ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Characterisation of the potentially diverse transcriptional regulatory mechanisms of human coagulation PZ gene

Mohamed Hussain A. Ben-Hasan

A thesis submitted in partial fulfilment of the requirements
of Sheffield Hallam University
for the degree of Doctor of Philosophy

Date: May 2011

Abstract

Protein Z (PZ) is a vitamin k-dependent plasma glycoprotein produced mainly in the liver, which circulates as a complex with PZ-dependent protease inhibitor. Human PZ has been reported as having both procoagulant and anticoagulant activity, but its anticoagulant role appears to be more physiologically relevant. PZ levels vary widely among healthy individuals (normal range of 0.6 to 5.7 μg/mL) with the average being higher in men than women. The aim of this study was to investigate the mechanisms by which PZ expression is controlled at the level of transcription, including whether hormonal and/or inflammatory signals modify expression and, therefore, possibly account, at least in part, for the wide normal range.

Bioinformatic analysis enabled the identification of two age stable elements (ASEs) in the PZ gene promoter showing close homology to those found in two evolutionarily related proteins, human factor IX and protein C. These regions were found to be bound specifically by potentially the same or a similar protein; but the identity of the protein could not be confirmed as either PEA3 or Ets1, although transactivation data from HepG2 cells suggested PEA3 is important in controlling expression of endogenous PZ.

Two further regions, proximal and distal to the PZ transcription start site, were identified by bioinformatic analysis for binding of hepatocyte nuclear factor 4 alpha (HNF4α), with the proximal site also reported by Sugawara and colleagues. Although both sites showed specific binding of protein only the protein binding the proximal site was confirmed as being HNF4α. The importance of HNF4α was further demonstrated by overexpression and siRNA knockdown producing a respective increase and decrease in PZ mRNA levels.

Furthermore, the role of steroid hormones and inflammatory signals as potential modifiers of PZ transcription were also assessed in the HepG2 model system.

This study has demonstrated, to varying degrees, the importance of tissue-specific factors (HNF4α), more ubiquitous factors (PEA3, Ets1), steroid hormones (oestrogen and glucocorticoid) and inflammatory pathways (bacterial LPS induction) on the regulation of PZ gene transcription.

Abbreviations

AE	Age-regulatory element
AF	Ligand-independent activation function
AP1	Activator protein 1
APC	Activated protein C
APP	Acute phase protein
AR	Androgen receptor
ASE	Age-stable element
AT III	Antithrombin III
ATF-1	Activating transcription factor 1
Bp	Base pair
C/EBP	CCAAT-enhancer-binding proteins
Cp	Crossing point
Cq	Quantification cycle
C1q	Subcomponent 1 which recognises and binds to immunoglobulins
CREB	Cyclic AMP response element-binding protein
Ct	Cycle threshold
Cys	Cysteine
DBD	DNA-binding domain
DMSO	Dimethyl sulfoxide
dNTPs	Deoxyneocleoside triphosphates
dPZ	Distal protein Z
dsRNA	Double-stranded RNA
EGF	Epidermal growth factor-like
eIF	Eukaryotic initiation factor
EMSA	Electrophoretic mobility shift assay
Ets	E- twenty six
FII	Factor II
FIIa	Factor II activated
FV	Factor V
FVa	Factor V activated
FVi	Factor V inactivated

FVII	Factor VII
FVIIa	Factor VII activated
FVIIIa	Factor VIII activated
FVIIIi	Factor VIII inactivated
FIX	Factor IX
FIXa	Factor IX activated
FX	Factor X
FXa	Factor X activated
FXIIa	Factor XII activated
FXIIIa	Factor XIII activated
FDU	1 unit of the FastDigest enzyme
Gla	Gama-carboxyglutamic acid domain
GR	Glucocorticoid receptor
GRE	Glucocorticoid response element
GTP	Guanosine 5'-triphosphate
HepG2	Human liver carcinoma cell line
hFIX	Human factor IX
HNF4	Hepatocyte nuclear factor 4
hPC	Human protein C
HRE	Hormone responsive elements
IκB	kappa light polypeptide gene enhancer in B-cells inhibitor
IL	Interleukin
Kb	Kilo base pair
LPS	Lipopolysaccharide
miRNA	Micro-RNA
mMol	Millimole
mRNA	Messenger RNA
NCBI	National Centre of Biotechnology Information
N-CoR	Nuclear receptor corepressor
NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
nt	Nucleotides
ORE	Oestrogen response element
PC	Protein C
PCR	Polymerase chain reaction

PEA3	polyomavirus enhancer activator 3
PKA	protein kinase A
pMol	Pico mole
Pol II	RNA polymerase II
POU	Pituitary-specific Octamer neural Unc-86
pPZ	proximal protein Z
PS	Protein S
PZ	Protein Z
qPCR	quantitative real-time polymerase chain reaction
RISC	RNA-induced silencing complex
RNAi	RNA interference
rRNA	ribosomal RNA
siRNA	small interfering RNA
SMRT	silencing mediator of retinoid and thyroid receptor
SNPs	Single-nucleotide polymorphism
SP	Serine protease-like
TAFI	Thrombin-activatable fibrinolysis inhibitor
TBE	Tris-borate EDTA
TF	tissue factor
TFII	general transcription factor II
TFPI	Tissue factor pathway inhibitor
TOP	take-off point
TNFα	Tumor necrosis factor alpha
TRAF	Tissue necrosis factor receptor-associated factor
UBC	Ubiquitin C
UTR	Untranslated region
VKD	vitamin-k dependent
vWF	von Willebrand factor
ZPI	Protein Z-dependent protease inhibitor

Acknowledgments

I am heartily thankful to my supervisors, Dr Adrian Hall, Dr Peter Winship and Dr Caroline Dalton, whose encouragement, guidance and support enabled me to develop an understanding of the subject and completing of the project. I have very much appreciated this throughout my study.

I would also like to express my gratitude to Dr Christine Le Maitre and Dr Neil Cross for their advice in biostatistics of this thesis and to Dr Malcom Lock, Dr Verena Kriechbaumer, BMRC students and post-docs who supported me in many respects during my PhD project.

Lastly, I would like to express my love to my mother and my wife for their encouragement and support and my children who continue to enrich my life.

Meeting abstracts

Ben-Hasan, M., et al. (2007). Characterisation of ASE-like elements in the human protein Z gene promoter. *Journal of Thrombosis Haemostasis* ; **5** Supplement 2: P-S-107

Ben-Hasan, M., et al. (2009). Transcriptional regulation of protein Z. *Journal of Thrombosis and Haemostasis* ; **7**, Supplement 2: Abstract PP-WE-161

Contents

Abstract I
Abbreviations II
Acknowledgments V
Meeting abstracts VI
List of figures XI
List of tables XVII

Chapter 1-Introduction 1

1.0. Introduction 2
1.1. The Coagulation Process 2
1.2. Protein Z 6
1.2.1. Function of Protein Z 6
1.2.2. Structure of Protein Z 8
1.2.4. Protein Z levels 10
1.3. The regulation of gene expression 10
1.3.1. Basal promoter element 11
1.3.2. Transcription factors 12
1.3.2.1. Structure of transcription factors 13
1.3.2.2. Function of transcription factors 14
1.3.3. Tissue-specific regulation 15
1.3.4. Inducible regulation 16
1.3.4.1. Steroid hormone receptors 16
1.3.4.2. Inflammatory response 18
1.3.5. Role of chromatin structure in controlling expression 20
1.3.6. Post-transcriptional regulation via the RNAi pathway 23
1.4. Aim of the study 25

Chapter 2-Materials and Methods 26

2.1. DNA sequencing 27
2.1.1. PCR of sequencing template 28
2.1.2. Preparation of the 1% agarose gel 29
2.1.3. Purification of amplicons 29
2.2. Protein-DNA binding studies 30
2.2.1. Nuclear extract preparation 30
2.2.2. Bradford Protein Assay 30
2.2.3. Double stranded DNA Probe preparation (Fluorescent and biotin labelled) 32
2.2.4.1. Electrophoretic Mobility Shift Assay (EMSA) 33
2.2.4.2. EMSA confirming the identity of the bound protein 34
2.3. Gene expression studies 34
2.3.1. RNA preparation from HepG2 cells 34
2.3.2. Analysis of RNA by agarose gel electrophoresis 35
2.3.3. cDNA synthesis 35
2.3.4. Real-time PCR 36
2.3.4.1. Primer design and efficiency 37
2.3.4.2. Validation of reference gene 38

2.3.4.3. Different chemistries in real time PCR....39
2.4. Cell culture, validation, transfection and stimulation....42
2.4.1. Culturing of HepG2 cells....42
2.4.2. Determination of relevant protein expression levels in HepG2 cells....44
2.4.3. Preparation of transcription factor expression plasmids for transactivation studies....44
2.4.3.1. Transformation of E. coli competent cells....44
2.4.3.2. Growing up and purification of expression plasmids....45
2.4.3.3. Digestion of plasmids by restriction enzymes....46
2.4.4. Transfection of HepG2 cells by expression vectors....46
2.4.5. Transfection of HepG2 with siRNAs....47
2.4.6. Stimulation of HepG2 with Lipopolysaccharide....48

Chapter 3-Optimisation Of Electrophoretic Mobility Shift Assay....49
3.1. Introduction....50
3.2. Optimisation of EMSAs....51
3.3. Materials and Methods....54
3.3.1. Nuclear extract Preparation....54
3.3.2. EMSA Probe Preparation....56
3.3.3. EMSA Binding Reactions and Size Fractionation....56
3.3.5. IRDye Detection....59
3.3.6. EMSA analysis....60
3.4. Results....61
3.4.1. Biotinylated Probe EMSA....61
3.4.1.1. Effect of nuclear extract, longer incubation and probe amount on protein-DNA binding....62
3.4.1.2. Effect of binding buffer on protein-DNA binding....64
3.4.1.3. Effect of temperature and incubation period on antibody-protein binding....67
3.4.1.4. Optimisation of supershift assay using PEA3 binding site consensus....72
3.4.2. Use of IRDye labelled Probes....74
3.4.2.1. Optimisation of antibody-protein binding using CREB binding site consensus....74
3.4.2.2. Effect of temperature and incubation period on antibody-protein binding using CREB binding consensus....75
3.4.2.3. Effect of new protocol with different buffers on CREB supershift..77
3.4.2.4. EMSA and supershift assays for HNF4α....78
3.4.2.5. Effect of temperature and incubation period on HNF4α supershift80
3.4.2.6. Effect of new binding buffer on EMSA using HNF4α binding site consensus....83
3.5. Discussion....85

Chapter 4-Liver specific transcriptional regulation of Protein Z....94
4.1. Introduction....95
4.1.1. Liver specificity....95
4.1.2. Hepatocyte nuclear factor 4....96
4.1.3. HNF4α binding to the PZ gene promoter....97
4.2. Aims....99

4.3. Methods **100**
4.3.1. Protein-DNA binding analysis 100
4.3.1.1. Electrophoretic Mobility Shift Assay (EMSA) 100
4.3.1.2. Supershift assays 100
4.3.2. Gene expression studies of the endogenous PZ gene in HepG2 cells 101
4.3.2.1. Comparison of PZ gene promoter sequences in HepG2 cells with the wild type promoter. 101
4.3.2.2. Expression plasmid verification 101
4.3.2.3. Extraction of RNA 102
4.3.2.4. Primer efficiency 102
4.3.2.5. Reference gene Validation 102
4.3.2.6. The effect of HNF4α overexpression in HepG2 cells on endogenous PZ transcript levels 103
4.3.2.7. The effect of HNF4α Knock down on the levels of endogenous PZ expression 103
4.3.2.8. Comparison of the relative HNF4α mRNA expression in HepG2 to normal liver tissue. 104
4.3.2.9. Data analysis 104
4.3. Results **106**
4.3.1. Protein binding to putative HNF4α sites in the PZ promoter 106
4.3.2. Confirmation of the identity of the protein bound to the proximal HNF4α site 108
4.3.3. Gene expression studies of the endogenous PZ gene in HepG2 cells 109
4.3.3.1. Sequencing of PZ gene promoter 109
4.3.3.2. Expression plasmid verification 110
4.3.3.3. RNA analyses 110
4.3.3.4. Efficiency of primers 111
4.3.3.5. Reference gene Validation 112
4.3.3.6. Relative PZ mRNA expression in HNF4α overexpressed HepG2 cells. 113
4.3.3.7. Knockdown of HNF4α using small interfering RNAs 115
4.3.3.8. Expression of PZ and HNF4α in HepG2 compared to normal human liver 117
4.4. Discussion **118**
4.5. Conclusion **123**

Chapter 5-The role of transcription factors PEA3, Ets 1 and CREB in regulating protein Z expression **125**
5.1. Introduction **126**
5.1.1. E twenty six transformation-specific (Ets) family proteins 128
5.1.2. The role of Ets family in the coagulation process 132
5.1.3. The cAMP response element-binding protein (CREB) 133
5.2. Aims **135**
5.3. Methods **136**
5.3.1. Protein-DNA binding analysis 136
5.3.1.1. Electrophoretic Mobility Shift Assay (EMSA) 136
5.3.1.2. Supershift assays 136
5.3.2. Gene expression studies of the endogenous PZ gene in HepG2 cells 137

5.3.2.1. Comparison of PZ gene promoter sequences in HepG2 cells with the wild type promoter. 137
5.3.2.2. Expression plasmid verification....137
5.3.2.3. Extraction of RNA 137
5.3.2.4. Primer efficiency 138
5.3.2.5. The effect of overexpression of PEA3 on the transcription levels of the endogenous PZ gene in HepG2 cells 138
5.3.2.6. The effect of PEA3 knock down on the expression of the endogenous PZ gene in HepG2 cells 138
5.3.2.7. Comparison of the relative PEA3 mRNA expression in HepG2 to normal liver tissue....138
5.3.2.8. Data analysis 139
5.3. Results....140
5.3.1. Protein binding to the putative age-stable elements and other potential PEA3 binding sites in the PZ promoter....140
5.3.2. Identification of the protein bound to the pPZ ASE and dPZ ASE 145
5.3.3. Gene expression studies of the endogenous PZ gene in HepG2 cells 148
5.3.3.1. Sequencing of PZ gene promoter....148
5.3.3.2. Verification of the PEA3 expression plasmid 148
5.3.3.3. RNA analysis 149
5.3.3.4. Reference gene validation 150
5.3.3.5. Efficiency of primers 150
5.3.3.6. Transactivation of PZ gene by PEA3 overexpression in HepG2 cells 151
5.3.3.7. The effect of PEA3 knockdown on PZ mRNA expression 151
5.3.3.8. The effect of CREB knockdown on PZ mRNA expression....153
5.3.3.9. Expression of PEA3 and CREB in HepG2 compared to normal human liver 155
5.4. Discussion....156
5.5. Conclusion 161

Chapter 6-Optimisation Of Footprinting and Chromatin Immunoprecipitation Assays....163
6.0. Introduction....164
6.1. Methods 170
6.1.1. Footprinting assay 170
6.1.1.1. Binding reaction using biotin-labelled primers....172
6.1.1.2. Binding reaction using fluorescence-labelled primers....173
6.1.2. Chromatin immunoprecipitation (CHIP) assay 174
6.1.3. Analysis....176
6.2. Results....177
6.2.1. Footprinting with biotinylated primers....177
6.2.2. Footprinting with IRD primers....179
6.2.3. Chromatin Immunoprecipitation....189
6.2.3.1. Optimisation of GAPDH primers for PCR....189
6.2.3.2. Optimisation of PZ -13/-163 region primers 189
6.2.3.3. PCR of the PZ -13/-163 with the immunoprecipitated chromatin. 193
6.3. Discussion....198
6.3.1. Footprinting assay using biotin-labelled primers....199

6.3.2. Footprinting assay using IRDye-labelled primers 200
6.3.3. Modified footprinting assay using IRDye-labelled primers 201
6.3.4. Chromatin immunoprecipitation assay 204

Chapter 7-Hormonal regulation and inflammatory response of the protein Z gene 207
7.1. Introduction 208
7.1.1. Steroid hormones 208
7.1.2. The role of steroid hormones in the regulation of protein Z and other coagulation factors 210
7.1.3. Inflammatory response 211
7.1.4. The effect of lipopolysaccharide on protein Z levels 212
7.2. Aims 213
7.3. Methods 214
7.3.1. Gene expression studies of the endogenous PZ gene in HepG2 cells 214
7.3.1.1. Verification of the expression plasmids 214
7.3.1.2. RNA analyses 214
7.3.1.3. Primer efficiency 215
7.3.1.4. Comparison of the HepG2 relative expression of steroid hormone receptors to normal liver tissue 215
7.3.1.5. The effect of steroid hormones on the endogenous PZ gene 215
7.3.1.6. The effect of LPS stimulation on PZ transcription 216
7.4. Results 217
7.4.1. Verification of expression plasmids 217
7.4.2. RNA analysis 218
7.4.3. Reference gene validation 218
7.4.4. Primer efficiency 219
7.4.5. Expression of AR, GR and OR in HepG2 compared to normal liver 220
7.4.6. The effect of androgen, glucocorticoid and oestrogen hormones on PZ transcription 221
7.4.7. The effect of LPS stimulation on PZ transcription 222
7.5. Discussion 223
7.6. Conclusion 225

Chapter 8-Discussion and future work 227
8.1. Discussion 228
8.1.1. The role of HNF4 in regulation of PZ transcription 228
8.1.2. The role of PEA3 and Ets 1 in the PZ regulation 231
8.1.3. Hormonal regulation and inflammatory response of the PZ gene 235
8.2. Conclusion 242
8.3. Future work 243
References 245
Appendix A 279
Appendix B 284
Appendix C 290

List of figures

Fig 1.1. The coagulation cascade model..... 3
Fig 1.2. The cell-based model of haemostasis..... 4
Fig 1.3. Modular organisation of evolutionarily conserved coagulation Proteins..... 9
Fig 1.4. Organisation of the human Protein Z gene..... 9
Fig 1.5. Organisation of the nuclear receptor structure.....18
Fig 1.6. The pathway leading to the translocation of NF-κb.....20
Fig 1.7. The basic structure of chromatin..... 21
Fig 1.8. The siRNA structure and the RNAi pathway..... 24
Fig 2.1. Standard curve for measuring the concentration of the HepG2 nuclear extract..... 31
Fig 2.2. Real time PCR amplification curve showing the four major Phases.....36
Fig 2.3. Real time quantitative PCR cycles using SYBR green I dye.....40
Fig 2.4. Melt curve of PZ and UBC with transfected and non transfected samples.....41
Fig 2.5. Real time PCR steps using TaqMan chemistry.....42
Fig 2.6. HepG2 cells after 5 days from passage..... 43
Fig 2.7. Physical Maps for the expression plasmids used in HNF4α and PEA3 transfection.....45
Fig 3.1. Comparison of the sensitivity of various EMSA Methods..... 52
Figure 3.2. IRDye® 700 EMSA..... 53
Fig 3.3. EMSA using pPZ and dPZ ASEs.....61
Fig 3.4. Competitive EMSA using Bio-pPZ ASE probe..... 62
Fig 3.5. EMSA using different Biotin-labelled ASEs and PEA3 consensus probes.....63
Fig 3.6. Competitive EMSA using Bio-dPZ ASE probe..... 63
Fig 3.7. EMSA using Bio-pPZ ASE with different binding buffers.....64
Fig 3.8. Competitive EMSA using the putative PZ proximal ASE binding site.....65
Fig 3.9. Competitive EMSA using the putative PZ distal ASE binding site.... 66
Fig 3.10. Supershift assay using biotinylated pPZ ASE binding sites.....67

Fig 3.11. Supershift assay at 4°C using biotinylated pPZ ASE binding sites..... 68
Fig 3.12. Supershift assay at 4°C using biotinylated pPZ ASE binding sites. 69
Fig 3.13. Supershift assay on ice using biotinylated pPZ ASE binding sites with Ets1.....70
Fig 3.14. Supershift assay at room temperature and 37°C using biotinylated dPZ ASE binding sites with PEA3 and Ets1.....71
Fig 3.15. Supershift assay at 4°C using biotinylated dPZ ASE binding sites with Ets1..... 72
Fig 3.16. Supershift assay using biotinylated PEA3 binding sites consensus with PEA3 antibody.....73
Fig 3.17. Supershift assay using biotinylated PEA3 binding sites consensus with PEA3 antibody and dialysed nuclear extract.....73
Fig 3.18. Supershift assay at room temperature using fluorescent pPZ ASE and dPZ ASE probes with Ets1..... 74
Fig 3.19. Supershift assay at room temperature using fluorescent CREB probe with CREB antibody.....75
Fig 3.20. Supershift assay at 4°C using fluorescent CREB probe with CREB antibody..... 76
Fig 3.21. Supershift assay at 37°C using fluorescent CREB probe with CREB antibody.....76
Fig 3.22. Supershift assay of the fluorescent CREB probe with CREB antibody using a new protocol..... 77
Fig 3.23. Competitive EMSA using the reported pPZHNF4α-fluorescent binding site.....78
Fig 3.24. Supershift assay and competitive EMSA using the pPZHNF4α-fluorescent binding site..... 79
Fig 3.25. Supershift assay using the consensus HNF4α-fluorescent binding site.....80
Fig 3.26. Supershift assay using the consensus HNF4α-fluorescent binding site at different conditions 81
Fig 3.27. Supershift assay using the consensus HNF4α-fluorescent binding site at 4°C.....82

Fig 3.28. Supershift assay using the consensus HNF4α-fluorescent binding site at 4°C.....82
Fig 3.29. EMSA assay using the consensus HNF4α-fluorescent binding Site.....83
Fig 3.30. Supershift assay using the consensus HNF4α-fluorescent binding site.....84
Fig 3.31. Supershift assay using the pPZHNF4α-fluorescent binding site.... 84
Figure 4.1. The previously reported (pPZHNF4α) and hypothetical (dPZHNF4α) HNF4α binding sites in the PZ gene promoter.....98
Fig 4.2. EMSA using the reported pPZHNF4α-fluorescent (cy5.5) binding site.....106
Fig 4.3. EMSA using the putative PZ distal HNF4α-fluorescent (IRD-700) binding site.....107
Fig 4.4. Supershift assay using the pPZHNF4α-fluorescent (cy5.5) binding site.....108
Fig 4.5. Supershift assay using the consensus HNF4α-fluorescent binding site..... 109
Fig 4.6. HNF4α encoded plasmid digested with restriction enzymes..... 110
Fig 4.7. Ribosomal RNA 28S and 18S bands on 1% w/v agaros gel..... 111
Fig 4.8. Determination of real time PCR efficiencies..... 112
Fig 4.9. Comparision of relative expression of some housekeeping genes for the stability..... 112
Fig 4.10. Amplification plot of PZ and UBC with transfected and non transfected cells.....113
Fig 4.11. The effect of HNF4α overexpression on PZ mRNA relative expression..... 114
Fig 4.12. The effect of mock transfection on PZ mRNA relative expression 114
Fig 4.13. GAPDH knockdown in HepG2 cells using siRNAs..... 115
Fig 4.14. HNF4α knockdown in HepG2 cells using siRNAs..... 116
Fig 4.15. The effect of HNF4α Knockdown on PZ mRNA expression.....116
Fig 4.16. The relative expression of HNF4α and PZ mRNA in HepG2 and normal liver tissue..... 117
Fig 5.1. Ribbon diagram shows the Ets family member Elk-1 binding to DNA.....129

Fig 5.2. The hypothetical PEA3 and CREB binding sites in the PZ gene promoter.....134
Fig 5.3. EMSA using the putative biotinylated PZ proximal ASE (pPZ ASE) binding site..... 140
Fig 5.4. IRDye pPZ ASE (-120/-101) competitive EMSA.....141
Fig 5.5. EMSA using the putative biotinylated PZ distal ASE (dPZ ASE) binding site..... 142
Fig 5.6. PZ-PEA3(1)-fluorescent (IRDye) competitive EMSA..... 143
Fig 5.7. PZ-PEA3(2)-fluorescent (IRDye) competitive EMSA..... 144
Fig 5.8. Supershift assay using biotinylated pPZ ASE binding sites.....145
Fig 5.9. Supershift assay on ice and room temperature using biotinylated pPZ ASE binding sites with Ets1..... 146
Fig 5.10. Supershift assay at room temperature using fluorescent CREB probe with CREB antibody..... 147
Fig 5.11. Supershift assay at room temperature using biotinylated dPZ ASE binding sites with PEA3 and Ets1..... 148
Fig 5.12. PEA3-encoded plasmid digested with restriction enzymes..... 149
Fig 5.13. Analysis of total RNA from transfected and untransfected HepG2.....150
Fig 5.14. The effect of PEA3 overexpression on PZ mRNA relative expression.....151
Fig 5.15. Determination of primers efficiencies used in PEA3 Knockdown analyses.....152
Fig 5.16. PEA3 knockdown in HepG2 cells using siRNAs.....152
Fig 5.17. The effect of PEA3 Knockdown on PZ mRNA relative expression.....153
Fig 5.18. CREB knockdown in HepG2 cells using siRNAs.....154
Fig 5.19. The effect of CREB Knockdown on PZ mRNA relative expression.....154
Fig 5.20. The realtive expression of CREB and PEA3 mRNA in HepG2 and normal liver cells..... 155
Fig 6.1. Schematic representation of the footprinting assay.....165
Fig 6.2. Schematic representation of various techniques of chromatin immunoprecipitation.....168

Fig 6.3. Location of the PZ gene promoter regions amplified by the designed primers in footprinting assay.....170
Fig 6.4. PCR of different regions of PZ gene promoter used for footprinting analysis..... 177
Fig 6.5. Footprinting of the Protein Z promoter region +86 to -213..... 178
Fig 6.6. Footprinting of the PZ region +86 to -213 with incubation at room temperature..... 179
Fig 6.7. Footprinting assay of the PZ region +86 to -213 without amplification using IRD-labelled fragment..... 180
Fig 6.8. Footprinting assayof the PZ region +86 to -213 post amplified by the 700-IRD primer.....181
Fig 6.9. Footprinting assay ofr the PZ region +86 to -213 post amplified by the 800-IRD primer..... 181
Fig 6.10. Footprinting assay of the PZ region +86 to -213 post amplified by the 700-IRD primer and with and without a 1:10 dilution..... 182
Fig 6.11. Footprinting assay of the PZ region +86 to -213 post amplified by the 800-IRD primer with and without 1:10 dilution.....183
Fig 6.12. Footprinting assay of the PZ region +86 to -213 post amplified by 700-IRD primer and 1:5 dilution..... 184
Fig 6.13. Footprinting assay of the PZ region +86 to -213 post amplified by the 800-IRD primer and 1:2 dilution..... 184
Fig 6.14. Footprinting assay of the PZ region +86 to -213 post amplified by the 700-IRD primer and 1:10 dilution.....185
Fig 6.15. Footprinting assay of the PZ region +86 to -213 post amplified by 800-IRD primer and 1:2 dilution..... 186
Fig 6.16. Footprinting assay using different binding buffers for the PZ region +86 to -213..... 187
Fig 6.17. Footprinting assay using different binding buffers and different incubation conditions..... 188
Fig 6.18. Optimisation of the GAPDH PCR with HepG2 genomic DNA..... 190
Fig 6.19. PCR reaction of PZ -13/-163 with input DNA (annealing temperature of 59°C)..... 191
Fig 6,20. PCR reaction of PZ -13/-163 with input DNA (annealing temperature of 61°C)..... 191

Fig 6.21. PCR reaction of PZ -13/-163 with input DNA (62°C)..... 192
Fig 6.22. PCR reaction of PZ -13/-163 with input DNA (63°C)..... 192
Fig 6.23. PCR of PZ -13/-163 with input DNA (62.5°C)..... 192
Fig 6.24. PCR of PZ -13/-163 with different immunoprecipitated chromatins..... 193
Fig 6.25. PCR of PZ -13/-163 with different immunoprecipitated chromatins..... 194
Fig 6.26. PCR of PZ -13/-163 with different immunoprecipitated chromatins..... 195
Fig 6.27. PCR of PZ -13/-163 with different immunoprecipitated chromatins..... 196
Fig 6.28. PCR of PZ -13/-163 with different immunoprecipitated chromatins..... 197
Fig 7.1. Restriction enzyme digestion of plasmids encoding steroid hormone receptors..... 217
Fig 7.2. Agarose gel (1% w/v) electrophoresis of total RNA prepared from control and expression plasmid transfected HepG2 cells..... 218
Fig 7.3. Relative expression of housekeeping genes with transfection..... 219
Fig 7.4. Determination of the YWHAZ, AR, GR and OR primers efficiency. 220
Fig 7.5. The relative expression of steroid hormone receptors in HepG2 cells and normal liver..... 220
Fig 7.6. The effect of steroid hormones on PZ mRNA levels in HepG2 cells..... 221
Fig 7.7. The effect of LPS on PZ mRNA levels in HepG2 cells..... 222
Fig 8.1. Transcription factors and proteins binding sites in the human protein Z gene promoter..... 239

List of tables

Table 2.1. Oligonucleotide primers used for PZ promoter sequencing.....27
Table 2.2 Annealing temperatures and $MgCl_2$ concentration used to amplify different PZ promoter regions for sequencing..... 28
Table 2.3. Fluorescent labeled and unlabeled oligonucleotides used in EMSA.....32
Table 2.4. Oligonucleotide primers used in reference gene validation and SYBR Green chemistry quantitative PCR..... 38
Table 3.1. Different variables used in optimisation of EMSA..... 58
Table 5.1. Haemostasis measurements for young, older and centenarians.126
Table 6.1. Oligonucleotide primers used to amplify different PZ promoter regions for footprinting.....171
Table 6.2. PCR template and cocktail used in amplifying different promoter regions.....171
Table 6.3. Different variables used in optimisation of footprinting assay.....199

Chapter 1-Introduction

1.0. Introduction

The haemostatic system is responsible for the prevention of bleeding from injured blood vessels. This system must be controlled to ensure normal blood flow by preventing blood coagulation within the vasculature (Smith 2009). Haemostasis can be considered as an elegant balance between procoagulant and anticoagulant activities, and any imbalance in one of these two forces may cause either bleeding or thrombosis, respectively (Aird 2003). Interactions of the three components of haemostasis, coagulation factors, platelets and blood vessels, are crucial in order to maintain the dynamic balance between the essential processes of clotting and bleeding (Riddel and Aouizerat 2006). Cross-talk between the coagulation system and the complement system, the innate immune system, is recognised as the first line of immune response against invaders and potentially injurious stimuli (Choi *et al.* 2006). Both systems are activated early after injury, and the corruption of their cross-talk may lead to haemorrhagic shock (Amara 2008). The two cascades share some serine protease activity, such as the activation of complement 1 subcomponent (C1q) by factor XIIa (FXIIa); C1q, in turn, leads to activation of the complement classical pathway. Furthermore, C1 esterase inhibitor inhibits the complement pathways as well as inhibiting kallikrein and factor XIIa (Davis 2004).

1.1. The Coagulation Process

Blood coagulation is a process to halt blood loss from a damaged vessel. Platelet adherence and aggregation at sites of vessel wall injury form the primary hemostatic plug . Plasma coagulation factors are then activated at the site of vessel injury and lead to the formation of a fibrin clot (Riddel *et al.* 2007). Upon healing, the fibrin deposits at the former site of injury are degraded and broken down by the serine protease plasmin.

The traditional coagulation cascade describes a sequence of steps involving the activation of coagulation factors. These clotting proteins circulate as inactive zymogens which are sequentially activated upon clotting initiation. The coagulation process consists of two pathways; the extrinsic, which involves the activation of factor X to factor Xa by a tissue factor: factor VII complex (TF:

FVIIa complex); and the intrinsic, that results in the activation of factor X via the contact activation pathway (Figure 1.1). Factor Xa, with its cofactor factor Va,

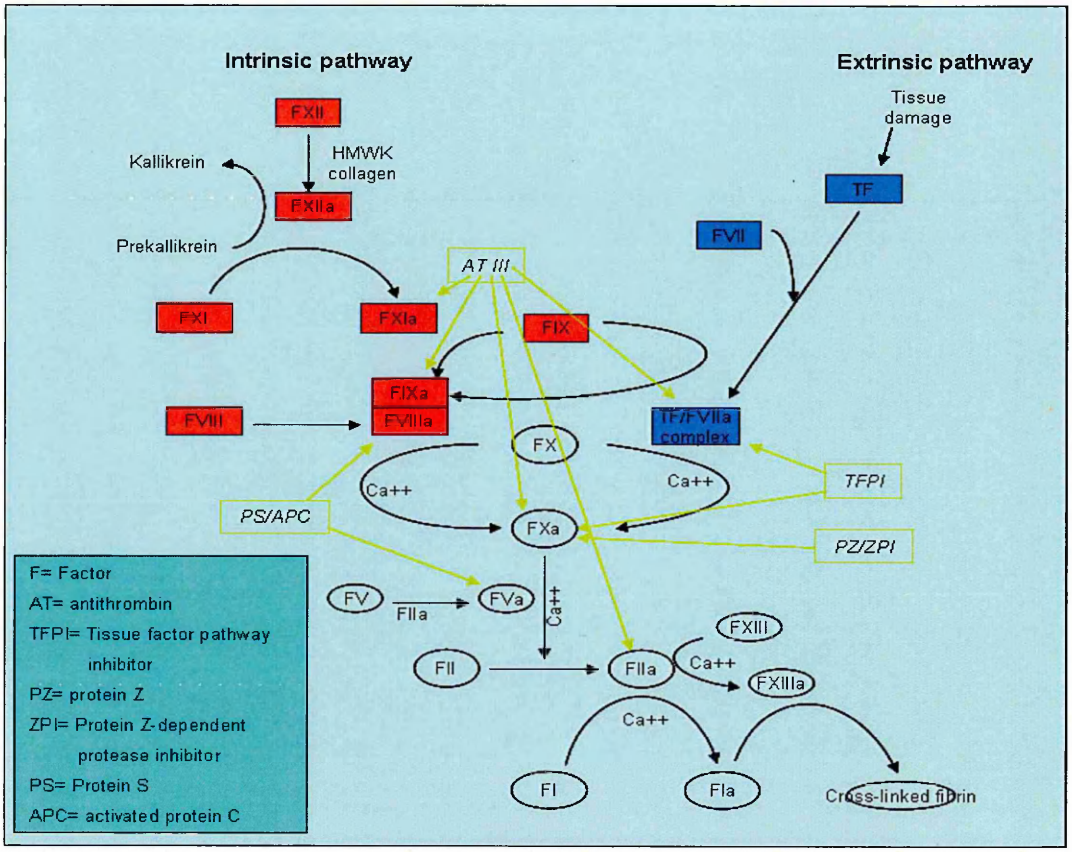


Figure 1.1. The coagulation cascade model adopted from Riddel *et al.* 2007
Both the extrinsic and the intrinsic pathways culminate in the activation of factor X with subsequent activation of thrombin to form the stabilized fibrin clot. Red color shows the intrinsic pathway. Blue color shows the extrinsic pathway. Green colour shows the anticoagulant action of a number of proteins.

then converts prothrombin to thrombin which in turn converts fibrinogen to fibrin (Davie and Ratnoff 1964, Macfarlane 1964). Tissue factor is expressed in many cell types and its deficiency has never been identified in humans (Mackman, Tilley and Key 2007, Bugge *et al.* 1996). This, and the fact that tissue factor knockout mice die in utero, are consistent with the vital role of this factor in haemostasis (Bugge *et al.* 1996).

The coagulation cascade model has failed to answer many questions *in vivo*. The flaws in this traditional cascade led to the remodeling of the coagulation cascade by Maureane Hoffman and her group to produce the cell-based model of coagulation. This model requires the participation of two types of cells; tissue factor-bearing cells (TF-bearing), where the procoagulant reactions are initiated,

and platelets for localization of the response. The coagulation process remains inactivated unless these two cells come into contact, after injury (Hoffman 2003). Three overlapping steps have been suggested; initiation, amplification and propagation (Fig 1.2). In the initiation step, a small amount of factor X is activated by the factor VIIa/TF complex to form factor Xa which with its cofactor factor Va forms the prothrombinase complex on the TF-bearing cell. Factor Xa activity is essentially localised to the TF-bearing cell under the control of the anticoagulant factors, tissue factor pathway inhibitor (TFPI) and antithrombin III (AT III), since its dissociation from this cell would result in its inhibition in the fluid phase by these anticogulants (Hoffman 2003). Factor Xa which escapes from the membrane surface may also be inhibited by the ZPI-PZ complex, that activated by heparin or heparin-like glycosaminoglycans at the site of injury (Huang *et al.* 2011).

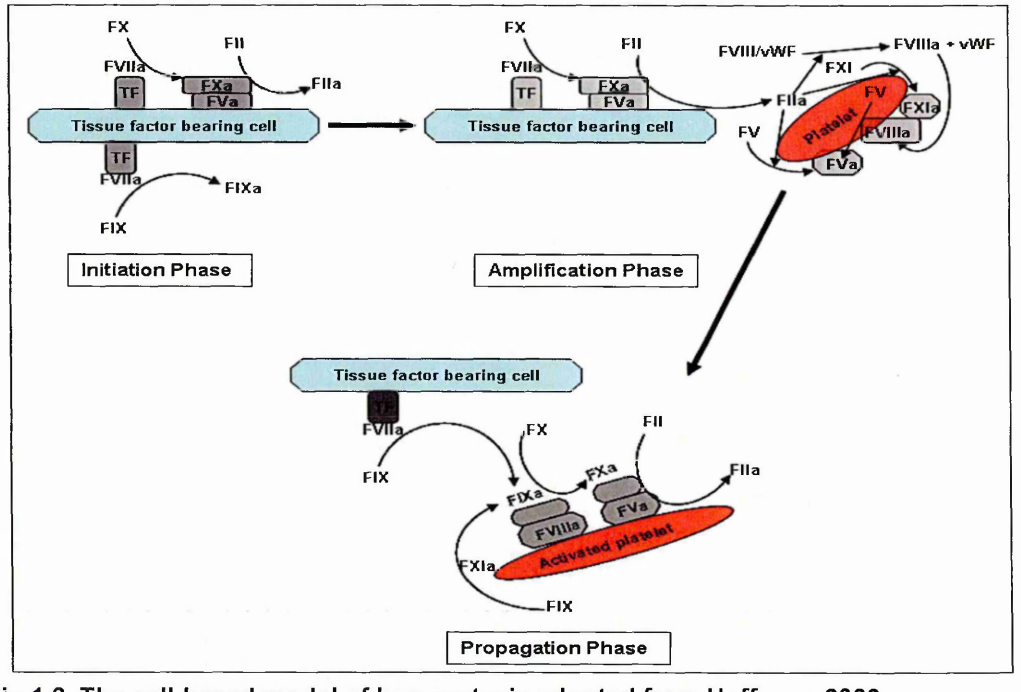


Fig 1.2. The cell-based model of haemostasis adopted from Hoffman 2003.
The initiation phase (a) where FXa is localized on the tissue factor-bearing cell and the prothrombinase complex is formed. The amplification phase (b) where platelets are activated exposing binding sites for activated factors, and cofactors V and VIII are activated on the platelets. Platelet aggregation is mediated by vWF at the site of vascular injury. The propagation phase(c) where the tenase and prothrombinase complexes are formed on the activated platelets.

The amplification phase takes place by the activation of the platelets which escape from the blood vessels and aggregate at the site of injury by thrombin

which has been generated and diffused from the TF-bearing cell. This thrombin, which has been generated in the initiation phase, also causes activation of factor XI (FXI) and factor V (FV) in addition to the activation of factor VIII which leads to its dissociation from von Willebrand factor (vWF). The dissociated vWF mediates more aggregation and adhesion of platelets while FVIII is further activated by thrombin (Smith 2009). The propagation phase occurs effectively only on an activated platelet surface. In this process factor IXa, activated in the initiation phase and from the action of factor XIa, binds to factor VIIIa to form the tenase complex, which then activates factor X. Factor Xa binds to factor Va to form the prothrombinase complex which leads to generation of sufficient thrombin for the conversion of fibrinogen to fibrin and covalent crosslinking by factor XIIIa (also activated by thrombin) (Hoffman 2003).

In order to maintain blood haemostasis, anticoagulant proteins such as antithrombin III (AT III), protein C (PC), protein S (PS) and tissue factor pathway inhibitor (TFPI) act to inhibit the coagulation process (Becker 2005).

AT III is a serpin, a serine protease inhibitor, essential in the inhibition of serine proteases with procoagulant activity in the intrinsic pathway of blood coagulation (Olson *et al.* 1992). This inhibitor is bound by heparin to increase its inhibition activity (Ersdal-Badju *et al.* 1997). In addition to the inactivation of thrombin, AT III inhibits other factors in the coagulation process, such as factors IXa, Xa, XIa, XIIa, plasmin and kallikrein (Fourrier *et al.* 1992). The protein C (PC) anticoagulant pathway involves thrombin bound to thrombomodulin which activates PC. Activated protein C (APC) is a vitamin K-dependent serine protease that cleaves factor Va and factor VIIIa to factor Vi and factor VIIIi, respectively (Esmon 2003). APC also promotes fibrinolysis by neutralisation of plasmin activator inhibitor-1 (PAI-1) and interfering with thrombin-activatable fibrinolysis inhibitor (TAFI) which inhibits the generation of plasmin (Ravindranath *et al.* 2007, Amara *et al.* 2008, Grailhe *et al.* 1993). The APC blocks the inhibition of the tissue plasminogen activator by forming a tight complex with PAI-1 (Esmon 2001). Moreover, FVa inactivation by APC leads to the restriction of thrombin formation and subsequently limiting TAFI activation (Bajzar, Nesheim and Tracy 1996). Protein S also plays an important anticoagulant role as it enhances the proteolytic inactivation of coagulation

factors Va and VIIIa to FVi and FVIIIi by APC (Persson *et al.* 2006) (see fig.1.1). TFPI, which is produced in the endothelial cells and stored in platelets, is responsible for the inhibition of FXa which is essential for the activation of prothrombin to thrombin and FIXa by preventing the formation TF-FVIIa complex (Ravindranath *et al.* 2007, Broze 2003).

The study of the protein motifs of gamma-carboxylated glutamic acid-epidermal growth factor-like-EGF2- serine proteases (Gla-EGF1-EGF2-SP) in factors VII, IX, X and protein C (PC), as well as the A domain in factors V and VIII, has revealed that blood coagulation components may have evolved from two main gene structures, a Gla-EGF1-EGF2-SP domain and an A1-A2-B-A3-C1-C2 domain through a series of duplications and diversification (Davidson, Tuddenham and McVey 2003).

1.2. Protein Z

Protein Z (PZ) is a vitamin K-dependent (VKD) plasma glycoprotein which was identified in bovine plasma in 1977 and named as it was the last VKD protein eluted by anion exchange chromatography (Prowse and Esnouf 1977). This protein was isolated in humans in 1984 (Broze and Miletich 1984). PZ is a 62 kilodalton (KDa) protein produced predominantly in liver cells and secreted into the plasma (Kemkes-Matthes and Matthes 1995). PZ circulates as a complex with protein Z-dependent protease inhibitor (Han, Fiehler and Broze 2000, Broze 2001) and is composed of 360 amino acids (Fujimaki *et al.* 1998). In normal pooled plasma all PZ was found to be bound to PZ-dependent protease inhibitor (Tabatabai, Fiehler and Broze 2001). Another minor site of PZ expression are endothelial cells (Vasse *et al.* 2006).

1.2.1. Function of Protein Z

The physiological function of protein Z is still not clear, although this protein was isolated in humans more than twenty years ago. On the one hand, Hogg and Stenflo (1991) observed that bovine thrombin, in the presence of bovine protein Z, binds to the phospholipid layer in order to stop bleeding and this process cannot be accomplished in the absence of protein Z. Therefore, a decline in protein Z should logically lead to a bleeding tendency (Hogg and Stenflo 1991).

However, human PZ showed a lower efficiency in promoting thrombin binding to the phospholipid layer than bovine PZ; this is due to the lack of the region 366 to 369 of the bovine orthologue which forms the thrombin binding region (Hogg and Stenflo 1991). This assumption was supported by work from Kemkes-Matthes and Matthes (1995) who demonstrated that the average level of protein Z and the bleeding time in 36 patients with bleeding tendency of unknown origin were significantly lower than in healthy persons (Kemkes-Matthes and Matthes 1995). The hypothetical procoagulant role of PZ was strengthened by Kobelt and colleagues when they found higher PZ levels in patients with ischemic stroke compared with normal controls (Kobelt *et al.* 2001). However, a further study demonstrated that the PZ level in 15 patients diagnosed with bleeding problems of unknown origin lay within the normal range (Gamba *et al.* 1998). Ravi et al (1998) also showed that protein Z levels in 48 patients with suspected bleeding were within the normal range compared to 200 healthy persons. Nevertheless, male patients in this study had lower protein Z than healthy males and one female with severe bleeding tendency of unknown cause was found to have the lowest level of protein Z (Ravi *et al.* 1998). On balance, these observations seem to support the procoagulant action of PZ to varying degrees. On the other hand, protein Z was found to enhance the inhibition of active factor X by protein Z-dependent protease inhibitor (ZPI) 1000 fold. This process involves formation of the Xa-PZ-\ZPI tertiary complex, in the presence Ca^{++} , at the phospholipid membrane surface (Han, Fiehler and Broze 2000, Han, Fiehler and Broze 1998). The ZPI reaction with FXa was found to be enhanced by heparin in the presence of calcium and this enhancement has a similar effect on free ZPI and ZPI-PZ complex (Huang *et al.* 2011). The location of the PZ interactive site with ZPI is within the C-terminal domain (Hogg and Stenflo 1991, Rezaie *et al.* 2008). The anticoagulant effect of ZPI, regulated by PZ, was confirmed by the observed formation of a high affinity stoichiometric complex of PZ with ZPI in the absence of phospholipid and Ca^{++}, and the finding that PZ enhances the inhibition of the activation of factor X by protein Z-protease inhibitor (ZPI) (Huang *et al.* 2008). Hence PZ acts as an anticoagulant and its deficiency would be expected to lead to a procoagulant state and hence cause thrombosis. The risk of thrombosis which is caused by protein Z deficiency has been shown in mice with the factor V Leiden mutation (FVL). Researchers

have indicated that the homozygote FVL genotype can lead to death of mice from intrauterine and perinatal thrombosis if combined with a deficiency in protein Z (Yin *et al.* 2000, Corral *et al.* 2007). In agreement with the results reported in the mice, thromboembolism was found to be experienced earlier in PZ deficient patients who also carried the FVL mutation (Kemkes-Matthes *et al.* 2002). The theory that protein Z deficiency will lead to a procoagulant state was also supported by Vasse and other researchers when they found that a deficiency in protein Z is associated with ischaemic stroke and thrombotic complications of atherosclerosis (Vasse *et al.* 2001) and also with deep vein thrombosis (Santacroce *et al.* 2006). A stronger association of ischemic stroke with low levels of PZ was also reported by Heeb and colleagues in males with other risk factors such as hypertension, diabetes, high cholesterol and smoking (Heeb *et al.* 2002). These data are consistent with the fact that the physiological function of PZ is not clear. However, the current weight of evidence in these recent studies suggests that the role of protein Z in the coagulation process seems to be anticoagulant rather than procoagulant.

1.2.2. Structure of Protein Z

The modular organisation of PZ comprises four domains, a gama-carboxyglutamic acid domain, two epidermal growth factor (EGF1 and EGF2)-like domains, and a serine protease (SP)-like domain. This domain structure has close similarity with other coagulation factors such as factor VII (FVII), factor IX (FIX), factor X (FX) and protein C (PC) (Ichinose *et al.* 1990, Sejima *et al.* 1990). The modular organisation of PZ and the evolutionarily related proteins is shown in figure 1.3. However, PZ differs from the above in that it lacks the serine and histidine residues of the catalytic triad and, therefore, lacks the proteolytic activity (Ichinose *et al.* 1990, Sejima *et al.* 1990, Chandrasekaran *et al.* 2008). The serine-195 and histidine-57 residues in the active site are substituted by threonine and alanine, respectively (Hojrup, Jensen and Petersen 1985). PZ also contains the highest content of Gla-residues (13) among the vitamin-K dependent proteins (Vasse 2008, Hojrup, Jensen and Petersen 1985). Compared with other VKD proteins, PZ displays slower membrane binding and dissociation kinetics which may correlate with the unique Gla residue placement at location 11 (McDonald *et al.* 1997).

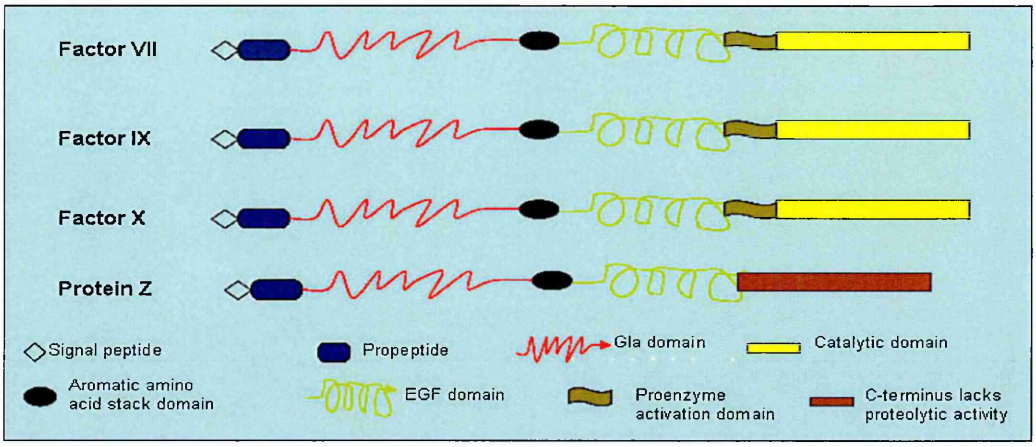


Fig 1.3. Modular organisation of evolutionarily conserved coagulation proteins. Protein Z has 39 and 22% homology to FVII, 45 and 25% to FIX and 47 and 24% to FX at the N-terminus and C-terminus, respectively; and, in common with these proteins, it features signal peptide and propeptide regions, a Gla domain, followed by an α-helical region containing a cluster of aromatic residues, two epidermal growth factor (EGF)-like domains and a serine protease-like domain. Protein Z, however, lacks catalytic activity and functions as a cofactor, unlike factors VII, IX and X which are all zymogens of serine proteases. Figure adapted from Hoffbrand *et al.* (Hoffbrand, Moss and Pettit 2006).

1.2.3. Organisation of the Protein Z gene

The gene encoding protein Z is located on chromosome 13 at band q34 (Fujimaki *et al.* 1998) and its organization is similar to other coagulation proteins such as coagulation factors VII, IX, X and protein C, again suggesting that they may have evolved from the same ancestral gene (Ichinose *et al.* 1990, Sejima *et al.* 1990). The PZ gene is approximately 14 kb in length, comprising eight regular and one alternative exons (Fujimaki *et al.* 1998); Fig 1.4).

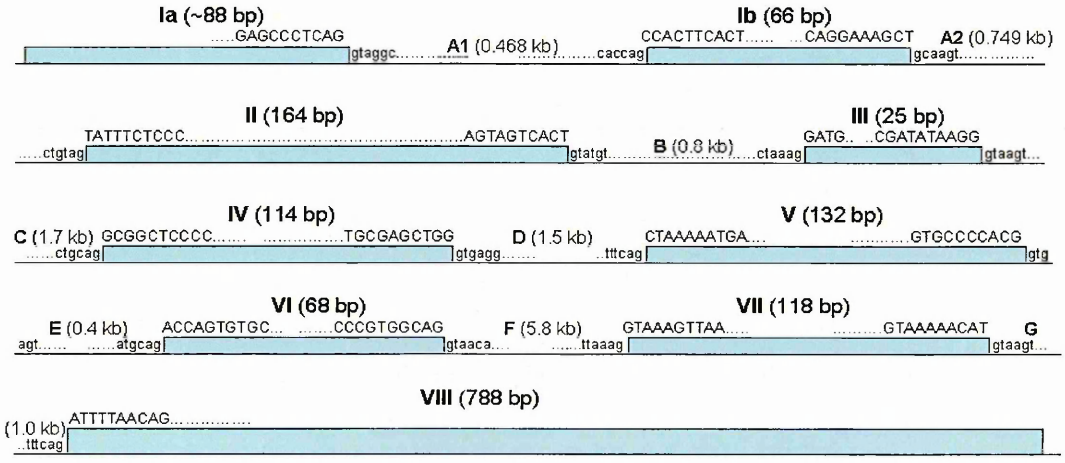


Fig 1.4. Organisation of the human Protein Z gene. The nine exons of PZ gene are shown with wide horizontal bars and are numbered with roman numerals. Nucleotides in capitals show the boundaries of the exons while nucleotides in small letters form the boundaries of the introns. The size and location of Introns A1 – G are shown.

1.2.4. Protein Z levels

Plasma PZ levels vary widely in different individuals with a correspondingly wide normal range (0.6-5.7 μg/mL) (Miletich and Broze 1987). The levels of plasma PZ in 200 healthy adults were found to be significantly higher in males than in females with no correlation with age (Ravi *et al.* 1998). Kobelt *et al.* (2001) supported this significant correlation of PZ levels with sex in addition to weak correlation with age (Kobelt *et al.* 2001). Furthermore, an earlier study also determined that there was no link between plasma PZ levels and age, although the correlation with sex was reported (Miletich and Broze 1987). This observed variation in PZ levels among normal individuals and sex may be explained, at least in part, by regulatory mechanisms controlling the rate of transcription from the PZ locus.

1.3. The regulation of gene expression

Gene expression is regulated through many subsequent stages that exist between the DNA and the production of specific protein. These stages are the transcription of the DNA to RNA, modification of the RNA transcript by the addition of a 5' cap, addition of adenosine residues at the 3' end of the RNA, removal of introns by RNA splicing, and transportation of the messenger RNA (mRNA) from the nucleus to the cytoplasm where the protein is translated (Latchman 1992). Regulation of transcription is vital in tissue-specific gene expression, as well as controlling expression of genes regulated in response to specific stimuli. These effects are mediated through the binding of specific transcription factors (which are themselves expressed in a tissue specific manner or activated in response to signalling pathways, respectively) to specific DNA sequences in the promoter regions of genes (Latchman 1997). Understanding of a genes promoter structure, where transcription factors bind their cognate sites in order to activate or inhibit transcription of the gene, is very important in knowing how that gene is regulated (Jiang, Jarrett and Haskins 2009). Once transcription is completed and mRNA is transcribed additional processes take place to determine the expression profiles of proteins. These processes control post-transcriptional regulation and include mRNA splicing, mRNA export, RNA stability and translation, adding more complexity to the

regulation of gene expression (Keene and Lager 2005, Mata, Marguerat and Bahler 2005). RNA-binding proteins (RBPs) and other mechanisms such as microRNAs and small interference RNAs are responsible for the post-transcriptional regulation (Keene and Lager 2005). The stablility of mRNAs, which is controlled by the poly(A) tails, is targeted for degradation. The degradation is accomplished by the exonucleolytic pathway which is initiated by deadenylation and in some transcripts is followed by removal of the 5'-methyl guanosine cap (Mata, Marguerat and Bahler 2005). mRNA decay can also be achieved by the endonucleolytic pathway including the degradation of a target transcript by RNA interference (Tomari and Zamore 2005). Translation of mRNA is the final process in the regulation of gene expression and is responsible for the adaptation of protein levels. This process can be divided into three steps, including initiation, elongation and termination (Jansen *et al.* 1995). During the initiation of this process mRNA is recruited to the ribosomal complex and the AUG initiation codon is selected. The recognition of the AUG-codon and initiation of protein synthesis are facilitated by binding of the initiator Met-tRNA to eukaryotic initiation factor (eIF2) and Guanosine 5'-triphosphate (GTP) (Calkhoven, Muller and Leutz 2002). The mRNA is decoded and the peptide bond is formed in the elongation step which is terminated once a termination codon is reached the active ribosomal site (Jansen *et al.* 1995). The accurate regulation of gene expression is essential, since its failure may lead to a disease development in the affected individual (Latchman 1992).

1.3.1. Basal promoter element

Gene expression is predominantly controlled at the level of the rate of transcription initiation. The transcription initiation complex is composed of RNA polymerase II and other basal proteins (also called general transcription factors) that assemble on the promoter region in order to initiate mRNA synthesis (Fickett and Wasserman 2000). RNA polymerase II (pol II) transcribes genes in eukaryotes that encode mRNAs and some small nuclear RNAs (Orphanides, Lagrange and Reinberg 1996). The genes that are transcribed by Pol II have basal promoter elements which can be identified by general transcription initiation factors and DNA elements recognised by transcription proteins (Roeder 1996). Pol II requires the presence of general transcription factors,

including TATA binding protein, transcription factor (TF) IIB, TFIIE, TFIIF and TFIIH in order to initiate transcription (Hampsey and Reinberg 1999). The structure and function of those factors are conserved among all eukaryotes (Orphanides, Lagrange and Reinberg 1996). Initiation of polymerase II transcription starts from the recognition of core promoter elements by the TBP/TFIID complex which is then supplemented with the binding of TFIIB and TFIIA. The TFIIF-pol II complex is then gathered and the preinitiation complex is completed by binding of TFIIE and TFIIH. Subsequently, the promoter is melted to give an open initiation complex. This is followed by synthesis of the phosphodiester bond of the mRNA transcript, promoter clearance and finally the RNA transcript elongation. TFIIA can stabilise the complex of initiation at any time after the binding of TFIID (Orphanides, Lagrange and Reinberg 1996). Therefore, basal transcription of genes encoding proteins is a complicated process which needs the interaction of many components including general transcription factors and the RNA pol II complex to be accomplished accurately.

1.3.2. Transcription factors

In addition to these basal factors the diversity of expression profiles seen in protein coding genes is produced by the action of a comprehensive array of general, tissue-specific and inducible transcription factors. These transcription factors are regulatory proteins which can bind to specific DNA sequence motifs within an enhancer or promoter region of a gene in order to regulate its transcription (Latchman 2005). Transcription factors interact with the general transcription factors, the RNA polymerase or intermediate complexes to activate or repress the efficiency of transcription initiation (Warren 2002). There are two essential roles, in general, for transcription factors to be effective; i) the ability to bind the DNA and, ii) to affect the transcription rate either positively or negatively (Latchman 1997). Binding of transcription factors to their specific DNA sequences is not sufficient; they also need to interact with other factors or the RNA polymerase complex itself in order to affect the rate of transcription (Latchman 1990). The synthesis of many transcription factors is cell-type specific (Latchman 1990) and these regulatory proteins are generally expressed at low levels in cells compared to house keeping proteins (Jiang *et al.* 2009).

Genes encoding transcription factors make up approximately 6% of all human genes (Venter *et al.* 2001).

1.3.2.1. Structure of transcription factors

Transcription factors have specific regions for DNA binding. These regions are called DNA-binding domains, which allow transcription factors to interact with their appropriate DNA sequence in a gene promoter in order to regulate the gene (Latchman 1997). The DNA binding activities of transcription factors are found in relatively small regions comprising approximately 60 to 100 amino acids (Mitchell and Tjian 1989). The characterisation of the different structures found in DNA binding domains has led to the classification of the transcription factors according to those domains (Latchman 1997). For instance, the helix-turn-helix binding domain in the homeobox-containing transcription factors, the two cysteine-two histidine zinc finger domain in the TFIID, the E26 (Ets) binding domain in polyomavirus enhancer activator 3 (PEA3), and other well-defined structures acting to efficiently and specifically bind to DNA (Latchman 2005).

The binding of transcription factors to DNA is significant; however, it is not sufficient for the activation of transcription. In order for transcription to be activated other regions in the transcription factor are required. These regions are called activation domains which enhance transcription by modulating the rate at which the RNA polymerase complex assembles on and transcribes the DNA template (Latchman 2005). There are three broad types of activation domain, an acidic activation domain with a high proportion of acidic amino acids (Ptashne 1988, Sigler 1988), a glutamine rich activation domain (Courey and Tjian 1988, Courey *et al.* 1989) and a proline rich activation domain (Mermod *et al.* 1989). Some transcription factors contain activation domains, but not a DNA binding domain. These factors instead form a complex with other binding proteins, without contacting the DNA, in order to modulate transcriptional activation. For example, the herpes simplex virus trans-activating protein VP16, which does not have a DNA binding domain, forms a complex with the cellular Octamer binding protein, Oct-1, to bind the viral promoter and transactivate it (Latchman 2005). Also, although the majority of transcription factors have activation domains some can have inhibitory domains instead. These domains

reduce gene transcription when they interact, either directly or indirectly, with the basal transcription complex (Latchman 1997).

1.3.2.2. Function of transcription factors

Transcription factors bind to sequence specific regulatory cis elements more frequently, located within the promoter region of the gene whose transcriptional activity they regulate. Indeed, these regulatory elements may be located several thousands of base pairs from the RNA transcription initiation site (Mitchell and Tjian 1989). As mentioned earlier many transcription factors activate transcription and are called activators, whilst others repress transcription and are known as inhibitory factors. Latchman has stated that activation of a gene by transcription factors occurs via different routes, the presence of the necessary active transcriptional protein in the correct tissue, activation of the pre-existing factor by binding to its ligand, dissociation of inhibitory protein from the activated protein, and activation of the protein by means of modification, such as phosphorylation, in response to a specific signal (Latchman 1990). Stargell *et al.* have reported that activators can induce transcription either directly by the interaction of the activation domain with the general transcription initiation factors, or indirectly by relieving nucleosome-mediated repression (Stargell and Struhl 1996). Activators can also initiate transcription by enhancing the assembly of the basal transcription complex or stimulating the activity and/or stability of the complex by altering the conformation of an already bound factor (Latchman 2005).

Inhibitory factors can act in a number of different ways to reduce transcription levels. They may block the activating effect of other transcription factors by preventing them from binding to the DNA. This prevention can be achieved by binding of the inhibitory factor to the DNA binding domain of the positively acting factor forming a non-binding complex, or interaction of the negatively acting factor with the positive one to block its activation domain (Latchman 1997).

The nature of the amino acids that make up the DNA binding domains of transcription factors are important for their functional activity. The stimulatory effect of a transcription factor can be drastically altered by just a single amino

acid substitution in its binding domain. For instance, replacement of isoleucine in the binding domain of Brn-3b, from the POU (Pit-Oct-Unc) family, with valine was sufficient to switch the activity of this transcription factor from a repressor to an activator (Dawson, Morris and Latchman 1996). Other transcription factors can also activate transcription of some genes and repress others depending on the DNA binding sites. For instance, glucocorticoid receptor binds the glucocorticoid response element (GRE), as a dimer, following hormone treatment and activates gene transcription. However, when this receptor binds as a trimer to a different but related GRE sequence, called a negative GRE (nGRE), this leads to repression of transcription (Latchman 2001).

1.3.3. Tissue-specific regulation

The regulation of transcription is often a key step in defining the expression pattern of a particular gene in a tissue-specific or developmental stage-specific manner (Latchman 1997). The expression levels of tissue enriched genes is higher in a particular tissue type, establishing the cell's specialized function, and being much lower or negligible in other tissues (She *et al.* 2009). Based on expressed sequence tag data analyses, the genomic length of tissue specific genes was found to be less and their expression level lower compared to house keeping genes (Zhu *et al.* 2008). Moreover, the majority of tissue specific genes lack both CpG-islands and TATA-boxes (Zhu *et al.* 2008). Many genes encoding tissue-specific proteins have been found to be regulated by tissue-specific transcription factors. The liver-enriched hepatocyte nuclear factor 4 (HNF4α) has been found to be essential in the regulation of many blood coagulation factors including PZ (Sugawara *et al.* 2007), Factor VII (FVII), factor VIII (FVIII), factor IX (FIX), factor X (FX), factor XI (FXI) and PS (Stauffer *et al.* 1998, Figueiredo and Brownlee 1995, Reijnen *et al.* 1992, Miao *et al.* 1992, Tarumi *et al.* 2002), which are normally produced in a tissue-restricted fashion in the liver. Sugawara and colleagues in 2007 reported the first step towards understanding the mechanism of PZ expression. They reported that the regulation of protein Z expression is controlled by HNF4α using electrophoretic mobility shift assays (EMSAs) and reporter gene assays (Sugawara *et al.* 2007). Moreover, the liver enriched factor HNF1 was found to activate the PS gene promoter (Hall, Peake and Winship 2006), and the liver

enriched factor GATA-4 and nuclear transcription factor Y (NF-Y) activate transcription of the gene encoding factor X (Hung *et al.* 2001). PZ, which is also predominantly produced in the liver, may, therefore, also be regulated by these transcription factors.

1.3.4. Inducible regulation

Regulation of the activity of transcription factors is important in the regulation of gene transcription (Latchman 1992). Transcription of a gene is sometimes not activated except in the presence of a particular stimulus. This stimulus activates the ability of a particular transcription factor, by conformational changes or a disruption of inhibitory protein-protein interaction, to bind the DNA and permit transcription of the gene (Latchman 1990). Examples of pathways that rely on signalling molecules to produce appropriate effects on gene expression include hormone and inflammatory responses.

1.3.4.1. Steroid hormone receptors

Steroid hormone receptors are crucial in the regulation of gene expression by steroid hormones, through their interaction with their DNA hormone responsive elements (HRE) (Beato *et al.* 1989). The steroid hormone receptors are related to a large family known as nuclear receptors and are bound by steroid hormones. (Laudet *et al.*1992). The major groups of steroid hormones, androgen, oestrogen, glucocorticoids, progestins and mineralocorticoids are extensively metabolised in the liver and regulate many physiological processes in cells that express their corresponding receptors (Beato and Klug 2000). The nuclear proteins are ubiquitous eukaryotic and structurally related transcription factors. These proteins bind to specific DNA cis-acting sequence and different coregulators, in response to their ligands, and regulate the transcription of many genes in a ligand-dependent manner (Raviscioni *et al.* 2006, Ishizuka *et al.* 2005). Many transcriptional cofactors, which are organized in multiprotein complexes, have been recognized to initiate chromatin remodeling and the modification of histones in order to enhance or repress the binding of the nuclear receptors and the RNA polymerase II with the basal factors to the target DNA (Ishizuka *et al.* 2005). Some of these cofactors can act as coactivators,

such as the CBP/P300 (Chakravarti *et al.* 1996) and P300/CBP associated factor (PCAF) / general control of amino-acid synthesis 5-like 2 (GCN5) (Blanco *et al.* 1998) whilst others act as corepressors, such as the silencing mediator of retinoid and thyroid receptor (SMRT) (Chen and Evans 1995) and the nuclear receptor corepressor (N-CoR) (Horlein *et al.* 1995). Transcription activation by steroid hormone receptors can occur directly by protein-protein interaction or indirectly by their synergistic action with adjacent transcription factors (Schule *et al.* 1988, Strahle, Schmid and Schutz 1988). A palindromic sequence composed of two sites interrupted by three bases has been identified as a hormone response element for nuclear receptors such as oestrogen, androgen and glucocorticoid receptors (Aumais *et al.* 1996). The DNA was found to be bent in response to oestrogen nuclear receptor binding, suggesting a change in the DNA conformation upon binding of the receptor is common to the steroid hormone family (Nardulli and Shapiro 1992). The structure of the nuclear receptors is organised as an N-terminal domain with ligand-independent activation function (AF-1) (Tora *et al.* 1989, Metzger *et al.* 1995), a C-terminal ligand binding domain, and a DNA binding domain (DBD) located between the two termini (Fig 1.5). The C-terminus domain is known as a ligand binding domain, with a complex structure composed of a ligand dependent activation function (AF-2), an interface for dimerisation as well as the ligand binding pocket (Steinmetz, Renaud and Moras 2001). The DBD, which binds to hormone response elements for regulation of a proposed gene (Khorasanizadeh and Rastinejad 2001), is a highly conserved region among nuclear receptors. It is composed of about 71 residues (Fig 1.5) containing two (cys 4) zinc finger domains (Evans 1988, Freedman *et al.* 1988, Hard *et al.* 1990). Increased levels of testosterone and estradiol were observed throughout pregnancy and of course lower levels of testosterone are normally found in women than in men. Furthermore, Low levels of androgen and/or oestrogen were reported in obese people. Therefore, higher levels of PZ during pregnancy (Quack Loetscher *et al.* 2005, Ramsay *et al.* 2005) as well as higher levels found in men compared to women (Ravi *et al.* 1998) and in obese pregnant women compared to lean pregnant women (Ramsay *et al.* 2005) suggest a role for steroid hormones in PZ gene regulation.

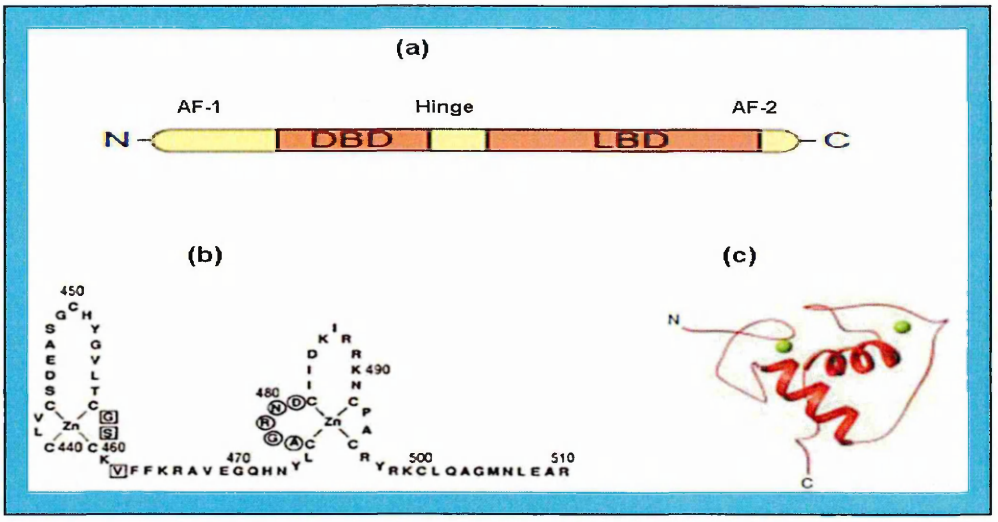


Fig 1.5. Organisation of the nuclear receptor structure. (a) The conserved architecture of nuclear receptors shows the locations of the ligand-independent activation function (AF1), DNA-binding domain (DBD), ligand-binding domain and the ligand-dependent activation function. (b) The DNA-binding domain of the rat glucocorticoid receptor (GR). The boxed amino acids are crucial for the distinction between glucocorticoid and oestrogen respose elements. The circled amino acids represent the region that essential for protein-protein interactions in the DBD-GRE complex. (c) The crystal structure of the DBD. The green spheres represent the zinc ions (Khorasanizadeh and Rastinejad 2001, Hard *et al.* 1990).

1.3.4.2. Inflammatory response

Inflammation is defined as a response of the tissue to injury and a defence mechanism for the organism against infections (Willoughby 1978). It is also defined as a responsive mechanism needed for re-establishing haemostasis (Medzhitov 2008). The inflammatory response, a part of the innate immune response, is a crucial step in restoring tissue damage, however, excessive inflammation can increase the severity of damaged tissue, and to avoid this many initiators of the inflammatory response can concurrently initiate reduction of inflammation (Anderson 2010). Inflammation is initiated exogenously by microbial and non-microbial inducers and endogenously by stressed, damaged or dysfunctioning tissues (Medzhitov 2008). The responses to inflammation can be regulated by cell-specific mechanisms which involve the regulation of the recruitment and activation of different cell types, signal-specific mechanisms which involve signalling pathways, and gene-specific mechanisms which function at individual gene and gene subset level (Medzhitov and Horng 2009). Cytokines, chemokines and other mediators of the inflammatory response which are produced by macrophages, mast cells (Medzhitov 2008), vessel

endothelial cells and fibroblasts (Ruminy *et al.* 2001) which are driven from microbial products (e.g. lipopolysaccharide (LPS)) and the proinflammatory cytokines including tumor necrosis factor α (TNFα), interleukin1(IL1), interleukin 17 (IL17) and interleukin 18 (IL18), lead to gathering of leukocytes at the site of damaged tissues. These proinflammatory cytokines and LPS bind to cell surface receptors in order to regulate inflammatory response genes at transcriptional and post-transcriptional levels (Kracht and Saklatvala 2002). The aggregation of proinflammatory cytokines on the receptors leads to oligomerisation of tissue necrosis factor receptor-associated factor (TRAF) and results in activation of the IκB and other kinase pathways (Kracht and Saklatvala 2002). Inflammation can induce activation of pre-existing inactive transcription factors. In this process, the inactive factor is dissociated from the inhibitory protein by means of different mechanisms such as ligand binding, alterations in protein-protein interaction or phosphorylation which then facilitates translocation to the nucleus to bind the appropriate response element and transactivate transcription (Perkins 1997). The promoters of inflammatory genes contain essential cis-elements for transactivating factors commonly NFκB, AP1, C/EBP and Ets (Kracht and Saklatvala 2002). A key transactivating factor which is important in the expression of inflammatory genes is NFκB. This transactivator, following activation of a relevant signal transduction pathway by LPS or interleukins, is dissociated from the inhibitory protein IκB in the cytoplasm. The dissociation of this factor is accomplished by degradation of IκB by proteases triggered by its phosphorylation and ubiquitination, allowing NFκB to translocate to the nucleus (Fig 1.6) and bind to its respective binding elements (Perkins 1997). Mutation or deletion of NF-κB or AP1 binding sites in the promoter of interleukin8 were found to impair its expression in response to inflammatory stimuli, indicating the essential role for the transactivating factors and their binding sites in the transcription of inflammatory genes (Mukaida *et al.* 1994). FVIII and tissue factor, proteins within the coagulation process, were reported to be induced by bacterial lipopolysaccharide (LPS) (Begbie et al 2000, Hall et al 1999). The NF-κB and AP1 binding sites in the tissue factor gene promoter and NF-κB and C/EBP binding sites in the FVIII gene promoter were found to be the LPS-responsive elements which are important in the expression of these genes. Previous

studies suggested that the PZ gene may also be regulated by inflammatory signals. For instance, levels of plasma protein Z were found to decrease in response to increasing levels of the cytokine interleukin 6 (IL-6) (Undar et al 1999). Interleukin-4 (IL-4) and tumour necrosis factor (TNF-α) were found to have a down-regulatory effect on PZ, again supporting the hypothesis that this protein may be a negative acute phase protein (Vasse et al 2002).

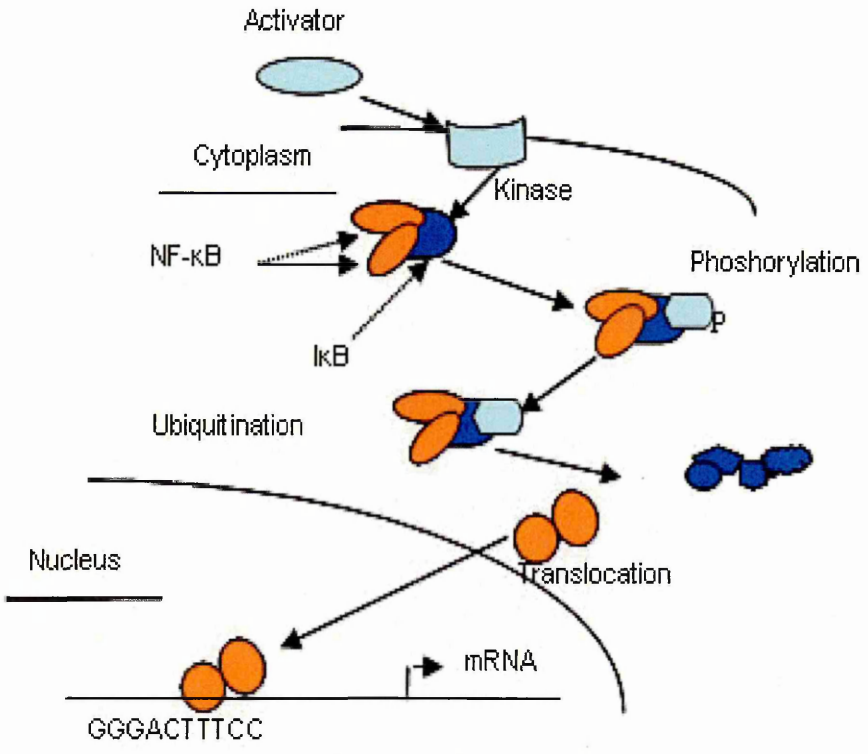


Fig 1.6. The pathway leading to the translocation of NF-κB.
After cellular stimulation the NF-κB-IκB complex is phosphorylated by relevant signal pathways leading to degradation of IkB and its dissociation from NF-κB. The free NF-κB, then, translocates to the nucleus to bind its DNA binding sites.

1.3.5. Role of chromatin structure in controlling expression

The architecture of chromatin plays an essential role in modulating the binding of transcription factors to the DNA and hence in regulating transcription (Wolffe and Guschin 2000). Chromatin is composed of a series of nucleosome core particles consisting of a histone octamer with about 160 base pairs of DNA wrapped around it which are separated by regions of nucleosome-free linker

DNA (Schrem, Klempnauer and Borlak 2002). The structure of chromatin is shown in figure 1.7. The histones H2A, H2B, H3 and H4 are the units that make up the histone octamer of the nucleosome core particle. Each octamer comprises two copies of an H3-H4 dimer which form a tetramer, with an H2A-H2B dimer on either side of the tetramer to form the octamer (Schrem, Klempnauer and Borlak 2002, Spencer and Davie 1999). A fifth class of histones, the linker histone (H1) with lysine-rich N and C-terminal domains, was found to maintain the stability of the DNA around the core histones by looping the nucleosomes on top of each other in order to further compact the DNA. This promotes formation of the 30 nm fiber characteristic of nucleosomeal organization (Spencer and Davie 1999, Stein *et al.* 1999, Hamon and Cossart 2008).

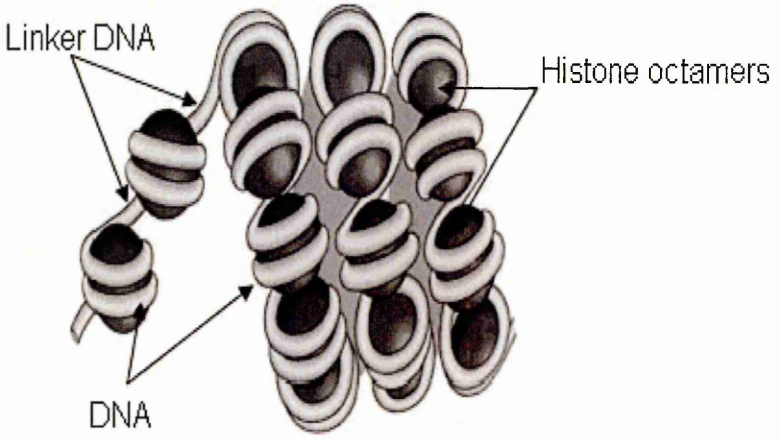


Fig 1.7. The basic structure of chromatin.
The black spheres represent the histone octamers which with the DNA wrapped around them form what is called a nucleosome core particle. The linker DNA regions connect the nucleosome core particles (Stein *et al.* 1999).

The chromatin structure with a solenoid-type arrangement of folded nucleosomes shortens the distance between regulatory elements of the promoter which can act to facilitate interactions and, therefore, response to signaling pathways, resulting in the transcriptional control of target genes (Stein *et al.* 1999). Since the chromatin structure works as a barrier that prevents transcription, this structure has to be opened up or remodeled to facilitate the

accession of transcription factors for the control of gene expression (Demeret, Vassetzky and Mechali 2001).

A chromatin environment that allows transcription factors to interact with their respective DNA cis-elements is crucial for efficient gene transcription (Schrem, Klempnauer and Borlak 2002). There are two main levels of mechanism responsible for the modification and regulation of chromatin; firstly the post-transcriptional modification of histone tails by processes such as phosphorylation, acetylation, and methylation; and secondly the ATP-dependent remodeling of chromatin (Falbo and Shen 2006). Chromatin kinetics, directed by chromatin remodeling and histone modification, are very important in the regulation of gene transcription (Schrem, Klempnauer and Borlak 2002, Spencer and Davie 1999, Stein *et al.* 1999, Hamon and Cossart 2008). Transcription is regulated by histone modifications through a two step mechanism, firstly the disruption of the connection between nucleosomes and then the encouragement of nonhistone proteins such as transcription factors to bind the DNA (Kouzarides 2007). The N-terminal tail is the most common site for histone modifications such as phosphorylation, methylation, acetylation and ubiquitinylation, however, these modifications might also occur in the histone core (Hamon and Cossart 2008). The chromatin remodeling is achieved by chromatin remodeling complexes which can alter the structure of the nucleosomes and their positioning (Demeret, Vassetzky and Mechali 2001). This change in the structure of the nucleosome affects the coiling of the DNA around the histone octamer, leading the DNA template to be accessible to its corresponding transcription factors (Schrem, Klempnauer and Borlak 2002, Falbo and Shen 2006).

Therefore, the modulation of chromatin structure by histone modifications and chromatin remodeling are the key mechanisms required for the transcriptional regulation of a genes. One of the advantages of using cell lines in studying gene expression, such as in overexpression and knock down techniques, is investigating the proposed endogenous gene within its intact chromatin context rather than using a purified *in vitro* system such as EMSA.

1.3.6. Post-transcriptional regulation via the RNAi pathway

RNA interference (RNAi) is a process by which gene expression may be regulated at a post-transcriptional level. The two key RNA molecules involved in RNA interference are microRNAs (miRNAs) and small interfering RNAs (siRNAs). siRNAs are produced from a double stranded RNA (dsRNA) which can be made from inverted-repeat transgenes, viral infection or abnormal transcription products (Ghildiyal and Zamore 2009). Among the first natural siRNAs reported in mammals were those specific to the long interspersed nuclear element retrotransposon and were detected in cultured human cells: HEK293 (human embryonic kidney) and Hela (human epithelial carcinoma) cell lines (Yang and Kazazian 2006). RNAi is an increasingly important mechanism for the regulation of gene expression which restricts the level of the transcript of a gene by post-transcriptional silencing through the degradation of a specific sequence of RNA or the blocking of translation (Agrawal *et al.* 2003). As stated earlier RNA silencing can be accomplished by two species of small RNAs, miRNAs and siRNAs. The biogenesis, molecular characteristic and effector functions of these two species are basically similar, whilst their molecular origin and target recognition mode are different (He and Hannon 2004). The miRNA, which is produced from its genome-encoded precursor, binds imperfectly to the target 3' UTR in order to negatively regulate expression at the translational level. By contrast, siRNA is cleaved from an exogenously introduced or endogenously bi-directionally transcribed (Hannon 2002) large double stranded DNA (dsDNA) and binds perfectly to its target mRNA causing cleavage (He and Hannon 2004).

RNAi, as well as being an important *in vivo* mechanism of post-transcriptional gene regulation, can also be utilised as a beneficial technique to study the function of genes in many organisms (Agrawal *et al.* 2003). The silencing of genes using siRNA can be exploited and used experimentally to modulate mRNA levels in cells for the analysis of most vertebral genes (McManus and Sharp 2002). The pathway of RNA interference by small double stranded RNAs (dsRNAs) was initially observed by Fire and his group (1998) when it was found that a mixture of sense and antisense RNA sequences of unc-22 gene had more potential interference with the endogenous gene activity than either RNA sequence alone (Fire *et al.* 1998). The trigger strands are required to form

double stranded RNA and must have homology with the target RNA in order to initiate the interference (Parrish *et al.* 2000). The double stranded RNAs which mediate RNAi, known as siRNAs, were found to have a specific structure of 5' phosphate, 3' hydroxyl, being 21-22 nucleotides in length and having two nucleotide overhangs at the 3' end (Fig 1.7a) (Elbashir, Lendeckel and Tuschl 2001). siRNA was later reported to interfere with endogenous and heterologous genes in some mammalian cell lines (Elbashir *et al.* 2001). Two steps are essential for the RNAi pathway (Fig 1.7b), the cleavage of the large dsRNA into siRNA of 21 to 25-nucleotide by the RNA nucleases and the degradation of the corresponding single stranded mRNA by RNA-induced silencing complex (RISC) (Agrawal *et al.* 2003). The cleavage of the dsRNA is carried out by Dicer, an endoribonuclease enzyme of the RNAse III family of nucleases which is conserved in neurospora, drosophila, caenorhabditis and mammalia (Bernstein *et al.* 2001). RISC, which leads to the degradation of the mRNA, contains a single-stranded antisense siRNA residue links the mRNA and at least one member of the argonaute protein family (Martinez *et al.* 2002).

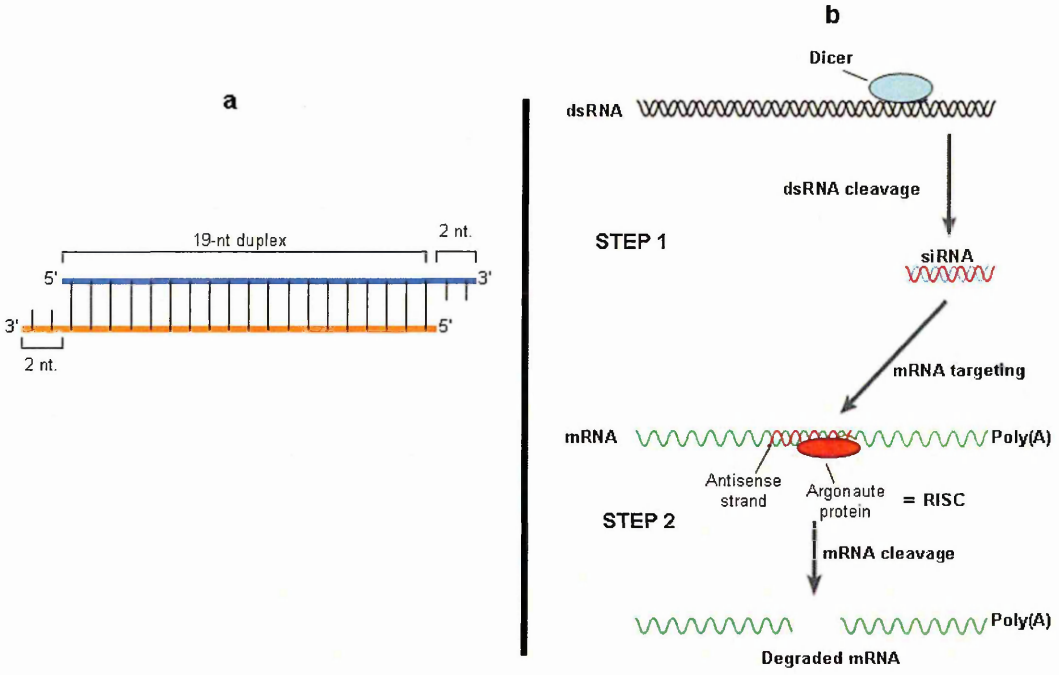


Fig 1.8. The siRNA structure and the RNAi pathway.
(a) siRNA shows its 23 nt with 2 nt overhanging at 3' end. (b) The RNAi pathway shows the two steps lead to the degradation of mRNA. Dicer is a nuclease enzyme involved in the formation of the siRNA in the first step. In the second step, The antisense strand of the siRNA binds the homologous sequence in the mRNA and leads to its cleavage and subsequently degradation. nt = nucleotides. RISC= RNA-induced silencing complex.

1.4. Aim of the study

The object of this study was to elucidate the mechanisms responsible for controlling the synthesis of Protein Z (PZ), a protein with reported pro and anticoagulant properties. PZ plasma levels vary widely in the normal population; despite this, lowered levels have been associated with a bleeding tendency in some individuals and thromboembolic complications in others. Hence, PZ level interacting with genetic and/or environmental background appears to determine pathological consequence.

The major aim of this work was, therefore, to investigate the proposed hypothesis that a significant degree of the observed variation in PZ level arises as a consequence of inter-individual differences in rate of transcription from the PZ gene locus. The mechanisms of transcriptional activation of the PZ gene were investigated toward possible regulatory processes in common with other evolutionarily related haemostatic factors and also those regulatory processes likely to account for the inter-individual variation. Firstly, it was investigated whether the PZ promoter contains the same liver-restricted and age-stable regulatory elements reported in other, related haemostatic factors, factor IX and protein C. Secondly, the wide inter-individual variation in PZ level suggests that it may be an acute phase reactant and, therefore, evidence of transcriptional elements associated with this physiological response in the PZ promoter were sought. Finally, the effect of steroid hormones, which themselves show inter-individual variation in level according to age and medication, were analysed for their potential effects in modulating transcriptional activity of the PZ gene.

Chapter 2-Materials and Methods

2.1. DNA sequencing

In order to validate the sequence of the PZ gene promoter in HepG2 (ECACC, UK) the promoter was compared with the sequence of the wild type PZ gene promoter. The wild type PZ gene promoter (Homo sapiens protein Z, AF440358, National Centre for Biotechnology Information) was divided into three regions overlapping each other which together spanned nucleotides -1884 to +106 relative to the translation initiation site at nucleotide +1. Region 1 spanned nucleotides -1883 to -1284, region 2 spanned nucleotides -1352 to -608 and region 3 spanned nucleotides -716 to +217. Primers were designed for each region in order to allow amplification by polymerase chain reaction (PCR). The primers for each region were designed using the webprimer website (http://www.yeastgenome.org/cgi-bin/web-primer) and synthesised by Eurofins MWG \ Operon (Eurofins MWG Operon, UK). The sequences of the designed primers are shown in table 2.1. The PCR reactions were performed in the Primus 96 plus PCR machine (MWG-BIOTECH, UK) using genomic DNA prepared from HepG2 cells using the GenElute Mammalian Genomic DNA miniprep kit (Sigma-Aldrich, UK), and the designed primers.

Primers description	Sequences
PZ promoter region-1-Forward	5'-CCTGTCTGCCTTTCCATTTT-3'
PZ promoter region-1-Reverse	5'-GATTGGGAGGATGGCAGTG-3'
PZ promoter region-2-Forward	5'-CCAAACCTCGCTTCCCTCT-3'
PZ promoter region-2-Reverse	5'-TTTCAGATGTAGGCCCTGTGT-3'
PZ promoter region-3-Forward	5'-CATGTGGAACTGGGAGTCAA-3'
PZ promoter region-3-Reverse	5'-CAAACAGCTTTTTCAGGGGA-3'

Table 2.1. Oligonucleotide primers used for PZ promoter sequencing.
The primers of the PZ gene (Homo sapiens protein Z, AF440358, National Centre for Biotechnology Information) promoter were designed using a webprimer website and synthesized by Eurofin MWG Operon. Region-1 primers amplify 600 bp fragment (-1883 to -1284), region-2 primers amplify 746 bp fragment (-1352 to -608), and region-3 primers amplify 934 bp fragment (-716 to +217).

2.1.1. PCR of sequencing template

To amplify the three regions of the PZ gene promoter for sequencing, PCRs were accomplished using the primers of each region and the prepared DNA from Human Caucasian hepatocyte carcinoma-derived (HepG2 cells) (ECACC, UK). PCR is an *in vitro* technique by which a specific DNA segment can be amplified using two specific primers that hybridise to opposite DNA strands (Sharma, Singh and Sharma 2002). Large quantities of specific DNA can be generated, in a single enzymatic reaction, from a complex DNA template within a matter of hours. PCR is performed in three clearly defined steps, DNA denaturation, primer annealing and primer extension (Sharma, Singh and Sharma 2002). The PCR cycle conditions used to amplify the regions of the PZ promoter for the sequencing were as follow: Firstly, one step at 95°C for 2 minutes, to ensure complete denaturation of the DNA template, followed by 30 cycles of 95°C for 1 minute for DNA denaturation, annealing temperature of the proposed primers for 1 minute, to anneal the primers to their complementary regions, and 72°C for 1 minute for the extension of the primers by the Taq polymerase in order to synthesis complementary strands. Finally, 5 minutes at 72°C was performed to ensure all the single-stranded DNAs were extended and formed double stranded DNAs, followed by a holding step at 4°C for short-term storage. The annealing temperature for the primers as well as $MgCl_2$ concentration used to amplify each region in the PZ gene promoter are shown in Table 2.2.

Primers of different PZ promoter regions	$MgCl_2$ concentration	Annealing temperature
promoter region-1 primers	1 mM	53 °C
promoter region-2 primers	1 mM	52 °C
promoter region-3 primers	2 mM	63 °C

Table 2.2 Annealing temperatures and $MgCl_2$ concentration used to amplify different PZ promoter regions for sequencing. The fragment of PZ gene promoter region-1 (-1883 to -1284) was amplified using the respective primers at annealing temperature of 53 ^{0}C and $MgCl_2$ concentration of 1 mM. Region-2 (-1352 to -608) was amplified using the designed primers at annealing temperature 52 ^{0}C and $MgCl_2$ concentration of 1 mM while region-3 (-716 to +217) amplified at annealing temperature 63 ^{0}C and $MgCl_2$ concentration of 2 mM. These amplified regions were used in the sequencing of PZ gene promoter. The genomic DNA template used to amplify the different PZ gene promoter was prepared from the HepG2 cells using GenElute Mammalian Genomic DNA miniprep kit (Sigma-Aldrich, UK).

The amplified fragments were electrophoresed, to check their quality, on 1% w/v agarose gel running in 1X Tris-Borate EDTA (TBE) buffer (Sigma Aldrich, UK) using 10 volt./cm and detected using 400 ng/ml ethidium bromide (Sigma Aldrich, UK) which was visualized using a UVP BioImaging system.

2.1.2. Preparation of the 1% agarose gel

The 1 % agarose gel was prepared by dissolving 0.5 g of agarose powder (Promega Corporation, UK) in 50 ml of 1X TBE, using a microwave. The solution was left to cool down to about 50 °C and 2 µl of ethidium bromide (10 mg/ml) was added. The gel tray was prepared by sealing the ends with tape and a comb was placed about 1 inch from the end of the gel tray before pouring the agarose solution to the tray to a depth of about 5 mm. The gel was left to solidify for 30 minutes at room temperature. The comb was then removed before placing the gel tray in the electrophoresis chamber.

2.1.3. Purification of amplicons

Prior to employing the PCR amplified DNA products from the different PZ promoter regions in sequencing, it is necessary to purify the products, removing any dNTPs and primer remnants which might interfere with the sequencing procedure. The Exonuclease I / Shrimp Alkaline Phosphatase (ExoSAP-IT) method was used to clean up the PCR products and remove these contaminants. The ExoSAP-IT (USB Europe GmbH, Germany) method has many advantages such as ease of use and speed, consistent performance from sample to sample and no sample loss. 2µl of the ExoSAP-IT reagent was mixed with each 5µl of the PCR product and incubated for 15 minutes at 37°C, to degrade remaining nucleotides and primers, followed by 15 minutes incubation at 80°C to inactivate the reagent. The purified products were then sent to the Genetic Core Facility, Sheffield Medical School for sequencing. BigDye v.3.1 terminator sequencing chemistry (Applied Biosystems, UK) was used and the results were analysed on the ABI3730 capillary electrophoresis system (Applied Biosystems, UK).

2.2. Protein-DNA binding studies

The study of Regulatory systems that control gene expression forms a major area of interest in modern biology. One of the essential steps in gene expression is the control of transcription (Nagaraj, O'Flanagan and Sengupta 2008). Protein-DNA binding is an essential step in the transcriptional regulation process and therefore, accurate recognition of transcription factor binding sites is very important in understanding complex gene expression patterns (Morozov *et al.* 2005). In order to perform protein-DNA binding studies, nuclear extracts should be prepared and incubated with a labelled double stranded DNA probe spanning the binding site of the proposed transcription factor.

2.2.1. Nuclear extract preparation

Nuclear extracts were prepared using a rapid micropreparation technique (Andrews and Faller 1991). Hep G2 cells (approximately between 2.0 to 2.5 x 10^7 cells) were scraped into phosphate buffered saline, pelleted and then resuspended in buffer A (10mM Hepes-KOH pH 7.9, 1.5mM $MgCl_2$,10mM KCl, 0.5 mM DTT, Protease inhibitor cocktail (Sigma-Aldrich, UK)). The Hep G2 cells were swollen in the hypotonic solution, on ice for 10 minutes, and vortexed for 10 seconds to extract the nuclei. The nuclear pellet was recovered by centrifugation at 16.1 x10^3 g for 10 seconds and resuspended in ice-cold buffer C (20mM Hepes-KOH pH 7.9, 2.5% glycerol, 420mM NaCl,1.5mM $MgCl_2$, 0.2mM EDTA, 0.5mM DTT, 1% (v/v) protease inhibitor cocktail) and incubated for 20 minutes on ice for high salt extraction of nuclei. The suspension was centrifuged again at maximum speed (16.1 x10^3 g) at 4°C for two minutes and the supernatant (nuclear extract) was stored at -80°C. The protein concentration was determined by the Bradford assay.

2.2.2. Bradford Protein Assay

Quantification of protein is necessary before processing protein samples for electrophoretic analysis. There are many methods for protein quantification such as Lowry (Lowry *et al.* 1951), Bicinchoninic acid (Smith *et al.* 1985) and Bradford protein (Bradford 1976) assays. One of the easiest and fastest protein assays with high performance is the Bradford protein assay. The Bradford

protein assay is a biochemical assay used for the determination of protein in a solution. Protein concentration is measured by a colour change which can be quantified using colourimetric techniques. Protein binding to the coomassie dye in an acidic environment causes a shift in spectrum from maximum absorbance at 456 nm (reddish/brown form) to maximum absorbance at 610 nm (blue form) with the greatest difference between the two forms at 595 nm. Different concentrations of Protein standards were made from a 10 mg/ml stock solution of bovine serum albumin (Promega Corporation, UK). In a 96 well plate (Thermo Fisher Scientific, UK), 5µl of each standard and sample (undiluted, 1:2, 1:5, and 1:10 dilutions) were pipetted in triplicate. 250µl of Bradford reagent (Sigma Aldrich), a brilliant Blue G in phosphoric acid and methanol, was added. The 96 well plate was left to incubate at room temperature for 30 minutes before absorbances were recorded at 595nm on a GENios Plus reader (Tecan Group Ltd.). The absorbances of the different albumin concentrations were measured to plot a standard curve (Figure 2.1), which was then used to determine the protein concentration of the sample.

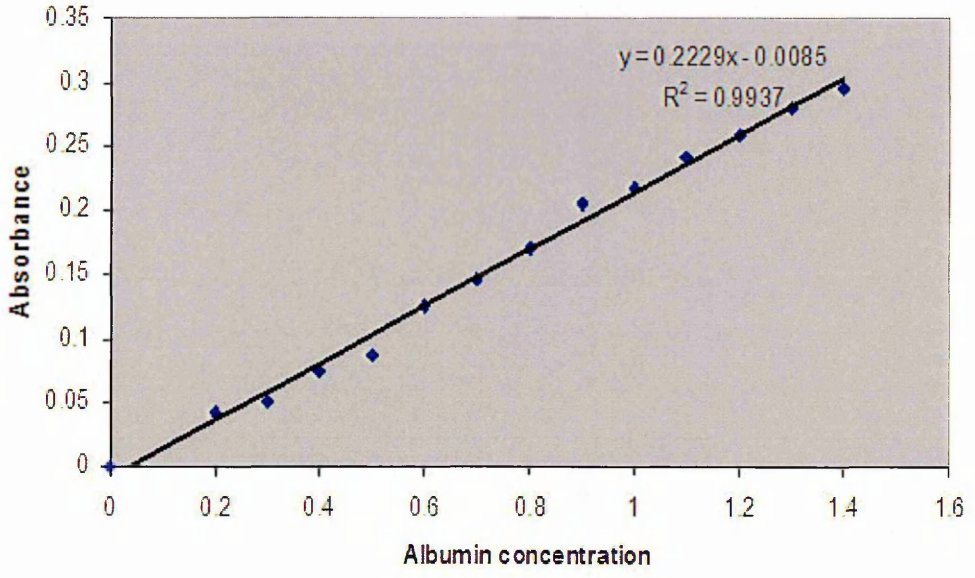


Fig 2.1. Standard curve for measuring the concentration of the HepG2 nuclear extract. Different concentrations of bovine albumin were prepared and the absorbance of each concentration was measured at 590 nm using GENios Plus reader. The absorbances were plotted against their respective albumin concentrations to generate a standard curve. The axis (x) represents different albumin concentrations while the axis (y) represents the absorbance of each concentration. The concentration of the the nuclear extract was measured from its absorbance using the equation derived from the standard curve, where x is the concentration and y is the absorbance of the unknown sample. The value of coefficient of determination (R^2) was high (0.99), revealing the close of the data points to the trend line.

2.2.3. Double stranded DNA Probe preparation (Fluorescent and biotin labelled)

Detection of protein-DNA binding using a probe encompassing the putative cis-element is one of the most important steps for the characterization of gene transcriptional regulation (Garner and Revzin 1981). The recommended probe length used in EMSA is between 20 to 40 base pairs and spans the putative protein binding site region (Kako *et al.* 1998).

The *in silico* prediction (ALGGEN-PROMO) of putative binding sites was used to determine the hypothetical binding sites for a number of transcription factors in the PZ gene promoter. The determined sequences for the hypothetical binding sites were then synthesised by Eurofin MWG \ Operon (Eurofins MWG Operon, UK), for the biotin and infra red 700 labelled probes, and by Sigma Aldrich (Sigma Aldrich, UK), for Cy5.5 labelled probes. Other unlabeled oligonucleotides used in competition reactions were synthesized by Eurofins MWG \ Operon, UK. The sequences of the probes with their labelling and the unlabeled oligonucleotides used in this study are shown in Table 2.3.

Probes	Sequences	Homology to consensus
pPZ ASE-Biotin	5'-AAACGT**CAGGAGG**GAAGCAC-3'	6/7
dPZ ASE-Biotin	5'-TTTTAT**AAGGATC**TGGCATT-3'	5/7
FIX-ASE	5'-TTCAGTC**GAGGAAG**GATAGGGT-3'	6/7
hPC-ASE	5'-TCGGAG**CAGGAAG**ACGCTTA-3'	7/7
PZ-PEA3(1)-IRD	5'-GACTGCAAACGTG**CTTCCTG**CTGCATGGCT-3'	7/7
PZ-PEA3(2)-IRD	5'-ATCAGCTTTCC**CTTCCTG**AAGCATCCA-3'	7/7
pPZHNF4α-cy5.5	5'-CCCAGC**CTGGACTTTGG**CTG-3'	9/11
dPZHNF4α-IRD	5'-AAGCTGCA**TGTGACTTTGT**CTG-3'	7/11
PEA3 consensus	5'-GATCTCGAG**CAGGAAG**TTCGA-3'	
HNF4α Consensus	5'-CTCAGC**TTGTACTTTGG**TACAACT-3'	
CREB consensus	5'-AGAGATTGCC**TGACGTCA**GAGAGCTAG-3'	
HNF3B consensus	5'-CAACTAAACCTCTTTTCTTAAATTA-3'	
AP1 consensus	5'-CGCTTGATGACTCAGCCGGAA-3'	
SP1 consensus	5'-ATTCGATCGGGGCGGGGCGAG-3'	

Table 2.3. Fluorescent labeled and unlabeled oligonucleotides used in EMSA.
The oligonucleotides were synthesized by Eurofins MWG Operon. The bold types indicate putative binding site for the proposed protein. Oligonucleotides without bold type represent the non-specific competitors used in the assay. (p) represents the proximal site relative to the translation start site while (d) represents the distal site.

2.2.4.1. Electrophoretic Mobility Shift Assay (EMSA)

The electrophoretic mobility shift assay has been used extensively in DNA-protein interactions studies (Fried and Crothers 1981, Revzin 1989). The principal of this technique relies on the fact that protein bound-DNA migrates more slowly than non-bound DNA in a native polyacrylamide gel. This assay is an easy, rapid, and very sensitive means of detecting protein-bound DNA. Proteins that bind to the probe fragment cause discrete bands of protein-DNA complexes (Buratowski and Chodosh 2001). Before starting the experiment, the forward and reverse oligonucleotides spanning the proposed binding site should be annealed. This was done by mixing 1pMol/µl of each oligonucleotide (forward and reverse) in 25 mM KCl and incubating at 95°C for 5 minutes in a heating block. The mixture was then left to cool slowly in the switched off heating block for the annealing of the oligonucleotides to form double-stranded DNA.

20µl binding reactions containing 10 fmole biotin labelled double stranded DNA or 100 fmole fluorescent labelled double stranded DNA flanking the putative binding site, approximately 4µg /12µg HepG2 nuclear extract (HepG2 NE), 1µg of non-specific competitor Poly dI-dC (Sigma Aldrich, UK) and binding buffer (10mM Tris-HCl, pH 7.5, 50mM NaCl, 1mM DTT, 1mM EDTA, 19%v/v glycerol, 10µl protease inhibitor cocktail) were incubated at room temperature for 20-30 minutes for protein-DNA binding. The samples were size-fractionated through a 5% w/v, 19:1 acrylamide/bis-acrylamide (Sigma Aldrich, UK) mini (8 x 8 x 0.1 cm) gel, using a voltage of ~ 12.5 V/cm, in 1x Tris-Borate EDTA (TBE) buffer (89 mM Tris-borate and 2 mM EDTA). The 5% polyacrylamide gel was prepared by mixing 2 ml of 5X TBE (Sigma Aldrich, UK) with 1.66 ml of the acrylamide/bis-acrylamide (30%), 100 µl of 10% w/v ammonium persulfate (Sigma Aldrich, UK) for polymerisation and 10 µl of Tetramethylethylenediamine (TEMED) (Sigma Aldrich, UK) to initiate the polymerisation in a total volume of 10 ml. The probes were then transferred (for biotin-labeled probes) to nylon membrane (Fluka Biochemika, UK) using an electrophoretic transfer unit (Bio-Rad, UK) using a voltage of ~ 12.5 V/cm for approximately 40 minutes at 4°C in 0.5X TBE. Detection of biotin-labelled DNA was carried out using a chemiluminescent EMSA kit (Perbio, UK) which depended on a streptavidin horseradish peroxidase conjugate (See EMSA optimisation chapter,Section

3.3.4) and the signal was visualised by the UVP BioImaging system (UVP, UK). Fluorescently labelled DNA-protein complexes were detected using the Li-Cor Odyssey Imaging System (Li-Cor Biotechnology, UK Ltd) (See EMSA optimisation chapter,Section 3.3.4). Competition reactions were carried out with the indicated fold molar excess of unlabelled competitor added concurrently with the labelled probe.

2.2.4.2. EMSA confirming the identity of the bound protein

To identify the transcription protein bound to the probe, all the steps in the previous section were performed with a specific antibody against the potential protein added to the binding reaction. The proposed antibody (Santa Cruz, Germany) was added to the binding buffer and HepG2 NE and incubated for 10 minutes at room temperature before adding the probe for a further incubation of 30 minutes. The antibody-protein-DNA complex, which is bigger in size than without the antibody, should normally form a supershift band with lower mobility.

2.3. Gene expression studies

The effect of selected transcriptional modulators on the expression of endogenous PZ was determined in the whole cell context by quantitative real-time PCR (qPCR), using appropriate transcription factor expression plasmids and small interfering RNAs (siRNAs).

2.3.1. RNA preparation from HepG2 cells

The preparation of RNA is one of the essential steps in gene expression studies using qPCR. RNA was extracted from HepG2 cells using a GenElute mammalian total RNA miniprep kit (Sigma-Aldrich, UK). The cells were lysed using the lysis buffer (500-250µl of lysis buffer and 5-2.5 µl β-mercaptoethanol depending on the number of the cells). The lysate was transferred to a GenElute filtration column and centrifuged at 13,300 rpm for 2 minutes to remove cellular debris and shear DNA. An equal volume of 70 % v/v ethanol was added to the filtered lysate for RNA precipitation before loading the sample on to a silica binding column. After binding the RNA to the column, the column

was washed with proprietary wash solution 1 and wash solution 2 from the kit to remove any residual proteins and salts. Finally, RNA was eluted in 50µl of elution buffer solution to release the RNA from the silica. Extracted RNA was treated with DNase1, Amp-D1 kit, (Sigma-Aldrich, UK) to prevent contamination of the sample with genomic DNA. 50µl (average ~ 50µg) of extracted RNA was treated with 5µl of 10X reaction buffer (200 mM Tris-HCl, pH 8.3, 20 mM $MgCl_2$) and 5µl (5 units) of amplification grade DNase 1. The sample was incubated at 37^0C for 30 minutes and then DNase 1 was inactivated by incubation at 70^0C for 10 minutes after adding 5µl of 50 mM EDTA stop solution to chelate the divalent cations required by DNase for activity. The RNA was then stored at -80^0C.

2.3.2. Analysis of RNA by agarose gel electrophoresis

The preparation of good quality RNA is one of the important steps in studying gene expression using qPCR. Degraded or impure RNA will affect the qPCR efficiency and give an inaccurate indication of gene expression (Applied Biosystems literature). Extensively degraded RNA samples should be excluded from real time PCR analysis (Miller *et al.* 2004). Nevertheless, RNA with moderate degradation might be used in gene expression analysis (Schoor *et al.* 2003) with the expression of genes in microarray analysis using a degraded RNA was found to be very constant when the cDNA synthesis primed properly (Xiang *et al.* 2003). Extracted RNA from HepG2 was checked for concentration and purity using the NanoDrop®ND-1000 spectrophotometer (NanoDrop Technologies, USA). The integrity of RNA samples was analysed by size fractionation on 1% w/v agarose gels (See section 2.1.2 for agarose gel preparation). The samples were electrophoresed (10 volt/cm) on the gel for approximately 40-50 minutes, the bands were detected using ethidium bromide (400ng/ml) and then visualized on the UVP BioImaging system. A good quality total RNA sample shows two bands of the ribosomal RNA, 18S and 28S.

2.3.3. cDNA synthesis

A single-stranded cDNA was transcribed using the iScript cDNA synthesis kit (Bio-Rad, UK). In each 20µl cDNA synthesis reaction 500 ng RNA, 10-14µl

ddH_20, 4µl of iScript buffer and 1µl of reverse transcriptase (1U) were mixed. Reaction tubes were incubated at 25^0C for 5 minutes, for efficient annealing of the random primers and then heated to 42^0C for 30 minutes, the optimum temperature for reverse transcriptase activity and finally heated to 85^0C for 5 minutes to stop the reaction. These steps were carried out in a thermocycler. All cDNA samples were stored at -20^0C before being used for real-time PCR.

2.3.4. Real-time PCR

Real-time PCR is becoming the major technique used in the detection and quantification of gene expression profiles (Deepak *et al.* 2007).

The great advantages of real-time PCR have enabled this technique to be one of the most broadly used methods in the field of gene expression quantification (Wong and Medrano 2005). The kinetics of this technique comprise four major phases: the early ground phase, early exponential phase, log-linear growth (exponential) phase and plateau phase (Fig 2.2). The quantitative information for the expression of a gene of interest is presented relative to the expression of a reference gene (Tichopad *et al.* 2003). The use of the terms cycle threshold (Ct), at which fluorescence for a gene is detectable above background, or Crossing point (Cp), or take-off point (TOP) depends on the instrument being used, but all have the same meaning, which should be referred to as quantification cycle (C_q) according to the Real-time PCR Data Markup Language (RDML) (Lefever *et al.* 2009).

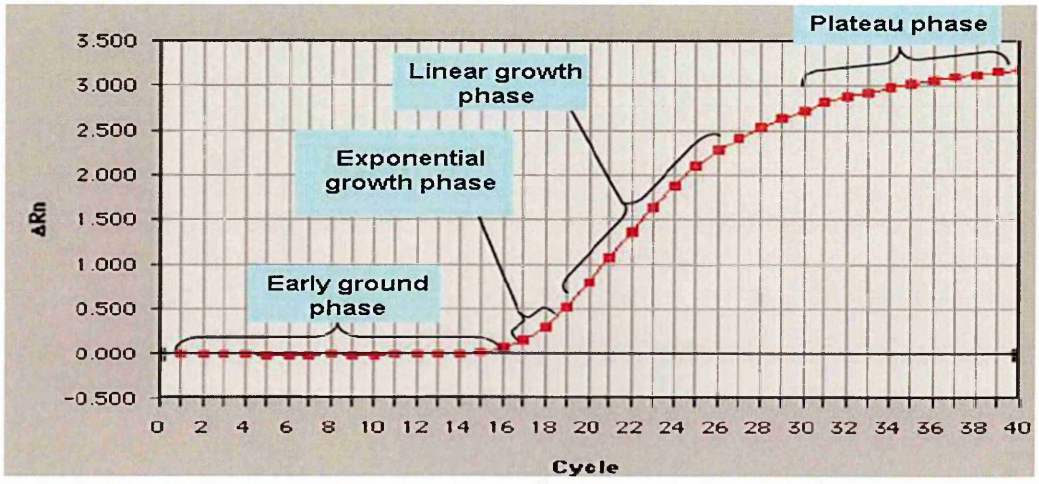


Fig 2.2. Real time PCR amplification curve showing the four major phases.
The graph shows the accumulation of the emitted fluorescence at each PCR cycle. The Y-axis indicates the magnitude of signal produced during PCR and the X-axis represents number of

2.3.4.1. Primer design and efficiency

The primers used to amplify the cDNA synthesized from the mRNA of the gene under investigation were all designed to span at least one intron-exon boundary in the genomic DNA locus encoding the gene of interest. This was to ensure amplification of the cDNA, not the genomic copy of the gene. Short amplicons are recommended as they are more likely to denature at 95°C, allowing the primers and probes to bind their complementary sequence and hence produce efficient amplification (Bustin 2000). All the primers used in the qPCR procedure were therefore designed to produce a product of 90 to 170 base pairs. The designed primers were synthesized by Eurofins MWG Operon (Eurofins MWG Operon, UK) for SYBR green PCR and by Applied Biosystems (Applied Biosystems, UK) for TaqMan PCR. The comparative Cq qPCR (qPCR) is not accurate unless the amplification efficiency of both target and reference genes is highly similar and close to 100% (Liu and Saint 2002). The primer efficiency was calculated by diluting the cDNA template with serial dilutions (undiluted, 1:10, 1:100, 1:1000). PCR reactions for the primer of interest were carried out in triplicate. The Cq values for each dilution were determined and plotted against the logarithm of the cDNA concentrations. The slope was derived from the standard curve and used in the equation:

Efficiency = $-1 + 10^{(-1/\text{slope})}$

The primers which were used in the current study are shown in Table 2.4.

Gene	Forward primer	Reverse primer	Accession number	Intron-exon boundaries	Primer efficiency
PZ	5'-TGGTGGTGTT ATAATACGGGA-3'	5'-GGTCTTGGCT CGTTCTGTTAA-3'	NM_003891	1	86%
PZ(2)	5-'AGGAAAAGAC TTCTGTGGTGG-3'	5'-TGCACATGG ACGTGCGTTAT-3'	NM_003891	1	98%
PEA3	5'-TGAGAAACCT CTGCGACCATT-3'	5'-CTCTCGAAA TGCACCGACC-3'	NM_001986	1	96%
AR	5'-CAATGTCAAC TCCAGGATGCT-3'	5'-TGATTTGGA GCCATCCAAAC-3'	NM_001011645	1	88%
GR	5'-TTGGAGGATC ATGACTACGCT-3'	5'-ATCTCCACC CCAGAGCAAAT-3'	NM_001018077	1	99%
OR	5'-CAAATGCTAC GAAGTGGGAAT-3'	5'-AAGGTTGGC AGCTCTCATGT-3'	NM_001122741	1	96%
UBC	5'-ATTTGGGTC GCGGTTCTTG-3'	5'-TGCCTTGACA TTCTCGATGGT-3'	M26880	1	95%
YWHAZ	5'-CCGCCAGGA CAAACCAGTAT-3'	5'-ACTTTTGGTACA TTGTGGCTTCAA-3'	NM_003406	0	95%
GAPDH	5'-GGCATGGACT GTGGTCATGAG-3'	5'-TGCACCACC AACTGCTTAGC-3'	NM_002046	1	96%
Beta Actin	5'-AAGGGACTTCC TGTAACAATGCA-3'	5'-CTGGAACGGT GAAGGTGACA-3'	NM_001101	0	94%

Table 2.4. Oligonucleotide primers used in reference gene validation and SYBR Green chemistry quantitative PCR. The primers were designed to produce amplicon crosses two exons using Web primer website and synthesized by Eurofin MWG Operon. The primers were designed using the accession numbers (National Centre of Biotechnology information) of each gene which indicated in the table. The primers of the house keeping genes were taken from Vandesompele *et al.* 2002. YWHAZ is tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide. GAPDH is Glyceraldehyde-3-phosphate dehydrogenase while UBC is a ubiquitin C. PZ, AR, GR and OR are protein Z, androgen receptor, glucocorticoid receptor and oestrogen receptor, respectively. Number 0 indicates the primers spanning within the same exon while number 1 indicates the primers which span an interon-exon boundary. The efficiencies of the primers were calculated by the equation shown above after deriving the slope from standard curve plotted for the logarithm of different cDNA concentrations against the Ct value for each concentration.

2.3.4.2. Validation of reference gene

Reference genes are needed to perform gene expression studies using qPCR and those genes should be selected carefully as their expression level may vary between different tissues and under different conditions (Cicinnati *et al.* 2008). The house keeping gene used in qPCR should be referred to as the reference gene according to the Real-time PCR Data Markup Language (RDML) (Bustin *et al.* 2009). The reference gene was selected from a set of four house keeping genes (β-actin, GADPH, UBC, and YWHAZ) with primer seqeunces taken from Vandesompele *et al.* (Vandesompele *et al.* 2002).

The initial assessment of stable reference gene expression was done by performing qPCR for the set of house keeping genes using transfected samples. This allowed the identification of the gene with the most stable

expression in the presence of an over-expression of the HNF4α transcription factor. This reference gene was also found to be stably expressed in response to other treatments by assessing the consistency of the raw Cq values. The location of the primers and the sequence of the house keeping genes and other genes used in qPCR are shown in Table 2.4.

2.3.4.3. Different chemistries in real time PCR

A selection of different fluorescent chemistries can be used to correlate PCR product concentration with fluorescent intensity (Higuchi *et al.* 1993). Two PCR chemistries were used in this project.

SYBR green I chemistry

SYBR green I, a DNA dye, is a fluorophore used in qPCR that is more cost-effective when compared with fluorescent probes. The SYBR green chemistry of qPCR is based on the fluorescent emission of SYBR green upon binding to double stranded DNA (Gibson 2006) which is proportional to the amount of PCR amplified product in every cycle (Fig 2.3) (Applied Biosystems). SYBR green dye interacts with the minor groove of the DNA and its high binding affinity to dsDNA is due to the two positive charges, from one of the aromatic systems of the cyanine, and the protonation of the amino group, that it possesses (Zipper *et al.* 2004). One advantage of SYBR green chemistry is not requiring a specific fluorescently labelled probe for each amplification target. However, the disadvantage is that all double stranded DNA products will contribute to the fluorescent signal, including non specific amplicons and primer dimers (Newby, Hadfield and Roberto 2003). However, although it is considered as a non-specific dye, SYBR green is ideal for use in real-time qPCR in the absence of interfering DNA such as primer-dimers which compete with the specific product and might lead to unreliable results (Kubista *et al.* 2006). In order to identify these potential non-specific products a melt curve analysis is carried out at the end of the amplification. The Melt curve is generated by monitoring the fluorescence from the reaction as a function of temperature. A signal from the desired amplicon would be seen as a single peak on a derivative plot of fluorescence against temperature (Ririe, Rasmussen and Wittwer 1997).

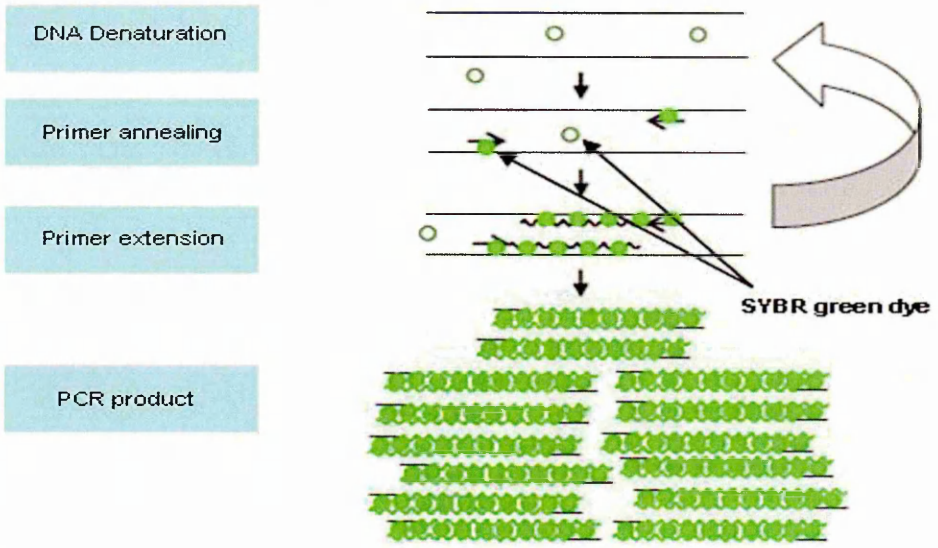


Fig 2.3. Real time quantitative PCR cycles using SYBR green I dye.
The schematic chart shows the different steps from denaturation of the dsDNA to amplification of the product. Free SYBR green dye is unfluorescent which starts fluorescing upon binding the DNA template. The more fluorescence emitted the much product amplified.

A relatively good melt curve, as an example, was shown in the current study for the reference gene UBC and the target PZ gene (Fig 2.4). This was achieved by increasing the temperature of the amplicon slowly above its melting temperature, which depends on the nucleotide composition. The amplification of artefacts normally melt at lower temperatures compared with the specific amplicon (Bustin 2000). The qPCR was performed by mixing 10µl of (2X) power SYBR green mix (Applied Biosystems, UK), 10 pMol (1µl) of each primer and 6µl of DNase free water. 2µl of the tested cDNA was added to the reaction mixture in the 48-well plate which was placed in the StepOne v2.1 PCR instrument (Applied Biosystems, UK) to measure the relative expression of PZ gene. The housekeeping gene UBC (See Table 2.4) was used as an endogenous reference to normalise data. The PCR cycles were as follows: 95°C for 15 minutes to activate the SYBR green mix followed by 40 cycles of the following steps; 95°C for 15 seconds to denature the cDNA, 55°C for 15 seconds to anneal the primers, and 72°C for 30 seconds to extend the primers by the Taq polymerase.

Melt curve: denaturation at 95°C for 15 seconds, annealing at 60°C for 1 minute increasing in 0.3°C increments.

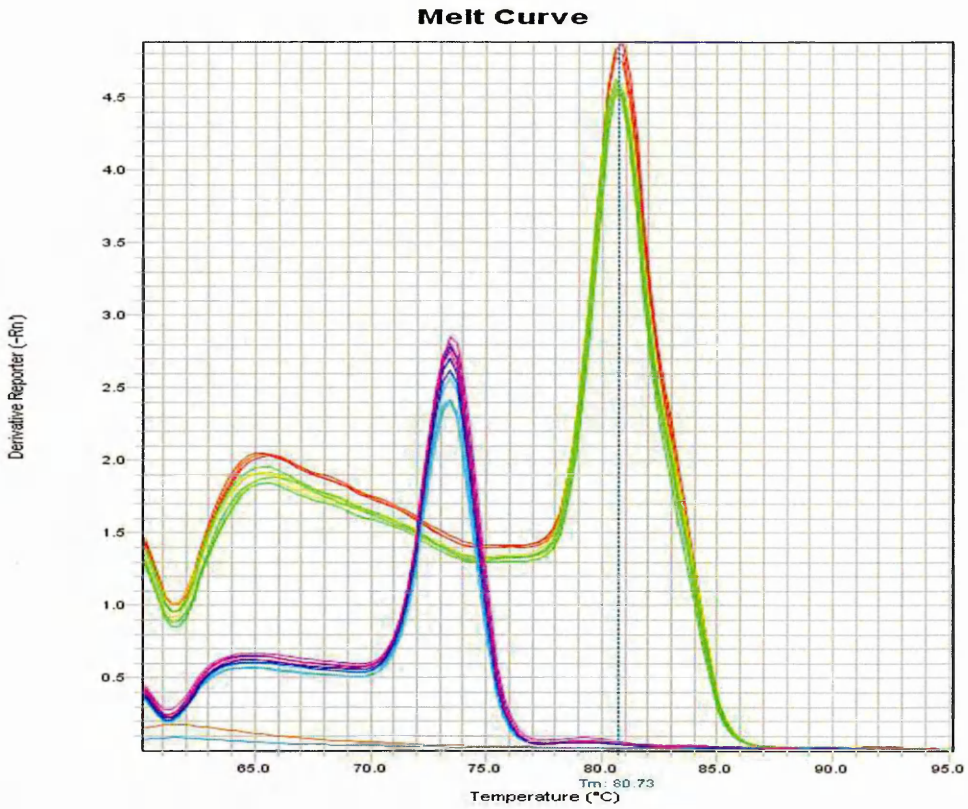


Fig 2.4. Melt curve of PZ and UBC with transfected and non transfected samples. Green and red lines represented the UBC melt curve with untransfected and transfected samples, respectively. Dark blue and light blue lines represented the PZ melt curve with untransfected and transfected samples, respectively. No contamination or primer dimers were detected which would be characterised by an additional shoulder in addition to the main peak.

TaqMan chemistry

In this chemistry an oligonucleotide (probe) labelled with fluorescent dye at the 5' end and a quencher dye at the 3' end are used. During PCR (Fig 2.5) the Taq DNA polymerase cleaves the fluorescent dye from the quencher in the hybridized probe causing an increase in the fluorescence intensity (Livak *et al.* 1995). The 5' endonucleolytic activity of Taq DNA polymerase is exploited during real time PCR amplification in the cleavage of the TaqMan probe causing a sequence-specific detectable signal from the separation of fluorophore and the quencher (Gibson 2006). Each 20µl of qPCR reaction contained 10µl of (2X) TaqMan® fast universal PCR master mix (Applied Biosystems, UK), 10 picomole of the probe with the two primers of the proposed gene, 7µl of Dnase free water and 2µl of the cDNA was added later to the mixture in the 48-well plate. An endogenous reference gene was used to normalise data.

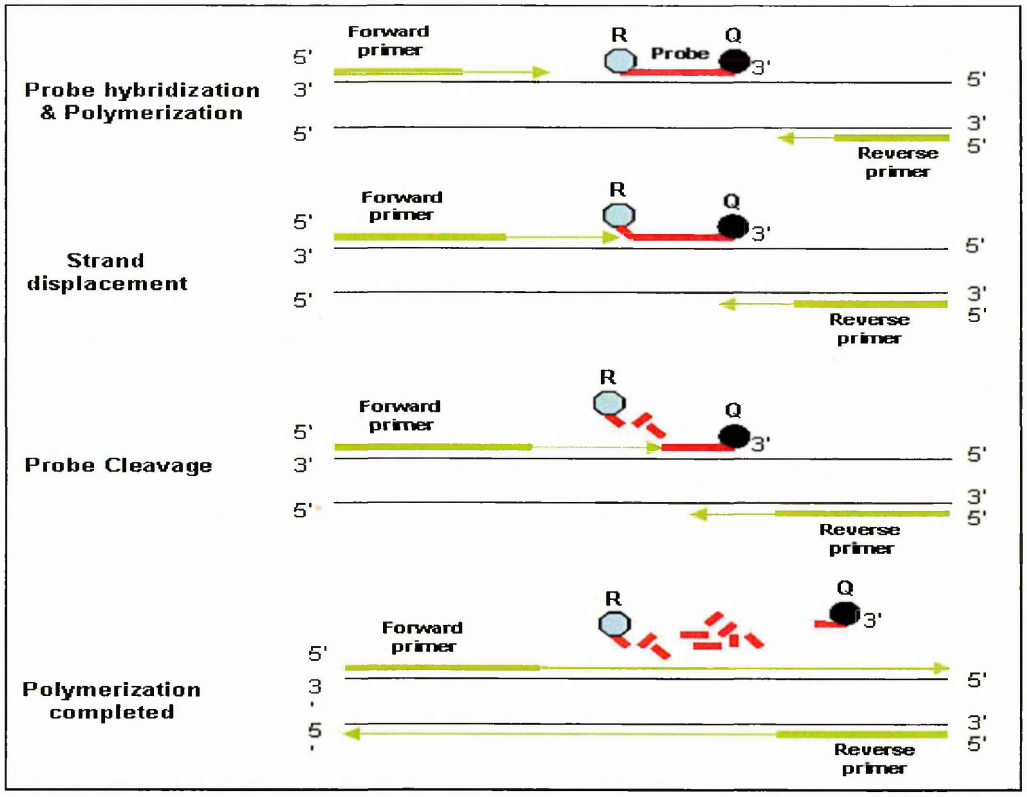


Fig 2.5. Real time PCR steps using TaqMan chemistry.
R is reporter dye and Q is quencher dye. After the probe hybridization and the polymerisation the Taq polymerase activity cleaves the probe from the quencher. The intensity of the fluorescence increases as the free reporter dye is no longer quenched.

The PCR cycles were as follows:

95°C for 20 seconds as a holding stage followed by 40 cycles of the following steps; denaturation at 95°C for 1 second and annealing at 60°C for 20 seconds.

2.4. Cell culture, validation, transfection and stimulation

2.4.1. Culturing of HepG2 cells

The restriction of using normal somatic human cells in research studies and biotechnology applications, in part, is due to their finite proliferative potential before becoming senescent (Hayflick and Moorhead 1961). For this reason and the difficulty in acquiring human tissue samples, immortalised human cell lines have been commonly used in research assays, despite the fact that they are abnormal cells and might show certain abnormal aspects. There are many means by which immortalized cell lines can be created including tumor-derived

cell lines, DNA tumour virus transduction (Katakura, Alam and Shirahata 1998) and induction of telomerase activity (Yeager and Reddel 1999). HepG2 cells (ECACC, UK), which have been used in this study, were grown in MEM medium with 10% v/v fetal calf serum, 2mM L-glutamine, 50 unit of penicillin and 50µg of streptomycin per ml and 1X non essential amino acids (Invitrogen limited, UK) and incubated in a humidity controlled environment at 37°C / 5% CO_2. The cells were passaged by rinsing first with 1X phosphate buffered saline (Invitrogen limited, UK) which contains monobasic potassium phosphate, sodium chloride and dibasic sodium phosphate. The cells were then detached using 5 ml of Trypsin/EDTA (0.5 g/l Trypsin and 0.2 g/l EDTA) by incubating at 37°C and 5% CO_2 for approximately 5 minutes. Cells were recovered by agitating the flask, transferring to a 50 ml centrifugation tube where an equal volume of complete medium was added to neutralize the trypsin, and centrifugation at 1000 rpm (400 g) for 5 min. The recovered cells were re-plated at a density of approximately $1.2x10^5$ cells/ml. The cells take approximately 7-9 days to form a confluent monolayer, until then the growth occurs in islands (Fig 2.6). Cells were passaged when 80 to 90 % confluent. HepG2 stocks were prepared by suspending 1-2 $x10^6$ cells in 1ml of fetal calf serum supplemented with 10% DMSO with subsequent storage in liquid nitrogen.

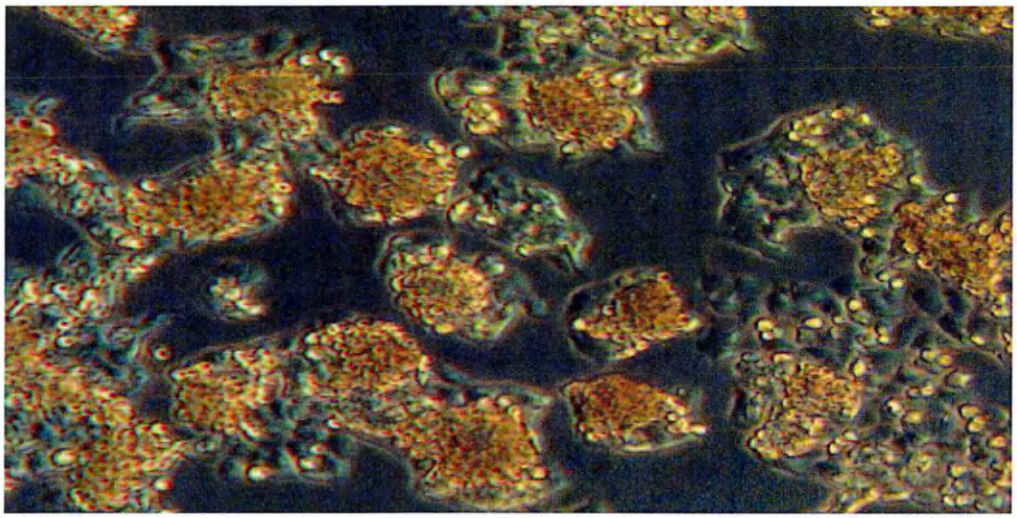


Fig 2.6. HepG2 cells after 5 days from passage.
The cells are growing in MEM medium supplemented with fetal calf serum, non essential amino acid, glutamine and pencillin streptomycin antibiotic. The cells are normally growing in islands before they become confluent.

2.4.2. Determination of relevant protein expression levels in HepG2 cells

To validate HepG2 cells as a suitable model for studying the liver specific expression of the proteins under investigation in this study, they were first compared with normal liver tissue for the expression of these proteins. RNA prepared (See Section 2.3.1) from HepG2 cells using a GenElute mammalian total RNA miniprep kit (Sigma-Aldrich, UK) and RNA from normal adult (64 years old male) liver tissue (AMS biotechnology Ltd, UK) were reverse transcribed (See Section 2.3.3) to cDNA using the iScript cDNA synthesis kit (Bio-Rad, UK). A qPCR was performed using the cDNA templates with designed primers for the candidate protein transcripts (section 2.3.4.3). The expression of candidate proteins in HepG2 was normalised to liver tissue and the expression of the UBC reference gene.

2.4.3. Preparation of transcription factor expression plasmids for transactivation studies

2.4.3.1. Transformation of E. coli competent cells

Competent cells were transformed using expression plasmids (Fig 2.7) encoding the appropriate transcription factors (OriGene Technologies Inc, USA). Expression plasmids for steroid hormone receptors were obtained as a gift from Dr. Peter Winship. 70µl competent cells (JM109, Promega, UK) were transferred to an ice chilled tube and 5 ng of the expression vector added. The tube was then returned immediately to ice for 10 minutes and then heat-shocked for 40 to 50 seconds at 42°C to internalise the plasmid. The mixture was then incubated on ice for a further 2 minutes to bring the temperature back down. 900 µl of L-broth medium was then added and the cells further incubated with shaking for 60 minutes at 37°C to allow the bacteria to recover and to express the antibiotic resistance marker encoded by the plasmid. The L-broth medium was prepared by dissolving 20 g of L-broth powder (Sigma Aldrich, UK) in one litre of distilled water, autoclaved and then stored at 4°C. 100µl of the transformation reaction was spread on a plate with L-broth agar medium containing 80 µg/ml ampicillin, to allow the growth of only the bacteria which are resistant to ampicillin, incubating at 37°C overnight. The L-broth agar plates with the 80 µg/ml ampicillin were prepared by suspending 35 g of the agar in

one litre of distilled water and the ingredients were dissolved and sterilised by autoclaving. 80 mg ampicillin powder (Sigma Aldrich, UK) were dissolved in 1 ml distilled water and sterilised by passing through a filter unit with a pore size of 0.22 µm. The prepared sterilised ampicillin was then added to one litre of the autoclaved agar after it had cooled down to about 50-55°C. The agar was then poured into the plates, left to cool and stored at 4°C to be used.

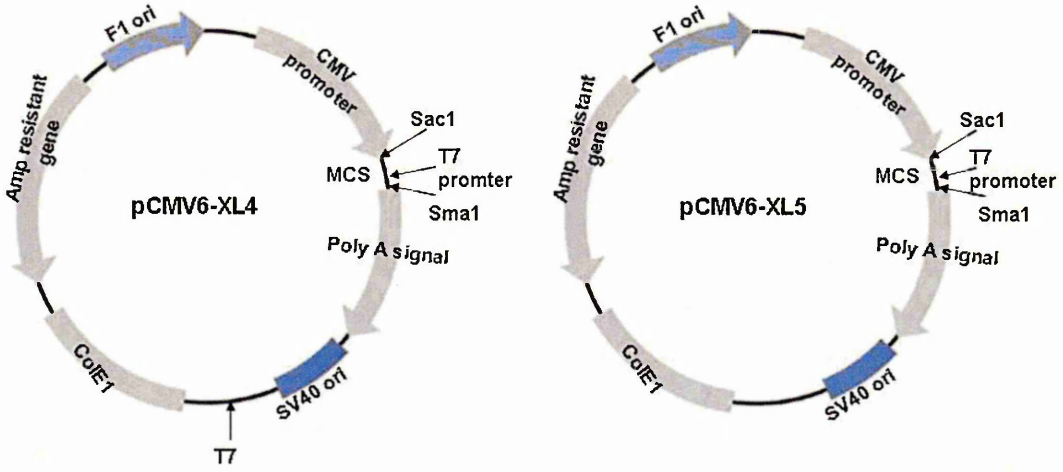


Fig 2.7. Physical Maps for the expression plasmids used in HNF4α and PEA3 transfection. The expression plasmid pCMV6-XL4 (~ 4.7 kb) with an inserted HNF4α DNA of 3.6 kb was used for HNF4α transfection of HepG2 cells while the expression plasmid pCMV-XL5 (~ 4.5 kb) with an inserted PEA3 DNA of 2.2 kb was used for PEA3 transfections. MCS, ColE1, F1 ori and CMV promoter are multiple cloning site, the bacterial origin of replication, the filamentous phage origin of replication and the cytomegalovirus promoter, respectively. The ampicillin resistant gene is used to confer the selection of the plasmid in *E. coli*.

2.4.3.2. Growing up and purification of expression plasmids

A single colony from the L-broth agar plate was inoculated in 100ml of L-broth medium containing 80µg/ml ampicillin (100 µl of the 80 mg/ml sterilised ampicillin were added to 100 ml of the medium) and incubated shaking at 37^{0}C overnight. The bacterial cells were harvested by centrifugation at 6000 x g and then purified using a midi kit for plasmid DNA purification (Qiagen, UK). The principle of this kit is based on the anion-exchange resin binding of plasmid after alkaline lysis. The harvested bacteria were resuspended in buffer P1, containing RNase to digest the bacterial RNA, then buffer P2 was added, containing sodium hydroxide to denature the chromosomal DNA and SDS to disrupt the cell membrane. Buffer N3 was then added which returned the pH to

neutral in the presence of guanidine hydrochloride, causing the precipitation of the denatured proteins, chromosomal DNA and the detergent. The lysate was then centrifuged at ≥ 20,000xg for 30 minutes at 4^0C to pellet the precipitated cell debris. The supernatant was applied to an equilibrated Qiagen-tip for DNA binding and the DNA then washed and finally eluted. The eluted DNA was precipitated with isopropanol and centrifugation at ≥ 15,000x g for 30 minutes at 4°C. The DNA pellet was washed using 70% v/v ethanol, air dried and redissolved in 50µl of DNase free water.

2.4.3.3. Digestion of plasmids by restriction enzymes

Plasmids used in transfections were verified using FastDigest™ restriction enzymes and 1X FastDigest™ buffer (Fermentas life science, UK). 1 unit of the FastDigest enzyme (FDU) was added to 2µl (up to 0.8µg) of the plasmid in 1X FastDigest buffer in a 20µl total reaction volume. FDU is the amount of the enzyme needed to digest 1µg of lambda DNA-HindIII in 5 minutes at 37°C in 1X FastDigest buffer. The mixture was incubated at 37°C for 5 minutes and then the enzyme was inactivated by heating at 65°C for 5 minutes. To confirm the digestion 3 to 5 µl of the mixture were size-fractionated (10 volt/cm) on a 1% w/v agarose gel and bands detected by ethidium bromide staining (See Section 2.1.1).

To determine the restriction sites, the inserted cDNA clone of the candidate expression plasmid was characterised for recognition sequences using the webcutter tool (http://bio.lundberg.gu.se/cutter2). The appropriate enzyme, which has one or two cut position sites, was then used to digest the candidate plasmid and the resultant bands were size-fractionated to verify the identity of the expression plasmid.

2.4.4. Transfection of HepG2 cells by expression vectors

Transfections of HepG2 cells using expression vectors encoding certain transcription factors were performed to check the effect of the overexpression of those transcription proteins on PZ mRNA expression levels. Tfx™-20 reagent (Promega corporation, UK) was used for this purpose which is based on the formation of multilamellar vesicles, which associate with nucleic acids, in the

transfer of the nucleic acids into cells by fusion with the cell membrane. This transfection reagent has been used extensively in the transfection of HepG2 cells (Carrasquillo *et al.* 2010, Takagi *et al.* 2008, Wong *et al.* 2007, Hattori and Maitani 2005). 1.25 µg of the proposed expression plasmid was added to 37^0C prewarmed cell growth medium and vortexed. Tfx-20, with a charge ratio 4:1 (reagent: DNA) was then added to the medium with the plasmid and vortexed immediately. The volume of medium, plasmid and the Tfx reagent was calculated using the online Promega transfection calculator. The transfection mixture (total volume 1 ml) was incubated at room temperature for 15 minutes and added to 80% confluent cells in a 6-well plate after removal of the old medium and incubated for a further hour at 37°C and 5% CO_2. 2 ml of fresh complete medium were added to the transfection mixture in the wells and incubated for a further 48 hours.

For determination of the PZ response to steroid hormones, 2 ml of fresh complete D-MEM phenol free medium (Invitrogen limited, UK) containing charcoal stripped fetal calf serum (Invitrogen limited, UK) was added to the transfection mixture as well as the proposed hormone (1x 10^{-8} M 17-β Oestradiol, 2.5 x 10^{-8} M Dihydrotestosterone, 1x 10^{-6} Dexamethasone final concentration in the HepG2 culture wells transfected with oestrogen, androgen and glucocorticoid receptors, respectively) and further incubated for 48 hours. Charcoal stripped foetal calf serum was used to remove any endogenous bovine steroid hormones that may have skewed the results.

2.4.5. Transfection of HepG2 with siRNAs

Small interfering RNAs are double stranded RNAs approximately 21 or 22 nucleotides in length with overhanging 3' ends involved in the RNA interference pathway and causing posttranscriptional gene silencing (Elbashir, Lendeckel and Tuschl 2001). Knocking down of the candidate transcription factor was achieved using the Thermo Scientific DharmaFECT® General Transfection Protocol (Dharmacon RNAi Technologies, UK). The HepG2 cells were trypsinized and diluted in antibiotic-free complete medium. 5 x 10^4 of cells were plated in 24-well plates and incubated overnight at 37^0C/ 5% CO_2. For each transfection, two tubes were used to mix the contents. In the first tube 25 µl of 2µM on-targetplus smart pool (siRNA) (Dharmacon RNAi Technologies, UK)

was added to 25 µl of serum-free growth medium while in the second tube 2 µl of the transfection reagent (DharmaEFCT 4) was added to 48 µl of the serum free medium. The contents of each tube were mixed gently before being combined together and mixed gently by careful pipetting up and down. The mixture was incubated for 20 minutes at room temperature. 400 µl of antibiotic-free complete medium was added to the mixture giving a total volume of 500 µl which was then added to the well after removing the culture medium. The cells with the transfection medium were incubated for 48 hours at 37^0C in 5% CO_2. The same was done for the negative (ON-Targetplus Non-Targeting pool) and positive (ON-Targetplus GAPD control pool) control samples (Dharmacon RNAi Technologies, UK). The efficiencies of the knocking down of the candidate protein and its effect on PZ were measured using qPCR (See Section 2.3.4.3).

2.4.6. Stimulation of HepG2 with Lipopolysaccharide

Lipopolysaccharide (LPS) is an outer membrane endotoxin glycoprotein which upon its identification by the body causes initiation of a variety of defence mechanisms (Schumann *et al.* 1990). Some coagulation factors have been found to be acute phase proteins and induced in response to LPS, such as factor VIII (Begbie *et al.* 2000) and tissue factor (Hall, Vos and Bertina 1999). To examine whether PZ is an acute phase protein or not, HepG2 cells were stimulated with lipopolysaccharide (LPS) from *Salmonella typhimurium* (Sigma Aldrich, UK). The cells were grown in 6-well plates to approximately 80% confluency and 10 µg/mL LPS was added to the cells and left for a time course of 30 minutes, 1 hour, 2 hours, 4 hours and 24 hours. RNA was extracted from HepG2 cells with each time point using a GenElute mammalian total RNA miniprep kit (See Section 2.3.1). Each RNA was then reverse transcribed to cDNA using the iScript cDNA synthesis kit (See section 2.3.3). The synthesised cDNA from each RNA sample was used in qPCR (2.3.4.3) to determine the effect of LPS on PZ at different time points.

Chapter 3-Optimisation Of Electrophoretic Mobility Shift Assay

3.1. Introduction

Transcriptional regulation is a process involving a diverse set of protein-nucleic acid interactions at multiple levels in the hierarchical organisation of the genome. One of the primary stages at which gene expression is regulated is by the binding of trans-acting factors to cis-regulatory elements within the promoter regions of genes. The binding of transcriptional regulatory proteins to their specific DNA binding sites needs to be assessed as part of any study of transcriptional regulatory mechanisms. However, the use of *in vitro* protein-DNA binding studies does inevitably lead to a certain loss of context in that some synergistic interactions between binding factors may be lost and/or the influence of chromatin structure and nucleosome positioning will no longer influence binding. However, these *in vitro* assays do provide important information in terms of the ability of particular factors to bind to specific sequences; information that can be corroborated in the higher order context by further analyses.

Therefore, protein-DNA binding studies, which lead to an accurate recognition of transcription factor binding sites, are essential in the understanding of complex gene expression patterns (Morozov *et al.* 2005). The electrophoretic mobility shift assay (EMSA) is one such technique, which has been used extensively in DNA-protein interaction studies (Fried and Crothers 1981, Revzin 1989). This assay, also referred to as the band shift, gel shift and gel retardation assay, provides information about the protein binding specificity and the structural consequences of the protein-DNA interaction by using small amounts of the candidate protein and very simple apparatus (Savery and Busby 1998). The EMSA analyses of the equilibrium binding of the protein to its specific binding sequence is based on physical separation of protein-double-stranded DNA (dsDNA) complexes from free dsDNA probes (Sidorova, Muradymov and Rau 2005). The electrophoretic mobility of a dsDNA is higher compared to the mobility of the dsDNA bound by protein (Carey 1988, Hellman and Fried 2007). The electrophoretic mobility of the protein-dsDNA complex is influenced by many factors, such as, the molecular weight of the protein, the size of the dsDNA (Fried and Crothers 1983), the pH and ionic strength of the electrophoresis buffer (Carey 1988) and the temperature and the concentration of the gel matrix (Taylor, Ackroyd and Halford 1994). Therefore, the

electrophoretic mobility shift assay should be optimized to observe protein-DNA binding.

3.2. Optimisation of EMSAs

EMSAs were first developed using radioactively labelled dsDNA probes and highly purified nuclear protein extracts (Fried and Crothers 1981, Garner and Revzin 1981). This procedure had the obvious advantage, with the use of high energy emitters, such as P^{32}, of a high degree of sensitivity (Hellman and Fried 2007). However, the drawbacks of the handling of radioactive reagents, the limited half-life of the isotopes and the cost of purchase and disposal of such labelled products (Jing *et al.* 2003) led to the development of alternative labelling systems. Two of these alternative methods were utilised and optimised throughout the current study.

The first method used was the Lightshift chemiluminescent EMSA system from Pierce, which relies on the use of biotinylated dsDNA probes, transfer of size fractionated complexes to a membrane, binding of streptavidin linked horseradish peroxidase and a chemiluminescent reaction to detect labelled products. The composition of the binding buffer, the length of the binding reaction incubation and the temperature at which the incubation is carried out are all parameters which can be modified to optimise the binding of the protein within a relatively broad range that encompasses physiological conditions. This chemiluminescent detection system has comparable sensitivity to a similar radiolabelled method (Fig. 3.1).

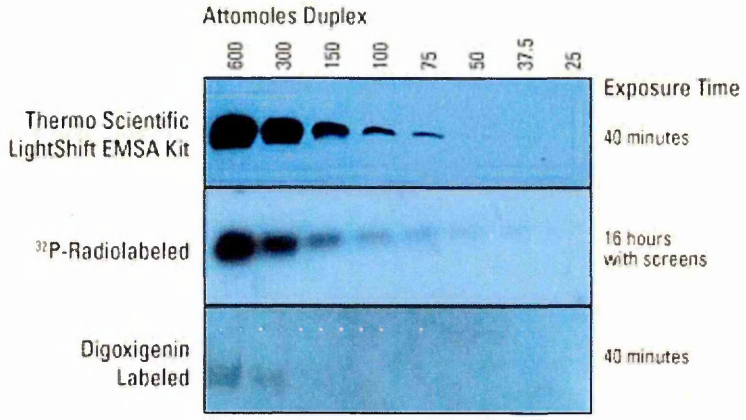


Fig 3.1. Comparison of the sensitivity of various EMSA Methods (Thermo Scientific, UK). Three different EMSAs were performed and detected according to the manufacturer's instructions using different concentrations of the labelled probes and electrophoresed on a 6% polyacrylamide gel. Comparable sensitivity was achieved using the biotin-labelled (chemilumescent kit) and ^{32}P-radiolabelled EMSAs. However, ^{32}P-radiolabelled EMSAs needed a significantly longer exposure time to achieve the comparable sensitivity to the biotin-labelled EMSAs. The signal detected using digoxigenin-labelled probes (another alternative labelling method) was very weak compared to the biotin-labelled probes with the same exposure time.

The second method of EMSA visualisation used in this study was the use of infra-red dye labelled (IRDye) DNA probes (see example assay Fig. 3.2). This system has a number of advantages over the use of radiolabelled DNA probes: similar to those mentioned above when discussing the chemiluminescent system, the IRDye system does not require the use of reagents subject to strict regulatory control with high purchase and disposal costs; IRDyes are non-hazardous and are stable compared to the limited half-life of radioisotopes (Ying, Fourmy and Yoshizawa 2007). In addition, following size fractionation through gel electrophoresis the products may be imaged directly on an IR scanner without the need to take the gel from between the electrophoresis plates; besides the obvious advantages of the time saving this also allows the electrophoresis to be restarted should further separation be required. This method, therefore, has a number of advantages over the radioisotope method and the chemiluminescent detection method with equivalent sensitivity.

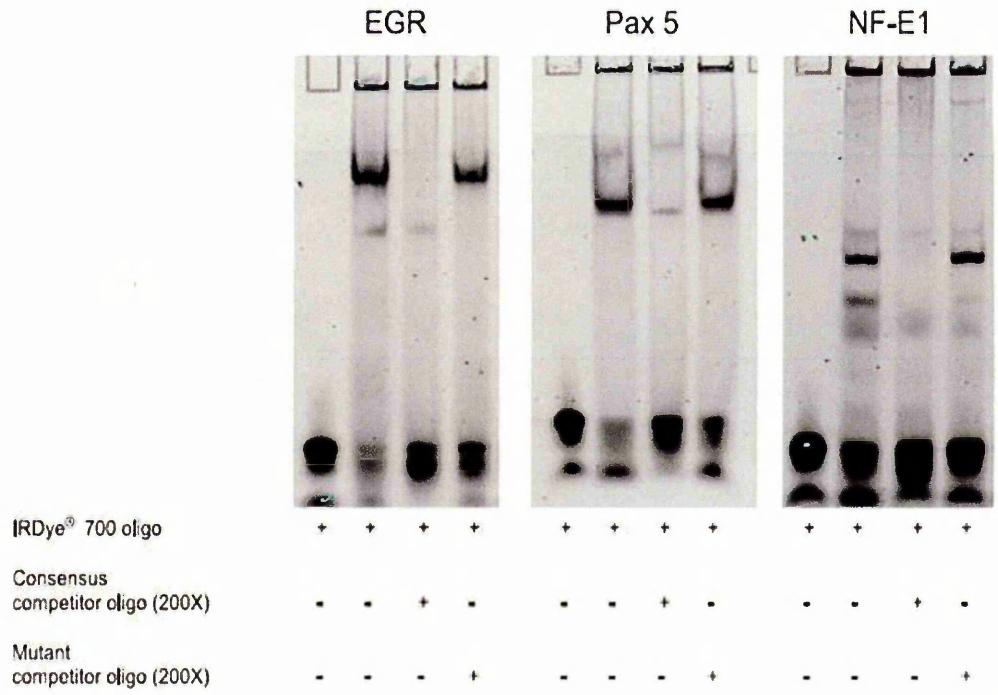


Fig 3.2. IRDye® 700 EMSA. Assays were carried out using consensus oligos for three different transcription factor targets, with and without the presence unlabelled competitor sequences (LI-COR Biosciences, 2006). EGR is the early growth transcription factor. Pax5 is a member of the paired pox family of transcription factors. NF-E1 is the nuclear factor erythroid 1.

In protein-DNA binding assays the quality of the nuclear extract is crucial in achieving optimal results and a number of precautions can be taken in the preparation dependent on the proteins being investigated. For example, in all methods, working on ice or at 4°C will maintain the integrity of the proteins (Holden and Tacon 2010) by limiting protease activity, as will the use of protease inhibitors such phenyl methylsulfonyl fluoride, aprotinin and papain (Shi and Sarna 2005, Weng *et al.* 2004, Stasiolek *et al.* 2000); although it is more usual nowadays to include a commercial protease inhibitor cocktail with the same broad range of action against the various types of protease activity present within the cell (Thermo Scientific and Sigma Aldrich literature). Also, if the activation state of the regulatory protein is being investigated phosphatase inhibitors may be included; this could be important if phosphorylation acts to modulate the DNA-binding domain of the transcription factor or if phosphorylation modulates transactivation potential and this is to be monitored with phospho-specifc antibodies in supershift assays (Hall, Vos and Bertina 1999). The original nuclear extract preparation method of Dignam *et al* (1983)

provides protein specific to the nucleus and of very good quality, when executed appropriately. However, the method is fairly time consuming and quicker and more straight forward methods have been developed that work well with cell culture samples; although for tissue the original Dignam method (Dignam *et al.* 1983) cannot be surpassed. These more time effective methods (Andrews and Faller 1991) may not produce protein that is quite as nucleus-specific, but the final sample is certainly nucleus enriched and provides protein that will function very well in the relatively crude environment of the *in vitro* EMSA (Personal communication with Hall AJ 2008).

There are, therefore, considerable number of variables to consider in optimising EMSA, from the labelling system to the binding reaction buffers, time and temperature of the binding reactions to electrophoresis parameters and detection methods. This chapter discusses the optimisation performed within the project.

3.3. Materials and Methods

3.3.1. Nuclear extract Preparation

The nuclear extract preparation method (Mirmira 2006) was based on a small scale development from the original Dignam (1983) procedure. The proteins were prepared by removing the growth medium from cells grown in culture using aspiration and then washing the cells with phosphate buffered saline (PBS) (Invitrogen, UK) (137 mM NaCl, 2.7 mM KCl, 8 mM sodium phosphate dibasic, 2 mM potassium phosphate monobasic pH 7.4) to remove residual growth medium prior to addition of 0.5 ml of buffer A (10 mM HEPES (Sigma Aldrich, UK) to buffer the solution at a near physiological pH 7.9, 10 mM KCl to facilitate the retention of proteins' functional conformations, 0.1 mM EDTA to chelate divalent cations, particularly magnesium, as these cations are necessary co-factors for a number of enzymes that could impact the quality of the final extract, 1mM dithiothreitol (DTT) (Sigma Aldrich, UK), a low concentration reducing agent, 20 µl protease inhibitor cocktail (Sigma Aldrich, UK) containing inhibitors with a broad specificity for aminopeptidase and serine, cysteine and acid proteases, and 200 µl of 10% of IGEPAL (Sigma Aldrich, UK), a detergent, to enhance cell lysis). The cells were incubated with

buffer A at room temperature for 10 minutes to allow the cells to swell in the hypotonic solution. The cells were then scraped from the surface of the flask and the lysate incubated on ice for 5 minutes in a microcentrifuge tube. The lysate was then centrifuged at 4°C to pellet the nuclei, placed on ice for 3-5 minutes and the supernatant containing the predominantly cytoplasmic fraction and cell debris was removed. 150 µl of hypertonic buffer B (20 mM HEPES (Sigma Aldrich, UK), 0.4 M Nacl, 1mM EDTA, 10% glycerol, 4 µl of the protease inhibitor cocktail (Sigma Aldrich, UK) and 1mM of DTT) was then added to the recovered nuclei and shaken for 2 hours at 4°C in 400 mM NaCl to perform a high salt extraction of the nuclear proteins from the nuclei. The sample was then centrifuged at maximum speed of 16.1 $x10^3$ g in a benchtop microcentrifuge to pellet the nuclear debris and the supernatant (nuclear protein extract) aliquoted into prechilled microcentrifuge tubes and snap frozen in liquid nitrogen and stored at -80°C.

An alternative small scale nuclear extract preparation method developed from the original Dignam method by Andrews and Faller was also used extensively in the current study (Andrews and Faller 1991). Hep G2 cells (approximately 2.38 x 10^7 cells) were scraped into PBS and resuspended in buffer A (10mM Hepes (Sigma Aldrich, UK)-KOH pH 7.9, 1.5mM $MgCl_2$,10mM KCl, 0.5 mM DTT, Protease inhibitor cocktail). The Hep G2 cells were swollen in the hypotonic solution, on ice for 10 minutes, and vortexed for 10 seconds to cause cell lysis. The nuclei were then recovered by centrifugation at 16.1 $x10^3$ g for 10 seconds and resuspended in ice-cold buffer C (20mM Hepes-KOH pH 7.9, 2.5% glycerol, 420mM NaCl,1.5mM $MgCl_2$, 0.2mM EDTA, 0.5mM DTT, 1% (v/v) protease inhibitor cocktail) and incubated for 20 minutes on ice for high salt extraction of the nuclear proteins. The suspension was centrifuged again at maximum speed (16.1 $x10^3$ g) at 4°C for two minutes and the supernatant (nuclear extract) was aliquoted, snap frozen and stored at -80°C. The protein concentration was subsequently determined by the Bradford assay (See Section 2.2.2). The effect of dialysis of the nuclear protein nuclear extract prepared by the Andrew and Faller 1991 method following the final extraction was also assessed. The dialysis was done by placing the HepG2 NE in semi-permeable cellulose dialysis tubing (Sigma Aldrich, UK) and dialysed against a dialysis buffer (20 mM Hepes pH 7.9, 0.1 M KCl, 0.2 mM EDTA, 0.5 mM DTT

and protease inhibitor cocktail) to reduce the salt concentration within the HepG2 NE.

3.3.2. EMSA Probe Preparation

The double-stranded DNA probes were prepared by annealing forward and reverse oligonucleotides, synthesised by Eurofins MWG Operon, UK, that spanned the putative cis-elements of the trans-acting factors. The designed putative cis-elements for candidate proteins used in the current study were predicted by bioinformatic analysis (ALGGEN-PROMO) of the PZ gene promoter (Homo sapiens protein Z, AF440358, National Centre for Biotechnology Information). The probes designed for EMSA in the current study were 20 to 30 base pairs in length and flanked the proposed binding sites; the recommended length of probes is usually between 20 and 40 base pairs (Kako *et al.* 1998). The forward and reverse oligonucleotides for a proposed protein binding site were annealed by mixing 1pMol/µl (final concentration) of each oligonucleotide (forward and reverse) in 25 mM KCl (to promote the annealing of the two negatively charged oligonucleotides) and incubating them at 95^0C for 5 minutes in a hot block, to denature the oligos, prior to switching off the block and leaving them to cool slowly to room temperature to enable the formation of double-stranded binding sites. The same solution (25 mM KCl) was then used for preparing the working concentration of the probe.

3.3.3. EMSA Binding Reactions and Size Fractionation

Once the proposed binding site probe (See Section 2.2.3) and nuclear extract were prepared (SeeSection 3.4.1), binding conditions were modified for optimum protein-DNA binding. Five binding buffers with different buffering agents (HEPES and TRIS) and differing ionic strength were used to optimise the protein-DNA interactions. Buffer-1(as per Fisher Scientific, UK): 10 mM Tris-Hcl (Sigma Aldrich, UK) pH 7.5, 50 mM Nacl, 1mM DTT, 1 mM EDTA, 19% glycerol and 10 µl/ml protease inhibitor cocktail (Sigma Aldrich, UK). Buffer-2 (Groupp and Donovan-Peluso 1996): 23 mM HEPES (Sigma Aldrich, UK) pH 7.9, 113 mM Nacl, 9 mM $Mgcl_2$, 0.23 mM EDTA, 19% glycerol, 4.5 mM DTT, 10 µl/ml protease inhibitor cocktail . Buffer-3 (Pierce, UK): 10 mM Tris, 50 mM

KCl, 1 mM DTT; pH 7.5. Santa Cruz binding buffer (Santa Cruz biotechnology, Germany): 10 mM Tris, pH 7.5, 50 mM NaCl, 1 mM DTT, 1 mM EDTA and 5% glycerol. Sugawara binding buffer (Sugawara *et al.* 2007): 20 mM Tris-HCl, pH 7.9, 4 mM $MgCl_2$, 1 mM EDTA, 0.1 mM DTT (Sigma Aldrich, UK), 10% glycerol (Sigma Aldrich, UK), 0.1% Nonidet P-40 (Sigma Aldrich, UK). Binding reactions also contained 1µg of the non-specific competitor Poly(dI-dC) (Sigma Aldrich, UK), which acts as a sink to bind the majority of the non-specific DNA-binding proteins (alternatives such as sonicated herring testes DNA could have also been used, however, polydI-dC is often chosen for its consistency).

The incubation time and temperature for the binding of the nuclear extract with the probes was also optimised, balancing optimal interaction conditions against the potential for protein degradation. Binding reactions were performed at 37^0C, room temperature, at 4^0C and on ice. These changes were also performed to determine optimal interaction for the antibody-protein binding reactions. The main variables used in optimisation of EMSAs are summarised in table 3.1. A total binding reaction volume of 20 µl was used for both biotinylated and fluorescent EMSAs. However, different amounts of probe and nuclear extract were used in the EMSAs with the different labelling systems. Approximately 4µg and 12 µg of nuclear extract were used with 10 fmole of the biotinylated probe and 100 fmole of the fluorescent probe, respectively. Moreover, different incubation periods for each probe with nuclear extract were performed to optimise the binding reaction. Having carried out the binding reaction a proprietary loading buffer (Pierce, UK) was added to each reaction and EMSAs performed using the biotinylated probes. The samples were predominantly size-fractionated through a 5% w/v, 19:1 acrylamide/bis-acrylamide (Sigma Aldrich, UK) mini (8 x 8 x 0.1 cm) gel, using a voltage of ~ 12.5 V/cm, in a Tris-Borate EDTA (TBE) buffer (Sigma Aldrich, UK) (89 mM Tris-borate and 2 mM EDTA). Although other buffer systems such as Tris-Glycine EDTA (Sigma Aldrich, UK) (50 mM Tris, pH 7.5, 0.38 M glycine and 2 mM EDTA) were tested in order to optimise separation conditions. In EMSAs using IRDye probes loading buffer was not added to conserve the composition of the binding buffer and to avoid the interference between the fluorescent signal of the labelled probes and the dyes within the loading buffer. The

samples were size-fractionated using the same parameters performed with the biotinylated probes.

Amount of probe	10 and 20 fmole of biotinylated probes and 100 fmole of IRDye probes.
Nuclear extract preparation	- Protocol adopted from Raghu Mirmira (Mirmira 2006). - Protocol adopted from Andrews and Faller (Andrews and Faller 1991), with and without dialysis. - 4 and 12 µg of the nuclear extract were used with biotintlated and IRDye probes, respectively.
Temperature	37°C, room temperature, 4°C and ice.
Incubation time	- Probe incubation: 15, 20 and 30 minutes. - Antibody incubation: 10, 20, 25, 30, 90, 120 minutes and overnight
Binding buffers	- Binding buffer-1 (Fisher Scientific,UK). - Binding buffer-2 (Groupp and Donovan-Peluso 1996). - Binding buffer-3 (Pierce, UK). - Santa Cruz binding buffer (Santa Cruz biotechnology, Germany). - Sugawara binding buffer (Sugawara *et al.* 2007)

Table 3.1. Different variables used in optimisation of EMSA. Five different binding buffers and three different prepared nuclear extracts with different amounts were used to optimise EMSAs. Moreover, four different temperatures with different incubation time of the antibody and different amounts of the probe were also performed.

3.3.4. Chemiluminescent Detection Of Biotinylated Probes

Following electrophoresis and size fractionation the probe and probe-protein complexes were transferred to a positively charged nylon membrane (Fluka Biochemika, Sigma Aldrich, UK) by electrotransfer. This was done by soaking the membrane in 1X Tris-borate EDTA (TBE) for at least 10 minutes and sandwiching it with the gel. The membrane-gel sandwich was placed in an electrophoretic transfer unit (Bio-Rad, UK) and electrophoresed in pre-cooled 1X TBE. Following transfer of the molecules they were cross-linked to the membrane using a uv transilluminator with a 312 nm wavelength bulb for 10-15

minutes. The biotin labelled DNA was then detected using the Lightshift chemilumescent detection kit (Pierce, UK) according the manufacturer's instructions. First the nylon membrane was blocked using 20 ml of a proprietary blocking buffer (Pierce, UK), to reduce the background signal resulting from the non-specific binding of the streptavidin complex to the membrane. This was followed by the addition of 20 ml of blocking buffer/stabilized streptavidin-horseradish peroxidise conjugate solution. In this step the streptavidin binds to the biotin-labelled probes with high affinity. The membrane was then washed three times with 20 ml of 1x washing buffer (Pierce, UK) with gentle shaking for 5 minutes at room temperature, again to reduce the non-specific background signal, before adding a substrate equilibration buffer to the membrane and further incubating for 5 minutes with gentle shaking. The membrane was then incubated according to the manufacturer's instruction in a substrate working solution containing luminol enhancer solution and stable peroxide solution for 5 minutes. In this step the enzyme horseradish peroxidase reacts with the substrate in the presence of the peroxide solution and produces light. The membrane was then wrapped in cling film to protect the membrane and the signals detected using a UVP BioImaging system (UVP, UK).

3.3.5. IRDye Detection

Following size fractionation the IRDye labelled probe and probe-protein complexes were detected using the Li-Cor Odyssey system (LI-COR Biosciences, UK). The electrophoresis gel, still between the glass plates, was placed on the Li-Cor Odyssey scanner which detects emission at both 700 nm and 800 nm from IRDye labelled probes. Detection of the IRDye labelled probes with the gel still within the plates has the advantage that, should it be required, the electrophoresis could be recommenced to separate the complexes further. The Li-Cor Odyssey system also allows for the quantification of the fluorescent signal.

3.3.6. EMSA analysis

Protein-DNA complexes are larger compared to free DNA pobes. Therefore protein-DNA complexes in EMSA were defined by bands with a lower mobility (shifted) compared to the mobility of the bands of the free probes. In competitive EMSA the decrease or loss of complex bands by unlabelled non-specific oligos (Oligos with different sequences) revealed non-specific binding of the protein to the DNA (protein-DNA complex). Maintenance of protein-DNA complex bands with the unlabelled non-specific oligos showed the specific binding of the protein to the DNA. Complexes which result from binding of antibody to protein-DNA complexes, are also larger than free DNA and than protein-DNA complexes. Therefore, bands with lower mobility in the presence of antibody compared to the bands in the absence of the antibody (supershifted) were considered as antibody binding to protein. The intensity of the bands in competitive EMSA and supershift assays was measured using the software of the BioImaging system, for biotinylated probes, and using the software of the Li-Cor Odyssey system for the IRDye probes. This was done by drawing a rectangle around one band which then was copied an pasted to the other proposed bands in order to have the same area around each band. The intensity of the bands was quantified by the system. The intensity of competed band (in the presence of unlabelled oligo) or antibody complex band were divided by the intensity of the band from the binding reaction (in the absence of competitor and antibody) to show the fold increase or decrease of band intensities.

3.4. Results

3.4.1. Biotinylated Probe EMSA

The detection system originally used in the EMSA analyses was the LightShift chemiluminescent assay, (SeeSection 3.3.4), with an incubation of probe with HepG2 nuclear extract (Mirmira 2006) for 20 minutes at room temperature. The EMSA using the original protocol from Pierce detected three high molecular weight bands (A-C) when HepG2 nuclear extract was present in the binding reaction. A fourth band (D) was observed using the pPZ ASE probe (lanes 2 and 3, Fig 3.3) which was absent when using the dPZ ASE probe (lanes 4 and 5, Fig 3.3).

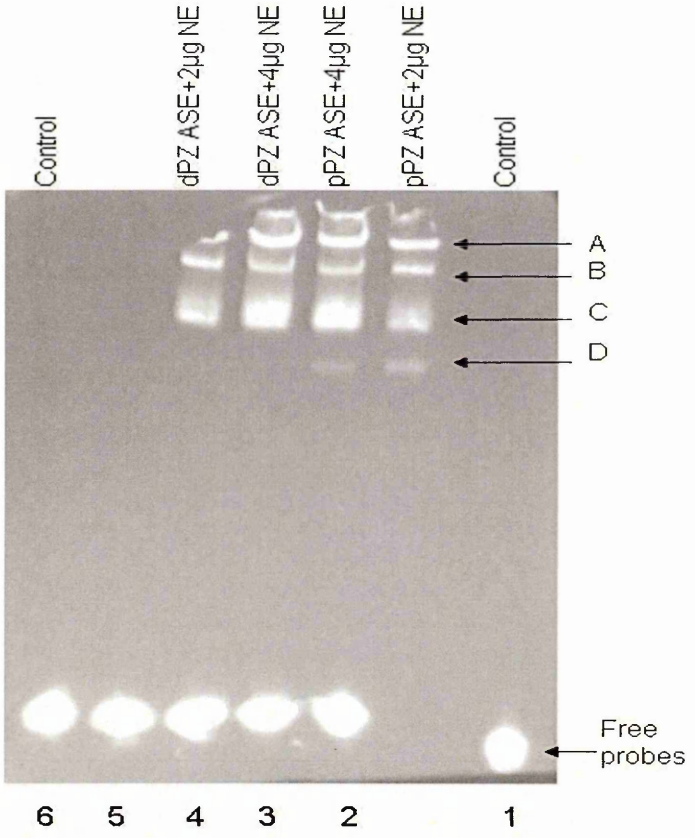


Fig 3.3. EMSA using pPZ and dPZ ASEs. Lane 1 (control) contains pPZ ASE probe without HepG2 nuclear extract (HepG2 NE). Lane 2 and 3 contain pPZ ASE probe with 2 µl and 4 µl HepG2 NE, respectively. Lane 4 and 5 contain dPZ ASE probe with 4 µl and 2 µl HepG2 NE, respectively. Lane 6 (control) contains dPZ ASE probe without HepG2 NE.

Competitive EMSA was carried out to confirm the specificity of binding following the standard LightShift protocol using the pPZ ASE probe (Fig 3.4).

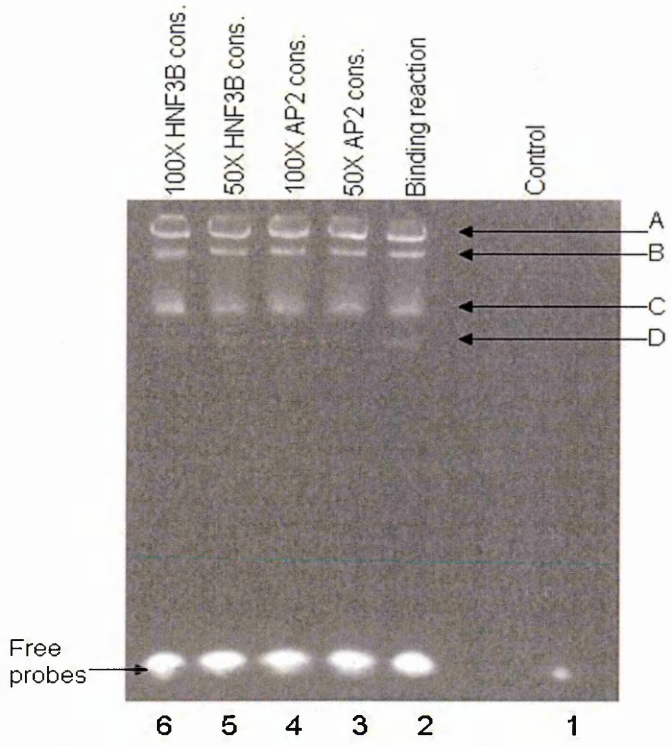


Fig 3.4. Competitive EMSA using Bio-pPZ ASE probe. Lane 1 (control) contains the probe without HepG2 NE. Lane 2 contains the binding reaction (pPZ ASE probe with HepG2 NE). Lane 3 and 4 contain the binding reaction with 50-fold and 100-fold excess unlabelled AP2 consensus (non-specific competitor), respectively. Lane 5 and 6 contain the binding reaction with 50-fold and 100-fold excess unlabelled HNF3B consensus (non-specific competitor), respectively.

All four previously observed bands (A-D) were observed in the binding reaction without competitors (lane 2, Fig 3.4). The addition of both unrelated competitor sequences (AP2 and HNF3B) produced a reduction in intensity of band D (lanes 3-6), with the AP2 competitor having a more marked effect (lanes 3-4). However, bands A-C were unaffected by either competitor.

3.4.1.1. Effect of nuclear extract, longer incubation and probe amount on protein-DNA binding

A further assay was carried out increasing the incubation period of the binding reaction from 20 minutes to 30 minutes at room temperature and, in addition, an alternative nuclear extract preparation procedure was employed (See Section 2.2.1 and 3.2.1). Band D of the EMSA appears to be more intense following this longer incubation and the modified nuclear extract preparation method (Fig 3.5).

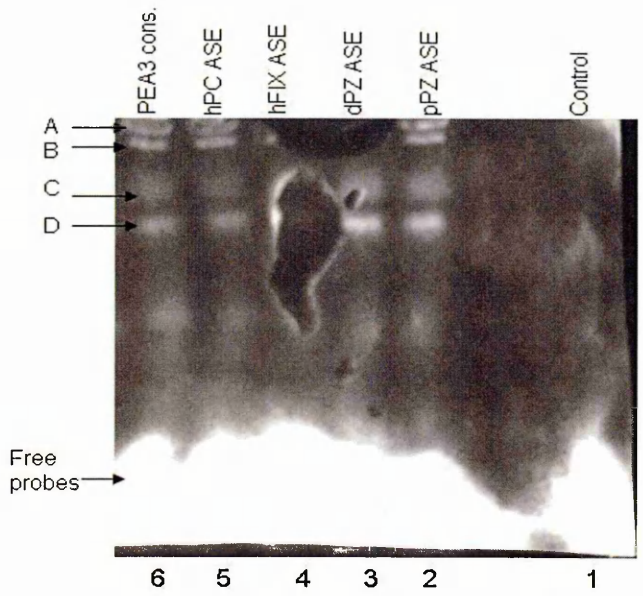


Fig 3.5. EMSA using different Biotin-labelled ASEs and PEA3 consensus probes. Lane1 (control) is the pPZ ASE probe without HepG2 NE. All other lanes contain Hep NE as well as probes as follows; pPZ ASE, dPZ ASE, human (h) FIX ASE, hPC ASE and PEA3 binding site consensus, respectively.

In order to reduce the proportion of free probe observed in the EMSA the amount of probe added to the binding reaction was reduced from 20 fmole to 10 fmole (Fig 3.6).

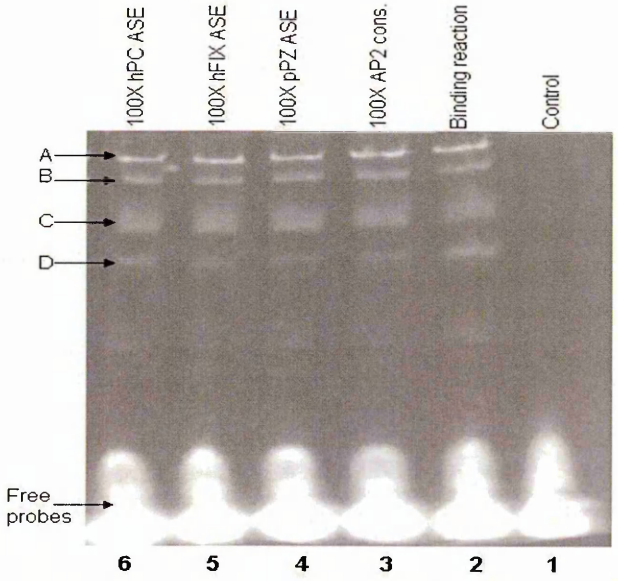


Fig 3.6. Competitive EMSA using Bio-dPZ ASE probe. Lane 1 (control) contains the probe without HepG2 NE. Lane 2 contains the binding reaction (probe with HepG2 NE). Lane 3, 4, 5 and 6 contain the binding reaction with 100-fold excess unlabelled AP2 consensus, pPZ ASE, hFIX ASE, hPC ASE (competitors), respectively. No clear competition was seen with all competitors.

Incubation was again carried out at room temperature for 30 minutes. The results showed less saturated signal of free probe. The binding complex D was relatively weak and showed negligible competition with 100-fold AP2 consensus and pPZ, hFIX, and hPC ASEs.

3.4.1.2. Effect of binding buffer on protein-DNA binding

Alternative binding buffers (See Section 3.3.3) were used in the following assay, again with 30 minutes incubation at room temperature (Fig 3.7). The EMSA using the pPZ ASE probe and 4µg of nuclear extract (Andrews and Faller 1991 method) with the two new binding buffers (buffers-1 and 2) were found to give more intense signals for the complexes than the original buffer (Pierce, UK), with buffer-1 giving the clearest and most intense signal. Moreover, bands A, B and C were detected in the second control lane (lane 2) where the Hep G2 NE and the binding buffer were incubated without the probe.

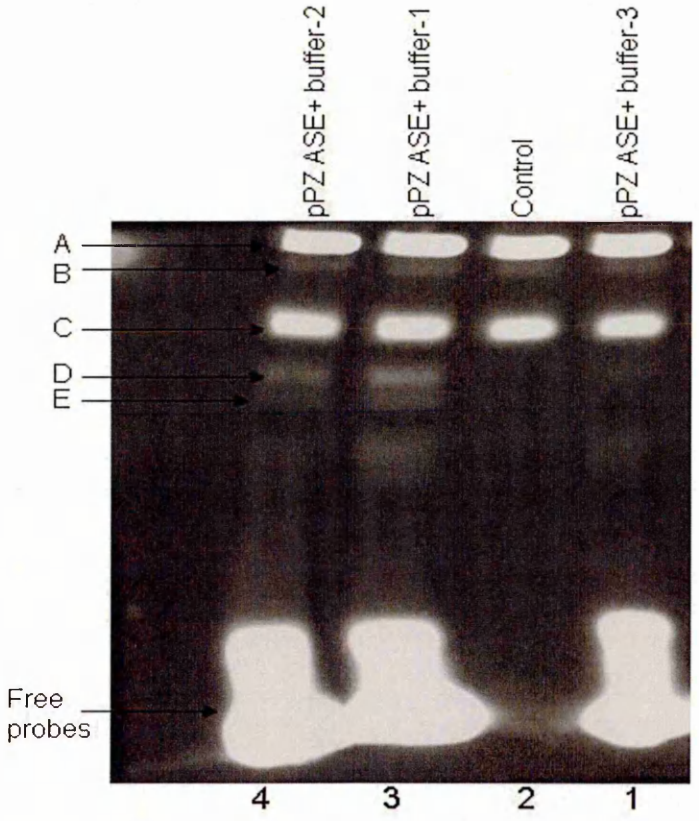


Fig 3.7. EMSA using Bio-pPZ ASE with different binding buffers. Lane 1 contains binding reaction (pPZ ASE probe and 4 µg HepG2 NE) with the original binding buffer (buffer-3). Lane 2 contains only a nuclear extract (new control). Lane 3 contains binding reaction with binding buffer-1. Lane 4 contains binding reaction with binding buffer-2.

Competitive EMSA were repeated using optimised binding buffer-1, 10 fmol of probe and 2µl of nuclear extract (4µg protein) in a 20µl total volume incubated for 30 minutes at room temperature. Bands A-C (Fig 3.8) appear in all lanes with Hep G2 NE present, indicating that non-specific proteins are present within the NE rather than being specific protein-biotinylated probe complexes. Bands 1 and 2 with the pPZ ASE (Fig 3.8) are, however, protein-biotinylated probe complexes. No competition was found with the AP2 or HNF3B competitors and only weak competition of 1.2-fold and 1.5-fold decrease was seen with the related sequence dPZ ASE in the intensity of complexes 1 and 2 respectively. However, with the PEA3 consensus a dramatic reduction in intensity of both complexes 1 and 2 was observed. A 4-fold decrease in the intensity of complex 1 and the disappearance of complex 2 was seen, as measured by the UVP software.

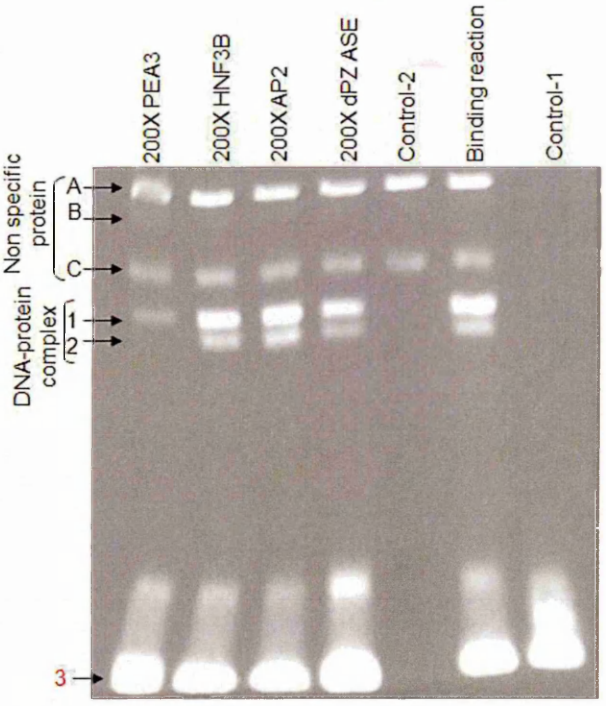


Fig 3.8. Competitive EMSA using the putative PZ proximal ASE binding site. Control-1 is the proximal ASE (pPZ ASE) probe (-120/-101) without HepG2 NE. Control-2 is the HepG2 NE without the probe and Binding reaction is the probe with the HepG2 NE. All other competition reactions contain the probe and the NE as well as unlabelled 200-fold excess dPZ ASE, 200-fold excess unlabelled AP2 consensus, 200-fold excess unlabelled HNF3B consensus and 200-fold excess PEA3 binding site consensus, respectively. Bands indicated by arrows A, B, and C represent a non-specific protein detected in the nuclear extract. Bands indicated by 1 and 2 represent protein-DNA complexes. Bands indicated by arrow 3 show unbound probe. Weak competition is seen with the dPZ ASE compared to strong competition with PEA3 binding site consensus. No competition is seen with AP2 and HNF3B consensus.

Competitive EMSA were again repeated with 10 fmole of dPZ ASE probe using optimised binding buffer 1 and 2µl of HepG2 NE (4µg protein) in a 20µl total volume incubated for 30 minutes at room temperature (Fig 3.9). Bands A-C (biotinylated protein) again appear in all lanes with nuclear extract present. Bands 1 and 2 with the dPZ ASE (Fig 3.9) showed protein-biotinylated probe complexes. No competition was observed with the AP2 or HNF3B competitors, however, strong competition was seen in both complexes 1 and 2 with the related sequence pPZ ASE and the PEA3 consensus. The intensity of both protein-DNA complexes were decreased 7-fold with pPZ ASE while complex 1 decreased 6-fold and complex 2 was abolished with the PEA3 consensus measured by the UVP software.

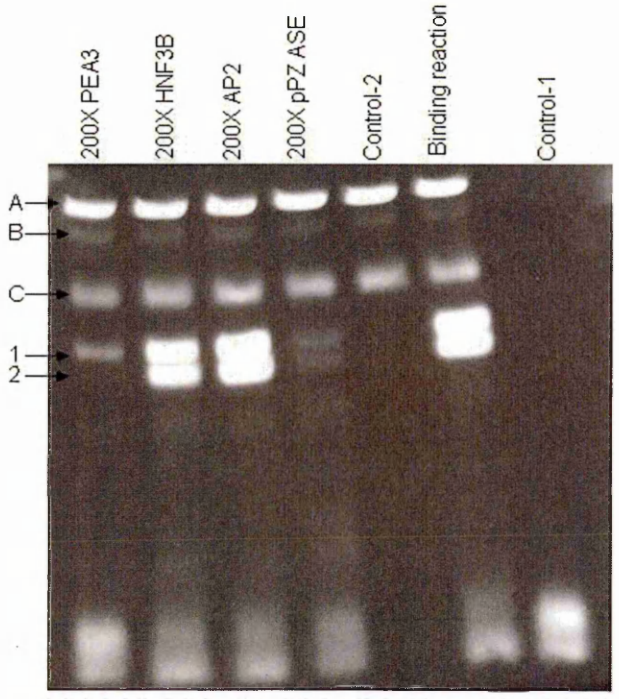


Fig 3.9. Competitive EMSA using the putative PZ distal ASE binding site.
Control-1 is the distal ASE (dPZ ASE) probe (-822/-803) without HepG2 NE. Control-2 is the HepG2 NE without the probe. Binding reaction is the probe with the HepG2 NE. All other competition reactions contain the probe and the NE as well as unlabelled 200-fold excess pPZ ASE, 200-fold excess unlabelled AP2 consensus, 200-fold excess unlabelled HNF3B consensus and 200-fold excess PEA3 binding site consensus, respectively. Bands indicated by arrows A, B, and C represent a biotinylated protein detected in the nuclear extract. Bands indicated by 1 and 2 represent protein-DNA complexes. Bands indicated by arrow 3 show unbound probe. Strong competition is seen with the dPZ ASE and PEA3 binding site consensus. No competition is seen with AP2 and HNF3B consensus.

3.4.1.3. Effect of temperature and incubation period on antibody-protein binding

Supershift assays were performed to identify the protein bound to the candidate probe. This was accomplished using the same binding reaction in the EMSA with the addition of a specific antibody against the putative binding protein. Antibodies against PEA3 and Ets1, candidate proteins bound to the pPZ and dPZ ASEs, were used and the reactions were optimised with those probes.

The supershift assay for the pPZ ASE binding site with the PEA3 antibody was firstly performed at room temperature (Fig 3.10). The antibody was incubated with the HepG2 NE and binding buffer-1 for approximately 10 minutes at room temperature followed by addition of the probe and a further incubation for 25-30 minutes, again at room temperature. The presence of the antibody produced a weak reduction in the intensity of bands 1 and 2, but no clear supershift.

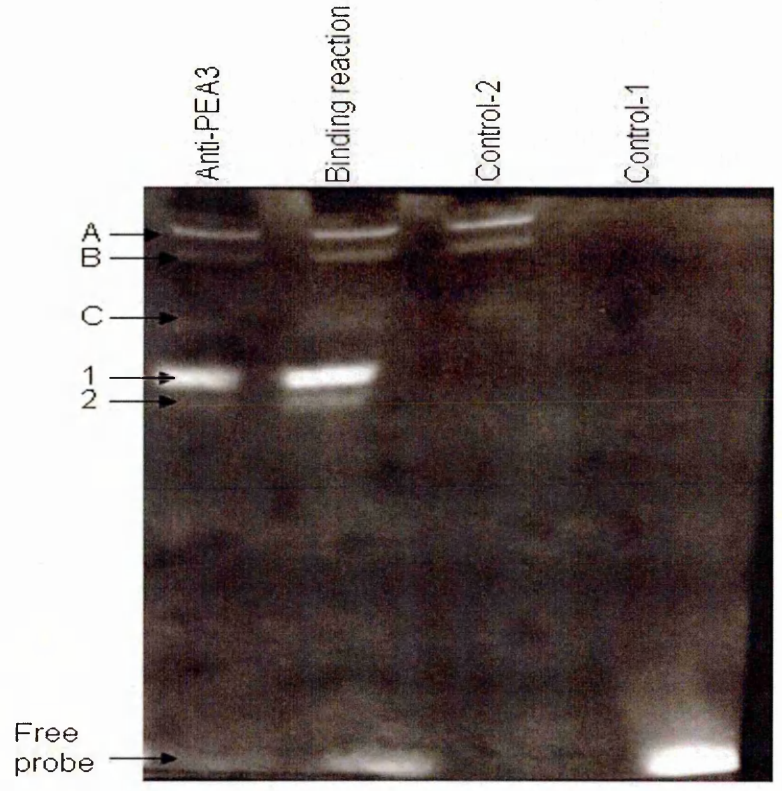


Fig 3.10. Supershift assay using biotinylated pPZ ASE binding sites.
Control-1 is the candidate probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE. Anti-PEA3 is an antibody against PEA3 protein together with the binding reaction. The antibody incubated with the HepG2 NE for 10 minutes and further incubated with the probe for 25 minutes at room temperature. No supershift was seen with the probe. A, B and C are control bands which come up whenever the HepG2 NE is added. 1 and 2 are the protein-DNA complexes.

The pPZ ASE binding site supershift was then performed with a prolonged incubation of the PEA3 antibody with the HepG2 NE and binding buffer-1 overnight (Fig 3.11.a) and for 90 minutes (Fig 3.11.b) at 4°C. The probe was then added and the incubation carried out for a further 25-30 minutes at room temperature.

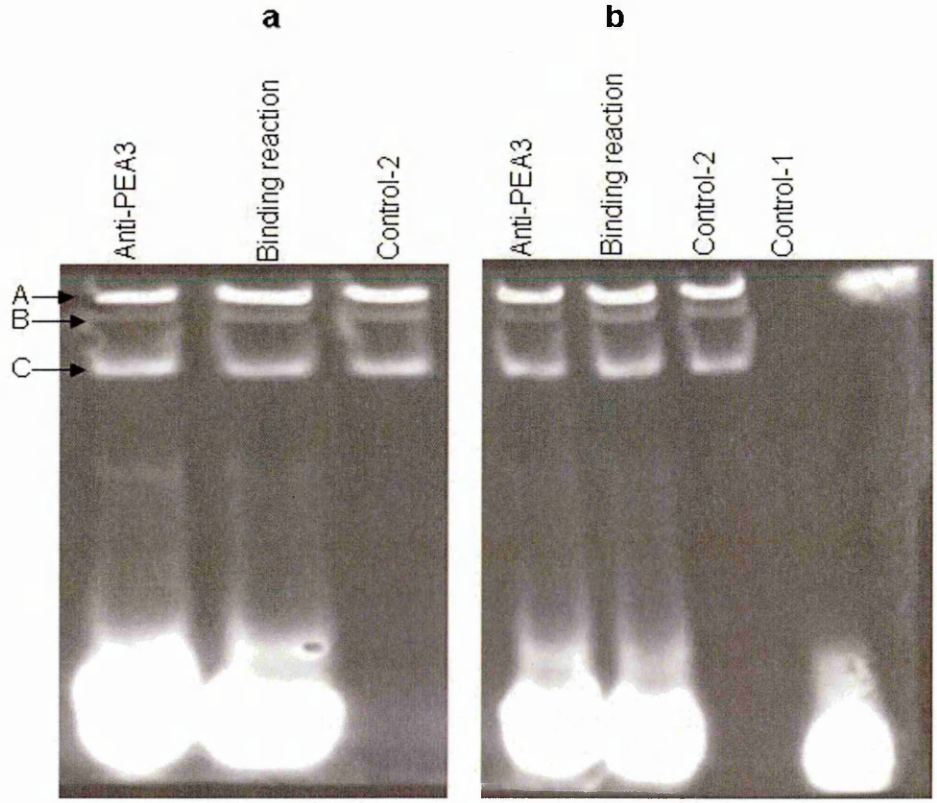


Fig 3.11. Supershift assay at 4°C using biotinylated pPZ ASE binding sites.
Control-1 is the pPZ ASE probe without HepG2 NE. Control-2 is the Hep NE without probe. Binding reaction is the probe with HepG2 NE. Anti-PEA3 is an antibody against PEA3 protein together with the binding reaction. (a)The antibody incubated with the HepG2 NE overnight at 4°C and further incubated with the probe for 25 minutes at room temperature. (b)The antibody incubated with the HepG2 NE 90 minutes at 4°C and further incubated with the probe for 25 minutes at room temperature. No supershift was seen with both conditions. A, B and C are control bands which come up whenever the HepG2 NE is added.

Both assays demonstrated the presence of the biotinylated proteins (bands A-C), but no protein-DNA complexes. To determine if the lack of protein-DNA complexes may be due to the degradation of the protein the supershift assay was repeated overnight at 4°C (Fig 3.12) with fresh protease inhibitor cocktail. Although protease inhibitor was added to prevent protein degradation, the binding reaction did not show protein-DNA complexes.

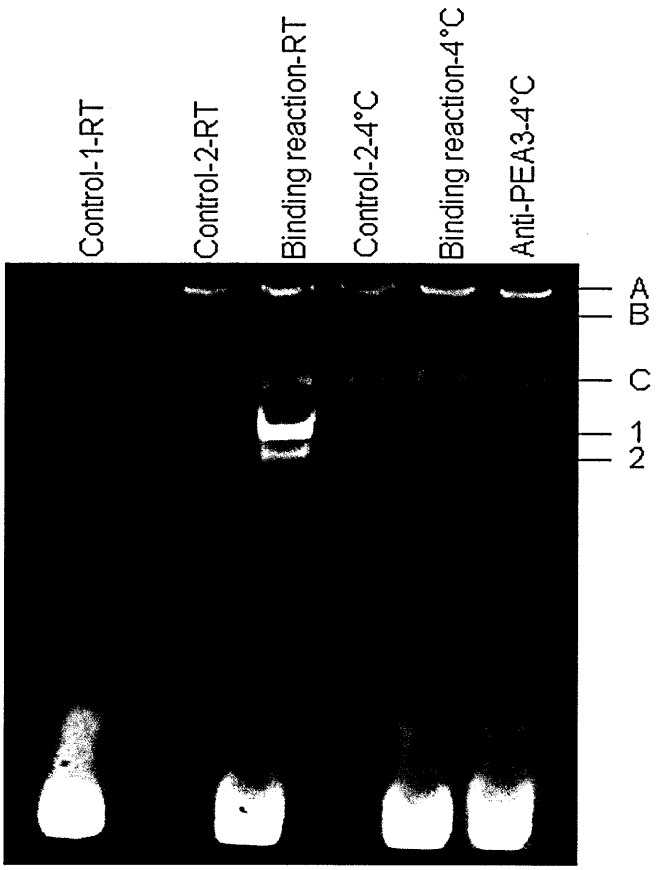


Fig 3.12. Supershift assay at 4°C using biotinylated pPZ ASE binding sites.
Control-1 is the pPZ ASE probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE. Anti-PEA3 is an antibody against PEA3 protein together with the binding reaction. From left to right, the first three lanes were performed at room temperature (RT) while the other following three lanes were performed at 4°C. Control-2-4°C is the HepG2 NE incubated with the binding buffer overnight at 4°C. Binding reaction-4°C is the HepG2 NE incubated with the binding buffer overnight at 4°C followed by incubation with the probe for 30 minutes at room temperature. Anti-PEA3-4°C is the antibody incubated with the HepG2 NE overnight at 4°C and further incubated with the probe for 30 minutes at room temperature. No supershift was seen. A, B and C are control bands which come up whenever the HepG2 NE is added. Bands 1 and 2 are the protein-DNA complexes. Protease inhibitor cocktail and the DTT were added to the binding buffer just before the assay.

Supershift assays using the pPZ ASE probe with the Ets1 antibody were carried out on ice and at room temperature (Fig 3.13). The antibody was incubated with the Hep G2 NE on ice or at room temperature for 10 minutes followed by an addition of the probe and further incubation for 30 minutes on ice and at room temperature, respectively. The results showed protein-DNA complexes (bands 1 and 2), but no antibody-protein-DNA complex with the pPZ ASE binding site at room temperature. On ice no complexes were observed in the presence or absence of the antibody.

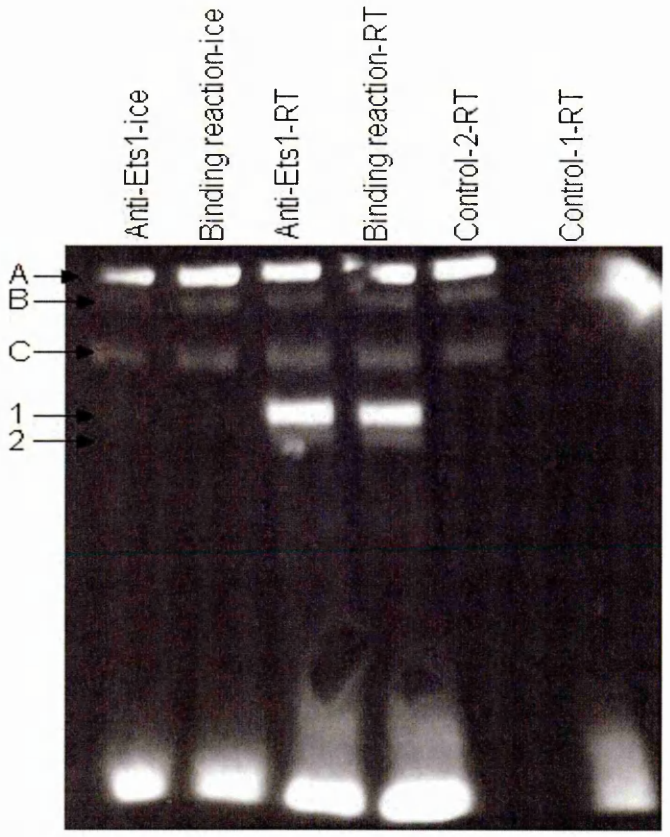


Fig 3.13. Supershift assay on ice using biotinylated pPZ ASE binding sites with Ets1. Control-1 is the pPZ ASE probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE. Anti-Ets1 is an antibody against Ets1 protein together with the binding reaction. From right to left, the first four lanes were performed at room temperature (RT) where the HepG2 NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. The following two lanes were performed on ice where the HepG2 NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. No supershift was seen at room temperature while on ice neither protein-DNA complex nor antibody-protein-DNA complex. A, B and C are control bands which come up whenever the HepG2 NE is added. Bands 1 and 2 are the protein-DNA complexes.

Supershift assays were also carried out using the dPZ ASE binding site probe with the PEA3 antibody (anti-PEA3) and the Ets1 antibody (anti-Ets1) at room temperature and 37°C. The results at room temperature were found to show a good protein-DNA binding complex (band 1) but no supershift band with either the Ets1 or the PEA3 antibodies (Fig 3.14b). However, the protein-DNA complex (band 1) in the presence of Ets1 antibody showed a very strong signal compared to the protein-DNA complex in the absence of the antibody. The intensity of band 1 in the presence of the Ets1 antibody was found to be

approximately four times more intense than in its absence, using the UVP system software (UVP, UK).

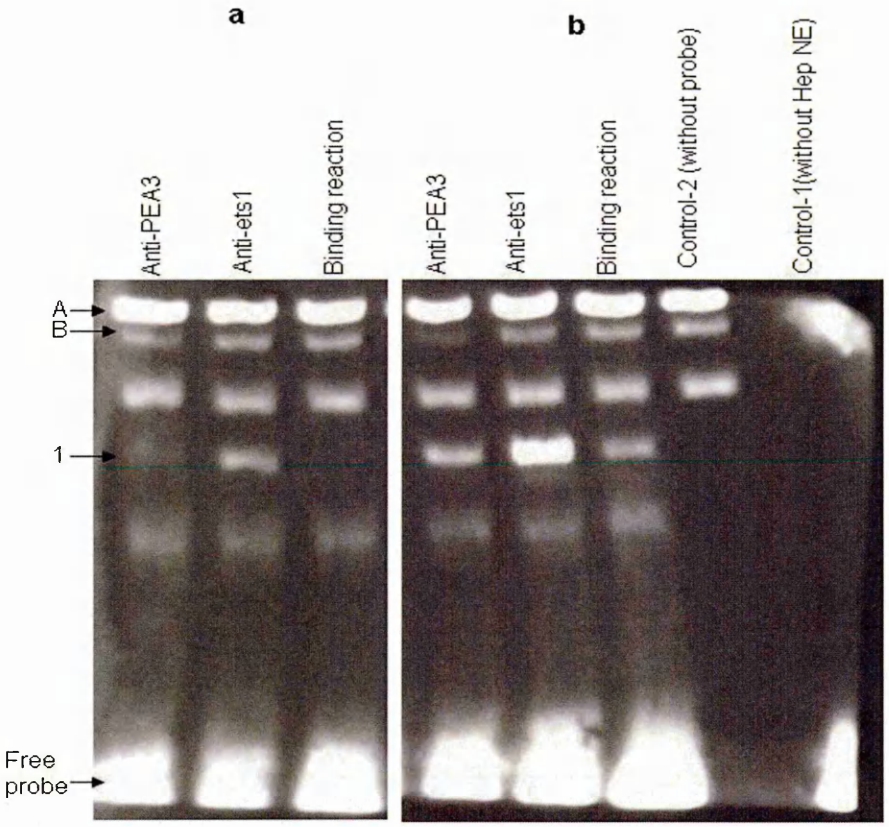


Fig 3.14. Supershift assay at room temperature and 37°C using biotinylated dPZ ASE binding sites with PEA3 and Ets1. Control-1 is the dPZ ASE probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE. Anti-PEA3 is an antibody against PEA3 protein together with the binding reaction. Anti-Ets1 is an antibody against Ets1 protein together with the binding reaction. (a) The reactions were performed at 37°C where the Hep NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. (b) The reactions were performed at room temperature where the HepG2 NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. No supershift was seen at room temperature with both anti-PEA3 and Ets1 although the intensity of the band with Ets1 increased which might be an indication for the antibody binding. However, at 37°C neither protein-DNA complex in the binding reaction nor antibody-protein-DNA complex with the antibodies were seen. A, B and C are control bands which come up whenever the HepG2 NE is added. Band 1 is the protein-DNA

The binding reaction of the dPZ ASE binding site probe with the HepG2 NE did not show protein-DNA binding at 37°C, although a clear band was detected in addition to the biotinylated protein bands A-C, in the presence of the Ets1 antibody (Fig 3.14a). Another EMSA for the dPZ ASE probe was done at 4°C to check the Ets1 supershift complexes. The binding mixture (binding buffer, HepG2 NE and Ets1 antibody) was incubated overnight at 4°C before the probe

was added and further incubated for 30 minutes. The results at 4°C showed neither protein-DNA complex in the binding reaction nor an additional band in the presence of the Ets1 antibody (Fig 3.15).

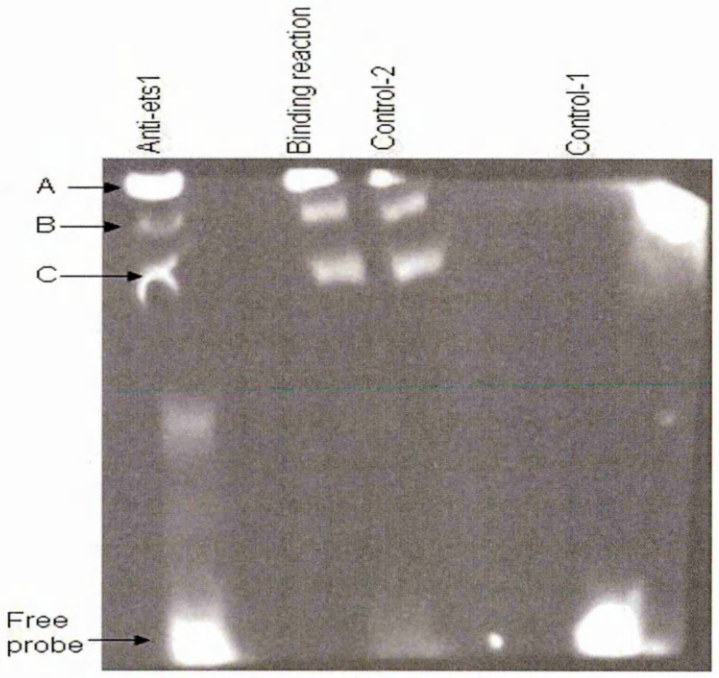


Fig 3.15. Supershift assay at 4°C using biotinylated dPZ ASE binding sites with Ets1. Control-1 is the dPZ ASE probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE incubated at 4°C overnight followed by incubation with the probe for 25-30 minutes. Anti-Ets1 is an antibody against Ets1 protein together with the binding reaction. The Ets1 reaction was performed by incubating the antibody with the HepG2 NE at 4°C overnight followed by incubation with the probe for 25-30 minutes. Neither protein-DNA complex in the binding reaction nor antibody-protein-DNA complex with the antibodiy was seen. A, B and C are the control bands.

3.4.1.4. Optimisation of supershift assay using PEA3 binding site consensus

A supershift assay was also performed using the PEA3 consensus binding site probe with the PEA3 antibody. The binding buffer was pre-incubated with the antibody at room temperature for 10 minutes before the addition of the probe and further incubation for 25-30 minutes. The results showed a clear protein-DNA complex but no supershifted band with the antibody (Fig 3.16). The same assay was repeated using a dialysed HepG2 NE (SeeSection 3.3.1). The results of the PEA3 binding site consensus probe with the PEA3 antibody using a dialysed HepG2 NE, again did not show the supershifted band which would be an indication of the antibody-protein-DNA complex (Fig 3.17).

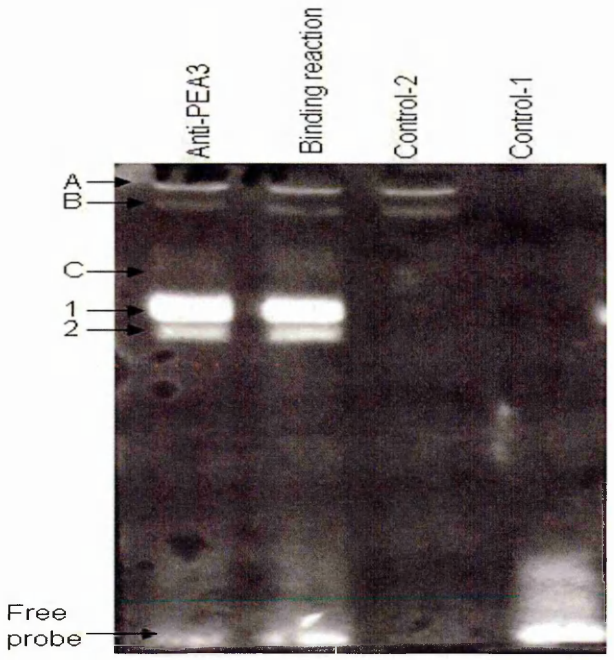


Fig 3.16. Supershift assay using biotinylated PEA3 binding sites consensus with PEA3 antibody. Control-1 is the PEA3 binding consensus probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE incubated at room temperature followed by incubation with the probe for 25-30 minutes. Anti-PEA3 reaction was performed by incubating the antibody with the HepG2 NE at room temperature followed by incubation with the probe for 25-30 minutes. No antibody-protein-DNA complex in the presence of the antibodiy was seen. A, B and C are control bands which come up whenever the HepG2 NE is added. 1 and 2 is the potential PEA3-DNA binding complex.

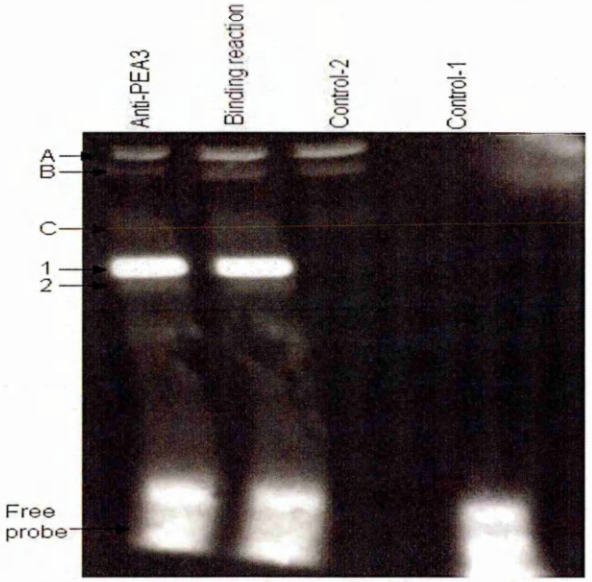


Fig 3.17. Supershift assay using biotinylated PEA3 binding sites consensus with PEA3 antibody and dialysed nuclear extract. Control-1 is the PEA3 binding consensus probe without HepG2 NE. Control-2 is the dialysed HepG2 NE without probe. Binding reaction is the probe with the dialysed HepG2 NE incubated at room temperature followed by incubation with the probe for 25-30 minutes. Anti-PEA3 reaction was performed by incubating the antibody with the dialysed HepG2 NE at room temperature followed by incubation with the probe for 25-30 minutes. No antibody-protein-DNA complex was seen in the presence of the antibodiy. A, B and C are control bands while 1 and 2 are the potential PEA3-DNA binding complex.

3.4.2. Use of IRDye labelled Probes

Supershift assays using 100 fmole pPZ ASE and dPZ ASE fluorescent probes were performed with 4µg of HepG2 NE and the Ets1 antibody. The antibody was pre-incubated with the HepG2 NE in binding buffer-1 for 10 minutes at room temperature prior to incubation with the probes for 30 minutes at room temperature. The results showed a protein-DNA complex band with the pPZ ASE probe but no supershifted band (Fig 3.18a). The dPZ ASE probe was not bound by protein (Fig 3.18b).

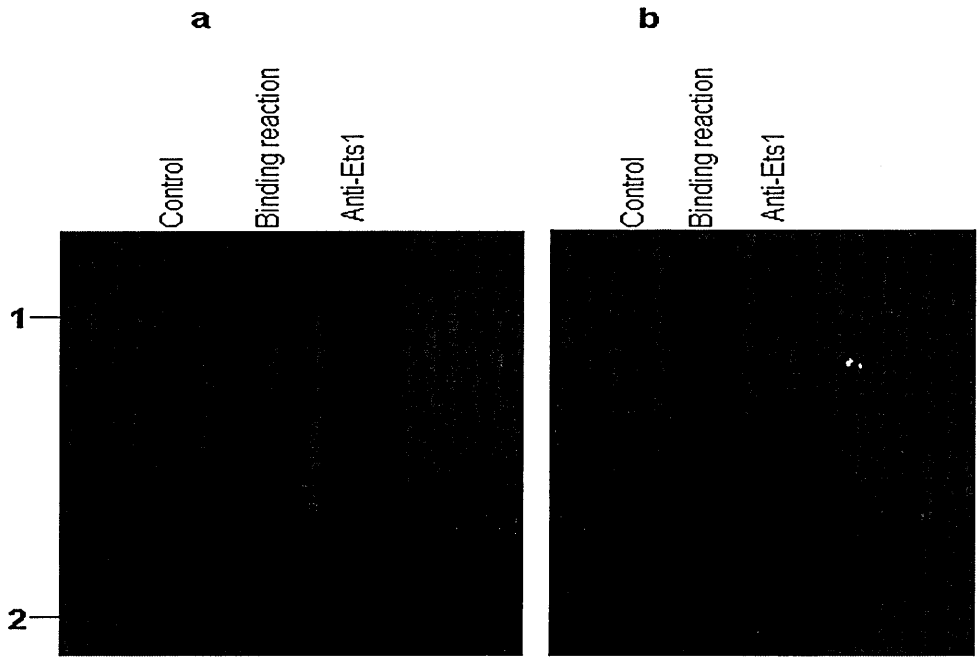


Fig 3.18. Supershift assay at room temperature using fluorescent pPZ ASE and dPZ ASE probes with Ets1. Control is probe without HepG2 NE. Binding reaction is the probe with HepG2 NE. Anti-Ets1 is an antibody against Ets1 protein together with the binding reaction. The reactions were performed at room temperature where the HepG2 NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. (a) is the reactions with pPZ ASE probe while (b) with the dPZ ASE probe. No supershift was seen at room temperature with both probes. dPZ ASE showed neither protein-DNA complex in the binding reaction nor antibody-protein-DNA complex with the antibody. Band 1 is the protein-DNA complex while band 2 is the free probe.

3.4.2.1. Optimisation of antibody-protein binding using CREB binding site consensus

Supershift assays were also performed using a CREB antibody with the fluorescently labelled CREB consensus binding site probe to optimise the formation of the antibody-protein-DNA complex. The HepG2 NE was pre-

incubated with the CREB antibody in a binding buffer lacking dithiothreitol (DTT) for 10 minutes and then further incubated with the probe for 25-30 minutes at room temperature. The results (Fig 3.19) showed a major protein-DNA complex (1), but did not show an antibody-protein-DNA complex.

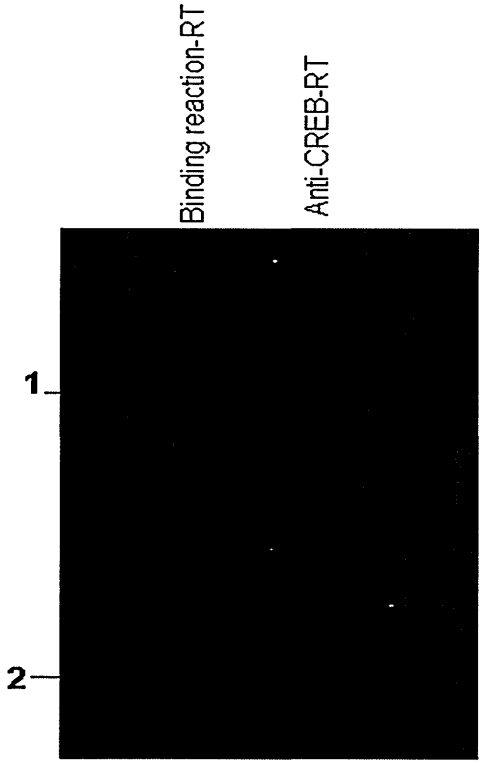


Fig 3.19. Supershift assay at room temperature using fluorescent CREB probe with CREB antibody. Binding reaction is the probe with HepG2 NE. Anti-CREB is an antibody against CREB protein together with the binding reaction. The reaction was performed at room temperature where the HepG2 NE incubated with the CREB antibody for 10 minutes and followed by 25-30 minutes incubation with the probe. The binding buffer was without DTT. No supershift was seen. Band 1 is the protein-DNA complex while band 2 is the free probe.

3.4.2.2. Effect of temperature and incubation period on antibody-protein binding using CREB binding consensus

Supershift analysis with the CREB-fluorescent consensus binding site was repeated by incubating the CREB antibody with the HepG2 NE for one hour at 4°C, in one experiment (Fig 3.20), and for 20 minutes at 37°C in another experiment (Fig 3.21), followed by the addition of the probe and further incubation for approximately 30 minutes. The results at 4°C, again, did not show any indication of antibody binding whilst at 37°C not even protein-DNA complexes were observed.

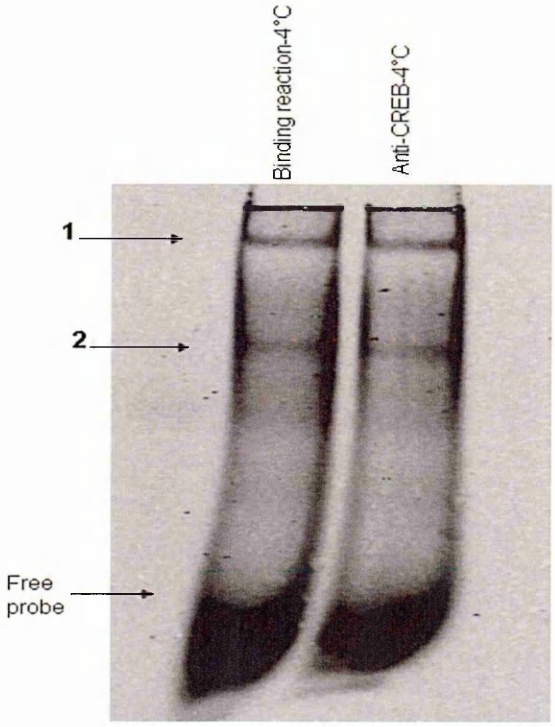


Fig 3.20. Supershift assay at 4°C using fluorescent CREB probe with CREB antibody. Binding reaction is the probe with one hour pre-incubated HepG2 NE at 4°C. Anti-CREB is the antibody against CREB protein together with the binding reaction. The antibody reaction was performed at 4°C where the HepG2 NE incubated with the CREB antibody for one hour and followed by 25-30 minutes incubation at room temperature with the probe. No supershift was seen. Bands 1 and 2 are the potential protein-DNA complexes.

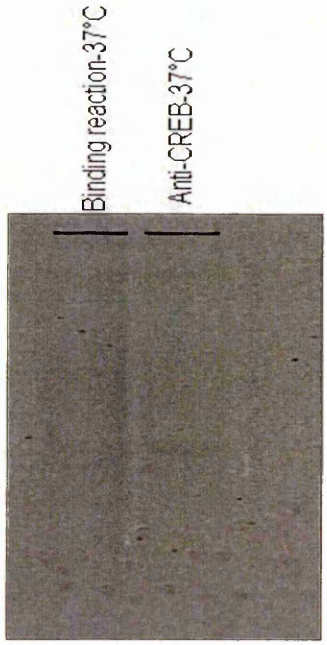


Fig 3.21. Supershift assay at 37°C using fluorescent CREB probe with CREB antibody. Binding reaction is the probe with 20 minutes pre-incubated HepG2 NE at 37°C. Anti-CREB is the antibody against CREB protein together with the binding reaction. The antibody reaction was performed at 37°C where the HepG2 NE incubated with the CREB antibody for 20 minutes and followed by 25-30 minutes incubation at room temperature with the probe. Neither protein-DNA nor antibody-protein-DNA complexes were seen.

3.4.2.3. Effect of new protocol with different buffers on CREB supershift

An alternative protocol (Santa Cruz biotechnology, Germany) was then used in an attempt to optimise the supershift assay. The binding reaction contained 10 µg of the HepG2 NE, 1 µg of poly dI-dC and 100 fmole of the CREB-fluorescent consensus binding site probe in a 20 µl binding reaction. The binding reaction was incubated for 20 minutes at room temperature. 4 µg of the CREB antibody were incubated in the binding mixture (SeeSection 3.3.3), without the probe, for 20 minutes at room temperature followed by addition of the CREB-fluorescent probe and further incubation for 30 minutes at room temperature. The binding reactions were size-fractionated through a 4% polyacrylamide gel containing 50 mM Tris, pH 7.5, 0.38 M glycine and 2mM EDTA and visualised by the Odyssey system (Li-Cor). Separation of the labelled products using this procedure was suboptimal and, therefore, difficult to interpret (Fig 3.22).

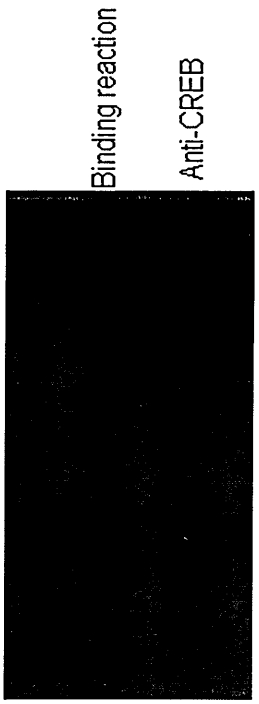


Fig 3.22. Supershift assay of the fluorescent CREB probe with CREB antibody using a new protocol (Santa Cruz, UK) . Binding reaction is the probe with 20 minutes pre-incubated HepG2 NE in the new binding buffer (See above) at room temperature. Anti-CREB is the antibody against CREB protein together with the binding reaction. The antibody reaction was performed at room temperature where the HepG2 NE incubated with the CREB antibody for 20 minutes and followed by 30 minutes incubation at room temperature with the probe. Neither clear protein-DNA binding nor antibody-protein-DNA complex was seen.

3.4.2.4. EMSA and supershift assays for HNF4α

Competitive EMSAs and supershift analysis were also performed for the potential proximal HNF4α binding site (pPZ HNF4α) in the PZ gene promoter. Competitive EMSA was accomplished by incubating approximately 12μg of HepG2 NE with 100-fold excess (compared to the probe) of competitor sites for 10 minutes at room temperature, followed by addition of 100 fmole of the probe and further incubation for 30 minutes. The results showed a strong competition with 100-fold unlabelled self and the HNF4α binding site consensus sequence and negligible competition with 100-fold of non-specific competitors, HNF3B, AP1 and SP1 (Fig 3.23).

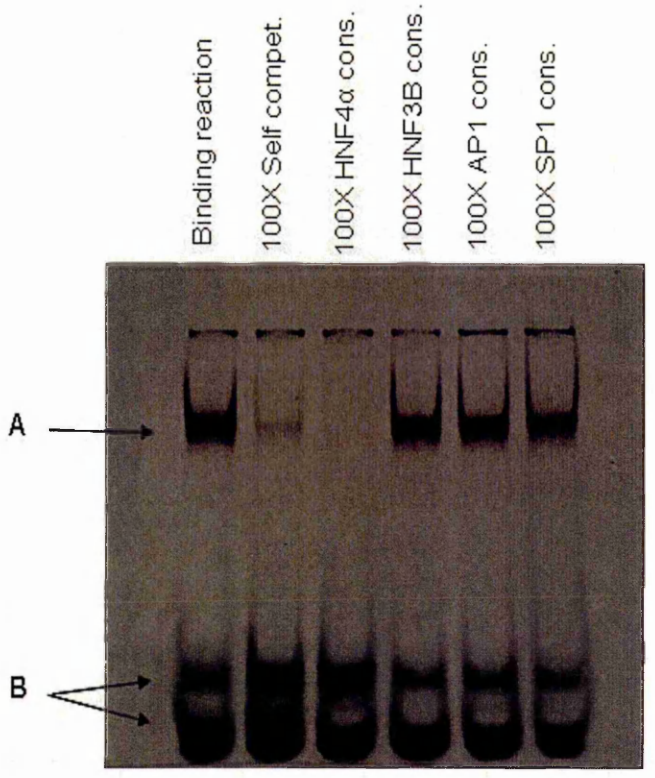


Fig 3.23. Competitive EMSA using the reported pPZHNF4α-fluorescent binding site. The binding reaction lane is the reported HNF4α (pPZHNF4α) probe (-48/-38) with HepG2 NE. Self compet. lane is the pPZHNF4α probe, the HepG2 NE and the unlabeled pPZHNF4α. HNF4α cons. lane is the pPZHNF4α probe, the HepG2 NE and unlabeled HNF4α binding consensus. HNF3B, AP1 and SP1 cons. lanes are the pPZHNF4α probe with the HepG2 NE and unlabeled HNF3B, AP1 and SP1 binding site consensuses added, respectively. The HepG2 NE was incubated with the competitor for 10 minutes at room temperature followed by 30 minutes incubation with the probe. Bands indicated by arrow A represent protein DNA complexes. Bands represented by arrow B show unbound fluorescent probe. Competition is seen with the self competitor and the HNF4α consensus sequence. Negligible competition is seen with the non-specific competitors HNF3B, AP1 and SP1 binding sites.

Supershift analyses were later performed for the pPZHNF4α together with a repetition of the competitive EMSA using the same parameters. In the supershift assay the antibody was incubated with the HepG2 NE for 10 minutes followed by incubation with the probe for 30 minutes. The results of the competitive EMSA again showed a dramatic reduction in the protein-DNA complex with the self competitor and loss of the complex with the HNF4α binding site consensus (Fig 3.24). Supershift analysis with the HNF4α antibody (Santa Cruz biotechnology, Germany) showed a loss of the protein-DNA complex and greater fluorescence intensity (approximately 3.7-fold) in the well with the antibody.

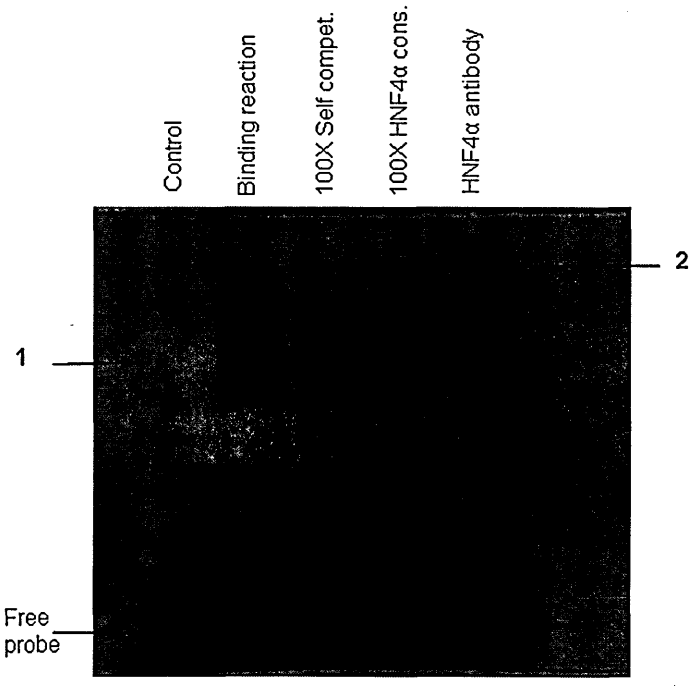


Fig 3.24. Supershift assay and competitive EMSA using the pPZHNF4α-fluorescent binding site. Control was the pHNF4α probe (-48/-38) without HepG2 NE. Binding reaction was the pPZHNF4α probe (-48/-38) with HepG2 nuclear extract (Hep NE). Self competitor was the pPZHNF4α probe, the HepG2 NE and the unlabeled pPZHNF4α. HNF4α consensus lane was the pPZHNF4α probe, the HepG2 NE and unlabeled HNF4α binding consensus. HNF4α antibody lane contained the pPZHNF4α probe, the HepG2 NE and the HNF4α antibody. Bands indicated by arrow 1 represent protein DNA complexes. Band 2 represents the potential antibody-protein-DNA complex. Competition is seen with 100-fold molar excess of self-competitor and the HNF4α consensus sequences. The reactions were at room temperature.

The assay was again repeated using the HNF4α binding site consensus probe with the same parameters as with the pPZHNF4α. The results showed a dramatic reduction in the protein-DNA complex and 4-fold increase in

fluorescent intensity, measured using the software of the Li-Cor system, in the well with the antibody compared to the well without the antibody (Fig 3.25).

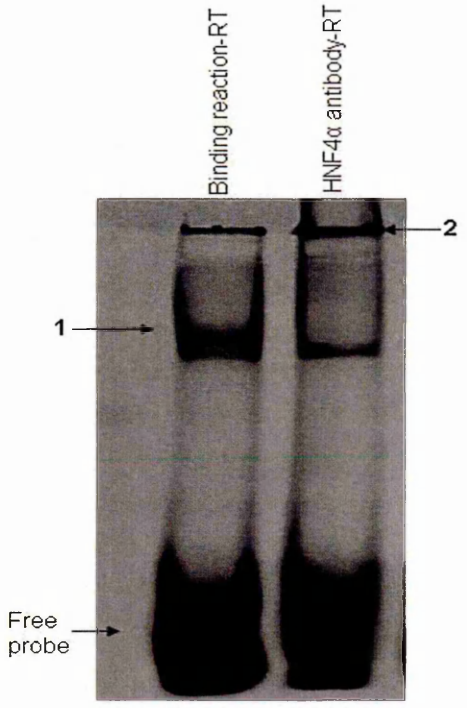


Fig 3.25. Supershift assay using the consensus HNF4α-fluorescent binding site. The binding reaction was the HNF4α binding site consensus probe with HepG2 NE. HNF4α antibody lane contained the HNF4α binding site consensus probe, the HepG2 NE and the HNF4α antibody. Band indicated by arrow 1 represents protein DNA complex. Band 2 represents the potential antibody-protein-DNA complex. The reactions were at room temperature.

3.4.2.5. Effect of temperature and incubation period on HNF4α supershift

Other supershift analyses were performed with different conditions to see whether a resolvable supershifted band could be obtained. One assay was done by incubating the HepG2 NE with the antibody on ice for two hours followed by further incubation with the probe for 30 minutes on ice (Fig 3.26a). A second assay was done by incubating the HepG2 NE with the antibody on ice for two hours followed by incubation with the probe for 30 minutes at room temperature (Fig 3.26b). The third experiment was performed by incubating the HepG2 NE with the antibody on ice for one hour followed by incubation with the probe for 30 minutes on ice, in one reaction, and at room temperature in another reaction (Fig 3.26c). The binding reaction showed two binding complexes. However, the reaction of the antibody with these three experiments

did not show supershift bands although changes in the strength of the protein-DNA complex bands were seen.

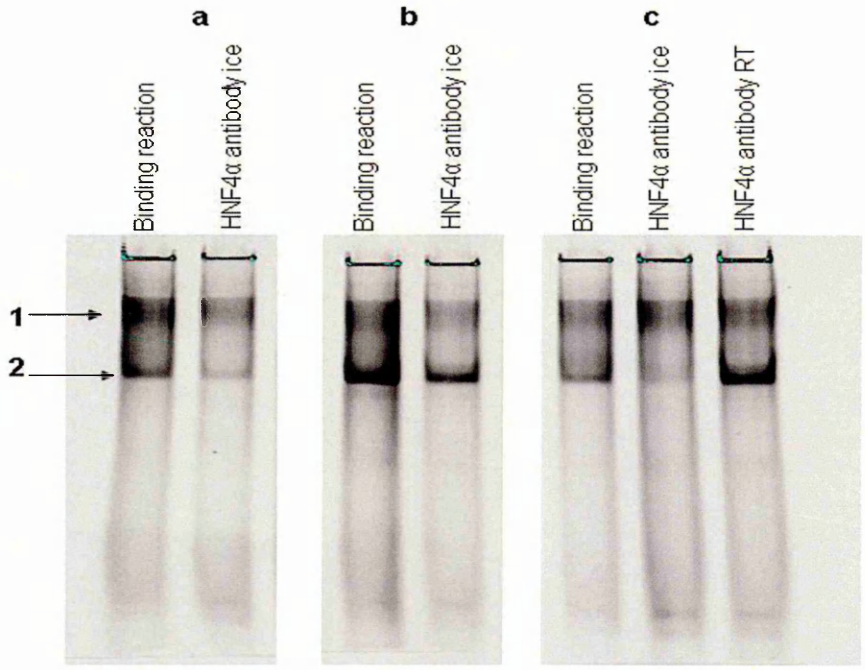


Fig 3.26. Supershift assay using the consensus HNF4α-fluorescent binding site at different conditions. The binding reaction was the HNF4α binding site consensus probe with HepG2 NE on ice. HNF4α antibody lane contained the HNF4α binding site consensus probe, the HepG2 NE and the HNF4α antibody. (a) The reactions were two hour incubation with the antibody on ice followed by 30 minutes incubation with the probe on ice. (b) The reactions were two hour incubation of the antibody on ice followed by 30 minutes incubation with the probe at room temperature. (c) The reactions were two hour incubation of the antibody on ice followed by 30 minutes incubation with the probe on ice and at room temperature. Bands indicated by arrow 1 and 2 represented potential protein-DNA complexes.

Other supershift experiments were done at 4°C. In the first experiment the HepG2 NE was incubated with the antibody at 4°C for 25 minutes and followed by further incubation with the probe for 20 minutes at room temperature (Fig 3.27). The second experiment was done by incubating the HepG2 NE with the antibody for one hour at 4°C followed by incubation of the probe for 20 minutes at room temperature (Fig 3.28). The first experiment (Fig 3.27) did not show any differences between the binding reaction and the supershift reaction while the second experiment with a longer incubation time of the HepG2 NE at 4°C did not show a clear discrete protein-DNA complex band (Fig 3.28). Three bands representing the protein-DNA complex were detected. Multiple bands instead of a single band may be due to protein degradation whilst maintaining DNA-binding capacity.

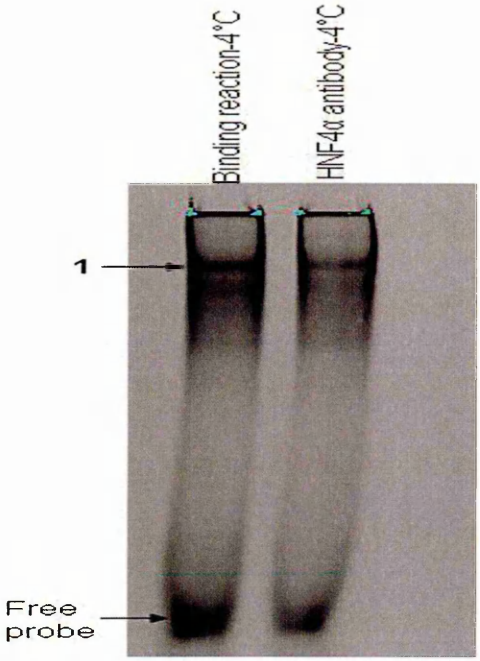


Fig 3.27. Supershift assay using the consensus HNF4α-fluorescent binding site at 4°C. The binding reaction was the HNF4α binding site consensus probe with HepG2 NE. HNF4α antibody lane contained the HepG2 NE, the HNF4α antibody and the HNF4α binding site consensus probe. The antibody was incubated with the HepG2 NE at 4°C followed by incubation with the probe at room temperature. The bands indicated by arrow 1 represent protein DNA complex.

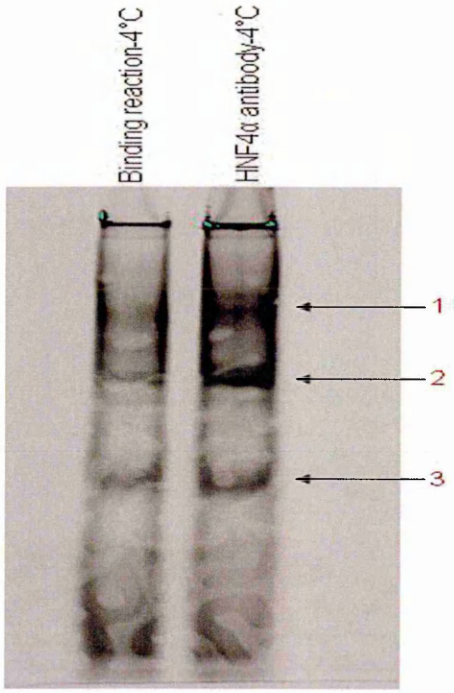


Fig 3.28. Supershift assay using the consensus HNF4α-fluorescent binding site at 4°C. The binding reaction was the HNF4α binding site consensus probe with HepG2 NE. HNF4α antibody lane contained the HepG2 NE, the HNF4α antibody and the HNF4α binding site consensus probe. The antibody was incubated with the HepG2 NE at 4°C for 1 hour followed by incubation with the probe at room temperature. Three bands (1,2 and 3) were shown which may be due to protein degradation. However, the bands were not clear in the binding reaction lane.

3.4.2.6. Effect of new binding buffer on EMSA using HNF4α binding site consensus

EMSA was also performed using the protocol published by Sugawara (Sugawara *et al.* 2007) for the HNF4α protein binding to the PZ gene promoter. The HepG2 NE was pre-incubated for 10 minutes at room temperature in a binding buffer (See Section 3.3.3) before addition of the probe and further incubation for 15 minutes. The results (Fig 3.29) did show protein-DNA binding, although the bands were not discrete.

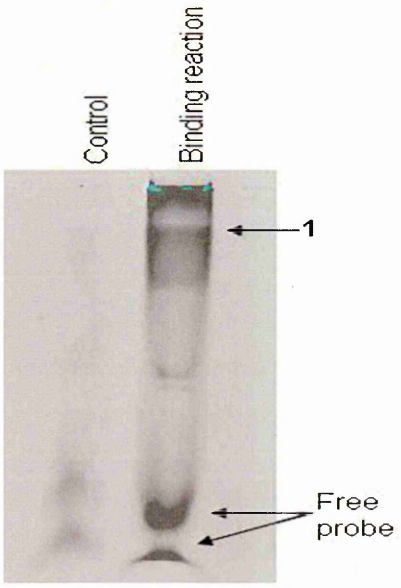


Fig 3.29. EMSA assay using the consensus HNF4α-fluorescent binding site.
Control lane contained the HNF4α probe without the HepG2 NE. The binding reaction was the HNF4α binding site consensus probe with HepG2 NE. The probe was added, after 10 minutes of the HepG2 NE incubation, and further incubated for 15 minutes at room temperature. Smear (1) instead of binding band was seen in the binding reaction lane.

To confirm that binding with the HNF4α binding site consensus sequence and the potential pPZHNF4α were optimum and consistent when performed at room temperature, the experiments (Fig 3.24 and Fig 3.25) were repeated (Fig 3.30 and Fig 3.31). The results of the repeated experiments again showed an abolition of the protein-DNA complex in the presence of the HNF4α antibody as well as a greater fluorescent intensity in the well compared to the binding reaction well without the antibody. The intensity was found to be approximately 3.2-fold and 4.0-fold greater in the presence of the antibody with the HNF4α

binding site consensus and pPZHNF4α probe, respectively, compared to the well without the antibody using the Li-Cor software.

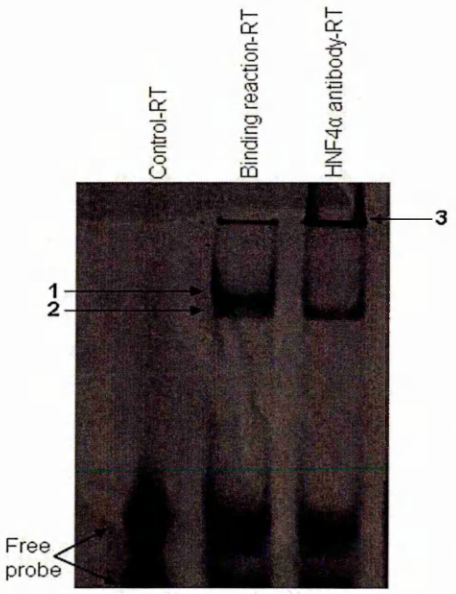


Fig 3.30. Supershift assay using the consensus HNF4α-fluorescent binding site.
The binding reaction was the HNF4α binding site consensus probe with HepG2 NE. HNF4α antibody lane contained the HNF4α binding site consensus probe, the HepG2 NE and the HNF4α antibody. Bands indicated by arrows 1 and 2 represent potential protein DNA complexes. Band 3 represents the potential antibody-protein-DNA complex. The protein-DNA complex (1) was competed with the addition of the antibody. The reactions were performed by the addition of the probe to 10 minutes pre-incubated binding mixture at room temperature followed by further 30 minute incubation at room temperature.

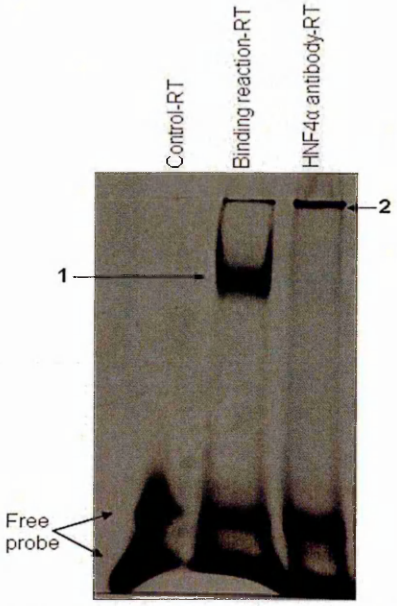


Fig 3.31. Supershift assay using the pPZHNF4α-fluorescent binding site.
The binding reaction was the pPZHNF4α binding site probe with HepG2 NE. HNF4α antibody lane contained the pPZHNF4α binding site probe, the HepG2 NE and the HNF4α antibody. Band indicated by arrow 1 represents the protein DNA complex. Band 2 represents the potential antibody-protein-DNA complex. The protein-DNA complex (1) was competed with the addition of the antibody. The reactions were performed by the addition of the probe to 10 minutes pre-incubated binding mixture at room temperature followed by further 30 minute incubation at room temperature.

3.5. Discussion

Binding of transacting proteins to their cis-regulatory elements within the promoter and/or enhancer is an essential stage in the transcriptional regulation of gene expression. Therefore, the accurate recognition of transcription factor binding sites and identification of the bound proteins is a crucial step in the characterisation of transcriptional gene regulation (Morozov *et al.* 2005). One of the techniques which has been used extensively in studying the interaction of proteins to their binding sites is the electrophoretic mobility shift assay (EMSA) (Fried and Crothers 1981, Revzin 1989). EMSAs were used in the study of the protein-DNA interactions in the current project. Since the electrophoretic mobility of the protein-DNA complex is affected by many factors, mentioned above, EMSA were optimised in this study to achieve optimum protein-DNA binding. Although the use of the radio-labelled probes is common in EMSA due to their sensitivity, alternative labelled probes were used in this study to avoid the many disadvantages encountered with the use of radioactive materials. All the probes used in this study (See Table 2.3) were composed of between 20 to 30 base pairs. There is no lower limit for the size of the probe as long as all the sequence required for binding of the proposed protein is included (Savery and Busby 1998). Short probes with approximately 25 base pairs are usually used in EMSA (Adams and Fried 2007). The disadvantage of using a very long probe is that its binding by a protein will cause only a small shift in the mobility compared to the free probe unless very large protein molecules bind (Taylor, Ackroyd and Halford 1994). Moreover, large size probes usually have many target sites which can cause difficulty in characterising the binding of a single component, in addition to their slow migration which may require prolonged electrophoretic time (Adams and Fried 2007).

The first alternative labelling system used in the current study was biotinylation of the probe. The EMSA using biotinylated probes was optimised for the amount of the probe, incubation time, temperature, nuclear extract and the binding buffer. EMSA was first optimised using the LightShift chemiluminescent kit according to the manufacturer's instruction (Pierce, UK). 20 fmole of the pPZ ASE and dPZ ASE probes were each incubated with 4 µg of nuclear extract in binding buffer-3 (SeeSection 3.3.3) for 20 minutes at room temperature. Three clear bands (A-C) with both probes and one faint band (D)

with only pPZ ASE were observed indicating protein-DNA binding (Fig 3.3). The experiment was repeated with the same parameters using non-specific competitors to confirm the specificity of the bound protein (Fig 3.4). Bands A-C were not affected in the presence of the non-specific competitors whilst band D was competed by both competitors. The low signal of the bands could have been due to numerous experimental parameters, such as poor transfer from the gel to the nylon membrane (Fig 5, band 4), insufficient time of incubation of the binding reaction (20 minutes) or the quality of the nuclear extract. To further optimise the binding reaction, the incubation time of the probes with the nuclear extract was increased from 20 to 30 minutes (Cheng *et al.* 2002) and a new nuclear extract (See Section 3.2.1) prepared with an alternative procedure (Andrews and Faller 1991). EMSA were performed with 20 fmole of different ASEs and the PEA3 consensus probes. EMSA with these parameters showed clear and more intense protein-DNA complex bands (A-D). However, the signal detected from the free probes was too strong, indicating that the probe was in significant excess (Fig 3.5). Therefore, the amount of the probe used was further optimized since using too much probe may exceed the dissociation constant of the protein affecting the equilibrium of the DNA-protein binding giving a false positive result (Rippe 1997).

Competitive EMSA using the 10 fmole of dPZ ASE probe with the new nuclear extract (Andrews and Faller 1991) and 100-fold excess of different competitors was performed at room temperature with 30 minutes incubation time (Fig 3.6). The competitive EMSA showed a relatively good binding reaction with very weak competition with the competitors. In order to further optimise the protein-DNA interaction alternative binding buffers (SeeSection 3.3.3) were used with the same parameters (Fig 3.7). Moreover, a new control containing the binding buffer with 4 µg of the HepG2 NE, but without the probe was included. The results showed better binding reactions using both new binding buffers (SeeSection 3.3.3) with buffer-1 giving the clearest and most intense signal. Moreover, this EMSA has revealed that the three bands A, B and C were not protein-DNA complexes, since these bands appeared whenever HepG2 NE was used, even without the probe present (Fig 3.7 lane 2). A similar artefact (band) was reported previously in chemiluminescent EMSA with HepG2 NE in the absence of probe and referred to as a non-specific band (Liang, Hassett

and Omiecinski 2005). Although these bands (A, B and C) were not due to the specific interaction of protein with the biotinylated probe, they were likely still due to a specific interaction of the streptavidin and therefore referred to as non-specific protein. Biotin has been identified to bind histones, such as H4, in human and other eukaryotes and plays a role in gene expression (Hassan and Zempleni 2008). Therefore, a control containing nuclear extract without probe was used to indicate the presence of non-specific proteins within the nuclear extract in subsequent assays. An advantage of these non-specific proteins is that they can be used as a good control for the transfer of the probes and the binding of streptavidin.

Competitive EMSA of the pPZ ASE was performed with the optimised parameters, binding buffer-1, 10 fmole of the probe and 4µg of the HepG2 NE with an incubation time of 30 minutes at room temperature (Moeenrezakhanlou, Nandan and Reiner 2008). The competitors were pre-incubated (Jing *et al.* 2004) for 10 minutes with the HepG2 NE in binding buffer-1 to improve the competition. The results showed the three bands (A, B and C) representing the biotinylated protein and clear strong intense bands (1 and 2), representing the protein-DNA complexes (Fig 3.8). The protein-DNA complexes were not competed with a 200-fold excess of non-specific competitors AP2 and HNF3B while weak competition with dPZ ASE and strong competition with PEA3 consensus were observed. Another competitive EMSA was again performed for the dPZ ASE probe with the same parameters to ensure that they were optimised for the protein-DNA interaction (Fig 3.9). The results again showed optimum protein-DNA complex bands with no competition with the 200-fold excess of AP2 and HNF3B non-specific competitors and strong competition with the pPZ ASE and PEA3 binding consensus. The results from the last two experiments (Fig 3.8 and 3.9) indicate that optimum protein-DNA binding can be accomplished using 10 fmole of the probe and 4 µg of HepG2 NE incubated in a 20 µl total volume with binding buffer-1 for 30 minutes at room temperature.

To identify the protein bound to the pPZ ASE and dPZ ASE, supershift assays were carried out using antibodies against the putative proteins bound to those sites. Bioinformatic analysis had revealed that the protein bound to the pPZ ASE and dPZ ASE could be either PEA3 or Ets1. The potential binding of these two proteins was confirmed to some extent by competitive EMSA using

the unlabelled PEA3 binding site consensus. No Ets1 binding site consensus was mentioned in competitive EMSA since it has the same sequences with the PEA3 binding site sequence (Santa Cruz literature). Supershift assays were performed for the pPZ ASE binding site biotinylated probe with the PEA3 antibody using the same parameters mentioned above for the optimum protein-DNA binding reaction. Moreover, the antibody was pre-incubated with the HepG2 NE for 10 minutes at room temperature (Han *et al.* 2002) in binding buffer-1 before adding the probe for a further 30 minutes at room temperature. The results showed only a weak reduction in the intensity of the protein-DNA complex bands 1 and 2, but no supershift band (Fig 3.10). The same parameters were used in other supershift assays again for the pPZ ASE with a prolonged incubation (overnight and for 90 minutes) of the PEA3 antibody with the HepG2 NE at 4°C (Poch *et al.* 2005, Boucher, Piechocki and Hines 1995, Zhang *et al.* 2001). The use of lower temperatures can avoid the potential degradation of the bound protein from the prolonged incubation. The results (Fig 3.11) with both conditions showed only the bands representing the biotinylated proteins (bands A-C). The absence of both supershift and protein-DNA complexes was considered to be due to protein degradation. To determine if the lack of protein-DNA complexes may be due to the degradation of the protein the supershift assay was repeated overnight at 4°C (Fig 3.12) with fresh protease inhibitor cocktail. Protease inhibitor was added to prevent protein degradation, suggesting that the problem was likely not due to protein degradation from protease activity.

Supershift assays were further performed for the pPZ ASE using Ets1 antibody on ice (Woo *et al.* 2002, Penner *et al.* 2002) and at room temperature. A protein-DNA complex was seen at room temperature, but no supershift band was observed. However, on ice the protein was not able to bind to the DNA (Fig 3.13). Further supershift assays were performed for the other proposed binding site, dPZ ASE, using the PEA3 and the Ets1 antibodies at room temperature and at 37°C (Fig 3.14). The antibody was pre-incubated with the HepG2 NE for 10 minutes and further incubated with the probe for 30 minutes at room temperature. A short incubation time at 37°C is to avoid potential degradation of proteins. Clear Protein-DNA complex bands were observed at room temperature, but again no supershift bands were seen in the presence of

the antibodies. However, although there was no supershift band, a much stronger complex band was seen in the presence of the Ets1 antibody (four times greater) compared to the complex in the absence of the antibody (Fig 3.14b). The assay at 37°C did not show protein-DNA complexes, however, a complex was seen in the presence of the Ets1 antibody (Fig 3.14a). The results from the supershift assays of the dPZ ASE at room temperature and 37°C has given some indication to a potential binding of the Ets1 protein. Further supershift assays were performed with an incubation of the HepG2 NE with the Ets1 antibody overnight at 4°C prior to the addition of the dPZ ASE probe followed by a further incubation for 30 minutes at room temperature. The results showed no protein-DNA complexes (Fig 3.15).

To determine if the antibody-protein binding reaction was optimised and PEA3 was not the protein which bound the ASE regions, a supershift assay was performed using the PEA3 binding site consensus with the PEA3 antibody at room temperature. In optimum reaction conditions this assay should show a supershift band. The assay again did not show a supershift band (Fig 3.16). The assay was repeated using dialysed (Kim and Nelson 1998) HepG2 NE to reduce the salt concentration which might affect the antibody-protein binding. The results again showed protein-DNA binding but no supershift band (Fig 3.17).

EMSA using biotinylated probes although they are sensitive (Fig 3.1), they are time consuming. In addition, the DNA needs to be transferred to the nylon membrane (Section 3.2.4) before the detection which sometimes may cause loss during this process (Fig 3.5). Therefore, infrared fluorescently-labelled probes, which can be detected directly with the Li-cor Odyssey system (SeeSection 3.3.5) without the transfer process, were used in EMSA in the current study. EMSA and supershift assays were performed using both pPZ ASE and dPZ ASE probes with the Ets1 antibody. The antibody was pre-incubated with 12 µg of HepG2 NE for 10 minutes at room temperature and further incubated with100 fmole of the probe. Protein-DNA complexes were observed with the pPZ ASE probe, but no supershifted band showing the antibody-protein-DNA complex (Fig 3.18a). However, the dPZ ASE probe did not show even protein binding (Fig 3.18b).

Further bioinformatic analysis revealed a potential binding site for the CREB protein to the pPZ ASE. Therefore, supershift assays were firstly performed using the CREB antibody with a CREB binding site consensus probe to optimize the formation of the antibody-protein complex before starting to use the antibody with the pPZ ASE probe. The same parameters were used as in the previous experiment, 10 minutes pre-incubation of HepG2 NE with the antibody and 30 minutes incubation with 100 fmole CREB binding site consensus probe. Binding buffer-1 was used without dithiothreitol (DTT). The assay was done in the binding buffer without DTT since this reducing agent disrupts disulfide bonds of proteins (Chen *et al.* 2003) and therefore may affect the activity of the antibody. The assay, again, showed a protein-DNA complex band, but no shifted band due to the antibody binding (Fig 3.19). The same results were seen when the antibody pre-incubated with the HepG2 NE for one hour at 4°C before their incubation with probe for 30 minutes at room temperature (Fig 3.20). The experiment was performed at 4°C to avoid potential degradation of the protein from prolonged incubation. The same supershift assay was performed with a 20 minute pre-incubation time of the antibody with the HepG2 NE at 37°C. The results showed neither protein-DNA complex band formation nor a complex in the presence of the antibody (Fig 3.21). As the supershift assay could not be optimised with different incubation periods and different temperatures, the optimisation of the antibody binding to the protein was carried out again using a new protocol modified from Santa Cruz biotechnology (SeeSection 3.4.2). The results with this protocol did not show any distinct bands (Fig 3.22).

Competitive EMSAs and supershift assays were also performed for the putative proximal HNF4α (pPZNHF4α) binding site. Competitive EMSA were first optimised using 100 fmole of the pPZHNF4α added to 10 minutes pre-incubated 12μg HepG2 NE with the 100-fold of the competitor in binding buffer-1 at room temperature and further incubated for 30 minutes at room temperature. The assay showed a typical protein-DNA complex with abolition of the complex with 100-fold excess of specific competitors (self and HNF4α binding site consensus) and negligible with the 100-fold excess of non-specific competitors (Fig 3.23). A supershift assay was then performed with the same parameters with a repetition of the competition with the specific competitors.

The HNF4α antibody was pre-incubated with the HepG2 NE for 10 minutes at room temperature and the pPZHNF4α probe was added for further 30 minute incubation at room temperature. Strong competition again was seen with the specific competitors in addition to abolition of the protein-DNA complex band in the presence of the HNF4α antibody (Fig 3.24). The loss of the protein-DNA complex rather than a complex of increased size that would have a lower mobility upon electrophoresis could be due to either the antibody binding to or sterically hindering the DNA binding domain of the protein or to the size of the complex being too large to efficiently migrate into the gel matrix. To confirm the latter hypothesis the intensity of the fluorescent band, indicated by arrow 2 (Fig 3.24), in the well with the antibody was compared to the fluorescent intensity in the well without the antibody. The fluorescence in the antibody well showed approximately 3.7-fold greater intensity compared with the other wells without antibody. The loss of the protein-DNA complex with the greater intensity of the fluorescence in the well in the presence of the antibody confirmed, to some extent, the binding of the antibody to the protein and the hindrance of migration rather than the antibody recognising the DNA binding domain of the protein and preventing its binding to the DNA (Carey and Smale 2001), resulting in a loss of the protein-DNA complex.

To confirm the previous result the supershift assay was performed using the HNF4α binding site consensus probe with HNF4α antibody. The results showed a similar picture to that with the pPZHNF4α (Fig 3.25). The protein-DNA complex was dramatically reduced in the presence of the antibody and the well showed fluorescent intensity of approximately 4-fold more than the intensity in the well without the antibody.

To determine whether a better supershifted complex could be obtained, more supershift assays were performed on ice (Ye, Zhang and Dong 1996) and at 4°C (Heltemes, Tuggle and Lamont 1999) with different incubation times with the antibody followed by further incubation with the probe at room temperature. However, although there were some changes in the intensity of the protein-DNA complexes in the presence of the antibody, those assays did not show any clear indications for antibody-protein-DNA complexes (Fig 3.26, Fig 3.27 and Fig 3.28). Sugawara and colleagues reported the binding of the HNF4α protein to the PZ gene promoter which was confirmed using the HNF4α antibody in a

supershift assay (Sugawara *et al.* 2007). The same binding buffer (See Section 3.4.2) used by the Sugawara group 2007 was used in performing EMSA in the current study. Nevertheless, results from EMSA using this binding buffer did not show distinct protein-DNA complexes (Fig 3.29); so this buffer was not used in subsequent supershift assays. The reason that no clear protein-DNA complex was seen may be due to optimisation of this buffer only for radio-labelled probes. The supershift assays with the parameters showing the best results (Fig 3.24 and Fig 3.25) were repeated to confirm their quality and consistency. The repeated assays again showed the best results (Fig 3.30 and Fig 3.31) compared to the assays performed with other different parameters throughout the optimisation. The fluorescence intensity in the antibody well with pPZHNF4α probe (Fig 3.31) was greater than the antibody well with the HNF4α binding consensus probe (Fig 3.30). This might be due to some of the protein bound to the HNF4α binding consensus probe was not bound by the antibody, which was indicated by the faint band of protein-DNA complex in the antibody well.

It can be concluded that optimum EMSA and supershift assays can be accomplished using 10 fmole or 100 fmole of the biotin-labelled or fluorescent-labelled probes, respectively, with 4 µg of HepG2 NE for the biotinylated probes and 12 µg for the fluorescent probes. The assays should be incubated at room temperature and use binding buffer-1 for optimal complex formation. These parameters were used to perform all EMSAs and supershift assays throughout the current study. Although EMSA is an easy and sensitive method in studying protein-DNA interactions, many limitations encountered in this assay were reported. For instance, transcription factors that require another protein for its binding to DNA may not be optimised by *in vitro* conditions (Wells and Farnham 2002), in addition different results may be obtained for the same DNA-protein system using different electrophoresis conditions (Sidorova, Muradymov and Rau 2005). Moreover, DNA secondary structure, such as looping out of intervening sequences which is required for bringing together of two distal protein binding sites into close proximity *in vivo*, cannot be created *in vitro* (Wells and Farnham 2002). The information of the DNA region location required for the protein binding can not be obtained directly by EMSA (Hellman

and Fried 2007) and needs to be determined indirectly by bioinformatic analysis or directly by footprinting (Brenowitz *et al.* 1986).

Chapter 4-Liver specific transcriptional regulation of Protein Z

4.1. Introduction

The regulation of gene product synthesis is essential for the expression pattern of a particular gene in a tissue-specific or developmental stage-specific way (Latchman 1997). Whilst the majority of genes differ in their expression from one tissue to another, the expression levels of tissue enriched genes is higher in a particular tissue type, controlling its specialised function, and much lower or negligible in other tissues (She *et al.* 2009). The genomic length of tissue specific genes was found to be shorter (less base pairs) and their expression was lower compared to house keeping genes according to the expressed sequence tag data analysis (Zhu *et al.* 2008). House keeping genes are expressed in all tissues and are needed to preserve basic cellular functions. Most house keeping genes have regulatory sequences termed CpG islands associated with them, but lack TATA-boxes in their proximal promoters whilst the majority of tissue specific genes lack both CpG-islands and TATA-boxes (Zhu *et al.* 2008).

4.1.1. Liver specificity

The liver is the organ responsible for many important functions such as synthesis and detoxification of plasma proteins, carbohydrates and fats, protein metabolism and many other processes. Moreover, this organ is considered as unique in its production of protein components required in coagulation, anticoagulation, fibrinolysis and the complement system. There have been 6,788 proteins identified in the human liver proteome database. Approximately 67% of the proteins taking part in the coagulation, anticoagulation and fibrinolysis systems were distinctly recognized and are significantly enriched in the liver (Jiang *et al.* 2009). Liver-specific expression is controlled by a network of liver-enriched transcription factors. The idea that these hepatic transcription factors set up a complex cross-regulatory network to ensure their abundance in liver cells might be attributed to the presence of binding regions for members of the hepatic transcription factors within their own promoters leading to a feedback loop (Kyrmizi *et al.* 2006). Kyrmizi *et al.* also suggest that the stability of expression of individual hepatic transcription factors is determined by the multiple regulatory interactions within the liver network (Kyrmizi *et al.* 2006). Six

families of liver-enriched transcription factors, the homeodomain containing hepatocyte nuclear factor 1 (HNF1) family, the leucine zipper domain family including CCAAT-enhancer-binding proteins (C/EBP), the winged helix domain family including hepatocyte nuclear factor 3 (HNF3), the zinc finger domain nuclear receptor family including hepatocyte nuclear factor 4 (HNF4), the onecut family (new class of homeodomain) including hepatocyte nuclear factor 6 (HNF6) and PAR of the leucine zipper domain family including D site-binding protein have been characterised and shown to be involved in liver-confined gene expression in adult hepatocytes (Schrem, Klempnauer and Borlak 2002, Schrem, Klempnauer and Borlak 2004). Furthermore, HNF4 expression was shown to be specific for hepatocyte-lineage cells in the heterogeneous cultured mouse embryonic stem cells (Kheolamai and Dickson 2009).

4.1.2. Hepatocyte nuclear factor 4

HNF4 expression is established at the very earliest stages of embryonic development. HNF4 messenger RNA (mRNA) can be detected in the endoderm of the blastocyst at embryonic day 4.5. The expression of HNF4 continues throughout embryonic development in the endoderm of the extraembryonic visceral cells of the yolk sac (Watt, Garrison and Duncan 2003). At embryonic day 8.5 of the developing mouse foetus the liver diverticulum and hindgut starts to express HNF4 mRNA followed by expression in the pancreas, stomach, intestine and kidney (Duncan *et al.* 1994). Saldek *et al.* found the expression of this factor is limited to the liver, kidney and intestine and not expressed in spleen, brain, lung and heart of rat (Sladek *et al.* 1990). The amino acid sequence of HNF4 revealed that this protein is a member of the nuclear receptor superfamily which includes steroid and thyroid hormone receptors and non steroid ligand-dependent transcription factors (Sladek *et al.* 1990). This transcription factor is different from many other group members as it localizes to the nucleus and its transcriptional activity occurs via homodimer formation. Other members of this group which show the ability to dimerize with themselves are present in the cytoplasm while others that show an ability to heterodimerize are found in the nucleus (Jiang *et al.* 1995). The proteins belonging to this superfamily contains small DNA-binding peptide motifs known as zinc fingers (Schrem, Klempnauer and Borlak 2002). These small motifs can

be used as modular building blocks to achieve the recognition of specific sequence of DNA (Klug 1999). The region of the two potential zinc fingers between amino acids 50 and 116 in the HNF4α protein was found to show 40-63% homology to the zinc finger domain found in other members of the nuclear receptor superfamily (Sladek *et al.* 1990). However, HNF4α was not identified to bind any ligand and, therefore, initially classified as an orphan nuclear receptor. The transactivation of HNF4α is seen in many cell types suggesting that its ligand is present in all cells or is not needed for its activation (Cereghini 1996). Hertz *et al.* went on to show that recombinant rat HNF4α is transcriptionally regulated by the naturally occurring ligands, fatty acyl-CoA thioesters (Hertz *et al.* 1998). This has been confirmed when the rat recombinant HNF4 ligand-binding domain was shown to be bound by fatty acyl-CoA by direct molecular interaction (Petrescu *et al.* 2002). Therefore, HNF4 should be considered as a nuclear receptor ligand-dependent transcription factor and not as an orphan nuclear receptor.

4.1.3. HNF4α binding to the PZ gene promoter

The HNF4α transcription factor was found to play a major role in the expression of a number of blood coagulation factors that are normally produced in the liver and will be discussed below. HNF4α has been found to play an important role in the transcription of PZ with the PZ gene promoter bound by HNF4α between nucleotides -70 and -30, relative to the translation start site (Sugawara *et al.* 2007). Bioinformatic analysis (ALGGEN-PROMO) identified proximal sequences within the PZ gene promoter that show homology to the HNF4α binding site consensus sequence spanning, nucleotides -48 to-38 (5'-CTGGACTTTGG-'3) and distal sequences from -310 to -300 (5'-TGTGACTTTGT-3') relative to the translation start site (Fig 4.1).

Other coagulation factors, which are believed to share a common evolutionary origin with PZ, and are suggested to have evolved from the same ancestral gene (Ichinose *et al.* 1990, Fujimaki *et al.* 1998), which include Prothrombin, Factor (F) VII, FIX and FX, have all been reported to be regulated by liver-enriched transcription factors, as would be predicted from their expression profile. FVII, FVIII, FIX, FX and FXI are activated by the binding of HNF4α to its

binding sites within their promoters (Stauffer *et al.* 1998, Figueiredo and Brownlee 1995, Reijnen *et al.* 1992, Miao *et al.* 1992, Tarumi *et al.* 2002).

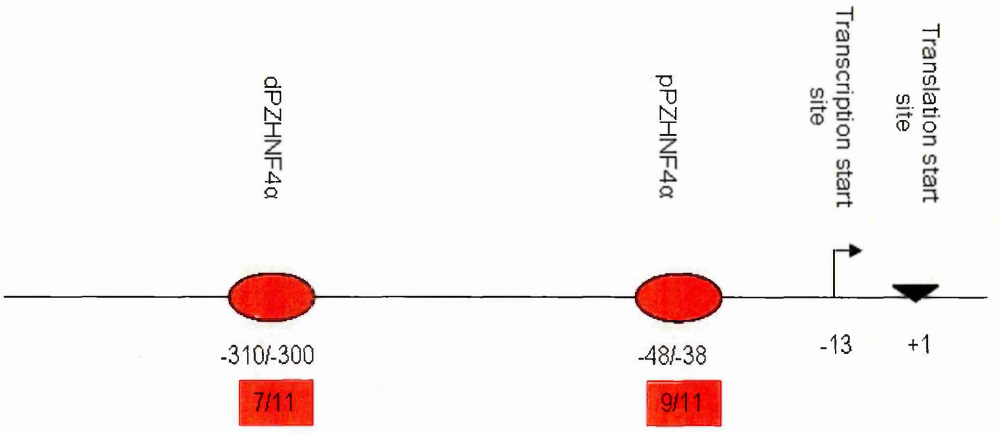


Figure 4.1. The previously reported (pPZHNF4α) and hypothetical (dPZHNF4α) HNF4α binding sites in the PZ gene promoter. Highlighted numbers show the degree of homology with the 11bp HNF4α consensus. Numbering is relative to the translation start site at +1.

A number of patients with haemophilia B, FIX deficiency, and patients with severe deficiency of FVII have been found to have mutations in critical HNF4α binding sites within their respective gene promoters (Reijnen *et al.* 1992, Arbini *et al.* 1997). However, the levels of FVII, FVIII and FX in HNF4α-null mice show no significant variation, suggesting evolutionary diversity and possible functional redundancy in regulatory sequences between mouse and human systems (Inoue *et al.* 2006). By contrast, FIX levels are reduced in HNF4α-null mice. Similarly, HNF4α-null mice do show an abolition of expression of FXII and FXIIIB expression in the adult liver and promoter activity has been found to be totally dependent on the binding of HNF4α to its binding sites (Inoue *et al.* 2006). Therefore, mutation in HNF4α binding site of factors XII and XIIIB might cause potential deficiency in these coagulation factors.

Protein S, which is a cofactor for activated protein C in the inactivation of FVa and FVIIIa and is expressed in a number of tissues including hepatocytes, vascular endothelial cells and megakaryocytes (Malm *et al.* 1994), has been reported to be activated by the liver-enriched factor HNF4α (Hall, Peake and Winship 2006). The characterised binding of HNF4α to the promoter of the protein S gene (nucleotides -415 to -375) in HepG2 cells may contribute to its

liver specific transcription (Hall, Peake and Winship 2006). As HNF4α has been demonstrated to regulate the transcription of many coagulation factors, it was hypothesised that, given its close ancestral origin to these factors, the transcription of PZ may be similarly regulated by HNF4.

4.2. Aims

To characterise the role of HNF4α in the regulation of PZ transcription by gene expression analysis of the endogenous PZ gene in the human liver-derived cell line HepG2 and to define such regulatory protein-DNA interactions.

4.3. Methods

4.3.1. Protein-DNA binding analysis

4.3.1.1. Electrophoretic Mobility Shift Assay (EMSA)

EMSAs were performed for the two potential binding sites of HNF4α identified using bioinformatic analysis (ALGGEN-PROMO) in the PZ gene promoter, the proximal site (pPZHNF4α) spanning nucleotides -48 to -38 and the distal site (dPZHNF4α) spanning nucleotides -310 to -300 relative to the translation start site, in order to demonstrate protein-DNA binding. 100 fmole of the candidate probe (See Table 2.3), 1 µg of poly dI-dC, 12 µg of HepG2 nuclear extract (HepG2 NE) (See Section 2.2.1) in 20 µl binding buffer (See Section 2.2.4.1) were used. The binding mixture was incubated for 30 minutes at room temperature and size-fractionated through 5% w/v, 19:1 acrylamide/bis-acrylamide (See Section 2.2.4.1) mini (8 x 8 x 0.1 cm) gel, using a voltage of ~ 12.5 V/cm, in 1x Tris-Borate EDTA (TBE) buffer. The protein-DNA complex was visualised (SeeSection 3.3.5) using the Odyssey Licor Imaging System (Applied Biosystems, UK) and identified by slower migration through the gel compared with the free probe. In competitive EMSAs, the indicated fold molar excess of unlabelled competitors was incubated for 10 minutes with the Hep NE at room temperature prior to addition of the labelled probe and further incubation.

4.3.1.2. Supershift assays

Specific binding of the protein to the probe, based on competitive EMSA, was further analysed using supershift assays for specific identification. 2µg of HNF4α antibody (H-171, Santa Cruz, Germany) was pre-incubated with the 12 µg HepG2 NE for 10 minutes at room temperature before the addition of 100 fmole of the probe and a further 30 minute incubation. The binding mixture was size fractionated through a 5% polyacrylamide gel (See Section 2.2.4.1). A typical antibody-protein-DNA complex is usually characterised by a shift in the protein-DNA complex due to its increased size.

4.3.2. Gene expression studies of the endogenous PZ gene in HepG2 cells

4.3.2.1. Comparison of PZ gene promoter sequences in HepG2 cells with the wild type promoter.

To sequence the PZ gene promoter present in HepG2 cells, the nucleotide sequence of the human wild type PZ gene promoter (Homo sapiens protein Z, AF440358, National Centre for Biotechnology Information) spanning -18845 to +106 relative to the translation site was divided into three overlapping regions (See Section 2.1). Primers for each region (See Table 2.1) were designed (See Section 2.1) to allow amplification by the polymerase chain reaction (PCR). PCR (See Section 2.1.1) were optimised to amplify each region using genomic DNA template prepared by the GenElute Mammalian Genomic DNA miniprep kit (Sigma-Aldrich, UK), from cultured HepG2 cells, and the respective primers. Each amplified fragment was size-fractionated on a 1% w/v agarose gel (See Sections 2.1.1 and 2.1.2), and visualised using a UVP BioImaging system. The fragments amplified for each region were then purified (See Section 2.1.3) and sent to the Genetic Core Facility of Sheffield Medical School for sequencing. The sequencing was performed using the Applied Biosystems BigDye v.3.1 ® Terminator sequencing chemistry and analysed by the ABI3730 capillary electrophoresis system (Applied Biosystems, UK). The sequence of the PZ gene promoter in HepG2 cells was compared with the sequence of the wild type promoter.

4.3.2.2. Expression plasmid verification

The identity of the expression vector pCMV6-XL4 (OriGene Technologies, USA), ~4.7 kb, with the HNF4α cDNA (3.6 kb), which was used in transfection of HepG2 cells, was confirmed by digestion with restriction enzymes (See Section 2.4.3.3). Sufficient quantities of plasmid were prepared for transfection (See Section 2.4.3). The webcutter 2.0 programme was used to determine the appropriate restriction enzyme for the HNF4α insert and the resultant fragments produced after digestion. The restriction enzyme Sac1 (Fermentas life science, UK) cuts the insert at nucleotides 721 and 3024 and was used to digest the plasmid encoding HNF4α for verification purposes. The resultant fragments were analysed using a1% v/w agarose gel.

4.3.2.3. Extraction of RNA

RNA was extracted using the GenElute mammalian total RNA miniprep kit (See Section 2.3.1) from untransfected HepG2 and HepG2 cells transfected with the HNF4α expression plasmid. The quality of the extracted RNA was determined by size fractionation through a 1% v/w agarose gel and assessment of the ribosomal 18S and 28S bands. Purity and concentration were measured using the NanoDrop®ND-1000 spectrophotometer (NanoDrop Technologies, USA). The RNA was treated with DNase 1 (Sigma-Aldrich, UK) to remove any contaminating genomic DNA. The RNA was then used to synthesise cDNA (See Section 2.3.3) using the iScript cDNA synthesis kit (Bio-Rad, UK) for use in the qPCR assays.

4.3.2.4. Primer efficiency

To calculate the efficiency of the primer pairs for qPCR, template cDNA was serially diluted (undiluted, 1:10, 1:100 and 1:1000) and each PCR reaction was carried out in triplicate. The C_q values obtained from each dilution were plotted against the logarithm of the cDNA concentrations and the slope of the line of best fit determined (See Section 2.3.4.1). The efficiency was then calculated by the equation: Efficiency = $-1 + 10^{(-1/\text{slope})}$

A slope of -3.32 indicates a primer efficiency of 100%.

4.3.2.5. Reference gene Validation

The reference gene for normalising qPCR quantification was selected by performing qPCR (See Section 2.3.4.3) on a set of housekeeping genes, β-actin, GADPH, UBC, and YWHAZ (See Table 2.4), taken from Vandesompele *et al.* (Vandesompele *et al.* 2002), using cDNA from untransfected and HNF4α transfected HepG2 cells. The most stable housekeeping gene, which showed the least variation in its relative expression between the untransfected and HNF4α transfected HepG2 cells, was then used as a reference gene to study the effect of HNF4α overexpression and knockdown on PZ gene expression in the HepG2 cell model.

4.3.2.6. The effect of HNF4α overexpression in HepG2 cells on endogenous PZ transcript levels

Tfx™-20 reagent (Promega corporation, UK) was used for the transfection (See Section 2.4.4) of HepG2 cells with HNF4α encoding plasmid (Fig 2.7) according to the manufacturer's instructions. The effect of HNF4α overexpression in HepG2 cells on the endogenous PZ gene was measured by qPCR using SYBR green chemistry (See Section 2.3.4.3).

The relative expression ratio of the PZ gene was normalised to the reference gene, UBC, and calculated using Pfaffl equation (Pfaffl 2001).

$$\text{Ratio} = \frac{(\text{E Target})^{\Delta \text{Cq target (Control-Sample)}}}{(\text{E Reference})^{\Delta \text{Cq reference (Control-Sample)}}}$$

E Target is the amplification efficiency of the primer pair specific for the gene of interest. E Reference is the amplification efficiency of the primer pair specific for the reference gene. ΔCq target is the difference between the quantification cycle (Cq) of the gene of interest in the control cells and the sample cells. ΔCq reference is the difference between the quantification cycle (Cq) of the reference gene in the control cells and the sample cells.

The data was analysed using the Pfaffl equation since the efficiency (86%) of the gene of interest (PZ) and the efficiency (95%) of the reference gene (UBC) primers showed some discrepancy. Melt curve analysis (See Section 2.3.4.3) was carried out to detect non-specific products and primer dimers.

4.3.2.7. The effect of HNF4α Knock down on the levels of endogenous PZ expression

HNF4α expression in HepG2 cells was knocked down using a smart pool siRNA (Dharmacon RNAi Technologies, UK) according to the protocol provided. 2μM of the siRNA and 2 μl of the transfection reagent in a total volume of 500 μl of medium were added to the cells and incubated for 24 hours (See Section 2.4.5). A positive control (ON-Targetplus GAPDH control pool) was used to optimise the protocol for transfection of the HepG2 cells and a negative control (ON-Targetplus Non-Targeting pool) was used to make sure the observed

effects were specific to the sequence being targeted by the particular siRNA being used in the experiment. The efficiencies of the knock down of the candidate gene expression and its effect on PZ were measured using qPCR (See Section 2.3.4.3).

4.3.2.8. Comparison of the relative HNF4α mRNA expression in HepG2 to normal liver tissue.

To compare the relative expression of HNF4α mRNA in HepG2 cells to normal liver tissue, RNA was prepared from the HepG2 cells (See Section 2.3.1) and a commercial preparation of total RNA from normal adult human liver was purchased (AMS biotechnology Ltd, UK) The two RNA samples were reverse transcribed to cDNA (See Section 2.3.3) and subjected to qPCR (See Section 2.3.4.3). The relative expression of HNF4α mRNA in HepG2 cells and normal liver was normalised to the UBC reference gene. For more details see section 2.4.2.

4.3.2.9. Data analysis

EMSA analysis: To determine protein-DNA binding specificity unrelated dsDNA competitors were used. Specific protein-DNA complexes show no diminution in the presence of a fold-excess of unrelated competitors. Competition was determined by measuring the intensity of the binding complex in the presence of a competitor and compared with the intensity of the complex in the absence of the competitor. The intensity of the protein-DNA complexes was measured using the Li-Cor instrument and quantitative software (Li-Cor, UK).

Real time PCR analysis: The fold increase of the target gene (PZ) after HNF4α overexpression was normalised by UBC and determined by Pfaffl equation (See Section 4.3.2.6), since the primer efficiency of the reference gene and the target gene were not close to each other. Basal gene expression was first expressed as the change in the quantification cycle (ΔCq) for the gene of interest and for the reference gene, where the quantification cycle (Cq) was the detectable fluorescence above the background for the candidate genes.

The fold decrease of PZ and HNF4α after HNF4α siRNA knockdown was determined using the formula $2^{-\Delta Cq}$, where 2 was the presumed amplification efficiency for the target and reference genes and the ΔCq was the Cq of target gene - the Cq of reference gene. To compare the relative expression of HNF4α and PZ mRNA in HepG2 to normal liver tissue, the formula $2^{-\Delta\Delta Cq}$ was used, where the relative expression in HepG2 was calculated after normalisation with the UBC reference gene and the liver tissue. These two formulas were used since the primers of both the target genes and the reference gene were 100% (Applied Biosystems, UK).

The significant difference between samples and controls was determined using REST 2009 software, developed by M. Pfaffl and Qiagen, taking into account the primers efficiencies and normalisation of reference gene. Randomization techniques are employed to determine whether any differences in expression are significant.

4.3. Results

4.3.1. Protein binding to putative HNF4α sites in the PZ promoter

EMSA reactions using fluorescently labelled dsDNA probes spanning the PZ HNF4α sites, pPZHNF4α (-48/-38) and dPZHNF4α (-310/-300), were shown to be bound by protein present within the nuclear extract from HepG2 cells (HepG2 NE) (Fig 4.2 & 4.3).

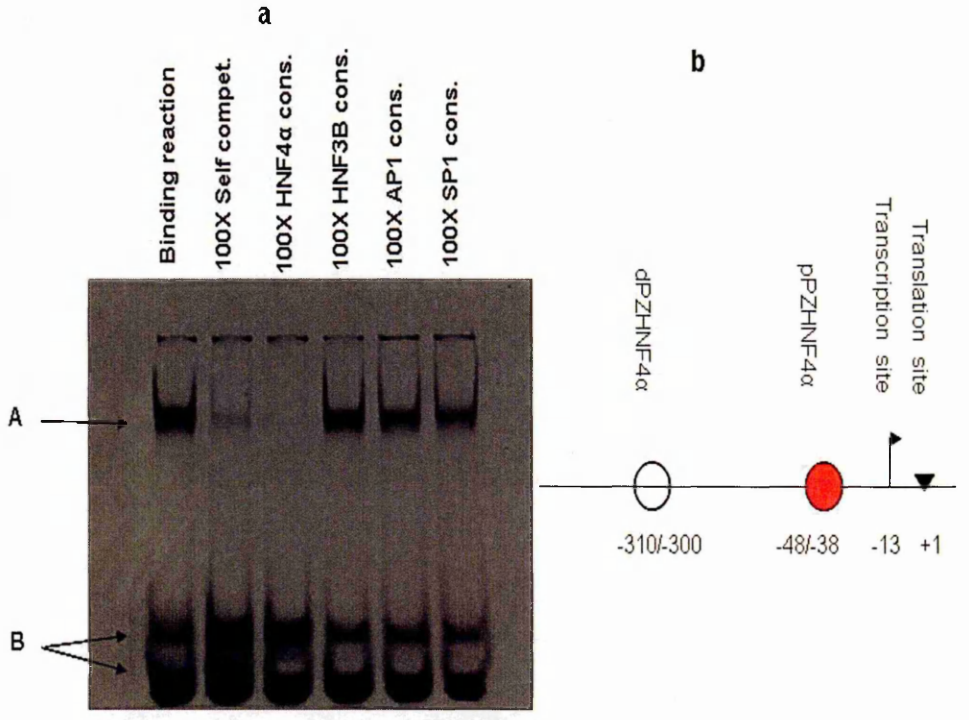


Fig 4.2. EMSA using the reported pPZHNF4α-fluorescent (cy5.5) binding site.
(a) Nucleotide sequences of the probes and competitors are shown in Table 2.2 (Section 2.2.3). The binding reaction lane is the reported HNF4α (pPZHNF4α) probe (-48/-38) with HepG2 NE. Self compet lane is the pPZHNF4α probe, the HepG2 NE and the unlabeled pPZHNF4α. HNF4α cons lane is the pPZHNF4α probe, the HepG2 NE and unlabeled HNF4α binding consensus. HNF3B, AP1 and SP1 cons lanes are the pPZHNF4α probe with the HepG2 NE and unlabeled HNF3B, AP1 and SP1 binding site consensuses added, respectively. Bands indicated by arrow A represent protein DNA complexes. Bands represented by arrow B represent unbound fluorescent probe. Competition is seen with the self competitor and the HNF4α consensus sequence. Negligible competition is seen with the non-specific competitors HNF3B, AP1 and SP1 binding sites. (b) The red circle shows the proximal HNF4α site in the PZ gene promoter while the numbers indicate the hypothetical spanning region.

The binding reaction with the pPZPZHNF4α probe showed one bound complex (Fig 4.2, band A) while dPZHNF4α showed two bound complexes (Fig 4.3, bands A and B). Competition reactions using the pPZHNF4α probe with both a

100-fold excess of unlabelled self and HNF4α binding site consensus sequence competitor showed significant competition of 60% and 100%, respectively. Competition with a 100-fold excess of non-specific competitor, HNF3B, AP1 and SP1, did not show any significant competition (reductions of bound complex of 9%, 17% and 21%, respectively).

Competition reactions using the dPZHNF4α probe demonstrated that the bound complex A (Fig 4.3) showed specificity to the probe sequence, with no observable competition with 50, 100 and 200-fold excess of HNF4α consensus binding site and non-specific competitors. The bound complex B, however, was demonstrated to be non-specific to the sequence, with 50, 100 and 200-fold excess of the HNF4α consensus and non-specific AP1 binding site consensus competitors showing a similar effect on bound complex levels (Fig 4.3).

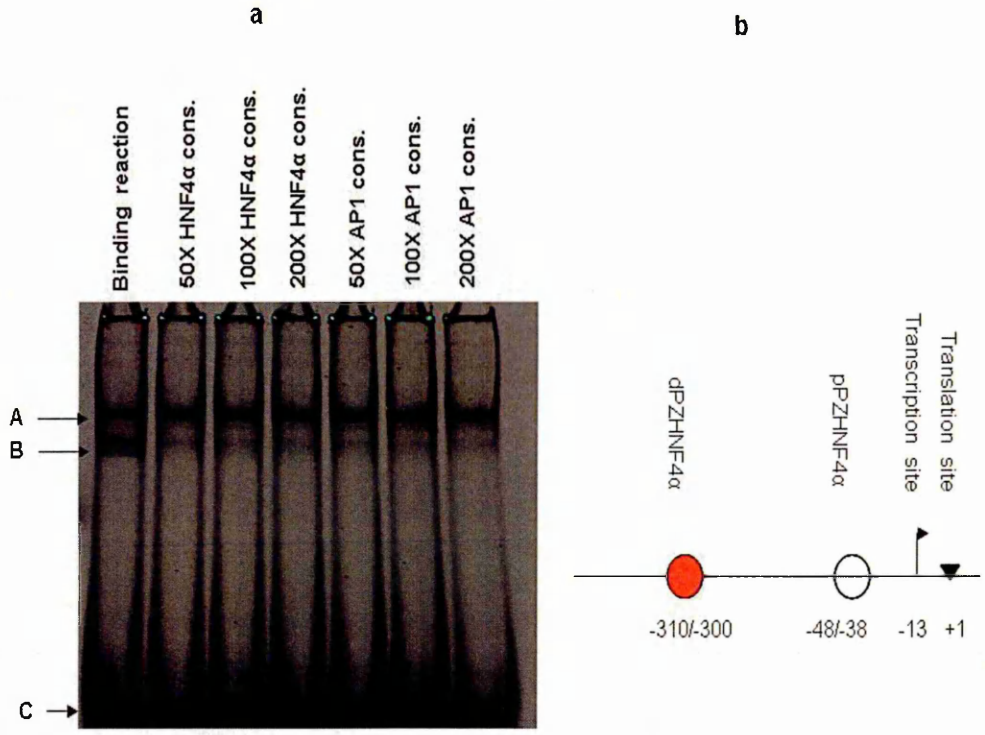


Fig 4.3. EMSA using the putative PZ distal HNF4α-fluorescent (IRD-700) binding site. (**a**) Nucleotide sequences of the probes and competitors are shown in table 2.2 (section 2.2.3). Binding reaction lane is the distal HNF4α (dPZHNF4α) probe (-310/-300) with HepG2 NE. HNF4α lanes are (dPZHNF4α) probe with HepG2 NE and 50, 100 and 200-fold molar excess unlabeled HNF4α consensus added, respectively. Lanes 5, 6 and 7 are (dPZHNF4α) probe with HepG2 NE and 50, 100 and 200-fold excess unlabeled AP1 added, respectively. Bands indicated by arrows A &B represent protein DNA complexes. Bands represented by arrow C represent unbound fluorescent probe. No competition of complex A is seen with increasing fold excess of the unlabeled HNF4α consensus and the unlabeled non-specific competitor AP1 binding site consensus. Competition of complex B is seen with both unlabeled HNF4α consensus and the unlabeled AP1. (**b**) The red circle shows the distal HNF4α site in the PZ gene promoter while the numbers indicate the hypothetical spanning regions.

4.3.2. Confirmation of the identity of the protein bound to the proximal HNF4α site

The assay described here used the HNF4α antibody to specifically identify the protein bound to the pPZHNF4α sequence as HNF4α (Fig 4.4). Again, arrow A indicates a protein-DNA complex and arrow B indicated the free probe. Competition with self competitor and the HNF4α competition both led to a dramatic reduction in the amount of bound probe present. The final reaction with the HNF4α antibody present also showed a loss of the protein DNA-complex and increased levels of non-migrating complex in the well (3.7-fold higher fluorescence intensity (Fig 4.4, band C)) compared with the other wells.

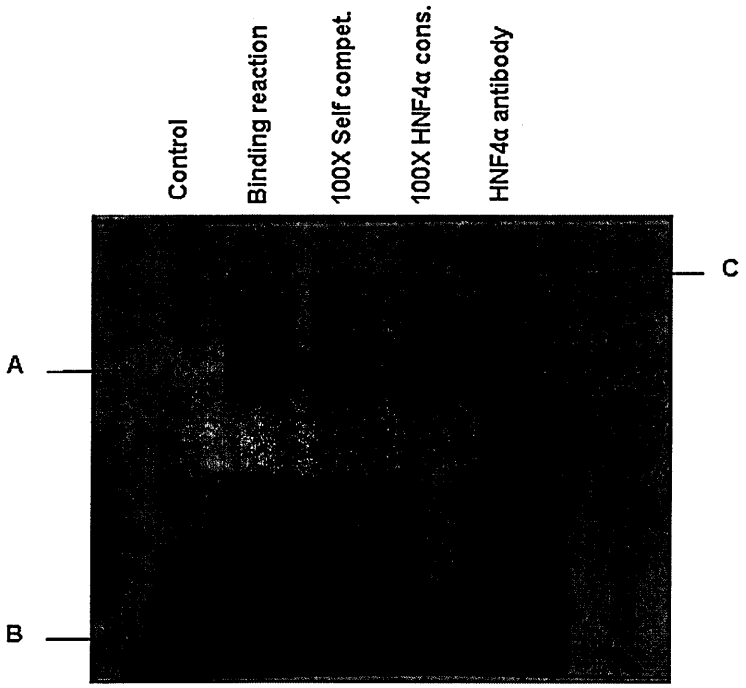


Fig 4.4. Supershift assay using the pPZHNF4α-fluorescent (cy5.5) binding site. Control is the pHNF4α probe (-48/-38) without HepG2 NE. Binding reaction was the pPZHNF4α probe (-48/-38) with HepG2 NE. Self competitor is the pPZHNF4α probe, the HepG2 NE and the unlabeled pPZHNF4α. HNF4α consensus lane is the pPZHNF4α probe, the HepG2 NE and unlabeled HNF4α binding consensus. HNF4α antibody lane contained the pPZHNF4α probe, the Hep NE and the HNF4α antibody. Bands indicated by arrow A represent protein DNA complexes. Bands indicated by arrow B represent unbound fluorescent probe. Band C represent the potential antibody-protein-DNA complex. Competition is seen with 100-fold molar excess of self-competitor and the HNF4α consensus sequences.

A supershift assay with the HNF4α antibody using the HNF4α binding site consensus sequence probe (Fig 4.5) showed very similar results, with a loss of protein-DNA complex (Fig 4.5, band A) and a greater fluorescence intensity

(approximately 4-fold) in the well containing the antibody (Fig 4.5, band C), compared to the well of the binding reaction without the antibody.

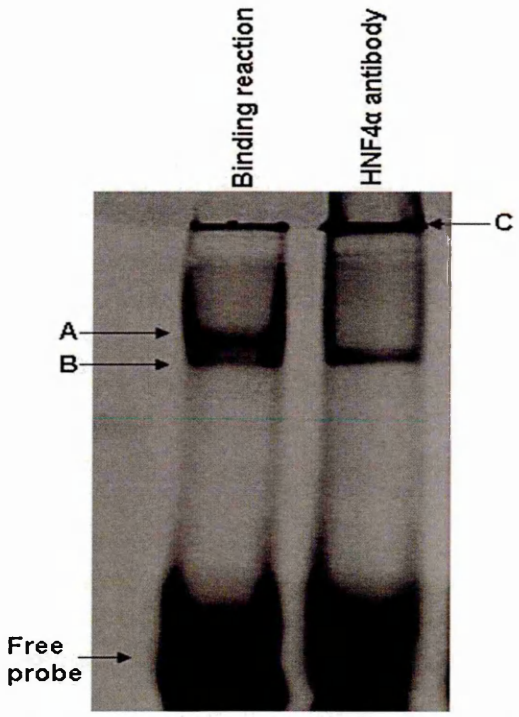


Fig 4.5. Supershift assay using the consensus HNF4α-fluorescent binding site. The binding reaction is the HNF4α binding site consensus probe with HepG2 NE. HNF4α antibody lane contains the HNF4α binding site consensus probe, the HepG2 NE and the HNF4α antibody. Bands indicated by arrow A and B represent protein DNA complexes. Band C represent the potential antibody-protein-DNA complex.

4.3.3. Gene expression studies of the endogenous PZ gene in HepG2 cells

4.3.3.1. Sequencing of PZ gene promoter

The results (See Appendix) of the PZ gene promoter sequencing indicated that the nucleotide sequence of the promoter region (residues -1884 to +106) of the endogenous PZ gene in HepG2 cells was identical to that reported as the reference sequence for human PZ (Homo sapiens protein Z, AF440358, National Centre for Biotechnology Information). The genotypes of the PZ promoter sequence single nucleotide polymorphisms (SNPs) in HepG2 cells were found to be GG at -13, CC at -336, GG at -861, GG at -984, CC at -1185, AA at -1252, TT at -1338, GG at -1498 and TT at -1621.

4.3.3.2. Expression plasmid verification

Restriction enzyme digestion of the expression plasmid encoding HNF4α (OriGene Technologies, UK) using the SacI restriction enzyme (Fermentas life science) showed three bands of 5.27 kb, 2.3 kb and 0.72 kb (Fig 4.6, lane 3) on 1% w/v agaros gel.

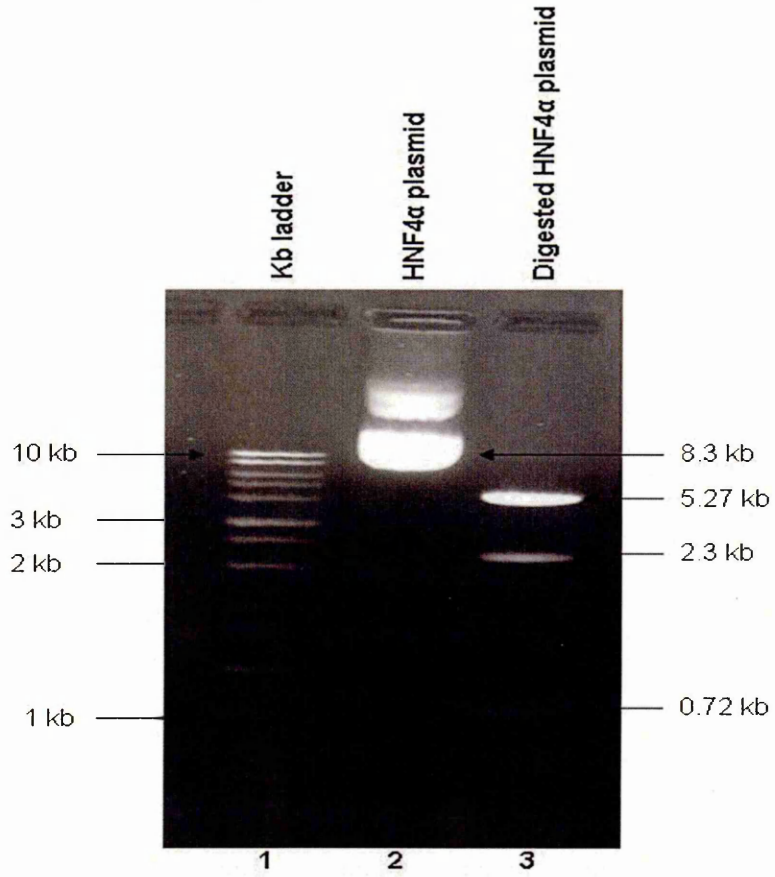


Fig 4.6. HNF4α encoded plasmid digested with restriction enzymes.
Lane 1 was a 1 Kb ladder, lane 2 the HNF4α plasmid and lane 3 the HNF4α plasmid digested with Sac1. The size of the HNF4α encoded plasmid before digestion was found to be approximately 8.3 kb (lane 2). Three bands of 5.72 kb, 2.3kb and 0.72 kb were detected as a result of the digested plasmid (lane 3).

4.3.3.3. RNA analyses

The purity of the RNA prepared from HepG2 cells before and after transfection with the expression plasmid encoding HNF4α was measured by the ratio of absorbance at 260 and 280 and found to be pure (>1.8). Moreover, no observable degradation was shown in the ribosomal bands 18S and 28S (Fig 4.7) upon assessment by gel electrophoresis.

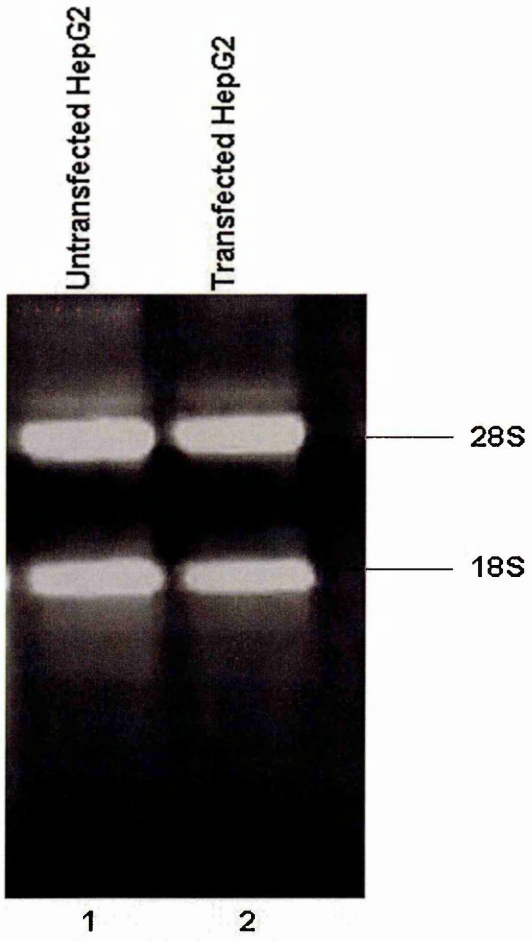


Fig 4.7. Ribosomal RNA 28S and 18S bands on 1% w/v agaros gel.
Lane 1 represented the 28S and 18S RNA bands from untransfected HepG2 cells. Lane 2 showed the RNA bands from the HNF4α transfected HepG2 cells. Both RNAs was shown to be with good quality.

4.3.3.4. Efficiency of primers

The Cq values were plotted against the log of template cDNA concentration and the slope of the line of best fit for both the reference gene (UBC) and the gene of interest (PZ) primer pairs were determined, both showinga high coefficient of determination (R^2>0.98), indicating the data points fitted very closely to the trend line (Fig 4.8). Primer efficiencies calculated were found to be 95% for UBC and 86% for PZ.

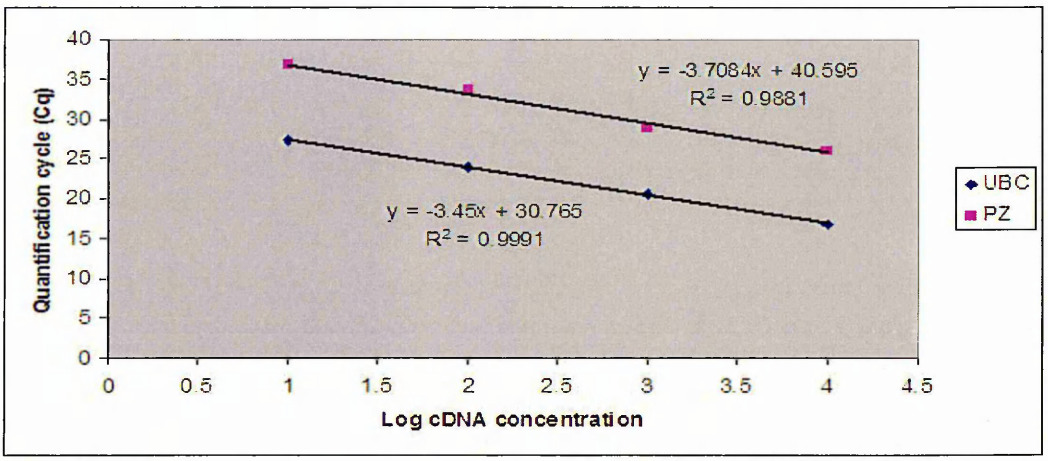


Fig 4.8. Determination of real time PCR efficiencies. The efficiencies of the reference gene (UBC) and the target gene (PZ) primers were measured by plotting Cq versus log cDNA for the serial dilutions. The slope was determined and used to calculate the efficiency of the primers. The efficiency of UBC was found to be 95% while PZ 86%.

4.3.3.5. Reference gene Validation

The UBC gene was found to be expressed at a constant level in untransfected and HNF4α transfected HepG2 cells with only a negligible change (2% decrease). However, the relative expression of GAPDH, Yahwaz and Bactin was found to be decreased in the transfected HepG2, 50%, 56% and 42%, respectively, compared to the untransfected HepG2 cells (Fig 4.9).

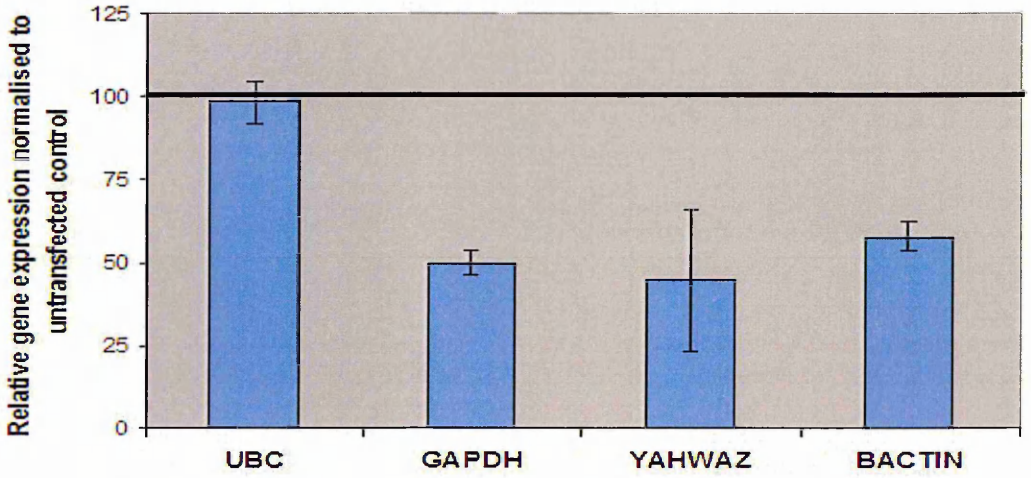


Fig 4.9. Comparision of relative expression of some housekeeping genes for the stability. Relative expression of housekeeping genes in HNF4α transfected HepG2 cells normalised to their expression in untransfected cells (100%). UBC gene relative expression was found to be the most stable in untransfected and HNF4α transfected cells. The relative expression of GAPDH, Yahwaz and Bactin was found to decrease 50%, 56% and 42%, respectively, in the transfected HepG2. The results shown are from four independent experiments.

4.3.3.6. Relative PZ mRNA expression in HNF4α overexpressed HepG2 cells.

The amplification plot (Fig 4.10) and Melt curve (Fig 2.4) showed neither the formation of non-specific products nor primer dimers, respectively.

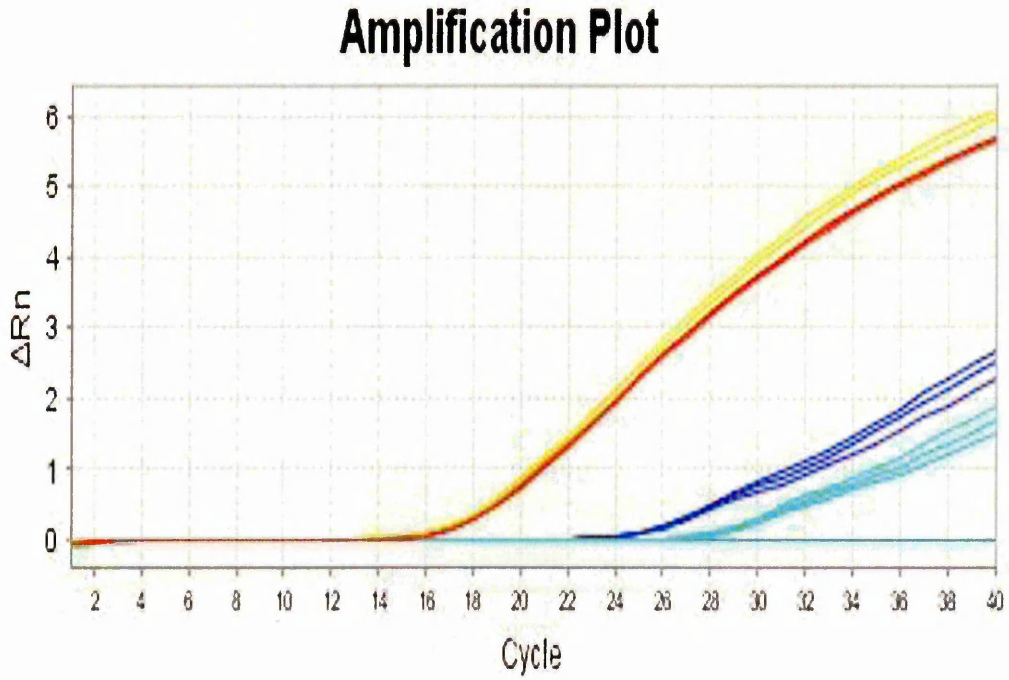


Fig 4.10. Amplification plot of PZ and UBC with transfected and non transfected cells. Green lines represent the UBC amplification with untransfected HepG2 cells while the red line with the transfected HepG2 cells. Dark blue colour represents the PZ amplification with the transfected sample while the light blue with the untransfected sample. The Cq values with PZ in the transfected samples (dark blue lines) decreased compared with non transfected (light blue lines). No amplification in no template wells indicates no non specific transcript or primer dimers. ΔRn, the magnitude of the signal generated, is the baseline signal from the initial stages of PCR subtracted from the real time data.

The level of endogenous PZ mRNA was found to be increased 2.5 fold after transfecting HepG2 cells with the expression vector encoding HNF4α (Fig 4.11). Statistical analysis of the data using the REST 2009 software showed that the increase was significant ($p < 0.05$, $n=6$). The relative expression of PZ mRNA after mock transfection of HepG2 cells (using an equivalent amount of the plasmid pCAT00) compared to untransfected cells (without transfection reagent or plasmid DNA), in the presence of the reference gene (UBC), was found to be unaffected (negligible increase 0.01 fold) (Fig 4.12).

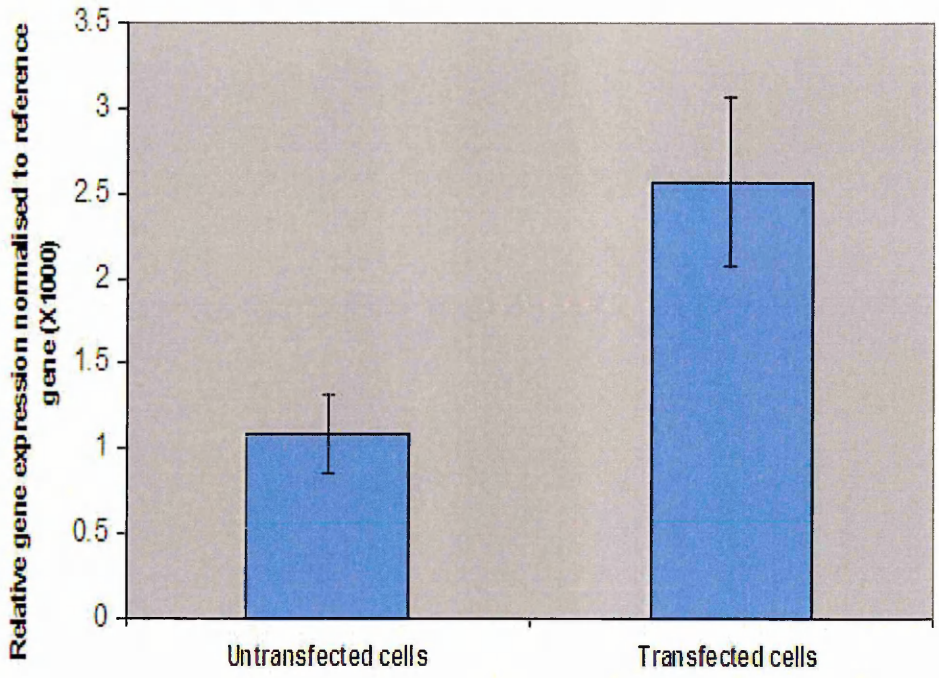


Fig 4.11. The effect of HNF4α overexpression on PZ mRNA relative expression. The HNF4α transfected cells showed 2.5 fold higher level of PZ mRNA expression (P value < 0.05. REST2009, n=6) compared with the untransfected (control) sample and normalised to UBC reference gene.

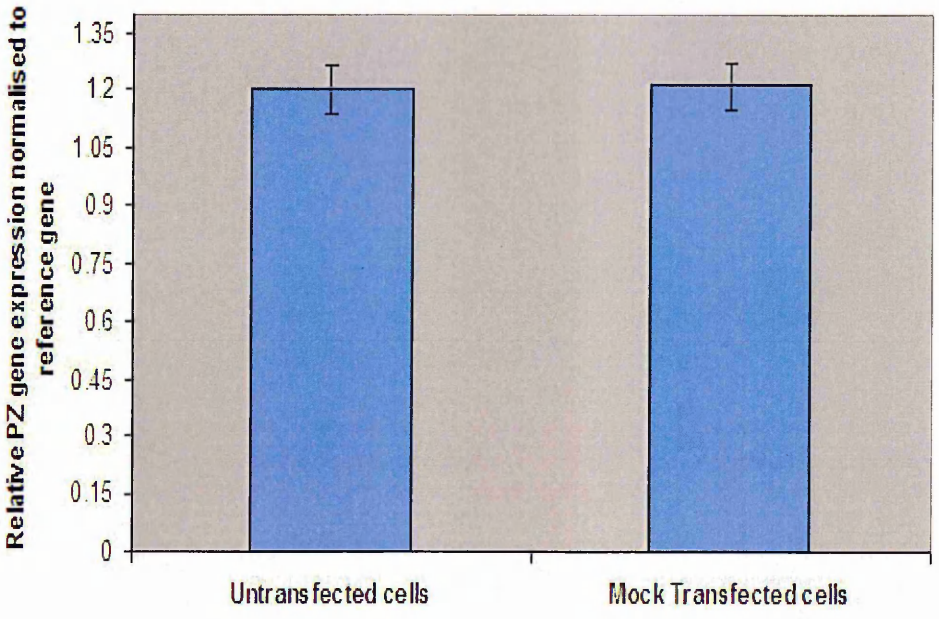


Fig 4.12. The effect of mock transfection on PZ mRNA relative expression. The relative expression of PZ in mock transfected (pCAT00) HepG2 cells showed negligible increase (0.01 fold change) compared to untransfected cells and normalised by UBC reference gene. The results were from four experiments.

4.3.3.7. Knockdown of HNF4α using small interfering RNAs

The positive control (ON-Targetplus GAPDH control pool) for GAPDH silencing resulted in an approximately 80% reduction (5-fold decrease) of GAPDH mRNA relative expression (Fig 4.13) compared to the negative control (ON-Targetplus Non-Targeting pool). The decrease in the GAPDH mRNA relative expression was found to be statistically significant ($P<0.05$, REST 2009, n=3)

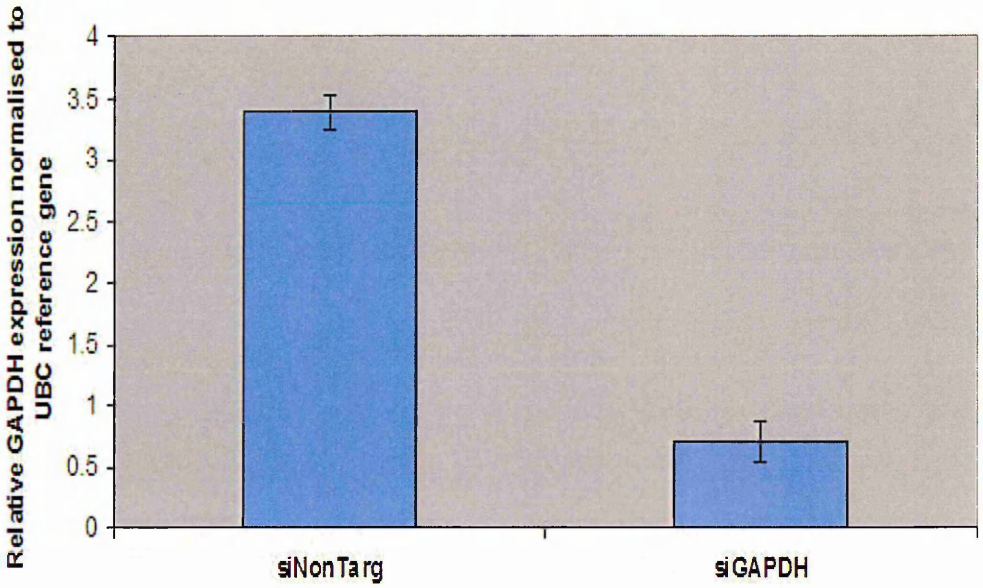


Fig 4.13. GAPDH knockdown in HepG2 cells using siRNAs.
The ON-Targetplus GAPDH control pool (siRNA) showed reduction of GAPDH mRNA to approximately 20% (5-fold decrease) compared with ON-Targetplus Non-Targeting pool (negative control) using qPCR. The reduction was statistically significant ($P<0.05$, REST 2009, n=3).

The HNF4α knockdown in HepG2 cells using an On-targetplus smart pool siRNA showed a 60% average reduction (2.5-fold decrease) of HNF4α mRNA relative expression compared to the ON-Targetplus Non-Targeting pool negative control (Fig 4.14). This reduction was found to be statistically significant using the REST 2009 software ($P< 0.05$, n=5).

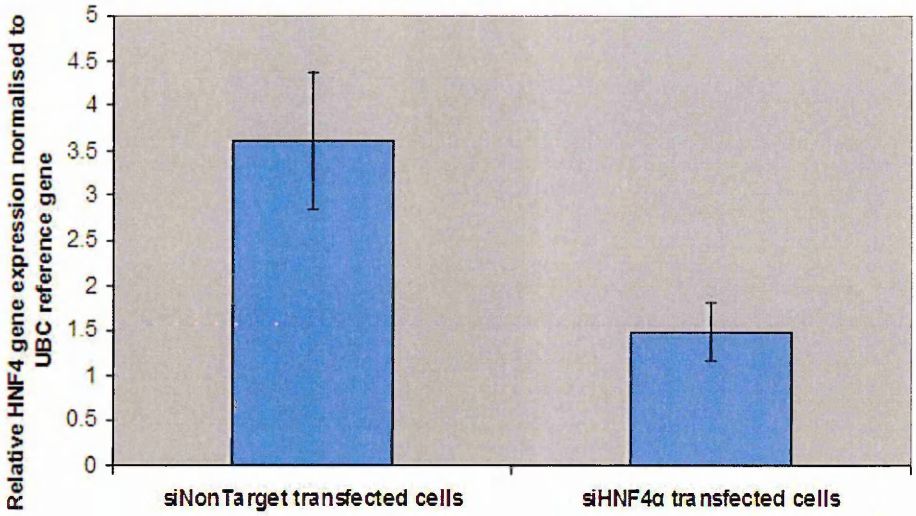


Fig 4.14. HNF4α knockdown in HepG2 cells using siRNAs.
HepG2 cells transfected with the On-targetplus smart pool siRNA for HNF4α (siHNF4α) showed a reduction in HNF4α mRNA expression to approximately 40% (~ 2.5 fold decrease) compared to the non targeting siRNA (siNonTarget) negative control ($P<0.05$, REST 2009, n=5) and normalised to UBC reference gene.

QPCR experiments using HNF4α mRNA knocked down HepG2 cells showed a significant effect on the level of PZ mRNA. Knockdown of HNF4α to 40% significantly reduced the relative expression of PZ mRNA (Fig 4.15) to approximately 36% (~ 2.8 fold decrease) (p value < 0.05, REST2009, n=6).

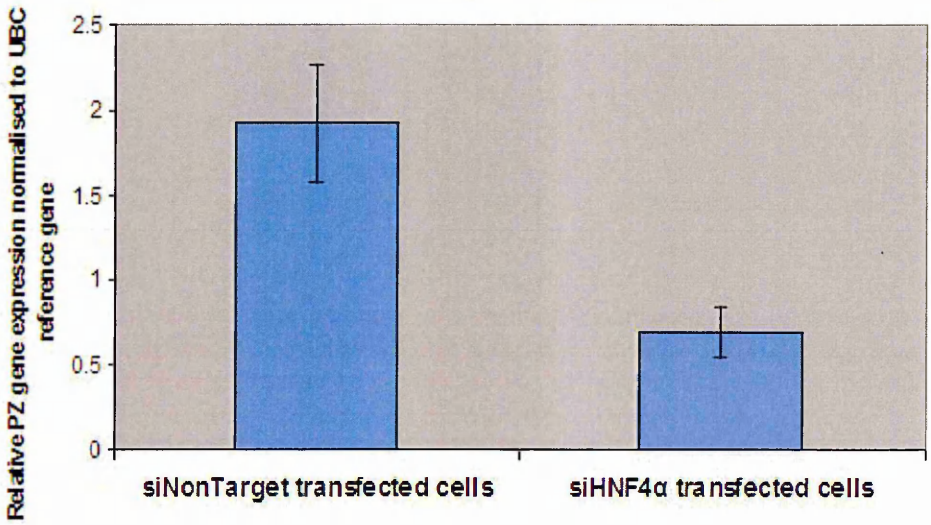


Fig 4.15. The effect of HNF4α Knockdown on PZ mRNA expression.
Knockdown of HNF4α in HepG2 cells decreased the PZ mRNA expression. 60% knockdown of HNF4α showed a decrease in the PZ mRNA expression to approximately 36% (P value< 0.05, REST 2009, n=6).

4.3.3.8. Expression of PZ and HNF4α in HepG2 compared to normal human liver

The relative expression of HNF4α mRNA in HepG2 was found to be approximately 2.3-fold more than in normal human liver tissue. The relative expression of PZ mRNA was found to be relatively expressed at very low levels in HepG2 cells, approximately 0.03-fold (~ 33-fold less) than that of the normal human liver tissue (Fig 4.16). The variation in the relative expression of PZ and HNF4α mRNA between HepG2 and normal human liver tissue were found to be statistically significant using the REST 2009 software ($P<0.05$, $n=3$).

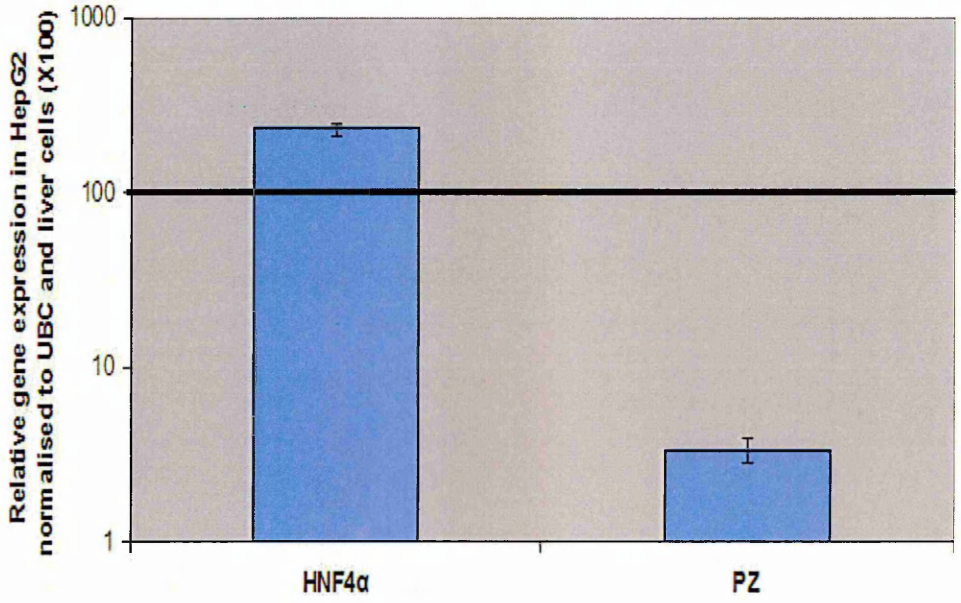


Fig 4.16. The relative expression of HNF4α and PZ mRNA in HepG2 and normal liver tissue. The figure showed a logarithmic scale for the relative expression of PZ and HNF4α in HepG2 cells compared to liver cells and normalised to UBC reference gene. The relative expression of HNF4α mRNA in HepG2 was found to approximately 2.3-fold more than in liver tissue. PZ mRNA expression was shown to be approximately 33-fold less in HepG2 than that of liver tissue. The horizontal thick line represented the levels of mRNA expression in liver tissue. The housekeeping gene UBC was used as a reference gene.

4.4. Discussion

The electrophoretic mobility shift assay is a gel electrophoresis technique which can be used to determine the ability of a protein to bind a linear double stranded DNA molecule, typically 20-25 base pairs and containing the binding site of interest (Jiang, Jarrett and Haskins 2009). This assay (EMSA) has been used extensively in DNA-protein interaction studies (Fried and Crothers 1981, Revzin 1989). Short sequence motifs which are found in the promoter and enhancer regions of genes and the transcription factors bound to those motifs are essential for the regulation of gene transcription. The *in silico* prediction of putative binding sites is of great importance to restrict the experiments and save the time used in protein-DNA interactions for characterisation of new genes (Messeguer *et al.* 2002).

As mentioned above two potential sequences for HNF4α binding (proximal -48 to -38 and distal -310 to -300) were identified using ALGGEN-PROMO. Probes were prepared that spanned the putative binding sites and also included flanking sequences (Section 2.2.3, Table 2.2). The EMSA probe spanning residues -54 to -35 of the PZ promoter, containing the potential binding site (-48 to -38) for HNF4α, was found to be bound by a protein. This protein showed binding specificity as it was dramatically competed with itself (60%) and the HNF4α binding site consensus sequence (100%), but showed negligible competition with non specific competitors such as the HNF3B consensus (7%), the AP1 consensus (19%) and the SP1 consensus (21%) sequences (Fig 3.2). The intensity of the protein-DNA complexes were quantified using the software associated with the Li-Cor odyssey imaging system (Li-Cor, UK). The specificity of the competition gave an indication that the protein bound to this site was indeed HNF4α.

To confirm HNF4α binding, a supershift binding assay was done using an antibody against HNF4α. The antibody added to the binding reaction caused a loss of the protein-DNA complex rather than a complex of increased size that would have a lower mobility upon electrophoresis. This could be due to the size of the complex being too large to efficiently migrate into the gel matrix (indicated by arrow C) where the well with the antibody showed the greatest fluorescence intensity (3.7-fold higher) compared with the other wells (Fig 4.4). To confirm this result and the electrophoretic properties of the HNF4α antibody-protein-

DNA complex, another EMSA reaction was done using the HNF4α antibody added to a binding reaction with the HNF4α binding site consensus sequence as a probe. The result showed that the DNA-protein-antibody complex using the HNF4α binding site consensus sequence was approximately the same as when the putative pPZHNF4α was used, with a 4-fold greater intensity in well with the antibody compared to the well with the binding reaction without antibody (Fig 4.5).

The binding of HNF4α to the PZ gene promoter was reported by Sugawara and colleagues (Sugawara *et al.* 2007) and, therefore, is consistent with the results in the current study. Moreover, evolutionarily related proteins with similar expression profiles, such as FVII, FIX and FX (Ichinose *et al.* 1990, Sejima *et al.* 1990) have also been found to be regulated by HNF4α (Stauffer *et al.* 1998, Reijnen *et al.* 1992, Miao *et al.* 1992).

The other potential HNF4α region identified (-310 to -300) by bioinformatic analysis was found to be bound by nuclear protein. This protein binding was found to be specific, since addition of various molar excesses of the non-specific competitor AP1 did not compete the bound protein. However, competition reactions with different molar excesses of the HNF4α binding site consensus also did not compete the bound protein, indicating that the protein bound to this region was not HNF4α (Fig 4.3, bands A).

Transient transfection is an important technique that can be used in the study of gene function (Miura and Yuan 2000) and has been used with hepatoma cell lines to analyse the promoters and enhancers of numerous genes in order to study their transcriptional regulation (Cereghini 1996). Therefore, to further analyse the role of HNF4α in PZ transcription, this transcription factor was overexpressed in HepG2 cells. Prior to the transfection, the HNF4α expression plasmid (pCMV6-XL4-HNF4α) was prepared (see section 2.4.3) and its identity was verified by digestion with SacI restriction enzyme and size-fractionated on a 1% w/v agarose gel (Section 2.4.3.3). Digestion of the expression plasmid with SacI showed the predicted fragments (Fig 3.6) of 5.27 kb, 2.3 kb and 0.72 kb according to the webcutter website (webcutter 2.0 website). Having verified the HNF4α encoding expression plasmid it was transfected transiently into HepG2 cells (See Section 2.4.4) and the effect on endogenous PZ mRNA levels measured using qPCR (See Section 2.3.4.3). The advantage of carrying

out experiments utilising the endogenous gene locus and qPCR rather than a reporter gene construct bearing a PZ promoter fragment is that the level of expression will be measured from the full endogenous gene promoter rather than a synthetic construct of a limited proportion of the promoter directing expression of the reporter gene. In addition, the DNA will be in the context of its normal chromatin structure wrapped around the histone octamers, rather than in a naked supercoiled plasmid. Also, synergistic interactions important for normal gene expression will occur within the intact nucleus when such interactions may be limited or absent in the reporter gene vector.

Preparation of good quality RNA is the first key step in the analysis of gene expression. Therefore, the RNA prepared from the HepG2 cells (See Section 2.3.1) before and after transfection was size fractionated using 1% w/v agarose gel. The prepared RNA showed no observable degradation of the ribosomal RNA bands 18S and 28S (Fig 4.7) and, therefore, was considered of good quality. Nevertheless, Auer et al in 2003 reported that degraded RNA samples could be used in gene expression studies, in this instance microarray analysis, if their integrity is comparable (Auer *et al.* 2003). The RNAs were also quantified and their purity was measured by the ratio of absorbance at 260 and 280 and found to be >1.8 using a NanoDrop®ND-1000. The RNAs from transfected and non-transfected HepG2 cells were then used to prepare cDNA, which used in qPCR.

A reference gene selected for gene expression studies should be expressed at a constant level and not be affected by experimental conditions. As it is difficult to find such a gene, the reference gene should be validated for expression and stability for every new transfection experiment (Schmittgen and Zakrajsek 2000). QPCR was performed on HNF4α transfected samples using four housekeeping genes (β-actin, GADPH, UBC, and YWHAZ). The housekeeping genes were analysed for amplification and stability after the transfection.

UBC was found to be the most appropriate reference gene for use in the calculation of expression ratios after HNF4α transfection by the SYBR green qPCR, as it was the gene whose expression was least affected by HNF4 overexpression (Fig 4.9).

Transfection of HepG2 cells by HNF4α was found to increase the relative expression of PZ mRNA (Fig 4.11). Although SYBR green I is normally used at

low concentrations which can lead to unreliable melt curves due to a dye distribution phenomenon (Varga and James 2006), the melt curve detected in this experiment was consistent with published data and showed no non-specific amplification products or primer dimers (Fig 2.4). The increase in the relative expression of PZ mRNA was detected by a decrease in the Cq value of PZ mRNA expression with transfected samples compared to the untransfected samples (control) in the presence of the reference gene (UBC). Since the primer efficiencies of the reference and target gene were not the same (Fig 4.8), the data were analysed using the Pfaffl equation (See Section 4.3.2.6). The increase of the relative PZ mRNA expression in the HNF4α overexpressed HepG2 cells was determined to be approximately 2.5-fold which is statistically significant ($p<0.05$, $n=6$).

An off-target effect is a potential biological response that could be caused by the introduction of foreign DNA and should be considered in studying any gene of interest by cell transfection (Jacobsen, Calvin and Lobenhofer 2009). This off-target effect could be due to the transfection reagent which is used in the transfection process. Tfx-20 is a highly efficient transfection reagent which has been used successfully for the transfection of HepG2 cells in studying gene transcription (Dongol *et al.* 2007). However, it was checked in this study for its potential effects on the transcription of the gene of interest. The ratio of PZ mRNA levels after mock transfection of HepG2 cells (using an equivalent amount of the plasmid pCAT00) to untransfected cells (without transfection reagent or plasmid DNA) was determined in the presence of the reference gene (UBC). The relative expression was found to be unaffected (a negligible increase of 0.01-fold) by the mock transfection (Fig 4.12), indicating the transfection reagent had no effect on the expression of the gene of interest.

To confirm the role of the HNF4α transactivational response in the transcriptional regulation of the endogenous PZ gene in HepG2 cells, the expression of the endogenous HNF4α was knocked down using an on-targetplus smart pool small interfering RNA (siRNA) and the effects on the endogenous PZ mRNA quantified using TagMan qPCR. Since UBC was found to be the most stable house keeping gene with HNF4α overexpression, it was used as reference gene assuming its stability with decreasing HNF4α levels. Off-target silencing has been observed with genes containing a limited

complement sequence with siRNAs (Jackson *et al.* 2003). Therefore, a positive control (ON-Targetplus GAPDH control pool) to optimise the protocol and transfection reagent with HepG2 cells and a negative control (ON-Targetplus Non-Targeting pool) to make sure the specific effects to the sequence targeting of the particular siRNA were used in the experiment. The smart pool technology was developed by Dharmacon Technologies for highly potent silencing and minimising off-target effects using reduced concentrations of four predesigned siRNA targeting one gene. On-targetplus smart pool siRNAs along with the positive GAPDH control pool and the negative Non-Targeting pool were used extensively in studying gene expression (Danes *et al.* 2008, Kwei *et al.* 2008, Hariparsad *et al.* 2009, Taniguchi *et al.* 2009, Rizzo *et al.* 2010, Lahkim Bennani-Belhaj *et al.* 2010)

Knock down of GAPDH mRNA expression level was first performed (Fig 4.13) to optimise the protocol and found to be reduced 80% ($P<0.05$, $n=3$) by the GAPDH positive control pool siRNA compared with the negative control, indicating an optimum protocol. Prior to determining the effect of HNF4α silencing on PZ the knockdown was confirmed using HNF4α primers. The level of HNF4α mRNA relative expression was found to be reduced to 60% ($P<0.05$, $n=5$) in the HepG2 cells transfected with the siRNA smart pool for HNF4α compared to the negative control (Fig 4.14). The HNF4α knocked down HepG2 cells were determined for their PZ relative expression compared to the negative control. The relative expression of PZ mRNA in HepG2 was found to be reduced to 36% (2.8-fold) in response to 60% HNF4α knock down compared to the negative control and normalised with the UBC reference gene (Fig 4.15). The decrease was statistically significant ($P<0.05$, $n=6$). These results were consistent with the HNF4α overexpression study and EMSA results in the current study in addition to the EMSA and reporter gene assays reported by Sugawara and colleagues (Sugawara *et al.* 2007). The current study, therefore, confirms the Sugawara data in that HNF4α has a significant role to play in expression of the PZ gene. However, the use of overexpression and knockdown of HNF4α, and analysis of the endogenous PZ locus adds a further level of characterisation to the promoter study, in that the importance of HNF4α activity is confirmed within the chromatin context of the intact nucleus.

The HepG2 cell line is a common model system for analysis of liver-specific gene expression. It was used to study the effect of the HNF4α in the transcriptional regulation of many coagulation proteins, such as, FVII (Stauffer *et al.* 1998), FVIII (Figueiredo and Brownlee 1995), FIX (Reijnen *et al.* 1992), FX (Miao *et al.* 1992), FXI (Tarumi *et al.* 2002) and FXII (Farsetti *et al.* 1998). However, as HepG2 is a hepatoma cell line, its genome are inevitably different to normal primary liver cells and the transcriptome may exhibit differences to normal liver cells that could impact on such studies. Therefore, it is important to assess whether or not the HepG2 cell line expresses PZ and HNF4α, and if so at what levels. Expression of HNF4α mRNA and PZ mRNA in the HepG2 cell line were, therefore, compared to their expression in normal liver cells using qPCR. Total RNA prepared from HepG2 cells and RNA from normal human liver cells (AMS biotechnology Ltd, UK) were reverse transcribed to cDNA and used to determine the relative expression for PZ and HNF4α by qPCR. The relative expression level of HNF4α in HepG2 was found to be 2.3-fold more than in liver tissue while PZ mRNA relative expression was shown to be approximately 33-fold less in HepG2 than that of liver tissue (Fig 4.16). The difference in the relative expression in HNF4α and PZ between HepG2 and liver tissue was found statistically significant ($P<0.05$, $n=3$). The high level of HNF4α and significantly low level of PZ in HepG2 compared to liver tissue may indicate that these cells were not the best model to study PZ transcriptional regulation and may affect the accuracy of the results. However, the RNA from the liver tissue that used in the comparison to HepG2 was prepared from only one male, and therefore, these differences may be not reliable.

4.5. Conclusion

The present study has defined an up-regulatory role for HNF4α in the regulation of the endogenous PZ gene within HepG2 cells. This was initially indicated by its specific binding to the proposed binding site in the PZ promoter spanning nucleotides -48 to -38 which was confirmed by supershift assay. The importance of HNF4α was then confirmed by its effect on the relative expression of PZ mRNA within the chromatin context in the HepG2 cell model using overexpression and knockdown techniques and measured by qPCR. The

effect was statistically significant ($P < 0.05$) in 6 independent experiments with each technique. These results were consistent with the study reported by Sugawara in 2007 (Sugawara *et al.* 2007). However, although these significant results were obtained within the chromatin context, they are still considered as *in vitro* and need to be confirmed *in vivo*.

Chapter 5-The role of transcription factors PEA3, Ets 1 and CREB in regulating protein Z expression

5.1. Introduction

The majority of blood coagulation proteins usually reach their adult levels by the age of six months (Andrew *et al.* 1987); with further increases in levels of a number of key factors and also their activation with subsequent increasing age (Hager *et al.* 1989); with analysis of centenarians showing increased levels of coagulation enzymes and activation peptides as well as enzyme-inhibitor complexes indicating activation of coagulation (Mari *et al.* 1995). Levels of the major procoagulants and anticoagulants, which were assessed by Mari *et al*, are shown in Table 5.1.

Measurments	Young controls	Old controls	Centenarians
FVIIa (ng/mL)	3.18 (2.76-3.64)	3.24 (2.86-3.67)	4.46 (3.91-5.06)
FIX activation peptide (pmol/L)	164 (144-185)	207 (187-229)	391 (351-435)
FX activation peptide (pmol/L)	86 (81-92)	99 (85-115)	248 (226-272)
Prothrombin fragment 1+2 (nmol/L)	0.56 (0.49-0.64)	0.69 (0.61-0.78)	1.15 (0.93-1.43)
Thrombin/antithrombin complex (ng/mL)	2.0 (1.6-2.4)	2.4 (2.0-2.8)	5.9 (4.6-7.6)
Fibrinopeptide A (nmol/L)	0.89 (0.74-1.06)	1.22 (1.0-1.50)	1.99 (1.72-2.31)
Fibrinogen (mg/dL)	263 (244-284)	312 (287-338)	347 (313-385)
F VIII (%)	97 (90-105)	134 (119-151)	165 (147-185)
Prothrombin (%)	104 (99-108)	98 (95-102)	88 (82-93)
F VII (%)	99 (93-105)	110 (104-116)	111 (103-119)
F IX (%)	96 (93-100)	104 (95-113)	107 (98-117)
Antithrombin (%)	98 (95-101)	97 (91-104)	91 (85-98)
Protein C (%)	87 (82-92)	88 (83-93)	90 (83-98)
Tissue factor pathway inhibitor (%)	95 (89-102)	110 (102-118)	109 (100-118)
Protein S (%)	106 (98-113)	115 (100-132)	109 (97-123)

Table 5.1. Haemostasis measurements for young, older and centenarians (Mari *et al.* 1995). Younger controls are 18-50 year old healthy individuals. Older controls are 51-69 year old healthy individuals. Centenarians are 100-102 year old healthy individuals. All the groups were including both males and males.

The increase in blood coagulation activity as people age may be due, in part at least, to the greater increase in expression of procoagulants of the coagulation system compared to anticoagulants, which in turn could lead to increased risk of thrombosis (Lowe *et al.* 1997). However, the risk of thrombosis does not

necessarily increase in response to increased activity of coagulation enzymes (Mari, Coppola and Provenzano 2008). The age-related imbalance between procoagulants and anticoagulants and the potential contribution to increasing thrombosis cases was also discussed by Massimo Franchini (Franchini 2006).

To understand why coagulation activity increases with age the genetic mechanisms responsible for age regulation of one of the key procoagulants, human factor IX (hFIX), was studied using transgenic mice (Kurachi *et al.* 1999, Kurachi *et al.* 2000, Kurachi and Kurachi 2000). The study of factor IX (FIX) using transgenic mice with a hFIX minigene vector provided evidence for two age-regulatory elements, AE5' and AE3'. AE5' is located in the FIX promoter region and was shown to be responsible for age-stable plasma levels, while AE3' is located in the 3' untranslated region (UTR) and is thought to be responsible for the increase in expression of FIX in the liver with increasing age (Kurachi *et al.* 1999). Both age-regulatory elements are important for the normal age-dependent regulation of FIX gene (Kurachi *et al.* 2000).

The minigene study using mice with a vector spanning the 5' end (-802 to -770) of hFIX promoter where the AE5' is located, showed age stable levels of plasma hFIX through the entire lifespan of the mice which correlated with stable levels of mRNA (Kurachi *et al.* 1999). These results defined the location of the sequence responsible for the age-stable expression of the hFIX gene, which was then supported by the presence of a DNA-protein complex detected within this region (-790/-784) using DNase I footprint analysis and electrophoretic mobility shift assays (EMSA) (Kurachi *et al.* 1999). The sequence spanning the AE5' (-790/-784) was found to have homology with the binding site consensus sequence ((G/C)AGGA(A/T)G) of the Ets family protein, polyomavirus enhancer activator 3 (PEA3). The consensus binding site of PEA3 competed for the protein bound to the AE5' region and gave an indication that the protein binding to the age-stable element was in fact PEA3 (Kurachi *et al.* 1999). Subsequent studies demonstrated an age-stable element with homology to the PEA3 motif within the human protein C (hPC) gene promoter. Minigene analysis of transgenic mice defined this element between -849 and -803 and determined that it was responsible for age-stable expression (Zhang, Kurachi and Kurachi 2002). The protein bound to the region spanning the age-stable element in the hPC gene had competition with the AE5' sequence of the

hFIX gene using EMSA, suggesting the same protein binds to both promoter sequences. However, Kurachi and his colleaques in 2009 reported that the age stable element (AE5' also referred as ASE) was bound by Ets1 and had an essential role in the spontaneous recovery of haemophilia B Leyden (Kurachi *et al.* 2009). Due to the close homology and suggested evolutionary relationship between PZ and PC and FIX (Ichinose *et al.* 1990), there is the possibility that this conservation extends to the regulatory mechanisms of expression including Ets family binding sites, such as those for Ets1 or PEA3. PEA3 and Ets1 are expressed in a number of different tissues, including the liver (Yanai *et al.* 2005) and may have a role in the constitutive expression of PZ.

5.1.1. E twenty six transformation-specific (Ets) family proteins

The E twenty six transformation-specific (Ets) gene was originally found to be transduced by the E twenty six (E26) avian leukaemia virus. Cellular sequences and transcripts in human and other organisms were subsequently identified by their homology to the Ets region of the E26 retrovirus (Watson, Ascione and Papas 1990), thus, defining the Ets family.

The proteins of the Ets family are transacting phosphoproteins which were found to play essential roles in the epithelial, endothelial, haemopoietic, neuronal and endocrine systems; particularly, in cell differentiation, proliferation, migration and oncogenesis (Seth and Watson 2005, Sharrocks 2001, Wasylyk, Hagman and Gutierrez-Hartmann 1998, Feldman, Sementchenko and Watson 2003, Gutierrez-Hartmann, Duval and Bradford 2007). The Ets family was classified as a winged helix-turn-helix family of DNA-binding domains by the analysis of the structure using nuclear magnetic resonance spectroscopy (NMR) (Shore *et al.* 1995, Werner *et al.* 1995, Donaldson *et al.* 1996). The distinctive characteristic of the Ets family is the presence of the Ets DNA-binding domain which is composed of approximately eighty five amino acids (Karim *et al.* 1990) and characterized by the presence of three conserved tryptophans separated by eighteen residues (Shore *et al.* 1995). Apart from the GA-binding protein (GABP)α, all Ets family proteins bind their DNA-binding site in the promoter of the target gene as monomers (Gutierrez-Hartmann, Duval and Bradford 2007, Graves and Petersen 1998). The core binding consensus for Ets family members was recognized as GGAA/T, with the interaction of each

protein in this family involving the bases flanking this core (Xin *et al.* 1992). Binding sites have been found in many genes such as those encoding transcription factors, regulators of apoptosis and cell cycle, protineases, transforming and tumor-associated products, ligands, receptors and tissue specific products (Sementchenko and Watson 2000). The bases adjacent to the core sequence may contain binding elements for other transcription factors which could interact synergistically with the Ets protein in order to activate or repress target genes (Li, Pei and Watson 2000). The Ets domain and the conserved arginine residues of the DNA recognition helix α3 have been shown to mediate the interaction with other transcriptional proteins, while the optimal contacts which result in a specific transcriptional response is set up by other specific domains (Verger and Duterque-Coquillaud 2002).

The GGA core is recognized by the α3 recognition helix which, mediated by two conserved arginine residues, embeds into the major groove (Fig 5.1), while the β3- β4 hairpin (wing) and α2 contacts the DNA phosphate backbone in the minor grooves (Mo *et al.* 2000, Mo *et al.* 1998).

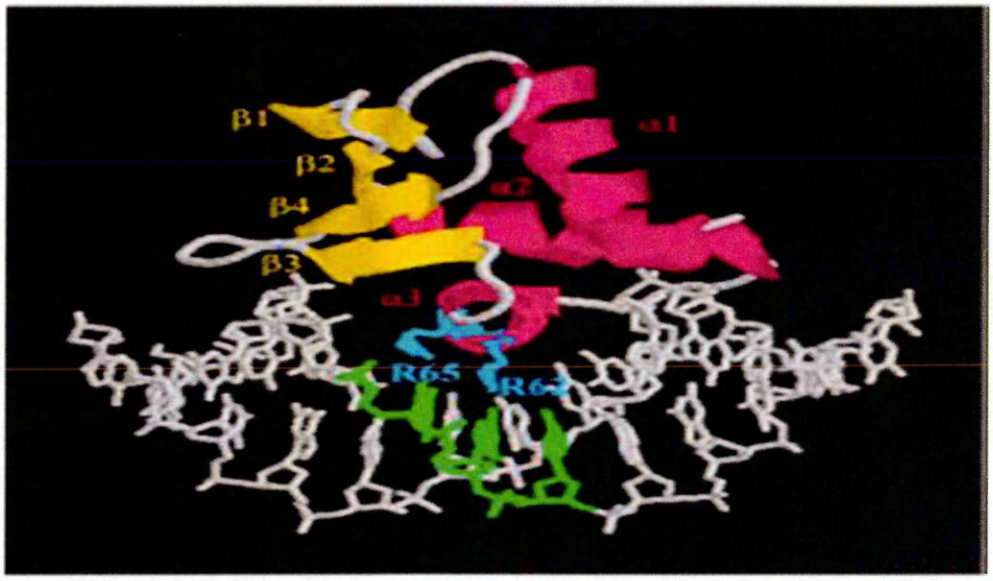


Fig 5.1. Ribbon diagram shows the Ets family member Elk-1 binding to DNA (Verger and Duterque-Coquillaud 2002). The DNA is white, the GGA core is green, the α-helices are pink and the β-strands are yellow. The two arginines which mediate the α3 recognition of the GGA core are greenish blue (R62 and R65).

Many proteins of this family are posttranslationally modified by signal transduction cascades, causing changes in their binding to DNA, transcriptional activities and association with cellular partners (Sharrocks 2001).

The 27 genes in humans and the 26 in mice that compose the Ets family can be subdivided into 11 subfamilies based on the structure composition; PEA3, ETS, TEL, ERG, TCF, ERF, ELG, ELF, ESE, SPI, and PDEF (Oikawa and Yamada 2003, Galang *et al.* 2004, Hollenhorst, Jones and Graves 2004). Although the majority of the Ets family members contain the Ets DNA-binding domain in their carboxyl-terminus many, such as the ternary complex factor (TCF) subfamily, were found to have their Ets domain in the amino-terminus. Furthermore, many members of the family have shown the presence of another evolutionarily-conserved domain in addition to the Ets DNA-binding domain. This domain is called the Pointed domain and has a role in protein-protein interactions by the presence of a helix-turn-helix structure (Kim *et al.* 2001). Some members of the Ets family, such as the Ets-related gene (ERG) subfamily, were discovered to have two activation domains, one in the amino-terminus and the other in the carboxyl terminus (Rao *et al.* 1993).

The DNA binding sequences which are bound by proteins of the Ets family are similar since the Ets binding domain of this family is highly conserved (Graves and Petersen 1998). Therefore, to ensure correct binding to the DNA site, many of the Ets family proteins contain auto-inhibitory domains providing increased specificity (Graves *et al.* 1998). The stressed conformation of the Ets domain, maintained by the interaction of the two inhibitory domains, is disrupted upon DNA binding. This disruption leads to changes in the N-terminal inhibitory region and a transient relaxed conformation of the Ets domain, stabilising the contact with the DNA (Jonsen *et al.* 1996). All proteins of the Ets family contain two inhibitory regions flanking the Ets domain responsible for their auto-inhibition (Gutierrez-Hartmann, Duval and Bradford 2007). Although the majority of Ets family proteins have been shown as transcriptional activators, some proteins such as Ets2 repressor factor (ERF), Ets variant 6 gene (ETV6, also called TEL), NET (a member of ternary complex factor) and YAN (the invertebrate ortholog of TEL, which appears to be regulated by MAPKs) have revealed activity as repressors (Mavrothalassitis and Ghysdael 2000). Several proteins of this family were also suggested to be able to activate and repress

transcription, depending on the cellular context, as they possess both activation and repression domains (Sharrocks 2001, Bojovic and Hassell 2001).

The PEA3 transcription factor is a member of the PEA3 group of the Ets family. The PEA3 group of transcription factors consists of PEA3 (also called ETV4 and E1AF) (Xin *et al.* 1992, Higashino *et al.* 1993), Ets related 81(ER81, also called ETV1) (Brown and McKnight 1992) and Ets Related Molecule PEA3-like (ERM, also called ETV5) (Monte *et al.* 1994). Members of the group show a structural similarity with 95% homology in the Ets domain, 85% in the acidic domain and 50% in the carboxyl-terminal tail. Their expression has been detected in many normal as well as cancerous tissues (de Launoit *et al.* 1997). The Ets domain of the PEA3 transcription factor spans residues 334 to 417 with deletions in this domain showing abrogation in the transcriptional activity of the protein (Bojovic and Hassell 2001). The PEA3 protein contains a strong activation domain spanning residues 42 to 85 and two negative regulatory regions flanking the activation domain that are located between residues 1 and 42 and residues 85 and 100, respectively, suggesting both negative and positive regulatory mechanisms (Bojovic and Hassell 2001). EMSA analysis has indicated that the DNA-binding consensus sequence for the PEA3 group members is GGAA/T (Xin *et al.* 1992, Brown and McKnight 1992, Monte *et al.* 1994, Laget *et al.* 1996). The PEA3 transcription factor messenger RNA (mRNA) in mice was found to be expressed in brain, epididymis and mammary glands (Xin *et al.* 1992). The expression of PEA3 was found in many human cancer cells. Higashino *et al.* detected the PEA3 mRNA in Hela cells, a cell line derived from the cervical cancer (Higashino *et al.* 1993). This transcription factor was also detected in carcinoma cells and stromal cells, suggesting its role in the progression of the tumor in patients with ovarian carcinoma (Davidson *et al.* 2003). PEA3 overexpression was detected in breast cancer and most HER2/Neu-mediated mammary tumors (O'Hagan and Hassell 1998, Shepherd *et al.* 2001), *although* Xing *et al.* (2000) reported that the expression of PEA3 might lead to the inhibition of HER2/Neu-mediated oncogenesis, which is contradictory to this observation (Xing *et al.* 2000). However, Shepherd *et al.* (2001) found that PEA3 is involved in mammary oncogenesis and the abolition of this transcription factor could suppress the oncogenesis (Shepherd *et al.* 2001). Other research demonstrated that there was no correlation between the

expression of PEA3 and HER2/Neu in the tumour tissues of breast cancer (Xia *et al.* 2006). The oncogenic role of PEA3 was also shown by its upregulation of multiple matrix metalloproteinase genes and induction of invasive and metastatic activities (Shindoh, Higashino and Kohgo 2004).

Ets1 is another protein of the Ets family. This protein was found to be overexpressed in tissues with pathological conditions requiring new formation of blood vessels and reported to have an important role in embryonic development, viral transformation, haemopoietic cell differentiation and regulation of invasive action of many tumour and normal cells (Dittmer 2003). The Ets1 DNA-binding domain was found to span 80 amino acids of the Ets domain and binds DNA as a monomer (Nye *et al.* 1992). The Ets domain of the Ets1 spans from 331 to 415 residues (Dittmer 2003). The full length Ets1 (p51) was shown to have an autoinhibition counteraction increasing its binding activity to a palindromic Ets binding site in the stromelysin-1 promoter to form an Ets1-DNA-Ets1 ternary complex (Baillat *et al.* 2002). The Ets1 isoform p42 was not able to form the same complex, implying the importance of the exon VII region in the autoinhibition counteraction.

5.1.2. The role of Ets family in the coagulation process

There is some evidence to indicate a role for the Ets family of transcription factors in the regulation of expression of certain coagulation factors. For instance, as mentioned earlier, the binding of Ets1 to the promoters of hFIX and the potential binding of the same protein to the hPC promoter was indicated by minigene constructs and EMSA competition assays (Zhang, Kurachi and Kurachi 2002, Kurachi *et al.* 2009). Another protein from the Ets family, Fli1, a member of the ERG subfamily, may also have an indirect effect on the coagulation process since its malfunction might lead to abnormal platelet production (Spyropoulos *et al.* 2000). The promoter activity of von Willebrand factor (vWF), the carrier and activity stabilizer for the coagulation factor VIII (Ruggeri and Ware 1993), was reported to be increased by the overexpression of Ets1 and Erg in HeLa cells (Schwachtgen *et al.* 1997). In line with this, bioinformatic analysis of the human protein Z gene promoter revealed the presence of two regions (Fig 5.2) showing close homology to the age-stable elements found in hFIX and hPC, two evolutionarily related proteins. The two

regions were termed proximal (spanning nucleotides -114 to -108 (CAGGAGG)) and distal (spanning nucleotides -816 to -810 (AAGGATC)). Two further sequences showing close homology to the PEA3 binding consensus were also identified by bioinformatic analysis. These upstream sites (Fig 4.2) span nucleotides -1190 to -1184 (CTTCCTG) and -1219 to -1213 (CTTCCTG).

5.1.3. The cAMP response element-binding protein (CREB)

The cAMP (cyclic adenosine-3',5'-monophosphate) response element-binding protein (CREB) is a member of the basic leucine zipper family of transcription factors (Smolik, Rose and Goodman 1992, Shaywitz and Greenberg 1999, Hoeffler *et al.* 1988) which binds to the cAMP response element (TGACGTCA). The factor binds as a dimer in order to regulate specific gene expression (Carlezon, Duman and Nestler 2005). CREB and other related proteins, such as activating transcription factor 1 (ATF-1) and cAMP response modulator (CREM), which form a subgroup of this family, contain a group of residues forming a basic region important for DNA binding followed by a zipper region of seven leucine repeats for homo- and hetero-dimerisation with other family members (Meyer and Habener 1993, Montminy 1997). The expression of CREB and ATF-1 was shown to be ubiquitous while CREM was limited to neuroendocrine tissues (Meyer and Habener 1993). The structure of CREB shows a conserved arrangement with a glutamine-rich domain at the N-terminus, a domain containing phosphorylation consensus sites for kinases known as the kinase-inducible domain followed by a region rich with glutamine residues responsible the basal activity and finally the basic leucine zipper domain at the C-terminus (Sands and Palmer 2008). The initiation of gene transcription by cAMP was shown to be activated by protein kinase A (PKA) phosphorylation of CREB at Ser-133 (Gonzalez and Montminy 1989). This phosphorylation, in turn, activates the interaction of CREB with the CREB binding protein in order to induce the transcription of the target gene (Brindle, Nakajima and Montminy 1995). CREB was also reported to be phosphorylated in response to adenosine 5'-monophosphate-activated protein kinase (AMPK) in order to regulate expression of genes that possess cAMP response elements in their promoters (Li *et al.* 2003, Thomson *et al.* 2008). Bittinger et al (2004) demonstrated a role for CREB coactivators, transducers of regulated CREB

activity (TORC) which activate CREB independently of Ser 133 phosphorylation, in the activation of CRE-mediated gene expression (Bittinger *et al.* 2004). They found that the nuclear translocation of these coactivators and PKA simultaneously is essential for the activation of CREB by cAMP. Shi *et al.* (2006) suggested a novel PKA-independent signal transduction pathway, with a cyclic AMP-dependent Epac-Rit signaling pathway being implicated in CREB activation (Shi, Rehmann and Andres 2006). CREB was also reported to be phosphorylated in response to other stimuli such as Ca2+/calmodulin-dependent kinase IV (Enslen *et al.* 1994) and mitogen/stress-activated kinases Msk1and Msk2 (Wiggin *et al.* 2002). CREB was found to upregulate pro-apoptotic proteins such as Puma and Bak and reduce survival of keratinocytes after CREB phosphorylation and activation via the extracellular-signal regulated kinase (ERK) pathway (Versteeg *et al.* 2008).

Although CREB has been reported to regulate many proteins, its regulatory role has not been associated, so far, with the coagulation factors. However, a sequence showing a high degree of homology to the cAMP response element, and overlapping a proposed PEA3 or Ets1 element, has been identified in the PZ gene promoter between -120 and -113 (Fig 4.2) by bioinformatic analysis.

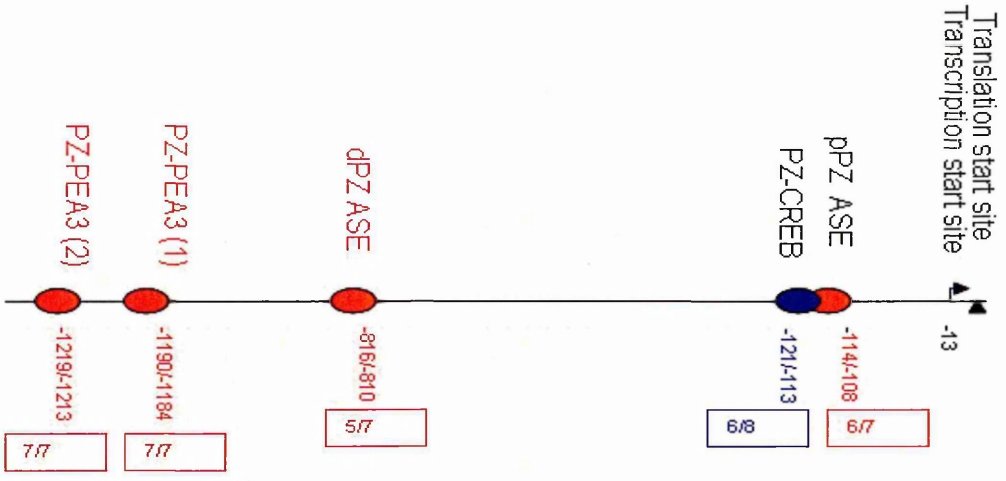


Fig 5.2. The hypothetical PEA3 and CREB binding sites in the PZ gene promoter. Sites are numbered relative to the translation start site and the numbers in boxes show the homology with the respective consensus sequence (Santa Cruz, Germany).

5.2. Aims

The aim of the present study was to characterise the role of the hypothetical age-stable elements identified in the PZ gene promoter *in silico* in regulating the transcription of PZ in a cell model system (HepG2 cells). Initial experiments looked to elucidate, by EMSA, whether nuclear proteins from a human liver-derived cell line bind to the putative PZ ASEs. Protein-DNA binding studies were also be used to test the hypothesis that the bound proteins were the Ets family members PEA3 and/or Ets1, which have previously been shown to bind to the hfIX ASE and potentially the PC ASE.. Also, to support any *in vitro* data,the role of these proteins was analysed in cell, in the context of the intact chromatin; with gene expression analysis of the endogenous PZ gene in the human liver-derived cell line HepG2.

5.3. Methods

5.3.1. Protein-DNA binding analysis

5.3.1.1. Electrophoretic Mobility Shift Assay (EMSA)

The DNA-protein binding reaction for EMSA was conducted in a 20 µl reaction volume (See Section 2.2.4.1), with the prepared (See Section 2.2.1) HepG2 nuclear extract (HepG2 NE) with 10 fmole of the biotinylated probe flanking the hypothetical proximal (pPZASE -114/-108) or distal (dPZASE -816/-810) age-stable element (See Table 2.3, Section 2.2.3). The binding mixture was incubated for 30 minutes at room temperature, size-fractionated through 5% w/v, 19:1 acrylamide/bis-acrylamide (See Section 2.2.4.1) mini (8 x 8 x 0.1 cm) gel, using a voltage of ~ 12.5 V/cm, in 1x Tris-Borate EDTA (TBE) buffer and binding complexes were detected (SeeSection 3.3.4) using the UVP BioImaging system (UVP, UK). Competitive EMSAs were performed by incubating the indicated fold molar excess of unlabelled competitors for 10 minutes with the Hep NE at room temperature prior to further incubation with the labelled probe for 30 minutes.

The same was done with IRDye probes using 100 fmole of the proposed probe and 12 µg of the HepG2 NE. The protein-DNA complexes were detected (SeeSection 3.3.5) using the Li-Cor Odyssey imaging system (LI-COR Biosciences, UK).

5.3.1.2. Supershift assays

Supershift assays were used to identify the protein binding specifically to the proposed probe. To perform the assay 2 µg of the candidate antibody (Santa Cruz, Germany) were pre-incubated with 4 µg of HepG2 NE for 10 minutes at room temperature followed by further incubation with 10 fmole of biotinylated probe for approximately 30 minutes. The binding mixture was size fractionated using a 5% polyacrylamide gel (See Section 2.2.4.1) and visualised by the UVP BioImaging system (SeeSection 3.3.4). The assay using IRDye probes was performed using 12 µg of the HepG2 NE and 100 fmole of the candidate probe and visualised by the Li-Cor Odyssey imaging system (SeeSection 3.3.5). The

incubation of the antibody with the HepG2 NE was performed on ice and at room temperature.

5.3.2. Gene expression studies of the endogenous PZ gene in HepG2 cells

5.3.2.1. Comparison of PZ gene promoter sequences in HepG2 cells with the wild type promoter.

The comparison of the wild type PZ gene promoter sequence to that present in the HepG2 cells has been described previously (See Section 2.1 and 4.3.2.1).

5.3.2.2. Expression plasmid verification

The expression vector encoding PEA3 was prepared (see Section 2.4.3) and verified using restriction enzyme mapping (Fermentas life science, UK). The appropriate restriction enzyme and its digestion sites in the PEA3 insert (2.2 kb) were determined using webcutter 2.0. The plasmid pCMV6-XL5-PEA3 (OriGene Technologies, USA), 4.5 kb, was digested (See Section 2.4.3.3) with Sac1. The expected size of the bands was confirmed by electrophoresis on a 1% w/v agarose gel for the verification of the expression plasmid to be used in the transactivation.

5.3.2.3. Extraction of RNA

RNA was extracted from untransfected and HepG2 cells transfected with the PEA3 expression vector using the GenElute mammalian total RNA miniprep kit (See Section 2.3.1). The quality of the extracted RNA was determined by a 1% w/v agarose gel and analysis of the ribosomal 18S and 28S bands. The purity and concentration of the RNA samples were measured using the NanoDrop®ND-1000 spectrophotometer (NanoDrop Technologies, USA). The RNA was treated with DNase 1 (Sigma-Aldrich, UK) to remove any genomic DNA. The RNA was then used to synthesise cDNA (See Section 2.3.3) using the iScript cDNA synthesis kit (Bio-Rad, UK) for subsequent use in qPCR.

5.3.2.4. Primer efficiency

Primer efficiencies were determined, as described previously (See Sections 2.3.4.1 and 4.3.2.4), using the efficiency equation Efficiency = $-1 + 10^{(-1/slope)}$. However, the efficiency of primers used in the TaqMan qPCR were considered as 100% according to the manufacturer (Applied Biosystems, UK).

5.3.2.5. The effect of overexpression of PEA3 on the transcription levels of the endogenous PZ gene in HepG2 cells

The Tfx™-20 transfection reagent (Promega corporation, UK) was used for the overexpression of PEA3 in HepG2 cells by transfection of a PEA3 expression plasmid (Fig 2.7) according to the manufacturer's instructions (See Section 2.4.4). The effect of PEA3 overexpression in HepG2 cells on the endogenous PZ gene was quantified by qPCR using SYBR green chemistry (See Section 2.3.4.3).

The relative expression ratio of the PZ gene was normalised using the reference UBC gene and calculated using the Pfaffl equation (Pfaffl 2001) (See Section 4.3.2.6). Melt curve (See Section 2.3.4.3) analysis was carried out to detect non-specific amplification products and primer dimers.

5.3.2.6. The effect of PEA3 knock down on the expression of the endogenous PZ gene in HepG2 cells

The target-specific smart pool siRNA (Dharmacon RNAi Technologies, UK) was used to knock down PEA3 according to the manufacturer's instructions (See Section 2.4.5 and 4.3.2.7). The effect of PEA3 knock down in HepG2 cells on PZ expression and the knock down efficiency were quantified by qPCR (See Section 2.3.4.3).

5.3.2.7. Comparison of the relative PEA3 mRNA expression in HepG2 to normal liver tissue.

To compare the relative expression of PEA3 mRNA in HepG2 to normal liver tissue, qPCR experiments were performed. Prior to the qPCR the RNA was prepared (See Section 2.3.1) from HepG2 cells. The HepG2 total RNA and the

commercial human liver tissue total RNA (AMS biotechnology Ltd, UK) were reverse transcribed to cDNA (See Section 2.3.3) and used in qPCR quantification (See Section 2.3.4.3). The relative expression of PEA3 mRNA in HepG2 was normalised to the UBC reference gene. For more details see section 2.4.2.

5.3.2.8. Data analysis

EMSA analysis: The specificity of the protein-DNA binding was determined by non-specific competitors (See Section 4.3.2.9). The intensity of the protein-DNA complexes for the biotinylated probes was measured using the UVP BioImaging software (UVP, UK), while for the IRDye probes measured by the Li-Cor software (Li-Cor, UK).

Real time PCR analysis: Basal gene expression was first expressed as ΔCq, the quantification cycle (Cq) at which fluorescence was detectable above background for the test gene minus the Cq of the reference gene (UBC). The ΔCq values were then standardised relative to the expression levels of the target in the HepG2 cells. The fold increase of the target gene (PZ) after PEA3 overexpression was normalised against UBC and determined using the Pfaffl equation (See Section 4.3.2.6) since the primer efficiency of the reference gene and the target gene were significantly different. The fold decrease of PZ and PEA3 after PEA3 knock down was determined using the formula $2^{-\Delta Cq}$, where 2 was the presumed amplification efficiency for the target and reference genes and the ΔCq was the Cq of target gene - the Cq of the reference gene. To compare the relative expression of PEA3 and CREB mRNA in HepG2 to normal liver tissue, the formula $2^{-\Delta\Delta Cq}$ was used, where the relative expression in HepG2 was calculated after normalisation with the UBC gene and the liver tissue. These two formulas were used to analyse the results after TaqMan qPCR, since the primers of both the target genes and the reference gene were 100% (Applied Biosystems, UK), and to analyse the PEA3 knock down experiments by SYBR green since the primer efficiencies of PEA3, PZ(2) and UBC highly similar (See Table 2.4). The significance of the difference between samples and controls was determined using REST 2009 software (See Section 4.3.2.9).

5.3. Results

5.3.1. Protein binding to the putative age-stable elements and other potential PEA3 binding sites in the PZ promoter

EMSA reactions using biotin labelled pPZ ASE probe, spanning nucleotides -120/-101, flanking the proximal age-stable-like element (-114/-108) showed protein binding using nuclear extract from HepG2 cells (Fig 5.3 bands 1 and 2). This detection system also visualised a potentially biotinylated protein (non-specific protein) in the HepG2 NE (bands A-C, Fig 5.3) in addtion to the biotinylated DNA probe.

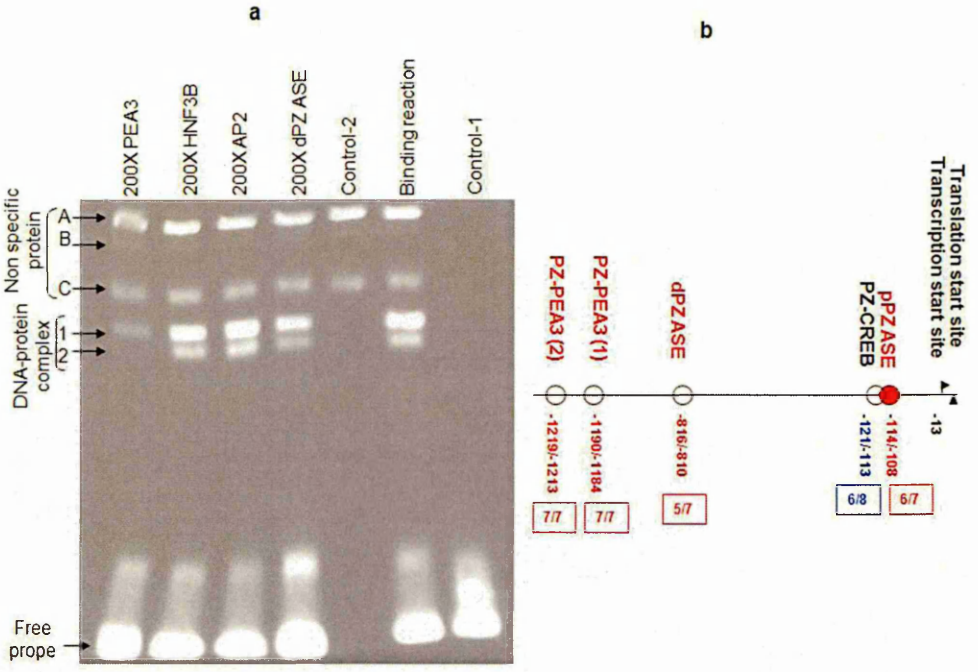


Fig 5.3. EMSA using the putative biotinylated PZ proximal ASE (pPZ ASE) binding site. (**a**) Nucleotide sequences of the probe and competitors are shown in Table 2.2 (Section 2.2.3). Control-1 is the biotinylated probe (pPZ ASE -120/-101) flanking the hypothetical proximal ASE (-114/-108) without HepG2 nuclear extract (HepG2 NE). Control-2 is the HepG2 NE without the probe. Binding reaction is the probe with the HepG2 NE. All other competition reactions contain the probe and the NE as well as unlabelled 200-fold excess dPZ ASE, 200-fold excess unlabelled AP2 (non-specific competitor) consensus, 200-fold excess unlabelled HNF3B and 200-fold excess PEA3 binding site consensus, respectively. Bands indicated by arrows A, B, and C represent a non-specific protein detected in the nuclear extract. Bands indicated by 1 and 2 represent protein-DNA complexes. No competition was found with the AP2 or HNF3B competitors and only weak competition of 1.1-fold and 1.2-fold decrease was seen with the related sequence dPZ ASE in the intensity of complex 1 and 2 respectively. However, complex 1 with the PEA3 consensus was dramatically reduced, with 4-fold decrease in the intensity, both complexes 1 and 2. However, PEA3 consensus showed a strong competition to both complexes, with 4-fold decrease in the intensity of complex 1 and abolishment of complex 2. The intensity was measured by the UVP software. (**b**) Red colour circle shows the proximal ASE site in the PZ gene promoter while the numbers indicate the hypothetical spanning region. Unfilled circles represent other potential binding sites. Numbers in the boxes show the similarity with the binding sites consequences in terms of number of nucleotides.

Competition reactions using the pPZ ASE probe showed no significant decrease in the formation of the probe-protein complex with the addition of a 200-fold excess of either unlabelled HNF3B or AP2 binding site consensus sequences, and negligible competition with a 200 fold excess of the dPZ ASE in the intensity of the protein-DNA complexes 1 and 2 (a 1.1-fold and 1.2-fold decrease, respectively). A 200-fold excess of the PEA3 binding site consensus showed a significant competition with a 4-fold decrease in the intensity of complex 1 and abolition of complex 2 (Fig 5.3). The intensity of the complexes were determined by the UVP software (UVP, UK).

Using the flourecent labelling system a single protein-probe complex was identified, as opposed to the two complexes seen using the chemiluminescent procedure. The intensity of this protein-DNA complex of the pPZ ASE was insignificantly competed (approximately1.3-fold decrease) with a 100-fold excess of CREB binding site consensus and with a 100-fold excess of SP1 binding site consensus (approximately 1.1-fold decrease) whilst no competition with a 100-fold excess of AP1 binding site consensus (Fig 5.4).

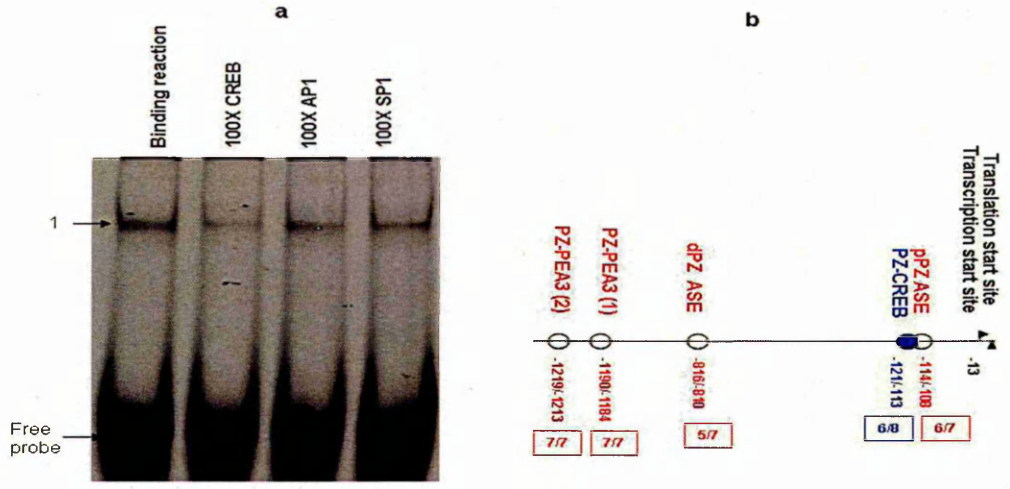


Fig 5.4. IRDye pPZ ASE (-120/-101) competitive EMSA

(**a**) Nucleotide sequences of the probe and competitors are shown in Table 2.2 (Section 2.2.3). Binding reaction is the probe spanning the region -120/-101 with HepG2 NE. Other reactions contained the probe, HepG2 NE, and a 100-fold excess of CREB, 100-fold excess AP1 or a 100-fold excess of SP1 consensus binding sitesrespectively. Bands indicated by arrow 1 represent DNA-protein complexes. The binding reaction showed the DNA-protein binding complex (1). With 100-fold CREB competitor a relatively weak competition of the bound protein with the 100-fold excess CREB binding site consensus is seen (1.3-fold). No competition is seen in the next reaction, with a 100-fold excess of the AP1 while a 100-fold excess of the SP1 binding sites consensus sequences negligibly competed the complex (1.1-fold). (**b**) The blue circle shows the potential CREB binding site in the PZ gene promoter while the numbers indicate the hypothetical spanning region. Unfilled circles represent other potential binding sites. Numbers in the boxes show the similarity with the binding sites consequences.

Competitive EMSA of the probe (dPZ ASE -822/-803) flanking the distal age-stable like element (-816/-810) showed strong reduction of the protein-DNA complexes with the addition of a 200-fold excess of the unlabelled pPZ ASE, with a 7-fold decrease in the intensity of both binding complexes 1 and 2. Addition of 200-fold excess of unlabelled HNF3B or AP2 binding site consensus sequences showed no effect on either protein-DNA complex. A 200-fold excess of the PEA3 binding site consensus showed a significant competition to both binding complexes with a 6-fold decrease in the intensity of complex 1 and abolition of complex 2 (Fig 5.5).

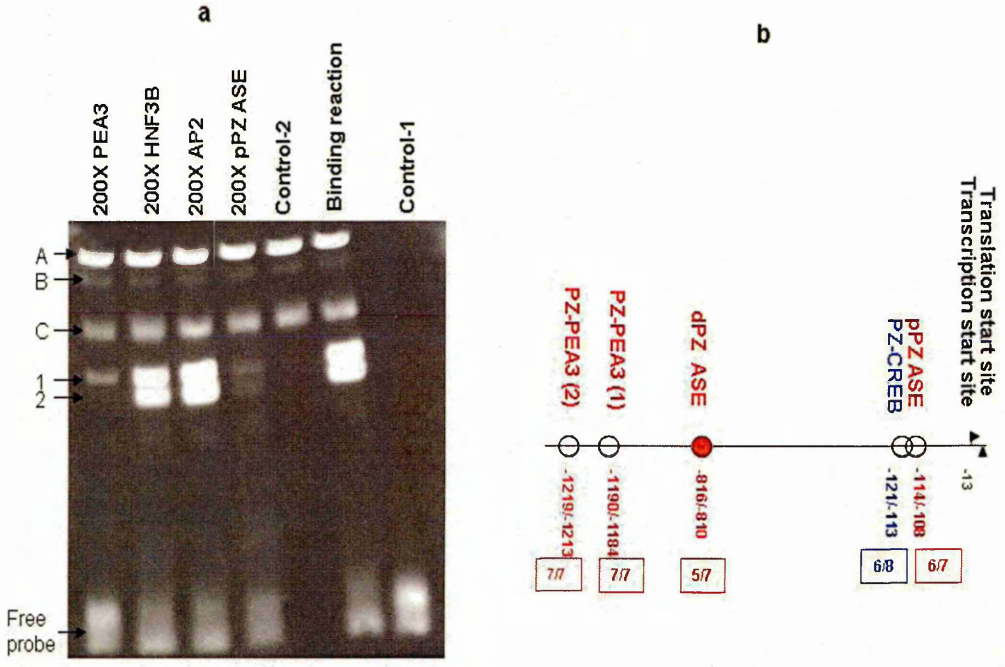


Fig 5.5. EMSA using the putative biotinylated PZ distal ASE (dPZ ASE) binding site.
(**a**) Nucleotide sequences of the probe and competitors are shown in Table 2.2 (Section 2.2.3). Control-1 is the biotinylated probe (dPZ ASE -822/-803) flanking the hypothetical proximal ASE (-816/-810) without HepG2 NE. Control-2 is the HepG2 NE without the probe. Binding reaction is the probe with the HepG2 NE. All competition reactions contain the probe and the NE as well as unlabelled 200-fold excess of pPZ ASE, 200-fold excess of unlabelled AP2 (non-specific competitor) consensus, 200-fold excess of unlabelled HNF3B and 200-fold excess of PEA3 binding site consensus, respectively. Bands indicated by arrows A, B, and C represent biotinylated protein detected in the nuclear extract. Bands indicated by 1 and 2 represent protein-DNA complexes. No competition was observed with the AP2 or HNF3B competitors, however, strong competition was seen in both complexes 1 and 2 with the related sequence pPZ ASE and the PEA3 consensus. The intensity of both protein-DNA complexes (1 and 2) were decreased 7-fold with the pPZ ASE. Protein-DNA complex 1 decreased 6-fold in intensity while complex 2 was abolished with the PEA3 consensus. The intensity of the complexes was measured by the UVP software. (**b**) The red circle shows the distal ASE site in the PZ gene promoter while the numbers indicate the hypothetical spanning region. Unfilled circles represent other potential binding sites. Numbers in the boxes show the similarity with the binding sites consensus sequences in terms of number of nucleotides.

The PZ-PEA3(1) probe (-1203/-1174) flanking a potential binding site for PEA3 (-1190/-1184) showed a protein-DNA binding complexes (Fig 5.6, bands 1 and 2) with the Hep NE. Competitive EMSA showed a reduction in the intensity of the protein-DNA complexes with increasing fold excess of either the PEA3 binding site consensus or the AP1 binding site consensus. The protein-DNA complex showed a reduction of 1.5-fold, 2.2-fold and 4-fold with a 50-fold, 100-fold and 200-fold excess of the PEA3 binding site consensus, respectively. The protein-DNA complex also showed a reduction of 1.5-fold, 1.7-fold and 3-fold with a 50-fold, 100-fold and 200-fold excess of the AP1 binding site consensus, respectively.

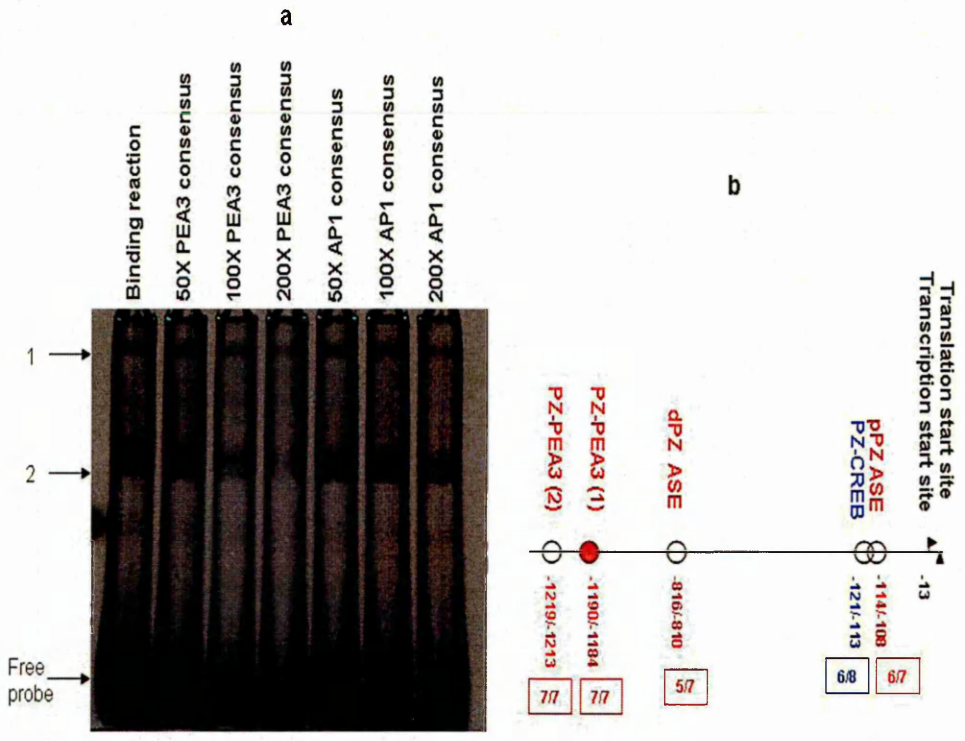


Fig 5.6. PZ-PEA3(1)-fluorescent (IRDye) competitive EMSA
(a) Nucleotide sequences of the probe and competitors are shown in Table 2.2 (Section 2.2.3). The binding reaction is the probe flanking the region -1190/-1184 with HepG2 NE and without competitor. Competition reactions contain the probe, HepG2 NE, and 50, 100 and 200-fold excess of PEA3 consensus, and 50, 100 and 200-fold excesses of AP1 consensus, respectively. Bands indicated by arrows 1 and 2 represent DNA-protein complexes. Both DNA-protein binding complexes represented by 1 and 2 did not show specificity, since they competed with increasing fold excess of PEA3 and AP1 binding site Consensuses. **(b)** The red circle shows the hypothetical PZ-PEA3(1) site in the PZ gene promoter while the numbers indicate the spanning region. Unfilled circles represent other potential binding sites. Numbers in the boxes show the similarity with binding sites consensus sequences.

The PZ-PEA3(2) probe (-1230 to -1204) flanking another potential PEA3 binding site (-1219 to -1213) was also shown to be bound by a protein (Fig 5.7). Competition analyses showed a reduction of the bound protein of 2-fold, 2.5-fold and 3-fold with a 100-fold, 200-fold and 300-fold excess of the PEA3 binding site consensus sequence, respectively, and a decrease of 1.8-fold, 2.8-fold and 3.5-fold with a 100-fold, 200-fold and 300-fold excess of the AP1 binding site consensus sequence, respectively (Fig 5.7).

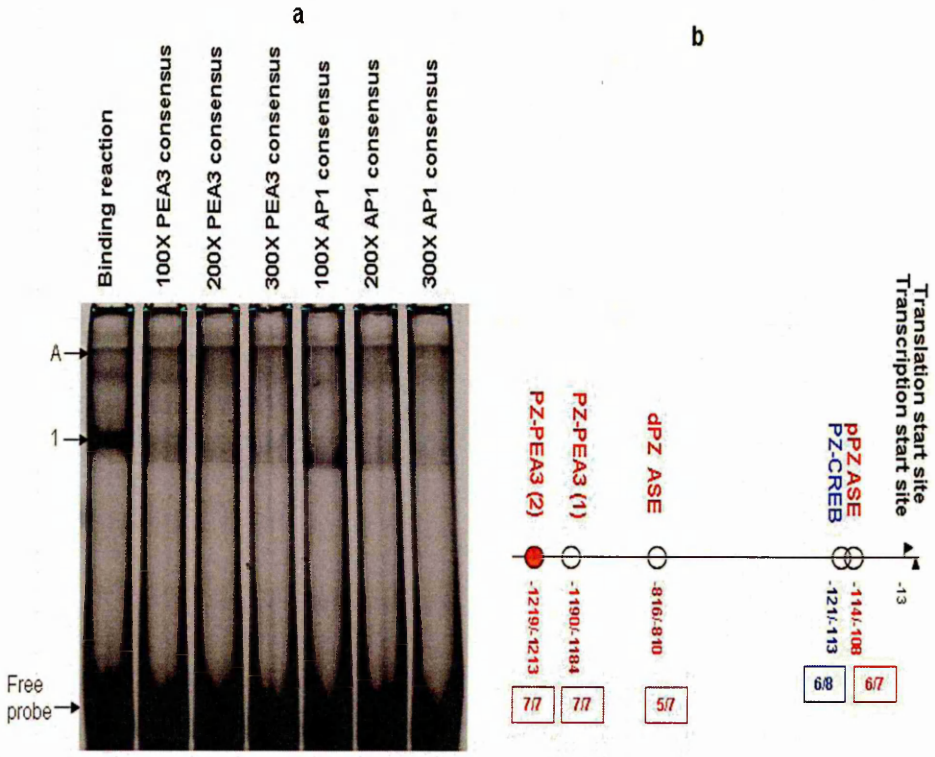


Fig 5.7. PZ-PEA3(2)-fluorescent (IRDye) competitive EMSA
(a) Nucleotide sequences of the probe and competitors are shown in Table 2.2 (Section 2.2.3). Binding reaction is the probe flanking the region -1219/-1213 with HepG2 NE and without competitor. Competition reactions contain the probe, HepG2 NE, and 100, 200 and 300-fold excess of PEA3 consensus and 100, 200 and 300-fold excess of AP1 consensus, respectively. Bands indicated by arrows A show an artifact from the running samples on the gel. B represents DNA-protein complexes. Protein bound to DNA which represented by arrow B is non-specific as it is competed with increasing fold excess of both PEA3 and AP1. **(b)** The red circle shows the hypothetical PZ-PEA3(2) site in the PZ gene promoter while the numbers indicate the spanning region. Unfilled circles represent other potential binding sites. Numbers in the boxes show the nucleotide similarity compared to binding sites consequences.

5.3.2. Identification of the protein bound to the pPZ ASE and dPZ ASE

Polyclonal antibodies raised against a PEA3 or Ets1 polypeptide were pre-incubated with the HepG2 NE for 10 minutes and further incubated with the pPZ ASE binding reaction at room temperature for approximately 30 minutes. No effect on the protein-DNA complex was observed with the addition of the Ets1 antibody (Fig 5.9). However, weak competition was observed with the addition of a PEA3 antibody (Fig 5.8). The intensity of the protein-DNA complex was decreased a 1.5-fold with the addition of the antibody. The incubation of the Ets1 antibody with the HepG2 NE for 10 minutes on ice and further incubated with the probe for approximately 35 minutes again did not show any effect on the protein-DNA complex (Fig 5.9).

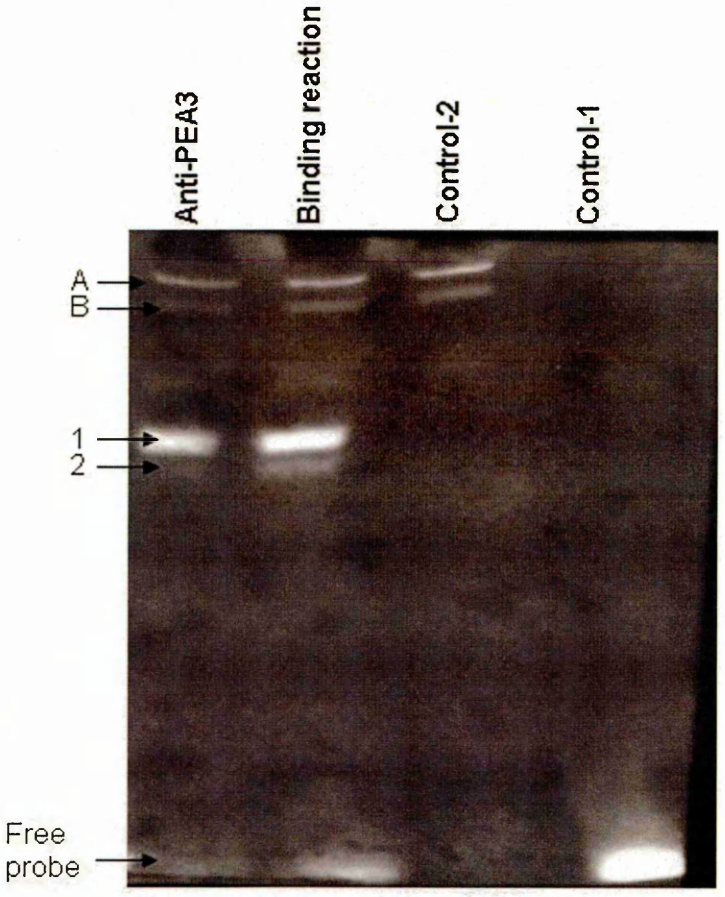


Fig 5.8. Supershift assay using biotinylated pPZ ASE binding sites.
Control-1 is the pPZ ASE probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE. Anti-PEA3 is an antibody against the PEA3 protein together with the binding reaction. The antibody was incubated with the HepG2 NE for 10 minutes and then further incubated with the probe for 25 minutes at room temperature. No supershift was seen with the probe. However, the intensity of the protein-DNA complex was reduced 1,5 fold. A and B are control bands which come up whenever the NE is added. 1 and 2 are the protein-DNA complexes.

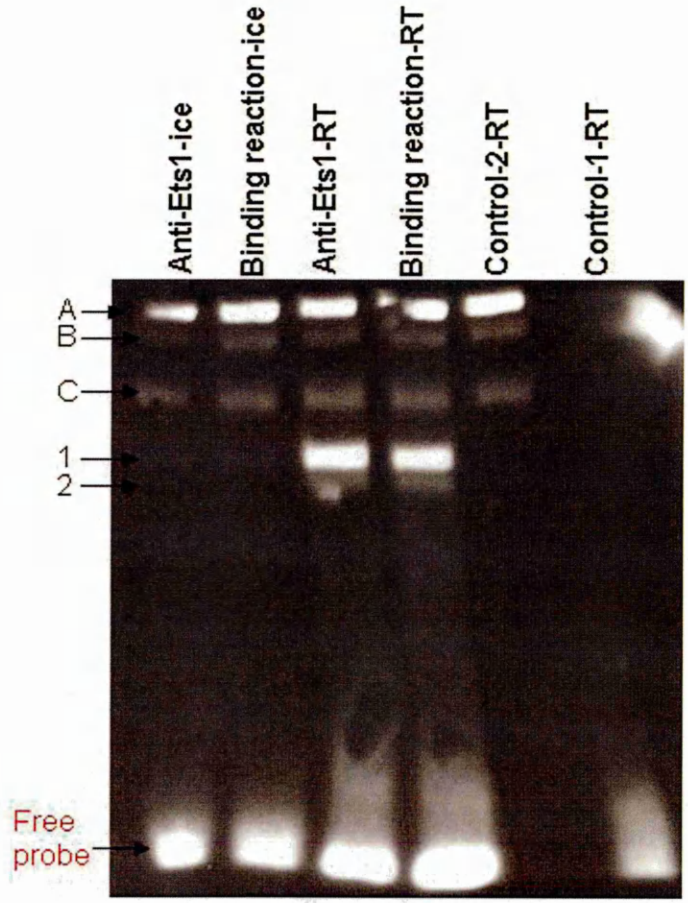


Fig 5.9. Supershift assay on ice and room temperature using biotinylated pPZ ASE binding sites with Ets1. Control-1 is the pPZ ASE probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE. Anti-Ets1 is an antibody against Ets1 protein together with the binding reaction. From right to left, the first four lanes were performed at room temperature (RT) where the HepG2 NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. The following two lanes were performed on ice where the HepG2 NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. No supershift was seen at room temperature while on ice neither protein-DNA complex nor antibody-protein-DNA complex. A, B and C are biotinylated protein which come up whenever the NE is added. Bands 1 and 2 are the protein-DNA complexes.

A supershift assay of the IRDye-labelled pPZ ASE probe using a CREB antibody was also performed. The antibody was pre-incubated with the Hep NE for 10 minutes followed by a further incubation with the probe for 30 minutes. No effect on the protein-DNA complex was observed (Fig 5.10).

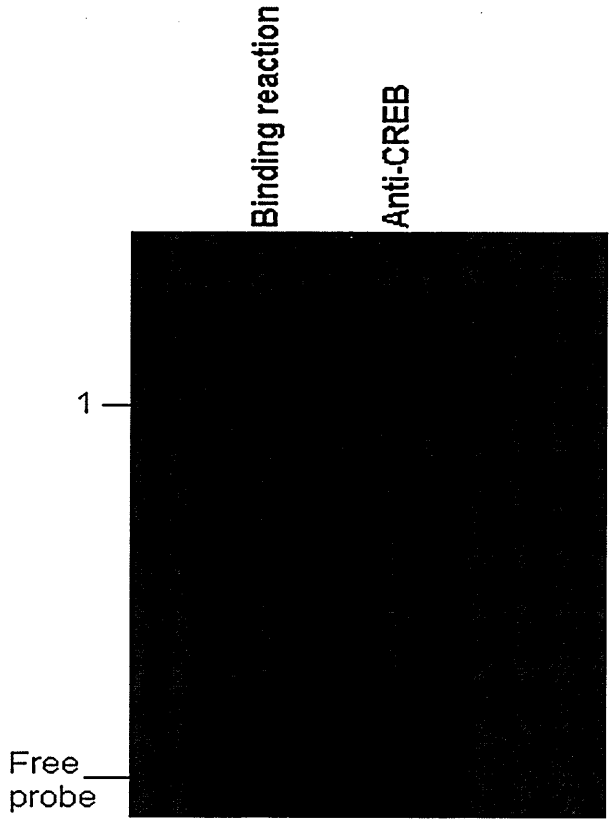


Fig 5.10. Supershift assay at room temperature using fluorescent CREB probe with CREB antibody. Binding reaction is the probe with HepG2 NE. Anti-CREB is an antibody against the CREB protein together with the binding reaction. The reaction was performed at room temperature where the HepG2 NE was pre-incubated with the CREB antibody for 10 minutes and followed by 25-30 minutes incubation with the probe. No supershift is seen. Band 1 is the protein-DNA complex.

Supershift assays of the dPZ ASE performed by pre-incubation of the HepG2 NE with either the PEA3 or Ets1 antibody at room temperature for 10 minutes followed by incubation with the probe for approximately 30 minutes also did not show a shift in the protein-DNA complex (Fig 5.11). However, in the presence of the Ets1 antibody the intensity of the protein-DNA complex increased four-fold.

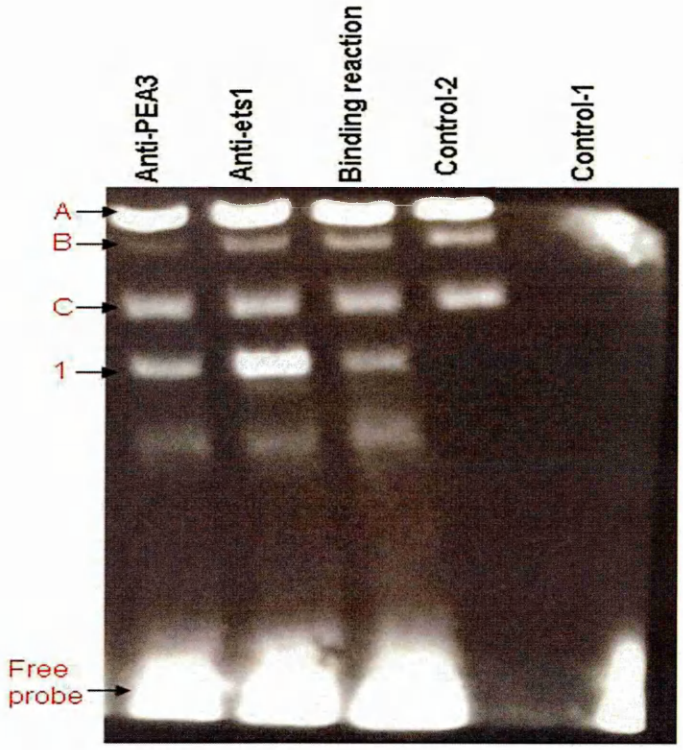


Fig 5.11. Supershift assay at room temperature using biotinylated dPZ ASE binding sites with PEA3 and Ets1. Control-1 is the dPZ ASE probe without HepG2 NE. Control-2 is the HepG2 NE without probe. Binding reaction is the probe with HepG2 NE. Anti-PEA3 is an antibody against PEA3 protein together with the binding reaction. Anti-Ets1 is an antibody against Ets1 protein together with the binding reaction. The reactions were performed at room temperature where the HepG2 NE incubated with the binding mixture for 10 minutes and followed by 25-30 minutes incubation with the probe. No supershift was seen with both anti-PEA3 and Ets1 although 4-fold increase in the intensity of the band in the presence of Ets1. A, B and C are control bands which come up whenever the NE is added. Band 1 is the protein-DNA complex.

5.3.3. Gene expression studies of the endogenous PZ gene in HepG2 cells

5.3.3.1. Sequencing of PZ gene promoter

The promoter sequence of the HepG2 endogenous PZ locus (See Appendix) was found to match the reported wild type PZ promoter sequence (Homo sapiens protein Z, AF440358, NCBI).

5.3.3.2. Verification of the PEA3 expression plasmid

The digestion of the PEA3 expression plasmid (pCMV6-XL5-PEA3) with Sac1 showed the expected restriction fragments of 1.91 kb and 4.78 kb (Fig 5.12).

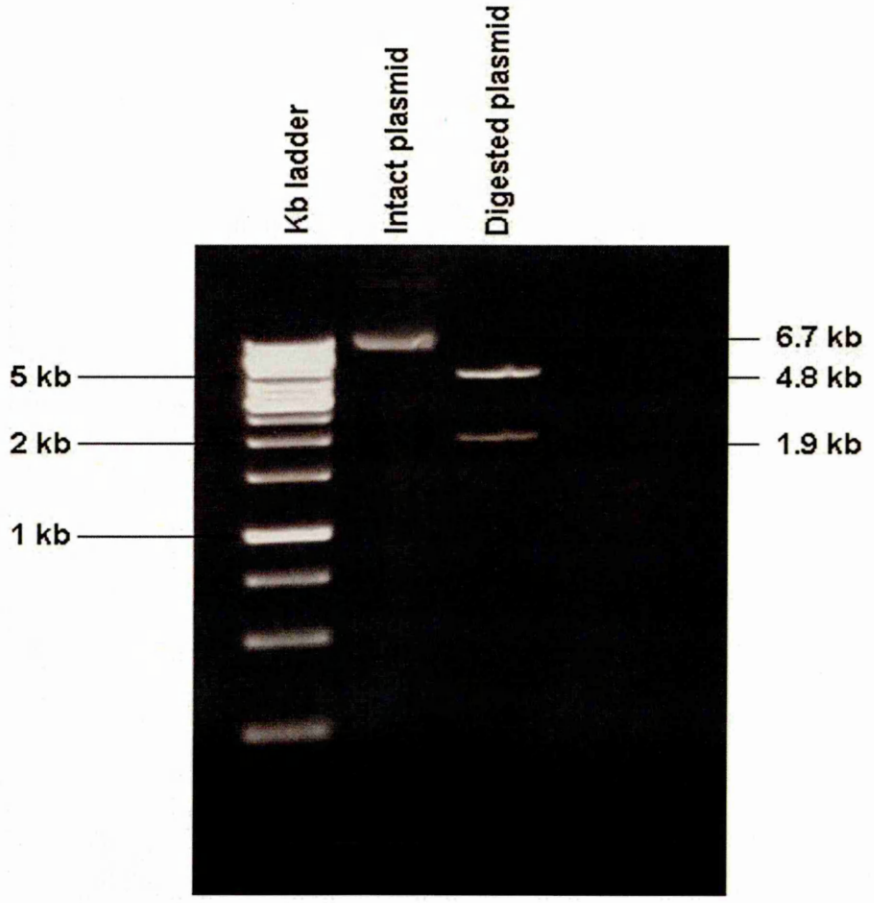


Fig 5.12. PEA3-encoded plasmid digested with restriction enzymes.
Kb ladder (fermentas, UK). The PEA3 plasmid was digested with Sac1, showing the expected bands 4.8 kb and 1.9 kb.

5.3.3.3. RNA analysis

The purity of the RNA samples was determined by the ratio of absorbance at 260 nm and 280 nm and found to be > 1.8 in all cases. The 18S and 28S rRNA bands from both transfected and untransfected HepG2 cells appeared without significant degradation upon analysis on a 1% agarose gel, indicating a good quality of RNA samples (Fig 5.13).

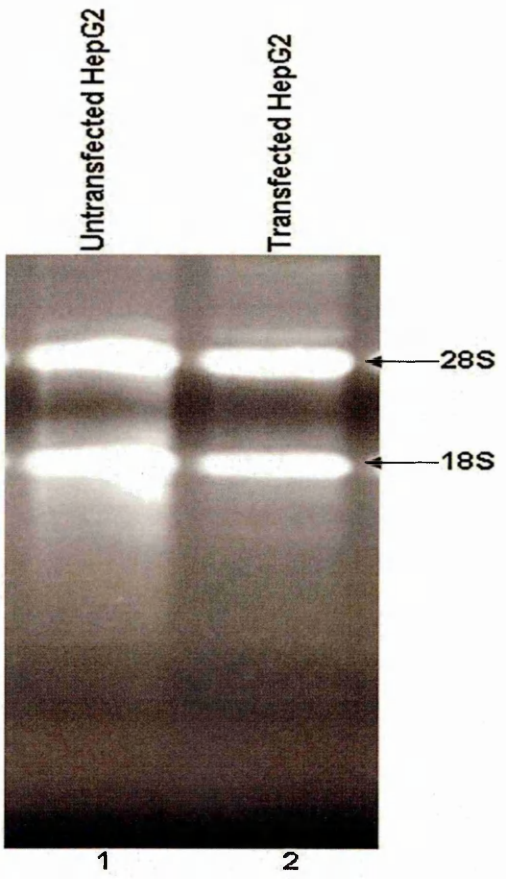


Fig 5.13. Analysis of total RNA from transfected and untransfected HepG2
The integrity of RNA was checked using a 1% w/v 1x TBE agarose gel. The 28S and 18S bands are the ribosomal RNA (rRNA). Lane 1 represents the rRNA bands from untransfected sample. Lane 2 shows the rRNA bands from the PEA3 transfected samples (500ng of total RNA was loaded). Bands from both samples show good quality RNA.

5.3.3.4. Reference gene validation

The UBC housekeeping gene in untransfected and PEA3 expression plasmid transfected HepG2 cells showed no change in its relative expression determined by SYBR green qPCR (2^delta Cq = 0.99, n=5).

5.3.3.5. Efficiency of primers

The primer efficiencies of the reference gene (UBC) and the gene of interest (PZ) were determined (See Section 2.3.4.1) before calculating relative gene expression: primer pair efficiencies were found to be 95% and 86% for UBC and PZ, respectively (Fig 4.8).

5.3.3.6. Transactivation of PZ gene by PEA3 overexpression in HepG2 cells

The relative expression ratio was determined on the basis of primer efficiencies and Cq values from the control (untransfected) and sample (PEA3 transfected) cells for the investigated transcripts using the Pfaffl equation (See Section 5.3.2.8 and Section 4.3.2.6). The level of endogenous PZ mRNA was found to be increased by an average of 1.9 fold in response to overexpression of PEA3 in HepG2 cells (Fig 5.14). This increase in PZ expression was found to be statistically significant ($p < 0.05$, n=6) using the REST 2009 software. No primer dimers or non-specific amplification was detected by melt curve analysis (See Fig 2.4).

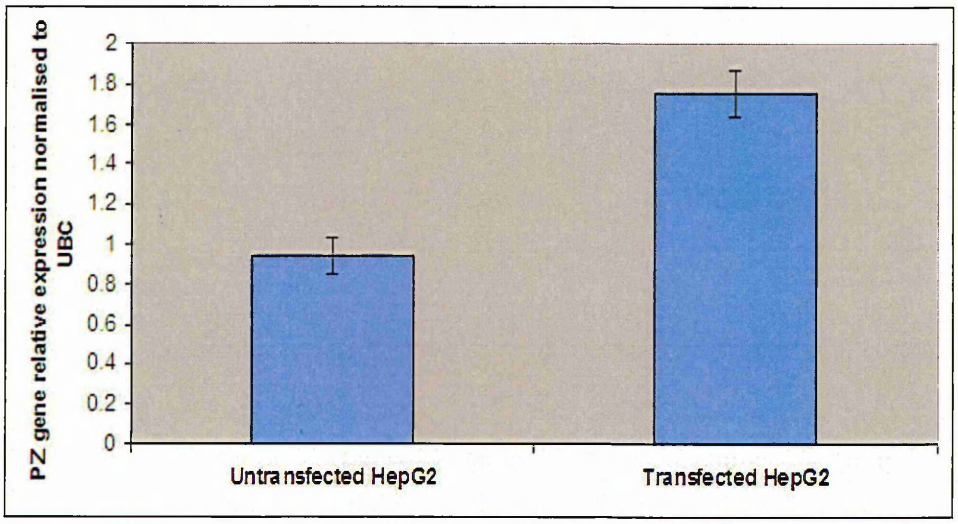


Fig 5.14. The effect of PEA3 overexpression on PZ mRNA relative expression.
The PEA3 transfected cells show a relative expression of 1.9-fold more PZ mRNA compared with the untransfected (control) sample. The increase in the relative expression was found statistically significant using the REST 2009 software ($P<0.05$, n=6).

5.3.3.7. The effect of PEA3 knockdown on PZ mRNA expression

The efficiencies of UBC, PEA3 and PZ(2) primers used in PEA3 knockdown experiments were found to be 95%, 96% and 98%, respectively (Fig 5.15). Knockdown of PEA3 in HepG2 cells using an On-targetplus smart pool siRNA led to an 80% average reduction of PEA3 mRNA compared to the negative control (ON-Targetplus Non-Targeting pool).

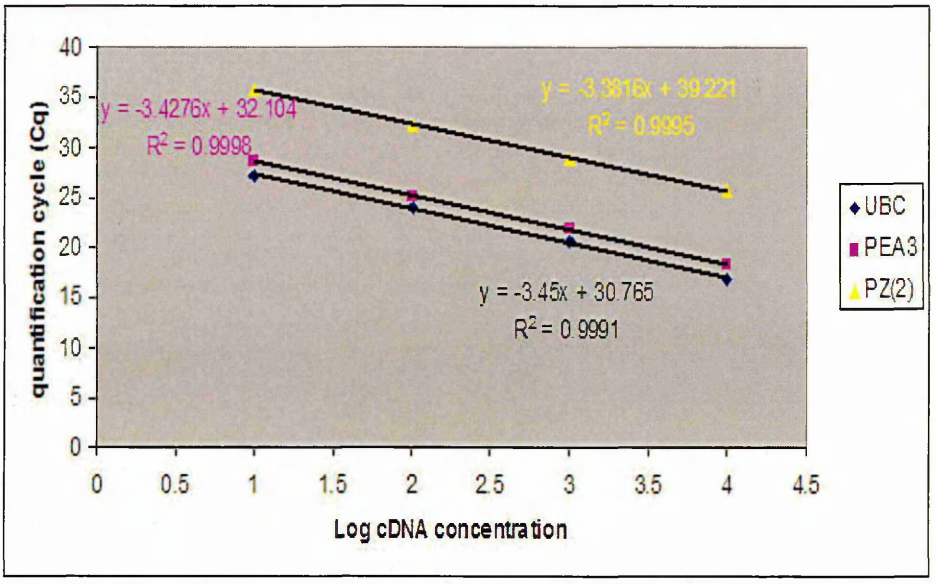


Fig 5.15. Determination of primers efficiencies used in PEA3 Knockdown analyses
The efficiency of the reference gene (UBC) and the target genes, PEA3 and PZ(2) primers were measured by plotting Cq cycles versus cDNA serial dilution to calculate the slope which used in the efficiency equation. Investigated transcripts showed high primers efficiencies of 95%, 96% and 98% for UBC, PEA3 and PZ(2), respectively, with high linearity.

However, this reduction (Fig 5.16) was not found to be statistically significant using the REST 2009 analysis package (P=0.1, n=5).

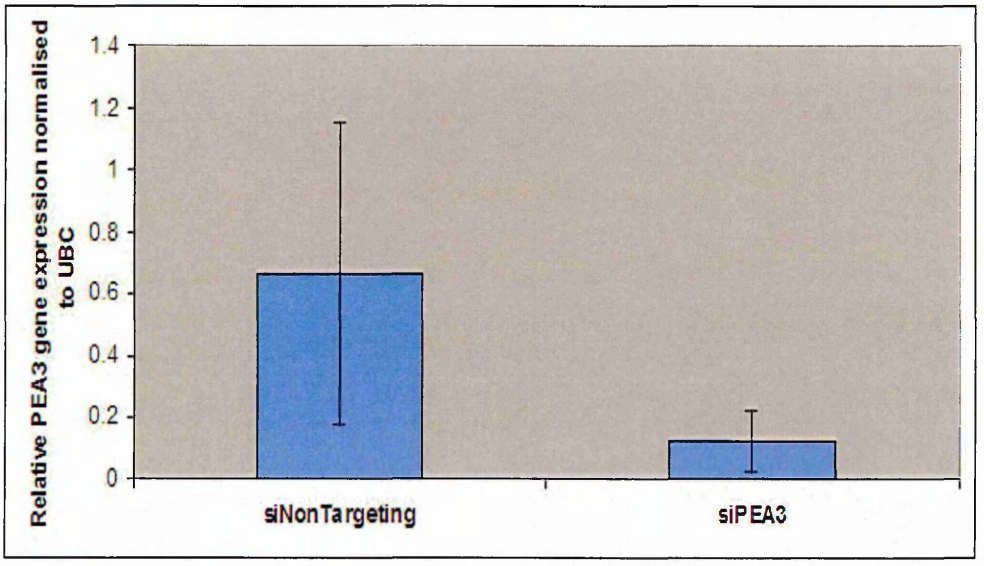


Fig 5.16. PEA3 knockdown in HepG2 cells using siRNAs.
The On-targetplus smart pool siRNA for PEA3 shows an 80% reduction in relative PEA3 mRNA levels compared to the non targeting negative control. However, the decrease was found statistically not significant (P=0.1, REST 2009, n=5).

The knockdown of the relative PEA3 mRNA expression increased the levels of relative expression of PZ mRNA to 1.9 fold (Fig 5.17). The REST 2009 software, however, showed that the increase in the PZ mRNA relative expression was not statistically significant (P=0.5, n=6).

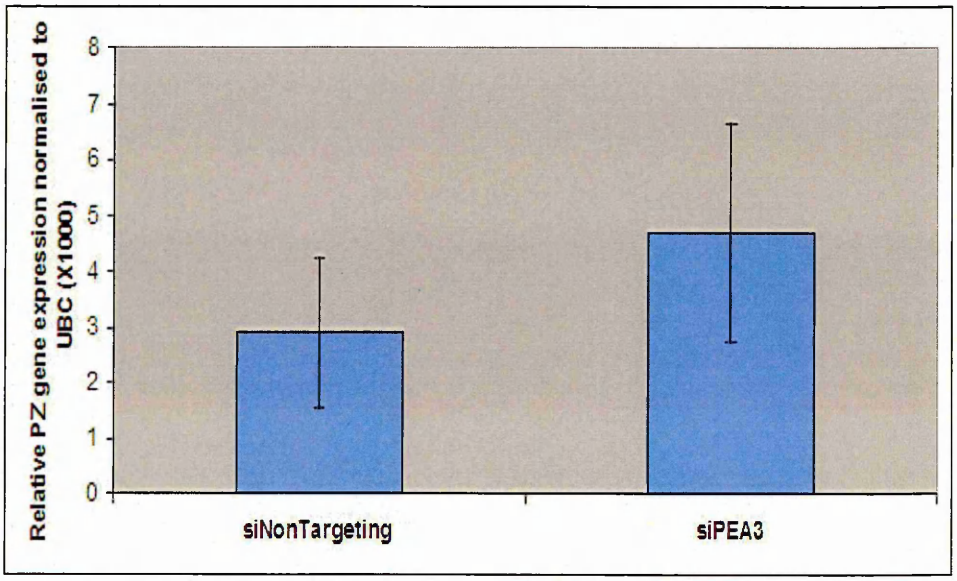


Fig 5.17. The effect of PEA3 Knockdown on PZ mRNA relative expression.
Knockdown of PEA3 in HepG2 cells increased the PZ mRNA levels. 80% reduction of PEA3 mRNA relative expression increased the relative expression of PZ mRNA by approximately 1.9 fold. The increase of the PZ mRNA relative expression was found statistically not significant using the REST 2009 software (P=0.5, n=6).

5.3.3.8. The effect of CREB knockdown on PZ mRNA expression

The relative expression of CREB mRNA in HepG2 cells was knocked down using the respective siRNA and found to be reduced by 75% (Fig 5.18). The reduction was found to be statistically significant using the REST 2009 software ($P < 0.05$, n=7).

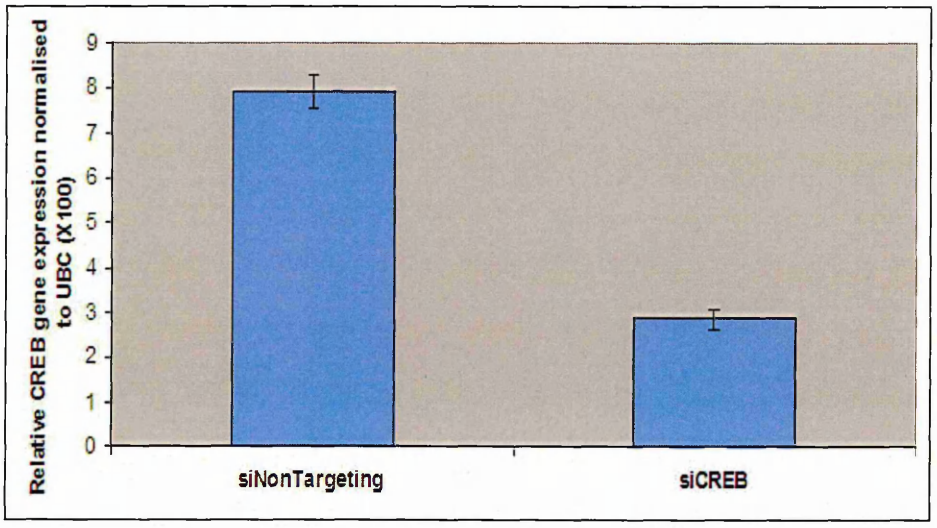


Fig 5.18. CREB knockdown in HepG2 cells using siRNAs.
The transfection of the siRNA for CREB showed 75% reduction in relative expression of CREB mRNA compared to the non targeting negative control. The Knockdown was found statistically significant using the REST 2009 software ($P<0.05$ t-test, n=7).

The 75% reduction of CREB mRNA in HepG2 cells increased the relative expression of PZ mRNA by 1.4 fold (Fig 5.19). This increase, however, was not found to be statistically significant using the REST 2009 software (P= 0.2, n=7).

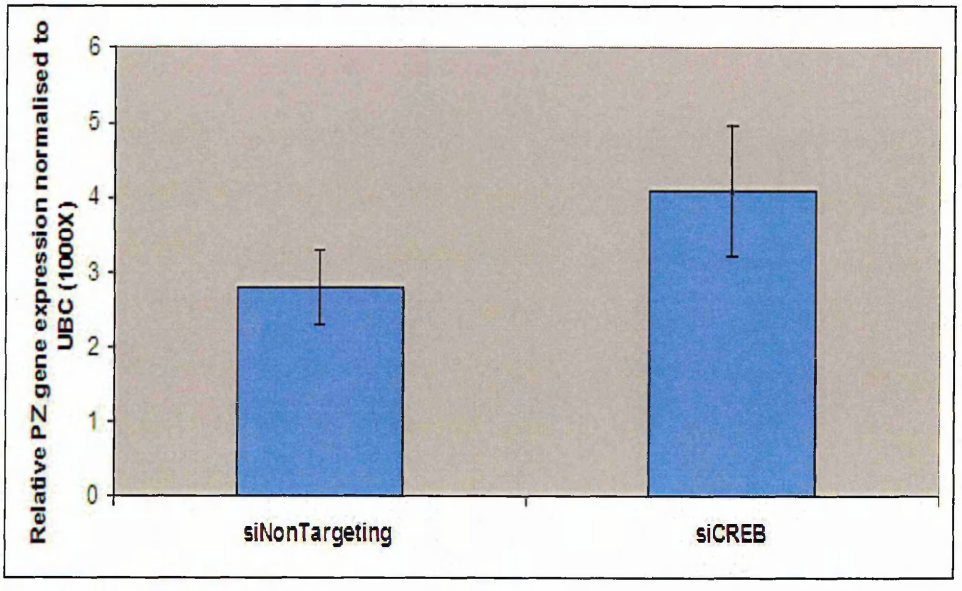


Fig 5.19. The effect of CREB Knockdown on PZ mRNA relative expression.
Knockdown of CREB in HepG2 cells increased the PZ mRNA levels. 75% reduction of CREB mRNA relative expression increased the relative expression of PZ mRNA to approximately 1.4 fold (P= 0.2, REST 2009, n=7).

5.3.3.9. Expression of PEA3 and CREB in HepG2 compared to normal human liver

The relative gene expression of PEA3 and CREB was found to be higher in HepG2 cells compared to normal liver tissue (Fig 5.20). PEA3 mRNA relative expression was found to be considerably higher in HepG2 cells with approximately a 1767-fold higher level than normal liver tissue. The relative expression of CREB mRNA was found to be 2-fold higher in HepG2 compared to the normal liver tissue. These differences of relative expression of PEA3 and CREB mRNAs between HepG2 and the normal liver tissue were found statistically significant using the REST 2009 software ($P<0.05$, $n=3$).

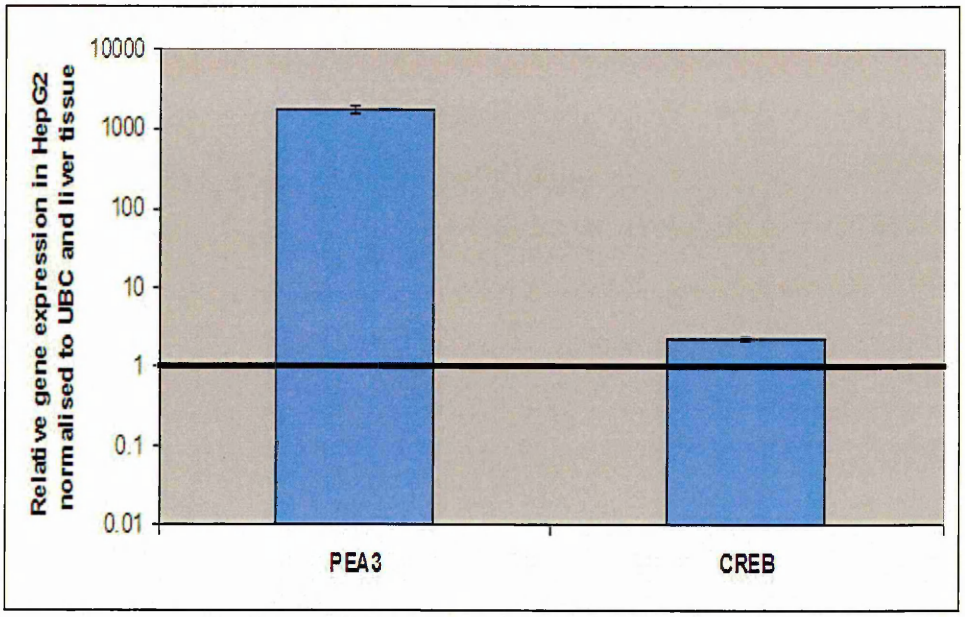


Fig 5.20. The realtive expression of CREB and PEA3 mRNA in HepG2 and normal liver cells. The figure shows a logarithmic scale for the relative expression of PEA3 and CREB in HepG2 cells compared to liver tissue and normalised to UBC reference gene. The relative expression of PEA3 mRNA in HepG2 was found to be approximately 1767-fold more than in liver tissue. CREB mRNA expression was shown to be approximately 2-fold more in HepG2 than that of liver tissue. The results were found statistically significant ($P<0.05$, REST 2009, $n=3$). The horizontal thick line represents the relative expression of mRNA expression in liver tissue. The housekeeping gene UBC was used as a reference gene.

5.4. Discussion

The accurate identification of transcription binding sites, which can be accomplished by protein-DNA interaction studies, is essential for understanding complex gene expression patterns (Morozov *et al.* 2005). EMSA has been one of the most common techniques extensively used in such DNA-protein interaction studies (Fried and Crothers 1981, Revzin 1989) and in the current study. The protein-DNA binding results from the EMSAs showed interaction between regions of the PZ gene promoter and nuclear proteins prepared from HepG2 cells. Two age-stable like elements, identified in the PZ gene promoter by bioinformatic analysis, show a close homology to those reported in hFIX and hPC (Kurachi *et al.* 1999, Zhang, Kurachi and Kurachi 2002). The age-stable like elements in hFIX and hPC, genes that are evolutionarily related to hPZ (Ichinose *et al.* 1990), have been proposed as PEA3 binding sites (Kurachi *et al.* 1999, Zhang, Kurachi and Kurachi 2002). However, a more recent study from Kurachi and colleagues in 2009 using minigenes and transgenic mice constructs, as well as EMSA, indicated that the protein bound to the age-stable element (ASE) in the hFIX promoter was in fact Ets1 (Kurachi *et al.* 2009). The two age-stable like elements identified in PZ, the proximal (pPZ ASE) and distal (dPZ ASE) sites, showed binding with protein in the nuclear extract of HepG2 cells using EMSA (Fig 5.3 and 5.5).

During chemiluminescent detection bands on the EMSA, which were not due to the interaction of the streptavidin with the biotinylated probes, were detected whenever HepG2 NE used. These protein bands were likely due to the interaction of the streptavidin with nuclear protein and referred to as non-specific protein (Bands A, B and C, Fig 5.3 and 5.4). A similar artefact was detected previously in chemiluminescet EMSA using Hep NE and referred to as a non-specific band (Liang, Hassett and Omiecinski 2005).

Competitive EMSAs showed no competition of the protein bound pPZ ASE with unlabelled 200-fold AP1 and 200-fold HNF3B excess binding consensus sequences, suggesting specificity of the bound protein (Fig 5.3). The bound protein was strongly competed by the PEA3 binding site consensus sequence, with a 4-fold decrease in the intensity of binding complex 1 and a loss of complex 2. Competitive EMSA of the dPZ ASE showed similar results, with no effect from the 200-fold AP1 or HNF3B, but a significant reduction of binding

complexes by 200-fold PEA3, with a 6-fold decrease in the intensity of binding complex 1 and loss of binding complex 2 (Fig 5.5). The strong competition of the bound protein to the pPZ ASE and the dPZ ASE by the PEA3 binding site consensus suggested that the bound protein was PEA3. However, there remained the possibility that the bound protein may also be Ets1, since the PEA3 binding site consensus used in the competitive EMSAs (Table 2.2) was the same binding site consensus as Ets1 (Santa Cruz literature). The interaction of different Ets family members to DNA is based on the bases flanking the core binding site consensus GGAA/T (Xin *et al.* 1992). However, different bases flanking the core binding site were reported for the binding specificity of PEA3 and Ets 1. PEA3 was found to bind the sequence CAGGAAG in the polyomavirus enhancer (Martin *et al.* 1988) and suggested to bind A/CGGAAGT in a study of the DNA binding specificity of PEA3 using nuclear extract from mouse FM3A cells and bacterial extract (Xin *et al.* 1992). Moreover, a PEA3 oligonucleotide flanking the AGGAAA core sequence was used to demonstrate the expression of PEA3 in HepG2 cells (Guo and Inagami 1994). The Ets 1 protein was also found to bind the AGCGGAAGCG sequence in the promoter of murine sarcoma virus long terminal repeat (MSV LTR) (Gunther *et al.* 1990), AGAGGATGTG in T-cell receptor α gene enhancer (Ho *et al.* 1990), GCA/CGGAAGT in the stromelysin promoter (Wasylyk *et al.* 1991) and GAGGAAG in the hFIX promoter (Kurachi *et al.* 2009). Therefore, competitive EMSAs could not provide a definitive answer as to the identity of the bound protein.

The region (-120/-101) flanking the pPZ ASE was also found to contain another hypothetical binding site (ALGGEN-PROMO), that of CREB (-120/-113) sharing two base pairs with the overlapping pPZ ASE (-114/-108); suggesting that the protein bound to the region (-120/-101) flanking the pPZ ASE could be CREB.

Competition reactions of the IRDye labelled pPZ ASE showed some competition of the bound protein with 100-fold excess of the CREB and SP1 binding site consensus sequences but no competition with 100-fold excess of AP1 binding site (Fig 5.4). The intensity of the protein-DNA complex was reduced to 1.3-fold and 1.1-fold with the CREB and SP1 binding site consensuses, respectively. However, another study found that the probe flanking this region was competed by a mutated CREB binding site consensus

(Sugawara *et al.* 2007). Therefore, the role of CREB in regulating expression of PZ remains unclear.

Bioinformatic analysis showed two other potential binding sites for PEA3 in the PZ promoter, PZ-PEA3(1) spanning the region -1190 to -1184 (CTTCCTG) and PZ-PEA3(2) from -1219 to -1213 (CTTCCTG). EMSA reactions using the probe flanking the PZ-PEA3(1) binding site and the probe flanking the PZ-PEA3(2) binding site indicated non-specific protein binding to both sites (Fig 5.6 and 5.7). The bound protein at both sites, using competitive EMSA, was competed with an increasing fold excess of either the PEA3 binding site consensus or the AP1 binding site consensus. These results suggest that the protein which is bound to both regions was not specific, despite the slightly better affinity of the PZ-PEA3(1) bound protein to the PEA3 binding site consensus compared to the non specific competitor AP1 (Fig 5.6).

The initial EMSAs suggested that the protein bound specifically to the pPZ ASE could in fact be PEA3, Ets 1 or, possibly, CREB, while the protein bound to the dPZ ASE could be PEA3 or Ets 1. These hypotheses were based on the strong competition of the bound protein at pPZ ASE with the PEA3 binding site consensus, which is also considered as an Ets 1 binding site consensus, and also some competition with the CREB binding site consensus (Fig 5.3 and 5.4); in addition to the strong competition of the protein bound to dPZ ASE with the PEA3 binding site consensus (Fig 5.5).

To identify the specific protein bound to the pPZ ASE and dPZ ASE supershift assays were performed. Supershift assay of the pPZ ASE using a polyclonal antibody raised against a PEA3 polypeptide in the binding reaction did not show a shift of the protein-DNA complex band (Fig 5.8). Nevertheless, the protein-DNA complexes were weakly competed in the presence of the antibody, with a decrease of 1.5-fold in the intensity of the complexes. The protein-antibody interaction in EMSA may cause a reduction of the protein-DNA complex, a complete supershift of the protein-DNA complex or a supershifted complex along with the protein-DNA complex (Carey and Smale 2001). Ets 1 and CREB antibodies were also used in supershift assays of the pPZ ASE and showed no effect on the protein-DNA complexes (Fig 5.9 and 5.10). CREB and PEA3 antibodies, however, did not show interactions with the protein bound to CREB and PEA3 consensus sequences, respectively.

Further supershift assays were also performed with the dPZ ASE using PEA3 and Ets 1 antibodies. The results showed no interaction of the PEA3 antibody, again, with the bound protein. However, the presence of the Ets 1 antibody caused a 4-fold increase in the intensity of the protein-DNA complex, which could potentially be due to a antibody-protein-DNA complex (Fig 5.11). Ets 1 antibody was found to increase the intensity of the binding complex in the supershift assay used to identify the protein bound ASE in hFIX promoter (Kurachi *et al.* 2009) and therefore, support the result of the Ets 1 antibody with the dPZ ASE in the current study.

Although EMSA is an easy and sensitive technique that has been used extensively in studying protein-DNA interactions *in vitro*, many disadvantages of this assay have been reported. *In vivo* conditions required to optimise transcription factor binding, such as the need for other proteins for their binding to the DNA may not be obtained *in vitro* (Wells and Farnham 2002). Moreover, the *in vivo* conformation of the DNA secondary structure, such as looping, which is needed to bring distal protein binding sites into close proximity can not be created *in vitro*. Different results from different electrophoresis condition may also be obtained for the same DNA-protein system (Sidorova, Muradymov and Rau 2005). Therefore, transcriptional analyses of the endogenous PZ gene within its chromatin environment were performed.

To further analyse the role of PEA3 in PZ transcription, this transcription factor was overexpressed in HepG2 cells. Prior to the transfection, the PEA3 encoded expression plasmid (pCMV6-XL5-PEA3) was prepared (see section 2.4.3) and its identity was verified by digestion with the SacI restriction enzyme and size-fractionated on a 1% w/v agarose gel (Section 2.4.3.3). Digestion of the expression plasmid with SacI showed the expected bands 1.91 kb and 4.78 kb (Fig 5.12) according to the webcutter website (webcutter 2.0 website). The sequence as well as SNPs of the PZ gene promoter in HepG2 was validated (See Sections 5.3.3.1, 4.3.2.1 and 2.1) and found to be identical to the reported wild type PZ promoter (Homo sapiens protein Z, AF440358, NCBI). Having verified the PEA3 expression plasmid it was transfected transiently into HepG2 cells (See Section 2.4.4) and the effect on the endogenous PZ mRNA relative expression measured using qPCR (See Section 2.3.4.3). Overexpression of PEA3 in HepG2 was found to increase the relative expression of PZ mRNA 1.9

fold (Fig 5.14). The increase of the PZ mRNA relative expression was found to be statistically significant ($P<0.05$, REST 2009, n=6). PEA3 has previously been reported to up-regulate the expression of many genes, such as, the vimentin gene in the mammary epithelial and tumour cells (Chen *et al.* 1996) and increase the human transmembrane mucin (MUC4) mRNA in pancreatic cancer cells (Fauquette *et al.* 2005). However, it was also reported to have down-regulatory effect on other genes, such as, HER-2/neu (Xing *et al.* 2000, Hung and Wang 1999).

The transfection reagent, Tfx-20, which has been used successfully in the transfection of HepG2 cell to study gene transcription (Dongol *et al.* 2007) and in the current study, was tested for the off-target effect which may be caused by the introduction of foreign DNA (Jacobsen, Calvin and Lobenhofer 2009) and found to have no effect on the expression of the gene of interest (Fig 4.12). This confirmed that the increase in the relative expression of PZ mRNA was due to the overexpression of PEA3. The moderate increase of the relative expression of PZ mRNA (1.9-fold) in response to PEA3 overexpression might be due to the high level of PEA3 relative expression in HepG2 (1767-fold more than normal liver tissue), which may be close to the effect threshold of PEA3 on PZ expression.

To confirm the role of PEA3 transactivation in the transcriptional regulation of the endogenous PZ gene in HepG2, the expression of endogenous PEA3 was knocked down using small interfering RNAs (siRNA) (see Section 2.4.5) and the effect on the endogenous PZ mRNA determined using qPCR. UBC was used as a reference gene assuming its stability with PEA3 overexpression would be the same as with PEA3 knock down. A new primer pair PZ(2) (Table 2.3) was also used which is different to that used in the PEA3 and HNF4α overexpression qPCR experiments. The reason for designing additional primers for the PZ transcript was to obtain a level of efficiency closer to that of the reference gene primers. Smart pool siRNAs, which have been used extensively in gene expression studies (Danes *et al.* 2008, Kwei *et al.* 2008, Hariparsad *et al.* 2009, Taniguchi *et al.* 2009, Rizzo *et al.* 2010, Lahkim Bennani-Belhaj *et al.* 2010), were used to knock down PEA3 in HepG2 cells. Smart pool siRNAs were developed and used to minimise the off-target silencing which has been observed with genes containing a limited

complementary sequence with siRNAs (Jackson *et al.* 2003), since they contain four predesigned siRNAs targeting one gene (Dharmacon Technologies Literature).

The observed PEA3 knockdown and the small increase of the relative expression of PZ mRNA in response to the PEA3 knockdown were both found not to be statistically significant (P=0.1, n=5 and P=0.5, n=6, respectively). The significant increase in PZ mRNA relative expression due to the overexpression of PEA3, along with the non significant results of the PEA3 knockdown (presumably because of the very high endogenous level of PEA3 expression in HepG2) may indicate, overall, that PEA3 may be a transcriptional activator of PZ, either directly or indirectly.

Study of the knockdown of CREB in HepG2 cells showed an increase in PZ mRNA. The 75% reduction of CREB relative expression (Fig 5.18), which was statistically significant ($P<0.05$, n=7), was found to increase the relative expression of PZ mRNA 1.4 fold (Fig 5.19). However, this increase in PZ expression was not found to be statistically significant (P=0.2, n=7). Although CREB protein is expressed in normal liver tissue and HepG2 cells (Fig 5.20), the knockdown results along with the competition of the protein-DNA complex in the potential CREB binding site by the mutated CREB binding site consensus (Sugawara *et al.* 2007) suggest that CREB has no regulatory effect on the PZ gene.

The expression of the endogenous PZ gene in the HepG2 cell model, is an in cell system with the gene in the context of the intact chromatin structure of the DNA rather than naked dsDNA. However such an in cell system is still considered an *in vitro* situation and such data derived from this model should be confirmed *in vivo.*

5.5. Conclusion

It can be concluded that PZ gene expression may be regulated by PEA3 transcription factors. Competitive EMSAs for the protein bound to the ASE-like regions (potential binding sites for PEA3 or Ets 1) suggest the potential role for PEA3 and/or Ets 1 in the regulation of the PZ gene (Fig 5.3 and 5.5). Ets 1 binding to the dPZ ASE was confirmed using the Ets 1 antibody in the

supershift assay (Fig 5.11). However, the regulation of this transcription factor to the endogenous PZ gene needs to be confirmed. The up-regulatory role of PEA3 on the endogenous PZ gene (Fig 5.14) was confirmed by over expressing PEA3 in the HepG2 cell model ($P<0.05$, $n=6$). However, this technique still considered as an *in vitro*, and therefore, *in vivo* studies should be further performed for confirmation.

Although CREB was initially suggested in the current study to bind the PZ gene promoter by competition of the protein-DNA complex by the CREB binding site consensus, a mutated CREB binding site consensus has shown the same competition in other studies, indicating that the bound protein was not CREB. This was investigated by the gene expression study of the endogenous PZ using the knockdown technique by siRNAs in HepG2 cells. Knockdown of CREB mRNA (75 %) showed a 1.4-fold increase in the relative expression of PZ mRNA; however this was not found to be statistically significant ($P=0.2$, $n=7$). Therefore, it can be concluded that CREB has no role in the regulation of the PZ gene.

Chapter 6-Optimisation Of Footprinting and Chromatin Immunoprecipitation Assays

6.0. Introduction

The binding of transcriptional regulatory proteins to their specific DNA binding sites which was studied by EMSA, as discussed in chapter three, can also be studied by DNase I footprinting. Although EMSA is more sensitive compared to footprinting and the protein-DNA complex can be detected with a lower concentration of the protein in the nuclear extract (Carey and Smale 2001, Hellman and Fried 2007), the footprinting assay can provide information which cannot be obtained by EMSA. This information is, for instance, the location of protein binding as well as multiple protein binding sites within a DNA fragment. In addition, detection of proteins requiring binding of factors to adjacent recognition sequences on the probe can also be visualized (Carey and Smale 2001). Examples of such interactions include the synergistic binding of a multiprotein Polycomb core complex (PCC) and Pleiohomeotic (PHO) to the Polycomb response elements (PREs), an epigenetic DNA element in Drosophilia (Mohd-Sarip *et al.* 2005), and the synergistic binding of the *Vibno fischeri* Luminescence gene activator (LuxR) and the RNA polymerase (RNAP) in the promoter of the luminescence gene coding for autoinducer synthase (LuxI) (Stevens, Dolan and Greenberg 1994).

The footprinting assay which is used for studying a wide variety of sequence-specific protein-DNA interactions and which can potentially allow characterisation of transcriptional regulatory mechanisms of a gene, was developed by Galas and Schmitz in 1978 (Galas and Schmitz 1978). The technique is based on the protection of the protein bound DNA regions from digestion by a cleavage agent. The undigested region of DNA, which is protected by the bound protein, shows a gap in the resulting ladder of DNA fragments on denaturing polyacrylamide gel electrophoresis (Hampshire *et al.* 2007). Alternatively, the bound protein may cause severe distortions to the DNA structure, resulting in an increase in the intensity of the bands rather than their decrease (Guille and Kneale 1997). The footprinting assay is summarised in figure 6.1. Footprints, where the DNA is protected by the bound protein are normally localised using G+A and/or C+T sequencing ladders (Maxam and Gilbert 1980) which are run alongside the protected and naked DNA ladders (Leblanc and Moss 2009). Cleavage agents such as hydroxyl radicals which are produced from the Fenton reaction between Fe^{2+} and H_2O_2 can be used in

the footprinting assay. Although the accuracy in the estimation of protein binding sites is high, hydroxyl radical footprinting has disadvantages such as being time consuming and it is sensitive to solvents that scavenge free radicals (Hampshire *et al.* 2007).

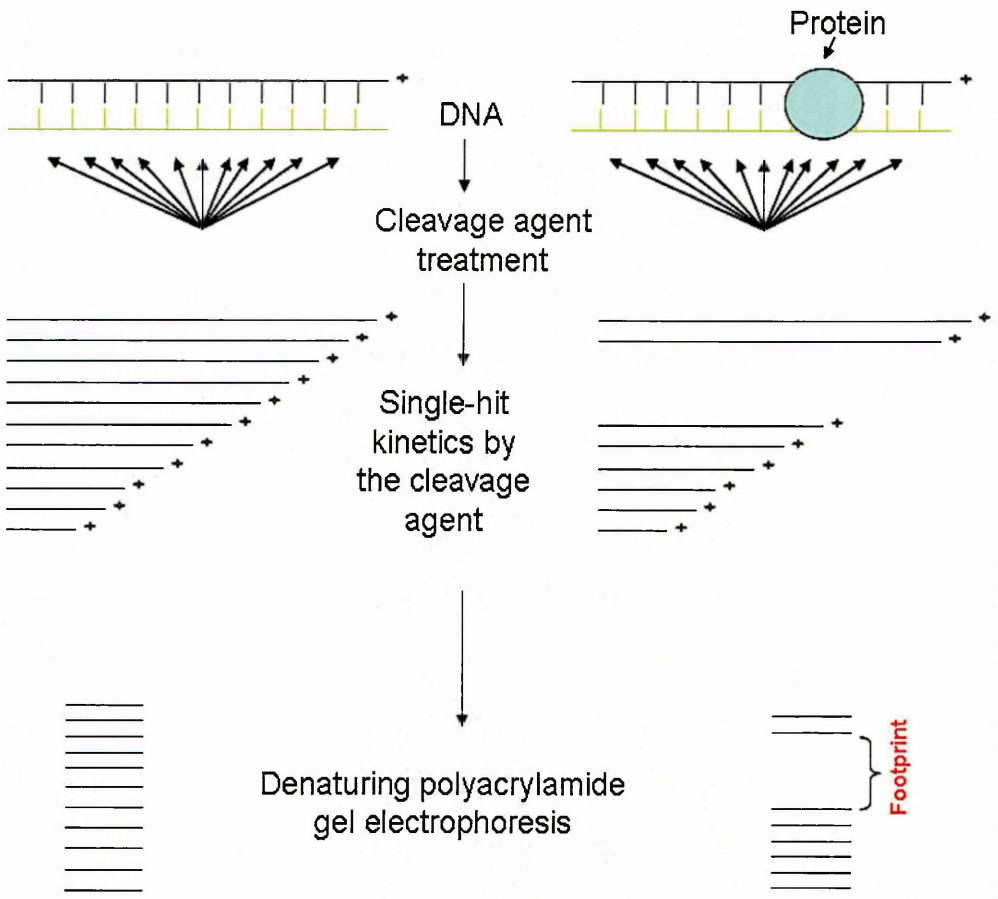


Fig 6.1. Schematic representation of the footprinting assay. The DNA fragment is labelled (star) at one end of one strand. The DNA is digested using a cleavage agent, such as, DNase I, The digested DNA is size fractionated using denaturing polyacrylamide and visualised. The DNA region bound by protein shows a gap (footprint) due to the protein protection of this region from digestion.

The cleavage agent DNase I is prone to steric hindrance when attacking DNA, because of its large size and, therefore, it has been the most popular choice in footprinting, producing clear interference patterns (Leblanc and Moss 2009). However, a disadvantage of its large size is that it prevents it from cutting the DNA immediately after the protein-DNA complex and, therefore, does not show the exact boundaries of the footprint (Guille and Kneale 1997). DNase I is a double-stranded DNA endonuclease which binds the minor groove of the DNA

and interacts with the backbone, cleaving the O3'-P of both strands independently (Hampshire *et al.* 2007, Suck, Lahm and Oefner 1988).

A development of this technology is, instead of using nuclear extracts which may not accurately represent the *in vivo* situation, to look at binding site occupancy '*in vivo*'. The *in vivo* techniques in protein-DNA analyses have many advantages over the *in vitro* assays. Such advantages include transcription factors, *in vivo*, interacting with the DNA in its chromatin context which can work as a barrier to the accession of proteins (Demeret, Vassetzky and Mechali 2001), rather than naked DNA *in vitro*, which may not represent the real situation in the body. Moreover, sequences which are similar or identical to those bound by tissue specific factors could be bound by ubiquitous factors *in vitro*, complicating analyses (Ikuta and Kan 1991).

The *'in vivo'* footprinting technique, although not strictly *in vivo*, has been commonly used to detect protein-DNA binding sites within intact cells (Hornstra and Yang 1993). Dimethyl sulphate (DMS) is a common agent used in '*in vivo*' footprinting. This agent preferentially reacts with guanine bases causing their methylation (Sclavi 2008) and the degree of reactivity between the DNA bound by protein and the same sequence of the naked DNA is compared. Although '*in vivo*' footprinting using DMS was the most common technique, because of its membrane permeability, the use of DNase I has been found to provide better information about the protein-DNA interaction. This is due to the better delineation of footprinting boundaries, since DNase I cleaves the DNA surrounding protein-DNA complexes and detection of complexes without guanine contacts is observed (Hornstra and Yang 1993). The permeability of the cells to be subjected to '*in vivo*' DNase I footprinting has been achieved using ethanol, described for use with *Escherichia coli* (Cassler *et al.* 1999), and lysolecithin in the footprinting of primary cultures of vascular umbilical smooth muscle cells and peripheral blood lymphocytes (Angers *et al.* 2005). Ligation-mediated PCR exponentially amplifies the genomic sequence ladder, following denaturation of the chemically or enzymatically cleaved genomic DNA, and has been used in '*in vivo*' footprinting to eliminate experimental signal-to-noise ratio and the use of large numbers of cells (Mueller and Wold 1989). The PCR-mediated ligation is performed using two primers complementary to the ends of the DNA fragments. One end of the DNA fragments is fixed by a primer or

restriction cut whilst the other end is unique as it is determined by chemical cleavage or chain termination (Mueller and Wold 1989).

Having localised the binding sites for a protein by EMSA using bioinformatic analysis or practically by footprinting the identity of the bound protein has to be determined. EMSA (supershift assay) is one of the assays which can be used for the identification of the potential protein using a protein-specific antibody. However, using a synthetic probe spanning the detected binding site in the EMSA analysis does not represent the normal situation of the regulatory DNA sequence within its complex chromatin environment. Therefore, the identification of the bound protein by the chromatin immunoprecipitation technique may be utilised since the protein is identified within that chromatin context, as mentioned above, in the discussion of '*in vivo*' footprinting. Although this assay has been used for several years to confirm the identity of the protein which was predetermined to bind a specific binding sequence (Narayanan *et al.* 2004, Collas 2009, Hill *et al.* 2011), the development of this technique has been able to identify target binding sites for specific transcriptional regulators throughout the genome (Collas 2009).

In the technique of ChIP (Fig 6.2), the protein and DNA are cross-linked in live cells with formaldehyde and then the sheared chromatin immunoprecipitated with an antibody against the putative protein to enrich the chromatin for regions to which the proposed protein binds. This is followed by chromatin reverse cross-linking and comparison of the immunoprecipitated enriched DNA with negative control DNA using PCR (Aparicio, Geisberg and Struhl 2004).

The development of the ChIP assay has led to its combination with microarrays in a technique known as ChIP-on-Chip to provide greater knowledge by mapping transcription factor binding throughout the genome. The ChIP-on-Chip experiment (Fig 6.2) was first reported by the group of Ren *et al* (2000) for their mapping of the binding sites of Gal4 and Ste12 transcription factors in the entire yeast genome (Ren *et al.* 2000). In this technique blunt ends are generated in the sheared reverse cross-linked DNA using DNA polymerase, ligated with common linkers and amplified with primers complementary to the linkers using PCR. This is followed by hybridisation of the recovered DNA fragments to a microarray containing probes for the transcription factor binding regions of interest on a genome-wide scale (Sikder and Kodadek 2005). However, bias

from the extensive amplification required for the large amounts of DNA necessary for the hybridisation, the hindrance to interpretation posed by repetitive sequence, the detection of allelic variants due to cross-hybridisation,

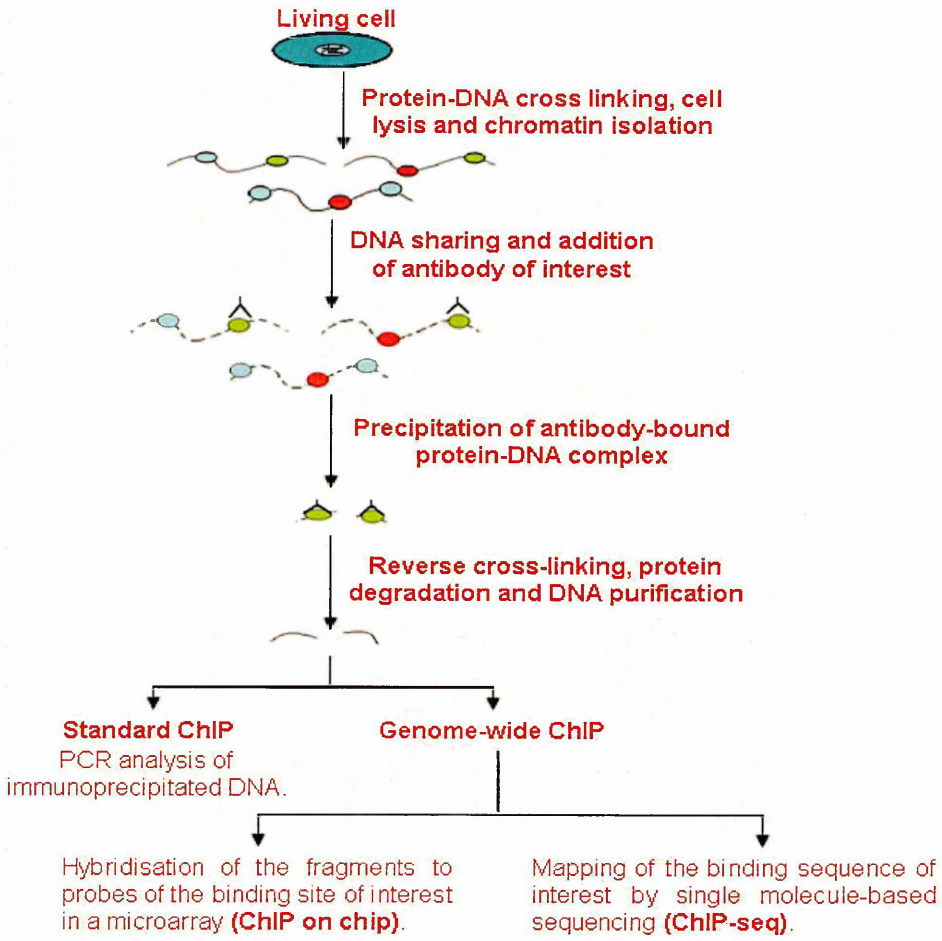


Fig 6.2. Schematic representation of various techniques of chromatin immunoprecipitation. Protein is cross-linked in living cells and the cells are breaked open and the chromatin is sheared. The antibody of interest is added and the antibody-protein complex is precipitated. The immunoprecipitated chromatin is reverse cross-linked, the protein is degraded and the DNA is purified. The immunoprecipitated DNA is analysed for the protein of interest in the proposed region by PCR analysis (Standard ChIP) or in genome-wide ChIP using hybridisation in a microarray (ChIP on ChIP) or mapping of the sequence of interest (ChIP-seq).

and cost are all considered practical limitations of the ChIP-on-Chip technique in analysing the mammalian genome in such a way (Mikkelsen *et al.* 2007). Another approach recently reported to characterise the immunoprecipitated DNA is high-throughput sequencing (ChIP-seq). This single molecule-based sequencing technology was reported by Robertson's group in Hela S3 cells to

map the binding sequences for the STAT1 protein (Robertson *et al.* 2007). In this technique (Fig 6.2) the molecules of DNA are arrayed across a surface and locally amplified, followed by single-base extension and determination of the added bases. Although the 'read length' is short (25 to 50 bases), simultaneous reads of millions of DNA fragments give sequence data across the full fragment lengths (Mikkelsen *et al.* 2007).

The time involved with the ChIP protocol (3-4 days) and the multi-millions of cells required for immunoprecipitation (Collas 2009), in addition to the involvement of extensive handling which may lead to material loss, technical error and contamination are all limitations of what is a powerful technique. However, to minimise these limitations, modifications to the ChIP protocol have been made to make it simpler, shorter and valid for small cell numbers. For instance, Nelson and his group in 2006 modified the ChIP protocol to take less time. They used the chelating resin Chelex-100 to reduce the time for the preparation of PCR-ready templates after immunoprecipitatin, to one hour instead of one day. Moreover, they incubated the antibody with pre-incubated sheared chromatin in an ultrasonic bath to increase the rate of antibody-protein binding, shortening the time of immunoprecipitation from several hours to only 15 minutes (Nelson *et al.* 2006).

Therefore, numerous techniques exist to map protein-DNA interactions all with their respective benefits and disadvantages. From a holistic regulatory perspective those techniques that allow increased characterisation of such interactions in the context of the three-dimensional organisation of the nuclear DNA, although technically more demanding, are preferable.

The Aims

The aims of this work are to confirm the binding sites which have been identified in the current study by bioinformatic analysis, and to identify additional binding sites in a longer DNA fragment using the footprinting technique. Moreover, to confirm the identity of the proteins, which have been characterised by *in vitro* EMSA in the current study, in their chromatin environment using an '*in vivo*' chromatin immunoprecipitation technique.

6.1. Methods

6.1.1. Footprinting assay

The PZ gene promoter (NCBI, AF440358) was divided into four overlapping regions spanning nucleotides -1099 to +86, relative to the translation start site at nucleotide +1; primers were designed for each region (Fig 6.3) in order to allow amplification by the polymerase chain reaction (PCR).

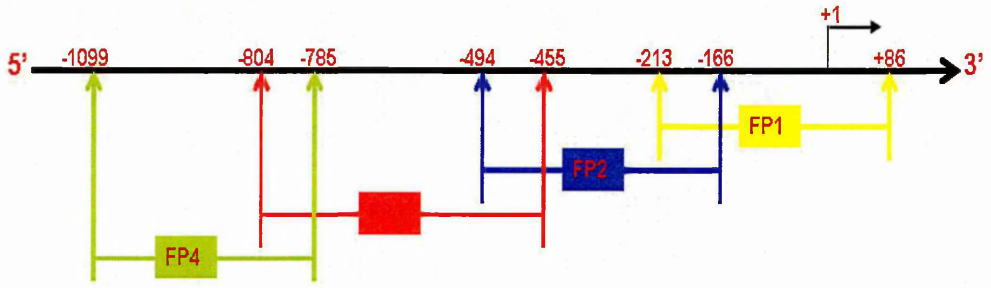


Fig 6.3. Location of the PZ gene promoter regions amplified by the designed primers in footprinting assay. Different colours show the different regions for footprinting assay. Arrows with different colours show the location of primers used to amplify the respective region. The sequences of the primers are shown in table 1. The translation start site is represented by +1. FP is an abbreviation for footprinting region.

The primers for each region were designed using the webprimer website (http://www.yeastgenome.org/cgi-bin/web-primer) and synthesised by Eurofin MWG/Operon (Eurofins MWG Operon, UK). The PCR reactions were performed by the Primus 96 plus PCR machine (MWG-BIOTECH, UK) using commercial human genomic DNA (50ng/µl) prepared from placenta (Sigma-Aldrich, UK), and the designed primer pairs (Table 6.1). The PCR cycle conditions used to amplify the footprinting regions in the PZ promoter were as follow: firstly, one step at 95°C for 2 minutes, to ensure complete denaturation of the DNA template, followed by 30 cycles of 95°C for 1 minute for DNA denaturation, annealing temperature of the proposed primers (Table 2.2) for 1 minute, to anneal the primers to their complementary regions, and 72°C for 1 minute for the extension of the primers by Taq polymerase in order to synthesise complementary strands. Finally, 5 minutes at 72°C was performed to ensure all the single-stranded DNAs were extended and formed double stranded DNAs, followed by a holding step at 4°C for short-term storage.

Primers description		Sequences	Amplicon size
Footprinting (Fp1) -Forward	region-1	5'-TGCTGCCTTTACCTTCTGTTC-3'	300 bp
Footprinting (Fp1)-Reverse	region-1	5'-TAAGCCAGATGCCTACCTGA-3'	
Footprinting (Fp2)-Forward	region-2	5'-AATGGTGCTGAGCTGTAGTGA-3'	329 bp
Footprinting (Fp2)-Reverse	region-2	5'-ACCCTGGGAGATATTTAGGGA-3'	
Footprinting (Fp3)-Forward	region-3	5'-TTTCCCCTGCTTCAGTCATT-3'	350 bp
Footprinting (Fp3)-Reverse	region-3	5'-AAGGGCCAAATAGTGTCTGTC-3'	
Footprinting (Fp4)-Forward	region-4	5-AGCCCAGTGATCCCTTCAGA-3'	315 bp
Footprinting (Fp4)-Reverse	region-4	5'-AATGACTGAAGCAGGGGAAA-3'	

Table 6.1. Oligonucleotide primers used to amplify different PZ promoter regions for footprinting. The primers were designed using webprimer (http://www.yeastgenome.org/cgi-bin/web-primer) and synthesized by Eurofin MWG Operon. The forward primers for each region were biotin-labelled for their detection by the UVP BioImaging system (UVP, UK). In the modified footprinting assay forward primers were labeled by IRDye-700 while reverse primers were labeled by IRDye-800 for their detection in the Licor 4200L electrophoresis system (Li-cor, UK). The annealing temperatures for Fp1, Fp2, Fp3 and Fp4 are 60, 58, 57 and 62.5 °C, respectively.

The PCR cocktail and different parameters used in the amplifications are shown below in table 6.2.

Genomic DNA (100 ng/µl)	1 µl
PCR buffer (10X)	2.5 µl
dNTPs (2 mM)	1.25 µl
$Mgcl_2$ (25 mM)	1-4 µl
Biotin-forward primer (5 pm/ µl)	1 µl
Reverse primer (5 pm/ µl)	1 µl
Nuclease-free water	Upto 25 total volume.
Taq DNA polymerase (5 units/ µl)	0.5 µl

Table 6.2. PCR template and cocktail used in amplifying different promoter regions. the various promoter regions were amplified with the same amounts and concentrations of the cocktail. $MgCl_2$ was optimised for the amplification of different regions.

Prior to the footprinting binding reaction the amplified PCR products were purified by adding the same volume of phenol-chloroform and vortexing, followed by microcentrifugation ($16.1x10^3$ g). The upper aqueous layer was

transferred to a clean microcentrifuge tube followed by addition of 1/10 and 3X volumes of 3 M sodium acetate and absolute ethanol, respectively. The presence of the salt and alcohol disrupts the interaction between the water molecules and the nucleic acid causing precipitation. The solution was chilled for 15 minutes at -70°C and centrifuged at $16.1x10^3$ g for 15 minutes at 4°C. The pellets were rinsed with 500 µl 70% ethanol and centrifuged at $16.1x10^3$ g for 15 minutes at 4°C. The pellets were dried at room temperature and resuspended in 10 µl of water to be used in the footprinting binding reaction.

6.1.1.1. Binding reaction using biotin-labelled primers

DNase I footprinting was performed according to the protocol reported in 2001 by Sambrook and Russell (Sambrook and Russell 2001). The experiment was carried out in 5 different tubes; each 25 µl reaction contained binding buffer-1 (Andrews and Faller 1991) (a final concentration of 16 mM Hepes-KOH pH 7.9, 20 % glycerol, 336 mM NaCl, 1.2 mM $MgCl_2$, 0.16 mM EDTA, 0.32 mM DTT, 1% (v/v) protease inhibitor cocktail (Sigma Aldrich, UK)), 1µg of poly (dI-dC) and 1.0 µl of the 10 µl total volume of the proposed purified biotin labelled DNA fragment (PZ promoter region +86 to -213). HepG2 nuclear extract (HepG2 NE), 10, 20 and 40 µg was added to tubes 2, 3 and 4, respectively, while no nuclear extract was added to tube 1 and 5 (controls). The reaction mixture was centrifuged for 5 seconds at 4°C and incubated for 20 minutes on ice. 50 µl of $MgCl_2$-$CaCl_2$ (10 mM $MgCl_2$ + 5 mM $CaCl_2$) solution, for DNase I activity, was added to the mixture and mixed gently and incubated for 1 minute at room temperature. Digestion with 0.01 units of DNase I (Sigma Aldrich, UK) was left for exactly 1 minute at room temperature and the enzyme was inactivated using 75 µl of stop mix solution (20 mM EDTA, 1% w/v SDS, 0.2 M $NaCl_2$, and 125 µg/ml yeast tRNA). The DNA was extracted using phenol:chloroform and centrifuged at 4°C for 10 minutes at 16.1 $x10^3$ g after incubation at -20°C overnight in 2.5 volumes of absolute ethanol to precipitate the DNA. The added tRNA acted as a carrier to assist in the precipitation of the small amounts of labelled DNA. The pellets were rinsed with 500 µl of 70% ethanol and dried at room temperature. The pellets were solubilised in 5 µl of formamide dye mix (10 µl dionised formamide+ 10 µl bromophenol blue-xylene cyanol), which was used as a loading buffer and to maintain the strands of DNA in their denatured

state, by vigorous vortexing and the DNA solution was denatured by boiling for 3 minutes. The denatured DNA was placed on ice for 5 minutes. 5 µl of each sample were electrophoresed on a denaturing (7M urea) 5%, 19:1 polyacrylamide gel. The gel was separated into two pieces (upper and lower), in order to fit in the electrotransfer unit (Bio-Rad, UK) and to reduce the amount of reagents used in the detection during optimisation, and the DNA was transferred to nylon membrane and detected using the chemiluminesent EMSA kit (SeeSection 3.3.4).

6.1.1.2. Binding reaction using fluorescence-labelled primers

The procedure described using biotin-labelled primers was repeated using fluorescent labelled primers. In order to improve sensitivity, a new modified protocol using infra-red (IRDye) labelled primers was developed in this study. The new protocol was performed as follows; the region of interest of the PZ gene promoter (See table 6.1) was amplified using unlabelled amplification primers. The amplification product was purified and then used in the binding reactions. The binding reaction with the unlabelled amplified region was performed (with and without nuclear extract) with the same parameters as with the biotin-labelled amplified region (See Section 6.1.1.1). Nevertheless, different concentrations of the unlabelled amplified region, different binding buffers and different incubation conditions were used for optimisation. The different binding buffers used in the optimization were as follows: binding buffer-1 (described in section 6.1.1.1), binding buffer-2 (Cordingley and Hager 1988) (final concentration of 12 mM HEPES (pH 7.9), 0.12 mM EDTA, 60 mM KCl, 12% glycerol and 0.5 mM DTT) and binding buffer-3 (Batista *et al.* 2007) (final concentration of 50 mM Tris-HCl, pH 8.0, 100 mM KCl, 12.5 mM MgCl, 1 mM EDTA, 1 mM DTT and 20% glycerol). Having carried out the binding reaction the DNA fragment was digested with DNaseI (0.01 units) for 1 minute at room temperature, extracted by phenol:chloroform, precipitated with alcohol and air dried at room temperature (See Section 6.1.1). The DNA was then re-suspended in 10 µl of DNase free water.

In order to increase assay sensitivity the digested DNA was then amplified in a linear fashion using single primers (an IRDye 700-labelled forward primer and an IRDye 800-labelled reverse primer). The use of a forward primer in one

PCR reaction and a reverse primer in a second PCR reaction allowed the mapping of binding sites along both DNA strands. Moreover, as the primers had different fluorescent labels the ultimate aim was to develop a multiplex fluorescent footprinting technique where the amplifications and analysis could be carried out in combination.

The amplifications were carried out using the same parameters shown in table 6.2 with the exception that 5 µl (from the 10 µl purified total volume) of amplified DNA used in the binding reaction were used instead of the genomic DNA and only one primer (700 or 800 IRD) was used. Having PCR amplified the DNA fragment it was again precipitated using sodium acetate and alcohol (See Section 6.1.1) and air dried at room temperature. The pellets were re-suspended in 4 µl microSTOP loading buffer (Li-Cor, UK), vortexed and denatured by boiling at 100 °C for 3 minutes. 2 µl of each sample were electrophoresed on a denaturing (7 M urea) 5%, 19:1 polyacrylamide gel (40 x 20 x 0.2 cm) and detected using a Li-Cor 4200 sequencing system (Li-Cor, UK). A 20 base pair ladder (Li-Cor, UK) was run alongside to serve as a size marker.

6.1.2. Chromatin immunoprecipitation (CHIP) assay

The chromatin immunoprecipitation assay was performed according the manufacturer's instructions (Active Motif, Belgium). 20 ml of 1% formaldehyde were added to each of three 75 cm^2 flasks with 70-80% confluency for 10 minutes to cross-link nuclear protein to genomic DNA and prevent their dissociation in subsequent steps. This was followed by washing with 1X PBS. The cross-linking was stopped by adding a glycine stop-fix solution for 5 minutes and then washing with1X PBS. The cells were scraped from the surface of the flasks into cell scraping solution and pooled together in a 15 ml conical tube. The cells were pelleted at 747 x g resuspended in lysis buffer and incubated on ice for 30 minutes. The nuclei were released by the aid of dounce homogenisation and then pelleted at 2988 x g for 10 minutes at 4°C. The pellet was re-suspended in 1 ml of pre-warmed digestion buffer and 50 µl of enzymatic shearing cocktail (200U/ml) was added and incubated at 37°C for 15 minutes (optimal time was determined for shearing). The reaction was stopped by addition of 1 µl EDTA (0.5 M) and incubated on ice for 10 minutes. The

sheared DNA (chromatin) was centrifuged, aliquoted and stored at -80°C if not used immediately.

The sheared chromatin was pre-cleared of immunoglobulins by adding protein G beads, to reduce non-specific background, and rotated at 4°C for 2 hours. The protein G beads were then removed by centrifugation. 10 µl of pre-cleared chromatin were stored at -20°C to be used in PCR analysis. 3 µg of the antibody of interest were added to 170 µl of the pre-cleared chromatin in a 0.65 ml tube. Positive and negative antibodies (Active Motif, Belgium) were also used in separate control tubes. The reactions were incubated overnight on a rotator at 4°C to allow maximal antibody binding. The antibody-bound protein-DNA complexes were collected using protein G beads followed by incubation on a rotator for 1.5 hours at 4°C. The beads were pelleted by centrifugation at 1912 x g for 2 minutes and allowed to settle completely by leaving them to stand for 30 seconds. The beads were washed, using the proprietary wash buffers provided, by vortexing for 10 seconds on a low setting, incubating on the rotator for 3 minutes, centrifugation and removal of the supernatant without disturbing the beads. The immunoprecipitated DNA was eluted from the washed protein G beads using elution buffer containing 50 mM $NaHCO_2$ and 1% SDS. The washed protein G beads were briefly vortexed for 10 seconds with the elution buffer and incubated for 15 minutes at room temperature. The eluted DNA (supernatant) was collected by centrifuging the mixture for 2 minutes at 1912 x g and leaving the beads to settle completely.

The chromatin immunoprecipitated (eluted DNA) along with the input DNA were reverse cross-linked and RNA was removed using 5 M NaCl and RNase incubating at 65°C overnight. The proteins were digested using 0.5 M EDTA, 1 M Tris-Cl pH 6.5 and proteinase K solution with incubation at 42°Cfor 2 hours. The eluted DNA was finally purified using the provided DNA purification mini-columns (Active Motif, Belgium). This was done by adding 500 µl of the supplied binding buffer to the eluted DNA and centrifugation of the mini-column for 30 seconds at 26890 x g. The flow-through was discarded and the mini-column was replaced in the tube to be washed using the provided DNA wash buffer. The DNA was washed on the mini-column by adding the wash buffer and centrifuging for 30 seconds at 26890 x g. The flow-through was, again, discarded and the DNA was washed a second time for 2 minutes at 26890 x g.

The purified DNA was finally eluted by adding 50 µl DEPC-treated water and centrifuging for 1 minute at 26890 x g. The purified DNA was stored at -20°C to be used in subsequent PCR analysis.

PCR was used to analyse the DNA fragments obtained by ChIP. The same amount of each PCR product was size fractionated on a 5%, 19:1 polyacrylamide gel. The signal of the products was detected and the signal intensity was measured using the UVP BioImaging system (UVP, UK).

Primers used in PCR analysis were complementary to the GAPDH template (Forward primer: 5'- TACTAGCGGTTTTACGGGCG-3', Reverse: 5'-TCGAACAGGAGGAGCAGAGAGCGA-3': which produces a fragment of 166 bp spanning nucleotides -222 to -55, and flanking the transcription start site at -102 (relative to the translation initiation site), as a positive control (Active Motif, Belgium) and the protein Z template (PZ -13/-163) primers (Forward primer: 5'-TCTGAGCCTTCACCGTTCATT-3', Reverse primer: 5-TGCTGGGACAGGGCTACAAA-3') designed using webprimer (http://www.yeastgenome.org/cgi-bin/web-primer) and synthesised by Eurofin MWG \ Operon.

6.1.3. Analysis

In the footprinting assay the samples electrophoresed on the denaturing polyacrylamide gel were visualised using the UVP Bioimaging system (UVP, UK), for the biotin-labelled probes, and the Li-Cor 4200 sequencing system (Li-Cor, UK), for the IRDye-labelled probes. A 20 base pair ladder (Li-Cor, UK) was run alongside the IRDye-labelled fragments to serve as a size marker and determine the location of protein-DNA binding. Moreover, a DNase I digested naked DNA (without the addition of nuclear extract) was also run along side to serve as a negative control.

In the ChIP assay the PCR products from the immunoprecipitated and input chromatins were visualised and their intensity were determined using the UVP bioimaging system. The intensity of the PCR product from the immunoprecipitated chromatin was compared with the input chromatin to determine the enrichment of the candidate region with the proposed antibody.

6.2. Results

6.2.1. Footprinting with biotinylated primers

The different regions of the PZ promoter were amplified by the respective biotinylated primers and PCR reagents (See Section 6.1.1) using the Primus 96 plus PCR machine (MWG-BIOTECH, UK). PCR of the FP1, FP2 and FP3 regions showed optimum amplified product using 1 mM $MgCl_2$ while 2 mM for the FP4 (Fig 6.4). The annealing temperatures were 60°C, 58°C, 57 and 62.5°C for FP1, FP2, FP3 and FP4, respectively.

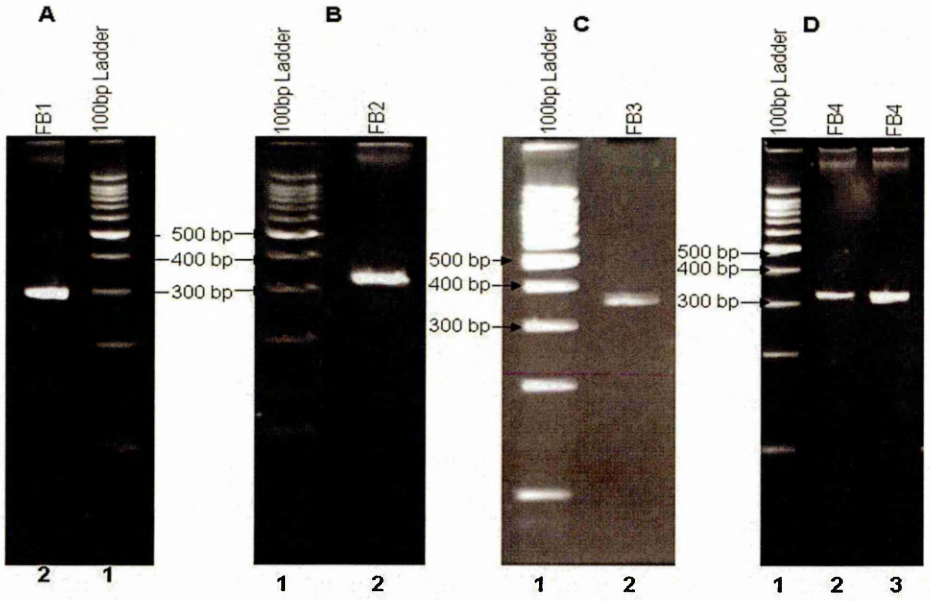


Fig 6.4. PCR of different regions of PZ gene promoter used for footprinting analysis. Number 1 in all images represents the 100bp ladder (Promega, UK). Number 2 in the images A, B, C and D represents the amplified fragments fp1 (+86 to -213), fp2 (-166 to -494), fp3 (-455 to -804), fp4 (-785 to -1099), respectively. The optimal concentration of $MgCl_2$ in the PCR reactions for fp1, fp2 and fp3 was 1mM. Although both 1mM and 2mM $MgCl_2$ showed good product with fp4, the 2mM was better than 1mM. Numbers 2 and 3 in the image D shows the PCR fragments of fp4 with 1mM and 2mM, respectively.

Footprinting using the region +86 to -213 with different amounts of nuclear extract showed a uniform ladder of the DNA as well as a lack of accumulated DNA in the upper or lower gel, indicating the appropriate amount of DNase I and time of digestion had been used . However, no protein binding was observed (Fig 6.5.A and B). A freshly prepared nuclear extract was used in another footprinting assay, but again no protein binding to the DNA was observed (Fig 6.5 C and D).

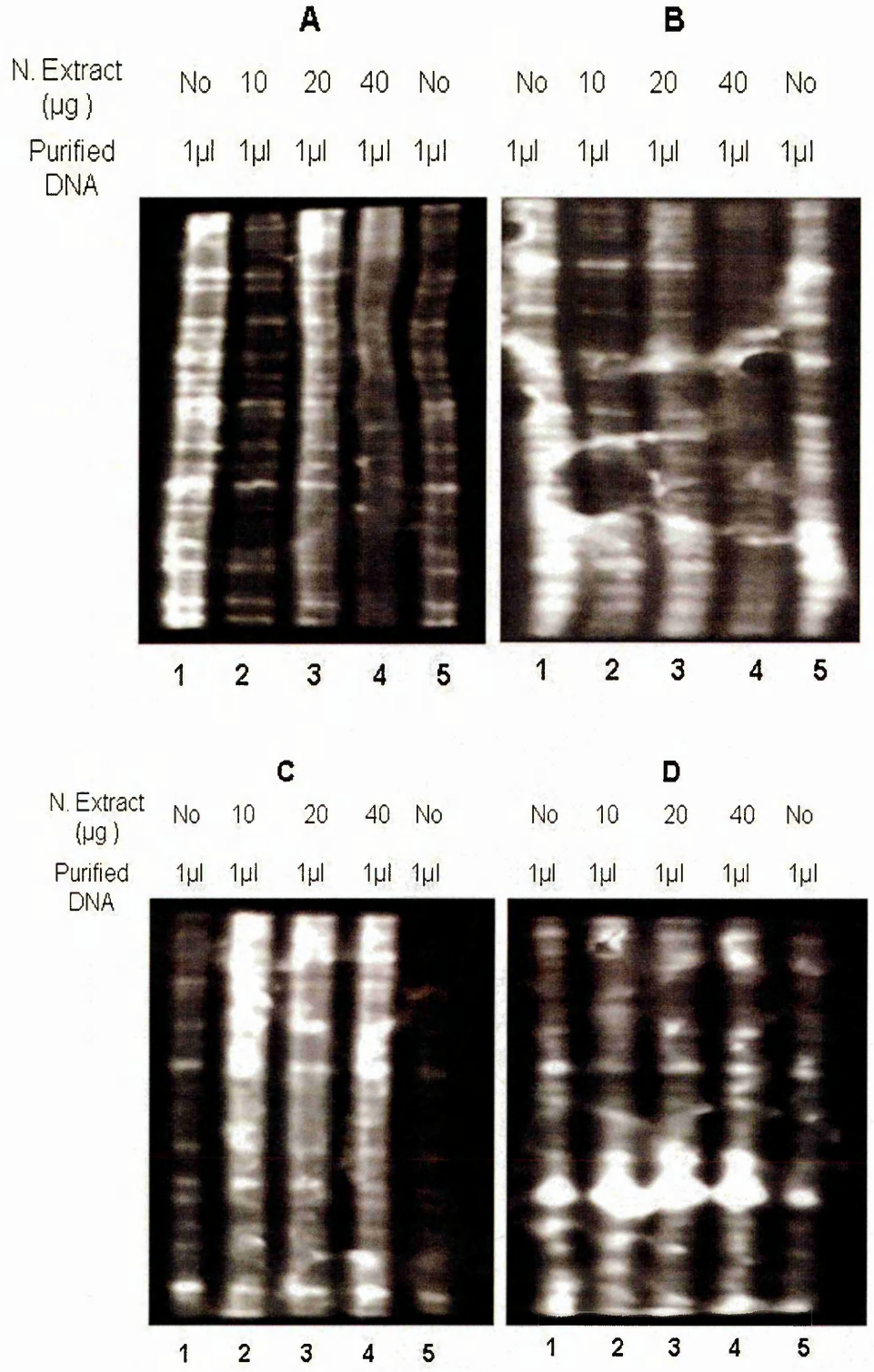


Fig 6.5. Footprinting of the Protein Z promoter region +86 to -213. Lanes 1 and 5 contain the DNA fragment without nuclear extract (control). Lanes 2, 3 and 4 contain the DNA fragment with 10µg, 20µg and 40 µg of HepG2 NE and 15 µl, 10 µl and no binding buffer, repectively (25 µl total volume). [A] is an upper region of the gel. [B] is the bottom region of the gel. No protein was found to bind the DNA fragment. [C] and [D] is an upper and lower region of the gel from another experiment for the same promoter region with the same parameters but with freshly prepared HepG2 NE.

Protein binding to DNA should be seen as a gap (footprint) in the DNA ladder, due to the protection of the region bound by the protein from the DNase I digestion. Another footprinting experiment was performed using a 20 minute binding incubation time at room temperature instead of on ice. Moreover, the amount of DNA template used was 0.5 µl instead of 1.0 µl. The results showed uniform DNA bands due to DNaseI digestion, but again did not show any protein binding to the DNA fragment as no footprinting was detected (Fig 6.6).

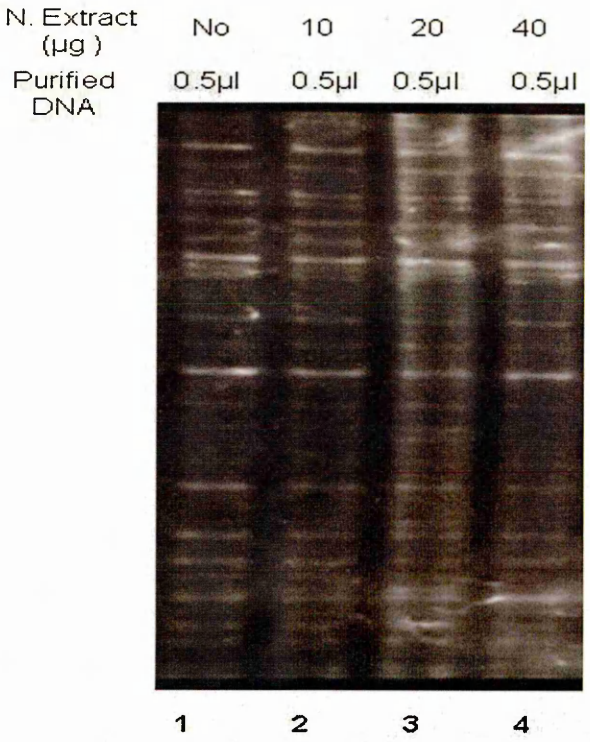


Fig 6.6. Footprinting of the PZ region +86 to -213 with incubation at room temperature. Lane 1 contains the DNA fragment without HepG2 NE (control). Lanes 2, 3 and 4 contain the DNA fragment with 10µg, 20µg and 40 µg HepG2 NE. The binding reaction mixture was incubated for 20 minutes at room temperature instead of ice. Approximately 0.5 µl of the purified DNA fragment was used.

6.2.2. Footprinting with IRD primers

The first footprinting experiment with the IRD primers was performed using the previous protocol (See Section 6.1.1.1). Ten fmole of the 700 and 800 IRD-labelled amplified fragments were incubated for 20 minutes on ice with different amounts (10, 20 and 40 µg) of dialysed HepG2 NE (SeeSection 3.3.1). The results showed that no visible fragments were detected at either 800 nm or 700 nm apart from the molecular weight ladder (Fig 6.7).

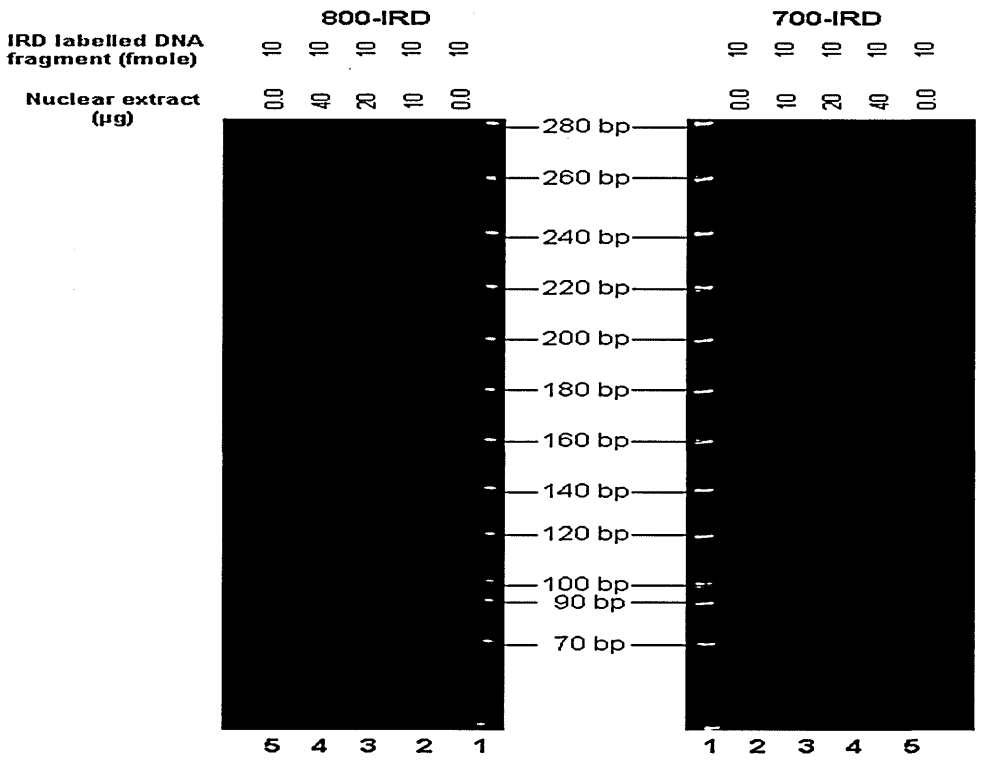


Fig 6.7. Footprinting assay of the PZ region +86 to -213 without amplification using IRD-labelled fragment. Left hand image represents the assay with the 800-IRD fragment while the right hand image is the assay with the 700-IRD fragment. Lane 1 is a 20 bp ladder. Lane 2 (control) is the 10 fmole IRD-labelled fragment without HepG2 NE. Lanes 3, 4 and 5 are 10 fmole of IRD-labelled DNA fragment with 10, 20 and 40 µg of dialysed HepG2 NE, respectively. No signal was detected at either 700 nm or 800 nm.

The footprinting assay performed by the modified protocol (See Section 6.1.1.2) using different concentrations of the unlabelled FP1 fragment and different concentrations of dialysed HepG2 NE showed too high a signal intensity when amplified by the 700-IRD forward primer (Fig 6.8). However, lower signal intensity was detected at 800 nm when the 800-IRD reverse primer was used for the post amplification (Fig 6.9). The assay using 0.1 and 0.5 fmole of the +86 / -213 unlabelled fragment followed by 50X post amplification with the 800-IRD reverse primer showed a uniform intensity of digested bands. Nevertheless, no protein-DNA binding (footprinting) was detected.

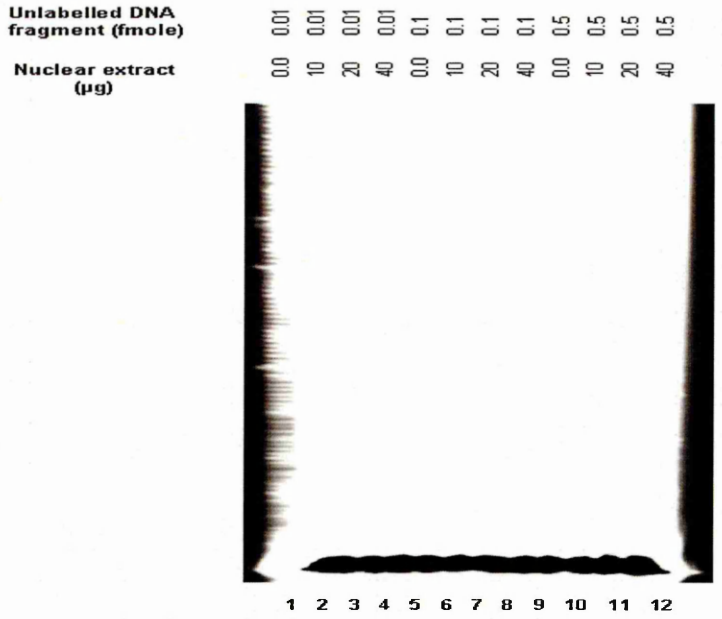


Fig 6.8. Footprinting assayof the PZ region +86 to -213 post amplified by the 700-IRD primer. The image shows a too high signal, detected at 700 nm, in all lanes with different concentrations of the DNA fragment and the dialysed HepG2 NE. However, the DNA fragment seems to have uniform digestion by the DNase I.

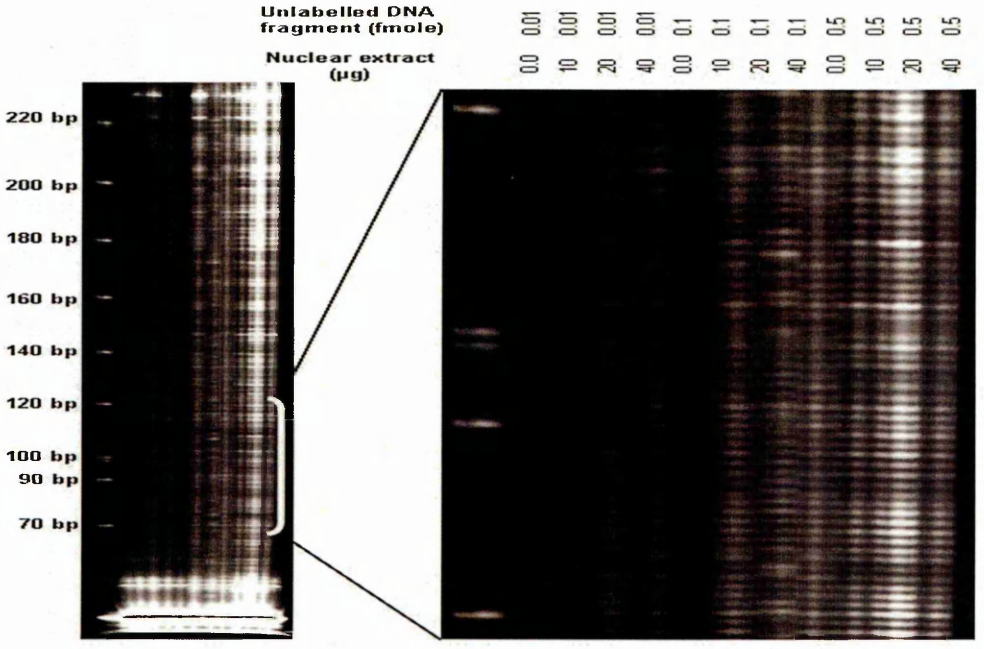


Fig 6.9. Footprinting assay ofr the PZ region +86 to -213 post amplified by the 800-IRD primer. The left hand image represents most of the DNA fraction with the 20 bp single-stranded ladder. The right hand image represents a magnified 120 bp section of the gel. The gel shows uniform bands as a result of DNaseI digestion. The ladder resulting from 0.01 fmole of DNA in the binding reaction showed a weak signal while better signals were shown by using 0.1 and 0.5 fmole of the DNA. However, no protein binding was shown with the different concentrations of the DNA using different concentrations of the dialysed HepG2 NE.

The same assay was repeated with the same concentrations of the +86 / -213 unlabelled fragment using 40 µg of the dialysed HepG2 NE followed by a 1:10 dilution of the product, using the loading buffer, after the post-treatment amplification with the candidate IRD primer. The results showed too high a signal detected at 700 nm from the undiluted samples with the different concentrations of the region of interest. However, samples which were diluted 1:10 after the post-treatment amplification showed relatively uniform signal with the use of 0.5 fmole of the unlabelled region and relatively weak signal with the use of 0.1 and 0.01 fmole (Fig 6.10). The experiment again did not show any protein-DNA binding.

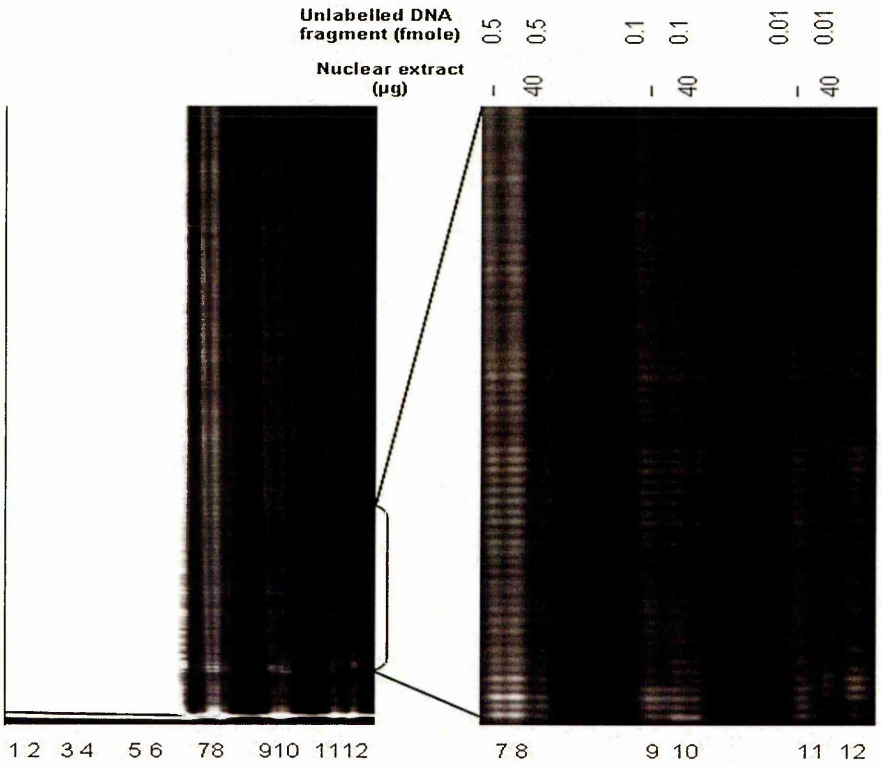


Fig 6.10. Footprinting assay of the PZ region +86 to -213 post amplified by the 700-IRD primer and with and without a 1:10 dilution. The left hand image represents most of the digested DNA fraction with a 20 bp single-stranded size standard. The right hand image represents a magnified DNA fragment, indicated by the bracket to the right of the left hand image. Lanes 1 to 6 contain different concentrations of the unlabelled region and show too high a signal after the post amplification step when loaded without dilution. The 20 bp ladder was hidden by the very strong signal. Lanes 7 to 12 show relatively good signal using 0.5 fmole of the proposed region in the binding reaction followed by the post amplification step and a 1:10 dilution with the loading buffer. The use of 0.1 and 0.01 fmole of the proposed region with the 1:10 dilution after the post amplification step showed relatively weak signal. Lanes 1, 3, 5, 7, 9 and 11 are controls which contain the DNA fragment without the HepG2 NE, while lanes 2, 4, 6, 8, 10 and 12 are the DNA fragment incubated with the HepG2 NE for 20 minutes on ice.

The results detected at 800 nm showed a relatively good signal without the 1: 10 dilution after the post amplification step, although some of the bands were quite weak. However, the dilution after post amplification resulted in loss of the signal (Fig 6.11). The results also showed some cross-lane contamination due to leakage from the sharkstooth comb.

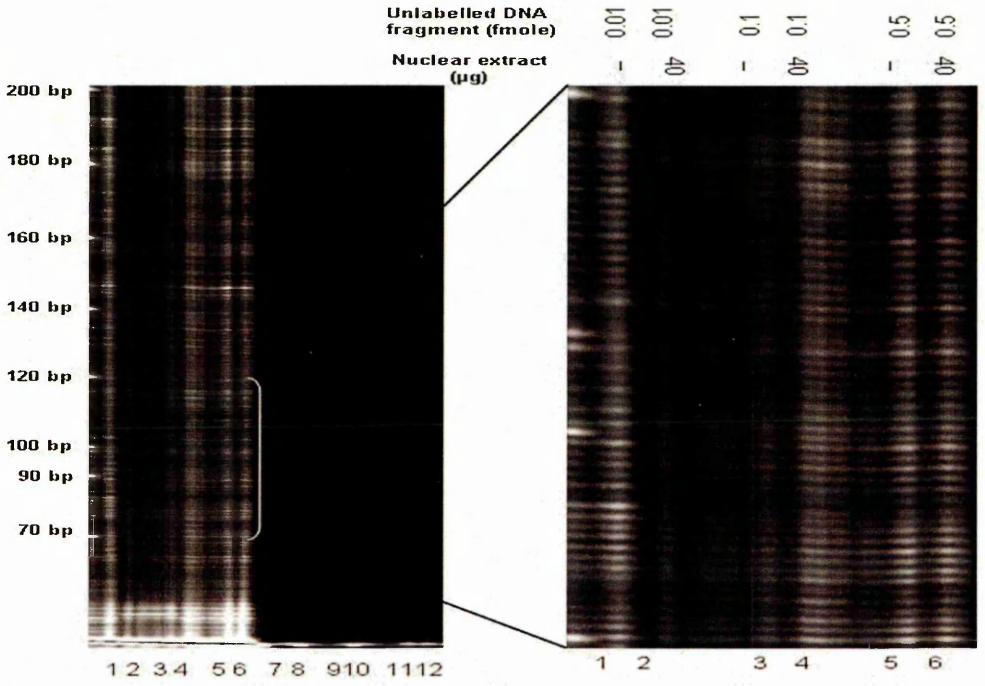


Fig 6.11. Footprinting assay of the PZ region +86 to -213 post amplified by the 800-IRD primer with and without 1:10 dilution. The left hand image represents most of the digested DNA fraction and the 20 bp single-stranded size standard. The right hand image represents a magnified DNA fragment which indicated by the right bracket on the left handed image. Lanes 1 to 6 contained different concentrations of the unlabelled region and showed a relatively weak to good signal after the post amplification step of the digested DNA and loading without dilution. Lanes 7 to 12 showed a loss of the signal using different concentrations of the proposed region in the binding reaction followed by the post amplification step and 1:10 dilution with the loading buffer. Lanes 1, 3, 5, 7, 9 and 11 were controls which contained the DNA fragment without the HepG2 NE while lanes 2, 4, 6, 8, 10 and 12 were the DNA fragment incubated with the HepG2 NE for 20 minutes on ice.

The same experiment was repeated with an incubation of the HepG2 NE for 30 minutes on ice with the different amounts of the unlabelled fragment (+86/-213) followed by 1: 5 dilution of the DNA fragments amplified by the 700-IRD forward primer and 1: 2 dilution of the DNA fragments amplified by the 800-IRD reverse primer after the digestion and purification. The results again showed too high a signal detected at 700 nm (Fig 6.12) while a relatively good signal was detected at 800 nm (Fig 6.13). No protein-DNA binding was detected.

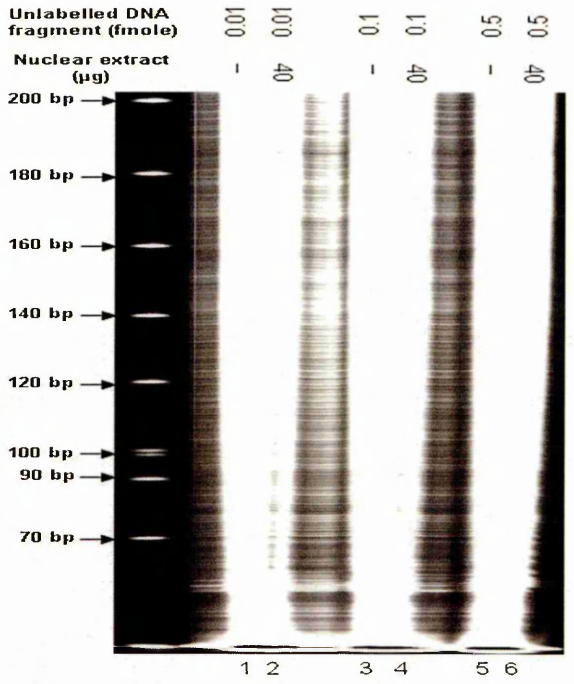


Fig 6.12. Footprinting assay of the PZ region +86 to -213 post amplified by 700-IRD primer and 1:5 dilution. Lanes 2, 4 and 6 contain different concentration of the unlabelled region with 40 µg of the HepG2 NE while lanes 1, 3 and 5 are controls without the HepG2 NE. All the different concentrations of the DNA fragment again showed a too high signal after the post amplification step of the digested DNA and 1:5 dilution with the loading buffer. The protein-DNA binding could not be determined due to the very high signal intensity.

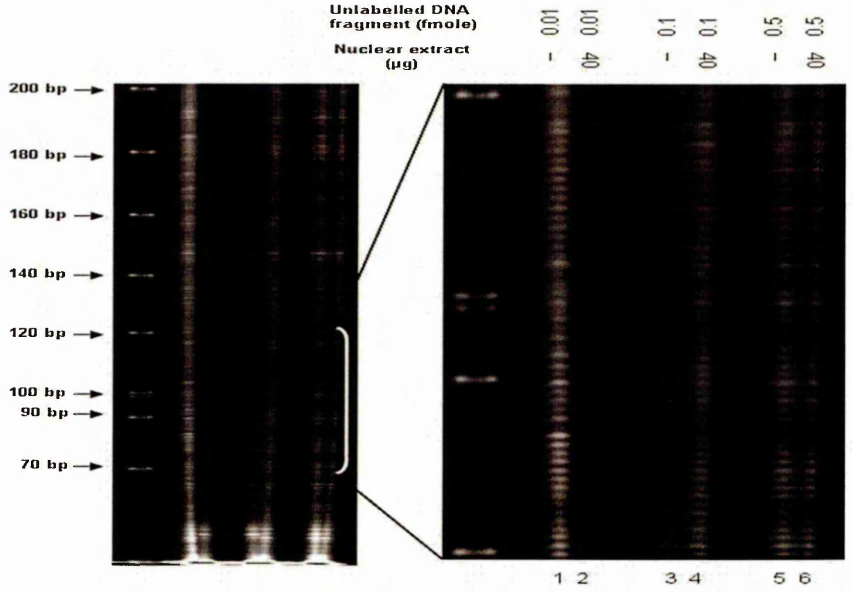


Fig 6.13. Footprinting assay of the PZ region +86 to -213 post amplified by the 800-IRD primer and 1:2 dilution. Lanes 2, 4 and 6 contain different concentrations of the unlabelled region with 40 µg of the HepG2 NE while lanes 1, 3 and 5 are controls without the HepG2 NE. All the different concentrations of the DNA fragment again showed relatively good signal after the post amplification step of the digested DNA and 1:2 dilution with the loading buffer. However, the binding reaction using 0.01 fmole of the DNA fragment showed a very weak signal. No protein-DNA binding was determined.

Another footprinting assay using different incubation times with DNaseI, 90 seconds and 2 minutes, was carried out. The fragment amplified with the 700-IRD forward primer was diluted 1:10 before loading on the denaturing gel while the fragment amplified with the 800-IRD reverse primer was diluted 1:2. The digestion of the DNA fragment with the DNaseI for 90 seconds and 2 minutes again showed a relatively uniform ladder of the fragments which was similar to that seen with the digestion with the DNaseI for 1 minute. The ladder produced from the DNA fragment (+86/-213) with the DNaseI digestion showed relatively good signal detected at both 700 nm and 800 nm (Fig 6.14 and Fig 6.15). However, some of the lanes showed very high signal while others showed very weak signal. The assay again did not show any protein-DNA binding.

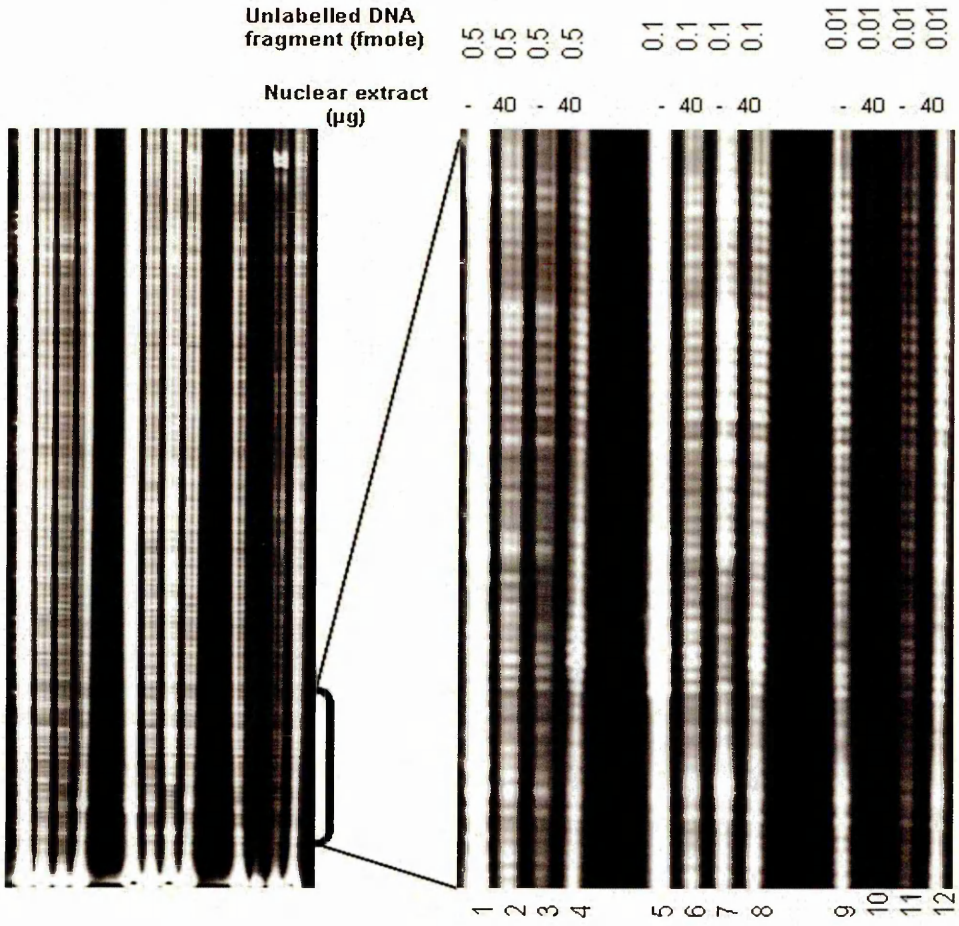


Fig 6.14. Footprinting assay of the PZ region +86 to -213 post amplified by the 700-IRD primer and 1:10 dilution. The right hand image represents a magnified gel fragment indicated by the bracket to the right of the left hand image. Lanes 1, 2, 5, 6, 9 and 10 show the DNA fragment incubated with DNaseI for 90 seconds while lanes 3, 4, 7, 8, 11 and 12 were incubated with the DNaseI for two minutes. Lanes 1, 3, 5, 7, 9 and 11 represent the controls without the HepG2 NE. The lanes show uniform ladders of the DNA fragment with both incubations with the DNaseI. However, no footprints were observed.

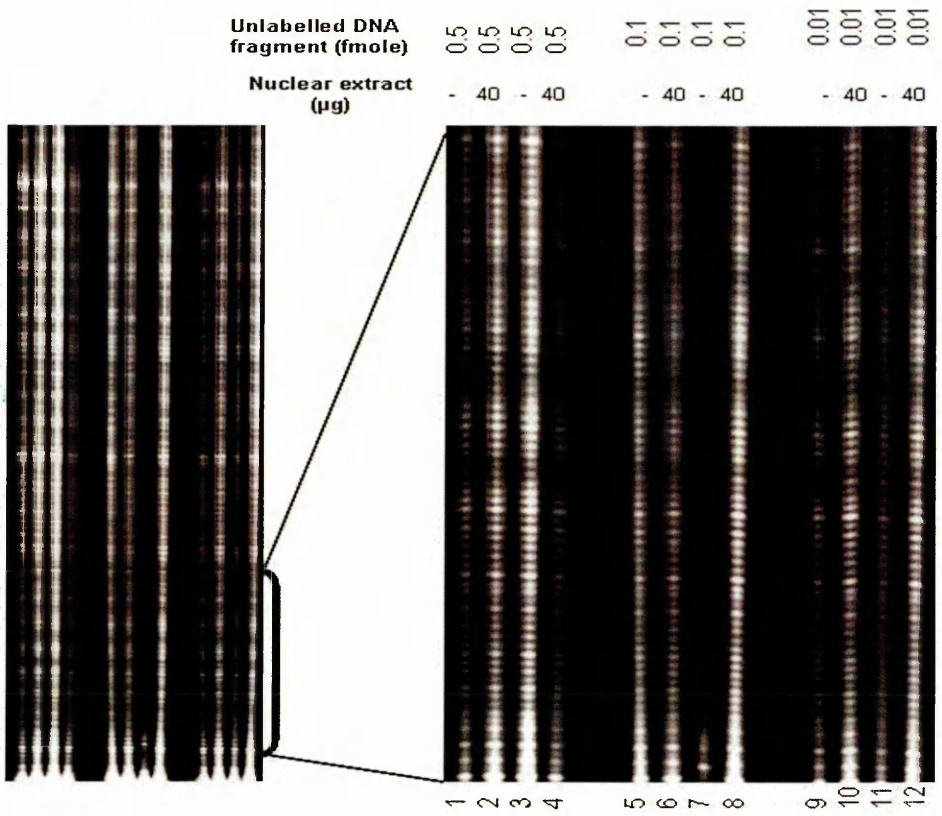


Fig 6.15. Footprinting assay of the PZ region +86 to -213 post amplified by 800-IRD primer and 1:2 dilution. The right hand image represents a magnified gel fragment indicated by the bracket to the right of the left hand image. Lanes 1, 2, 5, 6, 9 and 10 show the DNA fragment incubated with DNaseI for 90 s while lanes 3, 4, 7, 8, 11 and 12 were incubated with DNaseI for two minutes. Lanes 1, 3, 5, 7, 9 and 11 represent the controls without the Hep NE. The lanes show uniform ladders of the DNA fragment with both incubations with the DNaseI. However, no footprints were shown.

The footprinting assay was again carried out using three different binding buffers (See Section 6.1.1.2). 0.5 fmole of the unlabelled fragment incubated for 45 minutes on ice with 30 and 40 µg were used for the binding reaction. After the DNaseI digestion for one minute and the post amplification with the 700-IRD forward primer and post amplification with the 800-IRD reverse primer, 1:10 and 1:2 dilutions of the 700-IRD fragment and the 800-IRD fragment, respectively, were loaded on the denaturing gel. The results again showed a uniform ladder produced from the digested fragment although different signal intensities through the lanes were detected at both 700 and 800 nm with the best result observed with binding buffer-1. Again, no protein-DNA binding was detected (Fig 6.16 a and b).

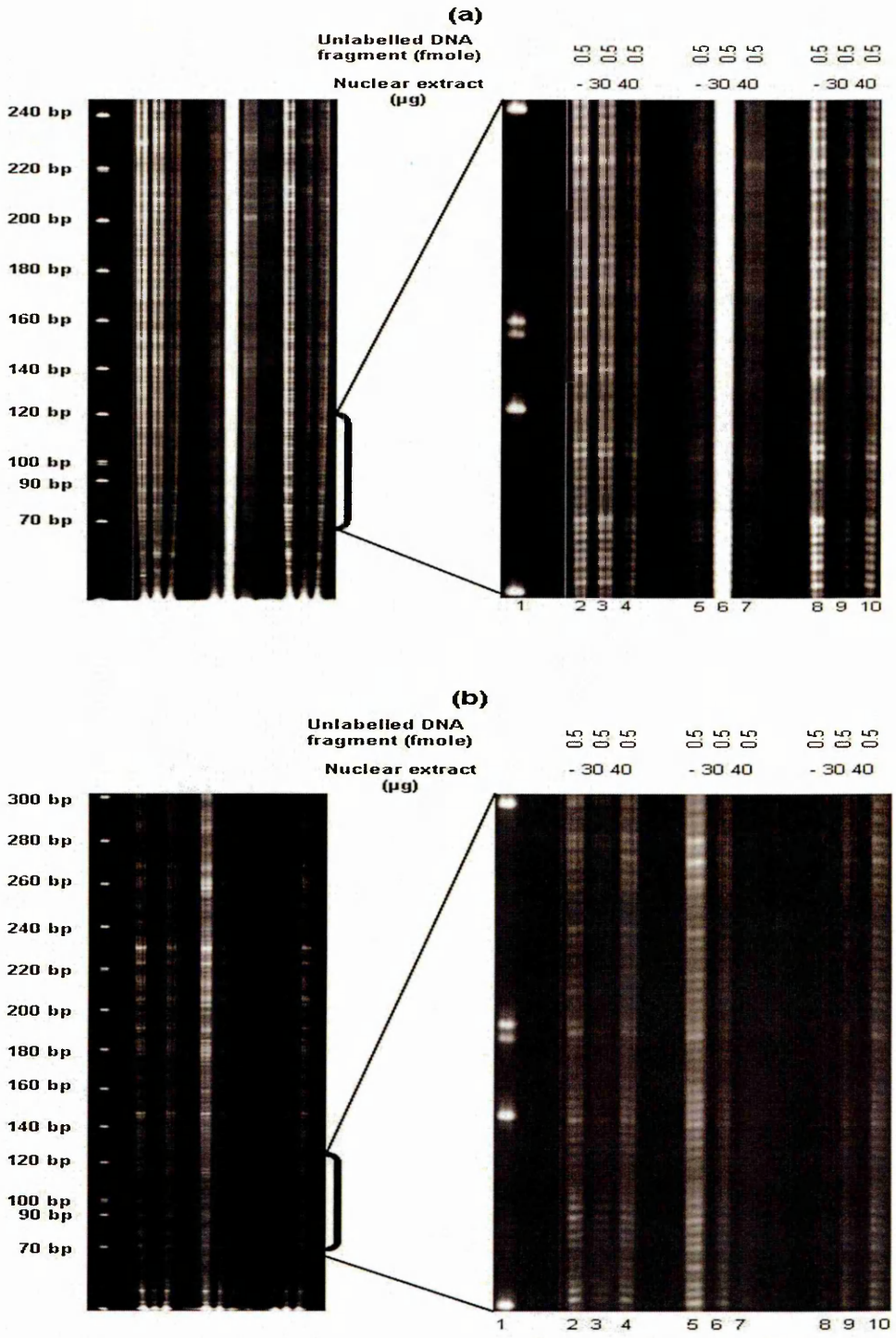


Fig 6.16. Footprinting assay using different binding buffers for the PZ region +86 to -213. (a) represents the digested fragment amplified by the 700-IRD forward primer. **(b)** represents the digested fragment amplified by the 800-IRD reverse primer. The right hand image represents a magnified gel fragment which is indicated by the bracket to the right of the left hand image. Lane 1 was a 20 bp ladder. Lanes 2, 5 and 8 were controls without Hep NE. Lanes 3 and 4 were 0.5 fmole of the fragment incubated for 45 minutes on ice with 30 µg and 40 µg of HepG2 NE, respectively, in binding buffer-1. Lanes 6 and 7 were 0.5 fmole of the fragment incubated for 45 minutes on ice with 30 µg and 40 µg of HepG2 NE, respectively, in binding buffer-2. Lanes 9 and 10 were 0.5 fmole of the fragment incubated for 45 minutes on ice with 30 µg and 40 µg of HepG2 NE, respectively, in binding buffer-3. No footprints were detected.

The footprinting assay using different binding buffers (See Section 6.1.1.2) was repeated with different incubation conditions (4°C, room temperature and 37°C) for the proposed DNA fragment with the Hep NE (Fig 6.17). The use of different incubation condition and different binding buffers again did not show any footprints.

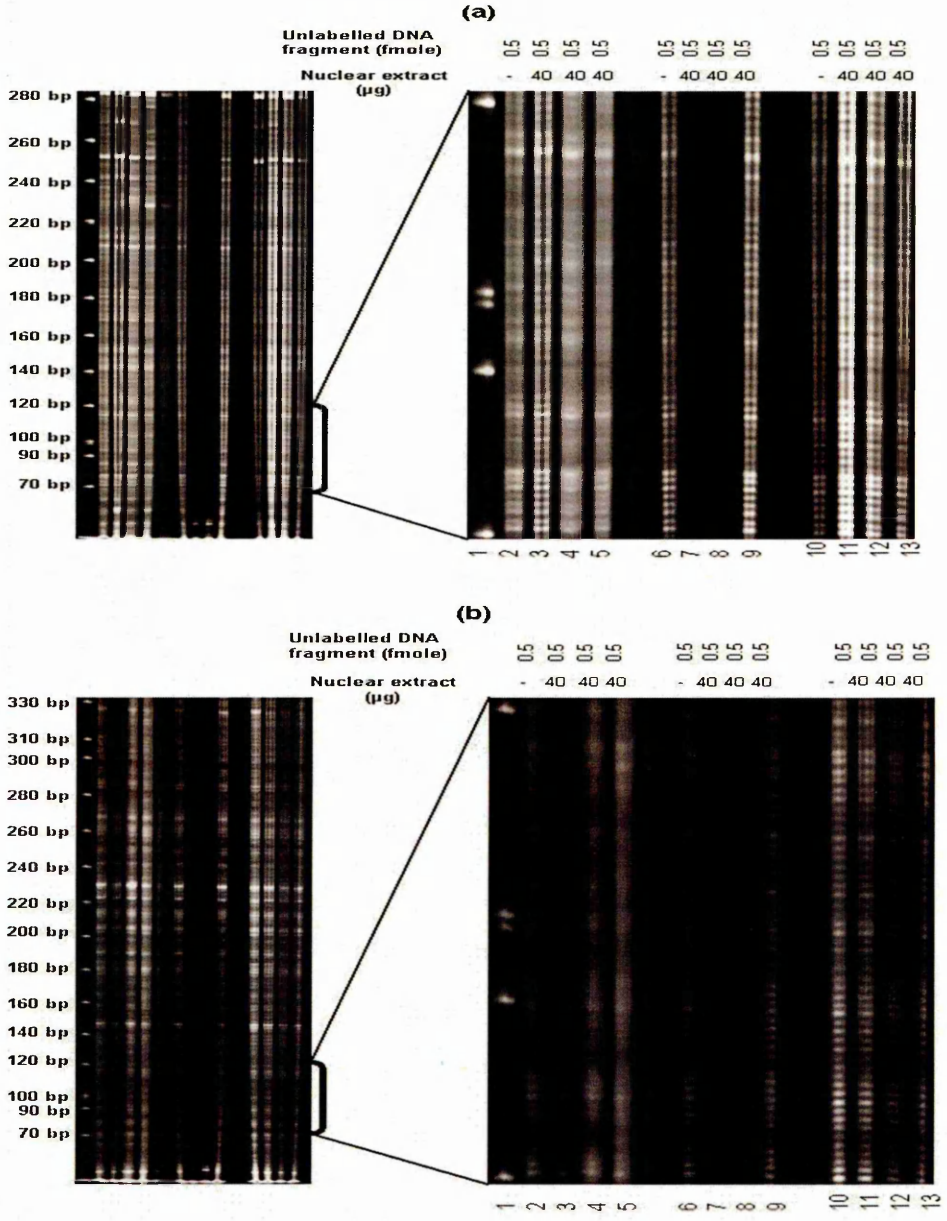


Fig 6.17. Footprinting assay using different binding buffers and different incubation conditions. (a) represents the digested fragment (+86/-213) amplified by the 700-IRD forward primer. **(b)** represents the fragment amplified by the 800-IRD reverse primer. Ther right hand image represents a magnified gel fragment which is indicated by the bracket to the right of the left hand image. Lane 1 was a 20 bp ladder. Lanes 2, 6 and 10 were controls without Hep NE. Binding buffer-1 was used in lanes 2, 3, 4 and 5. Binding buffer-2 was used in lanes 6, 7, 8 and 9. Binding buffer-3 was used in lanes 10, 11, 12 and 13. The DNA fragment was incubated with the Hep NE overnight at 4°C in lanes 3, 7 and 11 and at room temperature for 45 minutes in lanes 4, 8 and 12 and at 37°C for 45 minutes. No footprints were detected under all conditions.

6.2.3. Chromatin Immunoprecipitation

6.2.3.1. Optimisation of GAPDH primers for PCR

GAPDH primers were provided with the ChIP kit (Active Motif, Belgium). These act as a positive control in that they flank the start point of transcription of the GAPDH gene and should show enrichment when the chromatin is immunoprecipitated with the RNA pol II antibody. PCR conditions for this primer pair were optimised using genomic DNA prepared from HepG2 cells. The GAPDH primers were used to demonstrate the efficacy of the protocol and the reagents that were provided with the kit. The optimised PCR should show a specific product of 166 bp in size. The PCR of the GAPDH primers using different temperatures and different concentrations of $MgCl_2$ did not show its optimal specific product (Fig 6.18). The PCR of GAPDH primers at 58°C showed two bands of high molecular weight products (Fig 6.18,A). The temperature was then increased to 59°C and 1mM and 2mM MgCl used in the PCR. The results showed many non-specific amplification products (Fig 6.18,B). Therefore, the temperature, again, was increased to 60°C and different concentrations of MgCl (1mM and 3mM) were used in order to increase specificity of primer binding. The results, again, showed two high molecular weight products but not the specific 166 bp product (Fig 6.18,C).

6.2.3.2. Optimisation of PZ -13/-163 region primers

The primers for the region of interest in the PZ gene promoter (PZ -13/-163), which spans nucleotides -13 to -163 relative to the translation start site, were optimised for amplification. This region has been found, in the current study, to be bound by HNF4α and gave some indication of PEA3 or Ets 1 binding using EMSA analyses. The optimisation reactions were performed using the input DNA cross-linked, sheared, reverse cross-linked and purified at different temperatures and using different concentrations of $MgCl_2$.

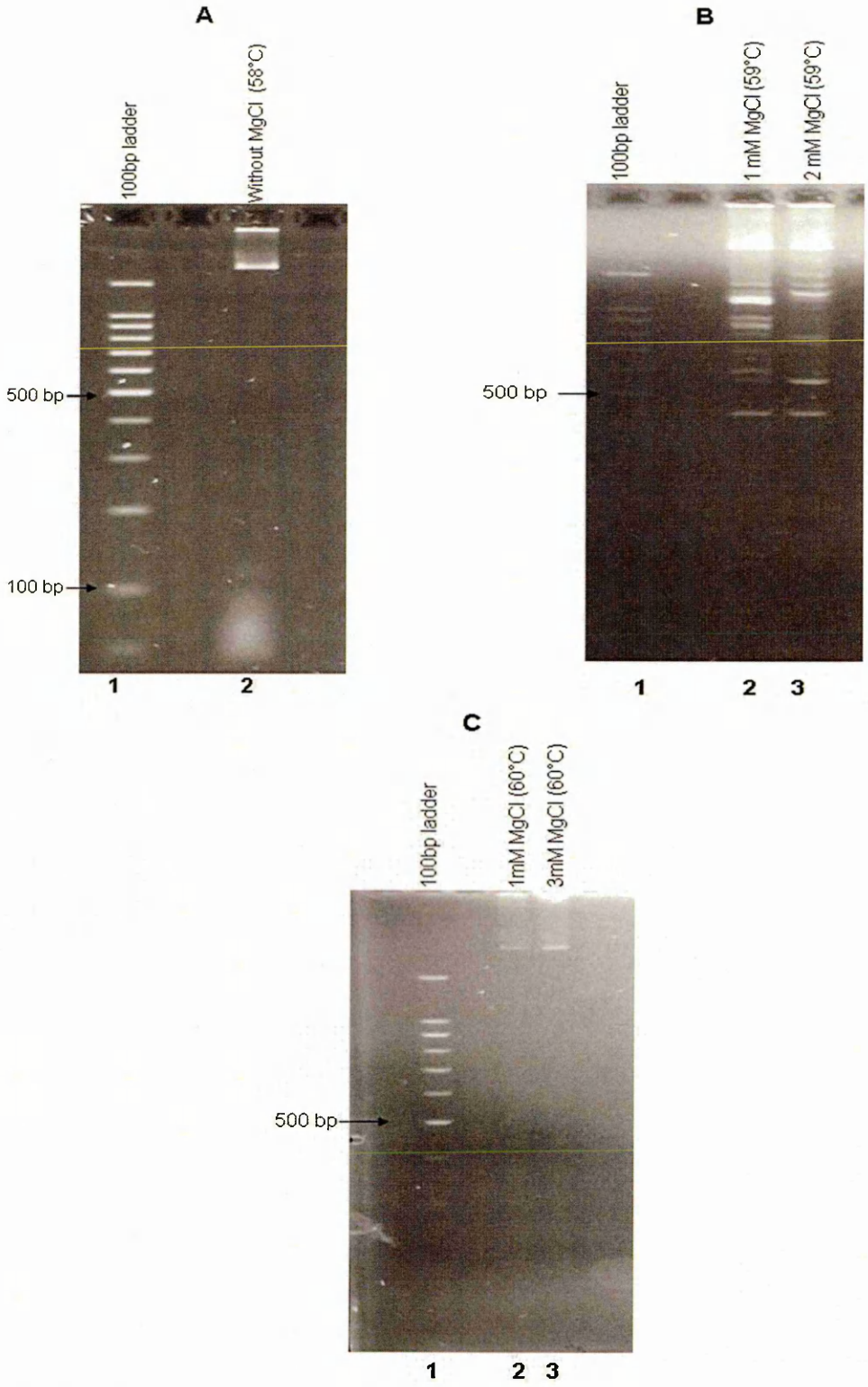


Fig 6.18. Optimisation of the GAPDH PCR with HepG2 genomic DNA. Lane 1 in all images represented a 100 bp ladder. Image [A] lane 2 shows high molecular weight PCR products at 58°C without adding $MgCl_2$. Image [B] shows PCR at 59°C using 1 mM and 2 mM $MgCl_2$ in lane 2 and 3 respectively. Multiple non-specific bands were shown with both concentrations. Image [C] shows PCR at 60°C using 1 mM and 3 mM $MgCl_2$ in lane 2 and 3, respectively. Two high molecular weight products were shown with both mM $MgCl_2$.

The first PCR reaction was carried out at an annealing temperature of 59°C using 1mM and 2mM $MgCl_2$. The results show the specific product along with additional non-specific products (Fig 6.19). The PCR was repeated with higher annealing temperatures (61°C, 62°C and 63°C) and with different concentrations of $MgCl_2$. The PCR results, again, showed non-specific products along with the specific product with annealing temperatures 61°C and 62°C (Fig 6.20 and Fig 6.21). However, with an annealing temperature of 63°C no products were amplified (Fig 6.22). Therefore, the annealing temperature was decreased to 62.5°C and another PCR was performed with 1mM $MgCl_2$, since it showed the best specific product with the previous temperatures. The specific product amplified by PCR for the PZ -13/-163 region at annealing temperature 62.5 °C and 1 mM $MgCl_2$ was found to be the most specific out of the tested variables (Fig 6.23). The size of the specific product was confirmed using a standard curve of the migration of the molecular weight ladder. This was done by measuring the migration distance of the PCR product to find out its size.

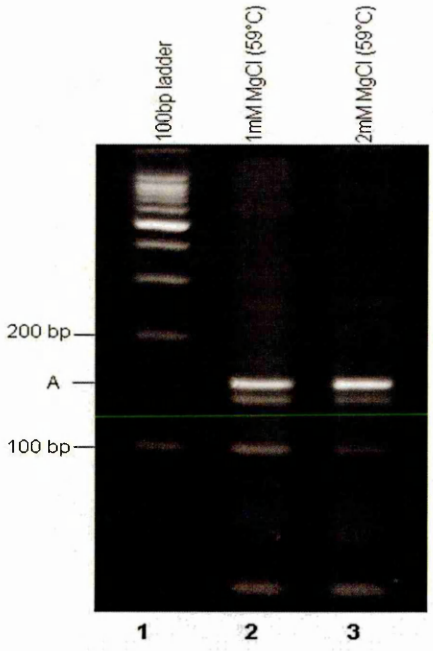


Fig 6.19. PCR reaction of PZ -13/-163 with input DNA (annealing temperature of 59°C). Lane 1 is a 100 bp ladder. Lane 2 and 3 are the reaction with 2 mM and 3 mM $MgCl_2$, respectively. Non-specific products were detected along with the specific product (A) of 151 bp.

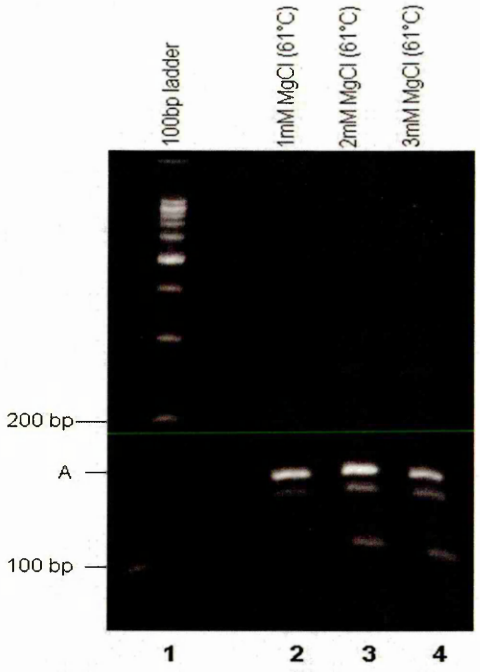


Fig 6,20. PCR reaction of PZ -13/-163 with input DNA (annealing temperature of 61°C). Lane 1 is 100 bp ladder. Lane 2, 3 and 4 are the reactions with 1, 2 and 3 mM $MgCl_2$ respectively. Non-specific products were detected with the specific product (A) of 151 bp.

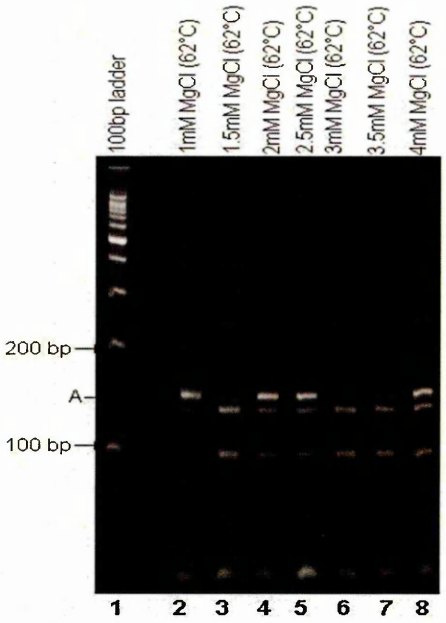


Fig 6.21. PCR reaction of PZ -13/-163 with input DNA (62°C). Lane1 is 100 bp ladder. Lanes 2, 3, 4, 5, 6, 7 and 8 are the reactions with 1, 1.5, 2, 2.5, 3, 3.5, 4.0 mM $Mgcl_2$, respectively. Non-specific products were detected along with the specific product (A) of 151 bp.

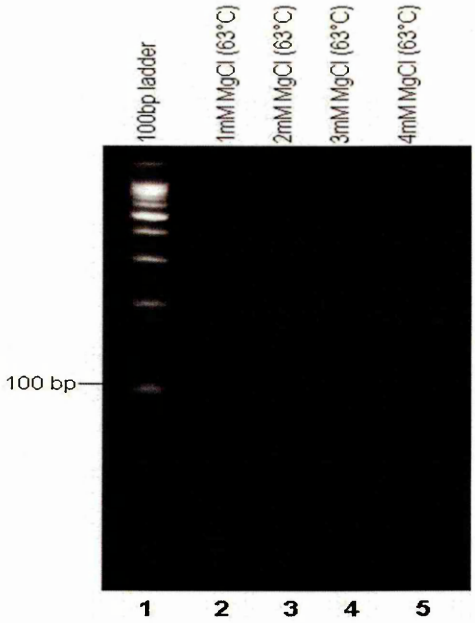


Fig 6.22. PCR reaction of PZ -13/-163 with input DNA (63°C). Lane1 is 100 bp ladder. Lane 2, 3, 4 and 5 are the reactions with 1, 2, 3 and 4 mM $Mgcl_2$ respectively. No amplified product was detected.

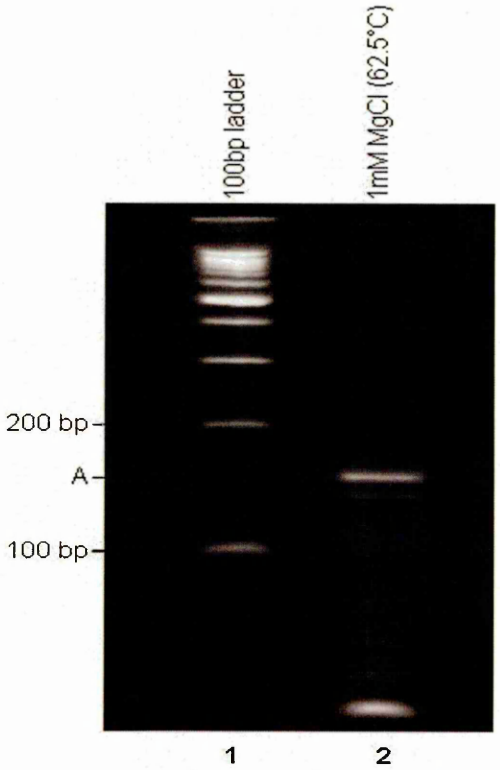


Fig 6.23. PCR of PZ -13/-163 with input DNA (62.5°C). Lane 1 is the 100 bp ladder. Lane 2 is the reaction with 1 mM $MgCl_2$. Band A indicates the specific product. These parameters were found to be the most specific in amplifying the PZ -13/-163 region with PCR.

6.2.3.3. PCR of the PZ -13/-163 with the immunoprecipitated chromatin.

PCR was performed with the primers of the region of interest spanning nucleotides -13 and -163, relative to the translation start site (+1). The optimised PCR (62.5°C annealing temperature and 1mM $MgCl_2$) was performed using the DNA (chromatin) immunoprecipitated by HNF4α antibody alongside the DNA immunoprecipitated by the IgG antibody (negative control), the DNA immunoprecipitated by the pol II antibody (positive control) and the input DNA. An amplified no template PCR negative control and a 100bp ladder were also run on the gel. A band of weaker intensity (20% lower) than the input DNA was observed with the HNF4α immunoprecipitated DNA suggesting no enrichment. This product was also of lower intensity (30% lower) than with the DNA immunoprecipitated by the pol II antibody (Fig 6.24). The no template negative controls showed contamination (Lanes 2 and 3) with the presence of the same product as with the other DNA templates.

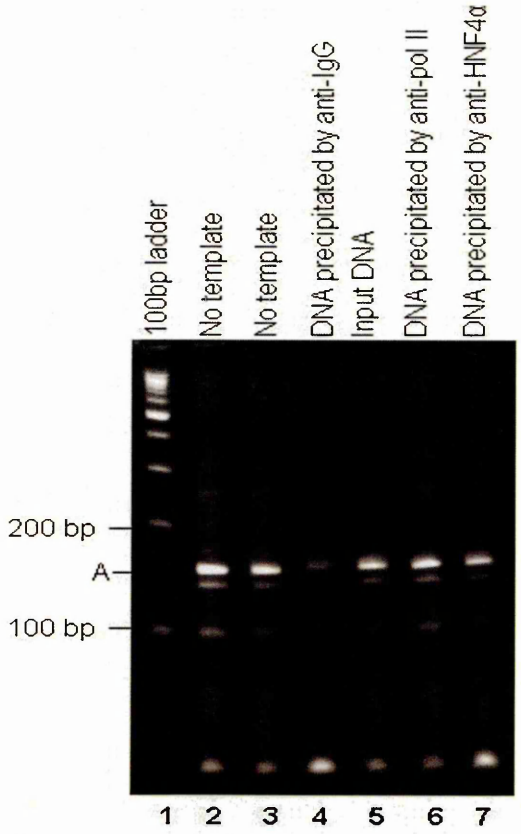


Fig 6.24. PCR of PZ -13/-163 with different immunoprecipitated chromatins. Lane 1 is a 100 bp ladder. Lanes 2 and 3 are no template negative controls containing DEPC and DNase free water, respectively. Lane 4 contains DNA immunoprecipitated by IgG. Lane 5 contains input DNA. Lane 6 contains DNA immunoprecipitated by the pol II antibody. Lane 7 contains DNA immunoprecipitated by the HNF4 antibody. [A] indicates the desired product (151 bp).

The same experiment was repeated using new aliquots of 10X PCR buffer and the primers to isolate the contamination. Stronger signals of the proposed product were detected with both pol II (1.8-fold) and HNF4α (2.2-fold) enriched DNA compared to the negative control (Fig 6.25). However, no signal was detected with input DNA and the no template control, again, showed contamination although it was of lower intensity than the previous experiment.

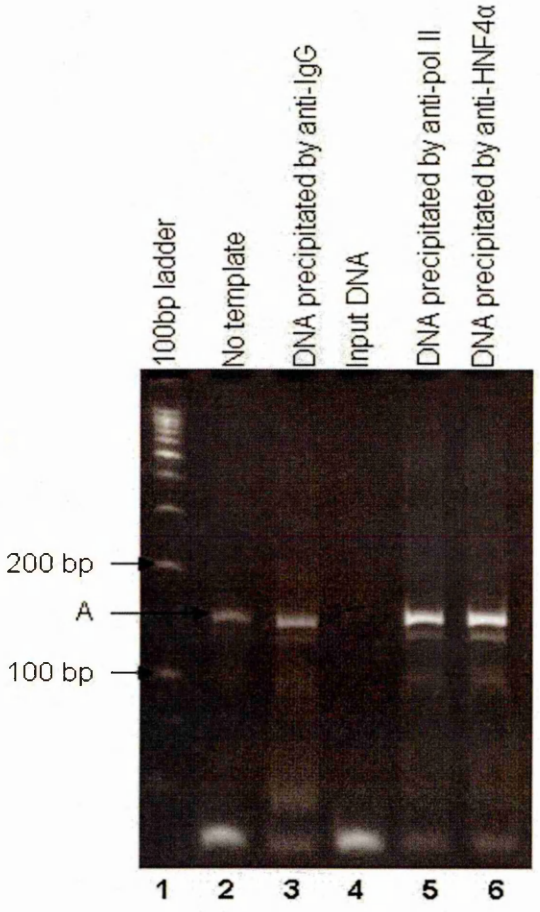


Fig 6.25. PCR of PZ -13/-163 with different immunoprecipitated chromatins.
Lane 1 is a100 bp ladder. Lane 2 is a no template negative control containing DEPC. Lane 3 contains DNA immunoprecipitated by IgG. Lane 4 contains input DNA. Lane 5 contains DNA immunoprecipitated by the pol II antibody. Lane 6 contains DNA immunoprecipitated by the HNF4 antibody. [A] indicates the desired product (151 bp). New aliquots of 10X PCR buffer and primers were used.

A further experiment for the PZ region -13/-163 was performed using new aliquots of all PCR reagents. The results showed the expected product with no contamination in the no template control. However, no increase in the signal strength was observed with the HNF4α enriched DNA (67%) compared to the input DNA (Fig 6.26).

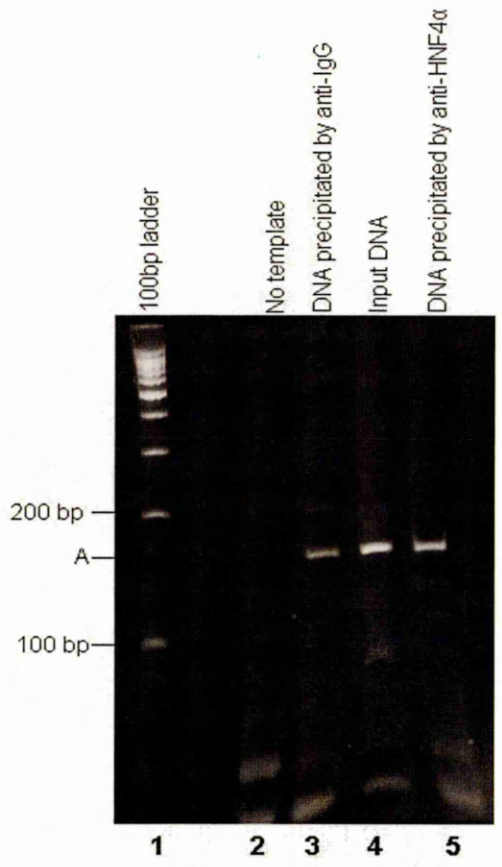


Fig 6.26. PCR of PZ -13/-163 with different immunoprecipitated chromatins. Lane 1 was 100 bp ladder. Lane 2 was no template negative controls containing DNase free water. Lane 3 contained DNA immunoprecipitated by IgG. Lane 4 contained input DNA. Lane 5 contained DNA immunoprecipitated by HNF4 antibody. A indicates the desired product (151 bp). New aliquots from all PCR reagents were used.

Chromatin immunoprecipitated with the CREB antibody, the PEA3 antibody and the HNF4α antibody were used in a further PCR. None of the enriched DNA samples showed any observable increase in the signal of the proposed product compared to the same amount of the input DNA. The signal intensity with CREB, HNF4α and PEA3 enriched DNA was 40%, 55% and 22%, respectively, compared to the same amount of the input DNA. Moreover, contamination was observed again in the no template control (Fig 6.27).

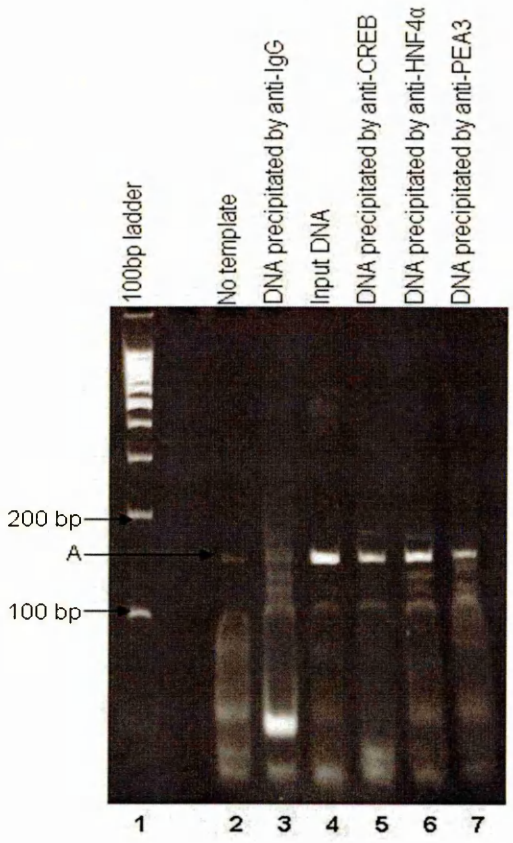


Fig 6.27. PCR of PZ -13/-163 with different immunoprecipitated chromatins.
Lane 1 was 100 bp ladder. Lane 2 is a no template PCR. Lane 3 contains DNA immunoprecipitated by IgG. Lane 4 contains input DNA. Lane 5 contains DNA immunoprecipitated by CREB antibody. Lane 6 contains DNA immunoprecipitated by HNF4α antibody. Lane 7 contains DNA immunoprecipitated by PEA3 antibody. [A] indicates the desired product (151 bp).

The experiment was repeated with new aliquots of the PCR reagents and new immunoprecipitated chromatin prepared from the HepG2 cells. However, similar results were observed, with contamination in the no template control and a stronger signal with the negative control (IgG immunoprecipitated) compared to the input DNA and the other immunoprecipitated chromatin (Fig 6.28). The signal intensity with the input DNA, CREB, HNF4α and PEA3 enriched DNA was 1%, 90%, 70% and 57%, respectively, compared to the same amount of the negative control.

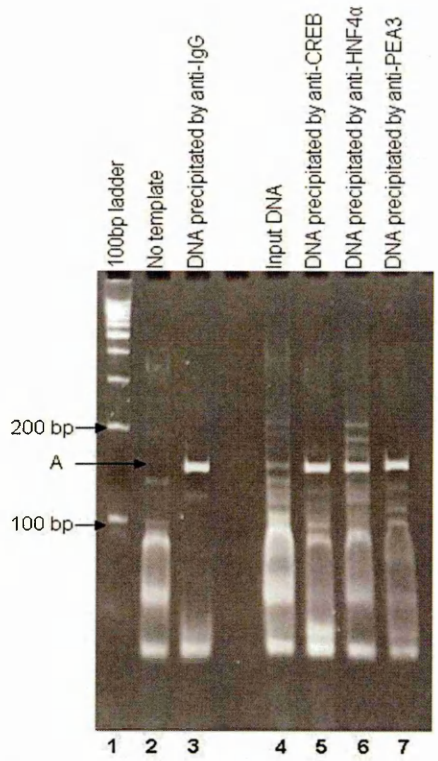


Fig 6.28. PCR of PZ -13/-163 with different immunoprecipitated chromatins. Lane 1 is a 100 bp ladder. Lane 2 is a no template PCR. Lane 3 contains DNA immunoprecipitated by IgG. Lane 4 contains input DNA. Lane 5 contains DNA immunoprecipitated by CREB antibody. Lane 6 contains DNA immunoprecipitated by HNF4α antibody. Lane 7 contains DNA immunoprecipitated by PEA3 antibody. [A] represents the desired product (151 bp).

6.3. Discussion

The study of protein-DNA interactions is a major area of research since it provides valuable biological information with regard to the regulation of genes (Jones and Thornton 2003). The accurate identification of binding sites and the identity of the transcription factors bound to these sites is a very important step in the characterisation of gene transcriptional regulation (Morozov *et al.* 2005). The electrophoretic mobility shift assay (EMSA) is a technique which has been used extensively in studying such protein-DNA interactions (Fried and Crothers 1981, Revzin 1989). This technique was used in the current project to study the proteins potentially binding the PZ gene promoter. The EMSA is a relatively fast and simple, well-established assay for studying a specific binding sequence for particular transcription factors in a region of DNA *in vitro* (Buratowski and Chodosh 2001). However, this assay has many limitations in studying the protein-DNA interactions that require *in vivo* conditions (Carey and Smale 2001). The assay is limited, for instance, in studying proteins that need looping of the DNA to bring distal binding sites closer together; a mechanism recently observed in the murine osteopontin gene promoter with the looping of DNA facilitated the interaction of NFκB and AP1 binding to *cis*-regulatory elements ~1700 bp apart (Zhao *et al.* 2011). In addition, the gel electrophoresis may cause dissolution of multiprotein complexes required for stabilising their interaction with the DNA, or inhibit the detection of transcription factors which require recruitment by an interaction with another factor for their cooperative DNA binding. The human β-interferon enhanceosome is an example of such complex formation with a requirement for the binding of the structural protein HMGA1 to facilitate the binding of NFκB, IRF and Jun/ATF in a highly cooperative manner, which then go on to recruit a coactivator, CBP (CREB binding protein) or p300 (Panne, Maniatis and Harrison 2007).

Footprinting is another assay which can be used in studying protein-DNA interactions. This technique is more powerful than EMSA and gives more information about the protein-DNA interactions, such as the number and position of protein-DNA interactions in longer fragments of DNA (Gilmartin and Bowler 2002). The main variables used in optimisation of footprinting assay are summarised in table 6.3.

DNaseI incubation	60, 90 or 120 seconds at room temperature
Nuclear extract	10 μg, 20 μg and 40 μg total protein
DNA fragment	0.5 μl and 1.0 μl from 10 μl of total volume of the proposed purified biotin labelled DNA fragment and 5 μl of the IRDye-labelled DNA fragment were loaded in the gel. 0.01 fmole, 0.1 fmole and 0.5 fmole of the purified unlabelled DNA fragment were used in the binding reaction and 5 μl were loaded in the gel after the amplification with one of the IRDye labelled primers.
Incubation of Binding reaction	Ice, room temperature or 37°C for 45 minutes. 4°C for overnight. 30 and 20 minutes on ice or 20 minutes at room temperature.
Binding buffers	Binding buffer-1 (Andrews and Faller 1991) Binding buffer-2 (Cordingley and Hager 1988) Binding buffer-3 (Batista et al. 2007)
Detection of DNA fragment	UVP Bioimaging system for biotin-labelled. Li-Cor 4200 sequencing system for IRDye-labelled.

Table 6.3. Different variables used in optimisation of footprinting assay.
Three different binding buffers and three different amounts of HepG2 NE with different incubations of the binding reaction were used to optimise the footprinting protocol. Moreover, different incubation periods were used to optimise the DNaseI digestion. The UVP bioimaging system and Li-Cor 4200 sequencing system were used to visualise the biotin-labelled and IRDye-labelled DNA digested fragment, respectively.

6.3.1. Footprinting assay using biotin-labelled primers

Although the footprinting technique traditionally uses radiolabelled probes the current study looked to employ alternative labelling methods. The optimisation of this technique was first done with biotin-labelled primers using the established protocol of Sambrook and Russell (2001) (Sambrook and Russell 2001). Different amounts of HepG2 NE was added to the purified DNA fragment of interest (PZ +86 to -213) in a binding buffer with the presence of a non specific competitor, poly (dI-dC), to bind non-specific DNA binding proteins. Binding reactions were carried out on ice, 4°C and at room temperature,

followed by DNase I digestion, recovery of the digested fragments and subsequent size fractionation on a 5% denaturing polyacrylamide gel (See Section 6.1.1.1).

Results showed a uniform ladder of the DNA fragment indicating optimisation of the DNase I digestion had been achieved. However, no protein binding to the DNA was shown with variation in any of the parameters assessed, such as temperature of the binding reaction or amount of protein (Fig 6.5 and Fig 6.6). The results also showed a non-uniform transfer of the DNA fragments to the nylon membrane.

6.3.2. Footprinting assay using IRDye-labelled primers

To avoid such issues with the transfer process, which may potentially lead to misinterpretation of the results, in addition to the amount of time to perform the experiment using biotin-labelled primers, IRDye-labelled primers were determined to be an attractive alternative. The majority of footprinting experiments have been successfully performed using radioisotope molecules. However, the time required in producing the autoradiograph and the hazardous nature of radioisotopes, not to mention their limited half-life has led to their substitution by fluorescent methodologies (Ying, Fourmy and Yoshizawa 2007).

The first experiment using the IRDye labelled DNA was performed using 10 fmole of purified IRDye-labelled fragment produced by PCR (+86 to -213) with the same protocol as that used for the biotin-labelled primers. Moreover, the HepG2 NE used was dialysed (SeeSection 3.3.1) to reduce its salt concentration. The DNA fragment was incubated with the HepG2 NE on ice for 20 minutes. The DNA was then digested by 0.01 units of DNase I for 1 minute, purified and electrophoresed on a 5% denaturing polyacrylamide gel alongside a 20 bp molecular weight size standard (Li-Cor, UK) to identify regions of protein-DNA binding. For more accurate identification of any protein-DNA binding regions a sequencing ladder can be used. Traditionally this has been G+A and C+T sequencing ladders (Maxam and Gilbert 1980), although a dideoxy sequencing ladder is easier to prepare and would provide the same, if not more information. No products were observed from this reaction at either 700 nm or 800 nm (Fig 6.7).

6.3.3. Modified footprinting assay using IRDye-labelled primers

Assuming that sensitivity may be an issue with fluorescently labelled products compared to radiolabelled products the protocol was then modified by adding an extra amplification step using IRDye-labelled primers, after the DNase I digestion (See Section 6.1.1.2). Different amounts of the unlabelled DNA fragment (+86 to -213) were incubated with different amounts of the Hep NE on ice for 20 minutes, digested, purified and amplified using the an IRDye-primer. The amplification of the DNA fragment was performed using a 700 IRDye labelled forward primer and then repeated with the 800 IR Dye-labelled reverse primer. An advantage of fluorescent labels is that reactions can be multiplexed and this method could allow the mapping of the protein binding sites on both strands of the double-stranded DNA simultaneously. The results showed a very strong signal with the 700-forward primer and a relatively good signal with the 800-reverse primer, suggesting that the amplification step was successful, when comparing signal intensity to the absence of amplification (fig. 6.7). However, no protein-DNA binding was seen (Fig 6.8 and Fig 6.9).

The same experiment was then repeated using 40µg of the Hep NE and the samples diluted 1:10 with the loading buffer before loading onto the gel, in an attempt to optimise the signal intensity. The results with the 700-labelled primer showed a relatively faint signal after the dilution, particularly when 0.01 and 0.1 fmole of the unlabelled fragment were used in the binding reaction, while too high a signal was detected from the undiluted samples (Fig 6.10). The signal detected from the 800-IRDye was relatively good with undiluted samples while it was lost after the dilution (Fig 6.11), suggesting either differential amplification between the two primers or a difference in their original labelling efficiencies during manufacture. However, again no footprints were detected with either IRDye-lablled primer. Moreover, although the use of the sharkstooth comb in order to run the samples close to each other is considered as an advantage for more accurate comparison, the samples showed some degree of cross-lane contamination due to seepage from the comb, which is a potential disadvantage of using such combs (Theophilus 1996).

A further experiment was conducted increasing the incubation time of the DNA with the HepG2 NE to 30 minutes to provide more time for the binding of the protein to the DNA to reach equilibrium. Moreover, based on the signal

intensities from the previous experiment the samples amplified by the 700-labelled primer were diluted 1:5 and those amplified by the 800-labelled primer diluted 1:2 to provide appropriate and balanced signals. The results, however, still showed too strong a signal with the 700-labelled primer (Fig 6.12) while a better signal was detected with the 800-labelled primer (Fig 6.13). However, again, no footprints were detected.

To optimise the signal with the 700-labelled fragment and to confirm the consistency of the signal detected from the 800-labelled fragment with the 1:2 dilution, the previous experiment was repeated using a 1:10 dilution with the DNA fragment amplified by the 700-labelled primer. Moreover, the digestion time with the DNase I was increased to 90 s and 2.0 minutes to determine how crucial the digestion time was to the uniform banding pattern of the digested fragments. The results showed good signal intensity at both 700 nm and 800 nm (Fig 6.14 and Fig 6.15). In addition, the incubation of DNase I with the DNA fragment for 90 s and 2.0 minutes showed the same pattern of banding obtained by the 1 minute incubation. However, again, footprints were not detected. Since the signal intensities from both 700 and 800-IRDye fragments were now optimised, as was the DNase I treatment, further studies were carried out in order to optimise the protein binding to DNA.

In an attempt to enhance protein-DNA interactions two further binding buffers, in addition to the original buffer (See Section 6.1.1.2) were assessed in a further footprinting assay. Moreover, the incubation time of the DNA fragment with the Hep NE was increased to 45 minutes on ice. Following post-digestion amplification the samples were diluted 1:10 and 1:2 for the 700-IRDye and 800-IRDye fragments, respectively, before they were size fractionated, since these dilutions had previously been shown to be optimal. The banding pattern still did not show any indication of protein-DNA binding with either labelled fragments.

To further optimise the protein-DNA binding other incubation conditions were used to perform the assay. The DNA fragment was incubated with the Hep NE overnight at 4°C, 45 minutes at room temperature and 45 minutes on ice. Nevertheless, none of these conditions showed footprints. Although most of the possible conditions to optimise the protein binding were tested, no protein-DNA interaction was achieved.

The DNase I footprinting assay, in order to be performed successfully, requires large amounts of the DNA-binding proteins to saturate the DNA binding sites (Guille and Kneale 1997). The majority of footprinting assays in previous studies were successfully accomplished using purified proteins (Wu, Urbanowski and Stauffer 1995, Maruyama *et al.* 2003), primary cells such as rat liver (Lee *et al.* 1997, Lee *et al.* 1999, Rath *et al.* 2008, Kim and Ahn 1998) or rabbit liver (Ip, Lee and Hammond 2000) , or nuclear extract from HepG2 cells (El-Sankary *et al.* 2002, Sugawara *et al.* 2007) prepared by the classical protocol reported by Dignam and colleagues (Dignam *et al.* 1983). The protocol used to prepare the nuclear extract (Andrews and Faller 1991), although derived from the large scale method developed by Dignam in 1983, was based on a limited number of cells and missed many steps which may affect the quality of the nuclear extract and the quantity of the DNA-binding proteins. Therefore, the current study has shown that, in principle, the novel technique of DNase I footprinting with a post-DNase I digestion strand-specific amplification step provides a sensitive technique for fragment detection, with a consistent and uniform banding pattern. However, the failure to optimise the protein-DNA interactions would seem to be due to the quality of the Hep NE and further work should be conducted using the original Dignam protocol and rat liver which may provide better quality and quantity of the DNA-binding proteins.

This '*in vitro*' footprinting assay can only show the interaction between the DNA sequence of interest and the binding proteins that are found in the nuclear extract and it does not demonstrate if these potential interactions actually occur within the cell (Hornstra and Yang 1993). Therefore, the use of '*in vivo*' footprinting would be useful to confirm the DNA-protein interactions within the cells of interest. DNase I footprinting, however, may lead to a limited interpretation of such interactions due to protein-DNA binding in a DNase I poorly cleaved region such as A or T tracts (Guille and Kneale 1997). The use of the footprinting assay also provides only the locations of the protein-DNA interactions and not the identity of the specific proteins. Therefore, the chromatin immunoprecipitation (ChIP) assay would be a useful way to confirm the location and the identity of the interacting protein, especially if the protein of interest was potentially known (Elnitski *et al.* 2006).

6.3.4. Chromatin immunoprecipitation assay

This *in situ* technique has become the assay of choice for fine mapping of transcription factors and providing an additional physiological representation of the events in the nucleus which are involved in DNA processing (Spencer *et al.* 2003). Moreover, the technique not only gives information about the protein bound to the DNA but also the proteins that their binding may depend on, such as other proteins which are not bound directly to the DNA (Shang *et al.* 2000). Although many agents could be used in the cross-linking step of this technique, such as methylene blue and UV, formaldehyde was chosen since it allows the study of both proteins that bind directly to DNA and those associated with chromatin via other proteins, due to its inducing both protein-DNA and protein-protein cross-links (Negre *et al.* 2006).

The ChIP assay was used in the current study to attempt to confirm the binding location of the DNA-binding proteins that had been identified in the PZ promoter by techniques such as EMSA that are less sympathetic to the chromatin conformation. First, the PCR amplification reactions of the DNA template of interest were optimised: (-13/-163) of PZ and (-222/-55) of GAPDH. It is useful to have the GAPDH control (Active Motif, Belgium) to be used to amplify the DNA region immunoprecipitated by the RNA pol II antibody (Active Motif, Belgium) as an indication of the optimisation of a quite technically demanding protocol. In spite of variation in the key amplification parameters the GAPDH primers failed to be optimised to amplify the proposed product (166 bp) (Fig 6.18). Ideally a positive control should be tested with another gene promoter known to be bound by the proposed protein. The chromatin immunoprecipitated by the RNA pol II antibody amplified by the primers used to amplify the -13/-163 region of PZ gene promoter was used as a positive control since RNA pol II binding should be found in any transcribed gene. The PZ gene region -13/-163 was then amplified by specific primers (See Section 6.1.2) and optimal amplification conditions determined (annealing temperature 62.5°C using 1mM $MgCl_2$ (Fig 6.23)). The primers amplifying the region of interest were then used with the DNA immunoprecipitated by the HNF4α antibody to confirm the binding of this protein to the region of interest, which had been detected *in vitro* by EMSA. The product amplified from the DNA immunoprecipitated by the HNF4α antibody was compared to the products

amplified by the DNA immunoprecipitated with the RNA pol II antibody and the IgG negative control antibody as well as the input DNA (sheared, pre-cleared and reverse cross-linked chromatin). Although a more intense signal was seen from the HNF4α enriched DNA compared to the negative control, no significant change was observed when compared to the input DNA (Fig 6.24). Moreover, contamination was found in the no template PCR control. In an attempt to isolate the contamination the same procedure was repeated using a new aliquot of the 10X PCR buffer and the same results were found with contamination in the no template control (Fig 6.25). Again, the experiment was repeated using new aliquots of all reagents for the PCR. Although the contamination disappeared, no significant change was observed in the intensity of the product amplified from the HNF4α enriched DNA and the input DNA (Fig 6.26).

As other factors had been identified to potentially bind within the PZ promoter region additional antibody enriched templates were prepared. CREB enriched DNA and PEA3 enriched DNA, in addition to the HNF4α enriched DNA were used in a further PZ-specific PCR. CREB and PEA3 were investigated since these proteins were thought, to some extent, to have potential binding sites within the candidate region (-13/-163). However, no enrichment was observed with any of the samples, with signal intensity being no greater than that with the input DNA (Fig 6.27). Although another PCR showed DNA enrichment with newly prepared immunoprecipitated chromatins compared to the input DNA, the intensity of the bands from these enriched DNAs was found to be similar to that of the negative control (Fig 6.28).

The ChIP technique has been used successfully in characterising the binding of many transcription factors, such as the binding of estrogen receptor to estrogen responsive promoters (Shang *et al.* 2000) and the binding of thyroid hormone receptor to mammalian T3-responsive promoters (Sharma and Fondell 2002). Although the basic principle was used in the current study, this technique failed to characterise any protein binding to the proposed region of PZ gene promoter. However, the immuno-basis of this technique may indicate that the problem was due to the use of an inappropriate antibody. Therefore, optimisation of the assay using different antibodies might have given better results. However, further work in optimising the key parameters would provide valuable information about the proteins that interact with the PZ promoter sequence.

Moreover, the end-point PCR used in the analysis of the bound DNA sequence by ChIP in the current study could have been replaced by real-time PCR (qPCR) to give more accurate quantification of the immunoprecipitated DNAs and, therefore, more reliable analysis (Mukhopadhyay *et al.* 2008).

In recent years a number of techniques have been developed to study protein-DNA interactions. For instance, the CHIP method to detect sequence-specific interactions of the bound protein and the footprinting technique in providing distinct results when the sequence-specific protein binding occur only in a certain fraction of cells, has led to a combination of these two methods as a developed technique to analyse protein-DNA interactions *in* situ within the nucleus (Kang, Vieira and Bungert 2002). This combination method is performed by immunoprecipitation of proteins by their specific antibodies followed by cleavage with DMSO and piperidine and analysis by LMPCR. Moreover, techniques such as chromosome conformation capture (3C) can be used to study long-range interactions of regulatory regions at the three dimensional level of chromatin architecture and are based on the capture of spatially related chromatin fragments by proximity ligation (Cullen, Kladde and Seyfred 1993). This can provide useful information when it is combined with ChIP. This ChIP-3C technique can reduce noise from non-specific interactions and provide important information about DNA-protein interactions and chromatin looping which is important for the binding of some proteins to mediate gene regulation (Fullwood and Ruan 2009).

The current study, therefore, attempts to take the initial data from relatively artificial *in vitro* experiments, such as EMSA, and to contextualise them. Such studies include the development of a new, highly sensitive footprinting methodology to accurately map the binding of proteins that may require adjacent binding site occupancy for their own binding; and the initial studies using ChIP to confirm protein-DNA interactions *in situ* within the nucleus. At the end of these studies both techniques do require further development and optimisation, but both offer the potential for relevant detailed data with regard to regulation of the PZ promoter.

Chapter 7-Hormonal regulation and inflammatory response of the protein Z gene

7.1. Introduction

7.1.1. Steroid hormones

Steroid and thyroid hormones are important molecules involved in many complex processes in the body such as cell development, differentiation and stimulatory response. These hormones are hydrophobic ligands for a large family of proteins known as nuclear receptors (Laudet *et al.* 1992). They act through binding to their respective receptors within cells and activating the expression of target genes (Evans 1988). In early studies this family was shown to have members known as orphan nuclear receptors with unknown ligands (Evans 1988, Moore 1990). The orphan receptors were suggested as the primitives of this family while the liganded receptors were considered as derivatives that had acquired their ability to bind ligands through the evolutionary process (Escriva *et al.* 1997).

The structure of the nuclear receptors is organised as an N-terminal domain with ligand-independent activation function (AF-1) (Tora *et al.* 1989, Metzger *et al.* 1995), a C-terminal ligand-binding domain, and a DNA binding domain (DBD), between the two termini, which binds to hormone response elements regulating expression of target genes (Khorasanizadeh and Rastinejad 2001). The domain in the C-terminus, known as the ligand-binding domain, has a complex structure including a region with a ligand-dependent activation function (AF-2), an interface for dimerisation and a ligand binding pocket (Steinmetz, Renaud and Moras 2001). The DBD, a highly conserved region among nuclear receptors, is composed of approximately 71 residues containing two zinc finger motifs matched tetrahedrally by conserved cysteines (Evans 1988, Freedman *et al.* 1988, Hard *et al.* 1990). The ubiquitous eukaryotic nuclear receptor proteins are structurally related transcription factors that bind to specific DNA cis-acting sequences and different coregulators, in response to binding of their respective ligand, and regulate the transcription of a multitude of target genes involved in such processes as reproduction, development, and metabolism of lipids and carbohydrates (Raviscioni *et al.* 2006, Ishizuka *et al.* 2005).

Many transcription cofactors, organized in multiprotein complexes, have been recognized as initiating chromatin remodeling and the modification of histones in order to expedite the binding of the nuclear receptors and RNA polymerase

II, with basal factors, to the target DNA (Ishizuka *et al.* 2005). These cofactors, for instance CBP/P300 (Chakravarti *et al.* 1996), P300/CBP associated factor (PCAF) and general control of amino-acid synthesis 5-like 2 (GCN5) (Blanco *et al.* 1998) act as coactivators, whereas the silencing mediator of retinoid and thyroid receptor (SMRT) (Chen and Evans 1995) and the nuclear receptor corepressor (N-CoR) (Horlein *et al.* 1995) act as corepressors.

Activation of transcription by steroid hormone receptors can be direct through protein-protein interaction or indirect by synergistic action with adjacent transcription factors (Schule *et al.* 1988, Strahle, Schmid and Schutz 1988). Interestingly, the target DNA is found to be bent in response to the binding of oestrogen nuclear receptors, suggesting the change of DNA conformation upon binding by steroid hormone receptors plays a part in their regulatory mechanism (Nardulli and Shapiro 1992). This bending may play a role in stabilising the receptor-DNA complex by forming a dimerisation interface and also may facilitate the interaction between transcription factors and between transcription factors and the basal transcription complex.

Hormone response elements are the DNA binding sites for the hormone receptors and are palindromic sequences composed of two half-sites separated by 3 bases (Aumais *et al.* 1996). The oestrogen response element has the sequence AGGTCANNNTGACCT in the vitellogenin A2 gene of *Xenopus laevis* (Klein-Hitpass *et al.* 1986). However, oestrogen was found to bind imperfect palindromic sequences (GGTCAACAT**A**ACC) in the chicken vitellogenin gene (VTGII) (Burch *et al.* 1988). Similarly, the palindromic sequences for glucocorticoid binding, with a half site of AGAACA (Aumais *et al.* 1996), were found to be imperfect (**TGT**ACANNNTGTTCT) in the rat tyrosine aminotransferase gene (Jantzen *et al.* 1987). It has been suggested that the two half sites are recognized by the steroid receptors as separate elements and that the integrity of the half sites, as well as the spacing between them, are critical for the binding of those receptors (Nordeen *et al.* 1990). Further studies also suggested that a widely spaced, repeated half site (RGGTCA) of the palindromic motifs can function as an oestrogen response element (Kato *et al.* 1992, Dana *et al.* 1994, Kato *et al.* 1995). The repeated half site of the glucocorticoid receptor palindromic sequences (TGTTCT) was also found to be

important for optimal glucocorticoid inducibility, independent of the other half site (Payvar *et al.* 1983, Chalepakis *et al.* 1988).

7.1.2. The role of steroid hormones in the regulation of protein Z and other coagulation factors

Recent work has suggested that PZ expression may be influenced by steroid hormone levels. The increased levels of testosterone, estradiol (Yoshida *et al.* 1982) and androstenedione (Soldin *et al.* 2005, Castracane *et al.* 1998) throughout pregnancy, where PZ levels were found to be increased from the first trimester and returned to normal after delivery (Quack Loetscher *et al.* 2005, Ramsay *et al.* 2005), suggest a potential role for androgen and oestrogen in the regulation of PZ. The level of 17-hydroxyprogesterone was also found to be increased after the thirty-third week of gestation (Tulchinsky *et al.* 1972). The increase of this hormone, which is produced during the synthesis of glucocorticoids and sex steroids, along with the increased levels of PZ during pregnancy (Quack Loetscher *et al.* 2005, Ramsay *et al.* 2005), also suggest the potential role of steroid hormones in the regulation of PZ. Moreover, the lower testosterone levels in women compared to men (Reckelhoff *et al.* 2005) which is mirrored by the levels of PZ (Ravi *et al.* 1998), in addition to the decrease in the amount of androgen and/or oestrogen with obesity (Mayes and Watson 2004), along with the increased levels of PZ in obese pregnant women compared to lean pregnant women (Ramsay *et al.* 2005), also suggest a role for the steroid hormones in PZ regulation. Furthermore, the decrease of androgen concentrations and the increase of PZ levels with the use of contraceptives (James 1993, Al-Shanqeeti *et al.* 2005) again support the putative role of the steroid hormones in PZ regulation.

In addition to the literature above which suggests the potential role of the steroid hormones in the regulation of PZ, the effect of steroid hormones in the regulation of FVII (Di Bitondo *et al.* 2002), which is suggested to be evolutionarily related to PZ, also predicts the possibility of similar regulatory mechanisms, such as hormonal regulation of the PZ gene. An oestrogen response element (ORE) detected in the gene promoter of FVII was shown to play an important role in reducing the transcriptional activity of this factor in the presence of oestrogenic factors. Although the ORE that is found in the FVII promoter has two half sites separated with two spacer nucleotides, which differs

from the classical ORE which comprises two half sites separated with three spacer nucleotides, the binding affinity for the oestrogenic factors was approximately the same (Di Bitondo *et al.* 2002). Furthermore, the age-related increase in the plasma level of FIX which is also believed to have evolved from the same ancestral gene as PZ is likely mediated through the activation of the androgen response element in the promoter (Crossley *et al.* 1992). The reported role of steroid hormones in regulating expression of these evolutionarily related factors raises the possibility that regulatory sequences may also have been conserved to some degree and that, therefore, hormonal regulation may account for the differences in PZ expression observed between the sexes and the various states affecting hormone levels.

The studies described above provide a complex picture of the relationship between steroid hormone levels and the levels of PZ, but all indicate to some degree that steroid hormones have a role to play in modulating expression of this haemostatic factor. This leads to two possible hypotheses; the first being that the balance between androgen and oestrogen in favour of androgen causes an increase in procoagulation by increasing procoagulant and decreasing anticoagulant levels. Therefore, the expression of PZ, a suggested anticoagulant, may decrease in response to androgen and increase in response to oestrogen. The second hypothesis being that in this state of higher androgen levels the haemostatic balance is maintained, to some degree, by an increase in both procoagulants and anticoagulants. Therefore, the expression of PZ may increase in response to androgen. The subsequent studies aimed to elucidate the exact role of steroid hormones in controlling the expression of PZ and, therefore, to determine if the level of PZ may be an additional factor that may contribute to levels of thrombotic risk in different hormonal states.

7.1.3. Inflammatory response

The relationship between inflammation and coagulation has been well documented, with the body having evolved a response that attempts to limit the invasion of infectious agents or aberrant cells through the activation of coagulation and the formation of fibrin. This response, however, must be localised to avoid the consequences of systemic procoagulant activation (Levi, van der Poll and Buller 2004). A variety of defence mechanisms against

infections have been delineated; one such mechanism that responds to infection by gram-negative bacteria relies on the identification of an outer membrane endotoxin glycoprotein known as a lipopolysaccharide (LPS) (Schumann *et al.* 1990). Lipopolysaccharides are normally composed of a hydrophobic endotoxin, also known as lipid A, followed by a core of non repeating oligosaccharide, and a distal polysaccharide, also referred to as O-antigen (Pugin *et al.* 1993, Raetz and Whitfield 2002). The defence mechanisms initiated against the endotoxin in many mammalian cells are activated through a pathway that includes the LPS binding glycoprotein (LBP) and Cluster of differentiation 14 (CD14) glycoprotein (Tobias *et al.* 1995). LPS initiates activation in cells such as monocytes, macrophages and neutrophils by its binding to glycerophosphoinositol tailed CD14 found in the membrane (mCD14) in the presence of LBP (Schumann *et al.* 1990, Wright *et al.* 1990). However, in other cells that do not express mCD14, including endothelial and smooth muscle cells, the LBP-LPS complex binds plasma circulating CD14 (sCD14); this complex then binds to sCD14-LPS receptors (Pugin *et al.* 1993, Loppnow *et al.* 1995). The effect of the endotoxin was found to be mediated by cytokines, such as tumour necrosis factor (Sherry and Cerami 1988) which upon increased secretion might cause endotoxic shock and lead to death (Wright *et al.* 1990).

7.1.4. The effect of lipopolysaccharide on protein Z levels

The levels of a number of coagulation factors have been reported as being responsive to inflammatory signals through modification of transcription levels. For instance, the transcription of factor VIII (FVIII), an acute phase coagulation protein mainly produced in the liver, responds to bacterial lipopolysaccharide (LPS) by increasing its level in human circulation. The FVIII gene promoter region responsible for the stimulatory effect of bacterial lipopolysaccharide (LPS) contains two binding sites for C/EBPβ and one for NF-κB. Mutation analysis of the disrupted down stream C/EBP binding site in HepG2 cells reduces the response of the promoter to LPS by approximately 50%, whilst mutation of the NF-κB binding site completely stops LPS responsiveness. However, the level of mouse factor VIII was found to be unchanged in response to inflammation. This may be explained by the functional importance of the

highly conserved sequence (GGG), at the 5' end of the NF-κB binding site, which is absent in the orthologous mouse sequence (Begbie *et al.* 2000). The expression of tissue factor, the primary initiator of coagulation, is also subjected to induction by LPS. Tissue factor was found to be induced in monocytes by LPS in certain diseases causing disseminated intravascular coagulation (Osterud and Flaegstad 1983). The gene promoter of tissue factor has a region called the LPS response element (LRS) which contains transcription factor binding sites, one for NFκB and two for AP1 (proximal and distal) which are crucial for LPS activation (Mackman, Brand and Edgington 1991). Mutation in any of these sites compromises LPS induction. The mechanism to control transcription of the tissue factor gene following LPS induction includes proteolysis of the cytoplasmic IκBα leading to translocation of p65 and c-Rel from the cytoplasm to nucleus and transient phosphorylation of JunD in the constitutive AP1 complexes bound to the LRE (Hall, Vos and Bertina 1999).

One explanation for the wide range of variation in PZ levels among normal individuals is that inflammatory response may modulate levels of expression. Raczkowski *et al.* (1987) reported the correlation between the negative acute phase protein (APP), transferrin, and the levels of PZ in the serum of patients with rheumatoid arthritis, suggesting PZ may be acting as a negative acute phase protein (Raczkowski *et al.* 1987). This prediction was supported by another group when it was demonstrated that levels of plasma protein Z decrease in response to increasing levels of the cytokine interleukin 6 (IL-6) in patients suffering from acute leukemia and non-hodgkin's lymphoma (Undar, Karadogan and Ozturk 1999).

7.2. Aims

The aim of this work was to investigate the roles of steroid hormones and inflammatory pathways in modulating PZ expression levels and, therefore, potentially account for some of the wide variation in PZ levels observed in the normal population.

7.3. Methods

7.3.1. Gene expression studies of the endogenous PZ gene in HepG2 cells

Prior to investigation of the importance of steroid hormones on transcription of the PZ gene in the HepG2 cell model, the identity of the expression plasmids encoding the receptors of those hormones were verified.

7.3.1.1. Verification of the expression plasmids

Prior to plasmids verification, large scale cultures of expression plasmid transformed *E. coli* were prepared and the plasmids purified and quantified (see section 2.4.3). The expression vector encoding the particular hormone receptor was digested using appropriate restriction enzymes (Fermentas life science, UK) (2.4.3.3) before size-fractionation through a 1% w/v 1x Tris borate EDTA (1x TBE) agarose gel for 45 minutes at 10 volt/cm. The sizes of the digestion products were determined using a one kilo base pair molecular weight ladder (Fermentas life science, UK). The androgen receptor expression plasmid (pKC3 rAR) was digested with BglII and HindIII; the glucocorticoid receptor expression plasmid (pRShGRα) was digested with *Eco*RI; and the oestrogen receptor expression plasmid (pSG5 HEO) was digested with *Eco*RI and *Dra*III. The expression plasmids were obtained as a gift from Dr. Peter Winship (Sheffield University).

7.3.1.2. RNA analyses

Total RNA was prepared from HepG2 cells after transfection with the expression plasmid encoding the hormone receptor and from control untransfected cells, using the GenElute mammalian total RNA miniprep kit (See Section 2.3.1). The purity and the quantity of the prepared RNA was determined by the NanoDrop®ND-1000 spectrophotometer and the integrity was assessed using a 1% w/v (1x TBE) agarose gel (see Section 2.3.2). The RNA was treated with DNase 1 (Sigma-Aldrich, UK) to remove any genomic DNA. The RNA was then reverse transcribed to cDNA (See Section 2.3.3) to be used in the qPCR.

7.3.1.3. Primer efficiency

The efficiencies of the primers used in qPCR were determined, as described earlier (See Sections 2.3.4.1 and 4.3.2.4), by the efficiency equation;
Efficiency = $-1 + 10^{(-1/slope)}$.

7.3.1.4. Comparison of the HepG2 relative expression of steroid hormone receptors to normal liver tissue

The relative expression of androgen, glucocorticoid and oestrogen receptors was measured using the cDNA reverse transcribed from the RNA prepared from HepG2 cells and the cDNA prepared from the commercial liver tissue total RNA (See Section 2.4.2). The relative expression of the particular hormone receptors in HepG2 cells was compared to normal liver tissue after normalisation to the UBC reference gene.

7.3.1.5. The effect of steroid hormones on the endogenous PZ gene

The Tfx™-20 transfection reagent (Promega corporation, UK) was used for the overexpression of each steroid hormone receptor in HepG2 cells using plasmids encoding the respective receptor, according to the manufacturer's instructions (See Section 2.4.4). After the one hour pre-incubation of the HepG2 cells with the transfection mixture containing the expression plasmids, the respective ligand was added and further incubated for 48 hours at 37°C and 5% CO_2. 1x 10^{-8} M 17-β Oestradiol, 2.5 x 10^{-8} M Dihydrotestosterone and 1x 10^{-6} Dexamethasone final concentration were added to the HepG2 culture wells transfected with oestrogen, androgen and glucocorticoid receptors, respectively. Fresh complete D-MEM phenol free medium (Invitrogen limited, UK) containing charcoal stripped fetal calf serum (Invitrogen limited, UK) was used in the transfection. The effect of the candidate hormone receptor overexpressed in HepG2 cells and stimulated with the respective ligand on the expression of the endogenous PZ gene was measured by qPCR using SYBR green chemistry (See Section 2.3.4.3).

Prior to performing qPCR a housekeeping gene was selected for its stability (See Section 2.3.4.2), after transfection the HepG2 cells with the proposed

receptor followed by stimulation with the respective hormone, to be used as reference gene.

7.3.1.6. The effect of LPS stimulation on PZ transcription

HepG2 cells were stimulated with 10 µg/ml LPS from *Salmonella typhimurium* (Sigma Aldrich, UK) over a 24 hour time-course (See Section 2.4.6). The relative PZ mRNA expression in the stimulated sample compared to unstimulated sample was measured by qPCR using the SYBR green method (See Section 2.3.4.3).

7.4. Results

7.4.1. Verification of expression plasmids

Digestion of the androgen receptor (AR) expression plasmid (pKC3 rAR) with BglII and HindIII restriction enzymes produced three bands of 3.5, 1.6 and 1.2 kb. The glucocorticoid receptor (GR) expression plasmid (pRShGRα) after the digestion with *Eco*RI enzyme produced approximately 2.9, 1.7, 1.4 and 0.75 kb products while the products obtained from *Eco*RI and *Dra*III digestion of the oestrogen receptor (OR) expression plasmid (pSG5 HEO) were 1.4kb, 1.5kb, 2.6kb, 0.3 kb. The digestion patterns for all three plasmids are shown in Fig 7.1.

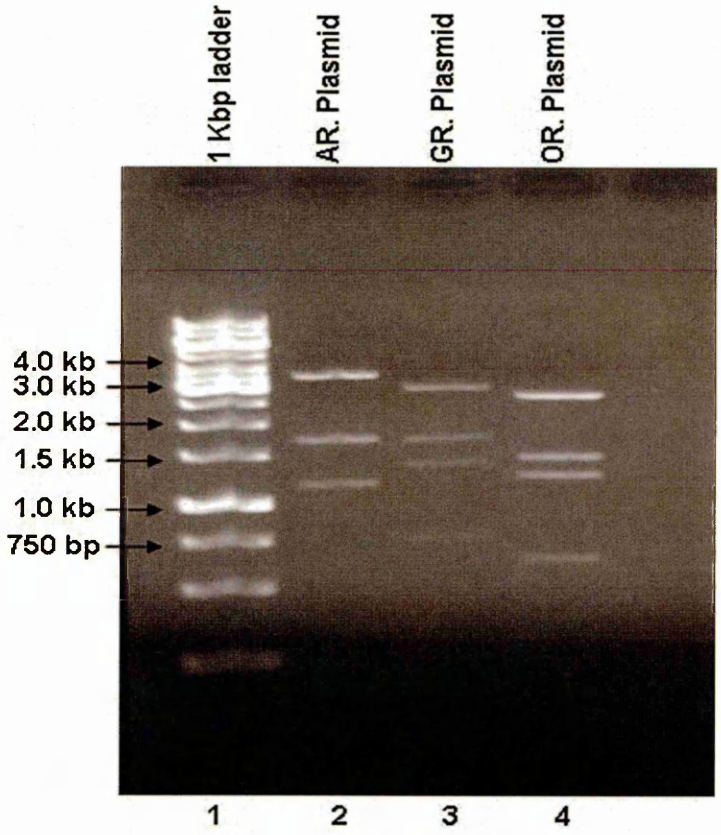


Fig 7.1. Restriction enzyme digestion of plasmids encoding steroid hormone receptors. Lane 1 a 1 Kb ladder. Lane 2 contains the androgen receptor (AR) expression plasmid after digestion with BglII and HindIII. Lane 3 contains the glucocorticoid receptor (GR) expression plasmid after digestion with *Eco*RI. Lane 4 contains the oestrogen receptor (OR) expression plasmid digested with *Eco*RI and *Dra*III. Digestions of the AR produced 3.5, 1.6 and 1.2 kb products, GR produced 2.9, 1.7, 1.4 and 0.75 kb products, and OR produced 1.4kb, 1.5kb, 2.6kb, 0.3 kb products.

7.4.2. RNA analysis

The purity of the prepared RNA, as determined by the absorbance ratio 260 nm/280 nm was found to be >1.8 by the NanoDrop®ND-1000 spectrophotometer. The 18S and 28S ribosomal RNA (rRNA) bands on 1% w/v agarose gel showed no significant degradation in all samples, indicating good quality RNAs. The rRNA bands from untransfected and transfected HepG2 cells with plasmids encoding the hormone receptors are shown in Figure 7.2.

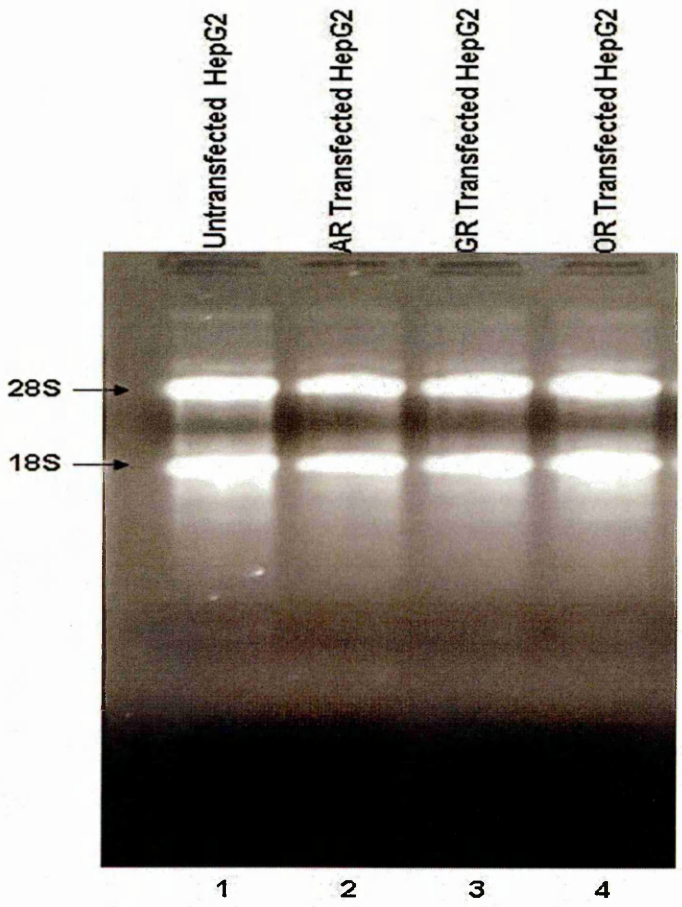


Fig 7.2. Agarose gel (1% w/v) electrophoresis of total RNA prepared from control and expression plasmid transfected HepG2 cells. Lane 1 shows the 28S and 18S rRNA bands from untransfected HepG2 cells. Lane 2, 3 and 4 show the 28S and 18S rRNA bands from samples transfected with expression plasmids encoding the androgen, glucocorticoid and oestrogen receptors, respectively. No significant degradation indicated good quality RNAs.

7.4.3. Reference gene validation

The tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (YWHAZ) gene showed stability of its relative

expression, by qPCR, in untrasfected and activated AR transfected HepG2 cells. UBC also showed stability in relative expression with and without the activated hormone receptor for either the glucocorticoid receptor or the oestrogen receptor (Fig 7.3).

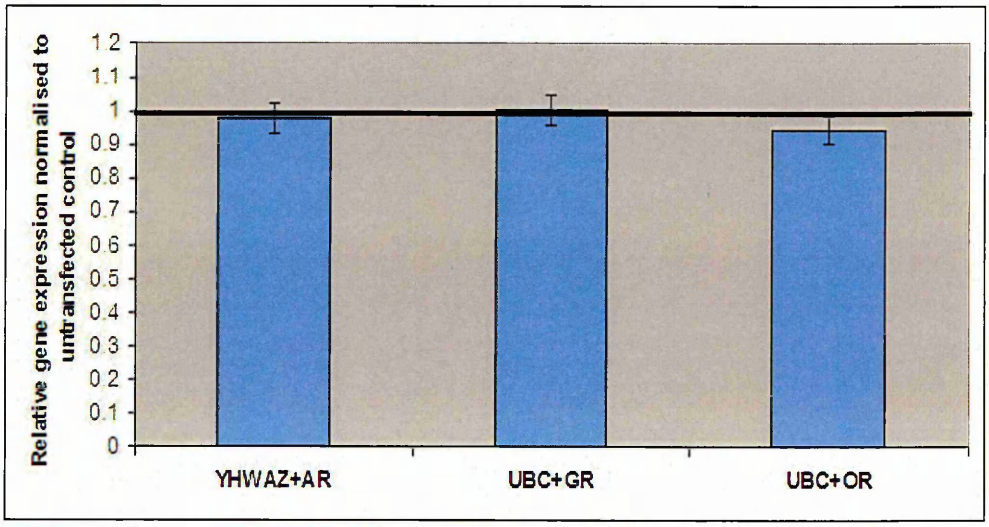


Fig 7.3. Relative expression of housekeeping genes with transfection.
The chart shows stability of YHWAZ mRNA relative gene expression in HepG2 trasfected with androgen receptor normalised to the untransfected cells, with only negligible change of 2% after trasfection. UBC mRNA relative expression of glucocorticoid receptor (GR) transfected HepG2 cells was found to be stable compared to untransfected cells while its relative expression in oestrogen receptor (OR) transfected HepG2 cells was not significantly affected, with a change of only 6%. The horizontal thick line represents the levels of mRNA relative expression in untransfected cells. The results shown were from three independent experiments.

7.4.4. Primer efficiency

The primers used in testing the effect of the steroid hormones and in comparing HepG2 relative expressions of steroid hormone receptors to liver tissue were determined. UBC and PZ(2) primer efficiencies were previously measured (See Section 5.3.3.7) and found to be 95% and 98%, respectively. The YWHAZ, AR, GR and OR primer efficiencies were found to be 102%, 88%, 99% and 96%, respectively (Fig 7.4).

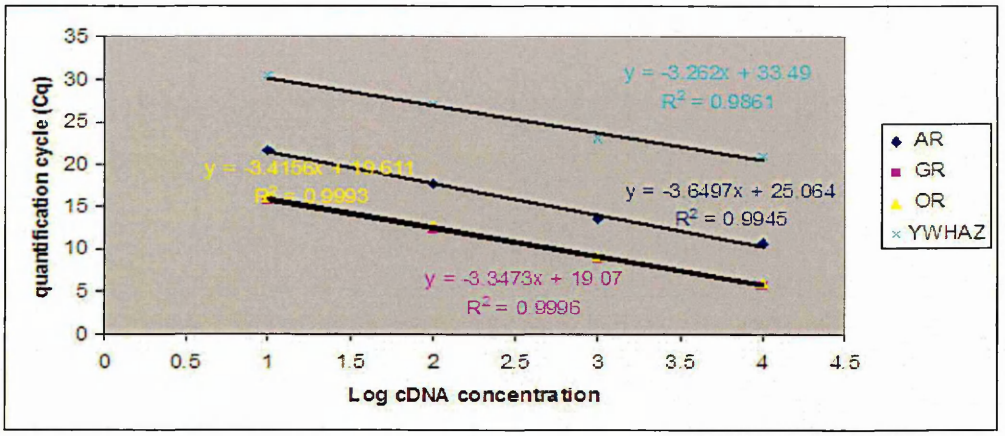


Fig 7.4. Determination of the YWHAZ, AR, GR and OR primers efficiency.
The efficiency of the reference gene (YWHAZ) primers was measured by plotting the Cq values versus the cDNA serial dilutions. The slope was used in the efficiency equation (Efficiency = $-1 + 10^{(-1/slope)}$). The efficiency was calculated and found to be 102%, 88%, 99% and 96% for YWHAZ, AR, GR and OR, respectively.

7.4.5. Expression of AR, GR and OR in HepG2 compared to normal liver

The expression levels of AR and OR were undetermined in HepG2 cells, suggesting no or very low levels of expression. The GR mRNA relative expression in HepG2 cells was found to be approximately 0.65-fold that of normal human liver (Fig 7.5).

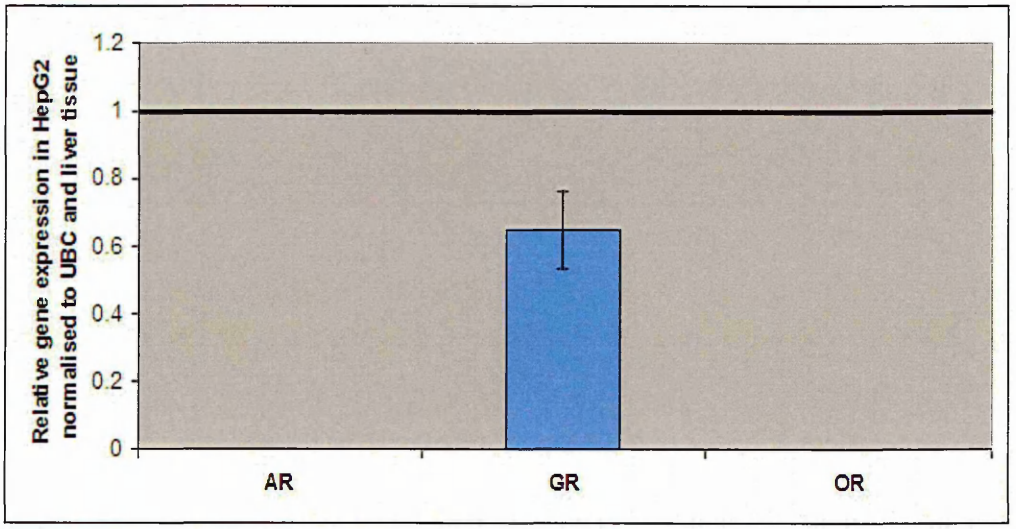


Fig 7.5. The relative expression of steroid hormone receptors in HepG2 cells and normal liver. The relative expression of AR and OR mRNA in HepG2 cells were found to be undetermined whilst the GR mRNA was found approximately 0.65-fold relatively expressed compared to liver tissue and normalised to UBC reference gene. The thick horizontal line represents the relative expression in liver tissue.

7.4.6. The effect of androgen, glucocorticoid and oestrogen hormones on PZ transcription

Transfection of HepG2 cells with the AR followed by stimulation with the ligand dihydrotestosterone did not affect the relative expression of PZ mRNA. GR transfected HepG2 stimulated by dexamethasone showed a 1.3-fold increase in PZ mRNA relative expression. A 1.2-fold increase in PZ mRNA relative expression was observed in HepG2 cells transfected with OR and activated by 17-β oestradiol (Fig 7.6). The results were not found to be statistically significant (P>0.05) using the REST 2009 software. These results were obtained from four, four and seven independent experiments for dihydrotestosterone, dexamethasone and 17-β oestradiol activation, respectively.

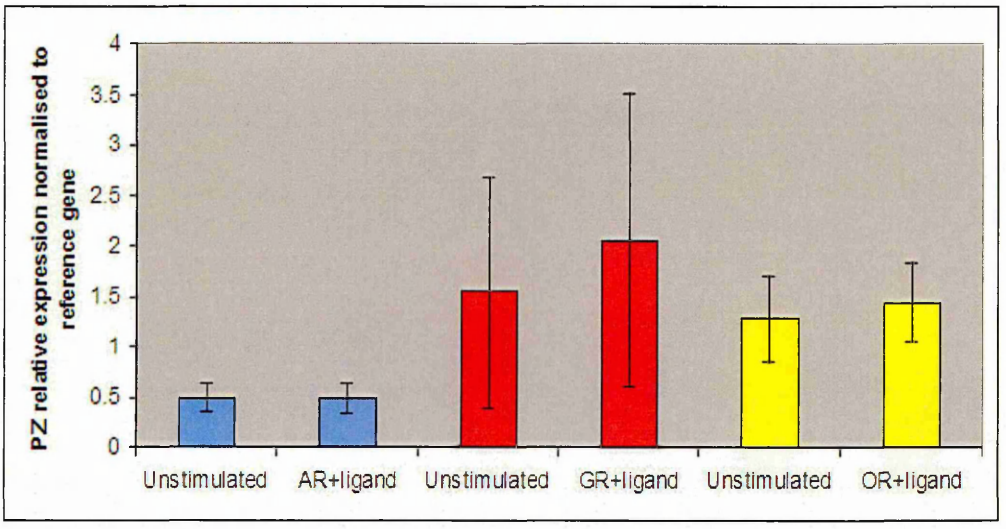


Fig 7.6. The effect of steroid hormones on PZ mRNA levels in HepG2 cells.
The androgen hormone shows no effect on the relative expression of PZ mRNA (blue columns), whilst glucocorticoid and oestrogen both produce a small increase in the relative expression of PZ mRNA, of 1.3-fold (red columns) and 1.2-fold (yellow columns), respectively. The increase with both glucocorticoid and oestrogen hormones was found not to be statistically significant using the REST 2009 (P>0.05). The effect of androgen stimulations were normalised by YWHAZ, whilst glucocoticoids and oestrogen stimulations were normalised by UBC. These results of adrogen, glucocorticoid and oestrogen were obtained from four, four and seven independent experiments, respectively.

The melt curves (Fig 2.4) of the amplfication products for the experiments described previously showed neither non-specific amplification nor the formation of primer-dimers in the transfected and control samples indicating a reliable quantification of the respective transcripts.

7.4.7. The effect of LPS stimulation on PZ transcription

The PZ mRNA relative expression quantified by qPCR was found to be decreased gradually over the first hour. Relative expression started increasing by 2 hours and had returned to pre-stimulation levels at four hours and 24 hours (Fig 7.7). Although a trend was clearly observed showing a drop in expression over the first hour of LPS stimulation, the results were not found to be statistically significant using the Kruskal Wallis test ($P>0.05$, $n=7$) (used for multiple time points that do not show a normal distribution).

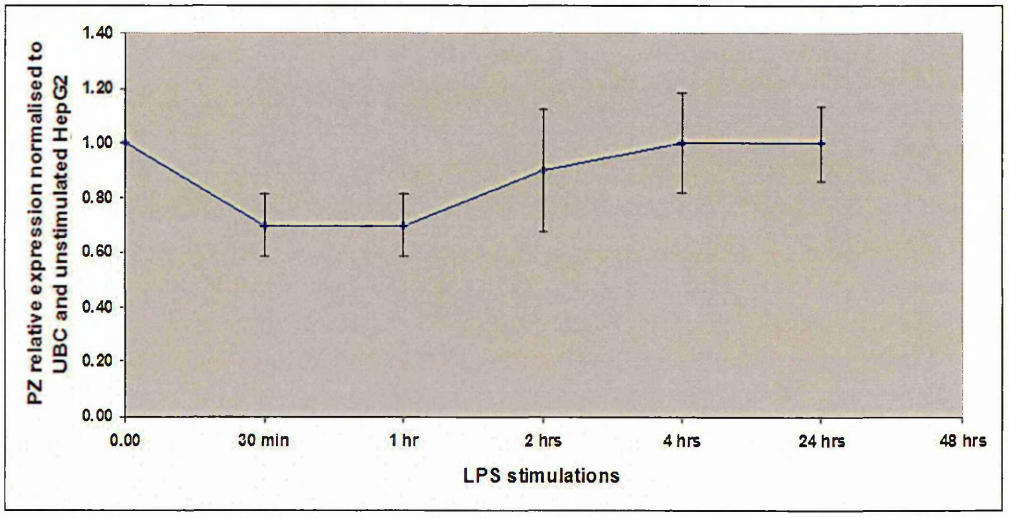


Fig 7.7. The effect of LPS on PZ mRNA levels in HepG2 cells.
HepG2 cells were stimulated with 10 µg/ml LPS for the times indicated. Total RNAs were prepared from unstimulated and the stimulated cells and the expression of PZ was measured at each point by qPCR. Down regulation of PZ mRNA relative expression was seen at 30 minutes and after one hour whilst expression started to increase at 2 hours and returned to pre-stimulation levels by four hours. The variation between points is not significant when analysed using the Kruskal-Wallis test ($P>0.05$, $n=7$)

7.5. Discussion

A potential effect of steroid hormones in regulating PZ expression was suggested, as mentioned previously, from the observations of the variation in protein Z levels in different hormonal states such as pregnancy (Quack Loetscher *et al.* 2005, Ramsay *et al.* 2005), oral contraceptive use (Al-Shanqeeti *et al.* 2005) and obese and lean pregnant women (Ramsay *et al.* 2005). Moreover, transcriptional regulation of FVII and FIX, evolutionarily related genes to PZ (Ichinose *et al.* 1990), were found to be regulated by oestrogen and androgen receptors, respectively (Di Bitondo *et al.* 2002, Crossley *et al.* 1992). The effect of steroid hormones on the PZ mRNA relative expression was investigated by studying the endogenous PZ gene locus in the model system, HepG2 cells. Nuclear receptors (NRs) are essential for the cellular action of steroid hormones. These lipophilic hormones bind the C-terminal binding domain of the NRs, causing their nuclear translocation and binding to their cognate sequences (He, Cheng and Xie 2010, Germain *et al.* 2006). HepG2 cells, the model used in the study, were tested for their expression of AR, GR and OR and compared to normal liver cells. The relative expression of AR and OR mRNAs was undetermined in HepG2 cells, suggesting no or very low levels of expression whilst relative expression of GR mRNA was 0.65-fold normalised to UBC and normal human liver cells (Fig 7.5). The HepG2 cells were reported previously for their lack of OR (Harnish *et al.* 1998, Barkhem *et al.* 2002) and AR expression (Nakagama *et al.* 1991) which is consistent with the current study. The lack of the relative expression of OR and AR in HepG2 cells may suggest limitations of these cells in being used as a model in studying the action of those steroid hormones. However, HepG2 cells have been used in numerous studies in order to study the effect of these hormones on the regulation of genes by using the transfection of the steroid hormone receptors (Yuen and Leung 2008, Tang, Norlin and Wikvall 2007, Farsetti *et al.* 2001, Harnish *et al.* 1998, Sabbah *et al.* 1998). Therefore, to monitor the regulation of the endogenous PZ by steroid hormones HepG2 cells were transfected by the respective hormone receptor. HepG2 cells were maintained in phenol free medium containing charcoal stripped fetal calf serum (Invitrogen limited, UK) during the process of transfection to remove any endogenous bovine steroid hormones that may have skewed the results (Yuen

and Leung 2008, Sabbah *et al.* 1998, Wiren *et al.* 1997). Prior to transfection, the identity of androgen receptor (AR), glucocorticoid receptor (GR) and oestrogen receptor (OR) expression plasmids were verified by restriction enzyme digestion and the expected digestion patterns for all three plasmids were observed (Fig 5.1).

The results of transfections showed that the levels of PZ mRNA relative expression were not affected in response to androgen, but showed a slight increase with both glucocorticoid (1.3 fold average) and oestrogen (average of 1.2 fold) (Fig 5.4). The effect of the glucocorticoid (n=4) and oestrogen (n=7) was, however, not found to be statistically significant using the REST 2009 software ($P>0.05$). UBC reference gene was used to normalise the relative expression with glucocorticoid and oestrogen, whilst YWHAZ reference gene with androgen due to their stability (Fig 7.3). The non-significant increase of glucocorticoid and oestrogen does not necessary exclude the biological significance of these hormones in the transcriptional regulation of protein Z. The increased levels of oestrogen and androstenedione, an intermediate in the production of androgen testerone and oestradiol, throughout pregnancy (Soldin *et al.* 2005, Castracane *et al.* 1998, Yoshida *et al.* 1982) may contribute to the high levels of PZ which were also found in pregnancy (Quack Loetscher *et al.* 2005, Ramsay *et al.* 2005). This is consistent with the current study where increased levels of PZ were found in response to oestrogen. Therefore, although this increase was not statistically significant, it may be worth further analyses to determine if oestrogen levels have a biological significance in controlling PZ expression. Furthermore, the role of oestrogen in reducing the transcriptional activity of FVII in addition to its role in increasing levels of mRNA in the bone marrow of ovariectomized mice (Moverare *et al.* 2004), may be mirrored in oestrogen's effects on other coagulation factors including PZ.

The increased levels of the intermediate of glucocorticoid synthesis, 17-hydroxyprogesterone, during pregnancy (Tulchinsky *et al.* 1972) may also suggest the role of glucocorticoids in elevating PZ levels which were found in pregnant women. Moreover, other coagulation factors such as FVII, which are suggested to have evolved from the same ancestral gene as PZ, and factors VIII and XI have shown increased levels in humans treated with dexamethasone (Brotman *et al.* 2006). The increase of these coagulation

factors in response to dexamethasone may support the biological effect of glucocorticoids on PZ which was not found to be statistically significant in the current study. Furthermore, the number of independent experiments (n=4) for dexamethasone treatment may not be sufficient to determine relatively small changes in expression to statistical significance.

Characterisation of the transcriptional regulation of the endogenous PZ gene, stimulated by steroid hormones in the HepG2 cell model, although studied in the context of the normal chromatin structure, is still considered to be *in vitro*. Therefore, further studies using animal models may give more reliable results.

Also, as described in the introduction to this chapter there is some evidence to suggest that PZ may be an acute phase protein. In this study the induction by LPS of the HepG2 endogenous PZ gene was analysed over a 24 hour period. The levels of PZ mRNA decreased in response to LPS by 30 minutes and reached a minimum level at one hour exposure. The PZ mRNA levels had increased by 2 hours and reached pre-stimulation levels by four hours. However, the results were not found to be statistically significant using the Kruskal Wallis test ($P>0.05$, $n=7$). The response of PZ to LPS was found to be similar to the time course response of other coagulation factors to LPS. For instance, tissue factor, which was suggested as an acute phase protein, had expression induced by LPS stimulation at time points of 30 minutes and 1 hour and started to return back to normal levels at a time point of 2 hours (Hall, Vos and Bertina 1999), which mirrors the response of PZ to LPS in this study, suggesting PZ may also be an acute phase protein. In addition, plasma PZ levels have been shown to be decreased in response to increasing levels of IL-6 (Undar, Karadogan and Ozturk 1999). Raczkowski has also suggested PZ behaves as a negative acute phase reactant (Raczkowski *et al.* 1987). These previous studies and the results from this study suggest that PZ may be a negative acute phase protein but confirmatory studies to statistical significance should be carried out.

7.6. Conclusion

It can be concluded that, although the expression data of the endogenous PZ gene in HepG2 cells demonstrated no statistically significant effect with

glucocorticoid ($P>0.05$, REST 2009, $n=4$) or oestrogen ($P>0.05$, REST 2009, $n=7$) hormones and inflammatory mediators ($P>0.05$, Kruskal Wallis test, $n=7$), the trend of PZ expression in response to these signalling mechanisms correlates with previous studies of PZ levels in various hormonal and inflammatory states. Furthermore, the limitations of the model system may account for effects that did not produce significant results in culture, but that do have a biological significance; with the time-course of LPS induction showing a high degree of correlation to predicted response profiles for other acute phase proteins.

Chapter 8-Discussion and future work

8.1. Discussion

The levels of plasma PZ are found to be highly variable among normal individuals (Miletich and Broze 1987). This wide variation in PZ levels may be explained, in part, by inter-individual differences in the rate of transcription from the PZ gene.

8.1.1. The role of HNF4 in regulation of PZ transcription

The HNF4α transcription factor was found to play a major role in the expression of a number of blood coagulation factors, such as, FVII, FVIII, FIX, FX and FXI (Stauffer *et al.* 1998, Figueiredo and Brownlee 1995, Reijnen *et al.* 1992, Miao *et al.* 1992, Tarumi *et al.* 2002). Moreover, although the mechanisms by which transcription of the PZ gene is regulated have not been defined in detail, a recent study by Sugawara and colleagues in 2007 looked at the role of HNF4α in controlling PZ transcription using protein-DNA binding studies (EMSAs) and reporter gene assays (Sugawara *et al.* 2007). Therefore, the role of HNF4α in PZ transcriptional regulation is defined in the current study using EMSAs and utilising the endogenous PZ gene locus rather than a reporter gene construct.

EMSAs using probes containing the hypothetical proximal binding site for HNF4α (pPZHNF4α), identified by bioinformatic analysis (-48/-38), showed specific protein binding as the binding was negligibly competed by a 100X fold excess of either HNF3B (9%), AP1 (17%) or SP1 (21%) consensus binding sites (non specific competitors), but was considerably diminished by a 100X excess of self competitor (60%) and abolished by a 100X fold excess of an HNF4α consensus binding site (100%) (Fig 4.2). These data suggest that the bound protein is HNF4α. To confirm this, an HNF4α antibody was used to confirm the identity of the bound protein. The addition of the antibody caused a loss of shifted complex and retention of fluorescent signal within the well of the gel, suggesting creation of a DNA-protein-antibody complex that was too large to efficiently migrate through the gel matrix (Fig 4.4). This was confirmed by the greatest fluorescence intensity shown (3.7-fold more) in the well containing the antibody compared with the other wells. This positive result and the mobility pattern of the HNF4α antibody-protein-DNA complex were confirmed by another EMSA using the HNF4α antibody with HNF4α binding site consensus as a probe (Fig 4.5). These results seem to confirm the binding of HNF4α to the PZ

gene promoter region between -48 and -38, which is consistent with the reported binding region for HNF4α in the PZ gene promoter reported by Sugawara and colleagues which was between -70 and -30 relative to the translation initiation site (Sugawara *et al.* 2007).

A second hypothetical binding site for HNF4α (-310/-300) in the PZ promoter, more distal to the translation start site, and again identified by bioinformatic analysis, was also explored by EMSA. A probe containing this hypothetical distal binding site (dPZHNF4α) showed protein binding. The bound protein showed a specificity to the hypothetical dPZHNF4α site since it was not competed with up to a 200- fold excess of either HNF4α or AP1 binding site consensus sequences (Fig 3.3). However, these results indicated that the hypothetical dPZHNF4α was not bound by the HNF4α transcription factor.

The binding of a transcription factor to a binding site, however, does not always correlate to a regulatory interaction, since transcription factors in eukaryotes have been reported to bind spurious binding sites without causing a functional effect (Wunderlich and Mirny 2009). These spurious sites were expected, using information theory and simulations, to be between 10^4 and 10^6 per genome and arise every four kb by chance. Moreover, the widespread binding of transcriptional regulators found genome-wide using ChIP-chip experiments also suggests that many of these interactions may be nonfunctional and occur by chance (Harbison *et al.* 2004). Furthermore, the use of *in vitro* EMSA is not ideal for the detection of regulatory protein-DNA binding regions (Hellman and Fried 2007). The naked DNA which is used in EMSA lacks its chromatin context that works *in vivo* as a barrier for the accession of transcription factors to their potential binding sites (Demeret, Vassetzky and Mechali 2001). The role of HNF4α was, therefore, confirmed by studying the transactivation of the endogenous PZ gene in HepG2 cells in response to overexpression and knockdown of this regulatory protein. Transfection of an HNF4α expression plasmid into HepG2 cells has been used to study the role of this transcription factor in transcriptional regulation of many genes, including coagulation factors (Solaas *et al.* 2010, Pineda Torra *et al.* 2002, Greenberg *et al.* 1995, Reijnen *et al.* 1992).

The study of the endogenous PZ gene showed that the levels of PZ mRNA increased 2.5 fold ($P<0.05$, $n=6$) by overexpression of HNF4α in HepG2 cells

(Fig 4.11). Increased transcription of HNF4α was not quantified, due to the presence of the expression plasmid, however, the significant effect of the presence of the HNF4α expression plasmid on PZ transcript levels suggests that functional protein was produced. The HNF4α transactivation of the PZ gene was confirmed by knockdown experiments using smart pool siRNAs. The knockdown of HNF4α mRNA relative expression to an average 40% (~ 2.5 fold decrease) (Fig 4.14), using the smart pool siRNA, decreased the relative expression of PZ mRNA (Fig 4.15) to approximately 36% (~ 2.8 fold decrease). These data for endogenous PZ gene transactivation by HNF4α, determined by qPCR, were found to be statistically significant (p value < 0.05, $n=6$).

Off-target silencing of unrelated mRNAs containing limited complementary sequence to the targeting siRNA has been reported (Jackson *et al.* 2003, McManus and Sharp 2002). However, Smart pool siRNAs which were used in the current study, contain four predesigned siRNAs with lower concentrations (pooling strategy) targeting one gene to minimise these off-target silencing effects (Dharmacon Technologies Literature) and have been used extensively and successfully to study transcriptional regulation of many genes (Danes *et al.* 2008, Kwei *et al.* 2008, Hariparsad *et al.* 2009, Taniguchi *et al.* 2009, Rizzo *et al.* 2010, Lahkim Bennani-Belhaj *et al.* 2010).

The role of HNF4α in the transcriptional regulation of PZ was reported by Sugawara *et al.* (2007). However, the role was determined by *in vitro* EMSA analysis and reporter gene assays where synthetic naked supercoiled DNA, lacking its chromatin context, was used. Taken together the *in vitro* protein-DNA binding experiments and analysis of the transcriptional strength of the endogenous gene within its chromatin environment in response to HNF4α in this current study seem to confirm the crucial role of HNF4α transcription factor in PZ gene regulation.

Other coagulation factors such as factor VII and factor IX, which are suggested to have evolved from the same ancestral gene as PZ (Ichinose *et al.* 1990) and have a similar tissue-restricted expression profile, have also been found to be regulated by HNF4α using protein-DNA interaction and reporter gene assays (Stauffer *et al.* 1998, Reijnen *et al.* 1992). A role for HNF4α was also reported in the regulation of the coagulation factor XI (Tarumi *et al.* 2002). The data from the current study supported the importance of HNF4α in controlling the

expression of PZ in a tissue-restricted manner, and mirror the regulatory mechanisms of other coagulation proteins that are also expressed predominantly in the liver.

8.1.2. The role of PEA3 and Ets 1 in the PZ regulation

Age-stable transcriptional regulatory elements (ASEs) have been reported in the promoter regions of both the human FIX and PC genes. Competition EMSA experiments indicated that these *cis*-regulatory elements were bound by PEA3, a member of the Ets family of transcription factors (Kurachi *et al.* 1999, Zhang, Kurachi and Kurachi 2002). However, further study revealed that the protein bound to the ASE in the FIX promoter was in fact Ets1 (Kurachi *et al.* 2009). Since members of the Ets family are detected in many cell types and their DNA binding site core consensus sequence is common, their specific binding to particular target sequences has been difficult to characterise (Spyropoulos *et al.* 2000). Proteins of this family show a structural similarity with a homology of 95%, 85% and 50% in the Ets domain, acidic domain and carboxyl-terminal tail, respectively (de Launoit *et al.* 1997).

Due to the level of homology between PZ and FIX and PC, some degree of common evolutionary ancestry is suggested (Ichinose *et al.* 1990, Fujimaki *et al.* 1998) and may also be reflected in the regulatory mechanisms controlling their expression. Therefore, the PZ promoter was studied for the presence of similar PEA3 and/or Ets1 *cis*-regulatory elements. Bioinformatic analysis (ALGGEN-PROMO) of the PZ promoter sequence identified two age-stable-like elements that showed close homology to the age-stable elements found in the human FIX (hFIX) and PC (hPC) gene promoters. These two elements in the PZ promoter, referred to as the proximal PZ ASE (pPZ ASE (-114/-108)) and the distal protein Z ASE (dPZ ASE (-816/-810)), were analysed for protein binding using EMSA analysis in the HepG2 cell model. EMSAs showed protein from the HepG2 NE binding to both ASEs. Competitive EMSA suggested specificity of the protein bound to both pPZ and dPZ ASEs since no competition of the bound protein was seen with a 200-fold excess of unlabelled non-specific competitors (AP1 and HNF3B binding site consensuses) (Fig 5.3 and 5.5). The bound protein was, however, strongly competed with a 200-fold excess of the unlabelled PEA3/Ets1 binding site consensus, with a 4-fold decrease in the intensity of the pPZ ASE complex 1 and loss of complex 2, and a 6-fold

decrease of the dPZ ASE complex 1 and loss of complex 2. These data are consistent with the previous studies on the ASEs of the evolutionary related genes, hFIX and hPC (Kurachi *et al.* 2009, Zhang, Kurachi and Kurachi 2002). Although, Kurachi *et al.* (2009) stated that Ets1 binds only to the DNA sequence G/CAGGAAG, this DNA binding sequence was found to be bound by PEA3 in the polyomavirus enhancer (Martin *et al.* 1988) and also suggested to be bound by PEA3 in a study of the DNA binding specificity of PEA3 using nuclear extract from mouse FM3A cells and bacterial extract (Xin *et al.* 1992). Moreover, the Ets 1 protein was also found to bind AGAGGATGTG in the T-cell receptor α gene enhancer (Ho *et al.* 1990) and GCA/CGGAAGT in the stromelysin promoter (Wasylyk *et al.* 1991), which are both different from the Ets1 binding sequence seen in the hFIX promoter (Kurachi *et al.* 2009).

The protein bound to dPZ ASE was also found to be strongly competed (7-fold decrease) with a 200-fold excess of the pPZ ASE, suggesting the same protein binding to both sites (Fig 5.5). However, the protein bound to the pPZ ASE showed a negligible competition with the dPZ ASE, with only a 1.1-fold and 1.2-fold decrease of the binding of complex 1 and complex 2, respectively (Fig 5.3). Further bioinformatic analysis of the region spanning the pPZ ASE revealed a hypothetical binding site for CREB located in this region. The site overlaps with the putative PEA3 site, sharing two base pairs. Competitive EMSA using the probe spanning the pPZ ASE showed relatively strong competition with a 100-fold excess of the CREB binding site consensus compared to negligible competition with a 100-fold excess of non-specific competitors (AP1 and SP1) (Fig 5.4), suggesting the protein bound to the region flanking the pPZ ASE may be CREB. However, a study found that a mutated CREB binding site consensus competed the protein bound to the probe flanking this region (Sugawara *et al.* 2007). The mutation in the CREB binding site consensus used in Sugawara's study, AC to TG, was in two highly conserved bases (TFSEARCH) and, therefore, suggested that the protein bound to this region may not in fact be CREB.

To specifically identify whether the protein bound to the pPZ ASE was PEA3, Ets1 or CREB, antibodies raised against these proteins were used in supershift EMSA analyses. The results (Fig 5.8, Fig 5.9 and Fig 5.10) indicated that the protein bound to the pPZ ASE-like elements did not interact with any of these

antibodies, although a decrease in the intensity of the protein-DNA complex was shown in the presence of the PEA3 antibody (Fig 5.8). However, the use of PEA3 and CREB in supershift analyses with their respective binding site consensus probes could not be optimised for protein-antibody interactions (See Section 3.3). The results from the supershift analyses with the PEA3 and CREB binding site consensus sequences suggest the lack of a supershift was due to the lack of antibody binding and did not necessarily indicate that the protein bound to the pPZ ASE was something other than PEA3, since competitive EMSA showed a very strong competition of the bound protein with the PEA3 binding site consensus.

Supershift EMSA analyses for the dPZ ASE probe were also performed using the PEA3 and Ets1 antibodies (Fig 5.11). The results again did not show an interaction of the PEA3 with the dPZ ASE probe, however, the intensity of the protein-DNA complex increased 4-fold in the presence of the Ets1 antibody. Kurachi *et al.* (2009) reported that the interaction of the Ets1 antibody with the ASE in the hFIX promoter, in supershift EMSA, was shown by a dramatic increase in the intensity of the protein-DNA complex without shift in the position of the complex (Kurachi *et al.* 2009), which was consistent with and supports the results of the Ets1 supershift EMSA in the current study.

Bioinformatic analysis identified other hypothetical binding sites for PEA3 spanning the regions -1190 to -1184 (PZPEA3(1)) and -1219 to -1214 (PZPEA3(2)) relative to the translation start site in the PZ gene promoter. These two regions showed DNA-protein interaction using EMSAs. These complexes, however, showed no specific binding as they were competed with both PEA3 and AP1 consensus sequences (Fig 5.6 and Fig 5.7).

To further confirm these findings 'in cell' PEA3 and CREB were studied for their role in the transactivation of the endogenous PZ gene within its chromatin context. The overexpression of PEA3 in HepG2 cells increased the relative expression of PZ mRNA 1.9 fold, suggesting a role for this protein in the control of transcription of PZ, either directly or indirectly. This increase of PZ mRNA relative expression, which was determined by qPCR, was found to be statistically significant ($P<0.05$, $n=6$). The introduction of foreign DNA in the transfection may cause an off-target effect on the expression of the gene of interest. Therefore, the pCAT00 plasmid was transfected into the HepG2 cells

using the Tfx-20 transfection reagent, which was used throughout the current study and has been used successfully in the transfection of HepG2 cell to study gene transcription (Dongol *et al.* 2007). The transfection of pCAT00 was found to have no effect on the expression of PZ mRNA (Fig 4.12), and therefore, confirm that the increase in the relative expression of PZ mRNA was due to the overexpression of PEA3 rather than a non-specific effect of transfecting plasmid DNA. The increase in PZ mRNA suggests a transactivational role for PEA3 in PZ expression. The upregulatory effect of PEA3 has previously been reported on many genes, such as, vimentin in mammary epithelial and tumour cells (Chen *et al.* 1996) and human transmembrane mucin mRNA in Pancreatic cancer cells (Fauquette *et al.* 2005). The relatively moderate increase of PZ mRNA in response to PEA3 overexpression, although it is statistically significant, may, however, be due to the very high baseline levels of PEA3 mRNA in HepG2 cells masking the true transactivatory effect of this protein on PZ gene transcription. The high levels of PEA3 in HepG2, which were found to be 1767-fold that of normal human liver cells (Fig 5.20), may already exceed the effect threshold on the PZ promoter. The high level of PEA3 mRNA in HepG2 cells compared to that in normal liver tissue may reflect the role of this protein in the progression of tumours, since it is expressed in many cancerous cells (Higashino *et al.* 1993, Davidson *et al.* 2003, Shepherd *et al.* 2001).

Knockdown of PEA3 was performed in HepG2 cells, using a smart pool siRNA, to confirm its role in the transactivation of PZ transcription The knockdown of PEA3 , although the relative expression of PEA3 mRNA was reduced 80%, was not statistically significant ($P=0.1$, $n=5$). Moreover, the mean change in PZ expression upon PEA3 knockdown was also not statistically significant ($P=0.5$, $n=6$).

The knock down of CREB in this study, using small interfering RNA (siRNA) techniques, showed an increase in the PZ mRNA by 1.4 fold (Fig 5.19) in response to a 75% reduction of CREB mRNA (Fig 5.18). The reduction of the CREB mRNA by the siRNA was found to be statistically significant ($P<0.05$,, $n=7$). However, the effect on the PZ mRNA was not found to be statistically significant ($P=0.2$,$n=7$). CREB was previously reported to repress the transcription of both the fos (Ofir *et al.* 1991) and c-jun promoters (Lamph *et al.* 1990). Moreover, the levels of CREB mRNA in liver tissue were found to be

half those found in HepG2 cells (Fig 5.20). Therefore, the knockdown of CREB in HepG2 may not have been significant enough to show its maximum effect on the levels of PZ mRNA. This hypothesis may explain the relatively small increase in PZ mRNA levels as a result of the CREB knockdown. The siRNA used to knockdown CREB, although designed for high potency and specificity (Applied Biosystems, UK), is only a single sequence, and therefore, targeting one region of the mRNA. The use of multiple separate siRNAs, targeting discrete regions of the mRNA, can help to eliminate off-target knockdown effects which may occur using individual siRNAs (Saxena, Jonsson and Dutta 2003). The fact that an individual siRNA was used in the transfection, rather than a pool of siRNAs increases the possibility of off-target silencing. Therefore, off-target silencing of another protein, in addition to CREB, which has an upregulatory effect on the PZ gene, may be a potential, if remote, cause of the non-significant down regulatory effect of CREB on PZ mRNA relative expression. The expression of CREB is fairly common across different tissues, including the liver (Yanai *et al.* 2005), but its regulatory role has not been reported with any other coagulation factors; therefore, it may not be involved in the regulation of PZ gene expression. However, more techniques, such as, transactivation by CREB overexpression could be done to confirm any role for CREB in the regulation of PZ gene expression.

8.1.3. Hormonal regulation and inflammatory response of the PZ gene

One possible source of variation in transcription levels of PZ might be as a result of transcriptional modifiers such as inflammation and/or response to hormones. The modulation of PZ gene expression by steroid hormones was predicted by a number of studies such as the reported high levels of PZ during pregnancy (Quack Loetscher *et al.* 2005, Ramsay *et al.* 2005), with the use of oral contraceptives (Al-Shanqeeti *et al.* 2005), and in obese pregnant women (Ramsay *et al.* 2005). Review of the literature mentioned above and in chapter seven suggests two possible models of steroid hormone regulation of PZ expression. Firstly, that upon higher levels of androgen/lower levels of oestrogen the expression of PZ, as a suggested anticoagulant, decreases, shifting haemostasis to a procoagulant state. Alternatively, with higher levels of androgen/lower levels of oestrogen PZ expression may increase to maintain the haemostatic balance with an increase observed in both procoagulant and

anticoagulant activity. The effect of steroid hormones on PZ was, therefore, studied in the current study by stimulation of the HepG2 cells by the candidate hormone after overexpression of the respective receptor. The HepG2 cells were tested for their expression of the proposed receptors of steroid hormones compared to normal human liver tissue. The relative expression of GR mRNA in HepG2 was found to be 65% compared to its relative expression in the normal liver tissue, while no expression of AR mRNA or OR mRNA was detected in HepG2, which is consistent with previous studies (Harnish *et al.* 1998, Barkhem *et al.* 2002, Nakagama *et al.* 1991). The lack of these receptors might suggest that HepG2 cells are not the appropriate model to study the effect of oestrogen and androgen on PZ gene transcription. However, studies to investigate the effect of androgen and oestrogen on the transcriptional regulation of genes in the HepG2 cell model, after the transfection of their receptors, have been reported previously in many studies (Yuen and Leung 2008, Tang, Norlin and Wikvall 2007, Farsetti *et al.* 2001, Harnish *et al.* 1998, Sabbah *et al.* 1998). In the current study, AR, GR and OR were overexpressed in HepG2 cells (See Section 2.4.4) using the respective expression plasmids and stimulated by Dihydrotestosterone, Dexamethasone and 17 β-estradiol, respectively.

In the current study, steroid hormones were found to have no significant effect on the expression of the endogenous PZ gene in HepG2 cells at the level of transcription. Stimulation with steroid hormones after the overexpression of their receptors in HepG2 cells resulted in no effect on PZ mRNA relative expression in response to androgen (n=4), a 1.3 fold average (n=4) increase in response to glucocorticoid and a 1.2 fold average (n=7) increase in response to oestrogen using qPCR (Fig 7.6). The modest increase of PZ mRNA in response to glucocorticoid and oestrogen was not found to be statistically significant ($P>0.05$). However, these small, if non-significant increases, when taken together with the data on PZ levels in different hormonal states, may still suggest a role for oestrogen and glucocorticoid hormones in regulating PZ transcription. The increased levels of 17 β-estradiol and androstenedione throughout pregnancy (Soldin *et al.* 2005, Castracane *et al.* 1998, Yoshida *et al.* 1982) may contribute to the high levels of PZ reported in pregnancy (Quack Loetscher *et al.* 2005, Ramsay *et al.* 2005). This is consistent with the current

study where increased levels of PZ were found in response to oestrogen, although the data was not statistically significant. Furthermore, oestrogen plays a role in reducing the transcriptional activity of FVII (Di Bitondo *et al.* 2002) and FV with increased levels of FV mRNA observed in the bone marrow of ovariectomized mice after long term treatment with 17β-estradiol (Moverare *et al.* 2004). These studies demonstrated that oestrogen is involved in regulating the expression of coagulation factors with some degree of homology to PZ and it is therefore possible that similar regulatory mechanisms controlling expression may be present in the PZ promoter. The lack of statistically significant results may be due, to some extent, to the number of samples used (n=7) which could be considered as relatively small given the small increase in PZ mRNA relative expression (1.2-fold). Moreover, the efficiency of the transfection, which was not quantified, may also contribute to the small effect on PZ mRNA relative expression and, therefore, to the lack of statistical significance.

17-hydroxyprogesterone is produced during glucocorticoid synthesis with levels increasing during pregnancy (Tulchinsky *et al.* 1972). As PZ levels increase during pregnancy and a modest, if non-significant increase in PZ mRNA is observed in the current transfection data, a role could be suggested for glucocorticoids in elevating PZ levels in pregnant women. Moreover, other coagulation factors such as FVII, which are suggested to have evolved from the same ancestral gene as PZ, in addition to FVIII and XI, have all been shown to increase in humans treated with dexamethasone (Brotman *et al.* 2006). The increase of these coagulation factors in response to dexamethasone may also suggest a biological effect of glucocorticoids on PZ which in the current study were not observed to be statistically significant. Furthermore, the relatively small number of samples (n= 4) used to determine the effect of glucocorticoid on PZ may contribute to the lack of statistical significance.

Another transcriptional modifier which may have a role in the wide range of variation in PZ levels among normal individuals could be the effect of the inflammatory response on this protein. The gene transcription of acute phase proteins contributes to most of the increase or decrease of their biosynthesis (Steel and Whitehead 1994). The role of inflammation in modulating the level of expression of PZ was suggested by Raczowski and colleagues when they

found a strong correlation between the levels of this protein with the negative acute phase protein, transferrin in the serum of patients with rheumatoid arthritis (Raczkowski *et al.* 1987). The decrease in plasma PZ levels in response to increasing levels of interleukin-6 (IL-6) in acute leukaemia and non-Hodgkin's lymphoma (Undar, Karadogan and Ozturk 1999) also suggested that PZ may be a negative acute phase protein. IL-6 has been found to regulate genes of many acute phase proteins (Baumann and Gauldie 1994) and causes thrombotic diathesis, by increasing the fibrinogen levels in plasma in humans with inflammation (Undar, Karadogan and Ozturk 1999). Moreover, other coagulation factors have been reported as acute phase proteins such as TF and FVIII, due to their induction by LPS in monocytes and in the human blood circulation, respectively (Hall, Vos and Bertina 1999, Begbie *et al.* 2000), raising the possibility that PZ, as a component of the coagulation system, may also be an acute phase protein.

LPS is an outer membrane endotoxin glycoprotein which upon detection triggers a defence mechanism in both vertebrates and invertebrates (Schumann *et al.* 1990). This endotoxin stimulates mononuclear phagocytes to produce cytokines that are important in the inflammatory response and initiation of the immune system (Herrmann and Mertelsmann 1989). LPS was tested for its effect on the transcriptional regulation of the endogenous PZ gene in HepG2 cells (See Section 2.4.6). Stimulation of HepG2 cells with bacterial (*Salmonella typhimurium*) LPS (10µg/ml) for 30 minutes was found to decrease the levels of PZ mRNA. The levels of PZ mRNA were further reduced at the 1 hour time point, suggesting that PZ may be a negative acute phase protein. The levels of PZ mRNA increased at 2 hours of stimulation and had returned to normal by 4 and 24 hours (Fig 7.7). Although the time course profile of the change in expression mirrors that of other acute phase proteins the changes in PZ mRNA were not found to be significant by the Kruskal Wallis test ($P>0.05$, $n=7$). A bidirectional relationship has been reported between inflammation and coagulation. This relationship is exemplified by the activation of the coagulation system by inflammation and the effect of coagulation on the inflammatory response (Levi, van der Poll and Buller 2004). The effect of coagulation on inflammation was, for instance, shown by the anti-inflammatory effect of activated protein C. Deletion of one allele in the protein C gene in mice was

shown to increase the levels of circulating proinflammatory cytokines (Levi, de Jonge and van der Poll 2001). The induction of TF and FVIII by LPS and the reduction of PZ levels by increasing levels of IL-6, as mentioned above, also show the inflammatory effect on the coagulation system. The effect of LPS on PZ in the current study shows a similar pattern of regulation to that shown by TF (Hall, Vos and Bertina 1999). TF, which was suggested to be an acute phase protein, was induced by LPS stimulation at 30 minute and 1 hour and started to return back to normal after 2 hours. This time course effect correlates with the response of PZ to LPS in this study in which PZ decreased at time points 30 minutes and 1 hour and started to return back to normal after 2 hours. The increase of TF and the decrease in PZ in response to LPS correspond to TF's role as a procoagulant and PZ's suggested role as an anticoagulant. Therefore, the results from the current study, although not statistically significant, still maintain the possibility that PZ may be a negative acute phase protein.

The current state of knowledge regarding the transacting factors that bind *cis*-regulatroy elements within the PZ gene promoter and may regulate gene transcription is summarised in Fig 8.1.

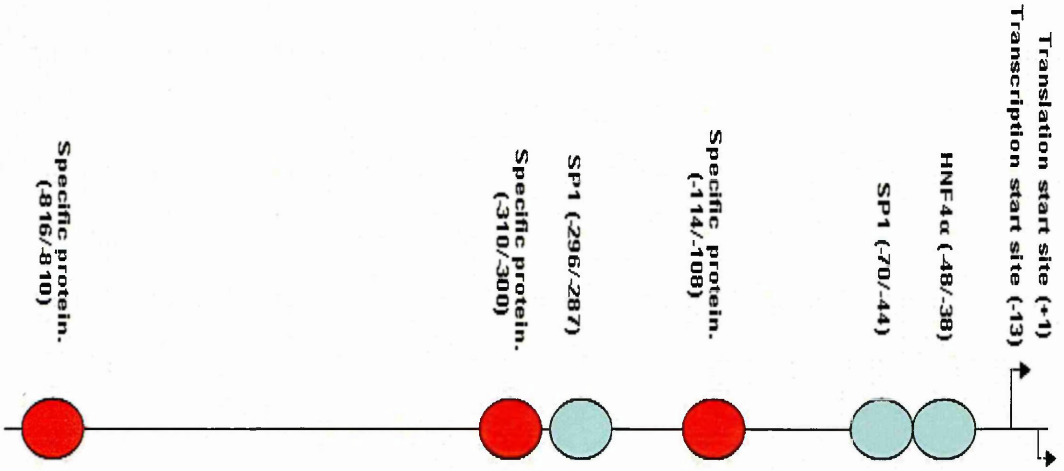


Fig 8.1. Transcription factors and proteins binding sites in the human protein Z gene promoter. The binding sites are defined in relation to the translation start site. Blue circles represent previously reported transcription factor binding sites, including the HNF4α site which has been confirmed in the present study. Red circles are binding sites that have been identified in the current study, to be bound by specific proteins.

Although the current study has provided new information and corroborative evidence regarding the transcriptional regulation of PZ, such as the contribution of HNF4α to controlling the liver-restricted expression profile, all the techniques used are considered as *in vitro* and may not accurately represent the *in vivo* situation. For instance, EMSA, which was used extensively in the current study to define protein-DNA interactions, was performed using synthetic probes interacting with the nuclear proteins extracted from a liver carcinoma cell line (HepG2). This may affect the accuracy of the specific protein binding, since these probes lack the *in vivo* barrier for the accession of transcription factors represented by the chromatin context (Demeret, Vassetzky and Mechali 2001). Nor do isolated sites allow the synergistic interactions of adjacent bound factors EMSA is also not able to define the location of protein binding (Carey and Smale 2001) without initial information based on *in silico* analysis or previous studies reported in the literature. The footprinting assay is one of the techniques which may overcome some of these limitations by allowing a greater proportion of promoter sequence to be analysed for binding site occupancy. A new development of this assay with a greater degree of sensitivity was performed in the current study, although optimisation of the DNA-protein interaction remains to be completed.

The nuclear extract which was used throughout this body of work was prepared from a cell line (HepG2 cells) which is genetically modified, and therefore, the protein constitution of the nuclear extract in HepG2 is different from that of normal liver cells. The variation in the protein constitution, which was confirmed in this study (Fig 4.16, Fig 5.20, Fig 7.5), may have contributed to the difficulties in the optimisation of the antibody-protein interaction in some supershift EMSAs using binding site consensus sequences as well as protein-DNA binding in the footprinting assay. However, the footprinting assay is not able to identify the protein bound to the DNA. Therefore, the footprinting assay is mainly used to define the location of protein-DNA interaction sites as a first step to studying transcriptional regulation of a gene and the protein binding to the DNA can be identified by other techniques, such as the supershift assay which was used in the current study. The identity of the bound protein could also be identified using other techniques, such as the chromatin immunoprecipitation assay (ChIP). This assay is more relevant in a cellular context than the supershift

assay since it is used to define the identity of the bound protein in the chromatin context. The assay was performed in this study to confirm and identify of proteins found to bind specifically regions of the PZ gene promoter by EMSA; however, the optimisation of the assay remains to be completed.

Finally, transient transfection of HepG2 cells which has been used extensively in the study of transcriptional regulation of many genes (Cereghini 1996) was also used in the current study. The data of gene expression studies of the HepG2 endogenous PZ gene, in this study, by overexpression and knockdown of the candidate protein using qPCR showed the role of both PEA3 and HNF4α in the transcriptional regulation of PZ. These data, refer, of course, to the relative expression of mRNA rather than protein levels, and therefore, it would be prudent to substantiate these findings in the future by techniques such as western blotting or ELISA to define protein levels. Moreover, the efficiency of transfection in order to investigate the effect of the overexpressed candidate protein was not determined in this study and, therefore, the non-significant data obtained from the overexpression technique, such as, with the steroid hormone receptors, may lack significance due to inefficient transfection and, therefore, overexpression in a small percentage of cells. qPCR was used to monitor the efficiency of transfection, however, the efficiency of the overexpression was not quantified due to the presence of the expression plasmid in the reverse transcriptase negative control. Many other techniques to monitor the transfection efficiency which were reported previously, such as, the use of the green fluorescent protein (Moore *et al.* 2010, Tang *et al.* 2009, Moede *et al.* 2001, Ha *et al.* 1996) and the use of the beta glycosidase enzyme (Sun, Zhang and Zhang 2004, Bode *et al.* 2001) could have been used to confirm whether the steroid hormones investigated in this study have an effect. However, the efficiency of the transfections for the other investigated proteins, HNF4α, PEA3 and CREB was suggested by the positive effect of the candidate protein on the endogenous PZ and/or confirmed by qPCR of the silenced protein using siRNAs.

8.2. Conclusion

It can be concluded from the data shown in this thesis that HNF4α is able to bind to the PZ gene promoter and has an upregulatory role in the expression of the PZ gene when assessed in the model system HepG2. The binding of this transcription factor was shown between -48 and -38 of the promoter by competitive EMSAs and supershift assay while its role in the regulation of the endogenous PZ gene was confirmed by both overexpression ($P<0.05$, n=6) and knockdown (p value < 0.05, n=6) techniques. Another region (-310/-300), which by *in silico* prediction was expected to be bound by HNF4α, was found to be bound specifically by a protein using EMSA. However, competitive EMSA indicated that the protein bound to this region was not HNF4α.

Other proteins were also detected to bind specifically to regions within the PZ promoter. An initial hypothesis suggested that two of these regions (pPZ ASE and dPZ ASE) could have been PEA3 or Ets1 binding sites which are similar to those found in other conserved factors. Competitive EMSA initially indicated that both regions were bound by PEA3 or Ets1. Moreover, the protein bound to dPZ ASE (-816/-810) was identified as Ets1 using the supershift assay.

Further study using competitive EMSAs suggested that a region spanning the pPZ ASE (-120/-101) could have been a CREB binding site. The protein bound to this region was found to be competed by CREB, however, the identity of the bound protein is still not confirmed, although a study by Sugawara (Sugawara *et al.* 2007) in the same region, using a mutant CREB binding site consensus, suggested it is not CREB binding. Gene expression studies of the endogenous PZ gene, in this study, showed the role of PEA3, by its overexpression ($P<0.05$, n=6) in the HepG2 cell model. However, CREB knockdown indicated no role for this protein in the transcriptional regulation of the endogenous PZ gene ($P=0.2$, n=7) which supports the hypothesis that the protein bound to the pPZ ASE (-120/-101) is not CREB. The current study also suggested that steroid hormones, androgen, glucocorticoid and oestrogen have no significant role in the regulation of the endogenous PZ gene. However, the potential role for steroid hormones in controlling PZ expression suggested in previous studies in addition to the unconfirmed transfection efficiency of the steroid hormone receptors, the mild increase of PZ mRNA in response to glucocorticoid (1.3 fold, n=4) and oestrogen (1.2 fold, n=7) shown in this study may still suggest a

potential role of these hormones in the regulation of PZ gene expression, and therefore, further study is needed to confirm the results.

It has also been suggested that PZ may be an acute phase protein (Raczkowski *et al.* 1987, Undar, Karadogan and Ozturk 1999). The results of the current study, although not statistically significant ($P>0.05$, $n=7$), would seem to correlate with this hypothesis with a transient decrease in PZ mRNA levels observed in response to bacterial LPS within the HepG2 model system. This response of PZ to LPS is similar to the time course response of other coagulation factors to LPS and, along with the literature, does suggest PZ may be a negative acute phase protein.

8.3. Future work

There are a number of areas that the present study has been able to provide corroborative evidence regarding the regulatory mechanisms of transcriptional regulation of PZ, such as the contribution of HNF4α to controlling the liver-restricted expression profile. However, there are a number of other areas where the current data has given glimpses into potential regulatory mechanisms that could account for observed phenotypic data without conclusive answers; such as the role of steroid hormones and the potential of PZ being an acute phase protein. Although the current data with regard to these aspects are not significant, the observed trends may prove to be real effects when studied further. Transient transfections with monitoring of the efficiency in further study, especially with the steroid hormone receptors, may provide more conclusive results for the effect of the steroid hormones on the transcriptional regulation of the PZ gene. In addition, specific protein binding sites have been identified that were characterized for the potential binding of PEA3 or Ets1, again without definitive answers as to the proteins' identity. Further studies, using other techniques, such as, ChIP should enable confirmation of the proteins occupying these regions. Moreover, optimisation of the footprinting technique to be used in such study can reveal additional proteins binding to the PZ gene promoter. These proteins, then, can be identified by the ChIP technique, and in turn, give more information about the regulation of the PZ gene.

Expanding the potential regulatory mechanisms investigated, the thyroid hormone T3 has been found to induce expression of coagulation factors such as thrombin and factor X and is, therefore, suggested as playing a significant role in the blood coagulation process (Shih *et al.* 2004). The identification of the role of this hormone in controlling coagulation components suggests the possibility that PZ could also be regulated by thyroid hormones, as it has already shown some hormonal response; further work must therefore be done to examine the role of both the thyroid and steroid hormones in controlling PZ gene regulation.

In addition, a number of SNPs have been reported in the PZ gene promoter and such variation in potential cis-regulatory regions has previously been demonstrated as being a potential transcription modifier (NCBI). Although initial data using reporter gene assays and a model system suggest that the -13 polymorphism does not affect transcription levels (Winship *et al.* 2005), *in vivo* studies of this polymorphic variation found that it did affect the levels of plasma PZ in patients with acute stroke (Staton et al, 2005). Patients with acute stroke were found to have significantly higher PZ concentrations compared with normal controls and the patients with the A-13G AA genotype had the highest PZ concentrations while those with the GG genotype showed the lowest concentrations. It would therefore, be of benefit to characterise the effects of such allelic variation.

Insights into the transcriptional regulation of PZ have been provided, but much remains to be done in trying to define the exact mechanisms by which PZ expression is controlled.

ADAMS, C. and FRIED, M. G. (2007). Analysis of protein-DNA equilibria by native gel electrophoresis. In: SCHUCK, P. (ed.). *Protein interactions: biophysical approaches for the study of component reversible systems.* New York, Academic Press, 5, 417-446.

AGRAWAL, N., *et al.* (2003). RNA interference: Biology, mechanism, and applications. *Microbiology and molecular biology reviews : MMBR,* **67** (4), 657-685.

AIRD, W. C. (2003). Hemostasis and irreducible complexity. *Journal of thrombosis and haemostasis : JTH,* **1** (2), 227-230.

AL-SHANQEETI, A., *et al.* (2005). Protein Z and protein Z-dependent protease inhibitor. determinants of levels and risk of venous thrombosis. *Thrombosis and haemostasis,* **93** (3), 411-413.

AMARA, U., *et al.* (2008). Interaction between the coagulation and complement system. *Advances in experimental medicine and biology,* **632** , 71-79.

ANDERSON, P. (2010). Post-transcriptional regulons coordinate the initiation and resolution of inflammation. *Nature reviews.immunology,* **10** (1), 24-35.

ANDREW, M., *et al.* (1987). Development of the human coagulation system in the full-term infant. *Blood,* **70** (1), 165-172.

ANDREWS, N. C. and FALLER, D. V. (1991). A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. *Nucleic acids research,* **19** (9), 2499.

ANGERS, M., *et al.* (2005). *In vivo* DNase I-mediated footprinting analysis along the human bradykinin B1 receptor (BDKRB1) gene promoter: Evidence for cell-specific regulation. *The biochemical journal,* **389** (Pt 1), 37-46.

APARICIO, O., GEISBERG, J. V. and STRUHL, K. (2004). Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences *in vivo. Current protocols in cell biology / editorial board, juan S.bonifacino ...[et al.],* **Chapter 17** , Unit 17.7.

ARBINI, A. A., *et al.* (1997). Severe factor VII deficiency due to a mutation disrupting a hepatocyte nuclear factor 4 binding site in the factor VII promoter. *Blood,* **89** (1), 176-182.

AUER, H., *et al.* (2003). Chipping away at the chip bias: RNA degradation in microarray analysis. *Nature genetics,* **35** (4), 292-293.

AUMAIS, J. P., *et al.* (1996). Function of directly repeated half-sites as response elements for steroid hormone receptors. *The journal of biological chemistry,* **271** (21), 12568-12577.

BAILLAT, D., *et al.* (2002). ETS-1 transcription factor binds cooperatively to the palindromic head to head ETS-binding sites of the stromelysin-1 promoter by

counteracting autoinhibition. *The journal of biological chemistry,* **277** (33), 29386-29398.

BAJZAR, L., NESHEIM, M. E. and TRACY, P. B. (1996). The profibrinolytic effect of activated protein C in clots formed from plasma is TAFI-dependent. *Blood,* **88** (6), 2093-2100.

BARKHEM, T., *et al.* (2002). pS2 gene expression in HepG2 cells: Complex regulation through crosstalk between the estrogen receptor alpha, an estrogen-responsive element, and the activator protein 1 response element. *Molecular pharmacology,* **61** (6), 1273-1283.

BATISTA, W. L., *et al.* (2007). Identification of transcription elements in the 5' intergenic region shared by LON and MDJ1 heat shock genes from the human pathogen paracoccidioides brasiliensis. evaluation of gene expression. *Fungal genetics and biology : FG & B,* **44** (5), 347-356.

BAUMANN, H. and GAULDIE, J. (1994). The acute phase response. *Immunology today,* **15** (2), 74-80.

BEATO, M., *et al.* (1989). DNA regulatory elements for steroid hormones. *Journal of steroid biochemistry,* **32** (5), 737-747.

BEATO, M. and KLUG, J. (2000). Steroid hormone receptors: An update. *Human reproduction update,* **6** (3), 225-236.

BECKER, R. C. (2005). Cell-based models of coagulation: A paradigm in evolution. *Journal of thrombosis and thrombolysis,* **20** (1), 65-68.

BEGBIE, M., *et al.* (2000). The factor VIII acute phase response requires the participation of NFkappaB and C/EBP. *Thrombosis and haemostasis,* **84** (2), 216-222.

BERNSTEIN, E., *et al.* (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature,* **409** (6818), 363-366.

BIECHE, I., *et al.* (2004). Expression of PEA3/E1AF/ETV4, an ets-related transcription factor, in breast tumors: Positive links to MMP2, NRG1 and CGB expression. *Carcinogenesis,* **25** (3), 405-411.

BITTINGER, M. A., *et al.* (2004). Activation of cAMP response element-mediated gene expression by regulated nuclear transport of TORC proteins. *Current biology : CB,* **14** (23), 2156-2161.

BLANCO, J. C., *et al.* (1998). The histone acetylase PCAF is a nuclear receptor coactivator. *Genes & development,* **12** (11), 1638-1651.

BODE, J. G., *et al.* (2001). The MKK6/p38 mitogen-activated protein kinase pathway is capable of inducing SOCS3 gene expression and inhibits IL-6-induced transcription. *Biological chemistry,* **382** (10), 1447-1453.

BOJOVIC, B. B. and HASSELL, J. A. (2001). The PEA3 ets transcription factor comprises multiple domains that regulate transactivation and DNA binding. *The journal of biological chemistry,* **276** (6), 4509-4521.

BOUCHER, P. D., PIECHOCKI, M. P. and HINES, R. N. (1995). Partial characterization of the human CYP1A1 negatively acting transcription factor and mutational analysis of its cognate DNA recognition sequence. *Molecular and cellular biology,* **15** (9), 5144-5151.

BRADFORD, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry,* **72** , 248-254.

BRENOWITZ, M., *et al.* (1986). Quantitative DNase footprint titration: A method for studying protein-DNA interactions. *Methods in enzymology,* **130** , 132-181.

BRINDLE, P., NAKAJIMA, T. and MONTMINY, M. (1995). Multiple protein kinase A-regulated events are required for transcriptional induction by cAMP. *Proceedings of the national academy of sciences of the united states of america,* **92** (23), 10521-10525.

BROTMAN, D. J., *et al.* (2006). Effects of short-term glucocorticoids on hemostatic factors in healthy volunteers. *Thrombosis research,* **118** (2), 247-252.

BROWN, T. A. and MCKNIGHT, S. L. (1992). Specificities of protein-protein and protein-DNA interaction of GABP alpha and two newly defined ets-related proteins. *Genes & development,* **6** (12B), 2502-2512.

BROZE, G. J.,Jr (2001). Protein Z-dependent regulation of coagulation. *Thrombosis and haemostasis,* **86** (1), 8-13.

BROZE, G. J.,Jr (2003). The rediscovery and isolation of TFPI. *Journal of thrombosis and haemostasis : JTH,* **1** (8), 1671-1675.

BROZE, G. J.,Jr and MILETICH, J. P. (1984). Human protein Z. *The journal of clinical investigation,* **73** (4), 933-938.

BUGGE, T. H., *et al.* (1996). Fatal embryonic bleeding events in mice lacking tissue factor, the cell-associated initiator of blood coagulation. *Proceedings of the national academy of sciences of the united states of america,* **93** (13), 6258-6263.

BURATOWSKI, S. and CHODOSH, L. A. (2001). Mobility shift DNA-binding assay using gel electrophoresis. *Current protocols in molecular biology / edited by frederick M.ausubel ...[et al.],* **Chapter 12** , Unit 12.2.

BURCH, J. B., *et al.* (1988). Two functional estrogen response elements are located upstream of the major chicken vitellogenin gene. *Molecular and cellular biology,* **8** (3), 1123-1131.

BUSTIN, S. A. (2000). Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of molecular endocrinology,* **25** (2), 169-193.

BUSTIN, S. A., *et al.* (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry,* **55** (4), 611-622.

CALKHOVEN, C. F., MULLER, C. and LEUTZ, A. (2002). Translational control of gene expression and disease. *Trends in molecular medicine,* **8** (12), 577-583.

CAREY, J. (1988). Gel retardation at low pH resolves trp repressor-DNA complexes for quantitative study. *Proceedings of the national academy of sciences of the united states of america,* **85** (4), 975-979.

CAREY, M. and SMALE, S. T. (2001). *Transcriptional regulation in eukaryotes: Concepts, strategies, and techniques.* 2nd ed., Cold Spring Harbor Laboratory Press.

CARLEZON, W. A.,Jr, DUMAN, R. S. and NESTLER, E. J. (2005). The many faces of CREB. *Trends in neurosciences,* **28** (8), 436-445.

CARRASQUILLO, M. M., *et al.* (2010). Concordant association of insulin degrading enzyme gene (IDE) variants with IDE mRNA, abeta, and alzheimer's disease. *PloS one,* **5** (1), e8764.

CASSLER, M. R., *et al.* (1999). Drunken-cell footprints: Nuclease treatment of ethanol-permeabilized bacteria reveals an initiation-like nucleoprotein complex in stationary phase replication origins. *Nucleic acids research,* **27** (23), 4570-4576.

CASTRACANE, V. D., *et al.* (1998). Maternal serum androgens in human pregnancy: Early increases within the cycle of conception. *Human reproduction (oxford, england),* **13** (2), 460-464.

CEREGHINI, S. (1996). Liver-enriched transcription factors and hepatocyte differentiation. *The FASEB journal : Official publication of the federation of american societies for experimental biology,* **10** (2), 267-282.

CHAKRAVARTI, D., *et al.* (1996). Role of CBP/P300 in nuclear receptor signalling. *Nature,* **383** (6595), 99-103.

CHALEPAKIS, G., *et al.* (1988). Differential gene activation by glucocorticoids and progestins through the hormone regulatory element of mouse mammary tumor virus. *Cell,* **53** (3), 371-382.

CHANDRASEKARAN, V., *et al.* (2008). Computational study of the putative active form of protein Z (PZa): Sequence design and structural modeling. *Protein science : A publication of the protein society,* **17** (8), 1354-1361.

CHEN, J. D. and EVANS, R. M. (1995). A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature,* **377** (6548), 454-457.

CHEN, J. H., *et al.* (1996). PEA3 transactivates vimentin promoter in mammary epithelial and tumor cells. *Oncogene,* **13** (8), 1667-1675.

CHEN, L., *et al.* (2003). G3 domains of aggrecan and PG-M/versican form intermolecular disulfide bonds that stabilize cell-matrix interaction. *Biochemistry,* **42** (27), 8332-8341.

CHENG, C. K., *et al.* (2002). Characterization of a new upstream GnRH receptor promoter in human ovarian granulosa-luteal cells. *Molecular endocrinology (baltimore, md.),* **16** (7), 1552-1564.

CHOI, G., *et al.* (2006). The relationship between inflammation and the coagulation system. *Swiss medical weekly : Official journal of the swiss society of infectious diseases, the swiss society of internal medicine, the swiss society of pneumology,* **136** (9-10), 139-144.

CICINNATI, V. R., *et al.* (2008). Validation of putative reference genes for gene expression studies in human hepatocellular carcinoma using real-time quantitative RT-PCR. *BMC cancer,* **8** , 350.

COLLAS, P. (2009). The state-of-the-art of chromatin immunoprecipitation. *Methods in molecular biology (clifton, N.J.),* **567** , 1-25.

CORRAL, J., *et al.* (2007). Protein Z/Z-dependent protease inhibitor (PZ/ZPI) anticoagulant system and thrombosis. *British journal of haematology,* **137** (2), 99-108.

CORDINGLEY, M. G. and HAGER, G. L. (1988). Binding of multiple factors to the MMTV promoter in crude and fractionated nuclear extracts. *Nucleic acids research,* **16** (2), 609-628.

COUREY, A. J., *et al.* (1989). Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell,* **59** (5), 827-836.

COUREY, A. J. and TJIAN, R. (1988). Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell,* **55** (5), 887-898.

CROSSLEY, M., *et al.* (1992). Recovery from hemophilia B leyden: An androgen-responsive element in the factor IX promoter. *Science (new york, N.Y.),* **257** (5068), 377-379.

CULLEN, K. E., KLADDE, M. P. and SEYFRED, M. A. (1993). Interaction between transcription regulatory regions of prolactin chromatin. *Science (new york, N.Y.),* **261** (5118), 203-206.

DANA, S. L., *et al.* (1994). Novel estrogen response elements identified by genetic selection in yeast are differentially responsive to estrogens and

antiestrogens in mammalian cells. *Molecular endocrinology (baltimore, md.),* **8** (9), 1193-1207.

DANES, C. G., *et al.* (2008). 14-3-3 zeta down-regulates P53 in mammary epithelial cells and confers luminal filling. *Cancer research,* **68** (6), 1760-1767.

DAVIDSON, B., *et al.* (2003). PEA3 is the second ets family transcription factor involved in tumor progression in ovarian carcinoma. *Clinical cancer research : An official journal of the american association for cancer research,* **9** (4), 1412-1419.

DAVIDSON, C. J., TUDDENHAM, E. G. and MCVEY, J. H. (2003). 450 million years of hemostasis. *Journal of thrombosis and haemostasis : JTH,* **1** (7), 1487-1494.

DAVIE, E. W. and RATNOFF, O. D. (1964). Waterfall sequence for intrinsic blood clotting. *Science (new york, N.Y.),* **145** , 1310-1312.

DAVIS, A. E.,3rd (2004). Biological effects of C1 inhibitor. *Drug news & perspectives,* **17** (7), 439-446.

DAWSON, S. J., MORRIS, P. J. and LATCHMAN, D. S. (1996). A single amino acid change converts an inhibitory transcription factor into an activator. *The journal of biological chemistry,* **271** (20), 11631-11633.

DE LAUNOIT, Y., *et al.* (1997). Structure-function relationships of the PEA3 group of ets-related transcription factors. *Biochemical and molecular medicine,* **61** (2), 127-135.

DEEPAK, S., *et al.* (2007). Real-time PCR: Revolutionizing detection and expression analysis of genes. *Current genomics,* **8** (4), 234-251.

DEMERET, C., VASSETZKY, Y. and MECHALI, M. (2001). Chromatin remodelling and DNA replication: From nucleosomes to loop domains. *Oncogene,* **20** (24), 3086-3093.

DHEDA, K., *et al.* (2004). Validation of housekeeping genes for normalizing RNA expression in real-time PCR. *BioTechniques,* **37** (1), 112-4, 116, 118-9.

DI BITONDO, R., *et al.* (2002). Oestrogenic repression of human coagulation factor VII expression mediated through an oestrogen response element sequence motif in the promoter region. *Human molecular genetics,* **11** (7), 723-731.

DIGNAM, J. D., *et al.* (1983). Eukaryotic gene transcription with purified components. *Methods in enzymology,* **101** , 582-598.

DITTMER, J. (2003). The biology of the Ets1 proto-oncogene. *Molecular cancer,* **2** , 29.

DONALDSON, L. W., *et al.* (1996). Solution structure of the ETS domain from murine ets-1: A winged helix-turn-helix DNA binding motif. *The EMBO journal,* **15** (1), 125-134.

DONGOL, B., *et al.* (2007). The acyl-CoA thioesterase I is regulated by PPARalpha and HNF4alpha via a distal response element in the promoter. *Journal of lipid research,* **48** (8), 1781-1791.

DUNCAN, S. A., *et al.* (1994). Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proceedings of the national academy of sciences of the united states of america,* **91** (16), 7598-7602.

ELBASHIR, S. M., *et al.* (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature,* **411** (6836), 494-498.

ELBASHIR, S. M., LENDECKEL, W. and TUSCHL, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & development,* **15** (2), 188-200.

ELNITSKI, L., *et al.* (2006). Locating mammalian transcription factor binding sites: A survey of computational and experimental techniques. *Genome research,* **16** (12), 1455-1464.

EL-SANKARY, W., *et al.* (2002). Glucocorticoid-mediated induction of CYP3A4 is decreased by disruption of a protein: DNA interaction distinct from the pregnane X receptor response element. *Drug metabolism and disposition: The biological fate of chemicals,* **30** (9), 1029-1034.

ENSLEN, H., *et al.* (1994). Characterization of Ca2+/calmodulin-dependent protein kinase IV. role in transcriptional regulation. *The journal of biological chemistry,* **269** (22), 15520-15527.

ERSDAL-BADJU, E., *et al.* (1997). Identification of the antithrombin III heparin binding site. *The journal of biological chemistry,* **272** (31), 19393-19400.

ESCRIVA, H., *et al.* (1997). Ligand binding was acquired during evolution of nuclear receptors. *Proceedings of the national academy of sciences of the united states of america,* **94** (13), 6803-6808.

ESMON, C. T. (2001). The normal role of activated protein C in maintaining homeostasis and its relevance to critical illness. *Critical care (london, england),* **5** (2), S7-12.

ESMON, C. T. (2003). The protein C pathway. *Chest,* **124** (3 Suppl), 26S-32S.

EVANS, R. M. (1988). The steroid and thyroid hormone receptor superfamily. *Science (new york, N.Y.),* **240** (4854), 889-895.

FALBO, K. B. and SHEN, X. (2006). Chromatin remodeling in DNA replication. *Journal of cellular biochemistry,* **97** (4), 684-689.

FARSETTI, A., *et al.* (1998). Orphan receptor hepatocyte nuclear factor-4 antagonizes estrogen receptor alpha-mediated induction of human coagulation factor XII gene. *Endocrinology,* **139** (11), 4581-4589.

FARSETTI, A., *et al.* (2001). Inhibition of ERalpha-mediated trans-activation of human coagulation factor XII gene by heteromeric transcription factor NF-Y. *Endocrinology,* **142** (8), 3380-3388.

FAUQUETTE, V., *et al.* (2005). The antagonistic regulation of human MUC4 and ErbB-2 genes by the ets protein PEA3 in pancreatic cancer cells: Implications for the proliferation/differentiation balance in the cells. *The biochemical journal,* **386** (Pt 1), 35-45.

FELDMAN, R. J., SEMENTCHENKO, V. I. and WATSON, D. K. (2003). The epithelial-specific ets factors occupy a unique position in defining epithelial proliferation, differentiation and carcinogenesis. *Anticancer research,* **23** (3A), 2125-2131.

FICKETT, J. W. and WASSERMAN, W. W. (2000). Discovery and modeling of transcriptional regulatory regions. *Current opinion in biotechnology,* **11** (1), 19-24.

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. *The journal of biological chemistry,* **270** (20), 11828-11838.

FIRE, A., *et al.* (1998). Potent and specific genetic interference by double-stranded RNA in caenorhabditis elegans. *Nature,* **391** (6669), 806-811.

FOURRIER, F., *et al.* (1992). Septic shock, multiple organ failure, and disseminated intravascular coagulation. compared patterns of antithrombin III, protein C, and protein S deficiencies. *Chest,* **101** (3), 816-823.

FRANCHINI, M. (2006). Hemostasis and aging. *Critical reviews in oncology/hematology,* **60** (2), 144-151.

FREEDMAN, L. P., *et al.* (1988). The function and structure of the metal coordination sites within the glucocorticoid receptor DNA binding domain. *Nature,* **334** (6182), 543-546.

FRIED, M. and CROTHERS, D. M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic acids research,* **9** (23), 6505-6525.

FRIED, M. G. and CROTHERS, D. M. (1983). CAP and RNA polymerase interactions with the lac promoter: Binding stoichiometry and long range effects. *Nucleic acids research,* **11** (1), 141-158.

FUJIMAKI, K., *et al.* (1998). The gene for human protein Z is localized to chromosome 13 at band q34 and is coded by eight regular exons and one alternative exon. *Biochemistry,* **37** (19), 6838-6846.

FULLWOOD, M. J. and RUAN, Y. (2009). ChIP-based methods for the identification of long-range chromatin interactions. *Journal of cellular biochemistry,* **107** (1), 30-39.

GALANG, C. K., *et al.* (2004). Changes in the expression of many ets family transcription factors and of potential target genes in normal mammary tissue and tumors. *The journal of biological chemistry,* **279** (12), 11281-11292.

GALAS, D. J. and SCHMITZ, A. (1978). DNAse footprinting: A simple method for the detection of protein-DNA binding specificity. *Nucleic acids research,* **5** (9), 3157-3170.

GAMBA, G., *et al.* (1998). Bleeding tendency of unknown origin and protein Z levels. *Thrombosis research,* **90** (6), 291-295.

GARNER, M. M. and REVZIN, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: Application to components of the escherichia coli lactose operon regulatory system. *Nucleic acids research,* **9** (13), 3047-3060.

GERMAIN, P., *et al.* (2006). Overview of nomenclature of nuclear receptors. *Pharmacological reviews,* **58** (4), 685-704.

Ghildiyal, M. and Zamore, P. D. (2009). Small silencing RNAs: An expanding universe. *Nature reviews.genetics,* 10 (2), 94-108.

GIBSON, N. J. (2006). The use of real-time PCR methods in DNA sequence variation analysis. *Clinica chimica acta; international journal of clinical chemistry,* **363** (1-2), 32-47.

GILMARTIN, P. M. and BOWLER, C. (2002). *Molecular plant biology: A practical approach.* 1st ed., New York, Oxford University Press. 2.

GONZALEZ, G. A. and MONTMINY, M. R. (1989). Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell,* **59** (4), 675-680.

GRAILHE, P., *et al.* (1993). Neutralization of plasminogen activator inhibitor-1 (PAI-1) by activated protein C is species-dependent. *Fibrinolysis,* **7** (2), 123-133.

GRAVES, B. J., *et al.* (1998). Autoinhibition as a transcriptional regulatory mechanism. *Cold spring harbor symposia on quantitative biology,* **63** , 621-629.

GRAVES, B. J. and PETERSEN, J. M. (1998). Specificity within the ets family of transcription factors. *Advances in cancer research,* **75** , 1-55.

GROUPP, E. R. and DONOVAN-PELUSO, M. (1996). Lipopolysaccharide induction of THP-1 cells activates binding of c-jun, ets, and egr-1 to the tissue factor promoter. *The journal of biological chemistry,* **271** (21), 12423-12430.

GUILLE, M. J. and KNEALE, G. G. (1997). Methods for the analysis of DNA-protein interactions. *Molecular biotechnology,* **8** (1), 35-52.

GUNTHER, C. V., *et al.* (1990). Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the moloney murine sarcoma virus. *Genes & development,* **4** (4), 667-679.

GUO, D. F. and INAGAMI, T. (1994). Epidermal growth factor-enhanced human angiotensin II type 1 receptor. *Hypertension,* **23** (6 Pt 2), 1032-1035.

GUTIERREZ-HARTMANN, A., DUVAL, D. L. and BRADFORD, A. P. (2007). ETS transcription factors in endocrine systems. *Trends in endocrinology and metabolism: TEM,* **18** (4), 150-158.

HA, D. S., *et al.* (1996). Use of the green fluorescent protein as a marker in transfected leishmania. *Molecular and biochemical parasitology,* **77** (1), 57-64.

HAGER, K., *et al.* (1989). Blood coagulation factors in the elderly. *Archives of gerontology and geriatrics,* **9** (3), 277-282.

HALL, A. J., PEAKE, I. R. and WINSHIP, P. R. (2006). Regulation of the human protein S gene promoter by liver enriched transcription factors. *British journal of haematology,* **135** (4), 538-546.

HALL, A. J., VOS, H. L. and BERTINA, R. M. (1999). Lipopolysaccharide induction of tissue factor in THP-1 cells involves jun protein phosphorylation and nuclear factor kappaB nuclear translocation. *The journal of biological chemistry,* **274** (1), 376-383.

HAMON, M. A. and COSSART, P. (2008). Histone modifications and chromatin remodeling during bacterial infections. *Cell host & microbe,* **4** (2), 100-109.

HAMPSEY, M. and REINBERG, D. (1999). RNA polymerase II as a control panel for multiple coactivator complexes. *Current opinion in genetics & development,* **9** (2), 132-139.

HAMPSHIRE, A. J., *et al.* (2007). Footprinting: A method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods (san diego, calif.),* **42** (2), 128-140.

HAN, S. S., *et al.* (2002). Curcumin suppresses activation of NF-kappaB and AP-1 induced by phorbol ester in cultured human promyelocytic leukemia cells. *Journal of biochemistry and molecular biology,* **35** (3), 337-342.

HAN, X., FIEHLER, R. and BROZE, G. J.,Jr (1998). Isolation of a protein Z-dependent plasma protease inhibitor. *Proceedings of the national academy of sciences of the united states of america,* **95** (16), 9250-9255.

HAN, X., FIEHLER, R. and BROZE, G. J.,Jr (2000). Characterization of the protein Z-dependent protease inhibitor. *Blood,* **96** (9), 3049-3055.

HANNON, G. J. (2002). RNA interference. *Nature,* **418** (6894), 244-251.

HARBISON, C. T., *et al.* (2004). Transcriptional regulatory code of a eukaryotic genome. *Nature,* **431** (7004), 99-104.

HARD, T., *et al.* (1990). Solution structure of the glucocorticoid receptor DNA-binding domain. *Science (new york, N.Y.),* **249** (4965), 157-160.

HARIPARSAD, N., *et al.* (2009). Identification of pregnane-X receptor target genes and coactivator and corepressor binding to promoter elements in human hepatocytes. *Nucleic acids research,* **37** (4), 1160-1173.

HARNISH, D. C., *et al.* (1998). Estrogen regulation of the apolipoprotein Al gene promoter through transcription cofactor sharing. *The journal of biological chemistry,* **273** (15), 9270-9278.

HASSAN, Y. I. and ZEMPLENI, J. (2008). A novel, enigmatic histone modification: Biotinylation of histones by holocarboxylase synthetase. *Nutrition reviews,* **66** (12), 721-725.

HATTORI, Y. and MAITANI, Y. (2005). Folate-linked nanoparticle-mediated suicide gene therapy in human prostate cancer and nasopharyngeal cancer with herpes simplex virus thymidine kinase. *Cancer gene therapy,* **12** (10), 796-809.

HAYFLICK, L. and MOORHEAD, P. S. (1961). The serial cultivation of human diploid cell strains. *Experimental cell research,* **25** , 585-621.

HE, J., CHENG, Q. and XIE, W. (2010). Minireview: Nuclear receptor-controlled steroid hormone synthesis and metabolism. *Molecular endocrinology (baltimore, md.),* **24** (1), 11-21.

HE, L. and HANNON, G. J. (2004). MicroRNAs: Small RNAs with a big role in gene regulation. *Nature reviews.genetics,* **5** (7), 522-531.

HEEB, M. J., *et al.* (2002). Low protein Z levels and risk of ischemic stroke: Differences by diabetic status and gender. *Blood cells, molecules & diseases,* **29** (2), 139-144.

HELLMAN, L. M. and FRIED, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nature protocols,* **2** (8), 1849-1861.

HELTEMES, L. M., TUGGLE, C. K. and LAMONT, S. J. (1999). Tissue expression patterns of chicken octamer-binding proteins. *Poultry science,* **78** (4), 574-578.

HERRMANN, F. and MERTELSMANN, R. (1989). Polypeptides controlling hematopoietic cell development and activation. I. *in vitro* results. *Blut,* **58** (3), 117-128.

HERTZ, R., *et al.* (1998). Fatty acyl-CoA thioesters are ligands of hepatic nuclear factor-4alpha. *Nature,* **392** (6675), 512-516.

HIGASHINO, F., *et al.* (1993). Isolation of a cDNA encoding the adenovirus E1A enhancer binding protein: A new human member of the ets oncogene family. *Nucleic acids research,* **21** (3), 547-553.

HIGUCHI, R., *et al.* (1993). Kinetic PCR analysis: Real-time monitoring of DNA amplification reactions. *Bio/technology (nature publishing company),* **11** (9), 1026-1030.

HILL, J. T., *et al.* (2011). Novel computational analysis of protein binding array data identifies direct targets of Nkx2.2 in the pancreas. *BMC bioinformatics,* **12** , 62.

HO, I. C., *et al.* (1990). Sequence-specific binding of human ets-1 to the T cell receptor alpha gene enhancer. *Science (new york, N.Y.),* **250** (4982), 814-818.

HOEFFLER, J. P., *et al.* (1988). Cyclic AMP-responsive DNA-binding protein: Structure based on a cloned placental cDNA. *Science (new york, N.Y.),* **242** (4884), 1430-1433.

HOFFBRAND, A. V., MOSS, P. A. H. and PETTIT, J. E. (2006). *Essential haematology.* 5th edition., Blackwell.

HOFFMAN, M. (2003). Remodeling the blood coagulation cascade. *Journal of thrombosis and thrombolysis,* **16** (1-2), 17-20.

HOGG, P. J. and STENFLO, J. (1991). Interaction of vitamin K-dependent protein Z with thrombin. consequences for the amidolytic activity of thrombin and the interaction of thrombin with phospholipid vesicles. *The journal of biological chemistry,* **266** (17), 10953-10958.

HOJRUP, P., JENSEN, M. S. and PETERSEN, T. E. (1985). Amino acid sequence of bovine protein Z: A vitamin K-dependent serine protease homolog. *FEBS letters,* **184** (2), 333-338.

HOLDEN, N. S. and TACON, C. E. (2010). Principles and problems of the electrophoretic mobility shift assay. *Journal of pharmacological and toxicological methods,* .

HOLLENHORST, P. C., JONES, D. A. and GRAVES, B. J. (2004). Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic acids research,* **32** (18), 5693-5702.

HORLEIN, A. J., *et al.* (1995). Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature,* **377** (6548), 397-404.

HORNSTRA, I. K. and YANG, T. P. (1993). *In vivo* footprinting and genomic sequencing by ligation-mediated PCR. *Analytical biochemistry,* **213** (2), 179-193.

HUANG, X., *et al.* (2008). Kinetic characterization of the protein Z-dependent protease inhibitor reaction with blood coagulation factor xa. *The journal of biological chemistry,* **283** (44), 29770-29783.

HUNG, H. L., *et al.* (2001). Regulation of human coagulation factor X gene expression by GATA-4 and the sp family of transcription factors. *Blood,* **97** (4), 946-951.

HUNG, M. C. and WANG, S. C. (1999). Transcriptional repression in cancer gene therapy: Targeting HER-2/neu overexpression as an example. *Gene therapy and molecular biology,* **3** , 91-101.

http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3

http://users.unimi.it/~camelot/tools/cut2.html

ICHINOSE, A., *et al.* (1990). Amino acid sequence of human protein Z, a vitamin K-dependent plasma glycoprotein. *Biochemical and biophysical research communications,* **172** (3), 1139-1144.

IKUTA, T. and KAN, Y. W. (1991). *In vivo* protein-DNA interactions at the beta-globin gene locus. *Proceedings of the national academy of sciences of the united states of america,* **88** (22), 10188-10192.

INOUE, Y., *et al.* (2006). Role of hepatocyte nuclear factor 4alpha in control of blood coagulation factor gene expression. *Journal of molecular medicine (berlin, germany),* **84** (4), 334-344.

IP, Y. C., LEE, W. M. and HAMMOND, G. L. (2000). The rabbit sex hormone-binding globulin gene: Structural organization and characterization of its 5-flanking region. *Endocrinology,* **141** (4), 1356-1365.

ISHIZUKA, M., *et al.* (2005). A zinc finger protein TZF is a novel corepressor of androgen receptor. *Biochemical and biophysical research communications,* **331** (4), 1025-1031.

JACKSON, A. L., *et al.* (2003). Expression profiling reveals off-target gene regulation by RNAi. *Nature biotechnology,* **21** (6), 635-637.

JACOBSEN, L., CALVIN, S. and LOBENHOFER, E. (2009). Transcriptional effects of transfection: The potential for misinterpretation of gene expression data generated from transiently transfected cells. *BioTechniques,* **47** (1), 617-624.

JAMES, W. H. (1993). Rheumatoid arthritis, the contraceptive pill, and androgens. *Annals of the rheumatic diseases,* **52** (6), 470-474.

JAMES, W. H. (1993). Rheumatoid arthritis, the contraceptive pill, and androgens. *Annals of the rheumatic diseases,* **52** (6), 470-474.

JANSEN, M., *et al.* (1995). Translational control of gene expression. *Pediatric research,* **37** (6), 681-686.

JANTZEN, H. M., *et al.* (1987). Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. *Cell,* **49** (1), 29-38.

JIANG, D., JARRETT, H. W. and HASKINS, W. E. (2009). Methods for proteomic analysis of transcription factors. *Journal of chromatography.A,* .

JIANG, G., *et al.* (1995). Exclusive homodimerization of the orphan receptor hepatocyte nuclear factor 4 defines a new subclass of nuclear receptors. *Molecular and cellular biology,* **15** (9), 5131-5143.

JIANG, Y., *et al.* (2009). First insight into human liver proteome from PROTEOMESKY-LIVERHu 1.0, a publicly-available database. *Journal of proteome research,* .

JING, D., *et al.* (2003). A sensitive two-color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels. *Proteomics,* 3 (7), 1172-1180.

JING, Z., *et al.* (2004). Identification of the DNA binding element of the human ZNF333 protein. *Journal of biochemistry and molecular biology,* **37** (6), 663-670.

JONES, S. and THORNTON, J. M. (2003). Protein-DNA interactions: The story so far and a new method for prediction. *Comparative and functional genomics,* **4** (4), 428-431.

JONSEN, M. D., *et al.* (1996). Characterization of the cooperative function of inhibitory sequences in ets-1. *Molecular and cellular biology,* **16** (5), 2065-2073.

KAKO, K., *et al.* (1998). Examination of DNA-binding activity of neuronal transcription factors by electrophoretical mobility shift assay. *Brain research.brain research protocols,* **2** (4), 243-249.

KANG, S. H., VIEIRA, K. and BUNGERT, J. (2002). Combining chromatin immunoprecipitation and DNA footprinting: A novel method to analyze protein-DNA interactions *in vivo. Nucleic acids research,* **30** (10), e44.

KARIM, F. D., *et al.* (1990). The ETS-domain: A new DNA-binding motif that recognizes a purine-rich core DNA sequence. *Genes & development,* **4** (9), 1451-1453.

KATAKURA, Y., ALAM, S. and SHIRAHATA, S. (1998). Immortalization by gene transfection. *Methods in cell biology,* **57** , 69-91.

KATO, S., *et al.* (1995). Widely spaced, directly repeated PuGGTCA elements act as promiscuous enhancers for different classes of nuclear receptors. *Molecular and cellular biology,* **15** (11), 5858-5867.

KATO, S., *et al.* (1992). A far upstream estrogen response element of the ovalbumin gene contains several half-palindromic 5'-TGACC-3' motifs acting synergistically. *Cell,* **68** (4), 731-742.

KEENE, J. D. and LAGER, P. J. (2005). Post-transcriptional operons and regulons co-ordinating gene expression. *Chromosome research : An international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology,* **13** (3), 327-337.

KEMKES-MATTHES, B. and MATTHES, K. J. (1995). Protein Z, a new haemostatic factor, in liver diseases. *Haemostasis,* **25** (6), 312-316.

KEMKES-MATTHES, B., *et al.* (2002). Protein Z influences the prothrombotic phenotype in factor V leiden patients. *Thrombosis research,* **106** (4-5), 183-185.

KHEOLAMAI, P. and DICKSON, A. J. (2009). Liver-enriched transcription factors are critical for the expression of hepatocyte marker genes in mES-derived hepatocyte-lineage cells. *BMC molecular biology,* **10** , 35.

KHORASANIZADEH, S. and RASTINEJAD, F. (2001). Nuclear-receptor interactions on DNA-response elements. *Trends in biochemical sciences,* **26** (6), 384-390.

KIM, C. A., *et al.* (2001). Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *The EMBO journal,* **20** (15), 4173-4182.

KIM, J. W. and AHN, Y. H. (1998). CCAAT/enhancer binding protein regulates the promoter activity of the rat GLUT2 glucose transporter gene in liver cells. *The biochemical journal,* **336 (Pt 1)** (Pt 1), 83-90.

KIM, S. and NELSON, P. G. (1998). Transcriptional regulation of the prothrombin gene in muscle. *The journal of biological chemistry,* **273** (19), 11923-11929.

KLEIN-HITPASS, L., *et al.* (1986). An estrogen-responsive element derived from the 5' flanking region of the xenopus vitellogenin A2 gene functions in transfected human cells. *Cell,* **46** (7), 1053-1061.

KLUG, A. (1999). Zinc finger peptides for the regulation of gene expression. *Journal of molecular biology,* **293** (2), 215-218.

KOBELT, K., *et al.* (2001). Protein Z in ischaemic stroke. *British journal of haematology,* **114** (1), 169-173.

KOUZARIDES, T. (2007). Chromatin modifications and their function. *Cell,* **128** (4), 693-705.

KRACHT, M. and SAKLATVALA, J. (2002). Transcriptional and post-transcriptional control of gene expression in inflammation. *Cytokine,* **20** (3), 91-106.

KUBISTA, M., *et al.* (2006). The real-time polymerase chain reaction. *Molecular aspects of medicine,* **27** (2-3), 95-125.

KURACHI, K. and KURACHI, S. (2000). Genetic mechanisms of age regulation of blood coagulation: Factor IX model. *Arteriosclerosis, thrombosis, and vascular biology,* **20** (4), 902-906.

KURACHI, K., *et al.* (2000). Genetic and molecular mechanisms of age regulation (homeostasis) of blood coagulation. *IUBMB life,* **49** (3), 189-196.

KURACHI, S., *et al.* (1999). Genetic mechanisms of age regulation of human blood coagulation factor IX. *Science (new york, N.Y.),* **285** (5428), 739-743.

KURACHI, S., *et al.* (2009). An age-related homeostasis mechanism is essential for spontaneous amelioration of hemophilia B leyden. *Proceedings of the national academy of sciences of the united states of america,* **106** (19), 7921-7926.

KWEI, K. A., *et al.* (2008). Genomic profiling identifies GATA6 as a candidate oncogene amplified in pancreatobiliary cancer. *PLoS genetics,* **4** (5), e1000081.

KYRMIZI, I., *et al.* (2006). Plasticity and expanding complexity of the hepatic transcription factor network during liver development. *Genes & development,* **20** (16), 2293-2305.

LAGET, M. P., *et al.* (1996). Two functionally distinct domains responsible for transactivation by the ets family member ERM. *Oncogene,* **12** (6), 1325-1336.

LAHKIM BENNANI-BELHAJ, K., *et al.* (2010). BLM deficiency is not associated with sensitivity to hydroxyurea-induced replication stress. *Journal of nucleic acids,* **2010** , 319754.

LAMPH, W. W., *et al.* (1990). Negative and positive regulation by transcription factor cAMP response element-binding protein is modulated by phosphorylation. *Proceedings of the national academy of sciences of the united states of america,* **87** (11), 4320-4324.

LATCHMAN, D. S. (1990). Eukaryotic transcription factors. *The biochemical journal,* **270** (2), 281-289.

LATCHMAN, D. S. (1992). Gene regulation. *BMJ (clinical research ed.),* **304** (6834), 1103-1105.

LATCHMAN, D. S. (1997). Transcription factors: An overview. *The international journal of biochemistry & cell biology,* **29** (12), 1305-1312.

LATCHMAN, D. S. (2001). Transcription factors: Bound to activate or repress. *Trends in biochemical sciences,* **26** (4), 211-213.

LATCHMAN, D. S. (2005). *Gene regulation.* 5th ed., Abingdon, Taylor and Francis Group.

LAUDET, V., *et al.* (1992). Evolution of the nuclear receptor gene superfamily. *The EMBO journal,* **11** (3), 1003-1013.

LEBLANC, B. and MOSS, T. (2009). DNase I footprinting. *Methods in molecular biology (clifton, N.J.),* **543** , 37-47.

LEE, M., *et al.* (1997). A 100 kDa protein binding to bHLH family consensus recognition sequence of RAT p53 promoter. *J. biochem. mol. biol.,* **30** (3), 205-210.

LEE, M., *et al.* (1999). *In vitro* transcription assay with the purified 40kDa NF1-like protein binding to the rat p53 promoter. *Biochemistry and molecular biology international,* **47** (3), 427-434.

LEFEVER, S., *et al.* (2009). RDML: Structured language and reporting guidelines for real-time quantitative PCR data. *Nucleic acids research,* **37** (7), 2065-2069.

LEVI, M., DE JONGE, E. and VAN DER POLL, T. (2001). Rationale for restoration of physiological anticoagulant pathways in patients with sepsis and disseminated intravascular coagulation. *Critical care medicine,* **29** (7 Suppl), S90-4.

LEVI, M., VAN DER POLL, T. and BULLER, H. R. (2004). Bidirectional relation between inflammation and coagulation. *Circulation,* **109** (22), 2698-2704.

LI, R., PEI, H. and WATSON, D. K. (2000). Regulation of ets function by protein - protein interactions. *Oncogene,* **19** (55), 6514-6523.

LI, Y., *et al.* (2003). Homogeneous assays for adenosine 5'-monophosphate-activated protein kinase. *Analytical biochemistry,* **321** (2), 151-156.

LIANG, S. H., HASSETT, C. and OMIECINSKI, C. J. (2005). Alternative promoters determine tissue-specific expression profiles of the human microsomal epoxide hydrolase gene (EPHX1). *Molecular pharmacology,* **67** (1), 220-230.

LIU, W. and SAINT, D. A. (2002). A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics. *Analytical biochemistry,* **302** (1), 52-59.

LIVAK, K. J., *et al.* (1995). Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR methods and applications,* **4** (6), 357-362.

LOPPNOW, H., *et al.* (1995). Endotoxin activates human vascular smooth muscle cells despite lack of expression of CD14 mRNA or endogenous membrane CD14. *Infection and immunity,* **63** (3), 1020-1026.

LOWE, G. D., *et al.* (1997). Epidemiology of coagulation factors, inhibitors and activation markers: The third glasgow MONICA survey. I. illustrative reference ranges by age, sex and hormone use. *British journal of haematology,* **97** (4), 775-784.

LOWRY, O. H., *et al.* (1951). Protein measurement with the folin phenol reagent. *The journal of biological chemistry,* **193** (1), 265-275.

MACFARLANE, R. G. (1964). An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier. *Nature,* **202** , 498-499.

MACKMAN, N., BRAND, K. and EDGINGTON, T. S. (1991). Lipopolysaccharide-mediated transcriptional activation of the human tissue factor gene in THP-1 monocytic cells requires both activator protein 1 and nuclear factor kappa B binding sites. *The journal of experimental medicine,* **174** (6), 1517-1526.

MACKMAN, N., TILLEY, R. E. and KEY, N. S. (2007). Role of the extrinsic pathway of blood coagulation in hemostasis and thrombosis. *Arteriosclerosis, thrombosis, and vascular biology,* **27** (8), 1687-1693.

MARI, D., COPPOLA, R. and PROVENZANO, R. (2008). Hemostasis factors and aging. *Experimental gerontology,* **43** (2), 66-73.

MARI, D., *et al.* (1995). Hypercoagulability in centenarians: The paradox of successful aging. *Blood,* **85** (11), 3144-3149.

MARTIN, M. E., *et al.* (1988). Activation of the polyomavirus enhancer by a murine activator protein 1 (AP1) homolog and two contiguous proteins. *Proceedings of the national academy of sciences of the united states of america,* **85** (16), 5839-5843.

MARTINEZ, J., *et al.* (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell,* **110** (5), 563-574.

MARUYAMA, S., *et al.* (2003). Identification of IFN regulatory factor-1 binding site in IL-12 p40 gene promoter. *Journal of immunology (baltimore, md.: 1950),* **170** (2), 997-1001.

MATA, J., MARGUERAT, S. and BAHLER, J. (2005). Post-transcriptional control of gene expression: A genome-wide perspective. *Trends in biochemical sciences,* **30** (9), 506-514.

MAVROTHALASSITIS, G. and GHYSDAEL, J. (2000). Proteins of the ETS family with transcriptional repressor activity. *Oncogene,* **19** (55), 6524-6532.

MAXAM, A. M. and GILBERT, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods in enzymology,* **65** (1), 499-560.

MAYES, J. S. and WATSON, G. H. (2004). Direct effects of sex steroid hormones on adipose tissues and obesity. *Obesity reviews : An official journal of the international association for the study of obesity,* **5** (4), 197-216.

MCDONALD, J. F., *et al.* (1997). Comparison of naturally occurring vitamin K-dependent proteins: Correlation of amino acid sequences and membrane binding properties suggests a membrane contact site. *Biochemistry,* **36** (17), 5120-5127.

MCMANUS, M. T. and SHARP, P. A. (2002). Gene silencing in mammals by small interfering RNAs. *Nature reviews.genetics,* **3** (10), 737-747.

MEDZHITOV, R. (2008). Origin and physiological roles of inflammation. *Nature,* **454** (7203), 428-435.

MEDZHITOV, R. and HORNG, T. (2009). Transcriptional control of the inflammatory response. *Nature reviews.immunology,* **9** (10), 692-703.

MERMOD, N., *et al.* (1989). The proline-rich transcriptional activator of CTF/NF-I is distinct from the replication and DNA binding domain. *Cell,* **58** (4), 741-753.

MESSEGUER, X., *et al.* (2002). PROMO: Detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics (oxford, england),* **18** (2), 333-334.

METZGER, D., *et al.* (1995). Characterization of the amino-terminal transcriptional activation function of the human estrogen receptor in animal and yeast cells. *The journal of biological chemistry,* **270** (16), 9535-9542.

MEYER, T. E. and HABENER, J. F. (1993). Cyclic adenosine 3',5'-monophosphate response element binding protein (CREB) and related transcription-activating deoxyribonucleic acid-binding proteins. *Endocrine reviews,* **14** (3), 269-290.

MIAO, C. H., *et al.* (1992). Liver-specific expression of the gene coding for human factor X, a blood coagulation factor. *The journal of biological chemistry,* **267** (11), 7395-7401.

MIKKELSEN, T. S., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature,* **448** (7153), 553-560.

MILETICH, J. P. and BROZE, G. J.,Jr (1987). Human plasma protein Z antigen: Range in normal subjects and effect of warfarin therapy. *Blood,* **69** (6), 1580-1586.

MILLER, C. L., *et al.* (2004). Evaluating RNA status for RT-PCR in extracts of postmortem human brain tissue. *BioTechniques,* **36** (4), 628-633.

Mirmira, R. (2006). Preparation of Nuclear Extracts for Gel Shifts and Westerns Blots. Accessed on 9/2006, *http://faculty.virginia.edu/mirmira/resources_files /Protocols/ Nuclear%20Extracts.htm.*

MITCHELL, P. J. and TJIAN, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science (new york, N.Y.),* **245** (4916), 371-378.

MIURA, M. and YUAN, J. (2000). Transient transfection assay of cell death genes. *Methods in enzymology,* **322** , 480-492.

MO, Y., *et al.* (1998). Structures of SAP-1 bound to DNA targets from the E74 and c-fos promoters: Insights into DNA sequence discrimination by ets proteins. *Molecular cell,* **2** (2), 201-212.

MO, Y., *et al.* (2000). Structure of the elk-1-DNA complex reveals how DNA-distal residues affect ETS domain recognition of DNA. *Nature structural biology,* **7** (4), 292-297.

MOEDE, T., *et al.* (2001). Online monitoring of stimulus-induced gene expression in pancreatic beta-cells. *Diabetes,* **50 Suppl 1** , S15-9.

MOEENREZAKHANLOU, A., NANDAN, D. and REINER, N. E. (2008). Identification of a calcitriol-regulated sp-1 site in the promoter of human CD14 using a combined western blotting electrophoresis mobility shift assay (WEMSA). *Biological procedures online,* **10** , 29-35.

MOHD-SARIP, A., *et al.* (2005). Synergistic recognition of an epigenetic DNA element by pleiohomeotic and a polycomb core complex. *Genes & development,* **19** (15), 1755-1760.

MONTE, D., *et al.* (1994). Molecular cloning and characterization of human ERM, a new member of the ets family closely related to mouse PEA3 and ER81 transcription factors. *Oncogene,* **9** (5), 1397-1406.

MONTMINY, M. (1997). Transcriptional regulation by cyclic AMP. *Annual review of biochemistry,* **66** , 807-822.

MOORE, D. D. (1990). Diversity and unity in the nuclear hormone receptors: A terpenoid receptor superfamily. *The new biologist,* **2** (1), 100-105.

MOORE, J. C., *et al.* (2010). Efficient, high-throughput transfection of human embryonic stem cells. *Stem cell research & therapy,* **1** (3), 23.

MOROZOV, A. V., *et al.* (2005). Protein-DNA binding specificity predictions with structural models. *Nucleic acids research,* **33** (18), 5781-5798.

MOVERARE, S., *et al.* (2004). Estrogen increases coagulation factor V mRNA levels via both estrogen receptor-alpha and -beta in murine bone marrow/bone. *European journal of endocrinology / european federation of endocrine societies,* **151** (2), 259-263.

MUELLER, P. R. and WOLD, B. (1989). *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science (new york, N.Y.),* **246** (4931), 780-786.

MUKAIDA, N., *et al.* (1994). Molecular mechanism of interleukin-8 gene expression. *Journal of leukocyte biology,* **56** (5), 554-558.

MUKHOPADHYAY, A., *et al.* (2008). Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in caenorhabditis elegans. *Nature protocols,* **3** (4), 698-709.

NAGARAJ, V. H., O'FLANAGAN, R. A. and SENGUPTA, A. M. (2008). Better estimation of protein-DNA interaction parameters improve prediction of functional sites. *BMC biotechnology,* **8** , 94.

NAKAGAMA, H., *et al.* (1991). Expression of androgen receptor mRNA in human hepatocellular carcinomas and hepatoma cell lines. *Hepatology (baltimore, md.),* **14** (1), 99-102.

NARAYANAN, K., *et al.* (2004). The CCAAT enhancer-binding protein (C/EBP)beta and Nrf1 interact to regulate dentin sialophosphoprotein (DSPP) gene expression during odontoblast differentiation. *The journal of biological chemistry,* **279** (44), 45423-45432.

NARDULLI, A. M. and SHAPIRO, D. J. (1992). Binding of the estrogen receptor DNA-binding domain to the estrogen response element induces DNA bending. *Molecular and cellular biology,* **12** (5), 2037-2042.

NEGRE, N., *et al.* (2006). Mapping the distribution of chromatin proteins by ChIP on chip. *Methods in enzymology,* **410** , 316-341.

NELSON, J. D., *et al.* (2006). Fast chromatin immunoprecipitation assay. *Nucleic acids research,* **34** (1), e2.

NEWBY, D. T., HADFIELD, T. L. and ROBERTO, F. F. (2003). Real-time PCR detection of brucella abortus: A comparative study of SYBR green I, 5'-exonuclease, and hybridization probe assays. *Applied and environmental microbiology,* **69** (8), 4753-4759.

NIKOLAKAKI, E., *et al.* (1993). Glycogen synthase kinase 3 phosphorylates jun family members *in vitro* and negatively regulates their transactivating potential in intact cells. *Oncogene,* **8** (4), 833-840.

NORDEEN, S. K., *et al.* (1990). Structural determinants of a glucocorticoid receptor recognition element. *Molecular endocrinology (baltimore, md.),* **4** (12), 1866-1873.

NYE, J. A., *et al.* (1992). Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes & development,* **6** (6), 975-990.

OFIR, R., *et al.* (1991). CREB represses transcription of fos promoter: Role of phosphorylation. *Gene expression,* **1** (1), 55-60.

O'HAGAN, R. C. and HASSELL, J. A. (1998). The PEA3 ets transcription factor is a downstream target of the HER2/Neu receptor tyrosine kinase. *Oncogene,* **16** (3), 301-310.

OIKAWA, T. and YAMADA, T. (2003). Molecular biology of the ets family of transcription factors. *Gene,* **303** , 11-34.

OLSON, S. T., *et al.* (1992). Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. resolution of the antithrombin conformational change contribution to heparin rate enhancement. *The journal of biological chemistry,* **267** (18), 12528-12538.

ORPHANIDES, G., LAGRANGE, T. and REINBERG, D. (1996). The general transcription factors of RNA polymerase II. *Genes & development,* **10** (21), 2657-2683.

OSTERUD, B. and FLAEGSTAD, T. (1983). Increased tissue thromboplastin activity in monocytes of patients with meningococcal infection: Related to an unfavourable prognosis. *Thrombosis and haemostasis,* **49** (1), 5-7.

PANNE, D., MANIATIS, T. and HARRISON, S. C. (2007). An atomic model of the interferon-beta enhanceosome. *Cell,* **129** (6), 1111-1123.

PARRISH, S., *et al.* (2000). Functional anatomy of a dsRNA trigger: Differential requirement for the two trigger strands in RNA interference. *Molecular cell,* **6** (5), 1077-1087.

PAYVAR, F., *et al.* (1983). Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. *Cell,* **35** (2 Pt 1), 381-392.

PENNER, G., *et al.* (2002). C/EBP DNA-binding activity is upregulated by a glucocorticoid-dependent mechanism in septic muscle. *American journal of physiology.regulatory, integrative and comparative physiology,* **282** (2), R439-44.

PERKINS, N. D. (1997). Achieving transcriptional specificity with NF-kappa B. *The international journal of biochemistry & cell biology,* **29** (12), 1433-1448.

PERSSON, K. E., *et al.* (2006). Binding of calcium to anticoagulant protein S: Role of the fourth EGF module. *Biochemistry,* **45** (35), 10682-10689.

PETRESCU, A. D., *et al.* (2002). Ligand specificity and conformational dependence of the hepatic nuclear factor-4alpha (HNF-4alpha). *The journal of biological chemistry,* **277** (27), 23988-23999.

PFAFFL, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research,* **29** (9), e45.

PINEDA TORRA, I., *et al.* (2002). Characterization of the human PPARalpha promoter: Identification of a functional nuclear receptor response element. *Molecular endocrinology (baltimore, md.),* **16** (5), 1013-1028.

POCH, M. T., *et al.* (2005). Molecular mechanisms regulating human CYP4B1 lung-selective expression. *Drug metabolism and disposition: The biological fate of chemicals,* **33** (8), 1174-1184.

PTASHNE, M. (1988). How eukaryotic transcriptional activators work. *Nature,* **335** (6192), 683-689.

PUGIN, J., *et al.* (1993). Lipopolysaccharide activation of human endothelial and epithelial cells is mediated by lipopolysaccharide-binding protein and soluble CD14. *Proceedings of the national academy of sciences of the united states of america,* **90** (7), 2744-2748.

QUACK LOETSCHER, K. C., *et al.* (2005). Protein Z in normal pregnancy. *Thrombosis and haemostasis,* **93** (4), 706-709.

RAETZ, C. R. and WHITFIELD, C. (2002). Lipopolysaccharide endotoxins. *Annual review of biochemistry,* **71** , 635-700.

RAMSAY, J. E., *et al.* (2005). Protein Z in pregnancy: Exaggerated rise in obese women. *Journal of thrombosis and haemostasis : JTH,* **3** (11), 2584-2586.

RAO, V. N., *et al.* (1993). Analysis of the DNA-binding and transcriptional activation functions of human fli-1 protein. *Oncogene,* **8** (8), 2167-2173.

RATH, B., *et al.* (2008). Molecular characterization of senescence marker protein-30 gene promoter: Identification of repressor elements and functional nuclear factor binding sites. *BMC molecular biology,* **9** , 43.

RAVI, S., *et al.* (1998). Protein Z in healthy human individuals and in patients with a bleeding tendency. *British journal of haematology,* **102** (5), 1219-1223.

RAVINDRANATH, T. M., *et al.* (2007). Plasma thrombin activatable fibrinolysis inhibitor and tissue factor pathway inhibitor changes following sepsis. *Clinical and applied thrombosis/hemostasis : Official journal of the international academy of clinical and applied Thrombosis/Hemostasis,* **13** (4), 362-368.

RAVISCIONI, M., *et al.* (2006). Evolutionary identification of a subtype specific functional site in the ligand binding domain of steroid receptors. *Proteins,* **64** (4), 1046-1057.

RECKELHOFF, J. F., *et al.* (2005). Testosterone supplementation in aging men and women: Possible impact on cardiovascular-renal disease. *American journal of physiology.renal physiology,* **289** (5), F941-8.

REIJNEN, M. J., *et al.* (1992). Disruption of a binding site for hepatocyte nuclear factor 4 results in hemophilia B leyden. *Proceedings of the national academy of sciences of the united states of america,* **89** (14), 6300-6303.

REN, B., *et al.* (2000). Genome-wide location and function of DNA binding proteins. *Science (new york, N.Y.),* **290** (5500), 2306-2309.

REVZIN, A. (1989). Gel electrophoresis assays for DNA-protein interactions. *BioTechniques,* **7** (4), 346-355.

RIDDEL, J. P.,Jr and AOUIZERAT, B. E. (2006). Genetics of von willebrand disease type 1. *Biological research for nursing,* **8** (2), 147-156.

RIDDEL, J. P.,Jr, *et al.* (2007). Theories of blood coagulation. *Journal of pediatric oncology nursing : Official journal of the association of pediatric oncology nurses,* **24** (3), 123-131.

RIPPE, K. (1997). Analysis of protein-DNA binding at equilibrium. *B.I.F. FUTURA,* (12), 20-26.

RIRIE, K. M., RASMUSSEN, R. P. and WITTWER, C. T. (1997). Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. *Analytical biochemistry,* **245** (2), 154-160.

RIZZO, G., *et al.* (2010). Functional characterization of the semisynthetic bile acid derivative INT-767, a dual farnesoid X receptor and TGR5 agonist. *Molecular pharmacology,* **78** (4), 617-630.

ROBERTSON, G., *et al.* (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature methods,* **4** (8), 651-657.

ROEDER, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends in biochemical sciences,* **21** (9), 327-335.

RUGGERI, Z. M. and WARE, J. (1993). Von willebrand factor. *The FASEB journal : Official publication of the federation of american societies for experimental biology,* **7** (2), 308-316.

RUMINY, P., *et al.* (2001). Gene transcription in hepatocytes during the acute phase of a systemic inflammation: From transcription factors to target genes. *Inflammation research : Official journal of the european histamine research society ...[et al.],* **50** (8), 383-390.

SABBAH, M., *et al.* (1998). Oestrogen receptor facilitates the formation of preinitiation complex assembly: Involvement of the general transcription factor TFIIB. *The biochemical journal,* **336 (Pt 3)** (Pt 3), 639-646.

SAMBROOK, J. and RUSSELL, D. W. (2001). *Molecular cloning: A laboratory manual.* 3rd ed., New York, CSHL press. 2.

SANDS, W. A. and PALMER, T. M. (2008). Regulating gene transcription in response to cyclic AMP elevation. *Cellular signalling,* **20** (3), 460-466.

SANTACROCE, R., *et al.* (2006). Low protein Z levels and risk of occurrence of deep vein thrombosis. *Journal of thrombosis and haemostasis : JTH,* **4** (11), 2417-2422.

SAVERY, N. J. and BUSBY, S. J. W. (1998). Mobility shift assays. In: RAPLEY, R. and WALKER, J. M. (eds.). *Molecular biomethods handbook.* 1st ed., Totowa, Humana Press Inc, 121-129.

SAXENA, S., JONSSON, Z. O. and DUTTA, A. (2003). Small RNAs with imperfect match to endogenous mRNA repress translation. implications for off-target activity of small inhibitory RNA in mammalian cells. *The journal of biological chemistry,* **278** (45), 44312-44319.

SCHMITTGEN, T. D. and ZAKRAJSEK, B. A. (2000). Effect of experimental treatment on housekeeping gene expression: Validation by real-time, quantitative RT-PCR. *Journal of biochemical and biophysical methods,* **46** (1-2), 69-81.

SCHOOR, O., *et al.* (2003). Moderate degradation does not preclude microarray analysis of small amounts of RNA. *BioTechniques,* **35** (6), 1192-6, 1198-201.

SCHREM, H., KLEMPNAUER, J. and BORLAK, J. (2002). Liver-enriched transcription factors in liver function and development. part I: The hepatocyte nuclear factor network and liver-specific gene expression. *Pharmacological reviews,* **54** (1), 129-158.

SCHREM, H., KLEMPNAUER, J. and BORLAK, J. (2004). Liver-enriched transcription factors in liver function and development. part II: The C/EBPs and D site-binding protein in cell cycle control, carcinogenesis, circadian gene regulation, liver regeneration, apoptosis, and liver-specific gene regulation. *Pharmacological reviews,* **56** (2), 291-330.

SCHULE, R., *et al.* (1988). Many transcription factors interact synergistically with steroid receptors. *Science (new york, N.Y.),* **242** (4884), 1418-1420.

SCHUMANN, R. R., *et al.* (1990). Structure and function of lipopolysaccharide binding protein. *Science (new york, N.Y.),* **249** (4975), 1429-1431.

SCHWACHTGEN, J. L., *et al.* (1997). Ets transcription factors bind and transactivate the core promoter of the von willebrand factor gene. *Oncogene,* **15** (25), 3091-3102.

SCLAVI, B. (2008). Time-resolved footprinting for the study of the structural dynamics of DNA-protein interactions. *Biochemical society transactions,* **36** (Pt 4), 745-748.

SEJIMA, H., *et al.* (1990). Primary structure of vitamin K-dependent human protein Z. *Biochemical and biophysical research communications,* **171** (2), 661-668.

SEMENTCHENKO, V. I. and WATSON, D. K. (2000). Ets target genes: Past, present and future. *Oncogene,* **19** (55), 6533-6548.

SETH, A. and WATSON, D. K. (2005). ETS transcription factors and their emerging roles in human cancer. *European journal of cancer (oxford, england : 1990),* **41** (16), 2462-2478.

SHANG, Y., *et al.* (2000). Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell,* **103** (6), 843-852.

SHARMA, D. and FONDELL, J. D. (2002). Ordered recruitment of histone acetyltransferases and the TRAP/Mediator complex to thyroid hormone-responsive promoters *in vivo*. *Proceedings of the national academy of sciences of the united states of america,* **99** (12), 7934-7939.

SHARMA, R., SINGH, M. and SHARMA, A. (2002). Polymerase chain reaction: An emerging tool for research in pharmacology. *Indian journal of pharmacology,* **34** , 229-236.

SHARROCKS, A. D. (2001). The ETS-domain transcription factor family. *Nature reviews.molecular cell biology,* **2** (11), 827-837.

SHAYWITZ, A. J. and GREENBERG, M. E. (1999). CREB: A stimulus-induced transcription factor activated by a diverse array of extracellular signals. *Annual review of biochemistry,* **68** , 821-861.

SHE, X., *et al.* (2009). Definition, conservation and epigenetics of housekeeping and tissue-enriched genes. *BMC genomics,* **10** , 269.

SHEPHERD, T. G., *et al.* (2001). The pea3 subfamily ets genes are required for HER2/Neu-mediated mammary oncogenesis. *Current biology : CB,* **11** (22), 1739-1748.

SHERRY, B. and CERAMI, A. (1988). Cachectin/tumor necrosis factor exerts endocrine, paracrine, and autocrine control of inflammatory responses. *The journal of cell biology,* **107** (4), 1269-1277.

SHI, G. X., REHMANN, H. and ANDRES, D. A. (2006). A novel cyclic AMP-dependent epac-rit signaling pathway contributes to PACAP38-mediated neuronal differentiation. *Molecular and cellular biology,* **26** (23), 9136-9147.

SHI, X. Z. and SARNA, S. K. (2005). Transcriptional regulation of inflammatory mediators secreted by human colonic circular smooth muscle cells. *American journal of physiology.gastrointestinal and liver physiology,* 289 (2), G274-84.

SHIH, C. H., *et al.* (2004). Thyroid hormone receptor-dependent transcriptional regulation of fibrinogen and coagulation proteins. *Endocrinology,* **145** (6), 2804-2814.

SHINDOH, M., HIGASHINO, F. and KOHGO, T. (2004). E1AF, an ets-oncogene family transcription factor. *Cancer letters,* **216** (1), 1-8.

SHORE, P., *et al.* (1995). Characterization of the elk-1 ETS DNA-binding domain. *The journal of biological chemistry,* **270** (11), 5805-5811.

SIDOROVA, N. Y., MURADYMOV, S. and RAU, D. C. (2005). Trapping DNA-protein binding reactions with neutral osmolytes for the analysis by gel mobility shift and self-cleavage assays. *Nucleic acids research,* **33** (16), 5145-5155.

SIGLER, P. B. (1988). Transcriptional activation. acid blobs and negative noodles. *Nature,* **333** (6170), 210-212.

SIKDER, D. and KODADEK, T. (2005). Genomic studies of transcription factor-DNA interactions. *Current opinion in chemical biology,* **9** (1), 38-45.

SLADEK, F. M., *et al.* (1990). Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. *Genes & development,* **4** (12B), 2353-2365.

SMITH, P. K., *et al.* (1985). Measurement of protein using bicinchoninic acid. *Analytical biochemistry,* **150** (1), 76-85.

SMITH, S. A. (2009). The cell-based model of coagulation. *Journal of veterinary emergency and critical care (san antonio, tex.: 2001),* **19** (1), 3-10.

SMOLIK, S. M., ROSE, R. E. and GOODMAN, R. H. (1992). A cyclic AMP-responsive element-binding transcriptional activator in drosophila melanogaster, dCREB-A, is a member of the leucine zipper family. *Molecular and cellular biology,* **12** (9), 4123-4131.

(Solaas *et al.* 2010, Pineda Torra *et al.* 2002, Greenberg *et al.* 1995, Reijnen *et al.* 1992)

SOLDIN, O. P., *et al.* (2005). Steroid hormone levels in pregnancy and 1 year postpartum using isotope dilution tandem mass spectrometry. *Fertility and sterility,* **84** (3), 701-710.

SPENCER, V. A. and DAVIE, J. R. (1999). Role of covalent modifications of histones in regulating gene expression. *Gene,* **240** (1), 1-12.

SPENCER, V. A., *et al.* (2003). Chromatin immunoprecipitation: A tool for studying histone acetylation and transcription factor binding. *Methods (san diego, calif.),* **31** (1), 67-75.

SPYROPOULOS, D. D., *et al.* (2000). Hemorrhage, impaired hematopoiesis, and lethality in mouse embryos carrying a targeted disruption of the Fli1 transcription factor. *Molecular and cellular biology,* **20** (15), 5643-5652.

STARGELL, L. A. and STRUHL, K. (1996). Mechanisms of transcriptional activation *in vivo*: Two steps forward. *Trends in genetics : TIG,* **12** (8), 311-315.

STASIOLEK, M., *et al.* (2000). Inhibitory and stimulatory effects of lactacystin on expression of nitric oxide synthase type 2 in brain glial cells. the role of ikappa B-beta. *The journal of biological chemistry,* 275 (32), 24847-24856.

STAUFFER, D. R., *et al.* (1998). Characterization of transcriptional regulatory elements in the promoter region of the murine blood coagulation factor VII gene. *The journal of biological chemistry,* **273** (4), 2277-2287.

STEIN, G. S., *et al.* (1999). Nuclear structure/gene expression interrelationships. *Journal of cellular physiology,* **181** (2), 240-250.

STEINMETZ, A. C., RENAUD, J. P. and MORAS, D. (2001). Binding of ligands and activation of transcription by nuclear receptors. *Annual review of biophysics and biomolecular structure,* **30** , 329-359.

STEVENS, A. M., DOLAN, K. M. and GREENBERG, E. P. (1994). Synergistic binding of the vibrio fischeri LuxR transcriptional activator domain and RNA polymerase to the lux promoter region. *Proceedings of the national academy of sciences of the united states of america,* **91** (26), 12619-12623.

STRAHLE, U., SCHMID, W. and SCHUTZ, G. (1988). Synergistic action of the glucocorticoid receptor with transcription factors. *The EMBO journal,* **7** (11), 3389-3395.

SUCK, D., LAHM, A. and OEFNER, C. (1988). Structure refined to 2A of a nicked DNA octanucleotide complex with DNase I. *Nature,* **332** (6163), 464-468.

SUGAWARA, H., *et al.* (2007). Regulation of human protein Z gene expression by liver-enriched transcription factor HNF-4alpha and ubiquitous factor Sp1. *Journal of thrombosis and haemostasis : JTH,* **5** (11), 2250-2258.

SUN, X., ZHANG, H. W. and ZHANG, Z. R. (2004). Transfection efficiency of pORF lacZ plasmid lipopolyplex to hepatocytes and hepatoma cells. *World journal of gastroenterology : WJG,* **10** (4), 531-534.

TABATABAI, A., FIEHLER, R. and BROZE, G. J.,Jr (2001). Protein Z circulates in plasma in a complex with protein Z-dependent protease inhibitor. *Thrombosis and haemostasis,* **85** (4), 655-660.

TAKAGI, S., *et al.* (2008). Post-transcriptional regulation of human pregnane X receptor by micro-RNA affects the expression of cytochrome P450 3A4. *The journal of biological chemistry,* **283** (15), 9674-9680.

TANG, J., *et al.* (2009). pGreen-S: A clone vector bearing absence of enhanced green fluorescent protein for screening recombinants. *Analytical biochemistry,* **388** (1), 173-174.

TANG, W., NORLIN, M. and WIKVALL, K. (2007). Regulation of human CYP27A1 by estrogens and androgens in HepG2 and prostate cells. *Archives of biochemistry and biophysics,* **462** (1), 13-20.

TANIGUCHI, F., *et al.* (2009). TAK1 activation for cytokine synthesis and proliferation of endometriotic cells. *Molecular and cellular endocrinology,* **307** (1-2), 196-204.

TARUMI, T., *et al.* (2002). Cloning and characterization of the human factor XI gene promoter: Transcription factor hepatocyte nuclear factor 4alpha (HNF-4alpha) is required for hepatocyte-specific expression of factor XI. *The journal of biological chemistry,* **277** (21), 18510-18516.

TAYLOR, J. D., ACKROYD, A. J. and HALFORD, S. E. (1994). The gel shift assay for the analysis of DNA-protein interactions. *Methods in molecular biology (clifton, N.J.),* **30** , 263-279.

THEOPHILUS, B. D. (1996). Preparation and analysis of DNA sequencing gels. *Methods in molecular biology (clifton, N.J.),* **65** , 1-9.

THOMSON, D. M., *et al.* (2008). AMP-activated protein kinase phosphorylates transcription factors of the CREB family. *Journal of applied physiology (bethesda, md.: 1985),* **104** (2), 429-438.

TICHOPAD, A., *et al.* (2003). Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic acids research,* **31** (20), e122.

TOBIAS, P. S., *et al.* (1995). Lipopolysaccharide binding protein-mediated complexation of lipopolysaccharide with soluble CD14. *The journal of biological chemistry,* **270** (18), 10482-10488.

TOMARI, Y. and ZAMORE, P. D. (2005). Perspective: Machines for RNAi. *Genes & development,* **19** (5), 517-529.

TORA, L., *et al.* (1989). The human estrogen receptor has two independent nonacidic transcriptional activation functions. *Cell,* **59** (3), 477-487.

TULCHINSKY, D., *et al.* (1972). Plasma estrone, estradiol, estriol, progesterone, and 17-hydroxyprogesterone in human pregnancy. I. normal pregnancy. *American journal of obstetrics and gynecology,* **112** (8), 1095-1100.

UNDAR, L., KARADOGAN, I. and OZTURK, F. (1999). Plasma protein Z levels inversely correlate with plasma interleukin-6 levels in patients with acute leukemia and non-hodgkin's lymphoma. *Thrombosis research,* **94** (2), 131-134.

VANDESOMPELE, J., *et al.* (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome biology,* **3** (7), RESEARCH0034.

VARGA, A. and JAMES, D. (2006). Real-time RT-PCR and SYBR green I melting curve analysis for the identification of plum pox virus strains C, EA, and W: Effect of amplicon size, melt rate, and dye translocation. *Journal of virological methods,* **132** (1-2), 146-153.

VASSE, M., *et al.* (2006). Human endothelial cells synthesize protein Z, but not the protein Z dependent inhibitor. *Thrombosis and haemostasis,* **95** (3), 519-523.

VASSE, M., *et al.* (2001). Frequency of protein Z deficiency in patients with ischaemic stroke. *Lancet,* **357** (9260), 933-934.

VASSE, M. (2008). Protein Z, a protein seeking a pathology. *Thrombosis and haemostasis,* **100** (4), 548-556.

VENTER, J. C., *et al.* (2001). The sequence of the human genome. *Science (new york, N.Y.),* **291** (5507), 1304-1351.

VERGER, A. and DUTERQUE-COQUILLAUD, M. (2002). When ets transcription factors meet their partners. *BioEssays : News and reviews in molecular, cellular and developmental biology,* **24** (4), 362-370.

VERSTEEG, H. H., *et al.* (2008). TF:FVIIa-specific activation of CREB upregulates proapoptotic proteins via protease-activated receptor-2. *Journal of thrombosis and haemostasis : JTH,* **6** (9), 1550-1557.

WARREN, A. J. (2002). Eukaryotic transcription factors. *Current opinion in structural biology,* **12** (1), 107-114.

WASYLYK, B., HAGMAN, J. and GUTIERREZ-HARTMANN, A. (1998). Ets transcription factors: Nuclear effectors of the ras-MAP-kinase signaling pathway. *Trends in biochemical sciences,* **23** (6), 213-216.

WASYLYK, C., *et al.* (1991). The c-ets oncoprotein activates the stromelysin promoter through the same elements as several non-nuclear oncoproteins. *The EMBO journal,* **10** (5), 1127-1134.

WATSON, D. K., ASCIONE, R. and PAPAS, T. S. (1990). Molecular analysis of the ets genes and their products. *Critical reviews in oncogenesis,* **1** (4), 409-436.

WATT, A. J., GARRISON, W. D. and DUNCAN, S. A. (2003). HNF4: A central regulator of hepatocyte differentiation and function. *Hepatology (baltimore, md.),* **37** (6), 1249-1253.

WELLS, J. and FARNHAM, P. J. (2002). Characterizing transcription factor binding sites using formaldehyde crosslinking and immunoprecipitation. *Methods (san diego, calif.),* **26** (1), 48-56.

WENG, M. W., *et al.* (2004). Benzo[a]pyrene diol epoxide up-regulates COX-2 expression through NF-kappaB in rat astrocytes. *Toxicology letters,* 151 (2), 345-355.

WERNER, M. H., *et al.* (1995). The solution structure of the human ETS1-DNA complex reveals a novel mode of binding and true side chain intercalation. *Cell,* **83** (5), 761-771.

WIGGIN, G. R., *et al.* (2002). MSK1 and MSK2 are required for the mitogen- and stress-induced phosphorylation of CREB and ATF1 in fibroblasts. *Molecular and cellular biology,* **22** (8), 2871-2881.

WILLOUGHBY, D. A. (1978). Inflammation. *Endeavour,* **2** (2), 57-65.

WINSHIP, P. R., *et al.* (2005). Functional analysis of A -13 G/A dimorphism at the major transcription initiation site of the protein Z gene. *Journal of thrombosis and haemostasis,* **3** (1), P1259 Abstract .

WIREN, K. M., *et al.* (1997). Transcriptional up-regulation of the human androgen receptor by androgen in bone cells. *Endocrinology,* **138** (6), 2291-2300.

WONG, M. L. and MEDRANO, J. F. (2005). Real-time PCR for mRNA quantitation. *BioTechniques,* **39** (1), 75-85.

WONG, S. H., *et al.* (2007). Zinc deficiency depresses p21 gene expression: Inhibition of cell cycle progression is independent of the decrease in p21 protein level in HepG2 cells. *American journal of physiology.cell physiology,* **292** (6), C2175-84.

WOO, A. J., *et al.* (2002). A proteomics approach for the identification of DNA binding activities observed in the electrophoretic mobility shift assay. *Molecular & cellular proteomics : MCP,* **1** (6), 472-478.

WRIGHT, S. D., *et al.* (1990). CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. *Science (new york, N.Y.),* **249** (4975), 1431-1433.

WU, W. F., URBANOWSKI, M. L. and STAUFFER, G. V. (1995). Characterization of a second MetR-binding site in the metE metR regulatory region of salmonella typhimurium. *Journal of bacteriology,* **177** (7), 1834-1839.

WUNDERLICH, Z. and MIRNY, L. A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. *Trends in genetics : TIG,* **25** (10), 434-440.

XIA, W. Y., *et al.* (2006). Expression of PEA3 and lack of correlation between PEA3 and HER-2/neu expression in breast cancer. *Breast cancer research and treatment,* **98** (3), 295-301.

XIANG, C. C., *et al.* (2003). A new strategy to amplify degraded RNA from small tissue samples for microarray studies. *Nucleic acids research,* **31** (9), e53.

XIN, J. H., *et al.* (1992). Molecular cloning and characterization of PEA3, a new member of the ets oncogene family that is differentially expressed in mouse embryonic cells. *Genes & development,* **6** (3), 481-496.

XING, X., *et al.* (2000). The ets protein PEA3 suppresses HER-2/neu overexpression and inhibits tumorigenesis. *Nature medicine,* **6** (2), 189-195.

YANAI, I., *et al.* (2005). Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics (oxford, england),* **21** (5), 650-659.

Yang, N. and Kazazian, H. H.,Jr (2006). L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nature structural & molecular biology,* 13 (9), 763-771.

YE, J., ZHANG, X. and DONG, Z. (1996). Characterization of the human granulocyte-macrophage colony-stimulating factor gene promoter: An AP1 complex and an Sp1-related complex transactivate the promoter activity that is suppressed by a YY1 complex. *Molecular and cellular biology,* **16** (1), 157-167.

YEAGER, T. R. and REDDEL, R. R. (1999). Constructing immortalized human cell lines. *Current opinion in biotechnology,* **10** (5), 465-469.

YIN, Z. F., *et al.* (2000). Prothrombotic phenotype of protein Z deficiency. *Proceedings of the national academy of sciences of the united states of america,* **97** (12), 6734-6738.

YING, B. W., FOURMY, D. and YOSHIZAWA, S. (2007). Substitution of the use of radioactivity by fluorescence for biochemical studies of RNA. *RNA (new york, N.Y.),* 13 (11), 2042-2050.

YOSHIDA, N., *et al.* (1982). Testosterone-estradiol binding globulin and sex steroids during pregnancy and at delivery. *Nippon sanka fujinka gakkai zasshi,* **34** (11), 1833-1838.

YUEN, Y. M. and LEUNG, L. K. (2008). Genistein and daidzein induced apoA-1 transactivation in hepG2 cells expressing oestrogen receptor-alpha. *The british journal of nutrition,* **99** (5), 1007-1012.

ZHANG, K., KURACHI, S. and KURACHI, K. (2002). Genetic mechanisms of age regulation of protein C and blood coagulation. *The journal of biological chemistry,* **277** (6), 4532-4540.

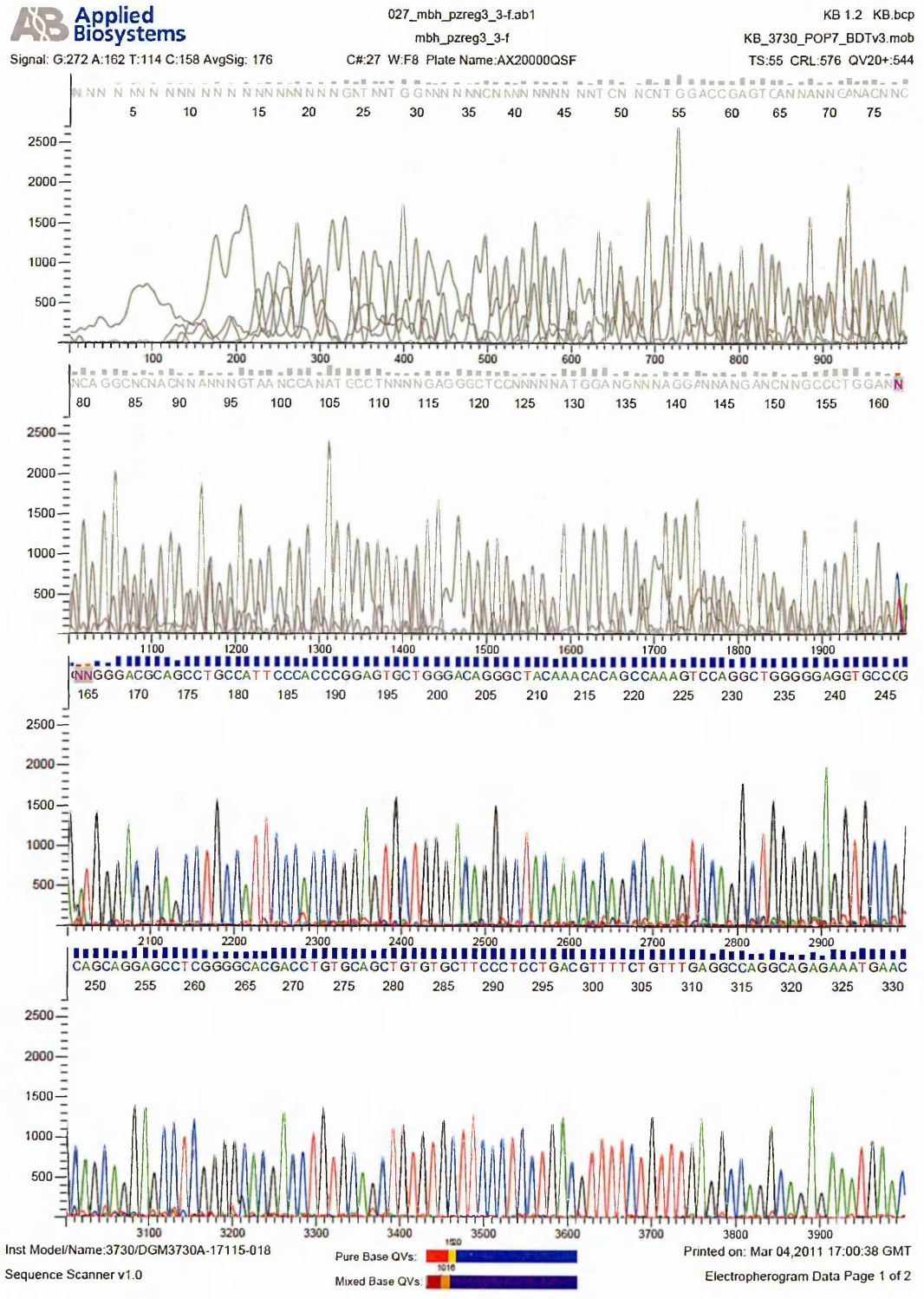
ZHANG, S. B., *et al.* (2001). Function of GATA transcription factors in hydroxyurea-induced HEL cells. *Cell research,* **11** (4), 301-310.

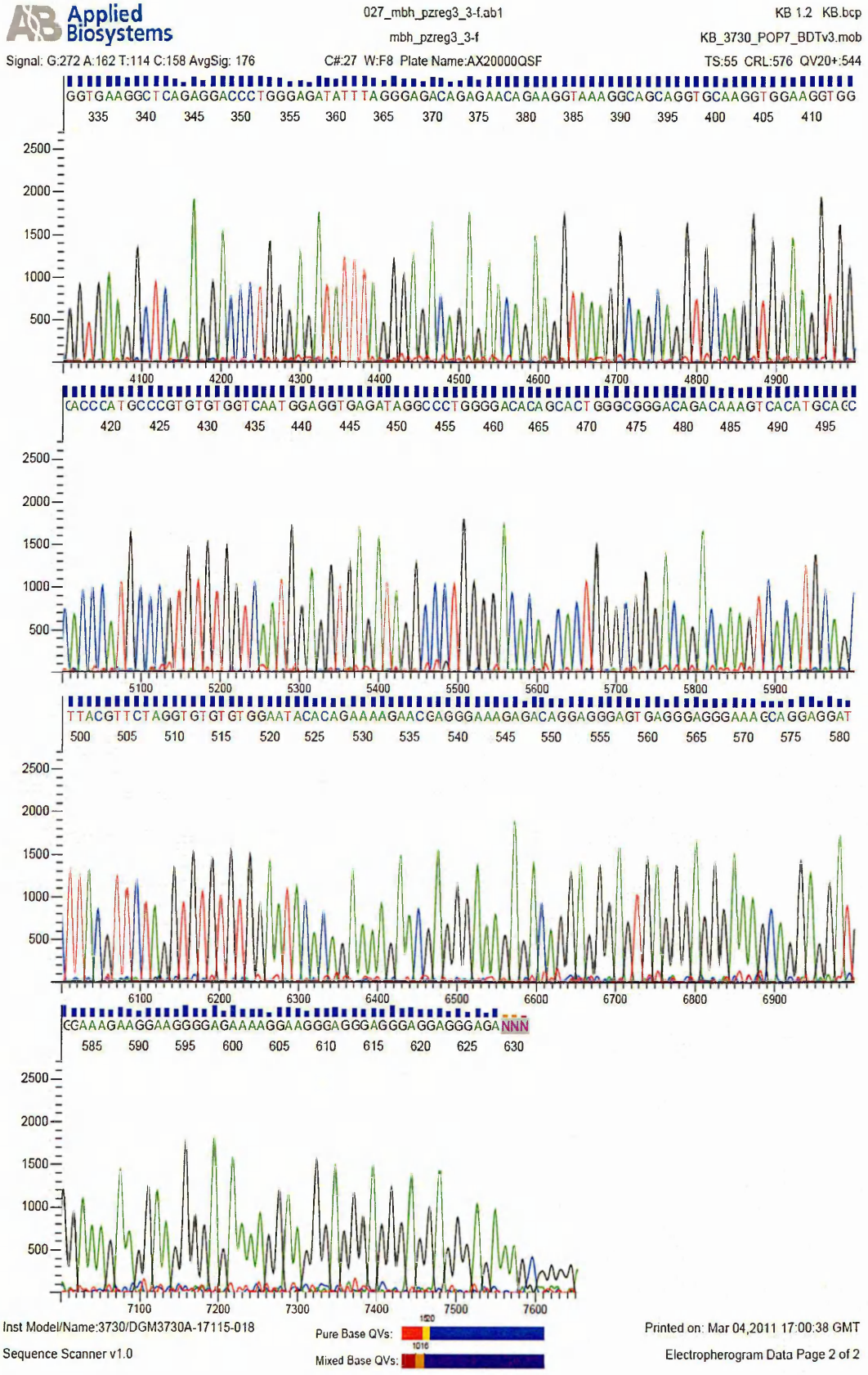
ZHAO, W., *et al.* (2011). NF-kappaB- and AP-1-mediated DNA looping regulates osteopontin transcription in endotoxin-stimulated murine macrophages. *Journal of immunology (baltimore, md.: 1950),* **186** (5), 3173-3179.

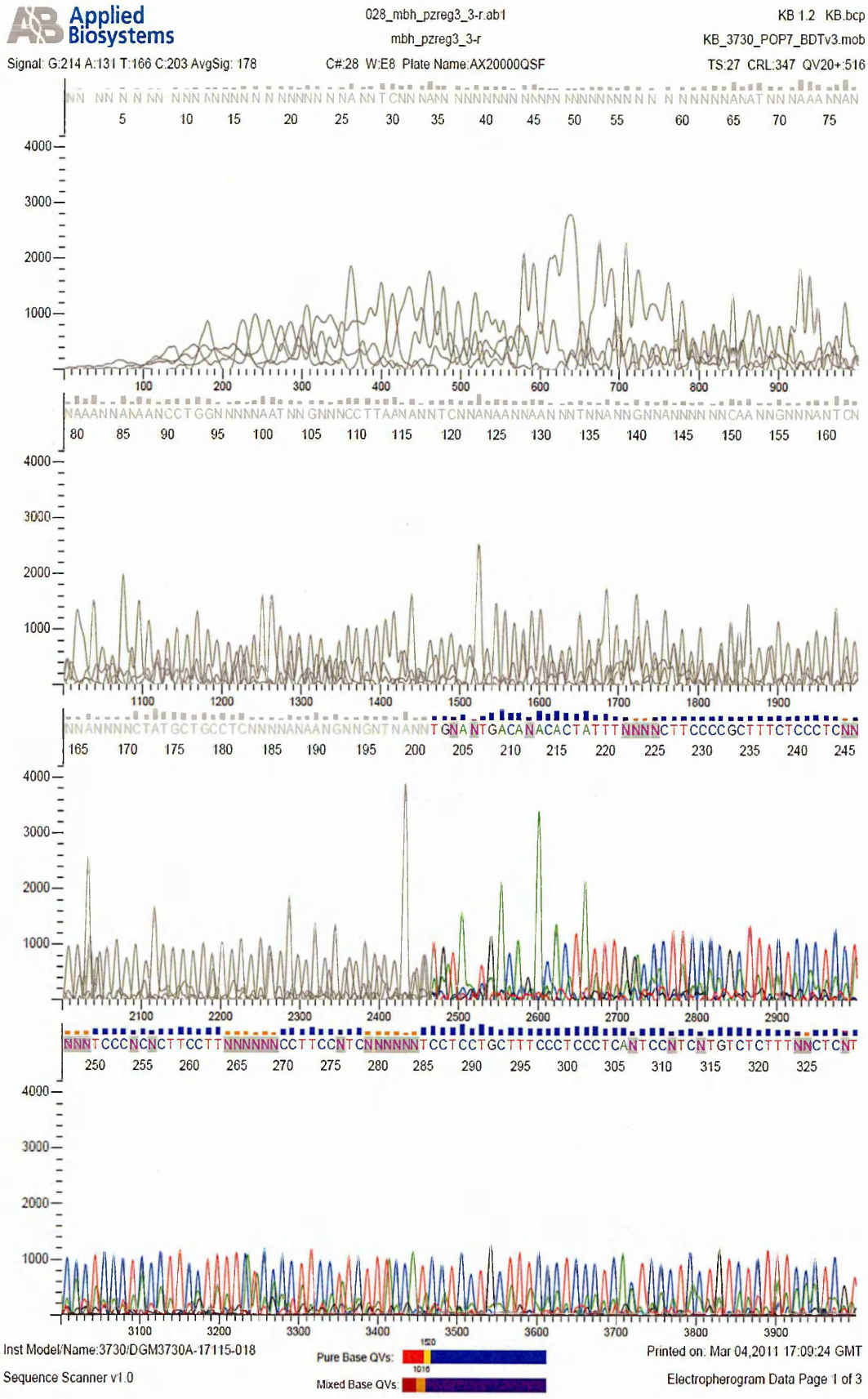
ZHU, J., *et al.* (2008). On the nature of human housekeeping genes. *Trends in genetics : TIG,* **24** (10), 481-484.

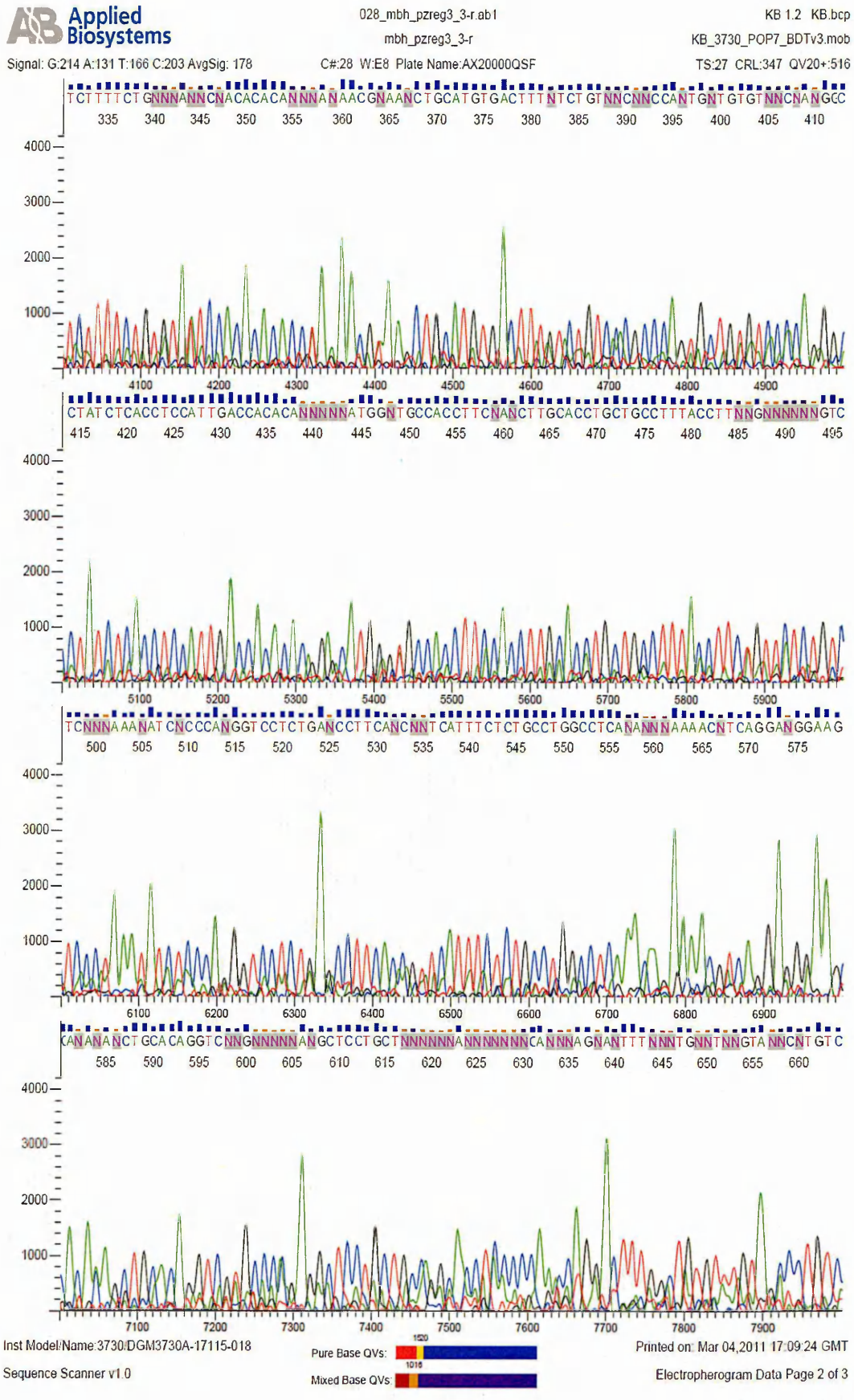
ZIPPER, H., *et al.* (2004). Investigations on DNA intercalation and surface binding by SYBR green I, its structure determination and methodological implications. *Nucleic acids research,* **32** (12), e103.

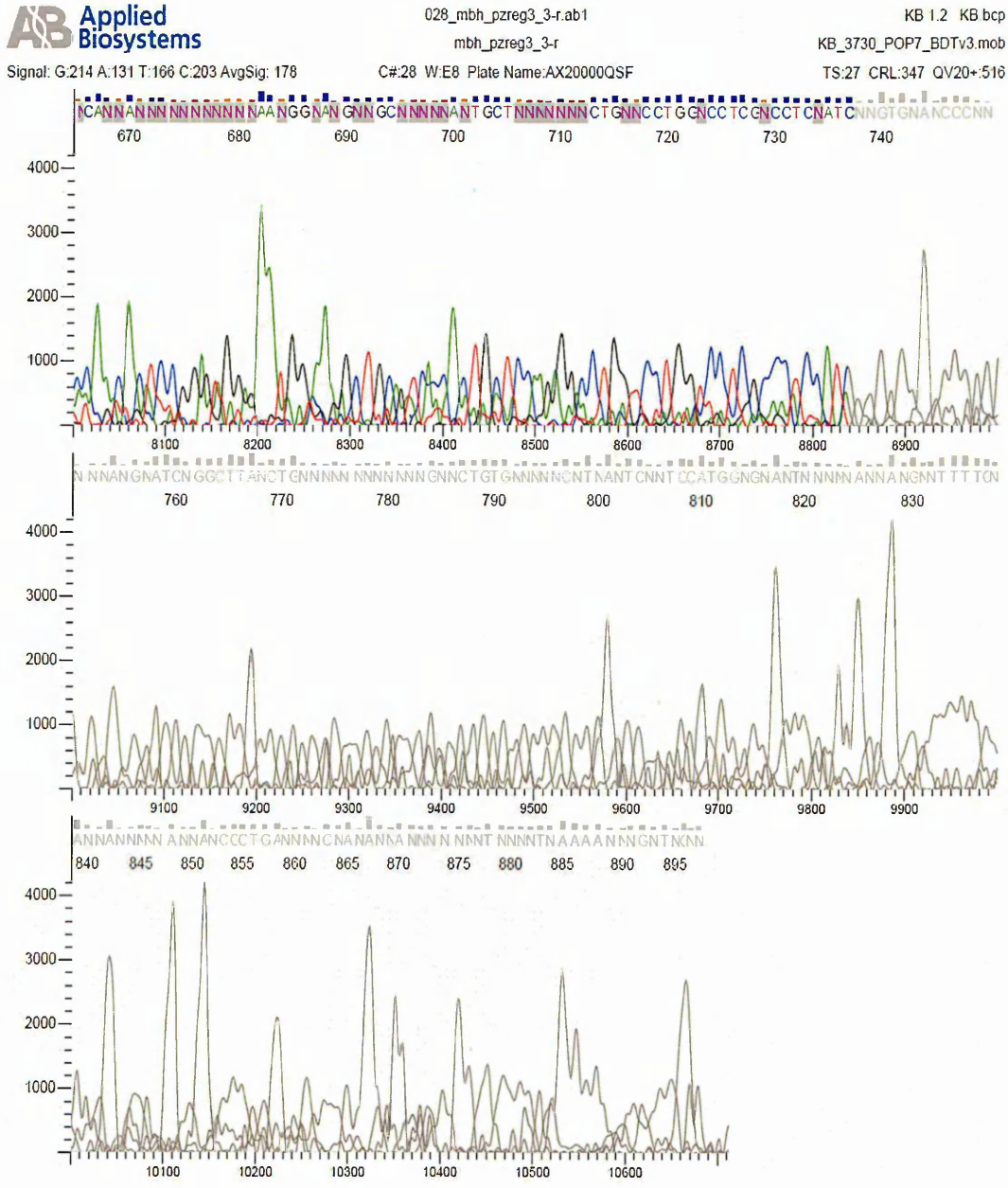
Appendix A: Diagrams show the sequences of the PZ promoter region spanning nucleotides -716 to +217.



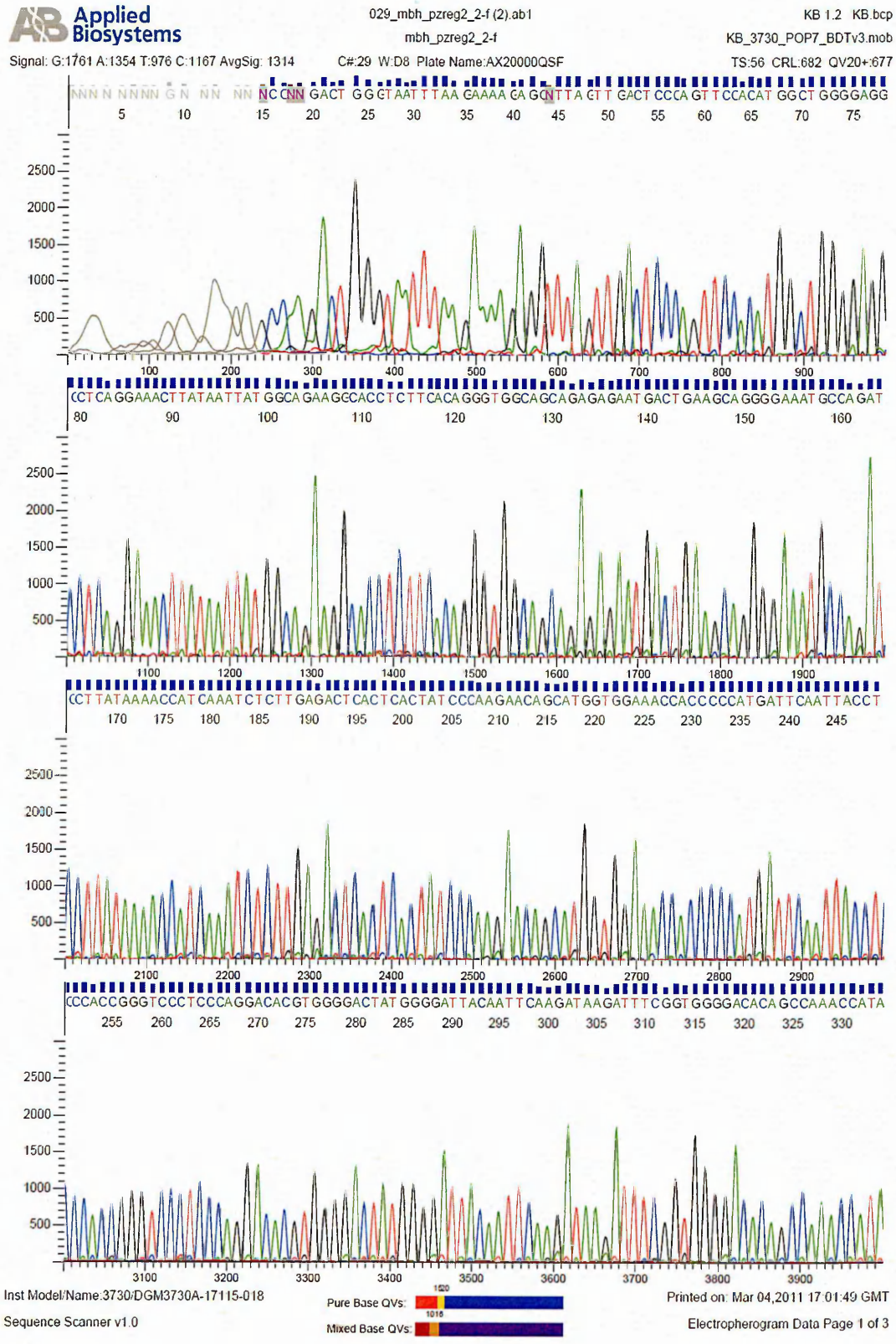

Applied Biosystems
027_mbh_pzreg3_3-f.ab1
mbh_pzreg3_3-f
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:272 A:162 T:114 C:158 AvgSig: 176
C#:27 W:F8 Plate Name:AX20000QSF
TS:55 CRL:576 QV20+:544
GGTGAAGGCTCAGAGGACCCTGGGAGATATTTAGGGAGACAGAGAACAGAAGGTAAAGGCAGCAGGTGCAAGGTGGAAGGTGG
335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410
2500 2000 1500 1000 500
4100 4200 4300 4400 4500 4600 4700 4800 4900
CACCCATGCCCGTGTGTGGTCAATGGAGGTGAGATAGGCCCTGGGGACACAGCACTGGGCGGGACAGACAAAGTCACATGCACC
420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495
2500 2000 1500 1000 500
5100 5200 5300 5400 5500 5600 5700 5800 5900
TTACGTTCTAGGTGTGTGGAATACACAGAAAAGAACCAGGGAAAGAGACAGGAGGGAGTGAGGGAGGGAAAGCAGGAGGAT
500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580
2500 2000 1500 1000 500
6100 6200 6300 6400 6500 6600 6700 6800 6900
CCAAAGAAGGAAGGGGAGAAAAGGAAGGGAGGGAGGGAGGAGGGAGANNN
585 590 595 600 605 610 615 620 625 630
2500 2000 1500 1000 500
7100 7200 7300 7400 7500 7600
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
1520
1016
Mixed Base QVs:
Printed on: Mar 04,2011 17:00:38 GMT
Electropherogram Data Page 2 of 2

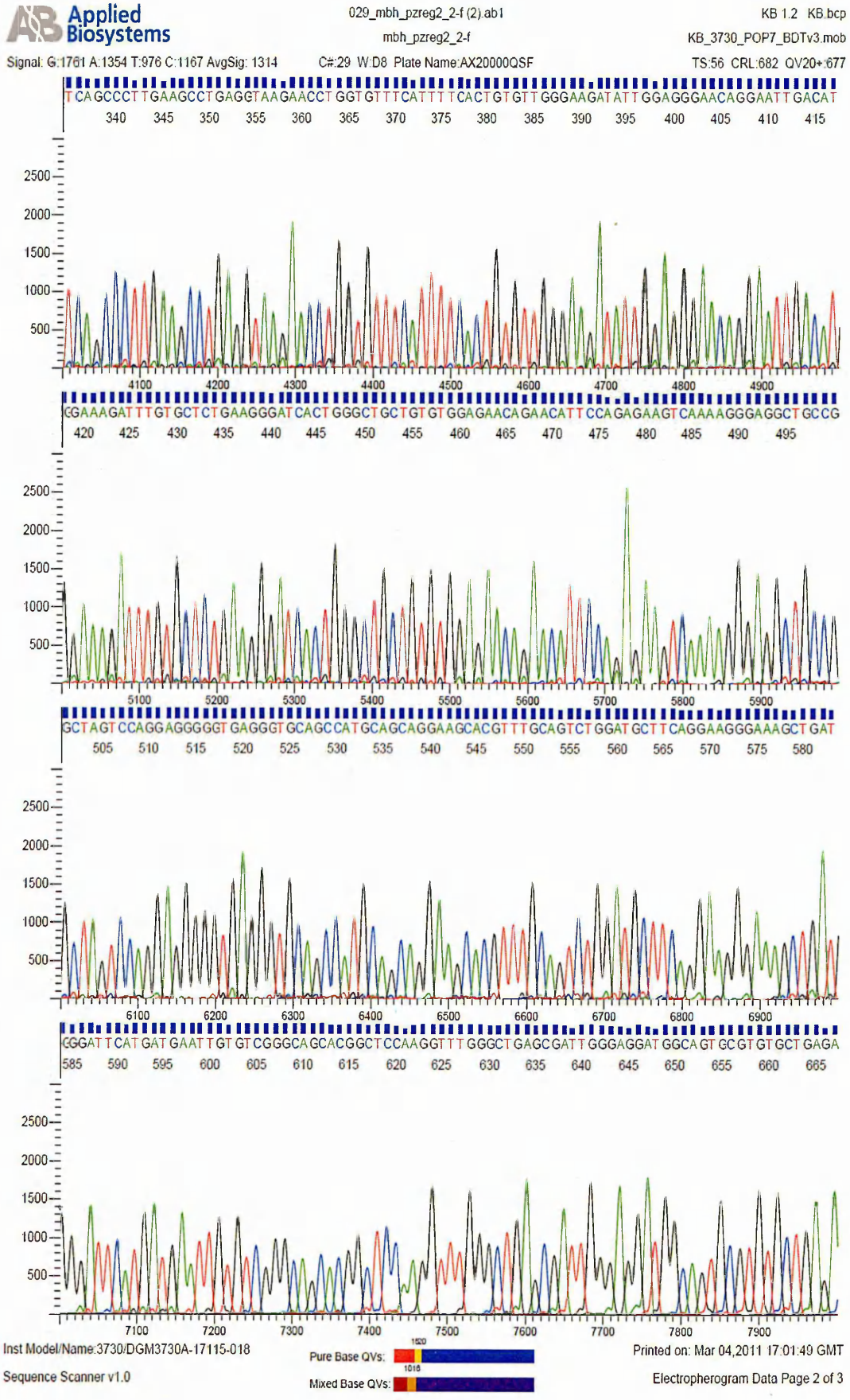

Applied Biosystems
028_mbh_pzreg3_3-r.ab1
mbh_pzreg3_3-r
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:214 A:131 T:166 C:203 AvgSig: 178
C#:28 W:E8 Plate Name:AX20000QSF
TS:27 CRL:347 QV20+:516
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Mar 04,2011 17:09:24 GMT
Electropherogram Data Page 1 of 3


Applied Biosystems
028_mbh_pzreg3_3-r.ab1
mbh_pzreg3_3-r
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:214 A:131 T:166 C:203 AvgSig: 178
C#:28 W:E8 Plate Name:AX20000QSF
TS:27 CRL:347 QV20+:516
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Mar 04,2011 17:09:24 GMT
Electropherogram Data Page 2 of 3


Applied Biosystems
028_mbh_pzreg3_3-r.ab1
mbh_pzreg3_3-r
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:214 A:131 T:166 C:203 AvgSig: 178
C#:28 W:E8 Plate Name:AX20000QSF
TS:27 CRL:347 QV20+:516
670 680 690 700 710 720 730 740
8100 8200 8300 8400 8500 8600 8700 8800 8900
760 770 780 790 800 810 820 830
9100 9200 9300 9400 9500 9600 9700 9800 9900
840 845 850 855 860 865 870 875 880 885 890 895
10100 10200 10300 10400 10500 10600
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Mar 04,2011 17:09:24 GMT
Electropherogram Data Page 3 of 3

Appendix B: Diagrams show the sequences of the PZ promoter region spanning nucleotides -1352 to -608.




Applied Biosystems
029_mbh_pzreg2_2-f (2).ab1
mbh_pzreg2_2-f
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:1761 A:1354 T:976 C:1167 AvgSig: 1314
C#:29 W:D8 Plate Name:AX20000QSF
TS:56 CRL:682 QV20+:677
TCAGCCCTTGAAGCCTGAGGTAAGAACCTGGTGTTTCATTTTCACTGTGTTGGGAAGATATTGGAGGGAACAGGAATTGACAT
340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415
2500 2000 1500 1000 500
4100 4200 4300 4400 4500 4600 4700 4800 4900
GGAAAGATTTGTGCTCTGAAGGGATCACTGGGCTGCTGTGTGGAGAACAGAACATTCCAGAGAAGTCAAAAGGGAGGCTGCCG
420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495
2500 2000 1500 1000 500
5100 5200 5300 5400 5500 5600 5700 5800 5900
GCTAGTCCAGGAGGGGGTGAGGGTGCAGCCATGCAGCAGGAAGCACGTTTGCAGTCTGGATGCTTCAGGAAGGGAAAGCTGAT
505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580
2500 2000 1500 1000 500
6100 6200 6300 6400 6500 6600 6700 6800 6900
GGGATTCATGATGAATTGTGTCGGGCAGCACGGCTCCAAGGTTTGGGCTGAGCGATTGGGAGGATGGCAGTGCGTGTGCTGAGA
585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665
2500 2000 1500 1000 500
7100 7200 7300 7400 7500 7600 7700 7800 7900
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
1520
1016
Mixed Base QVs:
Printed on: Mar 04,2011 17:01:49 GMT
Electropherogram Data Page 2 of 3

Applied Biosystems

029_mbh_pzreg2_2-f (2).ab1

mbh_pzreg2_2-f

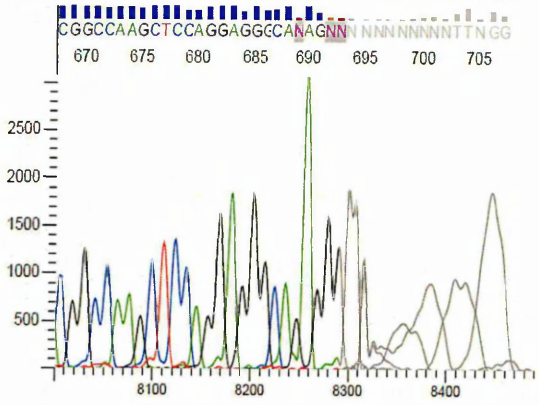
KB 1.2 KB.bcp

KB_3730_POP7_BDTv3.mob

Signal: G:1761 A:1354 T:976 C:1167 AvgSig: 1314

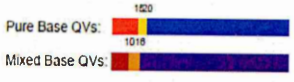
C#:29 W:D8 Plate Name:AX20000QSF

TS:56 CRL:682 QV20+:677



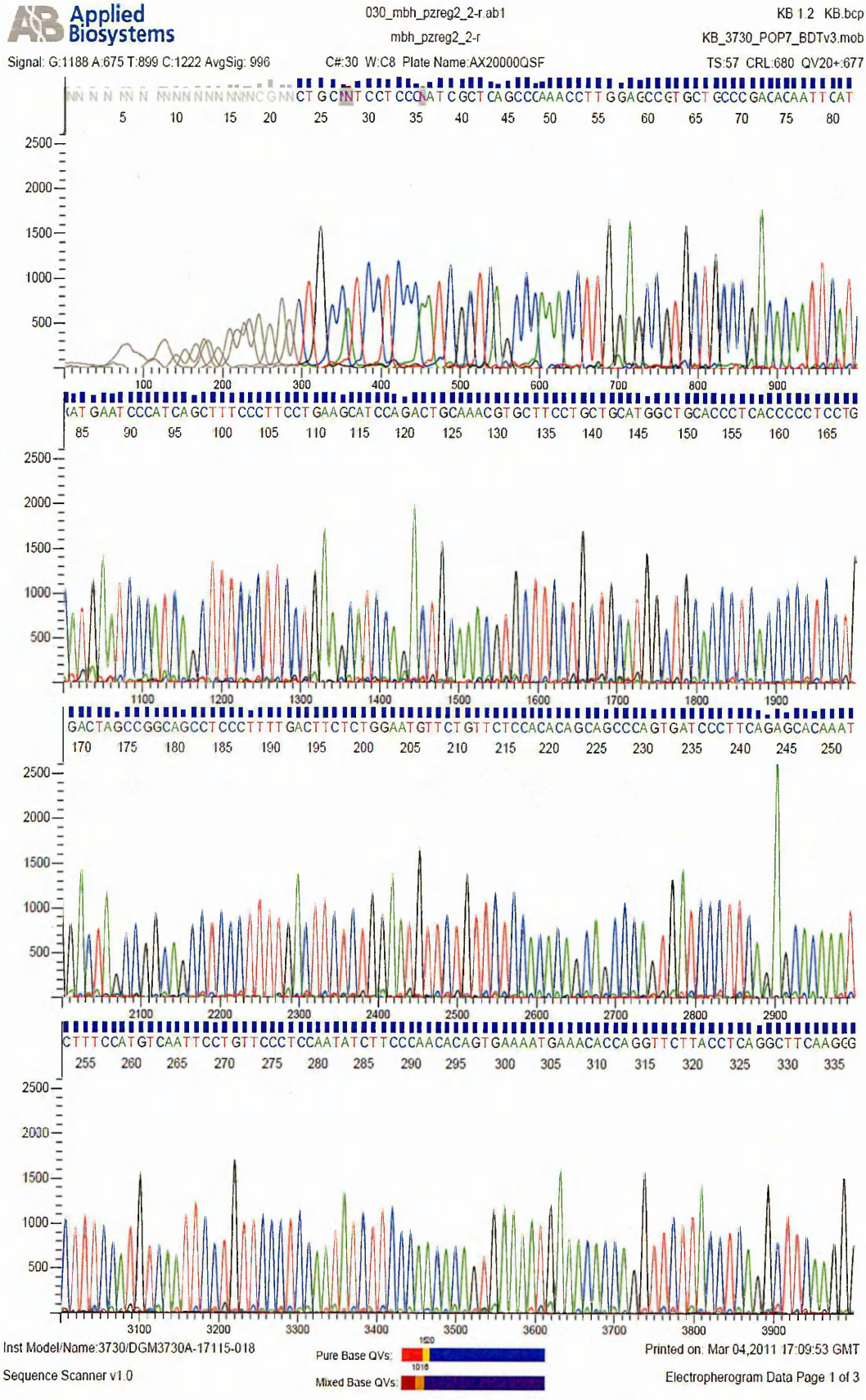
Inst Model/Name:3730/DGM3730A-I7115-018

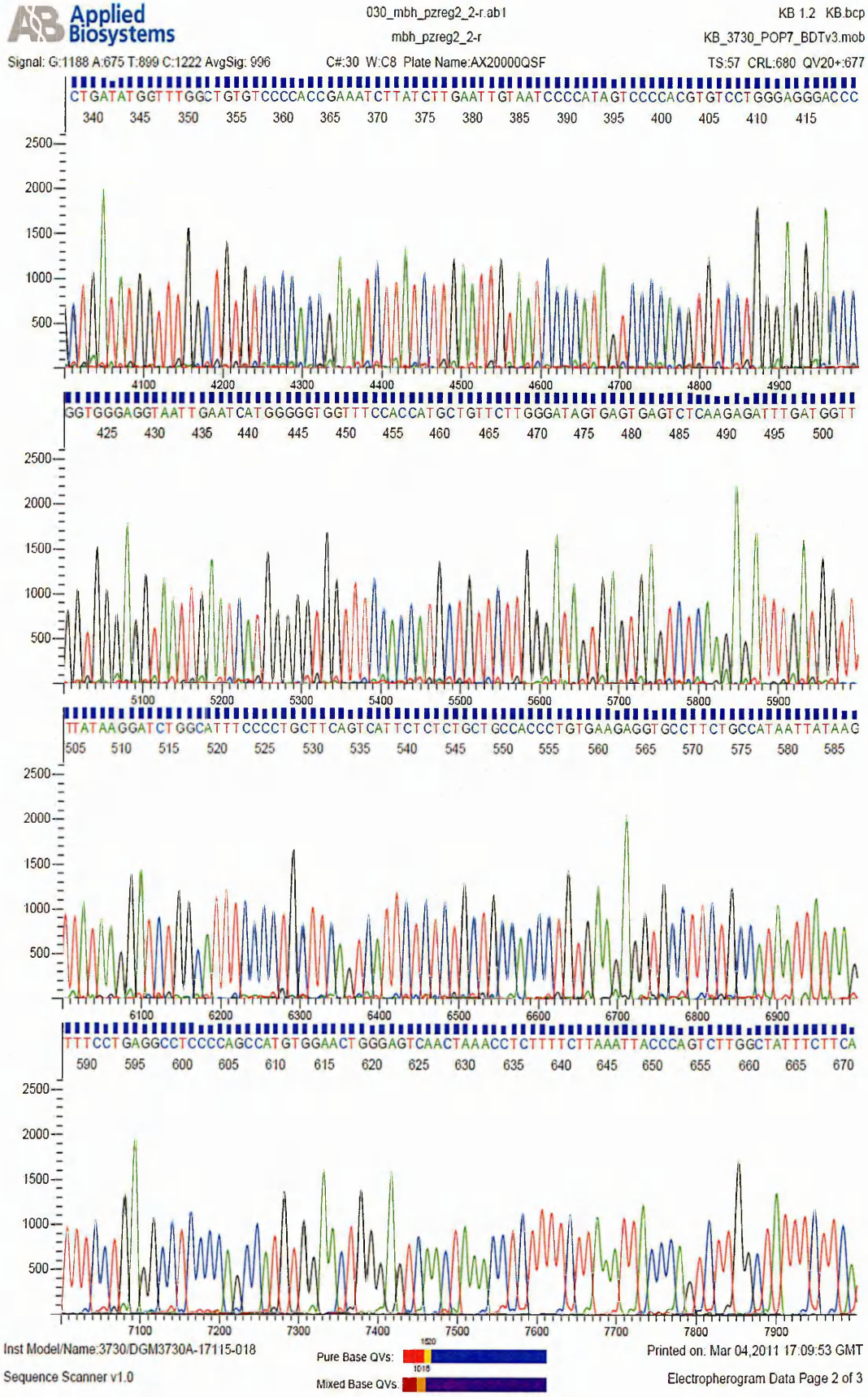
Sequence Scanner v1.0

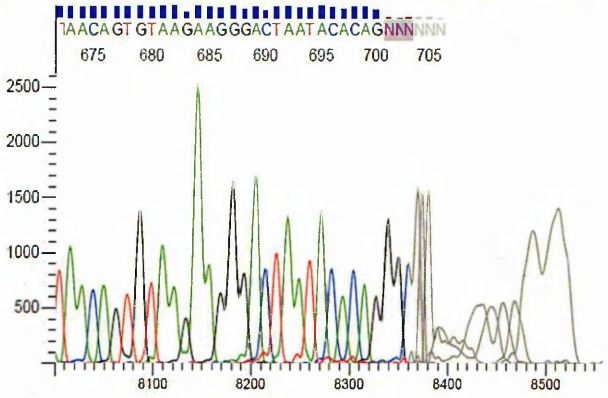


Printed on: Mar 04,2011 17:01:49 GMT

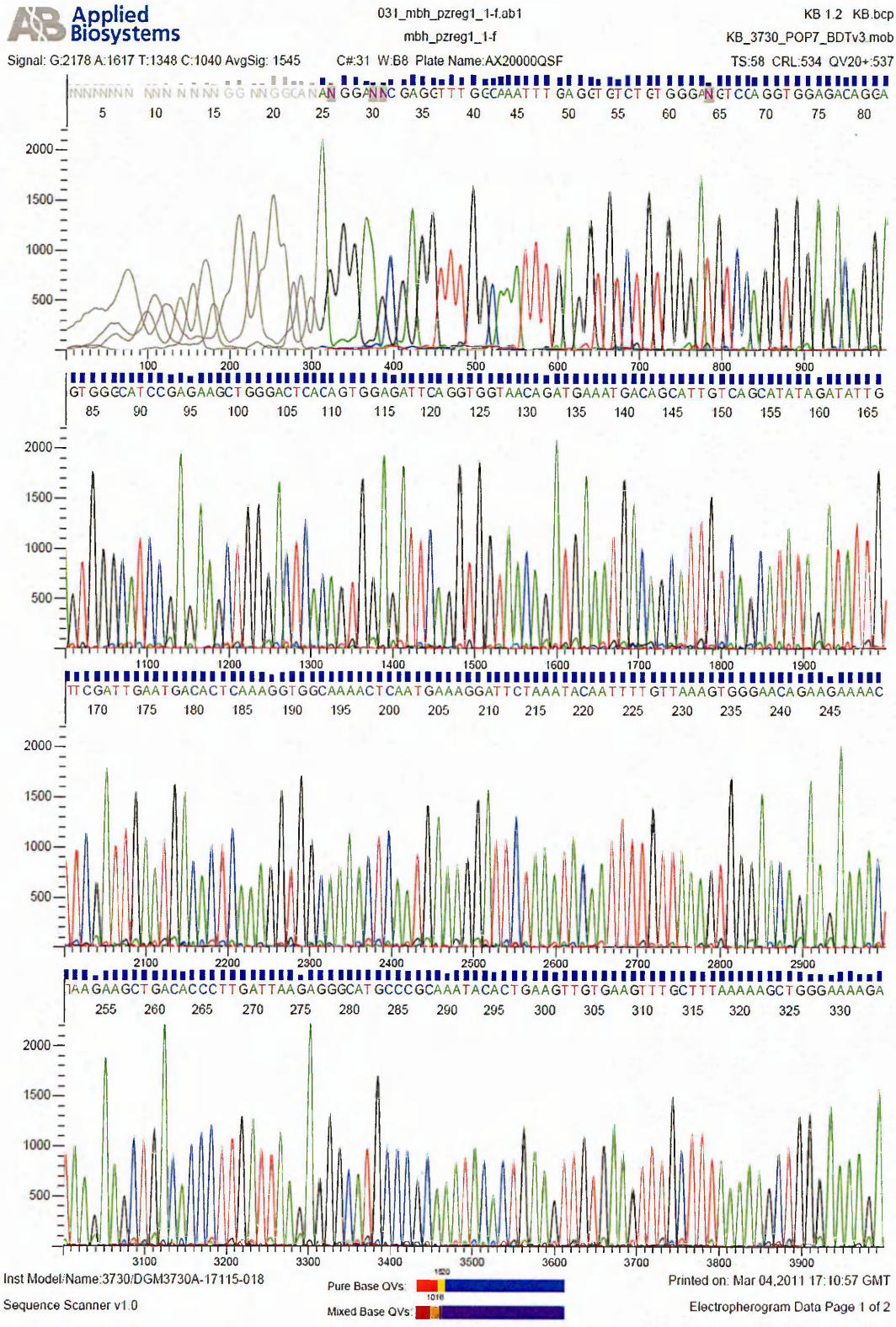
Electropherogram Data Page 3 of 3

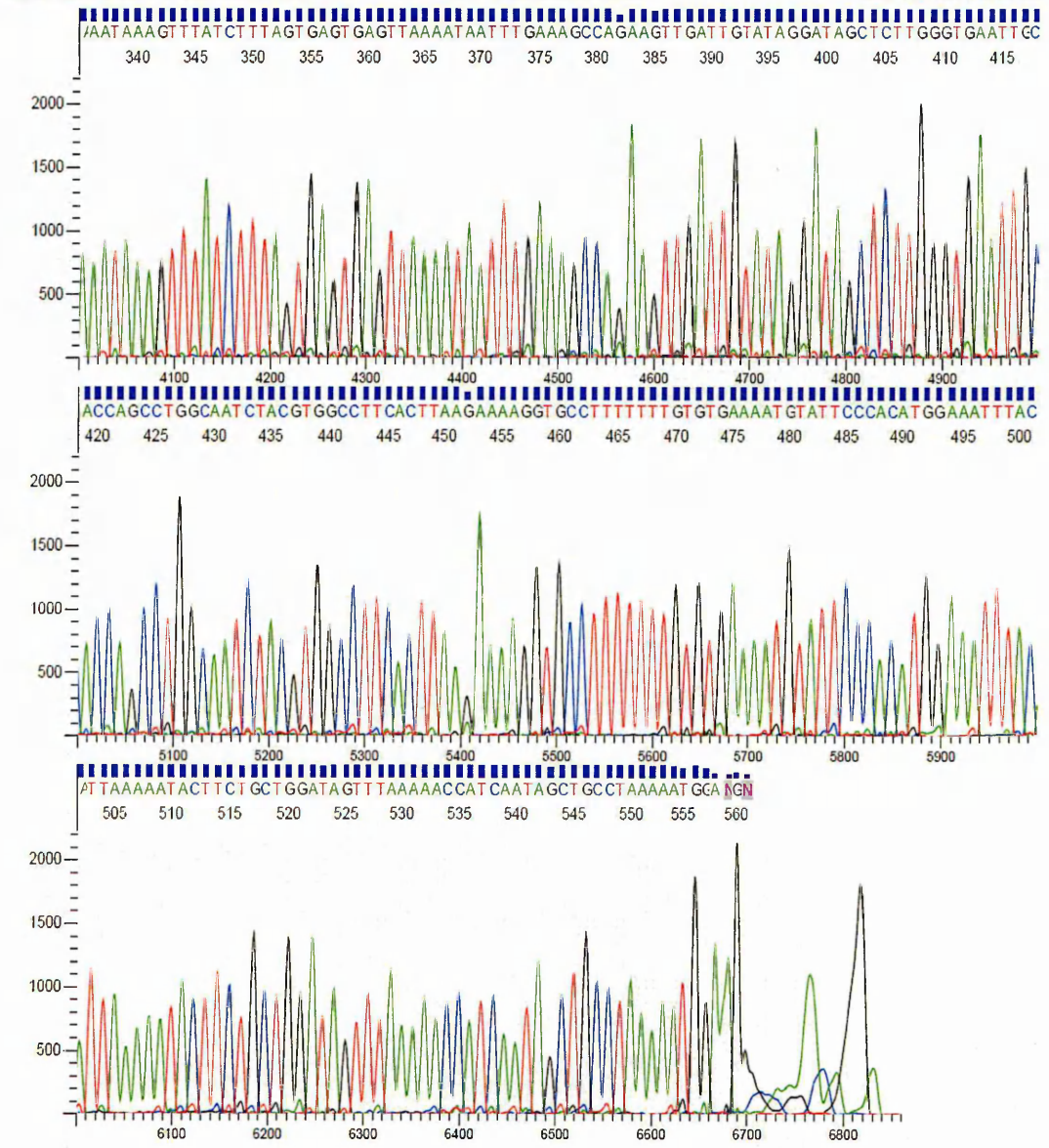

Applied Biosystems
030_mbh_pzreg2_2-r.ab1
mbh_pzreg2_2-r
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:1188 A:675 T:899 C:1222 AvgSig: 996
C#:30 W:C8 Plate Name:AX20000QSF
TS:57 CRL:680 QV20+:677
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Mar 04,2011 17:09:53 GMT
Electropherogram Data Page 1 of 3


Applied Biosystems
030_mbh_pzreg2_2-r.ab1
mbh_pzreg2_2-r
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:1188 A:675 T:899 C:1222 AvgSig: 996
C#:30 W:C8 Plate Name:AX20000QSF
TS:57 CRL:680 QV20+:677
CTGATATGGTTTGGCTGTGTCCCCACCGAAAATCTTATCTTGAATTGTAATCCCCATAGTCCCCACGTGTCCTGGGAGGGACCC
GGTGGGAGGTAATTGAATCATGGGGGTGGTTTCCACCATGCTGTTCTTGGGATAGTGAGTGAGTCTCAAGAGATTTGATGGTT
TTATAAGGATCTGGCATTTCCCCTGCTTCAGTCATTCTCTCTGCTGCCACCCTGTGAAGAGGTGCCTTCTGCCATAATTATAAG
TTTCCTGAGGCCTCCCCAGCCATGTGGAACTGGGAGTCAACTAAACCTCTTTTCTTAAATTACCCAGTCTTGGCTATTTCTTCA
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Mar 04,2011 17:09:53 GMT
Electropherogram Data Page 2 of 3


TAACAGT GTAAGAAGGGACT AATACACAGNNNNNN
675
680
685
690
695
700
705
2500
2000
1500
1000
500
8100
8200
8300
8400
8500

Appendix C: Diagrams show the sequences of PZ promoter region spanning nucleotides -1883 to-1284.





Inst Model/Name:3730/DGM3730A-17115-018

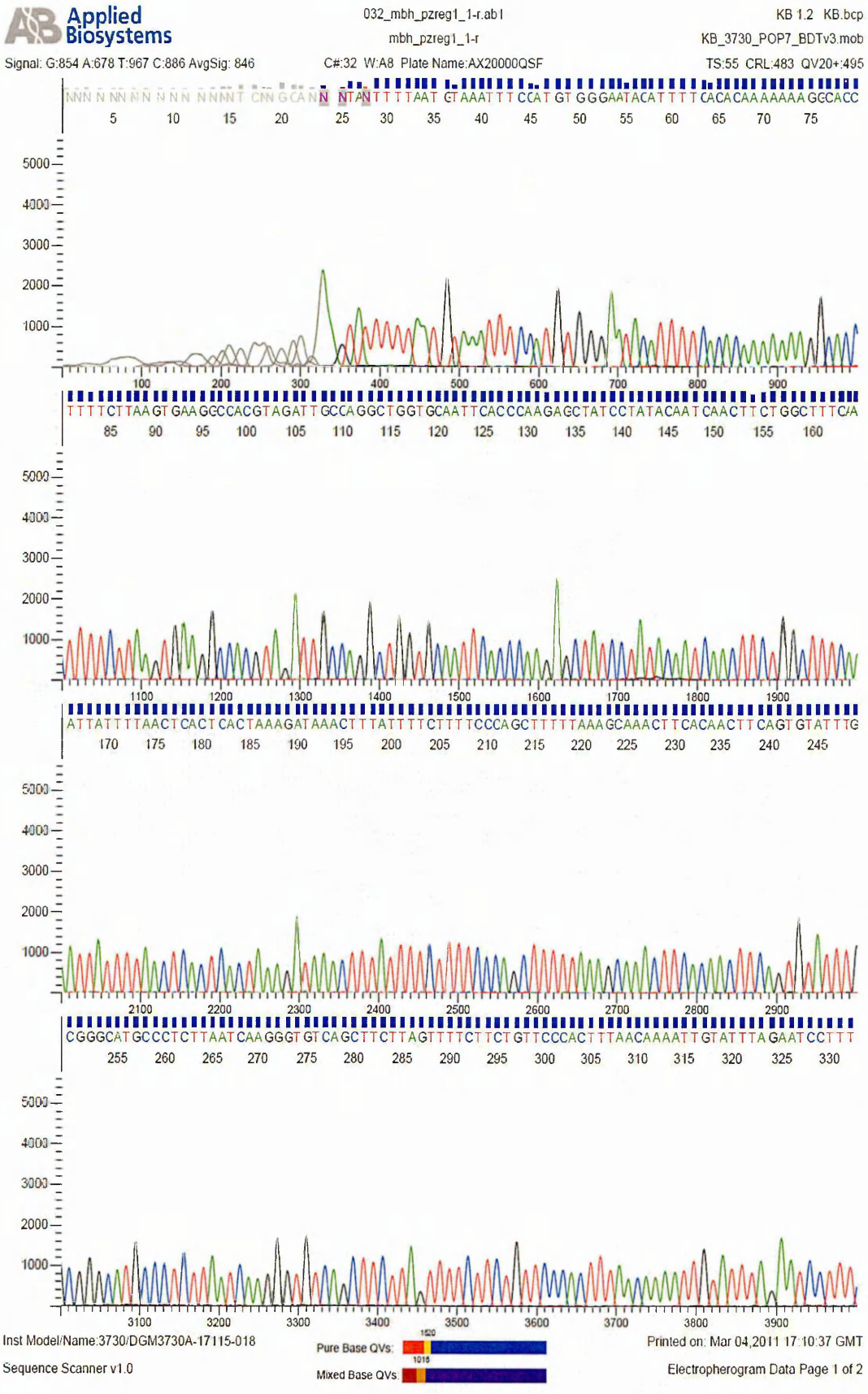
Sequence Scanner v1.0

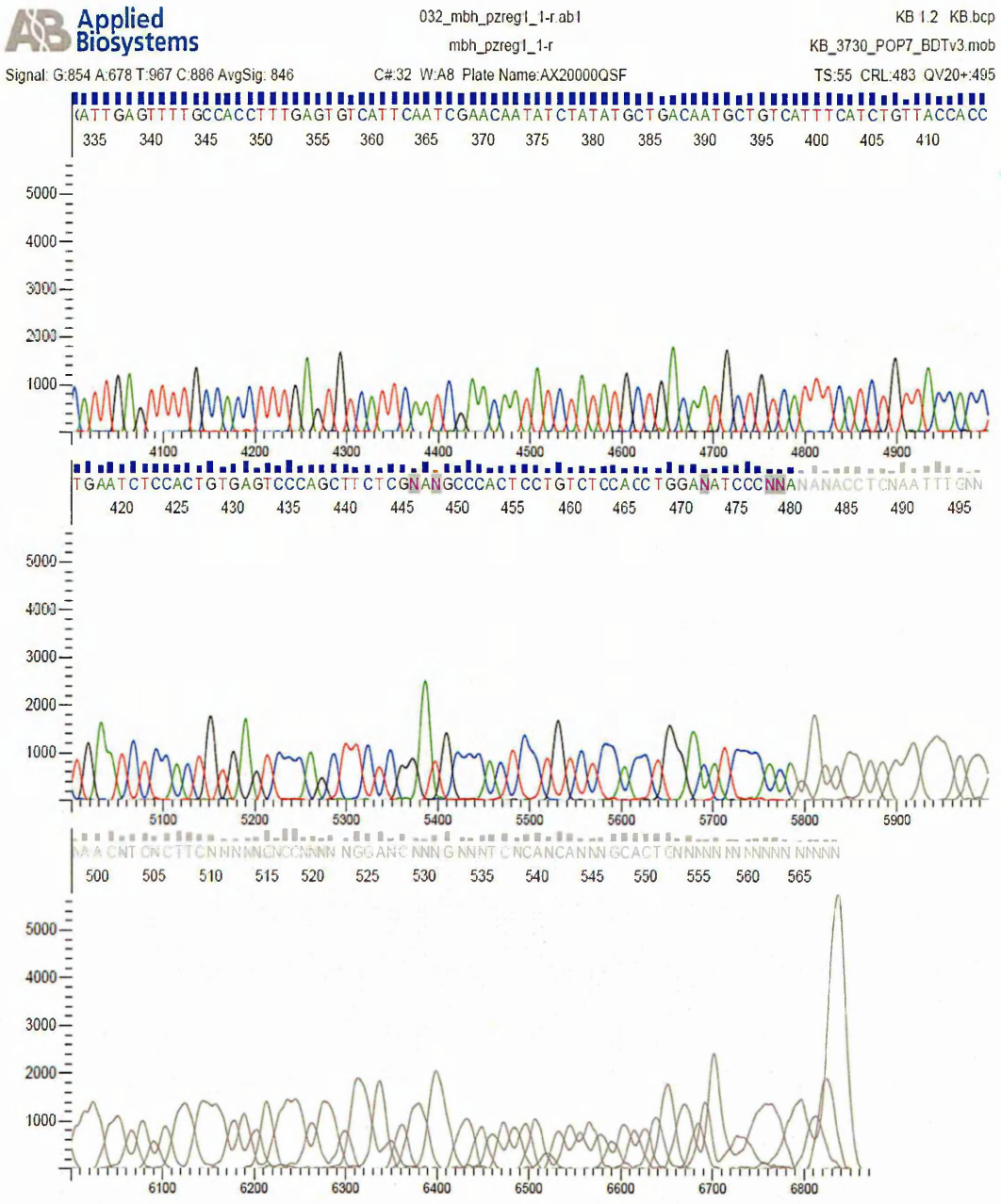
Pure Base QVs:

Mixed Base QVs:

Printed on: Mar 04,2011 17:10:57 GMT

Electropherogram Data Page 2 of 2


Applied Biosystems
032_mbh_pzreg1_1-r.abl
mbh_pzreg1_1-r
KB 1.2 KB.bcp
KB_3730_POP7_BDTv3.mob
Signal: G:854 A:678 T:967 C:886 AvgSig: 846
C#:32 W:A8 Plate Name:AX20000QSF
TS:55 CRL:483 QV20+:495
Inst Model/Name:3730/DGM3730A-17115-018
Sequence Scanner v1.0
Pure Base QVs:
Mixed Base QVs:
Printed on: Mar 04,2011 17:10:37 GMT
Electropherogram Data Page 1 of 2



Inst Model/Name:3730/DGM3730A-17115-018

Sequence Scanner v1.0

Pure Base QVs: 1520

Mixed Base QVs: 1018

Printed on: Mar 04,2011 17:10:37 GMT

Electropherogram Data Page 2 of 2