *Danazol*, was able to increase the secretion of immunoreactive AT by HepG2 cells (data not shown). Whether this observation was directly related to transcriptional induction is not known.

Preliminary transfection experiments with a reporter plasmid including AT sequences located at -1100/+68 nt did not enable us to observe an increase in reporter activity (luciferase expression) in the presence of *Danazol* (unpublished observations). Further investigations into other regions of the gene or into putative responsive elements isolated from the -1100/+68 nt was not attempted.

A few reports have addressed transcriptional control mechanisms directly underlying AT gene expression. An attempt was made to compare the 5'upstream sequences of AT genes from species other than humans but little information is available in the literature. Analysis of data available for the mouse, chicken and sheep sequences, even if very partial, did not show obvious homologies and similarities with their human counterparts with perhaps the exception of a G,A,C-rich region, 40 nt upstream of the first translation codon in the chicken upstream sequences (Amrani,1993). Mapping of the transcriptional start site of the AT gene has been done in tissues and transformed cell lines. In human liver, primer extension experiments have located a major single product corresponding to a C, 72 nt upstream from the first translation codon (Prochownik, 1984). Similar experimental approaches by Bock located the start site further downstream, 47 nt before the initiation codon, but this location was not confirmed in further studies (Bock,1983; Bock,1988). The trancriptional start site of the AT mRNA has also been determined in two types of cultured cells (HeLa and Cos1 cells) transfected with an minigene containing 304 nt of 5 'upstream sequences, followed by the coding sequences, a truncated portion of the first intervening sequences, and the 3' untranslated region (Prochownik,1984). In these two cell lines, multiple products initiating 5 nt and 8 nt

downstream from the presumed liver start site were found. Further experiments with these minigenes also showed an accurate splicing for the transcripts. Interestingly, S1 nuclease mapping experiments revealed an alternate splice acceptor site within IVS1 which was utilized at 30% efficiency of the true splice site. The truncated polypeptide generated did not, nevertheless, contain the hydrophobic region of the AT signal sequences.

The regulatory mechanisms underlying constitutive transcription have been addressed in only two studies. The first study, using CAT-reporter assays, showed that the transcriptional effects mediated by the SV40 minimal promoter were enhanced by AT sequences (Prochownik,1985); the sequences tested were the first 340 nt of the 5'upstream region, the first exon and 304 nt of the first IVS. 5'deletion of 120 nt in the AT upstream region did not modify enhancer strength or orientation independent activity. The responsive cells tested in Prochownik's study were Cos1, Alexander human hepatoma and M2 myeloma cells. Also in this study, two putative AT regulatory elements were found by database homology; the first element, located in the human or the murine Jk-Ck intron of the immunoglobulin -K chain gene, is homologous to a putative initiator element at -43 nt in the AT gene; the second element, in the murine Jk-Ck intron also, is homologous to the AT region rich in GAGA nucleotides and located just upstream from the presumed transcriptional. Both Jk-Ck intron regions also bear similarities to viral enhancer elements (Prochownik,1985). The second report was provided by a group studying the regulation of the human transferrin proximal promoter (Ochoa,1989). An element of this promoter, TF-DR I (TCTTTGACCT) at -480/-451 nt was found to be homologous to the -92/-68 nt region of the AT gene. Both elements were shown to interact with liver-specific nuclear

extracts in footprints and in gel retardation assays but the actual transcription factors were not identified in this report.

2 Transcriptional Machinery in Eukaryotes

2.1 Introduction

The attempt in this thesis was to identify mechanisms able to influence AT gene expression. More precisely, the focus of the investigative efforts was on the regulation of gene transcription. In eukaryotes, transcription of mRNA encoding genes is a multistep process. Early chromatin events with unwinding of the DNA make RNA polymerase II DNA dependent and associated factors accessible to regulatory DNA sequences (or *cis*-acting elements) (Lewin,1994). This step is followed by transcription *per se*, which is mediated by the binding of initiation factors to *cis*-acting core promoters located in the immediate vicinity of the mRNA start site(s). The factors for initiation are RNA Polymerase II (which cannot initiate transcription alone in eukaryotes) and an array of general initiation factors, TFII A to J, or GTF (Buratowski,1994; Koleske,1995). The efficiency and the rate of initiation are also modulated by transcriptional activators, coactivators and repressors, all of which may interact directly either in the vicinity of the start site in the promoter, or at a distance, in enhancers often but not always located upstream from the promoter elements (Drapkin,1993; Tjian,1994). The relationship between the general and auxiliary factors is only partly understood and will be discussed further. This relationship results in silent, constitutive or induced transcriptional states.

Classically, inducible transcription is the response to developmental or extracellular signals, often involving enhancer sites. Constitutive transcription (often improperly called basal) refers to a steady-state of transcription. Nevertheless, as Drapkin states, "the cell must overcome a derepressed state in order to activate transcriptional mechanisms, even under basal conditions" (Drapkin,1993). In addition, more and more frequent reports of links between transcription, cell-cycling and DNA-repair suggest that transcription is not an isolated event (Drapkin,1993). The following sections review some recent aspects of transcription regulation and current knowledge on the proximal and distal *cis*- and *trans*-acting elements implicated in this process.

2.2 Core Promoter Elements

Traditionally, two types of basal elements have been described in eukaryotic promoters. This dichotomy is based upon the fact that transcription is most often triggered at specific DNA sequences or TATA boxes. This also allows for the separation of promoters in two classes: "TATA plus" or "TATA-less". TATA boxes have the general consensus TATANNNN (N= adenine rich). Nevertheless, and particularly in mammalian cells, TATA-box motifs are quite heterogenous; it has also been shown that this heterogeneity exists at the functional level (Bucher,1990; Wefald,1990). TATA boxes are located approximately 30 nt upstream of the start site(s) of transcription and determine the site and the directionality of mRNA initiation as well as the basal levels of mRNA synthesis.

2.2.1 Architecture of Regular ("TATA plus") Promoters

Core elements in regular promoters have been mostly defined *in vitro*. They are sufficient to initiate the basal transcription process and can function independently in the absence of regulators. Deletion of the TATA box often results in spurious initiation and low levels of transcription. TATA boxes are DNA binding targets for the initiation factor TBP, or TATA Binding Protein (Buratowski,1994). Early binding of TBP nucleates the formation of the initiation complex by making protein-protein contacts with other associated factors. The complex of TBP and associated factors (TAF's) is named TFII-D (Gill,1992; Goodrich,1994). The X-ray crystal structure of TBP has been resolved by Nikolow in 1992. It resembles a saddle, with the DNA binding domain in the inner part and the TAF and GTF interaction-domain in the outer surface. Binding of TFII-D induces topological changes and distortion in the DNA. Many mechanistic models have been proposed to explain the assembly of the various GTFs around TBP and DNA. The classical view, as recently reviewed by Buratowski, is based upon *in vitro* reconstitution experiments (Buratowski,1994). An initial complex is formed by binding of TFII-D to the TATA box, which then recruites TFII-B, polymerase II, TFII-F and then TFII-E and H. After the initial complex is assembled, an ATP-dependent activation preceeds transcription *per-se*. TFII-A associates with TBP at an early stage, and even if not essential, this factor can counteracts the effect of negative regulators of transcription. Some promoters necessitate all or only some GTF to function *in vitro* (Parvin,1994). In addition, the nature of these factors and their order of assembly also influence transcript elongation (Drapkin,1993). A more dynamic view of transcription, based upon *in vivo* observations,

is also emerging (Koleske,1995). This view suggests that the assembled complex, Polymerase II, TFII-F, B, H and SRB (Suppressor of RNA Polymerase B) is recruited by promoters at which the TFII-D complex is already bound. Therefore, complex assembly in this latter model is no longer considered as a step-by-step process. "TATA-plus" promoters also include other control sequences, such as CAAT boxes (located around -70 to 100 nt) and GC rich regions. These elements target transcriptional regulatory factors which regulate the efficiency of transcription initiation such as CTF/NF1, C/EBP, NFY and SP1 (Johnson,1989; Chodosh,1988).

2.2.2 Architecture of "TATA less" Promoters

"TATA less" promoters, which are not well-known, are also able to specify the site of initiation and to control transcription initiation. They form a very heterogenous group including initiator elements, AT-rich nucleotide regions with no real consensus for a TATA box, GC rich regions or sequences with none of the above features (Weis, 1992). The initiator element (or Inr) has been originally identified in the terminal-deoxynucleotidyl-transferase gene as a 17 nt element able to position transcription initiation *in vitro* and *in vivo* (Smale,1989). Inr usually encompass the transcriptional start site regions or are less frequently found in close vicinity to the start site. They have the consensus PyPyCANT(A)PyPy in mammalian genes. However, this consensus is not absolute and a classification in families has been proposed based upon Inr sequence variations and requirements for general and specific transcription factors (Weis,1992). Inr can be found alone or in association with a TATA box or a GC box; the protein C

promoter, for example, contains a TATA-box and a consensus for an Inr (Plutzky,1986). When both types of initiating elements co-exist in a gene, a functional hierarchy TATA>Inr is often present (Weis, 1992; Martinez, 1994). Inr elements have been shown to interact with the general factors TFII D, B, F, TAF and RNA Polymerase II (Pugh,1991; Martinez,1994). In addition, specific factors binding at the initiator site have been identified, such as the the factor binding to the site of capping (or CBF) or the factors YY1 (or NF-E1), HIP-1, or TFII-I (Weis,1992; Roy,1993; Martinez,1994). TFII-I is a 120 kDa protein immunologically related to USF (Upstream-Stimulating-Factor) (Aso,1994). Therefore, many pathways and mechanisms explaining transcription initiation have been proposed, and it has even been shown that TFII-D, at least *in vitro,* can bind to non-consensus TATA-boxes (Pugh,1991; Zhou,1992).

2.3 Enhancers, Activators, Coactivators, Repressors

Activation of transcription involves interactions of regulators (activators, repressors, coactivators) at the promoter and enhancer sites. This directly influences the assembly of the general and auxiliary factors at the site(s) of initiation and the rate and the efficiency of transcription.

2.3.1 *Cis*-Acting Elements (Enhancers)

Enhancer elements are located by definition at a distance from the promoter (Atchison,1989). In fact, enhancers can be part of a modular promoter --such a case is found for many genes expressed in the liver-- or they can be located very far from a

promoter. In the latter case, they are called upstream-activating-sequences (or UAS). An enhancer effect is defined in reporter assays by the ability to increase transcription mediated by an endogenous or a heterologous promoter. Up to recently, this effect was believed to be orientation independent (no directionality). However, architectural constraints are more and more often shown to be critical for proper function of regulators (Eissenberg, 1991; Tjian,1994). Most enhancers are modular, containing binding sites for various factors (which can in addition homo- and/or heterodimerize). They can synergize or cooperate with adjacent sites or other promoter elements. The interaction with initiation elements mediated by regulators and coactivators often induces topological changes in DNA structure (such as DNA bending and loop-out to allow contact with the initiation machinery). It is classically admitted that extracellular signals and developmental switches target preferentially enhancers.

2.3.2 *Trans*-Acting Regulators

(a) Coactivators -- Coactivators provide a link between the initiation factors and the activators. Two classes of coactivators have been defined, based upon whether or not they interact with TBP (Tjian,1994). For example, the activators SP1, CTF/NF1 interact with TBP by the TAF(s) (Goodrich,1994). Coactivators of several members of the nuclear-hormone-receptors have also been described (Tsai,1994; Teng,1995). These coactivators can provide bridging molecules and/or change conformation, phosphorylation or properties and activity of a factor to allow transcription (Tsai,1994; Gerber,1995).

(b) Activators-- Activators form a very heterogenous group of proteins with a

modular structure (Tjian,1994). They are classically composed of a DNA binding domain, a multimerization domain and a transactivation domain (Ptashne,1988; Pabo, 1992). Some of these domains may be located in more than one site of the molecule. This modular structure multiplies the possibilities of interaction with other factors, ligands and DNA. The DNA binding domain confers binding specificity but it has been proven many times that binding is not sufficient for activation of transcription. Site specific recognition always involves interactions with bases of the major groove and with the DNA backbone. Hydrogen bonding and side-chain contacts are critical for these mechanisms. Many types of DNA binding domains have been described and will be detailed further for the liver specific factors involved in this study. Factor multimerization allows for the formation of homo- and/or heterodimers and can have either a positive or negative effect on transcription. This property creates unlimited possibilities for new regulatory patterns important for tissue specific responses and responses to developmental and external stimuli. It also directly influences proper binding; such a case is found, for example, in the superfamily of nuclear hormone receptors. The activation domains are often amphipathic α-helixes rich in negative charges. They can be acidic, glutamine rich, or proline rich regions, any of which can directly or indirectly influence transcription initiation upon interaction. Activators with strong transactivation potential often control enhancer function. Transactivation domains are not as specific as the DNA-multimerization domains and can be interchanged to create hybrid activators with new DNA binding or dimerization domains (Ptashne,1988; Mitchell,1989). In addition they are strongly conserved throughout evolution. Many routes for transcription activation have been

reported, such as *de novo* synthesis of a given activator, ligand binding, dissociation of an inhibitory protein, dimerization, DNA bending capacity, direct interaction with the preinitiation complexes and post transcriptional control (Ptashne,1990; Drapkin,1993). In addition, *in vivo* and *in vitro* differences are often noticed in activation potential. The case later discussed of the nuclear hormone receptors will illustrate these differences.

(c) Repressors -- Repression of transcription can occur at various levels (Cowell,1994). First, an active role in this process is played by chromatin, the physiological template for transcription. Chromatin packaging into nucleosomes as well as the core histones of nucleosomes is an obstacle for transcription (Lewin,1995). However, the rearrangement of nucleosomes, the involment of scaffold-associated regions and of topoisomerases, the presence of nucleosome-free regions (nuclease hypersensitive sites) and the ability of some classes of activators to directly modify chromatin structures, provide examples of "open windows" that allow transcription (Laemmli,1992; Lewin,1995). In fact, the displacement of histones or the competition between histones and transcription factors (such as the nuclear hormone receptors, the factors of the NF family and the heat-shock transcription factors) for DNA accessibility -two mechanisms of derepression- are required for initiating transcription in certain genes (Felsenfeld,1992). Second, genuine repressors can play either a passive or an active role in the transcription process. Passive repressors down-regulate the potential of activators either by direct competition on DNA binding sites, by steric hindrance, by heterodimerization (for example, between members of the nuclear hormone receptors), by binding of an inhibitory cofactor or by post-transcriptional modifications. Active repressors have a direct effect on

transcription by interfering with formation of the initiation complexes. In addition, they have been shown to inhibit the activation potential of promoter and enhancer bound factors. YY1, a factor associated with initiator elements, has such a role (Roy,1993; Martinez,1994).

Repression or activation of transcription are also often a matter of abundance of a given factor at an activation site and depends on promoter context (binding sites content and architecture). Furthermore, both the synthesis of a transcription factor and its activity are regulated by independant mechanisms; regulation of synthesis, for example, is achieved at the transcriptional and post-transcriptional levels. As well, regulation of activity can be obtained by a myriad of mechanisms, including dimerization, allosteric interactions, phosphorylation, or cofactors. As a result of the combination of both regulation pathways, a given factor will often have a dual potential for gene expression involving repression or activation of transcription.

3 Liver Specific Promoters

3.1 Introduction

The mechanisms that generate high levels of gene expression in the liver are still poorly known. The hepatocytes contribute in majority to the transcriptional activity of the liver, and it is well recognized that the levels of expression via polymerase II are mostly controlled at the mRNA levels (Herbts,1990; Derman,1981). The control of transcription in the liver depends on the binding of transcription factors (those identified to date are

mostly activators) to *cis*-acting elements, enhancers and promoters located most of the time at the immediate vicinity of the transcriptional start(s) site(s). Enhancers located far upstream of the reading frame have also been identified, but more rarely; for example, in the albumin and the α_1-antitrypsin genes (Pinkert,1987; Tripodi,1991). In most cases, such elements have been identified in transgenic experiments. Two types of factors can bind to liver *cis*-regulating elements: general factors at the vicinity of the TATA-boxes (or initiator elements) and regulators of transcription (mostly transactivators). These transcriptional regulators are ubiquitous or are expressed at high rate in the liver. For example, the ubiquitous factors AP1, NF1 and NFY are often found in liver promoters at the vicinity of the transcription initiation complex (Ciliberto,1993). A recent list of the various families of liver-enriched factors is given in Table I.1. This table was reprinted from a recent review by Sladek (1994). To date, known activators are able to bind to their cognate elements under constitutive and/or induced conditions of transcription. Well known examples of transcriptional induction in the liver are the acute-phase response and the hormonal control by steroid hormones, glucocorticoids, growth-hormone, somatostatin, insulin, or cyclicAMP.

3.2 Liver-Enriched Factors

3.2.1 HNF1

HNF1 was originally characterized by its ability to bind to the β-fibrinogen and α_1-antitrypsin promoters (Courtois, 1987). Its consensus for DNA binding is the palindrome

		LIVER	INTEST	LUNG	STOM	KIDNEY	BRAIN	SPLEEN	HEART	FAT	OVARY	TESTES
HNF1	α	++	++	-	+	++	-	+/-	-		-	
	β	+/-	+	+	+	++	-	-	-		+	
C/EBP	α	++	+	+		+/-	+	+/-	+/-	++		-
	β	++	++	++	+/-	+/-	+/-	+/-	+/-	++	+	+/-
	δ	+	++	++		+	+/-	+	+			
HNF3	α	++	+	+	+	-	-	-			-	-
	β	++	+	+	+	-	-	-			+	-
	γ	++	+	-	+	-	-	-			+	+
HNF4		++	++	-	-	++	-	-	-	-		

Table I.1 : ***Tissue distribution of liver-enriched transcription factors.***
(from Sladek, 1994).

GTTAATNATAAC; as expected, this factor binds to DNA as a dimer. The two major HNF1 isoforms, α (or LFB1, or APF) and β (or LFB3, or vHNF1) have been extensively characterized; they share dimerization and homeodomains but differ in their transactivation domains. Their tissue distribution is also different as seen in Table I.1 (Cereghini,1990; Mendel,1991; Xanthopoulos,1991). HNF1α is primarily present in the liver, intestine and kidney whereas HNF1β is predominant in the kidney. HNF1α is a 88 kDa protein with an amino-terminal DNA binding domain (an unusually long homeobox, XL-HD, recognizing various A,T-rich nucleotide sequences and half-palindromic sites). The latter domain is followed by a region slightly homologous to a POU domain (Φ-pou) and essential for the specificity of DNA binding to the HNF1 consensus. Two transactivation domains have been mapped in the HNF1 molecule: the first is located in the carboxy terminus (Frain,1989; Mendel,1991-1) and the second is adjacent to the XL-HD domain (De Simone,1993). HNF1 (either α or β) is described mainly as a transcriptional activator (Ryffel,1989; de Simone, 1993). Dimerization occurs by interaction of leucine residues aligned on the same side of a kinked α helix and located in the extreme amino terminus of the HNF1 molecule. The formation of both homodimers and 1α-1β heterodimers is documented (De Simone,1992). In the liver (and likely in relation with tissue specificity), high levels of α homodimers are associated with no α-β heterodimer formation. Very recent progress in the field of HNF1 research include the identification of a 11 kDa cofactor, DcoH, which regulates heterodimerization and, to a certain extent, transactivation of HNF1 (Mendel,1991-2; de Simone,1993). This cofactor has no direct binding to DNA or transcriptional activity *per se*. Very recently, additional members of

the HNF1 family have been identified (Bach, 1993). Nevertheless, binding, heterodimerization, and transactivation properties of these new HNF1 isoforms remain to be determined. Whether or not these new factors are also involved in development by implementing an "epithelial" program, similar to HNF1α or β, is unknown to date (Ciliberto,1993).

3.2.2 HNF3

HNF3 was initially described as a 49 kDa hepatocyte-enriched transcription factor interacting with the transthyretin gene promoter and with the α_1-antitrypsin proximal enhancer (Costa,1989). To date, up to 14 members of this family have been reported in the literature (Clevidence,1993; Pierrou,1994). The three isoforms α, β and γ have been cloned and extensively characterized (Lai,1991; Costa,1991). Their DNA binding domain, located in the amino terminus, is highly conserved. This domain remains constant upon cell dedifferentiation-redifferentiation. It is also highly homologous to the Drosophila homeotic gene product *forkhead* (Lai, 1991) with a bundle of three helices and a short β-strand between helices H1 and H2 (helix-turn-helix motif). Similar to other members of this homeodomain family, HNF3 binds DNA as a monomer. Upon binding, HNF3 interacts with the sugar-phosphate backbone on either side of the DNA. In contrast to their binding profile, the tissue distribution of the three HNF3 isoforms is variable (Table I.1); nonetheless, the expression of this protein family is restricted to cells derived from the primitive gut. HNF3 members have been mainly --but not exclusively-- described as transcriptional activators, both *in vitro* and in co-transfection experiments (Ciliberto,1993).

Very recently, seven new human *forkhead* proteins (or FREAC) were obtained after reverse PCR (Pierrou,1994); within these new proteins, FREAC-3 and -5 are liver-enriched. They are highly homologous to the known HNF3 members and slightly differ from each-other in their binding site specificity (RTAAAYA) and their DNA bending capacity. Similarly to many liver-enriched factors HNF3 has been shown to play an important role in organ and tissue differentiation (Ciliberto,1993).

3.2.3 HNF4

HNF4 was identified originally in rat liver nuclear extracts as a 54 kDa factor critical for the expression of the transthyretin and apolipoprotein C-III promoters (Costa,1989; Leff,1989). HNF4 mRNA and protein -whose levels are parallel- are found mostly in liver, kidney and intestine. A detailed tissue distribution of this factor is shown in Table I.1. HNF4 is phosphorylated *in vivo* but the implications of this modification have not been evaluated. HNF4 is a member of the nuclear hormone receptor superfamily, a highly frequent and diversified family of transcriptional regulators (Sladek,1990). Similarly to receptors of this family, HNF4 displays a DNA binding domain encoded by two cysteine containing zinc-fingers and a carboxy-terminus ligand binding module with hydrophobic repeats important for ligand interaction, protein dimerization, and transactivation. Moreover, additional transactivation regions map to the extreme ends of the molecule. A hinge-region, located between the DNA binding and the transactivation domain, supposedly plays a role in nuclear translocation. HNF4 binds as a homodimer and has not been reported yet to bind as an heterodimer with other members of the nuclear

hormone receptor superfamily. Nevertheless, direct interactions between HNF4 and transcription factors from families other than nuclear receptors (leucine-zipper for example) have been mentioned in the literature (Metzger, 1993). The ligand of HNF4 (if any and if not an endogenous cell metabolite) is still unknown. HNF4 is therefore described as an orphan receptor such as COUP-TFI or PPAR for other examples (Sladek,1994). It has been strongly suggested that the factors LFA1 and HNF2 were either identical, subtypes or isoforms of HNF4. Supershift-assays have shown that LFA1 and HNF2 were HNF4 related (Ramji,1991; Rangan,1990; Sladek,1994) but these findings are nevertheless still debated by some authors (Ciliberto,1993). Three rat HNF4 isoforms α, β, and γ have been isolated: they are believed to have originated from alternative splicing. Their distribution varies in relation to the stage of differentiation and the type of tissue (Sladek,1994). The differences between the various isoforms are present either in the N-terminus ot the C-terminus of the molecule. The impact of these differences for DNA binding or transactivation has not been evaluated. Very recently, cloning and sequencing of cDNA's encoding human HNF4 have also indicated the presence of two isoforms in human liver (Chartier,1994). Chartier's work has indicated that the main differences between rat and human HNF4 are present in the dimerization domain and the C-terminus part of the molecule. His work has shown also that the largest isoforms of human and rat HNF4 have a high level of similarity in the DNA binding domain.

Many HNF4 responsive genes have been identified: they are involved in fatty acid and cholesterol metabolism, gluconeogenesis, amino-acid metabolism, metal detoxification, hepatitis B infections and blood coagulation. HNF4 responsive elements

with homology to the AT 5'upstream region are located, for example, in the apolipoproteins, the coagulation factors IX, VII, and X, the transferrin, the α_1-antitrypsin and the transthyretin genes. The consensus binding site for HNF4 contains two direct motifs of the AGGTCA motif separated by one nucleotide. In fact, the compilation of binding sites to this factor gives a quite variable consensus and we will rediscuss HNF4 binding properties latter in this thesis. HNF4 is highly conserved from insects to mammals; it plays a vital role in early development of the liver, kidney and gut (Zhong,1993). Owing to an early embryonic lethality with ectodermal cell death, no HNF4 knockout-mouse is available as a model of HNF4 function *in-vivo* (Chen,1994). DNA binding activity of HNF4 is influenced by the stage of differentiation of hepatoma cells to a certain extent and this phenomenon has been observed also for other members of the nuclear-hormone and leucine-zipper superfamilies. Therefore, the extinction of many hepatic functions and their reexpression is correlated to variations in the expression of HNF4 (Sladek,1994).

3.2.4 C/EBP

Initially identified as a factor binding to the CCAAT consensus, the first member of this family isolated, C/EBPα, was a heat stable factor able to bind to the dyad symmetrical ATTGCGCAAT motif present in animal virus genes (Landschulz,1987). Subsequently, this factor was also implicated in liver transcription (Lichtsteiner,1988; McKnight,1992). In fact many members of this family have been cloned and to date the 4 C/EBP subtypes α, β, δ, and γ have been characterized (de Simone,1992). C/EBPα and

C/EBPβ are present in liver, adipocytes (Table I.1) and in other tissues that metabolize at high rate lipids and cholesterol components (gut, lung, adrenal glands and placenta, for example). DBP, which is not beleived to be a genuine C/EBP family member but a member of the TEF/HLF family (Mueller,1991; Falvey,1995) is to date the only factor related to this family that shows mostly a liver specificity; DBP is also unable to heterodimerize with other C/EBP family members and its expression follows a strong circadian rhythm. It is now well admitted that the consensus binding site of all these factors is not limited to perfectly symmetrical CAAT motifs and the proposed sequence A(G)TTGCGC(T)AAT(G,C) reflects better the promiscuity of binding of C/EBP and related factors. In addition high and low affinity CCAAT binding sites have been identified for C/EBPα (McKnight,1992). The C/EBP family members are all structurally related and share extensive similarity in their DNA binding and dimerization domains. They can also homodimerize and heterodimerize with each other and with factors of other families. C/EBP members are synthesized by different genes located on different chromosomes. To further complicate the identification of these factors, a number of literature reports mention the same factor under a variety of names. In the case of C/EBP β, for example, related (or identical) factors have been described under the name of LAP, NF-IL6, IL6-DBP, VBP, CRP2, AGP/EBP, Ig/EBP-1, NF-IL6B, CELF, CRP1 and CRP3 (Descombes,1991; Ciliberto,1993). Most factors enumerated in this list play a regulatory role not only in the expression of genes constitutively expressed in the liver but also in the acute-phase response mediated by interleukin 6 (IL 6) for genes such as α_1-antitrypsin, fibrinogen and α_2-macroglobulin. Of interest, the possibility of alternative translation of a given gene, for

example C/EBP α and β, results in the formation of products truncated of their N-terminus. A C/EBP α 30-kDa and LIP are two such isoforms for C/EBP α and β, respectively (Descombes, 1991; Lin, 1993).

The N-terminus region of C/EBP, a region distinctive of the various C/EBP isoforms and subtypes, encodes a complex transactivation domain. Mapping-experiments have identified several active regions --with either a positive or a negative effect on transactivation-- near the extreme NH2-terminus as well as from amino-acid 150 to 196 (Friedman,1990; McDougald,1995). Interaction between the various regions and phosphorylation also complicate transactivation events. It is believed that most forms of C/EBP behave like transcriptional activators in reporter assays but this is not always the case (Friedman,1990; Cilberto, 1993).

The C-terminus region of C/EBP supports binding-dimerization and is one of the first model of "b-ZIP" proteins binding to DNA characterized. Its dimeric binding to DNA is mediated *via* a leucine zipper important for dimerization, and a direct binding domain. The leucine zipper, located close to the C-terminus, is formed by heptad repeats of leucines within a 30 amino-acid sequence capable of forming an α-helix (Landschulz 1988). This domain supports not only dimerization, but also affects the sequence-specific recognition of DNA. The DNA contact surface is highly basic and located immediately upstream to the leucine repeats. Phosphorylation appears to modulate dimerization and binding to DNA (Ciliberto, 1993). Negative cofactors for heterodimerization have also been identified for the C/EBP family, such as the truncated factor mentioned above, LIP (or liver inhibitory protein) for C/EBP β (Descombes,1991) and CHOP (Ron,1992). LIP

is a repressor of LAP (C/EBP β) activation by formation of LIP/LAP heterodimers. CHOP, also a C/EBP family member, can heterodimerize with C/EBP α and β. As a result, the binding of the later factors to their responsive units is prevented; CHOP-C/EBP heterodimers can nevertheless bind to non consensus binding sites.

Deletion of Dmc\EBP in *Drosophila* (the equivalent of C/EBP) results in late embryonic lethality (Montel,1992). Moreover, the very recent "cut-through" findings in C/EBPα knockout-mice have provided evidence for the vital requirement of this factor in development, by establishing and maintaining energy homeostasis in neonates. In mutant mice (lacking C/EBP) the hepatocytes and adipocytes do not accumulate lipids (Wang, 1995). In cell culture systems, C/EBP plays a role in establishing and maintaining the cellular differentiated state (De Simone,1992; Ciliberto,1993; MacDougald, 1995).

3.2.5 Other Liver Enriched Factors

A number of factors have been identified in extracts from crude liver or hepatoma cell lines and do not appear to be identical to the HNF1, HNF3, HNF4 and C/EBP families. For example, the apolipoprotein C-III promoter region -86/-74 is recognized by two new transcriptional activators: CIIIB1, a 41-kDa heat stable protein binding to CAGGTGAC motifs (Ogami,1991) and CIIIB2 (NF-BA1), a 60-kDa labile protein which shares the central motif GTGAC for binding. The latter factor has also been identified in other apoliproteins promoters (Kardassis,1990).

BAP is an ubiquitous transcriptional 58-kDa activator binding to the B-activator binding site of the *Xenopus* B1 vitellogenin gene 5'TGCACATGCGC3' (Kugler,1990).

This factor is different from YY1 and from the Upstream Stimulating Factor, USF/MLTF, also binding to this site (consensus CACGTG).

GATA binding factors have been identified in rat hepatocytes (GATA(T)). They are homolog of GATA 1 and 4 and have been shown to bind the rat serine dehydratase gene promoter (Matsuda,1994). Preliminary studies suggest that some of these factors could be negative regulators of liver transcription (Matsuda,1994).

The sterol responsive element binding factors 1 and 2, SREB1 and SREB2, are basic loop-helix leucine-zipper proteins that control transcription of the low density lipoprotein receptor gene. These proteins, mainly studied in adipose tissue, are newly identified transcription factors also frequently found in liver specific control sequences (Yokoyama,1993).

3.3 Single Stranded Binding Proteins and Liver Expression

The factors described in the two previous subheadings bind to double-stranded DNA. In addition to these families, a number of reports have detailed the identification and cloning of transcriptional regulators binding to single-stranded DNA and to RNA. Targets for this class of factors have been reported in a number of liver expressed genes.

PYBP (or Pyrimidine Binding Protein) was isolated during the identification in rat nuclear extracts of proteins binding to the the human transferrin promoter (Brunel,1991). One element of this promoter, located in the antisense-strand of PR-I, TCTTTGACCT, is highly homologous with a site of the human AT gene at -72 nt. PYBP is an ubiquitous, highly conserved 58-kDa protein (a doublet of 57 and 59-kDa in SDS-PAGE). More recent

reports have shown PYBP homologies with a mouse nuclear protein (Bothwell, 1991) and with pPTB (in fact PYBP and pPTB show the same antibody specificity; Brunel, personal communication). Primary structure indicated that PYBP contains four sequence repeats, each with a typical recognition motif found also in a number of RNA binding proteins such as the human ribonucleoprotein (hnRNP-L) (Gil, 1991). Once bound, PYBP is able to disrupt the two strands of DNA and therefore to unwind DNA. pPTB (or P62) was detected in HeLa cells by UV cross-linking to pre-mRNAs by Garcia-Blanco (1989); this protein interacts specifically with the polypyrimidine tract of introns of pre-mRNAs. It has been suggested that pPTB play a role in the early stages of splice-site recognition and pre-spliceosome assembly (Garcia-blanco,1989). Cloning of pPTB allowed the isolation of 3 classes of cDNA's which probably reflect alternative splicing (Gil,1991). The rat homolog of pBTB has also been identified as a protein binding a liver specific enhancer element of the rat amino transferase gene, TCTTTGATCT (Jansen-Durr,1992). This factor was detected by an antibody raised against PYBP.

A clear role for these single stranded binding factors in transcription has not been demonstrated to date but a number of reports suggest their importance. Brunel has suggested a role of activator for PYBP in a G-free cassette assay (Brunel,1991). These proteins have a high sequence specificity for DNA recognition, unlike other RNA binding proteins. They have been isolated in the enhancers of tissue specific genes, liver, muscle and adipocyte-specific genes, for example, and in the SV40 early promoter. In this later case, the binding protein, H16 (highly homologous to a RNA-binding protein of hnRNP particles) has also been shown to be able to stimulate strongly the activity of RNA

polymerase II (Gaillard,1994). Furthermore, proteins of this family modulate binding of steroid hormone receptors notably in the ovalbumin and vitellogenin genes (Mukherjee,1990; Nordstrom,1993).

3.4 Specificity of Gene Expression in the Liver

Up to the present time, no liver-enriched factor has been shown to be exclusively expressed in the liver. It is believed therefore that the interplay of various liver-enriched and ubiquitous factors in a given regulatory region is critical to the efficiency of transcription in the liver. But qualitative and quantitative differences in the factor's levels of expression and a different distribution among the subtypes and isoforms of a multigene family could also contribute to the specificity of expression at the hepatic site. In the liver also, the *cis*-elements critical for transcriptional activity are clustered in the first 500 nt upstream from the transcriptional start site: in all genes analysed to date, at least one binding site for a liver-enriched factor is present in the first 100 nt upstream from the TATA box or the initiating region. The orientation of these binding sites can vary without affecting transcriptional efficiency. Table I.2, based upon a recent review by Ciliberto (1993), summarizes findings for a number of liver specific promoters showing some degree of homology with the AT promoter region. This table also illustrates the modular nature of these promoters. After dissection by deletional analysis in reporter assays, some elements often do not function independently but require cooperative interactions (Ciliberto,1993). As well, additive effects between the various sites have also been described. Furthermore, the same element can be occupied by various factors, for

	HNF1α	HNF1β	HNF3	HNF4	C/EBP	ubiquit	other liver
human albumin	+	+			+	NFY NF1	
mouse albumin	+	+			+ LAP,DBP	NFY CRP	
rat aldolase B			+			NFY	
human α_1 antitrypsin	+	+		+		AP1	
human apolipoprotein AI				+	+	ARP1 RXR	
human apolipoprotein A2	+			+	+		
human apolipoprotein B	+	+	+	+		ARP1	
human apolipoprotein CIII				+	+	COUP ARP1	
rat CPS					+		
hepatitis B virus enhancer				+			HNF5
rat PEPCK	+			+	+	NF1 AP1	AP2
human transferrin			+	+	+	NF1 COUP	
mouse transferrin	+		+	+	+	AP1	
human β fibrinogen	+						
rat Spi 1.2			+		+		
human factor VII				+			
human factor VIII	+			+	+	NFκB	
human factor IX	+			+	+		
human factor X				+			
human prothrombin	+						

Table I.2: ***Selected examples of liver specific promoters.***
(From Ciliberto, 1993).

example, members of the nuclear hormone receptors or proteins binding to CCAAT motifs (McKnight,1992; Sladek,1994). In addition, the same factor can also occupy various binding sites. Numerous possibilities of heterodimerization also multiply the possibilities of interaction, activation and repression. Finally, it is also evident that interactions exist between the liver enriched factors and the general transcriptional machinery at the vicinity of the TATA box or the initiating elements. But the mechanisms underlying these interactions and the bridging factors between these sites are not well known at the moment.

Another control mechanism of liver expression is exerted at the level of expression regulation of the liver enriched factors. The dissection of the promoters of these factors has resulted in the discovery of regulatory loops critical for expression. The HNF1, C/EBP α, HNF3 β promoters contain, for example, binding sites for HNF1, C/EBP and HNF3 members respectively (Christy,1988; Pani,1992), and self regulation of expression has been postulated for these factors (Xanthopoulos,1991; Jones,1991; Kritis,1991). Several studies also suggest a hierarchy in the activation cascades by the various transcription factors. The activation of the HNF1 promoter provides an example in this regard; this promoter contains elements binding HNF3 and HNF4 (Kuo, 1992). Deletion of the HNF3 binding site only slightly impairs transcription of the HNF1 promoter whereas deletion of the HNF4 binding site suppresses it almost totally. In addition, dedifferentiated hepatoma cell-lines lacking HNF1 and HNF4 are induced to express HNF1 mRNA after stable transfection with an HNF4 expression plasmid. The mechanisms which regulate HNF4 expression are to date unknown and it will be critical to determine whether or not this activator is a key factor in tissue specificity and in hepatocyte differentiation .

The characterization of *cis*- and *trans*-acting regulatory elements in genes whose expression is liver enriched liver is described in the literature mainly with *in vitro* techniques (footprints, protein-DNA gel retardation assays and reporter assays). The development of novel techniques such as *in vivo* footprinting or actualized techniques for detection of DNA hypersensitive sites or the delineation of locus control regions does not always confirm findings given by *in vitro* assays (Mueller,1989; Mirkovitch,1991; Ciliberto,1993). The above techniques are utilized to identify regulatory regions within native chromatin structures in intact cells and therefore take into account the DNA's structural parameters and its topological changes during transcription. In the case of genes highly expressed in the liver, it is often the case that the *in vivo* protective strength of members of the HNF1, C/EBP and HNF4 families is different from that obtained by *in vitro* footprinting techniques (Mirkovitch,1991; Ciliberto,1993). The actual signification of these differences is to date not known.

In addition, most investigative reports rarely take into account cell-cell interactions in the hepatic tissue as well as extracellular matrix-hepatocytes interactions. Both mechanisms increase transcriptional efficiency in the liver (De Simone,1992). Nevertheless, the dissection of regulatory elements and the identification of their cognate factors are prerequisites to approach liver-enriched expression. As well, insights into regulation of gene expression in this tissue have been given by transgenic experiments and somatic cell genetics (Antoine,1990; Tremp,1995; Bulla,1994). The former approach has, for example, shown that several promoters of genes preferentially expressed in the liver target expression at this location and this include the 5' upstream region of the AT gene

(Antoine,1990; Tremp,1995). The latter approach has, for example, shown that the extinction of α_1-antitrypsin expression in cell hybrids is linked to the extinction of HNF1 and HNH4 expression (Cereghini,1990; Bulla,1994).

4. General Features of Regulatory Elements found in Proteinase Inhibitors and Clotting Factors Involved in Coagulation, with a Liver Enriched Expression

This section is intended to review only a few well-known examples of proteinase inhibitors and clotting-factor genes with high levels of expression in the liver and structurally or functionally related to the AT gene. The probability of a common ancestor gene has been proposed for the serpin genes (Hunt,1980; Twining,1994).

4.1 Serpin Genes other than Human Antithrombin, α_2-Macroglobulin

Serpins, and more precisely serpins of the AT group, include high molecular weight inhibitors of the serine proteases involved in coagulation, fibrinolysis, inflammation, complement, and hormone transport (Twining,1994). Such is the case for heparin-cofactor II, the plasmin inhibitors, the plasminogen-activator inhibitors, α1-antitrypsin and antichymotrypsin, C1-inhibitor and angiotensin. The expression of these genes is generally inducible (they are mainly positive or negative acute-phase reactants) and data on their constitutive expression and their tissue restricted expression is often lacking. The acute phase response induces or represses expression (with consequent variations in the plasmatic levels of the secreted proteins) of many serpins and coagulation factors following infection,

inflammation and pathological conditions. Serpin synthesis has been induced (or repressed) either *in vivo* or in cell culture systems by cytokines (interleukins 1 and/or 6, γ-interferon for example), hormones (insulin, dexamethasone and other glucocorticoids), growth factors and thrombin, proteinase-inhibitor complexes, peptides and bacterial by-products (TNF and LPS) or lipoproteins (VLDL) and thrombin. Cytokine and hormone regulated mechanisms (signaling and transcriptional) are reviewed extensively elsewhere in the literature and are not the subject of this thesis (Ciliberto,1993; Darnell,1995); hence, we have not attempted to redetail these mechanisms in this introduction. The links between inducible pathways and constitutive expression mediated by the various liver enriched transcription factors reviewed above will be discussed in the appropriate section of this thesis in chapter IV.

4.1.1 *Cis*-Acting Features for Transcription Initiation

The prototype of the serpin genes, human $\alpha_1$antitrypsin --the main inhibitor of trypsin, neutrophil-elastase, plasmin and eventually thrombin-- is interesting at the level of tissue specificity; effectively two promoters separated by 2000 nt have been identified in this gene (Perlino,1987). The upstream promoter is only active in macrophages whereas the downstream promoter is only active in the liver. Antitrypsin is an acute phase reactant, and studies have shown that the gene uses different initiation sites in different cell types under basal but also under IL6 induced transcription (Hafeez, 1992).

Data for transcription initiation of other serpin genes expressed at high rates in the liver are scarce. Very often, the constitutive expression of these genes shows a

heterogeneous initiation with multiple start sites. TATA boxes are also often lacking, and the $\alpha_2$plasmin inhibitor and the plasminogen-activator inhibitor 1 provide two of the few examples of serpin genes in which regular TATA boxes, GC rich regions and CCAAT boxes have been identified (Hirosawa,1988; Andreasen,1990).

α_2-plasmin inhibitor (α_2-PI) is a key player of fibrinolysis; congenital deficiency in this serpin has as a consequence a severe hemorrhagic tendency. The α_2-PI gene is expressed at high rate in the liver as one major mRNA transcript (Saito,1982). The general and specific cis- and trans-acting elements which regulate transcription of the gene are poorly known at present (Hirosawa,1988).

Plasminogen-activator inhibitor-1 (PAI-1) is a serpin critical for vascular pathology. Elevated levels of this inhibitor have been associated with the development of coronary artery disease and venous thrombosis. This effect is believed to be mediated by a decrease in plasmin generation (Egelberg,1994) secondary to an inhibition of tissue plasminogen-activator (tPA and uPA). Expression of PAI-1 --this protein is one of the more dramatic example of an acute phase reactant-- has been obtained in various tissues in addition to the liver (Thornton,1994; Chomiki,1994; Healy,1994). In hepatic endothelial cells two species of mRNA have been identified, due to alternative cleavage and polyadenylation (Healy,1994). The expression of this gene is regulated by transcriptional and post-transcriptional mechanisms primarily studied in inducible settings (Van Zonneveld,1988; Tuddenham,1994). In the constitutive setting the PAI-1 promoter sites initiating transcription have be found to be heterogeneous although this promoter contains classical features for a TATA box (Tuddenham,1994).

The example of heparin-cofactor II (HC II) illustrates the frequent non classical features of the serpin promoters. HC II, also a chemoattractant, has structural and functional similarities to AT. The 5'upstream region of the HCII gene contains several start sites clustered over 90 nt which suggest an *in-vivo* heterogeneity for transcriptional initiation (Ragg,1988).

Another example is provided by the protein C inhibitor (or plasminogen-activator inhibitor 3, PAI-3); 800 nt of 5' flanking sequences of the gene have been characterized by the presence of two transcriptional start sites in HepG2 cells and the absence of TATA and CAAT boxes (Meijers,1991).

C_1.inhibitor (C1-Inh), a multitarget inhibitor, has a gene structure strongly similar to that of AT. In addition, hereditary deficiency in C1-Inh is responsible for angioneurotic oedema. The promoter of the C1-Inh gene has interesting features (Zahcoi,1994) with the absence of a TATA box compensated for by the presence of a strong transcriptional start site, as well as an initiator-element, a CAAT box and a GC rich region. Under basal conditions, the sole region -3/+210 nt is necessary to direct transcription.

α_2.macroglobulin (α_2.MG) is not a serpin but a wide spectrum inhibitor able to "trap" proteases and to carry growth-factors and cytokines. Mainly synthesized by the liver, this inhibitor is also present in many other tissues. In hepatic cells, the basal expression of the human gene depends upon a -144/-104 nt region homologous to the 5'upstream region of the AT gene and including a TATA-less promoter with no apparent GC rich region (Matthijs,1994).

4.1.2 Responsive Elements and *Trans*-Acting Factors

(a) In relation to constitutive expression in the liver -- The α_1-antitrypsin gene promoter-proximal enhancer region, as presented in Table II.2, is composed of a cluster of binding sites for liver enriched transcription factors in a region encompassing the transcriptional start site; included are binding sites for HNF4 (LF-A1), for HNF1 (α and B), for C/EBP (3 sites) and for HNF3 (2 sites). Deletion of the two sites HNF4 and HNF1, *in vitro* in reporter assays and *in vivo* in transgenic mice, has critical consequences for the efficiency of α_1-antitrypsin transcription (Cereghini,1990). In addition, the region binding HNF4 and HNF1 in the liver (or the immediate upstream promoter) is the target for negative regulation in HeLa cells.

For PAI-1, eight regulatory sequences of the rat gene promoter have been identified in H4IIC cells under basal conditions of expression; within these sequences, binding sites for PEA-3, Sp1- and CTF/NF1 participate in transcriptional efficiency (Johnson,1992). Hence, the precise *trans*-acting elements which underly expression of the PAI-1 gene in the liver have not been identified to date.

Basal expression of human α_2-macroglobulin in hepatic cells depends upon a -144/-104 nt region with putative but not yet identified binding sites for nuclear hormone receptors (Matthijs,1994). Putative sites for the same receptors have been also postulated for the α_1-plasmin inhibitor (Hirosawa,1988).

(b) In Relation to Inducible Expression -- The basis for induced transcription of human antitrypsins (α_1-antitrypsin and $\alpha_1$antichymotrypsin) during the acute phase

response in HepG2 cells is the activation of IL6 responsive elements also located at the vicinity of the promoter (Bartalena, 1992; Hafeez,1992). Whether or not human α_2-macroglobulin is a acute phase reactant like its rat counterpart is still debated. The basis for this property in the rat gene is the presence of an IL6 responsive element CTGGGA (Hattori,1990). Such elements have also been identified in the C1 inhibitor gene (Zahedi,1993). HC II is believed to behave like an acute phase reactant,but the molecular basis for this phenomenon is unknown (Toulon1991). The acute phase responsive elements identified for the PAI-1 promoter in HepG2 cells are type I responsive elements (APRE) IL1+IL6 dependent (Healy,1994).

Other ubiquitous factors such as SP1-like proteins (Egr 1) and proteins of the AP-1 and AP-2 complexes bind to the promoter regions of α_1-AT, PAI-1 and PAI-3 upon induction by TGFβ, phorbol esters and cAMP (Grayson,1988; Ciliberto,1993; Descheemaeker,1992; Sandler,1994). Glucocorticoid responsive elements have been identified in the PAI-1 promoter and the α_2-macroglobulin 5'upstream region (Hocke, 1992; Anfosso,1993). Responsive elements for γ-interferon (IFN) have been mapped at -300 nt of the C1-inh gene 5'upstream region (γINF activated regions TTc/ACNNNAA binding the factors ISGF-2 and -3) and in the first intron (IFN stimulated responsive elements with the consensus GAAANN located in two half sites separated by approximately 20 nt). γIFN responsive elements are also present in the rat α_2-macroglobulin promoter (Yan,1994). The consensus for these responsive elements resembles a region of the AT promoter.

4.2 Expression of Liver Enriched Clotting Factors

4.2.1 *Cis*-Acting Features for Transcription Initiation

The regulatory mechanisms underlying liver expression of most coagulation factors are not well known. In some cases, the hepatic cell lines available do not express or express a given factor at very low rate; for example, factors V, VIII, or XI (Tuddenham,1994). This property suggests a role for inducible regulation in the expression of such factors. In other cases multiple species of mRNA are found; for example, for the α and γ fibrinogen and the protein C genes (Fu,1992; Chodosh,1988). Moreover, as with the serpin genes, the absence of TATA boxes or strong initiating elements (factors VII, X) and/or the presence of multiple or spurious transcriptional start-sites (factors XII, II, X, IX, plasminogen) often complicate the experimental approaches to the study of gene expression regulation (Cool, 1987; Bancroft, 1990; Malgaretti, 1990; Salier, 1990; Huang, 1992). Less often, regulatory features other than proximal regions have been mapped; for example, the factor IX gene (Salier,1990) contains a first non-proximal element, a promoter located 500 nt upstream of the proximal promoter and oriented in a reverse direction (Salier,1990). The latter gene contains also a repressor, ATCCTCTCC, located 1700 nt downstream of the first exon.

4.2.2 Responsive Elements and *Trans*-Acting Factors

(a) in Relation to Constitutive Expression in the Liver -- The basal regulatory elements involved in liver-enriched transcription, that are known for a number of

coagulation-factor genes, are presented in Table I.2. Prothrombin contains an upstream proximal enhancer binding HNF1. This enhancer is flanked by inverted CCTCCC repeats and a GC box, suggesting interaction with SP1-like proteins (Bancroft,1991; Chow,1991; Friezner,1992). HNF1 binding sites have been also identified in fibrinogen α and β, factors II, VIII, and IX promoter-proximal enhancer regions (Courtois,1987; Tuddenham,1994). Interestingly, the fibrinogen promoter contains SP1 and CBF binding sites --GC and CCAAT boxes binding factors-- in the vicinity of the HNF1 binding region (Courtois,1987).

Binding sites for HNF4 have been identified in clotting factors VII, VIII, IX, X, and Protein C promoters (Miao,1992; Huang,1992). The factor IX promoter, for example, contains a consensus for HNF4 binding,TGGACC, in two regions; one of these regions, at -21 nt, is highly homologous to a -92/-68 nt element of the AT gene 5'upstream region (Salier,1990).

Binding sites for C/EBP members are present in the factor VIII and IX and the α fibrinogen human genes (Hu,1995; Figueiredo,1995; Salier,1990). A recent report on constitutive factor VIII liver-enriched expression also mentions binding sites in the proximal promoter for NFκB (Figueiredo,1995). Fibrinogen and factor X also have CAAT binding sites that are nevertheless not occupied by members of the C/EBP family (Huang,1992).

(b) in Relation to Inducible Expression -- The known mechanisms of up- or down-regulation of expression in the coagulation factor genes suggest again similarities with

responsive elements supporting induction of serpin gene expression. If most of these genes are "house keeping genes" and code for proteins critical for coagulation (their inherited deficiency has severe pathological consequences), their expression appears also to respond to multiple external stimuli. One of the better known examples of genes from this family is provided by the various fibrinogen promoter regions which contain responsive elements for IL6 (+/-IL1) and glucocorticoids. Of interest, the regulatory regions of the various genes of the fibrinogen family are different even if their expression is coupled. Furthermore, species differences have also been well documented (Courtois,1987; Dalmon,1993; Roy,1994; Tuddendham,1994; Hu,1995; Mizuguchi,1995). The activity of these inducible elements is modulated by the presence of constitutive liver transactivators (HNF1 and C/EBP isoforms for example) and by neighbouring enhancer and repressor elements (Dalmon,1993; Hu,1995; Mizugushi,1995). Regulation by hormonal mechanisms also intervenes in some coagulation genes. For example estrogen responsive elements have been identified in the factor XII gene 5' upstream region (Bartalena,1992).

5. Pathology and Gene-Expression Regulation of the Clotting Factors and the Inhibitors of Coagulation

5.1 Antithrombin

A literature review of the frequent and multifactorial acquired AT deficiencies is beyond the scope of this study, as the emphasis of this work is on the constitutive

expression of the AT gene. Hence, it is likely that agents modifying plasma AT levels may directly influence transcriptional control mechanisms; for example, the effects of steroid hormones on AT expression could be in part transcriptional.

In contrast, the inherited cases of deficiencies in AT have a prevalence of only 1/500 to 1/5000 in the population (Egeberg,1965; reviewed by Blajchman,1994). These defects are almost exclusively associated with mutations in a single AT allele. The absence of both alleles has been described only rarely and is likely not compatible with life. Most kindreds are affected by venous thrombosis and its complications. However, the detection of an abnormal AT mutant is not necessarily associated with a thrombotic diathesis, especially in the cases of mutations affecting the heparin binding domain of the molecule (Sheffield,1995). The clinical manifestations often follow trauma, delivery, the use of oral contraceptives and injury. As well, the risks increase with age. These inherited deficiencies are classically separated in three types. In type 1, the gene product of one allele is not circulating in plasma. In this case, deletions or various types of mutations interrupt transcription, translation, post-translational events, or secretion of the affected allele (Blajchman,1994). Most modifications reported to date include small rearrangements in one gene, and only a small number of partial or extensive deletions are known (Olds,1994; Lane,1994). Table I.3 describes the known type I defects (Lane,1994). To date, there have been no reports of mutations in the regulatory sequences of the gene with direct pathological consequences. 2/Type 2 and 3 deficiencies are associated with the secretion of an allele with a single point mutation either in the AT thrombin binding domain coded by the cDNA and located toward the carboxy terminus (type 2 mutants), or in the heparin

SMALL REARRANGEMENTS	
SINGLE BASE MUTATIONS, SMALL INSERTIONS-DELETIONS. (CODON CHANGE 46-427) 59 MUTANTS (39 UNIQUE MOLECULAR EVENTS)	LANE 1993, 1994
PARTIAL AND WHOLE DELETIONS	
DELETES 2761BP, INCLUDING EXON 5, FROM 9060-9104 TO 11118	OLDS 1992
DELETES 5' END , EXON 1,2.	FERNANDEZ 1992
REMOVES 3' END OF GENE, INCLUDING EXON 6	PERRY ET AL,UNPUBLISHED
UNCHARACTERIZED	OLDS 1992
WHOLE GENE DELETION	WINTER 1983
WHOLE GENE DELETION	PROCHOWNIK 1983
WHOLE GENE DELETION	BOCK 1987

Table I.3: ***Classification of type 1 antithrombin hereditary deficiency.***
(From D. Lane, 1993 and 1994).

binding domain of the cDNA composed of several regions in the 4 first exons (type 3 mutants). In addition, a number of deficiencies involving point mutations in the carboxyl-terminus part of the molecule have pleiotropic effects (Lane,1993; Blajchman,1994). Combined deficiencies of AT with other inhibitors have also been described (with protein C and/or heparin cofactor II)(Jobin,1991). The molecular basis for these defects remains to be discovered.

5.2 Other Coagulation Factors and Serpins

5.2.1 Factor IX

A few examples of defects directly linked to an abnormal regulation of gene transcription have been reported. The best documented example is the pathology of the factor IX promoter with the various types of hemophilia-B Leyden (Robert,1993). These mutations, all clustered around the start site, from -26 to +13 nt according the start site determination by Brownlee et Pang (1990), have allowed for a better understanding of transcription regulation of the factor IX gene in disrupting binding of activators. Hence, co-transfection experiments have shown that the mutant promoter is unable to be co-activated (Crossley,1990). More precisely, downstream mutations at +8 and +13 nt disrupt C/EBP α binding (consensus CTTT/GCACAAT). Very recently, DBP has also been shown to bind to this region (Reijnen,1994). The factor IX activity associated with such mutations increases in patient's plasma often after puberty (Hiroshawa,1990). The factors binding to the -1 to -15 nt region have not been fully identified yet; they include

liver enriched and ubiquitous factors. The effects of a mutation at -5 nt have been recently shown to be compensated *in vitro* by C/EBP and DBP (Picketts,1993); moreover, compensation of the Leyden phenotype is thought to be correlated with induction of DBP expression around puberty. Mutations at -21, -20 (T>A) disrupt binding for HNF4 and their effect is corrected by puberty. In contrast, a mutation at -26(G>C) has been shown to affect overlapping binding of the androgen receptor (-36 AGNaCANNNTGTNCT-22) and of HNF4 (-27 TGgACTTTGGccC-15) (The bases that are unmatched are written in lower case letters). As a consequence, the effects of such a mutation are permanent (Crossley,1992).

5.2.2.Protein C

The second example of abnormal transcription associated with impaired expression is provided by the analysis of kindreds with protein C deficiency. First, it has been shown that genetic variation in the protein C promoter was associated with variation in plasma protein C levels and thrombotic risk (Spek,1995). The basis of this observation could be a polymorphism at +25 nt (in the untranslated region), but whether or not this polymorphism accounts for differences in transcription regulation is to date unknown. Second, mutations in the putative promoter region of protein C have been described; their location is at -33, -28, -11, and -14 nucleotides. Mutation at -14 nt disrupts binding with HNF1 and reduces transcription mediated by the protein C promoter in reporter assays (Tsay,1991; Berg,1994). Other mutations disrupt or reduce HNF3 binding and have an effect on the transactivation potential of this factor (Berg,1994; Spek,1995-2).

5.2.3 Fibrinogen

As with protein C, an association has been hypothesized between a common polymorphism in the 5'flanking region of the β fibrinogen gene and the fibrinogen plasma levels. In the latter case, this association is actually strongly debated (Humphries,1995; Connors,1992).

5.2.4 Plasminogen-Activator Inhibitor 1

This last example also provides evidence that a single base pair insertion/deletion polymorphism in the promoter region of a gene involved in coagulation modulates transcription regulation. More precisely, a polymorphism of the PAI-1 promoter affects the response of this promoter to IL1 in Hep G2 cells as well as the levels of induction (Dawson,1993). Of interest, comparable observations can be made for apolipoprotein A-1 (Apo A1), a gene which contains a promoter highly homologous to that of plasminogen and the plasminogen-activator inhibitor 1. Effectively, an Apo-A1 polymorphism at -78 nt has been associated with increased basal transcription efficiency in reporter assays and to higher plasma levels of Apo A1 and HDL-C. Of interest, this polymorphism has also been associated with impaired binding of a 90-kDa transcription factor interacting with a GCC(A/G)GGG motif. The factor binding to the GC rich element is believed to be a factor similar to a repressor of the epidermal growth factor receptor promoter and of other house-keeping genes (Dawson,1993).

6. This Project

The review of the literature detailed above illustrates the contrast that we observed between the abundance of data and the new insights into the general or the auxiliary factors (liver-enriched and ubiquitous) involved in transcription, and the partial knowledge for the regulatory mechanisms involved in the expression of the serpins' and coagulation-factors' genes. Obvious similarities between these two gene families, both highly expressed in the liver, suggest common pathways for regulation of expression. However, with the exception of a few well-characterized regulatory regions, little is known about the mechanisms underlying constitutive expression of these genes or those linking constitutive transcription with up- or down-regulated expression. These observations might be linked in part to the frequent reports of non-classical features for the promoters of this gene family, which include the absence of TATA boxes or strong initiation as well as ambiguities in the determination of transcription initiation sites. In some instances, sophisticated mechanisms of expression are also present, complicating the advances in this field. Some examples include, for example: multigene-families (the fibrinogen genes with gene-to-gene differences in the nature of the responsive elements and the mechanisms allowing transcription), alternate start-sites of translation, differential splicing or alternative promoter usage. Moreover, transcriptional responses for these genes are mainly studied in the induced setting (with induction of expression by cytokines (IL1, IL6) and other mediators of the acute phase response in most cases).

AT is a serpin that has been and is still extensively studied. Nevertheless, the investigative efforts have been targeted towards the characterization of the protein and the

molecular basis of the relation structure-activity. Until very recently, only little information on the regulation of AT gene transcription could be found. The desinterest on this facet of AT expression may have been due to the belief that the AT gene is a "house keeping gene" with only general features for transcription.

Nevertheless, many observations justified the study of regulatory features for the constitutive expression of the AT gene. Following the characterization of a kindred with congenital AT deficiency, it was indirectly observed that a natural deletion including in part the 5' region of the gene had as a consequence the absence of AT expression. More important were the suggestion of non-classical features in the 5' AT upstream region of the gene, the incertitude for the transcriptional start site location, the possibilities of induction for expression of the gene, the suggestion of the involvment of developmental factors in AT expression and the enriched liver expression of this protein. All these observations suggested a complex behavior of the mechanisms regulating AT expression. This project was therefore designed to better understand the mechanisms underlying the expression of the gene. The mapping of *cis*-elements in this region of the gene had not been studied by deletional analysis and was related only in one literature report. Systematic footprint-analysis of the promoter region had never been performed either. In addition, the factors binding to the element similar to the TF-LF1 consensus of the transferrin gene had not been identified (Ochoa,1989). We therefore decided to approach the study of the *cis*-acting elements and *trans*-acting factors determinant for constitutive expression of the AT gene.

CHAPTER II

MATERIALS AND EXPERIMENTAL PROTOCOLS

1. MATERIALS

1.1 Chemicals

Chemicals were of the highest quality available. The following list details the reagents mostly used and the commercial sources from which they were obtained.

Acrylamide	Bio-Rad Laboratories
Agarose (DNA grade)	Gibco/BRL
Ampicillin	Sigma Chemical Company
Anion exchange columns	Quiagen Laboratories
Bacto-agar	Difco Laboratories
Bio-Rad protein assay	Bio-Rad Laboratories
BIS	Bio-Rad Laboratories
Bovine serum albumin (DNA grade)	New England Biolabs
CsCl	Gibco/BRL
Deoxyribonucleotides	Pharmacia Inc.
DNA, denaturated:	
Salmon sperm	Sigma Chemical Company
Calf thymus	Sigma Chemical Company
Dulbecco modified Eagle Medium	McMaster University
DTT	Gibco/BRL
IPTG	Gibco/BRL
Lipofectin	Gibco/BRL
Luciferase detection kit	Promega Corp
Luciferin	Sigma Chemical company
Long-Ranger	IBI Chemicals

Molecular weight standards :	
1 kbp DNA ladder	Gibco/BRL
λ Hind III DNA ladder	Gibco/BRL
LMW Protein marker	Bio-Rad Laboratories
mRNA (human liver)	Clontech Laboratories
Nitrocellulose membranes	Schleicher and Schuell Inc
NA45 (DEAE membrane)	Schleicher and Schuell Inc
Oligonucleotides	Institute for Molecular Biology and Biotechnology, McMaster University
	VetroGen Corp
ONPG	Promega Corp
PBS	McMaster University
PEG 6000	Sigma Chemical Company
Penicillin G	Gibco/BRL
PMSF	Gibco/BRL
Poly(dI-dC)-poly(dI-dC)	Pharmacia Inc.
Poly (dI), Poly (dC)	"
Protein A Sepharose-4B	Sigma Chemical Company
RNAse inhibitor (RNAsin)	Promega Corp
Sephadex G25, G50	Pharmacia Inc
Spermidine	Sigma Chemical Company
Spermine	"
Streptavidin magnetic beads (M 280)	Dynal SA
Streptomycin	Gibco/Brl
TEMED	Gibco/BRL
Triton X-100	Sigma Chemical Company
tRNA (calf liver)	Boehringer Mannheim
Trypsin-EDTA	Gibco/BRL
X-gal	Gibco/BRL
Zetaprobe transfer membranes	BioRad Laboratories

1.2 Radiochemicals

[α-^{32}P]dATP (3000 Ci/mmol, 10 μCi/μl)	Dupont/NEN Canada Inc.
[γ-^{32}P]ATP (3000 Ci/mmol,10 μCi/μl)	Amersham; ICN
L-[^{35}S]-methionine (2800 Ci/mmol,10 μCi/μl)	Amersham; ICN

1.3 Restriction Enzymes, Other Enzymes

These enzymes were used according to manufacturer's instructions unless specified.

DNA ligase (T4)	New England Biolabs
DNA polymerase I Klenow fragment	New England Biolabs
DNA polymerase I (Nick translation)	Gibco/BRL
DNA polymerase, modified T7	United States Biochemicals (Sequenase)
DNA polymerase (*Thermus aquaticus*)	Perkin Elmer Cetus
DNAse I (bovine pancreas)	Pharmacia Inc
	United States Biochemicals
Polynucleotide kinase (T4)	Pharmacia Inc.
Proteinase K	Boehringer Mannheim
Restriction endonucleases	Gibco/BRL
	New England Biolabs
	Boehringer Mannheim
	Promega Corp
Reverse transcriptase	Gibco/BRL
RNAse A	Quiagen Corp
Terminal deoxynucleotidyl transferase (TdT) (calf thymus)	Gibco/BRL Pharmacia Inc

1.4 Reagents Kits

Cell-free transcription-translation	Promega Corp
Dideoxysequencing kit (Sequenase)	US Biochemicals Corp
Maxam and Gilbert sequencing kit	Boehringer Mannheim
Nick translation kit	Gibco/BRL
RACE kit and anchor primers	Gibco/BRL
Random primer labelling system	Gibco/BRL

1.5 cDNA's and Genomic Probes

Fig.II.1 describes the various probes utilized for the restriction analysis of the AT alleles in the kindred with type I AT deficiency as well as for mapping the normal AT gene, more particularly its 5' region.

1.5.1 cDNA Probes

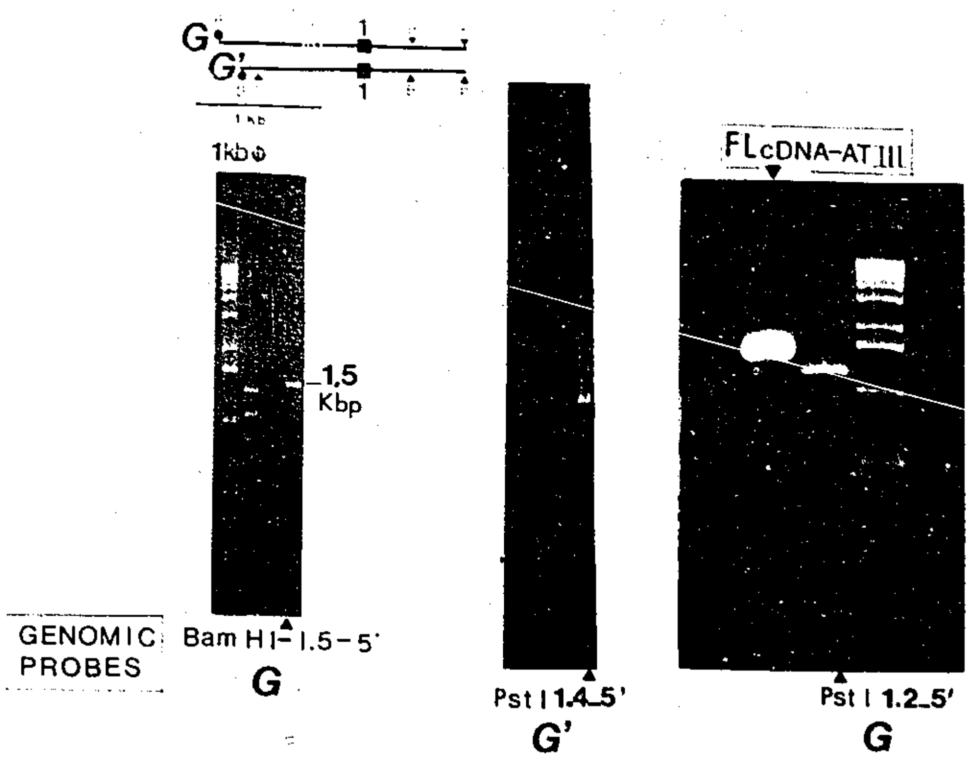
The initial probe used was a full-length 1500 nt human AT cDNA subcloned in the PstI site of the plasmid pTK218, a generous gift of Dr.E.V.Prochownik, University of Michigan, Ann Arbor, MI (Prochownik,1984); this probe, FL cDNA-ATIII in Fig.II.1, coded for 8 nucleotides of the 5' untranslated region, for the presequence, the 432 aminoacids of AT and 87 nucleotides of the 3' untranslated region.

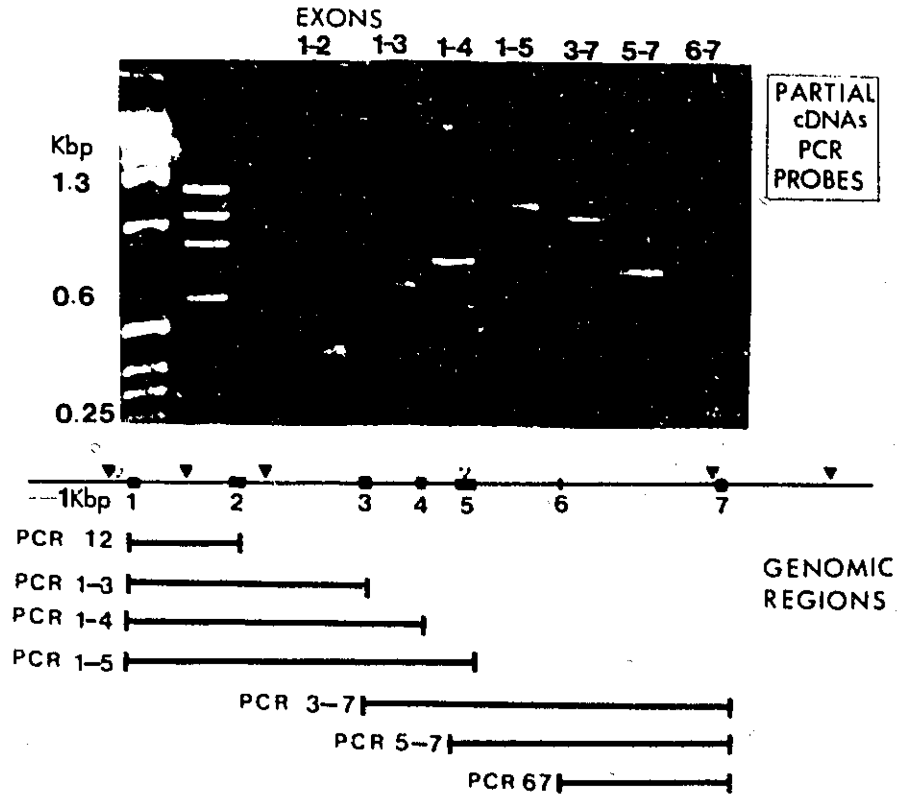
Truncated probes generated from this cDNA were also synthesized by the polymerase chain reaction (Partial cDNA PCR probes in Fig.II.1). The primers used were 20 mers specific for the beginning or the end of each exon, as detailed in the legends to Fig.II.1.

1.5.2 Genomic Probes

The full-length cDNA provided was unable to encompass the 5'upstream region of the gene; therefore, the following probes were utilized to study this region specifically: two BamHI and PstI restriction fragments encompassing exon 1 excised from normal genomic EMBL3 clones (see map in Fig.II.1), and a 1200 nt PstI fragment provided by Dr.S.Bock,Temple University, Philadelphia, PA.

Figure II.1: ***Genomic inserts, full length and partial cDNA antithrombin probes.*** The region encompassed by each genomic probe is indicated in the restriction map on the top left corner of the figure. B, BamHI; P, PstI. FL cDNA-AT III, full length cDNA antithrombin insert. The partial cDNA probes were generated by amplification of the full length cDNA. The antithrombin gene regions encompassed by each probe are indicated in the bottom of the figure. The conditions for PCR amplification of the various truncated cDNA, and the primer sequences were as described in Fernandez-Rachubinski, 1992.


G
G'
1
1
1kb
GENOMIC PROBES
Bam HI-1.5-5'
G
1,5 Kbp
Pst I 1.4_5'
G'
FL cDNA-AT III
Pst I 1.2_5'
G


EXONS
1-2
1-3
1-4
1-5
3-7
5-7
6-7
PARTIAL cDNAs PCR PROBES
Kbp
1.3
0.6
0.25
1Kbp
1
2
3
4
5
6
7
PCR 12
PCR 1-3
PCR 1-4
PCR 1-5
PCR 3-7
PCR 5-7
PCR 67
GENOMIC REGIONS

1.5.3. Probe Labelling

All inserts were eluted from polyacrylamide gel and ethanol precipitated according to standard protocols (Sambrook,1989). Radiolabelling was carried out by primer extension with p(dN)6 random primers (Gibco/BRL) and the Klenow fragment of *E. Coli* polymerase I, or via nick-translation with PolymeraseI and DNAseI. The radioisotopes used were [α ^{32}P]-dATP or [α ^{32}P]-dCTP. Unincorporated dinucleotides were removed by a chromatographic step on Sephadex G50. Specific activities of the probes used for hybridization ranged from 0.5 to 2.0×10^7 cpm/ng.

1.6 Cloning Vectors and Bacterial Strains

1.6.1 General Vectors

The EMBL3 phage DNA, the packaging extracts (Gigapack), the hosts LE-392, NM-538 and NM-538 were obtained from Promega Inc. Restriction fragments of genomic clones encompassing the 5'region of the gene were subcloned into pUC18 or 19 (Pharmacia Corp.), pGEM (Promega Inc.) or pBluescript SK+ (see details in the transfection assays section). In most cases, *E.Coli* DH5α competent cells (Gibco/BRL) were used for transformation. Growth medium was Luria Bertani [1% bacto-tryptone, 0.5% bacto yeast extract, 15 mM NaCl, pH 7.4] with 100 μg/ml ampicillin. Minipreparations of plasmid DNA were done using the alkaline lysis method followed by phenol-chloroform extraction and ethanol precipitation (Birnhoim,1983; Sambrook,1989).

1.6.2 Reporter Plasmids

(a) Basic Plasmids -- pSVOA-LΔ5' was a pBR322 and SV40 derived reporter plasmid containing an origin for replication, a β-lactamase gene to allow for selection in *E.Coli* (by ampicillin resistance) and a luciferase expression system derived from the firefly *Photinus Pyralis* gene. This plasmid, a modification of a plasmid built by de Wet (1987), was a generous gift of Dr.R.Rachubinski, Department of Anatomy and Cell Biology, University of Alberta, Edmonton, and Dr.S.Subramani, Department of Biology, University of California at San Diego. The schematic map of this vector is given in Fig.II.2. Some characteristics of this plasmid include the deletion of the 5'initiation codon of the luciferase gene (de Wet,1987; Gould,1988) and the insertion of SV40 polyadenylation signals upstream and downstream of the luciferase cDNA. The modifications above provide a further increase in luciferase reporter activity and a reduction in luciferase background expression in CV-1 cells and their derivatives (Cos) transfected with promoter-less plasmids. The polyadenylation signal also decreases background from cryptic promoters within the pBR322 core. A similar vector but with higher background activity in HepG2 and Cos1 cells was pGL-Basic, derived from pSV232ALΔ5'(Promega Inc.); the only advantage of the latter vector was the presence of multiple cloning sites upstream and downstream of the luciferase cDNA.

(b) Reporter Control Plasmids -- pCPS-Luc was a luciferase expression vector driven by 600 nt of the proximal promoter of the gene encoding rat liver carbamoyl phosphate synthetase, CPS (Howell,1989). This plasmid was engineered from the

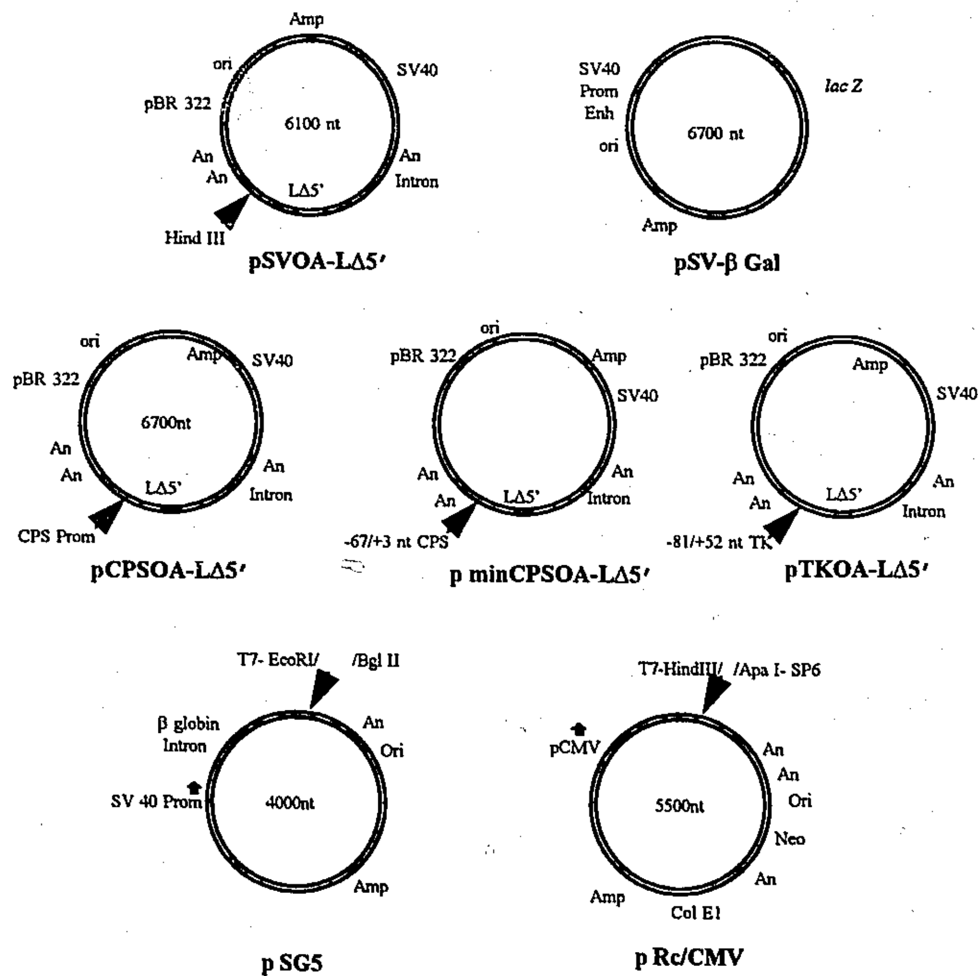


Figure II.2: ***Reporter plasmids and expression vectors.*** An, polyadenylation signal; Amp, ampicillin; ori, origin;Prom, promoter;Enh, enhancer; Neo, neomycin; pCMV, cytomegalovirus promoter; -67/+63 nt CPS, minimal promoter of the carbamoyl phosphate synthetase gene;-81/+52 nt TK, minimal promoter of the thymidine kinase gene.

promoter-less reporter basic plasmid pSVOA-LΔ5'. It was a generous gift from Dr.G. Shore, McGill University, Montreal, QB. The CPS promoter has been shown to bind to liver-enriched factors such as C/EBP (Lagace,1992). A schematic map of CPS-Luc is given in Fig.II.2. p minCPS -67/+3 nt-Luc was a truncated version of pCPS-Luc with only the CPS minimal promoter; this plasmid was obtained from Dr.G.Shore. A second plasmid with a minimal promoter, pTK -81/+52 nt-Luc, was also used; it was a generous gift from Dr.R.Rachubinski. In this plasmid luciferase expression was driven by a viral thymidine kinase -81/+52 nt promoter.

pSV2-LΔ5' was obtained from Dr.S.Subramani. This plasmid derived from the parent basic plasmid pSVOA-LΔ5' contained, in addition to the SV40 regulatory sequences derived from the early promoter region, the enhancer region as well as intervening sequences and polyadenylation sites. This vector also contained the SV40 origin for episomal replication (Subramani,1983). pGL-Control (Promega) was the commercial equivalent of the previous plasmid.

pGL-Prom (Promega) was a luciferase expression plasmid driven only by the SV40 derived minimal promoter. pGL-Enh (Promega) was a luciferase expression plasmid containing only a SV40 derived enhancer.

pSV40-βGal (Promega) was a β-galactosidase expression plasmid under the control of the SV40 promoter-enhancer complex. This vector was used as an internal control for transfection efficiency.

1.6.3 *In Vitro* and *In Vivo* Expression Vectors for Transcription Factors

(a) Human COUP-TF1 2200 nt cDNA in pGEM7Zf(+) was a generous gift of Dr.M.J.Tsai, Baylor College of Medecine, Houston, Texas (Cooney,1992). The mammalian expression vector pRc/CMV containing a HindIII/XbaI of human COUP-TF1 was a generous gift of K.Miyata and Dr.J.Capone, Department of Biochemistry, McMaster University, Hamilton, Ontario. pGEM-5Zf(+) containing a truncated version of human COUP (amino terminus minus 51 aminoacids) was also a generous gift of Dr.J. Capone (Miyata,1993).

(b) Rat HNF4 cDNA into pSG5 (Green,1988) was a generous gift of Dr.F.M. Sladek, University of California, Riverside, California, and C. Winrow in Dr R. Rachubinski's laboratory (Winrow 1994).

(c) Rat PPARα cDNA in pBluescript II SK(+) was obtained from Dr.D.Noonan, Ligand Pharmaceuticals, San Diego, CA. The PPAR cDNA subcloned into the mammalian expression vector pRc/CMV (In Vitrogen) was a generous gift of S.Marcus in Dr.R.Rachubinski's Laboratory (Marcus,1993).

(d) Human RXRα cDNA subcloned into pSG5 was a generous gift of R.M. Evans Salk Institute, San Diego, CA, and S. Marcus.

(e) Rat TRα cDNA subcloned into pRc/CMV was a generous gift of Dr. V. Nikodem, Clinical Endocriniology Branch, National Institute of Health, Bethesda, Maryland, and A.Kassam in Dr.R.Rachubinski's laboratory.

(f) Mouse C/EBPα full length cDNA and 420 nt upstream sequences subcloned in pGEM3 and in pSG5 were designed by K.G.Xanthopoulos and were a

generous gift of Dr.P.Hoodless and Dr.J.Darnell, Rockefeller University, New York, NY.

(g)Rat HNF3α cDNA in pGEM2 and HNF3β in pBluescript(SK)+ were generous gifts of Dr.P.Hoodless. The same cDNA's in the mammalian expression vector pCMV were also provided by Dr.Hoodless and were designed by Dr.W.Chen, Rockefeller University, New York, N.Y.

1.7 Antiseras and IgG Fractions

Rabbit IgG anti human COUP-TFI was a generous gift of Dr.M.J.Tsai, Baylor College of Medecine, Houston, Texas. Two rat antisera against HNF4 were tested; an antisera directed against a peptide in the carboxyl-terminus of the α isoform and obtained from Dr. F.M.Sladek; a total antisera obtained from Dr.J.E.Darnell, also specific of HNF4α. Rabbit anti human HNF3 α, β, and γ antiseras were provided by Dr.P.Hoodless and Dr.J.E.Darnell. Rabbit anti-human C/EBPα antiserum was a generous gift of Dr.G.Shore, McGill University, Montreal, Quebec. Rabbit anti-human PYBP antiserum was a generous gift of Dr.F.Brunel, Rockfeller University, New York, NY. Rabbit anti human RXRα and anti mouse PPARα were a generous gift of S.Marcus, Dr.J.Capone and Dr.R.Rachubinski. The other antibodies were obtained from Santa Cruz Biotechnology, Santa Cruz, CA. IgG fractions were purified from the COUP-TF1, RXR and PPAR antiseras by protein A Sepharose-4B affinity chromatography according to manufacturer's instructions (Pharmacia). These fractions were pooled, neutralised, dialysed in PBS buffer and concentrated to 5 mg/ml.

2.EXPERIMENTAL PROTOCOLS

2.1 Restriction Mapping of the Deleted Allele in the Kindred with Type I Antithrombin Deficiency. PCR Confirmation of Deletion in the 5'Upstream Region of the Gene.

The detailed experimental protocols for Southern analysis of genomic DNA samples isolated from various family members have been reported (Fernandez-Rachubinski,1992). Briefly, 10 μg samples of DNA (Poncz,1982) were restricted with various enzymes, subjected to 0.8% agarose gel electrophoresis and transferred to Zetaprobe modified nylon membranes according to manufacturer's instructions (BioRad Laboratories). A variety of probes encompassing selected regions of the gene were then used for hybridization (see Fig.II.1 for probe description and for the regions encompassed by each probe tested). The absence of the 5'upstream region in the deleted allele was further confirmed by direct amplification of a BamHI/BamHI 1400 or 1500 nt fragment flanking exon 1. The primers were 5'-TCGGATCCAGGGTCTGAATCAAG-3' and 5'-GGATCCCGTGAGTGCTGACTT-3'. Each reaction used 1 μg of DNA, 10 pmol of primer, in buffer containing 67 mM Tris-HCl (pH 8.8), 6.7 μM EDTA, 10 mM dithiothreitol, 17 mM $(NH_4)SO_4$, 5 mM $MgCl_2$ and 1mM each dNTP(s). The conditions for amplification were a precycle at 95oC-5 min, 72°C-2 min (to add Taq polymerase and overlay each sample with mineral oil), and 30 cycles at 92°C-2 min, 56°C-2 min, 72°C -2 min+ 5" auto-extension. The resulting products were analysed on 2% agarose gel electrophoresis.

2.2 Cloning of the Breakpoint of the Deleted Allele

2.2.1 Preliminary Step

For this purpose, 300 μg of genomic DNA from the mother's proposita were digested with PstI and sized on sucrose density gradient. The resulting fractions of 9-11x10^3 nt (containing the abnormal allele) and 4-6x10^3 nt (containing the normal allele) were selected. Two protocols were then followed to clone the breakpoint (Fernandez-Rachubinski,1992).

2.2.2 Subcloning and Branch-Capture PCR

Briefly, each fraction was subcloned into pGEM3ZF+, then, 1/10 of the ligated material was amplified with one primer (primer 1) biotinylated at its 5'end and complementary to the 3' boundary of exon 3. Isolation of the biotinylated amplified product was accomplished via streptavidin coated magnetic beads as described by Rosenthal (1990). This step was followed by a second amplification reaction, this time with two primers, a first specific of the SP6 region of the vector (primer 2) and a second complementary to exon 3 (primer 3, nested to the first exon 3 specific primer). The products were analysed on agarose gel electrophoresis and their specificity was confirmed with truncated cDNA probes.

2.2.3 Inverse PCR

To confirm the previous reaction, PstI 9-11x10^3 nt fractions were self-ligated. This

was then followed by a two step amplification reaction; the first step used the 5'biotinylated primer (primer 1) complementary to exon 3 and the self-ligated material. In a second step, the products bound to streptavidin were amplified with a primer located at the end of IVS6 (primer 4) and with a nested primer specific of exon 3 (primer 3).

2.2.4 Additional Steps

The PCR products obtained with either technique were sequenced either directly or after subcloning into pGEM 3ZF+ and pGEM 7ZF+ using the dideoxy-chain termination method. To confirm the presence of the deleted allele in the various members of the family available, PCR amplification of the breakpoint was directly tested in genomic DNA samples. The buffer system and general conditions were as described above. The primers were 5'-CAGGAGTTAGAGGCCAGCCT-3' (located 124 nt upstream to the breakpoint in the abnormal allele) and 5'-ATGGTGTCAAACTTAAATA[illegible]3' (complementary to exon 3). The conditions for amplification were 30 cycles at 92°C -2 min, 53°C -2 min, and 72°C -2 min + 1" auto extension.

2.3 Cloning of Genomic Inserts for the 5' Region of the Antithrombin Gene

Genomic DNA was extracted from anticoagulated whole blood of a normal individual after red cell lysis (Poncz,1982). 100ug of DNA was restricted with 100 units of Sau3A for O/N at 37°C. DNA fragments of 20 kbp were selected after DNA sizing on sucrose gradient and then ligated to EMBL3 phage arms BamHI restricted (500ng/reaction), with T4 DNA ligase according to manufacturer's instructions (Promega,

Madison, WI). Following ligation, the recombinant phages were packaged in Gigapack-II extracts and plated on the host strain LE-392. The library was screened with the cDNA's and 5'upstream inserts described above in subheading 1.5 (see also Fig.II.1). Positive clones were isolated by polyethylene-glycol 6000 phase separation in high salt, followed by CsCl gradient centrifugation as described (Ausubel,1989). Subsequent restriction analysis allowed the selection of clones encompassing the 5' part of the gene.

2.4 Mapping of 5' cDNA Ends

Three protocols were followed to map the transcriptional start site; S1 nuclease protection, primer extension and rapid amplification of cDNA ends (RACE) by the polymerase chain reaction (PCR). The sources of messenger RNA were human liver PolyA+(mRNA) and PolyA+(mRNA) extracted from HepG2 cells according to manufacturer's instructions (In Vitrogen, Canada).

2.4.1 S1 Nuclease Protection Assay

Two single stranded probes were prepared by asymmetric PCR amplification of a genomic DNA insert (Gyllensten, 1988). The 5' end of these probes was either complementary to the 3' boundary of exon 1, 5'-CTTTTCCAGAGGTTACAGT-3', or to intervening sequences located in the first intron, 5'-GGACGTCTTCCAAACAGGTCT-3'. The 3' end of both probes was located 216 nt upstream from the first ATG encoding the presequence of AT, 5'-CTTTAAAACTTCTCTACTAA-3'. After elution of the single-stranded PCR products from a polyacrylamide gel, each probe was 5' end labelled with

T4 kinase and [γ-^{32}P]ATP, then purified by Sephadex G25 column chromatography (Pharmacia) and stored with 5 mg/ml of tRNA. 20 pmoles of each probe were hybridized overnight at 47°C with 500 ng of human liver PolyA+(mRNA), after preincubation for 10 min. at 75°C in buffer containing 40 mM PIPES pH 6.4, 1 mM EDTA, 400 mM NaCl and 70% formamide. S1 nuclease digestion (100 units) was done at 37°C (see figure 1), in the presence of carrier DNA (7 ug) for 30 min or 60 min. Control experiments without enzyme were done concomitantly. After stopping the reaction, the products were extracted, ethanol precipitated and electrophoresed in pre-run 6% urea-polyacrylamide gel containing formamide. A sequencing ladder of the probe was used as a size marker. The gels were autoradiographed after drying and -70°C exposure.

2.4.2 Primer extension

For this reaction, two primers were tested; the first was identical to the first primer of the S1 nuclease assay, and the second, with a location nested to the former, was 5'-CATTGGAATACATGGCCGCTAATC-3'. After 5' end labeling and purification, 20 pmoles of each primer were hybridized at 48°C for 120 min to 1 ug of human liver PolyA+(mRNA) after a 10 min preincubation at 70°C. This was carried out in buffer containing 50 mM Tris.HCl, 75 mM KCl, 3 mM $MgCl_2$, I mM DTT and 40 units of RNAse inhibitor. After 5 min at 37°C, 100 units of Moloney Murine Leukemia Virus Reverse Transcriptase (BRL) were added. After 120 min at 37°C, the reaction was stopped, the nucleic acids extracted, ethanol precipitated and resuspended in 50 mM Tris 1mM EDTA pH 8.0. Aliquots of the reaction were run in 6% urea-polyacrylamide

denaturating gels and the gels were autoradiographed after drying and -70°C exposure. Sequencing ladders with 5' boundaries identical to those of the probes were used as markers.

2.4.3 Rapid amplification of the 5' cDNA end (RACE protocol)

This approach was also attempted in order to assess precisely the transcriptional start site location. The experimental approaches followed for amplification were as described by Frohman (1988) and Shuster (1992). Briefly, first strand cDNA synthesis from human liver poly(A) mRNA was performed with the antisense primer 5'-CTCCATCAGTTGCTGGAGGG-3', complementary to the second exon of the AT gene. After RNA removal, cDNA purification, homopolymeric tailing with dCTP as well as terminal transferase, amplification was done using a nested primer complementary to AT exon 2, 5'-GCTTGGCTGTGCAGATGTCCACAGG-3', and an anchor primer with multiple cloning sites. The conditions for amplification were a 5 min -95°C precycle of denaturation, followed by a 5 min -72°C incubation for adding Taq polymerase. 35 cycles of amplification (92; 2 min, 55; 2 min, 72; 2 min), were followed by a 15 min incubation at 72°C. The PCR products were subcloned in pBluescriptIIKS+ (Stratagene, Canada) and sequenced by the dideoxy-chain termination method with T7 DNA polymerase and [α^{32}P]dATP (Tabor,1987).

2.5 Plasmids and Antithrombin Constructs Tested in Reporter Assays

The AT gene constructs tested were derived from a 9500 nt BglII/BglII insert of

a normal genomic human AT clone. A Hind III/Hind III fragment of this clone, starting 4800 nt upstream to the first exon and ending 2100 nt downstream to the first exon, in the first intervening sequence of the gene, was subcloned in the plasmid pUC19 (Stratagene, Canada). This subclone and deletions derived therefrom were inserted into pSVOA-LΔ5'at a blunt HindIII site; these AT constructs were inserted in pSVOA-LΔ5' between the upstream (nt +9) and the presumed initiation codon (nt +52) of the luciferase cDNA. Constructs of AT deletions downstream of -1100 nt were made following subcloning of -1100/+68 nt BamHI/EaeI into pGEM-3Zf(+) and of -150/+68 nt DraI/EaeI into pGEM-7Zf(+). A detailed list of the restriction sites used in the reporter assay mutants is given below in Fig. III.11 and III.12. Some constructs were also synthesized by the polymerase chain reaction - for example the constructs ending at +11 nt of the AT gene in Fig. III.12. In this latter case the amplified material was the initial genomic clone and the 3' primer was 5'-CTGAAAACTGGTTCTTTCC-3'. After colony hybridization with a radiolabelled AT probe and mini-preparations of plasmid DNA, the orientation of the sequences tested was confirmed by restriction analysis, and if necessary by sequencing by the dideoxy-chain termination method. Large scale plasmid preparations used either two CsCl gradient centrifugations (Sambrook,1989) or anion exchange chromatography according to manufacturer's instructions (Quiagen,Canada). DNA concentrations were measured by fluorimetry (Sambrook,1989).

To confirm promoter or enhancer activity of the elements mapped in the plasmid pSVOA, AT deletions were also inserted in a luciferase reporter plasmid assay system (pGL, Promega) containing -in order to detect enhancing activity- the SV40 derived

minimal promoter (pGL-promoter), or -in order to confirm promoter activity- an enhancer, also SV40 derived (pGL-enhancer). The AT sequences tested in this fashion were the -640/+68 nt and the -150/+68 nt fragments of the promoter. These regions were tested upstream or downstream from the luciferase coding sequences in the pGL vectors.

Control plasmids for reporter assays were luciferase plasmids containing the SV40 early promoter-enhancer complex, or the liver-specific proximal promoter from the carbamoyl-phosphate synthetase, pCPS-Luc. A SV40 β-galactosidase derived reporter plasmid (Promega) was also used as an internal control for transfection efficiency. The parent luciferase plasmid pSVOA-LΔ5'was also included for background luciferase expression.

Single and multiple copies of the binding sites for the areas -92/-65 nt and +1/+37 nt (with added restriction sites) were inserted upstream to the minimal promoters of CPS and TK as described (Ausubel,1989). Briefly, single copies into pUC19 were digested with KpnI/SalI and inserted between the KpnI/XhoI sites of both reporter plasmids. Multiple copies of the elements obtained from self-ligation were subcloned into pBluescript(KS)+ (Stratagene), digested with KpnI/SacI or HindIII/SacI, and inserted into the KpnI/SacI sites of pCPS-Luc or the HindIII/SacI sites of pTK-Luc.

2.6 Cell Culture and Maintenance, Transfection Protocols

Cell lines tested for transfection were obtained from the American Type Culture Collection. Four transformed cell lines were used: the human hepatoma cell line HepG2, the simian kidney cell lines Cos1 and BSC40, and the human fibroblastic cell line Hela

(Aden,1979; Gluzman,1981; Darlington,1987). The cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 100 μg streptomycin/ml and 100 μg penicillin G/ml. DNA transfections were mediated either by calcium phosphate precipitation or by lipofection. Calcium phosphate transfection was was performed as described (Graham,1973; Zhang,1992) with the following modifications; cells at 50-60% confluency were transfected with 25 μg DNA (20 μg plasmid plus 5 μg carrier DNA) by calcium phosphate precipitation (10), followed by a 10% glycerol shock treatment 16 hours post transfection (with the exception of Cos1 cells). Transfection by lipofection was performed according to manufacturer's instructions (Gibco/BRL) and to Carlsson (1989) with minor modifications -DNA and liposomes were incubated 6 hours for Cos1 cells or 20 hours for HepG2 cells in the absence of carrier DNA.

In cotransfection experiments, all cells were incubated for 24 hours before and during transfection in medium without phenol red and with 5% charcoal-stripped fetal bovine serum as described (Marcus,1993). Transfection was mediated by the calcium phosphate method followed by a glycerol shock for HeLa and HepG2, and by a DMSO shock for BSC40. Cells were cotransfected with 5 μg of the AT promoter -150/+68 nt -luc and 1-8 μg of expression vector. The amount of plasmid transfected was normalised to 20 ug per plate with the parents expression vectors pSG5 or pRc/CMV plus carrier DNA (10 μg of salmon sperm DNA/per plate).

Reporter activity was measured 48 hours post-transfection by luminescence detection upon luciferase expression. Photon emission was quantified after direct addition of cell lysates to a luciferyl-CoA substrate, in the presence of Mg^{++} and ATP according

to manufacturer's instructions (Promega) and using a model 1253 Bio-Orbit luminometer (Oy, Turku, Finland). Variations in luciferase expression were normalized to the protein content of the lysates (Bradford assay) and to β-galactosidase activity driven by the SV40 promoter and determined by ONPG hydrolysis and spectrophotometry according to manufacturer's instructions (Promega).

2.7 Nuclear Extracts

Total extracts of nuclear proteins were prepared according to the protocol described by Dignam (1983) with minor modifications. Briefly, nuclei were isolated at low ionic strengh (in 0.01 M KCl) in a teflon Dounce homogenizer (pestle B). Cell lysis was monitored under the microscope by trypan blue exclusion. Soluble proteins were extracted from nuclei in 0.42M NaCl or, alternatively, by ammonium sulphate precipitation (Parker,1984). It was not intended in this study to test cytosolic fractions in gel retardation assays (S100 fractions). Rat liver nuclear extracts were prepared as described by Hattory (1990).

2.8 Footprint Analysis

The probe used to generate the footprints was a -150/+68 nt DraI/EaeI fragment of the AT gene, made blunt and subcloned in both orientations into the SmaI site of the plasmid pGEM-3Zf(+). The plasmid was digested with EcoRI, 3'end-labelled with the Klenow fragment and [α-^{32}P]dATP, and digested again with HindIII. Following removal of unincorporated radionucleotides, each restriction fragment was intercepted on a NA45

membrane during agarose gel electrophoresis. Approximately 5×10^5 cpm (1 to 5 ng) were used for each reaction. Sequencing ladders were also generated with each labelled probe according to the protocol of Maxam and Gilbert using a commercial kit (Boehringer Mannheim). G+A reactions, which do not necessitate the use of dimethylsulfate or hydrazine, were generally used as negative controls for the footprints. Protection by total extracts from the three cell lines studied (HepG2, Cos1, and HeLa) were performed on ice for 15 min. in the presence of BSA, poly(dI-dC)-poly(dI-dC), carrier DNA as described (Graves,1986). The initial method was modified as follows; Tris-HCl pH 7.5 buffer was substituted to Hepes-KOH pH 7.5. Preliminary experiments were done to determine the concentrations of non specific competitors, of extract, and of DNAseI to optimize footprint detection. Subsequent standard conditions per reaction used a mixture of salmon sperm DNA and poly(dI-dC)-poly(dI-dC) (2.5 μg each), and 100 μg of proteins unless indicated otherwise (extract+\-BSA)in a final volume of 50 μl. After preincubation of the above mix for 15 min on ice, DNAse was added for 2 min. The concentration range for the enzyme was from 2 to 15 units, based upon the probes lengh and sequences, and the protection strengh of the extracts. The final concentrations in $MgCl_2$ and $CaCl_2$ were 5 and 1 mM, respectively. The reactions were terminated by addition of an equal volume of 20 mM tris-HCl pH 8.0, 20 mM EDTA, 250 mM NaCl, and 0.5% SDS. After digestion with proteinase K for 30 min at 37° C in the presence of 50 μg tRNA, phenol-chloroform extraction, and ethanol precipitation, the products (after heat denaturation) were separated in pre-run polyacrylamide-8M urea sequencing gels and autoradiographed after drying.

2.9 Gel Retardation Assays with Total Nuclear Extracts

Electrophoretic mobility shift assays (EMSA) used total nuclear extracts prepared from monolayer cultures of HepG2, Cos1, HeLa, and BSC40 cells. The probes were the original inserts used for footprinting (see sequence in the previous paragraph), a -304/+68 nt HincII/EaeI fragment, as well as oligonucleotides corresponding to the three areas protected in the footprints. In the latter case BamHI/BglII cloning sites were added to facilitate subcloning. A list of the oligonucleotides used in retardation assays is given in Table II.2. This table provides also the sequences of the mutant probes tested, and the sequences of the binding sites for known transcription factors tested in competition assays. Additional oligonucleotides for binding of CTF/NF1, NFKB, TFIID, SP1, AP1, AP2, and AP3 were obtained from Promega Inc. Complementary strands of each probe to be radiolabelled were annealed and subcloned into the plasmid pUC18. Following restriction enzyme digestion, the inserts were purified by agarose gel electrophoresis with NA45 on 4% NuSieve, and 3' end-labelled with the Klenow fragment and [α-^{32}P]dATP. After phenol-chloroform extraction the unincorporated radionucleotides were removed with Sephadex G25. Probe B encompassed only an area upstream from the start site strongly protected in the three cell lines tested in the footprints (+1/+37 nt). In order to encompass also sequences upstream to the start site at -14/+3 nt -an area protected with cell source associated differences, a -28/+47 nt NlaIV/BstX1 fragment of the AT promoter into pGEM-7f(+) was also used for gel retardation after 3' end labelling and restriction enzyme digestion.

CONSENSUS	OLIGONUCLEOTIDE SEQUENCES
ELEMENT A	
Wild type	GGTCATCAGCCTTTGACCTCAGTTCCG
5' Truncated	CATCAGCCTTTGACCTCAGTTCCG
Mutant 1	GGTCATCAGCCTGAggTCAGTTCCG
Mutant 2	GGTCATCAGggaaaGACCTCAGTTCCG
Mutant 3	GGTCATCAGCCTTTcAggTCAGTTCCG
Mutant 4	GGTCATCAGggaaaTGAggTCAGTTCCG
PYBP-Pyr Strand	CTTTGACCTG
PYBP-Pur Strand	CAGGTCAAAG
ELEMENT B	
Wild type	CCAGTTTTCAGGCGGATTGCCTCAGATCAGATCAC
Mutant 1	CCAGggggCAGGC
Mutant 2	GGccccccCTCAGATC
Mutant 3	CCAGggggCAGGCccccCTCAGATC
ELEMENT C	
Wild type	CAACACTGGGCTCTACTTTGCTTAACCG
COMPETITION	
HNF4	TCGAGGCAAGGTTCATATTTGTGTAGG
COUP-TF1	TTTCTATGGTGTCAAAGGTCAAACTG
HNF3	CCTGATTCTGATTATTGACTTAGTCAACG
C/EBP	AGATTGCGCAATCTG
CTF/NF1	CCTTTGGCATGCTGCCAATATG
NFκB	AGTTGAGGGGACTTTCCCAGGC
TFIID	GCAGAGCATATAAGGTGAGGTAGGA
AP1	CGCTTGATGAGTCAGCCGGAA
AP2	GATCGAACTGACCGCCCGCGGCCGT
SP1	ATTCGATCGGGGCGGGGCGAGC
AP3	CTACTGGGACTTTCCACACATC
HD-PPRE	TCTCCTTTGACCTATTGAACTATTACCTACATT

Table II.1: ***Oligonucleotides used in gel retardation assays.*** The sequences are given in the 5' to 3' direction and for the upper-strand only.

Standard binding reactions were incubated with each probe 15 min. at 37°C, utilizing 10 μg of extracts, in the presence of poly(dI-dC).poly(dI-dC), bovine serum albumin and carrier DNA (Gillman,1986). After titration experiments, EMSA with the -304/+68 nt probe were done in the presence of 5 μg of poly(dI-dC).poly(dI-dC), 5 μg of a mixture of *E. coli* and salmon sperm DNA, and 5 μg BSA for 10 μg of extract. EMSA with the -150/+68 nt probe were done in the presence of 2.5 μg of poly(dI-dC).poly(dI-dC), 2 μg BSA, and 5 μg carrier DNA. The oligonucleotides probes A, B, and C, and the mutants of probe A and B, were assayed in the presence of of 1 μg of poly (dI-dC).poly(dI-dC), 2 μg BSA, and 0.5 μg carrier DNA. Buffer conditions were as described (Gillman,1986). In competition assays, unlabelled oligonucleotides was were used at a 100-molar excess of the test oligonucleotide and added 5 min before the latter. In supershift assays, 1 μl of a 1/10 dilution of each antiserum or 0.5 μg IgG was used unless indicated otherwise. Antisera were added just after the labelled probe and the reaction was incubated for a further 30 min. Electrophoresis was carried at 4°C on pre-run 4% 30:1 acrylamide:N-N'-methylenebisacrylamide in 1mM EDTA, 6.7 mM Boric acid, and 6.7 mM Tris-HCl pH 7.5. The gels were then autoradiographed overnight at -70°C.

2.10 *In Vitro* Transcription - Translation and Gel Retardation Assays with Transcription Factors

2.10.1 *In Vitro* Transcription Translation

In vitro transcription-translation assays were carried out with a rabbit reticulocyte

lysate system according to the manufacturer's instructions (Promega). This simplified approach of the original protocol of Krieg and Melton (Krieg 1987) coupled the two steps of mRNA synthesis and of translation in one reaction. Briefly, 1 μg of plasmid was incubated at 30°C for 90 to 120 min with 25 μl of rabbit reticulocyte lysate in the presence of SP6 or T7 RNA polymerase, amino acid mixture, and ribonuclease inhibitor (RNAsin, Promega) in a final reaction volume of 50ul. Reactions were done concomitantly with or without ^{35}S-methionine. In the first case, 1 μl of 1 mM amino acid mixture minus methionine and 40 μCi of ^{35}S-methionine (1000 Ci/mmol) were added. In the second case, 1 μl each of amino acid mixture minus methionine, and of amino acid mixture minus leucine were added. The first reaction was used to confirm the efficiency of the translation. The second reaction was used in EMSA assays. A plasmid expressing luciferase under the control of T7 RNA polymerase was used as a positive control for transcription-translation efficiency.

2.10.2 Analysis of the Translated Products

An aliquot of the previous reactions (2.5 μl) was treated at 95°C for 5 min in the same volume of 2×SDS-PAGE denaturating loading buffer (4% SDS, 15% glycerol, 62.5 mM Tris-HCl pH 6.8, 0.005% bromophenol blue, 200mM DTT). After cooling on ice, samples were loaded into mini slab gels (Bio-Rad Laboratories) with separation into 10% bis-polyacrylamide gels (1:30) in 0.1% APS, 0.1% TEMED, and 0.4 mM Tris-HCl pH 8.8 (Laemmli,1970). The electrophoresis was carried out into 25 mM Tris-HCl pH 6.8, 250 mM glycine, and 1% SDS. Low molecular weight protein markers (Bio Rad Labs)

were utilized. The gels were dried, then autoradiographed overnight on Kodak SK+ film.

2.10.3 Gel Retardation Assays

EMSA assays with translated products used the experimental conditions and the non-specific competitors described in the reactions with total extracts. 1-2 μl of translation mixture was used to test the individual binding profile of each factor. 1 μl of translation mixture/per factor was used to test their combinatorial effects. Unprogrammed lysate and the translation product of the luciferase expression plasmid provided by the kit were used as negative controls. A standard total amount of 4 μg of translation product was maintained by adding unprogrammed lysate to the translation products tested in each reaction. Competition assays and addition of antibodies were carried out as described for the total extracts.

CHAPTER III

RESULTS

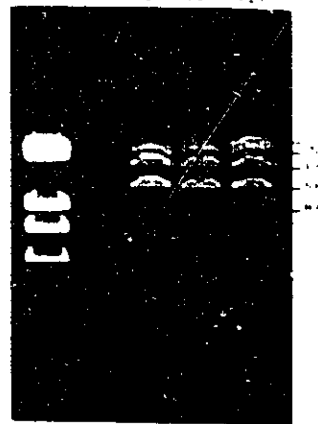
1- Genomic Clones Encompassing the 5' Upstream Region of the Antithrombin Gene and Restriction Mapping of the Antithrombin Gene

The experimental approach to obtain genomic clones corresponding to the 5'end of the AT gene is shown in Fig.III.1, Panels I-III. Panel I illustrates the sizing of Sau3A digested DNA; the 20Kbp fragments of lane 1 (Lane 2 is the undigested DNA) were used for ligation with EMBL3 arms. Panel II illustrates the ligation steps; lane 1 is the Bam H1 restricted phage DNA. Lanes 2 to 4 are aliquots of the ligation reactions at a 1:1, 2:1 and 4:1 molar excess of insert to EMBL3 arms. Reactions 3 and 4 were used to generate the library which contained 1.48×10^6 independent clones (insert size: 7-20 Kbp). 7 AT specific clones were found; 2 of those mapped to the 5' region of the gene whereas the remaining clones encompassed other regions of the gene. Panel III illustrates primary, secondary and tertiary screens; the positive signal seen in plaque A was one of two clones containing the 5'upstream region of the AT gene. The first 5' clone isolated encompassed a 7000 nt region upstream of exon 1, IVS 1, exon 2, IVS 2, exon 3, and ended in IVS3; the region mapped by this clone is indicated in Panel IV of Fig.III.1.The second 5' clone encompassed a large region upstream of exon 1, exon 1,IVS 1, exon 2 and ended in IVS 2 (Panel 4 of Fig.III.1). Restriction analysis of these two clones was facilitated by using the truncated cDNA probes encoding different regions of the gene downstream of exon 1 and by using genomic probes upstream from this region. Restriction analysis allowed us

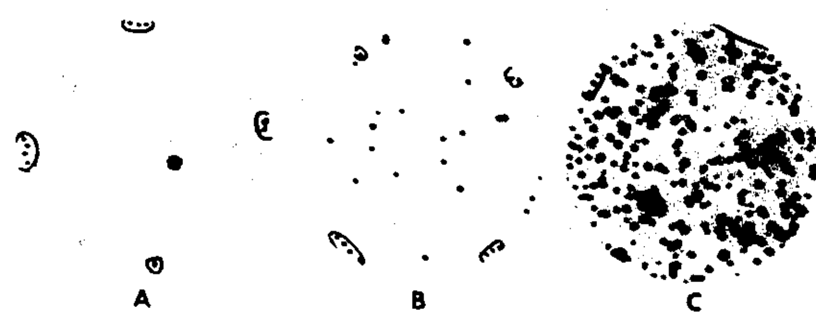
Figure III.1: ***Cloning of genomic inserts for the 5' region of the antithrombin gene.***
Panel I : Sau3A restricted DNA sucrose gradient sizing (agarose gel to the top): Selection of 20 Kbp sized fragments (left lane, agarose gel to the bottom). The DNA markers are a Hind III digest of the λ phage. **Panel II :** Ligation reactions. LA, left arm; SF, stuffer fragment; RA, right arm; INS, insert; L-INS, ligated insert. The map and gel digest to the left show EMBL3 restriction. The figure to the right shows the products of ligation at different molar ratios of insert to vector (1:1 to 4:1). **Panel III :** Primary, secondary, and tertiary screens with an antithrombin genomic probe. **Panel IV :** Restriction map of clones encompassing the 5' region of the AT gene. The two clones are indicated by the arrows under the AT gene restriction map; the 3' boundary of each clone is indicated by the arrowhead.

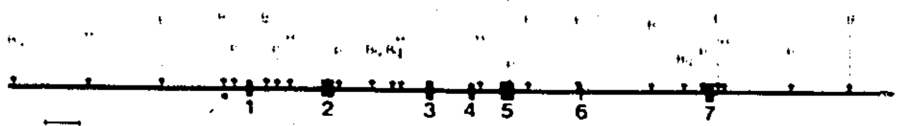
I

II

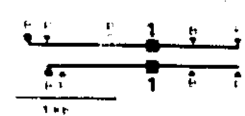


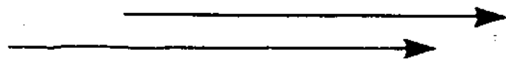
III





IV





also to generate the normal AT gene maps given in Fig.I.1 and in Panel IV of Fig.III.1. The second 5' genomic clone was used for sequencing and to generate the constructs to be tested subsequently in reporter assays. This clone included also the 108 nt length polymorphism upstream from exon 1 (Bock,1985).

Fig.III.2 and III.3 provide the nucleotide sequences of the regions immediately flanking exon 1; in Fig.III.2 those sequences start at a BamHI site 1100 nt before the translational start site in exon 1 and end at a second BamHI site 304 nt after the end of first exon. Sequence determination from a normal clone of the library was performed by L Bardossi. Fig.III.3 provides the nucleotide sequence of the +391/+895 nt BamHI-PstI region of IVS 1; the data-base was from a published reference (Olds,1993). Fragments of the sequences in Fig.III.2 and III.3 will be shown in the following paragraphs to contain active *cis*-acting elements.

2. Type I Antithrombin Deficiency : Restriction Mapping of the Abnormal Allele and PCR Confirmation of Deletion in the 5' Upstream Region

Fig.III.4 and III.5 illustrate the approaches followed for Southern analysis. In Figure III-4, Panel I is intended to describe the genetic traits of the family. This family, with a history of recurrent thromboembolic episodes, showed the biological characteristics of a type I kindred with lack of expression of one AT allele (immunological and functional AT levels reduced by 48-52% for the individuals shown by the black symbols). Also in this panel, each family member has been given a coded number used in all subsequent figures. The results of Southern analysis for the various enzymes tested have been reported in

```
BamHI     10          20          30          40          50          60
           |           |           |           |           |           |
   1 GATCCAGGGT CTGAATCAAG CCAAGAACTT AGACACAGCT TCAGTTCAGG AGAGATGTCC
     CTAGGTCCCA GACTTAGTTC GGTTCTTGAA TCTGTGTCGA AGTCAAGTCC TCTCTACAGG

  61 TCTCTCATGA AACGAAGAGC TCTCGGAAAA TGCCTATGAC TGAAACAACA ACTACAAGGA
     AGAGAGTACT TTGCTTCTCG AGAGCCTTTT ACGGATACTG ACTTTGTTGT TGATGTTCCT

 121 GTCCTTGATC ACACAGCAGG AGGCACATGC GCCCTGTGAA CTTGGTGGCT GCAGTCCTTG
     CAGGAACTAG TGTGTCGTCC TCCGTGTACG CGGGACACTT GAACCACCGA CGTCAGGAAC

 181 TGATGGTGCT TTGGCTATTT TCAGAAGAGG GAAATGTAAA CAAACCTGCC CTTTATCTCC
     ACTACCACGA AACCGATAAA AGTCTTCTCC CTTTACATTT GTTTGGACGG GAAATAGAGG

 241 ACTAGTTTGA ATTTACCCAA TCTCCCCAAT AACTAATAAA AATAACGATA CACATAAATA
     TGATCAAACT TAAATGGGTT AGAGGGGTTA TTGATTATTT TTATTGCTAT GTGTATTTAT

 301 TAATTTAGTT TGTTTTTATT TCAGCAGTTG TACGTGGGTT TTTCACCTGC CCCGCATTCT
     ATTAAATCAA ACAAAAATAA AGTCGTCAAC ATGCACCCAA AAAGTGGACG GGGCGTAAGA

 361 GTCTCCTGAT CCCCCAGTAG AGTTTTGCTA AGTATTTCCC AGTGCTGACA CCCCTTAGAA
     CAGAGGACTA GGGGGTCATC TCAAAACGAT TCATAAAGGG TCACGACTGT GGGGAATCTT

 421 ACGCGCTTGG CATGCACCCC GAGGCCCTGC TCTTCTCTCC CTGTCCACCA CTTCAGGGCT
     TGCGCGAACC GTACGTGGGG CTCCGGGACG AGAAGAGAGG GACAGGTGGT GAAGTCCCGA

 481 GCTGGGGAAT GGGTCTCTCT GTGGGCCACA GGTGTAACAT TGTGTTTTTC CTTGTCTGCT
     CGACCCCTTA CCCAGAGAGA CACCCGGTGT CCACATTGTA ACACAAAAAG GAACAGACGA

 541 GCCAGGGACA CCTTGGCATC AGATGCCTGA AGGTAGCAGC TTGTCCCTCT TTGCCTTCTC
     CGGTCCCTGT GGAACCGTAG TCTACGGACT TCCATCGTCG AACAGGGAGA AACGGAAGAG

 601 TAATTAGATA TTTCTCTCTC TCTCTCCCTC TCTCCATAAA GAAAACTATG AGAGAGGGAA
     ATTAATCTAT AAAGAGAGAG AGAGAGGGAG AGAGGTATTT CTTTTGATAC TCTCTCCCTT

 661 TTACACAGGT AGAGGGCTAG AAGTTTTTGG ACATTAACTA TTTCTATCTT CTGATTTAGT
     AATGTGTCCA TCTCCCGATC TTCAAAAACC TGTAATTGAT AAAGATAGAA GACTAAATCA

 721 TAACGAGAAA CAAAAAATCC TGCAGACAAG TTTCTCCTCA GTCAGGTATT TCCTAACCAA
     ATTGCTCTTT GTTTTTTAGG ACGTCTGTTC AAAGAGGAGT CAGTCCATAA AGGATTGGTT

 781 GTTTGAGGGT ATGAACATAC TCTCCTTTTC CTTTTCTATA AAGCTGAGGA GAAGAGTGAG
     CAAACTCCCA TACTTGTATG AGAGGAAAAG GAAAAGATAT TTCGACTCCT CTTCTCACTC

 841 GGAGTGTGGG CAAGAGAGGT GGCTCAGGCT TTCCCTGGGC CTGATTGAAC TTTAAAACTT
     CCTCACACCC GTTCTCTCCA CCGAGTCCGA AAGGGACCCG GACTAACTTG AAATTTTGAA

 901 CTCTACTAAT TAAACAACAC TGGGCTCTAC ACTTTGCTTA ACCCTGGGAA CTGGTCATCA
     GAGATGATTA ATTTGTTGTG ACCCGAGATG TGAAACGAAT TGGGACCCTT GACCAGTAGT

 961 GCCTTTGACC TCAGTTCCCC CTCCTGACCA GCTCTCTGCC CCACCCTGTC CTCTGGAACC
     CGGAAACTGG AGTCAAGGGG GAGGACTGGT CGAGAGACGG GGTGGGACAG GAGACCTTGG
                                  +1
                                  ->
1021 TCTGCGAGAT TTAGAGGAAA GAACCAGTTT TCAGGCGGAT TGCCTCAGAT CACACTATCT
     AGACGCTCTA AATCTCCTTT CTTGGTCAAA AGTCCGCCTA ACGGAGTCTA GTGTGATAGA

1081 CCACTTGCCC AGCCCTGTGG AAGATTAGCG GCCATGTATT CCAATGTGAT AGGAACTGTA
     GGTGAACGGG TCGGGACACC TTCTAATCGC CGGTACATAA GGTTACACTA TCCTTGACAT
                   |IVS 1
1141 ACCTCTGGAA AAAGGTAAGA GGGGTGAGCT TTCCCCTTGC CTGCCCCTAC TGGGTTTTGT
     TGGAGACCTT TTTCCATTCT CCCCACTCGA AAGGGGAACG GACGGGGATG ACCCAAAACA

1201 GACCTCCAAA GGACTCACAG GAATGACCTC CAACACCTTT GAGAAGACCA GGCCCTCTCC
     CTGGAGGTTT CCTGAGTGTC CTTACTGGAG GTTGTGGAAA CTCTTCTGGT CCGGGAGAGG

1261 CTGGTAGTTA CAGTCAAAGA CCTGTTTGGA AGACGTCATT TCAAGTGCTC TCCCTCCCAC
     GACCATCAAT GTCAGTTTCT GGACAAACCT TCTGCAGTAA AGTTCACGAG AGGGAGGGTG

1321 CCCACCTCTT GGGGTAAGGC CTTTCCTAAG CTACCCCTTG GGTCCCTAGC CTAAGAAACA
     GGGTGGAGAA CCCCATTCCG GAAAGGATTC GATGGGGAAC CCAGGGATCG GATTCTTTGT

1381 AGGGGGATGT CATCCCTGGT GTAAAGATGC TGTGCAGGAA GTCAGCACTC ACGGGATC
     TCCCCCTACA GTAGGGACCA CATTTCTACG ACACGTCCTT CAGTCGTGAG TGCCCTAG
                                                              BamHI
```

Figure III.2: ***Nucleotide sequence of the antithrombin proximal region flanking exon1.*** The presumed transcriptional start site is indicated by the arrow; the first translation codon is boxed; the exon1/IVS1 boundary is also indicated.

```
BamH I
GGA TCC AGG GGA CGC TCC AAG GGG AAT CCC CAG GGC CTG CCA TCC
CCT AGG TCC CCT GCG AGG TTC CCC TTA GGG GTC CCG GAC GGT AGG

ATC CGG GAA GAG AGC AAA TGC TAC CCA TGA GGA CCT CCT CAC TCC
TAG GCC CTT CTC TCG TTT ACG ATG GGT ACT CCT GGA GGA GTG AGG

CTT TTT GCT CTT TCT TCC ACT CAG ATC CAC CCC ACT CCA CCC CCA
GAA AAA CGA GAA AGA AGG TGA GTC TAG GTG GGG TGA GGT GGG GGT

CCC AAA TCC CAG TGA CCT TTG ACT AAA GGG CCA AAA CTG CTT CCT
GGG TTT AGG GTC ACT GGA AAC TGA TTT CCC GGT TTT GAC GAA GGA

TTT CTC ACA ATG AGA GTT GTC CCT CCC TCA ATG CCA CAC ACA CTC
AAA GAG TGT TAC TCT CAA CAG GGA GGG AGT TAC GGT GTG TGT GAG

CCT TCT TCA TCT GAG TTG TCA CAG GAG GCT AGA AAC GGG GTG GTG
GGA AGA AGT AGA CTC AAC AGT GTC CTC CGA TCT TTG CCC CAC CAC

GCA CAA CTG TCT TGG TTT TAA TTT GTG CTT CAT AGC CCT CCC AGG
CGT GTT GAC AGA ACC AAA ATT AAA CAC GAA GTA TCG GGA GGG TCC

TCC TCT CAG CCT CAA ATT GCA TTT CCA AAT GTA GTT GAA GGG ACA
AGG AGA GTC GGA GTT TAA CGT AAA GGT TTA CAT CAA CTT CCC TGT

GAG TGG GCA ACC GAA GCA GCA GTG GAG AGA TGG GAA GAT GAA TGG
CTC ACC CGT TGG CTT CGT CGT CAC CTC TCT ACC CTT CTA CTT ACC

CAG GGT CCT CTC CTC TCT CTC TCT GCT TCT TCA GCC TGC CTT CCA
GTC CCA GGA GAG GAG AGA GAG AGA CGA AGA AGT CGG ACG GAA GGT

CAT CTC CCT TGG TGC CGC TGC TTC TCT CCG GCT TTG CAC CTC TGT
GTA GAG GGA ACC ACG GCG ACG AAG AGA GGC CGA AAC GTG GAG ACA
                          Pst I
TCT TGA AAG GGC TGC AG
AGA ACT TTC CCG ACG TC
```

Figure III.3: ***Nucleotide sequence of the +391/+895 nt BamHI-PstI region of AT-IVS1.*** *(from Olds and Lane, 1993).*

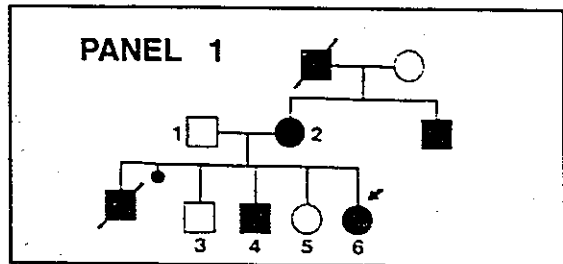
greater detail (Fernandez-Rachubinski, 1992). For the purpose of this thesis, only two examples have been selected, one for PstI restriction analysis, the other for HindIII+PstI. A summary of all data is also provided in Fig.III.6.

PstI analysis is shown in Fig.III.4, Panel 2. Using the full length or truncated cDNA probes, we wereable to detect an abnormal 9.5 Kbp piece in all affected individuals (indicated by a star in panel 2). No other variations from the normal pattern were noticed. In addition, RFLP analysis allowed the association of the normal AT allele from the mother of the proposita (number 2) to an intragenic polymorphism in exon 5 shown in the map of Panel 2, bottom left). This provided the basis for cloning the breakpoint deletion: effectively, in individual 2, the affected AT allele was associated with the 9.5 Kbp fragment whereas the normal allele was associated with two fragments of 5.0 and 5.5 Kbp. Restriction analysis in the 5'upstream region used the genomic probe e in the bottom right of panel 2, Fig.III.4.This last analysis strongly suggested that the deleted allele was missing this region of the AT gene; the detailed explanations are given in the legends to Fig.III.4. The direct confirmation of these findings was provided by PCR amplification of the 5'upstream region (BamHI/BamHI $1.4/1.5 \times 10^3$ nt) in genomic samples of the various family members (Panel 3 of Fig.III.4). Furthermore, restriction analysis with HindIII+PstI showed that the 3'end of the affected allele was normal and that the HindIII site normally located just 5' from exon 3 was missing (Fig.III.4). Based upon PstI, EcoRI and BamHI RFLP analysis, it was originally thought that the deletion breakpoint was close to exon 2. In fact, analyses with HindIII+PstI and with other enzymes as well indicated that the deletion breakpoint was located upstream to exon 3 in the IVS 2 region. The combined data and a proposed map for the partially deleted allele are shown in Fig.III.6.

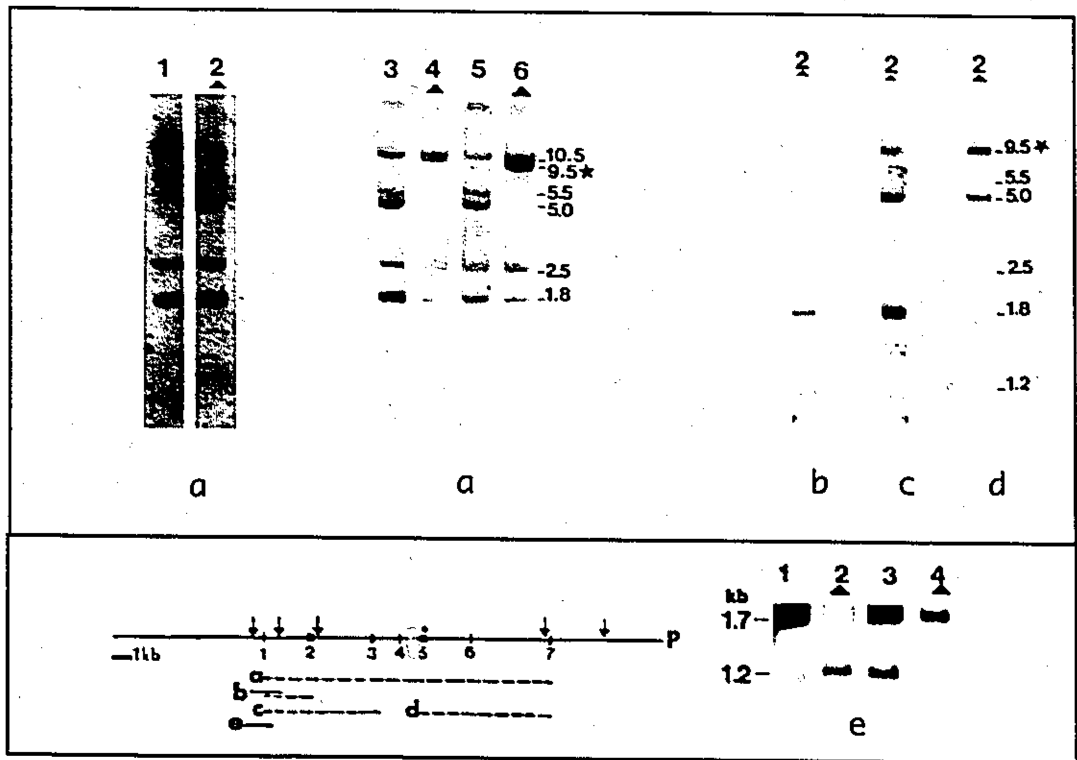
FigureIII.4: ***Type I antithrombin deficient kindred:Analysis of the 5'region of the gene.***

Panel 1: Genetic traits of the family. The arrow indicates the proposita. The code number of each family member is also used in subsequent figures. Filled boxes represent affected individuals; Open boxes represent normal individuals.
Panel 2: PstI Southern analysis. The number on top of the restricted fragments is the individual's numbers. The triangle indicates an affected individual. The star indicates an abnormal fragment. The lower case letters identify the probes. The region encompassed by these probes is indicated in the map at the bottom left part of the panel. This map also gives the normal PstI restriction pattern.Probe a was the full length cDNA; probe b was a mixture of a 5' genomic probe and of a truncated probe encompassing exons 1-2; probe c was a truncated cDNA encoding exons 1 to 3; probe d was also a truncated probe encoding exons 5 to 7; probe e was a genomic probe for the 5'upstream region.
Panel 3: PCR amplified products of the BamHI region flanking exon 1.The map in the left part of this panel details the length polymorphism in this region. (A) contains PCR products for various members of this family. (B) are PstI restriction digests of the previous products.ϕ HaeIII digest of phage 174.

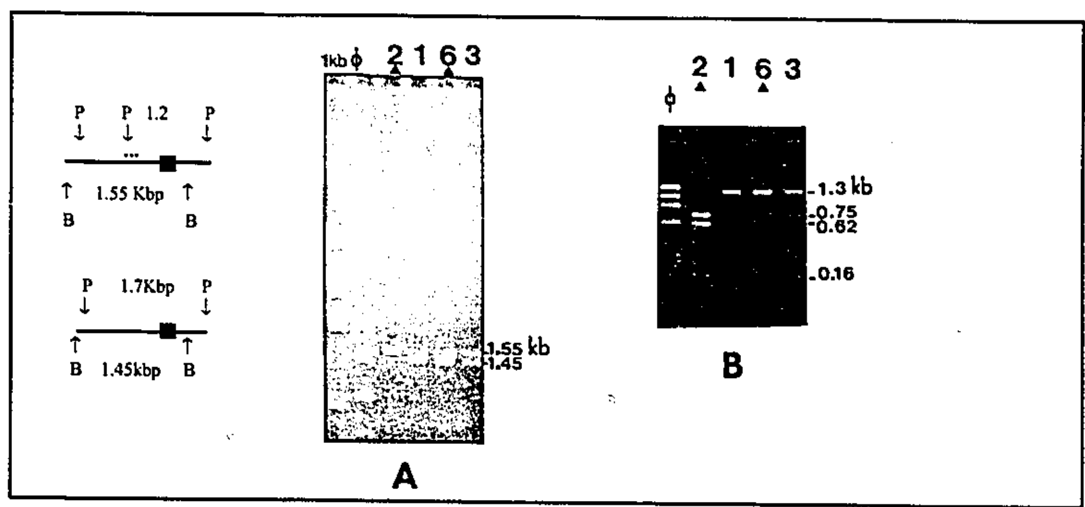
The PstI analysis of the 5' upstream region (with probe e in panel 2) and the PCR analysis showed that the region flanking exon 1 was missing in affected individuals. PstI analysis for the father (1) and the mother (2) of the proposita showed one 1.7- and one 1.2 Kbp fragment, respectively. In affected siblings, only a 1.7 Kbp fragment was seen. In unaffected siblings, 1.7- and 1.2 kbp fragments were present. If both alleles had been detected by a probe for region e, a 1.2 Kbp fragment (inherited from the mother) should have been seen in all siblings. This analysis indicated the presence of only a fragment inherited from the father in affected siblings. The PCR results shown in panel 3 confirmed that the 5' region was missing in affected individuals. For the father, a 1.45 Kbp fragment amplified, with no PstI site; for the mother, a 1.55 Kbp fragment was seen, with an internal PstI site. The amplified products, in affected siblings only, showed products of one allele, inherited from the father.



PANEL 2



PANEL 3



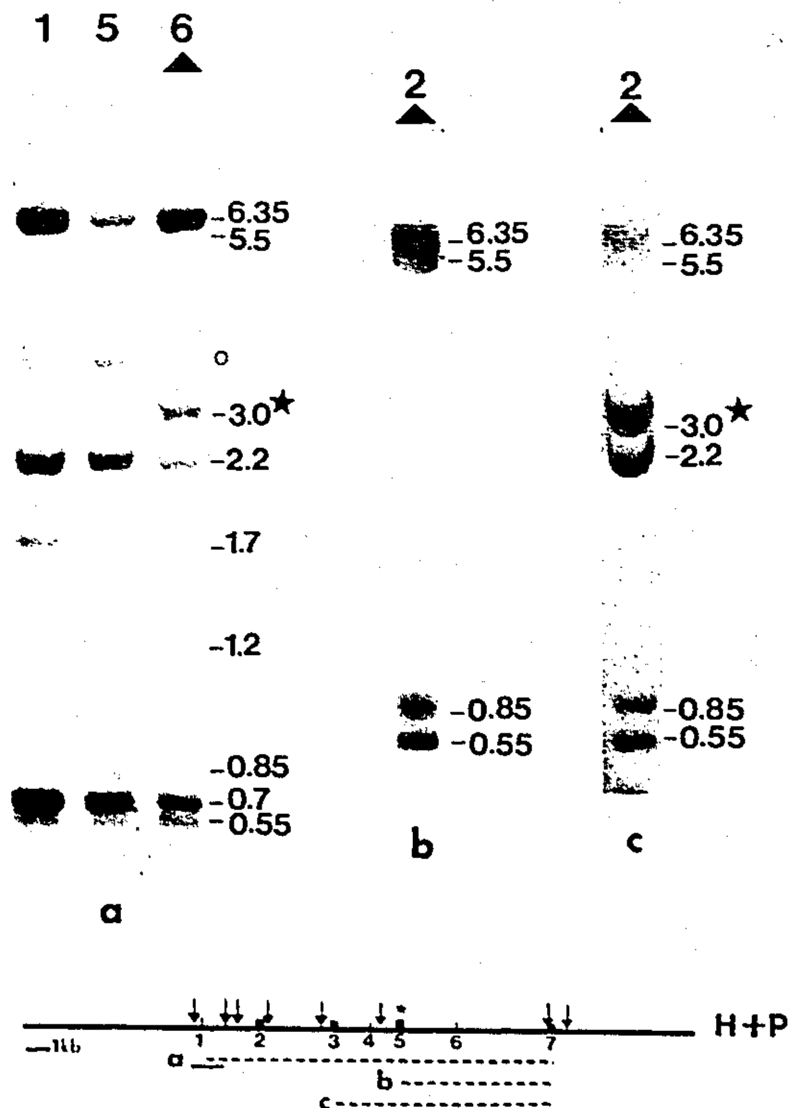
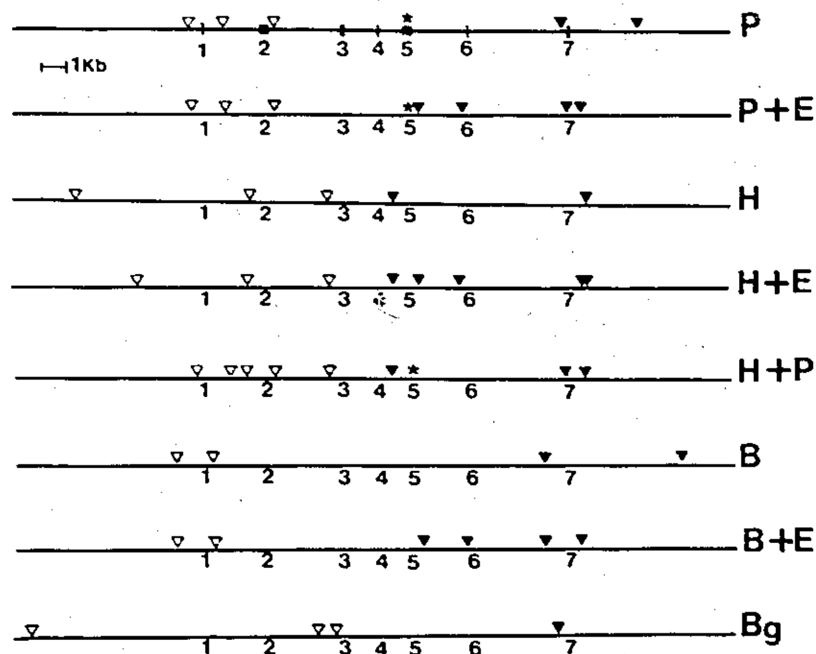
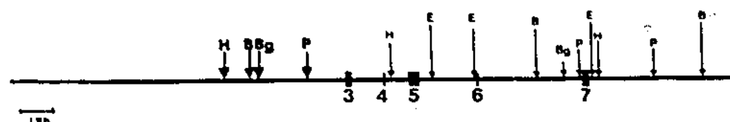


Figure III.5 ***Type I antithrombin deficient kindred: Pst I + Hind III Southern analysis.*** For this analysis, the mother of the proposita was selected, being the only affected individual for whom it was possible to differentiate the wild-type and the mutant antithrombin alleles, using an intragenic Pst I polymorphism in exon 5 . As expected, her normal allele, linked to this polymorphism,generated a 5.5 Kbp Pst I fragment (encompassing exons 5 to 7), a 0.85 Kbp Hind III-Pst I fragment (IVS 4-exon 5), and a 0.55 Kbp fragment (exon 7 and downstream). Her abnormal allele generated a 6.35 Kbp fragment. A truncated probe (c) for the regions between exons 3 and 7, showed, in addition to a probe (b) encompassing only exons 5 to 7, the presence of a normal fragment of 2.2 Kbp (Hind III exons 3 to 4), and of an unexpected fragment of 3.0 Kbp. This abnormal fragment was generated by the absence of a normal Hind III site immediately upstream to exon 3.
(see code legend in previous figure)



A



B

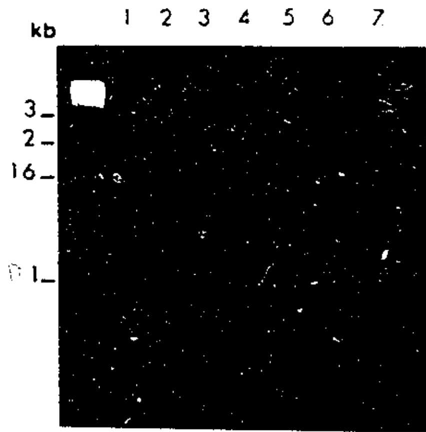
Figure III.6: ***(A) Summary of RFLP analysis in this kindred.***
(B) Proposed restriction map of the deleted allele.
▾, normal restriction site ; ▿, site missing or rearranged ; ★ intragenic PstI exon 5 polymorphism . P, Pst I; B, Bam HI; Bg, Bgl II; H, Hind III; P+E, PstI + EcoRI; H+E, Hind III + Eco RI; H+P, Hind III+ Pst I; B+E, BamHI + EcoRI.

3. Type I Antithrombin Deficient Kindred : Determination of the 5' Breakpoint of the Deleted Antithrombin Allele

Fig.III.7 illustrates the products obtained after completion of the various steps of the cloning procedure. Panels A (normal allele) and B (mutant allele) relate to branch-capture PCR whereas Panel C relates to inverse PCR. The legends to Fig.III.7 identify the intermediate products (lanes A1-6, B1-5 and C1-4) and emphasize the efficiency of the magnetic tag in increasing the specificity and the yield of PCR products. After the second round of PCR, a 2900 nt fragment was amplified in the normal allele (Fig.III.7, lane A7) whereas a 1400 nt fragment was amplified in the mutant allele (Fig.III.7, lanes B6 and C5). The specificity of the products obtained was confirmed by Southern analysis with truncated cDNA probes and restriction with various enzymes (data shown in Fernandez-Rachubinski,1992). Both alleles were subsequently sequenced at the breakpoint location. Furthermore, the presence of the mutant allele was confirmed to be present in the affected individuals genomic DNA samples by PCR amplification using a primer specific to the mutant allele located upstream to the breakpoint (Fig.III.7, Panel D). Fig.III.8 shows the nucleotide sequences of the wild-type and mutant alleles from and upstream of exon 3 (sequence in the left insert) and at the breakpoint location (illustration to the right of Fig.III.8).

Figure III.7: ***Type I antithrombin deficient kindred: PCR products generated during cloning of the breakpoint of the deleted allele.***

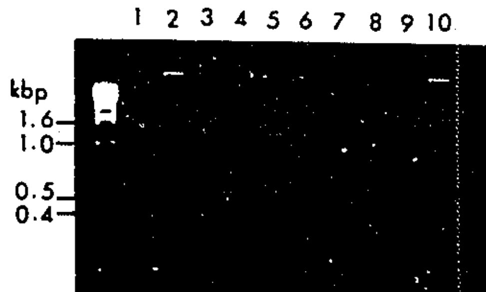
Panel A: Branch capture PCR of the normal allele.Lanes 1 and 2, control primers; lanes 3 and 4, products amplified from the ligation mixture and primer 1 (non biotinylated in lane 3, biotinylated in lane 4); lane 5, products amplified from the ligation mix, primer 1 non biotinylated, and primer 2; lane 6, captured product reamplified with primer 3; lane 7, captured product reamplified with primers 2 and 3. The primers' numbering code has been detailed in chapter II of this thesis.
Panel B: Branch capture PCR of the deleted allele.Lane 1, products amplified from the ligation mix and primer 1, unbiotinylated; lane 2, products amplified from the ligation mix and primer 2; lane 3, products amplified from the ligation mix, primer 1 unbiotinylated, and primer 2; lane 4, captured product reamplified with primers 1 and 2; lane 5, control primer 2; lane 6, captured product reamplified with primers 2 and 3.
Panel C: Inverse PCRof the deleted allele.Lane 1, control primer 1; lane 2, products from the self-ligation mix and unbiotinylated primer 1; lane 3, products from the ligation mix, primers 1 and 2 (non biotinylated); lane 4, captured products reamplified with primers 2 and 3; lane 5, control primer 2.
Panel D: Breakpoint amplification in genomic DNA samples. Lanes 1, 3, and 4, affected individuals (family members 2, 4, and 6 of figure III.4); lanes 2, 5, and 6, normal I ndividuals (family members 1, 3, and 5 of figure III.4); lane 7, control breakpoint primer; lane 8, control DNA; lane 9, control DNA and breakpoint primer; lane 10, control DNA and exon 3 specific primer.



—A—

—B—

—C—



—D—

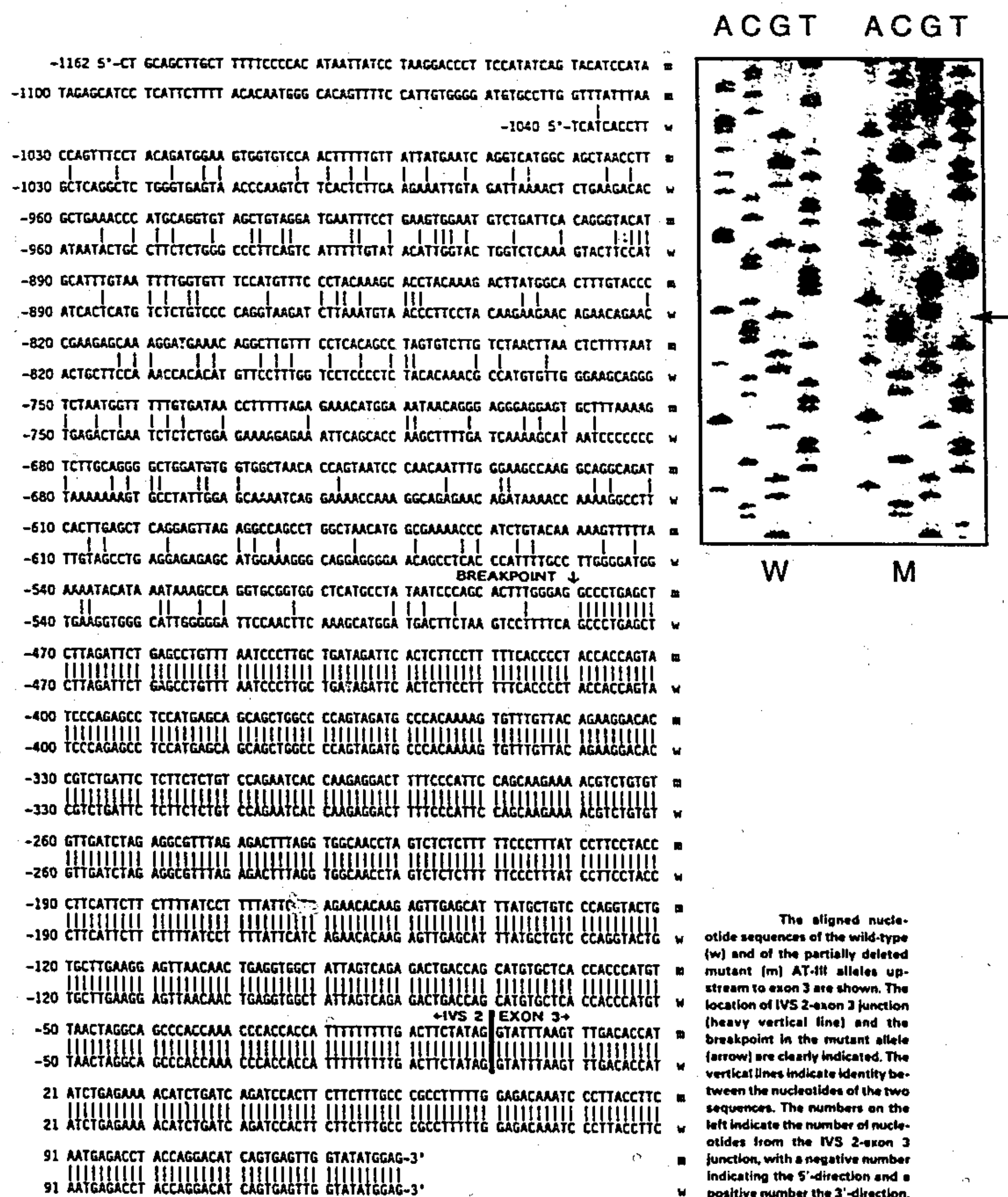


Figure III.8 :*Type I antithrombin deficient family*. Nucleotide sequences of the wild-type (w) and mutant (m) alleles from exon 3 and upstream (left panel) and at the breakpoint location (right panel).

4. Mapping of the 5' Antithrombin cDNA Ends

4.1 Primer Extension and S1 Nuclease Protection

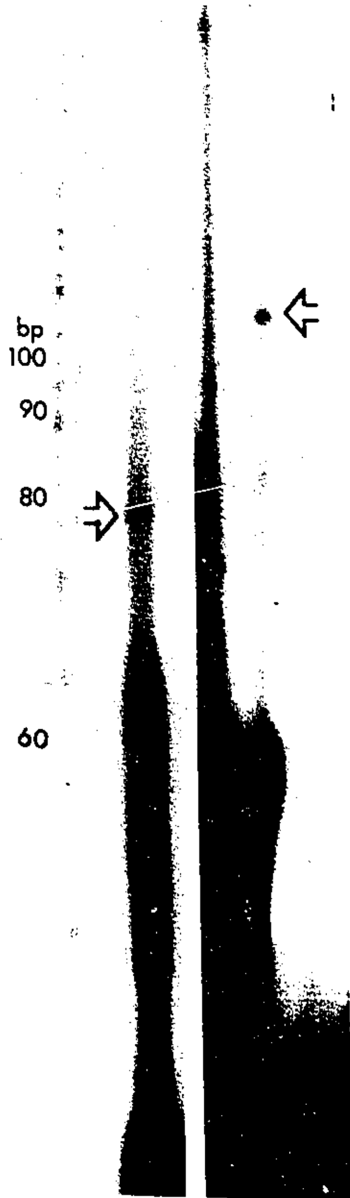
Fig.III.9 illustrates the mapping of the AT mRNA 5'ends using primer-extension and S1 nuclease-protection in human liver mRNA. Primer extensions are shown in Panel I of Fig.III.9; a major product was extended with both primers tested as illustrated in Panel 1, lanes A and B of Fig.III.9. The size of the latter product was 79 nt with the primer complementary to the first translated AT sequences and 105-110 nt with the primer complementary to the 3' boundary of exon 1. No other major bands were clearly seen upstream of the extended product even after long exposure at -70°C.

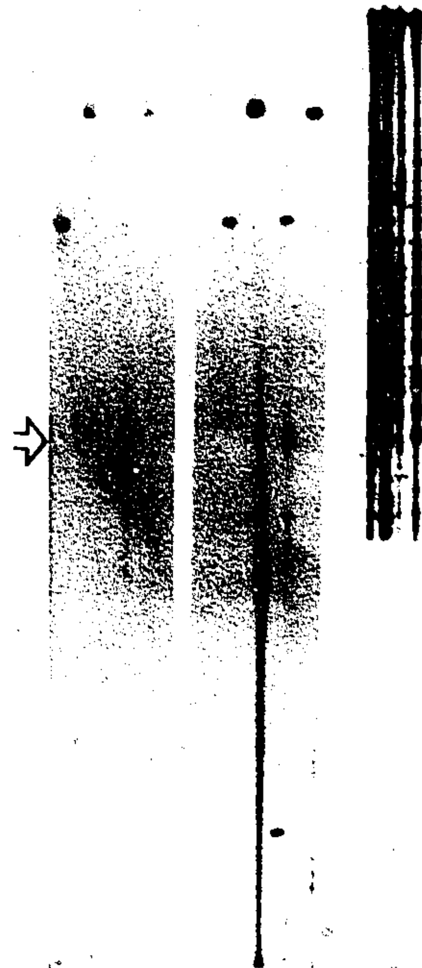
The results of S1 nuclease protection confirmed the results of primer extension. Effectively, a single 108 nt product was protected from digestion, following incubation with S1 nuclease for both 30 and 60 min (Fig.III.9, panel II). The 270 nt upstream of this product were digested. This placed the start site of transcription with both methods 71 nt upstream from the A of the translational start-site (first ATG in exon1) and was in agreement to within one nucleotide with the findings of Prochownik (1985,1). As expected, no product was protected when a probe was placed in IVS 1.

4.2 cDNA RACE

The determination of the start-site by rapid amplification of cDNA ends is shown in Fig.III.10. With human liver mRNA, a major product of 230 nt was seen when either of the two 5' anchoring primers were used after tailing of the cDNA (lanes 2 and 5 in

Figure III.9: ***Mapping of the antithrombin 5' cDNA ends by primer extension and S1 nuclease protection.*** **Panel I:** Primer extension. The arrow indicates the products extended. In lane A, the primer was 5'-CATTGGA ATA CATGGCCGCTAATC-3'. In lane B, the primer was 5'-CTTTT CCAGAGGTTACAGT-3'. A RNA ladder is shown to the left of lane A. **Panel II** : S1 nuclease protection. 30 and 60 min. indicate the incubation time with S1 nuclease. +/- refers to the presence/absence of S1 nuclease.The sequencing ladder is shown to the right.Lanes 1 and 3 correspond to the probe complementary to the 3' end of exon 1.Lanes 2 and 4 correspond to the probe complementary to IVS 1. The arrow indicates the product protected from digestion.


I
A
B
bp
100
90
80
60


II
60 min.
30 min.
1 2 3 4
1 2 3 4
− − + +
− − + +

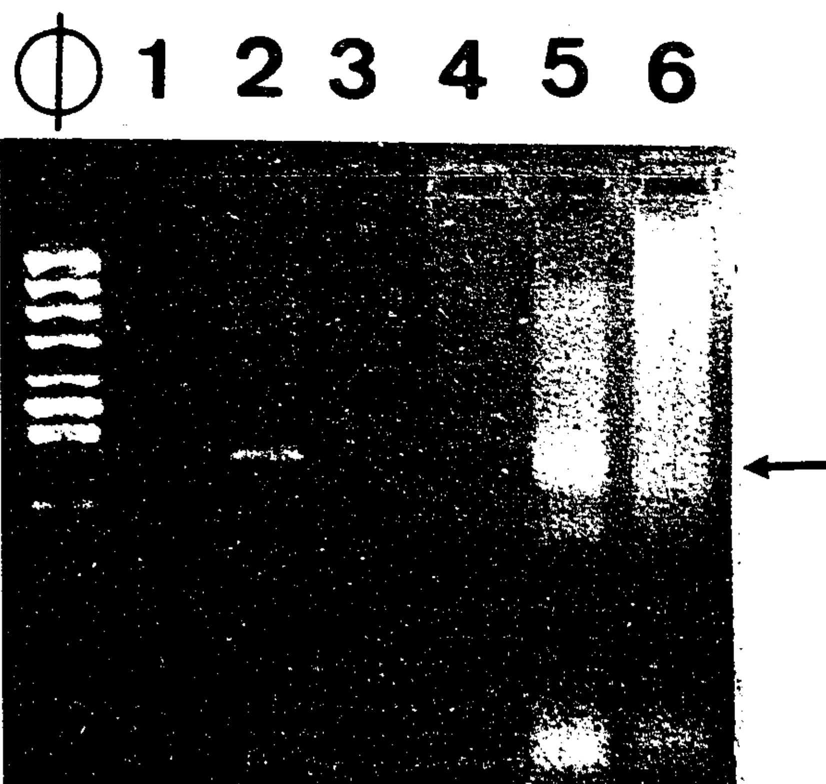


Figure III.10 : ***Mapping of the 5' Antithrombin cDNA ends by RACE.***
ϕ: Phage X174, HincII digest from: top to bottom, (bp) 1057, 770,612,495,392,335-345,291-297,210,162 and 79. Lanes 1, and 6, HepG2 mRNA; lanes 2, and 5, Human liver mRNA.
The PCR reactions used a primer complementary to exon 2, and a 5' anchoring primer, according to Shuster et al, in lanes 1-3, and to Frohman et al, in lanes 5 and 6. Lanes 3 and 4 are human liver control reactions without exon 3 specific primer (lane 3) or without anchoring primer (lane 4).

Fig.III.10). The size of the product obtained, 230 nt, was indicative only of the start site location because of tailing. The exact location of the mRNA 5' ends and the specificity of the product obtained were confirmed after amplification with nested primers for the AT cDNA region overlapping exon1-exon2 and by sequencing: two types of products of the same size were identified, one related to the AT cDNA (70% of the clones) and the other unrelated to the gene. The specific product placed the start-site of transcription in the same region as that of the two previous approaches, 2 nt upstream from the previously reported start site (Prochownik,1985). The determination of the cDNA ends by RACE was also attempted with mRNA extracted from HepG2 cells. As seen in Fig.III.10, two bands were obtained (lane 1). One had a size similar to that of liver mRNA, the second was larger, approximately twice the size of the first. This second product, once isolated, reamplified with the original primers utilized in the RACE reaction previously (the 5'anchoring primer and the primer complementary to exon 2) as a product with a size identical to that of the first product. It was therefore believed that the larger product was derived from the first and no further characterization was continued.

5. Mapping of *Cis*-Elements Flanking Exon 1 in Reporter Assays

5-1. Deletional analysis in the plasmid pSVOA-LΔ5'

Transfection experiments were carried out in a variety of mammalian transformed cells lines: HepG2 and H4IIA, two hepatoma derived cell lines; Cos1 and BSC40, two monkey kidney lines; and HeLa cells, a human epithelioid carcinoma derived cell line.

Cos1 and Hep G2 enabled the deletional analysis of the AT sequences flanking exon 1 in reporter assays. HeLa and BSC40 cells generated only 4-10% of the luciferase expression obtained with HepG2 cells. The conditions for transfection (described in the previous chapter) included the addition of a cell-shock treatment following the addition of plasmid DNA; this treatment consisted of 15% glycerol in 1XHBSS applied for 2 min at room temperature. Other alternatives, such as sodium butyrate (which increased efficiency for pSV2-LΔ5' as compared to glycerol alone) were inefficient for any AT construct tested (data not shown).

5.1.1 Elements Flanking Exon 1

Fig.III.11 illustrates the results of deletional analysis of *cis*-elements flanking exon 1 of the AT gene in the reporter plasmid pSVOA-LΔ5'. The data shown were generated following calcium phosphate-mediated transfection in the two cell-lines HepG2 and Cos1. As seen in Fig.III.11, two elements were mapped by this approach: the first element was located immediately upstream to exon 1 whereas the second was in the first intervening sequence of the AT gene.

(a) Mapping of the 5' upstream element: The larger construct assayed for reporter activity was a -6900 nt HindIII fragment encompassing exon 1 of the AT gene (HindIII -4800/+2117 nt construct in Fig.III.11). The latter fragment was unable to promote luciferase activity. The removal of exon 1 and downstream intronic sequences (HindIII/EaeI -4800/+68 nt construct) allowed detectable luciferase activity; the EaeI site was located 3 nt upstream of the translational start-site of the AT mRNA in exon 1.

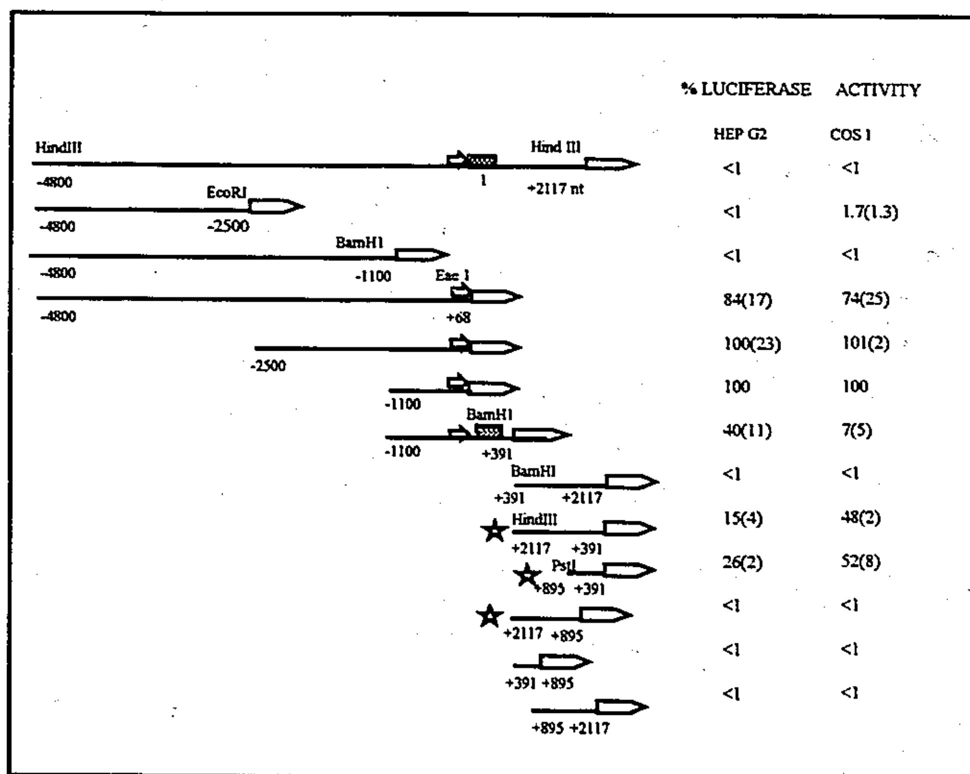


Figure III.11: ***Deletional analysis of the antithrombin sequences flanking exon 1.*** % Luciferase activity (mean +/- standard deviation ; n = 6 in duplicate). relative to the -1100 / +68 nt antithrombin region. ▭▷, luciferase cDNA; ⇨ , transcriptional start-site ; ☆ , constructs inserted in a reverse orientation ; ▩ , exon 1.

Further 5' deletions at -2500 and -1100 nt did not noticeably modify luciferase values (-EcoRI/EaeI -2500/+68 nt and BamHI/EaeI -1100/+68 nt constructs in Fig.III.11). As a result, and also to facilitate deletional analysis, an arbitrary value was given to the construct -1100/+68 nt; this construct was included in all further deletions as the 100% control. As slightly lower luciferase activity was obtained with the -4800/+68 nt fragment, the effect of sequences between -4800 and -1100 nt was tested directly in the pGL-SV40 promoter plasmid system; insertion of these AT sequences did not reduce luciferase expression driven by the SV40 minimal promoter, confirming the absence of a negative element in this region of the AT gene (data not shown).

Additional deletions directly in the exon 1 region were tested; the inclusion of exon 1 and of the first 304 nt of the first intron (BamHI -1100/+395 nt) significantly decreased luciferase activity of the -1100/+68 nt AT control sequences. This effect was seen in HepG2 (60% reduction) and was marked even by more in Cos1 cells (90% reduction). Whether the decrease in luciferase activity was due to the presence of translated sequences and/or to the first 304 nt of IVS1 was not determined.

(b) Mapping of the IVS1 element: Data provided in Fig.III.11 supports the the possibility of a second regulatory element downstream from exon 1 and within the first intervening sequence. This element was found following subcloning of the +391/+2117 nt BamHI/HindIII region of IVS1 in an inverse orientation upstream of the luciferase gene. Effectively, luciferase expression was detected only when insertion of the latter element, upstream to the luciferase gene in pSVOA-LΔ5', was in a HindIII/BamHI +2117/+391 nt orientation (constructs indicated by a star in Fig.III.11). The presence of

this second regulatory element was confirmed by further deletional analysis, narrowing the active sequences to a +895/+391 nt PstI/BamHI fragment. As expected, when the initial construct or its deletions were inserted in their natural orientation, no effect on luciferase expression was observed.

5.1.2 Fine Mapping of the 5' Minimal Promoter

Fig.III.12 details mapping of the minimal 5'upstream promoter in the same reporter plasmid: pSVOA-LΔ5'. 5' deletions of the -1100/+68 nt up to -150 nt did not noticeably alter luciferase activity in either HepG2 or Cos1. Further deletions were selected to encompass the area protected in DNA footprints (see below). Deletions to -101 nt, -70 nt, and -28 nt progressively abolished luciferase activity in HepG2 or Cos1 cells, and a -28/+68 nt NlaIV/EaeI fragment was unable to promote transcription. The same approach to map the upstream promoter was followed in the 3' direction upstream of +68 nt. Deletions up to +11 nt did not markedly altered luciferase activity. In contrast, a decrease in activity was seen upstream of -28 nt; at -28 nt, residual activities of 35% and 12% were obtained in HepG2 and Cos1 cells, respectively. Experiments for HepG2 cells were repeated with three different plasmid preparations ending at -28 nt. A construct ending at -150 nt (-1100/-150 nt) was unable to promote transcription.

The elements of the 5' upstream promoter mapped in DNA footprint analysis were studied further in reporter assays. Fig.III.13 presents data for these experiments. Single copies of the elements located at -124/-101, -98/-65, and +1/+37 nt were unable to promote luciferase expression when subcloned into pSVOA-LΔ5' in their natural

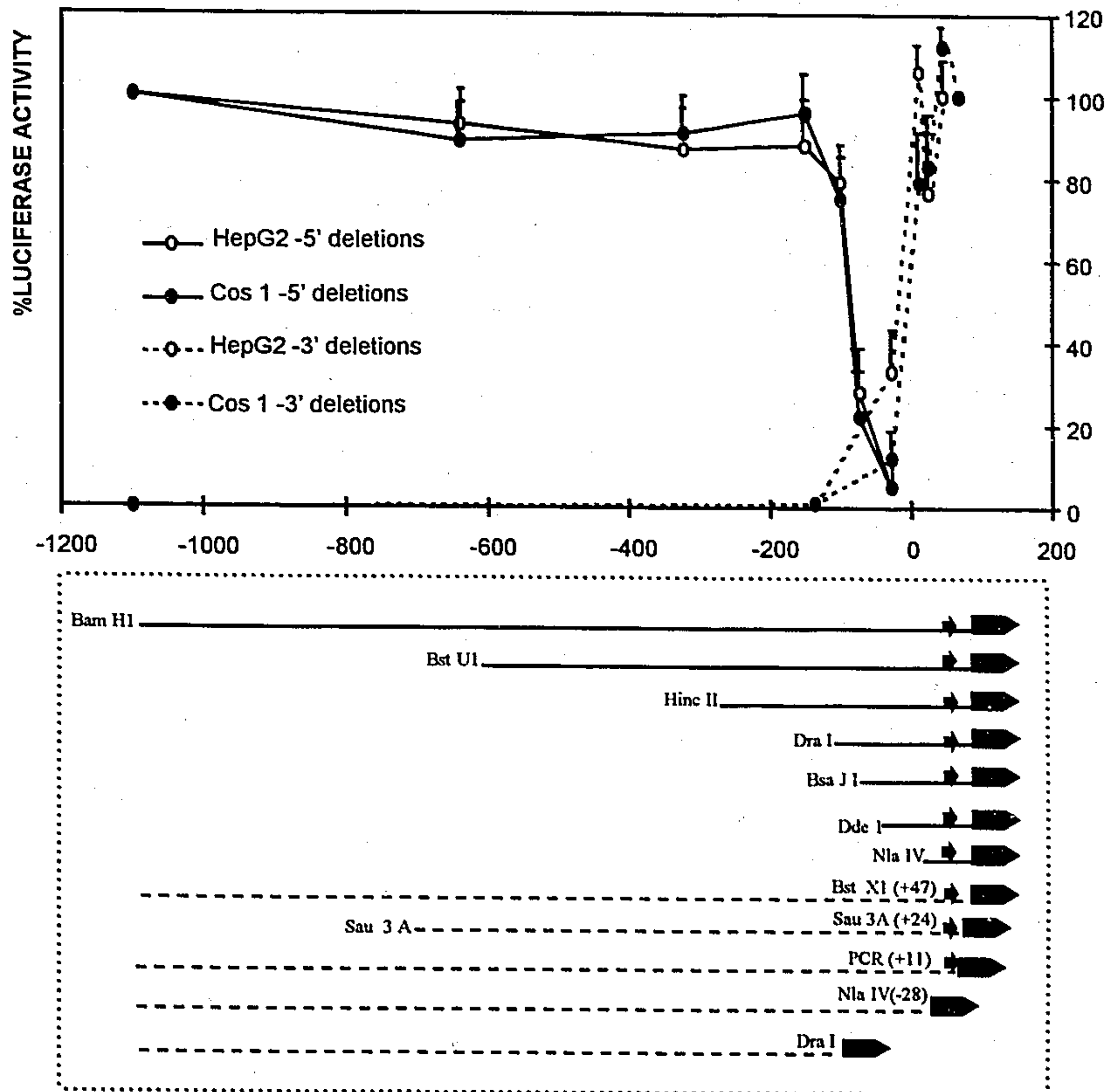


Figure III.12: ***Deletional analysis of the -1100/+68 nt antithrombin region.*** (% mean luciferase activity +/- standard deviation). The solid lines represent 5' deletions. The dashed lines represent 3' deletions. ▲, Hela cells ; ●, Cos1 cells; o, HepG2 cells.

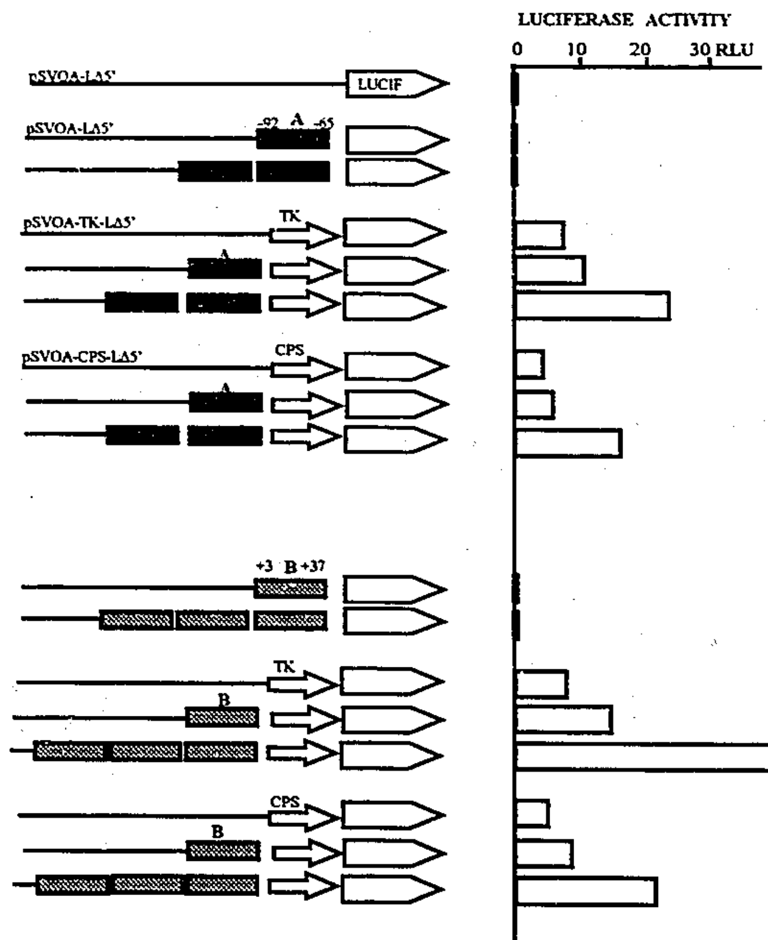


Figure III.13: ***Activity of antithrombin elements A (-92/-65 nt) and B (+1/+37 nt) on heterologous promoters.*** Values (RLU) reported are the mean percent luciferase activity in arbitrary light-units of pSVOA-LΔ5' (n=6).Black box, element A ; grey box , element B ; arrow , thymidine kinase (TK) or carbamoyl - phosphate synthetase (CPS) minimal promoter.

orientation. In contrast, two tandem-copies of the -98/-65 nt element or three copies of the +1/+37 nt element enhanced luciferase expression driven by the heterologous minimal promoters of the thymidine kinase (TK-luc) or the CPS (CPS-luc) genes.

5.1.3 Strength of the 5'Upstream Promoter

Fig.III.14 illustrates the results of transfection experiments evaluating promoter strength of the AT 5'upstream sequences. Luciferase expression driven by increasing doses of the AT -1100/+68 nt construct was compared to that of increasing doses of the SV40 early promoter pSV_2 , and of the liver-specific promoter pCPS-LΔ5'. The efficiency of the AT 5'upstream promoter in HepG2 was comparable of that of the CPS promoter and was about 10% of that of the SV40 promoter. Luciferase values were approximately 3 fold higher in Cos1 cells than in HepG2 cells for plates transfected with 20 μg of plasmid. A greater luciferase expression was obtained with the calcium phosphate protocol of transfection than with lipofection. Nevertheless, comparative analysis showed an identical trend with the latter protocol.

The IVS1 element was less efficient in promoting luciferase activity than the 5' upstream promoter; effectively, the levels of luciferase expression obtained with the IVS1 element were only 50% and 25% of that of the upstream promoter in Cos1 and HepG2 cells respectively. The lower efficiency in HepG2 cells compared to Cos1 cells was seen not only with the original IVS1 construct but also with its derived deletions. These results are illustrated in Fig.III.11.

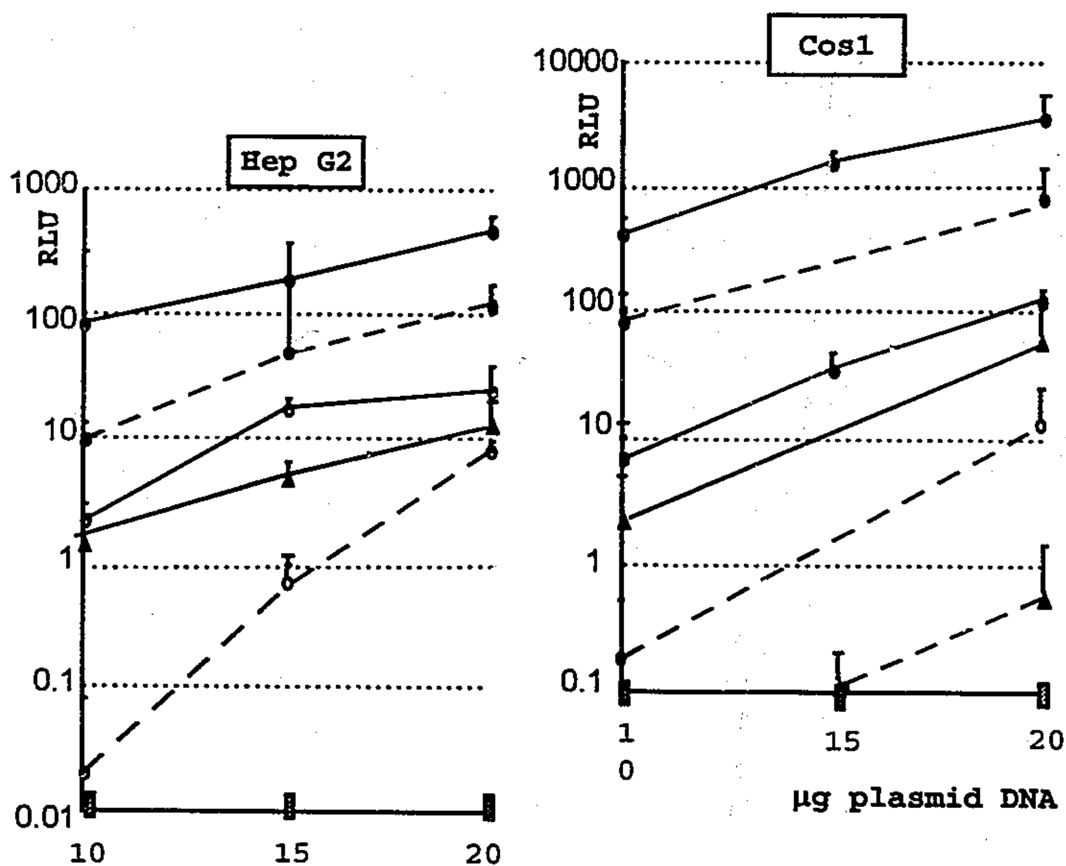


Figure III.14: ***Strength of the antithrombin 5'upstream promoter (-1100/+68 nt) in HepG2 and Cos1 cells.***

o, pAT-LΔ5'; ▲, pCPS-LΔ5' ; ●, pSV2-LΔ5'.

——— Calcium phosphate mediated transfection

- - - - - Liposomes mediated transfection

RLU, relative light units (mean +/- standard deviation).

5.2 Effects of the Antithrombin *Cis*-Elements on the SV40 Minimal Promoter and the SV40 Enhancer

To further confirm the transcriptional effects of the mapped elements, constructs including the -150/+68 nt or the +895/+391 nt regions were tested in the pGL system. These results are illustrated in Fig.III.15 and III.16 respectively, for HepG2 and Cos1 cells. Similar results were obtained for both elements: a marked increase in luciferase activity was seen when either promoter was inserted upstream of the SV40 enhancer in the regular orientation or when inserted alone in the basic reporter plasmid (an equivalent of pSVOA-LΔ5'). These results supported the hypothesis that both the AT sequences tested contained elements able to promote transcription. For the 5' upstream promoter there was a 47 fold increase in luciferase expression in HepG2 and a 53 fold increase in Cos1. For the IVS1 element, there was a 23 fold increase in HepG2 (Cos1 cells were non tested). These results confirmed that the IVS1 promoter was less active in HepG2 cells than the 5'upstream promoter.

When either AT element was placed upstream or downstream of the SV40 minimal promoter followed by the luciferase gene, there was no enhancement in luciferase activity in HepG2 cells. In contrast, in Cos1 cells, a slight increase in luciferase activity was observed for the 5' upstream promoter. This could suggest a dual role for this element, including promoter and enhancer specificities. Nevertheless, the luciferase values obtained with the SV40 promoter -minus or -plus AT sequences were not statistically different. The actual meaning of the differences observed is therefore unknown.

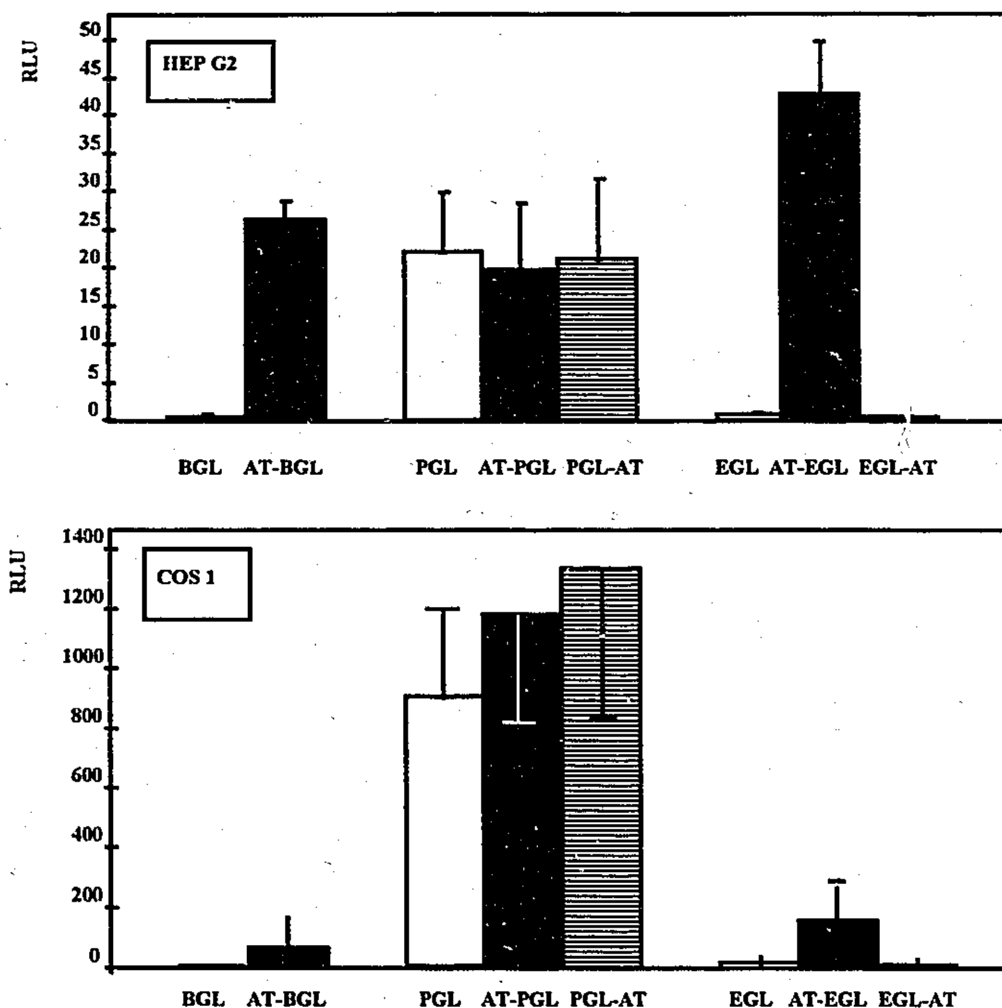


Figure III.15: ***Effects of the antithrombin 5'upstream proximal element (-150/+68 nt) on SV40 promoter or enhancer activity.*** RLU, relative light units (mean +/- standard deviation; n= 6 in duplicate). Top panel: HepG2 cells; bottom panel : CosI cells. BGL is the basal vector containing luciferase sequences only. PGL contains the SV40 minimal promoter upstream of the luciferase sequences. EGL contains the SV 40 enhancer downstream of the luciferase sequences. In AT-BGL, the -150/ +68 nt element was inserted upstream of the luciferase sequences; In AT-PGL or PGL-AT, the -150/ +68 nt element was inserted upstream or downstream of the SV40 promoter and the luciferase sequences, respectively. In AT-EGL or EGL-AT, the -150/+68 nt element was inserted upstream or downstream of the SV40 enhancer and the luciferase sequences, respectively.

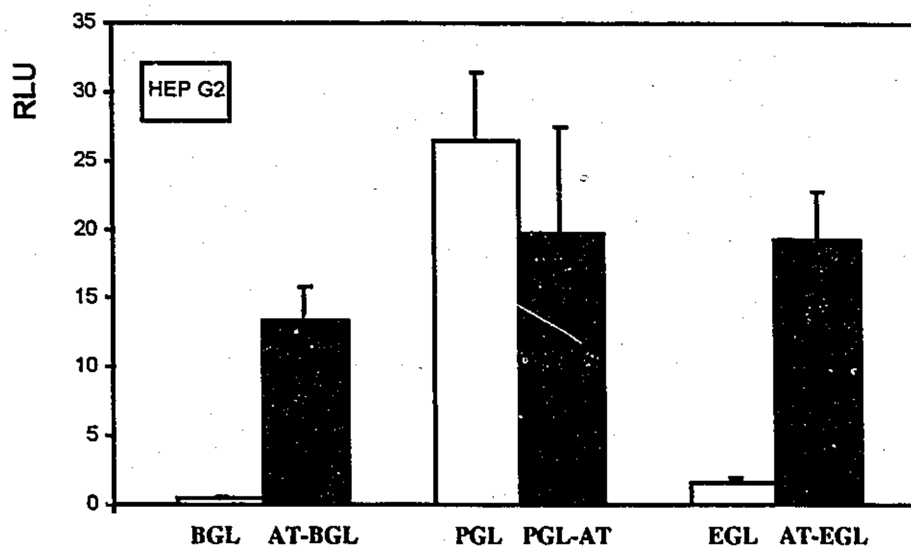


Figure III.16: ***Effects of the antithrombin +895/+391 nt IVS1 element on SV40 promoter or enhancer activity in HepG2 cells.*** BGL is the basal vector containing luciferase sequences only. PGL contains the SV40 minimal promoter upstream of the luciferase sequences. EGL contains the SV40 enhancer downstream of the luciferase sequences.In AT-BGL, the +895/+391 nt AT element was inserted upstream of the luciferase sequences; in PGL-AT, the AT element was inserted downstream of the SV 40 promoter and the luciferase sequences; in AT-EGL, the AT element was inserted upstream of the luciferase sequences and the SV40 enhancer.

6. Footprint Analysis of the 5' Upstream -150/+68 nt AT Promoter

Fig.III.17 to III.19 delineate the areas protected from DNAse I cleavage by nuclear extracts in the 5'upstream promoter. Three elements designated A, B, and C, in order of protective strentgh, were identified. Element A was strongly protected in both the upper and the lower DNA strands. This element also gave a very similar pattern of protection for the three cellular sources of nuclear extracts: HepG2, HeLa, and Cos1. It extended from -89 to -65 nt in the upper DNA strand (Fig.III.17) and from -92 to -68 nt (Fig.III.18) in the lower DNA strand.

Element B was protected in all cellular sources of nuclear extract tested, but stronger protection was observed with extracts from Cos1 and HepG2 cells. These results were apparent for the upper DNA strand when increasing amounts of DNAseI or of nuclear extracts were added; as illustrated in Fig.III.17, no protection was obtained in HeLa extracts at 12 units of DNAseI whereas protection was still visible for Cos1 and HepG2 extracts at this DNAse I concentration. Titration with increasing amounts of nuclear extracts also showed the same trend (data not shown). For the upper DNA strand also, differences in the boundaries of element B were observed as indicated in in Fig.III.17-lanes 2 (top dashed line), 8 (top solid line) and 10 (top double line, top solid line). In brief, element B extended from +3 to +37 nt in Hela cells, from -10 to +37 nt in HepG2 cells (area -10/+3 slightly protected), and from -14 to +37 nt in Cos1 cells. In contrast, the protection of area B in the lower strand by HeLa, Cos1 and HepG2 cells extracts, as detailed in Fig.III.18-lanes 13 to 15, was similar (-8/+30 nt) in the three cell lines.

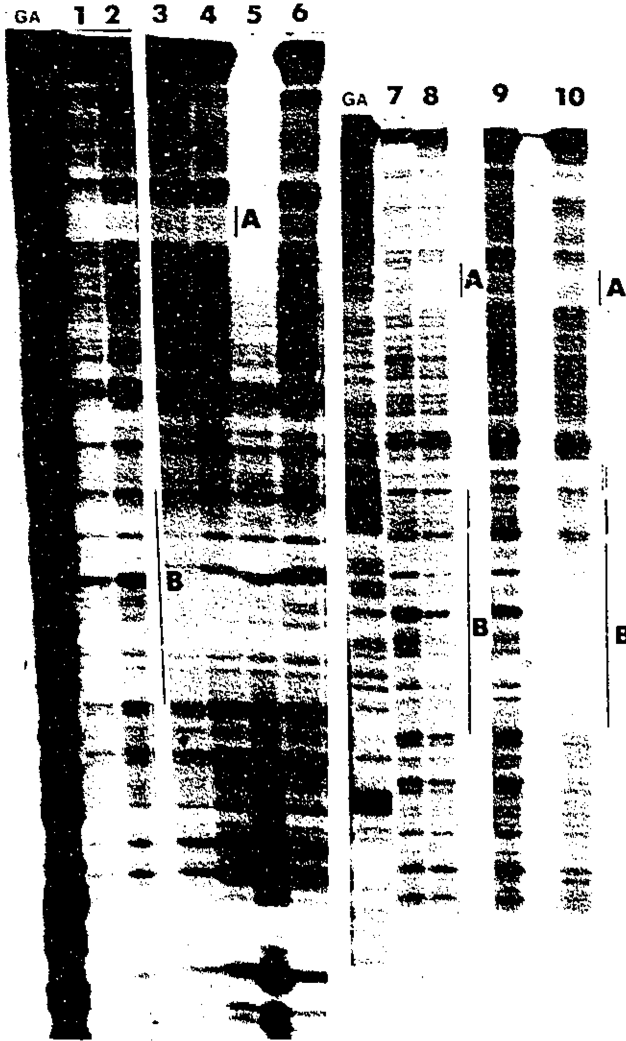


Figure III.17: ***Footprint analysis of the -150/+68 nt antithrombin upper strand.*** Lanes 1 and 2, 100 μg of HeLa extract and 8 and 12 units of DNAse I, respectively; Lanes 3, 4, 5, and 8, 100 μg of HepG2 extract and 8, 12, 15, and 5 units of DNAse I, respectively; Lanes 6 and 7, control reactions containing probe but no nuclear extract; Lane 10, 100 μg of Cos 1 extract and 8 units of DNAse I, respectively. Letters designate the location of elements A and B.

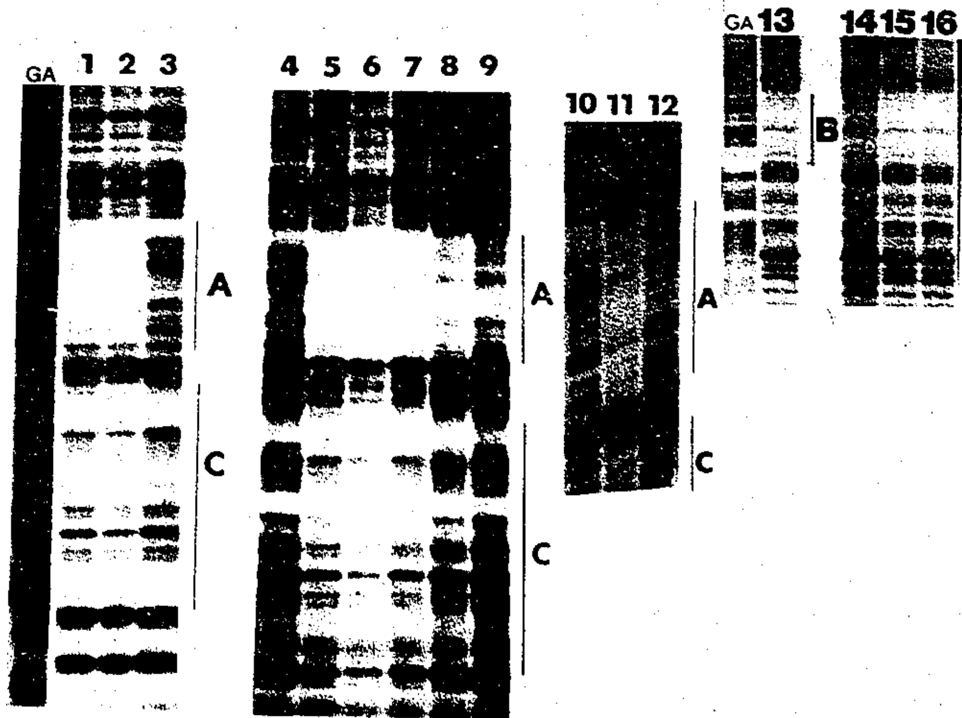


Figure III.18 : ***Footprint analysis of the -150/+68 nt antithrombin lower strand.*** Lanes 1, 2, 15, and 16, 100 μg of HepG2 extract and 8, 12, 8, and 12 Units of DNAse I, respectively; Lanes 3, 4, 10, 12, and 14, control reactions containing probe but no nuclear extract; Lanes 5 and 13, 100 μg of HeLa extract and 12 and 8 units of DNAse I, respectively; Lane 11, 100 μg of Cos 1 extract and 12 units of DNAse I; Lanes 6, 7, 8, and 9, 12 units of DNAse I and 200, 100, 50, and 10 μg of HepG2 extract, respectively. Letters designate the location of elements A, B, and C.

Element C at -124/-101 nt was seen clearly only in the lower DNA strand again with extracts from all three cell lines (Fig.III.18). Variation in the DNAse I concentrations or the amount of nuclear extract did not result in protection of the upper strand.

Fig.III.19 summarizes footprint data and details the boundaries of each element protected from DNAse I cleavage.

7. Gel Retardation Assays with Total Nuclear Extracts and the Elements -304/+68 and -150/+68 nt

Fig.III.20 presents the mobility shifts obtained with these two probes. Three complexes were retarded in extracts from HepG2 or Cos1 cells with either probe. In contrast, a different pattern was seen in extracts from HeLa cells with two areas of complexes, the upper complex being present as a doublet.

8. Gel Retardation Assays with Elements A, B, and C, and Total Nuclear Extracts

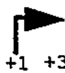
Incubation of double-stranded oligonucleotides corresponding to elements A,B, or C with nuclear extracts from HepG2, Cos1 or HeLa cells, resulted in mobility shifts (Fig.III.21). In this figure and for element A, panel A illustrates that a complex retarded as a doublet was present in extracts from either cell line. Binding was competed with an excess of unlabelled probe A but not with an excess of probe B or C. For element B, panel B illustrates the shifts obtained which, once again, were present as multiple bands in extracts from the three cell lines tested (with a slightly different pattern). These shifts were competed by an excess of unlabelled probe B but not A or C.For element C, panel C

-124
AAAACTTCTCTACTAATTAAACAACACTGGGCTC
TTTTGAAGAGATGATTAATTTGTTGTGACCCGAG

C
-101 -89
TACACTTTGCTTAACCCTGGGAACTGGTCATCAG
ATGTGAAACGAATTGGGACCCTTGACCAGTAGTC
-92

A
-65
CCTTTGACCTCAGTTCCCCCTCCTGACCAGCTCT
GGAAACTGGAGTCAAGGGGGAGGACTGGTCGAGA
-68

CTGCCCCACCCTGTCCTCTGGAACCTCTGCGAGA
GACGGGGTGGGACAGGAGACCTTGGAGACGCTCT



-14 -11 +1 +3 B
TTTAGAGGAAAGAACCAGTTTTCAGGCGGATTGC
AAATCTCCTTTCTTGGTCAAAAGTCCGCCTAACG
-8

+37
CTCAGATCACTATCTCCACTTGCCCAGCCCTGTG
GAGTCTAGTGATAGAGGTGAACGGGTCGGGACAC
+30

GAAGATTAGCGGCCATG
CTTCTAATCGCCGGTAC

Figure III.19: ***Summary of footprinting studies.*** The three protected elements A, B, and C are highlighted. The nucleotide numbering is from the presumed start site indicated by the arrow. The first ATG is boxed.

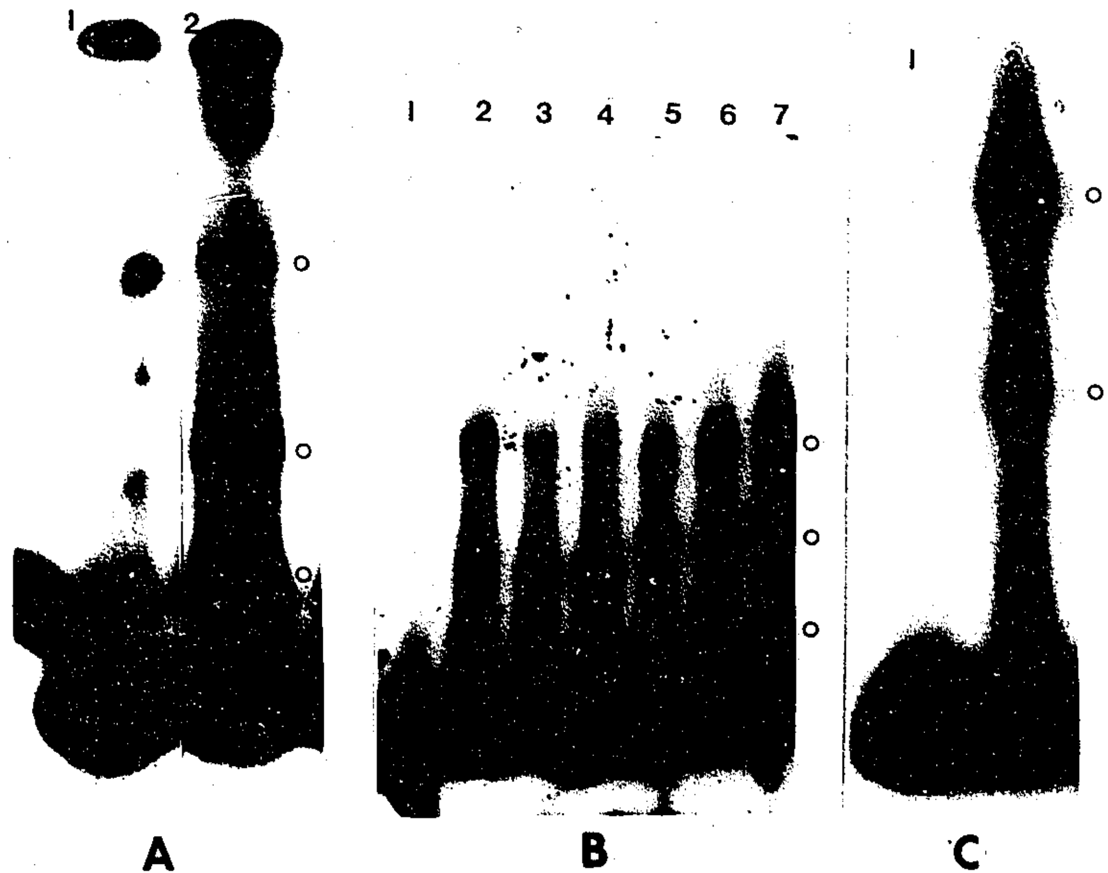


Figure III.20: ***EMSA with nuclear extracts and a -330/+68 nt or a -150/+68 nt antithrombin element.*** **Panel A** contains: in lane 1, the -330/-68 nt element alone;and in lane 2, HepG2 extracts with the same element. **Panel B** contains: in lane 1, the -150/+68 nt element alone; lanes 2-4, HepG2 extracts; Lanes 5-7, Cos1 extracts. In lanes 2,3,4 and 5,6,7; 2, 5, and 10μg poly (dI-dC) . poly (dI-dC) were used, respectively. **Panel C**: Lane 1, -150/+68 nt element alone; Lane 2, HeLa extracts. o indicates a shift.

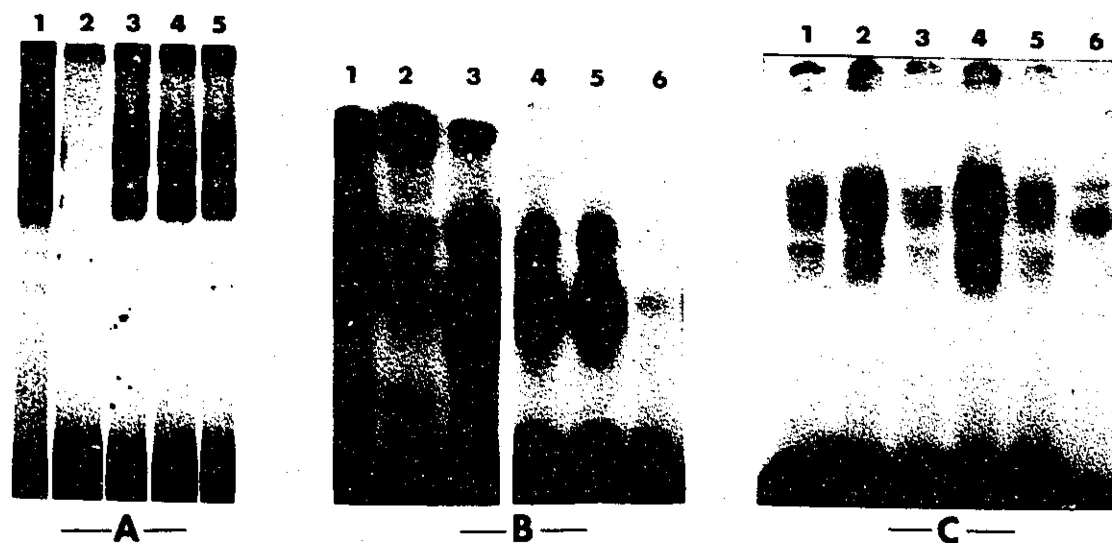


Figure III.21: ***EMSA with nuclear extracts and elements A, B, and C.*** **Panel A**: Element A with 10 μg of extracts from HepG2 cells (lanes 1-3), HeLa cells (lane 4) and Cos 1 cells (lane 5). In lanes 2 and 3, a 100-fold excess of unlabelled element A and B were added, respectively. **Panel B**: Element B with 10 μg of extracts from HeLa cells (lane 1), Cos1 (lane 2) and HepG2 cells (lanes 3-6). In lanes 4, 5, and 6, a 100-fold excess of unlabelled element A, C, and B were added, respectively.**Panel C**: Element C with 10 μg of extracts from HepG2 (lanes 1-4) and HeLa cells; Lane 6 shows element A with 10 μg of extracts from HepG2 cells, In lanes 1-3, a 100-fold excess of unlabelled element A, B, and C were added, respectively.

in Fig.III.21 illustrates the shifts obtained in extracts from the three cell lines as well as competition with unlabelled probes A, B, and C. Binding competition was observed with with an excess of unlabelled probe C and to some extent with probe A. Probe B was unable to compete. Non-specific competition with element C was difficult to block totally without altering shift formation. These preliminary data suggested that area A and C were possibly interacting with related factors and were homologous. Nevertheless, competition of probe A was not seen with cold probe C, even after several repetitions of competition experiments with these two elements. The hypothesis of competition with the added cloning sites was unlikely; these also being present in probe B, for which no competition with A or C was seen. In addition, sequence analysis for element A identified GGTCA and TGACC motifs, which evoked half-sites consensus for binding of nuclear hormone receptors, the larger family of transcription factors. The same analysis of element C gave again partial homologies with consensus to the orphan receptor HNF4, and the two liver enriched factors HNF3, and C/EBP.

9. Identification of Factors Acting in Trans with Element A (-92/-68 nt):

9.1. Competition Assays

9.1.1. Nuclear Hormone Receptors, HNF3

Fig.III.22 illustrates competition assays in extracts from HepG2 cells between element A and an excess of unlabelled oligonucleotides with consensus binding

sites for either COUP-TF1, HNF4, or HNF3. In the panel to the left, the residual shifts were obtained at respectively a 20- and a 2-fold molar excess of competitor. The graph to the right of Fig.III.22 relates the percentage of inhibition of complex formation to the molar excess of competitors. The results presented in the latter figure showed clearly that COUP-TF1 (primarily) and HNF4 (less strongly) were putative transacting candidates for element A. Unexpectively, a slight competition was also seen with a HNF3 oligonucleotide. These results are illustrated in both panels of Fig.III.22. Binding of element A was also effectively competed by the peroxisome proliferator responsive element of the rat hydratase dehydrogenase gene, HD-PPRE, an element very similar to element A (Table II.1) and known to interact with several nuclear hormone receptors under induced conditions of transfection (Zhang,1992). The results for competition with HD-PPRE are presented below in Fig.III.25. Finally, competitor oligonucleotides for general transcription factors were tested, partly to provide control of elements unrelated in sequence with element A. They were the following: AP1, AP2, AP3, SP1, CTF-NF1, TFIID, and NFκB. Their addition, as seen below in Fig.III.25, panel a, lanes 3-5 for CTF/NF1, AP2, and SP1, did not alter shift formation.

9.1.2 Pyrimidine Binding Protein, PYBP

Fig.III.23 details the binding profile of element A in the presence of PYBP oligonucleotides as competitors. As detailed in Table II.1, the site of interaction with PYBP, a single-stranded DNA binding protein interacting with the TF-DRI element of the transferrin promoter , was highly homologous with the pyrimidine-rich strand of the

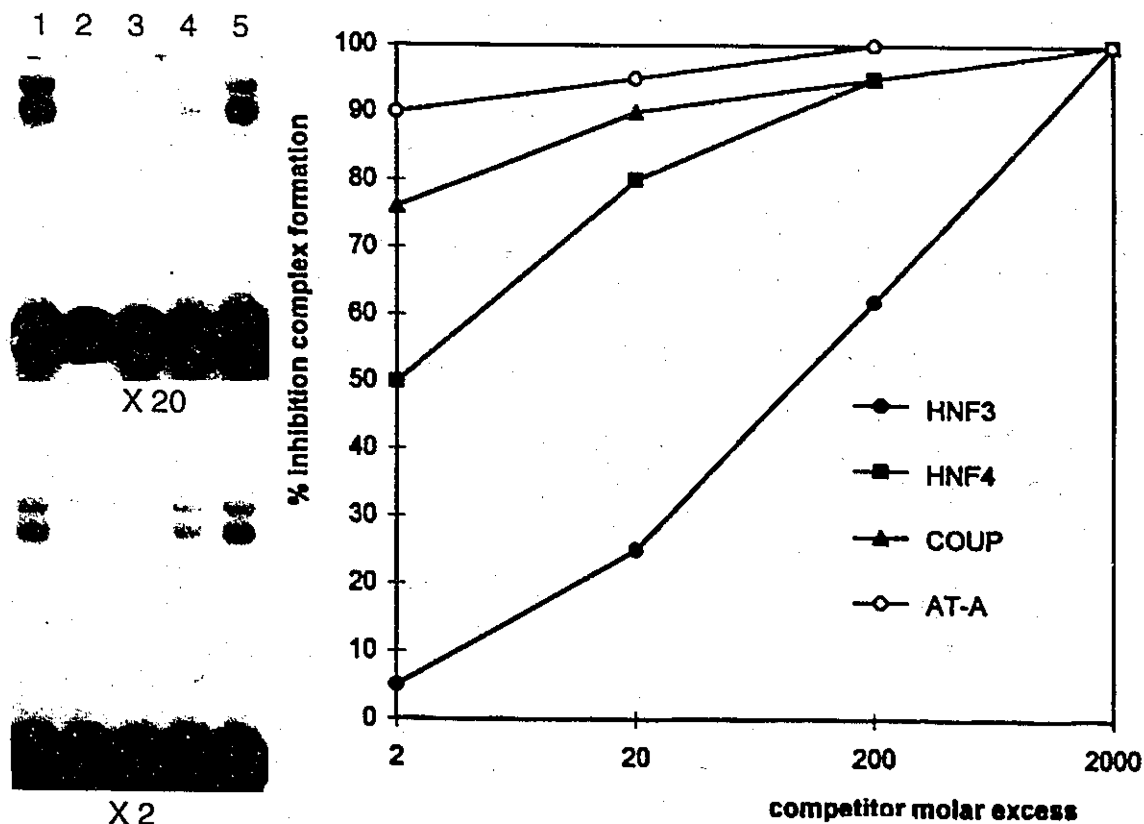


Figure III.22: ***Competition of element A with oligonucleotides for COUP, HNF4, and HNF3 in HepG2 nuclear extracts.*** The two panels to the left illustrate competition at 20- and 2-fold molar excess of competitor, respectively. Lane 1,control; lanes 2, 3, 4, and 5 contain competitor -oligonucleotides for : element A (-92/-65 nt), COUP-TF1binding-site, HNF4 binding-site and HNF3 binding site, respectively. The figure to the right shows the relation between the % inhibition of complex formation (as compared to the control) and the molar excess of competitor.

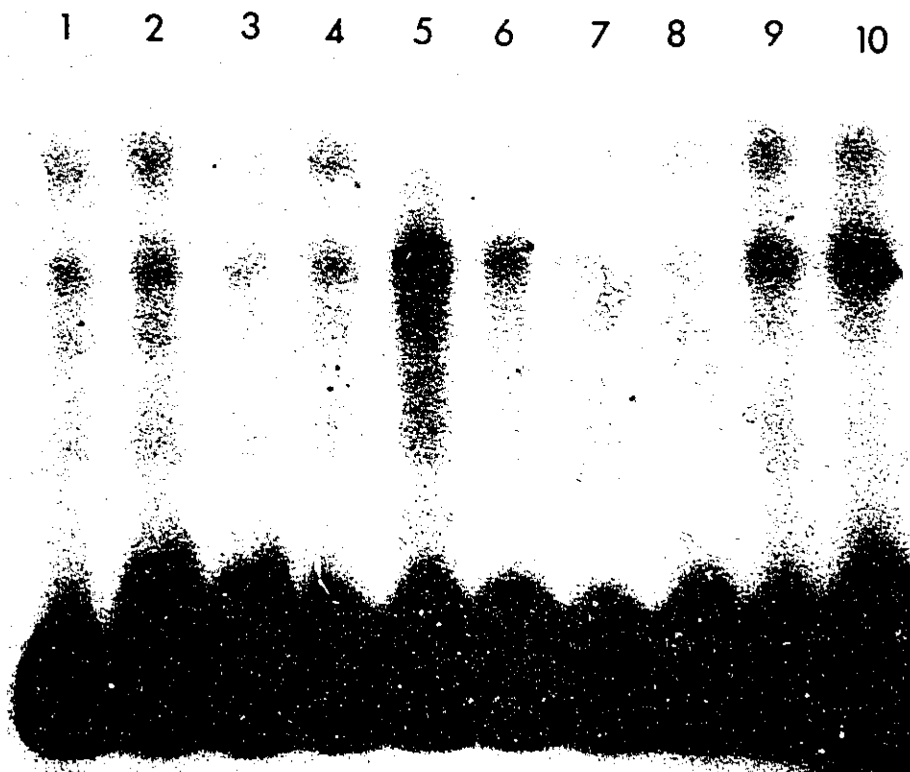


Figure III.23 : ***Competition for binding of element A with -82/-71 nt single- and double-stranded PYBP oligonucleotides.*** In lanes 1, 3, 5, 7 and 9 a 100-fold excess of competitor was used. In lanes 2, 4, 6, 8 and 10 a 10-fold excess of competitor was used. The competitor oligonucleotides were: in lanes 1,2, the purine-rich strand of the-82/-71 nt element, in lanes 3,4, the purine-rich strand of the -92/-65 nt element, in lanes 5, 6, the pyrimidine-rich strand of the -82/-71 nt element (PYBP binding site), in lanes 7, 8, the double-stranded -82/-71 nt element, and in lanes 9,10 a non specific competitor (SP1). 10 μg of nuclear extract from HepG2 cells were added in each lane.

-82/-71 nt sequence of element A (Brunel, 1989). Competition experiments were done with double- and single-stranded DNA probes, using either poly(dI-dC)-poly(dI-dC) or poly(dI)+poly(dC) as non-specific competitor respectively, and extracts from HepG2 cells. A partial competition with a double-stranded oligonucleotide encompassing only the PYBP binding site (-82/-71 nt) was seen for the two bands of the shift generated by element A (lanes 7 and 8). The addition of an unlabelled single-stranded oligonucleotide for the pyrimidine-rich strand of the -82/-71 nt element was able to abolish the formation of the upper-band of the shift. These results are shown in lanes 5 and 6. In contrast, less competition was obtained after addition of the complementary oligonucleotide (purine-rich strand of the -82/-71 nt element) as shown in lanes 1-4. Lanes 9 and 10 illustrate the absence of competition of an oligonucleotide unrelated to element A. These results strongly suggested that proteins were interacting with the pyrimidine-rich strand of element A. This hypothesis was tested directly by studying the formation of retarded complexes with labelled single-stranded oligonucleotides for element A. Fig.III.24 illustrates the shifts obtained: in nuclear extracts from HepG2 cells, both strands of element A separately radiolabelled generated shifts. Interestingly, as seen in Fig.III.24, the shifts generated were more abundant with the pyrimidine-rich strand of element A (upper-strand). As expected, shift formation with either single-stranded probe was prevented by addition of the corresponding unlabelled element (panel A, lanes 4 and 5), therefore demonstrating specificity.

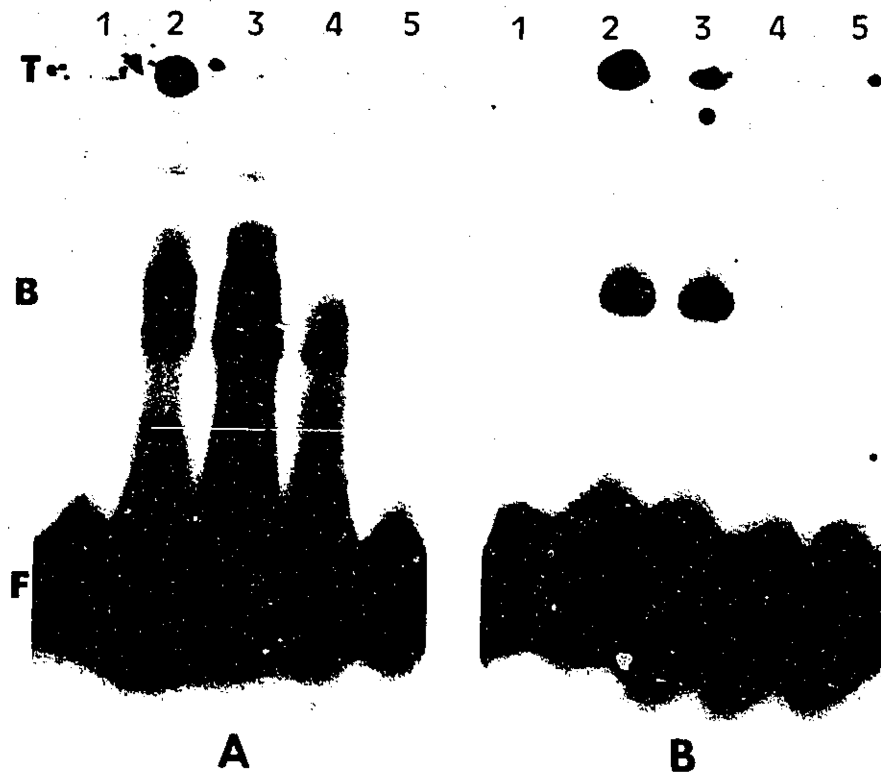


Figure III.24 : ***EMSA with nuclear extracts from HepG2 cells and single-stranded -82/-71nt PYBP elements.*** **Panel A**: Binding profile of the pyrimidine-rich strand of the -82/-71 nt element. **Panel B**: Binding profile of the purine-rich strand of the -82/-71 nt element. In panels A and B, lane1 contains the binding element alone; lanes 2 and 3 contain nuclear extracts as well as 250 μg and 2500 μg of non-specific competitor; lanes 4 and 5, identical to lanes 2,3, contain in addition a 10-fold excess of specific competitor (unlabelled single-stranded binding element: pyrimidine-rich strand or purine-rich strand in panel A and B, respectively. T, top of the gel (well); B, complexes retarded; F, free probe.

9.2 Supershifts

Supershift analysis with antibodies against the factors interacting with the competitor binding sites confirmed interaction of element A with a number of ubiquitous and liver-enriched factors. As seen in Fig.III.25, interaction with COUP-TF1 was confirmed in HepG2 (panel A), rat liver (panel B), HeLa (panel C) and Cos1 (panel D) nuclear extracts. Interestingly, less reactivity with the COUP-TF1 antibody was seen in rat liver extracts than in cellular extracts. Supershifts were also generated with an antibody against rat HNF4 in nuclear extracts of hepatic origin (Fig.III.26). In contrast, no supershifts were obtained with the latter antibody in nuclear extracts from HeLa and Cos1 cells (data for HeLa cells shown in Fig.III.26).

Antibodies for nuclear hormone receptors interacting with the HD-PPRE element were also tested for supershift formation. As shown in Fig.III.26, supershifts were seen with antibodies for RXRα and PPARα. These supershifts were obtained in extracts from the three cell-lines tested (see figure legends), but extracts from HeLa and HepG2 were the most reactive (see panels b and c). Similarly, an antibody to TRα, another ubiquitous nuclear hormone receptor shown to bind to the HD-PPRE, reacted with element A and extracts from HeLa cells (data not shown).

We wanted very much to test antibodies for PYBP with single-stranded wild-type and mutant probes, but were unable to do so, the antibody having been misplaced in the laboratory. A preliminary experiment detected, after addition of PYBP antibody, a slight supershift with a -92/-68 nt probe --denaturated with sodium hydroxide-- in extracts from HeLa cells (data not shown).

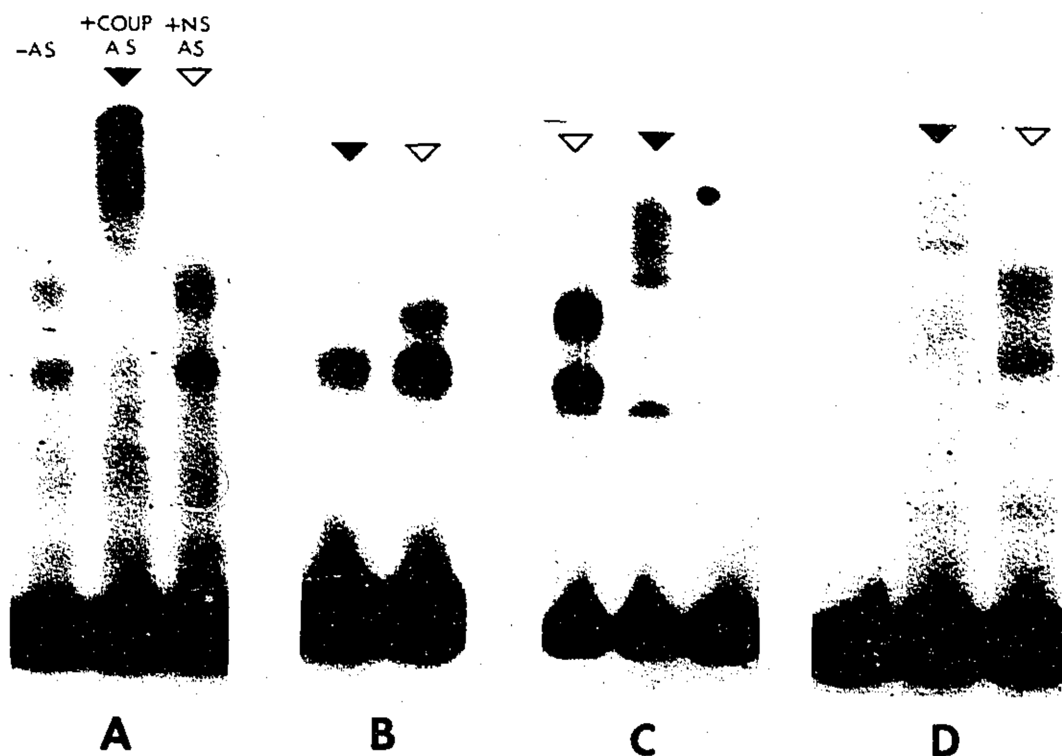


Figure III.25: ***COUP-TF1 supershifts with element A and nuclear extracts from HepG2 cells (panel A), rat liver (panel B), HeLa cells (panel C) and Cos 1 cells (panel D).*** The closed triangles indicate the presence of antibody for COUP-TF1. The open triangles indicate the presence of a non specific antibody.

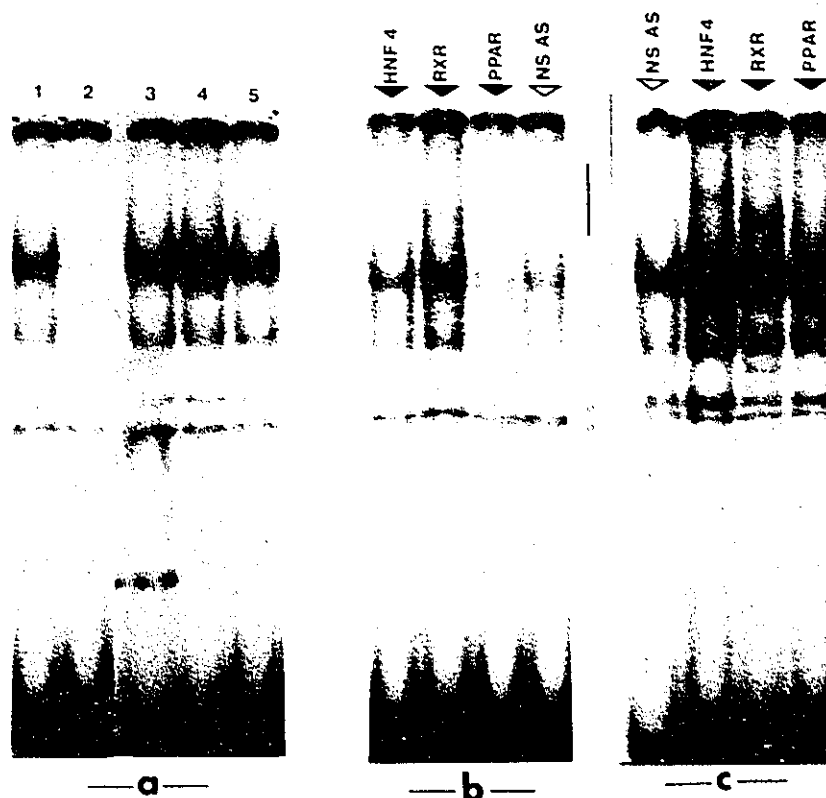


Figure III.26: ***HNF4, RXR, PPAR competition and supershifts with element A and nuclear extracts from HepG2 (panels a and c) and HeLa cells (panel b).*** **Panel a,** competition assays. Lane 1 contains no competitor; lane 2, HD-PPRE competitor; lane 3, CTF/NF1 competitor; lane 4, AP2 competitor; lane 5, SP1 competitor. **Panels b and c,** Supershifts. The closed triangles indicate the presence of antibodies to HNF4, RXRα and PPARα, respectively. The open triangles indicate the addition of a non specific antibody. The open circles indicate non specific bands.

9.3 Effects of Mutant Probes A on the Binding to Nuclear Extracts

The effects of various mutations in element A were determined in EMSA assays with nuclear extracts from HepG2 cells. Table II.1 in Chapter II details the oligonucleotide sequences of the various mutants. The mutations were targeted to the more obvious nuclear hormone receptor half-site, the TGACC motif at -75/-79 nt. In addition, the nucleotides immediately upstream of this motif -the PYBP binding site- were also modified. The choice of these mutations will be discussed later in this thesis. Fig.III.27 shows the competition experiments in the left panel as well as the binding capacity of the mutant probes in the right panel. Various degrees of partial competition were seen for all mutant probes; the bottom band of the shift remained present despite the addition of competitor (panel to the left, Fig.III.27). These results, which suggested that the mutations selected did not totally suppress binding, were confirmed by direct binding experiments in extracts from HepG2 cells (Fig.III.27, panel to the right); all mutants were able to generate partial shifts, although little shift formation was obtained with mutant 3. The effect of a deletion in element A was also tested: truncation of the -92/-88 nt GGTCA half-site in palindrome did no affect binding of element A. Effectively, total competition and an identical complex formation to the wild type probe was obtained (data not shown). These results will be discussed further below.

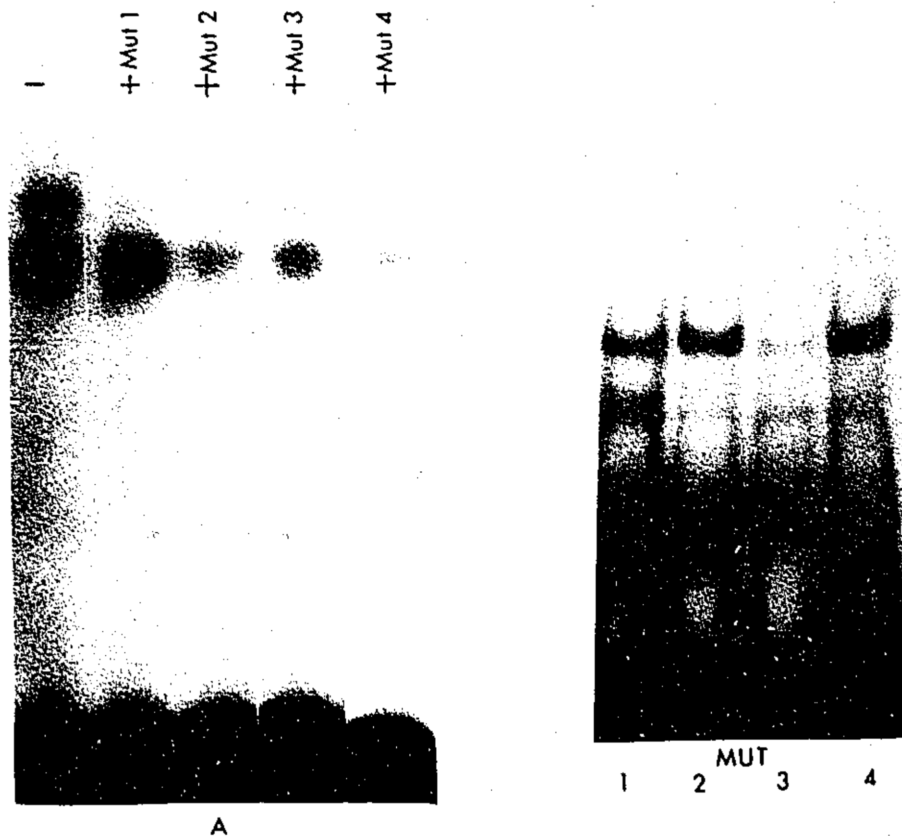


Figure III.27 : ***Competition assays with wild-type element A and mutants A 1-4 in nuclear extracts from HepG2 cells. Binding profile of mutants A 1-4 .*** In the panel to the left, the binding profile of element A (A) was evaluated. Element A was alone (lane 1) or co-incubated with unlabelled mutants A1,2,3 and 4 (lanes 2,3,4 and 5, respectively). On top of the panel Mut 1, 2, 3, 4 identify the competitor.In the panel to the right, the binding profile of mutated elements A 1-4 is shown. The mutants tested are indicated under the panel.

10.Identification of Factors Acting in Trans with the -14/+37 nt AT Element

10.1 Competition Assays and Supershifts

As seen in Fig.III.21, mobility shifts with multiple bands were generated with element B and nuclear extracts from the three cell lines tested. As the region +3/+37 nt was strongly protected in the three cell-lines, we tested this region separately. Competition experiments with a battery of oligonucleotides (in addition to the general competitors described for element A) suggested interaction with HNF3, C/EBP and CTF/NF1. An oligonucleotide binding the Epstein-Barr virus nuclear factor 1-- a protein interacting with CAAT boxes-- was also shown to strongly compete with element B (Pharmacia Inc.). This oligonucleotide, EBNA-1/NF1 was obtained from Pharmacia Inc. The results for competition are shown in Fig.III.27 for HepG2 extracts in panel a, lanes 1 to 12. Other general competitors listed above were tested and did not interfere with shift formation. The oligonucleotides shown to compete with area +1/+37 nt were also able to compete with the -28/+47 nt entire region (data not shown). Nevertheless, in the region -28/+3 nt, competitors for CCAAT sites binding factors, for C/EBP, for GA-binding factors, for TFIID, or factors of the NFκB complex were not tested.

Supershift assays confirmed interaction of C/EBPα with element B (Fig.III.27, panels B and C.The antibody was a 1:10 antiserum dilution. This supershift was seen in nuclear extracts from rat liver or HepG2 cells.In the latter case, a slight reactivity was seen after overnight incubation of the reaction mixture and the antibody before probe addition

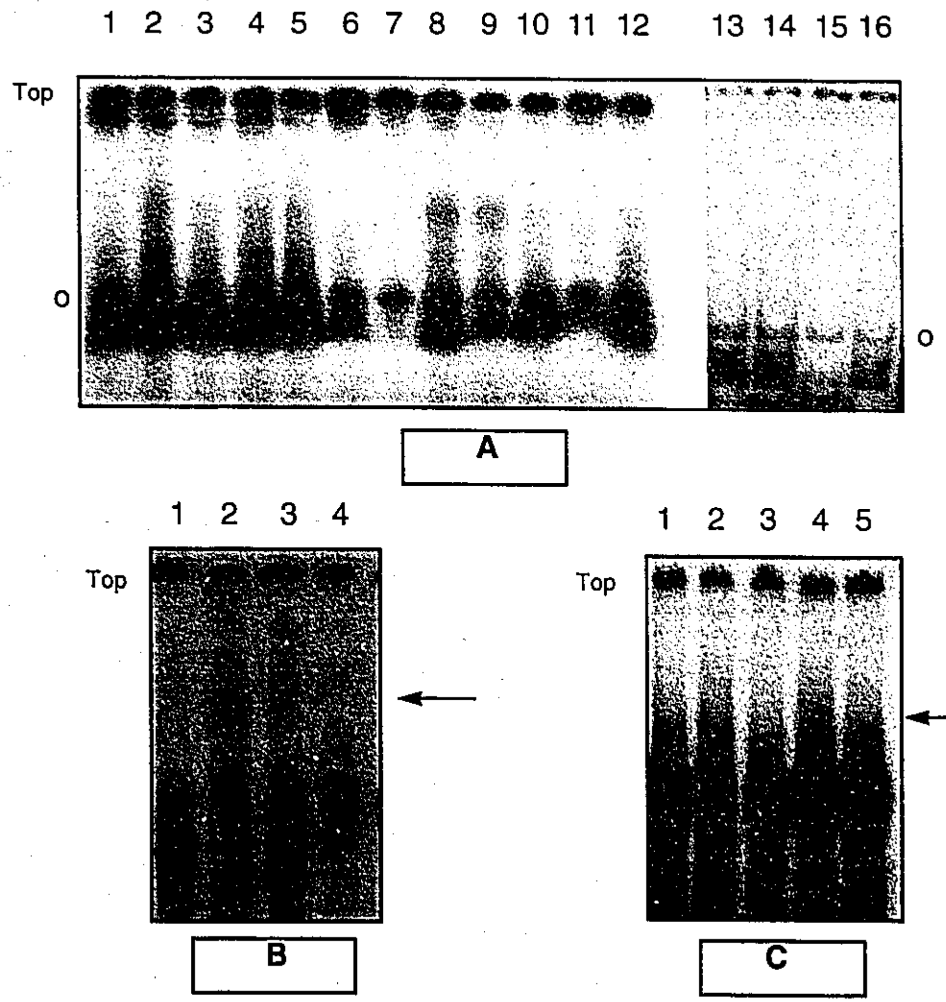


Figure III.28: ***Competition assays and supershifts with element B at +1/+37 n and nuclear extracts from HepG2 cells and rat liver*****.** **Panel A** :Competition assays in HepG2 cells. The competitors used were oligonucleotides binding to : in lane 1, SP1; in lane 2, AP1; in lane 3, AP2; in lane 4, AP3; inlane 5, HNF4; in lane 6, HNF3; in lane 7, C/EBP; in lane 8, NFKB; in lane 9, CTF/NF1; in lane 10, TFIID; in lane 11, EBNA/NF1.The competitors listed above were used at a 10-fold excess over element B. No competitor was added in lanes 12 and 13. In lanes 14 to 16, the competitors were Mutant B1,C/EBP binding site,and Mutant B2 ,respectively. **Panels B and C**: Supershifts in extracts from HepG2 cells (panel B) and rat liver (panel C). The arrow indicates the presence of a supershift. The antibodies tested were : in panel B, lane 2, C/EBP α; lane 3, HNF3 α; in panel C, lane 1, HNF3 γ; lane 2, HNF3 β; lane 3, HNF3 α, in lane 4, C/EBP. Panel B, lanes 1 and 4, and Panel C, lane 5, contain no antibody. In panel B, lane 4, a 10 fold-excess of competitor unlabelled element B was added.Panels A-C show only the material shifted by element B.The arrow indicates a super-shift; o indicates a shift non-specific.

(Fig.III.27). Antibodies for HNF3, α, β, and gamma did not generate supershifts in any cell line. No reactivity for C/EBP or HNF3 isoforms was detected in HeLa cells.

10.2 Effects of mutants B

Fig.III.28, panel A, lanes 13-16, illustrates also competition assays with C/EBP mutant probes. Lane 13 is the control wild-type probe. A first mutant in which the +4/+9 nt GTTTTCA motif was replaced by a GggggCA was unable to compete with binding of element B, as seen in lane 14. A second mutant in which the +10/+19 nt motif GGCGGATTGC was replaced by GGCggggggC only slightly competed with binding of element B (lane 16). A third mutant, AGggggCAGGCccccctCTCAGATC at +3/+27 nt, could not compete for binding to element B; in this last mutant, the first and second mutation detailed above were combined. In contrast, strong competition was seen between probe B and the consensus binding site for C/EBP (lane 15).

11. Identification of Factors Acting in Trans with Element C at -124/-101 nt

11.1 Competition Assays

Fig.III.29 illustrates competition assays and supershift assays with element C in panels A and B, respectively. Competition experiments suggested possible interaction with HNF4, HNF3, C/EBP and cross-competition with the HD-PPRE element. Nevertheless, very little competition was seen for an excess of unlabelled oligonucleotide binding site of COUP-TF1. None of the general ubiquitous factors binding sites listed with element A was able to prevent shift formation generated by element C.

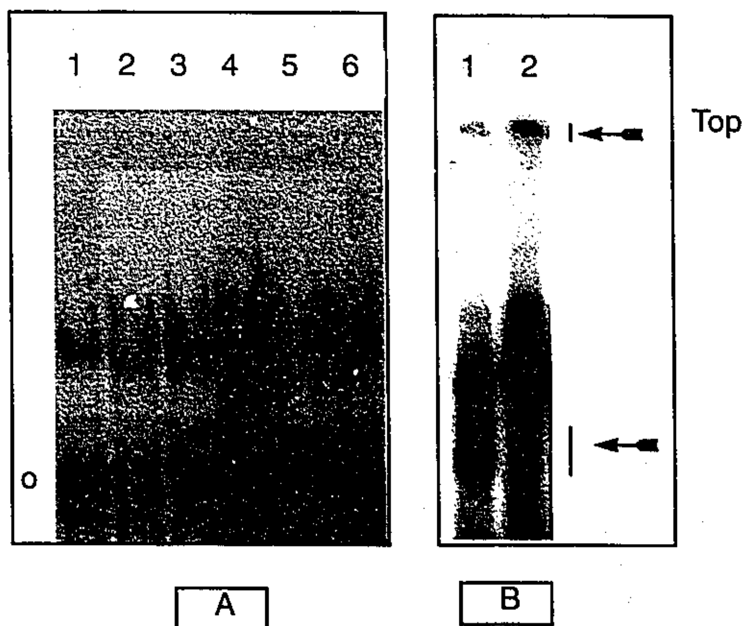


Figure III.29: ***Competition assays and supershifts with element C at -124/-101 nt and nuclear extracts from HepG2 cells.*** **Panel A** : Competition assays. The competing oligonucleotides were ; in lane 1, HD-PPRE; in lane 2, HNF4; in lane 3, C/EBP; in lane 4, COUP-TF1; in lane 5, HNF3; in lane 6, none. (o) indicates a non-specific shift. **Panel B** : Supershift analysis of element C with anti-HNF4 antibodies. Lane 1 is the control while lane 2 contains anti-HNF4. The arrow indicates the shift-species reactive with the antibody. Panels A and B show the material shifted by element C.

11.2 Supershifts Assays

These assays confirmed only partially an interaction with HNF4, as seen in Fig.III.29, panel B. The reactivity of the antibody for HNF4 was manifested by a decrease of the shifted species and an increase in labelled material in the well of the gel. The HNF4 interaction was seen only for extracts of HepG2 cells. The interaction of element C with HNF3, C/EBP, or COUP-TF1 was not confirmed by supershift assays, even after overnight incubation of the nuclear extracts with the respective antibodies. Attempts to vary quantitatively and qualitatively the amount of non specific competitors or the buffer conditions did not improve detection.

12. *In Vitro* Transcription-Translation of Individual Factors

Fig.III.30 illustrates SDS-PAGE of 2μl aliquot for the [^{35}S]-methionine labelled products obtained after translation. SP6 polymerase was used for the reactions with the vector for COUP-TF1, its truncated version Tr-COUP-TF1 and for HNF3α. T7 polymerase was used for all the other receptors, C/EBPα, and HNF3β. As seen in Fig.III.30, with the exception of TRα, Tr-COUP, and C/EBPα, which migrated in the 35-45 kDa range respectively, all the other receptors and HNF3 migrated in the 50-55 kDa range. In most cases, only a major single product for translation was seen, with the exception of C/EBP, which showed two close bands in SDS-PAGE. Alternate translation is known for the latter factor, but whether or not this explain the presence of two translated bands by this protocol is unknown, as we had only little information on the C/EBP plasmid provided. Similar reactions, done concomitantly with cold methionine, were used for EMSA assays.

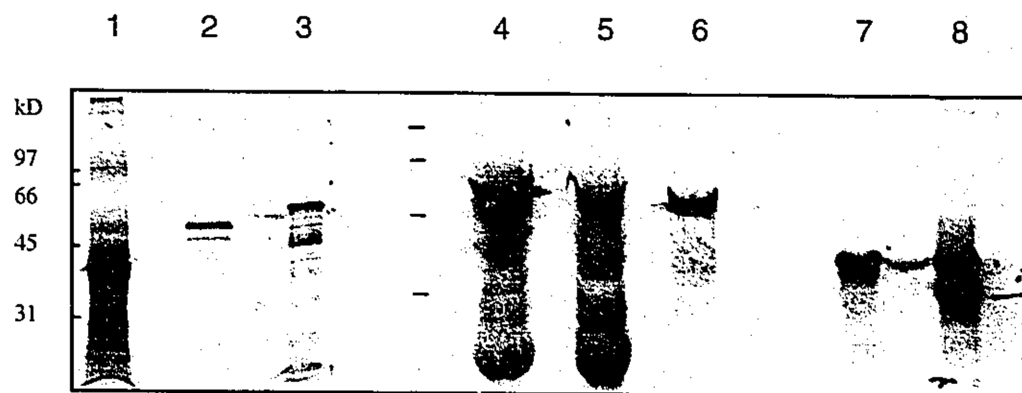


Figure III. 30 : ***SDS-PAGE of in vitro translated transcription factors*** **.** Lane 1 contains truncated COUP-TF ; lane 2 , HNF4 ; lane 3, luciferase; lane 4, PPAR; lane 5, RXR; lane 6, COUP; lane 7, HNF3α and lane 8, C/EBPα .

13. Interaction of Individual Transcription Factors with Element A

In unprogrammed lysate or in a control plasmid expressing a non-receptor (luciferase), there were no background products directly compromising the interpretation of the results. When non specific products were seen, these are indicated in the following figures by an open circle.

13.1 COUP-TF1

Fig.III.31 presents the binding profile of COUP-TF1 with wild-type and the various mutations in the TGACC half-sites of the -92/-68 nt element (A). As seen in lane 2, a strong interaction with this receptor was present. Additional bands on top the main shift were also seen in the presence of element A; whether or not these species were homodimers of COUP is unknown. As presented above in Table.II.1, the mutations selected were targeted within the two most visible COUP half-sites, i.e., the TGACCT motif at -79/-74 nt, and the motif in palindrome, GGTCA, at -92/-88 nt. The wild type sequence as seen in Table.II.1, was from -92 to -65 nt, GGTCATCA GCCTTTGACC TCAGTTCCG (the nucleotides targeted by the mutational analysis are underlined). Deletion of the -92/-88 nt element in palindrome (Tr-A) did not suppress COUP binding (Fig.III.31). A TGAgg mutant at -76/-75 nt (Mut 1) had reduced binding ability. A mutation at -83/-80 nt (destroying the PYBP binding site), ggaaTGACC (Mut 2) did not affect notably the COUP binding profile excepted for a slightly different position of the shift. A third mutation associating mutation 1 to the TGAgg transversion abolished binding. A last mutant (Mut 4), with a TcAgg modification at 79/-75 nt was also unable

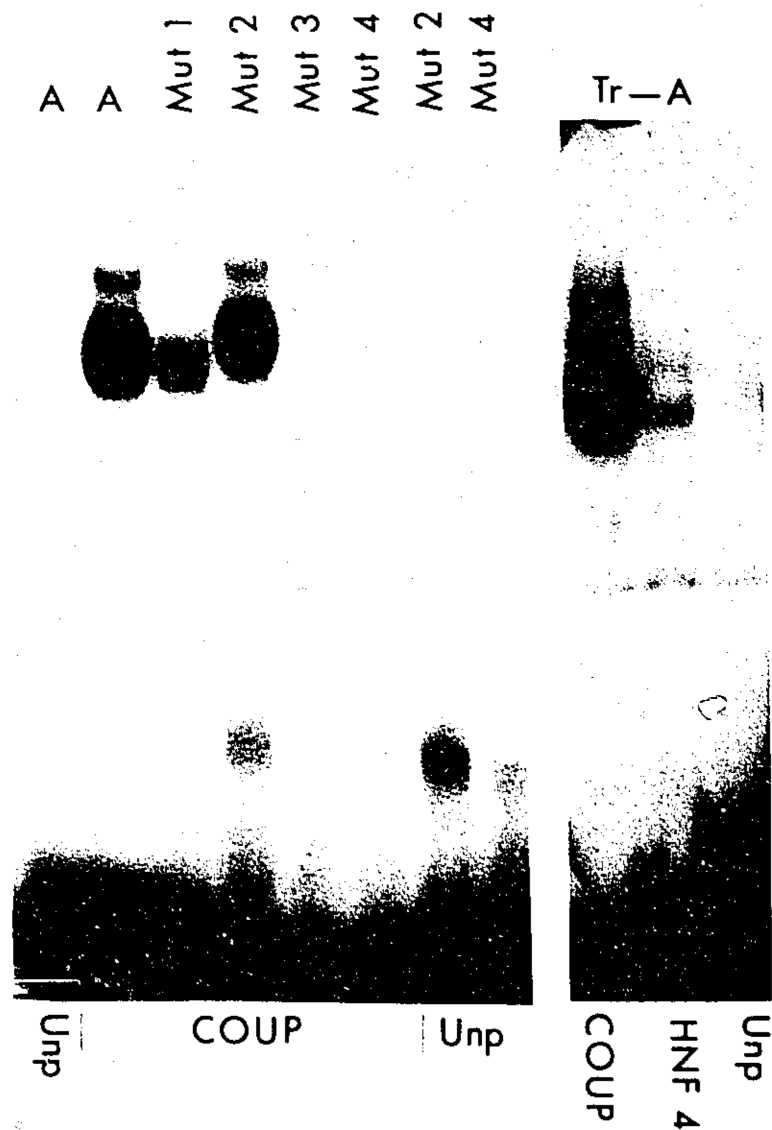


Figure III.31: ***Direct binding of element A (wild-type and mutants) with COUP-TF1.*** The labels under the figure identify the translated products; Unp, unprogrammed lysate; COUP, COUP-TF1; HNF4, HNF4. The symbols above each lane identify the binding element (A wild-type, mutants 1-4, or the truncated element A, Tr-A).

to bind. It was postulated from these results that the -79/-75 nt TGACC motif was implicated in COUP-TFI binding; the CCTTT motif immediately upstream was not as critical but influenced binding of COUP-TF1 also.

13.2 HNF4

Interaction of element A with HNF4 resulted in the formation of a mobility shift showing a single band after electrophoresis (Fig.III.32). Binding was not quantified with kinetics studies and the efficiency of transcription-translation with the expression vector for HNF4 was not determined; nevertheless, the shifts obtained were less intense than the one seen for COUP-TF1 and attempts to standardize loading based upon [^{35}S]-methionine incorporation showed again the same trend. Similar to binding of COUP, truncation of the second palindrome at -89 nt did not suppress binding to HNF4, as seen in Fig.III.31. None of the mutants tested for COUP was able to interact with HNF4 (Fig.III.32). Effectively, the slight mutation in the first half-site motif, TGAgg at -76-75 nt or the more drastic TcAgg mutation at the same location destroyed binding. Similarly, when the sequence immediately upstream was mutated (ggaaaTGACCT), no binding was observed. It was concluded that the integrity of the -79/-75 nt TGACC half-site was critical for binding to HNF4. Furthermore, and unlike the case of COUP, the CCTTT motif at -83/-79 nt is likely to be directly involved in this interaction.

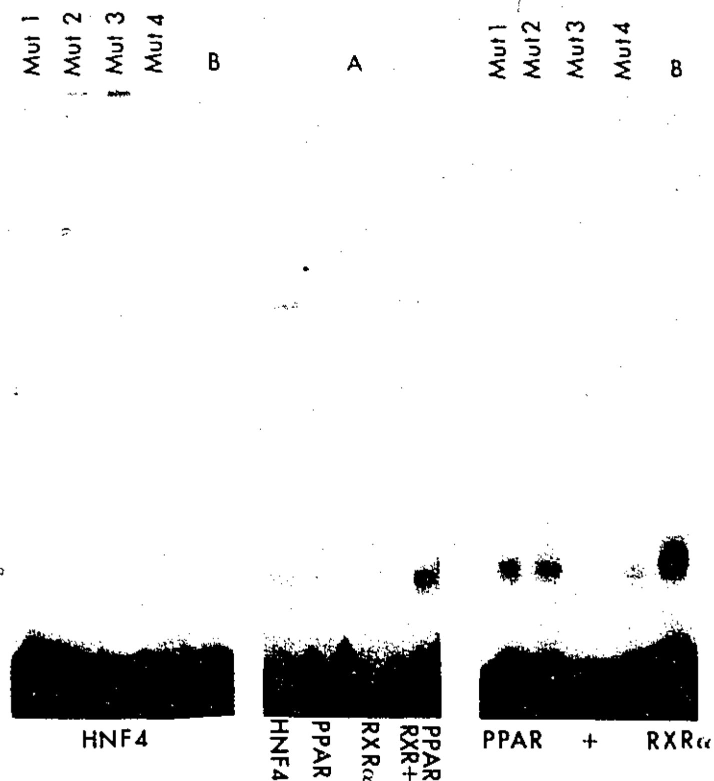


Figure III.32 : ***Direct binding of element A (wild type and mutants) with HNF4, RXR, and PPAR.*** The labels under the figure identify the translated products. The labels on top of the figure identify the binding elements (A, B, or Mut 1-4).

13.3 Other Receptors

A shift with a single band was also seen when the two nuclear receptors PPARα and RXRα were co-incubated in the EMSA reaction (Fig.III.32). When these two receptors were tested separately, no direct binding could be detected with probe A. Similarly to HNF4, no mutant tested was able to bind to the combination PPAR+RXR. Again, the truncated probe A was able to bind both receptors added together (data not shown). Finally, the thyroid-hormone receptor, TRα, was also to bind to element A directly (Fig.III.33, panel A). All the mutants tested for element A interacted with TRα although mutant 3 bound only very weakly (Fig.III.33, panel A).

13.4 Receptor Interactions

The possibility of receptor heterodimerization on element A was also approached by EMSA. Above data already showed that COUP-TF, HNF4 and TR could directly bind to element A whereas PPAR had to be complexed with RXR to be able to bind to this element. This property of the two receptors has been extensively documented, but in most cases for induced conditions of transcription (Mangelsdorf,1994).

Heterodimer formation for COUP-TF1 could be easily monitored using the truncated form of this receptor; effectively, as detailed in Chapter II, removal of the first 51 amino-acids in the native form of COUP-TF1 was shown previously to unaffect the binding properties of this receptor (Miyata,1994). Dimer formation with the truncated (Tr-COUP) and native forms of COUP-TF1 was assessed for wild-type, mutant 1, and mutant 2 element A. As illustrated in Fig.III.33, panel B, the addition of wild-type

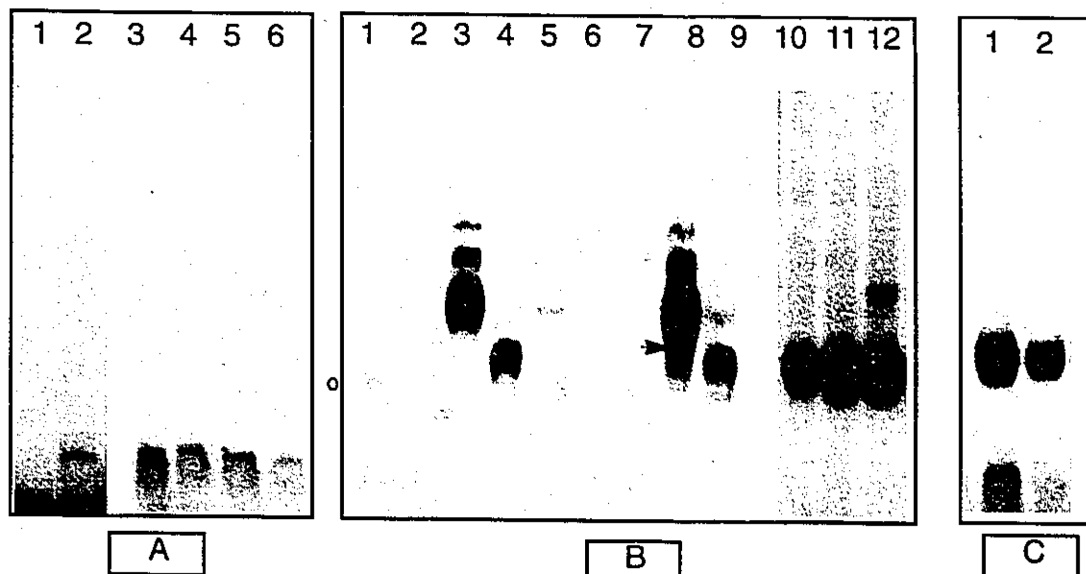


Figure III.33: ***Interaction of TR α with wild-type and mutant elements A (Panel A) ; Nuclear hormone receptor interactions on wild-type element A (Panels B and C).*** Panels A, B and C show the material shifted only . (o) indicates a non-specific shift.The arrow indicates COUP-TF1/Truncated COUP-TF1dimers. **Panel A** : Lane 1, unprogrammed lysate and wild-type A; lane 2,TRα and wild-type A; lanes 3, 4, 5 and 6, TRα and Mutants 1, 2, 3, and 4, respectively. **Panel B** : Lane 1, unprogrammed lysate; lane 2, luciferase expression plasmid ; lane 3, COUP-TF1; lane 4, truncated COUP-TF1 (Trunc); lane 5, HNF4; lane 6,RXRα; lane 7,PPARα; lane 8,Trunc-COUP-TF1+ COUP-TF; lane 9, Trunc-COUP-TF1 + HNF4; lane 10, RXR α +Trunc-COUP-TF1; lane 11, PPAR α +Trunc-COUP-TF 1; and lane 12, RXRα + PPARα +Trunc-COUP-TF1. **Panel C** : Lane 1 TRα + Trunc-COUP-TF1; lane 2, Trunc-COUP-TF1.

element A and of both forms of COUP, Tr-COUP and native COUP, resulted in the generation of a new species of shift migrating above the Tr-COUP mobility shift and below the native COUP mobility shift. The dimers are indicated by the arrow in Fig.III.33, panel B. Similar results were obtained with mutant 1 or 2 (data not shown). In contrast, no evidence of COUP-TF1 heterodimerization could be seen with any additional receptor tested. For example, Fig.III.33, panel C, shows the absence of heterodimerization between Tr-COUP and TRα. Similarly, Fig.III.33, panel B shows the absence of heterodimerization between Tr-COUP and the receptors HNF4, RXRα and PPARα.

Likewise, HNF4 did not heterodimerize with TRα (data not shown). Interaction between HNF4 and PPARα-RXRα was more difficult to study. Effectively, no truncated form of these receptors was available and the molecular weight and mobility of the full-length products of transcription were almost identical; the PPARα-RXRα complex had a mobility slightly slower than HNF4 but it was impossible to clearly identify heterodimers in the area between the mobility shifts of each receptor. In addition, PPARα and RXRα were unable to bind individually to element A. An attempt was made to use supershift assays to address this question and Fig.III.34 summarizes our findings. Addition of antibodies resulted in the disappearance of the mobility shift generated with element A; these shifts are indicated by a triangle in Fig.III.34. In several cases (but not constantly), supershifted species were also apparent in the top part of the gel; these are shown by the figure's arrows. HNF4 reactivity, for example, manifested by the presence of a double supershift illustrated by the arrow in Fig.III.34 for lanes 3, 10 and 13).

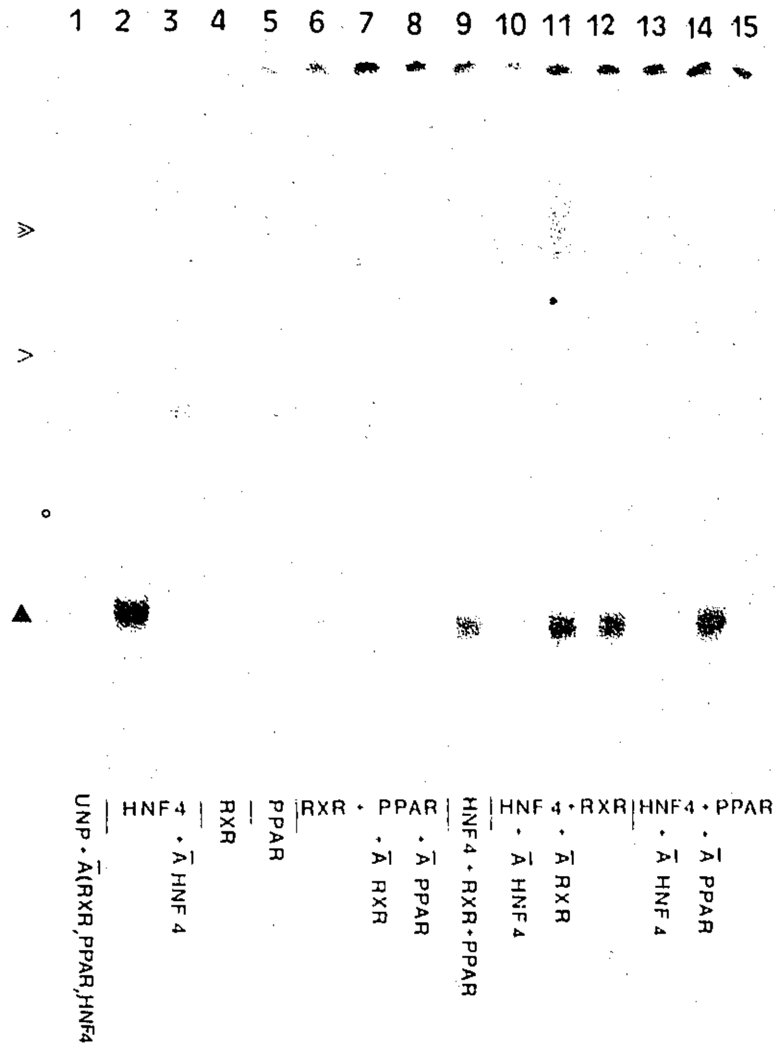


Figure III. 34: ***In vitro dimerization of HNF4, PPAR, RXR, on element A.*** The legends under the figure identify the translated products. Ā HNF4, RXR, PPAR indicates the presence of antibodies for these receptor. The closed triangle indicates the HNF4 or the PPAR+RXR shift. The single arrow indicates the supershift generated by addition of antibody for HNF4; the double arrow indicates the supershift generated by addition of antibody for RXRα. The open circle represents a shift non-specific in unprogrammed lysate (unp).

RXRα reactivity --when PPARα and RXRα were co-incubated with element A-- was shown by the presence of the upper supershift indicated by the double arrow (lanes 7, 11). PPARα supershifts -- when PPARα and RXRα were co-incubated with element A-- were not generated with *in-vitro* translated products, using either dilutions of IgG or a total antiserum for PPARα. As explained above, the antibody was reactive because its addition resulted in the disappearance of the PPAR-RXR shift (lane 8 in Fig.III.34). The mixture of HNF4 with RXRα reacted with the antibody for RXRα via formation of a supershift (lane 11) identical to that seen for PPAR+RXR (lane 8). Because RXRα was unable to bind individually with element A, it was postulated that heterodimers or at least a complex of HNF4 and RXRα could form on element A. It was also ensured that the antibodies for RXRα and for HNF4 were not cross-reacting; effectively, these two receptors have DNA binding domains very similar. Control experiments in EMSA showed that the HNF4 shift was unreactive with RXRα antibodies; similarly, the PPARα-RXRα shift was unreactive with HNF4 antibodies (data not shown).

14. Interaction of Individual Factors with Element B at +1/+37 nt

Fig.III.35 illustrates direct binding of HNF3 α and β and of C/EBPα. As seen, only the latter factor was able to bind to element B with a mobility shift with multiple bands. C/EBP homodimerization has been well documented for this factor (Lamb,1991); nevertheless, in our case, two products of translation of C/EBP were present after electrophoresis in SDS-PAGE gels (Fig.III.29). It may be likely that both products of translation could bind and be mistaken with homodimers in this case; all the shifted species

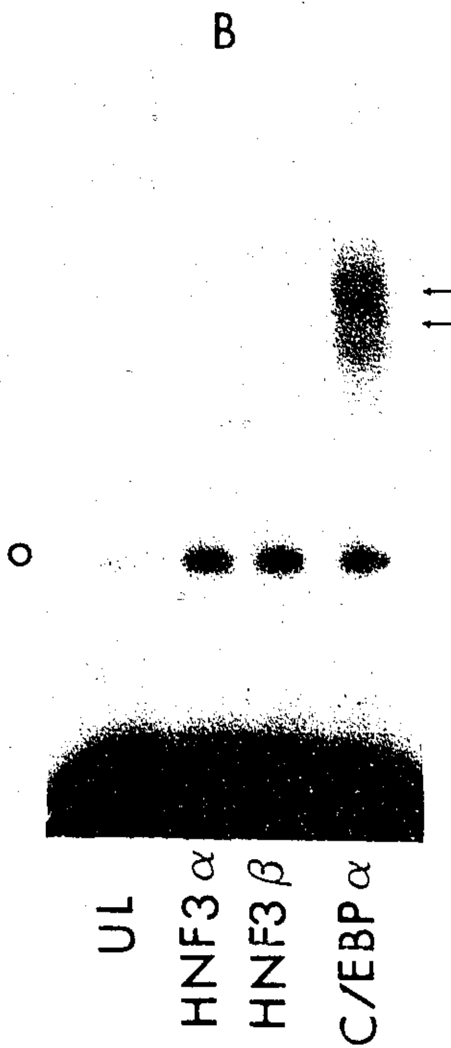


Figure III.35 : ***Direct binding of HNF3 and C/EBP with element B.***
The labels under each lane identify the translated product. UL, unprogrammed lysate.The circle indicates a non specific band. The two arrows indicate multiple bands in the shift.

were reactive with the antibody to C/EBP. Moreover, migration of translated C/EBPα with a pattern similar to the one seen herein has been reported previously (Metzger,1993). None of the two mutant elements B tested was able to bind directly *in-vitro* translated C/EBP (data not shown). Nevertheless, the binding capacity of mutants in which only a CAAT motif was modified in the whole +3/+37 nt region was not assessed.

15. Interaction of Individual Factors with Element C (-124/-101 nt)

Fig.III.36 illustrates binding of element C by various nuclear hormone receptors, HNF3, and C/EBP. Binding of element A was seen with HNF4 only. In a attempt to optimize the binding profile of element C, the amounts and nature of non specific competitors as well as the buffer conditions were modified. Nevertheless, these experimental attempts were insufficient to demonstrate a direct interaction of element C with putative receptors or transcription factors other than HNF4.

16. Effects of *in vivo* Expression of Nuclear Hormone Receptors, C/EBP, and HNF3 on the Transcriptional Efficiency of the AT 5'Upstream Promoter

The factors identified in the experiments detailed in the above sections were able to modulate the activity of the AT 5' upstream promoter, as presented in Fig.III.37 to III.39. The transcription factors were tested individually or in combination, following transfection of the three cell lines, HepG2, BSC40, and HeLa.

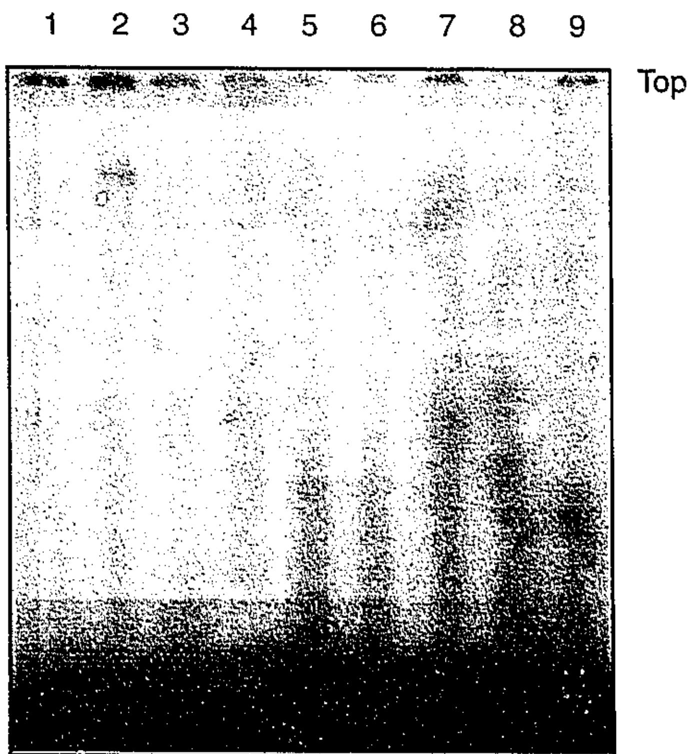


Figure III.36 : ***Interaction of individual transcription factors with element C at -124/-101 nt.*** The factors tested were : in lane 1, none (unprogrammed lysate); in lane 2, HNF4; in lane 3, COUP-TF1; in lane 4, RXR α; in lane 5, RXRα+PPARα ; in lane 6, PPARα; in lane 7, HNF3α; in lane 8, HNF3 β and in lane 9, C/EBP α.

16.1 Individual Effect

When nuclear hormone receptors were co-transfected, similar results were seen for either cell lines; this is illustrated in Fig.III.37 to III.39. HNF4 and RXRα activated the efficiency of the AT promoter; the effect of HNF4 was marked whereas RXRα gave only a moderate increase in AT reported luciferase activity. The activity of RXRα was weaker in HepG2 cells than in BSC40 or HeLa cells (Fig.III.37). The three other receptors tested, COUP-TF1, PPARα, and TRα, all had a repressive effect. In contrast to the strong repression observed with COUP-TF1 and TRα, the effect of PPARα was moderate; in fact, no effect of the latter receptor was clearly seen in HepG2 cells. The activity of the two subtypes of HNF3 was moderate and repressive. These findings were repetitively found in all cell lines tested. C/EBPα increased AT mediated luciferase activity in HepG2 and BSC40 cells (HeLa cells were non tested).

16.2 Combined Effects of Nuclear Hormone Receptors

The aim of these experiments was to determine possible modulation of the effects of HNF4, the factor with the stronger individual transactivation potential in our experiments. Fig.III.37 to III.39 illustrate that activation by HNF4 was reduced by co-expression of COUP-TR, TRα, PPARα, and PPARα+RXRα. PPARα had only a very slight negative effect on HNF4 activity. The effect of RXRα varied in the cell lines tested. Effectively, the latter receptor had no further effect on HNF4 activity in HepG2 cells (Fig.III.37); in contrast, coexpression of RXRα with HNF4 in BSC40 or HeLa cells resulted in a further increase in luciferase activity (Fig.III.38 and III.39, respectively). The increase in luciferase expression was up to 3- and 7-fold higher than the increase expected

by the sum of activities of both factors in HeLa and BSC40 cells, respectively. These results suggested synergistic effects between RXRα and HNF4. The modulation of RXR activity by other factors was not studied. The only combination tested, RXR+PPAR, was to determine a putative endogenous activation of the retinoid pathway via these receptors in HepG2 cells and gave no support for this mechanism. Fig.III.37-39 illustrate these findings for the various cell lines tested. Addition of all receptors combined resulted in a marked decrease of AT gene transcriptional efficiency.

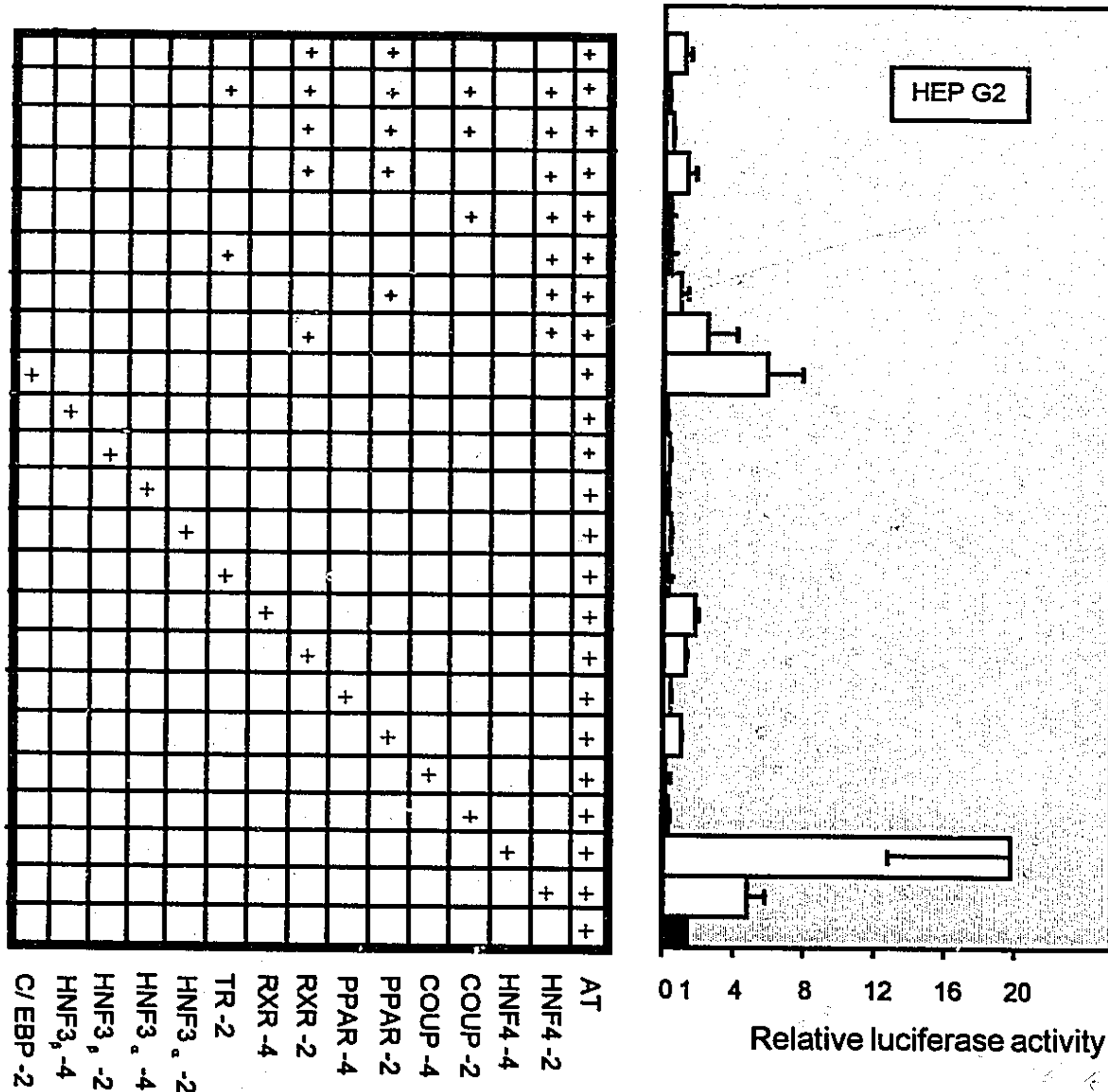


Figure III.37: ***Effects of expression of transcription factors on the activity of the antithrombin 5'upstream minimal promoter in HepG2 cells.*** The table to the left identifies the factors tested. The numbers 2 and 4 refer to the amounts of expression plasmid in µg. The illustration to the right reports the values (mean +standard deviation of 6 plates in duplicate) of luciferase activity relative to the activity of the 5' -upstream minimal promoter (black box with an arbitrary luciferase value of 1).

BSC 40

AT
HNF4 -2
HNF4 -4
COUP -2
COUP -4
PPAR -2
PPAR -4
RXR -2
RXR -4
TR -2
HNF3α-2
HNF3α-4
HNF3β-2
HNF3β-4
C/EBP -2

01 10 20 30 40 50

Relative luciferase activity

Figure III.38: ***Effects of expression of transcription factors on the activity of the antithrombin 5'upstream minimal promoter in BSC 40 cells.*** The table to the left identifies the factors tested. The numbers 2 and 4 refer to the amounts of expression plasmid in μg. The illustration to the right reports the values (mean +standard deviation of 6 plates in duplicate) of luciferase activity relative to the activity of the 5' -upstream minimal promoter (black box with an arbitrary luciferase value of 1).

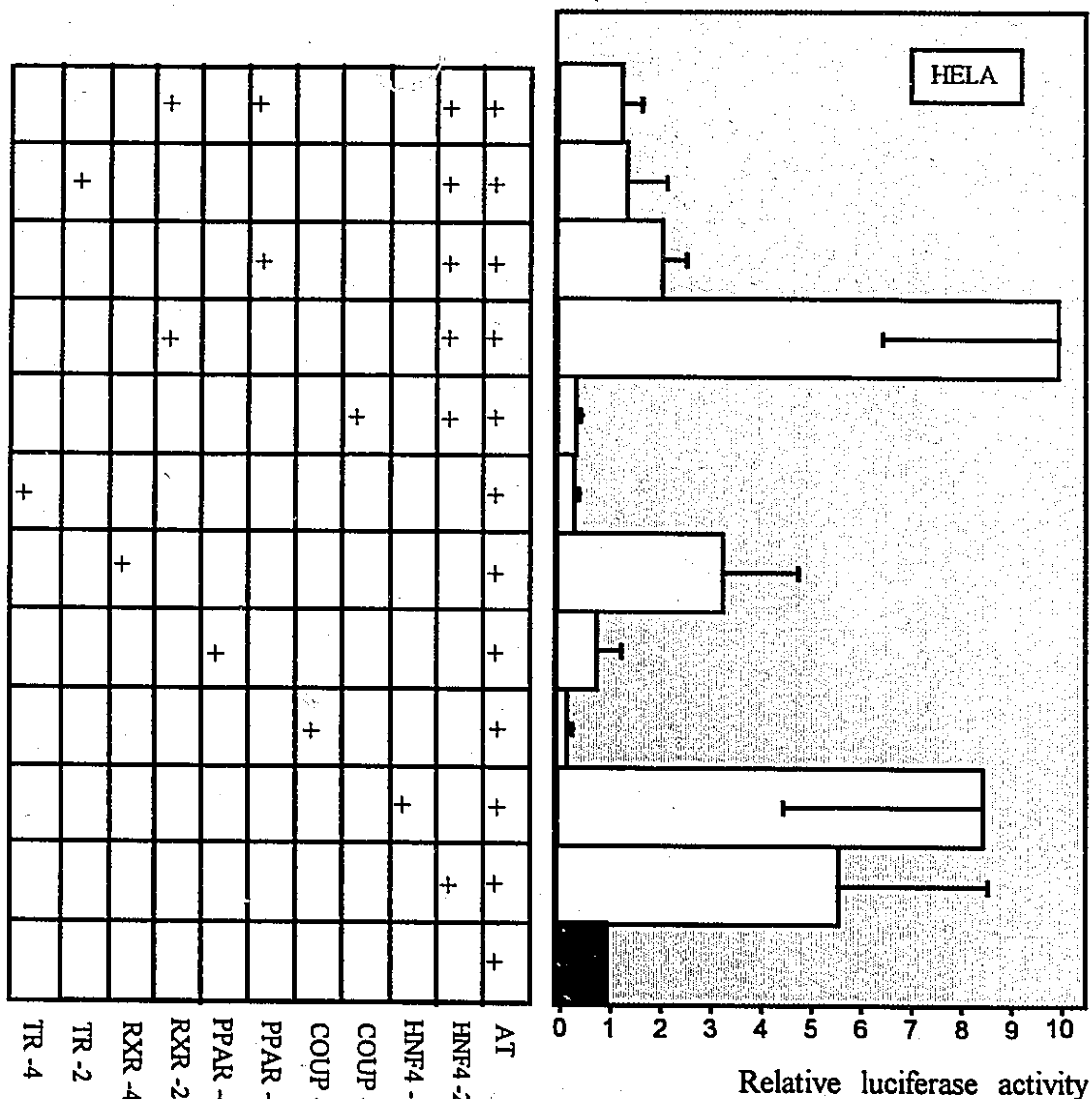


Figure III.39 : ***Effect of expression of transcription factors on the activity of the antithrombin 5'upstream minimal promoter in HeLa cells.*** The table to the left identifies the factors tested. The numbers 2 and 4 refer to the amounts of expression plasmid in µg. The illustration to the right reports the values (mean +/- standard deviation - n = 6 plates in duplicate) of luciferase activity relative to the activity of the 5'upstream promoter (black box with a luciferase value of 1).

CHAPTER IV

DISCUSSION

1 Type I Antithrombin Deficient Kindred

The results shown in Fig.III.4 to III.8 of the previous chapter suggested that a natural deletion in an AT allele could be responsible for the lack of expression of the gene. In addition, the absence of the 5' region of the AT gene in the mutant allele supported the presence of regulatory sequences for gene expression at this location. However, the lack of expression of the abnormal allele cannot be associated with a sole defect in this part of the AT gene. Effectively, only the 3' breakpoint of the deletion was determined in this study. The 5' breakpoint of the deleted sequences will have to be located by a more efficient method of mapping, for example by pulse-field electrophoresis with markers for the region of chromosome 1 neighbouring the AT locus.

Restriction analysis with a battery of probes and restriction enzymes, as well as PCR amplification of the 5'upstream region of the AT gene, located the 3'breakpoint of the deleted sequences in the 5' region of the AT gene. Enrichment in sized fractions containing the normal or the mutant allele, branch capture PCR, inverse PCR and subcloning-sequencing located this breakpoint 480 nt upstream from the third AT exon. Sequence homologies were searched at this location, as presented in Table IV.1. The database comparative analysis utilized in this table was based on the FASTN algorithm of Lipman and Pearson (Lipman,1983). The best scores found, as highlighted in the table,

sequence name	score	position	position in deleted allele	% homology
apolipoprotein C1 pseudogene	684	1716–2376	21-682	53.6
apolipoprotein CII gene	806	1633-2298	13-682	56.1
apoCIIgene for preproapolipo	822	1354-2014	21-682	56.5
cad gene 3' end	866	207–877	3-682	58.6
c-sis protooncogene intron 5	820	8-584	94-682	56.9
HPRT gene	866	2154-2826	2-682	57.4
cellular tumor antigen P53	746	6049-6738	1-682	53.6
HGM-CSF-mRNA	750	987-1658	5-682	54.9
C-FES/FPS proto-oncogene	796	9763-1043	94-682	55.8
proliferating cell nucleolar protein P120 gene	784	4467-5127	1-682	55.3
L-C acyltransferase	718	4927-5605	1-682	53.9
thymidinekinase gene5'region	666	1296-1966	16-682	52.6
γ-B-crystallin	836	20-394	2-682	55.9
ornithine decarboxylase gene	840	2191-2867	1-682	56.4
c-mos pseudogene with Alu repeats	528	3-237	438-682	71.8
oncogene c-mos(locusGP5,NV1)	528	3-237	438-682	71.8
C1-inhibitor	768	2320-2973	8-682	55.3
G6PD gene	770	720-1392	2-682	55.3
plasminogen activator-inhibitor gene	704	6941-7621	4-682	53.5
insulin receptor gene	842	3801-4463	4-682	56.7
gastric H,K-ATPase catalytic subunit gene	752	906-1563	2-682	55.2

Table IV.1: ***Homologies between the deleted antithrombin allele upstream to the 5' breakpoint and the eukaryotic data bank .***

indicated 71% homology with the human oncogene and kinase *c-mos*, as well as its pseudogene. This homology was intriguing but its actual relevance for the lack of expression of the deleted AT allele is unknown; more precisely, the possibility of an insertion by transposition at this location was not evaluated. The homology found was between nt +3/+237 of *c-mos* and the first 250 nt of the mutant AT allele directly upstream of the 3' breakpoint (Olds, 1993). All additional homologies listed in Table IV.1 were in the 50-60% range.

2 Transcriptional Start Site and General Control Sequences for Polymerase II Mediated Transcription in the Antithrombin Gene

2.1 Transcriptional Start Site

The location of the transcriptional start site was reassessed before embarking into transfection assays; effectively, the two previous determinations reported in the literature are in disagreement (Bock,1983; Prochownik,1985). Three approaches were followed for this purpose, including primer extension, S1-nuclease protection and RACE. As shown in Fig.III.9 and III.10, one major start site of transcription was detected using the three above approaches; the location of this start site was in agreement by one to two nucleotides with Prochownik's report (1985). Regular transcriptase (Super-Script, GIBCO/BRL) was utilized for primer extension in all cases; no attempts were made to utilize Taq polymerase, an enzyme supposedly less susceptible to pausing and to DNA bending during reverse transcription (Tse,1990). Consequently, whether or not the location of the 5'AT

mRNA ends found is the true transcriptional start site of the AT mRNA remains to be confirmed by other groups. In this regard, new data on general features for initiating elements will be further discussed below. As reviewed in chapter 1, multiple and / or alternative start-sites are often reported for liver expressed serpins and coagulation-factor genes. In addition, transcription from alternative promoters, which is also known to occur frequently for this gene family, has been related to tissue specificity, cytokine induction or developmental switches (Perlino,1987; Schibler,1987;Hafeez,1992).

2.2 Eukaryotic Control Sequences for General Initiating Elements of Polymerase II Mediated Transcription

The results of database searches for TATA boxes, GC-rich regions, CAP sites, and coding frames are shown in Table IV.2. This table is informative for the regions of the two promoters found, i.e., (1) the BamH1/BamH1 region flanking exon 1 and (2) the BamH1/Pst1 region of IVS1 (searched in a reverse orientation). Whether the sequences presented in Table IV.2 are meaningful for transcription initiation is unknown.

In the 5' upstream region, the closest TATA box to the initiation codon was found 155 nt upstream from the presumed start site of transcription. No consensus for general initiation elements was found immediately upstream from the 5'mRNA ends. As seen on the table, the other motifs for putative TATA boxes were found further upstream. In most cases, CAP site(s) motifs were also present 25 to 30 nt further dow155nstream from the TATA elements but this was not always the rule (see table IV.2). In one case, at -612 nt, an open reading-frame of 246 nt was also found. Nevertheless, this open reading-frame

5' UPSTREAM REGION	Position	Score	Cap Site
TAATAAAAATAACG	-155	-7.46	NO
ACATAAATATAATT	-229	-7.326	+28
CCATAAAGAAAACTA	-411	-7.08	+28 +31
CTATAAAGCTGAGGA	-753	-5.08	NO
CTTTAAAACTTCTCT	-771	-7.66	+26+38
IVS 1 REGION			
AAATTAAAACCAAGA	-229 from Pst 1	-8.05	+22+ 25

Table IV.2 : ***Signal sequences for eukaryotic promoter TATA elements in the antithrombin- gene regions flanking exon 1.***

was predicted to be non coding as determined by a PC-GENE/EUK-PROM program. Furthermore, we were unable to identify in the same region typical features significant for GC rich regions or putative CAAT boxes even if a number of CAAT motifs were present in the nucleotide sequences.

As a consequence of the probable absence of classical features for a "TATA plus" promoter in the close vicinity of AT exon 1, non-classical features for a "TATA less" promoter were also searched (Smale,1989); as a result, perfect matching with an initiator sequence, CCACCC, was found 43 nt upstream from the presumed 5' mRNA ends of AT (Weis,1992). As well, the sequence encompassing the presumed start site of the AT gene, according to Prochownik and to this study also, ACCAGTTT at -1/+7 nt, is also homologous to the general consensus for a mammalian initiator, PyPyCANT/APyPy (Weis,1992).

In addition, the region immediately upstream of the presumed start site, -11/+10nt, showed homology with a GAGA box. GAGA boxes are recently described control elements for growth-hormone induced transcription and have also been shown to be determinant for basal promoter activity in non hormonal responses (Le Cam,1994). They have been found in serpin genes, notably the rat serine protease inhibitor gene 1.2 (Spi 1.2), in a number of oncogenes promoters and introns such as *c-fos*, *c-myc*, and in the insulin promoter (Bossone, 1992; Kennedy,1992; Le Cam,1994). In fact the homologous element found in AT bears closer resemblance to the GAGA box of this latter gene. These boxes contain two half-sites, often encompassing the start-site region of a gene and do not

possess intrinsic enhancer activity. Mutations in these boxes inhibit basal promoter function (Le Cam,1994). Even if not properly considered true initiating elements, they exert a strong control on initiation (Le Cam,1994). In addition, transcriptional activators have been shown to bind specifically to these elements, for example, the 50 kDa zinc-finger protein Pur.1 (Purine rich activator) also known as MAZ.1 (Myc associated zinc finger protein) (Bossone,1992; Kennedy,1992). These boxes are also involved in the growth hormone response *via* the signal-transduction "JAK-STAT" pathway, with surface receptor dimerization, tyrosine kinase activation (JAK), phosphorylation and nuclear translocation of cytoplasmic proteins (STAT), and binding to responsive elements (or ISRE/GAS elements). In addition, it has been found in the rat Spi1.2 promoter that this response was also dependent on a C/EBP like activity overlapping with the response to GAGA box elements (Le Cam ,1994).

Motifs with consensus for eukaryotic signal sequences were also searched in the IVS1 region which generated luciferase expression (+895/+391 nt). Database analysis revealed a low score for a TATA box in the coding strand with an opposite orientation, 219 nt upstream of the Pst1 site (Fig.I.1). CAP signals were also detected 22 and 25 nt downstream of the TATA box (see Table IV.2). No other general feature for signal sequences was found. The search for an open reading frame in the lower strand, for long terminal repeat like sequences (LTR) as well as for LINE and SINE elements and Alu repeats was unsuccessful.

3 *Cis*-Acting Elements Mapped in Reporter Assays

3.1 Cell Lines Tested

The greater luciferase activity under the control of the AT 5'upstream promoter or the IVS element was generated in Cos1 cells and to a lesser degree in HepG2 cells. These results, shown in Fig.III, do not fully reflect the tissue distribution of AT expression. Effectively, Northern studies have detected mRNA AT primarily in the liver. The presence of constitutive AT expression in the kidney, the brain and at the vascular site is still debated in the literature. Whether or not the high transfection efficiency of Cos1 cells was related to the fact that this cell line is SV40 transformed is unknown; effectively, Cos1 cells express the SV40 large T antigen and allow for the replication of plasmids with a SV40 origin. Nonetheless, this is also the case for BSC40 cells, and these, even if of kidney origin also, had a very low efficiency for transfection. It is likely that the cellular transcription factor milieu differs between these cell lines. For example, Fig.III.27 in the previous chapter has given evidence for interaction of the -92/-68 nt element with nuclear extracts from Cos1 cells but not BSC40 cells. It is also known that the latter cell-line does not intrinsically express several members of the nuclear hormone receptors, such as PPAR and RXR (Miyata,1993).

3.2 5' Upstream Promoter

Data presented in Fig.III.11 allowed for the delineation of a first *cis*-acting element located directly upstream of exon 1. This element was able to promote luciferase

expression in the pSVOA-LΔ5' basic reporter assay system. The AT mediation of luciferase expression was confirmed by primer extension from a primer specific of the luciferase sequences in mRNA samples extracted from transfected HepG2 cells according to the manufacturer's instructions (In Vitrogen) (data not shown). In reporter assays, luciferase activity was obtained by deletion of the sequences located in and downstream of the translated sequences of AT in exon 1. Effectively, the presence of exon 1 and of the first 304 nt of IVS 1 in the constructs greatly decreased reporter activity. The interference of translated AT sequences with luciferase activity, a frequent finding, was not directly evaluated. Moreover, the actual role of putative regulatory elements in exon 1 and/or in the next 304 nt of the beginning of IVS 1 was not studied. It could be also possible as recently reported for the factor IX gene that this effect was due to splicing sequences (Kurachi,1995).

Deletional analysis pinpointed only one type of element located directly in the transcriptional start site region (-150/+68 nt) as presented in Fig.III.11. The region between -4800 and -1100 nt was also tested in reporter assays to ensure the absence of a silencer element at this location. Therefore, the AT proximal element showed the classic features of a liver specific promoter, including mainly components clustered at the transcriptional start site. Three enhancer elements at -124/-101 nt, -92/-68 nt, and +1/+37 nt were identified in this promoter in footprint and reporter assays (Fig.III.12 and Fig.III.13). The efficiency of the AT 5' upstream promoter was comparable to that of another liver specific promoter from the carbamoyl-phosphate synthetase gene (Howell,1989; Lagace,1992). These results were confirmed in HepG2 and Cos1 cells

(Fig.III.14). The activity of this *cis*-element was confirmed in the composite plasmid system containing either minimal promoter or enhancer SV40 derived elements (Fig.III.15 and III.16). These results differ from an initial 1985 report by Prochownik; in this report, using Cos1 cells, Alexander cells, and CAT assays in a SV40 reporter derived system, enhancer activity was shown in an AT region encompassing the -304 nt region upstream to exon 1, exon 1, and IVS1 sequences up to +395 nt (see Prochownik,1985,2; and Fig.I.1 for a map of this region); this activity manifested in Cos 1 cells by a 3 fold-increase in SV40 directed CAT expression. In Prochownik's study, the location of the active sequences was not determined, and a similar construct was not tested in our study in the composite system; these results are therefore difficult to discuss. The only similar construct tested in our case was inserted in the basic plasmid system (Fig.III.11). The results presented in Fig.III.15 and III.16 showed that AT sequences located at -150/+68 nt had no effect on the activity of the SV40 promoter in HepG2; in contrast, a slight increase in luciferase activity driven by the SV40 promoter could be seen in Cos1 cells (1.2 to 1.4 fold) even if these results were not statistically significant. As reported by Prochownik, this could indicate in Cos1 cells the duality of the 5' upstream region, containing mostly elements able to not only promote transcription but also to enhance it. It could also be that the permissivity of Cos1 cells to SV40 derived expression interfered with AT reported activity. Moreover, differences between HepG2 and Cos1 cells may also reflect differences in transcription factors' content, function and transactivation potential.

During 5'deletional analysis, sequences including the region polymorphic in length 303 nt upstream of the presumed start site of transcription were deleted (map in Fig.I.1).

Whether this polymorphism is associated with differences in transcriptional efficiency or in AT, such as has been recently suggested for other serpin promoters recently, is unknown (see literature review, paragraph 4 of chapter I). Deletion of this region (S allele) in the luciferase reporter assay system did not modify promoter strength, but direct studies are warranted to address this question.

3.3 IVS 1 Element

The detection of a promoter activity in IVS 1 was unexpected. Nevertheless, the results of deletional analysis for this region of the gene indicated features for such an element (see Fig.III.11). In addition, promoter activity was confirmed when this element was tested in the composite plasmid system in HepG2 cells (Fig.III.16). This second promoter was not as powerful as the 5' upstream promoter (25%), particularly in HepG2 cells (Fig.III.11 and III.16). What was also unexpected was the inverse directionality of this element with regard to the upstream promoter. This directionality was confirmed by deletional analysis, which narrowed down the active sequences to a +895/+391 nt PstI/BamHI region. The absence of reported activity of the full length construct including both the 5'upstream and IVS1 elements (-4800/+2117 nt) could be due to the fact that both elements function in opposite directions and that promoter activity is quenched, at least in reporter assays. It could also suggest that the IVS1 element could silence expression of the upstream promoter. We showed also that the presence of exon 1 and of the next IVS1 304 nt reduced -without nevertheless abolishing --luciferase expression as seen for the -4800/+2117 nt construct. In any event, our findings to date must be considered preliminary.

The presence of regulatory elements in introns has been documented for a number of serpins as well as for liver enriched coagulation factors. In the serpin family, examples of intronic elements with a role in transcription are the rat Spi 1.2 gene (Le Cam,1994) and the rat kallikrein binding protein gene (KBP), expressed in response to growth-hormone, steroids and the thyroid-hormone (Chai,1991). The KBP gene product is absent in spontaneous hypertensive rats. The rat Serp 1.2 gene contains, in addition to a classical upstream promoter-enhancer region, an internal promoter in the first intron downstream from the initiation site, with TATA and CAAT box consensus. This promoter also has a lower efficiency than the upstream promoter (7%) and functions in the same directionality. Antisense promoters in introns have also been described in tumor suppressor genes (Malik,1995). Enhancer and repressor activities have been reported in intronic elements; the apolipoprotein genes A-II, B, and E, for example, are known to contain regulatory elements in IVS1 and 2 (Bossu,1994). Within these genes, the IVS1 of the human apo A-II gene downregulates the activity of its 5'upstream promoter (Bossu,1994). In contrast, the first intron of the rat aldolase B gene is located a strong activator of the upstream region (Gregori,1991). For the coagulation proteins, Salier et al (1990) have reported the presence of a second promoter in the factor IX gene 500 nt upstream to the proximal promoter and with an opposite directionality. This second element was mapped by deletional analysis in reporter assays and its actual role is to date unknown.

IVS elements devoid of a direct transcriptional role have also been found in the factor VIII and IX genes (Levinson,1990; Kurachi,1994,1995).Factor VIII contains a transcribed gene in intron 22; this gene is transcribed in the opposite orientation to factor

VIII (Levinson,1990). Factor IX contains an IVS1 complex regulatory element; Initially though to be an enhancer augmenting factor IX expression, this element has no direct transcriptional role. Its effect on expression is due to an increase in mRNA stability mediated by its splicing sequences (Kurachi,1995).

4 Footprints and Deletional Analysis of the Minimal 5' Upstream Promoter

The footprint analysis detailed in chapter III delineated three areas of interaction with transcriptional regulators (Fig.III.17 to III.19). A previous report by Ochoa (1989) mentions protection in a region corresponding to element A in this study (-92/-65 nt). The results presented in Fig.III.19 in part confirmed Ochoa's findings; in our case, the protection was not liver specific as present in the three sources of nuclear extracts tested, i.e., HepG2, Cos1 and HeLa cells.

The second element protected, B in this study, flanked the start-site region, at -14/+37 nt, and has not been previously reported. Once again, the weaker protection seen in this region was not liver specific. In addition, in the upper strand, the upstream boundary was at -14 nt in Cos1 extracts, -10 nt in HepG2 extracts, and +3 nt in HeLa cells. The protection of the lower-strand was not different between cell lines (-8/+37 bp protection). Whether or not the differences in the upper-strand boundary reflect the protective strentgh of the various sources of extracts or are linked to actual differences in interaction with transcription factors is unknown; effectively, shift formation for the -14/+3 nt region was not studied directly in EMSA. It was also pointed out in Fig.III.20

that the shift's pattern was different for nuclear extracts from HepG2 (or Cos1 cells) and from HeLa cells with a -150/+68 nt probe.

Finally, this study detected a last footprint immediately upstream to area A, at -124/-101 nt. Very recently, the team investigating the activity of the transgene AT 5'upstream region-apolipoprotein A-II gene (Tremp,1995) identified in extracts from mouse liver four footprints of the AT region. Two of those, named regions II and III, are identical to element C and A respectively in this study. Region II is at -112/104 nt in the upper strand and at -105/-94 nt in the lower strand. Additional footprints, not seen in our study, were also located by Tremp at -138/-123 nt (region I) and at -48/-22 nt (region IV). Region I was missed in our study due to the fact that the 5'upstream boundary of the probe (-150 nt) was placed too closely to this protected area; the three last bases of footprint I were the first readable bases of our footprints. This was insufficient to conclude that protection occurred in this region. For the same reason, Tremp et al. may have not observed protection downstream from the start site; effectively, their 3' probe's boundary was at +24 nt; this region was protected in our study. Region IV (Tremp) is at the location of a putative initiating element at -48/-22 nt. Careful reexamination of all the footprints presented in Chapter III did not indicate protection in this area of the gene with either cell line tested. Detection of area IV may be due to the stronger protective strength of total liver extracts. Differences between total extracts and transformed cell lines for protection patterns have also be hypothesized (Angrand,1992).

Deletional analysis of the 5'upstream minimal promoter was designed in view of the footprint results presented in Fig.III.17 and III.18. Progressive 5'deletions of the -150/+68 nt element showed a progressive reduction in promoter efficiency (Fig.III.12). Removal of element C (area I from Tremp) only slightly decreased promoter strength. In contrast, further deletion of element A at -70 nt resulted only in 25% promoter activity. Elements including the start-site region exclusively (-28/+68 nt) or element B alone (+1/+37 nt) were unable to generate luciferase expression. Similarly, 3'deletional analysis of the minimal promoter showed the same trend. Removal of element B sequences up to +11 nt did not significantly reduce promoter reported expression. In contrast, further deletion at -28 nt almost totally abolished luciferase expression in Cos1 cells and reduced the same expression to 30% in HepG2 cells. No additional activity was detected upstream of -150 nt. These results favor a modular nature for the AT 5'upstream proximal promoter.

The elements mapped in the footprint were classified as enhancers in reporter assays and were unable to promote transcription individually (Fig.III.13). This observation and the results for deletional analysis of the proximal promoter seen in Fig.III.12 suggest the presence of elements critical for initiation in the -101/+11 nt region. Because of the absence of a visible TATA box in this region, the presence of initiator elements in this region can be discussed in view of our deletional analysis data. It was shown in Fig.III.12 that a -70/+68 nt construct retained 20 to 30% luciferase activity of the -150/+68 nt 5' upstream promoter in HepG2 cells, whereas a -28/+68 nt construct was unable to promote luciferase expression. Residual activity of 15 to 35% was also present for a 3' deletion

ending at -28 nt. In addition, it was mentioned above that a first consensus for an initiator CCACCC was found at -43 nt. These data therefore suggest the presence of an element able to initiate transcription in a region between -70 to -28 nt. Moreover, a very recent abstract has also hypothezised the presence of such an element at this location and reports, similarly to our study, deletional analysis of the AT minimal promoter in HepG2 cells, but by CAT assays (Rosenberg,1995). This abstract mentions the importance of an area encompassing footprints A and C for promoter activity by mutational analysis. In addition, it was mentioned that a construct including only the -67/+49 nt region supported CAT expression. This favors the presence of an Inr element in this region.

In the presence of a functional initiator, it is agreed that the start site of transcription should be located at the C or A position of the Inr consensus (Weis,1992). Moreover, initiators have been shown to interact with a number of general transcription factors such as TFIID, TAF's, or the more initiator-specific factors, for example, YY1, TFII I, or Maz1 (Weis,1992; Roy 1993; Aso,1994; Martinez,1994). In our study and in agreement with a previous determination in the literature (Prochownik,1985), the 5' ends of the mRNA ends were located downstream of the putative -43 nt initiator. Protection of this putative initiator region was not seen in nuclear extracts from cultured cells either (Fig.III.19). Protection at this location was nevertheless obtained in Tremp's study even if the transacting factors interacting with this AT region were not identified (Tremp,1995). It is worth mentioning that CACC-rich motifs are also present in sterol responsive-elements, binding factors such as SREB-P1 and 2, for example. As well, as detailed above, a second consensus for an initiating element was found at the presumed start site

location. As well, the minimal construct with residual activity in Rosenberg's abstract includes not only the first putative initiator region but also sequences downstream to the presumed start site (-69/+49 nt). In addition, protection in the downstream region was observed herein in footprints (Fig.III.19) and preliminary data upon EMSA assays suggested interaction with liver-enriched and ubiquitous factors; more precisely the CCAAT motifs present at the start site location -- putative targets for factors binding to these motifs such as NF/CTF family members -- are often found in close vicinity to the sites of transcription initiation (McKnight,1989). This was emphasized by the observation that deletion of this region at -28 nt also greatly affected the efficiency of the entire -150/+ 68 nt promoter. Nevertheless, unless a experimental mistake was made, a -28/+68 nt and a +1/+37 nt constructs were unable to generate individually detectable luciferase expression. The direct participation of this region of the gene to transcription initiation therefore remains elusive. The effect of sequences located between -28 and +1 nt was nevertheless not assessed directly. Furthermore, the eventuality of a GAGA box at -14/+37 nt was also not evaluated. Whether or not two regions containing consensus for initiators exist functionally and could cooperate remains to be determined.

5 *In Vitro* Effects of Transcription Factors Interacting with the Antithrombin Minimal Promoter

EMSA assays confirmed binding between elements A, B, and C mapped in the footprints with liver-enriched and ubiquitous transcriptional regulators.

5.1 Liver Enriched Factors

5.1.1 HNF4

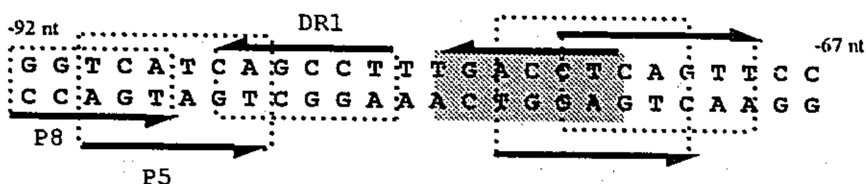
The interaction of HNF4 was observed with the two elements A and C of the upstream promoter; these results were given in Fig.III.22, III.26, and III.32 for element A, and in Fig.III.29 and III.36 for element C. The antibody directed against HNF4 reacted only in nuclear extracts from HepG2 and rat liver; no reactivity was observed in extracts from HeLa and Cos1 cells. The absence of HNF4 in Cos1 cells was unexpected; effectively, the kidney has been shown to express this factor at high rates, and the recent study by Tremp et al has furthermore shown that nuclear extracts from mouse kidney were partially supershifted with an antibody specific of HNF4 and the -89/-68 nt element of the AT promoter (Sladek,1994; Tremp,1995; Table I.1). Nevertheless, the absence of this factor in Cos1 cells has been documented previously; in particular, Mietus-Snyder et al (1992) have mentioned the absence of endogenous HNF4 in CV1 derived cells such as Cos1 or Cos7. Binding studies with translated material confirmed direct interaction with HNF4. Mutational analysis of element A strongly suggested that the -75/-83 nt region was involved in binding. The general consensus of HNF4 obtained from compilation of numerous HNF4 binding sites by Sladek, T/g/c/G/aA/gCCT/cT/cT/cG/aA/c/g/C/ACC/T (the bases in lower case letters are those which occur less frequently) matches 10/13 nt of a region of element A incompassing -86/-74 nt (consensus in upper case letters); this consensus is matched perfectly if the comparison between the two elements is based upon the oligonucleotides in lower case letters.

An attempt was made to identify possible interaction of HNF4 with other nuclear receptors (Fig.III.33 and III.34). No evidence for heterodimerization between HNF4 and either COUP-TF1, PPAR or TRα (data not shown for this latter receptor) could be seen. In contrast, preliminary observations presented in Chapter 3 suggested a possible interaction between HNF4 and RXRα (Fig.III.34). One explanation for our observations could have been the presence of cross-reactivity between antibodies for HNF4 and RXR; effectively, these two receptors are highly similar (Sladek,1994). Two approaches were taken to study a possible cross-reactivity; first, in EMSA assays, the supershifts obtained with HNF4 and RXR antibodies were specific to HNF4 or RXR respectively; to generate RXR shifts, a mixture of RXR and PPAR had to be used, RXR being unable to bind individually to DNA. For this reason, we ensured the absence of cross-reactivity directly through immunoprecipitation assays: *in vitro* translated and radiolabelled HNF4 and RXR were immunoprecipitated with IgG fractions against RXR. After binding to *Staphylococcus A* cells (Pansorbin, Calbiochem), the immune complexes were washed extensively, released by SDS-PAGE buffer at 95°C., and electrophoresed in 10% SDS-PAGE (Ausubel,1989). Following autoradiography, it was observed that RXR but not HNF4 was detected with the RXR antibody. Conversely, it was also shown that antibodies against HNF4 reacted only with HNF4 and not RXR (data not shown). Further experiments are needed to ensure the possibility of a RXR/HNF4 association. Effectively, HNF4 is beleived to homodimerize exclusively to DNA (Sladek,1994); the absence of heterodimerization HNF4/RXR should be ensured in solution and/or after binding to DNA, using deleted forms of these factors retaining intact binding, heterodimerization,

and activation dependent properties. Moreover, it may be that another other type of association allows for the interaction of both receptors; Forman has described very recently for example allosteric interactions between members of this family (Forman,1995).It may be also that RXR, which does not bind to element A directly, stabilizes HNF4 homodimer formation. HNF4 has already been shown to interact by an unknown mechanism with C/EBPα, a factor which is not a nuclear hormone receptor (Metzger,1993). Besides, in our preliminary studies, reticulocyte lysate was used for translation; it is possible that additional proteins provided by the lysate (co-factors for example) allowed for bridging between both receptors. In addition, another subtype of RXR (β and/or γ) could give more general proof for this mechanism, keeping in mind nevertheless that the reactivity of the various subtypes of a transcription factor is not always constant (Bach,1992; Marcus,1993; Giguere,1994). These studies are warranted because no reports of such an interaction have been found in the literature.

Binding of HNF4 with element C was suspected after cross-competition studies in total extracts with element A (Fig.III.21). Further studies with element C (competition, supershifts, and direct binding) confirmed that HNF4 bound strongly to this site. Many motifs in area C resemble TGACC, ACTGG, or bear resemblance with the HNF4 consensus. These hypothetical half-sites are illustrated in Fig.IV.1. In the binding area tested, -129/-101 nt, a first imperfect direct-repeat (DR1) was found through homology search, matching only 9/13 nt of the HNF4 consensus proposed by Sladek, and located at -104/-116 nt; this first hypothetical direct repeat was centered on a CTTTG motif. A

ELEMENT A



ELEMENT C

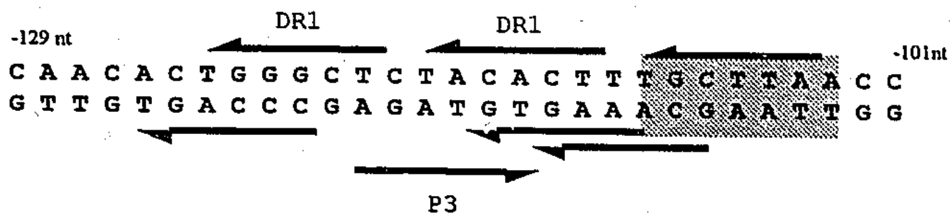


Figure IV. 1: ***Hypothetical half-site direct repeats and palindromes for elements A at -92/-67 nt and C at -124/-101 nt.***
DR, direct repeat ;P, palindrome. Numbers refer to the spacing between half-sites.The directionality of the repeats is in regard to the half-site highlighted.

second DR1 consensus with a higher score, 11/13 nt, was detected at -111/-123 nt (see figure). Mutational analysis in the three putative DR1 half-sites should determine the precise binding area. The recent abstract by Rosenberg on the deletional analysis of the AT proximal promoter mentions a putative FXP1 binding site at this location (-107/-112 nt). FXP1 was previously reported as a LFA1 binding site of the factor X promoter (Miao,1992; Rosenberg,1995). In fact, whether or not HNF4 and LFA1 are one unique factor is still debated (Ramji and Cortese,1991; Sladek,1994). The data presented in Chapter III strongly suggest that HNF4 could occupy a LFA1 binding site.

5.1.2 C/EBP

Element B was found to bind C/EBPα. Fig.III.30 and III.35 illustrated these findings. Competition assays with a C/EBP consensus oligonucleotide (table II.1) suggested that this factor was a binding candidate on element B (Fig.III.30). The addition of antibodies against mouse C/EBPα resulted in the formation of a supershift in hepatic nuclear extracts (Fig.III.30), and *in vitro* translated mouse C/EBPα bound as a homodimer with element B (Fig.III.35). Because of the low content of HepG2 in C/EBPα, interaction of element B with the latter factor was also confirmed in rat-liver nuclear extracts. The two mutant elements B tested alone or the third mutant including the addition of both previous mutations suggested that the GTTTTCCAGGC sequence at +4/+11 nt was involved in C/EBP binding. The participation of the ATTGCCT sequence at +16+21 nt in binding remained elusive (this mutant competing partially with binding on element B). In this regard, not only half-site mutants, but a mutant +1/+2 nt with mutation at +1/+11 nt

similar to Mut 1 or at +16/+21 nt similar to Mut 2 will have to be designed and directly tested for direct binding of C/EBPα. Both homo- and heterodimers of C/EBPα have been identified (Lamb,1991; Thomassin,1992). Antibodies or expression plasmids for other C/EBP subtypes or isoforms were not available, but other family members and apparented members such as DBP, C/EBP β (or LAP), δ and γ, have been shown (with the exception of DBP) to heterodimerize with each other (Descombes,1990; Mueller,1990; Thomassin,1992). Therefore, direct binding or heterodimer formation with other members of the C/EBP family on element B (under constitutive and induced conditions of transfection) remains to be investigated.

5.2 Ubiquitous factors

5.2.1 Nuclear Hormone Receptors Acting in Trans- with Element A

The ubiquitous factors binding to element A and identified in this study were potential modulators of HNF4 activity for transcription of the AT gene. The choice of the factors studied was based upon literature reports and mainly the availability of antibodies and plasmids.

General Outline: The interaction of other receptors with elements A and C, in addition to HNF4, had to be expected. Effectively, many receptors of this family have been shown to interact as monomers, homo- and heterodimers with TGACC responsive elements (Green,1993; Giguere,1994; Tsai,1994). The list of the known

receptors of this family is still growing; they are classically and arbitrarily classified on the basis of their P-Box sequence (amino acid sequence between the two zinc fingers) (Sladek,1994). For the steroid-receptor superfamily group, binding to DNA depends upon the orientation of the TGACC motifs, their number and spacing; these receptors bind as homodimers and recognize elements in palindromes separated by 3 to 5 nucleotides upon the cases. The members of the second group of nuclear-receptors include diverse subfamilies, for example, the retinoid-receptors (*all trans*-retinoic acid receptor or RAR, 9-cis retinoic acid-receptor or RXR, the thyroid hormone-receptors (TR)), the orphan receptors such as COUP-TF's, PPAR and other recent members detailed below, and the vitamin D receptors (VDR). The receptors of this second group have less stringent rules for binding to their responsive elements, accomodating variable DNA configurations (Green,1993; Giguere,1994; Tsai,1994). In addition, various possibilities of dimerization and other receptor-receptor interaction --through both the DNA binding domain and the ligand binding domain -- added to newly described allosteric interactions can influence binding specificity (Tsai,1994; Forman,1995). Most of the retinoid-receptors' and some orphan receptors' binding sites reported to date include at least two half-sites of various orientations. Direct-repeats (DR), palindromes (P), everted-repeats (ER) and, very rarely, inverted-palindromes (for TR and COUP) with variable spacing (0-10 nt) constitute the responsive elements. Some receptors, nevertheless, are believed to have preferential arrangements (DR 3,4,5 rule for VDR, TR, RAR) but this is not well agreed. Monomeric binding has been reported only in very limited cases (TR and some of the newest orphan receptors, see below). A limited number of PuGGTCA elements preceded by an AT rich

sequence have also been reported to be potential responsive elements (Giguere,1994; Retnakaran,1994). Fig.IV.1 in this discussion illustrates hypothetical half- site motifs for these receptors in element A and C; these sites were found through homology search for COUP-TF1 binding sites. The content of Fig.IV.1 is discussed immediately below for element A. For element C, the hypothetical repeats drawn in the figure suggested DR1 and DR4 arrangements (putative responsive elements for orphan receptors, thyroid hormone, and retinoid receptors), and palindromes (particularly a P3 motif which partially matches a steroid responsive element) and which is centred on a CTTTG motif.

Ubiquitous Receptors Identified Fig.III.22, III.25, III.31, and III.33 presented the interaction of element A with the orphan receptor COUP-TF1. This receptor was initially described in the liver as an activator of the chicken ovalbumin upstream promoter (Wang,1989). Four highly homologous members of this sub-family are known at present: COUP-TF1 (or Ear-3), COUP-TFII (or ARP-1), Ear-2, and RVR (Tsai,1994; Retnakaran,1994). Rev-Erb A, ROR, and EGON are also ubiquitous orphan receptors related to COUP (Sladek,1994; Giguere,1994). COUP-TF's have a very promiscuous dimeric binding ability and accommodate various orientations and spacings of their responsive elements (up to 10 nt, such as in the HIV type I long terminal repeat); it is recognized that the only arrangements they do not bind to are GGTCA half-sites (Cooney,1991; 1992). These findings are supported by conformational evidence which has shown that COUP can wrap around its responsive element or just make contact with one face of DNA (Tsai,1994). The data for binding of COUP-TFI by wild-type and mutant

elements A presented in Chapter III further emphasizes these findings. As presented in Fig.III.31, mutations were targeted at the vicinity of the most obvious half-site, the TGACCT motif at -74/-79 nt. Other hypothetical half-site's motifs shown in Fig.IV.1 suggested also a GGTCA motif at -92/-88 nt, a motif in palindrome (P8) with the main half site, another TCATCA motif at -90/-86 nt (also in palindrome (P5)), and finally a CAGCCT -87/-82 nt imperfect direct-repeat (DR1). The mutational analysis selected for our study suggested that mutation of the 5 nt upstream of the main consensus at -79/-74 nt or a TGAgg mutation of the same consensus was not sufficient, alone, to suppress COUP binding; both mutations had to be combined to suppress binding. It is likely that the nucleotides upstream of the main half-site are important to stabilize binding. Mutational analysis could also suggest the presence of additional half-sites downstream of -74 nt, and further mutations upstream and downstream of the main half-site at -74/-79 nt could be designed. In particular, more drastic mutations of the DR1 element or mutations in the putative P5 element could determine if these half-sites are implicated directly in COUP binding. It may be likely that a complex arrangement of half-sites is present in this region and that more than one binding arrangement is possible (for example, deletion of the P8 palindrome would only have suppressed one COUP binding area). Downstream from the main half-site, other motifs were also found by homology search, such as the palindrome (P-3) overlapping the -74/-79 nt motif, ACCTCA at -77/-72 nt; the latter sequence matches a half-site consensus, but such an arrangement has never been reported for COUP binding. Nevertheless, the results for mutational analysis presented in the previous chapter tended to suggest that a TGAgg mutant could bind COUP-TF. The proof

for this hypothesis could be obtained by testing mutations between nucleotides -74 and -72 excluding the main half-site. A -73 C to G transversion in this motif, shown to affect COUP binding of the main half-site, should also impair binding of this putative half-site (Wang,1989).

Evidence was also found for interaction of element A with three members of the retinoid receptors family, RXR, PPAR and TR. Competition with the COUP-TF1 and the HD-PPRE oligonucleotides gave important clues in this regard; effectively, these two elements were known to bind this family of receptors (Tran,1992; Zhang,1992; Miyata,1993; Zhang, 1993). The first element, in addition to being a COUP-TF1 binding site, is also a retinoic-acid responsive element (RARE) and a thyroid-receptor responsive element (TR). The second element has been shown to interact with RXR, PPAR, RAR, LXR, and TR. The results for element A were presented in Fig.III.32 and III.33. Interaction with these receptors was confirmed with antibodies for RXR, PPAR, and TR. The reactivity of these antibodies was manifested by a decrease in the shifted species and/or the presence of supershifts upon the cases. Other nuclear hormone receptors' antibodies were tested and gave negative results (data not shown). Element A did not bind to RXR or PPAR individually - i.e. as a monomer or a homodimer (Fig.III.32). The presence of RXR homodimers has been documented for high concentrations of this receptor and is believed to be uncommon (Mangelsdorf,1991; 1994). In contrast, the presence of heterodimers between PPAR and RXR was seen, which migrated slightly above the HNF4 mobility shift (Fig.III.32). Binding of PPAR as an heterodimer was in

agreement with the publications in this regard (Kliewer,1992; Marcus,1993; reviewed by Green,1993). The same observation could be made for HNF4 and the PPAR/RXR heterodimers, i.e., a lesser qualitative binding was observed than that for COUP-TF1. When the mutations described above for COUP-TF1 and HNF4 were tested with the RXR/PPAR heterodimers, no evidence of binding for any of the mutants tested was found. These results are very similar to the results obtained with HNF4. The oligonucleotides shown to compete with element A (COUP and HD-PPRE) have the sequence CCTTTGACCT in common with this element. HD-PPRE mutational analyses have determined that the first repeat of this element (the TGACCT of the CCTTTGACCT sequence) is a key for RXR heterodimerization (Miyata,1993). In addition, our results for mutational analysis suggested that this region of element A was implicated in the binding of the retinoid receptors tested. In other words, the imperfect half-site in direct-repeat, one nucleotide upstream of the main half-site motif at -74/-79 nt, is beleived to constitute the probable second half-site of the responsive element for the retinoids receptors RXR and PPAR (DR1 motif of Fig.IV.1). In addition, DR1 elements have been reported to allow binding of COUP-TFI and II, HNF4, RXR, of heterodimers of RXR with RAR (all-trans retinoic acid receptor), VDR, and PPAR (Umesomo 1991; Cooney,1992, Giguere,1994; Tsai,1995).

The wild-type element A or an element A truncated of its -92 nt palindrome bound to the translated TRα product with the presence of a single band which immunoprecipitated with specific antibodies (Fig.III.33). Mutational analysis for TRα binding to element A gave a new demonstration of binding promiscuity by this receptor (Umesomo,1991;

Desvergne,1994). Effectively, evidence of binding was obtained with all mutants tested, even though binding was decreased strongly by the third mutation (a TcAgg mutation in the main half-site presented in Fig.III.33). This approach did not allow us to delineate the TRα binding site; this receptor has been shown to effectively bind as a dimer (homo and heterodimer) or as a monomer to DNA half-sites. In addition, it has also been shown that monomers, homodimers and heterodimers contact different nucleotide sequences in thyroid-responsive elements (Ikeda,1994). Monomer binding is believed to include recognition of a half-site extended by two bases in its 5' end, giving a TAAGGTCA consensus (Desvergne,1994). In this case also, the TR binding area in element A might not even overlap the RXR, PPAR, HNF4, and even COUP binding sites. Monomeric motifs searched within this element matched imperfectly the sequences at -79/-73 nt, -72/-77 nt, and -86/-79 nt respectively; this area encompassed the other receptor's binding sites. It may also be that more drastic mutations had to be designed to destroy TR binding in this area. For this receptor (and for the other members of the retinoid family as well), binding as a monomer, a homo or an heterodimer, as specified by the half-sites' consensus, their type of arrangement, their orientation and spacing, is still a very debated area (Forman,1992; Ribeiro,1992; Glass,1994; Ikeda,1994).

5.2.2 Single-Stranded Binding Protein, PYBP

Preliminary data based upon competition assays and supershifts in HeLa cells suggested PYBP interaction with element A of the AT 5' upstream promoter. However, the reactivity of this protein was not tested under the appropriate conditions,

i.e., with single-stranded oligonucleotides. Attempts to obtain a second aliquot of antibodies were unsuccessful. It has been reported that pPTB, the mouse homolog of PYBP, plays a role during splicing. A role for this protein family in replication (their DNA unwinding property is just one aspect illustrating this point), or in transcription has also been postulated. Chapter I reviewed examples of the interaction of several steroid-hormone receptors with single-stranded binding proteins for binding to their responsive elements, and PYBP or other homologous proteins' transactivation potential on basal transcriptional elements (Brunel 1991; Flavin,1991; Mukherjee,1990; Gaillard,1994). Therefore, the possible interaction of PYBP with the nuclear hormone receptors identified in element A should be investigated. According to the results presented in Chapter III, binding sites for PYBP and the receptors binding to the DR1 motif were overlapping. It may be that the homo- or heterodimerization profiles by the nuclear hormone receptors binding to element A could be influenced by PYBP. Interaction outside the two classical heterodimer's conditioning domains (the DNA and the ligand-binding domains) have already been reported for some receptors (Tsai,1994). This hypothesis is based on the observation that a mutation which destroyed the pyrimidine strand also destroyed HNF4 binding or RXR/PPAR heterodimerization on the DNA. In contrast, TR and COUP binding were not suppressed (the position of the COUP shift was nevertheless modified by this mutation).

5.2.3 Ubiquitous Factors Acting in Trans with Element B at +1/+37 nt

In the course of the identification of factors binding to element B,

competition between oligonucleotides for EBNA-1/NF1 and CTF/NF1 (ATTTGGCTTTGAAGCAATATG) was observed. This could be due to the resemblance of the CTF/NF binding motifs with the core sequence of C/EBP and the many other proteins binding to CCAAT boxes. Nevertheless, antibodies or expression vectors specific to this protein family were not available and not tested. Whether or not factors of the CTF/NF are directly implicated in interaction with element B remains to be demonstrated.

6 Effect of *in Vivo* Expression of Nuclear Hormone Receptors, C/EBP, and HNF3 to the Transcriptional Efficiency of the AT 5' Upstream Promoter

It was critical to determine whether or not the factors interacting *in vitro* with the elements A-C were able to modulate the activity of the AT 5'upstream promoter directly. For this purpose, two alternatives were possible; the first choice was to perform co-transfection experiments with isolated elements A-C subcloned in tandem repeats and placed upstream of a reporter gene under the control of an heterologous promoter (the more often reported SV40 or the CPS or TK minimal promoters); effectively, elements A-C were unable to promote transcription individually but were enhancers of transcription. The second alternative was to carry-out co-transfection of expressed factors with the AT promoter itself, an element only partly analyzed for transcription factor binding sites' content. It has been shown, especially for the nuclear hormone receptors, that transfection of an expressed factor alone can modify transcriptional activity of a minimal promoter (SV40 and TK) in the absence of any added element (Kimura,1993; Malik,1995). In addition, our study suggested binding by the same factor to different sites of a promoter:

HNF4, for example, interacted with elements A and C; other ubiquitous nuclear hormone receptors shown to interact with element A could be binding candidates for element C also. In addition, this study suggested occupancy of a unique site by various factors, certainly not all identified. For all the reasons above, and to better reflect AT gene transcription, it was decided to approach these in the context of the AT minimal promoter.

The cell-lines used to map the two promoters were tested in the co-transfection assays. These cell lines, as previously described, were shown to endogenously contain some or all of the transcription factors tested. HNF4, at least the isoform tested, was present in HepG2 but not in HeLa or Cos1 cells (based upon supershifts experiments). COUP-TF was present in all cell lines. PPAR/RXR were present in HepG2, HeLa and Cos1 cells. This was in disagreement with reports stating that HepG2 cells do not contain significant endogenous amounts of RXR mRNA (Rottman,1991). In addition, this cell line has also been shown to contain endogenous RARα (VuDac,1994). We therefore tested the cell line BSC40 known to contain endogenous COUP-TF but not RXR,PPAR, or HNF4 (Miyata,1993). As a preliminary, an attempt was made to confirm these findings by supershift assays but element A (and C as well) was not able to bind properly to nuclear extracts from BSC40 cells in various experimental settings (data not shown). Another concern was also the presence of putative endogenous ligands able to modulate the transcriptional responses; effectively, whether nor not the effects that we observed are really unliganded endogenously is unknown. Therefore, HepG2 cells, which have been shown to contain endogenously retinoic acid derivatives and most of the factors that we identified, were not sufficient to perform these co-transfection experiments (Vu-

Dac,1994); these cells were nevertheless transfected, to provide an example of a cell-line containing these factors endogenously. In contrast, the effects on transcriptional efficiency of the AT 5'upstream promoter observed in HeLa cells and furthermore in BSC40 cells should better reflect the effect of expressed factors added exogenously.

To facilitate the interpretation of the co-transfection experiments, the results of which are summarized in Fig.III-37 to 39, an arbitrary luciferase value of 1 was attributed to the minimal AT 5'upstream promoter. In fact, in the conditions applied for co-transfection (detailed in Chapter II) and as seen previously in the setting used initially to map the AT regulatory elements in pSVOA-LΔ5', luciferase expression obtained with the AT promoter in HeLa and BSC40 cells was 20% and 7% respectively of the luciferase expression obtained with HepG2 cells.

6.1 Factors with a Positive Effect

The common theme of the cotransfection experiments presented in Fig.III.37 to III.39 was the positive effect of HNF4, C/EBP, and more slightly RXR, in both liver and non liver cell-lines. The effect of this last receptor was barely seen in HepG2 cells.

HNF4 effects were at a maximum in BSC40, and were slightly less marked in HepG2 cells and Hela cells. It should be mentioned also that 8 μg of HNF4 co-transfected in BSC40 cells gave values lower than for 4 μg, likely due to squelching by this receptor. The positive effects of HNF4 and C/EBP in co-transfection, have been abundantly documented (Ladias,1992; Zakin,1992; Mietus-Snyder,1992; Kimura,1993; Schaeffer,1993; see Sladek,1994 for review, Nitsch,1994; Vaulont,1994; Hall,1995). For

C/EBP, different and sometimes opposite effects by various subtypes and isoforms have been reported (Shaeffer,1993; Kowentz, 1994). The recent consensus is that C/EBPα is more active in cell cycle arrested and differentiated hepatocytes, whereas the other subtypes are implicated in signal transduction (via cytokines for example) in various dividing or non dividing cell types (Ciliberto,1993). In addition, phosphorylation of C/EBPβ has been recently implicated as a mechanism able to control transcription by activation through derepression (Kowentz, 1995).

RXR had only a slight and a variable positive effect (in the cell-lines) on the activity of the AT promoter. An attempt was made to find information on the effect of co-transfected RXRα. Such a task was difficult because the effect of unliganded RXRα intervenes mostly as a control for exogenous induction and liganded receptor activity studies. In most cases, no mention of "RXR-minus" activity is made in the constitutive context. The partial information found for the case of an uninduced RXRα response shows variable and very slight effect on minimal promoter strength: this effect is positive for the activity of the ovalbumin promoter (which contains the COUP-TF1 binding site used in our competition experiments) but negative in the nuclear hormone responsive elements of the apolipoprotein genes and of the rat hydratase dehydrogenase gene (Kliewer,1992; Tran,1992; Widom,1992; Marcus,1993). Similarly, Rottman el al. (1991) did not find a positive effect of RXRα on the activity of the apolipoprotein A1 minimal promoter. In our experiments, the variable positive effect of RXRα alone and the RXRα-HNF4 combination could be related to the cellular milieu; effectively, HepG2 cells contained endogenously HNF4 and RXRα whereas BSC40 cells did not. As well, HNF4 was absent in HeLa cells.

Further co-transfection experiments with mutants of the AT promoter should determine the basis of the synergistic effects seen in BSC40 or HeLa cells. In addition, EMSA assays and supershifts with nuclear extracts prepared from BSC40 co-transfected cells should identify the factors which intervene for synergy. Whether or not these observations reflect the liver enriched expression of HNF4 is unknown. It is tempting to speculate that in BSC40 and Hela, two cell lines that gave a barely detectable AT reported activity, RXR, on its own, or by its synergistic effects with exogenous HNF4, could be an adaptative receptor which will attempt to force AT expression. The results of the co-transfection experiments presented in this work at least illustrated that the interplay of transcription factors on the AT promoter could modulate constitutive AT expression.

6.2 Factors with a Negative Effect

6.2.1 COUP-TFI

The results of the co-transfection experiments indicated a marked repressive effect of COUP-TF1 on the activity of the AT promoter; repression was seen when COUP-TF1 was expressed alone or with HNF4 whether or not in combination with the other receptors RXR, PPAR and/or TR (Fig.III.37 -III.39). It remains to demonstrate if pathways that force COUP expression could potentially antagonize the activation of the AT promoter by HNF4 (and perhaps RXR) physiologically. The repressive effects of COUP have been documented for HNF4 in many genes involved in lipid transport and metabolism, as well as amino acid and glucose metabolism (Sladek,1994). In this study,

competition of COUP with HNF4 did not involve heterodimerization. In fact, similar observations with other promoters have incriminated direct competition for binding of these two receptors on identical responsive elements, COUP-TF1 having the higher binding affinity (reviewed by Sladek, 1994). The observations reported in Chapter III also support these findings. COUP-TF1 also antagonizes the effect of unlinganded ubiquitous members of this family such as the estrogen receptor in the lactoferrin gene, RXR homodimers, RAR/RXR and RXR/PPAR heterodimers in various retinoids, vitamin D, and thyroid hormone responsive elements (Kliewer,1992; Liu,1992; Mietus-Snyder,1992; Widom, 1992; Cooney,1993; Miyata,1993; reviewed by Sladek,1994; and Tsai,1994). Again, most of these co-transfection studies involve responsive elements linked to minimal promoters of heterologous genes and exogenous hormonal induction. In this study, the effect of COUP-TF1 combinations with RXR or TR and with 3 or more ubiquitous receptors identified was not evaluated. The mechanisms implicated for the repressive effect of COUP-TF1 on the unlinganded effect of ubiquitous receptors of this family are either the formation of COUP heterodimers (with RXR or TR) or once again the direct competition for binding with factors interacting as homodimers or heterodimers (mainly of RXR) (Cooney,1993; Miyata,1994; Tsai,1994). The presence of co-factors and the possible modifications of the cytosolic forms of these receptors (phosphorylation events and nuclear translocation), have been implicated in at least some of COUP modulatory effects (Tsai,1994). Therefore, further *in vivo* and *in vitro* investigations with translated material and nuclear extracts from co-transfected cells are warranted in explaining the combinatorial effects of these receptors on the activity of the AT promoter.

Less frequently but of interest, COUP-TF1 and its close parent Arp1 can also activate responsive elements and proximal promoters. Such is the case for the LFA1 site of the apolipoprotein AI gene, the ovalbumin promoter, the PEPCK promoter and the transferrin promoter (Wang,1989; Ladias,1991; Shaeffer,1993; Malik,1995). Tran et al (1992) have also reported that COUP increases basal expression levels of unliganded TRα or of TRα/RXR combinations in DR1 elements. The problematic positive effect of COUP on the minimal promoter of the TK gene --in the absence of any putative responsive element-- is also another example of such an activity (Kimura,1993). The case of the LFA1 element is interesting because direct interaction between Arp1 activation domains with components of the basal transcription machinery is strongly suspected and is currently under investigation (Malik,1995). The opposite situation, in which a repressive COUP effect is seen, is suspected to involve the activation domain, which maps the domains for repression as well.

This duality in COUP effects (which resembles the case of TR) also favours the presence of a ligand mediated effect for this receptor, which in our case of constitutive expression, will have to be endogenous. Such ligands, previously reported for the retinoid receptors, have been long suspected for HNF4. As well, the antagonistic effects of COUP with HNF4 suggest cross-talk of responsive elements for these receptors with within and outside the CCTTTGA/GCC core sequence; many proximal promoters with this core sequence and reported to bind HNF4, COUP, TR, and retinoid receptors, are effectively modulated differently by COUP under constitutive conditions. Within these genes, the phosphoenol pyruvate carboxykinase (PEPCK) promoter is certainly the best known; the

element AF1 of this gene has been shown to bind (directly or after heterodimerization upon the cases) RAR, PPAR, RXR, TRβ, HNF4, and COUP, a situation very reminiscent of AT and all the other genes listed (Mitchell,1994). This AF1 is also one of the four motifs implicated in the glucocorticoid response of PEPCK (Mitchell,1994; Hall,1995). Furthermore, the glucocorticoid-responsive region interacts with proteins binding to CAAT motifs (C/EBP, and OTF1) as well as with NFκB and TR (Mitchell,1994). Overlapping cAMP responsive elements in the C/EBP responsive regions have also been identified, which are connected to the glucocorticoid responsive module. All the elements in this gene are located in close vicinity to one another. In conclusion, if an element AF1 (i.e. a nuclear hormone responsive element) is or is not part of such a complex responsive unit, it might be that part of COUP effects could be different as a result of differences in cross-talk with this modular responsive unit. This implies that interactions between responsive units should take place in the constitutive setting. The PEPCK gene and its glucocorticoid responsive module is not the only example for such interactions. A recent report has involved the early growth response factor Egr-1 in apolipoprotein AI (Apo AI) gene transcription (Kilbourne,1995). This liver enriched activator has been shown to overcome Arp1 mediated repression of Apo A1 element A (another nuclear hormone receptor responsive unit shown to bind all the factors we identified in element A of the antithrombin promoter) in HepG2 cells. Part of the molecular basis for these observations is the presence of an Egr 1 binding domain formed by two sites overlapping each end of Apo A1 element A.

6.2.2 TRα

The repressive effect of TRα observed in this study confirms the numerous reports on the effect of the unliganded receptor (reviewed by Desvergnes,1994; and Tsai,1994). The negative effect of this receptor on transcriptional activation by HNF4 was less marked in HeLa cells than in HepG2 and BSC40 (Fig.III.41). COUP and RXR have been shown to have variable effects on TR activity and on its subtypes, effects further modified by hormonal induction (Tran,1992; reviewed by Desvergnes,1994). The high binding plasticity of TR, which can accommodate even more binding sites than COUP, might explain some of these observations. Repression by unlinganded TR has been explained trough interference with the formation of the preinitiation-complex. More precisely, direct interactions of TFIIB with the ligand binding domain of TR have been reported by Fondell (1993). Moreover, recent data have suggested that unlinganded TR suppress TFIIB dependent activation when complexed with RXR (Forman,1995).

6.2.3 PPAR

In this study, this orphan receptor, modulator of intracellular and extracellular lipid metabolism and transport, had a slightly repressive effect on luciferase expression mediated by the AT promoter. No report on the activity of this factor in the constitutive situation for minimal promoters was found. For tandem repeats of DR1 responsive elements inserted upstream of the TK minimal promoter, PPAR has either no effect or a positive effect on constitutive transcription (Dreyer,1992; Marcus, 1993; VuDac,1994). Additionally, differential effects by subtypes of PPAR are also known (Marcus,1993).

6.2.4 HNF3

The most surprising observation made was the negative effect of the two HNF3 isoforms tested. This factor, at least the two subtypes that we tested, α and β, has been described almost exclusively as an activator in co-transfection experiments (reviewed by Ciliberto,1993). Nevertheless, a recent report on the apolipoprotein A1 gene has shown cell-type differences on HNF3 β activity: repression was seen for the whole minimal promoter in HepG2 whereas activation was seen in CV1 cells; moreover, a synergistic effect between HNF4 and HNF3 were seen in CV1 whereas antagonism was seen in HepG2 (Harnish,1994); when the HNF3 responsive element was tested alone, only transactivation was seen. Harnish's results suggest also that cross-talk takes place between HNF3 and other factors, a case likely to be relevant in our study. Cross-talk between HNF3 and NF1 has already been documented in the albumin promoter (LeviWilson,1993). Blocking of transcriptional activation by HNF3 has also been reported to explain the function of an apolipoprotein B gene reducer in non hepatic cells (Paulweber,1993). Moreover, the direct involvement of this factor has not been proven by our study. Effectively, competition of a HNF3 binding site with binding of elements A-C was not confirmed by EMSA with antibodies or translated products specific of the α and β HNF3 subtypes.

CHAPTER V

CONCLUSION AND FUTURE DIRECTIONS

The work presented in this thesis attempted to demonstrate the importance of the region flanking exon 1 of the AT gene in transcriptional control of gene expression. Within a large 6.9 kbp region, two cis-acting regulatory elements were found; both able to promote transcription after deletional analysis of the region analysed in a luciferase reporter assay. The first element was clustered immediately upstream and downstream of the presumed transcriptional start site. The location and the nature of this first element were in good agreement with the description of those for liver enriched promoters (Ciliberto, 1993). The second element, which functioned in a reverse direction in regard to the first in reporter assays, was located in the first intervening sequence.

1. 5' Upstream Promoter

1.1 Summary of the results presented in this thesis

The first regulatory element was narrowed down by deletional analysis to a -150/+68 nt region. Footprints and EMSA assays allowed for the partial characterization of three areas of interaction with factors acting in trans with the -150/+68 nt promoter. These data are summarized in Fig.V.1; the architectural data on the promoter in the latter figure include additional information provided by the recently published report of Tremp

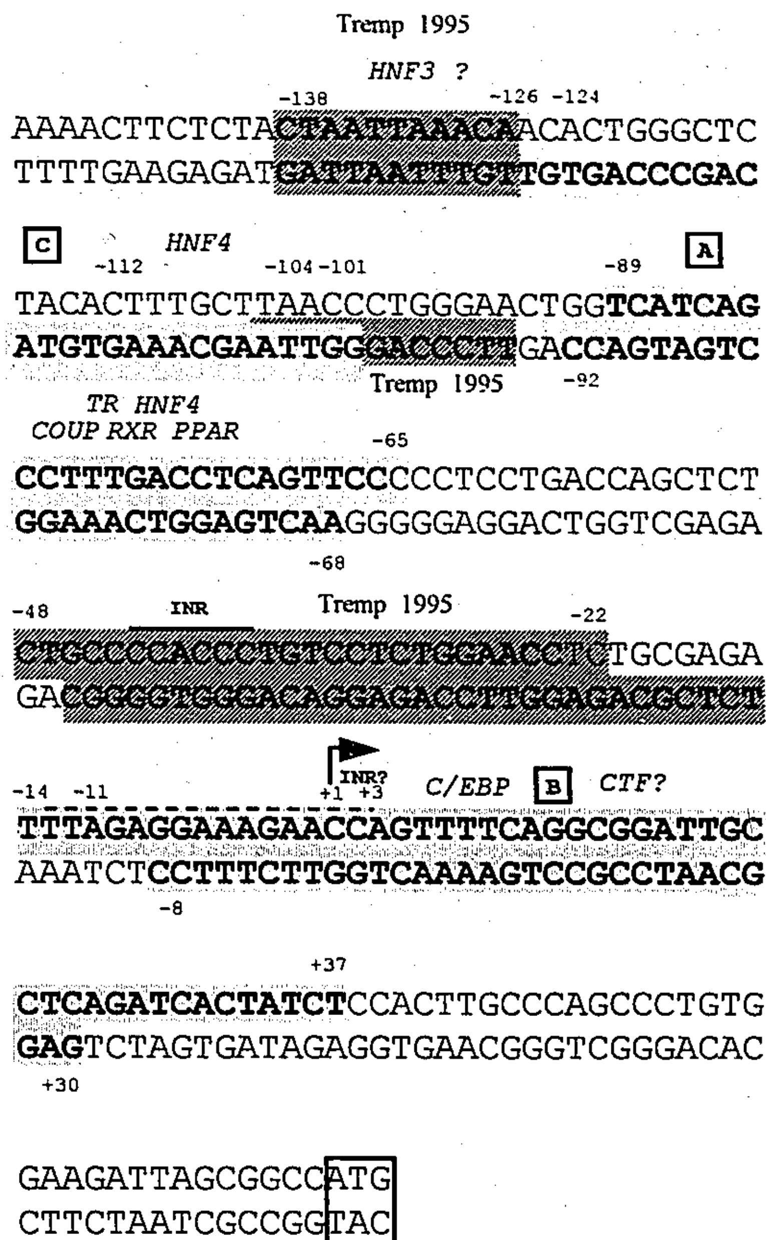


Figure V.1: ***Proposed architecture of the antithrombin 5' upstream promoter*.** The areas of interaction with transcription factors are shaded (this work) or hatched (Tremp et al). INR, initiator element. The first ATG is boxed.

et al. (1995). Our study determined that the elements A, B, and C were enhancers; Fig.V.1 also illustrates the interaction of the AT proximal promoter with the liver-enriched factors HNF4 (also reported in Tremp's study) and C/EBPα as well as with the ubiquitous factors RXRα, PPARα, COUP-TF1 and TRα. *In vitro* expression of the above factors showed activation by HNF4, C/EBPα and RXRα on the basal activity of the AT 5'upstream promoter. In contrast, TRα, COUP-TF1, and HNF3 α or β had a repressive effect on the activity of the same promoter. Our results suggest that the constitutive expression of AT depends upon the interplay of these factors; whether or not these results are physiologically relevant remains to be demonstrated. It is likely that promoter efficiency in the constitutive setting depends upon activation of HNF4 and C/EBPα. The marked decrease in promoter strength observed after deletion of the regions binding either of these two factors strongly suggests their importance. The relative participation of the sites binding HNF4 in element A and C and of the sites possibly binding C/EBP in element B and C will have to be evaluated separately through further co-transfections with deletional mutants of the 5'upstream minimal promoter. Nevertheless, co-transfection assays with mutations directly and only in the binding sites of these two factors should further confirm these observations. However, these studies might be complicated by the binding of several factors to the same site, by heterodimerization of these factors on the DNA, by possible modulatory influence by co-factors and allosteric interactions, and by cross-talk between responsive units. It is may be that a more appropriate approach in this case will be to suppress the endogenous expression of HNF4, C/EBPα or RXRα in cultured cell-lines and to determine the consequences of such a lack of expression on AT synthesis.

An interesting observation made in our study was the suggested synergy between HNF4 and RXR for activation of the AT 5'upstream promoter in BSC40 and HeLa cells. Whether or not this latter mechanism mediates transcription of the AT gene and/or is related to cell specific AT expression remains to be determined. The absence of synergy in HepG2 cells could have been influenced by the fact that this cell line produces HNF4 endogenously; this endogenous production of HNF4 (and of RXR as well) might interfere with the effects of exogenous factors expressed *in vivo*. It simply might also be that a wider selection of HNF4 and RXR dose combination had allowed for the detection of synergistic effects between the two factors. Moreover, several regulatory elements in the promoter (either binding HNF4 and/or RXR or in cross-talk with responsive elements for these two factors) might be involved in this effect. In addition, the *in vitro* data, supported by the fact that the two receptor's antibodies did not cross-react, also suggested a HNF4-RXR interaction; heterodimerization is only one of the many pathways for protein interaction and the fact that HNF4 is beleived to form only homodimers on the DNA might not exclude other forms of interaction, either directly or through co-factors and other unidentified mechanisms. These observations are interesting in that they provide the first suggestion of such a mechanism in constitutive transcription. Of interest, new evidence in the literature has shown that synergy between activators could be explained by stimulation of TFIIA-TFIID complex formation *via* interaction with TFIIB and the TAF's and not only by cooperative binding or interaction between activators (Chi,1995).

1.2 Prospective studies on Elements Acting *in Cis* with the 5'Upstream Minimal Promoter

The characterization of the active *cis*-acting elements constituting this promoter was incomplete in our study; effectively, the precise location of the sequences which initiate transcription in this promoter was not determined; based upon the deletional analysis carried out in this study, these sequences are likely located in the -70/+15 nt region. A more precise analysis of this region should determine whether or not the two putative initiators centred at -43 nt and +1 are functionally important.

1.3 Prospective Studies on the Factors Acting *in Trans* with the AT 5'upstream promoter

In this study, antibodies or expression vectors were usually tested for one isoform only of each transacting factor identified. In fact, several subtypes and isoforms are already known for the factors identified herein; these subtypes and isoforms exhibit *in vitro* and *in vivo* differences for tissue distribution, developmental stage, or transactivation potential (reviewed by Giguere, Sladek, and Tsai (1994); Honenberg, 1995). In addition, differences in transcriptional response pathways (for C/EBP for example) have also been linked to intrafamily diversity. Further studies with available isoforms of the factors identified in this study will be informative.

In addition, the identification of the factors acting in trans with the AT 5'upstream promoter is largely incomplete. It is likely that other liver-enriched and ubiquitous factors will interact with regulatory elements to be discovered or already identified.

1.3.1 Liver-Enriched Factors

The interaction with additional liver enriched factors such as members of the HNF3 family was also strongly suspected from the EMSA competition assays but remains to be demonstrated. Effectively, competition of an HNF3 consensus oligonucleotide was present for the three elements to a certain extent even if competition with element A was only partial. Competition with this element was not confirmed by direct binding of *in vitro* translated HNF3 α or β or by antibodies for HNF3 (data not shown). This could be explained by the fact that the HNF3α binding consensus contains two TGACTT in palindrome and a direct repeat of TGATT; both motifs above resemble the TGACCT consensus of members of the nuclear hormone superfamily. In addition, the sequence CCTTTG at -83 nt in element A also resembles the TGTTTGC motif of a transferrin enhancer and of an apolipoprotein B reducer (Auge-Guillou,1993; Paulweber,1993). Both elements, which were believed to be putative C/EBP binding sites, have been shown to directly bind HNF3α. Noone has ever investigated whether transcription factors other than HNF3, such as nuclear hormone receptors or members of the C/EBP family for example, could interact with an HNF3 consensus oligonucleotide. Competition with an HNF3 consensus binding site was also seen in element B. Nevertheless, no supershifts or direct binding could be seen with antibodies or expressed HNF3 α or β (Fig.III.30). It was noted above that the -11/+21 nt region of the AT promoter contains sequences with homologies to the binding sites of proteins interacting with CCAAT motifs. Within this region, the same element (TGGTTCTTTCC at +4/-7 nt) also resembles the HNF3α binding site found in the transferrin and apolipoprotein B elements (TGTTTGCTTTTC). Whether or

not the latter element at +4/-7 nt binds to HNF3 members remains to be investigated. Finally, no confirmation of HNF3 competition was obtained with element C by supershifts and direct studies with HNF3 α and β. In the recent report by Tremp, it was mentioned concerning the -100 nt region (area II) that "competition experiments suggested that this sequence could be a weak HNF3 related protein binding site". First, this factor may have not been identified for experimental reasons. Second, another HNF3 isoform could be binding; it was noted above that 14 additional members related to the HNF3 family were recently identified (Clevidence,1993; Pierrou,1994). Third, studies on HNF3 binding have reported strong and weak affinity binding sites (Costa,1991; Lai 1991); the classification of the latter sites is beleived to be unrelated to a clear difference in consensus but rather depends upon the affinity of the isoforms and also upon the nature of neighbouring binding sites. Le Cam et al (1994) have also reported in their study on GAGA boxes a CTGAGAAAT core binding site in the rat Spi 1.2 promoter with putative binding of C/EBP and HNF3, in which HNF3 competition was not confirmed by direct binding. Tremp et al (1995) have also mentioned the identification of HNF3 in the 138/-123 nt footprint of the AT promoter. Nevertheless, this assumption was based upon competition EMSA assays only. In addition, the oligonucleotide used in their study for this area, a -145/-111 nt element, did more than encompass the -138/-123 nt area, presumed to be the HNF3 putative binding site in their study.

Possibilities of C/EBP binding by AT elements other than B were not investigated in this work. Initial publications describe C/EBP as a factor not only binding CCAAT motifs but also TGTGGA/TA/TA/TG motifs in viral enhancers (Landschultz,1987).

Prochownik, in 1985, reported an homology between the -11/-3 nt region of the AT gene, AGAGGAAAG, with such a viral enhancer. We have studied only the binding areas immediately downstream of the presumed start site. Whether or not C/EBP could also bind to the region immediately upstream from the presumed start site has not been evaluated; the footprint of the start site region encompassed this area. Other sites in the AT promoter could also be C/EBP binding targets. For example, possible C/EBP and HNF4 overlapping binding sites could be present in the promoter in elements A or C such as those found in a highly homologous region of the apolipoprotein B promoter (Metzger,1993). In element A, the sequence CTTTG in CTTTGACCT resembles an inverted half-site of a C/EBP consensus, and in element C, the sequence CTTTGTTAAC also resembles the general binding consensus of this factor. In addition, we have observed competition of a C/EBP oligo with element C, even if we could not confirm, by the presence of supershifts, or by the detection of direct binding of *in vitro* translated material, the direct involvment of C/EBPα (Fig.III.29 and III.36). Either we have missed interaction with this factor for experimental reasons, or it could be that another subtype or isoform of C/EBP is involved at this site.

As well, no attempt was made to study interactions with members of the HNF1 family; this was due to the fact that the consensus binding site, G/AGTTAA/TNNTT/CAACC/A, could not be matched with area A, B, or C. Nevertheless this does not imply that this factor does not bind to any of these elements or to additional regulatory elements not identified in our study. The element in the -150 nt region, for example, reported by Tremp to bind HNF3 shows a partial homology with a

HNF1 consensus; overlapping HNF1-HNF3 sites have been described on several occasions (Jackson,1992; reviewed by Ciliberto,1993).

1.3.2 Ubiquitous Factors

Ubiquitous factors were also shown to bind or interact at sites also occupied by liver-enriched factors in the 5'upstream promoter. Whether such a situation exists *in vivo* is unknown, but it is now well recognized that tissue specific expression depends upon the interplay of factors enriched and non-enriched in a given tissue (Ciliberto, 1993).

The identification of ubiquitous factors binding to the AT promoter was carried out for element A only. Moreover, these studies were targeted mostly to an eventual modulation of HNF4 binding and activity. The RXR/PPAR association (and possibly HNF4/RXR) were the only two cases suggesting heterodimerization; no heterodimers were seen between COUP-TF1 and any of the receptors tested. Nevertheless, heterodimer formation between TR and other retinoid receptors was not studied. In fact, heterodimer formation between RXR, COUP-TF, RXR, RAR, and VitDR has been abundantly documented (Tran,1992; Berrodin,1992; Cooney,1993; reviewed by Giguere,1994). Heterodimerization of RXR with all the above receptors as well as recently described orphan receptors, for example, hNGF1-B/Nurr1/RNR1 (which binds on half sites), RLD-1, Ubiquitous Receptor, and ROR, is believed to be a key mechanism for retinoids signalling, both exogenously and endogenously (Zhang,1992; Allenby,1993 Giguere,1994; Appel, 1994; Retnakaran, 1994; Song, 1994; Forman,1995). In addition, there is now

multiple evidence showing that all these receptors bind to their responsive elements with higher affinity as heterodimers with RXR than as homodimers. Whether or not heterodimerization with RXR is also preponderant under constitutive conditions of transcription in the absence of exogenous ligand remains to be demonstrated. As well, the structural study recently published by Rastinejad (1995) should provide new insights into the structural arrangement of the heterodimer formed by RXRα: RXRα and TRβ DNA binding domains on a DR4 responsive element. In this arrangement, the DNA binding domains of the two receptors engage into the major DNA groove of the successive direct repeats with a head to tail orientation, giving an asymmetrical arrangement. The RXR binding domain occupies the upstream AGGTCA element and the TR binding domain the downstream element; this confirms previous observations that heterodimers of RXR with TR or RAR do not bind to DNA if RXR is forced to interact with the 3' half site. These interactions are further stabilized by interaction of the TR receptor with the minor groove of the half-site and spacer (this confirms the extended recognition consensus by TR 5'to its half site). In this study also, modelling in differently spaced RXR heterodimers (DR3 with VDR, DR5 with RAR) suggests spacer discrimination with "adaptation" to a different conformation or orientation that would not cause steric interference between subunits. It will be of major interest to follow further development in structural arrangements of RXR heterodimers on DR1 elements.

Elements A at -92/-65 nt and C at -124/-101 nt include potential half-site consensus binding sites for nuclear hormone receptors other than the one identified. We just mentioned above that several of these putative receptors such as RAR, VDR, or the latest

members, for example, Ubiquitous Receptor, LXR, ROR, Nur-1, or RLD 1, heterodimerize with RXR (Apfel,1994; Desvergnes,1994; Song,1994; Forman,1995). Most of these receptors have been described as activators but receptors repressing the activity of TR and RXR activity have also been reported (Retnakaran,1994; Hirose,1995). Whether or not these new receptors can modulate the constitutive expression of AT remains to be demonstrated. The molecular basis of the unlinganded activity of these factors is still being debated. As suggested by several reports, the repressive effects could be linked to direct interactions with the initiation machinery (Fondell,1993; reviewed by Tsai,1994). As well, bridge proteins between activators and the general factors have now been implicated in constitutive transcription; N-CoR, Trip-1 and Sug-1 are examples of such proteins for TR, RAR and RXR (Lee, 1995; Horlein, 1995; Kurokawa, 1995).

As well, footprints and EMSA assays in nuclear extracts of various cellular origin suggested the interaction of element B with ubiquitous factors. The very preliminary studies carried out in this regard incriminate binding by proteins binding to CAAT motifs such as CTF/NF. The involvment of factors of this family should be confirmed with respective antibodies and *in vitro* and *in vivo* expression vectors for these proteins. The presence of CTF/NF1, NFY, and Oct1 has been abundantly documented for liver specific promoters at the vicinity of their initiation transcriptional machinery (see Table I.1). Of interest, these factors often also participate to the mechanisms of replication. As well, we have discussed in the previous chapter putative GAGA binding sites, in the region of the promoter immediately upstream of the presumed start-site of transcription.

1.4 Constitutive *versus* Inducible AT Expression; Cross-Talk Between Responsive Elements

The factors identified and acting in *Trans* with the AT 5'upstream promoter have been also implicated in inducible transcriptional responses. For example, in element A, the interactions with RXR, TR, and PPAR suggest that expression by the AT promoter could be modulated by these receptors not only under constitutive conditions but also in case of induction by hormones, retinoids, peroxisome-proliferators, thyroid-hormone and pharmacological agents. These responses have been implicated in developmental and fundamental metabolic and physiological programmes. Responsive elements to retinoids have already been described in the tissue plasminogen activator and the apolipoprotein genes (Bovard-Houpermans, 1994; Vu-Dac, 1994, 1995; Bulens, 1995): the responsive elements in these genes resemble elements A and C of the AT gene. In addition, other putative responsive elements for retinoid and steroid receptors should be investigated in element C. As well, element B, in the region -14/+ 37 nt, contains putative responsive elements for hormonal and signalling pathways involved, for example, in the acute-phase or the growth-hormone response. These responses are known to involve in part C/EBP subfamilies (mainly β and δ) interaction and cross-talk. Furthermore, mention was made concerning the resemblance of sequences in the -14/+ 37 nt of the AT gene with consensus for NFκB, γ interferon (GAS/ISRE elements), HSRE binding factors and GA-boxes binding factors, as well as growth hormone responsive elements; most of the above responses have mediators in the JAK-STAT families. A recent report has identified *cis*-acting elements in the factor VIII gene. One of the features of the minimal promoter is a

GATAA box, located in a C/EBP binding region, and at the vicinity of the NFκB region (Figueiredo,1995): this case of the factor VIII gene is very reminiscent of that of the AT gene. These hypotheses were not investigated in this study, mainly because the focus was made on constitutive expression. Nevertheless, these putative pathways suggest that AT is not entirely a house-keeping gene with a steady-state expression. Our studies could also suggest that constitutive expression of AT involve dynamic mechanisms of transcription, also intervening in inducible responses. In the nuclear hormone receptor superfamily for example, the presence of endogenous ligands, of cross-talk with second messengers, growth factors, or kinase pathways is currently suspected to play a critical role in mechanisms of transcription independent of ligands (Tsai,1994).

An aspect not approached by these studies was the interplay and the role of various responsive units in the AT promoter to constitutive or inducible responses. We were able to show that HNF4, C/EBP and RXR activated basal expression of the promoter whereas COUP-TF1, TR, and HNF3 repressed this expression. The only association of factors tested were that of nuclear hormone receptors on element A. Other eventual associations, or the relative participation of diverse elements of the promoter to the global effects observed, were not evaluated. The literature describes, for example, the interplay of HNF4 and C/EBPα responsive elements in overlapping regions of the apolipoprotein B promoter (in an AF1 region). *In vivo* in the latter element these two factors have a synergistic effect (Metger,1993). *In vitro*, they compete for binding on their overlapping binding site, but participation of additional factors to this interaction is beleived to explain the synergy seen

in vivo. In contrast, the HNF4-C/EBP interaction described in the transferrin promoter is not synergistic *in-vivo* (Schaeffer,1993). In the latter promoter, the HNF4 binding site has been recently shown to be occupied by COUP-TF and PYBP also; this site is totally homologous with the -89/-68 nt CTTTGACCTC element of the AT gene and only partly homologous with the whole A element (14/22 matches). We have mentioned above C/EBP interactions with ubiquitous factors; some generate heterodimers, such as the p65 subunit of NFκB or the thyroid-hormone receptor (Liao,1992; Mitchell,1994; Ray,1995). These interactions are also believed to specify the cell-type of C/EBP mediated gene expression, especially in the case of NFY or of the glucocorticoid receptor (Kowentz,1994). Similarly, other interactions of HNF4 with factors not a part of the nuclear hormone receptor superfamily but binding to overlapping sites have also been reported, e.g., C/EBP and cAMP responsive elements (Vaulont,1994; Nitsch,1994).

2. IVS 1 Element

The fact that the characterization and the actual role of the IVS1 element of the AT gene are to date preliminary has been discussed extensively in Chapter IV. Whether or not this element has a true role in transcription has to be proven directly. In this regard, the -766/-416 nt promoter of the factor IX gene provides an example of an intronic element relevant to our study: the latter promoter, with an inverse directionality also, has a moderate silencer activity on the 5'upstream promoter; as well, it has been shown recently to be located within the 3'end region of a Line-1 sequence (Kurachi,1994). Whether or not the sequences in IVS1 could silence AT expression by the 5'upstream promoter is

unknown; the full-lentgh AT gene construct including 4800 nt of upstream sequences, exon 1, and 2117 nt of IVS1 was unable to promote luciferase expression. To determine whether or not the absence of reporter activity is related to the opposite directionality of the two elements mapped or to silencer activity by the IVS1 elements, successive deletions of the two elements in regular or inverse orientation should be made and tested for luciferase expression. These studies should determine if this element is a true silencer. Our preliminary studies have shown that this element promotes transcription in the SV40-derived pGL-reporter system. Nevertheless, the IVS1 element was never tested directly in tandem or in opposition with the AT 5'upstream promoter. An additional aspect which should not be forgotten is the observation that the regions in exon 1 and the immediate 304 nt after this exon also decreased activity of the 5'upstream promoter; whether this effect is linked to regulatory sequences for transcription or to translated and/or spliced sequences of AT remains to be investigated also. However, unlike the case of a construct including the whole IVS1 region, exon 1 and the immediate 3' 304 nt did not totally suppress 5'upstream promoter activity.

3. Impact of this Study For The Molecular Basis of AT Deficiency

To date, the only suggestion of the involvment of the regions regulating expression of the AT gene to inherited defects of AT expression was provided by the 5' deletion of one AT allele described in the introductory part of this work. It was discussed that the presence of this deletion was not directly associated with a defect in transcription of the AT gene. Of the many families characterized to date with type I AT deficiency, only

mutations in the coding or splicing sequences of the gene are known. Nevertheless, a number of type I AT deficient families with abnormal AT expression exhibit no mutation of the sequences involved in the coding-splicing events. It would be of interest to establish whether mutations of the elements mapped in our study, at least in the 5'upstream promoter, could explain the lack of expression of the affected AT allele in these families. These studies should take into account possible polymorphisms that could also influence AT gene expression. They also should be informative for transcriptional knowledge of this serpin promoter.

REFERENCES

Aden,D.P., Fogel,A., Plotkin,S., Damjanov,I., and Knowles,B.B (1979). Controlled synthesis of HBsAg in a differentiated human liver carcinoma derived cell line. *Nature* **282**, 615-618.

Allenby,G., Bocquel,M.T., Saunders,M., Kazmer,S., Speck,J., Rosenberg,M., Love,A., Kastner,P., Grippo,J.F., and Chambon, P. (1993). Retinoic acid receptors and retinoid X rceptors: interactions with endogenous retinoic acids. *Proc. Natl. Acad. Sci.USA* **90**, 30-34.

Amrani,D.L., Rosenberg,J., Samad,F., Bergstrom,G., and Banfield,D.K.(1993). Developmental expression of chicken ATIII is regulated by increased RNA abundance and intracellular processing. *Biochem. Biophys. Acta* **1171**, 239-246.

Anderson,K.P., and Lingrel,J.B.(1989). Glucocorticoid and oestrogen regulation of the rat T-kininogen gene. *Nucleic Acids Res.* **17**, 2835-2898.

Angotti,E., Mele,E., Costanzo,F., and Avedimento,E.V.(1994). A polymorphism in the -70 position of the apolipoprotein AI promoter increase transcription efficiency. *J. Biol. Chem.* **269**, 17371-17377.

Angrand,P.O., Rousset,J.P., and Weiss,M.C. (1992). Cell phenotype, binding affinity and promoter structure modulate transactivation by HNF1 and LAP. *Journal of Cell Science* **103**, 1083-1092.

Antoine,B, Levrat,F., Vallet,V., Berbar,T., Cartier,N., Dubois,N., Briand,P., and Kahn,A. (1992). Gene expression in hepatocyte-like lines established by targeted carcinogenesis in transgenic mice. *Exper. Cell Res.* **200**, 175-185.

Apfel,R., Benbrook,D., Lernhardt,E., Ortiz,M.A., Salbert,G., and Pfahl,M.(1994). Anovel orphan receptor specific for a subset of thyroid hormone-responsive elements and its interaction with the retinoid-thyroid hormone receptor subfamily. *Mol. and Cell. Biol.* **14**, 7025-7035.

Aso,T., Conaway,J.W., and Conaway,R.C.(1994). Role of core promoter structure in assembly of the RNA polymerase II preinitiation complex. *J. Biol. Chem.* **269**, 26575-26583.

Auge-Guillou,C., Petropoulos,I., and Zakin,M.M. (1993). Liver enriched HNF3α and ubiquitous factors interact with the human transferrin gene enhancer. *F.E.B.S. Letters* **223**, 4-10.

Atchison M.L.,(1989) Enhancers; mechanisms of action and cell specificity. *Ann. Rev. Cell Biol.* **4**, 127-153.

Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A.,and Struhl,K. (1989). *Current protocols in Molecular Biology. Vol 1 to 3,* Green Publishers Associates & Wiley Interscience, New York.

Bach,I., and Yaniv,M. (1993). More potent transcriptional activators or a transdominant inhibitor of the HNF1 homoeodomain family are generated by alternative RNA processing . *E.M.B.O. J.* **12**, 4229-4242.

Bancroft,J.D., Schaeffer,L.A., and Degen,S.J.F. (1990). Characterization of the Alu rich 5'flanking region of the human prothrombin encoding gene. *Gene* **95**, 253-260.

Berg,L.P., Grundy,C.B., Thomas,F., Millar,D.S., Green,P.J., Kiawczak,M., Reiss,J., Kakkar,V.V., and Cooper,D.N. (1992). *De novo* spice site mutation in the antithrombin gene causing recurrent venous thrombosis. Demonstration of exon skipping by ectopic transcript analysis. *Genomics* **13**, 1359-1361.

Berg,L.P., Scope,D.A., Alhaq,A., Kakkar,V.V., and Cooper,D.W. (1994). Disruption of a binding site for hepatocyte nuclear factor I in the protein C promoter is associated with hereditary thrombophilia. *Hum. Mol. Gen.* 3, 2147-2153.

Berrodin,T.J., Marks,M.S., Ozato,K., Linney,E., and Lazar, M.A. (1992). Heterodimerization among thyroid hormone receptor, retinoic acid receptor retinoid X receptors, chicken ovalbumin upstream promoter transcription factor, and an endogenous liver protein. *Mol. Endocrin.* **6**, 1468-1478.

Blajchman,M.A. (1994). An overview of the mechanism of action of antithrombin and its inherited deficiency states. *Blood Coag. Fibrin.* **5**, Suppl 1, 5-11, 59-64.

Bock,S.C., Wion,K.L., Vehar,G.A., and Lawn,R.M. (1982). Cloning and expression of the cDNA for human antithrombin III. *Nucl. Acids Res.* **10**, 8113-8125.

Bock,S.C., and Levitan,D.J. (1983). Characterization of an unusual DNA lengh polymorphism 5'to the human antithrombin III gene. *Nucl. Acids Res.* **11**, 8569-8583.

Bock,S.C., Harris,F.J., and Balazx,I. (1985). Assignment of the human antithrombin structural gene to chromosome 1q 23-25. *Cytog. Cell Genetics* 39, 67-69.

Bock,S.C., Marrinan,J.A., and Radziejewska,E. (1988). Antithrombin III Utah: proline 407 to leucine mutation in a highly conserved region near the inhibitor reactive site. *Biochemistry* **27**, 6171-6178.

Bossone,S.A., Asselin,C., Patel,A.J., and Marcu,K.B. (1992). MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences regulating transcriptional initiation and termination. *Proc. Natl. Acad. Sci.USA* **89**, 7452-7456.

Bossu,J.P., Chartier,F.L., Vu-Dac,N., Fruchart,J.C., and Laine,B. (1994). Transcription of the human apolipoprotein A-II is down regulated by the first intron of its gene. *Biochem. Biophys. Res. Commun.* **202**, 822-829.

Brunel,F., Alzari,P.M., Ferrara,P., and Zakin,M.M. (1991). Cloning and sequencing of PYBP, a pyrimidine-rich specific single strand DNA-binding protein. *Nucleic Acids Res.* **19**, 5237-5245.

Bucher,P.(1990). Weight matrix description of four eukaryotic RNA polymerase II promoter elements derived from 500 unrelated promoter sequences. *J. Mol. Biol.* **212**, 563-578.

Bulens,F., Ibanez-Talon,I., Van Acker,P., Nells,L., Belayew,A., and Collen,D.(1995). Retinoic acid induction of human tissue type plasminogen activator gene expression via a direct repeat element located at -7.0 Kbp. *J. Biol. Chem.* **13**, 7165-7175.

Bulla,G.A., and Fournier,R.K. (1994). Genetic analysis of a transcriptional activation pathway by using hepatoma cell variants. *Mol. and Cell. Biol.* **14**, 7086-7094.

Buratowski,S.(1994).The basics of basal transcription by RNA polymerase II.*Cell* **77**,1-4

Carrell,R.W., Stein,P.E., Fermi,G., and Werdell,M.R. (1994). Biological implications of a 3A structure of dimeric antithrombin. *Structure* **2**, 257-270.

Carslson,P., and Bjursell,G. (1989). Negative and positive promoter elements contribute to tissue specificity of apolipoprotein B expression. *Gene* **85**, 8998-9002.

Cereghini,S., Yaniv,M., and Cortese,R. (1990). Hepatocyte dedifferentiation and extinction is accompanied by a block in the synthesis of mRNA coding for the transcription factor HNF1/LFB1. *E.M.B.O. J.* **9**, 2257-2263.

Chai,K.X., Ma,J.X., Murray,S.R., Chao,J., and Chao,L. (1991). Molecular cloning and analysis of the rat kallikrein binding protein gene. *J. Biol. Chem.* **265**, 16029-16036.

Chartier,F.L., Bossu,J.P., Laudet,V., Fruchart,J.C., and Laine,B. (1994). Cloning and sequencing of cDNA encoding the human hepatocyte nuclear factor 4 indicate the presence of two isoforms in human liver. *Gene* **147**, 269-272.

Chen,J.D., and Evans,R.M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* **377**, 454-458.

Chen,W.S., Manova,K., Weinsflin,D.C., Duncan,S.A., Plump,A.S., Prezioso,V.R., Bachvarova,R.F., and Darnell,J.E.Jn. (1994). Disruption of the HNF4 gene expressed in the visceral endoderm leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryo. *Gene Dev.* **8**, 2466-2477.

Chi,T., Lieberman,P., Ellwood,K., and Carey,M. (1995). A general mechanism for transcriptional synergy by eukaryotic activators. *Nature* **377**, 254-257.

Chodosh,L.A., Baldwin,A.S., Carthew,R.W., and Sharp,P.A. (1988). Human CCAAT binding proteins have heterologous subunits. *Cell* **53**, 11-24.

Chomiki,N., Henry,M., Alessi,M.C., Anfosso,F., and Juhan Vague,I. (1994). Plasminogen activator inhibitor-1 expression in human liver and healthy or atherosclerotic vessel walls. *Thromb. Haemost.* **72**, 44-53.

Chow,B.K., Ting,W., Tufaro,F., and McGillavray,R.F.A. (1991). Characterization of a novel liver specific enhancer in the human prothrombin gene. *J. Biol. Chem.* **266**, 18927-18933.

Christy,R.J., Kaestner,K.H., Geiman,D.E., and Lane,M.D. (1988). CCAAT/enhancer binding protein gene promoter. *Proc. Natl. Acad. Sci. USA* **88**, 2593-2597.

Ciliberto,G., Colantuoni,V., De Franscesco,R., and Cortese,R. (1993). Transcriptional control of gene expression in hepatic cells. *Gene Expression* (Karin,M.,Ed), Birkauser, Boston, pp162-242.

Citarella,F., Smith,T., Felici,A., Aiuti,A., La Porta,C., and Fantoni,A. (1993). The 5' sequence of human factor XII gene contains transcription regulatory elements typical of liver specific, estrogen modulated genes. *Biochem. Biophys. Acta* **1172**, 197-199.

Clevidence,D.E., Overdier,D.G., Tao,W., Qian,X., Pani,L., lai,E., and Costa,R.H.(1993). Identification of nine tissue specific transcription factors of the hepatocyte nuclear factor 3/forkhead DNA binding domain family. *Proc. Natl. Acad. Sci. USA* **90**, 3948-3952.

Connor,J.M., Fowkes,F.G.R., Wood,J., Smith,F.B., Donnan,P.T., and Lowe,Gd. (1992). Genetic variation at the fibrinogen loci and plasma fibrinogen levels. *J. Med. Genet.* **29**, 480-482.

Cool,D.E., and MacGillavray,R.T. (1987). Characterization of the human blood coagulation factor XII gene. Intron exon gene organization and analysis of the 5' flanking region. *J. Biol. Chem.* **262**, 13662-13673.

Cooney,A.J., Tsai,S.Y., O'Malley,B.W., and Tsai,M.J. (1991). Chicken ovalbumin upstream promoter transcription factor binds to a negative regulatory region in the human immunodeficiency virus type I long terminal repeat. *J. Virol.* **65**, 2853-2860.

Cooney,A.J., Tsai,S.Y., O'Malley,B.W., and Tsai,M.J. (1992). Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different GGTCA response elements, allowing COUP-TF to repress hormonal induction of the vitamin D3, thyroid hormone, and retinoic acid receptors. *Mol. Cell. Biol.* **12**, 4153-4163.

Cooney,A.J., Leng,X., Tsai,S.Y., O'Malley,B.W. and Tsai,M.J. (1993). Multiple mechanisms of chicken ovalbumin upstream promoter transcription factor dependent repression of transactivation by the vitamin D, thyroid hormone, and retinoic acid receptors. *J. Biol. Chem.* **268**, 4152-4160.

Costa,R.H., Grayson,D.R., and Darnell,J.E. (1989). Multiple hepatocyte enriched nuclear factors function in the regulation of the transthyretin and the α1 antitrypsin genes. *Mol. Cell. Biol.* **9**, 1416-1425.

Costa,R.H., and Grayson,D.R. (1991). site-directed mutagenesis of hepatocyte nuclear factor binding sites in the mouse transthyretin promoter reveal synergistic interactions with its enhancer region. *Nucl. Acids Res.* **15**, 4139-4145.

Courtois,G., Morgan,J.G., Campbell,L.A., and Crabtree,G.R. (1987). Interaction of a liver specific nuclear factor with the fibrinogen and α1 antitrypsin promoters. *Science* **238**, 688-692.

Cowell,I.G. (1994). Repression versus activation in the control of gene transcription. *TIBS* **217**, 38-42.

Crossley,M., and Brownlee,G.G. (1990). Disruption of a C/EBP binding site in the factor IX promoter is associated with haemophilia B. *Nature* **345**, 444-446.

Crossley,M., Ludwig,M., Stowell,K.M., de Vos,P., Olek,K., and Brownlee,Q.G.(1992) Recovery from haemophilia Leyden; an androgen responsive element in the factor IX promoter. *Science* **257**, 377-379.

Dalmon,J., Laurent,M., and Courtois,G. (1993). The human β fibrinogen promoter contains a hepatocyte nuclear factor 1 dependent, IL6 responsive element. *Mol. Cell. Biol.* **13**, 1183-1193.

Darlington,G.L. (1987). Liver cell lines. *Met. Enzymol.* **151**, 19-51.

Dawson,S., Wiman,B., Hamsten,A., Green,F., Humphries,S., Henney,A.M. (1993). The two alleles sequences of a commun polymorphism in the promoter of the plasminogen activation inhibitor-1 (PAI-1) gene respond differently to IL1 in HepG2 cells. *J. Biol. Chem.* **268**, 10739-10745.

Derman,E., Kreuter,K, Walling,L., Weinberger,C., Ray,M., and Darnell,J.E. (1981). Transcriptional control in the production of liver-specific mRNA's. *Cell* **29**, 731-739.

Descheemaeker,K.A., Wyns,S., Nelles,L., Auwers,J., Ny,T., and Collen,D.(1992).Interaction of AP1, AP2, and Sp1 like proteins with two distint sites in the upstream regulatory region of the PAI-1 gene mediates the phorbol ester response. *J. Biol.Chem.* **267**, 15086-15091.

Deschepper,C.F., Bigornia,V., Berens,M.E., Lapointe,M.C. (1991). Production of thrombin and antithrombin III by brain and astroglial cell cultures. *Mol. Brain Res.* **11**, 355-358.

Descombes,P., Chojkier,M., Lichsteiner,S., Falvery,E., and Shibler,U. (1990). LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator. *Gene and Dev.* **4**, 1541-1551.

De Simone,V., and Cortese,K. (1992). Mechanisms of liver-specific gene expression. *Curr. Opin.Cell Biol.* **3**, 96-965.

De Simone,V., Ciliberto,G., Hardon,E., Paonessa,G., Palla,F., Lundberg,L, and Cortese,R. (1987). *Cis*- and *trans*-acting elements responsible for the cell specific expression of the human α1 antitrypsin gene. *E.M.B.O. J.* **6**, 2759-2766, 1987.

Desvergnes,B. (1994). How do thyroid hormone receptors bind to structurally diverse response elements? *Mol. Cell. Endocrin.* **100**, 125-131.

De Wet,J.R., Wood,K., Helinski,D.R., and DeLuca M. (1985). Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **82**, 7870-7873.

De Wet,J.R., Wood,K.V., De Luca,M., Helinski,D.R. and Subramani,S. (1987). Firefly Luciferase gene: structure and expression in mammalian cells. *Mol. Cell. Biol.* **7**, 725-737.

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

Dombroski,B.A., Mathias,S.I., Nanthakumar,E., Scott,A.F., and Kazazian,H.H.Jn. (1991). Isolation of an active human transposable element. *Science* **254**, 1805-1808.

Drapkin,R., Merino,A., and Reinberg,D. (1993). Regulation of RNA polymerase II transcription. *Current Opinion in Cell Biology* **5**, 469-476.

D'Souza,S.E.,and Mercer,J.B.F.(1987).Antithrombin III mRNA in adult rat liver and kidney and in rat during development. *Biochem.Biophys.Res.Comm.* **142**,417-421.

Edelberg,J.M., Sana,D.C., and Pizzo,S.V. (1994). Vascular regulation of plasminogen activator inhibitor-1 activity. *Sem. Thromb. Haemost.* **20**, 319-323.

Egeberg,O.(1965). Inherited antithrombin deficiency causing thrombophilia.*Thromb.Diath.Haem.* **13**, 516-530.

Eissenberg,J.C., and Elgin, S.C.R. (1991). Boundary functions in the control of gene expression. *Trends in Genetics* **7**, 335-339.

Erikson,P., Kallin,B., Houf,F.M., Bavenholm,P., and Haastem,A.(1995).Allele specific increase in basal transcription of the plasminogen activator inhibitor 1 gene is associated with myocardial infarction. *Proc. Natl. Acad. Sci. USA* **92**, 1851-1855.

Eyster,M.E., and Parker,M.E. (1985). Treatment of familial ATIII with danazol. *Haemostasis* **15**, 119-125.

Falvey,E., Fleury-Olela,F., and Schibler,U.(1995). The rat hepatic leukemia factor encodes two transcriptional activators with distinct circadian rhythms, tissue distribution, and target preferences. *E.M.B.O. J.* **14**, 4307-4317.

Felsch,J.S.,and Owen,W.G.(1994). Endogenous antithrombin associated with microvascular endothelium. *Biochemistry* **33**, 818-822.

Fernandez-Rachubinski,F., Rachubinski,R., and Blajchman,M.A. (1992). Partial deletion of an antithrombin III allele in a kindred with a type I deficiency. *Blood* **80**, 1476-1485.

Figueiredo,M.S., and Brownlee,G.G. (1995). *Cis*-acting elements and transcription factors involved in the promoter activity of the human factor VIII gene. *J. Biol. Chem.* **270**, 11828-11838. .

Flavin,M., and Strauss,F. (1991). Multiple sequence specific single strand binding proteins for the promoter region of the rat albumin gene. *DNA and Cell Biology* **10**, 113-118.

Fondell,J.D., Roy,A.L., and Roeder,R.G. (1993). Unliganded thyroid hormone receptor inhibits formation of a functional preinitiation complex; implications for active repression. *Genes and Developm.* **7**, 1400-1410.

Ford,I., Li,T.C., Cooke,I.D., and Preston,F.E. (1990). Changes in the natural inhibitors of coagulation, ATIII, Protein C and Protein S, during treatment with Danazol. *J. Mal. Vasc.* **15**, 23-35

Forman,B.M., Umesomo,K., Chen,J., and Evans,R.E. (1995). Unique response pathways are established by allosteric interactions among nuclear hormonereceptors. *Cell* **81**, 541-550.

Frain,M., Swart,G., Monaci,P., Nicosia,A., Stampli,S., Rainier,F., and Cortese,R. (1989). The liver specific transcription factor LF-B1 contains a diverged homeobox DNA binding domain.*Cell* **59**, 145-157.

Friedman,A.D., and Mc Knight,S.L. (1990). Identification of two polypeptides segments of CCAAT/ enhancer binding protein required for transcriptional activation of the serum albumin gene. *Genes and Developm.* **4**, 1416-1426.

Friezner,S.J.(1992). The prothrombin gene and its liver-specific expression. *Semin.Thromb Haemost.***18**, 230-242.

Frohman,M., Dush,M., and Martin,G. (1988). Rapid production of full lengh cDNA from rare transcripts; amplification using a single gene specific oligonucleotide primer. *Proc. Natl. Acad. Sci.USA* **85**, 8998-9002.

Gaillard,C., Cabannes,E., and Strauss,F. (1994). Identity of the RNA binding protein K of hnRNP particles with protein H16, a sequence specific single strand DNA binding protein. *Nucleic Acids Res.* **22**, 4183-4186.

Galas,D.J., and Schmitz,A.D (1978). DNAse footprinting; a simple method for the detection of protein-DNA binding specificity. *Nucl. Acids Res.* **5**, *3157-3170.*

Garcia-Blanco,M.A., Jamison,S., and Sharp,P.A. (1989). Identification and purification of a 62kDa protein that binds specifically to the polypyrimidine tract of introns. *Genes and Dev.* **3**, 1874-1886.

Gerber,H.P.,Hagman,M.,Seiger,K.,andCorden,J.(1995).RNA polymerase II CTD required for enhancer driven transcription. *Nature* **374**, 660-662.

Giguere,V. (1994). Retinoic acid receptors and cellular retinoid binding proteins: complex interplay in retinoid signaling. *Endocrine Reviews* **15**, 61-79.

Gill,A., Sharp,P.A., Jamison,S.F., and Garcia-Blanco,M.A. (1991). Characterization of cDNA's encoding the polypyrimidine tract-binding protein. *Gene and Developm.* **5**, 1224-1236.

Gill,G., and Tjian,R. (1992). Eukaryotic coactivators associated with the TATA box binding protein. *Curr. Opin. in Genetics and Development* **2**, 236-242.

Gillman,M.Z., Wilson,R.N., and Weinberg,R.A. (1986). Multiple protein binding sites in the 5' flanking regions regulate c-fos expression. *Mol. and Cell. Biol.* **6**, 4305-4316.

Gluzman,Y. (1981). SV40 transformed simian cells support replication of early SV40 mutants.*Cell* **23**, 175-182.

Goodrich,J.A., and Tjian,R. (1994). TBP-TAF complexes: selectivity factors for eukaryotic transcription. *Current Opin. in Cell Biol.* **6**, 403-409.

Gould,S.J., and Subramani,S. (1988). Firefly luciferase as a tool in molecular and cellular biology. *Anal. Biochem.* **175**, 5-13.

Graham,F.L., and Van Der Heb. (1973). A new technique for the assay of infectivity of human adenovirus IV DNA. *Virology* **52**, 456-471.

Graupner,G., Wills,K.N., Tzukerman,M., Zhang,X.K., and Pfahl,M. (1990). Dual regulatory role for thyroid hormone receptors allows control of retinoic acid receptor activity. *Nature* **340**, 653-655.

Graves,B., Johnson,P.F., and Mc Knight,S.L. (1986). Homologous recognition of a promoter domain common to the MSV-LTR and the HSV TK gene. *Cell* **44**, 565-576.

Green,S., Isseman,I. and Sheer,E. (1988). A versatile *in-vivo* and *in-vitro* eukaryotic expression vector for protein engineering. *Nucl. Acids Res.* **16**, 369.

Green,S.(1993).Nuclear hormone receptors. *Nature* **361**,590-591.

Gregori,C., Ginot,F., Decaux,J.F., Weber,A., Berbar,T., Khan,A., and Pichart,A.L. (1991). Expression of the rat aldolase B gene; a liver specific proximal promoter and an intronic activator. *Biochem. Biophys. Res. Commun.* **176**, 722-729.

Gyllesten,U.B., and Erlich,H.A. (1988). Generation of single stranded DNA by the polymerase chain reaction and its applications to direct sequencing of the HLA-DQA locus. *Proc. Natl. Acad. Sci. USA* **85**, 7652-7656.

Hafeez,W., Ciliberto,G., and Perlmutter,D.H. (1992). Constitutive and modulated expression of the human α1 Antitrypsin gene. Different transcriptional initiation sites used in three different call types. *J. Clin. Invest.* **89**, 1214-1222.

Hall,R.K., Sladek,F.M., and Granner,D.K. (1995). The orphan receptors COUP-TF and HNF4 serve as accessory factors required for induction of phosphoenolpyruvate carboxykinase gene transcription by glucocorticoids. *Proc. Natl. Acad. Sci.USA* **92**, 412-416.

Harnish,D.C., Malik,S., and Karathanassis,S.C. (1994). Activation of apolipoprotein -A1 gene transcription by the liver enriched HNF3. *J. Biol.Chem.* **269**, 28220-28226.

Hatton,M.W.C., Moar,S.L., and Richardson,M. (1988). Evidence that rabbit ^{125}I-antithrombin binds to proteoheparan sulfate at the subendothelium of the rabbit aorta *in-vivo. Blood Vessel* 67, 878-886.

Hattori,M., Abraham,L.J., Northemann,W., and Fey,G.H. (1990). Acute phase reaction induces a specific complex between hepatic nuclear proteins and the IL6 response element of the rat α2 macroglobulin gene. *Proc. Natl. Acad. Sci.* USA **87**, 2364-2368.

Hattory,M., Tugores,A., Veloz,L., Karin,M., and Brenner,D. (1990). A simplified method for the preparation of transcriptionally active liver nuclear extracts. *DNA Cell Biol.* **9**, 777-781.

Healy,A.M., and Gelether,T.D. (1994). Induction of PAI-1 in HepG2 cells by mediators of the acute phase response. *J. Biol. Chem.* **269**, 19095-19100.

Herbst,R.S., Nielsh,U., Sladek,F., Lai,E., Babiss,L.E., and Darnell,J.E. (1991). Differential regulation of hepatocyte enriched transcription factors explains changes in albumin and transthyretin gene expression among hepatoma cells. *New Biology* **3**, 289-296.

Hirosawa,S., Fahner,J.B., Salier,J.P., Wu,C.T., Lovrien,W.W., and Kurachi,K.(1990) Structural and functional basis of the developmental regulation of human coagulation factor IX gene: factor IX Leyden. *Proc. Natl. Acad. Sci. USA* **87**, 4421-4425.

Hirosawa,S., Nakamura,Y., Miura,O., Sumi,Y., and Aoki,N. Organization of the human α_2 plasmin inhibitor gene.(1988). *Proc. Natl. Acad. Sci. USA* **85**, 6836-6840.

Hirose,T., Apfel,R., Pfahl,M., and Jetten,A.M. (1995). The orphan receptor TAK-1 acts as a repressor of RAR,RXR, and T3R mediated signalling *pathways.Biochem. Biophys. Res. Commun.* **211**, 83-91.

Hocke,G.M., Barry,D., and Fey,G.H. (1992).Synergistic action of IL6 and glucocorticoids is mediated by the IL6 response element of the rat α2 macroglobulin gene. *Mol. and Cell. Biol.* **12**, 2282-2294.

Honenberg,A.H., Monden,T., and Wondisford,F.E. (1995). Ligand independent and dependent functions of THR isoforms depend upon their distinct amino termini. *J. Biol. Chem.* **270**, 14274-14280.

Horlein,A.J, Naar,A.M., Heinzel,T., Torchia,J., Gloss,B., Kurokawa,R., Ryan,A., Kamei,Y., Soderstrom ,M., Klass,C.K., and Rosenfeld,M.G. (1995). Ligand independent repression by the thyroid hormone receptor mediated by a nuclear co-repressor. *Nature* **377**, 397-403.

Howell,B.W., Lagace,M, and Shore,G.C. (1989) Activity of the carbamoyl phosphate synthetase promoter in liver nuclear extracts is dependent on a *cis*-acting C/EBP recognition element. *Mol. Cell. Biol.* **9**, 2928-2933.

Hu,C.H., Harris,J.E., Davie,E.W., and Chung,D.W. (19950. Characterization of the 5'-flanking region of the gene for the alpha chain of human fibrinogen. *J. Biol. Chem.* **270**, 28342-28349.

Huang,M.W., Hung,H.L., Stanfield-Oakley,S.A., and High,K.A.(1992).Characterization of the factor X promoter. *J. Biol. Chem.* **267**, 15440-15446.

Huber,P., Laurent,M., and Dalmon,J. (1990). Human β fibrinogen gene expression. *J. Biol. Chem.* **265**, 5695-5701.

Hunt,L.T. and Dayoff,M.O. (1980). A surprising new protein superfamily containing ovalbumin, antithrombin III and α1 proteinase inhibitor. *Biochem. Biophys. Res. Commun.* **95**, 864-869.

Ikeda,M., Rhee,M., and Chin,W.W. (1994). Thyroid hormone receptor monomer, modimer, and heterodimer (with RXR) contact different nucleotide sequences in thyroid hormone response elements. *Endocrinology* **94**, 1354-1358.

Isseman,I., Prince,R., Tugwood,J. and Green,S. (1990). Activation of a member ofthe steroid hormone receptor superfamily by peroxisomes proliferators. *Nature* **347**, 645-650.

Jansen-Dur,P., Boshart,M, Lupp,B., Bayerhoff,A., Frank,R.W., and Schutz,G (1992). The rat polipyrimide (1992). The rat polypyrimidine tract binding protein (PTB), interacts with a single-stranded DNA binding motif in a liver specific enhancer. *Nucl. Acids Res.* **20**, 1243-1249.

Jackson,D.A., Rowader,K.E., Stevens,K., Jiang,C., Milos,P., and Zaret,K.S. (1993). Modulation of liver specific transcription by interactions between hepatocyte nuclear factor 3 and nuclear factor 1 binding DNA in close opposition. *Mol. Cell. Biol.* **13**, 2401-2410.

Jobin,F., Vu,L., and Lessard,M.(1991). Two cases of inherited triple deficiencies in a large kindred with thrombotic diathesis and deficiencies of ATIII, HCII, Protein C, and Protein S. *Thromb. Haemost.* **66**, 295-299.

Johnson,M.R., Bruzdzinski,C.J., Winograd,S.S., and Gelehter,T.D. (1992). Regulatory sequences and protein-binding sites involved in the expression of the rat plasminogen activator inhibitor 1 gene. *J. Biol. Chem.* **267**, 12202-12210.

Johnson,P.F., and McKnight,J.L. (1989). Eukaryotic transcriptional regulatory proteins. *Ann. Rev. Bioch.* **58**, 799-839.

Kallaria,R.N, Golde,T., Kroon,S.N., and Perry,G. (1993). Serine protease inhibitor antithrombin III and its messenger RNA in the pathogenesis of Alzheimer's disease. *Am. J. Pathol.* **143**, 886-893.

Kardassis,N., Zannis,V.J., and Cladaras,C. (1990). NFBA1 purification and characterization of the nuclear factor BA1, a transcriptional activator of the human apoB gene. *J. Biol. Chem.* **265**, 21733-21740.

Kennedy,G.C., and Rutter,W. (1992). Pur-1, a zinc finger protein that binds to purine-rich sequences, transactivates an insulin promoter in heterologous cells. *Proc.Natl. Acad. Sci. USA* **89**, 11498-11502.

Kilbourne,E.J.,Widom,R.,Harnish,D.,Malik,S., and Karathanasis,S.K.(1995). Involvment of early growth response factor Egr-1 in apolipoprotein AI gene transcription. *J. Biol. Chem.* **270**, 7004-7010.

Kimura,A., Nishiyori,A., Murakami,T., Tsudamoto,T., Hata,S., Osumi,T., Okamura, R., Mori,M. and Takigushi, M. (1993) Chicken ovalbumin Upstream Promoter transcription factor (COUP-TF) represses transcription from the promoter of the gene for ornithine transcarbamylase in a manner antagonistic to hepatocyte nuclear factor-4 (HNF4). *J. Biol. Chem.* **268**, 11125-11133.

Kliewer,S.A., Umesomo,K., Noonan,D.J., Heyman,R.A., and Evans,R.M. (1992, A). Convergence of 9-cis retinoic acid and peroxisome proliferator signalling pathways trough heterodimer formation of their receptors. *Nature* **358**, 771-774.

Kliewer,S.A., Umesomo,K., Noonan,D.J., Heyman,R.A. and Evans R.M. (1992, B) Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature* **355**, 446-449.

Kluger,W., Kaling,M., Ross,K., Wagner,U, and Ryffel,G.U. (1990). BAP, a rat liver protein that activates transcriotion through a promoter element with similarity to the USF/MLTF binding site. *Nucl. Acids Res.* **18**, 6943-6949.

Koleske,A.J., and Young,R.A. (1995). The RNa polymerase II holoenzyme and its implications for gene regulation. *TIBS* **20**, 1133-1136.

Kowenz-Leutz,E.,Twanley,G.,Ansieau,S., and Leutz,A.(1994). Novel mechanism of C/EBPβ transcriptional control; activation through derepression. *Gene Developm.* **8**, 2781-2791.

Krieg,P., and Melton,P. (1984). Functional mRNA's are produced by SP6 *in-vitro* transcription of cloned cDNA's. *Nucl. Acids Res.* **12**, 7075-7070.

Kritis,A.A., Ktistaki,E., Barda,D., and Zannis,V.I., and Talianidis,I. (1993). An indirect negative autoregulatory mechanism involved in hepatocyte nuclear factor-1 gene expression. *Nucl. Acids Res.* **21**, 5882-5889.

Kuo,C.J., Conley,P.B., Chen,L., Sladek,F.M., Darnell,J.E., and Crabtree,G.R.(1992). A transcriptional hierarchy involved in mammalian cell-type specification. *Nature,* **355**, 457-461.

Kurachi,S., Hitomi,Y., Furukawa,M., and Kurachi,K. (1995). Role of intron 1 in expression of the human factor IX gene. *J. Biol. Chem.* **270** , 5276-5281.

Kurokawa,R.,Sodestrom,M.,Horlein,A., Halahmi,S., Brown,M.,Rosenfeld,M.G., and Glass,C.K. (1995). Polarity specific activities of RAR determined by a co-repressor. *Nature* **377**, 451-454

Ladias,J.A., and Karathanasis, S.K. (1991). Regulation of the Apolipoprotein A1 gene by Arp1, a novel member of the steroid receptor superfamily. *Science* **251**, 561-565.

Ladias,J.A., Hadzopoulou-Cladaras,M., Kardassis,D., Cardot,P., Cheng,J., Zannis,V., and Cladaras,C. (1992) Transcriptional regulation of human apolipoprotein genes apoB, apoCIII, and apoAII by members of the steroid hormone receptor superfamily HNF4, ARP-1, EAR-2, and EAR-3. *J. Biol. Chem.* **267**, 15849-15860.

Laemmli,U.K., Kas,E., Poljak,L., and Adachi,Y. (1992). Scaffold-associated regions. *Current Opinion in Genetics and Development* **2**, 275-285.

Lagace,M., Goping,I.S., Mueller,C.R., Lazzaro,M., and Shore,G.C. (1992). The carbamyl phosphate synthetase promoter contains multiple binding sites for C/EBP related proteins. *Gene* **118**, 231-238.

Lai,E., Prezioso,V.R., Smith,E., Litvin,O., Costa,R.H., and Darnell,J.E. (1990). HNF3-A, a hepatocyte enriched transcription factor of novel structure is regulated transcriptionally. *Genes and Developm.* **4**, 1427-1436.

Lamb,P., and McKnight,S.L. (1991) Diversity and specificity in transcriptional regulation; the benefits of heterotypic dimerization. *T.I.B.S.* **16**, 417-422.

Landschutz,W.H., Johnston,P.F., Adashi,E.Y., Graves,S., and McKnight,S.L. (1988). Isolation of a recombinant copy of the gene encoding C/EBP. *Genes and Developm.* **2**,786-800.

Lane,D.A., Ireland,H., Olds,R.J., Thein,S.L., Perry,D.J., and Aiach,M. (1991). Antithrombin III; a data base of mutations. *Thromb. Haemost.* **66**, 657-661.

Lane,D., Prentki,P., and Chandler,M. (1992) Use of gel retardation to analyze protein -nucleic acid interactions. *Microb. Rev.* **56**, 509-528.

Lane,D.A, Olds,R.J., Boisclair,M., Chowdury,V., Thein,S.L., Cooper,D.W., Blajchman,M., Perry,D., Emmerich,J., and Aiach,M. (1993). ATIII mutations database, first update. *Thromb. Haemost.* **70**, 361-369.

Lane,D.A., Olds,R.J., and Chandler,M. (1992). Antithrombin III; summary of the first database update. *Nucl. Acids Res.***22**, 3556-3559.

Lao,T.T., Yin,J.A.,and Ynen,P.M. (1990).Coagulation and anticoagulation systems in newborns.Correlation with their mothers at delivery. *Gynecol. Obstetr. Invest.* **29**, 181-187.

Laurell,C.B., and Rannevik,G. (1978). A comparison of plasma protein changesi induced by Danazol, pregnancy, and oestrogens. *J. Clin. Endocrino. and Metab.* **49**, 719-726.

Le Cam,A., Pantescu,V., Paquereau,L., Legraverend,C., Fauconnier,G., and Asins,G.(1994).*Cis*-acting elements controlling transcription from rat serine protease inhibitor 2.1 gene promoter. *J. Biol. Chem.* **269**, 21532-21539.

Lee,J.W., Ryan,F., Swaffield,J.C., Johnston,S.A., and Moore,D. (1995) Interactions of thyroid hormone receptor with a conserved transcriptional mediator. *Nature*

Leff,T., Reue,K., Culver,H., and Breslow,J.L. (1989). A regulatory element in the ApoCIII promoter that directs hepatic specific transcription binds to proteins in expressing and non expressing cell types. *J. Biol. Chem.* **264**, 16132-16137.

Levinson,B., Kenwrick,S., Lakich,D., Hammond,G., and Gitschier,J. (1990). A transcribed gene in an intron of the human factor VIII gene. *Genomics* **7**, 1-11.

Lewin B.(1994) Chromatin and gene expression: constant questions, but changing answers. *Cell* **79**, 397-406.

Liechsteiner,S., Wuarin,J., and Schibler,U. (1987). The interplay of DNA binding proteins on the promoter of the mouse albumin gene. *Cell* **51**, 963-973.

Lipman,D,and Pearson,W.R.(1983).Rapid and sensitive protein similaritiessearches. *Science* **227**, 1435-1441.

Liu,Y., Yang,N., and Teng,C.T. (1993). COUP-TF acts as a competitive repressor for estrogen receptor-mediated activation of the mouse lactoferrin gene. *Mol. Cell. Biol.* **13**, 1836-1846.

Malgaretti,V., Bruno,l., Pontaglio,M., Canchiari,G., Meroni,G., Cereghini,S., and Tonelli,V.S. (1990). Definition of the transcription initiation site of the human plasminogen gene in liver and non hepatic cell lines. *Biochem. Biophys. Res. Com.* **173**, 1013-1018.

Malik,S., and Karathanasis,S.C.(1995). Transcriptional activation by the orphan nuclear receptor Arp1. *Nucl. Acids Res.* **23**, 1536-1543.

Mangelsdorf,D.J., Borgmeyer,U., Heyman,R.A., Zhou,J.Y., Ong,E.S., Oro,A.E., Kakizuka,A. and Evans,R.M. (1992). Characterization of three RXR genes that mediate the action of 9-cis retinoic acid. *Gene and Developm.* **6**, 329-344.

Mangelsdorf,D.J., Umesomo,K, and Evans,R.M. (1994). The retinoid receptors. In Sporn MD, Roberts, AB., and Goodman, D.S., eds. The retinoids ; biology, chemistry and medecine. New York: Raven press, 319-349.

Marcus,S.L., Miyata,K.S., Zhang,B., Subramani,S., Rachubinski,R.A., and Capone, J.P. (1993) Diverse peroxisome proliferators activated receptors bind to the peroxisome proliferatior responsive element of the rat hydratase-dehydrogenase and fatty acyl-CoA oxidase genes but differentially induce expression. *Proc Natl. Acad. Sci. USA* **90**, 5723-5727.

Martinez,E., Chiang,C.M., Ge,H., and Roeder,R. (1994). TATA binding-protein associated factor(s) in TFIID function trough the initiator to direct basal transcription from a TATA-less class II promoter. *E.M.B.O. J.* **13**, 3115-3126.

Matsuda,K., Kobune,Y., Noda,C., and Ichihara,A. (1994). Expression of GATA binding transcription factors in rat hepatocytes. *FEBS Letters* **353**, 269-272.

Matthijs,G., Cassiman,J.J., Van Den Bergue,H., and Marynen,P.(1994). Characterization of the human $\alpha 2$ macroglobulin gene promoter: Identification of a novel, triple TRE/RARE/ERE response element. *Biochem. Biophys. Res. Commun.* **202**, 65-72.

Maxam,A.M., and Gilbert,W. (1980) Sequencing end-labelled DNA with base specific chemical cleavage. *Methods Enzymol.* **65**, 499-517.

Meijers,J.C.M., and Chung,D.W. (1991). Organization of the gene coding for human protein C inhibitor (Plasminogen activator inhibitor 3). *J. Biol. Chem.* **266**, 15028-15034.

Mendel,D.B., and Crabtree,G.R. (1991,1). HNF-1, a member of a novel class of dimerizing homeodomain proteins. *J. Biol. Chem.* **266**, 677-680.

Mendel,D.B., Khavary,P.A., Conley,P.A., Graves,M.K., Hansen,L.P., Admon,L.P. and Crabtree,G.R. (1991,2) Characterization of a cofactor that regulates dimerization of a mammalian homeodomain protein. *Science* **254**, 1762-1767.

Metzger,S., Halaas,J.L., Breslow,J.L., and Sladek,F.M. (1993) Orphan receptor HNF4 and bZIP protein C/EBPα bind to overlapping regions of the apolipoprotein B gene promoter and synergistically activate transcription. *J. Biol. Chem.* **268**, 16831-16838.

Miao,C.H., Leytus,S.P., Chung,D.W., and Davie,E.W. (1992). Liver specific expression of the gene coding for human factor X, a blood coagulation factor. *J. Biol. Chem.* **267**, 7395-7401.

Mietus-Snyder,M, Sladek,F.M., Ginsburg,G.S., Kuo,C.F., Ladias,J.A., Darnell, J.E., and Karathanasis, S.K. (1992). Antagonism between apolipoprotein AI regulatory protein 1, Ear3/COUP-TF, and hepatocyte nnuclear factor 4 modulates Apolipoprotein CIII gene expression in liver and intestinal cells.*Mol. and Cell. Biol.* **12**, 1708-1718.

Mitchell,J., Noisin,E., Hall,R., O'Brien,R., Imai,E., and Granner,D. (1994). Integration of multiple signals through a complex hormone response unit in the Phosphoenolpyruvate Carboxykinase Gene Promoter. *Mol. Endocrin.* **8**, 585-594.

Miyata,K.S., Zhang,B., Marcus,S.L., Capone,J.P., and Rachubinski,R.A. (1993). COUP-TF binds to a peroxisome proliferator responsive element and antagonizes peroxisome proliferator mediated signaling. *J. Biol. Chem.* **268**, 19169-19172.

Mirkovitch,J. and Darnell,J.E. (1991). Rapid *in-vivo* footprinting technique identifies proteins bound to the TTR gene in the mouse liver. *Gene and Developm.* **5**, 83-93.

Mitchell,P.M., and Tjian,R. (1989) Transcriptional regulation in mammalian cells by sequence specific DNA binding proteins. *Science* **245**, 371-378.

Montell,D.J., Roth,P., and Sprading,A.C. (1992). Slow border cells, a locus required for a developmentally regulated cell migration during oogenesis, encodes Drosophila C/EBP. *Cell* **71**, 51-62.

Mourey,L., Samama,J.P., Delarue,M., Petitou,M., Choay,J., and Moras,D. (1993). Crystal structure of cleaved bovine antithrombin III at 3.2 A resolution. *J. Mol. Biol.* **232**, 223-241.

Mueller,P.R., and Wold,B. (1989). *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246**, 780-786.

Mueller,C.R., Maire,P., and Schibler,U. (1990). DBP, a liver-enriched transcriptional activator, is expressed late in ontogeny, and its tissue specificity is determined posttranscriptionally. *Cell* **61**, 279-291.

Mukherjee,R., and Chambon,P. (1990). A single stranded DNA binding protein promotes the binding of the purified oestrogen receptor to its responsive element. *Nucl. Acids Res.* **18**, 5711-5716.

Nerlov,C., and Ziff,E.B. (1995). CCAAT/enhancer binding protein-alpha amino aci d motifs with dual TBP and TFIIB binding ability co-operate to activate transcription in both yeast and mammalian cells. *E.M.B.O. J.* **14**, 4318-4328.

Nicholow,D., Hu,S.H., Lin,J., Gash,A., and Burley,S. (1992). Crystal structure of TF-IID TATA-box binding protein. *Nature* **30**, 40-46.

Niessen,R.W., Sturk,A., Hordijk,P.L., Michiels,F., and Peters,M. (1992). Sequence characterization of a sheep cDNA for antithrombin III. *Biochem. Biophys. acta* **1171**, 207-210.

NIH/CEPH collaborative mapping group (1992). A comprehensive genetic linkage map of the human genome. *Science* **258**, 67-86.

Nitsch,D., Boshart,M., and Shultz,G. (1993). Activation of the aminotransferase gene on synergy between liver-specific and hormone responsive elements. *Proc. Natl. Acad. Sci. USA* **90**, 5479-5483.

Nitsch,D., Boshart,M., and Shultz,G. (1994). Extinction of Aminotransferase gene activity in somatic cell hybrids involves modification and loss of several essential transcription factors. *Genes Developm.* **7**, 308-319.

Nordstrom,L.A., Dean,D.M., and Sanders,M.M. (1993). A complex array of double stranded and single stranded DNA binding-protein mediates induction of the ovalbumin gene by steroid hormones. *J. Biol. Chem.* **268**, 13193-13202.

Ochoa,A., Brunel,F., Mendelzon,D., Cohen,G.N., and Zakin,M.M (1989). Different liver nuclear proteins bind to similar DNA sequences in the 5'flanking regions of three hepatic genes.*Nucl. Acids Res.* **17**, 116-133.

Ogami,R., Kardassis,D., Cladaras,C., and Zannis,V. (1991). Purification and characterization of a heat stable nuclear factor CIIIB, involved in the regulation of the ApoCIII gene.*J. Biol. Chem.* **266**, 9640-9646.

Olds,R.J., Lane,D.A., Chowdury,V., De Stephano,L., Leone,G., and Thein, S.L. (1993). Complete nucleotide sequence of the antithrombin gene; evidence for homologous recombination causing thrombophilia. *Biochemistry* **32**, 4216-4224.

Olds, R.J. (1994). The molecular genetics of Antithrombin deficiency. *Brit. J. Haemat.* **87**, 221-226.

Pabinger,I., Scheider,B.,and the GTH study group. (1994). Thrombotic risk of women with antithrombin III, protein C,and protein S deficiency taking oral contraceptive medications.*Thromb. Haemost.* **5**, 548-552.

Pabo,C.O., and Sauer,R.T. (1992) Transcription factors; structural families and principles of DNA recognition. *Ann. Rev. Biochem.* **61**, 1053-1095.

Pang,C.P., Crossley,M., Kent,G., and Brownlee,G.G. (1990). Comparative sequence analysis of mammalian factor IX promoter. *Nucl. Acids Res.* **18**, 6731-6732.

Pani,L., Quian,X.B., Clevidence,D., and Costa,R.H. (1992). The restricted promoter activity of the liver transcription factor hepatocyte nuclear factor 3 involves a cell-specific factor and positive autoactivation. *Mol. and Cell. Biol.* **12**, 552-562.

Parker,C.S., and Topol,J. (1984). A Drosophila RNA polymerase II transcription factor contains a promoter region specific DNA binding activity. *Cell* **36**, 356-369.

Parvin,D.P., Timmers,H.T.M., and Sharp,P.A. (1992). Promoter specificity of basal transcription factors. *Cell* **68**, 1135-1144.

Paulweber,B., Sanhofer,F., and Levy-Wilson,B. (1993). The mechanism by which the human apolipoprotein B gene reducer operates involves blocking of transcriptional activation by hepatocyte nuclear factor 3. *Mol. Cell. Biol.* **13**, 1534-1546.

Perlino,E., Cortese,R., and Ciliberto,G.(1987). The human α1 antitrypsin gene istranscribed from two different promoters in macrophages and hepatocytes.*E.M.B.O. J.* **6**, 2767-2771.

Perlmulter,D.H., Glover,G.L., Rivetna,M., Schasteen,C.S., and Fallona,R.J. (1990).Identification of a serpin-enzyme complex receptor on human hepatoma cells and human monocytes. *Proc. Natl. Acad. Sci. USA*, **87**, 3753-3757.

Perry,D.J. (1995). Ectopic transcript analysis in human antithrombin deficiency. *Blood Coag. and Fibrin.* **6**, 531-536.

Petropoulos,I., Auge-Guillou,C., and Zakin,M.M (1991). Characterization of the active part of the human transferrin gene enhancer and purification of two liver nuclear factors interacting with the TGTTTGC motif present in this region. *J. Biol.Chem.* **266**, 24220-24225.

Picketts,D.J., Lillicrap,D.P., and Mueller,C.R. (1993). Synergy between transcription factor DBP and C/EBP compensates for a haemophilia B Leyden factor IX mutation. *Nature Genetics* **3**, 175-179.

Pierrou,S., Mellquist,M., Samuelson,L., Enerback,S.,and Carlsson,P. 1994).Cloning and characterization of seven human forkhead proteins; binding site specificity and DNA bending.*E.M.B.O. J.* **13**, 5002-5012.

Pinkert,C.A., Ornitz,D.M., Brinster,R.L., and Palmiter,R.D. (1987). An albumin enhancer located 10kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice. *Genes and Developm.* **1**, 268-276.

Pizzo,S.V. (1989). Serpin receptor 1; a hepatic receptor that mediates the clearance of antithrombin III-proteinases complexes. *Am. J. Med.* **87**, 10-14 Sup.

Pizzo,S.V. (1994).The physiologic role of antithrombin III as an anticoagulant. *Seminars in Hematology* **31**, 4-7.

Plutzki,J., Hoskins,J.H., Long,G.L., and Crabtree,G.R. (1986). Evolution and organization of the human protein c gene. *Proc. Natl. Acad. Sci. USA.* **83**, 546-550.

Prochownik,E.V.,Markam,A.F., and Orkin,S.H. (1983). Isolation of a cDNA clone for human antithrombin. *J. Biol. Chem.* **258**, 8389-8394.

Prochownik,E.V., and Orkin,S.H. (1984). *In-vivo* transcription of a human ATIII minigene. *J. Biol. Chem.* **259**, 15386-15392.

Prochownik,E.V., Bock,S.C., and Orkin,S.H.(1985,1). Intron structure of the human antithrombin III gene differs from that of other members of the serine protease inhibitor superfamily. *J. Biol. Chem.* **260**, 9608-9612.

Prochownik,E.V.(1985,2). Relationship between an enhancer element in the human antithrombin III gene and a immunoglogulin light chain gene enhancer. *Nature,* **316**, 845-848.

Prochownik,E.V., Smith,M.J., and Markham,A. (1987). Two regions, downstream of AATAAA in the human ATIII gene are important for cleavage polyadenylation. *J. Biol .Chem.* **262**, 9004-9010.

Ptashne,M.(1988). How Eukaryotic transcriptional activators work. *Nature* **335**, 683-689.

Ptashne,M.(1990). Activators and targets. *Nature* **346**,329-331.

Pugh,B.F., and Tjian,R.T. (1991) Transcription from a TATA less promoter requires a multisubunit TFIID complex. *Genes and Developm.* **5**, 1935-1945.

Ragg,H. and Preibish,G. (1988). Structure and expression of the gene coding for human serpin h;S2. *J. Biol. Chem.* **263**, 12129-12134.

Ramji,D.P.,Tadros,M.H., Hardon,E., and Cortese,R. (1991). The transcription factor LF-A1 interacts with a bipartite recognition sequence in the promoter region of several liver specific genes. *Nucleic Acids Res.* **19**, 1139-1146.

Rangan,V.S., and Das,G.C. (1990). Purification and biochemical characterization of HNF2 involved in liver specific transcription of the human a1 antitrypsin gene. *J. Biol. Chem.* **265**, 8874-8879.

Rastinejad,F., Perlman,T., Evans,R.M., and Sigler,P.B. (1995). Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature* **375**, 203-211.

Reijnen,M.J., Maasdam,D., Bertina,R.M.,and Reistma,P.H. (1994). Haemophilia B Leyden ; the effect of mutations at position +13 on the liver specific transcription of the factor IX gene. *Blood Coag. and Fibrin.* **5**, 341-348.

Roberts,H. (1993). Factor IX. *Thromb. Haemost.* **70**, 1-9.

Retnakaran,R., Flock,G., and Giguere,V. (1994). Identification of RVR, a novel orphan nuclear receptor that acts as a negative transcriptional regulator. *Mol. Endocrin.* **8**, 1234-1244.

Ribeiro,R.C., Kushner,P.J., Apriletti,J.W., West,B.L., and Baxre,J.D. (1992). Thyroid hormone alters *in-vitro* DNA binding of monomers and dimers of thyroid hormone receptors. *Mol. Endocrinol.* **6**, 1142-1152.

Ron,D., Habener,J.F. (1992). CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene transcription. *Genes and Developm.* **6**, 439-453.

Rosenberg,R.D., and Damus,P.S. (1973). The purification and mechanism of action of human antithrombin - heparin cofactor. *J. Biol. Chem.* **248**, 6490-6505.

Rosenberg,J.B., Bergtrom,G., and Amrani,D.L. (1995). Identification of FXP1 and LF1 regulatory sites in the 5'upstream region of the human antithrombin gene. *Thromb. and Haemostasis* **73**, 1249 (abstract communicated by Dr M. Blajchman, June 1995).

Rosenthal,N. (1987). Identification of regulatory elements of cloned genes with functional assays. *Methods in Enzymol.* **152**, 704-720.

Rosenthal,A., and Jones,D.S. (1990).Genomic walking and sequencing by oligocassette mediated polymerase chain reaction. *Nucl. Acids Res.* **18**, 3095.

Roy,A.L., Malik,S., Meisterenst,M., and Roeder,R.G. (1993). An alternative pathway for transcription initiation involving TFII I. *Nature* **365**, 355-359.

Rottman,J.N., Widom,R.L., Nadal-Ginard,B., Mahdavi,V., and Karathanassis,S.K.(1991). A retinoic acid responsive element in the apolipoprotein A1 gene distinguishes between two different retinoic acid response pathways. *Mol. Cell. Biol.* **11**, 3814-3820.

Ryffel,G.U., Kugler,W., Wagner,U., and Kaling,M. (1989). Liver specific gene transcription *in-vitro*: the promoter elements HP1, and TATA box are necessary and sufficient to generate a liver-specific promoter. *Nucl. Acids Res.* **17**, 939-953.

Saiki,R.K., Gelfand,D.H., Stoffel,S., Sharf,S.J., Higushi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988). Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**, 487-491.

Salier,J.P., Hirosawa,S., and Kurachi,K. (1990). Functional characterization of the 5' regulatory region of the human factor IX gene. (1990). *J. Biol. Chem.* **265**, 7062-7068.

Sambrook,J., Fritsch,E.T., and Maniatis,T. (1989). Molecular cloning, a laboratory manual. Second Ed. Cold Spring Harbor Laboratory Press.

Schaeffer,E., Guillou,F., Part,D., and Zakin,M. (1993). A different combination of transcription factors modulates the expression of the human transferrin promoter in liver and Sertoli cells. *J. Biol. Chem.* **31**, 23399-23408.

Schibler,U., and Sierra,F. (1987). Alternative promoters in developmental gene expression. *Ann. Rev. Genet.* **21**, 237-257.

Schreuder,H., de Boer,B., Pronk,S., Hol,W., Oijkema,R., Mulders,J., and Theunissen, H. (1993). Crystallization and preliminary X ray analysis of human ATIII. *J. Mol. Biol.* 229, 249-250.

Sheffield,W.P., Brothers,A.B., Wells,M.J., Hatton,M.W.C., Clarke,B.J., and Blajchman,M. (1992). Molecular cloning and expression of rabbit antithrombin III. *Blood* **79**, 2330-2389.

Sheffield,W.P., Wu,Y.J., and Blajchman,M.A. (1995). Antithrombin: Structure and Function. in Molecular Basis of Thrombosis and Haemostasis. K.A. High, and H.R. Roberts. Marcel Dekker, Inc. 355-377.

Shuster,D., Bachman,G., and Raschchian,A. (1992). A simple and efficient method for amplification of cDNA ends using 5'RACE. *Focus* **14**, 46-52.

Sladek.F.M., Zhong,W., Lai,E., and Darnell,J.E.(1990). Liver enriched transcription factor HNF4 is a novel member of the steroid hormone receptor superfamily. *Gene Developm.* **4**, 2353-2365.

Sladek,F.M.(1994). in *Transcriptional regulation of liver specific genes* (Tronche,F., and Yaniv, M., Eds) , R.G. Landes Company, Austin, TX, p 1-32.

Smale,S.T., and Baltimore,D. (1989). The Initiator as a transcription control element. *Cell* **57**, 103-113.

Song,C., Kokontis,J.M., Hiipaka,R.A., and Liao,S. (1994). Ubiquitous receptor: a receptor that modulates gene activation by retinoic acid and thyroid hormone receptor.*Proc. Natl. Acad. Sci. USA* **91**, 10809-10813.

Spek,C.A, Koster,T., Rosendaal,F.R., Bertina,R.M., and Reistma,P.H. (1995). Genotypic variation in the promoter region of the protein C gene is associated with plasma protein C levels and trhombotic risk. *Arteriosclerosis, Thrombosis and Vascular biology* **15**, 214-218.

Spek,C.A., Greengard,J.S., Griffin, J.H., Bertina,R.M., and Reitsma,P.H. (1995). Two mutations in the promoter region of the human protein C gene both cause type I protein C deficiency by disruption of two HNF3 binding sites. *J. Biol. Chem.* **270**, 24216-24221.

Stackhouse,R., Chandra,T., Robson,K.J.H., and Woo,S.L.C. (1983). Purification of antithrombin mRNA and cloning of its cDNA. *J. Biol. Chem.* 258, 703-708.

Subramani,S., and Southern,P.J. (1983). Analysis of gene expression using Simian virus 40 vectors. *Anal. Biochem.* **135**, 1-15.

Tabor,S., and Richardson,C.C. (1987). DNA sequence analysis with a DNA modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci USA* **84**, 4776-4761.

Thomassin,H., Hamel,D., Bernier,D., Guertin,M., and Belanger,L. (1992). Molecular cloning of two C/EBP related proteins that bind to the promoter and the enhancer of the α-foetoprotein gene. Further analysis of C/EBPβ and C/EBPγ. *Nucl. Acids Res.* **20**, 3091-3098.

Tjian,R., and Maniatis,T.(1994).Transcriptiona l activation: a complex puzzle with few easy pieces. *Cell* **77**, 5-8.

Toulon,P., Vitoux,J.F., Fiessinger,J.N., Sicard,D., and Aiach, M. (1991) Heparin cofactor II; an acute phase reactant in patients with deep vein thrombosis. *Blood Coag. and Fibrin.* **2**, 435-439.

Tran,P., Zhang,X.K., Salbert,G., Herman,T., Lehman,J.M., and Pfagl,M. (1992). COUP orphan receptors are negative regulators of retinoic acid response pathways. *Mol. Cell. Biol.* **12**, 4666-4676.

Tremp,G.L., Duchange,N., Branellec,D., Ceregheni,S., Tailleux,A., Berthou,L., Fievet,C., Touchet,N., Shombert,B., Fruchart,J.C., Zakin,M., and Denefle,P. (1995). A 700bp fragment of the human antithrombin III promoter is sufficient to confer high, tissue specific expression on human apolipoprotein a-II in transgenic mice. *Gene* **156**, 199-205.

Tripodi,A., Abott,C., Vivian,N., Cortese,R., and Badge,R.L. (1991). Disruption of the LFA1 and LFB1 biding sites in the human α1 antitrypsin gene has a differential effect during development in transgenic mice. *E.M.B.O. J.* **10**, 3177-3182.

Tsai,W., Greegard,J.S., Montgomery,R., and Griffin,J.H. (1991). Five previously undescribed mutations in the protein C gene that identify elements critical for gene expression and protein activity. *Blood* **78**, suppl. 1, 849.

Tse,W.T., and Forget,B.G. (1990). Reverse transcription and direct amplification of cellular RNA transcripts by Taq polymerase.*Gene* **88**, 293-296.

Tuddenham,E.G.D. and Cooper,D.N. (1994). The molecular genetics of haemostasis and its inherited disorders. Oxford University Press.

Twining,S. (1994). Regulation of proteolytic activity in tissues. *Critic. Rev. Bioch. and Mol. Biol.* **29**, 315-383.

Van-Zonneveld,A.J., Curriden,S.A., and Loskutoff,D.J. (1988). Type I plasminogen activator-inhibitor gene; functional analysis and glucocorticoid regulation of its promoter. *Proc. Natl. Acad. Sci.USA* **85**, 5525-5529.

Vaulont,S., and Khan, A.(1994). Transcriptional control of metabolic regulation genes by carbohydrates. *FASEB J.* **8**, 28-35.

Vinson,C.R., Sigler,P., and McKnight,S.L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. *Science,* **246**, 911-916.

Vu-Dac,N., Schoonjans,K., Laine,B., Fruchart,J.C., Auwerx,J., and Staels,B.(1994). Negative regulation of the human apolipoprotein A-1 promoter by fibrates can be attenuated by the interaction of the peroxisome proliferator activated receptor with its response element. *J. Biol. Chem.* **269**, 31012-31018.

Vu-Dac,N., Schoonjans,K., Kosykh,H., Dallongeville,J., Fruchart,J.C., Staels,B., and Auwerx, J. (1995). Fibrates increase human apolipoprotein A-II expression through activation of the peroxisome proliferator activated receptor. *J. Clin. Invest.* **96**, 741-750.

Wade,D.P., Clark,J.G., and Lindarht,G.E. (1993). 5' control regions of the apolipoprotein (a) genes and members of the related plasminogen gene family.*Proc. Natl. Acad. Sci. USA* **90**, 1369-1373.

Wang,L.H., Tsai,S.Y., Cook,R.G., Beattie,W.G., Tsai,M.J., and O'Malley,B.W. (1989). COUP transcription factor is a member of the steroid receptor superfamily. *Nature* **340**, 163-166.

Wang,N.D., Finegold,M.J., Bradley,A., Ou,C.N., Abdelsayed,S.V., Wilde,M.D., Taylor,L.R., Wilson,D.R., and Darlington,G.J. (1995). Impaired energy homeostasis in C/EBP alpha knockout mice. *Science* **269**, 1108-1112.

Wefald,F.C., Devlin,B.H., and Williams,R.S. (1990). Functional heterogeneity of mammalian TATA box sequences revealed by interaction with a cell-specific enhancer. *Nature*, **344**, 260-263.

Weis,L., and Reinberg,D. (1992). Transcription by RNA polymerase II, initiator directed formation of transcription competent complexes. *FASEB J.* **6**, 3300-3304.

Weiss,D. (1994). A novel hepatocyte transcription factor that binds the α-foetoprotein promoter linked coupling element. *Mol. and Cell. Biol.* **14**, 6616-6626.

Widom,R.L., Rhee,M., and Karathanassis,S.K. (1992). Repression by Arp-1 sensitizes apolipoprotein AI gene responsiveness to RXRα and retinoic acid. *Mol. and Cell. Biol.* **12**, 3380-3389.

Winrow,C.J., Marcus,S.L., Miyata,K.S., Zhang,B., Capone,J.P., and Rachubinski, R.A. (1994). Transactivation of the peroxisome proliferator-activated receptor is differentially modulated by hepatocyte nuclear factor 4. *Gene Expression* **4**, 53-62.

Winter,J.H., Bennet,B., and Watt,J.L.(1982). Confirmation of linkage betweenantithrombin III and Duffy blood group and assigment of AT3 to 1q22-25. *Annals Hum. Gen.* **26**, 29-34.

Winter,J.H., Fenech,A., Bennett,B., and Douglas,A.S.(1984). Prophylactic antithrombotic therapy with stanozolol in patients with familial ATIII deficiency. *Brit. J. Haemat.* **57**, 527-537.

Wu,J.K., Sheffield,W.P., and Blajchman,M.A. (1992). Molecular cloning and cell-free expression of murine ATIII. *Thromb. Haemost.* **68**, 291-296.

Xanthopoulos,K.G., Preciozo,V.R., Chen,W.S., Sladek,F.M., Cortese,R., and Darnell,J.E. (1991). The different tissue transcription patterns of genes for HNF1,C/EBP, HNF3, and HNF4 protein factors that govern liver specific transcription. *Proc. Natl. Acad. Sci. USA* **88**, 3807-3811.

Yokoyama,C., Wang,X., Brigg,M., Admon,A., Wu,J., Hua,X., Goldstein,J.L. and Brown, M.S. (1993). SREB1, a basic helix loop helix leucine zipper protein that control transcription of the LDL receptor gene. *Cell* **75**, 187-197.

Zahcoi,K., Prada,A.E., and Davis,H.E. (1994). Transcriptional regulation of the C1 inhibitor gene by γInterferon. *J. Biol.Chem.* **269**, 9669-9674.

Zakin,M.M. (1992). Regulation of transferrin gene expression. *FASEB J.* **6**, 3253-3258.

Zhang,B., Marcus,S.L., Sajjadi,F., Alvares,K, Reddy,J.K., Subramani,S., Rachubinski,R.A., and Capone,J.P. (1992). Identification of a peroxisome proliferator-responsive element uopstream of the gene encoding rat peroximal enoyl-CoA hydratase3-hydroxyl-CoA dehydrogenase. *Proc. Natl. Acad. Sci USA* **89**, 7541-7545.

Zhang,X.K., Hoffmamn,B., Tran,P.B., Graupner,G., and Pfahl,M. (1992). Retinoid X receptor is an auxiliary protein protein for thyroid hormone and retinoic acid receptors. *Nature* **355**, *441-446.*

Zhong,W., Sladek,F.M., and Darnell,J.Jr. (1993). The expression pattern of adrosophila homolog to the mouse transcription factor HNF4 suggest a determinative role in the gut formation. *E.M.B.O. J.* **12**, 537-544.

Zhou,Q., Lieberman,P.M., Boyer,T.G., and Bert,A.J. (1992). Holo TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Gene and Developm.* **6**, 1964-1974.