THE BEHAVIOR OF HUMAN GLOBIN GENE
RECOMBINANTS IN MAMMALIAN CELLS

Paul Montague

A thesis submitted to the Council for National
Academic Awards in partial fulfilment of the
requirements for the degree of Doctor
of Philosophy

Beatson Institute for Cancer Research,
Wolfson Laboratory for Molecular Pathology
Garscube Estate,
Switchback Road,
Bearsden, Glasgow G61 3BD

February 1986
Contents

Page No.

Acknowledgements   i

Abbreviations      ii

List of Figures    v

List of Tables     xii

Abstract           xiv

Main Text

CHAPTER 1: Introduction 1

CHAPTER 2: Materials and Methods 62

CHAPTER 3: Results 105

CHAPTER 4: Discussion 209

CHAPTER 5: References 241
I would like to thank my supervisor Dr. John Paul for his advice and guidance throughout the course of this project. In addition I am indebted to a large number of colleagues both past and present for their much appreciated assistance and encouragement. In particular I extend thanks to Dr. Maggi Allan, Dr. George D. Birnie, Mrs G. Joan Grindlay, Dr. Peter Humphries, Dr. Demetrios A. Spandidos and Dr. J. Keith Vass. Finally, thanks to Frank Lombos and Denise Adam for the typing. This work was supported by the Cancer Research Campaign.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine Phosphoribosyl Transferase</td>
<td>APRT</td>
</tr>
<tr>
<td>Alpha Globin</td>
<td>α-Globin</td>
</tr>
<tr>
<td>Avian Erythroblastosis Virus</td>
<td>AEV</td>
</tr>
<tr>
<td>Baby Hamster Kidney Cell Line</td>
<td>BHK</td>
</tr>
<tr>
<td>Bacterial Alkaline Phosphatase</td>
<td>BAP</td>
</tr>
<tr>
<td>Bases (Adenine, Cytosine, Guanine, Thymidine and Uracil)</td>
<td>A,C,G,T,U</td>
</tr>
<tr>
<td>Base Pairs</td>
<td>bp's</td>
</tr>
<tr>
<td>Beta Globin</td>
<td>β-Globin</td>
</tr>
<tr>
<td>Bovine Papilloma Virus</td>
<td>BPV</td>
</tr>
<tr>
<td>Bovine Serum Albumin Fraction V</td>
<td>BSA</td>
</tr>
<tr>
<td>Calf Intestinal Phosphatase</td>
<td>CIP</td>
</tr>
<tr>
<td>Chloramphenicol Acetyl Transferase</td>
<td>CAT</td>
</tr>
<tr>
<td>Counts Per Minute</td>
<td>cpm</td>
</tr>
<tr>
<td>Curies Per Millimole</td>
<td>Ci/mmol</td>
</tr>
<tr>
<td>Delta Globin</td>
<td>δ-Globin</td>
</tr>
<tr>
<td>Dihydrofolic Acid</td>
<td>DHFA</td>
</tr>
<tr>
<td>Dihydrofolic Reductase</td>
<td>DHFR</td>
</tr>
<tr>
<td>Dimethyl Sulphoxide</td>
<td>DMSO</td>
</tr>
<tr>
<td>Dithiothreitol</td>
<td>DTT</td>
</tr>
<tr>
<td>DNase I Hypersensitive Site</td>
<td>DHS</td>
</tr>
<tr>
<td>Epsilon Globin</td>
<td>ε-Globin</td>
</tr>
<tr>
<td>Ethidium Bromide</td>
<td>EtBr</td>
</tr>
<tr>
<td>Ethylenediamine Tetra-Acetic Acid</td>
<td>EDTA</td>
</tr>
<tr>
<td>Gamma Globin Alanine or Glycine at Codon 136</td>
<td>γA/γG-</td>
</tr>
<tr>
<td>Global Grodinothreitol</td>
<td></td>
</tr>
<tr>
<td>Global Grodinothreitol</td>
<td></td>
</tr>
</tbody>
</table>
Geneticin G418
Hepes Buffered Saline HBS
Hereditary Persistence of Foetal Haemoglobin HPFH
Herpes Simplex Virus Type 1 HSV-1
Hexamethylene Bis-Acetamide HMBA
Human Erythroleukaemic Cell Line HEL
4-(2-Hydroxyethyl)-1-Piperazine Ethane Sulphonate HEPES
100 μM Hypoxanthine/0.4 μM Aminopterin/ HAT
16 μM Thymidine Medium HAT
Hypoxanthine Guanine Phosphoribosyl Transferase HGPRT
Kilobase Pairs kb
Molecular Weight MW βME
β-Mercaptoethanol βME
Moloney Murine Sarcoma Virus MoMuSV
3-(N-Morpholino) propanesulphonic Acid MOPS
Murine Erythroleukaemic Cell Line MEL
Neomycin Gene Neo
Oligodeoxythymidylic Acid Oligo dT
Optical Density Wavelength OD260/280/600
Phosphate Buffered Saline PBS
Phosphate Buffered EDTA PE
Piperazine-N, N'-bis[2-ethanesulphonic Acid] PIPES
Plaque Forming Units pfu's
Polyethylene Glycol PEG
Polyvinyl Pyrolidone PVP
Pseudo-globin Gene ψ
Ribonucleoprotein Particle RNP
<table>
<thead>
<tr>
<th>Biological Term</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rous Sarcoma Virus</td>
<td>RSV</td>
</tr>
<tr>
<td>Simian Virus 40</td>
<td>SV40</td>
</tr>
<tr>
<td>Sodium Dodecyl Sulphate</td>
<td>SDS</td>
</tr>
<tr>
<td>Special Liquid Medium</td>
<td>SLM</td>
</tr>
<tr>
<td>Standard Sodium Citrate (0.15 M NaCl/0.015 M Sodium Citrate)</td>
<td>SSC</td>
</tr>
<tr>
<td>2-amino-2-(hydroxymethyl)-1, 3-propanediol</td>
<td>tris</td>
</tr>
<tr>
<td>Aminoglycoside Phosphotransferase (Neomycin)</td>
<td></td>
</tr>
<tr>
<td>Gene: Encoded in the Bacterial Transposon</td>
<td></td>
</tr>
<tr>
<td>Tn5</td>
<td>Tn5 aph(II)</td>
</tr>
<tr>
<td>T&lt;sub&gt;4&lt;/sub&gt; Polynucleotide Kinase</td>
<td>T&lt;sub&gt;4&lt;/sub&gt;PK</td>
</tr>
<tr>
<td>Tetrahydrofolic Acid</td>
<td>THFA</td>
</tr>
<tr>
<td>N, N, N', N'-Tetramethylethylenediamine</td>
<td>TEMED</td>
</tr>
<tr>
<td>Thymidine Kinase Gene</td>
<td>tk</td>
</tr>
<tr>
<td>Tris. Borate. EDTA Buffer</td>
<td>TBE</td>
</tr>
<tr>
<td>Tris. EDTA Buffer</td>
<td>T.E</td>
</tr>
<tr>
<td>Ultra Violet Light</td>
<td>u.v.</td>
</tr>
<tr>
<td>Xanthine Guanine Phosphoribosyl Transferase</td>
<td>XGPRT</td>
</tr>
<tr>
<td>Zeta Globin</td>
<td>ζ-Globin</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>Figure 1</td>
<td>A schematic representation of differential globin gene expression during ontogeny</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Molecular arrangement of the human globin gene loci on chromosomes 11 and 16</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Comparison of the 5' flanking region of a range of β globin genes</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Structure and expression of human globin genes</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Summary of Molecular Lesions causing β- Thalassaemia</td>
</tr>
<tr>
<td>Figure 6</td>
<td>dTMP synthesis via Salvage and de novo metabolic pathways</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Schematic Eco RI mapping of plaque purified recombinant lambda phage recombinants λ M2, 3, 5, 9 and 10 isolated from a partial AluI/HaeIII human genomic liver library</td>
</tr>
</tbody>
</table>
Figure 8  Subcloning of Eco RI genomic fragments of λM3 into pAT153

Figure 9  Subcloning of Eco RI genomic fragments spanning the human β globin gene cluster into pAT153

Figure 10  Structure of the human epsilon globin gene and restriction map

Figure 11  Construction of pTKHεG and pMX

Figure 12  Construction of pTKHδβε

Figure 13  Construction of pMX Promoter Deletion Mutants: pMXΔn

Figure 14  Structural Analysis of pMXΔn Series

Figure 15  Line Diagram of pTKHεG, pMX, pMXΔn and pTKHδβε

Figure 16  Southern blotting analysis of pTKHδβε DNA from stably transformed BHKtk-, mouse LAtk- and MELtk- cells

Figure 17  Hybridisation probes used in S1 Nuclease mapping experiments
Figure 18  Strand separation of hybridisation probes 132

Figure 19  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from non-induced K562 cells 134

Figure 20  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from non-induced K562 cells under limiting probe conditions 135

Figure 21  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BKHtk-cells stably transformed with pTKHεG and pMX 139

Figure 22  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from mouse LAtk- cells stably transformed with pTKHεG and pMX 142

Figure 23  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from human 143tk- cells stably transformed with pTKHεG and pMX 143

Figure 24  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk-cells stably transformed with pTKHεG 145
Figure 25  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk-cells stably transformed with pMX

Figure 26  Splicing of epsilon globin gene transcripts in stably transformed BHKtk-cells

Figure 27  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-cells stably transformed with pMX promoter deletion mutants Δ1, Δ4 and Δ5

Figure 28  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-cells stably transformed with pMX promoter deletion mutants Δ2, Δ3, Δ8 and Δ11

Figure 29  S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-cells stably transformed with pMXΔ11 under conditions of increasing probe

Figure 30  Splicing of epsilon globin gene transcripts in BHKtk-cells stably transformed with a series of pMX promoter deletion mutants
Figure 31 Schematic representation of the transcription of the human epsilon globin gene from pMX promoter deletion templates

Figure 32 S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-, mouse LAtk- and MELtk- cells stably transformed with pTKHδβε

Figure 33 Northern blot of RNA prepared from three pMX stably transformed MELtk- cell lines in the presence and absence of 3 mM HMBA

Figure 34 S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk- cells stably transformed with pTKHεG grown in the presence and absence of 3 mM HMBA

Figure 35 S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk- cells stably transformed with pMX grown in the presence and absence of 3 mM HMBA

Figure 36 Construction of pSVODε and pSVEDε
Figure 37 The effect of plasmid replication and SV40 enhancer sequences on the S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from COS 7 and CV-1 cells acutely transformed with pSVODε and pSVEDε

Figure 38 Construction of pSVODε-TK and pSVEDε-TK

Figure 39 The effect of the SV40 origin and enhancer on the S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk- and 143tk- cells stably transformed with pSVODεTK, pSVEDεTK and pMX

Figure 40 S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from mouse LAtk- cells stably and acutely transformed with pHR11 and pMX

Figure 41 Structural differences between the plasmid recombinants pMX and pHR11

Figure 42 Construction of the pMX 3' deletion mutant pMXΔH

Figure 43 Construction of pMXΔ4 and pSXΔB
Figure 44 Line diagram of pHR11, pTKHeG, pMX, pMXΔ4, pMXΔH, pSX and pSXΔB

Figure 45 S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from mouse LAtk- cells acutely transformed with pHR11, pTKHeG, pMX, pMXΔ4, pMXΔH, pSX and pSXΔB

Figure 46 Construction of pSVKNe-3

Figure 47 S1 Nuclease mapping of the 5' termini of epsilon globin mRNA from tk-, COS 7, MELtk- and K562 cells stably transformed with pSVKNe-3

Figure 48 Diagram of pTK-1 based recombinants: pSVεTK, pεV, pTMo and pTV

Figure 49 Construction of pNV

Figure 50 Construction of pNεV and pNTV

Figure 51 Construction of pNMβV

Figure 52 Construction of pNεV10 and pNεV11
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Estimated $pTKH\delta\beta$ plasmid copy No. per stably transformed cell</td>
<td>127</td>
</tr>
<tr>
<td>Table 2</td>
<td>The distribution of epsilon globin gene transcriptional initiation sites in K562 and stably transformed BHKtk- and mouse LAtk- cells</td>
<td>140</td>
</tr>
<tr>
<td>Table 3</td>
<td>Summary of S1 Nuclease mapping data showing the distribution of 5' termini of the epsilon globin gene transcripts in K562 and stably transformed cells</td>
<td>148</td>
</tr>
<tr>
<td>Table 4</td>
<td>The effect of the promoters of the HSV-1 tk, human epsilon globin and MoMuSV genes on HSV-1 tk expression in stably transformed mouse LAtk- and human tk- cells</td>
<td>191</td>
</tr>
<tr>
<td>Table 5</td>
<td>The influence of Eco RI globin genomic fragments on epsilon globin and MoMuSV promoter driven HSV-1 tk expression in mouse LAtk- cells</td>
<td>193</td>
</tr>
</tbody>
</table>
Table 6  The effect of the human epsilon globin and HSV-1 tk promoters on Tn5 aph(II) expression in the mouse LAtk- cell and the human 143tk- cell

Table 7  The effect of different promoters on Tn5 aph (II) expression in a variety of cell types

Table 8  The effect of globin gene fragments Nos. '10' and '11' on epsilon globin promoter driven Tn5 aph (II) expression in a variety of cell types
The human epsilon globin gene has multiple sites of RNA initiation. Of these, the most abundant are the transcripts originating from the -200 promoter giving a cap: -200 ratio of 20:1 in RNA prepared from the erythroleukaemic cell line K562 and 5-10 week embryos.

A series of epsilon globin gene recombinants were introduced into a range of cell lines to compare the transcriptional profile with that obtained in the K562 cell and to identify putative regulatory regions. This approach led to the following observations.

1) Epsilon globin gene activity in stably and acutely transformed non-erythroid and MELtk- cells display a distinct tendency towards greater use of the -200 promoter. Even recombinants with up to 6 kb of 5' flanking DNA, cap: -200 ratios varied from 4:1 in mouse LAtk-, 1:1 in BHKtk- and 0:1 in MELtk- cells.

2) The major epsilon globin promoter (5'- -100/CCAAT/ATA-cap-3') is not sufficient to support optimal levels of cap site expression in stably transformed BHKtk- cells.

3) However when this structure was extended by the cis -linkage of 350 bp's of endogenous DNA, optimal cap site expression was restored.

4) When this recombinant was further extended by 1468 bp's, the cap: -200 ratio dropped from 4:1 to 1:20 in acutely transformed mouse LAtk- cells. Interestingly, this fragment contains two upstream promoters at -900 and -1480 bp's.

Using a functional assay system another putative regulatory region has been identified over 2 kb upstream of the major promoter. A 6.67 kb Eco RI fragment which contains a potential Z-DNA element and the most distal promoter at -4500 bp's markedly increases the expression of the bacterial Tn5 aph(II) gene from the epsilon globin promoter.

Additionally, several lines of evidence has been presented which suggest that the major and -200 promoters can be regulated independently by both exogenous cis linked sequences (SV40 origin and enhancer) and viral trans acting factors.

Based on these findings it is proposed that these minor promoters particularly the -200 are transcriptionally active during early embryonic erythropoietic development. During this period canonical cap site activity may in part be repressed by some sort of transcriptional interference mechanism. The switch from the upstream promoters occurs in later development mediated by an embryonic erythroid trans acting factor as experimentally demonstrated by the switch in cap: -200 ratio from 1:20 in non-erythroid cells to 20:1 in K562 cells transformed with a mutated epsilon globin gene recombinant to distinguish it from the endogenous gene.
# CHAPTER 1: Introduction

1:1 **Biology of the Human Globin Genes**

1:1:1 Background

1:1:2 Globin Gene Expression During Development

1:1:3 Molecular Organisation of the Human Globin Genes

1:1:4 The Molecular Structure of the Human Epsilon Globin Gene

1:2 **Regulation of Globin Gene Expression**

1:2:1 Replication

1:2:2 Transcription

1:2:3 Nuclear Processing

1:2:4 Nuclear-Cytoplasmic Transport

1:2:5 Message Stability

1:2:6 Translation

1:2:7 Conclusion

1:3 **The Conformation of Globin Genes in Chromatin**

1:4 **Methylation of Globin Genes**

1:5 **cis Control Sequences**

1:5:1 The Globin Gene Promoter

1:5:2 Enhancer Sequences

1:5:3 Z-DNA

1:5:4 Intergenic DNA
1:6  **trans Control Sequences**  44

1:6:1  Globin Gene Expression in Somatic Cell Hybrid Lines  45

1:6:2  Induction of Exogenous Globin Gene Expression in Transformed MELtk^- Cells  47

1:6:3  **trans** Activation of Globin Genes by Viral Proteins  50

1:7  **Aims of Project**  53

1:7:1  Experimental Approaches to Study Globin Gene Expression  54

1:7:2  Human Globin Gene Expression in Erythroleukaemic Cell Lines  55

1:7:3  Human Globin Gene Expression in Transformed Animal Cells  57

1:7:4  Experimental Plan  60
To acquire information about patterns of differential gene expression, most efforts have centred on the study of individual genetic systems. The main focus of this thesis is the epsilon globin gene of the human β globin gene cluster which has a specific association with the tightly regulated developmental pathway of haematopoiesis. The INTRODUCTION will be concerned with the structure and expression of the human globin genes (for recent reviews see [1] and [2]). Despite the importance, interest and relevance of haematopoiesis, a detailed account of the development of the red blood cell which involves the switching on and off of batteries of genes governing the synthesis of membrane proteins, enzymes, non-haem proteins in addition to haemoglobins is outwith the scope of this thesis (for a comprehensive review see [3]).

Globin genes are a feature of all vertebrates and insects. Moreover their ubiquitous nature extends to the plant kingdom where a similar protein to haemoglobin called leghaemoglobin is found in the root nodules of
leguminous plants such as Soybean [4]. However this does not imply that the primordial globin gene was derived from pre-plant life forms, as it remains a possibility that the globin gene was introduced into plants by an insect or some ancestral animal form.

Globins are luxury proteins of crucial physiological importance. Their accessibility in the blood minimizes sampling problems. This has contributed in determining the structure of the haemoglobin molecule and in the establishment of detailed biochemical structure/functional relationships [5]. The availability of globin message of high abundance (> 90% total mRNA population in the reticulocyte) and purity has advanced the isolation of globin genes by recombinant DNA technology as discussed in 1:1:2 and 1:1:4.

The genetics of human globins is relatively well understood. The various haemoglobinopathies characterised by abnormal gene structure and function have been particularly useful in classical genetic analysis. Among the range of molecular lesions of the globin loci, the thalassaemias have by far been the most extensively studied (for a comprehensive review, see [6]).

The human globin gene system provides an excellent model to study some of the fundamental problems of the regulation of differential gene expression at the molecular level.

Four outstanding features of human globin gene expression invite investigation.
Firstly, although low levels of globin gene specific transcripts have been detected in the fibroblast [7], Rous Sarcoma Virus infected fibroblasts [8], a few human leukaemias [9] and in certain non-erythroid cell lines [10], the overwhelming majority of data suggests that normal globin gene activity is restricted to the red blood cell and a defined range of its precursors. The reticulocyte contains ~150,000 copies of globin mRNA which represents 90% of the message population, in sharp contrast non erythroid tissues such as brain and liver, have less than one copy per cell measured on a population basis i.e. tissue specific regulation.

Secondly the type of globin gene expressed varies with the stage of development i.e. embryo, foetus and adult. Thus the globin genes are subject to strict temporal regulation during normal development.

Thirdly in man there are six different globin genes located on two chromosomal loci termed α and β. The α family consists of α and ζ genes. The β locus members are the β, δ, γ and ε genes. As depicted in Fig. (1), this generates considerable phenotypic diversity resulting in the formation of six functional proteins to accomodate the changing physiological conditions experienced at different stages of development. The composition of haemoglobin tetramers independent of the developmental stage consists of two α type and two β type polypeptides. Since these loci are on different chromosomes a correct balance of chain production must operate, otherwise precipitation of
Figure 1

A schematic representation of differential globin gene expression during ontogeny showing the changing anatomical site of erythropoiesis and the type of haemoglobin synthesised. Reproduced and modified from [6].
the protein would result [11]. Therefore the regulated production of functional haemoglobin protein is entirely dependent on an effective coordinate control programme, i.e. coordinate regulation.

Finally the number of functional globin genes remains constant in all cell types throughout development. This is inconsistent with a gene amplification or diminution mechanism operating during ontogeny. Furthermore detailed restriction mapping of the α and β globin gene loci from erythroid and non-erythroid DNA do not show any pattern differences so arguing against any gene rearrangement or translocation as a mechanism to account for globin gene regulation [12] unlike the events leading to immunoglobulin expression during B-lymphoid cell differentiation [13].

1:1:2 Globin Gene Expression During Development

The globin genes have a well defined ontogeny (see [6] for review). In man, as illustrated in Fig. (1) the anatomical site of erythropoiesis changes with development from embryo to foetus to adult. These differences are reflected in the types of haemoglobin produced in the erythrocyte as depicted in Fig. (1).

The production of red blood cells is initially observed in the mesenchyme of the yolk sac around two weeks gestation. These cells remain nucleated throughout their life span of approximately twelve weeks. The
earliest detectable haemoglobin is Gower 1 (ζ₂ε₂) which contains two epsilon chains from the β globin gene locus and two zeta (ζ) polypeptides encoded in the α globin gene cluster. Around five weeks gestation the liver becomes the major site of haemoglobin synthesis, splenic erythropoiesis is also recognisable at this period. A few weeks after the onset of hepatic erythropoiesis embryonic chain synthesis is gradually replaced by the adult α and the foetal γ polypeptides to give HbF(α₂γ₂). Prior to this transition two minor haemoglobins are synthesised i.e. Gower 2 (α₂ε₂) and Portland 1 (ζ₂γ₂). Hepatic erythropoiesis progressively declines after twenty weeks gestation with a concomittant increase in red cell production in the bone marrow. During this period the level of splenic erythropoiesis markedly decreases. Prior to birth, production of γ chains falls and is replaced by β or δ polypeptides to form α₂β₂(HbA) and α₂δ₂(HbA₂) respective the latter accounts for less than 2% of total haemoglobin production.

Early theories seeking to explain these events proposed the existence of specific embryonic, foetal and adult haematopoetic cell lineages. Much of the basis for those ideas stemmed from the observation that certain globin genes were apparently restricted in their expression to cells of unique developmental origin such as epsilon globin gene activity in yolk sac tissue. Sophisticated analytical techniques have essentially disproven this and have shown that haematopoetic stem cells appear to have a developmental clock, the mechanism
Accordingly the main question arising from an account of differential globin gene expression can be divided into two separate but overlapping parts. First, what are the underlying mechanisms responsible for the generation of a haemoglobin synthesising erythrocyte from the haematopoietic stem cell? Secondly, what determines the type of haemoglobin produced in the erythrocyte? Implicit in our comprehension of this is the elucidation of the mechanisms involved in commitment, tissue specificity, temporal and coordinate regulation.

1:1:3 Molecular Organisation of the Human Globin Genes

Cell fusion experiments have located the α globin gene cluster to the short arm of chromosome 16 at coordinates (16p12 - 16pter) [15] [16]. Similar studies have mapped the β globin gene family to the short arm of chromosome 11 at coordinates (11p11 - 11p15), but its precise location is uncertain [17] [18] [19].

All members of both human globin loci have been molecularly cloned from human genomic lambda phage libraries; the α gene, the ζ gene [20], the adult β genes, [21] [22] the foetal genes [23] [24] and the embryonic epsilon globin gene [25] (See Fig. (2)). In addition, the entire β globin gene cluster has been isolated in a series of four overlapping cosmids [26].

The α globin gene cluster as depicted in Fig. (2) is
Molecular Organisation of the Human Globin Gene Loci on Chromosomes 11 and 16. The solid triangles show Alu I repetitive sequences and orientation. The hatched boxes denote Kpn I family members and orientation. Prototypic α and β globin gene structures are shown. solid and open boxes depict exons and introns respectively. The α-like globin genes contain introns of approximately 95 and 125 bp's between codons 31 and 32 and 99 and 100 respectively. The prototypic β globin gene contain introns of approximately 125-150 and 800-900 bp's between codons 30 and 31 and 104 and 105 respectively.
contained within a 30 kb stretch of DNA comprising of one embryonic $\zeta$ and two adult $\alpha_2$ and $\alpha_1$ genes. The physical linkage was determined by the structural analysis of lambda recombinants [20]. The organisation $5' - \zeta - \psi\zeta - \psi\alpha - \alpha_2 \alpha_1 - 3'$, parallels the order of sequential expression during ontogeny.

The organisation of the $\beta$ type genes was determined by the detailed structural analysis of the appropriate lambda recombinants, reviewed in [27] and by genomic blotting experiments [28] [29]. As shown in Fig. (2) their arrangement $5' - \epsilon - \gamma^A - \gamma^G - \delta - \beta - 3'$ reflects their normal order of expression during red blood cell development.

Fig. (2) shows that the human globin gene loci contain three pseudogenes, the physiological function of which, if any is not understood, for a brief review, see [30]. These elements are a structural feature of a number of animal globin gene clusters including mouse [31], goat [32] and rabbit [33]. A pseudogene has an obvious evolutionary link to a normal gene but does not produce a functional protein. It is widely accepted that pseudogenes arose by duplication followed by sequence divergence [34]. The nature of the alteration responsible for the aberrant expression is variable [30].

The protein coding information in the human globin genes like many other cellular genes is not contiguous but is interrupted by non-coding intervening sequences [35]. Both the $\alpha$ and $\beta$ globin type genes contain two introns and three coding exon regions. Fig. (2) shows the introns of
the five β globin genes mapping at identical positions as do the three α globin type genes. This gross structure is highly conserved, even the plant leghaemoglobin genes display this pattern [4]. Rather surprisingly however the twelve different globin genes of the midge Chironomus thummi thummi lack intervening sequences [36].

The complete nucleotide sequence for all members of the α and β globin gene clusters has been elucidated; α [37], ϶ [38], β [39], δ [40], γ [41] and epsilon [42] [43] [44] [45]. As depicted in Figs. (3) and (4) the proximal five flanking region of β genes like a number of other protein coding genes contain two highly conserved promoter signals [47]. The Hogness or TATA Box [48] (TATAAAAC) which is referred to in globin genes as the ATA Box is located~30 bp's from the major initiation site and is flanked by short GC stretches. The second conserved domain is the CCAAT Box [49] mapping~80 bp's 5' to the cap site. This is duplicated in the human γ globin genes and diverges in the hypo-functional δ globin gene i.e. CCAAC. Except for the δ globin gene and the α type globin genes all other globin transcriptional units contain a C rich motif around 100 bp's upstream of the cap site with the consensus sequence CACC. This is a feature of other animal β globin genes [50] [51].

As shown in Fig. (4) the prototypic globin gene, the 5' non-coding region consists of the stretch of DNA between the cap site which is well conserved (Fig. 3) and the ATG codon. This distance is approximately 50 bp's and
Figure 3

Comparison of the 5’ flanking region of a range of β-globin genes in relation to the human epsilon globin gene which is enclosed. Areas of homology shown are the Cap Site, the ATA Box, the CCAAT Box and the C-rich -100 motif. Reproduced from [2].
Figure 4

Structure and Expression of Human Globin Genes. A shows the various promoter signals. B illustrates the gross structure of the globin genes in addition to the highly conserved sequences for splicing and polyadenylation involved in nuclear processing. C depicts the different stages of the flow of information from gene to protein. Reproduced and modified from [46].
contains no other open reading frame. In common with a variety of other Pol II genes, the globin leader sequence has a short conserved region located - 10 bp's downstream of the cap site with the consensus sequence CTTPyTG [52].

Comparison of the exon/intron boundary on a number of divergent genes have identified 5' and 3' splice consensus sequences C/A AG+GT G/A AGT and (Py)6NCAG+GG/T [53].

In common with other cellular genes the 3' untranslated region of the globin genes contain the highly conserved AATAAA [54] which is a signal involved in the endo-nucleolytic cleavage of RNA and its subsequent polyadenylation.

The 3' flanking region of the globin transcriptional unit represents the most highly divergent area. This is well illustrated by the two γ globin genes which differ by around 15% in this region [27]. Much of the human genome is repetitive [55]. These sequences can be classified into two types. The first are short nucleotide runs, the physiological function of which is unknown [56]. The second class of repeat elements are longer in length and are interspersed throughout the eucaryotic genome as well as the β cluster. Two sequence families have been described in the β locus viz the AluI and Kpnl families. As depicted in Fig. (2) the β cluster contains eight AluI members [57] [58] [59]. The α globin locus features three AluI elements [60]. There are two Kpnl members, one between the epsilon and γ globin genes [61] and the other 3' to the β globin gene [62] as outlined in Fig. (2).
The human epsilon globin gene was originally isolated by Proudfoot and Baralle [25]. HindIII digested normal and thalassaemic DNA was probed with radiolabelled rabbit β globin DNA. This cross hybridised to all β globin type genes including the embryonic gene. DNA corresponding to these autoradiographic signals was eluted from the agarose gel and the enriched fractions were cloned into HindIII restricted lambda vector [63]. The gene was isolated as an 8kb fragment. Since then the epsilon globin gene has been cloned in a Charon 4A lambda vector [21] and in a cosmid [26]. Structural analysis as depicted in Fig. (2) shows that the epsilon globin gene has the same structural pattern as the prototypic β globin gene. As outlined in Fig. (3) it bears considerable homology to other β genes especially in the proximal 5' flanking region. The entire coding and flanking sequences upto - 16 kb from the cap site have been determined [45]. As shown in Fig. (4) the epsilon globin gene has a tripartite promoter structure (5' - 100/CCAAT/ ATA - cap-3') which it shares with the actively expressed γ and β globin genes. A notable structural landmark of the epsilon globin gene domain is the two AluI members separated by 916 bp's, the proximal 1 kb upstream of the site of initiation as depicted in Fig. (2). Another feature of the epsilon globin gene is a region of alternating purine-pyrimidine
residues within a 995 bp's Sst I fragment ~ -6.5 kb upstream of the cap site.

Under certain physiological conditions these can adopt a Z DNA form [64].

1:2 Regulation of Globin Gene Expression

Differential globin gene regulation can be regarded as a series of events involving the sequential unfolding of a complex genetic programme during haematopoiesis in which expression of the globin genes are regulated in a tissue specific, coordinate and temporal manner. Despite the wealth of information on the genetic organisation and molecular anatomy of the globin loci our understanding of the logic of control and the underlying molecular mechanisms involved in these regulatory phenomena remain rudimentary.

Globin genes are transcribed by RNA Polymerase II. In general, genes transcribed by RNA Polymerase II can be regulated at one or more defined levels during gene expression [65]. The major steps of the intracellular information transfer are mRNA biogenesis; transport of message to the cytoplasm, turnover in the cytoplasm and protein synthesis which can involve post-translational processing. Prior to the onset of any of these events DNA replication should be considered as having an important role in gene expression especially quantal cycle division [66]. Excluding replication, this integrative scheme of
regulation has formed the basis of the cascade regulation hypothesis [67]. Therefore it is important to determine which level or levels the expression of globin genes are controlled.

1:2:1 Replication

Replication plays an integral part in differential gene expression. Orthodox opinion is that quantal cell cycles direct cell populations along different developmental pathways [66]; replication would be expected to have a similar function in haematopoiesis. Template duplication per se can be envisaged to influence the mechanisms responsible for 'gene activation' and subsequent transcription. During S phase the nucleosomes show a distinct association with the continuously synthesised strand, this effectively leaves the nucleosome depleted lagging strand more accessible to cellular control factors and the transcribing machinery [68].

DNA synthesis appears to be essential during the late stages of red blood cell development. Conkie et al [69] have demonstrated a cell cycle dependence on the induction of haemoglobin synthesis in butyric acid treated MEL cells (Murine Erythroleukaemic cells transformed with the Friend virus [70]). A similar effect was reported by Gambari et al [71]. Smithies [72] has recently formulated a model emphasising the importance of replication on the regulation of globin gene expression. He has postulated
that β globin gene activity is mediated by the differential use of a series of replicons in the β globin gene cluster. The tenets of this hypothesis are that four β globin gene replicons are selectively utilised at different stages of development and in different cell types. Smithies [72] proposes that when upstream replicons are used the proximal globin gene is activated. These origins and structural sequences are termed homo-directional in that both DNA synthesis and transcription occur in the same direction. When heterodirection (downstream) replicons are favoured the proximal gene is repressed. For example, during embryonic development the 5' epsilon and the 3' β replicons are activated, this ensures ε expression and repression of the adult and foetal genes. This model has been applied with mixed success to various thalassaemic conditions where the appropriate replicon is mutated or absent so offering a physiological explanation for the abnormal expression. Smithies [72] however does not put forward any ideas on the regulation of this differential use of replicons.

A recent study has raised doubts about the necessity for DNA synthesis during cell differentiation. Chiu and Blau [73] examined the activation of muscle genes in human fibroblast x mouse muscle somatic hybrids. Activity was assayed in the presence and absence of the replication inhibitor cytosine arabinoside. Heterokaryons irrespective of whether a complete cell cycle had occurred had comparable levels of the muscle protein MM creatine

- 19 -
kinase. This finding questions the role of quantal division during differentiation.

1:2:2 Transcription

Groudine and Weintraub [74] reported that in 20-30 hour chick erythroblasts, the globin genes are transcriptionally silent as measured by the lack of haemoglobin synthesis and the insensitivity of the gene to DNase I treatment. Two generations later the chromatin adopts an active conformation implied by an increased susceptibility to nuclease digestion; in addition haemoglobin is detected at 36 hours. These observations infer that globin gene activation during red blood cell development initially involves an alteration in the configuration of the chromatin around the globin genes prior to the onset of transcription.

Hoefer et al [75] examined mouse β globin transcripts by analysing nascent labelled RNA from MEL nuclei which had previously been cultured in the presence of the inducer DMSO (Dimethysulphoxide). This treatment elevates haemoglobin production between 10 and 30 fold. This value correlates with a concomitant increase in the rate of β maj globin mRNA synthesis. Moreover the chromatin in the vicinity of this gene is DNase I sensitive.

Curtis and Weissman determined the length of the mouse β globin RNA precursor molecules by hybridisation studies with pulse labelled RNA and immobilised cDNA [76]. These
authors described a 15S precursor which is processed to the 9S mature message. In this scheme, the primary transcript initiates around 50 bp's upstream of the ATG site and terminates at the polyadenylation signal. In contrast Bastos and Aviv [77] using a similar experimental design strategy and Renyaud et al [78] detected larger precursors of around 28S for the mouse and avian β-globin genes respectively, the latter data are particularly convincing as those molecules were measured by heteroduplex mapping. This difference could be explained by 5' or 3' extension or both. Transcripts which extend beyond the 3' terminus, can be accounted for by a read through type of mechanism however extension at the 5' end is difficult to imagine since the 5' termini are capped. Capping requires the covalent linkage of the α-phosphate of GTP to the α and β groups of the 5' terminus [79]. Thus cleavage of the RNA at the phosphodiester bond can only give an α moiety which means the GTP donating both the α and β phosphates. An alternative to this has been described by Abraham et al [80]. This involves the formation of a pyrophosphate group by the addition of a phosphate group prior to capping. This appears to be a specialised case and the overwhelming majority of capped molecules appear to be synthesised as previously outlined.

An alternative to an abnormal capping mechanism would be chain initiation at site(s) upstream of the canonical cap site so yielding a mixed population of transcripts. Hoefer and Darnell [81] using pulse labelled RNA from induced MEL cells described such a 5' heterogenous
population with around 20% of mouse β globin RNA initiating upstream of the major cap site. These transcripts however were not accurately mapped, hence nucleolytic cleavage could not be discounted. In addition it was not known whether these molecules were capped. These authors also detected transcripts mapping downstream of the polyadenylation signal. In sharp contrast Grosveld et al [82] could not detect any rabbit β globin precursor larger than 16S. A possible explanation for these inconsistent findings is the different experimental techniques employed to assay for primary transcripts. Some systems maybe more sensitive in detecting non processed molecules. Moreover the data cannot rule out instant processing from some larger precursor(s) which eventually are processed to the 15S and 27S species.

Further evidence for the existence of large globin RNA molecules comes from a study by Ross et al [83] with the chicken β globin RNA and its association with the nuclear matrix. Over 90% of β globin specific RNA transcripts are found in the nuclear matrix complex. This structure has been proposed to be a platform for a number of important cellular events including nuclear processing [84]. High MW (Molecular Weight) chicken β globin RNA which extends into the embryonic domain has been shown to be complexed with the matrix structure.

*In vitro* run-off transcription assays involving both S1 Nuclease mapping and primer extension on cloned ζ, ε and β globin gene fragments demonstrated that
the transcripts exclusively originated from their respective major cap sites [85]. Furthermore these RNA species had the same 5' cap structure as authentic mRNA.

In contrast Allan et al [86] [10] described RNA initiation sites upstream from the main start site in the human epsilon globin gene. RNA was prepared from the human erythroleukaemic cell line K562 which expresses embryonic and foetal globins [87] [88] and human embryonic livers of 5-10 weeks gestation. Epsilon globin gene specific transcripts were assayed by a combination of S1 nuclease and primer extension mapping techniques. Over 10% of the RNA molecules originated from a specific site located 270 bp's upstream from the initiation codon (the -200 promoter). Extended studies using probes spanning over 5 kb of 5' sequence identified upstream initiation regions at -65, -270, -900, -1480 and -4500 bp's from the ATG codon. (see Fig10) These minor initiation sites account for around 10-15% of the total epsilon transcript population. In addition the transcripts which yielded enough material for analysis were shown to be capped arguing against endonucleolytic cleavage of a larger precursor. The physiological significance of these alternate cap sites is not understood. Of more importance is that they may reflect some local perturbation in the chromatin structure surrounding the epsilon globin gene as is generally accepted to be the case with DNase I hypersensitive sites. DNase I mapping of the epsilon gene in K562 cells shows a strong correlation between upstream initiation and nuclease sensitivity [89] [90]. Moreover
the major upstream start site at -270 bp's contains an S1 hypersensitive site [91]. Non-canonical cap site initiation has also been detected in the human \( \gamma \) and \( \beta \) globin genes. Using a similar experimental approach Grindlay et al [92] and Ley et al [93] mapped transcripts in the \( \gamma \) globin gene between 50 and 100 bp's upstream from the major cap site. \( \beta \) globin gene transcripts originated between 140 and 170 bp's from the canonical cap signal.

In many of the sequenced Pol II gene major cap sites are preceded by the conserved promoter signals TATA and CCAAT. Extensive computer sequence searches were undertaken to determine if any of the upstream regions of the \( \beta \) type globin genes had these promoter domains in their vicinity. Despite some close fits with certain upstream sites, no unifying pattern emerged. This would seem to rule out a pre-requisite for the TATA and CCAAT Boxes for Pol II transcription. This is the case with certain viral genes [94] [95], and the GC rich promoters of a range of house keeping genes such as the human adenosine deaminase [96] and mouse dihydrofolate reductase (DHFR) [97] genes which are characterised by one or more Sp1 binding sites with the recognition sequence GGGCGG [98]. Alternatively these upstream promoters may be recognised by RNA polymerase III

1:2:3 **Nuclear Processing**

Nuclear processing is regarded as a passive event in
the control of globin gene expression. Krainer et al [99] have looked at the splicing patterns in wild type and a series of mutant human β globin templates which had the normal acceptor/donor conserved splicing sequences [53].

When assayed in an in vitro HeLa cell extract system all RNAs were correctly processed. Thus it would appear that the globin primary transcript does not require the presence of authentic 5' and 3' or a polyadenylated 3' end to be faithfully spliced. Furthermore these workers [99] showed that a capped 5' terminus was not required for correct splicing, however without this modification the uncapped transcripts are spliced less efficiently and accurately.

1:2:4 Nuclear-Cytoplasmic Transport

Chan [100] looked at the transport of chick embryonic globin RNA from the nucleus to the cytoplasm. He showed that newly synthesised RNA was retained exclusively in the nucleus for around 24 hours. Complete transport to the cytoplasm occurs within the next 24 hour period. The relative amounts of globin RNA in the nuclei and the cytoplasm change in a reciprocal manner, while the total amount of cellular globin RNA does not vary much during erythroid maturation. However it should be emphasised that confirmation of these data is still lacking. This apart, these findings support a hypothesis proposing that during the late stages of erythroid differentiation
transcription of the globin genes precedes activation of a regulatory transport mechanism. Interestingly a recent study by Therwath et al [101] showed that Avian Erythroblastosis Virus (AEV) transformed immature chick erythroid cells have a different globin gene transcription pattern from adult erythroid tissue. Adult globin specific RNAs are exclusively confined to the nucleus and are unprocessed forming a 7-8 kb $\alpha^A$ globin transcript while the contiguous $\alpha^D$ globin gene specific RNA is essentially undetectable. In addition transcripts from the unlinked $\beta$ globin gene locus are difficult to detect. These findings can be interpreted to support those of Chan [100]. An alternative explanation is that they are not transported because they are not correctly processed.

1:2:5 Message Stability

Transcriptional regulation alone cannot account for globin message being the most abundant species in the red blood cell. This domination is reflected in the stability of globin mRNA compared to the non-globin population. An early study on the $t_2$ of globin mRNA in MEL cells showed a dramatic decrease from 60h to 17h after induction [102]. To explain this Aviv [102] proposed that non globin messages are degraded at a greater rate during the late stages of erythroid development. However Volloch and Houseman [103] using modified tissue culture conditions which physically stabilise the MEL cell during
chemical induction calculated the $t_\frac{1}{2}$ of globin was ~60h in the terminal differentiated cells, while non globin half lives were ~30h. Differential mRNA degradation has also been considered as a mechanism in potentiating the levels of different haemoglobins in the erythrocyte. This could possibly influence globin switching by varying the stability of one globin species relative to another. Ross and Sullivan [104] investigated this by estimating the $t_\frac{1}{2}$ value for γ and β mRNAs in cultured human reticulocyte. Both species had similar turnover rates of 27-29 hours, which is inconsistent with the above postulate.

1:2:6 Translation

Selective translation of globin specific messages at the expense of other mRNA can be envisaged as a feasible mechanism in regulating the level of haemoglobin in the erythrocyte. There are two size classes of ribonucleoprotein particles (RNPs) in the red blood cell. Globins are associated with a 20S species and the other RNAs belong to the larger 35S structure. The larger RNP group contains a small RNA species which inhibits protein synthesis in vitro [105]. Vincent et al [106] have suggested that translational repression of the 35S class occurs during the late stages of erythroid development so effecting an increase in the synthesis of haemoglobin at the expense of non-globin in proteins. Differential translation in a similar fashion to selective stability as previously outlined may feature in the switching of globin
genes. Sullivan and Ross [107] examined this in the foetal/adult switch. These workers found that the translation capacity of reticulocytes has a half life of 6 - 8 hours and γ and β protein synthesis parallel this decay. Furthermore it appears that the protein synthesising apparatus and not the abundance of message is quite limiting for globin production.

1:2:7 Conclusion

To conclude, the weight of experimental evidence suggests that globin genes are principally regulated at the level of initiation of transcription (1:2:2). However data exists which can be interpreted to support the notion that in addition to this transcriptional regulation, globin genes are sequentially regulated after initial synthesis of the primary transcripts. These putative control levels being cytoplasmic transport (1:2:4), message stability (1:2:5) and translation (1:2:6) is consistent with some of the ideas within the framework of the Cascade Hypothesis of gene regulation [67].

1:3 The Conformation of Globin Genes in Chromatin

DNase I Hypersensitive Sites (DHS's) within isolated nuclei or chromatin have greater sensitivity to enzymatic probes and are associated with the binding of the non histone proteins (high mobility group) HMG 14 and HMG 17
It is envisaged that these DHS's perturb the regular nucleosomal array which in turn alter the accessibility of the DNA of these regions to both cis and trans acting factors (for review see [109]). The approach of DNase I probing chromatin prepared from erythroleukaemic cell lines (see 1:7:2) and from erythropoietic tissue at various stages of ontogeny has yielded valuable information in the identification and mapping of putative regulatory regions of the globin transcriptional unit. Several examples illustrate this success.

Groudine et al [110] investigated the chromatin conformation of the $\gamma^A - \gamma^G - \delta - \beta$ cluster in foetal and adult erythropoietic tissue and in the HEL and K562 cell lines which produce $\gamma$ globin mRNA but not $\beta$. The DNase I profile obtained with foetal liver cells showed 5' DHS's located ~200 bp's upstream of the four genes. However this pattern was not observed in the analysis of adult bone marrow in this case. In addition a DHS was detected in the large intron ~800 bp's 3' to the polyadenylation site. The hypersensitive regions were restricted to the adult genes. The HEL cell displayed a similar DNase I digestion profile to the foetal liver tissue. Furthermore, although $\gamma$ globin mRNA levels increased in haemin treated K562 cells no difference in the DNase I digestion was observed. In a similar type of study Lachman and Mears [111] described DHS's in 100 bp's 5' to the cap site of the $\gamma^A$ and $\gamma^G$ globin genes of K562 cells. Although the $\beta$ globin gene was
DNase I sensitive compared to the insulin gene only the transcriptionally active foetal globin genes have DHS's. These studies as did the earlier work of Weintraub and Groudine [112] suggest that globin genes in erythroid cell nuclei exhibit an increase in overall sensitivity of the \( \beta \) globin gene domain may reflect a change in local higher chromatin structure from the repressed conformation of the 300 Å fibre (solenoid) to the activated 100 Å structure, the conformation supposedly of the globin gene domains in all erythropoietic tissue [113]. The chromatin structure around the epsilon globin transcriptional unit in the K562 cell has been independently determined by Zhu et al [89] and Tuan and London [90]. These workers found that the DNase I hypersensitivity profile corresponded with the major and upstream RNA initiation sites described by Allan et al [10] [86] at -270, -900, -1480 and -4500 bp's. In addition a further DHS is located ~6500 bp's upstream of the major cap site. This region has two notable structural features. The first is a long AT rich region which includes a consecutive stretch of 28Ts. The second structure consists of a potential \( Z \) DNA forming element of \((AC)_{12}(AT)_{6}\), see Fig 10.

These experiments and a similar pattern of differential DNase I sensitivity during globin ontogeny in chicken [114] suggest that the formation of a DNase I hypersensitive site within the derepressed 100 Å conformation represents a primary event in globin gene regulation. However the appearance of a DHS or the
pre-activation of the globin gene is not by itself a sufficient condition for the transcription of the globin gene, a further event(s) must occur. As a further illustration Groudine and Weintraub [115] demonstrated that chicken embryo fibroblasts infected with a Rous Sarcoma Virus (RSV) temperature sensitive (ts) mutant caused the formation of DHS's in the α and β globin gene domains and the production of globin specific transcripts when the temperature was reduced from 41°C to 36°C. On adjusting the temperature back to 41°C globin gene transcription ceased, however, the DNase I hypersensitivity remained. Moreover they were maintained in the fibroblast for up to 20 generations.

More detailed work on the association with DHS's in globin gene regulation has been carried out with the chicken globin gene system. Emmerson and Felsenfeld [116] have described a specific factor exclusively present in a 9 day old chick which is involved in the creation of DHS in the 5' region of the β globin gene. Partially purified extracts from erythrocyte nuclei bind to a plasmid recombinant containing the hypersensitive region. This protein was not detectable in brain or oviduct tissue and was almost completely absent in 5 day old chick which does not express this gene. McGhee et al [117] located the DHS within a 115 bp's MspI fragment mapping 60 to 250 bp's upstream of the cap site. This region corresponds to C rich - 100 region of the β globin gene promoter. This DHS also contains a S1 Nuclease hypersensitive site which
is also present in supercoiled recombinants and in in vivo reconstituted histone core particles. However this S1 nuclease site does not show the symmetry of strand cleavage expected for single stranded DNA. Fine mapping revealed the S1 hypersensitive site to be a long poly G stretch. Interestingly other S1 sites have also been shown to be either homo-purine or homo-pyrimidine in structure [119] [120].

1:4 Methylation of Globin Genes

Demethylation of certain CpG residues has been coupled to transcriptional activity of a number of genes including globin genes (for a recent short review see [121]).

The early experiments of van der Ploeg and Flavell [12] on the methylation pattern of the human β globin gene locus from erythroid and sperm tissue showed an inverse correlation between methylation and gene expression. Since then, Mavilio et al [122] reported on a similar relationship in the promoter regions of the human ε and γ globin genes in embryonic erythroid tissue. In another related study Busslinger et al [123] using the technique of in vitro methylation demonstrated that methylation of upstream sequences - 760 to + 100 bp's prevent transcription of the human β globin gene after transfection into mouse cells. Furthermore erythroid cells cultured in 5'-azacytidine leads to a general inhibition of methyl transferase activity [124]. This has
provided the basis of attempting to switch on the foetal globin genes in thalassaemic patients [125], however it is contentious whether demethylation per se caused an increase in HbF synthesis. In contrast to the above findings Weintraub and his colleagues [126] working with erythroblasts transformed with a temperature sensitive Avian Erythroblastosis Virus, demonstrated non-expression of the α globin gene at the permissive temperature despite the existence of a DHS and undermethylation of 5' CpG residues. Of further intrigue was the observation that the β globin gene remains methylated under these particular circumstances.

The studies on globin gene methylation profiles and those of active genes which remain methylated such as chicken vitellogenin [127] and mouse dihydrofolate reductase [128] make a general conclusion on whether demethylation of specific CpG residues is a pre-requisite for transcription difficult to evaluate.

1:5 cis Control Sequences

Globin gene regulation at the molecular level is controlled by the interaction of proteins and various cellular factors with specific DNA sequences. An initial step in the study of these interactions is the identification of putative control sequences of the globin transcriptional unit. The regulatory sequences can be broadly divided into two main types.
The first are involved in directing the synthesis of authentic message. Sequences involved in qualitative control are the various functional domains of the promoter which are concerned with mRNA initiation and the short, highly conserved signals involved with nuclear processing (capping, splicing and polyadenylation, see Fig. (4)).

The second group of controlling elements are those sequences which govern the quantitative regulation of globin gene expression. These await identification and seemingly involve both cis linked DNA and trans acting genes. These sequences are associated with the regulation of globin genes.

1:5:1 The Globin Gene Promoter

Grosveld et al [129] examined the transcription of a number of 5' rabbit β globin gene deletion mutants in vitro in a HeLa cell extract assay system. They found that the ATA Box was required for the generation of faithful 5' transcripts. In contrast mutations in the CCAAT Box had no pronounced effect on transcription. This pattern is not reflected in vivo. Grosveld et al [130] engineered a series of rabbit - β - globin gene promoter deletion mutants. These templates were covalently linked to the SV40 origin of replication region which also contains an enhancer. Transcripts were analysed in a HeLa cell transient assay system by quantitative S1 Nuclease mapping. In the absence of the SV40 DNA, faithfully
initiated rabbit β globin gene transcripts are barely detectable which is in agreement with an original observation by Banerji et al [131]. In contrast to the in vivo studies, both the ATA and CCAAT signals are required for the synthesis of optimum levels of correctly initiated rabbit β globin gene transcripts. In addition deletion of the ATA Box led to greater heterogeneity of transcription initiation sites. Corroborative evidence proposing an important role for the ATA Box in globin gene regulation comes from studies on thalassaemias. There are several forms of β-thalassaemias as depicted in Fig. (5B) characterised by mutations in the ATA Box which is seemingly responsible for decreased transcriptional activity [46]. In addition, Grosveld and co-workers showed that deletions, which extended beyond the CCAAT motif into the C-rich -100 region, resulted in a significant reduction in the level of transcription, these down mutations were estimated to be in the order of a fifty fold reduction. Furthermore sequences 3' to the ATA Box including the major cap site were found to be dispensable for transcription. Corroborative evidence for a functional tripartite prototypic globin gene promoter structure (5' -100/CCAAT/ATA-cap-3') has been presented by Dierks et al [132]. However in contrast to the findings discussed above, these workers found that deletion of the ATA Box reduced the level of transcription, moreover very little 5' heterogeneity in the start sites of the few transcripts which were made were seen.
Figure 5

A summary of the molecular lesions causing β Thalassaemia. A shows the large deletions represented by the solid black lines in relation to the β-globin gene cluster. Hatched boxes indicate that that end point of the deletions has not been accurately mapped. B represents mutations causing β Thalassaemia which are point mutations and small deletions or insertions. Reproduced and modified from [2].
A comparable study on the behaviour of the human β globin gene in a transient assay system shows a similar transcriptional pattern to the rabbit gene. Green et al [133] demonstrated that the detection of authentic transcripts depends on the presence of a viral enhancer. S1 Nuclease analysis of 5' mutants confirm that sequences between 75 and 128 bp's upstream of the cap site are necessary for optimal transcriptional activity. As with the rabbit gene the region of 5' flanking sequence contains a double repeat of the C rich -100 element. Further evidence for a crucial physiological role for the -100 region when assayed in stable transformed cells was reported by Spandidos [134]. A series of 5' deletion mutants were covalently linked to the HSV-1 tk gene (Herpes Simplex Virus Type 1 thymidine kinase gene). These were stably introduced into the MEL tk- cell line. Transcriptional levels were determined by Northern Blotting of total RNA. Adjusting for the variation in copy numbers of the donor DNA it was concluded that a region upstream of the CCAAT Box was required for optimal expression. This area corresponds to the location of the repeated C rich region. Detailed correlations were not possible as the deletion mutants were not sequenced.

Functional domains have been identified in the α globin type promoter. Using α-1 globin gene recombinants harbouring the SV40 origin of replication Mellon et al [135] transformed COS 7 cells and measured the transcripts after a 48 hour expression period. COS 7 cells
constitutively produce T antigen [136] and are permissive for SV40 viral replication so ensuring high copy numbers of the α gene recombinants. In this system authentic globin mRNA is produced at easily detectable levels. Deletion mutant studies show that the region between -55 and -87 which contains the CCAAT motif is required for optimal activity.

In contrast the embryonic α globin gene ζ is not transcribed from the SV40 replication vector in COS 7 cells or HeLa cells [137]. However low activity levels were detected when the SV40 72 bp's enhancer was covalently linked to the ζ globin gene recombinants 50 bp's upstream of the CCAAT Box. Moreover when this recombinant was introduced into Xenopus oocyte by microinjection, mRNA synthesis dramatically increased. In addition authentic transcripts are generated in the absence of the SV40 enhancer. These results are somewhat different from the behaviour of other globin genes which are non specifically transcribed in Xenopus oocyte. For example rabbit β globin transcripts originate exclusively from vector sequences [138]. These data suggest some embryonic factor in the Xenopus oocyte can overcome the apparent inactivity of the ζ globin gene in Hela or COS 7 cells.

To conclude, three endogenous cis acting functional domains have been identified in the transcriptional unit of human globin genes. The ATA and CCAAT signals are common to all α and β globin genes except the
hypofunctional δ globin gene which has a CCAAC signal compared to the canonical CCAAT motif as depicted in Fig. (3) These short conserved elements constitute the core promoter which is common to a number of Pol II genes [47]. In addition the epsilon γ and β globin genes share a C rich -100 element giving these genes a distinct tripartite promoter structure. The δ globin gene and the α globin locus members lack this element. Finally it is interesting that the human α type globin genes and the epsilon, γ and β genes of the β globin gene cluster which have similar transcriptional rates in mature erythrocytes respond differently to the presence of cis linked viral enhancers in short term expression assays. As to what significance these differences have in the differential expression of the globin genes is unclear.

1:5:2 Enhancer Sequences

Enhancers can be operationally defined as integral cis acting functional domains of certain viral and cellular transcriptional units which markedly potentiate transcription from heterologous promoters relatively independent of distance from the gene and orientation in relation to it (for a comprehensive review see [139]).

Comparative sequence analysis of a diverse range of enhancer elements reveals a short homologous region termed the "core" i.e. T(G) [140]. Studies with SV40 core element point mutations suggests the 5' G is essential for
optimum activity [140]. A further common structural element has been identified by Nordheim and Rich [141] who surveyed a variety of viral enhancers and found a pattern of two pairs of potential Z DNA elements (≥ 8 bp's) separated by between 50 and 80 bp's.

Enhancers display some interesting physiological properties.

The expression of enhancer dependent promoters can be attenuated by the insertion of other promoters between the enhancer and the promoter under study. For example De Villiers [142] found that T-antigen expression dropped when a rabbit β-globin promoter was positioned between the 72 bp's element and the SV40 promoter.

Although viral enhancers can substitute for one another in functional assays [143], Spandidos and Wilkie [144] showed that some display a distinct preference for a species specific host such as the enhancers of BPV (Bovine Papilloma Virus) and MoMuSV (Moloney Murine Sarcoma Virus). This preference or requirement for a specific host is a restriction on the activity of a number of cellular enhancers which act in a highly tissue specific manner such as those of the immunoglobulin genes [145] [146] [147], the chymotrypsin and insulin gene [148] indicative of trans factor(s) involvement. This tissue specific mode of action of these enhancers is particularly germane to globin gene expression.

Enhancers are recognition sequences for cellular factors. Schöler and Gruss [149] monitored the expression
of an SV40 - CAT (Chloramphenicol Acetyl Transferase) plasmid in a transient assay system. A reduction in CAT activity was observed in the presence of increasing amounts of competitor DNA in the form of an SV40 - Neo (Neomycin Gene) recombinant.

The mechanism of enhancer action is unknown. A favoured model is that of a RNA Polymerase bidirectional entry site, where on access the polymerase or some as yet unidentified associated factor scans the template and initiates at the first promoter it recognises.

The existence of tissue specific enhancers of the transcriptional unit of the immunoglobulin [145] [146] [147], chymotrypsin and insulin [148] genes has given rise to the intriguing notion that such sequences may be associated with the tissue specificity of globin gene expression. However, to date no compelling data has been presented supporting this idea. Somewhat indirect evidence for transcriptional enhancement by globin genomic DNA but not tissue specificity has been reported by Di Maio et al [150]. These authors have demonstrated that a sequence element 3' to the human β globin gene markedly increases the transformation efficiency of mouse cells when cis linked to a BPV vector.

1:5:3 Z-DNA

Although the major conformation of DNA in biological systems is the B-form, other structures arising from
unusual sequence arrangement have been identified. Prominent amongst these alternative structures is left-handed or Z-DNA (see [64] for a general review) characterised by a string of alternating purine and pyrimidine residues which under certain physiological conditions can adopt a left-handed form.

Potential Z-DNA regions have been identified throughout the human genome [151]. Nucleotide sequencing of the human β globin gene cluster has revealed a number of alternating purine and pyrimidine tracts of varying length from 20 - 60 bp's. These have been mapped ~6.5 kb upstream of epsilon globin gene [45], within the large intron of the γ globin gene [40] and ~2.7 kb and ~0.5 kb upstream of the cap site of the human β globin gene [152] [153].

It has been proposed that left-handed DNA has a regulatory role in the expression of Pol II genes [64] and Pol III genes [154]. This influence is thought to occur by changes in the supercoiling of the DNA due to the B → Z transition which in turn causes a perturbation in the conformation of chromatin. Sinden [155] has shown that DNA which is transcriptionally silent is not under superhelical tension. And although proteins have been identified which specifically bind Z-DNA [156] it is not known whether such transition alter the three dimensional structure of neighbouring DNA in vivo.

A role for Z-DNA involvement in the expression of globin genes awaits detailed examination. However Gilmour
et al [157] have identified a negative effective element in the 5' flanking region of the mouse β maj globin gene within which maps a 53 bp's tract of alternating purine and pyrimidine residues. When this segment between 344 and 1143 bp's upstream of the cap site was linked to the mouse β maj globin or the human epsilon globin promoter/HSV-1 tk fused gene, transforming activity in HAT medium (100 μM Hypoxanthine, 0.4 μM Aminopterin, 16 μM Thymidine) [158] was essentially abolished. In contrast the repression was relieved when the repressor fragment was covalently linked to the HSV-1 tk promoter. Difficulties in transforming MEL tk- cells halted further investigation into whether this phenomenon would be a feature of erythroid cells which actively expresses the endogenous gene (R.S. Gilmour Pers Comm).

1:5:4  Intergenic DNA

Much of the intergenic DNA of the human β globin gene locus is repetitive (see Fig. (2)). Do any of these repeat elements have any physiological role in globin gene expression?

The 300 bp's AluI members contain the recognition signal for RNA Polymerase III [159]. Several of the AluI elements are transcribed in vitro [160] [161] [162]. Allan and Paul [163] demonstrated the in vivo expression of the distal AluI element of the human epsilon globin transcription unit in the erythroleukaemic cell line K562.
Interestingly, the 350 - 400 bp's transcript originates from the opposite strand and corresponds to the in vitro transcript described by De Segni et al [164]. These RNA species are non polyadenylated and nucleus confined. Of possible regulatory significance is the observation that these transcripts have only been detected in the K562 cell. However it is unknown if Pol III activity is a consequence or cause of epsilon globin gene activity in the K562 cell.

Along similar lines, Carlson and Ross [165] have identified human $\beta$ globin RNA species which are synthesised in vitro from the 5' flanking region between -235 and the cap site by S1 Nuclease mapping of run off transcripts. These molecules which extend to the coding region of the gene were shown to be $\alpha$ amanitin sensitive. Although as depicted in Fig. (2) there is not an AluI repeat element in this region of the human $\beta$ globin transcription unit, there is a region of Pol III homology a feature of AluI members. Of particular interest was the observation that these transcripts were only present in RNA prepared from erythroid tissue.

1:6  trans Control Sequences

Although cis - operating mechanisms are of crucial importance in the transcriptional initiation of globin genes, the events leading from repression to gene activation would seemingly involve a series of trans acting factors. Despite our almost total ignorance of
these mechanisms or the regulatory genes involved in the sequential activation and expression of globin genes, there is considerable evidence supporting the idea of trans factor involvement during red blood cell development. Three distinct experimental strategies have been employed to investigate trans activation of the globin genes. These are somatic cell genetics, induction of stably transformed MEL cells and elevation of globin gene expression by defined viral gene products.

1:6:1 Globin Gene Expression in Somatic Cell Hybrid Lines

Deisseroth and Hendrick [166] were able to activate the human α globin genes from non-erythroid cells in man x mouse somatic cell hybrids. These were prepared by fusing diploid human mononuclear lymphocytes with tetraploid MEL cells which are APRT (Adenine Phosphoribosyl Transferase) deficient, this facilitates selection since the human APRT gene is also on chromosome 16 which carries the human α globin genes. Hybrid cell lines containing human chromosome 16 were treated with DMSO to induce endogenous globin gene expression. Upon induction those cell lines produced detectable quantities of α globin mRNA and protein although at lower levels than the mouse equivalent. An important issue from these studies is that the mechanisms responsible for the repression of the human α genes in a non-erythroid environment are
reversible. The mouse erythroid cell can apparently overrule this genetic restriction. However it is unclear whether this derepression is accomplished in a manner similar to the sequential activation which occurs during normal ontogeny. In a further study Deisseroth et al [167] constructed two series of man x mouse somatic cell hybrids each carrying human chromosome 16. One set was derived from bone marrow cells and the other from haematopoietic non erythroid tissue. In the former series, high levels of human α globin message and protein were produced. In contrast, the non-erythroid hybrid had a lower message content and no α globin chains were detected. This implies that exogenous expression in the MEL cell hybrids is governed not only by the murine erythroid environment but also by the differential state of the human donor cell. A further interpretation suggests the involvement of cis operating controls so leading to the conclusion that the chromatin arrangement in the erythroid cell is more readily recognised by the erythroid regulatory signals in the mouse.

A parallel study on the trans activation of the human β globin gene was performed by Pyati et al [168] with somatic cell hybrid lines derived from induced MEL cells containing human chromosome 11 from lymphoblast cells. An interesting observation on the differential activation of the human β globin type gene was reported by Willing et al [169]. These workers showed activation of both the adult α and β globin genes in MEL x human fibroblast cell
hybrids. However, the \( \gamma \) globin gene sequences remained transcriptionally quiescent. This failure to \textit{trans} activate the human globin foetal genes maybe explained by the absence of mouse foetal factor(s) equivalent to the adult factor(s) since there are no murine foetal globin genes.

\textbf{1:6:2 Induction of Exogenous Globin Gene Expression in Transformed MELtk\textsuperscript{-} Cells}

The introduction of cloned mutagenised globin genes into animal cells is a powerful analytical technique in the identification of cis acting sequences of the prototypic globin transcription unit. Despite this it must be seriously questioned as being a suitable experimental system to examine globin gene regulation.

When the donor DNA is introduced into the recipient cell line and subsequently integrated into the hosts chromatin it generally adopts a more open or relaxed conformation that it does in its endogenous environment where it is transcriptionally silent. Recipient animal cells such as the mouse L-fibroblast or the Hela cell are end stage cells and do not normally undergo any further development. In contrast, the MEL cell is an erythroid cell which is arrested in red blood cell development at the proerythroblast stage [70]. This blockage is affected by an integrated Friend virus. In response to a variety of chemical inducers including DMSO, HMBA (Hexamethylene
Bis-acetamide) and Butyric Acid this cell line can undergo differentiation which is marked by an increase in the expression of the endogenous adult globin genes by 50 - 100 fold. Subsequently the introduction of non-murine globin DNA into the MEL line affords an opportunity to define the sequences involved in the induction process. Two groups have independently investigated the behaviour of cloned human globin genes in MEL cells.

Chao et al [170] established transformed MELtk- lines containing the intact β gene and a 5' mouse/3' human β globin gene. Upon induction both recombinants gave elevated levels of transcription as determined by S1 Nuclease mapping and primer extension mapping. This increase in exogenous globin gene activity of 5 - 50 fold does not parallel the endogenous level of 50 - 100 fold. A possible explanation is that there may be important differences in the chromatin structure or more 5' sequences are required for this quantitative control.

In a similar, but more comprehensive study, Wright et al [171] introduced a series of overlapping tk cosmids spanning the entire 60 kb of the human β globin gene locus into MELtk- cells. Transformed lines were treated with 3mM HMBA and transcripts from the induced and non-induced cells were quantitated by S1 Nuclease mapping. Only the adult β globin gene had elevated levels of expression. The foetal and embryonic genes showed no difference between the induced and non-induced state. This observation closely resembles the behaviour of the γ globin gene in
man x mouse somatic hybrids as described by Willing et al [169]. These workers failed to detect induction of the foetal genes in induced MEL x human hybrids. In contrast, Spandidos and Paul [172] have reported induction of the human epsilon globin gene in transformed induced MELtk- cells on the basis of Northern blot quantitation using a 3' specific epsilon globin gene probe. However not all of the transformed lines gave elevated levels of epsilon globin gene expression.

At present the sequences involved in the induction process are being investigated by comparing the activity of in vitro mutagenised templates with wild type DNA. A study by Charnay et al [173] using a series of human hybrid genes showed that either the human β globin promoter or its coding sequences could confer the property of inducibility. Wright et al [174] showed that the rabbit β globin promoter could be deleted 58 bp's from the cap site without loss of induction, however the level of expression was lower in both induced and non induced cells compared to the gene with an intact promoter.

The induction phenomenon has been investigated in the human erythroleukaemic cell line K562 [87] which on treatment with 30 - 60 μM haemin results in a 3 - 5 fold increase in both the rate of transcription and the level of accumulation of ζ , epsilon, γ and α globin mRNA [175]. Adult β globin message has not been detected in the K562 cell. As outlined in (1:7:3) the fact that the K562 cell contains a functional β globin gene, its non-expression is
presumably explained by the lack of or presence of specific trans factors. This expression and induction pattern of the endogeneous globin genes is mimicked by the equivalent exogenous genes in induced stably transformed K562 cells. Although the K562 cell does not carry a selectable marker, stably transformed cell lines can be generated by co-transforming or covalently linking the globin gene recombinant to the aminoglycoside phosphotransferase gene Tn5 aph (II) [176] (see 1:7:3) which is resistant to the antibiotic geneticin. Kioussis et al [177] used a series of hybrid genes in which the 3' half of the rabbit β globin gene was covalently linked to the 5' half of the human epsilon γ and β globin gene. Upon induction only the epsilon and γ globin gene hybrids produced elevated levels of expression.

In contrast to the above findings Spandidos [178] reported not only expression of an exogenous human β globin gene recombinant but also its co-induction in the K562 cell. Transformed K562 cell lines carrying full length copies of the epsilon, δ and β globin genes were induced by treatment with 1 mM Butyric Acid. Expression was quantified by Northern Blotting [179] using 3' probes. All three recombinants gave increased levels of activity.

1:6:3 trans Activation of Globin Genes by Viral Proteins

The human α and β globin promoters differ in their response to a cis linked viral enhancer in a transient
assay system. A possible explanation for this is that a cellular factor interferes with the transcription of the β globin gene in a transient expression system. Possible candidates to release this repression are the products of the Adenovirus E1A [180] and the Herpes virus α4 gene [181]. These genes which can complement one another [182] in co-infection experiments are powerful modulators of transcription. The E1A protein is required for the activation of all other early adenoviral transcriptional units [180]. Nevins [183] envisages two possible mechanisms for this trans transcriptional activation process. The first proposes that the E1A protein somehow interferes with the cellular factor which otherwise blocks transcription of the adenoviral transcription units. The second mechanism suggests that the E1A product could itself be a positive factor or alternatively, it may activate a quiescent factor. Moreover its ability to induce a heat shock protein [184] has provided credible evidence that it may have a similar effect on the human β globin gene in a transient assay. Green et al [133] examined the effect of the E1A and HSV α4 gene products on the α and β human globin promoters in the presence and absence of the SV40 enhancer. It was observed that these viral proteins can circumvent the necessity of an enhancer for optimal levels of β globin gene expression. Treisman et al [185] showed that this effect is maintained in three different experimental conditions. The first involves co-transfecting the recipient cell line
with an E1A recombinant. The second is by viral transfection. The third is by transforming the human embryonic kidney line 293 [186]. This established cell line constitutively expresses the E1A protein and can complement E1A mutants. Although this protein produced from the 293 cell can trans activate the exogenous globin genes, it does not derepress endogenous transcriptional units. This indicates that there are other forms of control in the activation of the endogenous quiescent gene and that this repression cannot alone be alleviated by the E1A protein. A further interpretation is that at some stage, possibly in the latter events of haematopoiesis, a trans effector, which has E1A like properties, is involved in the sequential activation of the β globin gene.

Although the E1A protein bypasses the need for a recognised enhancer in the expression of the β globin gene in short term assays, the mechanisms of activation of the enhancer and E1A appear to be fundamentally different. To determine the nucleotide requirement of each of these mechanisms, Green et al [133] examined the activity of a series of 5' deletion mutants of the human β globin gene. The enhancer dependent (cis - control) promoter require sequences up to 128 bp's upstream of the cap site, i.e. 5' -100/CCAAT/ATA - cap - 3' for optimal activity. In contrast, the E1A dependent promoter (trans - control) is fully active on a template with 36 bp's of 5' flanking DNA i.e. 5' - ATA/ - cap - 3'.
Aims of Project

The broad aim of this project was to identify putative regulatory sequences within the human β globin gene cluster. With this in mind a Charon 4A human genomic library kindly supplied by Dr. T. Manaitis [22] was screened for human globin genomic sequences. However this library proved to be incomplete in terms of globin sequence representation. The human epsilon globin gene isolated by Proudfoot and Baralle [23] was made available to this laboratory and became the focus of study. The epsilon globin gene like other members of the β globin gene cluster, has a well defined ontogeny (see 1:1:2). It is the first β type globin gene to be expressed and its location at the 5' end of the locus Fig. (2) may provide an opportunity to map the 5' of the cluster. Moreover its expression can be monitored in vivo in the inducible human erythroleukaemic cell line K562 as discussed in (1:7:3).

The weight of experimental evidence suggests that globin genes, once activated during ontogeny are primarily regulated at the level of the initiation of transcription (1:2:2).

Consequently an understanding of the transcriptional control mechanisms may provide some clues to the elucidation of the complexities of tissue specificity, coordinate and temporal regulation of the globin genes.

At the outset of this project several methodologies were available to study the transcription of the globin
Experimental Approaches to Study Globin Gene Expression

The thalassaemias can be considered natural mutants of the human globin gene loci [6]. As shown in Fig. (5A), the molecular lesions responsible for β thalassaemias can be broadly divided into those extensive deletions resulting in the Hereditary Persistence of Foetal Haemoglobin (HPFH) in which the foetal genes are expressed in adult life [6] [46] and those local lesions due to point mutations or small deletions/insertions resulting in a range of promoter and nuclear processing mutants as listed in Fig. (5B). Although these mutants provide a useful source of material for the identification of putative regulatory signals and regions, the juxtapositioning effect of large scale deletions must be recognised when proposing a unifying hypothesis linking lesion and control. Furthermore it is interesting that no trans thalassaemic mutants have been identified, arguably indicative that the diffusible regulatory molecules are of a general type in their mode of action such as E1A [183].

Finally it must be extremely doubtful if the thalassaemia approach would be useful in the study of human epsilon globin gene transcription, since any lesion of physiological significance in the epsilon globin transcription unit would be very difficult to identify due
to the operational problems in such a screening programme. Moreover if the mutation was homozygous it would likely be lethal.

Another considered strategy was to examine the conformation of the chromatin surrounding the globin genes in erythroid and non-erythroid tissue and also in the erythroleukaemic cell lines. This, as previously discussed (1:3), can lead to the identification of putative regulatory regions in terms of differential sensitivity of chromatin to nuclease probes. This approach has the advantage that transcription can be monitored at all stages of development.

1:7:2 Human Globin Gene Expression in Erythroleukaemic Cell Lines

The human erythroleukaemic cell lines including the K562 [87], HEL [187] and the KMOE [188] in addition to the murine MEL cell [70] provide an opportunity to examine the influence of erythroid trans acting factors on the transcriptional behaviour of the human epsilon globin gene.

The K562 cell line established by Lozzio and Lozzio [87] was derived from a patient in the terminal blast crisis stage of chronic myeloid leukaemia. The K562 cell can be regarded as having a mixed phenotype, mainly myeloid with a few erythroid characteristics. All globin genes are expressed in the K562 cell except the adult δ
and β [88]. The mechanism responsible for this block is not understood. Two lines of evidence tend to discount mutation. K562 DNA gives the same restriction endonuclease pattern around the β globin gene as wild type erythroid and non-erythroid DNA [189]. Donovan-Peluso et al [190] isolated the β globin gene from a Charon 21A/K562 library and cloned into the expression vector pBSV which contains the SV40 enhancer [191]. This was transfected into HeLa cells and assayed after an expression period of 48 hours. S1 Nuclease mapping and Northern Blotting showed that the transcripts were identical from those of a normal β globin gene in terms of site of initiation and the size of the processed product. Although the ζ, α, epsilon and γ globin genes are readily expressed in the K562 cell, the level of activity increases 3 - 5 fold after treatment with a range of low MW inducers such as haemin, and butyric acid [175] [192] [193].

The HEL (Human Erythroleukaemic Line) was isolated by Martin and Papayannopoulou [187] with the peripheral blood of a Hodgkins disease victim who subsequently developed erythroleukaemia. This line mainly produces Αγ and Gγ chains and low amounts of epsilon and ζ embryonic polypeptides. No β globin chains have been detected [187]. As was the case with the K562 cell these is no evidence at the Southern Blotting level for mutation in the β globin genomic domain. The HEL line is capable of both spontaneous and chemical induction (20 μM haemin).

The K562 and HEL cells appear to be frozen at the
embryonic/foetal switch and consequently do not express the adult β globin gene. In contrast, the KMOE cell isolated by Okano [188] from a patient with acute erythremia, expresses the adult β globin gene is inducible in the presence of Cytosine-β-D-Arabinofuranoside (Ara-C).

Accordingly these three erythroleukaemic cell lines which span the three gene switches of normal ontogeny may provide a rewarding experimental system to study the molecular mechanisms which govern these events.

However the physiological suitability of these cell types must be questioned. The expression of the globin genes in the non-induced state is indicative that a number of regulatory events have already occurred. Even the silent β globin gene in the K562 cell is DNase I sensitive [194]. In addition concern must be recognised on the neoplastic nature of these cell lines - do these cell types mirror the structure and regulation of globin chromatin during normal differentiation?

1:7:3 Human Globin Gene Expression In Transformed Animal Cells

The combined technologies of reverse genetics and in vitro mutagenesis of cloned cellular sequences have provided a powerful assay system in the search for putative regulatory elements. This technology has been responsible for determining the physiological importance of a number of sequences including promoter regions,
enhancers and processing signals. In vivo transcriptional assays measure gene activity in the living animal cell where the foreign cellular gene is transcribed by the host's RNA Polymerase system.

A variety of methods are available to introduce foreign DNA into eucaryotic cells. The most extensively used and best characterised technique is the calcium phosphate precipitation of DNA (see 2:4:2) pioneered by Graham and van der Eb a number of years ago [195]. This technique can also be performed with whole chromosomes [196].

DNA can also be introduced into animal cells via liposomes [197], by DEAE-dextran [198] and by protoplast fusion [199]. More recently retroviral vectors have been constructed which package the donor DNA in vitro prior to transfection [200]. Genes can also be introduced directly into the cell nuclei by a variety of methods including microinjection [201], and by electrical stimulation of the membrane (electroporation) [202] [203] [204]. In addition DNA can be introduced directly into the oocyte [205] and germ cells [206], the latter resulting in transgenomes.

Independently of how the DNA enters the cell, the biological activity of the donor DNA can be assayed either after a short expression period of usually between 36 and 48 hours, this method is termed short term, transient or acute transformation. Alternatively expression can be determined after a longer period commonly in the region of 1 to 3 weeks, termed long or stable transformation.
In transient expression experiments, the transcriptional activity of the donor DNA is assayed directly by analysis of RNA or by measuring the protein product. When the exogenous DNA is fused to a gene coding for an enzyme, the amount of protein in the cell can be regarded as a measure of the regulatory sequences driving expression of the gene encoding for the enzyme. Suitable genes for this purpose include amongst others CAT [207], XGPRT (Xanthine Guanine Phosphoribosyl Transferase) [208] and HSV-1 tk [209].

In addition expression can be estimated by immunofluorescence of the SV40 T-antigen [210]. The physical form of the donor DNA is thought to be episomal some of which form mini chromosomes [211]. Immediately after uptake the copy number is high, around 20,000 to 100,000 copies per cell. This peaks around 48 hours after which the number drops dramatically. The mechanism responsible for this is unknown. When the foreign DNA encodes for a selectable marker such as tk, transformed cells can be generated by growth in the appropriate selective medium usually over a 1 - 3 week period.

The conformation of donor DNA in biochemically transformed cells is a debatable issue. Current opinion subscribes to the view that it becomes integrated into the host's genome. Although the details of this mechanism are essentially unknown, some sort of non-reciprocal homologous recombination process is thought to occur resulting in the formation of a series of head to tail
concatenate structures [212]. In contrast some authors have reported that foreign sequences can exist as free forms [213] [214]. This has been well documented for BPV derived recombinants but is considered to be a specialised case since BPV specifically encodes for PMS (Plasmid Maintenance Sequences) [215].

Long term transformation is possible using two alternative strategies. Firstly, the DNA of interest can be engineered into a plasmid already containing a selectable marker. Alternatively the cloned DNA may be co-transformed with marker DNA. Upto 90% of the transformed colonies contain the non selectable DNA [216]. Readily available deficient cell lines include APRT [217], HGPRT (Hypoxanthine Phosphoribosyltransferase) [218], Galactokinase [129] and DHFR (Dihydrofolate Reductase) [220].

The two markers used during the course of this project were HSV-1 tk [221] and Tn5 aph(II)[176] (see 2:4:3 I).

For technical details and for a comprehensive review of these gene transfer techniques except electroporation and the use of transgenomes, see Spandidos and Wilkie [222]

1:7:4 Experimental Plan

The experimental method chosen to study human epsilon globin gene expression was reverse genetics. Two different in vivo transcriptional assay systems were to be employed.
First, to compare directly epsilon globin gene activity in K562 cells and to identify sequences in the in vivo transcription of this gene, a series of recombinants derived from the lambda clone λ788ε were to be constructed. Such plasmids would then be used to stably transform tk-deficient cell lines and epsilon globin gene specific expression studied by S1 Nuclease mapping.

The second assay system was the hybrid selectable gene technique. In such circumstances it can be contended that the different levels of transcriptional activity of the linked gene is a function of the fused globin DNA.
CHAPTER 2: Materials and Methods

2:1 Materials

| 2:1:1 | General Suppliers  | 65 |
| 2:1:2 | Standard Buffers  | 65 |
| 2:1:3 | Equipment  | 66 |
| 2:1:3I | Centrifugation  | 66 |
| 2:1:3II | Spectrophotometry  | 67 |
| 2:1:3III | pH Measurement  | 67 |
| 2:1:3IV | Scintillation Counting  | 67 |
| 2:1:3V | Photography  | 67 |

2:2 Routine Procedures

| 2:2:1 | Purification and Precipitation of Nucleic Acids  | 68 |
| 2:2:2 | Deionisation of Formamide  | 69 |
| 2:2:3 | Nick Translation  | 69 |
| 2:2:4 | Filter Hybridisation Conditions  | 70 |
| 2:2:5 | Analytical Gel Electrophoresis  | 72 |
| 2:2:6 | Isolation of DNA from Low Melting Point Agarose Gels  | 72 |
| 2:2:7 | Isolation of DNA from Neutral Poly-acrylamide Gels  | 73 |

2:3 Molecular Cloning

| 2:3:1 | Screening of Human Charon 4A Genomic Library  | 73 |
| 2:3:2 | General Cloning Strategies  | 75 |
| 2:3:2I | Dephosphorylation of DNA  | 76 |
2:3:2II Ligations 77
2:3:2III Molecular Linker Ligations 79
2:3:2IV Preparation of Competent HB101 Bacterial Cells 80
2:3:2V Transfection of Competent HB101 Bacterial Cells 80
2:3:2VI Screening Bacterial Colonies 81
2:3:2VII Large Scale Preparation of Plasmid DNA 82
2:3:3 Specific Manipulations 84
2:3:3I Bal 31 Nuclease Treatment 84
2:3:3II Maxam and Gilbert DNA Sequencing 85
2:3:3III Tha I/Eco RI Digestion of Recombinant DNA 85

2.4 Transformation of Cultured Animal Cells 86

2:4:1 Cell Lines and Culture Conditions 86
2:4:2 Calcium Phosphate Co-precipitation of Donor DNA 87
2:4:3 Long Term Transformation 88
2:4:3I Selectable Biochemical Markers 89
2:4:3II Selection in 0.9% Methocel Medium 92
2:4:3III Selection in Liquid Medium 92
2:4:3IV Quantitation of Colony Numbers 93
2:4:4 Short Term Transformation 93

2:5 The Detection and Conformation of Donor DNA in Transformed Cell Lines 94

2:5:1 Preparation of High Molecular Weight DNA 94
2:5:2 Southern Blotting Analysis 95

2:6 The Transcription of Donor DNA in Transformed Cell Lines 96

2:6:1 Preparation of Total Cellular RNA 96
2:6:1I Caesium Chloride Gradient Method 97
2:6:1II  Lithium Chloride Precipitation Method  97
2:6:2   Preparation of PolyA⁺ RNA       98
2:6:3   RNA Dot Blotting Analysis      99
2:6:4   Northern Blotting Analysis    100
2:6:5   S1 Nuclease Mapping          100
2:6:5I  Preparation of 5' End Labelled Single Stranded Probe  101
2:6:5II Hybridisation and S1 Nuclease Treatment  102
2:6:5III Denaturing Polyacrylamide Gel Electrophoresis  103
2:1 **Materials**

2:1:1 **General Suppliers**

All chemicals used were of Analar grade and supplied by BDH Chemicals unless otherwise stated.

All radioisotopes including 'Nick-Translation' kits were obtained from the Radiochemical Centre, Amersham International.

Restriction endonucleases and enzymes employed in the manipulation of DNA were purchased mainly from BRL Gibco (Europe), P-L Biochemicals Inc., Biolabs and Boehringer-Mannheim.

Tissue culture medium, supplements and serum were obtained from Gibco (Europe) or Flow Laboratories. Sterile disposable plastic flasks of growth areas 25, 75 and 175 cm² the culture of animal cells were supplied by Falcon and 'Nunc.

Ingredients of bacterial and bacteriophage media viz Tryptone, Yeast and Agar were obtained from Difco Laboratories.

2:1:2 **Standard Buffers**

Common multipurpose buffers referred to throughout the Materials and Methods and Results Sections of this thesis include the following.

Tris-HCl was prepared by dissolving Tris base (2-amino
2-hydroxymethyl-1,3-propanediol) in dist H₂O (distilled water) followed by the addition of conc HCl (concentrated hydrochloric acid) until the desired pH was attained.

T.E. Buffer (Tris-EDTA) (Ethylene diamino trichloro-acetic acid) comprises of {1mM tris HCl pH 8.0 0.1mM EDTA}. T.B.E. Buffer (Tris-Borate EDTA) is {89mM tris base/89 mM boric acid/1mM EDTA pH 8.3}. 20 x SSC (Saline Sodium Citrate) is {3M NaCl/0.3M Sodium Citrate}.

These buffers and others mentioned throughout the text were sterilised by autoclaving at 15p.s.i. (Pounds Square Inch⁻¹) for thirty minutes. For small volumes and solutions containing heat labile compounds, filter sterilisation through an 0.2 μm Aerodisc (Gelman U.K.) was carried out.

2:1:3 Equipment

2:1:3(I) Centrifugation

Micro-centrifugation of 1.5ml Eppendorf tubes at ~12,000 rpm (8000g) were performed in a Eppendorf 5414, and IEC Centra-35 or a MSE-Micro Centaur. Centrifugation below 5,000rpm was carried out in a MSE Coolspin. A Sorvall RC5B model was used for spins up to 15,000rpm. For ultracentrifugation between 35,000 and 40,000rpm, the Diamon IEC 60, the MSE Superspeed 65 or the MSE Prepspin 50 models were used.
2:1:3(II) Spectrophotometry

Spectrophotometric determinations of nucleic acid concentration and the optical density of bacterial cultures were estimated using a Varian 634 with a Model 9176 Chart Recorder or a Cecil CE594.

2:1:3(III) pH Measurement

pH values were determined with a Kent EIL 7045/4C Model.

2:1:3(IV) Scintillation Counting

Radioactivity measurements were performed on a LKB 1215 Rackbeta Liquid Scintillation Counter.

2:1:3(V) Photography

Both agarose and polyacrylamide gels were photographed with a Polaroid MP4 Land Camera using Polaroid type 55 or type 57 film under a UV transilluminator (Ultra Violet Products Inc. Model TL-33). Autoradiographs were automatically developed in a Kodak-X-Omat Processor Model ME-1A.
2:2 Routine Procedures

2:2:1 Purification and Precipitation of Nucleic Acids

Distilled phenol was made 0.1% Hydroxyquinoline and 0.2% β-Mercaptoethanol (βME). This was neutralised by the sequential addition of 1M and 100mM tris HCl pH 8.0 to give a standard equilibrated phenol solution. Phenol:Chloroform (CHCl₃) was prepared by adding an equal volume of CHCl₃ : Isoamyl alcohol 24:1 (v/v) to distilled phenol/0.1% Hydroxyquinilone/0.2% βME. This was neutralised as outlined above. Nucleic acid solutions were extracted with an equal volume of phenol or phenol: CHCl₃. The upper aqueous phase was separated by centrifugation in a Microcentrifuge (2:1:3(I)). Ether extractions were carried out to remove traces of phenol or CHCl₃ in the aqueous phase. The sample was mixed with four volumes of diethyl ether. Micro-centrifugation for 30 seconds separated the upper ether layer from the nucleic acid solution. This was heated for ten minutes at 65°C to evaporate any remaining ether.

Extractions with butan-2-ol were employed to reduce the volume of the DNA solution by the sequential addition of an equal volume of butan-2-ol followed by vortexing and brief micro-centrifugation until the desired volume was attained. The sample was finally ether extracted to remove traces of butan-2-ol. It should be noted that the salt concentration increases with volume reduction.
Organic extractions were invariably followed by precipitation of the nucleic acid in 200mM NaAc (sodium acetate) and three volumes of ice cold ethanol. This was left either in dry ice for thirty minutes or overnight at -20°C. The nucleic acid was pelleted by micro-centrifugation for four minutes, washed in 70% ethanol and dried under vacuum in a Modulyo Edwards Freeze Drier.

The above procedures were followed where organic extractions or ethanol precipitations are referred to in the Materials and Methods.

2:2:2 Deionisation of Formamide

Formamide which reduces the melting temperature (Tm) of a hybrid was used extensively in hybridisation solutions involved in Southern Blotting, Northern Blotting, RNA Dot Blotting and S1 Nuclease Mapping, initially had to be purified as contaminants interfere with the hybridisation process. This was achieved by treating Formamide (Fluka) with AG 501 - x 8(D) ion-exchange resin (Bio-Rad) with constant agitation for two hours at 4°C. The formamide was filtered through a Whatman No.1 filter.

2:2:3 Nick Translation

Two methods were employed to radiolabel DNA by the
'Nick Translation' technique. The first protocol is based on a modification of the method of Rigby et al [223]. A 50 µl reaction mixture was prepared containing \{0.25µg DNA /2.5 µM of dATP, dGTP and dTTP/50µCi α32PdCTP (400 ci/mmol) /1 x Nick Translation Buffer (50mM Tris-HCl pH 7.2, 10mM MgSO4, 10mM DTT, 50 µg/ml BSA)\}. This was followed by the addition of 1 µl (0.1 µg/ml) DNase I and five units of E coli DNA Polymerase I (Boeringher-Mannheim) and incubated for two hours at 12°C. The mixture was then applied to a 1XSSC equilibrated G50 Sephadex (Pharmacia) column (0.8 x 40 cm). 1 ml fractions were collected and 10 µl aliquots counted in Packard insta-gel scintillation fluid (2:1:3(IV)). Yields varied from 2 - 8 x 10^7 cpm/µg DNA.

The second approach, used the Amersham International Nick Translation Kit N 6000. 0.25 - 0.5 µg of DNA was incubated with 4 µl of Solution 1 (Nucleotide/Buffer Solution), 1 µl of Solution 2 (Enzyme Solution), 50 µCi d CTP in a final volume of 20 µl. This was incubated and separated as detailed in the first method. Yields were of a similar order as above.

2:2:4 Filter Hybridisation Conditions

Nitrocellulose filters after Southern Blotting (2:5:2), RNA Dot Blotting (2:6:3) and Northern Blotting (2:6:4) were prehybridised and hybridised according to the following protocol which is a modification of the conditions reported by Wahl et al [224].
Prehybridisation was carried out overnight in a shaking waterbath at 42 °C in sealed plastic bags containing (50% Deionised Formamide (v/v)/6 x SSC/ 5 x Denhardt's solution [225] [1 x Denhardt's = 0.02% Bovine Serum Albumin Fraction V (BSA); 0.02% Ficoll 400; 0.02% Polyvinylpyrolidone (PVP) (all Sigma)]/50mM Sodium phosphate (NaHPO₄) pH 6.5/1% Glycine/250 μg/ml} heat denatured sonicated Salmon Sperm DNA (Sigma).

After prehybridisation the bag was drained and replaced with a solution of {50% Deionised Formamide/6XSSC /1 x Denhardts Solution/50mM NaHPO₄ pH 6.5/ 100μg/ml heat denatured sonicated Salmon Sperm DNA} Approximately 2-5 x 10⁶ counts per minute (c.p.m.) of nick translated probe was added for a Southern Blotted filter and 1 - 3 x 10⁷ c.p.m. for a RNA blot. The filters were agitated overnight at 42°C.

After draining the plastic bags, the filters were initially rinsed in 2XSSC. This was followed by two thirty minute washes at 65°C in 2XSSC/0.1% SDS (Sodium Dodecyl Sulphate). Depending on the expected degree of cross hybridisation the stringency was increased to finally two thirty minute washes at 65°C in 0.1 x SSC/0.1% SDS. The filters were finally rinsed in the appropriate SSC concentration, blotted on 3MM paper, air dried and exposed to Kodak XRS X-ray film in the presence of a Dupont Kronex intensifying screen at -70°C.
2:2:5 Analytical Gel Electrophoresis

Depending on the size of the expected diagnostic bands, restricted DNA was electrophoresed through an agarose or polyacrylamide gel.

DNA samples either in restriction enzyme buffer or in T.E. after ethanol precipitation were mixed with 1/10th volume of gel loading dye (0.1% bromophenol Blue/30% sucrose/10mM Tris - HCl pH 8.0/0.1mM EDTA). The samples were electrophoresed either in:

a) a horizontal 0.5 - 2% agarose (BRL) gel in 1 x {50mM Tris - HCl/20 mM NaAc/ 1 mM EDTA pH 7.8} or 1 x T.B.E. The DNA was visualised under u.v. after staining in a 0.1μg/ml solution of ethidium bromide (EtBr).

b) a vertical neutral polyacrylamide gel containing {6% acrylamide - 2% bis - acrylamide/1 x T.B.E./0.004% ammonium persulphate/0.001% N,N,N,N - tetramethyl - ethylene diamine (TEMED)}. The DNA was visualised as above.

Running conditions varied with convenience.

2:2:6 Isolation of DNA From Low Melting Point Agarose Gels

Low Melting Point Agarose (BRL) was solublised in 1 x
T.B.E. by heating, chilled and poured at 4°C. Migration of DNA is slower through low melting point agarose than standard agarose and accordingly was routinely electrophoresed in 1 x T.B.E. overnight. The bands were visualised under u.v., excised and placed in 2 volumes of T.E.

This was heated to 70°C for 10 minutes, extracted three times with tris equilibrated phenol, ether extracted and ethanol precipitated.

2:2:7 Isolation of DNA From Neutral Polyacrylamide Gels

Bands from a 4 - 6% neutral polyacrylamide gel were excised and transferred to a heat sealed blue tip plugged with siliconised glass wool. The gel slice was crushed with a nineteen inch gauge needle and incubated overnight at 37°C in 400 µl of gel elution buffer (0.1% SDS/0.2M NaAc). The DNA was eluted by piercing the tip and collected by gravity. A further 400 µl of gel elution buffer was added and collected. The volume was reduced by treatment with butan-2-ol, phenol and ether extracted and ethanol precipitated.

2:3 Molecular Cloning

2:3:1 Screening of Human Charon 4A Genomic Library

A partial HaeIII/AluI Charon 4A human genomic library was kindly supplied by Dr T. Maniatis [22]. This was screened with nick-translated human β and γ globin c
DNA probes [226] [227] using a modification of the procedure of Benton and Davis [228].

For an average insert size of ~20 kb, the number of pfu's (plaque forming units) required to give a 99% probability of sequence representation is \( \approx 10^6 \) pfu's as estimated from the Clarke and Carbon equation [229]. \( \approx 10^6 \) pfu's were mixed with \( \approx 3 \times 10^{10} \) of exponential phase LE 392 cells [230] in 1 mM MgSO\(_4\) and left at room temperature for ten minutes to allow phage absorption. The mixture was added to 45°C 0.7% L-agar and aliquotted onto 20 x 14 cm petri dishes containing 1.2% L-agar. The plates were incubated overnight at 37°C to give 50,000 pfu's per dish. The plates were refrigerated for at least two hours prior to blotting.

Autoclaved 0.4 \( \mu \)m filters (Millipore) marked for autoradiograph alignment were gently placed on correspondingly marked petri dishes and left for two minutes at room temperature to absorb the phage followed by a duplicate impression. The filters were denatured by blotting on 1.5M NaCl/0.5 M NaOH saturated 3 M paper for five minutes at room temperature. Neutralisation followed by soaking in 3 M NaCl/0.5 tris pH 8.0 for ten minutes. The filters were dried, baked, prehybridised, hybridised, washed and autoradiographed as detailed in (2:2:4).

After initial screening autoradiographs with a duplicate hybridisation signal were aligned to the corresponding petri dish and an area of 1 cm\(^2\) excised, resuspended in 2 ml of phage buffer (100 mM NaCl/8 mM
MgSO$_4$ 7H$_2$O/100 mM tris-HCl (pH 7.5) 0.01% Gelatin) and left overnight at room temperature. This was plated out and screened as previously described. This procedure was repeated until ≥ 90% of the plaques gave a corresponding hybridisation signal.

DNA was prepared from positive clones by a modification of the procedure of Yamamoto et al [231]. A single plaque was expanded until enough phage was generated to lyse 20 x 14 cm L-agar plates. The petri dishes were scrapped and the agar washed with 50 ml of phage buffer and agitated for fifteen minutes. The agar was pelleted by centrifugation (10,000 rpm/10 minutes/4°C) in a Sorvall RC5B. The supernatant was made 0.5 M NaCl and 10% (w/v) PEG 600 (Polyethylene glycol) and left to dissolve on ice for two hours. After centrifugation (6,000 rpm/30 minutes/4°C) the pellet was resuspended in 10 ml of phage buffer by gentle passage through a syringe and 19 gauge needle. 0.71 g of CsCl (caesium chloride) (BRL-Gibco) was added per ml of suspension and centrifuged (40,000 rpm/24 hours/4°C). The white phage DNA band was carefully removed by a Pasteur pipette and extensively dialysed against 5 mM tris/2 mM MgSO$_4$. The solution was extracted with phenol:CHCl$_3$ and ethanol precipitated.

2:3:2 General Cloning Strategies

The recombinants constructed for study in this thesis can be divided into three broad categories. The first as
described in (3:1) of the Results involves the construction of a series of plasmid chimeras derived from the phage recombinant λ788ε [25] containing the human epsilon globin gene and λM3 isolated from the human Charon 4A genomic library which contains the human δ and β globin genes. The other two class of plasmid recombinants are detailed in (3:3) of the Results.

The first concerns the insertion of human globin genomic fragments into vectors derived from pre-existing plasmids kindly supplied by Dr J. Lang [232] in this laboratory. The final series of recombinants involve the construction of a series of Tn5 aph (II) plasmids driven by different promoters, the engineering of which in some cases involved the use of synthetic molecular linkers. Accordingly due to the wide range of cloning procedures it has been decided to give an account of general or common recombinant DNA manipulations. Finally in cases where specific manipulations were undertaken these will be outlined in (2:3:3.)

2:3:2(I) Dephosphorylation of DNA

Plasmid or phage DNA was digested with the appropriate restriction endonuclease using conditions recommended by the suppliers. The reaction was generally heat-killed by incubation at 65°C for five minutes. This was followed by one phenol : CHCl₃ extraction and ethanol precipitation. Calf Intestinal Phosphatase (CIP) (Boehringer-Mannheim)
catalyses the removal of 5' phosphate groups from DNA. CIP is supplied as an ammonium sulphate suspension. This was centrifuged in a Microfuge and the supernatent diluted in distilled H₂O to 0.01U μL⁻¹. This concentration of enzyme will remove the terminal phosphates from 1 pmole of 5' end of DNA in a 50 μl reaction mixture containing {5 μl 10 x CIP Buffer (0.5 M tris pH 9.0, 10 mM magnesium chloride (MgCl₂), 1 mM zinc chloride (ZnCl₂), 10 mM spermidine (Sigma Co), 0.01 U CIP μg⁻¹DNA}. This was incubated at 37°C for thirty minutes. A second equivalent aliquot of enzyme was added and the above incubation conditions repeated. The reaction was terminated by heating the mixture at 68°C for fifteen minutes. This was followed by two phenol : CHCl₃ extractions and one diethyl ether extraction, ethanol precipitated, washed in 70% ethanol, dried and resuspended in TE, ready for ligation.

2.3.2(II) Ligations

A DNA molecule with identical cohesive or blunt-ended termini can undergo circularisation or concatenation in the presence of T₄ DNA Ligase. Two parameters termed j and i [233] determine which event will be favoured. j describes the effective concentration of two ends of the same DNA molecule and is derived from the equation j = (3/2πlb) ends/ml where l is the length of the DNA in cms, b is the length of randomly coiled DNA and varies with ionic strength. Parameter i is a measure of the
concentration of complementary termini and is given by the equation:

\[ \text{\textit{l}} = 2\text{No} \times 10^3 \text{ends/ml} \]

where No = Avogadro's No and M = molarity.

When \( j = i \) the probability of a DNA molecule colliding either with the other end of the same molecule or another is equal. Thus when \( j > i \) circularisation is favoured such as in the construction of pMX (see 3:1:2(II)). In conditions of \( i > j \) concatenation will be the more frequent event.

Cloning experiments with two or more different molecules, the vector plasmid was routinely phosphatased (2:3:2) to reduce the probability of self-ligation and the ligation reaction driven by insert DNA usually at a range of increasing molar ratios from 1:1 to 1:5. For around 5 pmoles of vector ends, the DNA's were incubated in ligation buffer (66mM tris - HCl pH 7.6, 6.6 mM MgCl\(_2\), 10 mM dithiothreitol (DTT), 0.4 mM adenosine triphosphate (ATP)) in the presence of 0.5 - 2.0 units of T4 DNA Ligase at 22°C for one hour, or overnight at 42°C. The reaction was heat killed by incubation for 15 minutes at 65°C.

For blunt-end ligations such as in the construction of the pMX promoter deletion mutant series (see 3:1:3), more T4 DNA ligase was used, usually a 10 fold increase than that required for "sticky" end ligations, in addition the incubation period was longer, routinely overnight.
Molecular Linker Ligations

It was sometimes necessary to change the restriction recognition site to another to permit cloning into a particular vector such as in the construction of pNβV (see 3:3:2(IV)). Protruding 5' ends were flush-ended by treatment with Klenow enzyme. This large fragment of bacterial DNA polymerase I has 5' -> 3' polymerase activity and 3' -> 5' exonuclease activity. Restricted DNA, usually ~1 μg in a 25 μl reaction volume was incubated at 22°C for 30 minutes in the presence of {1 μl of each 2 mM dNTP, 2.5 μl 10 x Nick Translation Buffer (0.5 M tris pH 7.2, 0.1 M MgSO₄, 1 mM DTT, 500 μg/ml BSA)} and 2 units of Klenow enzyme. The reaction was heat killed at 70°C for 5 minutes. This was extracted with phenol: CHCl₃, diethyl ether, ethanol precipitated and the DNA recovered from a gel to ensure the separation of dNTP's which can interfere with subsequent ligations. The blunt-ended DNA was converted to the desired restriction site by the ligation of kinased molecular linkers [234]. Linkers are small oligonucleotides of 8 - 12 bp's and accordingly it is easy to achieve a high concentration of ends which drive the blunt-end ligations. Optimum conditions are between 20 and 50 μg/ml. Around 0.5 μg of blunt-ended DNA was ligated to a 50 fold excess of linkers in the presence of 1 x Ligase Buffer and two units of T4 DNA Ligase. This was incubated for six hours at 22°C or overnight at 4°C, heat killed, phenol: CHCl₃ extracted and ethanol
precipitated. The linkered DNA was then digested with a large excess of the appropriate enzyme (10 - 20 fold). After several hours digestion the DNA was ethanol precipitated after the standard extraction procedure and the DNA separated from any remaining linker molecules by gel separation.

**2:3:2(IV) Preparation of Competent HB101 Bacterial Cells**

Cells competent for DNA uptake were prepared from the E.coli K12 strain HB101 [235] using a modification of the method of Mandel and Higa [236]. A 1L culture was prepared according to the following protocol. A single colony was seeded into 10 ml of L-broth and was incubated overnight with agitation at 37°C in a Gallenkamp Orbital Incubator until the OD 600 was ~0.6. The culture was chilled on ice, centrifuged in a Sorvall RC5B (5,000 rpm/10 minutes/4°C) and left on ice for thirty minutes. The cells were harvested as above and resuspended in 500 ml of 50 mM CaCl$_2$, left on ice for fifteen minutes and centrifuged as previously outlined. The pellet was resuspended in 50 ml of (50 mM CaCl$_2$ /15% glycerol), dispensed into 1 ml aliquots and stored in liquid nitrogen for up to six months without significant loss in a transformation efficiency of around 1 - 5 x 10$^6$ colonies per 1 μg pBR322.

**2:3:2(V) Transfection of Competent HB101 Competent Cells**

Ligated DNA, the amount of which was determined by the
The anticipated number of ampicillin-resistant colonies was made up to 200 µl with transformation buffer (50 mM CaCl₂/10 mM tris-HCl pH 8.0). Competent HB101 cells were removed from liquid nitrogen storage and thawed on wet ice. A 200 µl aliquot was added to the DNA and the mixture left on ice for 25 minutes to allow DNA uptake. The cells were then heat shocked for two minutes at 42°C in a water-bath. The sample was immediately transferred to a bijou containing 2 ml L-broth and incubated at 37°C for sixty minutes without shaking. 2 x 200 µl aliquots were spread over 1.2% L-agar on 9 cm petri dishes containing 100 µg/ml ampicillin (Sigma). The plates were dried, inverted and incubated overnight at 37°C.

2:3:2(VI) Screening Bacterial Colonies

Colonies were picked with a flamed loop and incubated overnight with vigorous agitation in 2 ml of L-broth containing 100 µg/ml ampicillin. Plasmid DNA was prepared from these cultures according to the method reported by Birboim and Doly [237]. 1 ml aliquots in 1.5 ml eppendorf tubes were microfuged for thirty seconds and the pellet resuspended by vortexing in 100 µl of lysis buffer {2 mg/ml lysozyme (Sigma)/25 mM tris - HCl pH 8.0/10 mM EDTA/50 mM glucose}. This was left on ice for thirty minutes followed by the addition of 200 µl alkaline/SDS solution (0.2 N NaOH/1% SDS) which was incubated on ice for a further five minutes. The clear viscous solution
was neutralised by the addition of 150 μl 3M NaAc pH 4.8, incubation was continued on ice for sixty minutes. The resultant heavy coarse precipitate was pelleted by microcentrifugation for four minutes. 400 μl of the supernatent was precipitated in the presence of 1 ml of cold ethanol at -20°C for thirty minutes. The pellet was harvested by microcentrifugation for three minutes and resuspended in 100 μl {0.1 M NaAc/0.05 M tris-HCl pH 8.0} and precipitated by the addition of 200 μl of ice cold ethanol for ten minutes at -20°C. This was microfuged, washed in 70% ethanol and dried. The pellet was dissolved in dist H₂O and was ready for restriction enzyme digestion analysis. Where diagnostic fragments were less than 600 bp's the samples were treated with pancreatic RNase. DNA in 36 μl of dist H₂O was incubated with 4 μl 1 mg/ml pancreatic RNase (heat treated at 100°C for ten minutes) at 37°C for thirty minutes. Following this the samples were subject to analytical gel electrophoresis.

2:3:2(VII) Large Scale Preparation of Plasmid DNA

This protocol is based on the method of Birboim and Doly [237] modified by Ish-Horowicz [238]. From an overnight 50 ml culture an aliquot was added to 500 ml L-broth containing 100 μg/ml ampicillin to give an initial OD600~0.05. Incubation at 37°C in an orbital shaker followed until the OD600 increased to~0.6. The culture
was amplified in the presence of chloramphenicol (Sigma) at 170 μg/ml, incubation was continued overnight.

The cells were harvested and washed in 50 mM tris-HCl pH 7.8 in a Sorvall RC5B (10,000 rpm/10 minutes/4°C). The pellet was resuspended by gentle passage through a 19" gauge needle in 25 ml of lysis buffer {5mg/ml of lysozyme in 25 mM tris-HCl pH 8.0/10 mM EDTA/50 mM glucose} and left on ice for thirty minutes. This was followed by the addition of 40 mls alkaline/SDS {0.2 M NaOH/1% SDS}, gently mixed and left on ice for five minutes. The solution was neutralised by the addition of 20 ml of 3M KAc pH 4.7, this was vigourously shaken and incubated on ice for a further fifteen minutes and centrifuged (8,500 rpm/5 minutes/4°C). The supernatent was filtered through two layers of sterile gauze. 0.6X volume of cold propan-2-ol was added to the clear supernatent and immediately centrifuged as above. The pellet was dried over nitrogen gas and resuspended in 6.8 ml T.E.

After the addition of 7.2 g CsCl and 0.6 ml 10 mg/ml EtBr the solution was centrifuged (40,000 rpm/40 hours/20°C). The plasmid band was removed by pipetting and extracted five times with an equal volume of propan-2-ol. The solution was extensively dialysed against T.E. in a collodion bag (Sartorius), ethanol precipitated, washed in 70% ethanol, dried and resuspended in T.E.
Specific Manipulations

Bal 31 Nuclease Treatment

Bal 31 Nuclease, which degrades both termini of linear DNA was employed to digest from the XbaI site of pMX (see Fig. 13) to create a series of promoter deletion mutants (see 3:1:3). Although the rate of removal of nucleotides can be calculated from a formula [238] it was decided to estimate this empirically by following the protocol outlined by Legerski et al [239] and suggested by the commercial suppliers of the enzyme (Gibco - BRL).

20 μg of pMX was digested with XbaI, phenol extracted, ethanol precipitated and resuspended in distilled H₂O. A reaction mixture containing the linearised DNA in 1 x {12 mM CaCl₂/12 mM MgCl₂/200 mM NaCl/20 mM tris - HCl pH 8.0/1 mM EDTA} was incubated at 30°C after the addition of an appropriate amount of enzyme. 2 μg aliquots were removed at different time points and quenched by adjusting the EDTA to a final concentration of 20 mM. Bal 31 Nuclease is dependent on the presence of calcium ions for catalytic activity. 1 μg aliquots of these time points were doubly digested with PvuII and Bam H1 and gel electrophoresed. Samples which retained the PvuII promoter site at coordinate 8557 bp's and extended beyond the Bam H1 promoter site at coordinate 8360 bp's (see Fig. 11) were blunt-end ligated and transformed into HB101 cells.
The pMX promoter deletion series generated as outlined above (2:3:3(I)) were sequenced by the method of Maxam and Gilbert [240]. The strategy was to $^{32}$P end label at the promoter PvuII site and sequence upstream and compare it to the published wild type sequence [42].

20 μg PvuII digests of pMX promoter deletion DNA's were phenol extracted, ethanol precipitated, washed in 70% ethanol and resuspended in 11 μl of dist H$_2$O. To this was added 15 μl of 2XT4 Polymerase Buffer (130 mM tris pH 8.0/13 mM MgCl$_2$/10 mM DTT), 3 x 1 μl of 2 mM dNTP's, 75 μCi α$^{32}$P dGTP and 1.5U T4DNA Polymerase. The reaction was incubated at 11°C for three hours. The end labelled DNA was separated from the non-incorporated nucleotides by a rapid ethanol precipitation, two minutes on dry ice. The DNA precipitate was rinsed in 70% ethanol and resuspended in dist H$_2$O. Depending on the restriction map of the deletion mutant (see Fig. 14) the PvuII labelled digest was treated with Smal, EcoRI, HindIII, PstI or SalI and electrophoresed on an agarose gel. The labelled fragment was electroeluted into a dialysis sac and ethanol precipitated. The DNA was resuspended in 60 μl of dist H$_2$O and aliquoted for the five sequencing reactions according to the Maxam and Gilbert [240] method described in [241].

The restriction endonuclease enzyme ThaI with the
recognition sequence GC+GC [242] has twenty three sites in pBR322. In contrast its frequency in eucaryotic DNA is very low. This difference was exploited in the molecular cloning of the EcoRI human globin genomic fragments of pHR10 and pHR11 in the neomycin based vector pNeV (see 3:3:2(VI)). The pAT153 (pBR322) derived recombinants pHR10 and pHR11 were initially treated with ThaI at 60°C for the appropriate time. The reaction mixture was transferred to a 37°C water bath and digested with EcoRI. This strategy essentially isolates the globin genomic fragment and consequently reduces the number of possible recombination events.

2:4 Transformation of Cultured Animal Cells

All transformation procedures and cell handling techniques essentially follow those extensively described by Spandidos and Wilkie [222].

2:4:1 Cell Lines and Culture Conditions

tk deficient cell lines used throughout this study were: the fibroblastic mouse LA tk⁻ [221], BHK tk⁻ (Baby Hamster Kidney) [243] were obtained from the Institute of Virology, University of Glasgow. The MEL tk⁻ clone F4-12B2 was a gift from Dr Ostertag of the University of Hamburg. The human fibroblast 143 tk⁻ line [244] derived from the Kirsten Ras Murine Sarcoma Virus transformed line
R970 [245] was a gift from Dr Bacchetti, McMaster University, Ontario. The other cell lines employed were the human erythroleukaemic line K562 [87] (1:7:2). The EJ cell line is a bladder carcinoma cell line [246] and was supplied by Dr A. Balmain (Beatson Institute).

All sterile tissue culture techniques were performed in a laminar flow hood. Cells were passaged by sucking off the medium, washing with P.E. (Dulbecco's Phosphate Buffered Saline (PBS)/1 mM EDTA. The cells were removed by the addition of 1 ml of 0.25% trypsin solution per 75 cm$^2$ flask. Fresh medium was added and the cells harvested by centrifugation in a bench MSE at 2,000 rpm for five minutes. The pellet was resuspended and grown in SLM (Special Liquid Medium - Gibco Europe) supplemented with 4 mM glutamine and 15% Foetal Calf Serum. The flasks were made 5% CO$_2$ by gassing and incubated in a hot room at 37°C.

Transformed and wild type MEL tk- cells were trypsinised and replated at $10^5$ cells/ml in a 75 cm$^2$ flask in 20 ml of SLM/HAT medium. A filter sterile 500 mM stock solution of HMBA (a gift from Boyd Hardesty) was prepared by heat dissolving (37°C) 1 g of powdered HMBA in 10 ml of serum free SLM. This was added to the reseeded cells at a final concentration of 3-5 mM. A red pellet indicative of a successful induction was obtained after a 3-4 day incubation period.
2:4:2 Calcium Phosphate Co-Precipitation of Donor DNA

Donor DNA was introduced into the previously described cell lines for both long and short term transformations by the $\text{Ca}_2^+\text{PO}_4^{3-}$ coprecipitation technique of Graham and van der Eb [195] as modified by Wigler et al [247]. Donor DNA and high MW salmon sperm DNA (if required) at 40 $\mu$g/ml in T.E. was mixed with CaCl$_2$ at a final concentration of 250 mM. This was gently added to an equal volume of 2 x HBS (Hepes Buffered Saline) ($1 \times \text{HBS} = 140$ mM NaCl/25 mM HEPES (Sigma) (N-2-Hydroxyethylpiper-azine-N-2-ethanesulphonic Acid)/0.75 mM Na$_2$HPO$_4$ (disodium hydrogen phosphate) pH 7.1) giving a final DNA optimum concentration of 20 $\mu$g/ml which co-precipitates in the presence of 125 mM CaCl$_2$. After this gradual addition, the mixture was immediately vortexed and left at room temperature for thirty minutes after which period a fine precipitate forms.

2:4:3 Long Term Transformation

Exponentially growing cells were harvested by trypsinisation, counted and reseeded at a density appropriate to that cell line. For a 25 cm$^2$ growth area the following densities were adopted; BHKtk$^- 0.5 \times 10^6$, mouse LAtk$^- 1 \times 10^6$, MELtk$^- 1 \times 10^6$, human 143tk$^- 0.3 \times 10^6$, K562 $1 \times 10^6$ and EJ $1 \times 10^6$. The cells were incubated overnight at 37°C and replaced with fresh medium the following day. After co-precipitation of the DNA, the
mixture in a volume of 500 µl was slowly added to the recipient cell line containing 5 ml SLM. The flasks were incubated overnight to allow absorption of the DNA-Ca\textsuperscript{2+/2-}PO\textsubscript{4}\textsuperscript{2-} precipitate. The following day the medium was replaced with pre-warmed medium and incubated for a further 24 hours to allow the cells to recover.

2:4:3(I) Selectable Biochemical Markers

Two selectable biochemical markers were employed throughout the study, viz HSV-1 tk [221] and Tn 5 aph (II) [176]. Thymidine kinase (tk) catalyses the conversion of thymidine to TMP (thymidine monophosphate) in the salvage pathway as depicted in Fig. (6). During de novo synthesis orotic acid is converted to dUMP. Thymidylate synthetase catalyses the production of TMP from this nucleotide precursor. dTMP production in the de novo pathway is blocked when the cells are cultured in medium containing HAT [158]. Aminopterin inhibits carbon metabolism by inhibiting DHFA (Dihydrofolic Acid) to THFA (Tetrahydrofolic Acid). Hypoxanthine serves as a substrate for the salvage pathway. Under these conditions both pathways are perturbed, the salvage by mutation and the de novo by inhibition. For a cell to survive, exogenous tk DNA must be present. This selection system has two potential drawbacks. Firstly, TMP levels are controlled to an extent by feedback inhibition. Secondly, TMP being a relatively small molecule can easily pass from cell to cell.
De novo
Orotic acid
\[ \downarrow \]
dUMP

\text{thymidylate synthetase}

\text{blocked by aminopterin}

\text{Salvage}
TdR

\[ \downarrow \]
thymidine kinase

THFA

\[ \downarrow \]
DHFA

\text{DHFA reductase}

\text{dTMP}

\[ \downarrow \]
dTTP

\[ \downarrow \]
DNA

\text{Figure 6}

Metabolic synthesis of dTMP via the Salvage and de novo pathways. TdR = Thymidine, DUMP = Deoxyuridine Mono- phosphate, dTTP = Thymidine Triphosphate, DHFA = Dihydro- folic Acid, THFA = Tetrahydrofolic Acid. Kindly reproduced from [232].
cell via gap junctions [248] which can lead to the formation of secondary colonies. Cells grown under HAT selection were replaced with fresh HAT/SLM medium every three days. Colony formation varied with the recipient cell line. BHK tk− and human 143 tk− cells required about twelve days, MEL tk− took about fourteen days while the mouse LA tk− line required eighteen to twenty one days.

The bacterial transposon Tn5 encodes for the aminoglycoside 3′- phosphotransferase gene aph(II) which phosphorylates aminoglycosidic residues [176]. This enzyme confers resistance to bacteria grown on the presence of the aminoglycosidic antibiotics such as kanamycin and neomycin. Geneticin, commonly referred to as G418 (BRL-Gibco) will also kill animal cells [176] so serving as a dominant selectable marker which in theory can transform any cell line. Unlike the HSV-1 tk selection system, Tn5 aph(II) activity is not subject to feed-back inhibition, neither should its protein product cross feed non transformed cells. Geneticin was stored at -20°C at a stock concentration of 20 mg/ml. The amount required to kill cultured cells was empirically determined. Effective killing concentrations were: mouse LA tk− 200 - 400 µg/ml, BHK tk− 400- 600 µg/ml, MEL tk− 800- 1000 µg/ml, human 143 tk− 400 - 500 µg/ml, K562 800-1000 µg/ml, EJ 600-800 µg/ml. The variation arose with changes in the commercial batches since geneticin is not available as a pure compound.


2:4:3(II) Selection in 0.9% Methocel Medium

0.9% Methocel SF12/HAT medium was prepared by dissolving 3 g of Methocel (Fluka) in 200 ml dist H$_2$O and sterilised by autoclaving giving a clear solution after storage at 4°C for twenty four hours. This was heated to 37°C prior to the addition of 22 ml 10 x Hams SF12, 4.0 ml 50 x Essential Amino Acids, 4 ml 0.1 M sodium pyruvate, 2.5 ml 200 mM Glutamine, 5.0 ml 7.5% sodium bicarbonate 10,000 units of Penicillin and 100 ml Foetal Calf Serum giving a final serum concentration of 30%.

Methocel medium has the advantage over selection in liquid medium in that the medium is not changed during selection which decreases the likelihood of secondary colony formation. Colonies can be directly picked by a Pasteur or pipette man.

2:4:3(III) Selection in Liquid Medium

After co-precipitation of the DNA, recovery and selection, the selection procedure was continued until the antibiotic resistant colonies were clearly visible. The medium was removed and the cells washed with P.E. The top of the flask was removed by incision with a heated scalpel. The marked colony was enclosed in a stainless steel cloning ring of 0.8 cm diameter. The cells were removed after the addition of 250 µl 0.25% typticin solution and transferred into a 25 cm$^2$ flasking containing
selection medium.

\textbf{2:4:3(IV) Quantitation of Colony Numbers}

As discussed in 3:3 of the RESULTS, globin genomic fragments and a variety of promoters were relatively assessed by their abilities to drive the expression of the selectable marker genes HSV-1 tk and Tn5 aph (II). Upon clear visualisation of the antibiotic resistant colonies the flasks were drained, washed with PBS to remove any remaining medium and fixed for thirty minutes at room temperature in cold methanol. The colonies were stained in 10\% Giemsa Stain (prepared in dist H$_2$O) for thirty minutes. The stain was removed and the flasks rinsed in tap water, drained and dried in the hot room. Colony numbers were expressed in terms of transformation efficiency as the No. of antibiotic resistant colonies $10^6$ cells$^{-1}$, $\mu$g$^{-1}$ DNA, or as a transformation ratio relative to a "control" recombinant which assumes a ratio of 1.000.

\textbf{2:4:4 Short Term Transformation}

Short term transformation experiments were carried out using the maximum amount of donor DNA which for a 75 cm$^2$ flask containing 20 ml of medium was 40 $\mu$g. After co-precipitation of 40 $\mu$g of donor DNA, the 2 ml solution at 20 $\mu$g/ml was slowly added to a $\frac{1}{4}$ confluent exponential growing recipient cells seeded twenty four hours earlier.
The following day the medium was replaced and the cells left for a further forty eight hours to transiently express the foreign sequences. After this period the cells were harvested by trypsinisation and resuspended in cold PBS. Upon centrifugation the pellet was resuspended in a guanidinium based buffer and stored at -20°C awaiting RNA preparation.

2.5 The Detection and Conformation of Donor DNA in Transformed Cell Lines

2.5.1 Preparation of High Molecular Weight DNA

High MW DNA from cultured animal cells was prepared using a modification of the method initially described by Gross-Bellard et al [249]. This protocol outlines the method used to prepare high MW DNA from 1-3 x 10^7 cells on a growth area of 75 cm^2. The medium was aspirated and the cells washed three times with cold PBS. This was followed by the addition of 10 ml {10 mM tris-HCl pH 8.0/10 mM EDTA/10 mM NaCl/0.5% SDS} and left for twenty minutes at room temperature after which the viscous solution was transferred to a 25 ml plastic universal container. Proteinase K made up in the above buffer was added to give a final concentration of 100 μg/ml and gently agitated at 37°C overnight.

The lysate was transferred to a 50 ml plastic conical centrifuge tube and extracted sequentially with tris equilibrated phenol, phenol : CHCl₃ and CHCl₃. The DNA was spooled and resuspended in 10 ml T.E. DNase I free
heat treated RNase was added to give a working concentration 50 μg/ml. The solution was incubated at 37°C for four hours. Thereupon it was made 0.5% SDS and digested overnight with proteinase K as described above. The solution was extracted as outlined above. The DNA was spooled, washed in 70% ethanol, redissolved in T.E., extensively dialysed against T.E. and re-precipitated in ethanol.

2:5:2 Southern Blotting Analysis

The transfer of DNA from agarose gels to nitrocellulose paper was carried out according to the method of Southern [250]. 10-20 μg samples of high MW genomic DNA prepared from cultured animal cells was digested with the appropriate restriction endonuclease under conditions of excess enzyme over a three to six hour period. In addition, plasmid DNA was also digested to provide a full length reference profile. The samples were electrophoresed on a 1% agarose gel in {1 x 50 mM tris-HCl/20 mM NaAc /1 mM EDTA pH 7.8} until the marker dye was off the bottom of the gel. All manipulations on the gel were carried out at room temperatures with gentle agitation. The gel was depurinated in 0.25 N HCl for fifteen minutes to facilitate the subsequent transfer of the higher MW fraction of DNA.

This step was repeated. The gel was briefly washed in dist H2O to remove the acid. This was followed by two
fifteen minute washes in denaturation buffer \(0.5\text{M NaOH/1 M NaCl}\). The gel was rinsed in dist \(H_2O\) and neutralised by two thirty minute periods in neutralisation buffer \(0.5 \text{ M tris-HCl pH 7.4/3 M NaCl}\).

The gel was placed onto a glass plate covered with two pieces of 20 XSSC saturated Whatman 3 MM paper. Saran wrap or aluminium foil was placed around this area to maximise transfer from the gel to the nitrocellulose and to ensure the 3 MM paper did not dry out. The nitrocellulose filter (Sartorius) presoaked in 2 XSSC was carefully placed onto the treated gel taking care no air bubbles were trapped beneath. Two dry sheets of Whatman 3 MM sheets were placed on the filter in addition to a six inch block of dry paper towels or tissues. This was weighted down and transfer occurred over a 4 - 6 hour period. The filter was gently teased away from the agarose gel, air dried and baked for 2 hours at 80°C. Prehybridisation, hybridisation and washing conditions were as detailed in (2:2:4).

2:6 The Transcription of Donor DNA in Transformed Cell Lines

2:6:1 Preparation of Total Cellular RNA

Two methods were employed to prepare total cellular RNA from cultured animal cells.
2:6:1(I) Caesium Chloride Gradient Method

A cell pellet from a confluent 75 cm$^2$ flask containing $1 - 3 \times 10^7$ cells was resuspended in 2.5 ml {5 M Guanidinium Thiocyanate (Fluka)/50 mM tris - HCl pH 7.0/50 mM EDTA/5% (v/v) βME}, vortexed and passed through a Dounce homogeniser. Alternatively it was stored at -20°C. The solution was made 2% sacrosyl and heated to 50°C for ≥2 minutes. This was loaded onto a 1.2 ml cushion of {5.7M CsCl/50 mM EDTA pH 7.0 - refractive index 1.3925 - 1.4025} in a 6 x 4.2 ml Diamon I.E.C. tube and centrifuged (35,000 rpm/15°C/24 hours). The supernatent containing the DNA was carefully removed and the RNA pellet resuspended in 250 μl sterile dist H$_2$O and mixed with an equal volume of {4 M LiCl (Lithium Chloride)/8 M urea} and left overnight at 4°C. This was microfuged for 30 minutes at 4°C. The supernatent containing any contaminating DNA was discarded and the RNA pellet resuspended in T.E., ethanol precipitated, washed in 70% ethanol, dried, redissolved in distilled water and lyophilised.

2:6:1(II) Lithium Chloride Precipitation Method

This technique is based on the method of Chirgwin et al. [251] modified by Karrin (Pers Comm). A packed cell volume of 100 μl was resuspended 800 μl {5 M GuSCN/50 mM tris-HCl pH 7.0/10 mM EDTA} and 80 μl β ME, stored at -20°C or homogenised by passage through a 19" gauge needle.
and left at room temperature for twenty minutes until the lysate was completely dissolved. The solution was transferred to a 15 ml corex tube, mixed with 5 ml 4M LiCl and left overnight at 4°C. The solution was centrifuged (6,500 rpm/90 minutes/4°C). The supernatent containing the DNA was carefully removed and the RNA pellet washed with 3 ml 3M LiCl by passage through a 19" gauge syringe and centrifuged (6,500 rpm/30 minutes/4°C). The pellet was redissolved in 400 µl of T.E. and extracted initially with 0.5 volume of phenol : CHCl₃ followed by extraction with an equal volume of phenol : CHCl₃, ethanol precipitated, washed in 70% ethanol, dried, resuspended in dist H₂O and lyophilised.

2.6.2 Preparation of Poly A⁺ RNA

Poly A⁺ RNA was prepared according to a modification of the method of Aviv and Leder [252] by two cycles of oligo dT cellulose chromatography.

0.3 g of powdered oligo dT cellulose (BRL-Gibco) was placed in a Biorad plastic column and equilibrated with 5 x 1 ml aliquots of 1 x Binding buffer {0.5M NaCl/10 mM tris - HCl pH 7.5/0.5% SDS/1 mM EDTA}. Lyophilised total cellular RNA was dissolved in sterile dist H₂O by heating to 65°C for five minutes. An equal volume of 2 x Binding buffer was added and the sample applied to the column. The column was washed with 6 x 1 ml aliquots of 1 x Binding buffer, these eluted fractions which contain
mainly ribosomal, tRNA and Poly A- RNA were discarded. The Poly A+ population bound to the column was eluted by the sequential addition of 6 x 1 ml aliquots of Elution buffer (10 mM tris-HCl pH 7.5/0.05% SDS/1 mM EDTA). Sample numbers 2 to 6 were pooled. The column was re-equilibrated with 5 ml 2 x Binding buffer. The Poly A+ solution was loaded followed by two series of 5 x 1 ml aliquot washes of 1 x Binding buffer and 1 x Binding buffer - SDS. The Poly A+ RNA was eluted by 6 x 1 ml aliquots of sterile dist H2O. The fractions were read at OD260. In most instances, numbers 2, 3 and 4 contained over 95% of the absorbing material. These samples were pooled, ethanol precipitated, washed in 70% ethanol, resuspended in sterile dist H2O and lyophilised. Yields varied from 1 to 4% of the total population. The column was sequentially washed with 5 ml volumes of 1 x Binding buffer, 0.1 N NaOH, sterile dist H2O and ethanol. The column was stored at -20°C.

2:6:3 RNA Dot Blotting Analysis

RNA dot blotting was essentially performed as described in Spandidos et al [253]. Lyophilised RNA was redissolved in dist H2O at a concentration of 4 mg/ml. 4 µl aliquots were applied to 1 cm² areas on nitrocellulose filters (Sartorius) which had been previously pre-treated with 20 X SSC and air dried. The filter was baked at 80°C for 3 hours. Prehybridisation, hybridisation and washing
conditions were as previously described (2:2:4).

2:6:4 **Northern Blotting Analysis**

As detailed in Spandidos and Paul [172] 20 µg aliquots of lyophilised RNA was taken up in {50% deionised formamide/2.2M formaldehyde/1 X MOPS buffer (40 mM morpholinopropanesulphonic acid, 10 mM NaAc, 1 mM EDTA pH 8.0)} and heated at 55°C for 15 minutes. This was made 10% glycerol and loaded onto a 1% agarose gel in {1 X MOPS/2.2 M formaldehyde}. After electrophoresis at ~3V/cm for ~4 hours the gel was sequentially soaked in 1 X MOPS for 2 x 15 minute periods. This was followed by two 30 minute periods in 20 X SSC. The RNA was blotted onto a nitrocellulose filter in a similar fashion as detailed in Southern transfer (2:5:2) except for the duration of the transfer which was around 4 to 6 hours. The filter was gently removed from the agarose, washed in 3 X SSC, air dried and baked at 80°C for 3 hours. Prehybridisation, hybridisation and washing conditions were as detailed elsewhere (2:2:4)

2:6:5 **S1 Nuclease Mapping**

S1 Nuclease mapping studies were carried out according to the method of Berk and Sharp [254] as modified by
Weaver and Weissmann [255]. The DNA probe was prepared by dephosphorylating with Bacterial Alkaline Phosphatase (BAP) and end-labelling with $^{32}$PydATP in the presence of T4 Polynucleotide Kinase (T4PK) as detailed in [256].

2:6:5(I) Preparation of 5' End Labelled Single Stranded Probe

The two probes used in the S1 Nuclease mapping studies were prepared from the recombinant pMX [172]. A 371 bp MboII fragment was isolated from a 6% neutral polyacrylamide gel following digestion of a gel isolated 1080 bp XbaI/BglII fragment. A 652 bp's BamHI fragment was prepared by gel isolation after a Bam HI pMX digest.

Gel isolated DNA as a dried ethanol precipitate (0.5 - 2 μg) was resuspended in 50 μl of 50 mM tris-HCl pH 8.8. 0.05 - 0.2 units of BAP were added and the mixture incubated at 65°C for thirty minutes. The reaction was stopped by making the solution 1 mM EDTA. The mixture was extracted twice with an equal volume of phenol : CHCl₃, once with ether, ethanol precipitated, washed in 70% ethanol and dried.

The dephosphorylated DNA was taken up in 2.5 μl 10 mM tris-HCl pH 8.0 and was followed by the addition of 17.5 μl of Kinase Buffer 1 {20 mM tris-HCl pH 9.5/1 mM Spermidine/0.1 mM EDTA}. This was denatured at 90°C for two minutes and immediately placed on ice. 2.5 μl of Kinase Buffer 2 {500 mM tris-HCl pH 9.5/100 mM MgCl₂/50 mM...
DTT/50% glycerol (v/v) was added, followed by 20 - 80 units of T4PK (P-L Biochemicals). This mixture was transferred to 125 - 250 μCi $^{32}$PyPATP (Specific Activity > 5000 Ci/mmol) which had been dried down to a volume of 2.5 μl. The solution was incubated for thirty minutes at 37°C and stopped by the addition of 10 μl 0.1 M EDTA and ethanol precipitated in the presence of 20 μg tRNA carrier (P-L Biochemicals) for two minutes on dry ice, microfuged and dried.

The dried pellet was resuspended in 50 μl of strand separation Buffer (30% DMSO/1 x TBE/bromophenol blue and xylene cyanol marker dyes). The sample was denatured at 90°C for three to four minutes and electrophoresed for eighteen hours through a 5% neutral polyacrylamide gel at 150 V/5 mA. The separated labelled strands were visualised by autoradiography. The -ve strand was exised and isolated as previously described (2:2:7).

**2:6:5(II) Hybridisation and S1 Nuclease Treatment**

RNA prepared from transformed cells 30 - 60 μg or from K562 cells 1-5 μg and the single stranded end labelled DNA were lyophilised under conditions of probe excess. The pellets were redissolved in 10 μl of Formamide Hybridisation Buffer (80% deionised formamide/0.4 M NaCl/1 mM EDTA/40 mM Pipes (Piperazine -N, N-bis [2-ethanesulphonic acid] pH 6.4)). This was transferred to a replicoted glass capillary, heat sealed and placed in a
test tube of water. This was heated at 90°C for ten minutes and quickly transferred to the hybridisation water bath, which for these experiments was at 58.5°C and left overnight. After hybridisation the capillaries were placed into tubes containing ice cold ethanol. The glass was opened by diamond cutting and flushed out with 235 µl of {250 mM NaCl/0.03M NaAc pH 4.6/0.001 M ZnSO₄ (zinc sulphate)/20 µg/ml sonicated Calf Thymus DNA (Sigma)} and transferred to an eppendorf 1.5 ml tube containing 3,000 - 5,000 units of S1 Nuclease (Boeringher-Mannheim) and incubated at 37°C for two hours. This was followed by extraction with phenol : CHCl₃ and ethanol precipitation on dry ice for thirty minutes in the presence of 20 µg tRNA carrier. Upon microcentrifugation the pellet was washed in 70% ethanol and dried.

2:6:5(III) Denaturing Polyacrylamide Gel Electrophoresis

The samples were resuspended in 5 µl of loading buffer {80% deionised formamide/1 x TBE/bromophenol blue and xylene cyanol marker dyes}, denatured by incubation at 90°C for 3 minutes and immediately placed on ice. The S1 Nuclease resistant hybrids were sized by electrophoresis on a 6% denaturing polyacrylamide gel.

A 60 ml gel mix containing 6% acrylamide, 2% - bis - acrylamide, 8 M urea in 1 x T.B.E. was polymerised as detailed in (2:2:5). This was immediately placed between 2 x (40 x 20cm) Sylglas sealed glass plates separated by
0.5 mm spacers. Prior to loading the samples, the gel was pre-run at 42 Watts constant power (1400 V and 30 mA) for 1 hour. The samples were electrophoresed for 3 hours by which time the bromophenol blue had travelled the length of the plate. In later experiments an aluminium plate was clipped onto the front plate which gave straight bands on autoradiography. The plates were separated, the gel covered with a thin plastic bag covering and set up for autoradiography.

The size of the S1 products were estimated by comparing their migration with that of end labelled markers. Routinely used markers throughout this study included pBR322/HinfI(1631, 517, 506, 396, 344, 298, 221/220, 154 and 75 bp's); pAT153/HinfI (1631, 517, 396, 298, 221/220, 154, 145 and 75 bp's); PhiX/HinfI(1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118 and 75 bp's); pBR322/HaeIII(587, 540, 504, 458, 434, 267, 234, 213, 192, 184, 124/123, 104, 89 and 80 bp's).
CHAPTER 3: Results

3:1 Cloning of Basic Human Globin Gene Recombinants

3:1:1 Subcloning of Eco RI Globin Genomic Fragment of the Human β Globin Gene Cluster into pAT153

3:1:2 Construction of Human Epsilon Globin Genomic Recombinants Derived from the Bacteriophage Recombinant λ788ε

3:1:2I Construction of pTKHeG

3:1:2II Construction of pMX

3:1:2III Construction of pTKHδβε

3:1:3 Construction of pMX Promoter Deletion Mutants (pMXΔn)

3:1:3I Restriction Enzyme Mapping of pMXΔn

3:1:3II Sequence Analysis of pMXΔn Series

3:2 Transcriptional Analysis of Human Epsilon Globin Gene Recombinants in Transformed Animal Cells

3:2:1 The Physical Status of Globin Donor DNA in Stably Transformed Animal Cell Lines

3:2:2 S1 Nuclease Mapping of Human Epsilon Globin Gene Transcripts of pTKHeG and pMX in a Variety of Stably Transformed Cell Lines

3:2:2I Preparation of End Labelled Single Stranded S1 Hybridisation Probes

3:2:2II S1 Nuclease Mapping of the 5' Termini of Human Epsilon Globin Gene Transcripts in Non-Induced K562 Cells
3:2:2III S1 Nuclease Mapping of 5' Termini of Human Epsilon Globin Gene Transcripts in Stably Transformed
  A. BHKtk- Cells
  B. Mouse LAtk- Cells
  C. Human 143tk- Cells
  D. MELtk- Cells
3:2:2IV Splicing of Human Epsilon Globin Gene Transcripts in Stably Transformed BHKtk- Cells
3:2:3 S1 Nuclease Mapping of Human Epsilon Globin Gene Transcripts of pMX Deletion Mutants in Stably Transformed BHKtk- Cells
3:2:4I The Influence of cis Linked Human Adult $\beta$ Globin Genes on the Expression of the Human Epsilon Globin Gene in Stably Transformed Animal Cells
3:2:4II Transcription of the Human Epsilon Globin Gene in HMBA Induced Transformed MELtk- Cells
3:2:4III The Effect of Viral Regulatory Sequences on Human Epsilon Globin Gene Expression in Acutely Transformed CV-1 and COS 7 Cells
3:2:4IV The Effect of Viral Regulatory Sequences on Human Epsilon Globin Gene Transcription in Stably Transformed BHKtk and Human 143tk- Cells
3:2:4IV A. Construction of pSVOD$\varepsilon$TK and pSVED$\varepsilon$TK
B. S1 Nuclease Mapping of 5' Termini of Human Epsilon Globin Gene Transcripts of pSVODcTK and pSVEDcTK in Stably Transformed Human 143tk- and BHKtk- Cells

3:2:5 A Comparison of the S1 Nuclease Transcriptional Profiles of pMX and pHRI in Acutely Transformed Mouse LAtk- Cells

3:2:5I Construction of pMXΔH

3:2:5II Construction of pSXΔB

3:2:5III S1 Nuclease Mapping of 5' Termini of Human Epsilon Globin Gene Transcripts of pTKHcG, pMX, pSX, pMXΔ4, pSXΔB, pMXΔH and pHRI in Acutely Transformed Mouse LAtk- Cells

3:2:6 trans Activation of the Human Epsilon Globin Gene in Stably Transformed K562 Cells

3:3 The Effect of Globin Genomic Fragments on Tn5 aph (II) Expression in a Variety of Stably Transformed Animal Cell Lines

3:3:1 Background Experiments on HSV-1 tk Based Recombinants

3:3:2 Construction of a Series of Tn5 aph (II) Based Recombinants

I. Promoterless: pNV

II. HSV-1 tk Promoter: pNTV

III. Human Epsilon Globin Promoter: pNVεV

IV. Mouse β Globin Promoter: pNMβV

V. Human β Globin Promoter: pNβV
VI. Insertion of Globin Genomic Fragments Nos 10 and 11 into pNeV:

- pNeV10
- pNeV11

3:3:3 The Effect of the Promoters of HSV-1 tk, Human Epsilon Globin, Mouse β Globin and Human β Globin Genes on Tn5 aph (II) Activity in a Variety of Cell Lines

3:3:4 The Effect of Eco RI Globin Genomic Fragments '10' and '11' on Human Epsilon Globin Gene Promoter Driven Tn5 aph (II) Expression in a Range of Cell Types
The results are subdivided into three sections. (3:1) deals with the molecular cloning and description of the basic globin genomic recombinants used throughout this study. The other two sections are concerned with the in vivo expression of these genomic clones and their derivatives in transformed animal cells. (3:2) gives an account of the S1 Nuclease mapping of human epsilon globin gene specific transcripts from a variety of templates. Finally (3:3) describes the development of a functional assay system in which the rate of transcription of a number of globin DNA's is measured by the ability of these fragments to drive expression of the selectable marker genes HSV-1 tk or Tn5 aph (II).

3:1 Cloning of Basic Human Globin Gene Recombinants

3:1:1 Subcloning of Eco RI Globin Genomic Fragments of The Human β Globin Gene Cluster into pAT153

A partial HaeIII/AluI human genomic λ bacteriophage library was kindly made available by Dr T. Maniatis [22]. This was screened for genomic globin sequences with radiolabelled β and γ globin cDNA probes [226][227] using a modification of the Benton and Davis method [228] as described in the Materials and Methods (2:3:1). The library was fully screened as calculated from the Clarke and Carbon equation [229] on three separate occasions.
From this five clones designated \( \lambda M2, 3, 5, 9 \) and 10 were plaque purified. As schematically shown in Fig. (7) EcoRI digests of these five \( \lambda \) clones confirm they are identical to or are derived from the previously published clone H\( \beta G \)-2 which was isolated from the same library [22]. This is a 14.45 kb Eco RI fragment containing the adult \( \beta \) and \( \delta \) globin genes. Around this period it was learned that this gene bank was incomplete in terms of globin sequence representation.

The five Eco RI fragments of \( \lambda M3 \) were subcloned into the Eco RI site of pAT153 [257]. As detailed in [57] the strategy is depicted in Fig. (8). \( \lambda M3 \) DNA was digested to completion with Eco RI and treated with alkaline phosphatase to minimise fragment concatenation. This mixture was then ligated to Eco RI treated pAT153. In expectation of a high background due to pAT153 self-ligation, positive recombinants were identified by a Grunstein-Hogness hybridisation [258] using radiolabelled \( \lambda M3 \) DNA (data not shown). Of the thousand or so colonies picked over fifty gave a hybridisation signal. This shotgun approach was adopted due to the problems experienced at that time in obtaining ligatable fragments from agarose gels.

The Eco RI fragments spanning the embryonic and foetal regions of the human \( \beta \) globin gene cluster were subcloned from a series of overlapping \( \lambda \) clones [21] into the Eco RI site of pAT153 by Mrs G. J. Grindlay in this laboratory as depicted in Fig. (9).
Figure 7

Schematic Eco RI mapping of plaque purified lambda phage recombinants λM2, 3, 5, 9 and 10 isolated from a partial AluI/HaeIII human genomic foetal liver library in relation to the previously published globin genomic recombinant H-β-G-2 [22], a 14.45 kb fragment containing the adult δ and β globin genes.
Subcloning of Eco RI genomic fragments of λM3 isolated from the human lambda genomic library [22] into the Eco RI site of pAT153 as detailed in [57].
Figure 9

Subcloning of Eco RI genomic fragments spanning the human β globin gene cluster into pAT153. A shows the Eco RI map of > 55 kb of the β globin gene locus. The Eco RI sites are indicated by vertical bars, the size of each fragment is given in kb. Each fragment has a designated pHR (Human Recombinant) Number from pHRI0 to pHRI26. B depicts the series of overlapping lambda genomic recombinants from which the Eco RI fragments were isolated [21]. C Each Eco RI fragment was subcloned into pAT153.
Construction of Human Epsilon Globin Genomic Recombinants Derived from the Bacteriophage Recombinant λ788ε

An 8 kb human genomic fragment containing the human epsilon globin gene within a phage recombinant termed λ788ε was made available to this laboratory [25]. Fig (10) shows a restriction endonuclease enzyme map of this fragment in addition to the two pAT153 human epsilon globin gene recombinants pHR10 and pHR11.

Three recombinant plasmids derived from λ788ε were constructed. The first, pTKHɛG Fig. (11) contains the entire 8 kb genomic fragment which is convalently linked to pTK-1 [209] which encodes for a 3.5 kb Bam H1 fragment containing the HSV-1 tk gene which allows the stable transformation of tk deficient animal cell lines. This recombinant was presumed to have all the cis-acting and trans recognition sequences required for regulated expression. The second recombinant, pMX Fig. (11) derived from pTKHɛG was designed to eliminate both AluI repeat elements to test for the significance of these structures. The third chimera in this series pTKHδβɛ Fig. (12) derived from pMX and the recombinant phage λM3 was constructed to test the influence of cis-linked γ and β adult globin genes on epsilon globin gene expression. The details of each of these constructions have been published [172].
Figure 10

Structure of the human epsilon globin gene and restriction map. A shows the three basic recombinants used throughout this study and the referred restriction sites. B depicts the structural features of the human epsilon globin genes viz promoters, AluI elements, DHS's and a potential Z-DNA forming region. C gives the relative positioning of the restriction enzyme sites relative to the 5' Eco RI site of pHR10.
A schematic outline of the cloning strategy involved in the construction of pTK-1 [209] based recombinant pTKHδε containing the human ε, δ and β globin genes as detailed in [172].
3:1:2(I) Construction of pTKHeG

The 8 kb HindIII fragment containing the epsilon globin gene was cloned into pTK-1 as schematically outlined in Fig. (11). HindIII digested, alkaline phosphatase treated pTK-1 was ligated to HindIII cut λ788ε. Positive colonies were identified by HindIII digestion. The orientation of the insert with respect to the HSV-1 tk gene was determined by Bam H1 digestion.

3:1:2(II) Construction of pMX

The construction of pMX derived from pTKHeG is schematically represented in Fig. (11). The two AluI repeat elements map within a 4.715 kb XbaI fragment. pTKHeG was digested with XbaI and self-ligated at a low j/i ratio to favour circularisation. Positive clones were identified as a single 10.45 kb linear XbaI fragment.

3:1:2(III) Construction of pTKHδβε

The Eco RI human recombinant phage λM3 contains the δ and β globin genes encoded within a ~10.8 kb XbaI subfragment. As depicted in Fig. (12) XbaI digested λM3 was ligated to XbaI, alkaline phosphatase treated pMX to generate the 21.3 kb triple globin gene recombinant pTKHδβε. The orientation of XbaI insert was determined by a Bam H1 digestion.
Construction of pMX Promoter Deletion Mutants

(pMXΔn)

A series of epsilon globin gene promoter mutants were engineered as a means to identify the putative functional domains of the epsilon globin gene promoter. Reverse genetics has been successfully used for the delineation of the promoter for a variety of cloned Pol II genes [47]. The cloning strategy employed in the generation of the pMX deletion mutants is depicted in Fig. (13). The putative functional domains of the epsilon globin gene promoter map within a 197 bp 5'-Bam H1/PvuII-3' fragment. Fig. (14A) shows the tripartite structure 5'-100/CCAAT/ATA-cap-3'. Bal31 Nuclease was used to generate deletions from XbaI linearised pMX. This enzyme displays two separate forms of catalytic activity [239]. It has a highly specific single stranded endodeoxyribonuclease and exonuclease activity that catalyses the removal of small oligonucleotides or mononucleotides from each terminus of double stranded DNA. In addition Bal31 is highly specific for single stranded DNA in a manner analogous to S1 Nuclease. Despite the non synchronous nature of these functions at any given time point during Bal31 treatment up to 70% of the digested DNA molecules will have flush ends at both termini. This strategy was chosen in preference to a combined approach of Exonuclease V and S1 Nuclease. The Bal31 nuclease reduces the number of enzymatic manipulation steps. After establishing the digestion conditions empirically, pMX was
Schematic representation of the construction of a series of a series of pMX promoter deletion mutants. The human epsilon globin gene promoter is contained with a 197 bp's 5'-BamH1/PvuII -3' fragment shown as a solid green box.
Figure 14

Structural Analysis of pMXΔn series. A shows the sequence break points of pMXΔs 1, 3, 4, 5, 6, 7 and 8 in relation to the putative functional promoter domains. B depicts the size of deletion of pMXΔ1 - pMXΔ12 as determined by restriction enzyme analysis. The black bars give an estimate of the size of deletion from the Xba I site. Restriction enzyme sites shown are E = Eco RI, H = HindIII, S = SmaI, X = XbaI, B = Bam HI and P = PvuII.
incubated with the enzyme at 30°C and samples removed at different times, quenched, blunt end ligated and transformed into HB101 cells. Alkaline extracted DNA from transformed bacterial colonies were assayed by restriction endonuclease analysis for digestion of the Bam H1 site at the distal end of the promoter and retention of the PvuII site at the cap site end. Fifteen clones were selected, designated pMXA2- pMXA16. In addition a control was chosen in which the deletion mapped just beyond the PvuII site, termed pMXA1.

3:1:3(I) Restriction Enzyme Mapping of pMXAn Series

As shown in Fig. (14B) the extent of Bal31 digestion was determined by treating the deletion DNA's with SmaI, HindIII or Eco RI. From this an approximate estimation on the size of the excised DNA could be determined. The deletions ranged from 700 bp's for pMXA14 to 900 bp's for pMXA13.

3:1:3(II) Sequence Analysis of pMXAn Series

To map these promoter deletions the Maxam and Gilbert DNA sequencing technique was employed. This was undertaken in collaboration with Dr J. K. Vaas and Ms M. A. Stinson of the Beatson Institute. The sequencing strategy was to end label at the PvuII site (except for pMXA1) and sequence upstream of this until the obtained sequencing
differed from that of pMX. Fig. (14A) shows the break points between plasmid and epsilon globin gene sequences for mutants pMXA2 to pMXA8. A3, A7 and A8 have the three putative functional domains, i.e. -100, CCAAT and ATA; A4 maps at the end of 'C rich' -100 region and A5 lacks a CCAAT motif. According to size estimates coupled to limited sequencing data pMXA1 maps just beyond the PvuII site.

3:2 Transcriptional Analysis Of Human Epsilon Globin Gene Recombinants In Transformed Animal Cells

Two sets of epsilon globin gene recombinants Fig. (15) (1) pTKHeG, pMX and pTKHδβɛ and (2) the pMX deletion mutant series were introduced into (BHKtk−), mouse LAtk−, human 143 tk− and MELtk− cells by the Graham and Van Der Eb [195] calcium phosphate precipitation technique as described in detail in (2:4:2). HAT resistant colonies were picked and expanded into clonal established transformed cell lines.

The notation describing the transformed cell lines used throughout the text is:-

- BHKtk− e.g. B/pTKHeG-1,2 etc
- Mouse LA tk− e.g. LA/pMX-1,2 etc
- Human 143tk− e.g. 143/pMX-1,2 etc
- MELtk− e.g. MEL/pTKHδβɛ-1,2 etc
Figure 15

Line diagram showing the pTK-1 based epsilon globin gene recombinants pTKHeG, pMX, pMXΔn and pTKHδβε. The open boxes represent the globin genes, the black line depicts human globin flanking DNA, the solid black line shows HSV-1 tk DNA. Restriction enzyme sites shown are H = HindIII, X = XbaI, B = BamH1 and P = PvuII.
The Physical Status Of Globin Donor DNA In Stably Transformed Animal Cells

To confirm that these colonies were not tk revertants (reversion frequency is low e.g. MEL < 10^{-6}) and to investigate the physical status of the foreign DNA the following series of experiments were undertaken.

As a representative example of one of those, high MW DNA [249] prepared from mouse LAtk^-, BHKtk^- and MELtk^- cells transformed with pTKH6βε was digested to completion with Eco RI, transferred onto nitrocellulose paper by the method of Southern [250] and hybridised to nick-translated pTKH6βε. Fig. (16) shows that all transformed cell lines give a hybridisation signal except B/pTKH6βε-1 Fig. (16) lane 1). The hybridisation profile corresponds to the Eco RI digestion pattern of the donor plasmid. The higher MW bands maybe indicative of sequences which have become integrated internally into the host's DNA.

The full length copy pattern is consistent with the majority of donor plasmid sequences in stably transformed cells as suggested by Folger et al [212] to be in the form of integrated concatenates generated by a homologous recombination mechanism. It should be recognised however that these experiments [212] were performed by direct microinjection of the DNA into the host's nuclei.

An estimate of the donor DNA copy number for these transformed cell lines is presented in Table (1). This was determined by comparing the band intensities in the
Southern blotting analysis of pTKHδβεDNA from stably transformed BHKtk-, mouse LAtk- and MELtk- cells. 10 μg of High MW DNA from transformed cell lines was digested with Eco RI, electrophoresed on a 1% agarose gel, transferred to a Sartorius nitrocellulose filter by the method of Southern [250] and hybridised to nick-translated pTKHδβε. Lane 1 (B/pTKHδβε-1), lane 2 (B/pTKHδβε-2), lane 3 (LA/pTKHδβε-1), lane 4 (LAtk-), lane 5 (MEL/pTKHδβε-1), lane 6 (MEL/pTKHδβε-2), lane 7 (MELtk-), lane P represents 500 pg of Eco RI digested pTKHδβε.
Table (1)

Estimated pTKHδβε plasmid copy no. per stably transformed cell.

<table>
<thead>
<tr>
<th>Transformed Cell Line</th>
<th>Copy No.</th>
<th>Fig. (16) coord.</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-pTKHδβε-1</td>
<td>&lt;1</td>
<td>Lane 1</td>
</tr>
<tr>
<td>B-pTKHδβε-2</td>
<td>6</td>
<td>Lane 2</td>
</tr>
<tr>
<td>LA-pTKHδβε-1</td>
<td>&gt;10</td>
<td>Lane 3</td>
</tr>
<tr>
<td>MEL-pTKHδβε-1</td>
<td>5</td>
<td>Lane 5</td>
</tr>
<tr>
<td>MEL-pTKHδβε-2</td>
<td>5</td>
<td>Lane 6</td>
</tr>
</tbody>
</table>

1 - Estimates based on comparative signal intensity. Lane P (fig. (16)) contains 500 pg Eco RI digested pTKHδβε.
Southern blots of the donor plasmid relative to the pattern in the transformed cell lines. Copy number ranged from 1 to 20 per cell which is in agreement with estimates reported by others [172] [253].

This variation in copy number and the possibility of the exogenous sequences forming a mixed population of integrated concatemates, rearranged forms, episomes makes the task of interpretation of recombinant DNA expression in transformed cells difficult. Does the copy number influence the level of transcriptional activity? - does the chromosomal location effect transcription?. Such points should be borne in mind.

3:2:2 S1 Nuclease Mapping Of Human Epsilon Globin Gene Transcripts Of pTKHeG and pMX In A Variety Of Stably Transformed Cell Lines

Total cellular [251] and PolyA⁺ RNA [252] was prepared from a variety of tk⁻ cell lines transformed with the recombinant plasmids pTKHeG, pMX and pHδβε. The presence of epsilon globin gene specific transcripts was determined by the RNA dot blot hybridisation technique [253] (data not shown). In addition nuclear processing of these transcripts was confirmed by Northern blotting [179] in which Poly A⁺ transformed RNA co-migrated with the fully processed 9S message from K562 cells (data not shown).

To obtain more detail on the structure of these RNA molecules and to compare this with the transcriptional
profile of the epsilon globin gene in the K562 cell, S1 Nuclease mapping studies were undertaken. The S1 Nuclease mapping procedure is a powerful analytical technique to fine map to the nucleotide level, the termini and splice junctions of transcripts. The S1 Mapping protocol used throughout this thesis as described in the Materials and Methods (2:6:5) essentially follows the method of Berk and Sharp [254] as later modified by Weaver and Weissman [255].

3:2:2(I) Preparation of End Labelled Single Stranded S1 Hybridisation Probes

As schematically illustrated in Fig. (17) two probes were earmarked for use in the S1 Nuclease mapping experiments. Each is derived from pMX. Probe I is the 371 MboII bp fragment spanning the 5' terminus of the gene. Fig. (17) shows RNA molecules which initiate at the canonical cap site i.e. 53 and 55 bp's upstream of the first codon give two protected bands at 129 and 131 bp's. Transcripts originating from the -200 region [86] generate a S1 hybrid of 344 bp's.

Probe II is the 652 bp Bam H1 fragment covering the promoter region to the second exon. This was used to map splicing regions. As depicted in Fig. (17) correctly spliced transcripts give a S1 product of 207 bp's - this is equivalent to the distance between the Bam H1 site of the second exon to the intron/exon boundary. Non
Figure 17

Hybridisation probes used in S1 Nuclease mapping experiments. Both probes were prepared from pMX. I, pMX was digested with BglII and XbaI which liberates a 1114 bp fragment which was gel purified. This was subsequently digested with MboII and the 371 bp's fragment isolated by polyacrylamide gel electrophoresis. II, the 652 bp Bam H1 fragment was isolated by 6% polyacrylamide gel electrophoresis after digesting pMX with Bam H1. M denotes Mbo II and B shows Bam H1 sites.
processed molecules will extend beyond this. These probes were 5' end labelled as described elsewhere (2:6:5(I)) and strand separated on 5% polyacrylamide gel as shown in Fig. (18). The position of the (−ve) strand which cannot be predicted is a function of the secondary structure of the DNA. For the MboII DNA this was confirmed by Maxam and Gilbert sequencing [240] and shown to be the top band. The (−ve) strand of the Bam H1 probe was identified by a series of pilot experiments including hybridisation of K562 RNA to both strands, this approach led to the bottom fragment being designated as the (−ve) strand. In addition to the strand-separated population, variable amount of the DNA reanneals as illustrated in Fig. (18). This was routinely re-cycled to increase yields.

3:2:2(II) S1 Nuclease Mapping of 5' Termini of Human Epsilon Globin Gene Transcripts in Non-Induced K526 Cells

Human epsilon globin gene transcription in the K562 cell has been extensively investigated by Dr Allan and her colleagues [10] [86] in this laboratory using the S1 Nuclease mapping technique. It was considered important to verify her findings and establish the S1 Nuclease mapping technique since all intended S1 Nuclease mapping experiments with epsilon globin gene recombinants were to be interpreted and analysed in relation to the transcription profile of that obtained in the K562 cell.
Figure 18

Strand separation of hybridisation probes. 5' end labelled double stranded fragments were denatured in 30% DMSO and separated on a neutral 6% polyacrylamide gel. The strands were identified by autoradiography, excised and extracted overnight at 37°C in Gel Elution Buffer. The -ve strand of the MboII probe (I) is the upper band while the lower signal is the -ve strand of the Bam H1 probe (II). Non-separated double stranded (ds) DNA migrate further than the strand-separated pair for each probe.
Total cellular RNA was prepared as described in the Materials and Methods (2:6:1) from non-induced K562 cells. Using experimental conditions already well established in this laboratory and outlined in the Materials and Methods (2:6:5) the K562 RNA was hybridised to the 5' end-labelled single stranded 371 bp MboII probe (Figs. (17) and (18)) to map the 5' termini of the epsilon globin gene transcripts. The hybrids were treated with S1 Nuclease and the products separated on a 5% polyacrylamide gel and directly autoradiographed as depicted in Figs. (19) and (20).

Fig. (19) lanes 1, 2 and 3 corresponds to 10 µg of RNA with increasing amounts of probe and lanes 4, 5, 6 and 7 represent a fixed input of probe with rising amounts of RNA viz 0.1, 1, 5, and 20 µg.

The transcriptional profile described by Allan et al [86] with the MboII probe is illustrated in lane 7. The full length protected fragment can be explained by hybridisation to transcripts which originate upstream of the MboII [86] or by the probe self-annealing to the contaminating +ve DNA strand. The remainder of the signals coincide with RNA species with different 5' termini. The strongest hybridisation band is the doublet of protected fragments at 129 and 131 bp's which correspond to the heterogenous canonical cap site initiation at +53 and +55 bp's respectively as originally predicted on the basis of sequence by Baralle et al [42]. The other S1 products represent RNA species which
S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from non-induced K562 cells. RNA which was made up to 60 μg with yeast tRNA was hybridised at 58.5°C to the end-labelled (-ve) MboII strand of approximate specific activity of 200,000 cpm μg. The hybrids were treated with 3,000 units S1 Nuclease for 2 hours at 37°C and the products analysed on a 6% polyacrylamide denaturing gel. Lane P gives the position of the 371 bp's MboII probe. Lane M is the end-labelled pAT153/HinfI markers. Lane 1, 10 μg K562 RNA and 10 ng probe. Lane 2, 10 μg K562 RNA and 20 ng probe. Lane 3, 10 μg K562 RNA and 30 ng probe. Lane 4, 0.1 μg K562 RNA and 30 ng probe. Lane 5, 1 μg K562 RNA and 30 ng probe. Lane 6, 5 μg K562 RNA and 30 ng probe. Lane 7, 20 μg K562 and 30 ng probe.
Figure 20

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from non-induced K562 cells under limiting probe conditions. Experimental conditions were as outlined in the legend to Fig. (19). Lane 1, 10 μg K562 RNA and 10 ng probe. Lane 2, 10 μg K562 RNA and 5 ng probe. Lane 3, 10 μg K562 RNA and 1 ng probe. Lane 4, 10 μg K562 RNA and 0.1 ng probe. Lane 5, yeast RNA and 1 ng probe.
originate upstream from the cap site. The predominant S1 hybrid at 344 bp corresponds to the -200 promoter and represents the major non canonical cap site transcript. Its ratio to the cap site of 18:1 (see Table (1)) is in accord with the figure reported by Allan et al [86]. The minor initiation sites are faintly detectable in lane 7 Fig. (19) and coincide with a series of RNA molecules which originate between the cap site and the -200 promoter.

In Fig. (19), lanes 1, 2 and 3 give an identical signal profile to lane 7 indicating that probe excess conditions have been realised. This is particularly important since the -200 hybrid having more sites of nucleation is more readily formed than the cap site hybrid implying that in circumstances when the probe is limiting a false cap site: -200 ratio may result as illustrated in Fig. (20). Fig. (19), lanes 4 to 7 representing the titration of RNA with a fixed amount of probe show that between 1 and 5 µg of non induced K562 RNA is readily detectable.

3:2:2(III) S1 Nuclease Mapping of 5' Termini of Human Epsilon Globin Gene Transcripts in Stably Transformed: A BHKtk- Cells, B Mouse LAtk- Cells, C Human 143 tk- Cells and D MELtk- Cells

The overall aim of these studies was to compare
the transcriptional profile of the epsilon globin gene in transformed cells with the previously described pattern in the K562 cell. Such an approach would possibly lead to the identification of putative regulatory sequences involved in the expression of the epsilon globin gene. This strategy could be divided into a series of separate questions.

(1) Is there any significant difference in the cap: -200 ratio in the recombinants compared to epsilon globin gene activity in the K562 cell?

(2) Is there variation in the activity of the 8 kb genomic plasmid pTKHeG compared to pMX from which the two AluI repeat elements have been excised?

(3) Can functional promotor domains be identified by comparing the expression of the deletion mutant series pMXΔn?

(4) Are these transcripts processed correctly?

To tackle some of these points a series of S1 Nuclease mapping experiments were performed on RNA prepared from various transformed cell lines.

For this series of S1 Nuclease mapping experiments the following conditions were employed. 60 μg of total cellular RNA was hybridised to the single stranded MboII
probe, the hybrids were treated with S1 Nuclease as detailed elsewhere (2:6:511).

A. BHKtk- Cells

Fig. (21) illustrates the activity of two transformed BHKtk- cell lines viz B/pTKHεG-1 and B/pMX-1 in comparison with epsilon globin gene expression in non-induced K562 cells. Lanes 1, 2 and 3 depict 1, 10 and 20 μg of K562 RNA with the same amount of MboII probe. The cap:-200 ratio at 1 μg suggests probe excess conditions have been attained. Lanes 4 and 5 show B/pTKHεG-1 at increasing probe amounts and the same number of counts are represented in Lanes 6 and 7 for B/pMX-1. A striking feature of this experiment is the marked difference in the cap:-200 ratio between the K562 cell and the BHKtk- transformed cell. To quantify this, the appropriate regions of the gel corresponding to the transcription signals were excised and Cherenkov counted. The data are presented in Table (2). The value for 1 μg K562 RNA (Fig. (21), lane 7) corresponds to the published findings of Allan et al [86]. In contrast, the transformants gave a cap:-200 ratio of around 1:1. However it should be emphasised that these ratios are only approximate as it is difficult to accurately excise these bands, especially the -200 signal due to its proximity to the full length probe signal.
Figure 21

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-cells stably transformed with pTKHecG and pMX. Experimental conditions were as outlined in the legend to Fig. (19). Lane 1, 1 μg K562 RNA and 30 ng of probe. Lane 2, 10 μg K562 RNA and 30 ng of probe. Lane 3, 20 μg K562 RNA and 30 ng of probe. Lane 4, 60 μg B/pTKHecG-1 RNA and 30 ng of probe. Lane 5, 60 μg B/pTKHecG-1 RNA and 50 ng of probe. Lane 6, 60 μg B/pMX-1 RNA and 30 ng of probe. Lane 7, 60 μg B/pMX-1 RNA and 50 ng of probe. Lane M is the end labelled pAT153/Hin fi markers.
Quantitation (as determined by Cherenkov counting of excised gel sections) of the distribution of epsilon globin gene transcriptional initiation sites in K562 cells and stably transformed BHKtk- and mouse LAtk- cells.

<table>
<thead>
<tr>
<th>Hybridisation Condtns</th>
<th>Figure Coordinate</th>
<th>Cherenkov Counts (cpm)</th>
<th>% Distribution RNA Initiation Sites</th>
<th>Cap: -200 Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RNA Initiation Sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-53       -55 -270</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-53       -55 -270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K562 Lit. Value [86]</td>
<td>-</td>
<td>-         - -</td>
<td>80</td>
<td>20:1</td>
</tr>
<tr>
<td>1µg K562/30 ng</td>
<td>Fig 21/Lne1</td>
<td>42        8 3</td>
<td>79</td>
<td>17.7:1</td>
</tr>
<tr>
<td>60µg B/pTKHeG-1/30 ng</td>
<td>Fig 21/Lne 4</td>
<td>46        12 70</td>
<td>36</td>
<td>1:1.2</td>
</tr>
<tr>
<td>60µg B/pTKHeG-1/50 ng</td>
<td>Fig 21/Lne 5</td>
<td>68        19 102</td>
<td>36</td>
<td>1:1.7</td>
</tr>
<tr>
<td>60µg B/pMX-1/30 ng</td>
<td>Fig 21/Lne 6</td>
<td>31        8 51</td>
<td>34</td>
<td>1:1.3</td>
</tr>
<tr>
<td>60µg B/pMX-1/50 ng</td>
<td>Fig 21/Lne 7</td>
<td>28        11 43</td>
<td>39</td>
<td>1:1.1</td>
</tr>
<tr>
<td>5µg K562/25 ng</td>
<td>Fig 22/Lne 1</td>
<td>682       40</td>
<td>17.05:1</td>
<td></td>
</tr>
<tr>
<td>60µg LA/pTKHeG-1/25 ng</td>
<td>Fig 22/Lne 2</td>
<td>194       48</td>
<td>4.04:1</td>
<td></td>
</tr>
<tr>
<td>60µg LA/pTKHeG-2/25 ng</td>
<td>Fig 22/Lne 3</td>
<td>138       43</td>
<td>3.21:1</td>
<td></td>
</tr>
<tr>
<td>60µg LA/pMX-1/25 ng</td>
<td>Fig 22/Lne 4</td>
<td>53        13</td>
<td>4.08:1</td>
<td></td>
</tr>
<tr>
<td>60µg LA/pMX-2/25 ng</td>
<td>Fig 22/Lne 5</td>
<td>92        21</td>
<td>4.38:1</td>
<td></td>
</tr>
</tbody>
</table>

1 - Due to the merger of cap site signals the -53 and -55 bands counted as one
2 - Lne denotes Lane.
B. Mouse LAtk\(^{-}\) Cells

Fig. (22) shows the S1 Nuclease product profile of the mouse LAtk\(^{-}\) transformed cell lines, LA/pTKHeG -1 and -2 (lanes 2 and 3) and LA/pMX -1 and 2 (lanes 4 and 5). Lane 1 shows the familiar K562 pattern which differs from the four transformants. In addition, compared to the BHKtk\(^{-}\) transformed cell lines (Fig. (21)), the mouse LAtk\(^{-}\) transformants appear to produce relatively more major cap site initiated transcripts. Again as with the BHKtk\(^{-}\) transformants there are no discernible differences in the activity of pTKHeG compared to pMX. The bands corresponding to the major and -200 promoters were excised from the autoradiograph and the signal intensity estimated by Cherenkov counting as depicted in Table (2) giving an approximate cap:- 200 ratio of 4:1 which is significantly higher than that recorded for equivalent BHKtk\(^{-}\) transformants.

C. Human 143 tk\(^{-}\) Cells

The distribution of the 5' termini of the epsilon globin gene transcripts in the 143tk\(^{-}\) cell transformed with pTKHeG and pMX is depicted in Fig. (23). Lanes 1 and 2 show 5 \(\mu\)g K562 RNA with increasing probe levels. Lanes 3, 4 and 5 correspond to the activity of the transformed cell line 143/pTKHeG-1 in the presence of increasing amounts of the MboII probe and lanes 6, 7 and 8 show
Figure 22

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from mouse LAtk-cells stably transformed with pTKHeG and pMX. Experimental conditions were as outlined in the legend to Fig. (19). Lane 1, 5 μg of K562 RNA and 25 ng probe. Lane 2, 60 μg LA/pTKHeG-1 RNA and 25 ng probe. Lane 3, 60 μg LA/ pTKHeG-2 RNA and 25 ng probe. Lane 4, 60 μg LA/pMX-1 RNA and 25 ng probe. Lane 5, 60 μg LA/pMX-2 RNA and 25 ng probe. Lane M depicts pBR322/HaeIII end labelled markers.
Figure 23

S1 Nuclease mapping of the 5' termini of epsilon globin globin mRNA prepared from human 143tk- cells stably transformed with pTKHεG and pMX. Experimental conditions were as outlined in the legend to Fig. (19). Lane 1, 5 µg K562 RNA and 25 ng probe. Lane 2, 5 µg K562 RNA and 50 ng probe. Lane 3, 60 µg 143/pTKHεG-1 RNA and 10 ng probe. Lane 4, 60 µg 143/pTKHεG-1 RNA and 20 ng probe. Lane 5, 60 µg 143/pTKHεG-1 RNA and 30 ng probe. Lane 6, 60 µg 143/pMX-1 RNA and 10 ng probe. Lane 7, 60 µg 143/pMX-1 RNA and 20 ng probe. Lane 8, 60 µg 143/ pMX-1 and 30 ng. Lane M depicts Phi X/Hin fI end labelled markers.
the same conditions for the transformed cell line 143/pMX-1. Only at the highest amounts of probe input (lanes 5 and 8) can the major cap signal be seen. The strong signal just below the band corresponding to initiation from the -200 promoter made it impossible to excise. However, by eye and with cross reference to another experiment (See Fig. (39)) the cap:-200 ratio appears to be around 1:1.

D. MELtk⁻ Cells

Total cellular RNA from two pTKHeG transformed MELtk⁻ lines (Fig. (24)) and three pMX MELtk⁻ lines (Fig. (25)) were S1 Nuclease mapped. Examination of these films shows a considerable shift towards the apparent exclusive use of the -200 promoter. This dramatic reduction in the cap:-200 ratio may be accounted for by the adult MELtk⁻ cell actively expressing an adult trans factor which leads to repression of the major promoter of the human epsilon globin gene. Alternatively this ratio may be an experimental artefact resulting from cross hybridisation of the endogenous β globin RNA with the epsilon globin MboII probe. Despite there being some homology in the 5' regions of these genes the stringent hybridisation conditions of 58.5°C/80% formamide would not favour the stable formation of such a cross-species hybrid. Even allowing for the excess of mouse β globin message over human epsilon globin in RNA, the hybrids would have to
Figure 24

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk-cells stably transformed with pTKHeG. Experimental conditions were as outlined in the legend to Fig. (19). Lane 1, 5 µg K562 RNA and 25 ng probe. Lane 2, 60 µg MEL/pTKHeG-1 RNA and 25 ng probe. Lane 3, 60 µg MEL/pTKHeG-1 RNA and 50 ng probe. Lane 4, 60 µg MEL/pTKHeG-2 RNA and 25 ng probe. Lane 5, 60 µg MEL/pTKHeG-2 RNA and 50 ng probe. Lane M shows pAT153/Hin fI end labelled markers.
Figure 25

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk-cells stably transformed with pMX. Experimental conditions were as outlined in the legend to Fig. (19). Lane P gives the position of the 371 bp's MboII probe. Lane 1, 60 \( \mu \)g MELtk- RNA and 25 ng of probe. Lane 2, 60 \( \mu \)g MEL/pMX-1 RNA and 25 ng probe. Lane 3, 60 \( \mu \)g MEL/pMX-1 RNA and 50 ng probe. Lane 4, 60 \( \mu \)g MEL/pMX-2 RNA and 25 ng probe. Lane 5, 60 \( \mu \)g MEL/pMX-2 RNA and 50 ng probe. Lane 6, 60 \( \mu \)g MEL/pMX-3 RNA and 50 ng probe. Lane M depicts Phi X/Hin fl end labelled markers.
resist the S1 Nuclease nicking activity at the mismatch sites which form "bubble structures". Moreover, as discussed in the co-induction studies (See 3:2:4(II)), and confirmed in Figs. (34) and (35), RNA from both induced and non-induced non-transformed MELtk^- cells fail to give any S1 hybrids.

The results from these series of S1 Nuclease mapping experiments are presented in summary in Table (3).

3:2:2(IV) Splicing of Human Epsilon Globin Gene Transcripts in Stably Transformed BHKtk^- Cells

The Bam H1 652 bp's probe shown in Fig. (17) was employed to investigate splicing of the exon regions of the epsilon globin RNA transcripts. Spliced RNA molecules will give a protected fragment of 207 bp's which corresponds to the distance between the Bam H1 site in the second exon to the exon/intron boundary.

Fig. (26) is a representative example of this processing and shows the S1 product profile of BHKtk^- cells transformed with pMX (lane 1) and pTKHeG (lane 2). Lane 3 confirms that the signals from the transformed RNA co-migrate with that obtained with K562 RNA giving a signal less than the 220 bp's marker.

These experiments suggest that the highly conserved splice acceptor/donor site of the epsilon globin gene is recognised by the splicing enzymatic machinery of the BHKtk^- cell to yield a correctly spliced transcript. This
Table (3)

Summary of S1 Nuclease mapping data showing the distribution of 5' termini of the epsilon globin gene transcripts in K562 and stably transformed cells.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Approximate cap: -200 ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>K562</td>
<td>20:1</td>
</tr>
<tr>
<td>K562</td>
<td>18:1</td>
</tr>
<tr>
<td>Transformed MELtk-</td>
<td>0:1</td>
</tr>
<tr>
<td>Transformed 143tk-</td>
<td>1:1</td>
</tr>
<tr>
<td>Transformed BHKtk-</td>
<td>1:1</td>
</tr>
<tr>
<td>Transformed Mouse LAtk-</td>
<td>4:1</td>
</tr>
</tbody>
</table>

1 - Literature value [86]

2 - Values do not significantly vary between pTKHerG and pMX
Figure 26

Splicing of epsilon globin gene transcripts in stably transformed BHKtk- cells. Total cellular RNA prepared from non-induced K562 cells and BHKtk- cells stably transformed with pTKHeG and pMX was hybridised at 58.5°C to the end-labelled (-ve) Bam H1 strand of approximate specific activity of 400,000 cpm μg⁻¹. The hybrids were treated with 3,000 units of S1 Nuclease for 2 hours at 37°C and the products analysed on a 6% polyacrylamide denaturing gel. Lane P depicts the 652 bp's Bam H1 probe. Lane 1, 60 μg B/pTKHeG-1 RNA and 50 ng probe. Lane 2, 60 μg B/pMX-1 and 50 ng probe. Lane 3, 5 μg K562 RNA and 50 ng probe.
is consistent with the processing of other Pol II genes in foreign cells such as the rabbit β globin in mouse Ltk- cells [259].

3:2:3 S1 Nuclease Mapping of Human Epsilon Globin Gene Transcripts of pMX Deletion Mutants in Stably Transformed BHKtk- Cells

Initially three deletion mutants were examined. pMXΔ4 retains the entire prototypic β globin gene tripartite promoter structure and pMXΔ5 which lacks the CCAAT motif (See Fig. (14A)). To determine the 5' termini of these templates total cellular RNA's from these BHKtk- transformed cell lines in conjunction with B/pTKHeG-1 and B/pMX-1 were hybridised to the 371 bp's MboII fragment. The range of S1 products are represented in Fig. (27). Lanes 1 and 2 show the previously described 1:1 cap:- 200 ratio for pTKHeG and pMX compared to 20:1 for K562 (lane 7). No protected fragments are detected from B/pMXΔ1 (lane 6) and B/pMXΔ4 (lane 5). The former result is not unexpected considering the extent of deletion in the Δ1 plasmid. The failure to detect any signal in the latter is consistent with the data on the Southern analysis of B/pMXΔ4 and shows that only the 2075 bp PvuII fragment harbouring the HSV-1 gene could be detected (data not shown).

In contrast B/pMXΔ5 (lane 4) gives two protected fragments. The main band running just above the 154 bp's
Figure 27

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-cells stably transformed with pMX promoter deletion mutants Δ1, Δ4 and Δ5. Experimental conditions were as outlined in the legend to Fig. (19). Lane P shows the positioning of the 371 bp's MboII probe. Lane 1, 60 μg B/pTKHeG-1 RNA and 25 ng probe. Lane 2, 60 μg B/pMX-1 RNA and 25 ng probe. Lane 3, 60 μg BHKtk- RNA and 25 ng probe. Lane 4, 60 μg B/pMXΔ5 RNA and 25 ng probe. Lane 5, 60 μg B/pMXΔ4 RNA and 25 ng probe. Lane 6, 60 μg B/pMXΔ1 RNA and 25 ng probe. Lane 7, 5 μg K562 RNA and 25 ng probe. Lane M depicts pBR322/Hin fI end labelled markers.
marker maps to the plasmid (pAT153)/epsilon globin gene boundary. From the sequence analysis of the pMXΔ5 recombinant (Fig. (14A)) this should be 155 bp's in length corresponding to the distance between the MboII site of the first exon to the deletion breakpoint. The second signal coincides with transcripts originating from the major cap site. But this is by far the minor species.

The 5' termini of four other mutants viz Δ2, Δ3, Δ8, and Δ11 was determined in an analogous manner. As depicted in Fig. (28) all the major S1 products as was the case for Δ5 map at a position from the 5' end of the MboII probe to the deletion boundary. These deletion boundary S1 products remain in conditions of increasing probe as illustrated with the constancy of such a band at ~280 bp's from B/pMXΔ11 as shown in Fig. (29). Fig. (30) shows that transcripts from the mutant templates give a spliced S1 product which at the resolution of the polyacrylamide gel is indistinguishable from that obtained from K562 RNA.

S1 Nuclease mapping of the 5' termini of a series of deletion mutants derived from pMX in BHKtk - cells suggests that under these particular assay conditions epsilon globin gene specific transcripts preferentially initiate from the plasmid moiety of the deletion recombinant (Figs. (27) and (28)). These S1 products map between the MboII site of the first exon and the epsilon globin gene plasmid boundary as schematically illustrated in Fig. (31).
Figure 28

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-cells stably transformed with pMX promoter deletion mutants Δ2, Δ3, Δ8 and Δ11. Experimental conditions were as outlined in the legend to Fig. (19). Lane 1, 5 μg, K562 RNA and 25 ng probe. Lane 2, 60 μg B/pMXΔ2-1 RNA and 25 ng probe. Lane 3, 60 μg B/pMXΔ3-4 RNA and 25 ng probe. Lane 4, 60 μg B/pMXΔ8-2 RNA and 25 ng probe. Lane 5, 60 μg B/pMXΔ11-1 RNA and 25 ng probe. Lane M shows Phi X/Hin fl end labelled markers.
Figure 29

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-cells stably transformed with pMXΔ11 under conditions of increasing probe. Experimental conditions were as outlined in the legend to Fig. (19). Lane P shows the position of the 371 bp's MboII probe. Lanes 1, 2 and 3 contain 60 μg B/pMXΔ11 with 25, 50 and 100 ng of probe respectively. Lane M depicts Phi X/Hin fI end labelled markers.
Figure 30

Splicing of epsilon globin gene transcripts in BHKtk- cells stably transformed with a series of pMX promoter deletion mutants. Experimental conditions were as outlined in the legend to Fig. (26). Lane P shows the 652 bp's Bam H1 probe. Lane 1, 60 μg B/pMXΔ11-1 RNA and 50 ng probe. Lane 2, 60 μg pMXΔ8-2 RNA and 50 ng probe. Lane 3, 60 μg pMXΔ3-4 RNA and 50 ng probe. Lane 4, 60 μg pMXΔ2-1 RNA and 50 ng probe. Lane 5, 5μgK562 RNA and 50 ng probe. Lane M shows Phi X/Hin fl end labelled markers.
Schematic representation of the transcription of the human epsilon globin gene from pMX promoter deletion templates depicted in A showing the S1 protected products which correspond to B initiation from the canonical cap site and C initiation from the plasmid moiety.
This preferred use of plasmid promoters could arise:

(a) the plasmid promoters are relatively strong and there is no termination signal between them and the epsilon globin gene, hence transcriptional readthrough or

(b) deletion of a promoter functional domain, this could only be the case if there was such an element located between the C-rich-100 region and the Bam H1 site or

(c) a requirement for sequences upstream of the Bam H1 site. This observation is strengthened considering the activity of pMX in the BHKtk^- cell. As shown in Fig. (21), up to half of the epsilon globin gene specific transcripts, synthesised from this template originate from the major promoter region.

Accordingly this experimental approach may not be the most suitable strategy to define the functional domains of the human epsilon globin gene promoter. It may well be more rewarding to engineer a recombinant which contains the upstream element, i.e. the 350 bp's 5'-XbaI/Bam H1-3' fragment from which a series of deletions could be generated from the Bam H1 site in an undirectional fashion towards the promoter. In addition to identify the critical sequences contained within the 5'XbaI/Bam H1-3' fragment a series of deletions gradually extending from the XbaI site to the Bam H1 recognition sequence may be
To test if this differential use of promoters in the transcriptional unit of the human epsilon globin gene is of any physiological significance, a series of experiments were designed with the purpose of attempting to manipulate the cap: -200 ratio, arguably a marker for epsilon globin gene expression in a terminally differentiated cell. Three experimental strategies were undertaken with a view to this aim:

(1) Would the presence of other \( \beta \) globin genes be influential?

(2) Could the exogenous epsilon globin gene be co-induced in transformed MELtk- cells? - if so, which promoter would respond?

(3) Can expression be directly effected by the presence of a cis linked regulatory element such as the SV40 origin or replication and enhancer?
The Influence of cis Linked Human Adult β Globin Genes on the Expression of the Human Epsilon Globin Gene in Stably Transformed Animal Cells

S1 Nuclease mapping of the 5' termini of the human epsilon globin gene transcripts was performed on BHKtk−, mouse LAtk− and MELtk− cells transformed with the triple globin gene recombinant pTKHδβε Fig. (12). The autoradiograph is depicted in Fig. (32).

Lanes 2 and 3 represent the BHKtk− transformants, B/pTKHδβε-1 does not show any hybridisation which is consistent with the Southern analysis which failed to detect any donor sequences Fig. (16). B/pTKHδβε-2 displays a profile similar to that of B/pMX-1 Fig. (21). This trend is also observed in the mouse LAtk− and MELtk− cells.

These studies seemingly rule out any obvious influence of the two adult β globin gene under these particular experimental conditions.

Transcription of the Human Epsilon Globin Gene in HMBA Induced Transformed MELtk− Cells

It was intended to examine whether MELtk− cells transformed with epsilon globin gene recombinants were co-inducible in the presence of HMBA and if they were, to ascertain which of the promoters responded to the inducer. Previous studies (this thesis) had established that
Figure 32

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk-, mouse LAtk- and MELtk- cells stably transformed with pTKHδβε. Experimental conditions were as outlined in the legend to Fig. (19). Lane P shows the positioning of the 371 bp's MboII probe. Lane 1, 5 µg K562 RNA and 25 ng probe. Lane 2, 60 µg B/pTKHδβε-1 RNA and 25 ng probe. Lane 3, 60 µg B/pTKHδβε-1 RNA and 25 ng of probe. Lane 5, 60 µg LA/pTKHδβε-2 RNA and 25 ng probe. Lane 6, 60 µg MEL/pTKHδβε-1 RNA and 25 ng probe. Lane 7, 60 µg MEL/pTKHδβε-2 RNA and 25 ng probe. Lane M shows the positioning of pBR322/HinfI end-labelling markers.
epsilon globin gene specific transcripts from non-induced transformed MELtk- lines almost exclusively originate from the -200 promoter (Figs. (24) and (25)).

Transformed MELtk- cell lines were treated with 3 mM HMBA (2:4:1). Total cellular RNA was Northern blotted (2:2:4) and hybridised against nick-translated pBR322/MbG plasmid [260] which contains the mouse β globin gene as depicted in Fig. (33). All three pMX lines show a marked increase in the expression of the mouse β globin gene indicative of induction.

In addition to these three cell lines total cellular RNA from MEL/pTKHeG-1 and 2 (+) inducer were S1 mapped as shown in Figs. (34) and (35). No major cap site initiated transcripts were detected in induced transformants, moreover the -200 promoter is not responsive to the presence of HMBA.

3:2:4(III) The Effect of Viral Regulatory Sequences on Human Epsilon Globin Gene Expression in Transformed CV-1 and COS7 Cells

It has been established in some cases that the use of upstream promoters of viral genes is coupled to replication. Osbourne and Berk [261] confirmed that a series of minor transcripts which extend from a site upstream of the major cap region of the E1A gene are directly dependent on the replication of the template. Cells cultured in the presence of the DNA synthesis
Figure 33

Northern Blot of RNA prepared from three pMX stably transformed MEL tk- cell lines in the presence and absence of 3 mM HMBA. 15 µg of total cellular RNA was electrophoresed on a 1% agarose gel, transferred onto a Sartorius nitrocellulose filter as detailed in (2:6:4) and hybridised to nick-translated mouse β-globin/pBR322 DNA. (+) and (-) indicate growth in the presence and absence of 3 mM HMBA as outlined in (2:4:1). Lanes 1 and 2, MELtk-, Lanes 3 and 4, MEL/pMX-1. Lanes 5 and 6, MEL/pMX-2. Lanes 7 and 8, MEL/pMX-3.
S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk-cells stably transformed with pTKHeG grown in the presence (+) and absence (-) of 3 mM HMBA. Experimental conditions were as detailed in the legend to Fig. (19). Lanes 1 and 2, 60 μg MEL/pTKHeG-1 RNA and 25 ng probe. Lanes 3 and 4, 60 μg MEL/pTKHeG-2 RNA and 25 ng probe. Lane 5, 5 μg K562RNA and 25 ng probe. Lane M depicts the positioning of pAT153/\texttt{H}infI end labelled markers.
Figure 35

S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from MELtk-cells stably transformed with pMX grown in the presence (+) and absence (-) of 3mM HMBA. Experimental conditions were as detailed in the legend to Fig. (19). Lane P shows the positioning of the 371 bp's MboII probe. Lanes 1 and 2, 60 µg MEL/pMX-1 RNA and 25 ng probe. Lanes 3 and 4, 60 µg MEL/pMX-2 RNA and 25 ng probe. Lanes 5 and 6, 60 µg MELtk-RNA and 25 ng probe. Lanes 7 and 8, 60 µg MEL/pMX-3 and 25 ng probe. Lane M shows the pAT153/ HinfI end labelled markers.
inhibitor Arabinoside C do not produce these RNA species.

As discussed in (1:5:2) viral enhancers increase the level of expression of the nearest promoter. Enhancer elements are a common structural feature of vectors employed in transient expression experiments being necessary for the detection of correctly initiated transcripts. This observation has led to a classification of some Pol II promoters being enhancer dependent or independent eg. the human $\beta$ globin and $\alpha$ globin promoters respectively [262].

Thus the presence of an origin of replication or an enhancer could possibly influence the use of the -200 and major promoters of the human epsilon globin gene particularly in short term transfection experiments.

With this in mind a series of recombinants containing the origin of replication and enhancer of SV40 were constructed. Fig. (36) shows the "enhancer" recombinants pSVEDE-1 and -2, and the "origin" plasmids pSVODE-3 and -4. These are derived from pHR11 (Fig. (9)) which contains the 3.751 kb Eco RI epsilon globin gene fragment. The details of these constructions have been published, see Allan et al [91]. The S1 Nuclease analysis of these recombinants in CV-1 and COS-7 short term transformations were performed by Dr Allan in this laboratory, for details see Fig. (37) which has been reproduced from [91].

This series of experiments suggest:

(i) the activity of the -200 promoter is influenced by
Figure 36

Schematic outline of the construction of pSVODE-3 and 4 containing the SV40 origin of replication and pSVED-1 and 2 encoding the SV40 enhancer and origin. pHR11 containing the 3.751 kb epsilon globin fragment was digested with Eco RI and treated with alkaline phosphatase. This was separately ligated to Eco RI digested pSVOD and pSVED to yield the above recombinants. Reproduced and modified from [91] which also gives details of construction.
The effect of plasmid replication and SV40 enhancer sequences on the S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from COS 7 and CV-1 cells acutely transformed with pSVODε and pSVEDε, reproduced from [91].

Cells were transfected with 40 μg per 75 cm flask of pSVEDε-1 and -2 and pSVODε-3 and -4. Total cellular RNA was prepared 48 hours after transfection and hybridised for 16 hours to the 371 bp single-stranded MboII probe at 55°C. Hybrids were treated with 1000 Units S1 Nuclease for 90 minutes at 37°C. The products were denatured and separated on 6% polyacrylamide gel. Lane 1, 0.5 μg Poly A RNA from induced K562 cells. Lane 2, 2 μg total RNA from COS 7/pSVODε-4. Lane 3, 2 μg total RNA from COS 7/pSVEDε-3. Lane 4, 2 μg total RNA from COS 7/pSVEDε-2. Lane 5, 2 μg total RNA from COS 7/pSVODε-1. Lane 6, 20 μg total RNA from CV-1/pSVODε-4. Lane 7, 20 μg total RNA from CV-1/pSVEDε-3. Lane 8, 20 μg total RNA from CV-1/pSVEDε-2. Lane 9, 20 μg total RNA from CV-1/pSVEDε-1. Lane M show PhiX/HaeIII end-labelled markers. All hybridisations were performed under conditions of probe excess.
replication in COS-7 cells when the origin is positioned in the opposite orientation from the globin gene.

(ii) only the major promoter is responsive to the presence of a cis linked viral enhancer.

(iii) the activity of these two promoter regions can be uncoupled.

(iv) the epsilon globin gene component of pHR11 does not contain an active equivalent endogenous enhancer element when assayed under these conditions.

3:2:4(IV) The Effect of Viral Regulatory Sequences on Human Epsilon Globin Gene Transcription in Stably Transformed BHKtk- and Human 143tk- Cell Lines

As discussed in the previous section and illustrated in Fig. (37), the presence of a viral enhancer increases transcription from the canonical cap site of the epsilon globin gene fragment in short-term transfection experiments.

To assess if a similar pattern would be obtained in stably transformed cells and to compare the transcriptional profile with pMX which gives a cap: -200 ratio of around 1:1 in BHKtk- cells (Table (3)) two recombinants were constructed.
Construction of pSVODεTK and pSVEDεTK

Previous long term transformation experiments involved the use of HSV-1 tk based vectors, eg. pTKHεG, pMX and pTKHδε which allow selection of transformed colonies by growth in HAT medium. pSVODε-3 and pSVEDε-2 lack this biochemical marker. To circumvent this problem as depicted in Fig. (38), the ~7 kb HindIII fragment of pMX containing the HSV-1 tk gene was isolated by agarose gel electrophoresis and inserted at the HindIII site of pSVODε-4 and pSVEDε-1 to give pSVODεTK and pSVEDεTK respectively.

S1 Nuclease Mapping of 5' Termini of Human Epsilon Globin Gene Transcripts of pSVODεTK and pSVEDεTK in Stably Transformed BHKtk- and Human 143 tk- Cell Lines

pSVODεTK, pSVEDεTK and pMX were stably introduced into BHKtk- and human 143tk- cells. Transformed colonies were not individually selected but were pooled and expanded into a transformed cell line as a mixture. This protocol was adopted to reduce the inconsistency which can arise using clonal populations. In these circumstances differences in copy number, rearrangement of exogenous sequences and site of integration can conceivably give rise to inconsistency in the behaviour of separate cloned lines. By using a pool
Figure 38

A schematic outline of the construction of pSVODε-TK and pSVEDε-TK by insertion of the 7.2 kb HindIII fragment of pMX containing the HSV-1 tk gene into pSVODε and pSVEDε respectively.
of transformed colonies this variation can be minimised.

The S1 product profile is presented in Fig. (39), ((A) represents a 4-day and (B) an overnight exposure respectively) (no data is available for 143/pSVODeTK as this transformation was unsuccessful). The pattern of "OD" and "ED" type of activity observed in short term transfections of CV-1 cells closely resembles that obtained in stable transformants. Lanes 2 and 3 of Fig. (39) denote pSVODeTK and pSVEDeTK in BHKtk- cells. These ratio differences are more obvious in the shorter exposed autoradiograph (Fig. (39B)). Of particular interest is the similarity in the usage of the major promoter of pSVEDeTK and pMX as illustrated in Fig. (39), lanes 3 and 4 (BHKtk-) and lanes 5 and 6 (143tk- cells). Two issues arise from this observation. First it would appear that pMX possesses an endogenous non erythroid enhancer like element. Second, the epsilon globin, gene component of pMX differs from the pHR11 derived recombinants pSVEDeTK and pSVODeTK.

3:2:5 A Comparison of the S1 Nuclease Transcriptional Profiles of pMX and pHR11 in Acutely Transformed Mouse LAtk- Cells

It was decided to investigate this difference in expression between pMX and pHR11 in acutely transformed mouse LAtk- cells since this line displays a relatively high cap: -200 for pMX (Table (3)). Fig. (40) depicts the
The effect of the SV40 origin and enhancer on the S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from BHKtk- and 143tk- cells stably transformed with psVODεTK, pSVEDεTK and pMX. Cells were transfected with 40 μg of plasmid DNA per 75 cm flask. After HAT selection the transformed colonies were pooled and expanded. Hybridisation conditions were as detailed in the legend to Fig. (37). Lane 1, 0.5 μg induced K562 RNA and 25 ng probe. Lane 2, 60 μg B/pSVODεTK RNA and 25 ng probe. Lane 3, 60 μg pSVEDεTK RNA and 25 ng probe. Lane 4, 60 μg B/pMX RNA and 25 ng probe. Lane 5, 60 μg 143/pSVEDεTK RNA and 25 ng probe. Lane 6, 60 μg 143/pMX RNA and 25 ng probe. Lane M shows the positioning of Phi X/HaeIII end labelled markers. A shows a 3 day exposure, B an overnight exposure.
S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from mouse LAtk-cells stably and acutely transformed with pHR11 and pMX. Stable transformations were as detailed in the legend to Fig. (39). Acute transfections and hybridisation conditions were as outlined in Fig. (37). Lane 1, 0.5 μg Poly A⁺ K562 RNA and 25 ng probe. Lane 2, 30 μg LA/pHR11 RNA (Stable) and 25 ng probe. Lane 3, 30 μg LA/pMX RNA (Stable) and 25 ng probe. Lane 4, 30 μg LA/pHR11 RNA (Acute) and 25 ng probe. Lane 5, 30 μg LA/pMX RNA (Acute) and 25 ng probe. Lane 6, 30 μg tRNA and 25 ng probe.
S1 product profiles between these plasmids in long term as well as acutely transformed mouse LAk- cells. The difference in activity is particularly highlighted in transiently expressed cells, compare lane 4 (pHR11) and lane 5 (pMX) and the stable transformants, lane 2 (pHR11) and lane 3 (pMX).

Fig. (41) shows in annotated form the four structural differences between these two plasmids. pHRI1 contains the 1468 bp's 5'-Eco RI/XbaI-3' fragment within which is located 5' to the proximal AluI element. Does this region have a repressor type which can be overridden by the presence of a cis linked enhancer? - compare the activity of pSVODe and pSVEDe (Fig. (37)). The second structural difference is in the 3' region. pHRI1 lacks the 550 bp's 5'-Eco RI/HindIII-3' fragment. It is interesting that some evidence now exists which subscribes to the view that regulatory signals might be located in the 3' flanking region of globin genes. Recently Rohrbaugh et al [263] proposed that transcription of the rabbit β globin gene may be regulated in part by sequences in the 3' flanking region of the gene by an attenuation mechanism which is dissimilar to the procaryotic model [264]. They argue for the involvement of two structural elements. An inverted repeat which can adopt a stem-loop conformation and a short RNA transcribed from the opposite strand by RNA Polymerase II are located within a region of declining transcriptional activity. These workers propose that the stem-loop structure and the short transcript may
Figure 41

Structural differences between the plasmid recombinants pMX and pHR11. Restriction enzyme sites shown are E = Eco RI, X = XbaI, B = Bam H1, P = PvuII and H = HindIII.

(1) pHR11 contains the 1468 bp's 5'-Eco RI/XbaI-3' fragment containing the AluI repeat element.

(2) pMX encodes the 466 bp's 5'-HindIII/XbaI-3' fragment

(3) pMX has an extended 3' region in the form of a 550 bp's 5'-Eco RI/HindIII fragment.

(4) pMX is cis-linked to the HSV-1 tk gene
physically block the movement of RNA Polymerase II molecules along the template. Of further interest is the observation that the 3' region of the human β globin gene displays these two structural features.

The third difference between pMX and pHR11 is the presence of the 5'-HindIII/XbaI-3' 466 bp's fragment at the 5' end of pMX. Finally, pMX contains the 3.5 kb Bam H1 fragment of the pTK-1 recombinant which contains the HSV-1 tk gene.

**3:2:5(I) Construction of pMXΔH**

To test if the 3' flanking region of the human epsilon globin gene was responsible for the different activity between pMX and pHR11, the 550 bp's 5'-Eco RI/HindIII-3' 3' fragment was excised from pMX.

As schematically outlined in Fig. (42) the 2675 bp's Eco RI fragment containing the human epsilon globin gene was gel isolated, flush-ended and the termini converted to HindIII recognition sites by linker ligation. In parallel with this pMX DNA was digested with HindIII, treated with alkaline phosphatase and the 6786 bp's HSV-1 tk fragment isolated by low melting point agarose gel electrophoresis. These HindIII DNA fragments were covalently linked to give the pMX 3' deletion recombinant termed pMXΔH.
Figure 42

A schematic outline of the cloning strategy adopted in the construction of the pMX 3' deletion mutant pMXΔH involving the removal of the 550 bp's 5'-Eco RI/HindIII-3' fragment from pMX.
As outlined in Fig. (43) pSX and pMX differ only in the orientation of the human epsilon globin gene with respect to the HSV-1 tk gene. pSX was constructed by Dr D. A. Spandidos in this laboratory by an analogous procedure to that described in the construction of pMX (3:1:2(II)). By comparing the transcriptional activity of pMX and pSX it might be possible to determine the influence of any tk read-through effect on the expression of the human epsilon globin gene.

Another possible way in which the HSV-1 tk gene could influence the expression of the human epsilon globin gene is over the competition for factors involved in the initiation of transcription. However in pSX (Fig. (43)) the distance between the two promoter regions of the genes is over 500 bp's. Thus it would seem that pSX is not a suitable recombinant to investigate this phenomenon, a plasmid in which the promoters were close to each other however would be purposeful. Fig. (43) shows such a recombinant, pSXAB, in which the 560 bp's Bam H1 fragment spanning the intergenic regions between the two genes has been excised leaving both promoters adjacent to one another. This plasmid was constructed by performing a partial Bam H1 of pSX and isolating by low melting point agarose gel electrophoresis the deleted fragment in which the two Bam H1 sites covering this intergenic DNA had been cut. This fragment was re-sealed at a low j/í ratio to
Cloning strategy employed in deleting sequences upstream of the epsilon globin promoter in the construction of A pMXΔ4 and BpSXΔB. pMX and pSX differ in the orientation of the HSV-1 tk gene.
favour circularisation, positive colonies were selected for on the disappearance of the Bam H1 560 bp's fragment to give pSXΔB.

3:2:5(III) S1 Nuclease Mapping of 5' Termini of Human Epsilon Globin Gene Transcripts of pTKHeG, pMX, pSX, pMXΔ4, pMXΔH and pHR11 in Acutely Transformed Mouse LATk- Cells

As depicted in the line diagram Fig. (44), pTKHeG, pMX, pSX, pMXΔ4, pMXΔH and pHR11 were transiently expressed in mouse LATk- cells and the 5' termini determined by S1 Nuclease mapping as shown in Fig. (45). An inconsistent feature regarding the cap: -200 ratio for pMX has arisen. This as depicted in Fig. (40) should be around 4:1. This cannot be readily explained by failing to attain probe excess conditions since this has been achieved for 5 μg of K562 RNA. Moreover the pMX tract Fig. (45), lane 3 also displays a strong protected probe band.

Despite this, pMX and pHR11 display the same general pattern differences as obtained in Fig. (40), lanes 3 and 5. Removal of the 5'-Eco RI/HindIII-3' 550 bp's fragment from pMX does not appear to significantly alter its activity, compare Fig. (45) lanes 3 (pMX) and 5 (pMXΔH).

The influence of any HSV-1 tk gene read-through on the expression of the human epsilon globin gene seems to be minimal as judged by the similarity between pMX and pSX as
Figure 44

Line diagram illustrating the structural differences between the recombinant plasmids pHR11, pTKHfG, pMX, pMXΔ4, pMXΔH, pSX and pSXΔB. The open box represents the epsilon globin gene, the thin black line depicts epsilon globin flanking DNA, the solid black line is the HSV-1 tk gene and flanking DNA, pAT153 DNA is represented by the broken line. The epsilon globin promoter is depicted in green, the HSV-1 tk promoter is blue. Restriction enzyme sites shown are H=HindIII, X = XbaI, E = Eco RI, B = Bam HI, P = PvuII and Bg = BglII.
S1 Nuclease mapping of the 5' termini of epsilon globin mRNA prepared from mouse LAtk-cells acutely transformed with pHR11, pTKHsG, pMX, pMXΔH, pSX and pSXΔB. Acute transfections and hybridisation conditions were as outlined in the legend to Fig. (37).

Lane 1, 5 μg K562 RNA and 25 ng probe.
Lane 2, 30 μg LA/pTKHsG RNA and 25 ng probe.
Lane 3, 30 μg LA/pMX RNA and 25 ng probe.
Lane 4, 30 μg LA/pHR11 RNA and 25 ng probe.
Lane 5, 30 μg LA/pMXΔH RNA and 25 ng probe.
Lane 6, 30 μg LA/pMXΔ4 RNA and 25 ng probe.
Lane 7, 30 μg LA/pSX RNA and 25 ng probe.
Lane 8, 30 μg LA/pSXΔB RNA and 25 ng probe.
Lane 9, 30 μg yeast tRNA and 25 ng probe.

Lane M depicts the positioning of Phi X/HinfI end labelled markers.
depicted in Fig. (45) (lanes 3 and 7).

Particularly striking is the difference between the mutant templates pMXΔ4 (lane 6) and pSXΔB (lane 8). With the former, cap site initiated transcripts are present although the vast majority of the RNA population extend from the plasmid moiety as determined by the strong plasmid/epsilon globin gene boundary signal (See Fig. (31)).

However with pSXΔB in which the promoters of the HSV-1 tk and human epsilon globin gene are in an adjacent back to back arrangement very little major cap site initiated transcripts are observed. In addition as would be predicted on the basis of the orientation of the two genes there is no signal corresponding to the epsilon globin gene/plasmid boundary. This suggests that the HSV-1 tk gene promoter is a more effective competitor for transcription factors than the epsilon globin gene counterpart under these conditions. But whether this phenomenon has any bearing on the central investigation of the difference between pMX and pHRI1 is unclear. Finally the 466 bp's 5'-HindIII/Xba I-3' fragment has been inserted into the XbaI site of pSVODc-3 (Fig. (38)) by Mrs G. J. Grindlay and Dr M. Allan in this laboratory (Pers Comm) and transiently expressed in COS 7 cells. No differences were observed in the S1 Nuclease product profiles between these two plasmids.

To conclude these series of S1 Nuclease mapping experiments by a process of elimination and implication
suggest that the 1468 bp's 5'-'Eco RI/XbaI-3' fragment of pHR11 acts as inhibitor or down regulator of major cap site expression of the human epsilon globin gene under these particular experimental conditions.

3:2:6 Trans Activation of The Human Epsilon Globin Gene in Stably Transformed K562 Cells

It has been established that the 3.751 kb epsilon globin gene fragment of pHR11 is responsive to the presence of a cis linked viral enhancer both in short term [91] and long term transformed mammalian cells (3:2:4(IV)). Does this result from the absence of an endogenous enhancer from pHR11? or does it in fact contain an enhancer which is tissue specific in its expression. Alternatively, does pHR11 have a recognition sequence for trans activators which are only expressed in the appropriate erythroid tissue? To investigate some of these possibilities required the introduction of pHR11 into cultured cells which are actively expressing the epsilon globin gene, i.e. K562 cells. This required using an epsilon globin gene recombinant which could be distinguished from the active endogenous counterpart in the K562 cell. This experimental strategy was formulated in collaboration with Dr M. Allan in this laboratory, the results of which have been published [265].

The PvuII site at the 3' end of the human epsilon globin gene of pSVODε-3 (Fig. (36)) was converted to a KpnI
site by Dr M. Allan and Mrs G. J. Grindlay to give pSVKε-3. The wild type and "marked" gene could be distinguished by S1 Nuclease mapping. With the 371 bp's MboII probe the endogenous gene will give a protected fragment of 109 bp's while the "marked" genes S1 product is 138 bp's. For long term transformation experiments it was decided to link pSVKε-3 to the bacterial neomycin gene which allows the selection of transformed colonies by growth in the presence of geneticin. For a separate study (3:3:2), the neomycin gene was available in the recombinant pNV (Fig.(49)). From this the 1.7 kb fragment containing the neomycin gene was cloned into pSVKε-3 as depicted in Fig. (46) to give pSVKNε-3.

pSVKNε-3 was stably introduced into human 143 tk-, COS 7, MELtk- and K562 cells. Geneticin resistant colonies were pooled and expanded as a mixed population for reasons already alluded to (3:2:4(IVB)). The S1 Nuclease mapping of the 5' termini of the transcripts were performed by Dr M. Allan and shown in Fig. (47) reproduced from [265]. Transformed human 143 tk-, COS 7 and MELtk- cells, Fig. (46), lanes A, C and E, produce transcripts which originate from the -200 promoter, lanes B, D and F represent the non-transformed cell lines. With transformed K562 cell, lane H in addition to the endogenous cap signal at 109 bp's there is a strong signal corresponding to a RNA molecule extending from the cap site of the "marked" gene. Moreover, the cap: -200 ratio of both the endogenous and exogenous epsilon globin genes
Figure 46

Schematic outline of the cloning strategy employed in the construction of pSVKNe-3, reproduced from [265]. The promoter PvuII site of the pHR11 derived recombinant pSVODc-3 containing the SV40 origin [91] was converted to a KpnI site. The neomycin resistance gene Tn5 aph (II) driven by the HSV-1 tk promoter is encoded in the plasmid pAGO.60 [176]. The BglII site adjacent to the structural gene was converted to a HindIII site. This 3.9 kb HindIII fragment was partially digested with PvuII and the isolated 1.7 kb HindIII/PvuII fragment inserted into pSVKc-3 as shown in the above diagram.
S1 Nuclease mapping of the 5' termini of epsilon globin mRNA from 143tk-, COS 7, MELtk- and K562 cells stably transformed with the 'marked' epsilon globin gene recombinant pSVKNe-3. Cells were transfected with 40 μg of plasmid DNA per 75 cm flask. After G418 selection the transformed colonies were pooled and expanded. 20 μg of total RNA was hybridised to the 371 bp MboII probe under conditions of probe excess for 16 hours at 57°C. The hybrids were digested with 2000 units S1 Nuclease for 1.5 hours at 40°C. Lane A, 20 μg 143/pSVKNe-3 RNA. Lane B, 20 μg 143tk- RNA. Lane C, 20 μg COS 7/pSVKNe-3 RNA. Lane D, 20 μg COS 7 RNA. Lane E, 20 μg MEL/pSVKNe-3 RNA. Lane F, 20 μg MELtk- RNA. Lane G, 20 μg K562 RNA. Lane H, 20 μg K562/pSVKNe-3 RNA. Lane M shows the positioning of Phi X/HaeIII end-labelled markers. Reproduced and modified from [265].
in K562 cells is virtually identical.

These findings indicate the 3.75 kb Eco RI fragment is recognised by an embryonic erythroid trans factor(s) which by some unknown mechanism increases major cap site expression. Equivalent factors are absent in the other cell lines examined.

To conclude it would appear that the human epsilon globin gene fragment in pHR11 derived recombinants can be transcriptionally activated by two fundamentally different mechanisms, i.e. a cis linked viral enhancer and by trans activation in the K562 cell.

3:3 The Effect of Globin Genomic Fragments on Tn5aph(II) Expression in a Variety of Stably Transformed Animal Cell Lines

3:3:1 Background Experiments on HSV-1 tk Based Recombinants

The aims of (3:3) were two-fold. First, to examine the comparative strength of the human epsilon globin gene promoter in relation to a defined set of other promoters and secondly to scan the human β globin gene locus for putative regulatory elements. The experimental strategy considered appropriate for both studies was the fused gene technique. This involves testing DNA fragments to support expression of a covalently linked selectable gene. Under these conditions it can be contended that the relative
increase or decrease in the activity of the linked gene is a function of the inserted fragment.

Initially the tk gene of HSV-1 [209] was chosen as the selectable marker. The recipient cell lines were the mouse LATk- and the human 143tk-, the latter to account for the possibility of species specificity. The first task in the development of a functional assay system was the construction of HSV-1 tk based vectors which would be suitable for long term transformation experiments. As shown in Fig. (9) the sixteen Eco RI fragments spanning the entire human β globin gene cluster were cloned into pAT153.

Fig. (48) diagramatically illustrates the two single Eco RI site vectors employed to assay these Eco RI globin genomic DNAs. One contains the human epsilon globin gene promoter fused to the HSV-1 tk gene viz peV, the other pTV is promotorless. A third recombinant termed pTMo in which the HSV-1 tk gene is driven by the (MoMuSV) was employed as a further internal control.

These recombinants were stably introduced into mouse LATk- and human 143tk- cells as described in (2:4:3(III)). The transforming abilities of these plasmids is expressed as a transformation ratio, i.e. the number of HAT resistant colonies relative to pTK-1 for the promoter series of experiments and peV for the Eco RI globin genomic recombinants transfections. Two interesting phenomena emerged from these studies. First as shown in Table (4), the activity of the human epsilon globin gene promoter
pTK-1 based recombinants: pSVeTK, pεV, pTMo and pTV. A pTK-1 contains a 3.5 kb Bam H1 fragment encoding the HSV-1 tk gene in addition to 5' and 3' flanking DNA cloned into the Bam H1 site of pAT153 [209]. B pSVeTK, the tk gene is driven by the 197 bp's epsilon globin promoter and contains 420 bp's SV40 fragment within which maps the enhancer region [231]. C pεV, the tk gene is driven by the epsilon globin promoter and contains a single Eco RI site was derived from a pre-existing plasmid pεP1 kindly donated by Dr. J. Lang [232]. D pTMo, the tk gene is driven by the 180 bp's MoMuSV promoter and was cloned into the Bam H1/HindIII sites of pTV. E pTV is a promoterless plasmid derived from a pre-existing plasmid pTK-10 supplied by Dr. Lang [231].

**KEY**

- 197 bps ε GLOBIN PROMOTER
- 180 bps MoMuSV PROMOTER
- 252 bps HSV-1 tk PROMOTER
- 420 bps SV40 ENHANCER
- HSV-1 tk GENE
- pAT153 DNA
- B= Bam H1
- Bg= Bgl II
- E= Eco RI
- P= Pvu II

---

*Figure 48*
The effect of the promoters of the HSV-1 tk, human epsilon globin and MoMuSV genes on HSV-1 tk expression in stably transformed mouse LAtk- and human 143tk- cells.

<table>
<thead>
<tr>
<th>Donor DNA</th>
<th>Recipient Cell Line</th>
<th>Transformation Ratio ± SD (No of Observations)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mouse LAtk-</td>
<td>Human 143tk-</td>
</tr>
<tr>
<td>pTK</td>
<td>1.000 (4)</td>
<td>1.000 (7)</td>
</tr>
<tr>
<td>pεV</td>
<td>0.46 ± 0.22 (5)</td>
<td>0.05 ± 0.05 (7)</td>
</tr>
<tr>
<td>pTMo</td>
<td>0.57 ± 0.15 (2)</td>
<td>1.25 ± 0.28 (2)</td>
</tr>
<tr>
<td>pTV</td>
<td>0.017 ± 0.008 (5)</td>
<td>0.03 ± 0.03 (7)</td>
</tr>
</tbody>
</table>

1 - Donor DNA's, 25 ng-1 μg for mouse LAtk- cells and 2 - 10 μg for human 143tk- (the human 143tk- cells have a relatively low transformation efficiency) were mixed with high MW carrier salmon sperm DNA at a final concentration of 20 μg/ml in calcium phosphate as detailed in (2:4:2). Transformations in duplicate or triplicate were performed in 25 cm flasks in 5 ml of medium (2:4:3) under HAT selection (2:4:3I). After the appropriate selection period, for mouse LAtk- cells 14 - 21 days, human 143tk- cells 10 - 16 days, the HAT resistant colonies were stained and quantitated as described in (2:4:3IV).

2 - Each observation refers to one experiment carried out in duplicate or triplicate.
recombinant \( p\epsilon V \) is relatively low in the human \( 143tk^- \) cell. Secondly, inspection of Table (5) suggests that \( p\epsilon V10 \) increases HSV-1 tk expression significantly above the value obtained for the promoter itself, i.e. \( p\epsilon V \). The transformation ratio of 3.45 is close to that observed for the activity of human epsilon globin gene promoter driven HSV-1 tk expression when covalently linked to SV40 enhancer i.e. \( pSV\epsilon TK-2 \) (Fig. (48)). In addition this enhancing effect was maintained when fragment 10 was inserted into the MoMuSV promoter recombinant \( pTMo \) (Table (5)).

Many of the transformations and the insertion of the Eco RI fragments into \( p\epsilon V \) and \( pTV \) were performed in collaboration with Mrs G. J. Grindlay. Experimental details have not been elaborated upon since these two observations formed the preliminary basis for a more detailed study.

These two phenomena prompted an investigation into the following questions:

(1) Was the apparent repression of the human epsilon globin gene promoter in the human \( 143tk^- \) cell a property common to all other non-erythroid cell lines?

(2) Would the repression be released in the K562 cell in which the endogenous promoter is active?

(3) Would other globin gene promoters behave in a similar
Table (5)

The influence of Eco RI globin genomic fragments on epsilon globin and MoMuSV promoter driven HSV-1 tk expression in mouse LAtk- cells.

<table>
<thead>
<tr>
<th>Donor DNA</th>
<th>Transformation Ratio ± SD (No of Observations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pcV</td>
<td>1.000 (4)</td>
</tr>
<tr>
<td>pTV</td>
<td>0.017 ± 0.008 (5)</td>
</tr>
<tr>
<td>pSVεTK</td>
<td>2.015 ± 0.18 (3)</td>
</tr>
<tr>
<td>peV10</td>
<td>3.45 ± 0.125 (3)</td>
</tr>
<tr>
<td>peV11</td>
<td>1.312 ± 0.095 (4)</td>
</tr>
<tr>
<td>peV14</td>
<td>1.509 ± 0.317 (4)</td>
</tr>
<tr>
<td>peV15</td>
<td>0.879 (1)</td>
</tr>
<tr>
<td>peV17</td>
<td>1.110 ± 0.365 (4)</td>
</tr>
<tr>
<td>peV18</td>
<td>0.558 (1)</td>
</tr>
<tr>
<td>peV19</td>
<td>0.753 ± 0.204 (4)</td>
</tr>
<tr>
<td>peV20</td>
<td>1.41 ± 0.310 (2)</td>
</tr>
<tr>
<td>peV21</td>
<td>1.273 ± 0.425 (3)</td>
</tr>
<tr>
<td>peV22</td>
<td>1.719 ± 0.279 (4)</td>
</tr>
<tr>
<td>peV23</td>
<td>1.801 ± 0.095 (4)</td>
</tr>
<tr>
<td>peV25</td>
<td>1.625 ± 0.105 (3)</td>
</tr>
<tr>
<td>peV26</td>
<td>1.73 ± 0.425 (2)</td>
</tr>
<tr>
<td>pTMo</td>
<td>1.000 (1)</td>
</tr>
<tr>
<td>pTMo10</td>
<td>4.5 (1)</td>
</tr>
<tr>
<td>pTMo23</td>
<td>1.28 (1)</td>
</tr>
</tbody>
</table>

1 - Experimental details were as described in Table (4)

2 - The discrepancy in the No. of observations reflects the relative availability of the recombinants
manner in the 143tk− cell to the human epsilon globin gene promoter?

(4) The series of experiments (Table (5)) in which fragment 10 elevates expression of the HSV-1 tk gene independent of promoter source could infer that this 6.67 kb stretch of DNA located ~2 kb upstream of the human epsilon globin gene possesses some kind of non-erythroid enhancer element. To extend these studies by examining its activity in the K562 cell and in other non-erythroid cell lines it became necessary to engineer recombinants based on the geneticin selection system.

To follow these lines of inquiry required changes in the assay system. The cell lines of interest to investigate points (1), (2) and (4) were not available as tk− mutants. To circumvent this the bacterial gene neomycin aph (II) encoded on the transposon Tn5 was employed as a selectable marker [176].

3:3:2 The Construction of a Series of Tn5 aph (II) Gene Based Recombinants

The initial aim was to engineer a promoterless neomycin vector into which promoters of interest could be cloned. As illustrated in Fig. (49) pAG060 [176] contains the neomycin structural gene driven by the HSV-1 tk promoter which maps in a 252 bp's 5'-PvuII/BglII-3'
Schematic outline in the construction of the promoterless neomycin vector pNV. The recombinant plasmid pAGO.60 encodes the bacterial gene Tn5 aph (II) [176] is driven by the 252 bp's 5'-PvuII/BglII-3' HSV-1 tk promoter represented as a blue box. HSV-1 tk also provides the signals for termination and polyadenylation as a 627 bp's Smal/PvuII-3' fragment shown as a hatched box. The BglII HSV-1 tk site was converted to a HindIII site and the gel isolated 3.9 kb HindIII fragment was partially digested with PvuII to liberate the promoterless 1.7kb fragment. This was cloned into the HindIII/PvuII sites of pTV to generate pNV. The solid black line is 5' flanking HSV-1 tk DNA. The broken line is pAT153 DNA.
fragment. The signals for eucaryotic termination and polyadenylation lie within a 5'-SmaI/PvuII-3' 627 bp's tk fragment. Since the promoters of immediate interest ie. from the human epsilon globin gene and HSV-1 tk were available in existing plasmids as HindIII fragments viz. pcV (Fig. (48)) and pTK-8 [232] it was planned to construct a basic neomycin vector which had a single HindIII site at the beginning of the bacterial gene which would serve as a cloning site for promoter DNAs.

I. Construction of the promoterless vector pNV

The cloning strategy is schematically outlined in Fig. (49). pAGO60 was linearised with BglII, flush-ended by treatment with Klenow enzyme and ligated to a HindIII linker. On HindIII digestion, the tk promoter is separated from the neomycin structural gene which is now bounded by a single HindIII site. This fragment was gel purified. The internal PvuII site necessitated the isolation of the 1653 bp's fragment by partial digestion. This HindIII/PvuII fragment was covalently linked to the HindIII/PvuII digested, alkaline phosphatase treated pTV (Fig. (48)) to give the 4837 bp's vector pNV.

II and III Construction of pNTV and pNeV

The cloning strategy for the insertion of the HSV-1 tk and human epsilon globin gene promoters into pNV is
diagramatically illustrated in Fig. (50). The 197 bp's 5'-Bam H1/PvuII-3' human epsilon globin gene promoter was available as a HindIII fragment in the HSV-1 tk recombinant pERV (Fig. (48)). The 252 bp's 5' PvuII/BglII-3' promoter fragment of HSV-1 tk was likewise available as a HindIII fragment in the HSV-1 tk plasmid pTK-8 [232]. The promoter fragments were isolated by 6% polyacrylamide gel electrophoresis. These were ligated to HindIII linearised alkaline phosphatased treated pNV giving pNTV and pNcV. The orientation of the promoter with respect to the neomycin gene was determined by Eco RI digestion of pNTV and HinfI treatment of pNcV.

IV Construction of pNMBV

The murine βmaj globin gene promoter was available as a HindIII/HincII 369 bp fragment in a recombinant constructed by Dr R. S. Gilmour in this laboratory. This fragment was isolated from a 6% polyacrylamide gel as depicted in Fig. (51). This was blunt ended by treatment with Klenow enzyme (HincII already has flush ends). The termini were extended by HindIII linkers and ligated to HindIII digested, alkaline phosphatase treated pNV to give pNMBV. To determine the orientation of the insert the recombinants were cut with HaeIII.

V Construction of pNBV

This recombinant was engineered by Mrs G.J. Grindlay
A schematic outline of the construction of neomycin based vectors driven by the epsilon globin promoter A pNeV and the HSV-1 tk promoter B pNTV. A The epsilon globin promoter contained within a 197 bp's HindIII fragment was isolated from pNeV and inserted into pNV. The orientation was determined by performing a HindIII digest. BpTK-8 contains the HSV-1 tk promoter as a HindIII 252 bp's fragment was supplied by Dr. J. Lang [232]. This fragment was isolated and cloned into pNV. The orientation was confirmed by an Eco RI digest.
Figure 51

A schematic outline of the construction of the neomycin vector pNMBV driven by the mouse β globin promoter. The mouse β globin promoter contained within a 369 bp's 5'-HindIII/HincII-3' fragment was kindly supplied by Dr. R.S. Gilmour. This was converted to a HindIII fragment and inserted into pNV. The orientation was determined by a HaeIII digest.
in this laboratory. A 258 bp's 5' - Sau3A /NcoI - 3' fragment was converted to a HindIII fragment and inserted into pNV to give pNβV.

VI Insertion of Globin Genomic Fragments Nos. 10 and 11 into pNeV

In addition to cloning fragment 10 into pNeV for reasons already elaborated on, fragment 11 was included as an internal control. As outlined in Fig. (52) these Eco RI fragments were liberated from the recombinants pHR10 and pHR11 by sequentially treating the plasmids with ThalI and Eco RI (2:3:3III). The former enzyme has the recognition signal 5'-CG+CG-3', a sequence which is not found often in eucaroyctic DNA but is frequent in procaryoytes. These double digestion mixtures were ligated to Eco RI, alkaline phosphatased pNeV to give pNeV10 and pNeV11.

3:3:3 The Effect of the Promoters of HSV-1 tk, Human Epsilon Globin, Mouse β Globin and Human β Globin Genes on Tn5 aph (II) Expression in a Variety of Cell Lines

Table (6) shows the collated data of a series of mouse LA tk⁻ and human 143 tk⁻ transformation experiments with pNV, pNTV and pNeV.

The data is expressed as a mean transformation ratio (±) standard deviation relative to the HSV-1 tk promoter.
A schematic diagram of the cloning strategy employed in the construction of pNeV 10 and pNeV11. pHR10 and pHR11 were sequentially digested with ThaI and Eco RI to liberate the globin genomic fragments and inserted at the Eco RI site of pNeV. The orientation in pNeV 10 was determined by performing a HindIII digestion and pNeV 11 by a XbaI/HindIII double digest.
Table (6)

The effect of the human epsilon globin and HSV-1 tk promoters on Tn5 aph (II) expression on the mouse LAtk- and human 143tk- cells.

<table>
<thead>
<tr>
<th>Donor DNA</th>
<th>Recipient Cell Line</th>
<th>Transformation Ratio ± SD (No of Observations)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mouse LAtk-</td>
<td>Human 143tk-</td>
</tr>
<tr>
<td>pNTV</td>
<td>1.000 (4)</td>
<td>1.000 (7)</td>
</tr>
<tr>
<td>pNV</td>
<td>0.080 ± 0.035 (4)</td>
<td>0.005 ± 0.015 (7)</td>
</tr>
<tr>
<td>pNeV</td>
<td>1.940 ± 0.475 (4)</td>
<td>0.205 ± 0.090 (7)</td>
</tr>
</tbody>
</table>

1 - Experimental details were as outlined in Table (4) except for the selection conditions which were G418
The general pattern initially observed for the tk based vectors viz pTK-1, pcV and pTV (Table (4)) is maintained for the equivalent neomycin recombinants. Tables (4) and (6) confirm that the low activity of the human epsilon globin promoter in the 143tk- cell is not a feature peculiar to HAT selection. Nevertheless there is definite shift in the overall efficiency between the different selection systems. It would appear the human epsilon globin gene promoter driven expression increases around four fold in the mouse LAtk- cell under geneticin selection, the transformation ratio increases from 0.46 (Table (4)) to 1.94 (Table (6)). This trend is also observed in the 143tk- cell which increases from ~0.05 (Table (4)) to 0.2 (Table (6)).

These differences are difficult to account for. A possible explanation is that the HSV-1 tk promoter has been uncoupled from its "natural" structural gene in the pNTV recombinant and accordingly can now be regarded as heterologous as is the human epsilon globin gene promoter. Subsequently both promoters maybe compared more directly in the neomycin based vectors than the HSV-1 tk plasmids.

The previous series of experiments suggests that the epsilon globin promoter contains some structural feature which is recognised by a cellular factor which of all the lines tested is only present in the human 143tk- cell. The measurable outcome of this is a dramatic reduction in the number of cells which survive the appropriate selection conditions. The human and mouse β globin genes share a
similar promoter structure to the epsilon globin gene promoter. Subsequently it seemed reasonable to ascertain if the 143tk⁻ cell "factor" would recognise the sequence counterpart in the adult globin genes.

To test this the human and mouse adult β globin promoter recombinants pNβV and pNMβV in addition to pNV, pNTV and pNeV were assayed for neomycin expression in mouse LAtk⁻, 143tk⁻, K562, HeLa and EJ cell lines. The results for a series of transformations are presented in Table (7) No consistent pattern emerged from these studies relating to the activity of the adult globin gene promoter plasmids. The mouse recombinant pNMβV is reasonably active in the human 143tk⁻ cell line as it is in the other cell types examined. This seemingly contradicts the sequence homology model. However the behaviour of the adult β globin gene promoter plasmid pNβV is somewhat difficult to interpret. Although its activity is as low as the epsilon globin gene promoter in the 143tk⁻ cell, it is also correspondingly low on all the other cell types and maybe a weak promoter per se, of possible relevance to this is the recent finding proposing that regulatory sequences of the human and rabbit β globin transcriptional unit are located within the body of the gene [173] [174]. But this of course does not directly explain the discrepancy between the activity of these two adult β globin genes.

It is interesting that the level of epsilon globin gene promoter driven expression in the K562 cell is not
Table (7)

The effect of different promoters on Tn5 aph (II) expression in a variety of cell types

<table>
<thead>
<tr>
<th>Donor DNA</th>
<th>Mouse LAtk-</th>
<th>Human 143tk-</th>
<th>K562</th>
<th>EJ</th>
<th>HeLa</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNTV</td>
<td>1.000 (6)</td>
<td>1.000 (9)</td>
<td>1.000 (5)</td>
<td>1.000 (4)</td>
<td>1.000 (3)</td>
</tr>
<tr>
<td>pNV</td>
<td>0.04 ± 0.041 (6)</td>
<td>0.003 ± 0.002 (9)</td>
<td>0.000 (5)</td>
<td>0.024 ± 0.014 (4)</td>
<td>0.000 (3)</td>
</tr>
<tr>
<td>pNeV</td>
<td>1.539 ± 0.32 (6)</td>
<td>0.115 ± 0.12 (9)</td>
<td>0.817 ± 0.126 (5)</td>
<td>1.699 ± 0.413 (4)</td>
<td>0.756 ± 0.312 (3)</td>
</tr>
<tr>
<td>pNßV</td>
<td>0.28 ± 0.12 (3)</td>
<td>0.142 ± 0.11 (3)</td>
<td>0.000 (2)</td>
<td>0.176 ± 0.05 (2)</td>
<td>0.000 (2)</td>
</tr>
<tr>
<td>pNMBßV</td>
<td>0.774 ± 0.31 (3)</td>
<td>0.841 ± 0.43 (5)</td>
<td>0.426 ± 0.31 (3)</td>
<td>1.04 ± 0.27 (2)</td>
<td>0.622 ± 0.31 (2)</td>
</tr>
</tbody>
</table>

1 - Experimental details were as outlined in Table (4) except for the selection conditions which was G418
significantly higher than in other lines. An obvious but possibly oversimplified interpretation is that sequences upstream of the promoter are required for optimal expression. With regards to this it would be interesting to test the effect of covalently linking the 350 bp's 5'-XbaI/BamH1-3' fragment into pNeV and examining its expression in K562 cells.

3:3:4 The Effect of Eco RI Globin Genomic Fragments Nos. 10 and 11 on Human Epsilon Globin Gene Promoter Driven Tn5 aph(II) Expression in a Variety of Cell Types

The transforming abilities of the recombinants pNeV10 and pNeV11 were tested against pNeV in the following cell lines, human 143tk⁻, mouse LAtk⁻, EJ, HeLa and K562. pNTV was included as an internal control to monitor the relative behaviour of pNeV. The transformation data for a series of experiments is summarised in Table (8).

These results imply that fragment 10 elevates Tn5 aph (II) expression relatively independent of the cellular environment. This effect ranges from about a four-fold increase in mouse LAtk⁻ cell to an eight-fold increase in the K562 cell. Moreover its presence apparently releases the block on the epsilon globin promoter repression in the human 143tk⁻ cell. This effect is much more dramatic than that recorded for the tk equivalent recombinant viz pEcV10 which although higher in tk expression than pEcV it...
Table (8)

The effect of globin gene fragments Nos. '10' and '11' on epsilon globin promoter driven Tn5 aph (II) expression in a variety of cell types.

<table>
<thead>
<tr>
<th>Donor DNA</th>
<th>Human 143tk-</th>
<th>Mouse LAtk-</th>
<th>BHKtk-</th>
<th>EJ</th>
<th>HeLa</th>
<th>K562</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNeV</td>
<td>1.000 (3)</td>
<td>1.000 (3)</td>
<td>1.000 (2)</td>
<td>1.000 (2)</td>
<td>1.000 (2)</td>
<td>1.000 (3)</td>
</tr>
<tr>
<td>pNV</td>
<td>0.254±0.208 (3)</td>
<td>0.003±0.02 (3)</td>
<td>0.008±0.02 (2)</td>
<td>0.002±0.001 (2)</td>
<td>0.000 (2)</td>
<td>0.000 (3)</td>
</tr>
<tr>
<td>pNTV</td>
<td>7.99 ± 1.94 (3)</td>
<td>0.81 ± 0.24 (3)</td>
<td>0.96 ± 0.36 (2)</td>
<td>0.84 ± 0.18 (2)</td>
<td>0.80±0.2 (2)</td>
<td>1.23 ± 0.42 (3)</td>
</tr>
<tr>
<td>pNeV10</td>
<td>12.86±2.95 (3)</td>
<td>3.8 ± 0.95 (3)</td>
<td>5.19±2.1 (2)</td>
<td>4.29 ± 0.68 (2)</td>
<td>5.36±0.65 (2)</td>
<td>8.26 ± 3.1 (3)</td>
</tr>
<tr>
<td>pNeV11</td>
<td>2.77±1.03 (2)</td>
<td>1.06 (1)</td>
<td>3.3 (1)</td>
<td>1.98 (1)</td>
<td>1.82 (1)</td>
<td>1.4 ± 0.62 (3)</td>
</tr>
</tbody>
</table>

1 - Experimental details were as outlined in Table (4) except for the selection conditions which was G418.
did not elevate tk expression to the level it does under geneticin selection. This observation is difficult to account for and represents a major inconsistency in the behaviour of an identical fragment under different selection systems. Despite this a definite pattern has emerged from these transformation studies concerning the activity of the 6.535 kb epsilon globin DNA fragment in a variety of cell types suggesting that it contains a structural element which under these assay conditions displays an enhancing type of activity.
CHAPTER 4: Discussion

4:1 The Transcription Profile of the Human Epsilon Globin Gene - The Differential Use of Promoters

4:2 Putative Control Sequences

4:2:1 The Major Promoter
4:2:2 Upstream Regulatory Regions
4:2:3 trans Acting Sequences
4:2:4 Conclusion - the Spatial and Functional Overlap of Regulatory Elements

4:3 The Regulated Expression of the Human Epsilon Globin Gene - A Hypothesis Involving Both Proximal and Distal Sequences

4:4 Tissue Specific and Developmental Expression of Adult β Globin Genes - A Critical Role for Sequences Close to and/or 5' to the Cap Site

4:5 Prospects
The human epsilon globin gene provides a suitable experimental system to study the various phenomena associated with differential gene expression. Two outstanding questions arise from such an investigation. Firstly, what are the regulatory events leading to the production of a haemoglobin synthesising erythrocyte from a pluripotential haematopoietic stem cell? Secondly, what are the mechanisms which ensure that the human epsilon globin gene is only expressed in embryonic erythroid tissue?

Globin genes are primarily regulated at the level of the initiation of transcription (1:2:2). Accordingly, by studying the transcriptional pattern of the human epsilon globin gene under a variety of conditions some insight into the controls governing its regulation may be obtained.

Detailed studies by Allan et al [10] [86] [163] have shown that the human epsilon globin gene has a defined number of transcription initiation loci located upstream of the canonical cap site. Of these minor RNA species, the transcripts originating from the -200 promoter located within a region of S1 hypersensitivity [91] are the most abundant comprising of around 5% of the total epsilon globin gene specific transcript population, giving a cap: -200 ratio of 20:1. This ratio was also observed
with RNA prepared from human embryos of 5-10 weeks gestation [86]. Thus despite the neoplasticity and the mixed myeloid and erythroid phenotype, the K562 cell can be considered a suitable model for regulated epsilon globin gene expression with the cap: -200 ratio of 20:1 indicative of maximal epsilon globin gene expression.

In contrast to epsilon globin gene expression in the K562 cell, its activity in stably and acutely transformed animal cells differs markedly with a distinct tendency towards greater use of the -200 promoter. These findings suggest that the major and -200 promoters can act in an independent manner. Further evidence for this uncoupling phenomenon comes from a study by Allan et al [91]. In a number of non-erythroid cell lines epsilon globin gene specific transcripts were identified by S1 Nuclease mapping.

Interestingly, all the RNA molecules were synthesized from the -200 promoter. No epsilon globin gene specific transcripts were found in RNA prepared from different non-erythroid tissue including 3 month foetal kidney, foetal brain and foetal muscle. Cultured animal cells have a greater RNA complexity than corresponding tissue possibly reflecting a relaxation of control of gene expression for certain sequences. Such a condition may account for the above 'leaky transcription' witnessed in a number of defined cell lines. This observation could mean that the -200 promoter is not as tightly regulated as the major promoter which for maximal expression may require
the presence of an embryonic erythroid specific trans factor(s).

Why these promoters should be governed by separate controls is unknown. It may well be the case that the -200 promoter is active during the early stages of embryonic erythropoiesis where the red blood cell precursor is essentially in a 'non-erythroid' environment, a condition which transformed cell lines and even some non-transformed cells appear to mimic to varying degrees.

Further evidence for this differential use of the promoters is seen in the way the major and the -200 promoters respond to exogenous cis linked viral regulatory sequences.

As detailed in [91] when we inserted the SV40 origin of replication in the opposite direction to the human epsilon globin gene, transcription solely increased from the -200 initiation site in transiently expressed COS 7 cells. In sharp contrast only the major promoter responds to the presence of the SV40 enhancer in an orientation independent manner. This somewhat contradicts the observations of de Villiers et al [142] who reported that enhancer containing vectors in transient assays will preferentially stimulate transcription from the nearest promoter.

During the preparation of this discussion an interesting finding has been reported by Kollias et al [266] on the activity of the -200 promoter in vitro. Using an in vitro HeLa cell whole extract transcription
run-off system, they have shown that transcripts originating from the -200 promoter are insensitive to αamanitin concentrations of 100 μg/ml, suggestive of RNA Polymerase III activity. In comparison RNA molecules initiating from the canonical cap site are sensitive to 2 μg/ml αamanitin indicative of RNA Polymerase II transcription. But as reported by Allan et al [86] there is no obvious Pol III recognition site around the -200 region. However such a correlation is witnessed in the in vitro expression of the human β globin gene [165] where the Pol III transcripts are initiated from a Pol III consensus promoter.

4:2 Putative Control Sequences

What sequences of the human epsilon globin transcription unit are involved in this differential use of promoters? Does the DNase I hyper-sensitivity map constructed by Zhu et al [89] and Tuan and London [90] provide regulatory landmarks to focus attention in an effort to identify the control regions of DNA? The possible significance of these sites which correspond to the upstream promoters at -4500, -1480, -900 and -270 bp's 5' to the canonical cap site in addition to a DHS at -6500 bp's which contains a potential Z-DNA element are discussed in light of the data obtained from this study.

The role of these cis acting sequences for the purpose of distinction is broadly divided into a discussion on the
sequences which influence the activity of the major promoter as defined by the 5' - Bam H1/PvuII - 3' fragment and those putative regulatory regions upstream from this.

4:2:1 The Major Promoter

Compared to the activity of pMX, the promoter mutant templates gave low levels of transcripts originating from the major cap site.

McCafferty [267] performed S1 Nuclease mapping experiments on circular templates of pMX and a series of deletion mutants using a whole cell extract in vitro assay system. In marked contrast to the in vivo data in addition to transcripts mapping to the epsilon globin gene/plasmid boundary all recombinant templates gave S1 products corresponding to RNA species extending from the canonical cap site, these transcripts were the most abundant of the human epsilon globin gene specific population. Moreover, McCafferty [267] observed that pMXΔ4 which lacks the -100 region showed a decrease in cap site expression compared to others which contained this putative functional domain. Of further interest was the finding that pMX itself gave the lowest level of cap site initiated transcripts of all the tested DNAs. This differs from its activity in the stably transformed BHKtk- cell where over half the transcripts are synthesized from the major cap site.

These differences in the activity of the human epsilon
globin promoter in stable transformed BHKtk" cells (this thesis) and in vitro [267] may not be peculiar to this gene but may be related to the different assay conditions. For example, the above differences are also reflected in some ways on the expression of the rabbit ß globin gene in long and short term transformed cells, where the presence of the CCAAT Box provides sufficient cis sequences for optimal expression in stably transformed mouse Ltk" cells [268] while in transiently expressed HeLa [130] or mouse 3T6 cells [132] in addition to a covalently linked enhancer, all three promoter elements are required.

On the physiological suitability of the transient assay system, it is interesting to note that the human ð and ß globin genes which are expressed at a ratio of approximately 1:40 in normal adult erythrocytes are transcribed at a similar ratio when each of these genes is covalently linked to a viral enhancer in short-term transfected monkey kidney cells [262]. It is presumed that the exogenous sequences under these conditions are in an episomal form and are considered to be fully 'open' and accessible to the cellular transcription machinery. This contrasts with the integrated donor DNA which may have structural constraints imparted on it.

Despite these difficulties in comparing different assay systems and different genes the data from the pMX promoter deletion mutant experiments tend to support the idea for the existence of a putative regulatory element located upstream of the major promoter.
Possible regulatory significance of this 350 bp's 5'-XbaI/Bam H1-3' fragment has been suggested from a study by Paul et al [269] and Allan et al [270]. This DNA which contains the -200 promoter was inserted into a CAT plasmid driven by the HSV -1 immediate early promoter. CAT activity in COS 7 cells was assayed 48 hours after transfection.

The HSV-1 immediate early/CAT recombinant gave an arbitrary CAT value of 10 which increases to 200 in the presence of co-transfected E1A. These respective values were maintained when the 5'- XbaI/Bam H1-3' fragment was covalently linked to the basic construct. However, in the opposite orientation, CAT activity was essentially abolished. A possible interpretation is that the -200 promoter may compete with the major promoter for transcriptional factors, the balance of which may change throughout embryonic erythropoiesis. Alternatively this phenomenon might arise by repressor binding at the 5' end of the '5-XbaI/Bam H1-3' fragment which when in the reverse orientation is positioned adjacent to the major promoter.

A possible prediction from the pMX promoter deletion S1 Nuclease mapping studies is that the '5-Bam H1/PvuII-3' fragment would be a weak promoter in non-erythroid cells. However when the promoters of the HSV-1 tk and mouse β globin genes in addition to the epsilon globin promoter were linked to the selectable Tn5 aph (II) gene no significant changes in the level of expression between
these promoters was observed in a variety of non-erythroid cell lines.

However rather surprisingly, this was not the case in stably transformed human 143tk- cells where the activity of the human epsilon globin promoter drops by around 80% compared to its expression in all other cell lines tested. Furthermore this was not a globin related effect since the mouse β globin promoter was active in the 143tk- cell. In addition this phenomenon is not peculiar to the long term assay system. In collaboration with Dr. Gary Sibbet (Beatson Institute), the activity of the human epsilon globin promoter and the HSV-1 tk promoter were relatively assessed by their ability to drive expression of the CAT gene in transiently expressed human 143tk- and mouse LAtk- cells. The same pattern emerged (data not shown). However, this effect was relieved by the presence of the 6.535 kb Eco RI fragment of pHR10 or by the covalent linkage of the SV40 enhancer.

That only the human 143tk- cell derived from an osteogenic sarcoma-transformed cell line represents a true non-erythroid cell type although appealing seems very unlikely to be of any physiological significance. Bearing this in mind it is difficult to assess if it is worthwhile to investigate this phenomenon any further. For example it may be feasible to, at least ascertain, if the effect is a result of repressor binding by performing DNase I mapping or footprinting studies [271]. In addition competition experiments could be undertaken using the CAT system by
co-transforming in the presence of increasing amounts of pNeV.

Caution should be exercised when comparing the S1 Nuclease mapping studies with the pMX deletion mutants with the activity of different promoters driving the expression of a selectable gene. S1 Nuclease mapping gives a measure of the steady state RNA levels, while the functional assay is a quantitative measure of the expression of a selectable gene but may only be proportional to transcriptional activity over a limited defined range.

4:2:2 Upstream Regulatory Regions

The difference in activity between pHR11 and its 5' deletion mutant pMX appears to be due to the presence of the 1468 bp's 5'-Eco RI/XbaI-3' fragment of pHR11 within which maps the -900 promoter and the proximal AluI repeat element which is transcribed from the -1480 bp's promoter. This stretch of DNA seems to repress transcription from the major promoter. This down regulation effect may be due to non-erythroid repressor binding, alternatively the transcripts originating from these two upstream promoters may repress RNA synthesis from the canonical cap site by some sort of read through inhibition mechanism.

Evidence for the existence of another upstream putative regulatory element comes from the functional studies performed with the 6.535 kb Eco RI fragment of
pHR10. The ability of this genomic DNA to increase expression of two selectable genes (HSV-1 tk and Tn5 aph (II)) from the promoters of two different genes (human epsilon globin and MoMuSV) in a variety of cell types provides evidence for the existence of a sequence with enhancer like properties.

In an extension of these studies in collaboration with Mrs G.J. Grindlay we attempted to localise this 'enhancer' effect. Initially as depicted in Fig. (10) the two large subfragments of pHR10, the distal 2.822 kb 5'-Eco RI/HindIII-3' and the proximal 3.713 kb 5'-HindIII/Eco RI-3' fragments were subcloned at the Eco RI site of pNeV and assayed under standard assay conditions. Surprisingly there was no increase in Tn5 aph (II) expression with any of the large sub-fragments. Furthermore and consistent with this finding, when the subfragments of the proximal 3.713 kb 5'- HindIII/Eco RI-3' fragment were molecularly cloned into the Eco RI site of pNeV (Fig. (50)) none of the recombinants mimicked pNeV10 activity (data not shown).

Of further concern was the lack of any difference between pNeV10 in transiently expressed mouse LAtk- cells. In a quantitative S1 Nuclease mapping assay performed by Mrs G.J. Grindlay with both 5' and 3' end labelled probes, no discernible difference in signal intensity was recorded. This failure to reproduce the 'enhancing' phenomenon in acutely transformed mouse LAtk- cells is of considerable concern and contrasts with an observation of
Spandidos and Anderson [272]. These workers using the expression vector pNcV (termed pNc1 in the publication) identified an 8 kb fragment containing the human \( \lambda \) immunoglobulin enhancer which increased expression of the Tn5 aph (II) gene upto 10 fold in a tissue specific manner. Moreover this increased level of activity was observed in both long and short termed transformed cells. (It is of interest regarding the credibility of the assay system that this recombinant did not increase the activity of pNcV in mouse LAtk\(^-\) and human 143tk\(^-\) cells while pNcV10 did).

Can these inconsistencies be adequately explained? As depicted in Fig. (10) the 6.535 kb fragment '10' contains two putative regulatory elements. Located within the distal 2.822 kb 5'-Eco RI/HindIII-3' fragment is 722 bp's SstI fragment which contains a stretch of alternating purine/pyrimidine residues indicative of a potential Z-DNA forming element. In the proximal 3.713 kb 5'-Eco RI/HindIII-3' fragment maps the -4.5 kb upstream promoter which arguably corresponds to the 5' end of the transcriptional unit. This promoter is contained within the 1.48 kb 5'-XbaI/TaqI-3' fragment which also has an enhancer 'core' sequence. Thus the failure to localise the 'enhancing' activity of pNcV10 may result from a requirement for the presence of each of these putative regulatory elements. In such a scheme the Z-DNA sequences may alter the superhelical density of the -4.5 kb promoter in such a way as to increase transcription from this region. Alternatively some other mechanism may be involved
in causing a change in the activity of the epsilon globin gene promoter or/and procaryotic sequences as a result to a change in the conformation of the upstream promoter. The failure to reproduce the '10' effect in short term transfections is difficult to explain. A possible case to argue follows along the lines previously raised on the differences between the activity of globin genes in stable and acutely transformed cells in which it was proposed that sequences may adopt a different conformation whether integrated or free form. This however is by no means a satisfactory account of the discrepancy between these two assay systems and accordingly must question the activity of fragment '10' under these particular experimental conditions.

4.2.3 \textit{Trans Acting Sequences}

In addition to the previously discussed \textit{cis} sequences which influence the expression of the human epsilon globin gene several lines of evidence show that \textit{trans} acting sequences also play an important part in its expression.

Allan \textit{et al} \cite{91} demonstrated \textit{trans} activation of the epsilon globin gene by co-transformation with a recombinant encoding the viral protein EIA. This product is thought to be involved in the activation of transcription of the late adenovirus genes by a catalytic mechanism possibly involving the assembly of stable transcriptional complexes \cite{273}. In these experiments,
transcription was reduced from the -200 promoter by around 20 fold with a corresponding increase in the number of transcripts synthesised from the major cap site involving some sort of redirection mechanism.

To summarise several lines of evidence as presented suggest that the major and the -200 promoters of the human epsilon globin transcriptional unit can be regulated independently in response to a variety of both cis and trans operating mechanisms. These studies propose that these upstream promoters particularly the -200 promoter is transcriptionally active during early embryonic erythrophoietic development. The switch from the upstream promoters including the -200 occurs in later development. This redirection mechanism is mediated at least in part by a specific embryonic trans factor, the action of which can be mimicked by EIA or the presence of a cis linked viral enhancer, albeit by a fundamentally different mechanistic process.

In a study by Allan et al [265] as described in detail in this thesis we demonstrated trans activation of the pHRI1 derived recombinant pSVKNε-3 in both acutely and stably transformed K562 cells. In all other cell lines tested the cap: -200 ratio remained 1:20 compared to the 20:1 value in the K562 cell. However this in many ways contrasts to the apparent failure of the human epsilon globin promoter to relatively increase the expression of selectable Tn5 aph (II) gene in the K562 cell. This possibly suggests that the trans operating embryonic
erythroid specific factors do not recognise the 5'-Bam H1/PvuII-3' promoter fragment alone, implying that further recognition sequences may be located upstream or possibly encoded within the structural gene. In addition, as previously discussed, the fundamental differences of the assay systems must be taken into account when comparing these phenomena.

A recent study by Young et al [274] demonstrated induction of an exogenous epsilon globin gene recombinant in K562 cells. In this series of experiments the epsilon globin gene was covalently linked to the SV40-neomycin vector pSV2-Neo [191] and stably introduced into the K562 cell line variant Bos. These cells display a very low activity of the endogenous epsilon globin gene. After induction by treatment with 20 μM haemin the transcripts were S1 mapped with the 371 bp's MboII probe. Only the level of transcripts originating from the major cap site increased. This finding is consistent with the observation made in this study on the failure to co-induce the human epsilon globin gene in stably transformed MELtk⁻ cells as was the case with two other independent groups [170] [171].

These various studies imply the human epsilon globin gene is subject to trans acting factors. In addition it would appear that the -200 region does not markedly respond to these diffusible molecules which may explain the non regulated expression of the human epsilon globin gene in transformed cells due to the lack of expression of those embryonic erythroid trans acting genes.
4:2:4 Conclusion - the Spatial and Functional Overlap of Regulatory Elements

What exactly these experiments say about the regulation of the human epsilon globin gene is difficult to evaluate coherently. Moreover how valid are these data in that can the study of one component of the transcriptional unit be assessed independently of other cis sequences, both coding and flanking DNA? Along these lines Serfling et al [275] have recently suggested that the promoter and other regulatory sequences may overlap both physically and functionally. They extend their idea that upstream regulatory regions are composed of some sort of modular arrangement of short sequence stretches which are recognised by trans acting factors which in the case of the human epsilon globin gene would be involved in conferring tissue specificity, developmental regulation, a general enhancement of transcription and induction during red blood cell maturation.

4:3 The Regulated Expression of the Human Epsilon Globin Gene - A Hypothesis Involving Both Proximal and Distant Sequences

The S1 Nuclease mapping studies on the human epsilon globin gene in K562 and transformed cells clearly imply
the existence of putative regulatory sequences upstream of
the major promoter. Moreover some of the work described
in this thesis has identified the possible loci of these
control regions. Does this information tell us anything
about how this gene is regulated? A somewhat speculative
model is presented.

The human β globin gene family in non-erythroid tissue
is in an inactive chromatin conformation generally
associated with the 300 Å solenoid higher order structure
[113]. In contrast this locus is more sensitive to DNase
I in the erythroleukaemic cell lines K562 and HEL than non
erythroid tissue [194], moreover in the haematopoietic
promyelocytic leukaemic cell line HL60 it is also DNase I
insensitive [194]. This difference in DNase I digestion
pattern possibly reflects a conformational change common
to all erythroid precursor cells which can be envisaged as
a transition from the inaccessible 300 Å solenoid
structure to the more open 100 Å 'beads on a string' fibre
form [113]. Accordingly this may represent the initial
event in globin gene activation and is only a feature of
haematopoietic cell populations which are committed to the
erythroid lineage.

It is not known whether the globin genes are repressed
by a general or specific mechanism. Considering that a
typical higher eucaryote contains around 10,000 tissue
specific genes and in an erythroid cell most of these
sequences will not be expressed it is likely that the
dominant repression mechanism will be of general type. A
favoured model involves the non-core histone H1 which has been proposed to function in the maintenance of the higher order 300 Å structure [276]. The mechanism responsible for the depression of the human β globin gene is unknown. However studies with activation of globin genes from fibroblasts in MEL somatic cell hybrids would argue that these trans acting factors are not restricted in their activity to early erythroid precursors. This constitutive type of expression may imply that they may also be involved in the maintenance of the derepressed state in non erythroid tissue. Tuan et al [194] have DNase I mapped the human β globin gene cluster in K562, HEL, HL60 and marrow tissue. They distinguish two types of DHS's. Minor sites reflect the particular ontogenic stage and vary in K562, HEL and bone marrow cells. In addition they have identified four major DHS's which are present in all erythroid tissue irrespective of the development period and are totally absent in the myeloid HL60 cells. Three of these are located 5' to the epsilon globin gene at ~6.5 kb, ~11 kb and ~17 kb. The other maps around ~20 kb downstream of the β globin gene.

Each of these major DNase I hypersensitive sites are located in regions in DNA containing one or more enhancer core elements and stretches of alternating purines and pyrimidines which under certain physiological conditions can form Z-DNA. A possible function of these major DNase I hypersensitive sites might be to organise and ensure the structure of the β globin gene locus in chromatin is in a
transcriptionally preactivated conformation which is receptive to erythroid specific signals which in turn modulates expression of the individual genes.

The next series of events envisaged involve local changes in the chromatin structure of the 10 kb or so of the epsilon globin gene domain. A possible candidate involved in this is the most proximal of the Z-DNA elements described by Tuan et al [194] located ~6.5 kb upstream of the major cap site. Despite the observation that the effect of such a B→Z transition becomes weaker with increasing distances from the target DNA it could be the case this transition only represents the initial event of a sequential mechanism involved in the transcriptional activation of the epsilon globin gene domain. Once unwound other regulatory cis acting sequences may be subject to regulatory interaction with trans acting factors. This Z-DNA element is located about 2 kb upstream of the most distal promoter at ~4.5 kb. It is interesting that the Eco RI genomic fragment of pHR10 which increases expression of the Tn5 aph (II) in a range of cell types encodes for both of these putative regulatory elements. Excluding the possible involvement of this Z-DNA element, the epsilon globin gene of an already derepressed human β globin gene cluster undergoes a conformational change, the result of which is the appearance of a constellation of DHS's which correspond to the siting of upstream promoters. These changes are presumed to be mediated by the action of embryonic early
erythroid trans factors, one of which possibly binds to the Z-DNA element.

It is not known if transcriptional activity immediately follows this or whether some further structural modification is required. At these stages of early development the upstream promoters are active especially the -200 promoter. Evidence supporting this notion as described throughout this text includes the lower cap: -200 ratio observed in transformed cells, 'leaky' transcription involving exclusive use of the -200 promoter in certain non-erythroid cell lines [10], its failure to be activated in induced K562 cells [273] and the influence of the cis linked SV40 origin of replication on transcription from the -200 start site [91]. This final point is particularly suited to this model since during early erythropoietic differentiation the cells are dividing rapidly up to the pro-erythroblast stage. Finally during late maturation the embryonic erythrocyte which for the purposes of this model can be represented as an erythro-leukaemic K562 cell alters its use of promoters resulting in the majority of transcripts initiating from the major cap site giving the characteristic cap: -200 ratio of 20:1. This transition is arguably mediated by the action of late embryonic erythropoietic specific trans acting factors as suggested by comparing the cap: -200 ratio of the 'marked' epsilon globin gene recombinant in the K562 cell and a variety of other cell types.

A possible mechanism of action responsible for this
differential use of promoters during ontogeny is the phenomenon of transcriptional interference. How can transcription per se exert a regulatory influence on epsilon globin gene expression?

Paul et al [269] have suggested a number of ways in which transcription could exert an influence. Transcription may increase the accessibility of the chromatin template to regulatory diffusible molecules. RNA molecules originating from upstream promoters could inhibit the binding of transcriptional factors at the major promoter, such a scheme has been proposed to account for the action of the bacterial IS1 insertion element [277]. A similar type of mechanism may be involved in the apparent repression of major cap site usage of the 1468 bp's 5'-Eco RI/XbaI-3' fragment in pHRL1 and an unpublished observation of Allan and Grindlay (Pers Comm) that the 1480 bp 5'-XbaI/TaqI-3' fragment containing the -4500 promoter down regulates the major promoter. Transcriptional interference of the longest epsilon globin gene transcript may arise from the expression of the distal AluI Pol III promoter which is transcribed from the opposite strand. This proposal is strengthened by the observation that this sequence is only expressed in K562 cells[163].

Continuing on this theme Johnson et al [278] have presented firm evidence that the putative termination signal of the sea urchin H2A histone gene can efficiently terminate the transcription of the human α globin gene
when inserted into the third exon in a HeLa cell transient assay system. Moreover when this fragment was placed between the two human α globin genes and assayed as above using S1 Nuclease mapping the activity level of the second gene increased by around three-fold. These results imply that transcriptional interference can occur between adjacent genes in the absence of an active terminator. Partially based on these findings these authors have argued a case on the possible role of termination in the developmental expression of the human β globin genes. In this positive control type of mechanism Proudfoot et al [279] have speculated on the existence of two terminator regions which are only activated in response to erythroid trans factors. It is proposed that one is positioned between the embryonic and foetal domains, the other between the foetal and adult regions. All genes downstream of the expressed gene are repressed in part by a transcriptional readthrough interference mechanism. At specific stages each terminator is activated so attenuating these long transcripts and concomittantly allowing expression of the downstream gene.

Lastly it has been suggested that these upstream regions may be RNA Polymerase II entry sites, a mechanism proposed to explain the action of enhancers. This scheme predicts that upon entry the transcriptional complex will scan the template until it recognises the major promoter, however because of a lack of late embryonic erythroid specific trans factors the complex does not recognise the
major promoter and subsequently will initiate at these entry sites.

4:4 **Tissue Specific and Developmental Expression of Adult β Globin Genes - A Critical Role for Sequences Close to and/or 5' to the Cap Site**

In the previous sections evidence was presented which strongly favoured an important role for sequences upstream of the major promoter in the regulated expression of the human epsilon globin gene. But how does this interpretation of a wide range of phenomena fit in with current thinking on how globin genes are regulated? Is it realistic to expect the upstream promoters of the foetal and adult globin gene to have an analogous physiological role to that proposed for these sequences in the epsilon globin gene? Do the mechanisms which regulate expression vary for different globin genes or are the complex genetic programmes responsible for commitment, tissue specificity, coordinate and temporal regulation highly conserved? Such conservation as indicated by homology in the 5' flanking DNA is not too evident. Sequence comparison of this region from a range of β globin type genes show little homology extending beyond the promoter region. Moreover the β globin promoter has been classified as enhancer dependent (1:5:1) and replication independent as shown by its lack of expression in the frog oocyte while the α globin gene
categorised as enhancer independent is actively expressed in the frog oocyte [138]. Interestingly the human epsilon globin gene displays both properties, albeit via different promoters. The -200 region responds to the presence of the origin of replication while only the major cap site activity increases when cis linked to a viral enhancer [91].

The transcriptional activity of the adult β globin genes of human [170], [173], rabbit [174] and mouse [280] with templates of various sizes and with hybrid templates were examined in stably transformed induced MEL tk- cells. These series of experiments point to at least two non-contiguous cis putative regulatory regions of the adult β globin transcriptional unit associated with the tissue specific expression of this gene. One is positioned within the tripartite promoter region and the other located 3' to the cap site possibly encoded in the structural gene.

Can long term transformation experiments such as the above co-induction studies be equated with how an adult β globin gene behaves during normal ontogeny considering our ignorance of the physical status of the foreign DNA in the animal cell?

An approach to counter the important criticisms against globin gene transformation experiments involves the use of transgenic mice in which the exogenous gene becomes a stable and inheritable component of the mouse's genotype [206]. Early experiments with plasmid and
bacteriophage recombinants failed to produce transgenic mice in which the exogenous DNA was correctly expressed [281] [282]. But in subsequent experiments in which the amount of procaryotic sequences were reduced, erythroid specific expression of a mouse/human adult β globin hybrid gene was demonstrated by Chada et al [283]. An extension of this study by Magram et al [284] provided evidence that this hybrid globin gene was also expressed at the correct developmental stage. Although these studies confirmed tissue specificity and developmental regulation of a foreign globin gene, the level of globin hybrid message was low, between 2 and 4% of endogenous activity.

Townes et al [285] introduced a cloned human β globin gene containing 4300 bp's of 5' flanking sequences and 1700 bp's of 3' flanking DNA up to the pronuclei of mice as a plasmid free linear fragment. Around 75% of the transgenic mice expressed the globin gene in a tissue-specific and developmental manner. In contrast to the previously described findings of Chada et al [283] and Magram et al [284] these workers found that the level of foreign β globin gene expression was comparable to the endogenous mouse adult β globin gene.

But in transgenic mice carrying the human β globin gene recombinant covalently linked to plasmid sequences the level of exogenous globin gene activity dropped to between 100 and 1000 fold of that of the pure globin DNA linear fragment. This phenomenon of procaryotic sequence interference may account for some of the earlier
transgenic experiments which failed to obtain regulated expression.

Of further interest from this series of transgenic experiments was the observation that only 48 bp's of 5' flanking sequence was required for tissue specific and temporal regulation of the human β globin gene. Although qualitative expression was observed the level of transcriptional activity was lower on average than measured for templates with flanking sequences beyond the promoter.

These examples clearly demonstrate that the use of transgenic mice can be a rewarding system in the identification of cis sequences in the regulated expression of globin genes. A popular criticism levelled against globin gene transfer experiments is that the normal events leading to gene activation and expression are bypassed, accordingly sequential regulation cannot be monitored. Despite the success of transgenic mice in terms of tissue specificity and developmental regulation of adult globin genes, as an in vivo expression system is not immune to criticism particularly on the chromosomal location of the foreign DNA.

A notable feature of these experiments is the significant proportion of transgenic mice with integrated copies of the donor globin genomic sequences which are transcriptionally silent. A possible explanation for this is that these exogenous globin genomic fragments are integrated into chromosomal regions which are incompatible
with expression. Presumably in those animals which readily express the foreign globin gene, the donor DNA is integrated into an active region of the chromatin possibly near a housekeeping gene. It is not known if the conformation of the neighbouring globin gene is significantly influenced by the presence of this 'active' gene. Transcriptional readthrough and changes in the degree of superhelicity may have profound effects on the expression of the exogenous DNA. Such an influence could conceivably perturb the sequential regulated expression of the foreign globin gene. It could now be out of phase with those mechanisms governing the expression of the endogenous globin gene. The events leading to the initial activation of the mouse β globin gene cluster may not recognise the foreign globin DNA due to topological influences of active genes in its vicinity.

The combined studies with MEL transformed cells and transgenic mice carrying foreign adult globin genes support the hypothesis that sequences both in the promoter region and downstream of the cap site are an integral component of the mechanism responsible for the regulated expression of the human β globin gene. An extension of this is that much of the DNA of the β globin domain is redundant. This is in striking contrast to the proposed explanations for some cases of the thalassaemic condition known as HPFH (hereditary persistence of foetal haemoglobin) which implicate distant sequences in the expression of the γ globin genes, however this may be non-
regulated or abnormal expression. In addition these findings possibly call for a reassessment of work of Emerson and Felsenfeld [116] on the existence of the chicken erythroid specific factor which is only expressed in adult erythroid cells and binds to a region upstream of the ATA box of the adult β globin gene.

4:5 Prospects

In many respects these results on the in vivo expression of adult globin genes issue a challenge to the findings reported in this thesis on the behaviour of the human epsilon globin gene in transformed cells. The identification of sequences near to or possibly encoded within the structural gene involved in the regulated expression of the human β globin gene differs markedly from the implied importance of sequences extending well beyond the major promoter of the human epsilon globin gene. The experiments on the human epsilon globin gene lag behind those performed on the adult β globin genes of man, rabbit and mouse where the fine mapping of sequences to the nucleotide level involved in tissue specificity and temporal regulation are likely to be determined in the near future. Mapping of equivalent regions within the transcriptional unit of the human epsilon globin gene seems far off at present. This rather pessimistic prediction is based on unpublished findings by Constantini (Oxford) [Pers Comm] who failed to produce transgenic mice
which expressed the human epsilon globin gene in a regulated fashion. This observation, albeit of a preliminary nature, is somewhat unexpected considering the mouse embryonic β globin genes are active at the appropriate time during normal ontogeny. A possible explanation is that the embryonic β type globin genes of man and mouse are expressed in mechanistically different ways. In this respect it would be interesting to know if the embryonic murine gene is highly expressed or inducible in the K562 cell. Perhaps the human foetal locus is more comparable to the mouse embryonic domain suggesting that throughout the course of β globin gene evolution different globin genes are expressed at different developmental stages. This hypothesis would possibly predict that the human γ globin genes would be expressed in the embryonic erythroid tissue of transgenic mice.

Is there another animal in which transgenic experiments could be performed with? Examples such as the goat [286] which have an embryonic globin gene domain, would in theory seem to be good candidates. However there may be technical and operational difficulties in establishing such a system at this juncture.

What experimental strategies should be followed to identify regulatory sequences of the human epsilon globin gene, particularly those involved in tissue specificity and developmental regulation?

The studies on human β globin gene expression in transformed induced MEL tk⁻ cells and transgenic mice
display a similar sequence requirement for co-induction and regulated expression. Similar induction experiments could be undertaken with the 'marked' epsilon globin gene recombinant pSVKNε-3 [265] by generating a series of linker scanner mutants and examining their expression in haemin treated K562 cells.

The chromosome transfer experiments discussed in (1:6:2) are indicative that globin genes from non-erythroid cells in which the globin gene loci are assumed to be in the depressed state, are activated and expressed in a regulated manner in the terminally differentiated erythroid cell. Presumably the presence of the whole chromosome is not required for regulated expression. These ideas inevitably lead back to gene transfer experiments. But as discussed throughout this thesis, gene transfer experiments represent an artificial system in which the exogenous globin gene may be influenced by other sequences. Therefore can the 'marked' epsilon gene be introduced into a cell such that the gene is effectively isolated from the interference of non-epsilon globin gene sequences? Stable transformation experiments require the presence of a selectable gene in addition to the test DNA. As an initial step in minimising this non-globin sequence interference, the test and selectable eg. neomycin gene should ideally be spatially separated in the recipient cell genotype. This may be attained by introducing each of these DNAs in the absence of carrier DNA to the recipient cell line in a
staggered fashion to permit them to integrate separately into the host's chromosome. Suitable transformed cell lines for study would be selected on the basis of a low epsilon globin gene copy number. This condition is desired since our knowledge of the physical form of DNA in transformed cells is somewhat limited. If a mixed population is envisaged and there is a high copy number of donor DNA it would be very difficult to determine which form of donor DNA was responsible for expression.

The recent work by Johnson et al. [278] with the sea urchin H2A histone terminator on the possible importance of transcriptional interference as a regulatory phenomenon, can be exploited by employing this terminator to act as a buffer against sequences which potentially interfere with the expression of the human epsilon globin gene in transformed cells. Thus the positioning of the terminator at both termini of the epsilon globin gene fragment might satisfy the aim of examining the gene in isolation. Experiments along these lines with varying template ranging in size from promoter deletion to the whole domain and possibly even the entire ε globin gene cluster in a variety of erythroleukaemic cell types such as K562, HEL and KMOE which represent all stages of developmental expression in addition to non-erythroid cell lines may be a rewarding approach in the identification of regulatory signals important to the in vivo expression of the human epsilon globin gene.

Another possible approach to circumvent this potential
problem of transcriptional interference is to direct the cloned epsilon globin gene into its normal chromosomal location at the expense of the endogenous equivalent by homologous recombination. Smithies et al [287] recently demonstrated using a rationale and technique initially developed in yeast [288], the homologous exchange between a human β globin test gene and the endogenous gene. The test plasmid in addition to carrying the neomycin gene to allow the selection of transformed colonies also encodes for a bacterial supF gene. This provides a means of selectively propagating the exchanged DNA in bacterial cells after rescue from the transformed animal cells [289]. This functions as an assay to check for homologous recombination among the transformed population carrying exogenous globin sequences. However the rate of homologous recombination was estimated to be around $10^{-3}$ of the transformed cells, this figure was independent of whether the endogenous globin gene was active or repressed. Such analogous experiments although theoretically possible with the human epsilon globin gene, the low transformation efficiency of the K562 cell would demand an initial series of long term transformations on a massive scale to generate a few thousand colonies before it would be worthwhile screening for homologous recombination.

ADDENDUM Since completing this Discussion, Chada et al [290] have introduced a cloned human $\gamma$ globin gene into the mouse gene line and demonstrated that this foetal gene is solely expressed during murine embryonic development in erythroid cells.
References


Activation of Cellular Genes by Avian RNA Tumour Viruses  
Proc. Nat. Acad. Sci. USA. 77: 5351-5354

Globin RNA Sequences in Human Leukaemic Peripheral Blood  
Nature 274: 610-612

Multiple Origins of Transcription in the 4.5 kb Upstream of the ε-Globin Gene  
Cell 35: 187-197

Excess Alpha Chain Synthesis Relative to Beta Chain Synthesis in Thalassaemia Major and Minor  
Nature 212: 1198-1200

DNA Methylation in the Human γδβ-Globin Locus in Erythroid and Nonerythroid Tissues  
Cell 19: 947-958

The Genetics of Antibody Diversity  
Sci. Am. (May) 72-83

Simultaneous Expression of Globin Genes for Embryonic and Adult Haemoglobins During Mammalian Ontogeny.  
Science 218: 1225-1229


Cell 19: 959-972

Cell 15: 1157-1174

Nucleic. Acid. Res. 7: 1147-1162

Science 202: 1284-1289

Proc. Natl. Acad. Sci. USA 76: 5435-5439

Gene 13: 227-237

Shoulders, C.C., and Proudfoot, N.J. (1980)  
The Structure and Evolution of the Human β-Globin Gene Family  
Cell 21: 653-668

Structure of the Human $\gamma$-$\delta$-β Globin Gene Locus  
Proc. Natl. Acad. Sci. USA 76: 4827-4831

Characterisation of Deletions Which Effect the Expression of Foetal Globin Genes in Man  
Nature 279: 598-603

Globin Psuedogenes  
Cell 28: 683-684

Mouse Globin System: A Functional and Evolutionary Analysis  
Science 209: 1136-1342

Two Related Psuedogenes are the-Result of a Gene Duplication in the Goat β-Globin Locus  
Cell 26: 181-190

The Nucleotide Sequence of a Rabbit β-Globin Pseudogene  
Cell 21: 545-553

A History of the Human Foetal Globin Gene Duplication  
Cell 26: 191-203

- 245 -
Why Genes in Pieces?
Nature 280: 162-164

Intron-less Globin Genes in the Insect *Chironomus thummi thummi*
Nature 310: 795-798

Cloning and Complete Nucleotide Sequence of Human 5' α- Globin Gene
Proc. Natl. Acad. Sci. USA 77: 7054-7058

The Structure of the Human Zeta-Globin Gene and a Closely Linked, Nearly Identical Pseudogene
Cell 31: 553-563

The Nucleotide Sequence of the Human β-Globin Gene
Cell 21: 647-651

Complete Nucleotide Sequence of the Human δ-Globin Gene
Cell 21: 639-646

Human Foetal Gγ- and Aγ- Globin Genes: Complete Nucleotide Sequences Suggest that DNA can be Exchanged between these Duplicated Genes
Cell 21: 627-638
The Primary Structure of the Human ε-Globin Gene
Cell 21: 621-626

[43] Baralle, F.E., Shoulders, C.C., Goodburn, S.,
The 5' Flanking Region of Human ε-Globin Gene
Nucleic. Acid. Res. 8: 4393-4404

Human Globin \( \psi \beta_2 \) is not a Globin Related Gene
Nucleic. Acids. Res. 10: 7809-7818

Nucleotide Sequence of 16- Kilobase Pairs of DNA 5'
to the Human ε Globin Gene
J. Biol. Chem. 260: 14901-14910

Advances in Thalassaemia Research
Blood 63: 738-758

Organisation and Expression of Eucaryotic Split
Genes Coding for Proteins

Sequence Analysis of Drosophila Histone Genes
Ph.D. Thesis, Stanford University, Palo Alto, California

[49] Benoist, C., O'Hare, K., Breathnach, R., and
Chambon, P. (1980)
The Ovalbumin Gene Sequence of Putative Control
Regions
Nucleic. Acid. Res. 8: 127-142
The Nucleotide Sequence of the Rabbit Embryonic Globin Gene $\beta_4$
J. Biol. Chem. 258: 8739-8744

Unusual Sequence Homology at the 5' Ends of the Developmentally Regulated $\beta^A$, $\beta^C$ and $\gamma$-Globin Genes of the Goat
Proc. Natl. Acad. Sci. USA 77: 7127-7131

AUG is the only Recognisable Signal Sequence in the 5' Non-Coding Regions of Eucaryotic in mRNA
Nature 274: 84-87

A Catalogue of Splice Junction Sequences
Nucleic. Acid. Res. 10: 459-472

3' Non-Coding Region Sequences in Eucaryotic Messenger RNA
Nature 263: 211-214

The Alu Family of Dispersed Repetitive Sequences
Science 216: 1065-1070

Sequence Organisation of the Human Genome
Cell 6: 345-358
Repetitive DNA Sequences Near Three Human β-Type Globin Genes
Nucleic. Acid. Res. 8: 3319-3334

Characterisation of Alu Family Repetitive Sequences Which Flank Human β-Type Globin Genes

A B1 Repetitive Sequence Near the Mouse β-Major Globin Gene
Gene 17: 113-116

Molecular Evolution of the Human Adult α-Globin Like Gene Region: Insertion and Deletion of Alu Family Repeats and Non-Alu DNA Sequences

Trans. Assoc. Am. Physiol. 25: 204-210

A Family of Long Reiterated DNA Sequences, One Copy of Which is Next to the Human Beta Globin Gene
Nucleic. Acids. Res. 8: 6113-6128

Lambdoid Phages that Simplify the Recovery of In Vitro Recombinants
Right-Handed and Left-Handed DNA: Conformational Information in Gentic Material

Variety in the Level of Gene Control in Eucaryotic Cells
Nature 297: 365-371

Lineages, Quantal Cell Cycles and the Generation of Cell Diversity
Quart. Rev. Biophys. 8: 523-557

[67] Scherrer, K, and Marcaus, L. (1968)
Messenger RNA in Avian Erythroblasts at the Transcriptional and Translational Levels and the Problem
J. Cell. Physiol. 72: 181-212

The Assymetric Segregation of Parental Nucleosomes During Chromosome Replication
Cell. 18: 439-449

Cell Cycle Dependence of Induced Haemoglobin Synthesis in Friend Erythroleukaemia Cells Temperature Sensitive for Growth
Proc. Natl. Acad. Sci. USA 78: 3644-3648

Erythroleukaemic Differentiation
Ann. Rev. Biochem. 47: 419-426


The Size of the Transcriptional Units of the Avian
Globin Genes Defined at the Pre-Messenger RNA level
J.Mol.Biol. 140: 481-504

Capping of Eucaryotic mRNAs
Cell 9: 645-653

[80] Abraham, G., Rhodes, D.P., and Banerjee, A.K.
(1975)
The 5' Terminal Structure of the Methylated mRNA
Synthesised In Vitro by Vesicular Stomatitis Virus
Cell 5: 51-58

The Primary Transcription Unit of the Mouse β-
Major Globin Genes
Cell 23: 585-593

A Transcription Map for the Rabbit-β-Globin Gene
Cell 23: 573-584

Association of Globin Ribonucleic Acid and its
Precursors with the Chicken Erythroblast Nuclear
Matrix
Biochemistry 21: 764-771

[84] Herlan, G., Eckert, W., Kaffenerberger, W., and
Wunderlich, F. (1979)
Isolation and Characterisation of an RNA-
Containing Nuclear Matrix from Tetrahymena
Macronuclei
Biochemistry 18: 1782-1788


Analysis of the Mouse DHFR Promoter Region: Existence of a Divergently Transcribed Gene
Mol. Cell. Biol. 5: 1838-1847

Promoter-Specific Activation of RNA Polymerase II Transcription by Sp1

Normal and Mutant Human β-Globin Pre-mRNAs Are Faithfully and Efficiently Spliced in Vitro
Cell 36: 993-1005

[100] Chan, L.L-N. (1976)
Transport of Globin mRNA from Nucleus into Cytoplasm in Differentiating Embryonic Red Blood Cells.
Nature 261: 157-159

Altered Globin Gene Transcription Pattern and the Presence of a 7-8 kb αA Globin Gene Transcript in Avian Erythroblastosis Virus-Transformed Cells
EMBO. J. 3: 491-495

Biosynthesis and Stability of Globin mRNA in Cultured Erythroleukaemic Friend Cells
Cell 8: 495-503

Stability of Globin mRNA in Terminally Differentiating Murine Erythroleukaemia Cells
Cell 23: 509-514
Abstract from Proceedings of the Fourth Conference on Haemoglobin Switching.

Evidence for a Translational Inhibitor Linked to Globin mRNA in Untranslated Free Cytoplasmic Messenger Ribonucleoprotein Complexes
FEBS. LETT. 72: 71-76

Differential Repression of Specific mRNA in Erythroblast Cytoplasm: A Possible Role for Free mRNP Proteins
EMBO. J. 2: 1869-1876

Half-Lives of Beta and Gamma Messenger RNAs in Cultured Human Reticulocytes.
Abstract from Proceedings of the Fourth Conference on Haemoglobin Switching

The Interaction of High Mobility Proteins HMG 14 and 17 with Nucleosomes
Nucleic. Acid. Res. 8: 3757-3778

Selected Topics in Chromatin Structure

Human Foetal to Adult Haemoglobin Switching:
Changes in Chromatin Structure of the β-Globin Gene Locus
Proc. Natl. Acad. Sci. USA 80: 7551-7555

DNase I Hypersensitivity in the γ Globin Gene Locus of K562 Cells
Nucleic. Acid. Res. 11: 6065-6077

Chromosomal Subunits in Active Genes Have an Altered Conformation.
Science 193: 848-858

Higher Order Structure of Chromatin: Orientation of Nucleosomes Within the 30 nm Chromatin Solenoid is Independent of Species and Spacer Length
Cell 33: 831-841

Hb Switching in Chickens
Cell 19: 973-980

Propagation of Globin DNase I - Hypersensitive Sites in Absence of Factors Required for Induction: A Possible Mechanism for Determination
Cell 30: 131-139

Specific Factor Conferring Nuclease Hypersensitivity of the 5' End of the Chicken Adult β-Globin Gene
A 200 Base Pair Region at the 5' End of the Chicken Adult β-Globin Gene is Accessible to Nuclease Digestion
Cell 27: 45-55

DNA Conformation at the 5' End of the Chicken Adult β-Globin Gene
Cell 35: 467-477

Endonuclease S1-Sensitive Site in Chicken Pro α2 (I) Collagen 5' Flanking Gene Region

Association of an S1 Nuclease-Sensitive Structure with Short Direct Repeats 5' of Drosophila Heat Shock Genes
Nature 304: 555-557

DNA Methylation - How Important in Gene Control?
Nature 307: 503-504

Molecular Mechanisms of Human Haemoglobin Switching: Selective Undermethylation and Expression of Globin Genes in Embryonic, Foetal and Adult Erythroblasts
Proc. Natl. Acad. Sci. USA 80: 6907-6911


Transcriptional Analysis of Human Zeta Globin Genes
EMBO. J. 3: 1533-1540

Gene Transfer in Amphibian Eggs and Oocytes

Eds. Enhancers and Eucaryotic Gene Expression
(Cold Spring Harbour Laboratory, Cold Spring Harbour, N.Y.)

Multiple Point Mutations Affecting the Simian Virus 40 Enhancer
Science 219: 626-631

Negatively Supercoiled Simian Virus 40 DNA Contains
Z-DNA Segments Within Transcriptional Enhancer Sequences
Nature 303: 674-678

Analysis of the Transcriptional Enhancer Effect

Activation of SV40 Genome by 72 Base Pair Tandem Repeats of Moloney Sarcoma Virus
Nature 295: 568-572
Host-Specificities of Papillomavirus, Moloney Murine Sarcoma Virus and Simian Virus 40 Enhancer Sequences
EMBO. J. 2: 1193-1199

A Tissue-Specific Transcription Enhancer Element is Located in the Major Intron of a Rearranged Immunoglobulin Heavy Chain Gene
Cell 38: 717-725

Immunoglobulin Gene Transcription is Activated by Downstream Sequence Elements
Cell 33: 741-748

A Lymphocyte-Specific Enhancer in the Mouse Immunoglobulin Kappa Gene
Nature 307: 80-82

Cell-Specific Expression Controlled by the 5' Flanking Region of Insulin and Chymotrypsin Genes
Nature 306: 557-561

Specific Interaction Between Enhancer-Containing Molecules and Cellular Components
Cell 36: 403-411

Bovine Papillomavirus Vector that Propagates as a Plasmid in Both Mouse and Bacterial Cells
Potential Z-DNA Forming Sequences are Highly
Dispersed in the Human Genome
Nature 298: 396-398

A Member of a New Repeated Sequence Family Which is
Conserved Throughout Eucaryotic Evolution is Found
Between the Human $\gamma$ and $\beta$ Globin Genes
Nucleic. Acid. Res. 9: 5931-5947

(1982)
The DNA Sequence of the 5' Flanking Region of the
Human $\beta$-Globin Gene: Evolutionary Conservation and
Polymorphic Differences
Nucleic. Acid. Res. 10: 2109-2120

Inhibition of Eucaryotic RNA Transcription by
Potential Z-DNA Sequences
EMBO. J. 3: 1553-1559

(1980)
Torsional Tension in the DNA Double Helix Measured
with Trimethylpsoralen in Living E. coli Cells:
Analogous Measurements in Insects and Human Cells
Cell 21: 773-783

[156] Nordheim, A., Tesser, P., Azorin, F., Kwon, Y.H.,
Isolation of Drosophila Proteins that Bind
Selectively to Left-Handed Z-DNA
A Negative Regulatory Sequence Near the Mouse β-maj
Globin Gene Associated with a Region of Potential
Z-DNA
EMBO. J. 3: 1263-1272

Genetics of Human Cell Lines, DNA Mediated
Heritable Transformation of a Biochemical Trait

Internal Control Regions for Transcription of
Eucaryotic tRNA Genes
Proc. Natl. Acad. Sci. USA 78: 6657-6661

Weissman, S.M. (1979)
RNA Polymerase III Transcriptional Units are
Interspersed Among Human Non-α-Globin Genes
Proc. Natl. Acad. Sci. USA 76: 5095-5099

[161] Duncan, C.H., Jagadeeswaran, P., Wang, R.R.C., and
Structural Analysis of Templates and RNA Polymerase
III Transcripts of Alu Family Sequences Inter-
spersed Among the Human β-Like Globin Genes
Gene 13: 185-196

Synthesis In Vitro of an Exceptionally Long RNA
Transcript Promoted by an Alu I Sequence
Nature 300: 376-379
Transcription *In Vivo* of an Alu Family Member
Upstream from the Human ε-Globin Gene.
Nucleic. Acid. Res. 12: 1193-1200

[164] Di Segni, G., Carrara, G., and Tocchini-Valenti, G.
Selective *In Vitro* Transcription of One of the Two
Alu Family Repeats Present in the 5' Flanking
Region of the Human ε-Globin Gene.
Nucleic. Acid. Res. 9: 6709-6722

Human β-Globin Promoter and Coding Sequences
Transcribed by RNA Polymerase III
Cell 34: 857-864

Human α-Globin Gene Expression Following
Chromosomal Dependent Gene Transfer into Mouse
Erythroleukaemia Cells
Cell 15: 55-63

[167] Deisseroth, A., Bode, U., Fontana, J., and
Expression of Human α-Globin Genes in Hybrid Mouse
Erythroleukaemia Cells Depends on Differential
State of Human Donor Cell
Nature 285: 36-38

[168] Pyati, J., Kuckerlapati, R.S., and Skoultchi, A.I.
(1980)
Activation of Human β-Globin Genes from
Nonerythroid Cells by Fusion with Murine Erythro-
leukaemia Cells
Proc. Natl. Acad. Sci. USA 77: 3435-3439
Selective Activation of Human β-But Not γ Globin Gene in Human Fibroblast x Mouse Erythroleukaemia Cell Hybrids
Nature 277: 534-538

The Regulated Expression of β-Globin Genes Introduced into Mouse Erythroleukaemia Cells
Cell 32: 483-493

Regulated Expression of the Human βGlobin Gene Family in Murine Erythroleukaemic Cell Hybrids
Nature 305: 333-336

Transfer of Human Globin Genes to Erythroleukaemic Mouse Cells
EMBO.J. 1: 15-20

Differences in Human α- and β-Globin Gene Expression in Mouse Erythroleukaemia Cells:
The Role of Intragenic Sequences
Cell 38: 251-263

DNA Sequences Required for Regulated Expression of β-Globin Genes in Murine Erythroleukaemia Cells
Cell 38: 265-273
Transcriptional Regulation of Globin Gene Expression in the Human Erythroid Cell Line K562
Science 220: 1281-1283

A New Dominant Hybrid Selective Marker for Higher Eucaryotic Cells.
J. Mol. Biol. 150: 1-14

Differential Expression of Human Globin Genes Introduced in K562 Cells
EMBO.J. 4: 927-931

Transfer of Human Globin Genes to Human Erythro-leukaemia Cells
Mol. Biol. Med. 2: 167-175

Hybridisation in the Analysis of RNA.
Nucleic Acid Hybridisation - A Practical Approach
Eds. Hames, B.D., and Higgins, S.J.
IRL PRESS

Pre Early Adenovirus 5 Gene Product Regulates Synthesis of Early Viral Messenger RNAs
Cell 17: 935-944

Control of Abundance of Immediate-Early mRNA in Herpesvirus (Pseudorabies)-Infected Cells
Virology 116: 250-263
Activation of Early Adenovirus Transcription by the
Herpesvirus Immediate Early Gene: Evidence for a
Common Cellular Control Factor

Mechanism of Activation of Early Viral
Transcription by the Adenovirus E1A Gene Product
Cell 26: 213-220

Induction of the Synthesis of a 70,000 Dalton
Mammalian Heat-Shock Protein by the E1A Gene
Product
Cell 29: 913-919

Cis and Trans Activation of Globin Gene Transcrip­
tion in Transient Assays

[186] Graham, F.L., Smiley, J., Russel, W.C., and Nairn,
R. (1977)
Characteristics of a Human Cell Line Transformed by
DNA from Human Adenovirus Type 5

HEL Cells: A New Human Erythroleukaemia Cell Line
with Spontaneous and Induced Globin Expression
Science 216: 1233-1235

Differential Induction of Adult and Foetal Globin
Genes in a Human Erythroleukaemic Cell Line
Abstract
Abstract from: Proceedings of the Fourth Conference
on Haemoglobin Switching
A β-Globin Gene, Inactive in the K562 Leukaemic Cell, Functions Normally in a Heterologous Expression System
Proc. Natl. Acad. Sci. USA 81: 4485-4489

Erythroleukaemia (K562) Cells Contain a Functional β-Globin Gene
Mol. Cell. Biol. 4: 2553-2555

β+ Thalassaemia: Aberrant Splicing Results from a Single Point Mutation in an Intron
Cell 27: 289-298

K562 Human Leukaemic Cells Synthesise Embryonic Haemoglobin in Response to Haemin
Nature 280: 164-165

Differential Expression of the Globin Genes in Human Leukaemia K562(s) Cells Induced to Differentiate by Haemin or Butyric Acid
Cancer Res. 41: 237-243

The "β-Like-Globin" Gene Domain in Human Erythroid Cells
Proc. Natl. Acad. Sci. USA 82:6384-6388


Gene Transfer into Mouse Lyoma Cells by Electroporation in High Electric Fields 
EMBO. J. 1: 841-845

Electric Field-Mediated DNA Transfer: Transient and Stable Gene Expression in Human and Mouse Lymphoid Cells 

Use of Electroporation to Introduce Biologically Active Foreign Genes into Primary Rat Hepatocytes 

Transcription of Xenopus tDNA and Sea-Urchin Histone tDNA\^\text{\textsuperscript{met}} Injected into the Xenopus Oocyte Nucleus 

Transgenic Mice 
Cell 41: 343-345

Recombinant Genomes Which Express Chloramphenicol Acetyltransferase in Mammalian Cells 
Mol. Cell. Biol. 2: 1044-1051

Expression of a Bacterial Gene in Mammalian Cells 
Science 209: 1422-1427


Mouse Cells Transformed by Bovine Papillomavirus 
Contain Only Extrachromosomal Viral DNA Sequences 
Proc. Natl. Acad. Sci. USA 78: 2727-2731

[216] Wigler, M., Sweet, R., Sim, G.K., Wold, B., 
Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979) 
Transformation of Mammalian Cells with Genes from 
Procaryotes and Eucaryotes 
Cell 16: 777-785

Urlaub, G., and Chasin, L. (1979) 
DNA-Mediated Transfer of the Adenine Phosphoribo-
syltransferase Locus into Mammalian Cells 
Proc. Natl. Acad. Sci. USA 76: 1373-1376

[218] Jonge, A.J.R.D., Abrahams, P.J., Westerveld, A., 
Expression of Human HPRT Gene on the Inactive X 
Chromosome After DNA-Mediated Gene Transfer 
Nature 295: 624-626

(1982) 
Efficient Expression of Escherichia coli Galacto-
kinase Gene in Mammalian Cells (1982) 

[220] Murray, M.J., Kaufman, R.J., Latt, S.A. and 
Construction and Use of a Dominant, Selectable 
Marker: A Harvey Sarcoma Virus - Dihydrofolate 
Reductase Chimera 
Deletion of Thymidine Kinase Activity from L Cells
Resistant to Bromodeoxyuridine

Expression of Exogenous DNA in Mammalian Cells
In vitro Transcription and Translation - A
Practical Approach (Hames, B.D., and Higgins, S.J., eds)
IRL Press, Oxford, pp 1-48

Labelling Deoxyribonucleic Acid to High Specific
Activity In Vitro by Nick Translation with DNA
Polymerase I
J. Mol. Biol. 113: 237-251

Efficient Transfer of Large DNA Fragments from
Agarose Gels to Diazobenzloxymethal-Paper and Rapid
Hybridisation by Using Dextran Sulphate
Proc. Natl. Acad. Sci. USA 76: 3683-3687

A Membrane-Filter Technique for the Detection of
Complementary DNA

[226] Wilson, J.T., Wilson, L.B., de Reil, J.K., Villa
Komaroff, L., Efstratiadis, A., Forget, B.G., and
Weissman, S.M. (1978)
Insertion of Synthetic Copies of Human Globin Genes
into Bacterial Plasmids
Nucleic. Acid. Res. 5: 563-581
Structure of Recombinant Plasmids Containing Synthetic Human Foetal Globin Gene Sequences

Screening λgt Recombinant Clones by Hybridisation to Single Plaques In Situ
Science 196: 180-182

Clarke, L., and Carbon, J. (1976)
A Colony Bank Containing Synthetic Col E1 Hybrid Plasmids Representative of the Entire E. coli Genome
Cell 9: 91-99

Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L.A.,
Charon Phages: Safer derivatives of Bacteriophage Lambda for DNA Cloning
Science 196: 161-169

Rapid Bacteriophage Sedimentation in the Presence of Polyethylene Glycol and its Application to Large-Scale Virus Purification
Virology 40: 734-744

Lang, J.C. (1983)
Characterisation of Eucaryotic Transcriptional Control Signals By Assay of Herpes Simplex Virus Type 1 Thymidine Kinase
Ph.D. Thesis, Beatston Institute, Glasgow

Ligation of Eco RI Endonuclease Generated DNA Fragments into Linear and Circular Structures
J. Mol. Biol. 96: 171-184
Chemical Synthesis of Restriction Enzyme Recognition Sites Useful for Cloning
Science 196: 177-180

Plasmids of Escherichia coli as Cloning Vectors
Methods Enzymol 68: 245-267

Calcium Dependent Bacteriophage DNA Infection
J. Mol. Biol. 53: 154-

A Rapid Alkaline Extraction Procedure for Screening Recombinant Plasmid DNA
Nucleic. Acid. Res. 7: 1513-1523

Molecular Cloning - A Laboratory Manual
Maniatis, T., Fritsch, E.F., and Sambrook, J.
Cold Spring Harbour

Extracellular Nucleases of Pseudomonas Bal 31 III
Use of the Double-Strand Deoxyriboexonuclease Activity as the Basis of a Convenient Method for the Mapping of Fragments of DNA Produced by Cleavage with Restriction Enzymes
Nucleic. Acid. Res. 5: 1445-1464

A New Method for Sequencing DNA
Sequencing End-Labelled DNA with Base-Specific Chemical Cleavages  
Methods Enzymol 65: 499-560

A Restriction Enzyme Thai from the Thermophilic Mycoplasma Thermoplasma acidophilum  
Nucleic. Acid. Res. 5: 1729-1740

Induction of Both Thymidine and Deoxycytidine Kinase Activity by Herpes Viruses  

Bacchetti, S., and Graham, F.L. (1977)  
Transfer of the Gene for Thymidine Kinase to Thymidine Kinase-Deficient Human Cells by Purified Herpes Simplex Viral DNA  

Non-Producer Human Cells Induced by Murine Sarcoma Virus  
Int. J. Cancer 15: 23-29

Markers of Neoplastic Transformation in Epithelial Cell Lines Derived from Human Carcinomas  

Transfer of Purified Herpes Virus Thymidine Kinase Gene to Cultured Mouse Cells.  
Cell 11: 223-232
Metabolic Co-operation Between Cells
Int. Rev. Cytol. 69: 45-104

[249] Gross-Bellard, M.l, Oudet, P. and Chambon, P.
(1973)
Isolation of High Molecular Weight DNA from
Mammalian Cells

Detection of Specific Sequences Among DNA Fragments
Separated by Gel Electrophoresis
J. Mol. Biol. 98: 503-517

[251] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J.l,
and Rutter, W. (1979)
Isolation of Biologically Active Ribonucleic Acid
from Sources Enriched in Ribonuclease
Biochemistry 18: 5294-5299

Purification of Biologically Active Globin
Messenger RNA by Chromatography on Oligothymidylic
Acid-Cellulose

[253] Spandidos, D.A., Harrison, P.R., and Paul, J.
(1981)
Transfer and Expression of Herpes Simplex Virus
Thymidine Kinase and Human Globin Genes in
Mammalian Cells Studied by Spot Hybridisation
Assays
Bioscience. Reports. 1: 911-920

Sizing and Mapping of Early Adenovirus mRNAs by Gel
Electrophoresis of S1 Endonuclease-Digested Hybrids
Cell 12: 721-732

- 278 -
Mapping of RNA by a Modification of the Berk-Sharp Procedure: the 5' Termini of 15S β-Globin mRNA Precursor and Mature 10S β-Globin mRNA have Identical Map Co-ordinates
Nucleic. Acid. Res. 6: 1175-1193

Transcription of Chromatin
In Vitro Transcription and Translation - A Practical Approach
(Hames, B.D., and Higgins, S.J., Eds)
IRL Press, Oxford, pp 142-145

Trans-Complementable Copy Number Mutants of Plasmid Col E1
Nature 283: 216-218

Colony Hybridisation: A Method for the Isolation of Cloned DNAs that Contain a Specific Gene

Rabbit β-Globin mRNA Production in Mouse L Cells Transformed with Cloned Rabbit β-Globin Chromosomal DNA
Nature 281: 40-46

Cloning Specific Segments of the Mammalian Genome: Bacteriophage λ Containing Mouse Globin and Surrounding Gene Sequences
Far Upstream Initiation Sites for Adenovirus Early Region 1A Transcription are Utilised After the Onset of Viral DNA Replication
J. Virol 45: 594-599

Differences in Human α-, β- and δ- Globin Gene Expression in Monkey Kidney Cells
Cell 30: 173-183

Transcription Unit of the Rabbit β1 Globin Gene

Attenuation in the Control of Expression of Bacterial Operons
Nature 289: 751-758

Tissue Specific Transcription of the Human ε-Globin Gene Following Transfection into the Embryonic Cell Line K562

α- Amanitin Insensitive Transcription of the Human ε- Globin Gene

In Vitro Expression of the Human ε-Globin Gene
Ph.D. Thesis. Beatson Institute, Glasgow
DNA Sequences Preceding the Rabbit β-Globin Gene are Required for Formation in Mouse L Cells of β-Globin RNA with Correct 5' Terminus
Proc. Natl. Acad. Sci. USA 78: 1411-1415

Distant Sequences Which Regulate Globin Genes
Experimental Approaches for the Study of Haemoglobin Switching pp 29-48 (eds. Stamatoyannopoulos, G., and Nienhuis, A.W.)

Multiple Elements Associated with the Human ε-Globin Gene
Edin. Int. Conference on Differentiation

DNase Footprinting: A Simple Method for the Detection of Protein - DNA Binding Specificity
Nucleic Acid Res. 5: 3157-3170

A Tissue-Specific Transcriptional Enhancer Element in the Human Immunoglobulin λ Light Chain Locus
FEBS. Letts. 175: 152-158

Cis-Acting Induction of Adenovirus Transcription
Cell 33: 683-693

[274] Young, K., Donovan-Peluso, M., Bloom, K., Allan, M., Paul, J., and Bank, A.
Stable Transfer and Expression of Exogenous Human Globin Genes in Human Erythroleukaemia (K562) Cells


Insertion of DNA Sequences into the Human Chromosomal β-Globin Locus by Homologous Recombination
Nature 317: 230-234

Yeast Transformation: A Model System for the Study of Recombination

Purification of Genomic Sequences from Bacteriophage Libraries by Recombination and Selection In Vivo
Nucleic. Acid. Res. 11: 2427-2445

An Embryonic Pattern of Expression of a Human Foetal Globin Gene in Transgenic Mice
Nature 319: 685-689