Post-Transcriptional Regulation of ALAS1
Expression by Haem

By

Sadie Jane Redding

A thesis submitted in candidature for the degree of Philosophiae Doctor

Department of Medical Biochemistry and Immunology
Wales College of Medicine
Cardiff University

September 2007
Declaration and Statements

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed .................................................. (candidate)  Date: .............................

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed .................................................. (candidate)  Date: .............................

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references.

Signed .................................................. (candidate)  Date: .............................

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .................................................. (candidate)  Date: .............................

STATEMENT 4 - BAR ON ACCESS APPROVED

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans after expiry of a bar on access approved by the Graduate Development Committee.

Signed .................................................. (candidate)  Date: .............................
Acknowledgements

Firstly I would like to thank my supervisors Dr. Andrew Roberts and Dr. David Llewellyn for their help and guidance throughout my PhD. Many thanks to Professor G.H Elder and Professor M Badminton for guidance on the direction of my studies. Thank you also to Professor B.P Morgan and the Department of Medical Biochemistry for supporting and funding my project.

I am most grateful for the technical advice on RNA-binding assays, given by those in Dr. Nancy Standart’s laboratory in the University of Cambridge, especially Dr. Lucy Colgrave. Also, thank you to Barrie Francis for help with DNA sequencing, and to Dr Sharon Whatley and Nicola Mason for general advice and making the conferences we attended most enjoyable.

Thank you to all the members of the Department of Medical Biochemistry and fellow UWCM PhD students. In particular, thank you to Tarnjit, Jon, Tom, Ben and Karolina for making my time there enjoyable and being readily available for coffee breaks.

I would like to thank my family: mum, dad and Paul, for all their encouragement and motivation to work hard and achieve my goal. Finally, thank you to Alistair for his constant support and for keeping me happy throughout the writing-up stage.
Publications Relevant to this Thesis


Unrefereed Abstracts

Redding SJ. Haem-mediated destabilisation of ALAS1 mRNA. The Tetrapyrrole Discussion group, University of Cambridge, 16th-17th September 2004.


Summary

Haem is the prosthetic moiety of numerous haemoproteins critical for the function of all aerobic cells. Its biosynthesis is a tightly controlled process since high intracellular haem concentrations are cytotoxic, whilst haem deficiency impedes the activity of essential haemoproteins. In the liver and probably all other non-erythroid cells, haem supply is regulated primarily through feedback regulation of the stability of the mRNA encoding aminolaevulinic acid synthase 1 (ALAS1), the first and rate-limiting enzyme in the haem biosynthetic pathway. However, the underlying mechanism of this destabilisation is unknown. Consequently, the primary aim of this thesis was to determine how haem regulates ALAS1 mRNA stability to control its own synthesis in non-erythroid cells, using the human hepatoma cell line, HepG2.

In humans, the ALAS1 exon 1b in the 5′-untranslated region (UTR) is alternatively spliced to produce a minor and major form of the enzyme. This thesis has demonstrated that unlike the major ALAS1 5′-UTR, the minor 5′-UTR causes a downstream heterologous RNA to be poorly translated. In addition, the minor ALAS1 isoform is relatively resistant to haem-mediated decay. Using reporter assays and RT-PCR, we have shown that the human ALAS1 mRNA contains a coding region determinant (CRD) that mediates its haem-sensitivity in HepG2 cells. This CRD can function independently of the ALAS1 5′- and 3′-UTR. Furthermore, this haem-mediated CRD has to be translated to function. RNA-electromobility shift assays (EMSAs) have defined two fragments of the ALAS1 coding region that can bind to HepG2 cytosolic protein. However, this binding does not seem to be affected by the addition or depletion of haem. From the data presented in this thesis, a mechanism into how haem destabilises the ALAS1 mRNA in humans has been proposed.
Contents

CHAPTER 1: General Introduction
1.1 Structure and Chemistry of Haem 1
1.2 Function of Haem 2
1.3 Haem Biosynthesis 3
  1.3.1 Aminolaevulinate synthase 8
  1.3.2 Aminolaevulinic acid dehydratase 8
  1.3.3 Porphobilinogen deaminase 9
  1.3.4 Uroporphyrinogen III synthase 9
  1.3.5 Uroporphyrinogen III decarboxylase 9
  1.3.6 Coproporphyrinogen oxidase 10
  1.3.7 Protoporphyrinogen oxidase 10
  1.3.8 Ferrochelatase 11
1.4 Mammalian Haem Transport 11
1.5 Role of Haem Biosynthesis in Disease 14
  1.5.1 Haem Excess 16
  1.5.2 Haem Deficiency 19
1.6 Homeostasis of Haem 20
  1.6.1 Haem Oxygenase 21
  1.6.2 Aminolaevulinate Synthase (ALAS) 22
    1.6.2.1 ALAS as a PLP-dependent enzyme 23
    1.6.2.2 The ALAS catalytic mechanism 24
    1.6.2.3 ALAS Regions 26
1.7 Comparison of the ALAS1 and ALAS2 Genes 26
  1.7.1 The Human ALAS1 Gene 29
  1.7.2 The Human ALAS2 Gene 31
1.8 Transcriptional Control of ALAS 32
  1.8.1 Regulation of ALAS2 32
  1.8.2 Regulation of ALAS1 32
    1.8.2.1 Insulin 33
    1.8.2.2 Proliferator-activated receptor γ coactivator 1α 33
  1.8.2.3 Nuclear respiratory factor-1 35
  1.8.2.4 Xenobiotic-sensing Nuclear Receptor 36
  1.8.4 Circadian Rhythms 36
1.9 Post-transcriptional Regulation of ALAS1 Expression by Haem 38
  1.9.1 Regulatory Haem Pool 39
  1.9.2 Haem Response Motifs 40
1.9.3 Haem-mediated destabilisation of ALAS1 mRNA 43
1.9.4 Control of Haem Synthesis in Other Cell Types 43
1.10 Post-transcriptional Regulation of Gene Expression 44
  1.10.1 Regulated mRNA Stability 45
  1.10.2 The 3' Untranslated Region 47
  1.10.3 The 5' Untranslated Region 47
  1.10.4 The Coding Region 48
1.11 Aims of the Thesis 50

CHAPTER 2: General Materials and Methods 51
2.1 Chemicals 52
2.2 Cell Culture 52
  2.2.1 Mammalian Cell Culture 52
  2.2.1.1 Cell Lines 52
  2.2.1.2 Media and Growth Conditions 52
  2.2.1.3 Seeding Cells 52
  2.2.1.3.1 Determining Cell Number 53
  2.2.1.4 Transfection 53
  2.2.1.5 Freezing Mammalian Cells 53
  2.2.1.6 Haem-Depleted Media 53
  2.2.1.7 Preparation of Cytosolic Protein Extracts 54
  2.2.1.8 Dual Luciferase Reporter Assay (DLRA) 54
  2.2.2 Bacterial Cell Culture 55
  2.2.2.1 Media and Growth Conditions 55
  2.2.2.2 Transformation 55
  2.2.2.3 Small Scale Plasmid Purification 56
  2.2.2.4 Large Scale Plasmid Purification 57
2.3 Protein Methodology 57
  2.3.1 Estimation of Protein Concentration 57
  2.3.2 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE) 58
  2.3.3 Coomassie Blue Staining 59
2.4 DNA Methodology 59
  2.4.1 Restriction Enzymes and Buffers 59
  2.4.2 Restriction Digests 60
  2.4.2.1 Single Digests 60
  2.4.2.2 Double Digests 61
  2.4.3 DNA Quantification 61
  2.4.4 Polymerase Chain Reaction (PCR) 62
2.4.4.1 Primers 62
2.4.4.2 Standard PCR Conditions 62
2.4.4.3 High Fidelity PCR Conditions 63
2.4.4.3.1 HotStarTaq Polymerase 63
2.4.4.3.2 Bio-X-Act Polymerase 64
2.4.5 Agarose-TAE Gel Electrophoresis 64
2.4.5.1 DNA Molecular Weight Markers 65
2.4.5.1.1 100bp Ladder 65
2.4.5.1.2 1Kb ladder 65
2.4.6 Gel Visualisation and Documentation 66
2.4.7 Gel Extraction 66
2.4.8 PCR Purification 66
2.4.9 Automated DNA Sequencing 67
2.4.10 Sequencing Clean-Up 68
2.4.11 In Vitro Transcription 68
2.4.12 Gel Purification 69

2.5 RNA Methodology 69
2.5.1 RNA Isolation 69
2.5.1.1 Total RNA Isolation 70
2.5.1.2 mRNA Isolation 70
2.5.2 RNA Quantification 71
2.5.3 Removal of DNA Contamination 71
2.5.4 First Strand Synthesis 72
2.5.5 RNase Protection Assay (RPA) 72
2.5.6 Detection of Non-Isotopic probes 73

2.6 Cloning 74
2.6.1 TA Cloning 74
2.6.2 Ligation 74
2.6.3 Screening Colonies for Inserts 75
2.6.4 Colony screening by PCR 75

2.7 RNA-Protein Interaction Assays 76
2.7.1 In Vitro Transcription 76
2.7.2 RNA Electrophoretic Mobility Shift Assay (R-EMSA) 76
2.7.2.3 Non-Denaturing Gel Electrophoresis 77
2.7.3 UV-Crosslinking 77
CHAPTER 3: Analysis of the major and minor ALAS1 5′-UTRs

3.1 Introduction
3.1.1 The Dual Luciferase Reporter Assay
3.1.1.2 Firefly and Renilla luciferases
3.1.2 Aims and Strategy

3.2 Materials and Methods
3.2.1 Semi-Quantitative RT-PCR
3.2.1.1 Optimising Semi-Quantitative RT-PCR
3.2.2 PCR and Plasmid Construction
3.2.3 Transfections and Dual Luciferase Reporter Assay

3.3 Results and Discussion
3.3.1 The ALAS1 minor isoform is resistant to haem-mediated decay.
3.3.2 The ALAS1 5′-UTR does not harbour a destabilising element, but the minor isoform inhibits expression of the downstream reporter gene.
3.3.3 The minor ALAS1 5′-UTR reduces translation of the downstream open reading frame.
3.3.4 Confirmation of ALAS1 5′-UTR Transcriptional Activity
3.3.5 Reduction in expression of downstream mRNA by the ALAS1 minor 5′-UTR is not restricted to hepatocytes.

3.4 Conclusions

CHAPTER 4: Delineation of the ALAS1 haem-mediated instability element

4.1 Introduction
4.1.1 mRNA regulatory elements in the 3′-UTR and coding region
4.1.2 Putative ALAS1 cis-acting elements
4.1.3 Aims and Strategy

4.2 Materials and Methods
4.2.1 PCR and Plasmid construction
4.2.2 Mammalian Cell Transfection
4.2.3 RNA Quantification by RNase Protection Assays (RPA)
4.2.3.1 Biotin-labelled antisense RNA probes
4.2.3.2 RPA
4.2.4 Semi-Quantitative RT-PCR
4.2.4.1 RT-PCR of pRALAS1 and pGL3P
4.2.4.2 RT-PCR of endogenous haem oxygenase-1
4.3 Results and Discussion

4.3.1 The cis-acting element in the 154bp region does not confer haem-mediated destabilisation.

4.3.2 The human ALAS1 coding region contains a haem-mediated cis-acting element.

4.3.3 The coding region determinant has to be translated to mediate instability of the mRNA in the presence of haem.

4.3.4 Overlapping deletion experiments to delineate the ALAS1 mRNA coding region determinant(s).

4.3.5 Confirmation of an ALAS1 CRD by RPA

4.3.6 Optimisation of the removal of haem and haemoproteins from FCS

4.3.7 Confirmation of an ALAS1 CRD by semi-quantitative RT-PCR

4.3.8 Haem destabilises ALAS1 mRNA in IMR32 cells.

4.4 Conclusions

CHAPTER 5: Identification of ALAS1 mRNA protein-binding regions

5.1 Introduction

5.1.1 Trans-acting regulatory factors

5.1.2 Tetrapyrrole binding to mRNA

5.1.3 Aims and Strategy

5.1.3.1 RNA-EMSAs

5.1.3.2 UV-Crosslinking

5.2 Materials and Methods

5.2.1 Radiolabelled ALAS1 RNA construction

5.2.2 RNA EMSA

5.2.3 UV-Crosslinking of cytosolic proteins to ALAS1

5.2.4 Binding of ALAS1 RNA to haem-agarose

5.2.4.1 RT-PCR of ALAS1 and GAPDH

5.3 Results and Discussion

5.3.1 HepG2 cytosolic protein binds to two distinct regions of the ALAS1 coding region.

5.3.2 Binding of HepG2 cytosolic protein to ALAS1 RNA probes with the addition of haemin in the EMSAs

5.3.3 Binding of cytosolic protein from haem depleted or supplemented HepG2 cells, to ALAS1 RNA

5.3.5 UV-Crosslinking of HepG2 cytosolic proteins bound to ALAS1 RNA

5.3.6 Optimisation of UV-crosslinking to fragment 2
5.3.7 The use of haemin-agarose beads to investigate whether haem binds ALAS1 mRNA directly 168

5.4 Conclusions 170

CHAPTER 6: Discussion 178

CHAPTER 7: References 189

APPENDIX 212
Appendix 1 213
Appendix 2 215
Appendix 3 219
Appendix 4 221
### Abbreviations and Symbols

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>Alpha</td>
</tr>
<tr>
<td>β</td>
<td>Beta</td>
</tr>
<tr>
<td>γ</td>
<td>Gamma</td>
</tr>
<tr>
<td>Δ</td>
<td>Delta</td>
</tr>
<tr>
<td>κ</td>
<td>Kappa</td>
</tr>
<tr>
<td>μg</td>
<td>Microgram(s)</td>
</tr>
<tr>
<td>μM</td>
<td>Micromolar</td>
</tr>
<tr>
<td>aa</td>
<td>Amino acid</td>
</tr>
<tr>
<td>ADH5/FDH</td>
<td>Human x-alcohol dehydrogenase/formaldehyde dehydrogenase</td>
</tr>
<tr>
<td>Ado-CBL</td>
<td>Adenosylcobalamin</td>
</tr>
<tr>
<td>ADRES</td>
<td>ALAS drug-responsive enhancer sequence</td>
</tr>
<tr>
<td>ALA</td>
<td>5′-aminolaevulinic acid</td>
</tr>
<tr>
<td>ALAD</td>
<td>ALA dehydratase</td>
</tr>
<tr>
<td>ALAS</td>
<td>ALA synthase</td>
</tr>
<tr>
<td>AONS</td>
<td>8-amino-7-oxononanoate synthase</td>
</tr>
<tr>
<td>APP</td>
<td>Amyloid precursor protein</td>
</tr>
<tr>
<td>ARE</td>
<td>AU rich element</td>
</tr>
<tr>
<td>ATP</td>
<td>5′-adenosine triphosphate</td>
</tr>
<tr>
<td>AU</td>
<td>Adenylate and uridylicate</td>
</tr>
<tr>
<td>AUBP</td>
<td>AU-binding protein</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair(s)</td>
</tr>
<tr>
<td>BCA</td>
<td>Bicinchoninic acid</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic AMP</td>
</tr>
<tr>
<td>CAR</td>
<td>Constitutive androstane receptor</td>
</tr>
<tr>
<td>CAT</td>
<td>Chloramphenicol acetyltransferase</td>
</tr>
<tr>
<td>CBL</td>
<td>Cobalamin</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CEB</td>
<td>Cytoplasmic extraction buffer</td>
</tr>
<tr>
<td>CMV</td>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>CNS</td>
<td>Central nervous system</td>
</tr>
<tr>
<td>CO</td>
<td>Carbon monoxide</td>
</tr>
<tr>
<td>CO₂</td>
<td>Carbon dioxide</td>
</tr>
<tr>
<td>CoA</td>
<td>Coenzyme A</td>
</tr>
<tr>
<td>cpm</td>
<td>Counts per minute</td>
</tr>
<tr>
<td>CPOX</td>
<td>Coproporphyrinogen oxidase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>CRD</td>
<td>Coding region determinant</td>
</tr>
<tr>
<td>CREB</td>
<td>cAMP element binding protein</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>dATP</td>
<td>2'-deoxyadenosine 5'-triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>2'-deoxycytosine 5'-triphosphate</td>
</tr>
<tr>
<td>dGTP</td>
<td>2'-deoxyguanosine 5'-triphosphate</td>
</tr>
<tr>
<td>DLRA</td>
<td>Dual luciferase reporter assay</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco's modified Eagle medium</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Equimolar mix of dATP, dCTP, dGTP, dTTP</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double stranded DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>dTTP</td>
<td>2'-deoxythymidine 5'-triphosphate</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetra acetic acid</td>
</tr>
<tr>
<td>eIF</td>
<td>Eukaryotic initiation factor</td>
</tr>
<tr>
<td>EMSA</td>
<td>Electromobility shift assay</td>
</tr>
<tr>
<td>ER</td>
<td>Endoplasmic reticulum</td>
</tr>
<tr>
<td>FCS</td>
<td>Foetal calf serum</td>
</tr>
<tr>
<td>FECH</td>
<td>Ferrochelatase</td>
</tr>
<tr>
<td>g</td>
<td>Gravity</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde phosphate dehydrogenase</td>
</tr>
<tr>
<td>Glc6P</td>
<td>Glucose-6-phosphatase</td>
</tr>
<tr>
<td>h</td>
<td>Hour(s)</td>
</tr>
<tr>
<td>HBS</td>
<td>HIF-1 DNA binding site</td>
</tr>
<tr>
<td>HCP1</td>
<td>Haem carrier protein 1</td>
</tr>
<tr>
<td>HIF-1</td>
<td>Hypoxia-inducible factor-1</td>
</tr>
<tr>
<td>HMB</td>
<td>Hydroxymethylbilane</td>
</tr>
<tr>
<td>HN3Fβ</td>
<td>Hepatocyte nuclear factor</td>
</tr>
<tr>
<td>HO-1</td>
<td>Haem oxygenase 1</td>
</tr>
<tr>
<td>IF</td>
<td>Initiation factor</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>Ins RE</td>
<td>Insulin responsive element</td>
</tr>
<tr>
<td>I.p.</td>
<td>Intraperitoneally</td>
</tr>
<tr>
<td>IRE</td>
<td>Iron response element</td>
</tr>
<tr>
<td>IRP</td>
<td>Iron response protein</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase(s)</td>
</tr>
</tbody>
</table>
KCl  Potassium Chloride  
$K_d$  Equilibrium dissociation constant  
kDa  Kilodalton  
KOH  Potassium Hydroxide  
LB  Luria-Bertani  
LDL  Low-density lipoprotein  
LMH  Leghorn male hepatoma cells  
m$^7$G  7-methyl-guanylate  
M  Molar  
MAPK  Mitogen-activated protein kinase  
MCS  Multiple cloning site  
MEL  Murine erythroleukaemia  
Mg  Milligram(s)  
MgCl$_2$  Magnesium Chloride  
MIP-1$\beta$  Macrophage inflammatory protein-1$\beta$  
MKK  MAP kinase kinase  
mL  Millilitre(s)  
mM  Millimolar  
mRNA  Messenger RNA  
mt-acon  Mitochondrial aconitase  
MW  Molecular weight  
NaCl  Sodium chloride  
NEB  New England Biolabs  
NF-L  Neurofilament light peptide  
g  Nanogram  
NGF  Nerve Growth Factor  
NM  Nanomolar  
NMD  Nonsense-mediated mRNA decay  
NMDA  n-methyl-D-aspartate  
NO  Nitric oxide  
NR  Nuclear receptor  
NRF-1  Nuclear respiratory factor-1  
NSAP1  NS1-associated protein  
n  Nucleotide  
OD  Optical Density  
ORF  Open reading frame  
OTC  Ornithine transcarbamoylase precursor  
PABP  Poly(A)-binding protein
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PAI-2</td>
<td>Plasminogen activator inhibitor type 2</td>
</tr>
<tr>
<td>PAIP</td>
<td>PABP-interacting protein</td>
</tr>
<tr>
<td>PB</td>
<td>Phenobarbital</td>
</tr>
<tr>
<td>PBG</td>
<td>Porphobilinogen</td>
</tr>
<tr>
<td>PBGD</td>
<td>PBG deaminase</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PC</td>
<td>Pyruvate carboxylase</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PEPCK</td>
<td>Phosphoenolpyruvate carboxykinase</td>
</tr>
<tr>
<td>PGC-1α</td>
<td>Proliferator-activated receptor γ co-activator 1α</td>
</tr>
<tr>
<td>PKC</td>
<td>Protein kinase C</td>
</tr>
<tr>
<td>PLB</td>
<td>Passive lysis Buffer</td>
</tr>
<tr>
<td>PLP</td>
<td>Pyridoxal 5-phosphate</td>
</tr>
<tr>
<td>pmol</td>
<td>Picomole</td>
</tr>
<tr>
<td>PPOX</td>
<td>Protoporphyrinogen oxidase</td>
</tr>
<tr>
<td>PXR</td>
<td>Pregnane X receptor</td>
</tr>
<tr>
<td>RBB</td>
<td>RNA binding buffer</td>
</tr>
<tr>
<td>RES</td>
<td>Reticuloendothelial system</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RNA-BP</td>
<td>RNA binding protein</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein</td>
</tr>
<tr>
<td>RPA</td>
<td>RNase protection assay</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RPMI</td>
<td>Roswell park memorial institute</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>SA</td>
<td>Succinyl acetone</td>
</tr>
<tr>
<td>SCN</td>
<td>Suprachiasmatic nuclei</td>
</tr>
<tr>
<td>SCSβA</td>
<td>Succinyl CoA synthase beta A subunit</td>
</tr>
<tr>
<td>S.D.</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single stranded DNA</td>
</tr>
<tr>
<td>Strep-AP</td>
<td>Streptavidin-alkaline phosphatase</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA buffer</td>
</tr>
<tr>
<td>TFR</td>
<td>Transferrin receptor</td>
</tr>
<tr>
<td>TGF-β</td>
<td>Transforming growth factor-β</td>
</tr>
</tbody>
</table>
TNF  Tumour necrosis factor

tRNA  Transfer RNA

TX  Triton X-100

U  Unit

uAUG  Upstream AUG

UHW  University Hospital of Wales

Unr  Upstream of N-ras

uORF  Upstream open reading frame

UROD  Uroporphyrinogen III decarboxylase

UROS  Uroporphyrinogen III (co-) synthase

UTR  Untranslated Region

UV  Ultraviolet

UWCM  University of Wales College of Medicine

VEGF  Vascular endothelial growth factor

v/v  Volume/volume

w/v  Weight/volume

XLSA  X-linked sideroblastic anaemia
CHAPTER 1

General introduction
1.1 Structure and Chemistry of Haem

Haem (ferroprotoporphyrin IX) is essential for life, playing a vital role in many ubiquitous and tissue-specific cellular processes. It serves as the prosthetic moiety of numerous haemoproteins, essential for the function of all aerobic cells (Ponka, 1999). The unique property of haem, in having an iron molecule coordinated within a tetapyrrole, allows it to function both as an electron carrier and a catalyst for redox reactions, as well as an oxygen carrier (figure 1-1). The haem molecule plays specific roles in oxygen binding, electron transport, reduction of oxygen and transfer of oxygen for hydroxylation reactions. These roles are in turn determined by the structure of the protein moiety of each specific haemoproteins, their substrates and the intracellular environment where it functions (Bottomley & Muller-Eberhard, 1988).

![Diagram of haem molecule]

Figure 1-1 Structure of haem.
Structurally, haem is a metallo-compound composed of an iron atom coordinated to the tetapyrrole (4 pyrroles) ring system known as the protoporphyrin moiety through its 4 nitrogen atoms (Kaplan, 1977). Haem bears 4 methyl, 2 propionate, and 2 vinyl groups attached as substituents to the pyrrole rings. The 4 pyrrole rings are joined to one another via methene bridges. Iron in haem is bound to a histidine residue of the globin chain and to oxygen that binds at other coordinated position of iron (Tsiftsoglou et al., 2006).
Haem is an iron-protoporphyrin complex consisting of four substituted pyrrole rings linked by via methene bridges. When the iron is in the ferrous state, the complex is called ferroporphyrin or haem, and the molecule is electrically neutral (figure 1-1). When the iron atom is in the ferric state, the complex is called ferriprotoporphyrin or haemin, and the molecule carries a unit positive charge (and is consequently associated with an anion). Haem is hydrophobic in nature and contains methyl groups in positions 1, 3, 5 and 8; vinyl groups in positions 2 and 4; and propionic acid groups in positions 6 and 7 (Kumar & Bandyopadhyay, 2005).

Four types of haem are known in eukaryotes, which differ in the composition of the side chains of the pyrrole rings: protohaem, haem a, b, and c. Protohaem is a pool of “free haem” that is not associated, or only loosely associated, with proteins, e.g. tryptophan pyrrolase and is the precursor for the other types of haem (reviewed in Atamna et al., 2002). The most ubiquitous type of haem is haem b, and constitutes the prosthetic group of all haemoproteins, except cytochromes types c and a. In these cytochromes, haem c and a, respectively, form their prosthetic moiety (Moraes et al., 2004; Kumar & Bandyopadhyay, 2005). Haem b and c are similar to protohaem, with minor modifications that take place as it is incorporated into each specific protein. The synthesis of haem a from protohaem requires two minor modifications: the farnesylation and addition of a formyl group to position 8 of the protohaem (Atamna et al., 2002).

1.2 Function of Haem

Haemoproteins are involved in a remarkable array of crucial biological functions including oxygen binding (haemoglobin, myoglobin), oxygen metabolism (oxidases, peroxidases, and catalases), oxidative phosphorylation (cytochrome c) and electron transfer (cytochrome P450’s) (reviewed in Ponka, 1999) (see table 1-1 for examples of haemoproteins).

Cytochrome P450 is the collective term for a large superfamily of haem-containing proteins that play an important role in the oxidative metabolism of numerous endogenous and foreign compounds (Nelson et al., 1996). Four of the P450-families, families 1 to 4, are involved in drug metabolism and are preferentially expressed in the liver. Drug-metabolizing P450 isoforms also occur in extra-hepatic tissue (Strobel et al., 1997), and several of the P450s expressed
in the brain are inducible by alcohol, neuroleptics, anticonvulsants, and endocrine factors (Volk et al., 1995; Hedlund et al., 1996).

<table>
<thead>
<tr>
<th>Haemoproteins (proteins containing haem as prosthetic group) and haem-based sensors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respiratory cytochromes - cytochrome c, cytochrome c oxidase, cytochrome reductase</td>
</tr>
<tr>
<td>Cytochromes P450</td>
</tr>
<tr>
<td>Catalases</td>
</tr>
<tr>
<td>Peroxidases</td>
</tr>
<tr>
<td>Globin-like coupled sensors</td>
</tr>
<tr>
<td>Histidine kinases</td>
</tr>
<tr>
<td>eIF2α-kinase (translational initiation factor)</td>
</tr>
<tr>
<td>Nitric oxide synthases (NOS): haem-based sensors</td>
</tr>
<tr>
<td>Soluble guanylcyctlases (SGC): haem-based sensors</td>
</tr>
<tr>
<td>Cyclic nucleotides phosphodiesterases</td>
</tr>
<tr>
<td>Haemoglobins (haem-based sensors)</td>
</tr>
<tr>
<td>Myoglobin (haem-based sensors)</td>
</tr>
<tr>
<td>Hydroxylamine oxidoreductase</td>
</tr>
</tbody>
</table>

### Table 1-1 Examples of haemoproteins and haem-based sensors in humans.

Table adapted from Tsiftsoglou et al., 2006.

Haem is also the prosthetic group of numerous haemoproteins that synthesise important regulatory or signalling molecules including cyclic GMP (guanylate cyclase), steroid hormones (hydroxylases) and nitric oxide (nitric oxide synthase). The haem molecule also plays an important role in controlling the expression of numerous proteins (globin, haem biosynthetic enzymes, cytochromes, myeloperoxidase, haem oxygenase-1 (HO-1), transferrin receptor (TfR)) and by providing carbon monoxide (CO), which may have a regulatory role akin to that of nitric oxide (NO) (Ponka, 1999).

CO is a biologically active product of the degradation of haem by the haem oxygenases 1, 2 and 3. Furthermore, this reaction produces the anti-oxidant biliverdin, Fe²⁺ and H₂O₂, which also have biological effects. CO is a signalling molecule with significant vasoactive and anti-inflammatory characteristics (Otterbein et al., 2000). CO selectively inhibits the expression of pro-inflammatory cytokines tumour necrosis factor-α (TNFα), interleukin 1β (IL-1β) and macrophage inflammatory protein-1β (MIP-1β), and the anti-inflammatory cytokine IL-10 (Otterbein et al., 2000). These effects are mediated independently of nitric oxide (NO) or cGMP which has been hypothesised to mediate the effects
of CO in cell culture systems (Morita et al., 1995). CO has been shown in vivo to exert its anti-inflammatory effects specifically through the MAP kinase kinase (MKK) 3/p38 pathway (Otterbein et al., 2000).

The requirements for haem vary significantly among various cells and tissues, with the highest amounts made in erythroid cells and hepatocytes (Ponka, 1999). Approximately 85% of haem in the body is synthesised by erythroid cells and utilised for haemoglobin formation (Harigae et al., 2003), whereas about 15% of the daily production takes place in the liver for the formation of haem-containing enzymes. Between 20 and 35% of newly formed haem is directly converted to bile pigments, suggesting a continuous demand for haem biosynthesis (Atamna, 2004). In senescing erythrocytes, haemoglobin undergoes degradation in the reticuloendothelial system of the liver, kidney, and especially the spleen (Ryter & Tyrrell, 2000). Haem biosynthesis in the brain varies according to the type of brain cell, and appears to be higher in non-neuronal cells and lower in neuronal cells (Atamna, 2004; Atamna et al., 2002).

Free haemoglobin and haem released from methemoglobin may also enter the bloodstream during haemolysis. Such extracellular haem circulates in the serum as complexes with the serum proteins, haemopexin, and albumin. Free vascular haemoglobin is often captured by its scavenger haptoglobin and transported to the reticuloendothelial system (RES) (Muller-Eberhard & Fraig, 1993). Haem and haemoglobin are then imported into the liver parenchyma by receptor-mediated endocytosis, and the haem subsequently degraded by HO activity in the liver microsomes. However, any intracellular “free” haem undergoes degradation in situ, rather than exportation to serum and subsequent hepatic degradation (Ryter & Tyrrell, 2000).

1.3 Haem Biosynthesis

The synthesis of haem involves eight enzymes, four of which are cytoplasmic and four are mitochondrial (reviewed in Ponka, 1999) (figure 1-2). The intermediate products formed in the haem biosynthetic pathway are the reduced forms of porphyrins – porphyrinogens. These undergo rapid oxidation in air to become porphyrins, which are the usual forms of the cyclic tetrapyrroles found in biological specimens. There are four basic processes of the haem biosynthetic pathway (Ajioka et al., 2006): (1) Formation of the pyrrole; (2) Assembly of the
tetrapyrrrole; (3) Modification of the tetrapyrrrole side chains; (4) Oxidation of protoporphyrinogen IX to protoporphyrin IX and insertion of iron.

All mammalian haem pathway enzymes have been cloned and the genes encoding these enzymes reside on different chromosomes (table 1-2).

<table>
<thead>
<tr>
<th>Haem Biosynthetic Enzyme</th>
<th>Chromosome</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aminolaevulinic acid synthase</td>
<td>ALAS1 3p21.1</td>
<td>Bishop, 1990; Cotter et al., 1995</td>
</tr>
<tr>
<td></td>
<td>ALAS2 Xp11.21</td>
<td>Cotter et al., 1992</td>
</tr>
<tr>
<td>Aminolaevulinic acid dehydratase</td>
<td>9q34</td>
<td>Potluri et al., 1987</td>
</tr>
<tr>
<td>Porphobilinogen deaminase</td>
<td>11q24</td>
<td>Namba et al., 1991</td>
</tr>
<tr>
<td>Uroporphyrinogen III synthase</td>
<td>10q25.2 - q26.3</td>
<td>Astrin et al., 1991</td>
</tr>
<tr>
<td>Uroporphyrinogen III decarboxylase</td>
<td>1p34</td>
<td>Dubart et al., 1986</td>
</tr>
<tr>
<td>Coproporphyrinogen III oxidase</td>
<td>3q12</td>
<td>Cacheux et al., 1994</td>
</tr>
<tr>
<td>Protoporphyrinogen III oxidase</td>
<td>1q22</td>
<td>Taketani et al., 1995</td>
</tr>
<tr>
<td>Ferrochelatase</td>
<td>18q21.3</td>
<td>Taketani et al., 1992</td>
</tr>
</tbody>
</table>

Table 1-2 Chromosomes on which the human haem biosynthetic enzymes reside.

In animal cells, anabolic and catabolic pathways are frequently confined to a single cellular compartment unless the process involves import or export of products or substrates. In contrast, the tetrapyrrrole biosynthetic pathway in animal cells has its first and its last three steps catalysed by mitochondrial enzymes, whereas the intermediate four steps are catalysed by cytosolic enzymes (Ponka, 1999). All pathway enzymes are nuclear-encoded and are synthesized in the cytosol with the mitochondrially located proteins being post-translationally translocated to their appropriate mitochondrial space (Dailey et al., 2005). To facilitate their high-level erythroid expression, the first four enzymes in the pathway are encoded by erythroid-specific transcripts generated by separate genes (ALAS; Bishop, 1990) or by genes with alternative housekeeping and erythroid promoters (ALA-dehydratase; Kaya et al., 1994; porphobilinogen deaminase; Grandchamp et al., 1987; and uroporphyrinogen III synthase; Aizencang et al., 2000). In contrast, each of the last four genes in the pathway has a single promoter that contains both erythroid-specific and housekeeping binding elements for housekeeping and enhanced erythroid expression (Delfau-Larue et al., 1994; Romana et al., 1987; Taketani et al., 1995; Tugores et al., 1994).
Figure 1-2 The haem biosynthetic pathway, from Ryter & Tyrrell, 2000.

Tetrapyrrole side chains: A=acetyl, M=methyl, P=propionyl, V=vinyl. ALAS\textsuperscript{a}=ALAS condenses glycine with succinyl-Co-A to form ALA and CO\textsubscript{2}. ALA dehydratase\textsuperscript{b}=ALAD condenses 2 molecules of ALA to PBG, eliminating 2 molecules of H\textsubscript{2}O. PBG deaminase\textsuperscript{c}=PBGD polymerizes PBG (4 molecules) to HMB, eliminating ammonia. UPG III cosynthase\textsuperscript{d}=UROS cyclizes HMB to uroporphyrinogen III. e=Non-enzymatic cyclization of HMB produces the useless uroporphyrinogen I. Further metabolism yields coproporphyrinogen I. Porphyrinogens undergo non-enzymatic oxidation to corresponding porphyrins, (uroporphyrin I, III; coproporphyrin I, III) which accumulate in porphyrias. UPG decarboxylase\textsuperscript{1}=UROD converts uroporphyrinogen III to coproporphyrinogen III by serial decarboxylation. CPG oxidase\textsuperscript{2}=CPOX converts coproporphyrinogen III to protoporphyrinogen IX (PGPIX) by decarboxylation and oxidation of 2 propionyl residues to vinyl groups. PPG oxidase\textsuperscript{1}=PPOX dehydrogenates protoporphyrinogen to protoporphyrin IX (PPIX). Ferrochelatase\textsuperscript{1} incorporates ferrous iron, Fe (II), into PPIX, generating protohaem.
1.3.1 Aminolaevulinate Synthase

Human 5-aminolaevulinate synthase [Succinyl-CoA: glycine C-succinyltransferase (decarboxylating); EC 2.3.1.37] catalyses the first committed step of haem biosynthesis in mammals (Bishop, 1990). This reaction occurs in the mitochondrial matrix and involves the condensation of succinyl coenzyme A (CoA) (from the TCA cycle) and glycine (from the mitochondrial pool of the amino acid) to form 5-aminolaevulinic acid (ALA) (Ponka, 1999). The biosynthesis of one molecule of haem requires eight molecules of glycine and eight molecules of succinyl CoA (Dailey, 1990). There are two forms of ALAS, a housekeeping and erythroid-specific enzyme, named ALAS1 and ALAS2 respectively (Bishop, 1990). Both of which are alternatively spliced (Roberts & Elder 2001; Conboy et al., 1992). A more detailed description of both the ALAS1 and ALAS2 gene is in section 1.7.

1.3.2 Aminolaevulinic acid dehydratase

Following its synthesis, ALA exits the mitochondria by an unknown mechanism. The next four steps of the synthetic pathway take place in the cytosol. ALA dehydratase (ALAD or porphobilinogen synthase) catalyses the asymmetric condensation of 2 molecules of ALA to form monopyrrole porphobilinogen (PBG). The human ALAD is encoded by a single gene located on chromosome 9q34 (Potluri et al., 1987). This gene consists of two alternatively spliced non-coding exons (1A and 1B) and eleven coding exons. The 1A and 1B genes follow housekeeping and erythroid-specific promoters, respectively, which have been demonstrated by CAT expression assays in HeLa and K562 cells. Similar to the ALAS1 gene, the housekeeping promoter of ALAD is active in both erythroid and non-erythroid cells, whereas the erythroid-specific promoter is active only in erythroid cells. As these alternatively spliced exons are present in the 5'-UTR, both transcripts encode identical 36kDa ALAS polypeptides in all cell types (Kaya et al., 1994). The human ALAD gene is similar to that of the mouse (Bishop et al., 1989). This suggests that an early mammalian evolutionary event of the separation of the housekeeping and erythroid-specific upstream exonic promoter sequences occurred to enable sufficient haem production for erythropoiesis (Kaya et al., 1994).
1.3.3 Porphobilinogen deaminase

The next two enzymatic steps convert four molecules of PBG into the cyclic tetrapyrrrole uroporphobilinogen III, which is then decarboxylated to form coproporphyrinogen III. Porphobilinogen deaminase (PBGD; hydroxymethylbilane synthase) is located in the cytosol and is the third enzyme of the haem biosynthetic pathway, which catalyses the stepwise condensation of four PBG units to form the linear tetrapyrrrole hydroxymethylbilane (HMB) (Battersby et al., 1980). Similar to ALAS and ALAD, two isoforms of PBGD exist which differ by approximately 2kDa, one present in all cells and the other erythroid-specific (Grandchamp et al., 1987). The PBGD gene is 10kb long, located at chromosome 11q24, and split into 15 exons (Namba et al., 1991; Yoo et al., 1993). Alternative splicing in the 5'-UTR results in two different primary transcripts, where the human PBGD gene is expressed in all tissues due to a housekeeping promoter, whereas the erythroid-specific promoter located 3kb downstream, is switched on during erythroid differentiation (Chretien et al., 1988). The isoforms differ only at their NH₂ ends, where the ubiquitous isoforms extend for an additional 17 residues, making a polypeptide of 361 amino acids (Grandchamp et al., 1987).

1.3.4 Uroporphyrinogen III synthase

Uroporphyrinogen III synthase (URO-synthase, UROS) then rapidly converts HMB to uroporphyrinogen III, the first cyclic tetrapyrrrole of the haem biosynthetic pathway, by an intramolecular rearrangement of the D-pyrrole group and ring closure. The UROS human gene is ~34kb and consists of 10 exons. Its expression is controlled by having two distinct erythroid and housekeeping promoters in its 5'-UTR, rather than the use of alternative splicing (Aizencang et al., 2000).

1.3.5 Uroporphyrinogen III decarboxylase

Human uroporphyrinogen III decarboxylase (UROD) catalyses the fifth step of the haem biosynthetic pathway. This is the sequential elimination of the carboxyl groups from each of the four acetate side chains of the uroporphyrinogen III substrate, yielding coproporphyrinogen III. The UROD gene has been assigned to the p34 band of chromosome 1 (Dubart et al., 1986), consisting of 10 exons, and is over 3kb long (Romana et al., 1987). RNase mapping and primer extension analysis has shown that transcription is initiated at two start sites, only 6 bases
apart. These initiation sites are used in all tissue types, with no specific promoter used for erythroid cells (Romana et al., 1987).

### 1.3.6 Coproporphyrinogen oxidase

The last three steps of the biosynthetic pathway take place in the mitochondria (Ponka et al., 1973). These are the oxidative decarboxylation of coproporphyrinogen III to protoporphyrinogen IX, followed by the six-electron oxidation of protoporphyrinogen to protoporphyrin IX, and finally the insertion of ferrous iron to form haem, catalysed by the enzymes coproporphyrinogen oxidase (CPOX), protoporphyrinogen oxidase (PPOX) and ferrochelatase respectively (reviewed in Dailey, 2002).

CPOX is a soluble mitochondrial protein that is localised in the intermembrane space within mammalian cells, and catalyses the sixth step in haem biosynthesis, the conversion of the two propionate groups at positions 2 and 4 of coproporphyrinogen III to two vinyl groups, thus producing protoporphyrinogen IX (Elder & Evans, 1978; Martasek et al., 1994). The CPOX gene has been mapped to chromosome band 3q12 (Cacheux et al., 1994), spanning 14kb and consisting of seven exons and six introns (Delfau-Larue et al., 1994). This protein possesses an unusually long N-terminal leader sequence of more than 120 amino acids, which is proteolytically removed (Delfau-Larue et al., 1994). The mature protein is a homodimer and possesses no known cofactor (Medlock & Dailey, 1996).

### 1.3.7 Protoporphyrinogen oxidase

PPOX is the penultimate enzyme of the pathway, and catalyses the six electron oxidation of protoporphyrinogen IX to form protoporphyrin IX. The enzyme is located in the inner mitochondrial membrane of various tissues including liver, lymphocytes and cultured fibroblasts. Its gene is mapped to chromosome 1q22, which spans approximately 8kb and consists of 13 exons and 12 introns (Taketani et al., 1995). Human PPOX contains 477 amino acids and has a molecular mass of 50.8 kDa (Nishimura et al., 1995). Eukaryotic mitochondrial PPOXs are strongly associated with the mitochondrial inner membrane, and the active site has been suggested to be on the cytosolic surface of the membrane (Deybach et al., 1985; Ferreira et al., 1988). The information required for efficient import of human PPOX into the mitochondria is contained within the first 250 amino acid
residues. Fully efficient targeting requires both a major targeting signal, which may be contained between residues 151 and 175, and involved in anchoring to the inner mitochondrial membrane, together with sequences within the first 150 residues (Morgan et al., 2004). Primer extension analysis has shown that PPOX mRNA synthesis is initiated at two major sites, used similarly by both erythroid and non-erythroid cells (Taketani et al., 1995).

1.3.8 Ferrochelatase

The terminal enzyme of the haem biosynthetic pathway is ferrochelatase (FECH), catalysing the insertion of ferrous iron into protoporphyrin IX, and yielding haem. In eukaryotic cells this enzyme is a mitochondrial inner-membrane associated protein, with the active site facing the matrix (Ferreira et al., 1995). The ferrochelatase gene is located on chromosome 18, at region q21.3. It contains a total of 11 exons with a minimum size of approximately 45kb. Primer extension analysis shows that transcription of ferrochelatase initiates at the same site in both hepatoma HepG2 and erythroleukemia K562 cell mRNA, suggesting that there is a single transcript which functions in both erythroid and non-erythroid cells (Taketani et al., 1992).

1.4 Mammalian Haem Transport

Protohaem made by ALAS1 is released into the mitochondrial matrix, where part of it is utilised for assembly of mitochondrial proteins, and an appreciable fraction is exported to the cytosol to be assembled into different cytosolic proteins (Ponka et al., 1973).

Newly formed haem is rapidly incorporated into cellular haem proteins, among which the microsomal cytochromes P-450 are quantitatively the most important in non-erythroid cells. In rats, turnover of these cytochromes appears to consume at least 50% of the haem produced in the liver (Grandchamp et al., 1981). After the translocation of haem across the mitochondrial membranes to varying locations, a proportion is then recruited to the endoplasmic reticulum to be degraded by haem oxygenase (HO) (Atamna et al., 2002).

Transmembrane transport of haem is essential both for the absorption in the gastrointestinal tract to maintain body iron homeostasis and for intracellular trafficking in other cells and tissues. Due to its lipophilic nature, haem can diffuse
into cell membranes; however in several cases transmembrane transport of haem is an energy-dependent process, requiring haem transport proteins (Latunde-Dada et al., 2006).

The carrier proteins and general mechanism that mediate this shuttling of haem between the different compartments of the cell have not been well understood. There have been a number of different putative haem-transport-associated proteins, of which their function and tissue localisation is listed in table 1-3.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Localization</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCP1</td>
<td>Duodenum, liver, kidney</td>
<td>Duodenal apical haem transporter, possible haem scavenging or recovery</td>
<td>Shayeghi et al., 2005</td>
</tr>
<tr>
<td>FLVCR</td>
<td>Haematopoietic cells, liver, kidney, pancreas</td>
<td>Haem efflux in erythroid progenitor cells</td>
<td>Quigley et al., 2004</td>
</tr>
<tr>
<td>ABCG2</td>
<td>Haematopoietic cells, fetal liver, pancreas, kidney, intestine</td>
<td>Haem porphyrin efflux</td>
<td>Krishnamuphy et al., 2004</td>
</tr>
<tr>
<td>SOUL</td>
<td>Eye</td>
<td>Putative haem-binding protein</td>
<td>Sato et al., 2004</td>
</tr>
<tr>
<td>ABCme</td>
<td>Erythroid cells, inner mitochondrial membrane</td>
<td>Possible transport of haem from mitochondria to cytosol</td>
<td>Lill &amp; Kispal, 2001</td>
</tr>
<tr>
<td>ABC3, ABC7, ABCme</td>
<td>Mitochondria</td>
<td>Transport of haem derivatives</td>
<td>Lill &amp; Kispal, 2001</td>
</tr>
<tr>
<td>Haemopexin receptor LRP or CD91</td>
<td>Hepatocyte</td>
<td>Haem endocytosis from circulation</td>
<td>Hvidberg et al., 2005</td>
</tr>
<tr>
<td>Haptoglobin receptor. CD163</td>
<td>Macrophages</td>
<td>Haemoglobin clearance from circulation</td>
<td>Kristiansen et al., 2001</td>
</tr>
<tr>
<td>Megalin and cubulin receptors</td>
<td>Kidney</td>
<td>Renal haemoglobin reabsorption</td>
<td>Gburek et al., 2002 &amp; 2003</td>
</tr>
</tbody>
</table>

Table 1-3 Function and tissue localisation of putative haem-transport-associated proteins; table adapted from Latunde-Dada et al., 2006.

HCP1 = haem carrier protein 1; FLVCR = Feline leukaemic virus receptor; ABC = ATP binding cassette transporter; LRP/CD91 = low-density lipoprotein receptor-related protein CD91.
Recently, Krishnamurthy et al (2006) identified the ATP-binding cassette transporter, ABCB6, as a mammalian mitochondrial porphyrin transporter. Their experiments showed that the Acbc6 mRNA and ABCB6 protein are strongly elevated by exogenously added haemin, in murine erythroleukaemia (MEL) cells. Treatment of MEL cells with ALA to endogenously induce haem biosynthesis caused a dose and time-dependent increase in ABCB6 protein and mRNA and intracellular protoporphyrin IX (PPIX) and ABCBC6 concentrations were correlated. Conversely, the inhibition of haem biosynthesis by the addition of succinyl acetone (SA) to MEL cells prevented an increase in intracellular PPIX and blocked upregulation of ABCB6 synthesis by ALA. Thus ABCB6 expression levels respond directly to the intracellular PPIX concentration. This has also been confirmed in vivo, using two mouse models. Moreover, not only does ABCB6 protein increase during erythroid differentiation and haem biosynthesis and is directly upregulated by increased intracellular porphyrin, ABCB6 overexpression also enhances porphyrin biosynthesis by upregulating the key genes for transferrin receptor, ALAD, ALAS and coporphyrinogen oxidase (Krishnamurthy et al., 2006).

Work published by Smith & Morgan (1984) and Morgan (1976) indicated that haem released from haemoglobin in plasma, following red blood cell degradation in spleen, is transported into liver cells via haemopexin, a carrier plasma protein. Haemopexin forms stable complexes with haem that are internalized inside the cells. Haemopexin returns intact to the extracellular space just after releasing exogenous haem into intracellular cytosolic compartment for reutilization. This carrier-mediated import of haem via haemopexin is a saturation-, time-, and energy-dependent as well as tissue-specific process. Iron released from haem by microsomal haem oxygenase is then stored into ferritin. This haemopexin-mediated import of haem does not account for all, but only a portion of transported haem (reviewed in Baker et al., 2003; Hvidberg et al., 2005).

The intestinal haem transporter HCP1 was identified by Shayeghi et al (2005), which is capable of mediating energy-dependent transmembrane uptake of the intact metal-porphyrin ring, and is used to absorb haem from the diet into the gut (Latunde-Dada et al., 2006). The identification of this transporter revealed the mechanism by which haem is absorbed from the diet by the gut, as an important source of dietary iron.
1.5 Role of Haem Biosynthesis in Disease

Both inherited mutations and environmental factors may affect the haem biosynthetic pathway and lead to diseases including X-linked sideroblastic anaemia (XLSA), lead poisoning and the porphyrias (table 1-4) (Anderson et al., 2000; Sassa, 2000; Jordan & Dailey, 1990). Porphyrias are mainly inherited in an autosomal dominant manner with incomplete penetrance, but autosomal recessive and more complex patterns of inheritance are also possible (reviewed in Kauppinen, 2005).

Porphyrias represent overproduction syndromes with the formation of toxic haem precursors, which are overproduced under particular circumstances, with acute porphyrias resulting in neurovisceral manifestations. Overproduction of porphyrins causes photosensitivity in cutaneous porphyrias (Kauppinen, 2005). Compelling evidence has been generated indicating that neurovisceral symptoms are due to neurotoxic affects of the pre-porphyrin precursors, which is primarily ALA. The photosensitivity is due to the fluorescent properties of the porphyrins (Sassa, 2006).

A convenient way to classify the porphyrias is to divide them according to the dominant clinical feature, although some authors prefer to divide the porphyrias based on whether excess substrate is generated in the liver (hepatic porphyrias) or the erythrocyte (erythropoietic porphyrias) (Sassa, 2006).

During an acute attack, ALAS1, the hepatic isoform of the first enzyme in the haem biosynthetic pathway is induced. The acute hepatic porphyrias are ALA dehydratase porphyria (ADP), acute intermittent porphyria (AIP), variegate porphyria (VP) and hereditary coproporphyria (HCP). ALAS1 formation in normal hepatocytes is repressed by feedback inhibition by the final product, haem. Metabolic inhibition along the pathway in the liver leads to reduced production of haem, resulting in de-repression of ALAS1. This leads to increased production of haem precursors in an effort to overcome the metabolic block, and contributes to the accumulation of intermediates prior to the deficient enzymatic step. This abnormality continues until sufficient haem synthesis is restored (Sassa, 2006).
<table>
<thead>
<tr>
<th>Enzyme in Haem Biosynthesis</th>
<th>Disease</th>
<th>Type</th>
<th>Symptoms</th>
<th>Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALAS2</td>
<td>X-linked sideroblastic anaemia (XLSA)</td>
<td>Erythroid</td>
<td>Microcytic anaemia</td>
<td>Sideroblasts</td>
</tr>
<tr>
<td>ALAD</td>
<td>ALA dehydratase porphyria (ADP)</td>
<td>Hepatic</td>
<td>Neurovisceral</td>
<td>Urinary ALA</td>
</tr>
<tr>
<td>PBGD</td>
<td>Acute intermittent porphyria (AIP)</td>
<td>Hepatic</td>
<td>Neurovisceral</td>
<td>Urinary ALA, PBG</td>
</tr>
<tr>
<td>UROS</td>
<td>Congenital erythropoietic porphyria (CEP)</td>
<td>Erythropoietic</td>
<td>Photosensitivity</td>
<td>Haemolytic anaemia</td>
</tr>
<tr>
<td>UROD</td>
<td>Porphyria cutanea tarda (PCT)</td>
<td>Hepatic</td>
<td>Photosensitivity</td>
<td>7-C porphyrin; faecal isocoprotoporphyrin</td>
</tr>
<tr>
<td></td>
<td>Hepatoerythropoietic porphyria (HEP)</td>
<td>Erythropoietic</td>
<td>Haemolytic</td>
<td>anaemia</td>
</tr>
<tr>
<td>CPOX</td>
<td>Hereditary coproporphyria (HCP)</td>
<td>Hepatic</td>
<td>Neurovisceral Photosensitivity</td>
<td>Urinary ALA, PBG, coproporphyrin</td>
</tr>
<tr>
<td>PPOX</td>
<td>Variegate porphyria (VP)</td>
<td>Hepatic</td>
<td>Neurovisceral Photosensitivity</td>
<td>Urinary ALA, PBG; faecal protoporphyrin</td>
</tr>
<tr>
<td>FECH</td>
<td>Erythropoietic protoporphyrria (EPP)</td>
<td>Erythropoietic</td>
<td>Photosensitivity</td>
<td>RBC protoporphyrin Faecal protoporphyrin</td>
</tr>
</tbody>
</table>

**Table 1-4 Classification of the porphyrias; adapted from Sassa, 2006.**

Enzymatic defects, associated diseases, major symptoms and principal accumulation products are shown. ALAS2 defect is responsible for XLSA but is not associated with any porphyria, since the enzymatic defect blocks production of ALA, the precursor for porphyrin formation. ADP and AIP are accompanied by acute hepatic porphyria but not by photocutaneous porphyria, because their enzymatic defects do not result in an increase in porphyrin synthesis. Enzymatic defects beyond UROS are all associated with photocutaneous porphyrias, because they produce excessive amounts of various porphyrins. HCP and VP are additionally associated with acute hepatic porphyria.

Transcription of ALAS1 can also be upregulated by the peroxisome proliferators-activated receptor γ-coactivator 1α (PGC-1α) (described in more detail in section 1.8.2.2). Under conditions of low glucose, PGC-1α production increases, leading to increased levels of ALAS1 and creating conditions for attacks of acute porphyrias (Handschin et al., 2005).
1.5.1 Haem Excess

Haem is absolutely required for aerobic life. However, haem has to be tightly regulated by cells as a deficiency or excess can have cytotoxic effects (figure 1-3), particularly in the presence of oxidants or activated phagocytes (Yoo et al., 1999; Kumar & Bandyopadhyay, 2005). Of all the sites of the body, the vasculature, and in particular the endothelial lining, may be at greatest risk of exposure to free haem. This is because erythrocytes contain haem in a concentration of 20mmol/L and are vulnerable to unexpected lysis. The extramolecular haemoglobin is easily oxidised to ferrihaemoglobin, which, in turn, will readily release haem. Given the hydrophobic nature of haem, it easily crosses cell membranes, then concentrates within the hydrophobic milieu of intact cells, and can synergistically enhance oxidant damage. Both in vitro and in vivo, cells will accumulate exogenous haem and synergistically amplify the cytotoxic effects of oxidants of reagent, enzymatic, or cellular origin (Balla et al., 1991).

Haem also acts as a catalyst for the oxidation of low-density lipoprotein (LDL), generating products toxic to endothelium (Camejo et al., 1998). These toxic effects may be important in a number of pathologies such as the acute conditions intravascular haemolysis as well as insidious processes such as atherogenesis. As a defence against such toxicity, normal cells upregulate HO-1 and ferritin (Jeney et al., 2002). High levels of free haem are found in pathological states of increased haemolysis, such as sickle cell disease, malaria and ischemia reperfusion. Under these conditions, the physiological mechanisms of removing free haem from the circulation, especially its binding to haemopexin, collapse, allowing non-specific haem uptake and haem-catalysed oxidation reactions (Arruda et al., 2004).
Figure 1-3 Toxic effects of free haem/haemin. Adapted from Kumar & Bandyopadhyay, 2005.

Haem, a ubiquitous iron-containing compound, is quite hydrophobic, readily enters cell membranes, and greatly increases cellular susceptibility to oxidant-mediated killing (Balla et al., 1991). Haem also acts as a catalyst for the oxidation of low-density lipoprotein (LDL), generating products toxic to endothelium (Balla et al., 1991, Miller et al., 1994, Camejo et al., 1998).

The accumulation of haem within the cell can ultimately lead to the accumulation of iron and the production of cell-damaging reactive oxygen species by the Fenton reaction. This is the reaction between ferrous iron and hydrogen peroxide to produce powerful oxidising intermediates, with \( \cdot \text{OH} \) being the most powerful oxidising species.

\[
\text{Fe}^{2+} + \text{H}_2\text{O}_2 \rightarrow \text{Fe}^{3+} + \cdot \text{OH} + \text{OH}^-
\]
(Wink et al., 1994)

Haem/porphyrin accumulation also leads to the collapse of mitochondrial function. The regulation of intracellular porphyrin levels is therefore fundamental to cell
survival. In erythroid cells, this regulation is especially important under conditions of low oxygen, when the cellular concentration of haem may increase. This is due to the increased ALAS2 transcription in a hypoxia-inducible factor-1 (HIF-1) independent manner, even though a single putative HIF-1 DNA binding site (HBS) has been identified within the promoter region of murine ALAS2 (Hofer et al., 2003; Kramer et al., 2000). Hypoxia has not been shown to in induce ALAS1 expression.

Once within the cell, haem can promote oxidative damage either directly or, perhaps more importantly, directly via the release of iron. This can occur through either non-enzymatic oxidative degradation of haem, or enzymatic, HO-catalysed haem cleavage. In either case the iron may initially lodge within the hydrophobic interstices of the phospholipid bilayer; within this highly oxidizable matrix, iron acts as an especially active catalyst of oxidation of cell membrane constituents (Balla et al., 2005).

Severe haemolysis occurring during pathological states, such as sickle cell disease, ischemia reperfusion, and malaria results in high levels of free haem (up to 20µM). Under these conditions the physiological mechanisms of removing free haem from the circulation collapse, allowing non-specific haem uptake and haem catalysed oxidation reactions. When large amounts of free haem proteins or haem accumulate, e.g. in a blood clot, the haem detoxification systems get overwhelmed, or are unable to reach them. This enables haem to exert its damaging effects (Kumar & Bandyopadhyay, 2005). These large amounts of exogenous haem, produced in a haemorrhage for example, may be toxic to neurons (Goldstein et al., 2003).

Cellular free haem may increase by pharmacological stimulation with porphyrinogenic agents, by accelerated degradation of intracellular haemoproteins, by impaired incorporation into apo-haemoproteins, or as a result of inhibition of cellular HO activity (Scriver et al., 1995). Cellular stress may also alter haemoprotein turnover. For example, oxidative stress may modify protein structure, causing fragmentation, crosslinking, and changes in charge, hydrophobicity, electrophoretic mobility, and solubility. Such modification, occurring in a haemoprotein, may either trigger the release of the haem, or alter the reactivity of hypervalent states of the protein towards organic molecules in the surrounding cellular context. Oxidative stress generated by UVA (Ultraviolet A) radiation increases degradation of cytochrome p-450 and causes accumulation
of free haem in fibroblasts. This free haem is a potentially dangerous molecule, and therefore, in order to preserve cellular haemostasis, its cellular concentration must be tightly regulated (Ryter & Tyrrell, 2000).

1.5.2 Haem Deficiency

Decreased haem biosynthesis in erythroid cells causes symptoms such as anaemia. The known consequences of decreased levels in non-erythroid cells are loss of mitochondrial complex IV, oxidative stress, accumulation of iron, and cell death whenever stimulation by growth factors occurs. The affects of changes of haem in vivo on the brain have not been fully elucidated (Atamna, 2004).

Haem deficiency prevents the assembly of complex IV, the terminal complex of the electron transport chain, by more than 95% and mildly affects complex II and cytochrome c (by 30%) in human fibroblasts (Atamna et al., 2001). The complex IV is tightly coupled to neuronal metabolic activity; therefore brain may be the most susceptible organ for a decrease in complex IV and for oxidative stress, which occur in ageing and Alzheimer's disease (Maurer et al., 2000). Levels of haem a in Alzheimer patients decrease by 22%, while haem b and c remain relatively unchanged, when compared to healthy controls (Atamna et al., 2002).

Haem deficiency obliterates the activation of the Ras-MAPK signalling pathway and its downstream transcription factor, which suggests that haem deficiency should ultimately cause neuronal gene expression. In addition to the genes affected by the Ras-MAPK signalling pathway and CREB, haem deficiency may affect the expression of numerous other neuron-specific genes (Zhu et al., 2002). Research by Zhu et al (2002) on haem-depleted PC12 cells, using succinyl acetone in the growth medium, greatly reduces NGF (nerve growth factor)-induced neurite outgrowth. The addition of exogenous haem to the succinyl acetone-treated cells largely restored neurite outgrowth, supporting the idea that the arrest of neurite outgrowth is caused by a lack of haem synthesis and suggesting that haem is critical for NGF signalling (Zhu et al., 2002). NGF is a neurotrophic factor important for both the peripheral nervous system and the CNS (Gage et al., 1990).

As described in section 1.2, haem is the prosthetic group in the P450 cytochromes (CYPs) (Rangarajan & Padmanaban, 1989). CYPs perform their manifold functions as functional holoenzymes saturated with haem (Correia &
Meyer, 1975; Jover et al., 1996). Thus, the regulation of haem availability is of major importance for CYP function. However, regulation of haem availability for CYPs is tissue-dependent. CYP1A1 is involved in the metabolic activation of polyaromatic hydrocarbons (Hankinson et al., 1991), and is expressed in most tissues. However, its function is unclear in the kidney and brain due to a much lower enzymatic activity. In the liver, partial haem deficiency leads to a less potent induction of CYP1A1 mRNA after β-NF treatment. In the brain, CYP1A1 protein is detected not only at the endoplasmic reticulum (ER), but also in the cytosol of transgenic mice with chronic impairment of haem biosynthesis – porphobilinogen deaminase-deficient (PBGD/-) mice. Therefore, haem deficiency in the brain leads to incomplete haem saturation of CYP1A1, which causes its improper incorporation into the ER membrane and persistence in the cytosol (Meyer et al., 2005).

In the liver, many drugs and chemicals affect haem synthesis, and this may also occur in the brain (De Matteis & Ray, 1982). As in the liver, ALAS1 activity in the brain declines with age, reducing the amount of haem available (Paterniti et al., 1978). Furthermore, the binding of haem to amyloid-β in Alzheimer’s disease may also limit its availability to cells (Atamna & Frey, 2004). Chernova et al (2006) have shown that inhibition of haem synthesis in neurons results in the earlier lowered expression of the neuron-specific genes N-methyl-D-aspartate (NMDA) receptor subunits and neurofilament light peptide (NF-L). Increased expression was then observed when haemin was exogenously added to the cells. In addition, lowered haem concentration caused by an increase in expression of ALAS1 and HO-1 is associated with the senscense of neurons. This may give an indication into how neurons functionality declines with age due to the reduction in haem (Chernova et al., 2006).

1.6 Homeostasis of Haem

Ultimately, the cellular concentration of haem in non-erythroid cells is controlled by its degradation and synthesis, controlled by the haem oxygenases (HOs) and ALAS1, respectively. Therefore, HO will be briefly discussed before moving on to both ALAS1 and ALAS2.
1.6.1 Haem Oxygenase

HO plays a critical role in physiological iron homeostasis, antioxidant defence, and, as previously mentioned, in signalling pathways that employ CO as a messenger (Gilles-Gonzalez et al., 2005). The HO system, oxidatively cleaves haem to produce CO, biliverdin, and free iron (figure 1-4). Three mammalian isoforms of HO have been identified: HO-1, an inducible enzyme that is most highly concentrated in tissues that are heavily involved in the catabolism of haem proteins (reviewed in Maines & Gibbs, 2005); HO-2, a non-inducible isoform that is present in highest concentrations in the brain and testes and is thought to be particularly involved in signalling pathways (Maines et al., 1986); and HO-3, a homologue of HO-2 with low catalytic activity and uncertain physiological role. HO-3 has been described in the rat brain (McCoubrey et al., 1997), however no homologue of this gene is found in either humans or mice and it is best characterized as a processed pseudogene (Hayashi et al., 2004; Maines et al., 2004).

HO-1 and HO-2 both have haem-binding pockets, a stretch of 23 amino acids, a basic sequence mostly made up of hydrophobic residues with a conserved histidine residue (McCoubrey et al., 1993). Outside the haem-binding pocket, the extent of sequence identity between HO-1 and HO-2 is limited to less than 50% (Maines & Gibbs, 2005).

![Haem degradation pathway](image)

**Figure 1-4 The pathway of haem degradation in mammals (reviewed in Maines & Gibbs, 2005).**

M=methyl, V=vinyl, P=propionic acid.

Oxidation of haem by the HO proteins requires concerted activity of NADPH-cytP450 reductase (figure 1-4) for providing reducing equivalents for supporting
the reduced state of iron (Fe\textsuperscript{2+}) and activation of molecular oxygen. The specific orientation of bound haem is crucial to specificity of cleavage of the porphyrin ring at the α-meso carbon bridge. The reaction requires 3 molecules of oxygen and 7 electrons. The rate-limiting step in oxidation of haem is the release of the product, biliverdin (Maines & Gibbs, 2005).

### 1.6.2 Aminolaevulinate Synthase (ALAS)

The reaction catalysed by ALAS involves the condensation of glycine and succinyl-CoA to produce ALA, CoA, and carbon dioxide. This occurs in a wide variety of species including animals, fungi, some protozoa and α-purple bacteria. In contrast, ALA in plants and most other bacteria is made from glutamate by an entirely different biosynthetic pathway involving three enzymatic steps and a glutamyl-tRNA intermediate (Avissar et al., 1989).

![Figure 1-5 Synthesis of ALA by ALAS, adapted from Ajioka et al., 2006.](image)

Decarboxylation of glycine followed by condensation with Succinyl-CoA is catalyzed by ALAS. Pyridoxal Phosphate is required as a co-factor. The products are ALA, CO\textsubscript{2} and CoA.
The crystal structure of the full-length ALAS has only been described in *R. capsulatus*, and so due to the homology between ALASRC and human ALAS2, a model has been derived for the core of human ALAS2 (Astner *et al.*, 2005).

### 1.6.2.1 ALAS as a PLP-dependent enzyme

ALAS requires pyridoxal 5'-phosphate (PLP, a vitamin B6 derivative) as an essential cofactor, and belongs to the α-oxoamine synthase subfamily of PLP-dependent enzymes (Zhang & Ferreira, 2002). Both forms of ALAS require PLP as a cofactor and both are expressed as homodimers. PLP binds to a specific lysine of ALAS. A PLP-glycine Schiff base complex is then formed which reacts with succinyl Co-A. ALAS catalyses the decarboxylation of glycine followed by condensation of Succinyl-CoA, producing CO₂, CoA and ALA (Ponka, 1999).

Almost all PLP-dependent enzymes, except glycogen phosphorylases are associated with biochemical pathways that involve amino compounds, mainly amino acids. These reactions include the transfer of the amino group, decarboxylation, interconversion of L- and D-amino acids, and removal or replacement of chemical groups at the β- and γ-carbon (reviewed in Percudani & Peracchi, 2003). Specifically, ALAS is in the α-oxoamine class of PLP-dependent enzymes, which is also known as class II of fold-type I. Other members include 8-amino-7-oxononanoate synthase (AONS), 2-amino-3-ketobutyrate CoA ligase, and serine palmitoyltransferase, which all possess similar catalytic mechanisms (Zhang & Ferreira, 2002). Fold-type I is the most common structure, and is found in a variety of aminotransferases and decarboxylases, as well as in enzymes that catalyse α-, β- or γ-eliminations (Percudani & Peracchi, 2003). The class II types typically catalyse Claisen condensations between amino acids and acyl-CoA thioesters with concomitant decarboxylation (Kerbarh *et al.*, 2006).

In the α-family of PLP-dependent enzymes, the covalency changes occur at the α-carbon of the amino acid that forms an aldimine linkage to the PLP cofactor through the amino group (Alexander *et al.*, 1994). The members of this enzyme family are believed to share a common evolutionary origin, with the absolute conservation of an aspartate residue approximately 30-40 residues to the N-terminus of the conserved lysine that forms the Schiff base linkage with the cofactor (Gong *et al.*, 1998).
1.6.2.2 The ALAS catalytic mechanism

To distinguish between the two possibilities for the active site arrangement of ALAS dimer (inter- and intrasubunit), Tan & Ferreira (1996) developed in vivo complementation stages in which two inactive mouse mutants were co-expressed in *Escherichia coli* (*E.coli*). Their results support the proposal that the active site of ALAS is shared by the two subunits and that therefore it resides at the subunit interface and dimerization is necessary for the functional ALAS enzyme (Tan & Ferreira, 1996).

ALAS exists in two conformational states, an open conformation in which the substrates bind and a closed conformation induced by the binding of the second substrate, succinyl-CoA. The reaction of glycine with ALAS involves a three-kinetic step (i.e. formation of the Michaelis complex and external aldimine and removal of the glycine C-α proton from the external aldimine). The presence of succinyl-CoA accelerates this third step and thus the intermediate formation of two quinonoid species. ALA release is the rate-limiting step in the ALAS catalytic pathway (Zhang & Ferreira, 2002).

Firstly, a transaldiminization reaction occurs between the substrate glycine and the internal aldimine, formed between the aldehyde group of PLP and an active site lysine residue. Following the transaldiminization reaction and formation of the external aldimine, the pro-R proton of glycine is removed, yielding a transient quinonoid species in the presence of succinyl-CoA. This quinonoid intermediate reacts with succinyl-CoA to form, upon removal of the CoA group, an aldimine to α-amino-β-ketoadipate. Subsequently, cleavage of the C-α-carboxylate bond leads to a stabilized quinonoid intermediate. A proton is then returned to the C-5 position of the aldimine between PLP and ALA. Finally, the aldimine dissociates to yield ALA and the holoenzyme (Zhang & Ferreira, 2002).
Figure 1-6 Scheme of the ALAS catalytic mechanism, taken from Zhang & Ferreira, 2002.

Radiolabeling studies using *Rhodobacter sphaeroides* ALAS provided information to develop a model of the ALAS catalytic reaction mechanism (Laghai & Jordan, 1977; Zhang & Ferreira, 2002).

The presence of a lysine residue in the ALAS2 active site, at position 313 appears to be essential to form the quinonoid intermediate in this reaction (Zhang & Ferreira, 2002). K313 of the murine mature ALAS2 acts as a general base to labilize both the pro-R proton of glycine and the pro-R position at C-5 of ALA. The lysine also acts as a general acid to donate a proton to the quinonoid intermediate formed from decarboxylation of the α-amino-β-ketoadipate intermediate (Hunter & Ferreira, 1999).

Also, K313 in mouse ALAS2 appears to be the residue involved in the binding of PLP cofactor, as mutagenesis of this residue to alanine and histidine abolishes the enzyme activity (Ferreira *et al.*, 1995). The arginine 149 (R149) in mouse ALAS2 had also been shown to be catalytically essential, which is highly conserved in all known ALAS enzymes (Gong & Ferreira, 1995).

The mechanism of the human ALAS enzymes have not been elucidated as yet. However, due to the high homology between species of the PLP-binding peptide, it is believed to be similar to the mechanism for murine ALAS2, as described
above. This PLP-binding peptide, consisting of the sequence surrounding the lysine involved in forming the Schiff base linkage with PLP, is present in all known ALAS sequences ranging from human to bacteria, and is present in the conserved catalytic domain (Ferreira et al., 1993). The human ALAS1 lysine residue that corresponds to the essential K313 of murine ALAS2 is K393 (Bawden et al., 1987). The essential lysine in human ALAS2 is K344 (Cox et al., 1991; Borthwick et al., 1985).

1.6.2.3 ALAS Regions

ALAS1 and ALAS2 are composed of three regions: region I is the mitochondrial targeting sequence, region II is the amino-terminal portion of the mature, proteolytically processed protein, and region III is the highly conserved catalytic region of the enzyme (Dailey et al., 2005). The bacterial ALAS enzymes characterised only consist of region III, confirming the catalytically-active domain in this region (Duncan et al., 1999). ALAS is a homodimer, and is translocated into the mitochondrion in a step that proteolytically removes the leader sequence (Region I), and subsequently the cofactor, PLP, is added. Region II, the N-terminal end of the mature protein, of ALAS1 is encoded by exons 2 and 3, and is not required for protein activity; therefore its function is unknown. The ALAS2 region II is encoded by exons 3 and 4 (Dailey et al., 2005).

ALAS is synthesised on cytoribosomes as a precursor form and is translocated to the mitochondria, where it accumulates in the extramitochondrial space. ALAS is then processed to give the mature enzyme (Yamauchi et al., 1980; Urban-Grimal et al., 1986). The capacity of the mitochondria to accommodate ALAS appears to be limited to a certain level, and this also favours the accumulation of ALAS in the extra-mitochondrial fraction in later stages of induction. It accumulates in large amounts in the liver cytosol fraction in chemically-induced porphyria animals, before it is translocated (Kikuchi & Hayashi, 1981).

1.7 Comparison of the ALAS1 and ALAS2 Genes

There are two genes for ALA synthase (ALAS), one of which is expressed ubiquitously with the other specific to erythroid cells. Consequently, there are two corresponding isoforms of the enzyme, named ALAS1 and ALAS2 respectively (Ponka, 1999). In humans, ALAS1 gene is on chromosome band 3p21.1, whilst the ALAS2 gene is in a distal subregion of Xp11.21 (Bishop, 1990; Cotter et al.,
1992) (table 1-2). The housekeeping enzyme, ALAS1, is expressed in all tissues to provide haem for respiratory cytochromes and other haemoproteins (May et al., 1995). The ALAS1 gene is expressed in both erythroid and non-erythroid cells, although its expression varies between tissues. Highest levels are found in the liver where ALAS1 transcription can be induced directly by compounds such as phenobarbital to co-ordinate an increased demand for haem with increased apo-cytochrome synthesis (Srivastava et al., 1988).

The ALAS2 pre-protein shows >75% homology to the ALAS1 protein, but the extreme 108 N-terminal residues of ALAS1 is missing from ALAS2. ALAS1 gene expression is also extremely low (2% of total enzyme) in erythroid cells. Nucleotide sequences of the ALAS2 and ALAS1 genes are approximately 60% similar, with the longest stretch of identical sequence being 21 nucleotides. There is very little homology, however, in the N-terminal region, while there is a high homology (~73%) after hepatic residue 197 (Scrivner et al., 1995). The full sequence alignment of the ALAS1 and ALAS2 proteins is shown in appendix 1.

About two-thirds of the sequence shows 60-80% amino acid conservation in protein alignments (Duncan et al., 1997). ALAS enzymes in unicellular and multicellular eukaryotes are closely related to those in the α-purple bacteria, which have been suggested to be the closest contemporary eubacterial relatives of mitochondria (Duncan et al., 1997; Gray, 1992). Sequence data of the ALAS gene has been obtained for many species, ranging from bacteria to mammals, including both forms of the human ALAS genes, rat, mouse, one protostome, five fungi, a protist, and some α-proteobacteria (table 1-5). These data conclude that the gene is highly conserved. The carboxy terminal two-thirds of the unprocessed ALAS shows 40-90% amino acid conservation in protein alignments across all taxa, with bacterial ALAS being the most highly conserved (Duncan et al., 1999). The two isoforms of the ALAS gene are suggested to have evolved from a common ancestral gene, by Duncan et al (1999). This gene duplication event created two genes that evolved to produce enzymes that perform the same function but are under separate regulatory mechanisms. This duplication event is hypothesised to have occurred between the branching of the cephalochordates and the earliest vertebrates (Duncan et al., 1999). Cephalochordates are lancelets, which have been traditionally viewed as the closest living relative of vertebrates (Delsuc et al., 2006).
The ALAS1 gene is highly conserved in species ranging from bacteria to mammals. The sequence encoded by the second to eleventh exons of rat ALAS1 contains essentially common sequence motifs to those of the corresponding exons of the ALAS2 gene (figure 1-7, Yomogida et al., 1993). The only exception to this structural conservation observed, is that the chicken ALAS1 gene lacks a small exon coding for the 5'-UTR, i.e. an exon which would correspond to the first exon of the other ALAS genes (Yomogida et al., 1993). The rat ALAS1 gene consists of 11 exons and spans more than 14kb in the rat genome, where as the chicken ALAS1 gene is 6.9kb and composed of 10 exons. The size of the exons of rat ALAS1 gene is homologous to that of the chicken ALAS1 gene. The second to eleventh exons of the rat ALAS1 gene correspond to the first to tenth exons of the chicken ALAS1 gene. The structure of the rat ALAS1 gene is also similar to that of the mouse ALAS2 gene which consists of 11 exons. In particular, the boundaries of the fifth to eleventh exons are highly conserved between the rat ALAS1 gene and the mouse ALAS2 gene, signifying that ALAS1 and ALAS2 genes evolved from a common ancestral gene (Yomogida et al., 1993).

Figure 1-7 Organisation of ALAS1 and ALAS2 genes is highly conserved in animals, figure modified from Yomogida et al., 1993.
The organization of the chicken ALAS1 gene (Maguire et al., 1986), mouse ALAS2 gene (Schoenhaut et al., 1989), and rat and human ALAS1 gene is highly conserved (Yomogida et al., 1993). ALAS1 genes are more compact than the mouse ALAS2 gene. All the genes, except that encoding chicken ALAS1, are composed of 11 exons.
In both chicken and rat, ALAS activity and mRNA is induced in the liver for the synthesis of drug-metabolising CYPs. It can also be induced, to a lesser extent, in the kidney but not in rat heart, brain, skeletal muscle, testis or spleen (May et al., 1995; Maguire et al., 1986; Srivastava et al., 1988). Analysis of the promoter sequences and mRNA transcripts of these genes has identified a single transcription start site about 25 nucleotides (nt) downstream from a TATA box that was used exclusively for basal transcription in all tissues and for induced transcription in the liver (Yomigida et al., 1993, Maguire et al., 1986; Srivastava et al., 1988).

<table>
<thead>
<tr>
<th>Species</th>
<th>Common Name</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Deuterostomes</strong></td>
<td></td>
</tr>
<tr>
<td>(both ALAS forms)</td>
<td></td>
</tr>
<tr>
<td><em>Homo sapiens</em></td>
<td>Human</td>
</tr>
<tr>
<td><em>Rattus norvegicus</em></td>
<td>Rat</td>
</tr>
<tr>
<td><em>Mus musculus</em></td>
<td>Mouse</td>
</tr>
<tr>
<td><em>Gallus gallus</em></td>
<td>Chicken</td>
</tr>
<tr>
<td><em>Opsanus tau</em></td>
<td>Toadfish</td>
</tr>
<tr>
<td><em>Delphinapterus leucas</em></td>
<td>Beluga whale</td>
</tr>
<tr>
<td><strong>Deuterostomes</strong></td>
<td></td>
</tr>
<tr>
<td>(single form)</td>
<td></td>
</tr>
<tr>
<td><em>Myxine glutinosa</em></td>
<td>Hagfish</td>
</tr>
<tr>
<td><em>Strongylocentrotus</em></td>
<td>Sea urchin</td>
</tr>
<tr>
<td><strong>Protostomes</strong></td>
<td></td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>Fruit fly</td>
</tr>
<tr>
<td><em>Glycera dibranchiata</em></td>
<td>Blood worm</td>
</tr>
<tr>
<td><em>Limulus polyphemus</em></td>
<td>Horseshoe crab</td>
</tr>
<tr>
<td><em>Sepia officinalis</em></td>
<td>Cuttlefish</td>
</tr>
<tr>
<td><strong>Fungi</strong></td>
<td></td>
</tr>
<tr>
<td><em>Agaricus bisporus</em></td>
<td>Mushroom</td>
</tr>
<tr>
<td><em>Schizosaccharomyces pombe</em></td>
<td></td>
</tr>
<tr>
<td><em>Emericells nidulans</em></td>
<td></td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td>Yeast</td>
</tr>
<tr>
<td><em>Kluyveromyces lactis</em></td>
<td></td>
</tr>
<tr>
<td><strong>Protists</strong></td>
<td></td>
</tr>
<tr>
<td><em>Plasmodium falciparum</em></td>
<td></td>
</tr>
<tr>
<td><strong>α-proteobacteria</strong></td>
<td></td>
</tr>
<tr>
<td><em>Rhodobacter sphaeroides</em></td>
<td>(2 forms, named A and T)</td>
</tr>
<tr>
<td><em>Bradyrhizobium japonicum</em></td>
<td></td>
</tr>
<tr>
<td><em>Agrobacterium radiobacter</em></td>
<td></td>
</tr>
<tr>
<td><em>Rhodobacter capsulatus</em></td>
<td></td>
</tr>
<tr>
<td><em>Paracoccus denitrificans</em></td>
<td></td>
</tr>
</tbody>
</table>

Table 1-5 ALAS sequences currently available, adapted from Duncan et al., 1999.

1.7.1 The Human ALAS1 Gene

The organisation of the human ALAS1 gene is similar to the orthologous rat, mouse and chicken genes, except that it contains two non-coding regions (1A and
1B) in its 5'-UTR, whereas the rat gene only has one, and the chicken has none (Yomogida et al., 1993; Maguire et al., 1986). The full sequence alignment between species of the ALAS1 gene is shown in appendix 2. All mammalian ALAS1 genes are composed of 11 exons and are homologous to each other.

![Diagram](image)

**Figure 1.8 Alternative splicing of the human ALAS1 mRNA.**

Alternative splicing of human ALAS1 exon 1B produces two isoforms, a major transcript comprising 90% total cellular ALAS1 mRNA and a minor transcript comprising of only 10% total cellular ALAS1 mRNA (Roberts & Elder, 2001).

Exon 1B of the human gene is alternatively spliced resulting in two mRNA species: a major one in which it is omitted and a minor one containing both exons 1A and 1B. This longer transcript is present in all human tissues and cell lines examined and therefore this alternative splicing does not appear to be specific to
particular cell types. Both ALAS1 transcripts have been detected in mRNA from human brain and liver and from K562, HeLa, HepG2 and B-lymphoblastoid cell lines with the percentage of minor transcript ranging from 7-15% of the total steady state concentration of ALAS1 mRNA (Roberts & Elder, 2001).

1.7.2 The Human ALAS2 Gene

Similar to ALAS1, ALAS2 is also alternatively spliced to produce two mRNA species. However, it is exon 4 which is alternatively spliced in ALAS2, with the larger product being 111bp longer. The deduced ALAS2 protein isoforms are identical except for the presence or absence of 37 amino acids, with structural heterogeneity in the N-terminal region. Preproerythroblasts, reticulocytes, adult bone marrow and three erythroid cell lines K562, HEL and KU812 all express the ALAS2 mRNA isoforms in relatively constant proportions. The strict transcriptional regulation of ALAS2 has been shown, as its mRNA is not present in any non-erythroid cells tested, such as HepG2 cells and B and T lymphocytes (Conboy et al., 1992). The alternative splicing of exon 4 in ALAS2 has only been shown in humans. PCR experiments on murine and canine erythroid cells have indicated that only the larger ALAS2 mRNA containing exon 4 is present, therefore this must be the essential form of ALAS2. The molecular explanation for this differential splicing in humans is unclear (Conboy et al., 1992).

Studies by Schoenhaut & Curtis (1989) have shown that there are two mouse ALAS2 species which differ by the insertion or deletion of 45 nucleotides in exon 3, produced by the utilization of alternate acceptor sites, with the major upstream site contributing 85% of the total ALAS2 cellular mRNA (Schoenhaut & Curtis, 1989). In human ALAS2 mRNA, the single base change of A to G inactivates the potential splice site, leaving only the upstream AG as a splice acceptor (Conboy et al., 1992). Both exons 3 and 4 are in region II of the ALAS2 enzyme. As these show species-specific splicing, there must be a specific role for these exons in the regulation, rather than enzymatic activity, of ALAS2.

In 2004, it was demonstrated that another three additional splice isoforms of ALAS2 exist in human erythroid cells. The principal alternatively spliced mRNA is that which lacks exon 4 encoded sequence, constituting around 35-45% of total ALAS2 mRNA in erythroid cells. However, the additional three human mRNA splice forms reported together account for approximately 15% of the total human ALAS2 mRNA in reticulocytes. One of these is characterized by the addition of
exon 1b. The other two mRNAs do not contain exon 2, which contains the normal translation initiation codon. However, it seems unlikely that these two ALAS2 mRNAs which lack exon 2 play a major role in ALAS2 activity in erythroid cells, as they do not contain the sequence necessary for direct translocation to the mitochondria (Cox et al., 2004).

1.8 Transcriptional Control of ALAS

1.8.1 Regulation of ALAS2

The activity of ALAS in erythroid cells is controlled mostly by the availability of iron in the cell, which regulates the interaction between iron-binding proteins and an iron responsive element in the 5'-UTR of ALAS2 mRNA to control its translation (Ponka, 1997; Hentze & Kuhn, 1996). Iron-responsive elements (IREs) are identified by the presence of a stem-loop structure in which the loop contains the sequence CAGUGH (H=A/C/U not G) in the 5-untranslated region of an ALAS nucleotide sequence. They are present only in the ALAS2 gene forms and not the ALAS1 gene (Dandekar et al., 1991; Thiel, 1994; Melefs et al., 1993).

Iron regulatory proteins (IRPs) are the regulators of iron metabolism, as this essential element used as a redox centre for proteins has to be tightly regulated in cells due to its toxicity to cells when it reacts with molecular oxygen. These IRPs regulate iron by controlling protein expression of those involved in iron metabolism by binding to the IREs of the mRNA under iron-depleted conditions. To date, two IRPs have been identified, IRP1 and IRP2, which differ in their regulatory mechanism (Ishikawa et al., 2005; Hentze et al., 2004).

1.8.2 Regulation of ALAS1

The liver ALAS1 gene is under multi-component control at the transcriptional level. Its transcription is stimulated by cAMP and respiratory uncoupling, and repressed by phorbol esters and insulin. This pattern of regulation is accomplished by CREB, nuclear respiratory factor-1, and AP-1 transactivation through the cis-acting elements CRE, nuclear respiratory factor-1 binding site, and 12-O-tetradecanoylphorbol-13-acetate-responsive element, respectively, found in the 5′—regulatory region of ALAS gene (Scassa et al., 2004).
The promoter of the ALAS1 gene contains a TATA box and 2 control elements located immediately upstream of the TATA box that are homologous to the binding site for the transcription factor nuclear respiratory factor-1 (NRF-1), which seems to be involved in the expression of some proteins involved in oxidative phosphorylation (Ponka, 1999).

1.8.2.1 Insulin

Insulin regulates many metabolic responses in a variety of mammalian tissues, principally liver, muscles, and adipose (Scassa et al., 1998). No unique consensus insulin-responsive sequence or element (Ins RE) has been reported. However, an Ins RE with a T(G/A)TTT(T/G)(G/T) core sequence has been associated with insulin-induced transcriptional repression of a number of metabolic genes, including those that encode phosphoenolpyruvate carboxykinase (PEPCK), insulin-like growth factor-binding protein 1, tyrosine aminotransferase (Scassa et al., 2004). Insulin has been reported to inhibit the activity of ALAS1 in rat and chicken hepatocytes cultures. It has been demonstrated that ALAS1 gene expression is stimulated by cAMP and inhibited by phorbol esters. The inhibitory effect of phorbol esters and diacylglycerol is mediated by the activation of protein kinase C (PKC). Scassa et al (1998) have provided evidence that insulin negatively regulates transcription of ALAS gene. The 870bp promoter region of ALAS1 sustains the inhibitory effect of insulin in HepG2 cells, and the signal transduction pathway involved is mediated by PKC activation. Treatment of rat hepatic cells with 5nM insulin decreases the amount of ALAS mRNA by five-fold in basal hepatocytes, and by nine-fold in cells exposed to phenobarbital, a known inducer of ALAS1 gene expression. However, it was found that the alterations in the ALAS1 mRNA levels following insulin and/or phenobarbital treatments were not due to changes in its stability and this insulin effect does not require protein synthesis (Scassa et al., 1998).

1.8.2.2 Proliferator-activated receptor γ coactivator 1α

Proliferator-activated receptor γ co-activator 1α (PGC-1α) is a co-activator of nuclear receptors and other transcription factors. PGC-1α controls mitochondrial biogenesis and oxidative metabolism in many tissues, including brown adipose tissue, skeletal muscle, heart, and liver. In the liver PGC-1α is induced during fasting, when the liver ceases using glucose as an energy supply and changes to the β-oxidation of fatty acids. This increase in fatty acid β-oxidation and elevation
of hepatic gluconeogenesis are both under control of PGC-1α. Therefore PGC-1α has a key role in liver energy homeostasis and many PGC-1α targets are haem proteins. PGC-1α is an important factor controlling the expression of ALAS1 in the fasted and fed liver. Hepatic PGC-1α is a major determinant of the severity of acute porphyric attacks in mouse models of chemical porphyria (Handschin et al., 2005).

Fasting can be a powerful stimulus to induce an acute porphyric attack and the liver is central to the fasting response in mammals. Experiments where Fao rat hepatoma cells, mouse primary hepatocytes, and rat liver cells were infected with adenovirus expressing PGC-1α, ectopic expression of PGC-1α increased ALAS1 transcript levels in a manner similar to that of glucose-6-phosphatase (Glc6P), a PGC-1α target gene involved in gluconeogenesis. In contrast to ALAS1, none of the other seven genes of the haem biosynthetic pathway were induced by PGC-1α in the liver (Handschin et al., 2005).

**Figure 1-9 Structure of the ALAS1 promoter, adapted from Handschin et al., 2005.**

PGC-1α activates the ALAS1 promoter by coactivating NRF-1 and FOXO1, both of which directly bind to the ALAS1 promoter. Insulin has been shown to activate the protein kinase Akt in the liver, and Akt in turn phosphorylates FOXO1 (Brunet et al., 1999; Nakae et al., 2001). Phosphorylation of FOXO1 results in disruption of its binding to PGC-1α and its export from the nucleus (Puigserver et al., 2003), thus inhibiting PGC-1α action.

HN3FB = hepatocyte nuclear factor 3β; Ins RE = Insulin responsive element; NRF-1 = nuclear respiratory factor 1.

Regulation of ALAS1 by fasting and feeding is mediated by the counter-regulatory hormones insulin and glucagons, which involves PGC-1α. The induction of PGC-1α has been shown to be a consequence of glucagons action and the transcription factor cAMP element binding protein (CREB), which binds directly to the PGC-1α
promoter (Herzig et al., 2001). In addition, CREB can also directly activate the ALAS1 promoter (Varone et al., 1999).

PGC-1α activates the ALAS1 promoter by coactivating NRF-1 and FOXO1, both of which directly bind to the ALAS1 promoter. It has been established that ALAS1 transcription is inhibited by the insulin pathway, involving Akt. Insulin activates the protein kinase Akt in the liver, and Akt in turn phosphorylates FOXO1. This results in disruption of its binding to PGC-1α and its export from the nucleus (Puigserver et al., 2003), thus inhibiting PGC-1α action (Handschin et al., 2005).

The ability of PGC-1α to positively regulate the ALAS1 gene and the requirement for PGC-1α in the fasting induction of ALAS-1 together provide a direct explanation for how fasting provokes an acute attack in an individual with hepatic porphyria (Handschin et al., 2005).

1.8.2.3 Nuclear respiratory factor-1

There are two binding sites for the nuclear respiratory factor-1 (NRF-1) in the ALAS1 promoter (Braidotti et al., 1993). NRF-1 is a transcription factor that increases expression of nuclear-encoded mitochondrial genes and is known to be activated by PGC-1 (Scassa et al., 2001). NRF-1 appears to have a coordinating role during mitochondrial biogenesis, and functional binding sites for this factor have been identified in the rat and *D. melanogaster* genes (Li et al., 1999; Braidotti et al., 1993; Ruiz de Mena et al., 1999) Braidotti et al (1993) established that the NRF-1 motifs in the rat ALAS1 promoter are critical for promoter activity. They used site-directed mutagenesis of the NRF-1 sites, and fused the 5' region of the ALAS1 gene to the chloramphenicol acetyltransferase (CAT) gene and measured relative expression levels. Mutagenesis of each NRF-1 motif gave substantially lowered levels of CAT expression, and mutagenesis of both NRF-1 motifs resulted in an almost complete loss of expression. Gene activation of the exon 1 in the rat ALAS1 is dependent on the presence of at least one functional NRF-1 binding site. It has been suggested that NRF-1 may coordinate the supply of mitochondrial haem with the synthesis of respiratory cytochromes by regulating ALAS1 expression. Further evidence to support this hypothesis is that ALAS2 lacks any NRF-1 binding sites in its promoter (Braidotti et al., 1993). Increased energy consumption increases expression of the NRF-1 gene, which therefore increases ALAS1 expression by interacting with the NRF-1 binding site (Li et al., 1999).
1.8.2.4 Xenobiotic-sensing Nuclear Receptor

The rate of haem biosynthesis responds to increased demands. For example, for the synthesis of drug-metabolising CYPs, the level of ALAS1 mRNA is quickly up-regulated to provide sufficient haem to the nascent apocytochromes. This increase is not a consequence of haem feedback regulation, but a direct activation of ALAS1 transcription by the drug (Jover et al., 1996; Srivastava et al., 1989). The induction of CYPs by drugs involves parallel increases in the transcription of ALAS1 and CYP genes, with subsequent de-repression of haem-mediated inhibition of synthesis of ALAS1 (Jover et al., 1996; Louis et al., 1998).

Fraser et al (2002) have characterised two drug-responsive elements isolated from the 5′-flanking region of the chicken ALAS1 gene in the hepatoma cell line LMH, of 176 and 167 base pairs in length. These regions respond to a wide range of drugs and are referred to as ALAS drug-responsive enhancer sequence (ADRES) elements (Fraser et al., 2002). These enhancer sequences respond to a wide variety of inducer compounds, and closely reflect the induction of ALAS1 mRNA.

Two distal 5′ regulatory elements in the human ALAS1 gene that mediate the drug response were identified by Podvinec et al (2004). The nuclear receptors (NR): human CAR (constitutive androstane receptor) and PXR (pregnane X receptor) can bind to DR4-type hexamer repeats found in both the ALAS1 regulatory elements and mediate direct transcriptional activation of ALAS1 mRNA in response to drugs. The two elements exhibit a high similarity in their responses to drugs even though there is not extensive similarity at the sequence level (Podvinec et al., 2004).

1.8.4 Circadian Rhythms

Circadian rhythms are an evolutionary conserved property of many biological processes in diverse life forms (Dunlap et al., 1999). In mammals, a master circadian clock resides in the suprachiasmatic nuclei (SCN) of the hypothalamus. In the absence of environmental clues such as light, this clock continues to operate with remarkable precision, stability and persisten with a cycle time that is approximately 24 hours. This internal clock can also be reset to external time cues (Zheng et al., 2001). The first clock mutants were identified in Drosophila, and orthologs of most of these circadian clock genes have been found in
mammals, highlighting a general conservation in the clock mechanism between insects and mammals. In particular, three mammalian Period genes (mPer1, mPer2 and mPer3), two Crytochrome genes (mCry1 and mCry2), and the genes Clock, Bmal1 and CK1ε have been identified (Zheng et al., 2001).

The mRNA transcripts and protein products for the mPer genes are expressed in a circadian manner in the SCN and in peripheral tissues. In vitro studies have shown that expression of mPers is driven by the CLOCK/BMAL1 transcription complex, and it was hypothesised that mPers, together with mCrys, serve to regulate the CLOCK/BMAL1 transcription complex negatively (Dunlap et al., 1999).

One of the numerous physiological processes that the circadian clock controls in mammals is haem biosynthesis, through the regulation of ALAS1. Several members of the core clock mechanisms are PAS domain proteins, one of which, neuronal PAS2 (NPAS2) has a haem-binding motif (Koudo et al., 2005). Also, haem controls activity of the BMAL1-NPAS2 transcription complex in vitro by inhibiting DNA binding in response to CO (Gilles-Gonzalez et al., 2004 & 2005).

To investigate the role of haem in the circadian clock mechanism, Kaasik & Lee (2004) injected mice intraperitoneally (i.p.) with haem or solvent during subjective day and night to determine the effect of haem on clock gene expression in vivo. They found that haem but not solvent reduced the expression of mPer2, but enhanced that of mPer1. These effects were most prominent during subjective night. Haem differentially modulates expression of the mammalian genes mPer1 and mPer2 in vivo by a mechanism involving NPAS2 and mPer2 (Kaasik & Lee, 2004).

Loss of both mPER1 and mPER2 function leads to deregulated Alas1 expression1. The BMAL1-NPAS2 transcription complex is a direct and the principal transcriptional regulator of the ALAS1 promoter in vivo. Kaasik & Lee (2004) used the ALAS1 gene cloned into a luciferase reporter vector and compared expression levels after incubation with BMAL1-NPAS2, with and without mPer2 and these proteins individually. A 3.5-fold stimulation of promoter activity was observed when NPAS2, BMAL1 and mPER2 were added together. Conversely, mPER2, BMAL1 or NPAS2 added individually did not stimulate Alas1-Luciferase activity (Kaasik & Lee, 2004).
These findings demonstrate a reciprocal regulation between circadian clock regulation of haem biosynthesis and the haem control of circadian clock-complex-regulated transcription. The expression of ALAS1 is dependent on BMAL1-NPAS2 transcription activity under positive control by mPer2. Furthermore, this demonstrates the only way in which haem indirectly controls transcription of the ALAS1 gene, and thereby controlling its own synthesis.

1.9 Post-transcriptional Regulation of ALAS1 Expression by Haem

The molecular mechanism by which haem regulates ALAS1 mRNA levels remains poorly characterized and, in particular, the down-regulation of ALAS1 transcription by haem has remained controversial (Srivastava et al., 1988 & 1990; Hamilton et al., 1991). It was thought that in mammals, but not in birds (Hamilton et al., 1991), haem may decrease transcription (May et al., 1995; Srivastava et al., 1998) although the mechanism of this effect is unknown. However, in 2005, Kolluri et al showed that haem actually represses transcription of the ALAS1 gene in the chicken hepatoma cell line LMH (leghorn male hepatoma). Furthermore, they identified separate haem and drug-responsive regions in the chicken ALAS1 promoter (Kolluri et al., 2005). Direct interaction of haem with an ALAS1 promoter has only been demonstrated in D. melanogaster (Ruiz de Mena et al., 2004).

In the liver, haem represses the expression of ALAS1 through multiple feedback mechanisms at the post-transcriptional levels. In contrast, the level of ALAS2 activity in erythroid cells is not directly related to the intracellular haem concentration. Two mechanisms have been proposed by which haem may regulate the expression of ALAS1:

1. Prevention of the transfer of the newly synthesised enzyme from cytosol to mitochondria
2. End product inhibition of newly synthesised ALAS activity.

Ryan & Ades (1991) have shown that an increase in ALAS1 mRNA due to the addition of translational inhibitors, cycloheximide or ricin, is attributed to the inhibition of protein synthesis. The cycloheximide significantly decreases the rate of degradation of the ALAS1 mRNA. Therefore it has been suggested that a labile protein is involved in the haem-stimulated degradation of ALAS1 mRNA (Hamilton et al., 1991).
1.9.1 Regulatory Haem Pool

Work on adult rat hepatocytes in primary culture provided direct evidence that a free pool of newly synthesised cytosolic haem appears to be central to formation of haem proteins in the liver and is the source of early labelled pigment (figure 1-10). Newly synthesised haem exists transiently in a cytosolic compartment before its transfer to membrane-bound or cytosolic acceptors or its degradation to bilirubin (Grandchamp et al., 1981). Under normal conditions, this regulatory intracellular free haem pool has been estimated to be at concentration of 10-100nM, by examining indirect evidence (Badawy, 1978).

![Diagram of haem metabolism]

**Figure 1-10 Effects of the regulatory haem pool on hepatic haem metabolism, adapted from Kolluri et al., 2005.**

Haem exerts regulatory effects on both ALAS1 and HO1 to control both its synthesis and degradation. ALAS1 catalyses the formation of ALA, from glycine and succinyl CoA. Haem is synthesized from ALA through a series of intermediary steps catalysed by the enzymes of the haem biosynthetic pathway. Several steps in the production of functional mature ALAS1 protein are down-regulated by exogenous or endogenous haem, as indicated by the minus signs, or by the plus sign indicating increased degradation of ALAS1 mRNA (Kolluri et al., 2005).
This level of haem is maintained by regulating the amount of haem synthesised within the cell thereby balancing the synthesis and degradation of haem within the cell. Cable et al (2000) found that when haem levels within the cell increase; the regulatory haem pool maintains cellular haem levels by first repressing ALAS1, in CWSV17 cells. When the haem levels in this regulatory haem pool are dramatically increased, HO is induced. Therefore there are two mechanisms of reducing the haem concentration, where the more sensitive of the two regulates ALAS1 levels (Cable et al., 2000).

1.9.2 Haem Response Motifs

Haem regulates ALAS1 post-transcriptionally by inhibiting the translocation of newly synthesised ALAS1 enzyme from the cytosol to the mitochondrion, with the use of haem response motifs (HRM)s (Srivastava et al., 1983). Murine ALAS is synthesised as an approximately 70kD cytosolic precursor (Pre-ALAS) that is post-translationally transported into the mitochondrial matrix, where it is proteolytically processed to form a 65kD mature enzyme. High haemin concentrations can inhibit the transport of pre-ALAS1 into mitochondria both in vitro and in vivo (Lathrop & Timko, 1995).

![Diagram of ALAS1 and ALAS2 presequence and mature enzyme](image)

**Figure 1-11** – HRM’s in human ALAS1 and ALAS2, with the amino acid sequence of each. Diagram adapted from Munakata et al., 2004 and Duncan et al., 1999. Numbers indicate the number of amino acids in each full protein (presequence + mature enzyme).
Both ALAS1 and ALAS2 contain three HRMs, which all contain invariant cysteine and proline residues. Other proteins that contain HRM sequences include haem lyase (which attaches haem to cytochromes), HO-2 and catalase (which degrades hydrogen peroxide). The first two enzymes bind haem as a substrate, and the third uses haem as a cofactor (Zhang & Guarente, 1995).

Generally, the HRM’s have been identified in the ALAS sequences in all the species listed in table 1-5, by the amino acid motif R/K/N-C-P-K/hydrophobic residue-L/M, although some minor alterations can be made to this motif (Lathrop & Timko, 1993; Duncan et al., 1999). As HRMs are present in almost all eukaryotic ALAS species, a functional role in the regulation of ALAS has been suggested. The motif is normally present three times near the N-terminus of the ALAS coding sequence, although the known fungal ALAS sequences and the ones for fruit fly and P.falciparum appear to be different from this pattern (Duncan et al., 1999).

The HRM motif consists of five amino acids, among which the second cysteine and third proline are completely conserved. Therefore HRM2 differs from the others by its first, fourth and fifth amino acids. The first is neutral in HRM2, but basic in the HRMs 1 and 3. The fourth and fifth amino acids are phenylalanine and leucine in HRM1 and 3, but lysine and methionine in HRM2 (Munakata et al., 2004). Further experiments are required to assess whether this sequence difference in amino acids reflects the HRM activity. Spectral and chromatographic studies by Zhang & Guarente (1995) have shown that peptides containing a HRM bind to haem and change its physical and chemical properties. The cysteine of the HRM serves the critical function of donating electrons to the haem iron. The basic residues which are usually found at the first position and the hydrophobic residues found in the fourth position may enhance binding by interacting with the carboxyl and vinyl side chains of haem, respectively. The proline in the HRM may serve a structural role by exposing the cysteine residue for bonding to the haem iron (Zhang & Guarente, 1995).

Two of these HRM’s are in region I and one is in region III. All three of these HRM’s in ALAS1 can function in situ as haem-responsive regulators of intracellular localisation. In the case of the mammalian ALAS1, haem probably binds to the HRMs in the leader when the protein is still in the cytoplasm, thus preventing unfolding and translocation across the mitochondrial membrane (Lathrop & Timko, 1993; Zhang & Guarente, 1995).
To study the ALAS2 HRM's in vitro, Lathrop & Timko (1993) used chimeric proteins transfected into mouse erythroid cells. These proteins consisted of the presequence and a portion of the mature peptide (encompassing the region containing the 15 amino acid deletion in the pre-ALAS2 minor form) linked to the mature ornithine transcarbamoylase precursor (OTC), a normally upregulated mitochondrial protein. Transport of preALAS2 major-OTC (which contains three copies of the HRM) and preALAS2 minor-OTC (which contains two HRMs) was inhibited 89% and 87%, respectively, by 25 μM haemin, an inhibition similar to that observed for the native preALAS2 proteins. The conserved cysteine within the HRM was shown by site-directed mutagenesis to be required for haemin inhibition (Lathrop & Timko, 1993).

Cysteines in the HRM’s are essential for their activity, as haem has been shown to bind in a stoichiometric manner to cysteines in synthetic peptides containing an HRM sequence (Zhang & Guarente, 1995). More recently, Munakata et al. (2004) have found that HRM1 is involved in the haem-mediated inhibition of the mitochondrial import of rat ALAS1, whereas HRM2 is not. Also, HRM3 contributes to the haem-mediated inhibition of ALAS1 transport, and seems to function together with HRM1. In vivo haem does not affect the mitochondrial import of ALAS2 (which is not affected by a wide range of haem concentrations). However, the import of ALAS2 is inhibited by haem using in vitro systems (Munakata et al., 2004; Lathrop & Timko, 1993). This suggests that ALAS2 and ALAS1 are controlled by different haem regulation mechanisms in vivo, and that the HRMs in the two isoforms of ALAS do not perform the same function (Munakata et al., 2004).

There have been controversial results of the affect of haem on the mitochondrial import of rat ALAS2. Munakata et al. (2004) found that a wide range of haem concentrations do not affect the import of ALAS2 in vivo. In contrast, Lathrop & Timko (1993) showed that the mitochondrial import of mouse ALAS2 was inhibited by haemin in vitro. Furthermore, this was mediated by the HRM within the pre-sequence, and only one HRM was required to cause the inhibition. However, it is difficult to compare these two contradictory results as they were performed in different species and Lathrop & Timko used in vitro systems (Lathrop & Timko, 1993; Munakata et al., 2004).
1.9.3 Haem-mediated destabilisation of ALAS1 mRNA

Haem has been shown to destabilise ALAS1 mRNA, experimentally, in a wide variety of cell types. For example, in the human hepatoma cell line, HepG2, the addition of haem decreases the ALAS1 mRNA half-life from 8.8 hours to 3.5 hours (Cable et al., 2001; Sassa et al., 1990). In the rat immortalised cell line, CWSV17, the addition of haem decreases the half-life from 2.5 hours to 1.3 hours, while conversely the addition of succinyl acetone (4,6-dioxoheptanoic acid), an inhibitor of haem biosynthesis, increases the ALAS1 half-life to 5.2 hours. In primary chick hepatocytes, haem decreases the mRNA half-life from 3.5 hours to 1.2 hours (Cable et al., 2000; Hamilton et al., 1991).

Haem-dependent decreases of ALAS1 mRNA levels occur more quickly and at lower concentrations than haem-dependent increases of HO. The ALAS1 mRNA remains at reduced levels for extended periods of time, whilst the increases in HO mRNA are much more transient (Cable et al., 2000).

Hamilton et al. (1991) confirmed that haem acts on ALAS1 post-transcriptionally by decreasing its mRNA stability, by using 2-propyl-2-isopropylacetamide (PIA), which increases the rate of transcription of the ALAS1 gene. The half-life of ALAS1 mRNA is approximately 3.5 hours in both control and PIA-induced primary cultures of chick embryo hepatocytes. Subsequent treatment of PIA-treated cells with either 1 or 10μM haem decreases the half-life of ALAS mRNA to approximately 1.2 hours. Cycloheximide completely blocks the haem-mediated decrease in ALAS1 mRNA half-life. In contrast, cycloheximide alone or in combination with PIA has no effect on the half-life of ALAS1 mRNA. These findings by Hamilton et al. have therefore suggested that there is labile protein that mediates the effect of haem on ALAS1 mRNA stability. Such a labile protein may function as a haem-binding protein to mediate the effect of haem at specific sites of the ALAS1 mRNA, a haem-induced or activated nuclease for ALAS1 mRNA, or a haem carrier within the cells (Hamilton et al., 1991).

1.9.4 Control of Haem Synthesis in Other Cell Types

Control of haem synthesis in cell types other than hepatic and erythroid cells has not been studied in such depth. A regulatory haem pool tightly regulates haem synthesis in the liver, but apparently less so in extra-hepatic tissue (Bissell and Hammaker, 1976; Giger & Meyer, 1983). ALA formation in these tissues is
presumed to be carried out by ALAS1, but regulation of haem biosynthesis in non-hepatic tissues seems different from that in the liver. For example, potent inducers of hepatic ALAS1 do not increase ALAS1 activity in the Harderian gland of mice, in the heart, adrenal gland, testes, brain and spleen in rats; or in cultured human amniotic cells. The effect of haemin on ALAS1 activity in non-hepatic tissues is also quite different from that in the liver. Haemin does not suppress ALAS1 in the Harderian gland, heart, or adrenal gland of mice, or in the testes of rats. ALAS1 in the foetal liver, which is largely erythroid during the foetal period, is also refractory to inhibition by haemin. Changes in ALAS activity in foetal guinea pig liver are correlated with the change in erythropoietic activity in this organ. Thus, regulation of the enzymatic activity of foetal liver reflects, by and large, that of erythroid cells (Scriver et al., 1995).

1.10 Post-transcriptional Regulation of Gene Expression

In many cases of controlling the flow of genetic information from DNA to protein, the regulation of protein production is principally achieved by controlling mRNA levels (Guhaniyogi & Brewer, 2001). In mammalian cells, physiological signals modulate the rate of mRNA turnover (Bevilacqua et al., 2003). Cytoplasmic mRNA levels represent a balance between the rates of nuclear RNA synthesis, processing, and export, and the rates of cytoplasmic mRNA degradation. The process of mRNA degradation is a major control point in gene expression, and this tightly regulated process dependent on specific cis-acting sequences and trans-acting factors (Guhaniyogi & Brewer, 2001).

Two major mechanisms of mRNA degradation are active in eukaryotic cells. In the first pathway, deadenylated transcripts are degraded by a complex of 3’-5’ exonucleases known as the exosome. Some mRNAs bearing AREs in their 3’-UTR are degraded by this pathway. The second pathway entails the removal of the seven-methyl guanosine cap from the 5’ end of the transcript by the DCP1-DCP2 complex, allowing 5’-3’ exonucleolytic degradation by XRN1 (reviewed in Meyer et al., 2004; Parker & Song, 2004). In yeast, components of this 5’-3’ decay pathway are concentrated at discrete cytoplasmic foci known as processing bodies (p bodies). Studies of mammalian cells have revealed similar structures that contain DCP1/2, XRN1, GW182, and LSM1-7 heptamer. Little is known about the signaling pathways and specific molecular events that govern p body assembly, although their size and number increase when 5’-3’ mRNA decay is blocked and vary throughout the cell cycle. Stress granules are thought to be
sites of mRNA sorting rather than decay, as p bodies are putative sites of 5'-3’ mRNA degradation (reviewed in Parker & Sheth, 2007).

1.10.1 Regulated mRNA Stability

The regulation of both mRNA stability and translation are essential in the control of gene expression and regulation of these two processes allows a cell to rapidly respond to changes in intracellular and extracellular stimuli (Barreau et al., 2005). The regulation of eukaryotic genes by the stability of mRNA was first suggested in 1976. Ouellette & Malt (1976) discovered that the half-life of some murine mRNAs were prolonged, when compared to the majority of murine mRNAs. Since then a large number of genes have found to be regulated by the stability of the mRNA. The mRNA level is controlled post-transcriptionally mainly by specific proteins binding to a specific cis-acting element. The majority of these mechanisms, especially in the 5'-UTR affect translation rather than stability of the mRNA. The regulation of mRNA stability constitutes a critical control step in cellular mRNA level. Often, regulatory cis-acting elements are found with the coding region of the mRNA, but the mechanism is unknown, for example the regulation of rat CYP2E1 by insulin (Moncion et al., 2002), described in section 1.10.4.

The degradation of mRNA is a highly controlled process in eukaryotic cells and provides a powerful means for controlling gene expression (Chen & Shyu, 1995). The rate of mRNA turnover not only determines the rate of disappearance of mRNA but also its induction. mRNAs with short half-lives respond to changes in transcription more rapidly than those that are relatively stable (Guhaniyogi & Brewer, 2001). Many clinically relevant mRNAs are regulated by differential RNA stability, and the aberrant control of mRNA stability has been implicated in disease states, including cancer, chronic inflammatory responses and coronary disease. The rates at which degradation processes can occur are dictated by both cis-acting elements within an mRNA and trans-acting factors that bind them. Cis-acting elements can regulate the rate of turnover of a transcript, either by promoting (destabiliser elements) or by inhibiting (stabiliser elements) decay (reviewed in Wilusz et al., 2001). The regulated decay of mRNA is achieved by the orchestrated interactions between an mRNA’s structural components and specific trans-acting factors. These components include the 7-methyl-guanylate (m7G) 5’ cap structure, the 5’-untranslated region (5'-UTR), the protein coding region, the 3’-UTR and the 3’ polyadenylate [Poly(A)] tail (Guhaniyogi & Brewer,
Figure 1-12 is a representation of the structure of a typical eukaryotic gene mRNA illustrating the possible types of cis-acting regulatory elements that exist. mRNA degradation can be preceded by shortening or removal of the poly(A) tail at the 3' end and/or by removal of the m7G cap at the 5' end. Degradation can also take place following endonuclease activity, independent of both deadenylation and decapping, e.g. the degradation of transferrin receptor (TFR) mRNA in response to iron involves an endonucleolytic rupture in the 3'-UTR (Pesole et al., 2000).

![Diagram of mRNA structure]

**Figure 1-12** The general structure of a eukaryotic mRNA, illustrating examples of cis-acting elements involved in mRNA stability mechanisms.

The vast majority of pre-mRNA molecules undergo several steps of processing before they become functional mRNAs. The introns are removed, a m7G cap structure is added at the 5' end of the first exon, and a stretch of 100-250 adenine residues is added at the 3' end of the last exon. The resultant mature mRNA has a tripartite structure consisting of a 5'-UTR, a coding region made up of triplet codons, and a 3'-UTR (Mignone et al., 2002). The 5' cap structure and 3' poly(A) tail protect the message from exonucleolytic decay and can also act to regulate translation (Wilusz et al., 2001).
1.10.2 The 3’ Untranslated Region

The 3’-UTR has emerged as having a major role in harbouring determinants that control mRNA decay. One of the most prevalent mRNA sequence elements involved in regulating mRNA stability is the adenylate and uridylate (AU)-rich element (ARE), found in the 3’-UTR of numerous mRNAs encoding cytokines, proto-oncogenes and growth factors such as c-fos, c-myc and interferon-α (Wilusz et al., 2001), which consequently have very short half-lives (Mitchell & Tollervey, 2000).

An AU-rich consensus sequence was first observed in the 3’-UTR of murine and human tumour necrosis factor (TNF) mRNAs by its ability to target the mRNA towards rapid degradation (Bevilacqua et al., 2003). AREs are sequence elements of approximately 50-150 nucleotides, and contain multiple copies of the pentamer AUUUA. AREs enhance rapid deadenylation rates and trigger the cleavage of the 5’ cap, leading to mRNA degradation (reviewed in Chen & Shen, 1995). In resting or unstimulated cells, the ARE-dependent degradation mechanism ensures a very low level of expression of these potent proteins. The importance of this repression in the resting state is testified by the observation that pathological states (cancers, chronic inflammations and auto-immune pathologies) are associated with a de-regulation of the stability of ARE-containing mRNAs. AU rich sequences are by far the most widespread and efficient determinant of RNA stability among those characterised in mammalian cells. It has been estimated that 5-8% of human genes code for ARE-containing mRNAs and the corresponding proteins perform a variety of functions implicated in numerous transient processes (Barreau et al., 2006).

Other 3’-UTR instability elements include the IREs present in the Tfr mRNA. When intracellular levels of iron are low, the IRE binds iron regulatory proteins (IRPs), which prevent endonucleolytic rupture and trigger mRNA degradation (Pesolet al., 2000). Histone mRNAs lack poly(A) tails but their cell-cycle-dependent decay is mediated by a 3’ terminal stem-loop motif (6bp stem and a 4 base loop) (Guhaniyogi & Brewer, 2001).

1.10.3 The 5’ Untranslated Region

The 5’-UTR can play significant roles in altering stability either in a translation-dependent manner, by containing translation-inhibiting stem-loops, or in a
manner independent of translation. An example of the latter is the stabilisation of the otherwise unstable c-myc mRNA by reciprocal translocations of immunoglobulin introns in the 5'-UTR of c-myc mRNA (Guhaniyogi & Brewer, 2001). Upstream initiation codons and open reading frames (ORFs) can cause mRNA decay through the nonsense-mediated mRNA decay (NMD) pathway. Upstream ORFs can also regulate mRNA stability through an NMD-independent mechanism by inhibiting ribosomal scanning, thereby promoting mRNA decay (Mignone et al., 2002). An example of a regulatory ORF occurs in tropoelastin mRNA that encodes the connective tissue protein, elastin, where a specific element within the ORF regulates elastin production in response to transforming growth factor β1 (TGF-β1) (Zhang et al., 1999).

IRE's can also be found in the 5'-UTR of mRNA molecules such as the Alzheimer’s Amyloid precursor protein (APP) transcript, ferritin heavy (H) and light (L) chain, ALAS2, mammalian mitochondrial aconitate (mt-acon) and D.melanogaster succinate dehydrogenase subunit b, which mediate the mRNA translational efficiency (Thomson et al., 1999; Rogers et al., 2002).

1.10.4 The Coding Region

Instability determinants have also been found within the coding regions of mRNA species, although these are less common than those in the UTR’s. The c-myc proto-oncogene contains two instability determinants, an ARE in its 3'-UTR and a 249-nucleotide coding region determinant (CRD) (Bernstein et al., 1992; Doyle et al., 1998). The CRD functions independently of the ARE to make the mRNA unstable (Herrick & Ross, 1994). Other mRNA species with CRDs include the plasminogen activator inhibitor type 2 (PAI-2) and c-fos (Tierney et al., 2001; Chen et al., 1992). The c-fos proto-oncogene mRNA is one of the least stable mammalian messages and is targeted for rapid degradation by at least two distinct mRNA degradation pathways, one involving an ARE, and the other a CRD. Two cellular proteins have been identified that interact with the c-fos CRD to form specific RNA-protein complexes that participate in the CRD-directed mRNA decay (Chen et al., 1992).

Recent experiments using the c-fos mRNA as a model system have demonstrated a clear role for translation in mammalian RNA turnover. Two destabilising regions within the c-fos protein-coding region, termed CRD-1 and CRD-2, have been identified, and CRD-1 is the major determinant (mCRD). Grosset et al (2000)
have found that a minimal distance of ~450bp spacer between the mCRD and the poly(A) tail is required for mRNA degradation. They revealed a novel role for the 3’ poly(A) tail, a dynamic interaction between the transcribed portion of the message and its poly(A) tail exists to influence the rate of deadenylation, therefore the poly(A) tail plays an active role. They have also provided evidence that the bridging complex contains RNA binding proteins can or have the potential to shuttle between the nucleus and cytoplasm, therefore shuttling RNA binding proteins have nuclear and cytoplasmic functions (Grosset et al., 2000).

Davis et al (2001) have identified a ~280 base element residing within the coding region of the MnSOD mRNA that confers cis-mediated destabilisation in the absence of ribosome transit. This element functions during basal and stimulus-dependent MnSOD gene expression (Davis et al., 2001).

Insulin controls the expression of rat CYP2E1 by altering its mRNA stability (De Waziers et al., 1995). However, the mechanism by which insulin achieves this has not been elucidated. A 16bp protein-binding site in the 5’-proximal region of CYP2E1 coding region has been identified as being involved. Although insulin does not modify the binding of proteins to this region, it has been shown to be involved in the mechanism by insertion of the sequence into a luciferase gene. This was sufficient to render the chimeric mRNA sensitive to insulin. There is indirect evidence that the human CYP2E1 mRNA is regulated by insulin similarly to the rat CYP2E1 mRNA. A hairpin loop is present in both the proposed regulatory regions with 88% sequence identity (Moncion et al., 2002). In addition, CYP2E1 is elevated in lymphocytes from patients with insulin-dependent diabetes mellitus (Hannon-Fletcher et al., 2001). As the binding of the regulatory binding protein to its target RNA in CYP2E1 is not regulated by insulin, it is possible that it instead regulates protein-protein interactions mediating the destabilisation of the mRNA (Moncion et al., 2002).
1.11 Aims of the Thesis

Haem regulates its own synthesis in non-erythroid cells through a regulatory haem pool, which is maintained by distinct haem-dependent control of ALAS1 and HO-1. However, the mechanism by which the expression of ALAS1 is post-transcriptionally regulated by haem has not yet been defined.

The overall objective of this thesis was to investigate the mechanism by which haem destabilises the ALAS1 mRNA in human non-erythroid cells. To elucidate this mechanism there were three main aims of the project:

1. To investigate the response of the alternative ALAS1 transcripts to haem.

2. To determine where the ALAS1 cis-acting element resides in the mRNA sequence.

3. To investigate protein binding to the ALAS1 mRNA, as this may be involved in its destabilisation in response to haem.
CHAPTER 2

Materials and Methods
2.1 Chemicals

All general laboratory chemicals were of at least AnalR grade and purchased from Sigma-Aldrich unless otherwise stated.

The full addresses of all suppliers are listed in Appendix 3.

2.2 Cell culture

2.2.1 Mammalian Cell Culture

2.2.1.1 Cell Lines

The human Caucasian hepatoma cell line, HepG2 (ECACC No. 85011430), the human Caucasian neuroblastoma cell line, IMR32 (ECACC No. 86041809), and the human Caucasian chronic myelogenous leukaemia cell line K562 (ECACC No. 89121407) were obtained from the European Collection of Cell Cultures (ECACC).

2.2.1.2 Media and Growth Conditions

HepG2 cells were grown and maintained in Dulbecco’s modified Eagle 1x medium (DMEM) supplemented with 10% (v/v) foetal calf serum (FCS) and 2mM L-glutamine. K562 and IMR32 cells were both grown in Roswell park memorial institute 1640 1x medium (RPMI 1640), supplemented with 10% and 5% FCS respectively and 2mM L-glutamine. All cells were grown at 37°C with 5% CO₂ in a humidified incubator.

2.2.1.3 Seeding Cells

Exponentially growing cells were split by washing with saline, followed by treatment with trypsin-EDTA (Sigma) for 5 minutes at 37°C. Cells were reseeded at the appropriate density in complete medium. For RNA extraction, cells were split into 6-well plates at a confluency of approximately 60%. For luciferase assays, cells were split into 12-well plates at a confluency of 60%.
2.2.1.3.1 Determining cell number

Mammalian cells in suspension were counted using a Z2 Coulter Particle Counter and Size Analyser (Beckman Coulter). Briefly, 0.5ml of cell suspension was added to 9.5ml of Isoton II diluent (Beckman Coulter) and analysed for particles in the range of 10 to 30µM. Cell counting was performed three times and the average output given as a number of cells per ml of the original cell suspension.

2.2.1.4 Transfection

Transfections of HepG2, K562 and IMR32 cells for RNA analysis were carried out using Lipofectamine 2000 (Invitrogen) in 6 well plates. One day prior to transfection, cells were seeded into 6-well plates in 2000µl of growth media so that they would be 90% confluent at the time of transfection. For each well, 3µg of DNA was mixed with 250µl of serum-free media (DMEM with 2mM L-glutamine). In a separate tube, 10µl of Lipofectamine 2000 was added to 250µl of serum-free media, and incubated for 5 minutes at room temperature. After the incubation, the DNA solution was combined with the Lipofectamine 2000 solution, mixed gently and incubated for 20 minutes at room temperature. The 500µl transfection mix was added drop-wise to the well and mixed gently. The cells were ready to be used in experiments, 24 hours post transfection.

When transfections were used for reporter assays, 12-well plates were used and the above method was scaled down.

2.2.1.5 Freezing Mammalian Cells

For long-term storage, cells were grown to approximately 2-4 x 10⁶ cells/ml in a 75cm² flask, washed with saline, trypsinised and resuspended in 10ml of freezing medium (90% FCS, 10% DMSO). 1ml aliquots of the cells were pipetted into cryotubes, and placed in an isopropanol container at -70°C, so that they were frozen at a cooling rate of between 1-3°C/minute. After 24 hours, the cells were transferred to liquid nitrogen for long-term storage.

2.2.1.6 Haem-Depleted Media

For haem-depletion experiments, haem was removed from the FCS, prior to its addition to DMEM. L-ascorbic acid was added to FCS at a final concentration of
10mM, and incubated at 37°C for approximately 12 hours, until the absorbance at 405nm decreased from 2.0 to 0.8. The medium was dialysed in phosphate-buffered saline (PBS) at 4°C overnight in dialysis tubing (Spectra/Por MWCO=2000). The medium was dialysed for a further 8 hours, changing the PBS every 4 hours. The haem-depleted serum was sterilised by filtration through a 0.2μM filter, and stored at −20°C until further use. FCS was added to the cell growth media, at concentrations as described.

2.2.1.7 Preparation of Cytosolic Protein Extracts

Ice-cold PBS (1.5mls) was added to each T75 flask and the cells removed from the surface using a cell scraper. The cell suspensions were transferred to 1.5ml eppendorf tubes and centrifuged at 13000 rpm for 2 minutes. The cell pellet was resuspended in 200μl of ice-cold cytoplasmic extraction buffer (CEB; 10mM Hepes, pH 7.5, 3mM MgCl₂, 14mM KCl, 5% glycerol, 0.8% Nonidet P-40, 1mM dithiothreitol [DTT]) with 10μl freshly added protease inhibitor cocktail (104mM AEBSF [4-[(Z-Aminoethyl)benzenesulfonyl fluoride hydrochloride], 80μM aprotinin, 4mM bestatin, 1.4mM E-64 [N-(trans-Epoxy succinyl)-L-leucine 4-guanidinobutyramide], 2mM leupeptin, 1.5mM pepstatin A in DMSO [dimethyl sulfoxide] [Sigma]). The tubes were vortexed for 10 seconds and incubated on ice for 20 min, with occasional vortexing. The cell membranes were pelleted by centrifugation for 5 minutes at 8500g. The resultant protein supernatant was removed and stored at −70°C after snap-freezing on dry ice.

2.2.1.8 Dual Luciferase Reporter Assay (DLRA)

To measure the luciferase activity of the transfected HepG2 and IMR32 cells, the cells were washed in saline and lysed in 250μl of 1X Passive Lysis Buffer (PLB) (Promega) with shaking for 15 minutes at room temperature. The assays for firefly luciferase and Renilla luciferase activity were performed sequentially from each well of lysed cells. 50μl of LAR II (Promega) and 50μl of the sample were mixed by vortexing in a small glass vial, and assayed in the Luminometer (Lumat LB9507 Luminometer, EG&G Berthold) to measure the firefly luciferase activity. 50μl of Stop & Glo Reagent (Promega) was then added, vortexed briefly and placed in the Luminometer to measure the Renilla luciferase activity. Each sample was assayed in duplicate, and the mean Renilla luciferase activity was normalised to the mean firefly luciferase activity, to compare the relative expression.
2.2.2 Bacterial Cell Culture

2.2.2.1 Media and Growth Conditions

Bacterial cell culture reagents were purchased from Oxoid and antibiotics from Sigma. All glassware was autoclaved at 126°C for 35 minutes prior to use.

<table>
<thead>
<tr>
<th>Media</th>
<th>Composition</th>
</tr>
</thead>
</table>
| Luria-Bertani (LB) | 1% w/v tryptone  
                      | 0.5% w/v yeast extract  
                      | 0.5% w/v NaCl          |
| LB agar          | 1% w/v tryptone  
                      | 0.5% w/v yeast extract  
                      | 0.5% w/v NaCl  
                      | 1.45% w/v agar         |
| SOC              | 2% w/v tryptone  
                      | 0.5% w/v yeast extract  
                      | 10mM NaCl  
                      | 2.5mM KCl  
                      | 20mM MgCl₂  
                      | 20mM glucose         |
| AX               | LB agar plus:  
                      | 100µg/ml ampicillin  
                      | 40µg/ml X-Gal         |

Table 2-1 Compositions of bacterial growth media

<table>
<thead>
<tr>
<th>Antibiotic</th>
<th>Stock Solution</th>
<th>Working Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>50mg/ml</td>
<td>100µg/ml</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>50mg/ml</td>
<td>50µg/ml</td>
</tr>
</tbody>
</table>

Table 2-2 Concentrations of antibiotics used during bacterial cell growth

2.2.2.2 Transformation

For routine subcloning, Library Efficiency™ Chemically Competent DH5α™ E. coli cells (Invitrogen) were used (expected tranformation efficiency: >1 x 10^8 cfu/µg supercoiled plasmid), of genotype:

F-Φ80lacZΔM15(lacZYA-argF)U169recA1 endA1 hsdR17(rK-,mK+)phoA supE44 thi-1 gyrA96 relA1 λ-.
The DH5α competent cells were thawed on ice before being gently mixed. An aliquot of 50μl was pipetted into a chilled 100mm polypropylene tube on ice. 1-5μl of the ligation reaction (approximately 1-10ng DNA) was added to the cells, and gently tapped to mix. The cells were incubated on ice for 30 minutes, heat-shocked at 42°C for 45 seconds and placed on ice for 2 minutes. 900μl of pre-warmed SOC medium was added to the transformation reaction and incubated at 37°C for 1 hour with shaking at 225rpm. The cells were pelleted by spinning at 6500rpm for 10 seconds, before being resuspended in 100μl of SOC media and spread onto a LB agar plate containing the appropriate concentration of antibiotic, and incubated overnight at 37°C.

2.2.2.3 Small Scale Plasmid Purification

For purification of up to 20μg of high-copy plasmid DNA from E. coli, a QIAprep Spin Miniprep Kit (Qiagen) was used. A single bacterial colony was picked from a LB agar plate, freshly streaked with DH5α cells transformed with plasmid DNA, and used to inoculate 5ml of LB with the appropriate antibiotic, and grown overnight at 37°C with shaking at 240rpm. The overnight culture (1.4ml) was transferred to a 1.5ml microcentrifuge tube and the bacteria pelleted by centrifugation at 13,000 rpm in a MSE Micro Centaur for 5 minutes.

The DNA was extracted from the bacterial cells using the alkaline lysis method, whereby SDS in the lysis buffer breaks down the cell membrane, and the NaOH loosens the cell wall and denatures DNA, releasing the plasmid. The bacterial DNA becomes linearised and the strands are separated, whereby the plasmid DNA is circular and remains topologically constrained. Adding potassium acetate to the DNA then renatures the plasmid DNA, whereas the sheared cellular DNA remains denatured as single stranded DNA (ssDNA) and is precipitated since it is insoluble in high salt.

The bacterial pellet was resuspended in 250μl of buffer P1 before 250μl of buffer P2 was added and the sample mixed thoroughly by inverting the tube 4-6 times. Buffer N3 (350μl) was added to the sample and mixed immediately by inversion, before being centrifuged for 10 minutes at 13,000rpm in a MSE Micro Centaur. The supernatant was transferred to a QIAprep spin column, centrifuged at 13,000rpm for 1 minute and the column washed with 0.5ml buffer PB to remove any trace nuclease activity. The spin column was washed with 0.75ml buffer PE and centrifuged twice at 13,000rpm for 1 minute, to remove residual wash buffer. The QIAprep column was placed in a clean nuclease-free microcentrifuge tube
and 50μl of sterile H₂O was added to the centre of the column. The column was left to stand for 1 minute at room temperature, and the DNA was eluted into a clean 1.5ml microcentrifuge tube by centrifugation at 13,000rpm for 1 minute.

2.2.2.4 Large Scale Plasmid Purification

For purification of up to 750μg of high-copy plasmid DNA from E. coli, a HiSpeed Plasmid Maxi Kit (Qiagen) was used. A single bacterial colony was picked from a freshly streaked selective LB agar plate and used to inoculate 5ml of LB with the appropriate antibiotic, and grown overnight at 37°C with shaking at 240rpm. 500μl of the starter culture was transferred into 250ml LB media plus any required antibiotics, and grown overnight at 37°C with shaking at 240rpm. The bacterial cells were harvested by centrifugation at 6,000g for 15 minutes at 4°C, and the supernatant was discarded.

The bacterial pellet was resuspended in 10ml buffer P1 by vortexing. Buffer P2 (10ml) was added to the bacteria, and mixed thoroughly by vigorously inverting the tube 4-6 times. The mix was then incubated for 5 minutes at room temperature to lyse the cells. Buffer P3 (10ml) was added to the lysate, and mixed immediately and thoroughly by vigorously inverting the tube. The lysate was poured into a QIAfilter Cartridge, and incubated at room temperature for 10 minutes. The cell lysate was filtered into the previously equilibrated HiSpeed Tip (10ml of buffer QT), and allowed to enter the resin by gravity flow. The HiSpeed Maxi Tip was washed with 60ml of buffer QC, and the plasmid DNA was eluted by the addition of 15ml of buffer QF. The DNA was then precipitated by the adding 10.5ml (0.7 volumes) of room temperature isopropanol, and incubated at room temperature for 5 minutes. The eluate/isopropanol mixture was filtered through a QIAprecipitator Maxi Module, and 2ml of 70% ethanol was passed through to wash the bound DNA. The filter was dried by passing air through the QIAprecipitator and the plasmid DNA was finally eluted into a 1.5ml microcentrifuge tube by passing 1ml of buffer TE through.

2.3 Protein Methodology

2.3.1 Estimation of Protein Concentration

Protein concentrations of samples were measured using the BCA™ Protein Assay Kit (Pierce). This is a detergent-compatible formulation based on bicinchoninic
acid (BCA) for the colorimetric detection and quantification of total protein. The method combines the reduction of Cu\(^{2+}\) to Cu\(^{+}\) by protein in an alkaline medium (the biuret reaction) with the highly sensitive colorimetric detection of the cuprous cation (Cu\(^{+}\)) using a reagent containing BCA. The purple-coloured reaction product of this assay is formed by two molecules of BCA chelating with one cuprous ion. This water-soluble complex exhibits a strong absorbance at 562nm that is nearly linear with increasing protein concentrations over a broad range (20-2,000 μg/ml). This method is not a true-end point method as the final colour continues to develop. However, following incubation the rate of continued colour development is sufficiently slow to allow large numbers of samples to be assayed together. Studies have shown that the macromolecular structure of protein, the number of peptide bonds and the presence of four particular amino acids (cysteine, cystine, tryptophan and tyrosine) are responsible for colour formation with BCA by the reduction of Cu\(^{2+}\) (Wiechelman et al, 1988).

The sample protein and the bovine serum albumin (BSA) standards (diluted with the sample buffer) were incubated at 37°C with the BCA working reagent (BCA\(^{TM}\) Reagent A mixed with BCA\(^{TM}\) Reagent B, 50:1) for 30 minutes. The tubes were then cooled to room temperature and the absorbencies measured at 562nm. The concentration of the protein was found with reference to a standard curve of the absorbance of the BSA standards, where a series of dilutions of known concentrations of BSA were prepared. These and the samples of unknown concentration were assayed alongside each other.

### 2.3.2 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

A resolving gel of 0.8mm thickness was used below a stacking gel to separate proteins according to their size. The stacking gel is only slightly acidic (pH 6.8), with a low acrylamide concentration thereby making it a porous gel. Under these conditions the proteins are poorly separated but form thin, sharply defined bands before entering the resolving gel, where the higher pH and acrylamide concentration causes the proteins to be resolved according to size. Samples were mixed with 4x loading buffer (125mM Tris, 20% (v/v) glycerol, 4% (w/v) 2-mercaptoethanol, 0.04% (w/v) bromo-phenol blue) and heated at 95°C for 2 minutes before loading into the gel. Samples were electrophoresed at a constant current in 1x SDS running buffer (250mM glycine pH8.3, 25mM Tris base, 0.1% SDS) alongside 5μl of kaleidoscope prestained protein standards (Bio-Rad), with a molecular weight range of 6.5–200 kdaltons.
Table 2-3 Compositions of stacking and separating gels used in SDS-PAGE

**2.3.3 Coomassie Blue Staining**

SDS-polyacrylamide gels were stained in Coomassie blue staining solution (20% methanol, 5% glycerol, 0.25% Coomassie brilliant blue R-250) for 10 minutes and destained in destaining solution (20% methanol, 10% acetic acid) for at least 30 minutes until the desired degree of background was achieved.

**2.4 DNA Methodology**

**2.4.1 Restriction Enzymes and Buffers**

Restriction enzymes were purchased from either Promega or New England Biolabs (NEB).

**Table 2-4 Enzymes from NEB**

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Restriction Site</th>
<th>Supplied NEB Buffer</th>
<th>Percentage Activity in NEB</th>
<th>Enzyme Assay Temp (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Kas I</td>
<td>G↓GCGCC</td>
<td>2</td>
<td>75</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table 2-5 Composition of NEB restriction enzyme buffers**

<table>
<thead>
<tr>
<th>NEB Buffer</th>
<th>Composition at 1x</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>10mM Tris-HCl, 10mM MgCl₂, 50mM NaCl, 1mM DTT, pH 7.9</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Restriction Site</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Apa I</td>
<td>GGGCC</td>
</tr>
<tr>
<td>EcoRI</td>
<td>G</td>
</tr>
<tr>
<td>Hind III</td>
<td>A</td>
</tr>
<tr>
<td>Kpn I</td>
<td>GGTAC</td>
</tr>
<tr>
<td>Nco I</td>
<td>C</td>
</tr>
<tr>
<td>Not I</td>
<td>GC</td>
</tr>
<tr>
<td>Pst I</td>
<td>C</td>
</tr>
<tr>
<td>Sac I</td>
<td>GAGCT</td>
</tr>
<tr>
<td>Spe I</td>
<td>A</td>
</tr>
<tr>
<td>Xba I</td>
<td>T</td>
</tr>
</tbody>
</table>

Table 2-6 Enzymes from Promega

<table>
<thead>
<tr>
<th>Promega Buffer</th>
<th>Composition at 1x</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6mM Tris-HCl, 6mM MgCl2, 6mM NaCl, 1mM DTT, pH 7.5</td>
</tr>
<tr>
<td>B</td>
<td>6mM Tris-HCl, 6mM MgCl2, 50mM NaCl, 1mM DTT, pH 7.5</td>
</tr>
<tr>
<td>D</td>
<td>6mM Tris-HCl, 6mM MgCl2, 150mM NaCl, 1mM DTT, pH 7.9</td>
</tr>
<tr>
<td>E</td>
<td>6mM Tris-HCl, 6mM MgCl2, 100mM NaCl, 1mM DTT, pH 7.5</td>
</tr>
<tr>
<td>H</td>
<td>90mM Tris-HCl, 10mM MgCl2, 50mM NaCl, pH 7.5</td>
</tr>
<tr>
<td>J</td>
<td>10mM Tris-HCl, 7mM MgCl2, 50mM KCl, 1mM DTT, pH 7.9</td>
</tr>
<tr>
<td>Multi-Core™</td>
<td>25mM Tris-Acetate, 100mM K-Acetate, 10mM Mg-Acetate, 1mM DTT, pH 7.5</td>
</tr>
</tbody>
</table>

Table 2-7 Composition of Promega restriction enzyme buffers

2.4.2 Restriction Digests

2.4.2.1 Single Digests

One unit of restriction enzyme activity is defined as the amount of enzyme required to produce complete digestion of 1μg of substrate DNA in 1 hour, at the appropriate assay temperature.

Generally, restriction digestions using a single restriction enzyme were set up as shown in table 2-8.
<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate DNA</td>
<td>1μg</td>
</tr>
<tr>
<td>Restriction Enzyme (10U/μl)</td>
<td>1μl</td>
</tr>
<tr>
<td>Restriction buffer 10x</td>
<td>1μl</td>
</tr>
<tr>
<td>H₂O</td>
<td>to 10μl final volume</td>
</tr>
</tbody>
</table>

**Table 2-8 Single restriction digest reaction**

The reactions were incubated at the appropriate temperature (table 2.4 & 2.5) for at least 3 hours.

**2.4.2.2 Double Digests**

When DNA needed to be digested by two restriction enzymes, the buffer that allows the highest percentage of activity for both enzymes was used (table 2-4 & 2.5)

Generally, restriction digestions using two restriction enzymes were set up as follows:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate DNA</td>
<td>1μg</td>
</tr>
<tr>
<td>Restriction Enzyme 1 (10U/μl)</td>
<td>1μl</td>
</tr>
<tr>
<td>Restriction Enzyme 2 (10U/μl)</td>
<td>1μl</td>
</tr>
<tr>
<td>Restriction buffer 10x</td>
<td>1μl</td>
</tr>
<tr>
<td>H₂O</td>
<td>to 20μl final volume</td>
</tr>
</tbody>
</table>

**Table 2-9 Double restriction enzyme digest**

The reactions were incubated at the appropriate temperature (tables 2-4 & 2-5) for at least 3 hours.

**2.4.3 DNA Quantification**

DNA was quantified using the GeneQuant DNA/RNA Calculator (Amersham Pharmacia Biotech), which measures absorbance at 260nm and 280nm. 3μl of the DNA sample to be quantified was added to 147μl of water in a quartz cuvette with a path length of 10mm. The GeneQuant was set to interpret either dsDNA or ssDNA as appropriate. The cuvette was placed in the spectrophotometer, and the
concentration of the sample recorded. The purity of the DNA was confirmed by checking the ratio of the absorbance at 260nm and 280nm. DNA with no impurities has a reading of 1.8.

2.4.4 Polymerase Chain Reaction (PCR)

2.4.4.1 Primers

For optimal PCR reactions, the primers were designed with several parameters in mind, such as length, %GC content, internal loop structures or the possibility of primer-dimer formation.

Primers were synthesised and purified by desalting, by Invitrogen’s custom primer service. For details of all primers, see appendix 4.

2.4.4.2 Standard PCR Conditions

For routine non-quantitative PCR to confirm ligations were successful during cloning, Taq Polymerase (Promega) was used, with its supplied buffers (table 2-8). Reactions were performed in a PCR Sprint machine (ThermoElectron) as described in table 2.10.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PCR Buffer</td>
<td>5</td>
</tr>
<tr>
<td>25mM MgCl₂</td>
<td>3</td>
</tr>
<tr>
<td>2mM dNTP mix</td>
<td>5</td>
</tr>
<tr>
<td>Sense Primer (10pmol/μl)</td>
<td>5</td>
</tr>
<tr>
<td>Antisense Primer (10pmol/μl)</td>
<td>5</td>
</tr>
<tr>
<td>DNA template</td>
<td>2</td>
</tr>
<tr>
<td>Taq Polymerase (4U/μl)</td>
<td>0.5</td>
</tr>
<tr>
<td>H₂O</td>
<td>24.5</td>
</tr>
</tbody>
</table>

Table 2-10 Composition of a standard 50μl PCR
<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>93°C</td>
<td>40 seconds</td>
<td>10</td>
</tr>
<tr>
<td>Annealing</td>
<td>1 minute</td>
<td></td>
</tr>
<tr>
<td>73°C</td>
<td>1 minute 40 seconds</td>
<td></td>
</tr>
<tr>
<td>93°C</td>
<td>40 seconds</td>
<td>10</td>
</tr>
<tr>
<td>Annealing</td>
<td>1 minute</td>
<td></td>
</tr>
<tr>
<td>73°C</td>
<td>1 minute 50 seconds</td>
<td></td>
</tr>
<tr>
<td>93°C</td>
<td>40 seconds</td>
<td>10</td>
</tr>
<tr>
<td>Annealing</td>
<td>2 minutes</td>
<td></td>
</tr>
<tr>
<td>73°C</td>
<td>10 minutes</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 2-11 Cycling protocol for standard PCR**

Annealing temperatures depended on the length of the primers and the %GC content, and usually ranged from 55°C-65°C.

**2.4.4.3 High Fidelity PCR Conditions**

For qualitative RT-PCR, or cloning, the polymerases Bio-X-Act and HotStarTaq were used respectively.

**2.4.4.3.1 HotStarTaq Polymerase**

HotStarTaq DNA polymerase (Qiagen) has a high PCR specificity with reduced non-specific amplification, and leaves an 'A' overhang. Reactions were set up and cycled as described in tables 2.12 and 2.13 respectively.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PCR buffer</td>
<td>5</td>
</tr>
<tr>
<td>2mM dNTP mix</td>
<td>5</td>
</tr>
<tr>
<td>Sense primer (5pmol/μl)</td>
<td>5</td>
</tr>
<tr>
<td>Antisense primer (5pmol/μl)</td>
<td>5</td>
</tr>
<tr>
<td>DNA template</td>
<td>2</td>
</tr>
<tr>
<td>HotStarTaq Polymerase (5U/μl)</td>
<td>0.5</td>
</tr>
<tr>
<td>H₂O</td>
<td>27.5</td>
</tr>
</tbody>
</table>

**Table 2-12 Composition of a standard 50μl PCR, using HotStarTaq polymerase**
Table 2-13 Cycling protocol for PCR with Taq polymerase

For PCR products longer than 1kb, an extension time of approximately 1 min per kb DNA was used.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>95°C</td>
<td>15 minutes</td>
<td>1</td>
</tr>
<tr>
<td>94°C</td>
<td>1 minute</td>
<td></td>
</tr>
<tr>
<td>Annealing</td>
<td>1 minute</td>
<td>35</td>
</tr>
<tr>
<td>72°C</td>
<td>1 minute 40 seconds</td>
<td></td>
</tr>
<tr>
<td>72°C</td>
<td>10 minutes</td>
<td>1</td>
</tr>
</tbody>
</table>

2.4.4.3.2 Bio-X-Act Polymerase

Bio-X-Act DNA Polymerases (Bioline) are a high-performance proprietary complex of enzymes and additives specifically designed for difficult or problematic applications requiring high processivity with high-fidelity. The unique composition of different enzymatic activities delivers a 17-fold increase in fidelity in comparison with regular Taq polymerase and is therefore suitable for reactions where accuracy is important. Bio-X-Act also leaves an ‘A’ overhang which can be utilised in TA cloning.

Bio-X-Act was used with its supplied buffers, as described in table 2.14 and cycled as described in table 2-11.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x Optibuffer</td>
<td>5</td>
</tr>
<tr>
<td>50mM MgCl2</td>
<td>1.5</td>
</tr>
<tr>
<td>2mM dNTP mix</td>
<td>5</td>
</tr>
<tr>
<td>Sense primer (10pmol/µl)</td>
<td>5</td>
</tr>
<tr>
<td>Antisense primer (10pmol/µl)</td>
<td>5</td>
</tr>
<tr>
<td>DNA template</td>
<td>2</td>
</tr>
<tr>
<td>Bio-X-Act (4U/µl)</td>
<td>0.5</td>
</tr>
<tr>
<td>H2O</td>
<td>26</td>
</tr>
</tbody>
</table>

Table 2-14 Composition of a high fidelity PCR using Bio-X-Act polymerase
2.4.5 Agarose-TAE Gel Electrophoresis

DNA samples were prepared by the addition of an appropriate volume of 6x loading buffer (Promega) and loaded onto a 0.7-2% electrophoresis grade agarose gel containing ethidium bromide at a final concentration of 0.5μg/ml. The gel was electrophoresed at a constant voltage of 100V in 1x TAE running buffer (pH 8.0; 40mM tris, 40mM acetic acid, 2mM EDTA) for 45-60 minutes. Sizes of DNA fragments were estimated by running alongside a standard DNA marker. The Promega 100bp and 1kb DNA ladders were used for DNA of sizes 100bp – 1kb and 1kb – 10kb respectively (see section 2.4.5.1).

2.4.5.1 DNA Molecular Weight Markers

2.4.5.1.1 100bp Ladder

The Promega 100bp ladder consists of DNA fragments that range in size from 100-1,000bp in 100bp increments, plus a fragment at 1,500bp. To allow easy identification the 500bp fragment is present at increased intensity (figure 2-1).

2.4.5.1.2 1Kb ladder

The Promega 1Kb ladder consists of 13 fragments ranging in size from 250-10,000bp. The 1,000 and 3000bp fragments are of increased intensity relative to the other bands to serve as reference indicators (figure 2-1).

![Figure 2-1 100bp and 1kb DNA ladders, adapted from www.promega.com](image-url)
2.4.6 Gel Visualisation and Documentation

DNA was visualised on a UV transilluminator and photographed using a Syngene gel documentation darkroom system and analysed by GeneSnap Version 2.60.0.14 and GeneTools Version 2.10.03 (Synoptics).

2.4.7 Gel Extraction

To extract and purify DNA from an agarose gel, a QIAquick Gel Extraction Kit (Qiagen) was used, where buffer QG is used to solubilise the agarose and provide appropriate conditions for the DNA to bind to the silica membrane of the column. The DNA is then washed with ethanol and eluted from the column.

The DNA fragment to be extracted was excised from the agarose gel with a clean scalpel. The gel slice was weighed; and three times gel volume of buffer QG was added to the sample and incubated at 50°C for 10 minutes until the gel slice was completely dissolved. One gel volume of isopropanol was added to the sample and mixed. To bind the DNA, the sample was applied to a QIAquick column in a 2ml collection tube and centrifuged for 1 minute at 13,000rpm. The flow-through was discarded and the QIAquick column was placed back in the same collection tube. 0.5ml of buffer QG was added to the column and centrifuged for 1 minute at 13,000rpm. To wash the DNA, 0.75ml of buffer PE was added to the QIAquick column and allowed to stand for 5 minutes at room temperature before centrifuging for 1 minute at 13,000rpm. The flow-through was discarded and the QIAquick column was centrifuged for an additional 1 minute at 13,000rpm to remove any residual ethanol. The QIAquick column was placed into a clean 1.5ml microcentrifuge tube and 30μl sterile H₂O was added to the column. The column was allowed to stand for 1 minute at room temperature and the DNA was eluted by centrifugation for 1 minute at 13,000rpm.

2.4.8 PCR Purification

To purify PCR reactions the DNA was desalted using the QIAquick PCR purification kit (Qiagen). 5 volumes of buffer PB was added to 1 volume of the PCR sample and mixed. To bind the DNA, the sample was applied to a QIAquick spin column in a 2ml collection tube and centrifuged for 1 minute at 13,000rpm. The flow-through was discarded and the QIAquick column was placed back in the same tube. To wash the DNA, 0.75ml of buffer PE was added to the QIAquick column
and centrifuged for 1 minute at 13,000rpm. The flow-through was discarded and the QIA-quick column was placed back into the same tube and centrifuged for an additional 1 minute at 13,000rpm to remove any residual ethanol. The QIAquick column was placed into a clean 1.5ml microcentrifuge tube and 30μl sterile H₂O was added to the centre of the QIAquick membrane and allowed to stand for 1 minute at room temperature. The DNA was eluted by centrifugation for 1 minute at 13,000rpm.

### 2.4.9 Automated DNA Sequencing

The sequences of all constructs were verified by DNA sequencing, using BigDye v3.1 in the Ready Reaction Premix with the supplied 5x buffer (Applied Biosystems). The reactions were set up as shown in table 2-15, using linearised DNA, and cycled as shown in table 2-16. The sequencing reactions were then cleaned by ethanol/sodium acetate precipitation and analysed on an ABI 377 sequencer (Applied Biosystems) by Central Biotechnology Services, Cardiff University.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ready Reaction Premix (2.5x)</td>
<td>4μl</td>
</tr>
<tr>
<td>BigDye Sequencing Buffer (5x)</td>
<td>2μl</td>
</tr>
<tr>
<td>DNA linear template</td>
<td>Single-stranded: 25-50ng</td>
</tr>
<tr>
<td></td>
<td>Double-stranded: 150-300ng</td>
</tr>
<tr>
<td></td>
<td>PCR template: 10-50ng</td>
</tr>
<tr>
<td>Primer</td>
<td>3.2pmols</td>
</tr>
<tr>
<td>H₂O</td>
<td>To 20μl</td>
</tr>
</tbody>
</table>

**Table 2-15 DNA sequencing reaction composition**

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Time</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>96°C</td>
<td>1 minute</td>
<td>1</td>
</tr>
<tr>
<td>96°C</td>
<td>10 seconds</td>
<td>25</td>
</tr>
<tr>
<td>50°C</td>
<td>5 seconds</td>
<td></td>
</tr>
<tr>
<td>60°C</td>
<td>4 minutes</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2-16 Cycling protocol for Big-Dye v3.1 sequencing**
2.4.10 Sequencing Clean-Up

To purify the extension product from any unincorporated dye-labelled terminators, ethanol/sodium acetate precipitation was used. The sequencing reaction was briefly microfuged to remove any condensation, followed by the addition of 2μl of 3M sodium acetate (pH5.5) and 50μl of 100% ethanol. The reactions were mixed followed by incubation at room temperature for 15 minutes, and centrifugation at 13,000rpm for 30 minutes in a MSE Micro Centaur. The supernatant was removed and the DNA was washed twice, by the addition of 200μl of 70% ethanol and centrifugation at 13,000rpm for 5 minutes. After the final wash, the supernatant was removed and the DNA was sequenced.

2.4.11 In Vitro Transcription

To make labelled RNA probes for use in the RPA, the sequence of the DNA of interest is firstly cloned into the pGEM-T vector, as previously described. The plasmid was linearised by digestion with a restriction enzyme that leaves a blunt or 5' overhang downstream of the DNA of interest. The products of the digestion were visualised using agarose-TAE electrophoresis (section 2.4.5) and fully cut plasmids were extracted using the Qiagen gel extraction kit (section 2.4.7).

To produce the antisense single stranded RNA transcript from the linearised DNA template, T7 phage RNA polymerase was used. The following transcription reaction was assembled at room temperature.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclease-free H₂O</td>
<td>to 20μl</td>
</tr>
<tr>
<td>Linear DNA template</td>
<td>1μg</td>
</tr>
<tr>
<td>10x Transcription Buffer</td>
<td>2μl</td>
</tr>
<tr>
<td>10mM ATP</td>
<td>1μl</td>
</tr>
<tr>
<td>10mM CTP</td>
<td>1μl</td>
</tr>
<tr>
<td>10mM GTP</td>
<td>1μl</td>
</tr>
<tr>
<td>10mM UTP</td>
<td>0.6μl</td>
</tr>
<tr>
<td>Biotin-16-UTP</td>
<td>0.4μl</td>
</tr>
<tr>
<td>T7 enzyme mix</td>
<td>2μl</td>
</tr>
</tbody>
</table>

Table 2-17 In Vitro transcription of biotin-labelled RNA
The reaction was mixed by gently pipetting up and down. The reaction was then microfuged briefly and incubated at 37°C for 2 hr. After which the DNA template was removed by the addition of 1µl DNase I (1U/µl), mixed and incubated at 37°C for 15 min. To stop the reaction, 1µl of 0.5M EDTA was added which chelates metal ions and thereby inhibits DNase I from functioning.

2.4.12 Gel Purification

As the probe was to be used for a ribonuclease protection assay (RPA) it was necessary to primarily have full-length probe by gel purification. After transcription, 20µl of gel loading buffer was added to the probe and incubated at 95°C for 4 minutes to denature any secondary structure, and placed on ice to prevent renaturation. The probe was loaded onto a 5% acrylamide/8M urea denaturing polyacrylamide gel, and run at 200V in 1x TBE buffer for approximately 1 hour until the blue dye front of the bromophenol blue reached the bottom of the gel.

The gel was stained with ethidium bromide in 1x TBE buffer for 5 min and visualised under UV light. The full length RNA probe was excised from the gel, cut into small pieces to increase the surface area and transferred into a nuclease-free microfuge tube. To elute the transcript, the gel fragment was incubated overnight in 350µl elution buffer (0.5M NH₄OAc/1mM EDTA/0.2% SDS). The probe was precipitated by the addition of 3x volume 100% ethanol and centrifugation at 13000rpm for 20 min at 4°C. The supernatant was removed, and the cell pellet consisting of the RNA transcript was resuspended in MMLV Reverse Transcriptase 1x buffer (Promega; 50mM Tris-Cl pH 8.3, 75mM KCl, 3mM MgCl₂, 10mM DTT) with 40U of RNasin ribonuclease inhibitor and stored at −70°C prior to its use in the RPA.

2.5 RNA Methodology

2.5.1 RNA Isolation

The RNA working area and pipettes were decontaminated using RNase Away (Merck) prior to handling RNA.
2.5.1.1 Total RNA Isolation

Total RNA was extracted from mammalian cell cultures using Ultraspec™ RNA isolation system (Biotecx USA, purchased from AMS Biotechnology). A monolayer of HepG2/IMR32 cells were scraped from a plate or flask into 1ml Ultraspec and transferred to an RNase/DNase free 1.5ml microcentrifuge tube. For non-adherent K562 cells, the cells were pelleted at 1500rpm for 3 minutes, before being lysed in 1ml Ultraspec. Cell lysates were homogenised by passing them through a pipette several times and incubated at 4°C for 5 minutes.

To extract the RNA, 200μl of chloroform were added per 1ml of Ultraspec and samples shaken for 15 seconds and incubated on ice for 15 minutes. The homogenate was centrifuged for 15 minutes at 13,000rpm in a MSE Micro Centaur at 4°C to separate the lower, organic phase containing the DNA from the upper, aqueous phase containing the RNA. The aqueous phase was transferred to a fresh RNase/DNase free 1.ml microcentrifuge tube and an equal volume of isopropanol added. Samples were incubated on ice for 10 minutes and centrifuged for 10 minutes at 13,000 rpm in a MSE Micro Centaur at 4°C to precipitate the RNA. The supernatant was carefully removed and the RNA pellet washed twice with 75% ethanol by vortexing and centrifugation for 5 minutes at 6,000rpm in a MSE Micro Centaur at 4°C. The RNA was air-dried and resuspended in 30μl nuclease-free water (Sigma) and stored at −70°C until use.

2.5.1.2 mRNA Isolation

To directly extract mRNA from HepG2 cells, the Dynabeads mRNA Direct Kit (Dynal) was used and the manufacturer’s instructions were followed.

A monolayer of HepG2/IMR32 cells were scraped from a plate or flask into 1ml ice-cold PBS, transferred to an RNase/DNase free 1.5ml microcentrifuge tube, and centrifuged at 1,500rpm for 3 minutes. The PBS was removed and the cell pellet was stored at -70°C until the mRNA was extracted.

The Dynabeads were prepared by transferring 250μl of beads into an RNase-free 1.5ml microcentrifuge tube placed on a magnet (Dynal MPC-S). After 30 seconds the supernatant was removed, and the beads were resuspended in 250μl of fresh lysis(binding buffer (100mM Tris-HCl pH 7.5, 500mM LiCl, 10mM EDTA pH 8, 1%
LiDS, 5mM DTT). The beads remained in this buffer until the sample lysate was ready for combination with the beads.

The cell pellet was lysed by the addition of 1250μl of lysis/binding buffer. To obtain complete lysis a repeated passage of the solution through a pipette tip was performed. The lysate was then forced through a 21-gauge needle 5 times using a 2ml syringe to shear the DNA. The lysate was transferred to the prepared Dynabeads Oligo (dT)$_{25}$. The beads were mixed with the sample lysate and incubated for 5 minutes at room temperature with continuous mixing to allow the polyA-tail of the mRNA to anneal to the oligo-dT on the beads. The vial was placed on the magnet for 2 minutes, and the supernatant was removed. The beads/mRNA complex was washed twice with 1ml of washing buffer A (10mM Tris-HCl pH 7.5, 150mM LiCl, 1mM EDTA, 0.1% LiDS) at room temperature, and then washed once in 1ml of washing buffer B (10mM Tris-HCl pH 7.5, 150mM LiCl, 1mM EDTA) at room temperature, using the magnet to separate the beads from the solution between each step.

To elute the mRNA from the beads, 25μl of 10mM Tris-HCl was added to the beads/mRNA complex and incubated for 2 minutes at 80°C. The tube was placed immediately on the magnet and the supernatant containing the mRNA was transferred to a new RNase-free microcentrifuge tube and stored at −70°C.

### 2.5.2 RNA Quantification

Total RNA was quantified using the GeneQuant DNA/RNA Calculator set to interpret RNA, as described in 2.3.3. The purity of the RNA was confirmed by checking the ratio of the absorbance at 260nm and 280nm. RNA with no impurities has a reading of 2.0.

### 2.5.3 Removal of DNA Contamination

Prior to first strand synthesis, RNA samples were treated with DNase I [Amplification Grade] (Invitrogen) to ensure that any remaining DNA was eliminated from the RNA preparation. Two units of DNase I was added to 1μg of RNA in DNase 10x reaction buffer (20mM Tris-HCl pH 8.4, 2mM MgCl$_2$, 50mM KCl) and incubated for 15 minutes at room temperature. The enzyme was inactivated by the addition of 1μl 25mM EDTA and heating to 65°C for 10 minutes.
2.5.4 First Strand Synthesis

For first strand synthesis of cDNA from RNA, MMLV Reverse Transcriptase was used with its supplied buffer, along with RNasin Ribonuclease Inhibitor to prevent RNA degradation and Oligo dT primer and dNTP mix, as described in table 2-18. The reaction was incubated for 5 minutes at 25°C, followed by incubation at 42°C for 1 hour and 70°C for 10 minutes to inactivate the reaction. The cDNA was stored at −20°C until further use.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5x M-MLV-RT Buffer</td>
<td>4</td>
</tr>
<tr>
<td>10mM dNTP mix</td>
<td>2</td>
</tr>
<tr>
<td>Oligo dT primer (50ng/μl)</td>
<td>1</td>
</tr>
<tr>
<td>RNasin Ribonuclease Inhibitor (40U/μl)</td>
<td>0.25</td>
</tr>
<tr>
<td>DNase I-treated RNA</td>
<td>2μg</td>
</tr>
<tr>
<td>M-MLV RT (200U/μl)</td>
<td>1</td>
</tr>
<tr>
<td>H₂O</td>
<td>To 20μl</td>
</tr>
</tbody>
</table>

Table 2-18 First strand synthesis of RNA

2.5.5 RNase Protection Assay (RPA)

For each experimental tube, labelled probe was mixed with mRNA extracted from the HepG2 cells, in a 1.5ml microcentrifuge tube. For each different probe used, two control tubes were included containing the same amount of labelled probe used for the experimental tubes plus yeast RNA equivalent to the highest amount of sample RNA. 5M NH₄OAc was added to a final concentration of 0.5M, followed by the addition of 2.5 volumes 100% ethanol, and mixed thoroughly. Samples were incubated at −20°C for 15 minutes. The RNA was pelleted by centrifugation at 13,000 RPM in a microcentrifuge for 15 minutes at 4°C. The supernatant was removed and the tubes were spun a second time for 5 seconds to collect any supernatant left on the sides of the tubes, and the remaining supernatant was removed. Pellets were air dried for 5 minutes and resuspended in 20μl of Hybridisation buffer. To denature the RNA and aid its solubilisation, samples were incubated for 4 minutes at 95°C. Samples were incubated overnight at 42°C to hybridise the probe to its complement in the sample RNA.

After hybridisation, the tubes were centrifuged briefly, and 200μl of the diluted RNase mixture was added to each tube, and to one of the two yeast RNA control
tubes that have been prepared for each probe in the experiment. To the remaining yeast RNA control tube, 200\( \mu \)l RNAse digestion buffer without RNAse was added. Samples were incubated for 30 minutes at 37\(^{\circ}\)C, during which time unprotected single-stranded RNA will be digested. 300\( \mu \)l of RNAse inactivation/precipitation solution was then added to the samples, with 2\( \mu \)l Yeast RNA to increase the size and visibility of the final pellets, and incubated at -20\(^{\circ}\)C for 15 minutes. The precipitated products were pelleted by centrifugation at 13,000 RPM at 4\(^{\circ}\)C, and the supernatant was removed. Pellets were resuspended in 6\( \mu \)l Gel Loading Buffer II and incubated for 3 minutes at 95\(^{\circ}\)C to completely solubilise and denature the RNA. RNA samples were loaded onto a 5% polyacrylamide gel and electrophoresed at 25mAmps constant current, until the leading dye band was near the bottom of the gel.

2.5.6 Detection of Non-Isotopic probes

To detect the biotinylated RNA probes used in the RPA, the BrightStar BioDetect kit (Ambion) with its supplied buffers was used, and the manufacturer's instructions were followed. This is a nonisotopic chemiluminescent detection system that does not suffer from the high backgrounds and low sensitivity typical of these systems, and is used to detect the biotinylated RNA on positively charged nylon membrane. All incubations/wash steps were done with gentle agitation at room temperature. The 100cm\(^2\) membrane was washed twice in 100ml 1x Wash buffer for 5 minutes. The membrane was then incubated twice in 50ml Blocking Buffer for 5 minutes each, before a final incubation in 100ml Blocking Buffer for 30 minutes. The membrane was then incubated in 10ml Blocking Buffer, containing 1\( \mu \)l Streptavidin-Alkaline Phosphatase (Strep-AP), for 30 minutes. The membrane was incubated for 10 minutes in 50ml Blocking Buffer, and washed three times in 1x Wash Buffer for 15 minutes each to decrease non-specific background. The membrane was then incubated twice in 50ml 1x Assay Buffer for 2 minutes each, and then in 5ml of CDP-Star. After 5 minutes, the membrane was blotted on filter paper to remove any excess CDP-Star, wrapped in a single layer of plastic wrap, and exposed to film at room temperature. The CDP-Star reaches peak light emission in 2-4 hours, and then emission falls to a plateau which persists for several days.
2.6 Cloning

2.6.1 TA Cloning

DNA fragments to be cloned were amplified using high fidelity PCR as described in 2.3.4. The PCR product was separated on TAE/agarose gel and the specific product band purified by gel extraction. The DNA insert was then ligated into pGEM®-T and the ligation mix used to transform DH5α cells which were plated onto LB-AIX plates for blue/white colour selection.

![Diagram of pGEM®-T Vector](image)

Figure 2-2 Vector map of pGEM®-T from www.promega.com

2.6.2 Ligation

To calculate the appropriate amount of insert DNA to be used in the ligation reaction containing 50ng of vector, the following equation was used:

\[
\text{ng of vector} \times \text{kb size of insert} \times \text{insert:vector molar ratio} = \text{ng of insert} \\
\text{kb size of vector}
\]
<table>
<thead>
<tr>
<th>Reagent</th>
<th>Standard Reaction</th>
<th>Positive Control</th>
<th>Background control</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x Rapid Ligation Buffer</td>
<td>5µl</td>
<td>5µl</td>
<td>5µl</td>
</tr>
<tr>
<td>pGEM®-T vector (50ng)</td>
<td>1µl</td>
<td>1µl</td>
<td>1µl</td>
</tr>
<tr>
<td>Insert DNA</td>
<td>As calculated by equation</td>
<td>0µl</td>
<td>0µl</td>
</tr>
<tr>
<td>Control insert DNA</td>
<td>0µl</td>
<td>2µl</td>
<td>0µl</td>
</tr>
<tr>
<td>T4 DNA Ligase (3U/µl)</td>
<td>1µl</td>
<td>1µl</td>
<td>1µl</td>
</tr>
<tr>
<td>H₂O</td>
<td>To 10µl</td>
<td>1µl</td>
<td>3µl</td>
</tr>
</tbody>
</table>

**Table 2-19 Ligation of insert DNA into the pGEM-T vector**

Generally, the insert:vector molar ratios used were 3:1 and 1:1. The ligation reactions were incubated overnight at 4°C, and then transformed into DH5α cells.

### 2.6.3 Screening Colonies for Inserts

pGEM-T produces β-galactosidase by α-complementation when transformed into DH5α cells, giving rise to blue colonies in the presence of X-gal. Successful cloning of an insert into pGEM®-T disrupts the β-galactosidase gene; therefore bacteria can be screened with the use of LB-AX plates (LB-agar containing ampicillin and X-Gal). DH5 cells do not require IPTG to induce expression from the lac promoter even though the strain expresses the lac repressor, as the copy number of most plasmids exceeds the repressor number in the cells. The insertion of DNA into the cloning site disrupts the lacZ gene, therefore preventing α-complementation and producing white colonies.

### 2.6.4 Colony screening by PCR

To verify if the bacterial colonies contain the correctly ligated construct in the correct orientation they were screened by PCR, using primers that would produce a PCR product that overlaps the plasmid and insert. Each colony was picked and placed into 30µl LB with the appropriate antibiotic and incubated for 1 hour at 37°C with shaking. PCR was then performed on the bacteria, as described in section 2.4.4.2, using Taq polymerase.
2.7 RNA-Protein Interaction Assays

2.7.1 In Vitro Transcription

The plasmid was linearised by digestion with a restriction enzyme that leaves a blunt or 5’ overhang downstream of the DNA of interest. The products of the digestion were visualised using agarose-TAE electrophoresis (section 2.4.5) and fully cut plasmids were extracted using the Qiagen gel extraction kit (section 2.4.7). For the EMSA’s and UV-crosslinking, radiolabelled RNA was made with α32-CTP. The method for in vitro transcription in section 2.4.11 was used, but with CTP as the limiting nucleotide. The method for radiolabelled RNA is slightly different to non-radiolabelled RNA, and is therefore described here.

Transcription reactions were performed using the Ambion T7 MAXIscript (similar to section 2.4.11). To 500ng of the linearised DNA template the reagents were added in following order: 1µl 10X transcription buffer, 0.5µl 10mM ATP, 0.5µl 10mM GTP, 0.5µl 10mM UTP, 1µl 0.1mM CTP, 2.5µl 32P CTP, 1µl T7 enzyme mix (15U/µl) and RNase-free H2O to a final volume of 10µl. The tube was gently flicked to mix and microfuged briefly. The reaction was incubated at 15°C for 1.5 hours and then stopped by the addition of 40µl TE buffer. The RNA was purified from the unincorporated nucleotides by spinning through a Sephadex G-50 column (Amersham-Pharmacia Biotech) at 4,000rpm for 2 minutes in a microfuge. The amount of radioisotope made in each reaction was quantified using a Wallac 1409 DSA liquid scintillation counter (Perkin Elmer Life Sciences) in the Department of Medical Physics, University Hospital of Wales, Cardiff.

To check the integrity and size of the probes, the RNA was electrophoresed on a 1x TBE, 4% polyacrylamide, 8M urea, denaturing gel. 2µl of each probe was incubated with 8µl RNA loading dye and incubated at 95°C for 2 minutes to destabilise any secondary structure, before loading onto the gel and separating at 200V in 1x TBE buffer. The gel was exposed to film for 5-10 minutes to visualise the RNA.

2.7.2 RNA Electrophoretic Mobility Shift Assay (R-EMSA)

The EMSA binding reactions were each carried out in a volume of 10µl using 5x-binding buffer (50mM hepes [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid] pH7.2, 15mM MgCl2, 25% glycerol, with freshly-added 5mM DTT). The binding
reaction contained 0.1mg/ml bovine liver rRNA, 25μg HepG2 cytosolic protein, 100mM KCl, and 100,000 cpm labelled RNA probe in 1x Binding Buffer and were incubated at room temperature for 15 minutes. Controls of each probe without any protein added were also used. The reactions were incubated for a further 10 minutes with 5μg/μl heparin. After the addition of 2μl of RNA loading buffer (48.5% glycerol, 0.5% bromophenol blue, 0.5% xylene cyanol) the samples were separated on a non-denaturing polyacrylamide gel.

2.7.2.3 Non-Denaturing Gel Electrophoresis

Samples were separated on a 4% non-denaturing polyacrylamide gel (60:1 acrylamide:bis-acrylamide) at 200V, for 5 hours in 0.5x TBE (44.5mM Tris-HCl, 44.5mM boric acid, 1mM EDTA, pH 8). This gel had been pre-run at 200V for 20 minutes to ensure a constant voltage for the EMSA. After electrophoresis, the gel was fixed for 20 minutes with destaining solution (20% methanol, 10% glacial acetic acid), dried for 1 hour at 70°C on 3mm Whatman paper, and exposed to film for up to 3 days at -70°C.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>40% acrylamide</td>
<td>2.95ml</td>
</tr>
<tr>
<td>2% bis-acrylamide</td>
<td>0.98ml</td>
</tr>
<tr>
<td>TBE 10x buffer</td>
<td>1.5ml</td>
</tr>
<tr>
<td>Ammonium persulphate (25%)</td>
<td>60μl</td>
</tr>
<tr>
<td>TEMED</td>
<td>60μl</td>
</tr>
<tr>
<td>H₂O</td>
<td>24.45ml</td>
</tr>
</tbody>
</table>

Table 2-20 4% polyacrylamide (60:1 acrylamide:bis-acrylamide) gel

2.7.3 UV-Crosslinking

As for the EMSAs, the UV-crosslinking reactions were carried out in a volume of 10μl, containing 0.1mg/ml bovine liver rRNA, 25μg HepG2 cytosolic protein, 50mM KCl, and 100,000 cpm labelled RNA probe in 1x binding buffer. The reactions were incubated at room temperature for 15 minutes. After which, heparin was added to a final concentration of 5μg/μl, and the reaction was incubated for 15 minutes at room temperature. The microtitre plate was placed on ice and exposed twice to 8600J of Ultraviolet (UV) radiation using a UVC 500 UV Crosslinker (Amersham Lifesciences), which cross-linked any RNA which had bound to the cytosolic proteins. The free RNA and the RNA not protected by cross-linked proteins were digested by incubating the reactions with 4μg of RNase
A and 1U of RNase T1 in RNase dilution buffer (10mM Tris-HCl pH7.5, 1mM MgCl₂, 100mM KCl) at 37°C for 30 minutes. Samples were denatured prior to SDS-PAGE analysis by the addition of 5µl of 4x loading buffer (125mM Tris, 20% (v/v) glycerol, 4% (w/v) 2-mercaptoethanol, 0.04% (w/v) bromo-phenol blue) and incubated for 5 minutes at 80°C. The reactions were loaded onto a 10% pre-run SDS-polyacrylamide gel, and electrophoresed at 200V in 1x SDS buffer (25mM Tris, 50mM glycine, 0.1% SDS) until the dye front had reached the bottom of the gel (see section 2.3.2). The gel was fixed by incubating in destaining solution (20% methanol, 10% glacial acetic acid) for 10 minutes and dried for 1 hour at 70°C on 3mm whatman paper. The gel was exposed to X-ray film overnight to up to 3 days at -70°C.
CHAPTER 3

Analysis of the human major and minor
ALAS1 5′-UTRs
3.1 Introduction

The organisation of the human ALAS1 gene is similar to the orthologous rat, mouse and chicken genes, except that it contains two non-coding exons (1A and 1B) in its 5'-UTR. This is unlike the rat gene, which contains one (Yomogida et al., 1993), and the chicken, which has no non-coding exons in its 5'-UTR, as explained in the general introduction (Yomogida et al., 1993; Maguire et al., 1986). Roberts et al have shown that exon 1B of the human gene is alternatively spliced resulting in two mRNA species: a major one, in which it is omitted, and a minor one containing both exons 1A and 1B (figure 1-8 in the general introduction). This longer minor mRNA is present in all tissues and cell lines that have so far been examined, and represents approximately 10% of the steady state concentration of ALAS1 mRNA (Roberts & Elder, 2001).

Intron 1 of the rat and mouse ALAS1 genes contains sequences that are very similar to the splice donor and acceptor sites of exon 1B in the human gene. However, there is a crucial difference in that the human gene contains an A nucleotide that creates a splice acceptor site in place of the C or T nucleotides in the other species. The inclusion of exon 1B in the minor transcript of the human gene extends its 5'-UTR by 177bp to a minimum of 304bp. Alternative splicing is unlikely to have a significant role in the basal expression of ALAS1, as the percentage of the minor transcript is similar in all tissues (Roberts & Elder, 2001). However, the possibility that it may modulate expression of the downstream mRNA has not yet been explored. Very little is known about the molecular mechanisms that determine either the basal or the drug-induced, tissue-specific regulated transcription of mammalian ALAS1 genes.

It has been reported for various genes, such as pyruvate carboxylase (PC), that increasing the length of the 5'-UTR alters translational efficiency (Jitrapakdee et al., 1998). PC catalyzes the ATP-dependent carboxylation of pyruvate to oxaloacetate, the first regulated step in the gluconeogenic pathway. There is only one copy of the pyruvate carboxylase PC gene (Webb et al., 1997), it generates several forms of PC mRNA, which diverge at their 5'-UTRs but share the same open reading frame encoding a 1178-residue polypeptide (Jitrapakdee et al., 1996). Two tissue-specific promoters are responsible for the production of two primary transcripts, which then are differentially spliced (Jitrapakdee et al., 1998).
Similar to the above example, translational efficiency may be affected by alternative splicing in ALAS1. If this is so, then any change in the proportion of transcripts during induction, or any effect of 5'-UTR lengthening on stability towards haem, provides a potential mechanism for the fine control of its regulated expression. Alternative splicing of exon 1B of ALAS1 does not appear to be specific to a particular tissue or cell line. Both transcripts have been detected by RT-PCR in mRNA from human brain and liver and from K562, HeLa, HepG2 and B-lymphoblastoid cell lines (Roberts & Elder, 2001).

The approach to assessing the differences between the ALAS1 mRNA isoforms involved the analysis of the two different 5'-UTRs in response to haem. Semi-quantitative RT-PCR was used to look at mRNA levels directly, whilst the dual luciferase reporter assay (DLRA) was used to compare ALAS1 expression.

3.1.1 The Dual Luciferase Reporter Assay

Genetic reporter systems are widely used to study eukaryotic gene expression and cellular physiology, and dual reporters are commonly used to improve experimental accuracy. Therefore this assay seemed particularly useful to assess the difference between the expression of the major and minor ALAS1 mRNA isoforms.

Dual luciferase reporter assays involve the simultaneous expression and measurement of two individual reporter enzymes within a single system. The 'experimental' reporter is correlated with the effect of specific experimental conditions, while the activity of the co-transfected 'control' reporter provides an internal control that serves as the baseline response. Normalising the activity of the experimental reporter to the activity of the internal control minimizes experimental variability caused by differences in cell viability or transfection efficiency.

In the Dual Luciferase Reporter Assay system the activities of firefly (Photinus pyralis) and Renilla (Renilla reniformis) luciferases are measured sequentially from a single sample. The firefly luciferase is measured first by adding Luciferase Assay Reagent II (LAR II) to generate a stabilised luminescent signal. After quantifying the firefly luminescence, this reaction is quenched, and the Renilla luciferase reaction is initiated by simultaneously adding Stop & Glo Reagent to the same tube. This reagent produces a stabilised signal from the Renilla luciferase.
Figure 3-1 Bioluminescent reactions catalysed by firefly and *Renilla* luciferases (adapted from www.promega.com).

a) Photon emission is achieved through oxidation of beetle luciferin in a reaction that requires ATP, Mg$^{2+}$ and O$_2$. The oxidation occurs through a luciferyl-AMP intermediate that turns over very slowly. As a result, this assay chemistry generates a 'flash' of light that rapidly decays after the substrate and enzyme are mixed.

b) The luminescent reaction of *Renilla* luciferase utilises O$_2$ and coelenterate-luciferin (coelenterazine). This reaction provides a stabilised luminescent signal that decays slowly over the course of the measurement.

### 3.1.1.2 Firefly and Renilla luciferases

Firefly and *Renilla* luciferases have distinct evolutionary origins and therefore have dissimilar enzyme structures and substrate requirements, making it possible to selectively discriminate between their respective bioluminescent reactions.

Firefly luciferase is a 61kDa monomeric protein that does not require post-translational processing for enzymatic activity. Therefore, it functions as a genetic reporter immediately upon translation. *Renilla* luciferase is a 36kDa protein, and similar to firefly luciferase, post-translational modification is not required for its activity. The enzyme may function as a genetic reporter immediately following translation.
Firefly and Renilla luciferase vectors may be used to co-transfect mammalian cells, and either firefly or Renilla luciferase may be used as the control or experimental reporter gene.

3.1.2 Aims and Strategy

It is not known how haem affects the expression of the two different ALAS1 isoforms. Therefore, the major aim of the work in this chapter was to determine whether both ALAS1 mRNA transcripts were destabilised by haem in HepG2 cells and the reasons for this.

This involved:
1. Assessing the stability of the minor and major 5'-UTR mRNA in response to haem.
2. Investigating whether the 5'-UTR affects expression of the downstream heterologous firefly luciferase gene.

In order to achieve these aims, semi-quantitative RT-PCR was used to assess the stability of the major and minor 5'-UTR in response to haem in HepG2 cells. ALAS1-Renilla luciferase fusion constructs were also used to determine how the differing 5'-UTRs affects expression of the downstream luciferase gene in response to haem.
3.2 Materials and Methods

This section details the specific materials and methods used in the experimental procedures adopted in the work for this chapter. General materials and methods are detailed in chapter 2.

The core promoter of the human ALAS1 gene contains multiple transcription start sites that are used in a tissue-specific fashion to produce transcripts of different lengths in brain and liver. There are six major transcription initiation start sites in the human and rat ALAS1 gene. However, in the liver, the majority of ALAS1 transcription is initiated from one specific initiation site, approximately 25bp downstream from the TATA element, demonstrated by primer extension analysis and S1 nuclease protection assay. This position is used for greater than 98% of both basal and induced transcription in rat liver. Transcription from upstream sites is less than 2% of the total and not significantly increased during induction. Basal transcription in human liver is also mainly from the same start site (Roberts & Elder, 2001). Therefore, the primers used throughout this chapter were designed to amplify the ALAS1 5'-UTR from the commonly used transcriptional start site in human liver cells.

3.2.1 Semi-Quantitative RT-PCR

To investigate the effect of haem upon the stability of the ALAS1 mRNAs, HepG2 cells were treated with/without haem for up to 8h. HepG2 cells were split into 6-well plates and when approximately 60% confluent, treated with 20μM haem or control medium with actinomycin D (5μg/ml) to inhibit transcription, and harvested at 0h, 1h, 2h, 4h and 8h when total RNA was isolated using Ultraspec, and its concentration and purity determined spectrophotometrically at 260 and 280nm. For RT-PCR, first strand cDNA synthesis was achieved by incubation at 42°C for 1h using 200U MMLV-reverse transcriptase with 2μg total RNA primed by oligo(dT)15. Changes in mRNA expression were quantified by PCR with HotStarTaq polymerase (table 3-2).

The two ALAS1 mRNAs was amplified specifically using different sense primers corresponding to the sequences across the exon 1A/1B boundary (primer P1) or the exon 1A/2 boundary (P2), to minimise competition for the two cDNA templates, and an antisense primer (P3) derived from the sequence in exon 3 (figure 3-2). The conditions were optimised for logarithmic amplification using
untreated control cells (section 3.2.1.1). To ensure results were normalised for equivalent concentrations of RNA, a 404bp GAPDH cDNA fragment was also amplified using sense (P4) and antisense (P5) primers (table 3-1). All primer sequences are listed in appendix 4.

<table>
<thead>
<tr>
<th>Target DNA</th>
<th>Primer Number</th>
<th>Product Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALAS1 minor 5'-UTR</td>
<td>Sense - P1, Anti-sense - P3</td>
<td>559bp</td>
</tr>
<tr>
<td>ALAS1 major 5'-UTR</td>
<td>Sense - P2, Anti-sense - P3</td>
<td>380bp</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Sense - P4, Anti-sense - P5</td>
<td>404bp</td>
</tr>
</tbody>
</table>

**Table 3-1 Semi-quantitative RT-PCR**
Primer sets used to amplify ALAS1 major, ALAS1 minor and GAPDH cDNA with their respective sizes.

**Figure 3-2 Primer Design to Differentiate ALAS1 5'-UTRs**
To detect and distinguish between the major and minor ALAS1 5'-UTRs, primers were designed which corresponded to sequences across the exon boundaries of exons 1A and 1B (P1) and exons 1A and 2 (P2). This minimised competition between the two cDNA templates. The same anti-sense primer (P3) in exon 3 was used for both. Only the 5'-UTR and the beginning of the ALAS1 coding region (exons 2 and 3) are shown.

The PCR of GAPDH was performed in a separate reaction, using primers P4 and P5, with HotstarTaq polymerase, and the conditions described in section 2.4.4.3.1. The cycling protocol for the PCRs was that as described in section 2.4.4.3.1, with 32 cycles, and annealing temperatures of 64°C (table 3-2).
<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PCR buffer</td>
<td>5</td>
</tr>
<tr>
<td>2mM dNTP mix</td>
<td>5</td>
</tr>
<tr>
<td>P1 5pmol</td>
<td>5</td>
</tr>
<tr>
<td>P2 5pmol</td>
<td>5</td>
</tr>
<tr>
<td>P3 5pmol</td>
<td>5</td>
</tr>
<tr>
<td>DNA Template (RT product)</td>
<td>5</td>
</tr>
<tr>
<td>HotStarTaq Polymerase (5U/µl)</td>
<td>1</td>
</tr>
<tr>
<td>H₂O</td>
<td>to 50µl</td>
</tr>
</tbody>
</table>

**Table 3-2 Composition of the multiplex RT-PCR**

The PCR products were resolved by agarose gel electrophoresis and the band intensities of the products were quantified as raw volume of the peak intensity using Gene Tools software (Syngene), and normalised to GAPDH to calculate the relative mRNA stability. Decay curves were plotted at equivalent GAPDH concentrations and half-life measurements calculated from the best-fit slope through the data. Similar experiments were also performed using IMR32 and K562 cells to establish if similar mechanisms are involved in other cells types.

**3.2.1.1 Optimising Semi-Quantitative RT-PCR**

The use of internal controls in semi-quantitative RT-PCR requires substantial optimisation for data to be meaningful. The PCR must be terminated when the products from both the internal control and the gene of interest are detectable and being amplified within the exponential phase. The exponential phase of amplification occurs when the reaction components are still in excess and the PCR products are accumulating at a constant rate, so that the quantity of products are proportional to the amount of starting target nucleic acid. To do this, PCRs were set up with a cycle number of 35, with reactions removed from the thermal cycler at cycle numbers ranging from 28 to 35, and resolved by agarose gel electrophoresis. Gene Tools software was used to quantify the intensity of the PCR product, and the PCR cycle number was then plotted against the natural log of the PCR intensity. When a straight line is obtained the samples are in the exponential phase of amplification, and at this cycle number the PCR can be used to quantify the products. This optimisation of PCR cycle number was used for all semi-quantitative PCRs within this thesis (data not shown).
3.2.2 PCR and Plasmid Construction

To investigate the role of the ALAS1 5'-UTRs in haem-mediated stability and to determine their effects upon translation, cDNA fragments for both the major and minor mRNAs were amplified and inserted into the firefly containing vectors, pGL3P and pGL3B (figures 3-3 and 3-4).

Both the major and minor ALAS1 5'-UTR sequences were amplified by PCR from HepG2 cDNA, from the major hepatic transcription start site in exon 1 to the initiating ATG in exon 2, using the sense primer P6 and the antisense primer P7, with 1U of Bio-X-Act Long Polymerase. These primers contained the Hind III (A↓GAGCT) and Nco I (C↓CATGG) restriction sites, respectively (table 3-3). The cDNAs were subsequently purified using a QIAGen PCR Clean-up Kit and cloned into the pGEM-T vector. Recombinant plasmid DNA was purified using a QIAprep spin kit and the identities of the clones for the long and short 5'-UTRs confirmed by automated fluorescent sequencing using Big Dye chemistry and analysis on an ABI Prism 3100 genetic analyser. Inserts from positive clones were removed by digestion with Hind III and Nco I and subcloned into Hind III/Nco I digested pGL3-Promoter and Basic vectors (figure 3-3). The Nco I site contains the initiating ATG of firefly luciferase thereby excluding any vector-derived 5'-UTR sequence.

<table>
<thead>
<tr>
<th>Target DNA</th>
<th>Primer Number</th>
<th>Restriction Site</th>
<th>Product Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALAS1 mRNA exon 1A - 2</td>
<td>Sense - P6,</td>
<td><em>Hind III</em></td>
<td>Major ALAS1 = 118bp</td>
</tr>
<tr>
<td></td>
<td>Anti-sense - P7</td>
<td><em>Nco I</em></td>
<td>Minor ALAS1 = 295bp</td>
</tr>
</tbody>
</table>

Table 3-3 Primers used to amplify the major and minor ALAS1 5'-UTRs, and the restriction enzyme sites within these primers.

To test for intrinsic transcriptional activity, these same ALAS1 fragments were ligated into Hind III/Nco I digested pGL3-basic vector, which lacks eukaryotic promoter and enhancer sequences (figures 3-3 and 3-4).
Figure 3-3 Vector maps of pGL3-Promoter and pGL3-Basic.

a) Vector map of the firefly luciferase expression vector pGL3-P from www.promega.co.uk

The pGL3-P vector contains an SV40 promoter upstream of the luciferase gene. DNA fragments containing putative enhancer elements can be inserted either upstream or downstream of the promoter-luc+ transcriptional unit.

b) Vector map of the firefly luciferase expression vector pGL3-B from www.promega.co.uk

The pGL3-Basic Vector lacks eukaryotic promoter and enhancer sequences, allowing maximum flexibility in cloning putative regulatory sequences. Expression of luciferase activity in cells transfected with this plasmid depends on insertion and proper orientation of a functional promoter upstream from luc+. Potential enhancer elements also can be inserted upstream of the promoter or in the BamHI or SalI sites downstream of the luc+ gene.
Figure 3-4 Generation of the ALAS1 5'-UTR – Firefly expression vectors.
The major and minor ALAS1 5'-UTRs were amplified from HepG2 cDNA from the major hepatic transcription site in exon 1A to the initiating ATG in exon 2. Following TA cloning into pGEM-T, the ALAS1 fragments were excised with appropriate restriction enzymes and cloned into both pGL3-P and pGL3-B at the 5'end of the firefly luciferase gene. The ALAS1 sequences essentially formed the 5'-UTR of the luciferase gene, with the initiating codon derived from the NcoI site in the antisense P7 primer.
3.2.3 Transfections and Dual Luciferase Reporter Assay

Transfections were carried out using Lipofectamine 2000 (Invitrogen). Briefly, HepG2 cells were trypsinized and seeded in 12-well plates at a density such that they were 80% confluent on the following day when they were transfected. pGL3-derived experimental constructs were used at 1μg per well and co-transfected with the Renilla luciferase encoding pRL-CMV plasmid (0.1μg) for normalisation purposes (figure 3-5). Transfected cells were lysed using Passive Lysis Buffer (PLB) and the luciferase activity was monitored in the lysed cell extracts by using the DLRA using a Lumat LB 9507 Luminometer (EG&G Berthold, Bad Wildbad, Germany).

![Figure 3-5 Vector map of pRL-CMV vector from www.promega.co.uk.](image)

The pRL-CMV Vector is a Renilla Luciferase encoding vector, and is intended for use as an internal control reporter and may be used in combination with any experimental reporter vector to co-transfect mammalian cells. pRL-CMV contains a cDNA (RLuc) encoding Renilla luciferase, which was originally cloned from the marine organism Renilla reniformis. The pRL-CMV Vector contains the CMV enhancer and early promoter elements to provide high-level expression of Renilla luciferase in co-transfected mammalian cells.
3.3 Results and Discussion

3.3.1 The ALAS1 minor isoform is resistant to haem-mediated decay.

The alternative splicing of the human ALAS1 gene generates two mRNA species: a major one, in which exon 1B is omitted, and a ubiquitously expressed minor one containing both exons 1A and 1B, accounting for only 10% of the steady-state concentration of ALAS1 mRNA (Roberts & Elder, 2001).

To determine whether the major and minor ALAS1 mRNA isoforms were destabilised similarly when exposed to haem, HepG2 cells were subjected to treatment with 5μg/ml actinomycin D to induce transcriptional arrest, and 20μM haemin (or control medium) for up to 8 hours. At specific time points, total RNA was extracted from the cells using Ultraspec. RT-PCR was quantified on all the samples to simultaneously amplify the major and minor ALAS1, and separately amplify GAPDH.

Figures 3-6 and 3-7 show that after the addition of exogenous haem, levels of the major ALAS1 mRNA (lower band) had fallen to 15% of their original levels by 8 hours. This is unsurprising, as it is known that ALAS1 mRNA is destabilised by haem in hepatocytes. In untreated cells, however, the level of the major ALAS1 mRNA had only dropped to 50%. The major ALAS1 mRNA half-life, calculated from the best-fit curve (figure 3-7), was reduced more than two-fold from 6.5h in untreated cells to 2.9h in cells exposed to haem.

In contrast to the major ALAS1 isoform, levels of the minor ALAS1 mRNA (upper band) remained essentially unchanged throughout the duration of the experiment in both the haem-treated and untreated control cells. Therefore, unlike the major ALAS1 form in HepG2 cells, the minor form is resistant to haem-mediated decay, under these experimental conditions. After 8 hours the mRNA of the minor and major ALAS1 isoforms had both decayed, even without haem. This is most probably due to the actinomycin D added to inhibit transcription, and therefore eventually becoming toxic to the cells. Furthermore, a peak in the mRNA concentration is observed in both isoforms at the timepoint t_1. Again, this effect may be attributed to the actinomycin D, rather than a haem effect.
Figure 3-6 ALAS1 Major mRNA is destabilised by haem, whereas the minor ALAS1 mRNA remains relatively stable.

a) Differential detection of endogenous ALAS1 mRNA was achieved by using different primer sets as shown in figure 3-2. HepG2 cells were either treated with 20μM haem or control medium for up to 8 hours, with actinomycin D (5μg/ml). Total RNA was extracted at the indicated time points and RT-PCR was used to assess the relative concentrations of ALAS1 mRNA transcripts. The bottom panel shows the results of cells exposed to haem and the upper panel those without haem.

b) Equivalent total cDNA concentrations between samples were confirmed by amplification of GAPDH cDNA. Experiments were carried out in triplicate on three separate occasions with similar results achieved each time. The figure is from one representative experiment.
Figure 3-7 The destabilisation of the ALAS1 major mRNA with and without the addition of haem for up to 8 hours.
The densitometry of the ALAS1 major mRNA levels at specific time-points during an eight hour period were normalised to the GAPDH levels, using arbitrary units. The half-lives of the major ALAS1 with and without haem were calculated to be approximately 2.9 and 6.5 hours respectively. mRNA decay, is a stochastic process, and therefore the change in mRNA concentration at any time point is a first-order process, depending on the amount of mRNA present at that time. The graph has two linear axes, showing that the rate of change of mRNA concentration changes over time and is proportional to the amount of mRNA present at that time-point. Data is from one representative experiment.

Decay curves for the major ALAS1 mRNA were plotted at equivalent GAPDH concentrations and $t_{1/2}$ measurements calculated from the best-fit slope through the data. No calculations were made for the minor ALAS1 mRNA since its stability was relatively unaffected by haem.

An explanation for this contrasting stability between the major and minor ALAS1 isoforms is that haem regulation is mediated through a cis-acting element located within exons 1A and 2 of the 5'-UTR, which is disrupted in the minor ALAS1 mRNA due to the inclusion of exon 1B. This hypothesis is shown diagrammatically.
in figure 3-8. An alternative explanation is that the haem-insensitivity of the minor ALAS1 mRNA might be a result of reduced translation, since inhibition of protein synthesis has been shown to prevent haem-mediated ALAS1 mRNA decay (Drew & Ades, 1989; Ryan & Ades, 1991), whilst alternatively spliced non-coding exons are known to inhibit downstream translation (Meijer & Thomas, 2002).

**Figure 3-8 Major and Minor ALAS1 5′-UTRs.**
Diagram of exons 1A - 2 in the major and minor forms of ALAS1, and a putative haem mediated cis-acting element across exons 1A and 2 in the major 5′-UTR. This would be disrupted by the exon 1B in the minor 5′-UTR.
3.3.2 The ALAS1 5'-UTR does not harbour a destabilising element, but the minor isoform inhibits expression of the downstream reporter gene.

It is hypothesised that exon 1B disrupts a cis-acting element, and therefore the minor ALAS1 is not destabilised by haem. To test this, heterologous luciferase reporter assays were used to determine if both the ALAS1 5'-UTRs were able to confer haem-mediated sensitivity on the downstream luciferase mRNA.

The major and minor 5'-UTRs were cloned immediately upstream of the firefly luciferase cDNA in the pGL3P vector to produce two reporter constructs, designated pGLALASmaj and pGLALASmin respectively, which were transfected into HepG2 cells. 24 hours after transfection, the cells were either exposed to 20μM haem or control medium for 24 hours, before being lysed to measure bioluminescent activity in the DLRA. An incubation of 24 hours with haem seemed a suitable time period to use, as the destabilisation of ALAS1 mRNA levels in cells by haem is long-lasting. Studies by Cable et al (2000) have shown that when CW5V17 cell cultures are treated with 10μM haem, a statistically significant reduction in ALAS1 mRNA levels are observed after 30 minutes, and maximal reduction is achieved by 45 minutes. These levels remain significantly reduced for periods longer than 18 hours (Cable et al., 2000).

Figure 3-9 shows the firefly luciferase activity of the pGL3P-ALAS1 constructs and the parental vector, normalized to the Renilla luciferase activity from the pCMV-RL vector, in HepG2 cells, after exposure to 20μM haem or control medium for 24 hours. Haem exposure did not significantly affect the normalised firefly luciferase activity produced by either of the two ALAS1 constructs or the parental pGL3P vector. These results demonstrate that a haem-mediated cis-acting element, which can function independently to the rest of the ALAS1 sequence, does not reside in the major or minor ALAS1 5'-UTR. However, the data shows that there were significant differences in firefly luciferase activity between the constructs. Normalised firefly luciferase activity in HepG2 cells transfected with the pGLALASmaj construct was 4.7-fold higher than cells transfected with the parental pGL3P vector. In contrast, normalised firefly luciferase activity in cells transfected with the pGLALASmin construct was nearly 2-fold less than those transfected with pGL3P. Therefore, the minor 5'-UTR reduced the expression of the downstream firefly luciferase approximately 9-fold, compared to the major 5'-UTR.
An explanation for this difference in expression between the major and minor constructs could be that either there was variation in the firefly luciferase expression or that the minor 5'-UTR inhibits translation of the downstream mRNA.

Figure 3-9 Effect of ALAS1 mRNA 5'-UTRs upon firefly luciferase activity and response to haem

The pGL3P, pGLALASmaj and pGLALASmin firefly luciferase constructs were separately co-transfected with the pRL-CMV Renilla luciferase vector into HepG2 cells. Luciferase activity was measured using the Dual Luciferase Reporter Assay, 48 hours after transfection. Haem was added for 24 hours before harvesting where indicated. The bars show normalised firefly luciferase activity to Renilla activity from pRL-CMV, relative to the pGL3P (minus haem) data, which was arbitrarily considered to be 1. Results represent mean (+ standard deviation) of at least three experiments each performed in triplicate. Statistical analysis was performed using a paired t test, and showed that that the differences plus and minus haem were not statistically significant.
3.3.3 The minor ALAS1 5'-UTR reduces translation of the downstream open reading frame.

RT-PCR was used to ensure that the differences between the firefly luciferase activities of the two ALAS 5'-UTR-firefly constructs were due to mRNA translation efficiency, and not variations in firefly luciferase mRNA expression. Cells were similarly transfected with the pGL3P constructs as in section 3.3.2, and treated with 20μM haem or control medium for 24 hours, after which the total RNA was extracted. Firefly luciferase and GAPDH mRNA were quantified by RT-PCR on cDNA samples derived from the total RNA extracted.

The RT-PCR analysis (figure 3-10) showed that there were no differences in firefly luciferase mRNA levels between cells transfected with the pGLALASmaj or pGLALASmin constructs. In both cases these were approximately 5-fold higher than those from the control pGL3-P-transfected cells. Therefore, insertion of either the major or minor ALAS1 5'-UTR does not affect transcription of the downstream firefly luciferase gene, despite there being a difference in the translation of the luciferase, as shown in the previous figure (figure 3-9).

![Diagram showing comparison of firefly luciferase and GAPDH expression in HepG2 cells transfected with pGL3P, pGLALASmaj, and pGLALASmin](image)

**Figure 3-10 Analysis of the firefly luciferase mRNA expression and activity in HepG2 cells transfected with the pGL3-P derived constructs.**
cDNA fragments for firefly luciferase (upper band) and GAPDH (lower band) were specifically amplified by RT-PCR using reverse transcribed (+) or control (no reverse transcription step) (-) DNase I-treated total RNA from transfected HepG2 cells. The image shown is a negative contrast for purposes of clarity. Expression of firefly luciferase mRNA increased from both the pGLALASmaj and pGLALASmin 5'-UTR constructs compared to the parental vector. Experiments were performed on three separate occasions in duplicate with similar results in each case.
Figures 3-9 and 3-10 demonstrate that in pGLALASmaj-transfected cells, the increased firefly luciferase mRNA expression results in a proportionate increase in the firefly luciferase activity. In contrast, however, the increased firefly luciferase mRNA expression in pGLALASmin-transfected cells does not result in an increase in the firefly luciferase activity. The firefly luciferase mRNA expression is similar in both the pGLALASmaj- and pGLALASmin-transfected cells, but the firefly luciferase activity is approximately 90% lower in the pGLALASmin-transfected cells. This data proves that the inclusion of exon 1B in the 5’-UTR of the minor ALAS1 mRNA inhibits the translational mechanism of the downstream firefly luciferase open reading frame. Therefore, in HepG2 cells, the minor ALAS1 mRNA is comparatively poorly translated under normal conditions.
3.3.4 Confirmation of ALAS1 5’-UTR transcriptional activity

Figure 3-10 showed that there is enhanced firefly luciferase mRNA expression observed with both the major and minor ALAS1 5’-UTR constructs, and not the parental pGL3-P vector. It was next examined if this was due to the 5’-UTRs both having intrinsic transcriptional activity, using the promoter-less pGL3-basic vector.

The major and minor 5’-UTRs were cloned into pGL3-basic vector, designated pBALASmaj and pBALASmin respectively. Total RNA was extracted from HepG2 cells transfected with the pGL3B constructs after 24 hours. Firefly luciferase and GAPDH RNA was quantified by RT-PCR of the cDNA samples derived from the RNA extracted. HepG2 cells transfected with these constructs and pRL-CMV for 24 hours were also used for the DLRA, to assess the firefly luciferase activity.

Densitometry derived from RT-PCR (figure 3-11a) showed that there was no substantial firefly luciferase mRNA production and consequently no luciferase activity in cells transfected with the parental pGL3-B vector (data not shown). However, figure 3-11b demonstrates that the ALAS1 5’-UTR constructs both produced firefly activity, but the activity from the pBALASmin construct was approximately 90% lower than that from the pBALASmaj construct.

This data confirms that translation of the minor ALAS1 transcript is inhibited compared to the major ALAS1. Furthermore, this demonstrates that transcription of the ALAS1 5’-UTR constructs in the promoter-less vector pGL3B must originate from sequences upstream of exon 1B, in exon 1A, which is common to both of the ALAS1 isoforms.
Figure 3-11 Analysis of the firefly luciferase mRNA expression and activity in HepG2 cells transfected with the pGL3-P derived constructs.

a) The cDNA fragments encoding for the major and minor 5'UTR were ligated into the pGL3-basic vector, to produce the pBALASmaj and pBALASmin constructs. In contrast to the parental pGL3B vector (not shown), both were able to drive transcription producing firefly luciferase mRNA. Values shown on the y axis are arbitrary units from densitometric analysis which are not shown. Results shown represent mean (+ S.D.) of at least three experiments performed in triplicate.

b) The pGL3B, pBALASmaj and pBALASmin firefly luciferase constructs were separately or separately co-transfected with the pRL-CMV Renilla luciferase vector into HepG2 cells. Luciferase activity was measured using the Dual Luciferase Reporter Assay, 48 hours after transfection. Haem was added for 24 hours before harvesting where indicated. The bars show the relative firefly activity of pBALASmaj (clear bar) and pBALASmin (solid bar), normalised to the activity of the Renilla luciferase vector, pRL-CMV. The activity of pBALASmin is reduced by nearly 90% compared to the pBALASmaj construct. Results shown represent mean (+ S.D.) of at least three experiments performed in triplicate.
3.3.5 Reduction in expression of downstream mRNA by the ALAS1 minor 5'-UTR is not restricted to hepatocytes.

To assess if this reduction of expression of the downstream firefly luciferase by the minor ALAS1 5'-UTR was restricted to HepG2 cells, the pGL3P derived constructs were also transfected into IMR32 and K562 cells. The results in figure 3-12 show that similar to the experiments in HepG2 cells, the expression of pGLALASmaj is approximately 4.3- and 6-fold higher than pGLALASmin in IMR32 and K562 cells respectively. Therefore it can be assumed that a similar mechanism that controls the ALAS1 expression in HepG2 cells occurs in these cell types. The intrinsic promoter activity in the major ALAS1 5'-UTR, derived from exon 1A, is active in both IMR32 and K562 cells, and not restricted to HepG2 cells.

![Graph showing relative firefly luciferase activity](image)

**Figure 3-12 A comparison of the firefly luciferase activity of the ALAS1 5'-UTR constructs in the IMR32 and K562 cell lines.**

The pGL3P, pGLALASmaj and pGLALASmin firefly luciferase constructs were separately co-transfected with the pRL-CMV *Renilla* luciferase vector into IMR32 and K562 cells. Luciferase activity was measured using the Dual Luciferase Reporter Assay, 48 hours after transfection. In contrast to the HepG2 experiments, no haem was added, as this was to compare the relative expression of the ALAS1 isoforms. The bars show normalised firefly luciferase activity relative to the pGL3P data, which was arbitrarily considered to be 1. Results represent mean (+ standard deviation) of at least three experiments each performed in triplicate. Statistical analysis was performed using a paired t test and shown not to be significantly different between the control and addition of haem.
3.4 Conclusions

In the liver and probably all other non-erythroid cells, haem supply is regulated primarily via feedback regulation of ALAS1 mRNA stability, yet the precise mechanism has not been elucidated.

There are three important observations from the data presented in this chapter:

1. Haem-dependent decay of ALAS1 is not dependent on the 5'-UTR of the mRNA.
2. The minor ALAS1 minor transcript is resistant to haem-mediated decay.
3. The inclusion of exon 1B of the minor 5'-UTR reduces translation of the downstream open reading frame.

The data produced confirms that the major human ALAS1 mRNA is destabilised by haem in HepG2 cells with half-life measurements of 2.9h and 6.5h, in the presence and absence of exogenous haem respectively. This two-fold decrease in the ALAS1 mRNA half-life is consistent with experiments by others. For example, quantitative RNase protection assays, following actinomycin D treatment, show that in HepG2 and Huh-7 cells, the addition of 20μM haem decreases the ALAS1 mRNA half-life from 8.8h to 3.5h and 6.0h to 3.6h respectively (Cable et al., 2001). In the rat immortalised cell line, CWSV17, the addition of haem decreases the half-life from 2.5h to 1.3h. In primary chick hepatocytes, haem decreases the mRNA half-life from 3.5h to 1.2h (Cable et al., 2000; Hamilton et al., 1991).

This data also demonstrates that the alternatively-spliced exon 1B in the 5'-UTR of human ALAS1 mRNA renders it resistant to haem-mediated decay. Our work with RNA assays consistently shows that the minor 5'-UTR inhibits downstream translation. This provides a crucial insight into the underlying mechanism by which haem destabilises ALAS1 mRNA.

Regulation of mRNA stability is an important mechanism in controlling gene expression, involving complex interactions between mRNA binding proteins and cis-acting sequence elements, most of which are located in the 3' or 5'-UTRs (Ross, 1995; Bevilacqua et al., 2003). The resistance of the minor ALAS1 mRNA to haem-mediated decay had suggested that the inclusion of exon 1B might disrupt an important regulatory element within the 5'-UTR. This was investigated and the data showed that the ALAS1 major 5'-UTR was unable to confer haem-
mediated destabilisation on its own. However, these results do not rule out there are elements within the 5′-UTR that interact with others elsewhere in the mRNA to bring about haem-mediated decay. Yet this appears unlikely, since there is poor sequence conservation between the 5′-UTRs of the major human, rat and chicken ALAS1 mRNAs (figure 3-13), and all of which are regulated by haem (Cable et al., 2001; Cable et al., 2000; Hamilton et al., 1991; Drew & Ades, 1989).

<table>
<thead>
<tr>
<th>Human</th>
<th>CAAGGCAGGAGGACGUUUCGUUGAACUUCGACUGUGCGCCUCUCUC---CUUCG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat</td>
<td>_______________________________________________________</td>
</tr>
<tr>
<td>Chicken</td>
<td>GCUGUUCGUUCCGCCCUCGUGGGGAGACAGCUGAGCGUCA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human</th>
<th>CCACGCACUCUCUGACUGCAAGCUUUCGACUUCACUGCACACCCUCUGGGACACCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat</td>
<td>_______________________________________________________</td>
</tr>
<tr>
<td>Chicken</td>
<td>CUUCCG----------------------------------------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human</th>
<th>CCACGGAGAAUCCUUCAGGACUGGACCCCCUCUGGACCCCCUCUGGGUUUA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat</td>
<td>_______________________________________________________</td>
</tr>
<tr>
<td>Chicken</td>
<td>G---UCGCCG------------------------------------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human</th>
<th>GGGGAUGUGGGGACCAGGAAAGUCAGAUCCCUAAAGAGUCUCCUCUGCCUGGAUGGA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat</td>
<td>_______________________________________________________</td>
</tr>
<tr>
<td>Chicken</td>
<td>_______________________________________________________</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Human</th>
<th>UGAGUGCUUCUUCUCCACCAGAUUUCUCCACAGGAGGCAAGCAGAUCUCUGGAACAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat</td>
<td>_______________________________________________________</td>
</tr>
<tr>
<td>Chicken</td>
<td>GU---AGCUCCGGAGGAGGAUGGA---------------------------------</td>
</tr>
</tbody>
</table>

**Figure 3-13 Sequences of the human, rat and chicken ALAS1 5′-UTRs.**
The sequences of the 5′-UTR of ALAS1 is poorly conserved between species, even though ALAS1 mRNA is always destabilised by haem. This indicates that a haem-mediated cis-acting element is not present in the 5′-UTR. Blue text represents exon 1B, which is only present in the minor form of human ALAS1 mRNA. Green text represents the first AUG of the coding region sequence of each species.

The minor 5′-UTR inhibited translation of the downstream heterologous mRNA and so an alternative hypothesis is proposed, where the resistance of the minor ALAS1 mRNA to haem-mediated decay results from its inhibited translation. This alternative theory is consistent with studies where inhibitors of protein synthesis, such as cyclohexamide, have been shown to stabilise ALAS1 mRNA and block
haem-mediated decay, leading to the suggestion that the haem effect requires a labile protein (Drew & Ades, 1989; Ryan & Ades, 1991). Whilst the data does not exclude the involvement of a labile trans-acting factor, it suggests strongly that haem-mediated ALAS1 mRNA decay requires translation of ALAS1 mRNA itself. Such regulated co-translational mRNA degradation has been observed for β-tubulin and c-myc (Bachurski et al., 1994; Lemm & Ross, 2002).

Translational control enables a cell to increase the concentration of a protein very rapidly and therefore appears to be especially suited to regulate genes implicated in cell proliferation and damage prevention. The majority of eukaryotic mRNAs are translated according to the scanning model of translation. A pre-initiation complex consisting of the ribosomal 40S subunit and several initiation factors (IFs) is assembled at the 5'-cap structure of mRNA (cap-dependent mechanism of initiation). This initiation complex then migrates along the 5'-UTR in search of an appropriate initiator AUG codon (scanning process). Recognition of an AUG leads to assembly of the 60S ribosomal subunit and initiation of protein synthesis (reviewed in Cazzola & Skoda, 2000).

The process of mRNA turnover is intimately linked to translation and numerous 5'-UTRs have been shown to regulate translation by interfering with efficient translation initiation at the physiological start site. These mRNA-specific characteristics include its size, GC content, secondary and tertiary structure, and the location of potentially active upstream open reading frames (uORFs) (Meijer & Thomas, 2002). For example, the translation of nitric oxide synthase has been shown to be regulated by an alternatively spliced 5'-UTR (Newton et al., 2003). It is thought that multiple UTR’s may regulate gene expression through differential expression with respect to cell type or stimulus, or their translational efficiency or stability may differentially modulate the amount of protein synthesised from their mRNAs.

The ALAS1 minor 5'-UTR sequence exhibits a number of features that might be responsible for its role in the reduced translation of the downstream mRNA. In most cases, 5'-UTRs that enable efficient translation are short, have a low GC content, are relatively unstructured and do not contain uAUG codons (Kochetov et al., 1998). The inclusion of exon 1B in the minor isoform creates a much longer 5'-UTR at 303nt than the 125nt major 5'-UTR. Whilst it is not particularly GC rich at 57% compared to exon 1A having a GC content of 64%, RNA folding algorithms predict that exon 1B forms a moderately stable putative stem-loop.
structure ($\Delta G = -38.3$ kcal), which might inhibit ribosomal scanning through it (figure 3-14). Secondary structures in the 5′-UTR of several mRNAs have been shown to repress translation of those mRNAs, presumably by stalling the ‘scanning’ of ribosomal initiation complexes (Kozak, 1989). Studies involving insertion of synthetic oligonucleotides that form stable hairpin structures within the 5′-UTR of pre-pro-insulin (Kozak, 1986) and thymidine kinase (Pelletier & Sonenberg, 1987) mRNAs and subsequent expression in COS cells demonstrated that excessive secondary structure in the 5′-UTR represses translation.

The extent to which scanning is affected is influenced greatly by the size or stability and position of the secondary structure relative to the 5′ cap. For example, a hairpin with a free energy of −30 kcal/mol situated close to the cap will impede scanning significantly. This is because secondary structure close to the cap can block access of the 43S pre-initiation complexes to the mRNA. Further from the cap, hairpins with a free energy stronger then −50kcal/mol are required to inhibit the progress of the ribosome. Such stem-loops are stable enough to resist the unwinding activity of eIF4A, the helicase component of the eIF4F complex. Although less stable stem loop can be melted by the helicase, such structures can be further stabilised by the interaction with RNA-binding proteins (Cazzola & Skoda, 2000). Therefore, the inclusion of exon 1B in the minor ALAS1 isoform may inhibit the translation of the ALAS1 protein.
Figure 3-14 The predicted secondary structure of exon 1b in the 5’-UTR

Exon 1B is only present in the minor form of the human ALAS1 gene, in its 5’-UTR. The RNA folding software mfold (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi) predicts that this exon 1B produces a stem-loop structure with a free energy of -38.3kcal/mol.

In addition to producing a stem-loop structure in the mRNA, exon 1B introduces four potentially active, upstream open reading frames (uORFs): uORFs 1, 2, and 4 are located exclusively within exon 1B whilst uORF3 initiates within exon 1B and terminates 2nt upstream of the ALAS1 initiating codon in exon 2 (figure 3-15).
Figure 3-15 Upstream Open Reading Frames (uORFs) located within the 5'-UTR of the minor ALAS1 mRNA

Exons 1A and 2 are shown in lower case, whereas Exon 1B is in upper case. Four uORFs are indicated, and have been designated U1 to U4. Potential initiating AUG codons and termination codons are given in bold letters. U1, U2 and U4 all initiate and terminate within exon 1B, whilst U3 starts in exon 1B and terminates 2 nucleotides upstream of the start of the major ALAS1 ORF in exon 2.

Of the uORFs, only 153nt U1 is in frame with the downstream major ALAS1 ORF separated by 54nt of mRNA. The U4 AUG is 1nt upstream of and overlaps the termination codon of U1. The ALAS1 ORF is shown at +1 with all nucleotide positions numbered accordingly.

A typical eukaryotic mRNA contains one major ORF in its 5'-UTR, the sequence between the methyl-guanosine cap and the initiation codon of the ORF, and is typically from 20 to several hundred nucleotides in length. Ribosomes generally begin translation at the first AUG in the mRNA, which encodes the major ORF. The sequence content around an AUG is important in determining whether it is utilised as a translation start site. Upstream open-reading frames are unusual in mammalian mRNAs. Fewer than 10% of the known mammalian genes have upstream open reading frames (uORFs) in their 5'-UTR (Kozak, 1987). Although uncommon, uORFs are found with higher frequency in certain gene families with regulatory importance, including homeobox genes, proto-oncogenes, growth factors, transcription factors, and genes encoding components of signal
transduction pathways. uORFs play important roles in translational regulation of gene expression, either increasing or inhibiting translation of the downstream ORF. The occurrence of uORFs in evolutionary divergent eukaryotes, including yeast, plants, mammals, and animal viruses, suggests that regulation by ORFs is an important mechanism (Morris & Geballe, 2000).

UORFs can have a number of effects on translation. For example, translation of a uORF may affect the translation of the downstream ORF by acting on the translation machinery or the RNA transcript. Both upstream AUGs and secondary structures within the 5′-UTR cause ribosome stalling, the ribosomes spend more time upstream or at the uORF, which may make fewer ribosomes available at the downstream ORF, as in the human ADH5/FDH gene (human x-alcohol dehydrogenase/formaldehyde dehydrogenase) (Meijer & Thomas, 2002; Kwon et al., 2001).

At present, it has not been determined which of these mechanisms might account for the reduced translation. Research in this laboratory (unpublished) has shown that mutagenesis of the uORFs 1, 2 and 4 does not inhibit the repression of translation. Translation of the downstream mRNA is still reduced when these uORFs are mutated. Therefore, it seems likely that it is the secondary structure that retards ribosomal scanning and affects the translation efficiency of the mRNA, rather than the presence of the uORFs.

The functional significance of alternative splicing in the human ALAS1 gene is not known. It is surprising that it has evolved in humans at all as data indicates that the minor mRNA appears to contribute very little of the steady-state concentration of ALAS1 protein, compared to the major mRNA. There is a high sequence homology of the 5′-UTRs of humans and chimpanzees, including the splice donor and acceptor sites, and the four potential uORFs. The presence of putative functional uORFs raises the prospect that regulated translation of the minor mRNA might be involved in developmental or tissue-specific expression of ALAS1.

As a final experiment to this chapter, we also investigated whether the reduction of expression of the downstream luciferase gene by the minor ALAS1 5′-UTR functioned in other cell types, as well as HepG2 cells. Our data indicate that the intrinsic promoter activity in the major ALAS1 5′-UTR, derived from exon 1A, is active in both IMR32 and K562 cells, and not restricted to HepG2 cells.
CHAPTER 4

Delineation of the ALAS1 haem-mediated instability element
4.1 Introduction

Haem primarily controls its own biosynthesis in hepatocytes, and probably all other non-erythroid cells, by post-transcriptional regulation of ALAS1, the first enzyme in its biosynthetic pathway. Haem inhibits the import of the ALAS1 enzyme from entering the mitochondria where it functions (Yamauchi et al., 1980; Hayashi et al., 1983), and also destabilises the ALAS1 mRNA (Drew & Ades, 1989; Cable et al., 2001). However, the mechanism of this destabilisation is unknown. The experiments in chapter 3 demonstrated that the 5'-UTR does not harbour any haem-mediated cis-acting elements. Therefore, it is likely that haem destabilises the ALAS1 mRNA by the use of cis-acting elements either in its coding region or 3'-UTR. In all cell lines examined so far the addition of haem causes the ALAS1 mRNA half-life (t_{1/2}) to reduce by approximately 2-fold. For example, in HepG2 cells, the cell line used throughout this thesis, the addition of 20μM haem causes the t_{1/2} to decrease from 8.8h to 3.5h (Cable et al., 2001). Similarly, RT-PCR experiments in the previous chapter showed that the t_{1/2} of HepG2 ALAS1 mRNA decreased from 6.5h to 2.9h, with the addition of 20μM haem.

The major focus of this chapter was to delineate the exact cis-acting element that causes haem-mediated destabilisation of the ALAS1 mRNA, in HepG2 cells. The approach to identify these sequences was to use indirect methods with reporter constructs, to look at the levels of ALAS1-Renilla luciferase fusion protein in response to haem. These results were confirmed by direct analysis of the mRNA levels, utilising the techniques RT-PCR and RNase Protection Assay (RPA).

4.1.1 mRNA regulatory elements in the 3'-UTR and coding region

Regulation of mRNA decay rate is an important control point in determining the abundance of cellular transcripts. Decay rates of individual mRNAs differ extensively, whereas some mRNAs decay with half-lives that are 100-fold shorter than cellular generation times, others have half-lives spanning several cell cycles. Moreover, although the decay rates of most transcripts, such as those encoding housekeeping genes, are invariant, the half-lives of numerous mRNAs change markedly in response to environmental cues. These differences in mRNA decay rates have notable effects on the expression of specific genes and upon alterations in the transcription rate; unstable mRNAs will attain steady-state levels in the least amount of time (reviewed in Wilusz et al., 2001).
<table>
<thead>
<tr>
<th>Localisation of determinant on mRNA</th>
<th>Decay determinant</th>
<th>Functional class/function of mRNA</th>
<th>mRNAs</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3'-UTR</td>
<td>ARE</td>
<td>Immune regulators</td>
<td>Interleukins (e.g. IL-1, IL-2, IL-3)</td>
<td>Stoecklin et al., 2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interferons (e.g. IFN-α)</td>
<td>Ma et al., 1997</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GM-CSF</td>
<td>Carballo et al., 2000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TNF-α</td>
<td>Lai et al., 1999</td>
<td></td>
</tr>
<tr>
<td>Proto-oncogenes</td>
<td></td>
<td>c-myc</td>
<td>Chen &amp; Shyu, 1995; Wilson &amp; Brewer, 1999; Brewer, 1991; Schiavi et al., 1992</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c-fos</td>
<td>Schiavi et al., 1994</td>
<td></td>
</tr>
<tr>
<td>B-adrenergic receptor</td>
<td></td>
<td>B1 and B2 receptors</td>
<td>Pende et al., 1996</td>
<td></td>
</tr>
<tr>
<td>Regulated</td>
<td></td>
<td>VEGF - response to hypoxia</td>
<td>Levy et al., 1998</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PTH - response to calcium</td>
<td>Sela-Brown et al., 2000</td>
<td></td>
</tr>
<tr>
<td>Iron-responsive element</td>
<td>Maintenance of iron homeostasis</td>
<td>TFR</td>
<td>Rouault &amp; Klausner, 1997</td>
<td></td>
</tr>
<tr>
<td>C-rich element</td>
<td></td>
<td>α-globin</td>
<td>Kiledjian et al., 1997</td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>Miscellaneous</td>
<td>Histone stem-loop</td>
<td>Whitfield et al., 2000</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>IGF II stem-loop</td>
<td>Schepfer et al., 1995</td>
<td></td>
</tr>
<tr>
<td>Coding region</td>
<td>CRD</td>
<td>Miscellaneous</td>
<td>c-myc</td>
<td>Yeilding &amp; Lee, 1997; Prokipcak et al., 1994.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c-fos</td>
<td>Chen et al., 1992</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>β-tubulin</td>
<td>Bachurski et al., 1994</td>
<td></td>
</tr>
</tbody>
</table>

**Table 4-1 Examples of decay determinants located in the 3'-UTR and coding region, adapted from Guhaniyogi & Brewer, 2001.**

One of the most prevalent is the A+U-rich element (ARE), found in the 3'-UTRs of some mRNAs. There are several classes of ARE with slightly different sequence determinants that are characterised by their abilities to promote the rapid deadenylation and subsequent decay of a transcript. A subset of AREs can also mediate mRNA stabilisation in the presence of certain stimuli, illustrating a dual role for these regulatory elements (Wilusz et al., 2001).
The turnover of mRNA is a highly controlled process. Several sequence elements can regulate the rate of turnover of a transcript, either by promoting (destabiliser elements) or by inhibiting (stabiliser elements) decay. These can be located in the 5'- and 3'-UTRs, as well as the coding region of the mRNA. However, those coding region elements are rare and usually specific to the particular mRNA, as opposed to the more common and general elements found in the 3'-UTR (reviewed in Guhaniyogi & Brewer, 2001). Examples of cis-acting elements and their mRNA locations are listed in table 4-1. We have looked for these specific elements in the human ALAS1 sequence, and found that none of these are present in the coding region or 3'-UTR.

<table>
<thead>
<tr>
<th>Group</th>
<th>Motif</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>WAUUUAW and a U-rich region</td>
<td>c-fos, c-myc</td>
</tr>
<tr>
<td>IIA</td>
<td>AUUUAUUUAUUUAUUUAUUUA</td>
<td>GM-CSF, TNF-α</td>
</tr>
<tr>
<td>IIB</td>
<td>AUUUAUUUAUUUAUUUA</td>
<td>Interferon-α</td>
</tr>
<tr>
<td>IIC</td>
<td>WAUUUAUUUAUUUAW</td>
<td>Cox-2, IL-2, VEGF</td>
</tr>
<tr>
<td>IID</td>
<td>WWWWAAUUWAUWW</td>
<td>FGF2</td>
</tr>
<tr>
<td>IIE</td>
<td>WWWWAAUUAWWW</td>
<td>u-PA receptor</td>
</tr>
<tr>
<td>III</td>
<td>U-rich, non-AUUUA</td>
<td>c-jun</td>
</tr>
</tbody>
</table>

Table 4-2 The classification of AREs and their associated motifs, from Wilusz et al., 2001.

Group II contains reiterations of the AUUUA pentamer. Given the large sequence variation of the ARE, an abundance of ARE-binding factors have been identified.

Cis-acting elements that modulate transcript stability can also be found in the coding region of mRNAs. In some cases these elements act in concert with 3'-UTR elements to regulate mRNA decay, and can only function when both these elements are present. For example, the c-fos mRNA contains both an ARE in the 3'-UTR and two destabilising sequences in its coding region. The most studied is known as the major protein-coding-region determinant (mCRD). This mCRD contains 320 nucleotides, is located near the centre of the mRNA, and encodes the basic and leucine zipper regions critical to c-fos protein function. This cis-acting element is position-specific, and needs to be at least 450 nucleotides proximal to the poly(A) tail. Furthermore, it requires continuing translation for its destabilising function (Chen et al., 1992).

As well as in c-fos, functional mRNA stability determinants have been detected within the coding region of a growing list of mammalian mRNAs including the
mRNAs of c-Myc, vascular endothelial growth factor (VEGF), β-tubulin and u-PAR. CRDs are relatively rare in mRNA molecules, in comparison to the number of instability elements found in the 3'- or 5'-UTRs.

4.1.2 Putative ALAS1 cis-acting elements

Sequence analysis of the human ALAS1 mRNA shows that it lacks any known regulatory stability motifs. However, this is unsurprising, since the cis-acting element may be present in the coding region. These are often specific to the mRNA, and regulation is related to the function of the encoded protein. As ALAS1 mRNA stability is controlled by haem, a specific haem-responsive stability element within the mRNA coding region or 3'-UTR is required.

Cable et al (2001) previously identified a specific sequence of the ALAS1 coding region that can bind to cytosolic proteins extracted from Huh-7 cells, a human hepatoma cell line. This 154bp protein-binding region was deleted from the ALAS1 sequence to make a deletion construct, and transfected into Huh-7 cells, along with the full-length ALAS1 as a control. Their results showed that the level of exogenously added deletion construct stayed relatively constant with the addition of haem for 4 hours. Conversely, the level of the full-length construct decreased with the addition of haem (Cable et al., 2001). Therefore, it was concluded that this region is involved in the haem-mediated destabilisation of ALAS1, and may contain a CRD. However, whether this sequence is the sole requirement for haem-mediated destabilisation of the ALAS1 mRNA has not been shown. The experiments in this chapter set out to establish if a haem-responsive element is indeed present in this region, and can function independently of the rest of the ALAS1 sequence.
4.1.3 Aims and Strategy

The major aim of the work in this chapter was to delineate which ALAS1 sequences are responsible for haem-mediated decay in HepG2 cells. We have also assessed the role of haem in the destabilisation mechanism in a non-hepatic cell line.

This involved:

1. Developing reporter assays to indirectly assess the stability of the ALAS1 mRNA in response to haem.
2. Confirming if the 154bp sequence was the only determinant of ALAS1 stability.
3. Identifying other regions within the coding region or 3'-UTR which may be involved in the destabilisation of ALAS1 mRNA by haem.
4. Assessing the role of translation in this mechanism.
5. Confirming the results generated by the DLRA, using RNA-analytical techniques.
6. Investigating the affect of haem on the stability of ALAS1 in neuronal IMR32 cells.
<table>
<thead>
<tr>
<th>Renilla-ALAS1 Construct</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Length of ALAS1 insert (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pR154</td>
<td>P8</td>
<td>P9</td>
<td>153</td>
</tr>
<tr>
<td>pRSTOP154</td>
<td>P10</td>
<td>P9</td>
<td>156</td>
</tr>
<tr>
<td>pRALAS1</td>
<td>P11</td>
<td>P12</td>
<td>1923</td>
</tr>
<tr>
<td>pRSTOPALAS1</td>
<td>P13</td>
<td>P12</td>
<td>1926</td>
</tr>
<tr>
<td>pRALAS3UTR</td>
<td>P11</td>
<td>P14</td>
<td>2073</td>
</tr>
<tr>
<td>pR3UTR</td>
<td>P15</td>
<td>P14</td>
<td>150</td>
</tr>
<tr>
<td>pRA1</td>
<td>P11</td>
<td>P16</td>
<td>480</td>
</tr>
<tr>
<td>pRA2</td>
<td>P17</td>
<td>P18</td>
<td>480</td>
</tr>
<tr>
<td>pRA3</td>
<td>P19</td>
<td>P20</td>
<td>480</td>
</tr>
<tr>
<td>pRA4</td>
<td>P21</td>
<td>P12</td>
<td>483</td>
</tr>
<tr>
<td>pRA5</td>
<td>P22</td>
<td>P23</td>
<td>480</td>
</tr>
<tr>
<td>pRA6</td>
<td>P24</td>
<td>P25</td>
<td>480</td>
</tr>
<tr>
<td>pRA7</td>
<td>P26</td>
<td>P27</td>
<td>480</td>
</tr>
</tbody>
</table>

Table 4-3 Primer combinations used for the cloning of ALAS1 sequences into the Renilla luciferase expression vector pRLuc-C2.

ALAS1 cDNA comprising the entire coding region of 1923bp, but lacking both 5’- and 3’-UTRs, was amplified using the forward primer P11, containing an Eco RI restriction site immediately prior to the initiating ATG codon, and the reverse primer P12, containing a Kpn I restriction site immediately after the stop codon utilised in the ALAS1 mRNA. As previously, an additional C nucleotide was added to the sequence in P11, after the Eco RI site, so that the ALAS1 coding region remained in frame with the Renilla luciferase gene in the pRLuc-C2 vector. In order to introduce a stop codon into the Renilla luciferase-ALAS1 cDNA fusion immediately prior to the ALAS1 moiety, a modified upstream sense primer, P13, was used with the reverse primer P12. This inserted a stop codon prior to the ATG, ensuring that the ALAS1 sequence is not translated when inserted into the pRLuc-C2 vector. The 2073bp DNA fragment, encoding the entire ALAS1 coding region together with the 3’-UTR was amplified using the P11 and P14 primers. To examine the effect of the 3’-UTR alone, a 150bp cDNA fragment corresponding to the ALAS1 mRNA 3’-UTR (positions 1997 - 2146; Genbank sequence, X56351) was amplified using the forward primer P15, and the reverse primer P14.

In order to further delineate the ALAS1 cis-acting element, seven smaller overlapping fragments of the ALAS1 coding region were also amplified by PCR, each of approximately 480bp in length, using the primers P11-P12 and P16-P27. Each forward primer contained the restriction site Eco RI, and the reverse primer contained the restriction site Kpn I, to aid cloning into the pRLuc-C2 vector.
All amplified ALAS1 products were purified from agarose gels and ligated into pGEM-T by TA cloning. The ligation mixture was used to transform subcloning efficiency, competent DH5α cells, and the plasmid DNA was isolated and purified. To produce the pRLuc-ALAS1 fusion constructs, the pGEM-T-ALAS1 plasmid DNA was digested with the restriction enzymes Eco RI and Kpn I. The ALAS1 insert was purified using the Qiagen gel extraction kit and subsequently ligated into pRLuc-C2 (figure 4-1) that had been previously digested with Eco RI and Kpn I. Unlike the other constructs, the sequence for the pR154 construct was amplified using the restriction site Eco RI in both the forward and reverse primers. Therefore, these were inserted into the pRLuc-C2 vector which had been digested with Eco RI alone. All cloned sequences were verified by DNA sequencing. The ALAS1 coding region and 3′-UTR constructs are shown diagrammatically in figure 4-2.

![Figure 4-1 Vector Map of pRLuc-C2 adapted from www.perkinelmer.com.](image)

The codon humanized pRLuc-C vectors contain a multiple cloning site (MCS) located downstream of the codon humanized Renilla Luciferase gene (Rluc(h)), for the subcloning of a gene of interest to create a *Renilla* luciferase fusion protein. The Rluc codons have been humanized to ensure higher expression levels of the fusion protein in mammalian cells. The fusion protein gene is placed under the control of the cytomegalovirus (CMV) promoter, ensuring a very high constitutive expression in a variety of cells.
Figure 4-2 Construction of the full-length and 3′-UTR Renilla-ALAS1 constructs.

- Renilla luciferase cDNA in pRLuc-C2
- ALAS1 coding region cDNA
- ALAS1 3′-UTR cDNA

E= EcoRI site; K=KpnI site; X=STOP site inserted by modified primer

The ALAS1 sequences were cloned in-frame and downstream of the Renilla luciferase gene into the Eco RI / Kpn I sites within the MCS of pRLuc-C2. For the pRSTOPALAS construct, a stop codon was inserted prior to the ALAS1 sequence, with a modified forward primer.

4.2.2 Mammalian Cell Transfection

For the DLRA experiments, confluent HepG2 and IMR32 cells were split into 12-well plates, with 1ml cells/well, in cDMEM 24 hours prior to transfection. The cells were transfected with 1µg/well of each the control vector pRLuc-C2 and all other constructs to be tested in triplicate, and co-transfected with 630ng/well of pGL3-P vector. 24 hours after transfection, the cells were treated either with 20µM haem or control medium (vehicle) in triplicate. Luciferase activity was measured 24 hours after incubation with haem/vehicle.

For RNA analysis, confluent HepG2 cells were split into 6-well plates in cDMEM, with 2ml cells/well, so that sufficient RNA was available for cDNA synthesis. Each well of cells were transfected with 2µg of pRALAS1 and 1.26µg of pGL3-P for normalisation purposes, where appropriate.
4.2.3 RNA Quantification by RNase Protection Assays (RPA)

To confirm the DLRA results, by directly analysing the ALAS1 mRNA levels in response to haem, RPA was used. This involved incubation of total RNA (extracted from transfected HepG2 cells, and treated with 20µM haem where necessary, for 6 hours) with labelled probe, so that the sense RNA, complementary to the antisense probe, will hybridise. After which, any free probe and other single-stranded RNA were digested with RNases. The remaining "RNase-protected" probes are purified, resolved on denaturing polyacrylamide gel, and visualised by blotting and UV-crosslinking onto a membrane. The quantity of each mRNA species in the original RNA sample can then be determined based on the intensity of the appropriately sized, protected probe fragment.

4.2.3.1 Biotin-labelled antisense RNA probes

Probe constructs were generated for RPA analysis of *Renilla* luciferase-ALAS1 fusion mRNA and GAPDH mRNA, as a control, by amplification of cDNA fragments from recombinant pRALAS1 plasmid DNA (P28 and P29 primers) and HepG2 RNA (P30 and P31 primers) respectively (figure 4-3). These were ligated into the pGEM-T vector and clones in an orientation to generate antisense RNA using the T7 promoter were selected. The probe for ALAS1 used a sequence that overlapped the *Renilla* luciferase and ALAS1 boundary, to ensure that only the pRALAS1 construct was detected by RPA, and not the endogenous ALAS1. These probes were made as antisense RNA, so that during the RPA they hybridised to the corresponding sense RNA sequence in the HepG2 total RNA extract, thereby protecting this region from degradation by RNase.

The *Renilla* luciferase-ALAS1 and GAPDH probe constructs were linearised by digestion with *Kas* I, which cuts in the *Renilla* luciferase sequence, and *Spe* I which cuts within the SP6 site of the pGEM-T vector, respectively. Antisense biotin-16-UTP probes for *Renilla* luciferase-ALAS1 mRNA (328nt) and GAPDH mRNA (258nt) were generated by *in vitro* transcription using the T7 maxiscript system (Ambion) in accordance with the manufacturer's protocol (section 2.4.11). Labelled RNA probes were gel-purified following electrophoresis through a 5% acrylamide/8M urea denaturing polyacrylamide gel.
Figure 4-3 Schematic diagram to show the generation of ALAS1 and GAPDH antisense RNA probes for RPA.

Renilla-ALAS1 and GAPDH sequences were amplified by PCR and inserted into the PGEM-T vector by TA cloning, in an antisense (3'-5') direction from the T7 start site. The constructs were excised with Kas I and Spe I for the ALAS1 and GAPDH probes respectively. The T7 maxiscript kit (Ambion) was used to produce antisense biotin 16-UTP probes by in vitro transcription. The labelled RNA probes were gel-purified following electrophoresis through a 5% acrylamide/8M urea denaturing polyacrylamide gel, to ensure that the probes were of the correct size and that any degraded probe or unincorporated dNTPs were removed.

4.2.3.2 RPA

HepG2 cells were co-transfected with pRALAS1 and pGL3P control vector, as in section 4.2.2. After 24h the medium was removed and replaced with DMEM, which contained 20µM haemin, or vehicle medium as appropriate. At t₀, cells in duplicate where trypsinised, harvested and stored at -70ºC. Following a further 6h incubation with either haem or control medium, these cells in duplicate were also trypsinised and harvested. A small fraction of the cells were lysed using Passive Lysis Buffer and luciferase activity monitored in lysed cell extracts by a DLRA, whilst the rest were treated with Ultraspec to extract total RNA. To account
for possible variations in transfection efficiency, the amount of RNA (between 3-10μg) used in the RPA was determined according to the firefly luciferase activity in each sample. The extracted RNA was then treated with DNase I to remove any residual contaminating plasmid DNA.

For each RPA sample, 250pg of each pRALAS1 and GAPDH probe were used, corresponding to a ~5-fold molar excess of each target RNA. GAPDH is a relatively abundant target mRNA that makes up approximately 0.1% of all mRNA, therefore 250pg was calculated to be enough for a 5-fold molar excess. Hybridisation and digestion steps were carried out using a RPA III kit (Ambion) in accordance with the manufacturer's instructions. Hybridised, digested samples were resuspended in 6μl gel loading buffer and heated to 95°C for 3 minutes to denature any secondary structure and solubilise the RNA. Samples were electrophoresed on a 5% acrylamide/8M urea denaturing gel at 25mA, and the RNA was electroblotted onto a BrightStar-Plus membrane (Ambion) using a Panther HEP-1 Semi-Dry Electrophotter (OWL Scientific, Portsmouth, NH, USA) for 30 minutes at 400mA, and immobilised by UV crosslinking at 120,000μJ/cm² for 30 seconds. Detection was performed following the BrightStar BioDetect protocol (Ambion).

4.2.4 Semi-Quantitative RT-PCR

4.2.4.1 RT-PCR of pRALAS1 and pGL3P

mRNA was isolated from pRALAS1 and pGL3P transfected HepG2 cells (section 4.2.2) using Dynal mRNA Direct Kit. cDNA was synthesised from the mRNA by first strand synthesis, and changes in mRNA expression were assayed by PCR using primer sets specific for cDNA fragments of ALAS1, with pGL3P as a control (tables 4-4 and 4-5). The primers used for ALAS1 amplified a sequence that overlapped the Renilla luciferase and ALAS1 boundary, to ensure that only the pRALAS1 construct was detected, and not the endogenous ALAS1. The cycling protocol for the pRALAS1 PCR was that as described in section 2.4.4.3.1, with 34 cycles, and an annealing temperature of 60°C. The pGL3P PCR was similarly of 34 cycles, with an annealing temperature of 65°C. Prior to the quantification by PCR, all the PCR conditions had been optimised for logarithmic amplification using untreated control cells (see section 3.2.1.1).
<table>
<thead>
<tr>
<th>Target DNA</th>
<th>Primer Number</th>
<th>Product Size</th>
<th>PCR Annealing Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRALAS1</td>
<td>Sense P28, Anti-Sense P29</td>
<td>537bp</td>
<td>60°C</td>
</tr>
<tr>
<td>pGL3P</td>
<td>Sense P32, Anti-Sense P33</td>
<td>267bp</td>
<td>65°C</td>
</tr>
</tbody>
</table>

Table 4-4 Primers used in semi-quantitative RT-PCR.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PCR buffer</td>
<td>2</td>
</tr>
<tr>
<td>2mM dNTP mix</td>
<td>2</td>
</tr>
<tr>
<td>Sense Primer 5pmol</td>
<td>5</td>
</tr>
<tr>
<td>Antisense Primer 5pmol</td>
<td>5</td>
</tr>
<tr>
<td>cDNA Template (RT product)</td>
<td>2</td>
</tr>
<tr>
<td>HotStarTaq Polymerase (5U/μl)</td>
<td>0.2</td>
</tr>
<tr>
<td>H₂O</td>
<td>to 20μl</td>
</tr>
</tbody>
</table>

Table 4-5 Composition of semi-quantitative RT-PCR.

PCR products were resolved by agarose gel electrophoresis. The band intensities of the products were quantified as raw volume of the peak intensity, using GeneTools software (Syngene), and normalised to GAPDH to calculate the relative induction.

4.2.4.2 RT-PCR of endogenous haem oxygenase-1

Total RNA was isolated from HepG2 cells using Ultraspec, and cDNA was synthesised from the mRNA by first strand synthesis. A change in haem oxygenase-1 (HO-1) mRNA expression was assayed by PCR using the forward primer P34, corresponding to a sequence in exon 2, and the reverse primer P35, corresponding to a sequence in exon 3. The PCR product, of 400bp, therefore spanned an exon boundary and ensured that only HO-1 cDNA was amplified, rather than any contaminating genomic DNA. The PCR was quantified as in table 4-5.
4.3 Results and Discussion

4.3.1 The cis-acting element in the 154bp region does not confer haem-mediated destabilisation.

Previous work by Cable et al has described a 154 nucleotide region that is able to bind cytosolic proteins, and therefore it was suggested that this region contains a haem-responsive cis-acting element (Cable et al., 2001). To test this hypothesis we used reporter assays, with a commercially available Renilla luciferase expression vector, pRLuc-C2. This vector permitted in-frame cloning of ALAS1 cDNA sequences downstream of the Renilla moiety, such that it could be determined whether these conferred haem-mediated destabilisation upon the fusion mRNA, by monitoring effects upon Renilla luciferase expression. In these studies, Renilla luciferase-ALAS1 fusion constructs were transfected into the human hepatoma cell line, HepG2.

A fusion construct was made with the putative cis-acting element of the coding region, to generate a sequence encoding 51 in-frame codons, as described in the materials and methods. Another construct was made with a stop codon (TAA) prior to this sequence to assess whether the sequence needs to be translated to function (pRstop154). The ALAS1 sequences were cloned in-frame to generate a fusion reporter gene construct such that the luminescent signal detected from the Renilla luciferase reports on the stability and translation efficiency of the ALAS1 sequence.

These constructs and the control pRLuc-C2 vector were transfected separately into HepG2 cells (1μg/well of cells each), together with the firefly luciferase expression plasmid, pGL3-promoter vector (630ng/well of cells) in order to normalise the Renilla luciferase data due to any differences in transfection efficiency. Cells were treated with 20μM haem or the vehicle medium, 24h after transfection. The luciferase activities were measured after 24h following exposure to 20μM haem.
Figure 4-4 The addition of haem does not cause the destabilisation of the ALAS1 153bp – Renilla luciferase fusion protein.

The putative cis-acting element of the ALAS1 coding region was cloned in frame to the pRLuc-C2 vector at the Eco RI site, with and without a stop codon prior to the sequence. Transfection and bioluminescent measurements were performed as previously described. There was no difference between relative Renilla expression of pR154 and pRstop154-transfected cells, treated and untreated with 20µM haem. Bars indicate plus standard deviation. Experiments were on two separate occasions in triplicate.
As expected, *Renilla* luciferase expression from the pRLuc-C2 vector alone did not significantly differ between untreated cells and those treated with haem, indicating that neither the activity of *Renilla* luciferase itself nor the CMV promoter driving its expression was affected by increased levels of intracellular haem. Consequently, data from the ALAS1 constructs were normalised to the value of the pRLuc-C2 treated cells, which was considered to be 100%. Data from these experiments showed that in both of these constructs the *Renilla* luciferase expression was not reduced with haem treatment (figure 4-4). Although this region has been shown to bind cytosolic proteins in hepatocytes, this data demonstrates that this region alone is not able to confer haem-mediated instability of the mRNA. This suggests that either the 154bp sequence is not significant in the haem destabilisation mechanism, or another sequence elsewhere in the ALAS1 coding region is required for it to function. It is possible that the ALAS1 CRD has a bi-partite structure, which may rely on specific mRNA sequences within the coding region to enable its destabilisation by haem.
4.3.2 The human ALAS1 coding region contains a haem-mediated cis-acting element.

The data in chapter 3 eliminated the 5'-UTR as a region harbouring haem-responsive cis-acting elements. As section 4.3.1 showed, the haem-response element is not present in the 154bp sequence of the coding region previously identified by Cable et al (Cable et al., 2001). Therefore, the next set of experiments set out to establish whether the instability determinant is indeed in the coding region or in the 3'-UTR, and whether it needs sequence elements in both of these regions to function.

Three ALAS1 constructs were generated; one with the coding region added in-frame at the 3' end of the Renilla luciferase (pRALAS1), another containing the ALAS1 3'-UTR as well as the coding region (pRALAS3UTR), and a final construct containing just the ALAS1 3'-UTR (pR3UTR) appended to the luciferase gene (figure 4-2).

The ALAS1 sequences were cloned in-frame to generate a fusion reporter gene construct such that the luminescent signal detected from the Renilla luciferase reports on the stability and translation efficiency of the ALAS1 sequence. As for the previous section, these constructs and the control pRluc-C2 vector were transfected separately into HepG2 cells (1μg/well of cells each), together with the firefly luciferase expression plasmid, pGL3-promoter vector (630ng/well of cells) for normalisation. Cells were treated with 20μM haem or the vehicle medium, 24h after transfection. The luciferase activities were measured after 24h following exposure to 20μM haem.

Figure 4-5 demonstrates that Renilla luciferase expression from the ALAS1-fusion constructs were much lower than that of the pRluc-C2 vector by approximately 15-fold, presumably due to the addition of the long ALAS1 sequence at the C-terminus of Renilla luciferase gene. However, Renilla expression remained sufficiently high for accurate analysis. Data from each ALAS1 construct was normalised to the value of the untreated cells transfected with the same construct, which was considered to be 100%. Renilla luciferase expression from the pRluc-C2 vector alone did not significantly differ between untreated cells and those treated with haem. In contrast, in cells transfected with pRALAS1, exposure to haem decreased Renilla luciferase activity consistently and significantly by 40% of that seen in untreated cells. Similarly, haem exposure also resulted in a
significant 33% decrease in Renilla luciferase activity in cells transfected with pRALAS3UTR. Paired sample t tests were used to establish that these differences of Renilla activity with and without haem were statistically significant.

![Graph showing the relative Renilla luciferase activity in different conditions](image)

**Figure 4-5 ALAS1 mRNA is destabilised by a coding region determinant.**
The Renilla-ALAS1 fusion constructs and the empty parental vector were co-transfected into HepG2 cells with the pGL3P vector as the internal control. Cells were treated with 20µM haem or the vehicle medium for 24h, and luciferase expression was measured using the DLRA. Data shows the relative Renilla expression after the addition of 20µM haem, compared to the expression without haem as 100%. Bars indicate plus standard deviation. Paired sample t tests were performed on the data for pRALAS1 and pRALAS3UTR, and both p values are less than 0.05, indicating a significant difference between Renilla expression with and without haem treatment. There was no significant difference between the Renilla expression from pR3UTR with or without the addition of haem. Experiments were on three separate occasions in triplicate. *= statistically significant, p value < 0.05.

There was no significant difference between haem-depressed Renilla luciferase activity in pRALAS1 cells and pRALAS3UTR cells indicating that the cis-acting haem responsive elements reside within the coding region rather than the 3' UTR. To confirm this, a fusion construct was generated in which the ALAS1 3'UTR alone (pR3UTR), with no coding region, was appended to Renilla luciferase. Unlike the pRALAS3UTR construct, its Renilla luciferase activity was not significantly affected by haem, which substantiated the previous data and conclusions.
4.3.3 The coding region determinant has to be translated to mediate instability of the mRNA in the presence of haem.

The data in chapter 3 demonstrated that the minor ALAS1 mRNA is resistant to haem-dependent degradation compared to the major form. Furthermore, our data suggested that translation of the minor mRNA is reduced approximately eight-fold compared to the major mRNA. This data indicates that ALAS1 mRNA stability is related to its rate of translation. Therefore, having established both that the haem-regulatory element(s) resided within the ALAS1 coding region and that the poorly translated minor ALAS1 mRNA was resistant to haem-mediated degradation, the translational mechanism through the ALAS1 coding region may be required for its haem-mediated destabilisation.

Accordingly, an additional fusion construct was made, designated pRSTOPALAS1, in which a stop codon (TGA) was inserted between the Renilla luciferase cDNA and the ALAS1 coding region such that the latter would constitute part of the 3’-UTR and not be translated.

This construct was transfected into HepG2 cells with pGL3P, as in the previous sections. In contrast to the reduced Renilla luciferase activity seen in haem-treated cells transfected with either pRALAS1 or pRALAS3UTR constructs, normalised Renilla luciferase activity in pRSTOPALAS1-transfected cells was not reduced in response to haem treatment but was in fact slightly elevated, though not significantly so, compared to that in untreated cells (paired samples t test, p > 0.05) (figure 4-6). This result indicates that translation through the coding region instability determinant is required for its haem-mediated destabilisation in HepG2 cells.
**Figure 4-6** The coding region determinant has to be translated to mediate instability of the mRNA in the presence of haem.

The pRstopALAS1 construct was made by introducing a TGA stop codon between the Renilla luciferase gene and the ALAS1 coding region by PCR. Transfection and bioluminescent measurements were as described in figure 4-5. Bars indicate plus standard deviation. There was no significant difference between the relative Renilla luciferase activity of pRSTOPALAS1 with and without haem. Experiments were on three separate occasions in triplicate.
4.3.4 Overlapping deletion experiments to delineate the ALAS1 mRNA coding region determinant(s).

To date, no sequences of ALAS1 mRNA have been shown to bind cytosolic protein, apart from the 154bp sequence at the 3’ end of the coding region. Section 4.3.1 demonstrated that this region is not destabilised by haem when used in reporter assays, and therefore this region cannot function alone as a haem-mediated CRD. Additional sequence elements adjacent to the previously identified 154 nucleotide region or other additional elements in the coding region may be needed for haem-mediated instability of the mRNA.

In order to define where other CRDs reside, seven Renilla luciferase-ALAS1 fusion constructs consisting of overlapping smaller fragments of the ALAS1 coding region were made (figure 4-7). These all contained ALAS1 sequences of approximately 480bp each, designated pRA1-pRA7. The putative cis-acting element, identified by Cable et al., fully resides in the construct pRA4, and partly resides in pRA7 (Cable et al., 2001). The sequences were cloned into the pRLuc-C2 vector, co-transfected with pGL3P into HepG2 cells as in the previous sections, and treated with either 20µM haem or the control medium.

![ALAS1 coding region = 1923bp](image)

**Figure 4-7** The seven overlapping sub-fractions of the ALAS1 coding region that were amplified by PCR, and inserted into the pRLuc-C2 vector.

Seven regions of the ALAS1 protein-coding region were amplified by PCR, located at the above nucleotides, designated pRA1 - pRA7. The seven overlapping fragments were amplified from HepG2 total RNA, using Bio-X-Act polymerase. These were cloned into the pGEM-T vector, digested with Eco RI and Kpn I, and cloned into the similarly digested pRLuc-C2 vector, to produce the constructs pRA1 - pRA7.
**Figure 4-8 The instability element of ALAS1 does not reside in a specific fragment of its coding region mRNA.**

Seven fragments of the ALAS1 coding region of approximately 480bp in length were cloned in frame to the pRluc-C2 vector. These were transfected into HepG2 cells, treated with 20μM haem or the vehicle medium for 24 hours, and assayed for bioluminescence as previously described. Unlike the previous graphs, the data was not normalised. Bars indicate plus standard deviation. Experiments were on two separate occasions in triplicate. The pRA4 sequence contains the 154 base pair region, which is able to bind cytosolic proteins.

The relative Renilla luciferase expression levels varied greatly between the seven ALAS1 constructs and repeated attempts to demonstrate haem sensitivity on these sequences was unsuccessful. The full-length ALAS1 construct, pRALAS1, was used as a positive control, and on each occasion was shown to be destabilised by haem (data for pRALAS1 not shown here, but demonstrated in figure 4-5). Therefore, these results cannot conclude which fragment the CRD is in.
Due to the unusually high differences between the *Renilla* luciferase expression of each construct, the actual values of relative *Renilla* luciferase expression data to pRLuc-C2 was plotted, showing the un-normalised data. Unlike the previous graphs in this study, figure 4-8 shows the *Renilla* expression of the constructs treated with the vehicle medium as its actual value, rather than 100%. This emphasises the variation of *Renilla* luciferase expression between each construct. Interestingly, the expression levels of the ALAS1 fragments closer to the 3’ end of the coding region are much lower that those at the 5’ end, with the *Renilla* expression of pRA4 being only approximately 5% of that of pRA1.

The data in this section confirms the results in section 4.3.1, which showed that the ALAS1 CRD does not fully reside within the 154bp region identified by Cable *et al* (2001). None of the seven ALAS1 deletion constructs were destabilised by haem in this assay, in which the full coding region is always destabilised. To define the specific haem-mediated *cis*-acting element in the coding region by reporter assay methods was not possible. Only the full-length coding region ALAS1 mRNA is unstable in the presence of haem, whereas the short truncated regions are not. The reason for this is unclear, but may be due to the ALAS1 coding region containing multiple sequence elements required for destabilisation by haem. These may require specific spatial and positional separation. Therefore, the methods used in this section are not able to further define the CRD. However, the differences in translational efficiency of these sequences may have a specific role in the haem-destabilisation mechanism.
4.3.5 Confirmation of an ALAS1 CRD by RPA

The previous experiments in this chapter, using reporter assays, have shown that the ALAS1 haem-responsive element is within its coding region. This needed to be confirmed by quantifying the mRNA directly, to establish that the effect of haem upon Renilla luciferase activity in the pRALAS1-transfected cells resulted from enhanced instability of the fusion mRNA, rather than an inhibitory effect upon its translation.

The first approach we used was to utilise RNA protection assays (RPA), using RNA extracted from pRALAS1-transfected HepG2 cells at $t_0$, and those at $t_6$, which were treated with and without 20μM haem, 24 hours post-transfection. The levels of the fusion mRNA at these time-points were assayed using RPA, with a biotin-UTP labelled 328nt probe that spanned the Renilla and ALAS1 mRNA moieties, along with a biotin-UTP labelled 258nt probe for the endogenous GAPDH mRNA as a control.

Figure 4-9 shows a typical result from this RPA experiment. This assay proved to be technically difficult, with the mRNA often being degraded too quickly for it to give reproducible and accurate results. Furthermore, no effect of the addition of haem on the stability of pRALAS1 mRNA was observed. It was thought that the media in which the HepG2 cells grow, already contain haem and haemoproteins in the FCS, which need to be removed so that the effect of haem on the ALAS1 mRNA stability is higher than two-fold, and can therefore be visualised.
Figure 4-9 RNase protection assay of pRALAS1 and GAPDH mRNA in response to 20μM haem for 6h.

To establish that the effect of haem upon Renilla luciferase activity in the pRALAS1 cells resulted from enhanced instability of the fusion mRNA rather than an inhibitory effect upon its translation, levels of the fusion mRNA were analysed using RNase protection analysis. The upper band is the protected fragment of the pRALAS1 Renilla-ALAS1 fusion mRNA (328nt) and the lower band that of GAPDH (258nt). The image shown is indicative of a typical experiment.

Lanes 1-2: No RNase control. These serve as a control for probe integrity, showing the gel migration of the full-length probe of pRALAS (1) and GAPDH (2). Any degradation or secondary structure would be seen here. These lanes show a single band for each probe, migrating at the correct size; 3-4: No target RNA control, used 5μg yeast RNA only. These serve as positive controls for the function of the RNases. There is no signal for the pRALAS1 probe (3), and a very faint signal for the GAPDH probe (4) as yeast mRNA also contains this gene; 5-6: RPA of mRNA extracted at t0 in duplicate; 7-8: RPA of mRNA extracted at t6, incubated with control medium only, in duplicate; 9-10: RPA of mRNA extracted at t6, incubated with 20μM haem, in duplicate.
4.3.6 Optimisation of the removal of haem and haemoproteins from FCS

Initial experiments using RPA to assess the stability of ALAS1 mRNA in HepG2 showed variable results. To obtain more robust data, it was thought that the media in which cells are grown should be depleted of haem, so that when haem is added a larger than two-fold difference of ALAS1 mRNA is seen. Furthermore, this would ensure that a controlled amount of haem is being added to the cells, giving consistency between experiments.

Foetal calf serum (FCS) is used to make cDMEM for the growth of HepG2 cells, and contains an unknown quantity of haem. A method was obtained from Dr Li Zhang (University of Columbia), to deplete FCS of haem by the addition of ascorbic acid and dialysis in phosphate-buffered saline (PBS). This method was optimised and confirmed to be working by measuring the expression of HO-1 mRNA in HepG2 cells in normal media, cells depleted of haem and cells depleted of haem for 24h prior to the addition of 20µM haem for a further 4h (the method for haem depletion is in the materials and methods chapter, section 2.2.1.6.).

Since the observed decrease in Renilla luciferase activity in the previous DLRA experiments was consistent and reproducible but less than two-fold, we endeavoured to effectively enhance the sensitivity of the experiments. This was by stabilising the fusion mRNA in cells through haem depletion using 100µM 4,6-dioxoheptanoic acid (succinyl acetone), a potent and specific inhibitor of ALA dehydratase, added to the haem-depleted media prior to haem treatment. Structurally, succinylacetone closely resembles ALA and acts as a competitive inhibitor of ALAD. ALAD activity in the liver and erythroid precursors is dramatically reduced and excess ALA is excreted in the urine (De Matteis & Marks, 1983). This ensured that the concentration of haem in the cells is completely reduced, and therefore a high concentration of ALAS1 mRNA is present in the cells.

To confirm that this method of haem-depletion was successful, RT-PCR was used to quantify the HO-1. HO is the rate-limiting enzyme in the degradation of haem, and induction of its isozyme, HO-1, is recognised as a protective response against haem protein-mediated and other insults to diverse tissues (Kanakiriya et al., 2003).
Figure 4-10 HO-1 mRNA expression is reduced in HepG2 cells grown in medium containing haem-depleted FCS.

HepG2 cells, in triplicate, were grown in either normal media or haem-depleted media with/without the addition of 20μM haem for 4 hours prior to the extraction of total RNA.

a) HO-1 and GAPDH mRNA was quantified by RT-PCR, and electrophoresed on a 1.4% agarose gel. Experiments were done on three separate occasions in duplicate. The above gel is an example of a typical experiment.

b) Densitometry of the HO-1 RNA levels normalised to the GAPDH RNA levels from figure 4-10a.
Cells were grown in cDMEM and split into 6-well plates (2ml cells/well). After 24h, the media was replaced with haem-depleted cDMEM, and 500µM succinyl acetone was added to inhibit the synthesis of haem. After a further 24h, 20µM haem was added to the cells, and the RNA was extracted after 4h, using Ultraspec. HO-1 and GAPDH mRNA was quantified by RT-PCR, using primers that spanned the introns so that genomic DNA could be separated from cDNA. PCR samples were electrophoresed on a 1.4% agarose gel.

The expression of HO-1 in HepG2 is reduced by approximately 25% when the cells are grown in haem-depleted media rather than normal media (figure 4-10). Therefore this method of dialysis of the FCS, after the addition of ascorbic acid, depletes the media of haem and haemoproteins. The addition of 20µM haem for 4h to cells grown in haem-depleted media increases the expression of HO-1 by approximately 4.5-fold. The method to deplete and then add an excess of haem to the HepG2 cells works well, and can therefore be relied upon in all further experiments to establish the effect of exogenously added haem on ALAS1 mRNA stability in HepG2 cells.
4.3.7 Confirmation of an ALAS1 CRD by semi-quantitative RT-PCR

Experiments using RPA to confirm the destabilisation of the coding region of ALAS1 by haem proved to be technically difficult and unsuccessful as the addition of haem causes only a two-fold decrease of ALAS1 mRNA levels. Therefore, a method to deplete cell media of haem was used, so that when HepG2 cells where grown in this media, the cellular content of ALAS1 mRNA would be high (and the level of HO-1 mRNA is low, as shown in figure 4-10). Consequently, when haem is added to the cells, the ALAS1 mRNA concentration drops significantly and can be analysed by RNA analysis. To quantify this, RT-PCR was used rather than the previous RPAs, as this is a quicker method, so that the mRNA had less time to degrade as in the RPAs. Furthermore, a two-fold difference is challenging to demonstrate on steady-state levels of mRNA, used in the RPA.

Similar to the DLRA experiments, HepG2 cells in 6-well plates were transfected with pRALAS1 and pGL3P and grown in haem-depleted media and treated with 100μM succinyl acetone. The cells were subjected to treatment with 5μg/ml actinomycin D to induce transcriptional arrest, and 20μM haemin (or control medium) for 6h. After 6h, mRNA was extracted from the cells using the Dynal mRNA direct kit (Ambion). The pRALAS1 and pGL3P cDNA were separately quantified by RT-PCR. As a control, pGL3P was quantified by PCR, so that any differences in transfection efficiency can be visualised.

The results of the RT-PCR in figure 4-11 demonstrate clearly that the addition of 20μM haem to HepG2 cells for 6h causes the destabilisation of the fusion mRNA containing the ALAS1 coding region. Therefore this is likely to be responsible for the decreased Renilla luciferase activity seen earlier. This confirms the data from the DLRA’s, that ALAS1 mRNA contains a haem-mediated CRD. The reporter assays were done in equilibrium and therefore showed a two-fold change in expression with the addition of haem. However, for the RT-PCR experiments, prior to the addition of haem, the cells were depleted of haem by using specific growth media and succinyl acetone. Furthermore, actinomycin D was used to inhibit transcription. The growth of HepG2 cells in this haem-depleted media, showed a larger than two-fold difference between ALAS1 mRNA with and without haem, visualised by semi-quantitative RT-PCR. We have calculated that the t1/2 of ALAS1 mRNA is reduced from 6.5h to 2.9h. Therefore, after incubation with haem for 6h, it is expected that the mRNA will go through at least two half-lives.
Figure 4-11 The addition of 20µM haem destabilises the fusion Renilla-ALAS1 coding region mRNA by more than 2-fold in 6h.

HepG2 cells were grown in haem-depleted media and transfected with pRALAS1 and pGL3P. mRNA was extracted at t₀, t₆ with the addition of control medium (t₆-haem), and t₆ with the addition of 20µM haem (t₆+haem), in duplicate. The addition of haem for 6h caused the ALAS1 mRNA levels to decrease by approximately 70%.

a) cDNA fragments for the pRALAS1 construct and pGL3P were specifically amplified by RT-PCR using reverse transcribed DNAse I-treated mRNA. Experiments were done on three separate occasions in duplicate. The above gel is an example of a typical experiment.

b) Densitometry of the pRALAS1 mRNA levels from figure 4-11a. Values were not normalised to pGL3P, as the densitometry was similar on each occasions, indicating similar transfection efficiency between samples.
4.3.8 Haem destabilises ALAS1 mRNA in IMR32 cells.

Using reporter assay techniques and semi-quantitative RT-PCR, it was found that in HepG2 cells, ALAS1 contains a CRD which functions to destabilise ALAS1 mRNA in response to haem. The majority of research on the regulation of haem biosynthesis in non-erythroid cells has been in hepatocytes. In order to assess if the CRD is used similarly in other cell types the DLRA technique was utilised on Renilla-ALAS1 transfected IMR32 cells.

The Renilla luciferase ALAS1 reporter constructs pRALAS1, pSTOPALAS1, pRALAS3UTR, along with the parental vector, pRLuc-C2 were transfected into IMR32 cells, each with pGL3P for normalisation purposes, under the same experimental conditions as for the HepG2 experiments.

Figure 4-12 demonstrates that similar to the reporter assays in HepG2 cells, the expression of the full length construct decreased by 35% in IMR32 cells, with the addition of haem. The addition of a stop codon, to make the coding region untranslated, does not completely prevent the haem effect. These results indicate that haem destabilises ALAS1 mRNA in non-hepatic tissues, but the mechanism may be different to the system used in hepatoma cells. However, the differences shown may be of experimental origin, rather than indicative of a different mechanism.
Figure 4-12 Relative *Renilla* luciferase expression of ALAS1 constructs in IMR32 cells in response to the addition of 20µM haem for 24h.

The *Renilla*-ALAS1 fusion constructs and the empty parental vector were co-transfected into IMR32 cells with the pGL3P vector as the internal control. Cells were treated with 20µM haem or the vehicle medium for 24h, and luciferase expression was measured using the DLRA. Data shows the relative *Renilla* expression after the addition of 20µM haem, compared to the expression without haem as 100%. Bars indicate plus standard deviation. Experiments were on two separate occasions in triplicate. Paired sample t tests were performed on the data for pRALAS1, pRSTOPALAS1 and pRALAS3UTR. p values were less than 0.05, indicating a significant difference between *Renilla* expression with and without haem treatment for all constructs.
4.4 Conclusions

The ability of haem to control its own synthesis in all cells is crucial. High intracellular concentrations of haem are cytotoxic, damaging cells through oxidative processes involving Fenton chemistry, whilst haem deficiency may impede the activity of certain haemoproteins with potentially harmful consequences, particularly with regard to neurodegeneration and ageing (Atamna, 2004; Atamna et al., 2002). In liver and probably all other non-erythroid cells, haem supply is regulated primarily via feedback regulation of ALAS1 mRNA stability, yet the precise mechanism has not been elucidated (Ponka, 1999).

There are a number of important observations from the data presented in this chapter:

1. The previously identified putative cis-acting element is not sufficient for haem-mediated destabilisation of the mRNA.
2. The haem-mediated element is present in the ALAS1 coding region, and the mechanism of instability does not involve sequence elements in the 5'-or 3'-UTR.
3. Destabilisation of the ALAS1 mRNA could only be demonstrated using the full-length coding region, and not in smaller deletion fragments.
4. In HepG2 cells, the haem-mediated destabilisation of ALAS1 mRNA requires the translation mechanism to occur through a CRD to function.

Neither the 5' nor 3'-UTR of the major ALAS1 mRNA conferred haem-responsive destabilisation upon a Renilla luciferase-ALAS1 fusion mRNA, but instead it was demonstrated by luciferase assays that the cis-acting elements responsible for haem-mediated destabilisation reside within the coding region. Methods were optimised to reduce the level of haem in FCS that was used in the medium in which HepG2 cells were grown. This enabled the RNA assays to be optimised, so that the destabilisation of the full-length ALAS1 coding region construct was seen with the addition of haem by using the technique RT-PCR. This confirmed the results from the DLRA's, that ALAS1 contains a haem-mediated CRD.

Cable et al (2001) identified a putative haem-mediated CRD due to its ability to bind cytosolic proteins, and that its deletion abolished the haem effect. However, the data in this chapter has shown that this sequence alone is not enough to confer the destabilisation of ALAS1 by exogenously added haem. The ALAS1
instability element resides solely in the coding region, although the 154bp region alone has not been shown to function as a CRD, under these experimental conditions. Thus, it is probable that additional elements surrounding this sequence or elsewhere within the coding region are required for haem-dependent regulation of ALAS1 mRNA. Indirect evidence supporting such a conclusion comes from erythroid cells in which two ALAS isoforms are expressed encoded by ALAS1 and ALAS2 genes (Bishop, 1990). Despite a homology of 61% between them over this 154nt region (figure 4-13), ALAS1 mRNA stability is reportedly regulated by haem in these cells (Fujita et al., 1991; Harigae et al., 2003), but ALAS2 mRNA is not.

| ALAS1  | GCAUCAAUUCUCCUAGCUCGGCAGGGAAAGACUACUCUCGCUUUCCUAGUAGGAAAGCAGAA |
| ALAS2  | GCCAUCAAUUCUCAUCUGCUUUCCUAGUAGGAAAGCAGAA |

| ALAS1  | CACCAACAGCCAGAUAUGAAGGACACUUCUCGCUUUCCUAGUAGGAAAGCAGAA |
| ALAS2  | CACCAACAGCCAGAUAUGAAGGACACUUCUCGCUUUCCUAGUAGGAAAGCAGAA |

| ALAS1  | GUGGGCGUGGAACUGAAGUCUAGUCUCUGCGUG |
| ALAS2  | GUGGGCGUGGAACUGAAGUCUAGUCUCUGCGUG |

**Figure 4-13 Homology between the 154bp sequence of human ALAS1 and ALAS2 mRNA.**
There is 61% homology of the putative cis-acting element between ALAS1 and ALAS2. Non-homologous nucleotides are indicated in red.

It is hypothesised that a larger region of the mRNA or other sequence elements within the coding region are required for destabilisation to occur, rather than the earlier identified 153bp sequence. Renilla-ALAS1 fusion constructs were made using seven overlapping regions of the coding region. Experiments using these constructs in the luciferase assays did not provide evidence to indicate that there is a specific region involved in this instability mechanism. Using reporter assays, we were only able to demonstrate the haem-mediated destabilisation of ALAS1 using the full-length coding region. The approach to define the CRD by using specific sequences on their own is not sufficient for the destabilisation by haem, suggesting that other regions are involved, but cannot function in isolation. Therefore, it is likely that the ALAS1 CRD is a bi- or tri-partite structure, similar to the destabilisation mechanisms of c-myc (Lemm & Ross, 2002), c-fos (Chen et al., 1992) and plasminogen activator type 2 (PAI-2) mRNA (Tierney & Medcalf,
2001). The specific secondary structure of the mRNA may be essential for this mechanism so that these sequence elements are close together in the mRNA molecule.

The data in chapter 3 showed that the minor ALAS1 form is not translated as well as the major form, and is relatively stable in the presence of haem. The data presented in this chapter has demonstrated that by inserting a stop codon prior to the ALAS1 coding region the destabilisation by haem was abolished. This indicates a mechanism in which the haem-mediated decay of ALAS1 mRNA requires translation through its coding region. Like most short-lived mRNAs, ALAS1 mRNA is stabilised in cells exposed to translational inhibitors. Inhibition of translation also prevents its haem-mediated decay (Drew & Ades, 1989; Cable et al., 2001). Although our data does not exclude the possibility that the mechanism involves a labile trans-acting factor, we have demonstrated a critical requirement for translation of ALAS1 mRNA itself.

The results showed that 6 of the 7 ALAS1 coding region fragments, when inserted into a Renilla luciferase-encoding vector, had high but variable expression. However, the pRA4 construct had much lower expression, which was at an equivalent expression to the full-length construct, pRALAS1, with a much longer ALAS1 insert of 1923bp. This 483bp sequence, at the 3' end of the coding region, seems to have a lower rate of translation, a factor which may be involved in the destabilisation mechanism. The different rates of translation may, in part, be due to different mRNA 3D structures that impede ribosomal scanning through them. The fact that fragment 4 is so badly translated may indicate a highly structured mRNA. Figure 4-14 indicates that this fragment 4, consisting of the last 483bp of human ALAS1 mRNA, contains multiple stem-loop structures and has a high free energy.
Figure 4-14 The predicted secondary structure of fragment 4 of the ALAS1 coding region

The RNA folding software mfold (http://frontend.bioinfo.rpi.edu/applications/mfold/cgi-bin/rna-form1.cgi) predicts that fragment 4, of 483bp in length, produces a highly structured mRNA with a free energy of -160.7 kcal/mol.

FCS used in the growth medium of HepG2 cells contains an unknown concentration of haem and haemoproteins, which may have affected the results of RNA analysis. Therefore, a method to deplete the FCS of haem was optimised and demonstrated to work by RT-PCR of HO-1 in HepG2 cells grown in this media. HO-1 is readily induced by haem proteins and their haem moiety, and in cells
grown in haem-depleted media, the concentration of HO-1 mRNA was markedly lower than in cells grown with 20µM haem in the media. We were then able to use this growth medium for RT-PCR analysis of pRALAS1-transfected HepG2 cells. Using RPA, we were unable to consistently demonstrate the haem-effect on ALAS1 mRNA steady state levels. The concentration of ALAS1 mRNA only decreases by 2-fold over 6 hours, which is challenging to show on steady-state levels. Therefore RT-PCR was used, in which the transcription of ALAS1 was stopped using actinomycin D. By this method, we were able to successfully show inductional repression of ALAS1 coding region mRNA by haem. This confirmed the results of the DLRA, that the human ALAS1 mRNA contains a cis-acting element which functions in HepG2 cells to destabilise the mRNA in the presence of haem.

Reporter assays on IMR32 cells demonstrated that the addition of haem destabilised the ALAS1 coding region construct, pRALAS1. However, in contrast to the results of similar experiments on HepG2 cells, the addition of haem also destabilised pRSTOPALAS1 and pRALAS3UTR in IMR32 cells. In hepatocytes, haem regulates its own synthesis post-transcriptionally, by either destabilising the ALAS1 mRNA, or preventing the pre-ALAS1 protein from entering the mitochondria where it functions. In IMR32 cells, the addition of haem seems to destabilise the ALAS1 coding region mRNA even when it is not translated, although not to the same level. Haem also destabilises the mRNA with and without the 3'-UTR. It is interesting to speculate that the ALAS1 mRNA is regulated by a different mechanism in human neuronal cells, and the mechanism by which the ALAS1 CRD has to be translated to function may be hepatocyte-specific.
CHAPTER 5
Identification of ALAS1 mRNA protein-binding regions
5.1 Introduction

The results in the previous chapter demonstrated that ALAS1 contains an element in its coding region that confers haem-mediated destabilisation on the mRNA, and requires translation to function. However, the cis-acting element could not be defined further with the use of smaller overlapping sequences of the coding region in reporter assays. An explanation for this is that there may be more than one destabilising element which are required to be spatially distinct to function. An alternative strategy was used to identify protein-binding regions within the human ALAS1 mRNA coding region, to discover whether more than one region binds protein. mRNA stability/instability determinants are also binding sites for proteins; therefore these protein-binding results would give an indication that other sites may be involved in haem-mediated destabilisation of ALAS1 and the coding region determinant could be further delineated. Furthermore, we sought to characterise the size of these ALAS1 binding proteins. In addition to this we also investigated if haem could bind to the ALAS1 mRNA directly.

5.1.1 Trans-acting regulatory factors

From the onset of transcription, RNA is associated with RNA-binding proteins. Various RNA-binding proteins are involved in the regulation of splicing and processing, export of mRNA to the cytoplasm, maintenance of mRNA in the cytoplasm for translation, and finally decay of the mRNA. Some RNA-binding proteins shuttle between the nucleus and the cytoplasm (Shyu & Wilkinson, 2000), and a subset of these proteins may stay with an mRNA molecule for its lifespan. Other RNA-binding proteins are transiently associated with the mRNA, and may bind and dissociate in response to different stimuli. Some RNA-binding proteins have been shown to have a role in the regulation of gene expression by stabilizing or destabilizing a particular mRNA. Multiple stimuli can alter the affinity of RNA-binding proteins for their target cis-acting elements, including hormones and cytokines, redox changes, UV-light, the cell cycle, and developmental stage (Hollams et al., 2002).

5.1.2 Tetrapyrrole binding to mRNA

In this chapter we examined if haem can bind directly to ALAS1 mRNA to regulate its own synthesis. The binding of haem to mRNA may be likely, as the tetrapyrrole cobalamin (CBL; vitamin B\textsubscript{12}), synthesised only in prokaryotes, has
been shown to bind to RNA. CBL is an essential cofactor for several important enzymes that catalyze a variety of transmethylation and rearrangement reactions (Martens et al., 2002). Expression of the Salmonella typhimurium cob operon, encoding the CBL biosynthetic pathway, and of the btuB gene of Escherichia coli and S. typhimurium, encoding the vitamin B12 transporter, is repressed by addition of vitamin B12 by a post-transcriptional regulatory mechanism (Lundrigan et al., 1991; Richter-Dahlfors & Andersson, 1992). The leader mRNAs of the cob and btuB genes contain an evolutionarily conserved sequence known as the B12-box. Adenosylcobalamin (Ado-CBL) is an effector molecule involved in the regulation of CBL genes (Nou & Kadner, 2000). The B12 RNA element is stabilized by direct binding of adenosylcobalamin. Thus, an adjacent regulatory hairpin terminator folds, which leads to transcriptional or translational repression of the genes. In the absence of the effector, the unstable vitamin-specific RNA element is replaced by an alternative anti-terminator RNA conformation allowing for transcription read-through or translation initiation (Vitreschak et al., 2003).

5.1.3 Aims and Strategy

The major aim of this chapter was to further delineate the ALAS1 CRD by identifying fragments of the mRNA coding region that bind protein. Secondly, it was investigated whether this protein-mRNA interaction is sensitive to haem. It was also examined whether haem is able to bind to the ALAS1 mRNA directly.

In order to achieve these aims, RNA-EMSAs were used to identify fragments of the ALAS1 coding region that bind to HepG2 cytosolic protein extracts. UV-crosslinking was then used to further characterise these proteins. Accordingly, the next part of the introduction details the salient features of the two RNA-binding methods used in this chapter (illustrated in figure 5-1).

To study RNA and haem interactions, haemin-agarose beads were incubated with HepG2 total RNA, and RT-PCR was used to establish if ALAS1 mRNA binds haem directly.

5.1.3.1 RNA-EMSAs

In order to identify if any regions of the ALAS1 mRNA coding region bind HepG2 cytosolic proteins, RNA-EMSAs (Electromobility Shift Assays) were used. EMSA's provide a rapid approach for both qualitative and quantitative analysis of nucleic
acid-protein interactions, and is based on the observation that the electrophoretic mobility of a nucleic acid through a non-denaturing polyacrylamide gel is retarded when a protein(s) is bound to it. This method was originally developed to assay DNA-binding proteins, and was subsequently modified to detect RNA-binding proteins. Whether a band of complexes is seen depends on the strength of the RNA-protein interaction, and for many specific reactions the binding may be quite strong and the complexes long-lived. In addition, the binding is generally enhanced by the relatively low-salt buffer used in the electrophoresis experiment. Any RNA from complexes which dissociate during the run will trail the free RNA band, forming a faint smear between the complex and unbound RNA that is not usually detectable (Walker et al., 1998).

All RNA-EMSA reactions were carried out in the presence of heparin and bovine liver ribosomal RNA (rRNA) at final concentrations of 5μg/μl and 0.1μg/μl respectively, using radiolabelled α²³P-CTP RNA. Heparin is a negatively charged anion that mimics the RNA phosphate backbone and is added to reduce non-specific protein binding to mRNA and background binding. Proteins that bind non-specifically to RNA simply because of its negative charge will bind to heparin rather than the RNA of interest. Ribosomal RNA is used as a competitor to bind to general RNA-binding proteins, therefore increasing the specificity of the assay. After incubation with protein, the RNA is electrophoresed on a non-denaturing gel, and exposed to film. Any RNA that has bound to protein will result in a shift on the gel, compared to the same RNA which had not been incubated with protein, run parallel on the gel.

5.1.3.2 UV-Crosslinking

To further characterise any trans-acting factors that bind to sequences of the ALAS1 coding region mRNA, UV-crosslinking experiments were carried out. UV-crosslinking of RNP (ribonucleoprotein) complexes takes advantage of the fact that UV light of sufficient intensity generates highly reactive species of RNA that react with molecules, including proteins, with which the RNA is in stable, direct contact. Crosslinking with UV light has several advantages over chemical reagents; it is relatively easy to carry out and the dose level can be accurately controlled. For example, irradiation at 254nm allows the identification of direct RNA-protein interactions since crosslink formation is observed only between molecular entities which are separated by no more than one covalent bond length (Walker et al., 1998). This covalent association of the RNA with proteins makes it
possible to isolate the UV-crosslinked complexes under strongly denaturing and disulphide reducing conditions that ensure that only proteins crosslinked to the RNA are isolated with the RNA. To remove RNA not covalently bound to protein, and to reduce the length of RNA bound to protein for subsequent size estimation, the irradiated reaction is treated with a mix of RNase A and T1. Therefore, the radiolabelled RNA probe is reduced to a stub covalently attached to the protein allowing the visualisation by autoradiography of the RNA-binding protein, and an estimation of its size, by electrophoresis on an SDS-polyacrylamide gel. Similar to the EMSAs, heparin and rRNA is also used in UV-crosslinking to reduce the non-specific binding to the RNA.
Figure 5-1 Schematic representation of the EMSA and UV-crosslinking reactions.

Cytosolic proteins from HepG2 cells are incubated with a uniformly labelled RNA probe at room temperature for 15 minutes. In the EMSA, the RNA is electrophoresed on a non-denaturing gel, and exposed to X-ray film. RNA-protein complexes are identified by a shift on the gel. In the UV-crosslinking reaction, the exposure to UV light covalently crosslinks closely associated proteins and RNA. The free RNA and the RNA not protected by UV-crosslinked protein are digested by an RNase A and T1 treatment. Finally, UV-crosslinked proteins are separated by SDS-PAGE, and exposed to X-ray film.
5.2 Materials and Methods

For the RNA-binding assays, the overlapping ALAS1 coding region sequences used in chapter 4 were cloned into the pGEM-T vector, in the T7 promoter direction. These were digested with the restriction enzyme, Not I to produce a 5' overhang. In Vitro transcription, using the Ambion T7 maxiscript kit, was used with 500ng of each linear construct, to produce the radioactive (\( ^{32} \)P-CTP) ALAS1 mRNA.

5.2.1 Radiolabelled ALAS1 RNA construction

The seven ALAS1 overlapping coding region sequences were amplified by PCR, using the primers listed in table 4-3 (chapter 4) with Bio-X-Act Long polymerase. These PCR products were gel extracted and inserted into the pGEM-T vector by TA cloning. These were used to transform subcloning efficiency, competent DH5α cells; and the plasmid DNA was isolated and purified. The orientation of each ALAS1 sequence in the pGEM-T vector was verified to be in the T7 sense direction by restriction digest.

![Figure 5-2 Schematic representation of the sub-cloning of ALAS1 DNA into pGEM-T vector.](image)

The ALAS1 inserts, each of approximately 480bp in length, were amplified by PCR with Bio-X-Act Long polymerase and inserted into the pGEM-T vector by TA cloning. The plasmids containing the inserts in the correct orientation were then linearised at the Not I site, for in vitro transcription from the T7 RNA polymerase start site.

To produce the radiolabelled RNA probes, each pGEM-T-ALAS1 plasmid was digested 5' of the SP6 promoter site with the restriction endonuclease Not I, 12bp downstream of the inserted ALAS1 sequence (figure 5-2). Not I cuts at the sequence 5'-GC$_3$GGCCGC-3' to produce a 5' overhang. The digested ALAS1-pGEM-T was then gel purified. T7 RNA Polymerase is an extremely processive
enzyme and will continue to transcribe around a circular template multiple times without disassociating, therefore the construct has to be digested, and specifically with an enzyme that does not produce a 3’ overhang. Enzymes which cleave the DNA and leave a 3’ overhang often give rise to aberrant transcripts.

Radiolabelled RNA was transcribed as described in section 2.7.1, using T7 polymerase and α³²P CTP as the limiting NTP at a final concentration of 3.125μM, with 10μM of cold CTP. The RNA was purified from the unincorporated nucleotides by spinning through a Sephadex G-50 column, and the amount of radioisotope made was quantified by liquid scintillation counting. Each radiolabelled RNA probe was electrophoresed on a denaturing gel to confirm that it was the correct size, no degradation had occurred, and that any unincorporated nucleotides had been completely removed. The RNA was stored at −70°C until further use, and used within 5 days so that only full-length and intact radiolabelled RNA was used in the experiments.

5.2.2 RNA EMSA

The seven ALAS1 overlapping coding region fragments were used in the RNA-EMSA reactions, as described in the materials and methods section, to establish if they bind to HepG2 cytosolic protein extracts. The protein used was extracted from confluent HepG2 cells, and the concentration measured using the BCA™ Protein Assay Kit (section 2.3.1). For each reaction, 100,000 cpm of the radiolabelled RNA was incubated with 25μg protein in a 10μl reaction for 15 minutes at room temperature. For a further 15 minutes, the samples were incubated with 5μg/μl of heparin and 0.1μg/μl of rRNA. The samples were then subjected to electrophoresis on a non-denaturing 4% polyacrylamide gel. In experiments where the affect of haem on the protein-binding was assayed, 20μM haem or control medium was added to the cells for 24 hours, after the cells had been grown in haem-depleted medium.

5.2.3 UV-Crosslinking of cytosolic proteins to ALAS1

The ALAS1 coding region fragments which were shown to bind to protein by EMSA were also used in cross-linking reactions with HepG2 cytosolic protein extracts. Similar to the EMSAs, 100,000 cpm of the radiolabelled RNA was incubated with 25μg protein in a 10μl reaction for 15 minutes. For a further 15 minutes, the samples were incubated with 5μg/μl of heparin and 0.1μg/μl of rRNA. As
described in the materials and methods section, the reaction was exposed twice to 8600 J of UV-light and unbound RNA was degraded by the addition of RNase A and T1. The samples were then subjected to electrophoresis by SDS-PAGE.

5.2.4 Binding of ALAS1 RNA to haemin-agarose

Haemin-agarose has previously been used to identify proteins that bind to haem in vitro. Therefore, this part of the chapter endeavoured to modify these protein-binding assays to assess if the ALAS1 mRNA binds to haem directly, thereby directly inhibiting its transcription. In these experiments, haemin-agarose beads were incubated with total RNA extracted from HepG2 cells, in the presence of 20 μg of glycogen. Glycogen was used as an inert material, allowing the precipitation of pg amounts of RNA. The method used was a modified version of the methods used by Trifilllis et al., 1999.

Haemin-agarose beads (60μl) were washed in 1ml of RNA binding buffer (RBB, 10mM Tris-HCl, pH 7.5, 1.5mM MgCl₂, 150mM KCl) by resuspending, centrifugation at 6000rpm at 4°C and removal of the supernatant. The beads were resuspended in 350μl of RBB and incubated with 10μg of HepG2 total RNA at 4°C for 1 hour. This incubation was followed by centrifugation at 4000rpm to remove unbound RNA and a wash in 1ml RBB/0.1% TX (Triton X-100), prior to further centrifugation at 4000rpm at 4°C. RNA interacting non-specifically with the haemin-agarose was competed off with 2 washes in RBB containing 1mg/ml heparin for 10 minutes at 4°C and centrifugation. The beads were subsequently rinsed four times in RBB/0.1% TX to remove unbound RNA, and the RNA was isolated from the drained beads by boiling for 5 minutes in 200μl TE/1% sodium dodecyl sulphate (SDS). The RNA was then phenol/chloroform (1:1) extracted, chloroform extracted, ethanol precipitated with 20μg glycogen and washed with 70% ethanol. The dried RNA was resuspended in 10μl sterile H₂O by heating to 65°C for 5 minutes. Washes were collected at varying stages of the procedure for RT-PCR analysis to assess how well specific RNA molecules bound to the haemin-agarose, and at what stage the ALAS1 RNA was eluted.

5.2.4.1 RT-PCR of ALAS1 and GAPDH

RT-PCR was used to amplify ALAS1 (primers P28, P29) and GAPDH (primers P32, P33), and compare elution of these RNA transcripts at varying stages of the assay. Prior to the quantification by PCR, all the PCR conditions had been
optimised for logarithmic amplification using untreated control cells (see section 3.2.1.1).

cDNA was synthesised from the RNA eluted from the haemin-agarose beads by first strand synthesis, and the RNA was quantified by PCR using primer sets specific for cDNA fragments of ALAS1 and GAPDH (table 5-1). The cycling protocol for the ALAS1 PCR was that as described in section 2.4.4.3.1, with 38 cycles, and an annealing temperature of 60°C. The GAPDH PCR used 34 cycles, with an annealing temperature of 65°C. The primers used for ALAS1 and GAPDH spanned exon boundaries, ensuring that only the mRNA was detected, rather than any contaminating genomic DNA in the RNA samples. The PCR products were resolved by agarose gel electrophoresis.

<table>
<thead>
<tr>
<th>Target DNA</th>
<th>Primer Number</th>
<th>Product Size</th>
<th>PCR Annealing Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALAS1</td>
<td>Sense P26, Anti-Sense P27</td>
<td>480bp</td>
<td>62°C</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Sense P32, Anti-Sense P33</td>
<td>267bp</td>
<td>65°C</td>
</tr>
</tbody>
</table>

Table 5-1 Primers used in semi-quantitative RT-PCR.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x PCR buffer</td>
<td>2</td>
</tr>
<tr>
<td>2mM dNTP mix</td>
<td>2</td>
</tr>
<tr>
<td>Sense Primer 5pmol</td>
<td>2</td>
</tr>
<tr>
<td>Antisense Primer 5pmol</td>
<td>2</td>
</tr>
<tr>
<td>cDNA Template (RT product)</td>
<td>2</td>
</tr>
<tr>
<td>HotStarTaq Polymerase (5U/µl)</td>
<td>0.2</td>
</tr>
<tr>
<td>H2O</td>
<td>to 20µl</td>
</tr>
</tbody>
</table>

Table 5-2 Composition of semi-quantitative RT-PCR.
5. 3 Results and Discussion

5.3.1 HepG2 cytosolic protein binds to two distinct regions of the ALAS1 coding region.

Cable et al. demonstrated that a specific 154bp sequence of the ALAS1 mRNA coding region binds to Huh-7 cytosolic proteins (Cable et al., 2001). RNA-EMSAs were used in this chapter to confirm that this region binds proteins from hepatocytes using HepG2 cells, and to investigate any other protein-binding regions.

Radiolabelled RNA that corresponded to the seven overlapping ALAS1 fragments of the coding region, designated RNA probes 1-7 (as indicated in figure 5-3b), were incubated with or without protein for 30 minutes, loaded onto a non-denaturing 4% polyacrylamide gel and electrophoresed at 200V for approximately 4 hours. The gel was dried and exposed to radiography film to visualise the RNA.

Figure 5-3a shows the EMSA of the RNA probes corresponding to the specific fragments of the ALAS1 coding region, incubated with and without HepG2 cytosolic protein. The lanes without protein demonstrate that the RNA had not decayed and was of the correct size. When the RNA probes were incubated with the HepG2 cytosolic protein extract, shifts were observed for probes 2, 4, 5 and 6 indicating that these regions of the RNA bind protein. No shifts were seen for probes 1, 3 and 7 of the ALAS1 coding region. These results suggest that no cis-acting elements are present in fragments 1, 3 and 7 as they do not bind cytosolic protein.

Interestingly, probe 4, which binds to protein, contains the 154bp sequence element previously identified by cable et al. However, our results from probe 2 indicate that an additional sequence towards the N-terminus (corresponding to 481-960bp) also binds protein. It is possible that the cis-acting element may still reside in this 154bp region, but it has a bi-partite structure, residing in both fragments 2 and 4. Two shifted bands are observed when probe 2 is incubated with protein. This may be the result of two different proteins binding to the RNA, or that additional secondary structures of the RNA are formed when binding occurs. The binding of protein to probe 2 is confirmed by the EMSAs of probes 5 and 6, which both overlap probe 2 and both bind to protein. This result suggests that there is extensive protein binding over multiple sites within fragment 2.
Figure 5-3 RNA-EMSA of overlapping probes of the ALAS1 coding region.

a) Radiolabeled RNA of 7 fragments of the ALAS1 coding region, each of 100,000 cpm, were incubated with 25μg HepG2 cytosolic protein extract or protein-free buffer in an EMSA reaction. The reactions were electrophoresed and visualised by autoradiography. Arrows indicate shifts in the electrophoresis, and indicate that RNA binds to protein. Lanes 1-2: probe 1; 3-4: probe 2; 5-6: probe 3; 7-8: probe 4; 9-10: probe 5; 11-12: probe 6; 13-14: probe 7.

Lanes 1,3,5,7,9,11,13: Protein-free EMSA; 2,4,6,8,10,12,14: Protein in EMSA reaction.

b) A schematic diagram to clarify the sequence of each of the ALAS1 probes. Binding of protein to the mRNA occurred only in fragments 2, 4, 5 and 6; not in 1, 3 and 7.
Probe 7 overlaps probe 4 by 240bp and did not bind protein, therefore it can be calculated that the protein-binding region is in the 243bp at the 3' end of the coding region. This fragment contains 141bp of the 154bp putative cis-acting element.

It can be clearly seen in the EMSA that the addition of protein extract to all the RNA probes causes the unbound RNA to shift slightly. This is most likely to be caused by the change in salt concentration of the sample, compared to the RNA probes with no protein extract added. Occasionally, the labelled RNA run in EMSA gels exhibits more than one conformation, resulting in multiple bands, in the absence of protein (Walker et al., 1998). This has occurred in the EMSA of RNA probe 3 (lane 5). This was confirmed by brief heating of the RNA for 3 minutes at 70°C followed by snap cooling on ice. This converts the slower migrating species into the faster migrating species (data not shown).

It has been reported that dyes such as bromophenol blue and xylene cyanol in the loading buffer can be inhibitory to binding and cause problems with complex formation (Walker et al., 1998). Therefore, EMSAs for fragments 1, 3 and 7 were repeated, using only glycerol to load the samples onto the gel, and electrophoresed alongside dye-containing buffer to monitor the extent of the electrophoresis. No shifts occurred for probes 1, 3 and 7 under these conditions, confirming that these regions do not bind protein (data not shown).

A faint smear is seen between the complex and unbound RNA of probes that were able to bind to cytosolic protein. RNA can dissociate from proteins during electrophoresis and therefore trail the free RNA band. This smearing may indicate that the binding of proteins to fragments 2, 4, 5 and 6 is not a strong association and the bonds formed can be easily disrupted. However, due to the large size of the RNA probes used, the unbound and bound RNA probes are not very well resolved which may add to the smearing in lanes 4, 8, 10 and 12.

These EMSA results show that there are two areas in the human ALAS1 coding region mRNA that can bind to Hepg2 cytosolic protein, corresponding to fragment 2 and the last 243bp of fragment 4. Probe 4 corresponds to the region previously identified by Cable et al (2001) as able to bind to proteins extracted from Huh-7 cells. This provides independent confirmation that the binding observed in the EMSA for probe 4 is correct.
5.3.2 Binding of HepG2 cytosolic protein to ALAS1 RNA probes with the addition of haemin in the EMSAs

Section 5.3.1 demonstrated that ALAS1 fragments 2 and 4 contain protein-binding sites. Therefore, it was next investigated whether the addition of haemin into the binding reaction of the EMSA affected this binding of cytosolic proteins to the ALAS1 RNA. Fragments 1 and 3 were also used in these EMSAs, to establish if the addition of haemin was able to induce protein binding to these sequences.

During the incubation stage of the EMSA, where the probe is incubated with protein extract for 30 minutes, either 20μM haemin solubilised in DMSO or DMSO alone was also added to the reaction. For each fragment, an EMSA with no DMSO or haemin added to the reaction was also, used as in section 5.3.1, as a further control.

This representative EMSA shown in figure 5-4 of probes 1 and 3 showed no demonstrable difference between samples with or without haemin added in the binding step. This confirms that the sequences corresponding to probes 1 and 3 are not involved in the haem destabilisation mechanism. No consistent effects of haemin could be demonstrated with repeated EMSAs, with haemin sometimes even appearing to increase the protein binding to both probes 2 and 4. However, the results of this EMSA in figure 5-4 show that the addition of haemin to the binding reaction of probe 2 caused a decrease in protein binding, whereas probe 4 decreased its protein binding with the addition of DMSO, and haemin had no effect. However, we were unable to consistently demonstrate effects of haemin on protein binding to probes 2 and 4.

These results confirm that the human ALAS1 mRNA contains two regions, corresponding to fragments 2 and 4 (figure 5-3), which can bind to cytosolic protein. Protein binding to either of these regions is not sensitive to the presence of haemin in the protein-binding step of the EMSA. Therefore, the binding to fragments 2 and 4 does not seem to be haem-sensitive under these experimental conditions.
Figure 5-4 RNA EMSA of the 4 ALAS1 coding region fragments incubated with HepG2 cytosolic protein, with haemin or DMSO.

An RNA-EMSA of the ALAS1 coding region fragments was carried out as previously. The radiolabelled RNA was incubated with HepG2 cytosolic protein, with either haemin, solubilised in DMSO, or DMSO alone as a control. For further controls, EMSAs were carried out using radiolabelled probe alone, or incubated with HepG2 cytosolic protein with no DMSO or haemin added. The above EMSA is a representative of the results obtained, when repeated three times.
5.3.3 Binding of cytosolic protein from haem depleted or supplemented HepG2 cells, to ALAS1 RNA

The previous section demonstrated that haemin added to the incubation step of the EMSA did not directly affect the binding of the HepG2 cytosolic protein to probes 2 and 4. However, it is possible that protein extracted from cells grown with the addition or depletion of haem may have an effect on the binding to the ALAS1 mRNA. The haemin added to the cells may cause any ALAS1 mRNA binding-proteins to be either up- or down-regulated within the cell, which could then be demonstrated by the EMSA. The different growth conditions for the HepG2 cells would have more physiological relevance and may have a different effect on protein-binding than the in vitro approach used in section 5.3.2.

To compare the affect of haemin added to the cells, prior to extraction of cytosolic proteins, the cells were treated in three different ways. HepG2 cells were either grown in normal media, as for the previous EMSA experiments as a control, or the cells were grown in haem-depleted media with the addition of 500µM succinyl acetone to inhibit haem biosynthesis. After 24 hours, haemin to a final concentration of 20µM or the vehicle medium was added to the cells grown in haem-depleted medium and the cytosolic protein was extracted after 4 hours. The RNA EMSA was performed as for the previous section.

This representative EMSA in figure 5-5 of probes 1 and 3 shows no binding occurred when incubated with the different protein samples. This confirms that the sequences corresponding to probes 1 and 3 are not involved in the haem destabilisation mechanism. The EMSA was repeated in triplicate, and was unable to demonstrate a consistent effect on protein-binding to fragments 2 and 4 by haem depletion/supplementing.

The representative EMSA in figure 5-5 shows that using protein from cells grown in medium with haem decreased the binding to probe 4. However, similar to the results seen in the previous section, although binding was decreased by haem the majority of times this EMSA was repeated, it did not occur each time. Haem depletion or add-back to cell medium had no effect on the protein binding to probe 2, compared to proteins from HepG2 cells grown in normal medium, on each occasion the EMSA was repeated. Therefore, this section confirms that fragments 2 and 4 bind to HepG2 cytosolic protein but we cannot conclusively prove that this binding is sensitive to haem.
Figure 5-5 RNA EMSA of the 4 ALAS1 coding region fragments incubated with cytosolic protein from cells grown in haem-depleted or supplemented medium.

An RNA-EMSA of the ALAS1 coding region fragments was carried out as previously. The 100,000cpm radiolabelled RNA was incubated with 25μg HepG2 cytosolic protein. As a control, EMSAs were carried out using radiolabelled probe alone. To test the affect of haem in the growth medium of cells on the protein binding, three types of medium were used: normal growth medium, haem-depleted medium, and haem-depleted medium with haem added for 4h before protein extraction. Differences in haem concentration had no effect on probes 1-3, but an excess of haem decreased the binding to probe 4 in the representative EMSA shown. However, this could not be repeated in further replicates of the experiment.
5.3.5 UV-Crosslinking of HepG2 cytosolic proteins bound to ALAS1 RNA

RNA-EMSAs identified two fragments of the ALAS1 mRNA that can bind to HepG2 cytosolic protein. UV-crosslinking assays were used to further characterise these proteins, as these may be important for the mechanism of haem-mediated ALAS1 mRNA destabilisation. Protein was only found to bind to the ALAS1 coding region probes 2 and 4 and consequently only these were used in the crosslinking assays.

Figure 5-6 shows the results of a typical experiment of the UV-crosslinking of fragments 2 and 4, in triplicate. There are a number of proteins that bind to fragment 2; those of the strongest intensity have the molecular weights of approximately 37, 40 and 70kDa (indicated by arrows). The other protein bands below these may be a result of protein degradation during the reaction and/or RNA degradation. However, the majority of the degraded RNA following the treatment with RNases is seen as a strong black band at the bottom of the gel. Probe 4 produced less protein bands than Probe 2. The bands at a higher intensity are of approximately 37 and 48kDa.
Figure 5-6 UV-crosslinking of ALAS1 radiolabelled RNA fragments 2 and 4 to characterise the binding proteins.

The UV-crosslinking reactions of ALAS1 coding region fragments 2 and 4 were carried out to identify the size of the RNA-binding proteins. For each fragment, the reaction was in triplicate, using 100,000 cpm radiolabelled probe incubated with 25μg HepG2 cytosolic protein, in the presence of 5μg/μl heparin and 0.1μg/μl rRNA to reduce non-specific binding. Samples were analysed by electrophoresis through an SDS-PAGE gel, and the dried gel was exposed to radiography film. The above figure is indicative of a typical gel. The experiment was repeated on three separate occasions.
Lanes 1-3: Radiolabelled ALAS1 coding region fragment 2 in triplicate;
Lanes 4-6: Radiolabelled ALAS1 coding region fragment 4 in triplicate.
5.3.6 Optimisation of UV-crosslinking to fragment 2

The UV-crosslinking of fragment 2 of the ALAS1 coding region produced numerous bands on the SDS-PAGE gel, many more than for fragment 4. Increasing concentrations of the negatively-charged anion, heparin, were used to investigate which of the proteins were specifically binding to probe 2, rather than non-specific proteins which bind to the RNA due to its negative charge. In the UV-crosslinking reaction, heparin was added to a final concentration of 5, 10, 25 and 50μg/μl into the binding reaction, with 25μg of HepG2 cytosolic protein, and the crosslinking assay was carried out as previously.

Figure 5-7 demonstrates that the increasing concentrations of heparin in the reactions increased the intensity of the proteins of approximately 140kDa and 80kDa, which did not seem to be significant on the previous gel (figure 5-6). The bands at approximately 40kDa and 37kDa also seemed to increase in intensity with the higher concentrations of heparin. Furthermore, the increased concentrations of heparin reduced the number and intensity of the smaller bands which are present at the bottom of the gel with low molecular weights. Therefore, using a higher heparin concentration in the UV-crosslinking caused less protein and RNA degradation to occur. It is difficult to analyse further the significance of these proteins which bind to probe 2. Only the molecular weights are known, therefore it cannot be deduced which proteins these actually are and whether the multiple bands are due to multiple proteins binding or numerous subunits of one protein which specifically binds to ALAS1 mRNA.
Figure 5-7 UV-crosslinking of RNA fragment 2 with HepG2 cytosolic protein extract and increasing concentrations of heparin.

Heparin at increasing concentrations was added to the binding reaction to compete out any unspecific binding of protein to the ALAS1 mRNA.

Final concentrations of heparin in the binding reaction (in duplicate):
Lanes 1-2: 5μg/μl; 3-4: 10μg/μl; 5-6: 25μg/μl; 7-8: 50μg/μl.

UV-crosslinking was carried out using 100,000cpm radiolabelled RNA probe with 25μg HepG2 cytosolic protein, as described in the methods. The reactions were analysed by SDS-PAGE, and the dried gel was exposed to radiography film.

The above gel is indicative of a typical experiment. The experiment was repeated on two separate occasions, with each heparin concentration in duplicate.
5.3.7 The use of haemin-agarose beads to investigate whether haem binds ALAS1 mRNA directly

To date, haem has not been shown to bind to any mRNA isoforms; therefore we endeavoured to find out if haem directly binds to ALAS1, the first enzyme in its biosynthetic pathway, thereby reducing the rate of its own synthesis. To assess if ALAS1 mRNA directly binds to haem, we utilized haemin-agarose beads incubated with total RNA extracted from HepG2 cells. Haemin-agarose has been previously used by others to investigate if proteins bind to haem; therefore a method was designed for RNA-binding analysis, by a type of affinity purification. RT-PCR was used to amplify the ALAS1 mRNA alongside GAPDH as a control for unspecific binding in the final eluted RNA sample. The RNA was also quantified from the total RNA samples prior to incubation with the haemin-agarose, as well as from two wash steps to establish when each RNA isoform was eluted from the beads.

Lanes 1 and 5 of figure 5-8 confirm that both ALAS1 and GAPDH mRNA are present in the HepG2 total RNA extract, used to incubate with haemin-agarose. The first wash of the beads eluted GAPDH mRNA (lane 2), and a small amount of ALAS1 mRNA (lane 6). In contrast, no GAPDH or ALAS1 mRNA was eluted after the beads were washed in heparin (lanes 3 and 7). The final step to elute the mRNA bound to the haemin demonstrates that both ALAS1 (lane 4) and GAPDH (lane 8) mRNA were able to bind the haemin-agarose, suggesting that the binding to ALAS1 mRNA is non-specific. However, it interesting that only a small quantity of ALAS1 mRNA was eluted from the beads during the first wash step. It may be that both the ALAS1 and GAPDH mRNAs are binding to the agarose rather than the haemin, thereby obscuring any genuine results from the experiment. Despite extensive modification of the wash steps, and even attempting to compete off the ALAS1 mRNA by adding haemin, no specific direct binding of haemin to ALAS1 mRNA was demonstrated in this RNA-binding assay.
Figure 5-8 RT-PCR analysis of GAPDH and ALAS1 mRNA binding to haemin agarose.

HepG2 total RNA (10μg) was incubated with 60μl haemin-agarose beads for 1h at 4°C. After incubation the sample was washed in RBB, followed by a wash in RBB containing 1mg/ml heparin. The beads were then rinsed four times in RBB and the RNA bound to the haemin-agarose was isolated by boiling for 3min in TE/1% SDS. RNA samples were used at various stages to assess if the ALAS1 and GAPDH mRNA binds to haemin. Each sample was ethanol precipitated and resuspended in 10μl 1x MMLV-RT buffer with 1U of Rnasin. Samples were subjected to RT-PCR and the ALAS1 and GAPDH mRNA were separately quantified by PCR.

Lane 1, 5: total RNA; 2, 6: RNA eluted from wash in RBB/0.1%TX; 3, 7: RNA eluted from wash with heparin; 4, 8: Final eluted RNA.
5.4 Conclusions

The data produced in the chapter demonstrates that cytosolic protein extracted from HepG2 cells is able to bind to radiolabelled RNA fragments of the ALAS1 coding region. Of the seven radiolabelled RNAs made, probes 2, 4, 5 and 6 were able to bind protein. Due to the overlapping nature of these sequences it can be deduced that protein binds in two distinct regions ranging from 481-960bp and 1681-1923bp of the coding region. A previous sequence at the 3’ end of the ALAS1 coding region of 154bp has previous been identified as a putative protein-binding site by Cable et al (2001). The results from the RNA-EMSAs have reduced this putative binding site by 12bp to a size of 142bp, as binding occurred only in fragment 4 and not fragment 7 (figure 5-9).

![ALAS1 coding region = 1923bp](image)

Figure 5-9 Putative cis-acting elements in the ALAS1 mRNA deduced from the protein-binding sites.

RNA-EMSAs were used to establish which of the seven overlapping fragments of the ALAS1 coding region can bind to HepG2 cytosolic protein. Fractions 2, 4, 5 and 6 are able to bind protein, whereas 1, 3 and 7 do not. Therefore, it is concluded that the protein-binding sites are in fragments 2 and at the end of 4. The red line indicates the 154bp cis-acting element identified by Cable et al. Probe 4 contained the full 154bp sequence and was able to bind protein, whereas probe 7 only contained 12bp of the putative element and did not bind protein. Therefore, this putative cis-acting element can be reduced to 142bp.

In these experiments, the addition of haemin in the binding reactions does not affect the binding to probes 2 and 4, and does not induce binding to probes 1 and 3. Furthermore, using protein from HepG2 cells grown in haem-depleted media or
grown in the presence of haem also did not conclusively affect binding to probes 2 and 4, or induce binding to probes 1 and 3. Therefore, fragments 2 and 4 are surmised to contain the haem-mediated cis-acting elements from the results of these EMSA experiments, as these bind to proteins, even though a haem-effect was not seen. It is notable that when Cable et al carried out EMSAs using Huh-7 protein extracts, they were also unable to show that haem affected the binding of protein to the 154bp sequence of the ALAS1 mRNA, corresponding to our fragment 4 (personal communication).

It has not been established if the putative elements in fragments 2 and 4 are both necessary for the haem-mediated destabilisation of the mRNA. Ways to investigate this may be to use a chimeric protein of fragments 2 and 4 in a DLRA, as used in chapter 4, to assess the stability in response to haem. Secondly, to use a chimeric radiolabelled RNA in the EMSA, with the sequences of regions 2 and 4 adjacent to each other. However, often the correct secondary structure of mRNA is vital, and altering the sequence of ALAS1 in this way by removing the sequence corresponding to region 3 may have aberrant results. As suggested in chapter 4, the cis-acting elements may have to be at a specific position in the mRNA coding region, at a certain distance from each other to function. Also, the RNA sizes of 480nt used in this chapter are very long for an EMSA. A radiolabelled RNA probe longer than this length would be difficult to be resolved from the protein bound RNA.

ALAS2 mRNA is not destabilised by haem. Rather, translation of the mRNA is controlled by an iron response element (IRE) in its 5'-UTR, where the higher the iron concentration within the cell, the more ALAS2 and therefore haem is produced. The IRE is bound to an IRP, and at high concentrations of iron this IRP is displaced, allowing translation through it (Dandekar et al., 1991). ALAS2 mRNA is not sensitive to haem concentration and is therefore unlikely to contain a cis-acting instability element, as is present in the ALAS1 mRNA. However, ALAS1 gene expression in murine Friend virus-transformed erythroleukemia (MEL) cells appears to be under the haem-mediated feedback control as it is in the liver. Expression of mRNAs encoding the ALAS1 and ALAS2 were examined using in situ hybridization. Following DMSO treatment, ALAS2 mRNA increased, whilst ALAS1 mRNA did not increase in wild-type MEL cells (Fujita et al., 1991; Mitani et al., 1992). Furthermore, it has been shown that ALAS1 is up-regulated in ALAS2-knockout erythroblasts, but down-regulated in wild-type erythroblasts (Harigae et al., 2003). Therefore, ALAS1 and ALAS2 mRNAs are under separate controls.
There must be a specific difference between the mRNA sequences of ALAS1 and ALAS2 to enable a haem-mediated destabilising cis-acting element to function only in the ALAS1 mRNA.

On comparison of the sequence of fragment 2 in human ALAS1 mRNA and the corresponding fragment of human ALAS2 mRNA, there is 61% homology (figure 5-10). This sequence homology suggests that there may be specific sequences in the fragment that are involved in haem-mediated destabilisation of ALAS1 mRNA.

The fragment 2 of the ALAS1 coding region contains exon 4, and interestingly this exon is often removed in ALAS2 by alternative splicing of the mRNA. There are five alternatively spliced forms of human ALAS2. However, the principal alternatively spliced mRNA is that which lacks exon 4 encoding sequence, constituting around 35-45% of total ALAS2 mRNA in erythroid tissues (Cox et al., 2004; Conboy et al., 1992). Similar proportions of the normal and the spliced exon 4 mRNA isoforms have also been found in foetal liver and normal bone marrow. Exclusion of exon 4-encoded sequence from the ALAS2 protein, which constitutes a major variant in vivo does not substantially affect mitochondrial import or its ability to associate with SCSβA (succinyl CoA synthase beta A subunit), but does reduce its catalytic activity in vitro. The exclusion of a definitive in vivo role for this region between the mitochondrial pre-sequence and the ‘catalytic’ core remains to be proven. Cox et al have postulated that exon 4 may either aid stability of the enzyme, determine whether it is further processed in mitochondria, or may be required for interaction with proteins other than SCSβA (Cox et al., 2004). It is unknown whether the corresponding exon in ALAS1 has a novel function involved in its expression.
ALAS1  GGGGAUCCCAUGGACUUGCUCAGAGAGACAGCACAGAGGAA
ALAS2  -----------------GAAAGCAUUUUCCGUGUCCCAAGGA-------------------GCAGGAGCAUUCUGC

ALAS1  AGAGAGUGCAUCUCUUCUAAGAACUUGCC-AAAAUCUGUUCACUCUUUCAGUAGA
ALAS2  AAGGUACACACACUCUACGACAAUAUGCCUGAAAACUAUGUC-------------------UUCAAGUUAUGA

ALAS1  UCUCUUCUUGAGAA-AAAAUGAUGGAAAGAAGUGAACACACCCUAGGAGUUGUUA
ALAS2  CCAGUUCUCCAGGACAAAGUAGUAGAAGAAACACUGACCCUAGCGUGUUA

ALAS1  AACUGUAGACCGGCGAGCGACACAUCCUCACUCUGACUGUAGACACACUGGCUAU
ALAS2  GCAGUAGACCGGCGAGCGACACAUCCUCACUCUGACUGUAGACACACUGGCUAU

<table>
<thead>
<tr>
<th>Fragment 5/6 boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALAS1  CACCAAAAGCAAGGUCAGUCUGGUGCAAGAAAGACACCAGGAGUCGACACC</td>
</tr>
<tr>
<td>ALAS2  CACCAAAAGCAAGGUCAGUCUGGUGCAAGAAAGACACCAGGAGUCGACACC</td>
</tr>
<tr>
<td>ALAS1  CACCGGCGUUGUGGCGAGUUAUGGCACACUAGAGCAAGGGCGAGGGAA</td>
</tr>
<tr>
<td>ALAS2  CACCGGCGUUGUGGCGAGUUAUGGCACACUAGAGCAAGGGCGAGGGAA</td>
</tr>
<tr>
<td>ALAS1  CUAGAAAUAUUCUGGACAGUAGUAAUACUCAUGUGACUAGACGCGAGCGYGACCC</td>
</tr>
<tr>
<td>ALAS2  CCACCCACAAUACUCAGGCACCAGUAGAUUCACUGAGCCAGCGAGCGYGACCC</td>
</tr>
<tr>
<td>ALAS1  UCCAGGGAAGAAGGCGCCACUCUUGUUCUCCUCUGUUGUGCACACAGACACCCC</td>
</tr>
<tr>
<td>ALAS2  UCCACCAAGAGGCAUCACCCUCUGUCUCUUUGUUGCCAAAUCACUCACCC</td>
</tr>
<tr>
<td>ALAS1  UC</td>
</tr>
<tr>
<td>ALAS2  UC</td>
</tr>
</tbody>
</table>

**Figure 5-10 mRNA sequence homology of ALAS1 fragment 2, and the corresponding nucleotides of ALAS2.**

Using EMSAs of overlapping fragments of the ALAS1 coding region, fragment 2 was shown to bind to HepG2 cytosolic protein. Fragments 5 and 6 which overlap fragment 2 were also shown to bind protein, and their boundary is shown. ALAS1 and ALAS2 have a sequence homology of 61% in this region. The non-homologous nucleotides are in red. The ALAS2 sequence used for this homology analysis contains exon 4 which is often alternatively spliced.
The protein-binding region of fragment 4 has been reduced to the last 243bp of this fragment, and of the full ALAS1 coding region. There is 56% homology between this 243bp region of ALAS1 and the corresponding sequence of ALAS2 (figure 5-11), and there is 59% homology of the 142bp putative cis-acting element. This reasonably low homology may suggest that this specific region is involved in the haem-destabilisation of ALAS1 mRNA, but it may not simply be this region alone that is required for the mechanism. Figure 5-11 shows that the longest length of non-homologous nucleotides is only 5 bases long. However, mapping of the required mRNA sequences of c-myc for binding by the CRD-BP demonstrate that essential sequences do not have to be directly next to one another, and only short nucleotide stretches are necessary for binding (Sparanese & Lee, 2007).

| ALAS1 | CCUCACUGUCCCCGGGAGAGAGCUCCUAAGUUGCCCCCACCCCCUGACCACACTCC |
| ALAS2 | CCACUGUCCCCGGGAGAGAGCUCCUGCGUGCCACCCCCACACCACACCCC |

| ALAS1 | CAGAUGAUGAACUUCCUUGAGAUCUGCUAGUCACAGGAAGCAAGUGGGCCUGAA |
| ALAS2 | CAGAUGAUGAACUUCCUUGAGAUCUGCUAGUCACAGGAAGCAAGUGGGCCUGAA |

| ALAS1 | CUGAGCCUCACUCCAGCUGAGCCACUCUCUGAGGCGACACUCCAUUUUGAAGUG |
| ALAS2 | CUGAGCCUCACUCCAGCUGAGCCACUCUCUGAGGCGACACUCCAUUUUGAAGUG |

| ALAS1 | AUGAGUGAAAGGAGAAUCCUACUUCUCAGCUGAG---CAUGUGUACUUCUG |
| ALAS2 | AUGAGUGAAAGGAGAAUCCUACUUCUCAGCUGAG---CAUGUGUACUUCUG |

| ALAS1 | GCCUGA |
| ALAS2 | GCCUGA |

**Figure 5-11** mRNA sequence homology of the protein binding-region of ALAS1 fragment 4 and the corresponding nucleotides of ALAS2.

The use of overlapping RNA sequences in EMSAs enabled a protein-binding region to be defined to a 243nt sequence at the 3’ end of the coding region. ALAS1 and ALAS2 have a sequence homology of 56% in this 243bp region of fragment 4. Underlined is the previous binding site identified by Cable et al., truncated to 142bp as we have ruled out the first 12bp of this sequence. This region has 59% homology between human ALAS1 and ALAS2 mRNA. Non-homologous nucleotides are in red.
In hepatocytes of the rat and chicken, ALAS1 is destabilised by haem, as in human cells (Cable et al., 2000; Hamilton et al., 1991; Drew & Ades, 1989; Cable et al., 2001). Therefore, a similar mechanism of haem-destabilisation of ALAS1 mRNA is surmised to occur amongst these species. The sequence homology of the 142bp putative cis-acting element of the rat, chicken and human ALAS1 gene was compared and shown to have 71% homology (figure 5-12). This high homology suggests that this region may contain specific sequences involved in the instability mechanism. The catalytic domain of the enzyme resides in its C-terminus, which may account for the high homology in this region (Dailey et al., 2005; Sadlon et al., 1999). However, the homology of this 142bp sequence between the human, rat and chicken ALAS1 with human ALAS2 is 57%.

| Human | CCUACGUUCCCGGCAGAGGCAGCUCCGUCCCGGCACCCUCACCCACACACCC |
| Rat   | CCACACGUUCCCGGCAGAGGCAGCUCCGUCCCGGCACCCUCACCCACACACCC |
| Chicken | CCACACGUUCCCGGCAGAGGCAGCUCCGUCCCGGCACCCUCACCCACACACCC |

| Human | CAGAUGAUGAAUCUCUCUGGAAUCUGCUAGUCACAUGGAAGCAGGAGGCGUGGA |
| Rat   | CAGAUGAUGAAUCUCUCUGGAAUCUGCUAGUCACAUGGAAGCAGGAGGCGUGGA |
| Chicken | CAGAUGAUGAAUCUCUCUGGAAUCUGCUAGUCACAUGGAAGCAGGAGGCGUGGA |

** ***** ******   ** ******  ***** ******

| Human | CUGAAACCUCAUCUCUCACUG |
| Rat   | CUGAAACCUCAUCUCACUG |
| Chicken | CUGAAACCUCAUCUCACUG |

** ***** ******   ** ******  ***** ******

**Figure 5-12 Sequence homology of the mRNA 142bp putative protein-binding sequence between species.**

Homologies: human and rat – 85%; human and chicken – 77%; rat and chicken – 75%; human, rat and chicken – 71%.

Non-homologous nucleotides of ALAS1 between the three species are shown in red, homologous nucleotides of ALAS1 are indicated by stars.

The EMSAs and UV-crosslinking used in this chapter required extensive optimisation for any protein-binding to the radiolabelled RNA probes to be visualised. In the context of this thesis, it was unfortunate that the crosslinking could not be optimised further to establish the specific proteins that bind and their role in haem-mediated destabilisation of ALAS1 mRNA. The use of increasing concentrations of heparin with rRNA in the UV-crosslinking assays was unable to determine a single band. Instead, it seemed that approximately four proteins bound to probe 2, and two main proteins bound to probe 4. These may be
different proteins or subunits of a single protein binding to each probe. Further analysis of protein-binding to ALAS1 RNA using EMSA and UV-crosslinking would have been useful to analyse the specificity of the binding, define the binding region and characterise the protein.

The majority of EMSAs normally use radiolabelled RNA varying in length from approximately 50 to 350nt. In general, however, short RNAs are easier to analyse since their retarded complexes can be well-resolved from the position of migration of naked RNA. The overall acrylamide content of the bandshift gels as well as their acrylamide:bis-acrylamide ratio can be empirically varied to suit the purpose; in practice it becomes very difficult to handle gels of less than 4% polyacrylamide. Therefore, further definition of the binding site is required in order to minimise the size of the RNA used in the assay. The assays in this chapter used probes of over 480nt in length, which meant that the difference between a shifted RNA and an unbound-RNA was not well resolved. It is necessary to further define the protein-binding region using a panel of truncated RNA fragments, therefore a larger shift would be visualised. This would allow the apparent equilibrium dissociation constant ($K_d$) for the RNA-binding protein to be calculated from quantifying the radioactivity in the free RNA band (or in the complex) by film densitometry or scintillation counting of the excised gel bands. An approximate $K_d$ corresponds to the protein concentration required to bind half the RNA, provided the RNA concentration is very low compared to the protein concentration at midpoint. Therefore, the $K_d$ for the RNA-binding protein could then be accurately compared with protein extracted from HepG2 cells grown in haem-depleted or haem-excess medium, to quantify the haem effect on protein binding (Walker et al., 1998; Ross, 1995).

Using probes of a shorter length than used in the EMSAs in this chapter would allow the specificity of the binding to be analysed, using specific and non-specific competitors. This is a crucial experimental factor, to demonstrate that the interaction is of a specific nature, as proteins can in fact bind to nucleic acids non-specifically. Specific binding can be achieved by carrying out the binding assay in the presence of appropriate competitor molecules, which can be specific or non-specific. Specific competitors would be unlabelled RNAs, identical in sequence to the radiolabelled RNA probes. When added in excess to a binding reaction, a specific competitor should displace the labelled RNA. Non-specific competitors are ideally random portions of RNA, or polyribonucleotides, and should not include the RNA sequence under study. These will not compete with the specific RNA
probe. In the EMSAs and UV-crosslinking experiments in this chapter, rRNA is included in the binding reaction which reduces the number of non-specific complexes formed with the labelled probe. Also, heparin was added to quench the non-specific binding of proteins to RNA. Further experiments using specific competitors with probes 2 and 4 are required. However, our experimental conditions, using heparin and rRNA as non-specific competitors, are highly stringent and provide independent confirmation of protein binding to the 154bp region as published (Cable et al., 2001).
CHAPTER 6
Discussion
The regulation of haem in erythroid cells has been researched extensively. In contrast, however, the mechanism by which haem is regulated in non-erythroid cells, for instance by regulating the ALAS1 mRNA, has been poorly characterised. Therefore, the aim of this thesis was to investigate this post-transcriptional mechanism to examine how haem destabilises ALAS1 mRNA in humans.

In summary, there are a number of important observations made from the data presented in this thesis:

1. The minor ALAS1 isoform is poorly translated and relatively resistant to haem-mediated decay, compared to the major form.
2. The human ALAS1 mRNA contains a CRD that mediates its haem-sensitivity. The 5’- and 3’-UTRs are not involved in this mechanism.
3. The ALAS1 mRNA CRD has to be translated to function.
4. The use of overlapping deletion fragments of the ALAS1 coding region is unable to further delineate the CRD.
5. RNA-EMSAs defined two fragments of the ALAS1 coding region that can bind to HepG2 cytosolic protein.

This thesis has shown that the destabilisation of ALAS1 by haem is not simply a trans-acting factor binding to a single cis-acting element in the mRNA. The mechanism involved may be more complicated than previously thought. Potential mechanisms of ALAS1 mRNA haem-mediated destabilisation will be discussed in this chapter.

A reason for the previous lack of understanding of the haem-mediated destabilisation of ALAS1 mRNA could be that it has proven difficult to establish continuous cell-culture models that retain high levels of mRNA, protein and activity of ALAS1 as occurs in intact organisms and also retain inductive or repressive responses to chemicals. Virtually all mammalian hepatoma cell lines exhibit very low levels of ALAS1, and these levels are only weakly inducible (1.5 to 2-fold induction), compared to primary cells (Iwasa et al., 1989). Low levels of ALAS1 mRNA in these cell lines have hampered work to understand the mechanism of destabilisation. Indeed, a similar problem was encountered in this project to enable a difference in ALAS1 mRNA levels in response to haem to be visualised in HepG2 cells. Therefore, we successfully used a method to deplete the cells of haem and haemoproteins, with the use of haem-depleted media and
succinyl acetone, to induce the levels of ALAS1 mRNA, prior to the addition of haem to the cells.

A further problem with the experiments throughout this thesis was the dose of haem used to treat the cells. Sassa (2004) identified that free haem within a normal cell is below 1μM. Free haem (0.1-0.3μM) added to the medium of chick embryo liver cell culture specifically inhibited the induction of ALAS1 by one-half. However, free haem concentrations of 1μM caused the induction of HO-1, thereby breaking down the haem. In this thesis a dose of 20μM haem was used to decrease the stability of the ALAS1 mRNA. Therefore, a lower dose may have induced a larger decrease in the mRNA stability with haem. However, the experiments by Sassa (2004) were looking at the endogenous genes rather than reporter assays. Research using reporter assays to monitor ALAS1 mRNA activity, such as that by Kolluri et al (2005), used 20μM haem, similar to the concentrations used throughout this thesis. Therefore, ALAS1 within a reporter construct may require a higher concentration of haem to be destabilised. However, in this situation the endogenous HO-1 gene will also be induced, and this may affect the mechanisms involved in ALAS1 mRNA destabilisation within the cell.

6.1 Translationally-dependent haem regulated decay

The major and minor ALAS1 mRNA isoforms were studied to establish if their expression differs in response to haem, and if so, would this offer a clue to the mechanism. The data in this thesis showed that the minor isoform is in fact resistant to haem-mediated decay compared to the major, and its 5’-UTR inhibits translation. We showed that sequence elements that are involved in this haem-stability mechanism are present in the coding region, and not in the 5’- and 3’-UTRs. Furthermore, by using reporter assays and inserting a stop codon prior to the ALAS1 coding region sequence, it was found that haem only destabilises the ALAS1 mRNA when it is translated.

The question arises as to why haem-regulated decay of ALAS1 mRNA should be translationally-dependent. ALAS1 protein is rapidly turned over in the mitochondria with a half-life of approximately 70 minutes in rat liver (Anderson et al., 2000). Cap-mediated translation drops sharply during periods of stress or starvation (Clemens, 2001), whilst translation rates are low during mitosis (Ross, 1997; Pyronnet & Sonenberg, 2001), with general stabilisation of mRNAs, which
may be a survival mechanism that ensures cells are not depleted of labile mRNAs as they enter G1 (Ross, 1997). Transcription and translation are inhibited by approximately 75% between pro-metaphase and telophase. In later telophase or shortly thereafter, transcription and translation rates return to normal (G1) levels (Prescott et al., 1976). Therefore, the CRD may have to be translated to cause haem-mediated decay to prevent any harmful consequences of ALAS1 depletion during periods when translation rates are low. This preserves the capacity of the cell to synthesise haem immediately upon the end of mitosis, when the cell cycle enters G1 phase again, or the cessation of stress. In addition, it is interesting to speculate that the minor ALAS1 mRNA might be translated under these conditions to ensure that there is sufficient haem for cell viability.

6.2 Delineation of the CRD

This thesis has shown that the destabilisation of ALAS1 by haem is not a simple mechanism of a trans-acting factor binding to a single cis-acting element in the mRNA. DLRA's showed that a haem-responsive element was present in the ALAS1 mRNA coding region, and was able to function independently of the 5' and 3'-UTR. However, reporter assays using dissected sequences of the coding region failed to identify a specific region that can confer haem-mediated destabilisation. This demonstrates that a CRD alone is not enough for destabilisation of the ALAS1 mRNA to occur. Furthermore, the ALAS1 mRNA requires active translation through its full-length coding region to be destabilised.

6.3 Cytosolic protein binding to ALAS1 mRNA

Cytosolic proteins from HepG2 cells bound to two separate fragments (numbered 2 and 4) of the ALAS1 mRNA coding region, using RNA-EMSAs. Fragment 4 contains the 154bp region previously identified sequence as a putative cis-acting element (Cable et al., 2001). By using overlapping ALAS1 coding region sequences we were able to define the putative cis-acting element further to a sequence of 142bp. An in vitro method of adding haemin to the binding reaction of the EMSA of fragment 4 did not have any affect on protein binding to the ALAS1 coding region. Therefore, HepG2 cells were grown in medium that was either depleted or supplemented with 20µM haem. The representative EMSA in chapter 5 shows that the addition of haem prior to the extraction of protein caused a decrease in protein binding to probe 4. However, this result was not demonstrated each time. Indeed, Cable et al were also not able to demonstrate a
reproducible effect of haem on the putative-binding site, even though removal of this site prevented haem-mediated decay (personal communication). A similar observation has been made with the regulation of the CYP2E1 gene by insulin. Insulin controls CYP2E1 expression by destabilising its mRNA. However, Moncion et al (2002) were unable to demonstrate that insulin affects protein binding to the CYP2E1 CRD. Therefore, it has been suggested that it may be protein-protein interactions mediating the destabilisation of the mRNA, or alternatively the RNA-binding proteins may constitute a scaffold upon which the actual target of insulin can bind and exert a RNA destabilisation function (Moncion et al., 2002).

Another explanation as to why haem does not seem to affect protein binding to the mRNA of the putative ALAS1 CRD in vitro may be that haem directly binds to the ribosomes or regulates endonucleolytic mRNAs. This may be similar to the destabilisation mechanism of c-myc, which will be described in section 6.4.1. Furthermore, haem-regulated binding of protein to ALAS1 mRNA may require active translation of the mRNA. Therefore, this is not observable as a simple RNA/protein interaction, and is not affected by haem in an EMSA.

6.4 Mechanisms of mRNA destabilisation involving CRDs

The emphasis on this thesis was to discover which sequence elements of ALAS1 mRNA are responsible for haem-mediated decay and to study the proteins that interact with them. The actual downstream reactions, including possible poly(A)-shortening, decapping and ribonuclease cleavage that result in its decay are not part of the study. The conclusions from this thesis allow a number of possible mechanisms of the destabilisation of ALAS1 mRNA to be proposed, based on a comparison of mechanisms from known CRDs. These include c-myc, c-fos, β-tubulin and plasminogen activator inhibitor type 2 (PAI-2), which are discussed below.

6.4.1 c-myc

The c-myc mRNA contains an instability sequence within its coding region, located in the last 249 nucleotides. It functions independently of an ARE to make the mRNA unstable (Wisdom & Lee, 1991; Yeilding et al., 1996). When the c-myc CRD is inserted in frame with the coding region of β-globin mRNA, the resulting chimeric mRNA is destabilised (Herrick & Ross, 1994). This CRD must be translated to destabilise the mRNA (Wisdom & Lee, 1991). The CRD interacts with
a 68-kDa CRD-binding protein (CRD-BP), which contains two RNA recognition motifs and four hnRNP K domains (Doyle et al., 1998). When the CRD-BP is bound to the c-myc mRNA, the CRD is shielded from endonucleolytic attack. The mRNA is then only degraded by an ARE-dependent deadenylation pathway. If the CRD-BP dissociates from the mRNA, the CRD becomes exposed to the endonuclease. The mRNA is then rapidly degraded by endonucleolytic cleavage within the CRD (Sparanese & Lee, 2007). Furthermore, ribosomes pause within the 5' segment of the CRD, at a translation pause site. Pausing generates a ribosome-free downstream region, between the pause site and the translation termination site. Pausing can be induced by several factors, including mRNA structure, the translation product itself, mRNA-binding proteins, signal recognition particle binding and tRNA abundance and rare codons. This unprotected site is attacked by an endonuclease, unless it is protected by CRD-BP during the pause interval. After elongation resumed, the ribosomes would displace the CRD-BP from the mRNA, and translation is completed. The c-myc CRD is a bipartite mRNA instability element, consisting of two juxtaposed sequence elements. The first 72 nucleotides contain the codons that induce pausing. The downstream 177 nucleotides contain sites for CRD-BP binding and endonuclease cleavage (Lemm & Ross, 2002).

6.4.2 c-fos

The c-fos proto-oncogene encodes a transcription factor that appears to play a role in cell growth and differentiation (Sheng & Greenberg, 1990). Transcription of the c-fos gene is transiently induced within minutes after stimulation by growth factors and other agents. Shortly after induction, c-fos transcription shuts-off again, returning to its very low basal level within 30-60 minutes after growth factor addition. The c-fos message persists in the cytoplasm for 30-45 minutes after the period of ongoing c-fos transcription because, upon arriving in the cytoplasm, this labile message is degraded extremely rapidly (Greenberg & Ziff, 1984). The multiple destabilising elements within the c-fos coding region work in concert to achieve the destabilising effect of the intact c-fos coding region. Translation of the c-fos coding region, or at least the assembly of ribosomes at the 5' end of the coding region, is necessary for these determinants to effect swift deadenylation and decay of c-fos mRNA (Schiavi et al., 1994). The requirement for translation could also be explained if there exists a ribosome-associated ribonuclease or degradative factor that is activated when the ribosome interacts with these destabilizing elements.
c-fos mRNA contains a 320-nucleotide CRD that functions as an instability element only when the translation mechanism has initiated (Shyu et al., 1991; Schiavi et al., 1994). However, the c-fos CRD differs in several ways from the c-myc CRD: firstly it is recognized by virtue of its sequence or structure, and secondly it promotes accelerated deadenylation, not endonucleolytic decay. A complex of five proteins (PABP, PABP-interacting protein (PAIP), AUF-1, NS1-associated protein (NSAPIP1) and Unr (Upstream of N-ras)) assembles on the mCRD and promotes rapid deadenylation and decay of the mRNA. One model proposes that transit of ribosomes through the mCRD element disrupts the complex and triggers the decay pathway. In this case, the mCRD complex might protect untranslated c-fos mRNA from the rapid deadenylation-dependent decay that is otherwise promoted by the ARE in its 3'-UTR (Grosset et al., 2000).

If the CRD is placed in frame with the globin mRNA coding region, the resulting 5'-globin-fos-globin-3' chimeric transcript is at least four-fold less stable than globin mRNA. If a frameshift mutation is introduced into the globin-fos-globin gene, such that the mRNA sequence is changed by only a single nucleotide but the peptide encoded by the c-fos region is entirely different, the frameshifted mRNA is just as unstable as the original globin-fos-globin mRNA. Therefore in this case the structure of the c-fos mRNA itself specifies the instability phenotype, independent of the encoded protein (Wellington et al., 1993). In contrast, the destabilisation of β-tubulin mRNA by tubulin monomers depends on the peptides encoded by their CRD’s (Yen et al., 1988).

6.4.3 β-tubulin

Microtubules are dynamic filamentous structures that are composed of tubulin monomers. Tubulin is a heterodimer of α- and β-subunits that exist in a dynamic equilibrium with the microtubule polymer (Guhaniyogi & Brewer, 2001). The expression of β-tubulin, one of the two principle constituents of microtubules, is regulated in higher eukaryotes by an autoregulatory mechanism that modulates the stability of the polysome-bound tubulin mRNAs (Cleveland, 1989). The stability of tubulin mRNA is specified by the level of its own translation products, the tubulin subunits themselves (Theodorakis & Cleveland, 1992; Gay et al., 1989). β-tubulin mRNA levels are sensitive to the intracellular concentration of tubulin subunits (Yen et al., 1988). The first 13 translated nucleotides of β-tubulin mRNA are sufficient to confer autoregulated instability onto heterologous RNAs if
present as the first translated nucleotides. Furthermore, for β-tubulin mRNAs to be substrates for autoregulation they must be associated with polysomes and that translation of them must proceed past 41 codons. The binding of cellular cofactors might also be involved (Bachurski et al., 1994). An endonuclease is suspected to be ribosome-associated which is activated to cleave the mRNA by β-tubulin binding to the nascent peptide, although this has not yet been demonstrated (Yen et al., 1988).

6.4.4 Plasminogen activator inhibitor type 2 (PAI-2)

The plasminogen activator system is an important proteolytic cascade that plays a role in the removal of blood clots from the circulation and the turnover of a variety of extracellular matrix proteins. PAI-2 is used in this system as a regulated serine protease inhibitor. Similar to ALAS1 mRNA, the PAI-2 transcript is unstable in the absence of the 3′-UTR (Maurer et al., 1999). Therefore, Tierney & Medcalf (2001) examined the possibility of the PAI-2 mRNA containing CRD’s. They demonstrated that individual exons or groups of exons of the PAI-2 gene could confer varying degrees of instability in a reporter assay, suggesting that multiple instability elements reside within the PAI-2 coding region.

To further define one of the protein-binding sites in PAI-2, Tierney & Medcalf (2001) used antisense DNA-masking experiments. In this approach, antisense DNA oligomers that were directed against overlapping sequences within the RNA probe were allowed to anneal. This inhibited binding to a specific cis-acting site, allowing this region to be identified. In the ALAS1 RNA-EMSAs in chapter 5, we identified two protein-binding regions of 480bp and 142bp in size. Therefore, it would be useful to use similar DNA masking experiments to further delineate the ALAS1 protein binding regions in human ALAS1 mRNA.

6.5 Potential mechanism of ALAS1 mRNA destabilisation by haem

A proposed mechanism of how haem destabilises the ALAS1 mRNA in hepatocytes needs to involve a number of observations made from this thesis. Firstly, fragment 4 alone is not sufficient to mediate the destabilisation; additional sequence elements in the ALAS1 coding region are required. Secondly, the destabilisation of ALAS1 mRNA is translation-dependent.
Similar to the mechanism used by c-myc, ribosomal pausing may occur in the ALAS1 mRNA. Ribosomal pausing is suggested to account for the following observations:

1. Placing a stop codon prior to the ALAS1 mRNA coding region inhibits its destabilisation in the presence of haem. This demonstrates that translation through the ALAS1 coding region is necessary for its destabilisation.
2. The region of ALAS1 mRNA containing a potential CRD is translated relatively inefficiently. Therefore, this is likely to cause pausing and result in a ribosome-free downstream region.

To establish if there is any ribosome-free regions in the ALAS1 mRNA coding region, ribosome toeprinting assays (translation-couple primer extensions) could be used, similar to those used on c-myc by Lemm & Ross (2002). These would confirm that translational inefficiency of particular fragments of ALAS1, such as fragment 4, is the result of ribosomal pausing.

Similar to the mechanism of c-myc and PAI-2 destabilisation, the ALAS1 CRD may be of a bi-partite structure. Of interest was the observation that fragment 4 dramatically reduced translation of the Renilla luciferase reporter gene. However, this region, by itself, is unable to confer haem-mediated instability on the heterologous protein, suggesting that other important elements in the coding region are also required. These may include another ALAS1 mRNA element that binds to protein, such as fragment 2, identified by the RNA-EMSAs in chapter 5. The downstream site in fragment 4 may be where the CRD-BP binds and cleavage can occur. It could be that an essential ribonuclease or mRNA binding protein is associated with the ribosome, or, that translational elongation is required for the displacement of a stabilising protein bound to a CRD, perhaps through melting of specific mRNA secondary structure.

Interestingly, using a translational-competent cell-free system, Yamamoto et al showed that haem rapidly inhibited ALAS1 translation by a mechanism that required translational elongation (Yamamoto et al., 1983). mRNA concentrations were not monitored however and so the contribution of mRNA decay to the reduced rate of ALAS1 production in those experiments cannot be evaluated. However, as suggested earlier, it may be that the initial regulated signal for mRNA decay involves haem-dependent inhibition or pausing of translation within a specific sequence of the ALAS1 coding region. The experiments on ALAS1 used a stop codon prior to the ALAS1 coding region sequence to establish that
translation is essential for the haem-mediated destabilisation of the mRNA. However, we did not investigate whether it is the encoded protein or the structure of the mRNA itself that specifies the haem sensitivity. To test this, as with the c-fos experiments, a frameshift mutation of ALAS1 could be used at the beginning of the coding region. This would establish if the haem-destabilisation mechanism is similar to c-fos or β-tubulin.

A mechanism of ALAS1 destabilisation by haem is proposed in figure 6-1 in which the tertiary structure of the mRNA is involved. This may either involve base pairing between two sequence elements within the coding region, for example between fragments 2 and 4; or involve protein interactions which cause the RNA to fold in a specific way. During translation ribosomes displace the RNA-binding proteins and cause a change in the folding of the mRNA and may expose a site that is degraded by a ribonuclease, or trigger a series of events for the mRNA to be degraded by an unknown mechanism. Furthermore, the ribosomes may have an associated protein such as a haemoprotein or endonuclease (Lee et al., 1998), which is induced by haem. It is not known whether fragment 2 of the coding region is involved in the mechanism. A further part to our proposed mechanism may be that ribosomal pausing occurs over a critical region within the 142bp sequence in the mRNA, allowing a ribosome-free downstream region to be degraded by a ribonuclease, similar to the c-myc mechanism. This would account for the reduced translation of the fragment 4 ALAS1 mRNA used in the reporter assays. Indeed, it is interesting to speculate that haem may regulate this mechanism by binding to the HRMs of the emerging peptide, analogous to the destabilisation mechanism of β-tubulin. As discussed earlier, Yamamoto et al. (1983) demonstrated that haem inhibits translation of the ALAS1 mRNA, and they also proposed that this might involve a specific interaction of haem with nascent peptide chains of ALAS1 on the polysomes.
Figure 6-1 A proposed mechanism of human ALAS1 mRNA destabilisation by haem in hepatocytes.

The two fragments (2 and 4) of the ALAS1 coding region are linked by a specific tertiary structure of the mRNA, which binds to a protein-scaffold. During translation, the ribosomes displace the RNA-binding proteins. The high haem concentration within the cell and the displacement of mRNA binding-proteins, trigger a sequence of events to degrade the ALAS1 mRNA by an unknown mechanism.

6.6 Recent literature on the haem control of ALAS1

Specific processes appear to have evolved to rapidly regulate haem concentration within the cell. Recently, Yoshino et al (2007) demonstrated that haem regulates its own synthesis by causing the degradation of the ALAS1 enzyme in the mitochondria. The homeostasis of haem is therefore controlled by three post-transcriptional mechanisms: by inhibiting the ALAS1 enzyme from entering the mitochondria; by destabilising the ALAS1 mRNA; and finally by proteolysis of the enzyme within the mitochondria. The homeostatic regulation of haem is likely to be maintained and finely tuned in the mitochondria. The increased instability of ALAS1 mRNA is therefore likely to be more physiologically relevant in reducing the high concentrations following transcriptional induction.
CHAPTER 7
References


I


J


---


APPENDIX
Appendix 1

Amino acid alignment of the human ALAS1 and ALAS2 proteins (using the amino acid letter code)

Homologous amino acids are starred and indicated in blue text. The amino acids are numbered at the end of each line.

* * * * * * * * * * * * * * * * *

ALAS1  ---MESVVRRCPLSRFPQAFILQAGKSLLFYAQNCFKMMEVGKPAPRALSTAAVHYQQ
ALAS2  MVTAAMLQCCFVLARGTSLGLGKVTHQFLFG------IGRCP---ILATQGFNCQ

* * * * * * * * * * * * * * * * *

ALAS1  IKTTPASEKDKTAKAVQQTPDGSQPSDGTQLPSHFLPSQTSGTSHKCPFGAAQMNQ
ALAS2  IHLKATKAGGDS-------------------------------PSWAKG---HCPMLSELDQ

* * * * * * * * * * * * * * * * *

ALAS1  RGSSVFCKASLQEDVQEMNAVRKEVAETSAGPSVSVKTDGGDFSSLKKNFDIQMKQ
ALAS2  GKSIVQKAAPEVQEDVKA------KTDLSSVSLVR-KFSGPQ---QE

* * * * * * * * * * * * * * * * *

ALAS1  RPRVSHLLQQDNLFKSVSTFQYDFRFEEKIDKEKRNDHTYRVTVRNRAHFFMADDYSD
ALAS2  ISGKVTTHLQNMFGNY-VFSYDQFFRDKIMEKQDHTYRVTVRNWADAYPFAQHFSE

* * * * * * * * * * * * * * * * *

ALAS1  SLITKQVSVWCSNDYLMGRSHRPVCAGMVLKQHGAGAGGTRNISGTKFHVLEREL
ALAS2  ASVASKDVWVCNLYMGSRHFPQLQATQETLQRHGAGGAGGTRNISGTKFHVLEQEL

* * * * * * * * * * * * * * * * *

ALAS1  ADLHGKDALLFSSCFVANDSTLTAKMVMPGCEISSDGSNGHSMIQGIRRNSRPKYIFR
ALAS2  AELHKGDSALLFSSCFVANDSTLTAKILPGCEISSDAGHSMIQGIRRNSGAARKFVFR

* * * * * * * * * * * * * * * * *

ALAS1  HNDVSHLRELLQRDPSVPKIVAFETVHSMDGAVCPLEELCDVAHEFGAIETFVDEVHAVG
ALAS2  HNDPDLKILKLEKSNPKIPKIVAFETVHSMDGACPLEELCDVSHQYGALTTFVDEVHAVG

* * * * * * * * * * * * * * * * *

ALAS1  LYGARGGGIGDRGVMKDIISGTGLKAFGCVCWGYIATSSILDSDVYAAAGFIPTTL
ALAS2  LYGSGAGIGERDGIMKDIISGTGLKAFGCVCWGYIATRDLVMVRSYAAAGFIPTTL
*** * ******** * * * ******** * * * *** *** *** * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

** ALAS1 **
PPMLLAGALESVRILKSAEGRVLRQPGRNVRKLRLQMLMDAGLPVHCPHSIHPVRA
537

** ALAS2 **
PPMVLSGALESVRLLKGEQQALRRAHQRNVRKHMRQLMDRGLFVIPCPSHIIIPRGN
483

** ALAS1 **
AKMTEVCODLMRSRNHYVQAINYPTVRGEEELLRIAPTGHPTQRMNYFLENILVTWQV
597

** ALAS2 **
ALNSKLCDLLSKHGTYVQAINYPTVRGEEEYLLALSPHSPQMEDFVEKLILMWTAV
543

** ALAS1 **
GLELKHSSAECNFCCRPLHEVMSERKFSGLS-KLVSAQA
640

** ALAS2 **
GLPIQDVSAACNFCCRPHFELMSWERSYFGNMGFQYVTYA
587
### Appendix 2

**Alignment of the full ALAS1 genes of the human, rat and chicken.**

Homologous nucleotides are indicated by stars.

<table>
<thead>
<tr>
<th></th>
<th>HUMAN</th>
<th>RAT</th>
<th>CHICKEN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HUMAN</strong></td>
<td>CTGTAATTAAAGCCGCGGCTGTGGATGCTGGGCTGGAAGAGCACAGGCGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>GTTTCTTTGGAGCTTCTGATGCTGGCCGCCCTGCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>CAGCGAGTTAGGCGCTGCTGGGAGACAGGAGAAGTCGCTGCTG CGTGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>GAAGGCACTGAGCTGCTGGTTCAGCTGCTGAGGCTGCTGCAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>TAGGTCTTTCCAGGAGGCGAGCGATATGAGTGGTGGTTCGGCCTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>CCCATTCCTATCCCGAGTCCCAGGCTTCTGAGAAAGCGAGGAAATCTGCTGTT</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>CTATGCCCCAACCTGCGAAAGATGTGGAAGTGGGCGCGCCAGCCGCTGCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>GCCCACTGCAGCAGTAGCTACACCACAGATCAAAGAADDCCCTCCGCCAGTGAGAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>AAAACTGTCCAAGGCGACCTGCTAGCTGGATCCAGCGGCAAGCAGGCATAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HUMAN</strong></td>
<td>AGATGGCAGACAGCTTCCGCTGGAACCCCTTGGCTGCCAAGGAGGCGACACTGCG</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CHICKEN</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3

Suppliers’ full addresses

Ambion
Applied Biosystems
Lingley House
120 Birchwood Boulevard
Warrington
WA3 7QH

Amersham Pharmacia Biotech
Amersham Biosciences UK Ltd
Amersham Place
Little Chalfont
Buckinghamshire
HP7 9NA

Beckman Coulter
Beckman Coulter (UK) Ltd
Oakley Court
Kingsmead Business Park
London Road
High Wycombe
Buckinghamshire
HP11 1JU

Bioline
Bioline Ltd.
16 The Edge Business Centre
Humber Road
London
NW2 6EW

Biotecx
AMS Biotechnology (Europe) Ltd
63B Milton Park
Milton
Abingdon
Oxon
OX4 3LY

ECACC
http://www.ecacc.org.uk/

Fisher Scientific
Fisher Scientific UK Ltd
Bishop Meadow Road
Loughborough
Leicestershire
LE11 5RG

Invitrogen
Invitrogen Ltd
3 Fountain Drive
Inchinnan Business Park
Paisley
PA4 9RF
**New England Biolabs**
New England Biolabs (UK) Ltd.
75/77 Knowl Piece, Wilbury Way
Hitchin, Herts.
SG4 0TY

**Oxoid**
Oxoid Ltd
Wade Road
Basingstoke
Hampshire
RG24 8PW

**Pierce Biotechnology, Inc.**
P.O. Box 117
Rockford
IL 61105

**Promega**
Promega UK Ltd
Delta House
Chilworth Research Centre
Southampton
SO10 7NS

**Qiagen**
Qiagen Ltd
Boundary Court
Gatwick Road
Crawley
West Sussex
RH10 9AX

**Sigma-Aldrich**
Sigma-Aldrich Company Ltd
The Old Brickyard
New Road
Gillingham
Dorset
SP8 4XT
<table>
<thead>
<tr>
<th>Orientation</th>
<th>5' - 3' Sequence</th>
<th>Restriction Enzyme Site</th>
<th>Target Sequence Primer Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>sense</td>
<td>CGACGCTGCTTATCAGAGTCAG</td>
<td>Eco RI</td>
<td>ALAS1 Fragments 2</td>
</tr>
<tr>
<td>antisense</td>
<td>CGGCTGCTTATCAGAGTCAG</td>
<td>Kpn I</td>
<td>ALAS1 Fragments 1</td>
</tr>
<tr>
<td>sense</td>
<td>CGACGCTGCTTATCAGAGTCAG</td>
<td>Eco RI</td>
<td>ALAS1 Start of 5' UTR</td>
</tr>
<tr>
<td>antisense</td>
<td>CGGCTGCTTATCAGAGTCAG</td>
<td>Kpn I</td>
<td>ALAS1 End of 3' UTR</td>
</tr>
<tr>
<td>sense</td>
<td>CGGCTGCTTATCAGAGTCAG</td>
<td>Eco RI</td>
<td>STOP</td>
</tr>
<tr>
<td>sense</td>
<td>CGGCTGCTTATCAGAGTCAG</td>
<td>Eco RI</td>
<td>ALAS1 Coding Region Exon 2 With STOP</td>
</tr>
<tr>
<td>sense</td>
<td>CGGCTGCTTATCAGAGTCAG</td>
<td>Eco RI</td>
<td>ALAS1 Coding Region Exon 2 From STOP</td>
</tr>
</tbody>
</table>

The table is continued on the next page.

The restriction enzyme sites within a primer sequence are shown in italics. Primers are labelled as 5'-35 and used throughout chapters.

Primers used in this thesis:

Appendix 4 - Primers
<table>
<thead>
<tr>
<th>Primer Name</th>
<th>5' - 3' Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer 1</td>
<td>GGATACGTTGATGAGCTAG</td>
</tr>
<tr>
<td>Primer 2</td>
<td>GGATACGTTGATGAGCTAG</td>
</tr>
<tr>
<td>Primer 3</td>
<td>GGATACGTTGATGAGCTAG</td>
</tr>
<tr>
<td>Primer 4</td>
<td>GGATACGTTGATGAGCTAG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Orientation</th>
<th>5' - 3' Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td>GGATACGTTGATGAGCTAG</td>
</tr>
<tr>
<td>Reverse</td>
<td>GGATACGTTGATGAGCTAG</td>
</tr>
</tbody>
</table>