The Structure and Expression of the Murine Glutathione Peroxidase Gene.

Ian Paul Chambers ©

Submitted for the degree of Doctor of Philosophy in the Faculty of Medicine, University of Glasgow.

The Beatson Institute for Cancer Research, Glasgow.

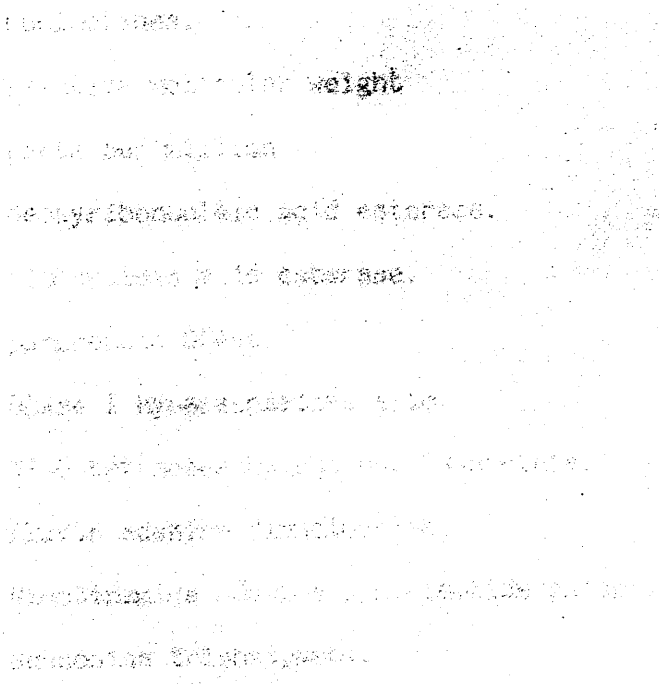
August 1987

For Mum, Dad and all the family.

CONTENTS

List of Figures.

List of Tables.



Abbreviations.

DNA	deoxyribonucleic acid.
RNA	ribonucleic acid.
GSH	glutathione.
GSSG	oxidised glutathione disulphide.
GSHPx	glutathione peroxidase.
fdhF	formate dehydrogenase selenopeptide.
mRNA	messenger RNA.
pre-mRNA	mRNA precursor.
tRNA	transfer RNA.
rRNA	ribosomal RNA.
RNP	ribonucleoprotein.
bp	base pairs.
kb	kilobase pairs.
kD	kilodaltons.
nt	nucleotides.
Mr	relative molecular weight
ppm	parts per million
DNAse	deoxyribonucleic acid esterase.
RNAse	ribonucleic acid esterase.
DNAse I	pancreatic DNAse.
DH site	DNAse I hypersensitive site.
TPA	12-O-tetradecanoylphorbol 13-acetate.
FAD	flavin adenine dinucleotide.
NADPH	Nicotinamide adenine dinucleotide phosphate.
ATP	adenosine triphosphate.

c.AMP	cyclic adenosine monophosphate.
IL-3	Interleukin 3.
epo	erythropoietin.
CFU-S	colony forming unit (spleen)
CFC-Mix	mixed colony forming cell.
GM-CFC	granulocyte/macrophage colony forming cell.
BFU-E	erythroid burst forming unit.
CFU-E	erythroid colony forming unit.
LTBM	long term bone marrow.
HMBA	hexamethylene bis-acetamide.
DMSO	dimethylsulphoxide.
A	Adenine.
C	Cytosine.
G	Guanine.
T	Thymine.
U	Uracil.
Pu	A or G.
Py	C or T.
R	A or T.
S	C or G.
V	A or C.
W	G or T.
N	A, C, T or G.

Acknowledgements.

I first of all wish to thank my supervisor, Paul Harrison, for his consistent support and encouragement throughout the period of this research and also for many stimulating discussions on both the work and wider aspects of bioscience. I am also grateful to Saveria Campo for providing a fresh invigorating critique of the work at regular intervals. Thanks are also due to Jon Frampton for his friendly practical support during the most barren stages of this work. My gratitude also extends generally to all the members of the Beatson and particularly to the members of the Harrison group for providing a pleasant atmosphere in which to work. The assistance of Margaret Blair in transcribing my scrawl into a form of English is also appreciated.

The financial support of the Medical Research Council for providing an essential studentship and Imperial Chemical Industries p.l.c. for providing a supplementary scholarship during both my post-graduate and undergraduate years is gratefully acknowledged.

I wish to take this opportunity to thank my parents for their constant loving support and encouragement : without their help this thesis would not exist. A special thank you also goes to Eleri Finnerty for her patient understanding and assistance in the production of this document : the final stages of this work were made the more bearable by her affectionate support.

I bet Indiana Jones didn't have this hassle getting his doctorate.

I.C. 17th August, 1987.

Summary

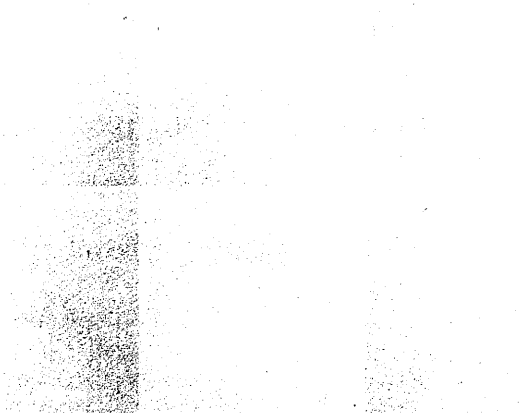
The literature relating to the cell biology of the murine haemopoietic system, particularly the erythropoietic lineage has been reviewed, as have current ideas relevant to the regulation of cell-specific gene transcription.

A series of experiments analysing the structure and expression of the murine gene encoding a polypeptide of approximately 19kD present in reticulocytes is then presented. These show that the gene encodes the selenium-dependent enzyme glutathione peroxidase (GSHPx) which contains the unusual amino acid selenocysteine in its polypeptide backbone. This amino acid has been shown by sequencing of the GSHPx gene and mRNA to be encoded by the "termination" codon, UGA. The significance of this finding is discussed.

The GSHPx mRNA is detectable in every cell type and tissue examined, although at varying levels. Thus in erythroid cells, liver and kidney, the GSHPx mRNA level is 20-50 fold higher than in other cells. These differences have been shown not to be due to nuclear retention of the mRNA or pre-mRNA in low expressing cells. Furthermore, the levels of pre-mRNA in a low and a high expressing cell line vary in the same way as the mRNA. It therefore seems that the different levels of GSHPx mRNA arise through differences either in the stability of the GSHPx mRNA or in the rate of transcription of the GSHPx gene in different tissues. Indirect support for the involvement of the latter mechanism has come from an analysis of the chromatin structure of the GSHPx gene in tissues expressing the mRNA at a high or low

level. Thus a broad DNAse I hypersensitive site is seen downstream from the gene. In addition, two erythroid specific DNAse I hypersensitive sites are also located 3' to the GSHPx gene.

Functional studies, involving short term transfection assays using fragments of the GSHPx gene linked to the chloramphenicol acetyltransferase (CAT) gene, showed that the region of DNA immediately upstream of the GSHPx gene transcription initiation site functions equally well as a promoter of transcription in both a low expressing and a high expressing cell line. In addition, a DNA fragment from the 3' end of the GSHPx gene, shown by sequence analysis to contain elements homologous to *cis*-acting viral and cellular DNA transcription regulatory sequences, did not have any effect upon the transcription from the GSHPx or other promoters when linked in *cis* to them.



Chapter 1

Introduction

1.1. Murine haemopoiesis

Differentiated cells within an adult eukaryote are continually dying and being replaced by differentiation of stem cell populations which also possess the ability to reproduce themselves (self-renewal). One of the best characterized differentiation systems is that forming the blood cells (reviewed in Till and McCulloch, 1980). Progress in this field has been steady since the development of the first assay for pluripotent stem cells, the spleen colony forming unit (CFU-S) assay (Till and McCulloch, 1961). By this method, mice which have been subjected to a lethal dose of radiation in order to destroy their haemopoietic stem cells are rescued by injection of haemopoietic cells from donor mice. Various types of colonies are seen in the spleens of the recipients; early colonies (7-8 days post-injection) consist entirely of differentiated erythroid cells and contain no bipotent precursor cells, whereas late colonies (10-14 days) consist of mixed mature myeloid cells and contain precursor cells with the potentiality of several lineages (Magli *et al*, 1982). Moreover, late colonies contain cells capable of forming spleen colonies in secondary irradiated recipients and the proportion of colonies containing such cells increases from day 10 to day 14 (Siminovitch *et al*, 1963).

Examination of the cellular composition of spleen colonies, together with *in vitro* culture methods used to examine the growth factor requirements and developmental potential of haemopoietic progenitor cells, has allowed lineage diagrams, such as Figure 1.1 to be drawn (Metcalf, 1981). These *in vitro* methods allow

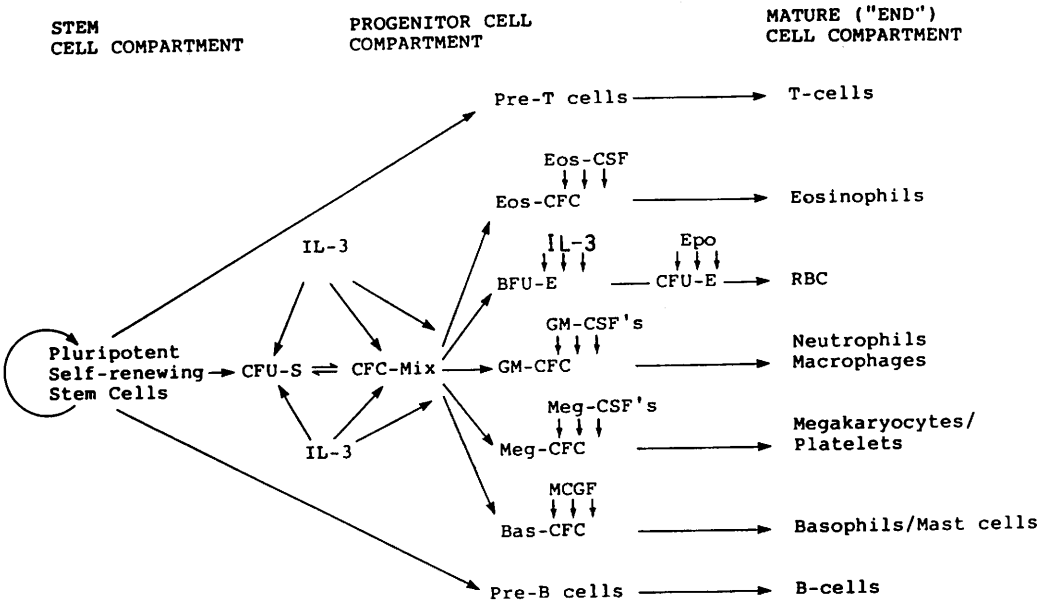


Figure 1.1. Scheme of haemopoiesis showing the various progenitor and committed cells and their growth factor requirements. Abbreviations: CFU-S, spleen colony forming unit; CFC-Mix, mixed colony forming cell; BFU-E, erythroid burst forming unit; GM-CFC, granulocyte/macrophage colony forming cell; CFU-E, erythroid colony forming unit; IL-3, interleukin 3; Epo, erythropoietin; GM-CSF, granulocyte/macrophage colony stimulating factor; MCGF, mast cell growth factor (adapted from Moore and Dexter, 1986).

haemopoietic progenitor cells to survive, proliferate and differentiate in semi-solid media provided that the culture is supplemented with the correct growth factor(s) (Metcalf, 1981). For example, this type of assay has identified a cell population in animals and humans termed the mixed colony-forming cells (CFC-Mix) since cells of all myeloid lineages (and perhaps also lymphoid lineages, although this remains controversial [Messner et al, 1982]) can be found in colonies derived from these cells. Because such macroscopic colonies of mixed lineages sometimes contain more CFC-Mix (as judged by replating experiments) some CFC-Mix must have a self-renewal capacity; in this respect they overlap with the in vivo operationally defined CFU-S, although most likely both labels relate to a heterogenous cell population with a range of potentialities (Humphries et al, 1981; Johnson et al, 1982). Because of this overlap between different operationally defined cell types, lineage diagrams typified by Figure 1.1 should be regarded as generalizations of cellular behaviour until a more rigorous molecular characterization of the illustrated cells is available.

The growth factors required in these assays can be purified from media conditioned by the growth of a variety of normal and leukaemic cell lines as well as from activated T-lymphocytes. Some of these growth factors have been purified to homogeneity and corresponding cDNAs cloned (reviewed in Metcalf, 1985; Whetton and Dexter, 1986). This has been, and will continue to be, important in order to assign a particular biological effect to a given growth factor, rather than to a low, but significant,

level of a contaminating factor in the growth factor preparation. Activities have been established for several growth factors, for example, interleukin 3 (IL-3) is lineage-indifferent, promoting survival, proliferation and development of all types of myeloid progenitor cells. In contrast, erythropoietin (epo) exerts its effects only upon committed erythroid cells (see next section). IL-3 influences cell lines absolutely dependent upon it for their survival by stimulating glucose transport into the cell and thereby maintaining intracellular ATP levels (Whetton *et al*, 1984). However, the *in vivo* role of IL-3 is unclear since it has not been detected *in vivo* or in long term bone marrow cultures supporting haemopoiesis (Dexter and Shadduck, 1980; Shadduck *et al*, 1983). Moreover, the low degree of conservation of the primary amino acid sequence of IL-3 between rats and mice (54% homology) argues against a conserved and irreplaceably important function for this molecule (Cohen *et al*, 1986).

The most controversial aspect of haemopoiesis concerns the mechanisms controlling the decision of a stem cell to become committed to a particular lineage. One model sees the haemopoietic microenvironment as having a deterministic role in this choice (Trentin, 1970), whereas a contrasting model views the commitment event as occurring stochastically, with only the probability of such an event being externally influenced (Till *et al*, 1964). However, the accum ulated evidence suggests that the commitment event occurs stochastically, with the microenvironment influencing the proliferation (or lack of proliferation) of particular types of committed precursor cells. For example, the

relative numbers of early granulocyte/macrophage (GM-CFC) and erythroid (BFU-E) committed precursors are strongly correlated in spleen colonies, irrespective of the overall composition of mature cells in the colonies (Gregory and Henkelman, 1977).

Haemopoiesis can also be encouraged in long term bone marrow (LTBM) cultures in which pluripotent stem cells can be maintained for many months (Dexter and Lajtha, 1974). Adherent cell layers, composed of similar kinds of stromal cells as are present *in vivo* (fat cells, reticular cells, endothelial cells and macrophages), are required for the maintenance of these cultures (Dexter *et al*, 1984). In addition, the pluripotent stem cells present in these cultures behave in the same way as CFU-S both *in vivo* and *in vitro* (Schofield and Dexter, 1985). When fresh bone marrow CFU-S are serially transplanted into irradiated recipients, there is a rapid decline in both the self-renewal capacity of the cells and their ability to rescue mice from death between the second and third transfers. When CFU-S present in the stromal layer of LTBM cultures are injected into irradiated mice the same effect is seen. Moreover, replating of a second aliquot of the same cells upon a secondary irradiated marrow stromal layer re-establishes haemopoiesis but only at a sub-optimal level and the CFU-S present have a poor self-renewal ability. Attempts to transfer these CFU-S onto a second irradiated stromal layer have not been successful. These results may be due to the loss or dilution of an "accessory" cell required for the maintainence of proliferation and self-renewal of the CFU-S or to the finite proliferation potential of individual CFU-S.

The behaviour of the CFU-S in LTBM cultures has been perturbed by infection of the cultures with a Moloney murine leukaemia virus-*src* [*src* (MoMuLV)] recombinant virus (Boettiger *et al*, 1984). Following infection, the number of mature cells is decreased and the number of CFU-S is increased dramatically. These CFU-S are not leukaemogenic and most surprisingly their self-renewal capacity is increased as judged by (1) their ability to protect mice from potentially lethal radiation over at least five serial transfers, (2) their ability to be serially passaged onto irradiated marrow stromal cells apparently indefinitely and (3) their ability to form mixed myeloid colonies *in vitro* that contain large numbers of CFC-Mix in response to IL-3 but not macrophage progenitor specific growth factor, macrophage colony stimulating factor (M-CSF) (Spooncer *et al*, 1984). Therefore infection of LTBM cultures produces an intrinsic alteration in the CFU-S such that they no longer require the presence of any putative "accessory" cell for maintenance of their self-renewal capacity. Moreover, cloned stem cell lines can be obtained from these infected cultures that can grow independently of stromal cells provided IL-3 is added to the cultures (Spooncer *et al*, 1986). The CFU-S in the infected cultures do not contain integrated copies of the *src* (MoMuLV) DNA : rather the intrinsic change they have undergone seems to have been mediated by interaction with a stromal cell whose activity has been altered by *src* integration and expression (Wyke *et al*, 1986). In fact, it is not inconceivable that the stochastic element in stem cell decision making is the probability that a CFU-S maintains its

contact with the putative "accessory" cell.

1.2 The erythropoietic lineage.

1.2.1. Normal erythropoiesis.

During development of the mouse, the anatomical site of erythropoiesis changes (Russell, 1979). The earliest erythroid cells are produced in the blood islands of the yolk sac between days 7 and 11 of gestation. These "primitive" nucleated erythrocytes are released into the embryonic circulation where they persist until day 15 -16. At about day 12 the site of erythropoiesis switches to the fetal liver, where smaller enucleated erythrocytes are produced until birth. A further switch in the site of erythropoiesis occurs at about day 16 when erythropoiesis begins in the spleen and the bone marrow. Splenic erythropoiesis is transient, ceasing at about birth, whereas bone marrow erythropoiesis continues throughout adult life. Under conditions of anaemic stress the adult spleen and liver can provide the appropriate microenvironmental niches for the expanded erythroid differentiation required to meet the increased demand for erythrocytes (Conkie *et al*, 1975).

The growth factor requirements of cells traversing the erythroid lineage have been reviewed by Harrison (1982). Briefly stated, bone marrow cells cultured *in vitro* in the presence of IL-3 and erythropoietin (epo) form colonies 10 - 15 days later containing cells of all the myeloid lineages derived from mixed colony forming cells (CFC-Mix; see previous section). In response to IL-3, CFC-Mix give rise to committed erythroid cells,

the so-called burst forming units-erythroid (BFU-E), which in turn can form large colonies containing 10^3 - 10^4 mature erythroblasts after 7 - 9 days in the presence of either high concentrations of epo or IL-3 plus low concentrations of epo (see Figure 1.1). BFU-E proliferate and differentiate in response to IL-3 to produce CFU-E (colony forming units - erythroid) which can be recognized by their ability to produce small colonies of appoximately 32 mature erythroblasts after only 2 days of culturing bone marrow cells in the presence of low concentrations of epo alone. The differentiation of BFU-E to CFU-E is characterized by a decreasing sensitivity to IL-3 coupled to an increasing sensitivity to epo (Eaves et al, 1979). This again illustrates the point made previously concerning overlap between different operationally defined cell populations. Recent work with a heterogeneous population of BFU-E-like cells has shown that these cells have an average of 500 epo-receptors on their surface, only 1-2% of which need bind epo in order to elicit a response (Krantz and Goldwasser, 1984). However, it is not known if either receptor number or receptor occupancy changes during maturation of these cells.

Erythropoiesis can also be stimulated in LTBM culture. Normally in these cultures, horse serum is used, the adherent layer is composed of macrophages, endothelial cells and fat cells and granulopoiesis is favoured (Testa and Dexter, 1977). Under these conditions BFU-E are present, although they do not differentiate further. Maturation to the CFU-E stage can be achieved by replacing the horse serum with foetal calf serum

supplemented with epo. Further maturation beyond the CFU-E stage can be achieved by mechanical agitation (Eliason *et al*, 1979) or by replacing the serum with either anaemic mouse serum or normal mouse serum plus epo (Dexter *et al*, 1981). An intriguing point of note here is that under conditions supporting erythropoiesis the areas of fat and endothelial cells normally seen in the stromal cell layer regress and macrophages predominate. In fact, electron microscopy has consistently shown developing erythroid cells present as a rosette around a central macrophage-like cell whose cytoplasmic membrane is extruded in interdigitating processes around the surrounding erythroid cells both *in vivo* and *in vitro* (Weiss, 1984; Allen and Dexter, 1983). Indeed a central role in regulating erythropoiesis has been suggested for the macrophage, not least because it produces epo (Rich and Kubanek, 1985).

As the CFU-E matures characteristic alterations occur in the cell (see Figure 1.2) : the membrane composition changes, haemoglobin accumulates, nuclear condensation and extrusion occur and mitochondria disappear. It should be stressed that all the processes described up till now occur in the erythropoietic organ (bone marrow, foetal liver or spleen); following nuclear extrusion the reticulocyte crosses the sinus wall and enters the circulation (some impressive electron photomicrographs of these processes can be seen in Weiss, 1984).

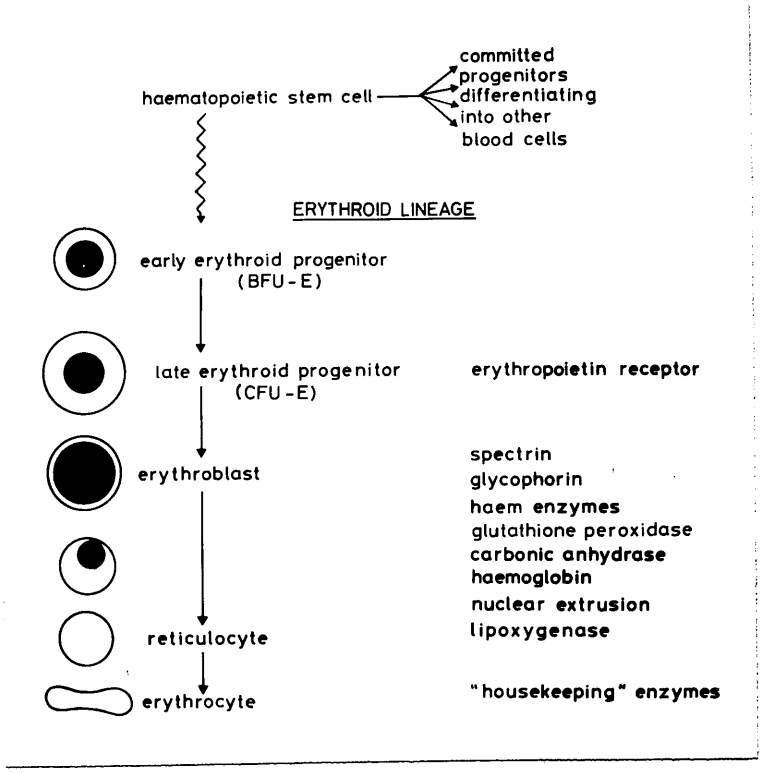


Figure 1.2. Morphological representation of the maturing erythroid cell indicating the presence of the erythropoietin receptor on BFU-E and CFU-E and the acquisition of differentiated red cell proteins in the post-CFU-E.

1.2.2. Erythropoiesis in murine erythroleukaemia cell lines.

Several cell lines have been established from tumours induced by infection of mice with the Friend virus complex. The aetiology of the diseases induced by these viruses and the derivation of cell lines has been reviewed (Harrison, 1977; 1982; Marks and Rifkind; 1978) and will not be considered here. Since foetal or adult erythroid organs comprise erythroid cells at various stages of differentiation, normal erythroblasts are difficult to purify in bulk. In contrast, Friend cells are readily available and represent a homogeneous population of erythroblast-like cells. These cells are malignant and their growth control is abberant, since they can survive and proliferate in culture in the absence of epo. Nevertheless they represent a useful *in vitro* cell population since after treatment with a number of chemicals such as hexamethylenebisacetamide and butyrate they undergo the same morphological and biochemical changes characteristic of normal erythroid maturation (Friend *et al*, 1971; these phenotypic alterations will be discussed in more detail in the next section). Indeed under optimal conditions of induction Friend cells can even extrude their nuclei, forming unstable reticulocyte-like cells (Volloch and Housman, 1981). The mechanism of action of these inducing chemicals is unknown, but two separate classes of inducer have been discerned on the basis of whether they act via an effect upon polyamine biosynthesis (Gazitt and Friend, 1980).

Friend cell differentiation can be resolved into two distinct processes; the first of these is a commitment process

which requires the presence of inducer, whereas the second, maturation, can occur in the absence of inducer. The commitment process includes a lag period during which irreversible cell commitment does not occur and a subsequent period in which individual cells become irreversibly committed to erythroid maturation in a stochastic fashion (Gusella et al, 1976). (It is important to emphasis that commitment in this context means the irreversible decision to erythroid-like maturation rather than the stable proliferation as proerythroblastic cells and is therefore distinct from the process of commitment of a normal pluripotent haemopoietic stem cell : any parallels which do exist between these two processes remain obscure.) The lag period of inducer treatment can be interrupted for up to 18h of inducer withdrawal without affecting the total time of inducer treatment required for initiation of commitment (Levenson and Housman, 1979a). Two separate processes occur during the lag period which require the presence of inducer. One of these, which is rate-limiting and requires protein synthesis (Levenson and Housman, 1979b) is the gradual accumulation of calcium ions (Levenson et al, 1980; Bridges et al, 1981). This requirement for protein synthesis is interesting in view of the recent demonstration that proteinaceous factor(s) from induced Friend cells can cause erythroid differentiation when introduced into Friend cells pretreated for only 6h with inducer (Nomura et al, 1986). The second process, which is not rate limiting is inhibited by sub-cytotoxic levels of 2'-deoxyadenosine, perhaps through an effect upon mRNA polyadenylation (Levenson et al,

1979). How a recently identified process which occurs early during the lag phase and can be inhibited by lithium ions fits into the above scheme is presently unclear (Zaricznyj and Macara, 1987).

Because Friend cell commitment occurs stochastically, an induced culture rapidly loses one of its most attractive features; its homogeneity. However, maturation of Friend cells can also be blocked following their commitment, to yield increasingly homogeneous, more mature populations of erythroid-like cells by the addition to the cultures of phorbol esters (Fibach et al, 1979) or dexamethasone (Santoro et al, 1978). The mechanisms by which these substances act is unknown, although it is possible that phorbol esters act via an effect upon protein kinase C (Nishizuka, 1984) whereas dexamethasone might, by binding to its receptor affect a crucial transcription unit in such a way that the result is incompatible with further erythroid maturation : the latter case might then be analogous to the way in which the avian erythroblastosis virus oncogene v-erb A (which is a member of the same gene family as the steroid receptors) blocks erythroid maturation (Bishop, 1986; Green and Chambon, 1986).

1.2.3. Phenotypic alterations occurring during erythropoiesis.

The phenotypic changes accompanying erythroid maturation of normal and erythroleukaemic cells are summarised in Figure 1.2. These have been discussed elsewhere (Harrison, 1984) and shall only briefly be reviewed here.

i. Early markers.

The earliest known molecular marker of a committed erythroid cell is the erythropoietin receptor, which is present on the BFU-E (Krantz and Goldwasser, 1984) and is probably absent from non-erythroid cells. As maturation proceeds other molecules accumulate; spectrin (a heterodimeric molecule forming part of the cytoskeleton) appears early in erythroblast maturation (Eisen, 1978) and the transmembrane sialoglycoproteins, the glycophorins appear in the post CFU-E with their synthesis reaching a peak at the mid-erythroblast stage (Kasturi and Harrison, 1985), that is, earlier than haemoglobin accumulation. The appearance of globin mRNA occurs normally at the basophilic erythroblast stage (Harrison et al, 1974) although under conditions of high erythropoietin, globin mRNA can be detected at the earlier proerythroblast stage (Conkie et al, 1975). In fact, there is even a temporal difference in the appearance of the different globin mRNAs, with α-globin mRNA being induced earlier during maturation than β-globin mRNA (Orkin et al, 1975).

ii. Enzymes of the haem biosynthetic pathway.

In order to supply sufficient haem for the globin chains being synthesized, the activities of the enzymes of the haem biosynthetic pathway must increase from their normal low levels required for cytochrome production. However, these activities do not increase synchronously, rather each activity in the pathway is induced sequentially (Sassa et al, 1978). Indeed, induction of the final enzyme in the pathway (ferrochelatase) has been suggested to mark the Rubicon in Friend cell maturation (Eisen et

al, 1978). This idea derives from studies of variant Friend cell clones which although uninducible for terminal maturation by DMSO nevertheless accumulate spectrin and other "early" markers when cultured in the presence of DMSO. This induction of the "early" erythroid programme is reversible since returning such DMSO treated cells to culture without DMSO causes reversion to the uninduced phenotype. If, on the other hand, haemin is added to such cultures then the terminal stages of erythroid maturation occur (Eisen et al, 1978). Similar phenotypes have been observed in cell hybrids formed between "normal" Friend cells and T-lymphoma cells (Harrison et al, 1977) and in DMSO-resistant Friend cell clones (Harrison et al, 1978). In the case of one of these variants (Fw), the defect was shown to result from a lack of inducibility of ferrochelatase (Rutherford and Harrison, 1979). Once again, treatment of these cells with DMSO plus haemin induces the terminal stages of erythroid maturation.

iii. Late events.

As well as haemoglobin formation the later events in the maturation programme include chromatin condensation and nuclear extrusion, the disappearance of mitochondria and ribosomes and the appearance of the typical biconcave disc morphology. The enzyme responsible for mitochondrial lysis, an apparently erythroid specific lipoxygenase becomes induced during reticulocyte maturation (Rapoport et al, 1979). A cDNA encoding this enzyme has recently been cloned (Thiele et al, 1987).

iv. GSHPx and other enzymes involved in oxidative defense.

In addition to these functions, RBCs contain a higher level

than normal of enzymes responsible for cellular protection against oxidative damage (superoxide dismutase, catalase and glutathione peroxidase). The conditions necessary for initiating peroxidative damage are nowhere better fulfilled than in the RBC; the cell is exposed to high oxygen tensions, there is an ample supply of transition metal catalysts for the formation of superoxide and there is a high degree of unsaturation in the lipid membrane (Chiu et al, 1982). Indeed, approximately 3% of the haemoglobin in the human body is oxidised to methaemoglobin each day. This involves the conversion of the ferrous iron to the ferric ion and concomitantly dioxygen is converted to the superoxide radical (Figure 1.3). This in itself is not thought to be particularly harmful, but through reaction with hydrogen peroxide, the superoxide radical can generate singlet oxygen and hydroxy radicals, both of which are extremely reactive and cytotoxic entities which may initiate free radical chain reactions in the lipid membrane leading to proteolytic damage (Chiu et al, 1982; Wolff et al, 1986). This is something which is particularly serious in the RBC, in view of its negligible capacity for protein synthesis; for this reason it is thought that the levels of superoxide dismutase, glutathione peroxidase and catalase are elevated in RBCs.

It was, in fact, within the RBC that glutathione peroxidase (GSHPx) activity was first detected (Mills, 1957), where it was demonstrated to catalyse the reaction:

$$ROOH + 2\ GSH \longrightarrow ROH + GSSG + H_2O$$

(where ROOH can be virtually any hydroperoxide). The enzyme is a

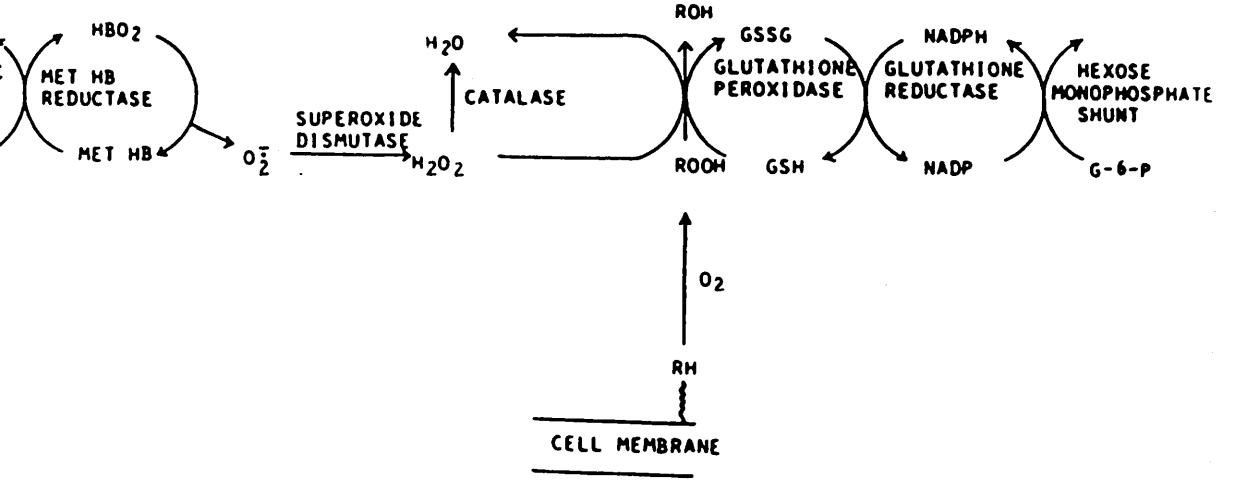


Figure 1.3. Oxy-protective mechanisms of the red cell (adapted from Chiu et al 1982).

homotetramer, with a subunit molecular weight of approximately 21kD (Flohe, 1982), each subunit containing an atom of the group VI element selenium (Rotruck et al, 1973; Flohe et al, 1973). The selenium has since been shown to be present in the enzyme as a residue of the unusual amino acid selenocysteine (a selenium substituted analogue of cysteine; Forstrom et al, 1978) which is inserted into the polypeptide backbone during translation (Hawkes and Tappel, 1983). A selenocysteyl-tRNA has also been identified, although neither its codon recognition nor its sequence have been determined (Hawkes et al, 1982).

Although glutathione peroxidase is known to be present at elevated levels in RBCs compared to other cell types (Tappel et al, 1982) it is not known how the level of the enzyme changes during erythropoiesis nor whether any changes which might occur parallel changes in the levels of superoxide dismutase or catalase. Superoxide dismutase mRNA is known to be expressed at a lower level in uninduced Friend cells than other murine non-erythroid cells, but whether or not its pattern of expression changes during erythroid maturation is not known (Sherman et al, 1983). Even less information exists concerning the pattern of expression of catalase mRNA (Quan et al, 1986).

How the activities of the genes encoding the diverse erythroid cell functions are modulated during erythropoiesis is an important question in terms of how a cell achieves a given pattern of gene transcription. To attempt to approach this problem, several laboratories have begun to isolate genes expressed in particular cell types. Apart from the globin

genes, the genes (or corresponding cDNAs) encoding other putatively erythroid functions have been isolated. These include murine erythroid band 3 (Kopito and Lodish, 1985; Demuth et al, 1986), human 2,3 bisphosphoglycerate mutase (Joulin et al, 1986), rabbit carbonic anhydrase type I isozyme (Konialis et al, 1985) and rabbit erythroid lipoxygenase (Thiele et al, 1987). In addition, there is a large group of genes which, while being expressed in other cells, are also expressed in erythroid cells at higher levels. These genes may therefore share aspects of their regulation with the strictly erythroid-specific genes. So far cDNAs isolated from this group include rat uroporphyrinogen decarboxylase (Romeo et al, 1984), rat porphobilinogen deaminase (Grandchamp et al, 1984), human 5-aminolevulinate dehydratase (Wetmur et al, 1986), murine 5-aminolevulinate synthase (Schoenhaut and Curtis, 1986), mouse and human carbonic anhydrase type II isozyme (Venta et al, 1985), human glycophorins A and B (Siebert and Fukuda, 1986a; 1986b) and human glycophorin C (Colin et al, 1986; High and Tanner, 1987). In addition, a gene encoding a 12S mRNA present in all tissues examined, but at elevated levels in erythroblasts, liver and kidney has been cloned (Affara et al, 1983; 1985; Goldfarb et al, 1983). When translated in vitro the corresponding mRNA produces a polypeptide of approximately 19kD (Affara et al, 1983) and so was termed ep19-erythroid protein 19kD (Affara et al, 1985). It is with the regulation of expression of this gene (which in fact proved to encode GSHPx) that this thesis is concerned.

1.3. Control of eukaryotic gene expression.

One particular type of differentiated cell within a eukaryotic organism can be distinguished from all the organism's other differentiated cell types by the structural and functional characteristics conferred upon it by its proteins. The first step in the production of these proteins is the transcription of the corresponding genes by RNA polymerase II. (The large ribosomal RNAs [rRNAs] are transcribed by polymerase I and the 5S rRNA and transfer RNAs [tRNAs] are transcribed by polymerase III : in this discussion transcription should be taken to refer to transcription by RNA polymerase II unless otherwise specified). Very shortly after the initiation of transcription (before the new RNA chain is 50 nucleotides [nt] long) the 5' end of the RNA is modified by the 7-methylguanosinetriphosphate "cap", which is thought to be involved in translational initiation (Shatkin, 1976; Darnell, 1982). Transcription then proceeds through the gene before terminating at heterogeneous sites hundreds to thousands of nucleotides 3' to the poly (A) addition site, which corresponds to the 3' end of the final mature mRNA (Birnsteil et al, 1985). An endonucleolytic cleavage of the RNA then occurs 10-30 nt 3' to a sequence homologous to the consensus AAUAAA (Proudfoot and Brownlee, 1976) and a poly (A) tail of 260-300 nt is then added to the RNA (Brawerman, 1981). The efficiency of this process depends upon additional sequences located in the vicinity of the AAUAAA motif (McLauchlan et al, 1985; Gil and Proudfoot, 1987). Another step in the pathway of mRNA biosynthesis is the removal of introns by splicing (Padgett et

al, 1986) after which the mRNA is transported from the nucleus to the cytoplasm. For the vast majority of pre-mRNAs, polyadenylation occurs before any of the introns have been removed by splicing, typical times required for these processes being 2min and 20min respectively.

Control of gene expression can be exerted at several stages within this pathway and could thereby cause a change in the levels of cytoplasmic mRNA, which might then be reflected in a change in the activity of the corresponding protein and an altered cell phenotype. For instance, cleavage of RNA downstream from alternate polyadenylation signals may lead to differentially spliced and polyadenylated RNAs from the same gene which can be translated to yield different proteins, for example soluble or membrane bound forms of IgM (Rogers et al, 1980; Early et al, 1980). Alternatively, differential splicing of the same pre-RNA can give rise to different proteins; γB fibrinogen is produced from the same pre-mRNA as γA fibrinogen, the difference being due to the non-splicing out of the seventh intron (Crabtree and Kant, 1982; see also Leff et al, 1986 for a review of alternate splicing). As well as controls leading to different protein products, controls at different stages in the pathway can lead to different amounts of the same protein. Within the nucleus there is the choice of whether to process or to destroy particular transcripts (Davidson et al, 1982; Darnell, 1982) while different cytoplasmic levels of a particular mRNA can result from a block in the transport of mRNA from the nucleus to the cytoplasm (Fulton et al, 1985). In addition, differences in the stability

of mRNAs can contribute to their steady state cytoplasmic levels. The level of casein mRNA in breast tissue explants increases markedly when the explants are cultured with prolactin and this is associated with a 20-fold increase in the stability of the casein mRNA (Guyette et al, 1979). Moreover, sequences have been identified in the 3' untranslated regions of particular mRNAs which are responsible for their instability (Shaw and Kamen, 1986). The level of a protein present in a cell can also be controlled by separating the mRNA into an actively translated fraction present on the polysomes and an inactive fraction present as free ribonucleoprotein particles : such control occurs during sea urchin embryogenesis (Davidson, 1976), duck erythroblast maturation (Vincent et al, 1981) and photo-responsive development of Volvox (Kirk and Kirk, 1985). In addition to mechanisms modulating the level of particular proteins, cellular phenotype can also be affected by altering the activities of particular proteins through changes in protein conformation or substrate availability induced by external signals. For example, cAMP dependent processes in liver and adipose tissue are activated through conformational change in the adenylate cyclase resulting from glucagon binding to its receptor (Newsholme and Start, 1973) and phospholipase A_2 dependent processes in platelets are activated through a conformational change in the lipid substrate mediated by thrombin binding (Rink and Hallam, 1984).

While all the above mentioned types of control are clearly important, further discussion shall focus upon those aspects of

cellular control which are most relevant to this thesis.

1.3.1. Sequences involved in promoter function.

A promoter is a site at the 5' end of a gene which allows polymerase binding and subsequent transcription initiation. Comparison of promoter sequences has been important in identifying elements involved in the control of transcription. The commonest sequence found in promoter regions is the "ATA" box (consensus TATARAR; Corden *et al*, 1980), which is important in determining the accuracy of transcription initiation, indeed the most 5' nucleotide of the "ATA" sequence is usually located 31nt upstream from the transcription initiation site (Breathnach and Chambon, 1981). Transcription factors have been identified which interact with "ATA" sequences and are required for accurate transcription *in vitro* (Davison *et al*, 1983; Parker and Topol, 1984). However, the chromatographic properties of the "ATA" binding factors identified by these authors differed from each other. This could be due to species differences or might reflect the existence of different types of "ATA" binding factors. In fact, evidence has recently been presented for the existence of distinct regulatory and constitutive classes of "ATA" sequences in yeast (Struhl, 1986; 1987). Moreover, distinct classes of "ATA" sequences also seem to exist in higher eukaryotes since substitution of the "ATA" sequence of the herpes simplex virus type 1 (HSV-1) early glycoprotein D gene for the "ATA" sequence of the SV40 early promoter causes the altered SV40 promoter to respond like the glycoprotein D promoter, and not like the

wild-type SV40 promoter when cells containing these promoters are infected with HSV-1 (R. Everrett, personal communication). Perhaps these differences in the physiological responses of "ATA" sequences might have been expected since many sequences commonly regarded as "ATA" sequences are only loose fits to the consensus.

That the "ATA" sequence is not an absolute prerequisite for transcription is demonstrated by its absence from the promoters of genes encoding certain ubiquitously expressed or "housekeeping" functions such as 3-phosphoglycerate kinase (Singer-Sam et al, 1984), 3-hydroxy-3-methylglutarylCoA reductase (Reynolds et al, 1984) and hypoxanthine phosphoribosyl transferase (Melton et al, 1984). Instead these three promoters are highly G-C rich and contain multiple binding sites for the transcription factor Sp1 (binding site consensus WGGGCGGPuPuPy; Kadonaga et al, 1986). This factor, which has been identified in a variety of different human tissues as well as chinese hamster ovary cells (Briggs et al, 1986), activates transcription from a promoter linked in cis following binding within the major groove of its recognition sequence (Gidoni et al, 1984). Stimulation of transcription by Sp1 is orientation independent (Gidoni et al, 1985) and can occur following binding of a single Sp1 molecule to a promoter (Jones and Tjian, 1985) although it is more common for a promoter to contain multiple Sp1 binding sites (Kadonaga et al, 1986). Moreover, while there is no cooperative nature to the binding of multiple Sp1 molecules to the SV40 promoter (Gidoni et al, 1985) Sp1 bound to DNA can facilitate the binding of additional transcription factors (Jones et al, 1985; Lee et al,

1987). In fact, evidence has been presented strongly suggesting that protein-protein interactions between Sp1 and the "ATA" box factor are required for efficient transcription from the SV40 early promoter (Takahashi et al, 1986a). Thus, insertion of oligonucleotides of multiples of 10bp between the "ATA" box and the 21bp repeats which left the Sp1 binding sites and the "ATA" box on the same side of the helix had only a slightly detrimental effect upon transcription initiation efficiency, whereas introduction of oligonucleotides of 15, 25bp etc. which placed the Sp1 binding sites and the "ATA" box on opposite sides of the helix caused a dramatic reduction in transcription efficiency.

In addition to the relatively widespread "ATA" and Sp1 binding sequences, others exist in a more restricted subset of promoters which can also bind transcription stimulatory factors. For example, in the immunoglobin gene promoters the so-called octamer sequence (ATTTGCAT) was first identified (Parslow et al, 1984; Falkner and Zachau, 1984) and was subsequently found in the promoters of the U1 and U2 small nuclear RNA genes as well as in transcriptional regulatory elements of other genes (Falkner et al, 1986). In addition, the sequence GCCACACCC, which is conserved amongst adult mammalian ß-globin genes was identified (Lacy and Maniatis, 1980) and shown by an analysis of plasmids carrying ß-globin promoter mutations to be required for efficient transcription of the ß-globin gene both in erythroid (Charnay et al, 1985) and in non-erythroid cells (Dierks et al, 1983; Charnay et al, 1985; Myers et al, 1986). A protein which binds to this sequence has recently been identified in HeLa cell nuclear

extracts by its ability to protect this DNA sequence from digestion with DNAse I *in vitro* (Jones *et al*, 1987). Another sequence important for efficient transcription of globin genes in non-erythroid cells (Mellon *et al*, 1981; Dierks *et al*, 1983; Charnay *et al*, 1985; Myers *et al*, 1986) and also found in the promoters of some non-globin genes is the "CAAT" sequence (consensus GGPyCAATCT; Benoist *et al*, 1980; Efstradiatis *et al*, 1980). This sequence has been shown to bind two distinct transcription factors, one isolated from HeLa cells (CAAT transcription factor; Jones *et al*, 1987) and one from rat liver (CAAT binding protein; Graves *et al*, 1986). In addition the sequence CNNGAANNTTCNNG has been identified in the promoters of genes which are activated in response to heat shock (Lindquist, 1986) and these sequences have been demonstrated to be bound by proteins in heat shocked cells (Wu, 1984). Receptor sites for steroid hormones also exist in the promoters of some genes : indeed by an analysis of plasmids carrying promoter mutations for their ability to stimulate transcription of a linked gene when introduced into cells and to bind glucocorticoid receptor *in vitro*, the consensus sequence for binding of glucocorticoid receptor (PyGGTNRCARNTGTPyCT) was deduced (Karin *et al*, 1984; see also Yamomoto; 1985 for a review of steroid receptor action).

No single promoter analysed to date contains all the sequence motifs referred to above, rather, different promoters appear to be built up in a modular fashion from elements containing recognition sites for various transcription factors. Perhaps the best illustration of this is the human

metallothionein IIA (hMT-IIA) gene which contains, in the 250bp immediately upstream from the cap site, (1) an "ATA" sequence,(2) an Sp1 binding sequence (3) a glucocorticoid responsive element (4) two copies of a basal level enhancer sequence (see section 1.3.4) which stimulates transcription by binding a protein which also interacts with the SV40 enhancer and finally (5) four sequences which stimulate transcription in response to heavy metal ions (Lee et al, 1987). Moreover, this is not the complete picture since expression of the hMT-IIA gene is also activated in response to interleukin 1 (Karin et al, 1985) and protein kinase C activators (Imbra and Karin, 1987). In this way the particular physiological requirements of a promoter are tailored from the elements it contains (Dynan and Tjian, 1985; Serfling et al, 1985). As hinted above, this also seems to be true of another class of cis-acting transcriptional regulatory elements, the enhancers (see section 1.3.4. below).

1.3.2. The role of DNA methylation in transcriptional control.

Methylation of cytosines (90% of which occurs at the dinucleotide sequence CpG) is associated with the transcriptional inactivity of many vertebrate genes. Moreover, there are numerous examples where the loss of methylated residues in the 5' flanking regions of genes has been correlated with the transcriptional activation of the corresponding genes either during differentiation or in response to a specific stimulus (for reviews see Doerfler, 1983; Bird, 1984). However, there are other instances where such a correlation is not observed

(Doerfler, 1983). This may well be a consequence of the method used in most of these studies for detecting methylation changes; restriction enzyme digestion with the methylation sensitive (Hpa II) or insensitive (Msp I) isoschizomers recognizing the sequences CCGG. Thus only a proportion of the possible methylation changes occuring in a given stretch of DNA would be detected by this approach. It is therefore possible that demethylation of specific sites not analysed by this method may be causally related to transcriptional activation. Furthermore, since nucleotides outwith the Msp I recognition site can strongly inhibit Msp I cleavage when the Msp I site is methylated at the CpG, such analysis occasionally give erroneous results (Keshet and Cedar, 1983; Busslinger et al, 1983). A point of note here is that stable maintenance of the methylation state of CpG dinucleotides following replication may maintain the transcriptional activation status of a given gene (Doerfler, 1983).

Much of the work alluded to above led to the presumption that demethylation altered the affinity of particular sequences within the 5' flanking regions of genes for trans-acting transcriptional regulatory molecules. Although experimental proof of this notion is lacking it is not unprecedented since in prokaryotes methylation of DNA in the binding site for the lac repressor alters the affinity of the repressor for its recognition site (Fisher and Caruthers, 1979). Moreover, as noted above, the effect of DNA methylation upon the action of certain restriction endonucleases is widely recognized.

More recent data obtained from the analysis of nuclear DNA following the transfection of totally methylated or unmethylated M13 DNA into mouse cells suggests that methylation may affect chromatin structure since totally methylated DNA was assembled into a compact, inactive, nuclease resistant structure whereas unmethylated DNA was formed into a more open, DNAse I sensitive structure (Keshet et al, 1986; see also next section for a fuller discussion of chromatin structure). Other evidence has shown widespread methylation changes to occur during differentiation. Thus, it was demonstrated that transient genome-wide hypomethylation occurred when teratocarcinoma or Friend cells were induced to undergo terminal maturation (Bestor et al, 1984). Furthermore, such hypomethylation was shown to occur by a mechanism independent of DNA replication whereby 5-methylcytosine was excised from DNA and replaced by cytosine (Razin et al, 1986).

Irrespective of the detailed mechanisms involved, it is clear that specific demethylation changes can affect cellular phenotype. For example, treatment of the embryonic fibroblasts (the cell line 10T1/2) with 5-azacytidine (an inhibitor of cytosine methylation; Santi et al, 1983) causes their conversion into myoblasts, chondrocytes and adipocytes (Konieczny and Emerson, 1984). In addition, treatment of HeLa cells with 5-azacytidine allows them to produce human muscle-specific gene products following their fusion to mouse muscle cells, an ability which they do not possess in the absence of 5-azacytidine treatment (Chiu and Blau, 1985).

Recently, a fraction of mouse DNA (representing only 1% of the genome) has been described which contains approximately 15% of the unmethylated genomic CCGG sequences (recognized by digestion with the methylation sensitive enzyme Hpa II). These sequences apparently never become methylated during development (Bird et al, 1985) and exist as islands of unique DNA within which Hpa II sites are concentrated : therefore they were termed Hpa II tiny fragments (HTFs). Four putative HTFs were identified in the G-C rich promoters of housekeeping genes, which suggested that these methylation free zones were responsible for determining that such housekeeping genes did not become transcriptionally silent (Bird et al, 1985). Indeed, it has also been suggested that a lack of methylation might favour the absence of nucleosomes from such DNA (Keshet et al, 1986). This is particularly interesting given the fact that HTFs generally occur at the 5' ends of genes since nucleosomes prevent transcription initiation by polymerase II but do not impede the progress of polymerase once transcription has been initiated (Lorch et al, 1987). This last point may explain the observations that the 5' end of the mouse β^{maj} globin gene is free from nucleosomes in erythroid cells where the gene is about to be or being transcribed, whereas the 3' end is wrapped in nucleosomes (Benezra et al, 1986; Cohen and Sheffery, 1985). In contrast the 5' end of the gene is constrained in nucleosomes in non-erythroid cells.

1.3.3. Chromatin structure and gene transcription.

The bulk of DNA in the nucleus is transcriptionally inert and is packaged in a compact structure which is generally considered to be inaccessible to trans-acting factors. In contrast, genes which are transcribed or are shortly to be transcribed within a given cell (approximately 10-20% of the total [Weisbrod, 1982]) exist in a more open, accessible chromatin configuration. This more open structure can be distinguished from bulk chromatin by its localisation to peripheral regions of the interphase nucleus (Hutchison and Weintraub, 1985) and by its preferential sensitivity to DNAse I digestion (Weintraub and Groudine, 1976). The maintenance of the preferentially DNAse I sensitive nature of chromosomal regions surrounding genes transcribed by both polymerase II and III requires the continuous activity of topoisomerase II (Ryoji and Worcel, 1984; Villeponteau et al, 1984) an ATP dependent enzyme which pumps negative supercoils into DNA (Earnshaw et al, 1985; Wang, 1985). The details of the specificity of recognition of a particular domain by topoisomerase II are unclear but an interaction through specific trans-regulatory factors is known to be involved (Yang et al, 1985; Kmiec and Worcel, 1985). In addition a consensus sequence (GTNAPyATTNATNNG) for recognition by topoisomerase II has been deduced (Sander and Hsieh, 1985).

Superimposed upon the region of DNAse I sensitivity, discrete DNAse I hypersensitive (DH) sites exist which are thought to be free from nucleosomes and bound by non-histone proteins (Saragosti et al, 1980; see also Weintraub, 1985;

Eissenberg *et al*, 1985 for recent reviews). DH sites are commonly formed over the promoter regions of genes; that these sites occur in regions important in gene transcription has been demonstrated most convincingly in an analysis of deletion mutants of the region 5' to the Drosophila *Sgs-4* gene (Shermoen and Beckendorf, 1982). Here, deletion of sequences covering one of the five DH sites upstream of the gene reduces the mRNA level 50-fold. A second mutant lacking sequences covering another two of these DH sites is unable to form the remaining three DH sites and does not produce any *Sgs-4* mRNA. Thus, not only is the presence of all five DH sites required for maximal gene expression but, at least in this instance, formation of a subset of the full complement of DH sites is an absolute prerequisite for formation of the remainder, implying cooperation between DH sites during formation of active chromatin.

However, DH site formation is not sufficient to allow transcription to proceed. For example, a hormone inducible DH site in the mouse mammary tumour virus long terminal repeat appears before maximal transcription (Zaret and Yamamoto, 1984). Moreover, the promoters of heat shock genes are DH sites even in non-heat shocked cells where they are not transcribed (Wu, 1980). A highly significant observation in this context is that RNA polymerase II is bound to the *hsp 70* promoter in non-heat shocked cells but is absent from the gene body : following heat shock, polymerase is still present at the promoter but in addition is detectable throughout the transcription unit (Gilmour and Lis, 1986). The clear implication here is that the rate limiting step

in transcription initiation occurs not only after formation of DH sites but also after binding of polymerase II to the promoter.

An erythroid cell clone transformed by avian erythroblastosis virus possesses DH sites around the globin genes but does not contain any cytoplasmic globin mRNA (Weintraub *et al*, 1982). A similar effect is seen after infection of chicken embryo fibroblasts with a Rous sarcoma virus containing a temperature sensitive *src* gene, at the non-permissive temperature (Groudine and Weintraub, 1982). Following a shift to the permissive temperature the cells become transformed, DH sites appear around the globin genes and they become transcribed. Significantly, after a shift back to the non-permissive temperature, the DH sites remained for 20 cell divisions in the absence not only of the original inducer of their formation, but also of detectable globin mRNA. However, it would appear that such persistence of DH sites in the absence of inducer is not indefinite, since hormonally induced DH sites at the 5' end of the chicken vitellogenin II gene are no longer detectable seven weeks after hormone withdrawal (Burch and Evans, 1986).

The presence of DH sites near a gene reflects the binding of non-histone proteins to the DNA, although the sites of protein binding do not necessarily occur exactly at the DH site (Wu, 1984). Furthermore, through a deletion analysis of DNA underlying the two closely spaced DH sites at the 5' end of the Drosophila heat shock gene *hsp 70* it was shown that sequences underlying one of the DH sites were not required for its formation (Costlow *et al*, 1985). However, further deletion into

the region between the two DH sites not only prevented formation of the DH sites from which the sequences had been deleted but also the formation of the second DH site. This is almost certainly due to deletion of a binding site for the *trans*-regulatory heat shock transcription factor (Wu, 1984). A similar closely spaced DH site doublet was seen 5' to the chicken β^H-globin gene and was shown to reflect the binding of a non-histone protein(s) to the region between the DH sites (Plumb *et al*, 1986).

Not all DH sites reflect the binding of positively acting *trans*-regulatory factors. During induced granulocytic differentiation of the human promyelocytic cell line, HL-60, c-*myc* expression decreases due to a transcriptional block at the 5' end of intron I; significantly, it is at this position that a DH site appears following induction (Bentley and Groudine, 1986; Eick and Bornkamm, 1986). A DH site is also seen 2.4kb upstream from the cap site of the chicken lysozyme gene in tissues where the gene is not expressed (Fritton *et al*, 1984). Recently a DNA fragment covering this DH site has been shown to act as a negative regulatory element during transient expression assays in cells in which the DH site is also present (Steiner *et al*, 1987).

In fact, different sets of DH sites are seen around genes which are expressed in different cell types or in the same cell type in response to differing stimuli. An excellent example of this is again the chicken lysozyme gene where different sets of DH sites are present depending upon whether the gene is not expressed, is expressed constitutively, is steroid inducible or

is steroid induced (Fritton et al, 1984).

1.3.4. Enhancers.

Sequence elements functionally defined as transcriptional enhancers have the ability to stimulate transcription from the normal cap site of a gene linked in cis, in an orientation independent manner (despite the absence of palindromic character) by increasing the frequency of transcription initiation (Weber and Schaffner, 1985; Treisman and Maniatis, 1985). Moreover, these elements can exert their effect over large distances (>1kb) and remain functional when positioned upstream or downstream from the transcription initiation site, preferentially stimulating transcription from the most proximal promoter (Wasylyk et al, 1983a; Kadesch and Berg, 1986).

Enhancers were first described in viral genomes but have since been found in association with several cellular genes (for reviews see Khoury and Gruss, 1983; Serfling et al, 1985). They often reside in regions of DNAse I hypersensitivity, but only in the cell types in which the enhancers are active (Herbomel et al, 1981; Saragosti et al, 1982; Parslow and Granner, 1982). Indeed, for polyoma and SV40, the enhancers are themselves the determinants of their hypersensitivty (Fromm and Berg, 1983; Jongstra et al, 1984; Bryan and Folk, 1986). Moreover, in SV40 minichromosomes the enhancer and DH regions both coincide with a nucleosome free gap known to bind non-histone proteins (Saragosti et al, 1980; Jongstra et al, 1984). It was therefore no surprise when it was demonstrated that enhancer activity both in vivo and

in vitro required specific interaction between the enhancer and cellular factors (Scholer and Gruss, 1984; Wildeman *et al*, 1984) and that such interactions occurred in a cell specific manner which reflected the cell-specificity of enhancer action (Scholer and Gruss, 1985).

The cell type specificity of transcription is known to be determined, at least in part, by the enhancers of the immunoglobulin, insulin, chymotrypsin, α-foetoprotein and lysozyme genes (reviewed by Voss *et al*, 1986). However, cell type specificity is not due solely to enhancer action since some promoters also show cell-specificity independently from enhancers (Foster *et al*, 1985; Mason *et al*, 1985; Edlund *et al*, 1985). Moreover, the degree of activation elicited by an enhancer depends upon the promoter to which it is linked and the cell type used for the analysis (Treisman *et al*, 1983; Berg *et al*, 1984; Garcia *et al*, 1986). A detailed comparison of the interaction between the SV40 enhancer and either the SV40 early or HSV-1 *tk* promoter in different cell types illustrates this promoter dependence (Robbins *et al*, 1986) : whereas the SV40 early promoter is stimulated by the SV40 enhancer in all cell types, the *tk* promoter is not, unless the *trans*-acting factors, T-antigen or the adenoviral immediate-early E1A proteins are present within the cell. It is thus clear that the cell specificity of gene transcription results from the interaction of the cell specificities of particular enhancer-promoter combinations.

Another conclusion to emerge from the analyses of Robbins *et*

al (1986) is that T-antigen stimulates SV40 enhancer activity not by binding to the enhancer but rather by altering the activity of cellular factors which can then bind to the enhancer. Phorbol ester tumour promoters can also stimulate the SV40 enhancer by a post-translational mechanism (Imbra and Karin, 1986) but whether this occurs via the same pathway as T-antigen stimulation is not known. A similar mode of action for the transcriptional induction of the adenovirus E2 transcription unit by adenovirus E1A proteins has been proposed, although it remains unclear whether E1A regulates the activity of cellular factor(s) by increasing their numbers or by increasing their activity through covalent or non-covalent modification (Kovesdi et al, 1986a; 1986b). It is worth noting at this juncture that one of the E1A proteins can repress certain enhancer activities in a cell specific manner (Jones, 1986), perhaps by the same mechanism as is used by some normal cellular repressors (Nir et al, 1986; Stein and Ziff, 1987).

Indeed sequences having a negative regulatory effect upon transcription have been identified near the genes encoding ß-interferon (Goodbourn et al, 1986), insulin (Nir et al, 1986), c-myc (Remmers et al, 1986), retinol binding protein (Colantuoni et al, 1987), albumin (Petit et al, 1986) and α-foetoprotein (Muglia and Rothman-Denes, 1986). Moreover, factors bound to the negative regulatory elements of the ß-interferon gene have been detected in vivo (Zinn and Maniatis, 1986). While it seems reasonable that negative regulation of ß-interferon transcription should occur in the absence of inducer in cells which can

potentially transcribe the gene, the physiological significance of specific DNA-binding repressors of transcription of retinol binding protein or insulin, present in cells which never transcribe these genes in less obvious. In these cells such genes would be expected to be in a compact inaccessible chromatin configuration and in fact this has been demonstrated to be the case for the insulin gene (Wu and Gilbert, 1981).

Comparison of the sequences of various enhancers led to the suggestion that a "core" sequence (GTGGRRRG; Weiher et al, 1983) has a key role in enhancer activity. Thorough analysis of the SV40 enhancer function in human (HeLa) or monkey kidney (CV-1) cells confirmed this suggestion but also identified other sequences which were just as crucial to enhancer function as the "core" (Herr and Clarke, 1986; Zenke et al, 1986). In fact three elements were identified within the enhancer, two within and one upstream of the 72bp repeat : two of these elements contained "core" sequences, but these formed only non-functional sub-domains. None of these elements alone could function as an enhancer but, surprisingly, tandem duplication of a single element restored enhancer activity (Herr and Clarke, 1986; Zenke et al, 1986). Each of these tandem duplications produces enhancers with differing cell type specificities and it was proposed that the broad cell specificity of the SV40 enhancer was due to the combination of smaller units with a more restricted cell specificity (Ondek et al, 1987).

The mechanism of enhancer action remains obscure although several different models have been proposed. Enhancers might act

to organize chromatin into an active conformation or they may alter the supercoil density of a DNA domain by acting as topoisomerase II cleavage sites. The SV40 enhancer can form a DH region over itself when transposed to another part of the chromosome (Fromm and Berg, 1983; Jongstra et al, 1984) and it also has the majority of topoisomerase II cleavage sites located in or near it (Yang et al, 1985). However, the fact that the enhancer seems to be required only to establish and not to maintain an active transcription complex (Wang and Calame, 1986) whereas the continuous activity of topoisomerase II is required to maintain an active chromatin conformation (Ryoji and Worcel, 1984; Villeponteau et al, 1984) argues against either of these mechanisms as being of primary importance in enhancer action.

Alternatively, enhancers may act by directing DNA to a particular nuclear locality (Hutchison and Weintraub, 1985) possibly the so-called matrix (Mirkovitch et al, 1984) but this is a controversial area of research (Jackson and Cook, 1985).

Another possibility is that enhancers act as entry sites for transcription factors which then move to the promoter to interact with other transcription factors causing an increase in transcription initiation. Such movement to the promoter could be caused by the looping out of DNA between the enhancer and the promoter or by the sliding of factors along the DNA from the enhancer to the promoter. Several instances have now been documented of cooperative binding of factors at enhancers and promoters (Takahashi et al, 1986; Lee et al, 1987; Jantzen et al, 1987) which tends to argue in favour of looping. However, other

experiments have shown that when the DNA interposed between the SV40 enhancer and the β-globin promoter is modified by psoralen adducts, enhancer stimulated transcription is decreased (Courey et al, 1986). These results seem incompatible with the idea of looping, suggesting rather that enhancer action involves sliding.

Regardless of the mechanism, one last point is worth emphasizing. Some B-cell lines exist which have deleted the IgH enhancer yet transcribe the IgH gene at normal levels (Wabl and Burrows, 1984; Klein et al, 1984). Recent experiments suggest that this may be explained by enhancer binding factors being required to establish a stable transcription complex which can then be maintained in the absence of such factors (Wang and Calame, 1986; Robbins et al, 1986). Such a mechanism could play an important role in development, as has been argued previously (Weintraub, 1985; Brown, 1984).

1.3.5. Trans-regulatory proteins.

The existence of positively acting trans-regulatory proteins capable of stimulating transcription following binding to particular consensus sequences has already been noted (section 1.3.1). As well as these other proteins have been identified which are converted from an inactive to an active form by post-transcriptional modifications (Sen and Baltimore, 1986b; Nabel and Baltimore, 1987). Still other factors have been inferred from in vivo footprinting (Church et al, 1985; Ephrussi et al, 1985) and transfection experiments (Gerster et al, 1987) yet remain elusive.

Some of the consensus sequences to which such factors bind (the "CAAT" and the octamer sequences) are actually recognized by more than one factor (Jones et al, 1987; Graves et al, 1986; Staudt et al, 1986; Singh et al, 1986; Bohmann et al, 1987); it therefore seems likely that these factors either have different affinities for the same sequence and/or cause different reactions when bound to the same sequence. For instance, oligomerized octamer sequences can cause transcriptional stimulation in the absence of other sequence motifs, but only in cells of the B-lineage (Gerster et al, 1987). This presumably reflects a crucial difference between the mechanisms of action of the B-cell specific and the ubiquitous octamer binding proteins.

The means by which trans-acting transcriptional stimulatory proteins act is exemplified by the steroid receptors. Each of the members of this family of related proteins can be broadly considered to be composed of three functional domains : the C-terminal region contains the hormone binding domain, the central portion the DNA binding domain, while the remaining N-terminal sequences somehow interact with the transcriptional machinery to stimulate transcription (Green and Chambon, 1986). Binding of ligand to these proteins (at least for the glucocorticoid receptor) has been deduced to cause an "unmasking" of the DNA binding domain, thus leading to binding of the receptor to its target sequence and so transcriptional stimulation, since mutant proteins lacking the hormone binding domain are constitutive activators of transcription in vivo (Godowski et al, 1987; Hollenberg et al, 1987). This unmasking

may be due to the removal of an additional protein complexed with the glucocorticoid receptor since purified receptor binds specifically to its target sequence *in vitro* in the absence of hormone (Willman and Beato, 1986) although *in vivo* such binding requires hormone (Becker *et al*, 1986).

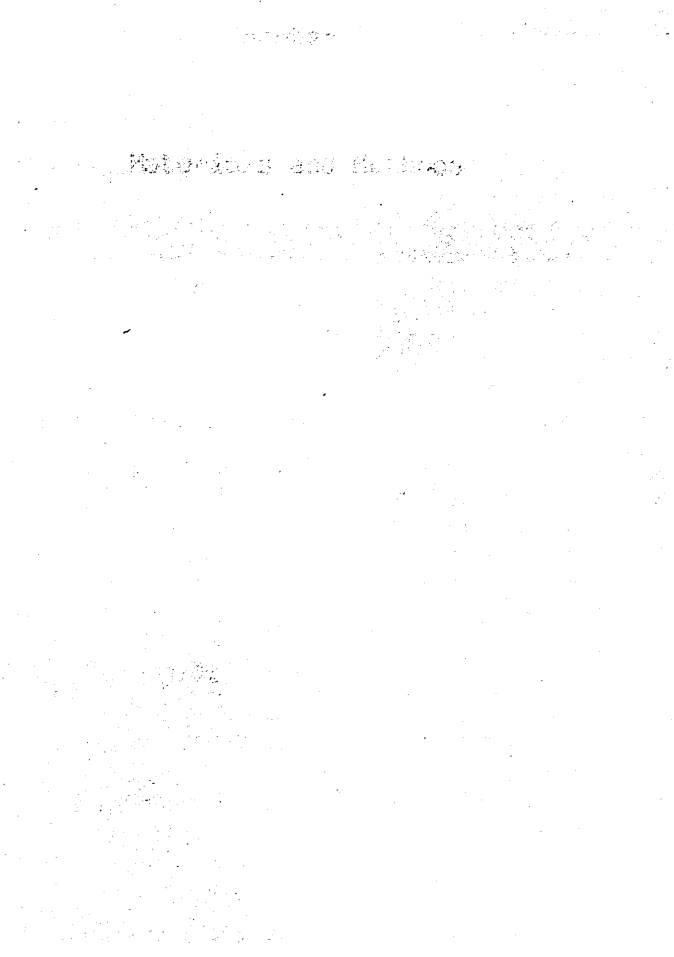
The steroid receptor DNA binding domain is thought to interact with its target sequence through DNA-binding fingers, short peptide loops tethered at the knuckle by coordination, through either two cysteines and two histidines or four cysteines, to a transition metal ion, commonly Zn^{2+} (Vincent, 1986). This type of structural motif was first identified in the *trans*-regulator of polymerase III transcription of the 5S rRNA genes (TFIIIA) and has since also been recognized in several *Drosophila* homeotic gene products as well as steroid and thyroid hormone receptors (Vincent, 1986; Green and Chambon, 1986). In contrast certain prokaryotic proteins contact their binding sites through an α-helix-β-turn-α-helix motif (Schevitz *et al*, 1985)

The N-terminal domain of the glucocorticoid receptor has been shown to increase the transcriptional stimulation caused by the receptor binding to its target sequence (Hollenberg *et al*, 1987) and this is probably the reason why mutant proteins lacking the N-terminal portion of the molecule have no physiological activity (Yamamoto, 1985). Indeed, such mutant proteins might act as repressors of transcription by binding to their target sequences in a sterile manner and this may form the basis of the inhibition of transcription of certain genes by steroids (Charron and Drouin, 1986). This would be analogous to the case of the

yeast transcriptional activator GAL4 which stimulates transcription of the GAL genes in response to galactose. This protein is composed of at least two domains, an N-terminal DNA binding domain and a C-terminal portion which causes transcriptional stimulation. Removal of the C-terminal portion of the molecule converts it into a DNA-binding transcriptional repressor (Keegan et al, 1986).

Trans-regulatory factors exist which can cause the repression of liver-specific functions following fusion of rat hepatoma cells with mouse fibroblasts. In the case of tyrosine aminotransferase and phosphenolpyruvate carboxykinase these trans-regulatory factors have been mapped to the same region of mouse chromosome 11 (Killary and Fournier, 1984; Chin and Fournier, 1987). For albumin the negative trans-regulatory factor has also been mapped to a single chromosome : it exerts its effect via a 400bp DNA region immediately upstream from the albumin cap site (Petit et al, 1986). Trans-dominant negative regulatory elements have also been demonstrated in cell hybrids, which act by abolishing the active chromatin structure of specific genes (Dyson et al, 1985; Affara et al, 1985). Given that the positively acting effector of polymerase III transcription of 5S rRNA genes (TFIIIA) has been shown to cause the formation of active chromatin through an interaction with topoisomerase II (Kmiec and Worcel, 1985; Kmiec et al, 1986) it may be possible for a negatively acting factor to act by a pathway in which it is converted to a topoisomerase II inhibitor upon binding to its target sequence. The associated chromatin

would then revert to an inactive conformation.



Chapter 2

Materials and Methods

2.1. Materials.

All chemicals and biochemicals were obtained from BDH Chemicals, Poole, Dorset or Sigma Chemical Co., St. Louis, Missouri, U.S.A., except those listed below.

Amersham International, England.

All radiochemicals, Nick translation kit N.5000.

Bethesda Research Laboratories, Rockville, Maryland, U.S.A.

All restriction endonucleases, DNA molecular weight markers (Hinf I digested φX174 DNA, HindIII digested λ DNA), agarose, low melting point agarose, urea, vanadyl-ribonucleoside complexes, NP-40, T4 polynucleotide kinase, 5-bromo-4-chloro-3-indolyl-ß galactoside (X-gal), isopropylthiogalactoside (IPTG).

Biorad Laboratories, Richmond, California, U.S.A.

Protein assay kit II.

Boehringer Mannheim, Lewes, East Sussex, England.

Calf intestinal phosphatase (special grade for molecular biologists). Klenow, large fragment of E. coli DNA polymerase I, AMV reverse transcriptase, oligo d(T) cellulose, proteinase K, S1 nuclease.

Camlab, Cambridge, England.

T.l.c. plates.

Difco Laboratories, Detroit, Michigan, U.S.A.

Bacto-tryptone, Bacto-agar, yeast extract.

Fisons p.l.c., Loughborough, England.

Ethyl acetate.

Fulka, Buchs, Switzerland.

Formamide, Guanidinium isothiocyanate.

Highland Stores, Switchback Road, Bearsden.

Marvel.

New England Biolabs, Beverly, Massachusetts, U.S.A.

HindIII linkers.

Pharmacia, Milton Keynes, England.

DNAse I, Sephadex G50, Ficoll 400, DEAE-dextran, dideoxynucleotide triphosphates.

Rathburn Chemicals Ltd., Walkerburn, Peebleshire.

Phenol (water saturated).

Schleicher and Schuell, Dassel, F.R.G.

Nitrocellulose paper.

Chemicals and biochemicals for use in cell culture were obtained from Flow Laboratories, Irvine; Gibco, Paisley or Northumbria Biologicals, Cramlington, England. Sera were from Gibco (horse) or Imperial Laboratories, Salisbury, England (foetal calf). Ham's SF12 and SLM were from Flow Laboratories.

2.2. Animals and cell cultures

2.2.1. Mice.

The mice used during the course of this work were of either the NIH or BALB/c strains. They were kept under 12 hour alternating light-dark cycles and fed ad libitum.

Reticulocytes were prepared by making adult mice anaemic by a single intra-peritoneal injection of acetylphenylhydrazine (0.2ml of a 2.5% w/v solution in Hank's BSS) 5 days prior to sacrifice. The buffy coat was removed by centrifugation (2000g; 20min; 4°C) prior to RNA preparation.

Tissues were dissected from adult mice, rinsed briefly in precooled Hank's BSS containing 10 units heparin/ml (0°C) and held at 0°C until further manipulation. During dissection gall bladders were removed from livers; no attempt was made to remove adrenal glands from kidneys.

2.2.2. Initiation of primary mouse embryo fibroblast cultures.

Embryos were removed from the uterus of pregnant females at 15 days post coitus (the presence of the vaginal plug was taken to indicate day 0 post coitus) and rinsed in Hank's BSS containing 10 units heparin/ml. Foetal livers were removed and the embryos bled by cerebral puncture. Embryos were then chopped finely using a scalpel and suspended in 1ml/embryo 0.025% w/v trypsin (in citrate buffer, pH7.8) precooled to 0°C and kept at 4°C overnight. Trypsinisation was then effected by 30 min incubation at 37°C, after which 5ml/embryo Ham's SF12 medium plus 10% v/v foetal calf serum (FCS) was added. [SF12 is Ham's F12 medium supplemented with

ten times the normal level of folic acid and with extra amino acids (equivalent to adding Eagle's minimum essential amino acids), but lacking thymidine and hypoxanthine]. Cells were then dispersed by repeated passage through a narrow mouthed 25ml pippette and, after allowing the larger fragments to settle, the supernatant was seeded into 2l roller bottles containing 500ml Ham's SF12 plus 10% v/v FCS and penicillin. Cultures were gassed with 10% CO_2/90% air and incubated at 37°C. Non-adherent cells were removed every 2 days and the cultures re-fed. After 1 week monolayers were close to confluence and were taken for RNA preparation.

2.2.3. Cell lines.

Unless otherwise stated all cells were obtained from the Beatson Institute stocks.

Friend cells (clones 707B10/1 and M707T, Harrison et al, 1978) were grown in Ham's SF12 supplemented with 10% v/v horse serum, in 10% CO_2/90% air at 37°C. Erythroid maturation was induced by the addition of 5mM hexamethylenebisacetamide (HMBA) to an exponentially growing culture set up at $5x10^4$ cells/ml, which was allowed to grow for a further 5 days. Friend cell clone F4-12B2 (Greiser-Wilke et al, 1981) was grown in Special Liquid Medium (SLM) (SLM - this is a modification of EM containing twice the normal concentration of all constituents except salts and glucose and with 0.11g pyruvate/litre) supplemented with 4mM glutamine and 10% v/v FCS in 10% CO_2/90% air at 37°C.

Neuroblastoma cells (N18-TG2) were grown in Ham's SF12

supplemented with 10% v/v horse serum, in 10% CO_2/90% air at 37°C.

BW1-J cells (Cassio and Weiss, 1979; Szpirer and Szpirer, 1975) a subclone of a cell line derived from a transplantable hepatoma were grown in Ham's SF12 supplemented with 10% v/v foetal calf serum (FCS) in 10% CO_2/90% air at 37°C. These cells were a gift from Mary Weiss.

RAG cells (Klebe et al, 1970) are a clone of an 8-azaguanine resistant mutant of a transplantable renal adenocarcinoma, a tumour of a kidney tubule cell. They were grown in Special Liquid Medium (SLM) containing 4mM glutamine and 10% v/v FCS in 10% CO_2/90% air at 37°C. These cells were purchased from the American Tissue Culture Collection.

LATK$^-$ cells (Kit et al, 1963) are a thymidine kinase negative subclone of a fibroblast derived cell line. These cells were grown in SLM containing 4mM glutamine and 10% v/v FCS in 10% CO_2/90% air at 37°C.

DeC1.15 cells (Spooncer et al, 1986); a cloned haemopoietic stem cell line, were grown in Fischer's medium supplemented with 20% v/v horse serum and 10% WEHI-conditioned medium at 33°C in 10% CO_2/90% air .

AD-3 cells (Dexter et al, 1980); a pregranulocytic cell line, were grown in Fischer's medium supplemented with 20% v/v horse serum and 10% WEHI-conditioned medium at 37°C in 10% CO_2/90% air.

Both DeC1.15 and AD-3 cells were grown at the Paterson Laboratories by Mike Dexter and colleagues and were shipped to the Beatson overnight when subconfluent.

2.2.4. Growth of cells in culture.

707B10/1, M707T and N18-TG2 grew in suspension and cell numbers were kept between $5x10^4$ and $2x10^6$/ml by subculturing. All other cells grew as adherent monolayers and were grown until subconfluent before harvesting or subculturing at $3x10^4$cells/cm^2.

To remove adherent cells from culture flasks, medium was first removed and the cells washed once in phosphate buffered saline (PBS). Trypsin (0.025% w/v in citrate buffer, pH7.8) was then added to the monolayer and left until the cells had become unattached from the flask. Three volumes of medium were then added and mixed thoroughly with the cells. Cell numbers were then measured in a coulter counter and the appropriate dilution of the cells made.

All cell lines were examined for the presence of mycoplasma using the Hoechst 33258 staining method of Chen (1977). Only cells in which mycoplasma were undetectable were used for experimentation. Any cell stocks in which mycoplasma were detected were removed from the Institute and destroyed.

2.2.5. Storage of cells in liquid nitrogen.

Cells were trypsinised (where necessary), pelleted and resuspended in the appropriate medium containing 10% v/v dimethylsulphoxide (or 10% v/v glycerol in the case of Friend cells) at 10^7 cells/ml. 1ml aliquots were placed in plastic 1.5ml Nunc tubes and these were wrapped in cotton wool inside a polystyrene box and placed at -70°C overnight. (Cooling of cells is at approximately 1°C per minute under these conditions).

Frozen cells were stored in liquid nitrogen.

Recovery from storage was affected by rapid thawing at 37°C. Cells were then incubated in medium at 10^6/ml overnight and then subcultured as described in section 2.2.4.

2.3. Gel electrophoresis.

2.3.1. Agarose gel electrophoresis of DNA.

Agarose gel electrophoresis of DNA was performed using a flat bed apparatus. Gels were made from 0.8% w/v to 2.0% w/v agarose dissolved and cast in buffer. When resolution of higher molecular weight DNA bands was required agarose was dissolved in 2mM EDTA; 40mM tris-acetate, pH7.8 (TAE) buffer; when resolution of lower molecular weight bands (<2kb) was preferred agarose was dissolved in 90mM tris; 90mM boric acid; 2.5mM EDTA, pH8.3 (TBE) buffer. Gels were cast by pouring agarose directly into the electrophoresis apparatus with a well-forming comb in place and the electrodes separated from the gel-forming chamber by removable plastic barriers. Once solidified the gel was submerged in the appropriate buffer and DNA (10-25µg for genomic DNA; 10-1000ng for plasmid DNA) solution (containing 1/10th volume of 0.25% w/v bromophenol blue; 0.25% w/v Orange G; 25% w/v Ficoll 400) loaded into the wells. The current applied was dependant upon the dimensions of the gel and the urgency of the immediate task but was generally in the range of 10-100mA for 0.5-18 h for gels cast and run in TBE and 25 - 200mA for 1-18 h for gels cast and run in TAE. Following electrophoresis, gels were soaked in ethidium bromide solution (0.5µg/ml) for 15 min and excess ethidium bromide

removed by soaking the gel in water for 15 min. DNA was then visualised by illumination on a short wave u.v. light box and photographed (when required) through a red filter using Polaroid type 57 high speed film.

2.3.2. Agarose gel electrophoresis of RNA.

Separation of RNA for "Northern" transfer (see section 2.7.2.) required a gel system which ensured that RNA remained denatured. The system chosen used gels containing formaldehyde whose fractionation properties have been well characterized by Lehrach et al (1977). Gels were prepared by dissolution of agarose in water at 100°C. Following cooling to 60°C, 10x running buffer and formaldehyde were added to give a final gel composition of 1.1% w/v agarose, 2.2M formaldehyde and 1x running buffer (40mM morpholinopropanesulphonic acid [MOPS] pH7.0; 10mM sodium acetate; 5mM EDTA). Gels were immediately cast as described in section 2.3.1. and submerged under 1x running buffer. If RNA was not being loaded immediately gels were stored for short periods of time by submersion in 1x running buffer supplemented with 2.2M formaldehyde (this prevented formaldehyde leaching out of the gel). RNA samples were prepared for electrophoresis by lyophilisation and resuspension in 35µl of 50% v/v formamide; 2.2M formaldehyde; 1xMOPS. These mixtures were then heated at 70°C for 10 min, cooled immediately on ice and 4µl 50% v/v glycerol; 1% w/v bromophenol blue; 10mM sodium phosphate buffer, pH7.0 added, before loading RNA onto the formaldehyde gel (50ml of 1M sodium phosphate buffer, pH7.0 is prepared by mixing 30.5ml 1M Na_2HPO_4

and 19.5ml 1M Na_2HPO_4). Electrophoresis was at 150mA with recirculation of buffer between electrode chambers until the dye had travelled 3/4 of the length of the gel. RNA size markers (mouse 18 and 28S rRNAs from N18-TG2 cells) were then stained in 0.5µg/ml ethidium bromide for 15 min, following separation of the lane containing the RNA markers from the rest of the gel, destained in 1x running buffer for several hours and visualised as detailed in section 2.3.1.

2.3.3. Native polyacrylamide gel electrophoresis.

Gels were prepared by catalysing the polymerization of a solution of 5.84% w/v acrylamide, 0.16% w/v NN' methylene bisacrylamide, 1xTBE, 0.2% v/v NNN'N' tetramethylethylenediamine by the addition of ammonium persulphate to 0.1% w/v. Immediately thereafter gels were cast by pouring between two glass plates separated by 1-2mm. Once solidifed, the gel was placed on a vertical apparatus with each end submerged in a buffer reservoir. DNA was loaded as detailed in section 2.3.1. and gels were run at 25-100mA for 1-4 h. DNA was visualised as described in section 2.3.1.

2.3.4. Denaturing polyacrylamide gel electrophoresis.

These were prepared and run as detailed in section 2.3.3. except that the gel mix contained 7M urea and samples were prepared by heating at 100°C for 2 min in 1xTBE; 50% v/v formamide; 0.25% w/v xylene cyanol; 0.25% w/v bromophenol blue. After running, gels were rinsed for 15 min in 10% v/v acetic acid;

10% v/v methanol and dried onto Whatman 3MM paper. Gels containing ^{35}S labelled DNA were exposed to Kodak X-S film at room temperature; those containing ^{32}P labelled DNA were exposed to Kodak XAR-5 film at -70°C with intensifying screens (Laskey and Mills, 1977).

2.4. Preparation of nucleic acids.

2.4.1. Eukaryotic RNA

Several precautions were adopted during RNA preparation to minimise degradation due to endogenous or contaminating RNAses. For preparation of total cellular RNA, guanidinium isothiocyanate and ß-mercaptoethanol were used to simultaneously disrupt cells and inactivate RNAses. When cells were lysed in the absence of chaotrophic agents, vanadyl-ribonucleoside complexes were added prior to cell lysis to inhibit RNAses. In addition all equipment and solutions to be used for RNA preparation were treated overnight with a 0.1% v/v aqueous solution of diethylpyrocarbonate (DEPC) to remove RNAses. DEPC was then removed from solutions by autoclaving and from equipment by rinsing in copious quantities of double distilled water.

i. Total Cellular RNA.

This was isolated using a modification of the method of Chirgwin et al, (1979).

Homogenized cells from mouse tissue or pelleted cells from a cell culture were washed once in PBS and centrifuged (2000g; 15 min; 4°C). 1ml packed cells were then lysed by vortexing in 8ml 5M guanidinium isothiocyanate; 10mM EDTA; 50mM tris-HCl, pH7.5 and

0.4ml ß-mercaptethanol. DNA was then thoroughly sheared by 10 passages through a wide bore hypodermic needle (gauge 19; Gillette surgical, Isleworth, Middlesex, U.K.) and 10 passages through a narrow bore hypodermic needle (gauge 23). After incubation at ambient temperature for 20 min, 50ml 4M LiCl were added and the mixture incubated at $4^{o}C$ for 20h to precipitate RNA. RNA was then pelleted by centrifugation (10 000g, 90 min, $4^{o}C$), washed thoroughly in 30ml 4M LiCl and again pelleted by centrifugation (10 000g, 30 min, $4^{o}C$). The pellet was then dissolved in 5ml 0.1% w/v SDS; 1mM EDTA; 10mM tris-HCl, pH7.5 and extracted successively against 2.5ml phenol once, 5ml phenol-chloroform twice and 5ml chloroform once (the preparation of phenol, phenol-chloroform and chloroform is described in section 2.9.3.) Phases were then separated by centrifugation (16 000g, 10 min, $20^{o}C$) and RNA precipitated (section 2.9.2.) and stored under ethanol at $-20^{o}C$.

ii. Total cytoplasmic RNA.

This was prepared essentially as described by Minty and Gros (1980). A 0.5ml cell pellet was suspended in 9.5ml 140mM NaCl; 1.5mM $MgCl_2$; 10mM tris-HCl, pH7.4 and 0.5ml vanadyl-ribonucleoside complexes added to inhibit RNAses. Lysis was then effected by addition of NP-40 to 0.5% v/v and incubation at $0^{o}C$ for 5 min. Nuclei were then pelleted by centrifugation (4 000g, 5 min. $4^{o}C$) and RNA precipitated from the supernatant by the addition of an equal volume of 8M urea; 4M LiCl and incubation at $4^{o}C$ for 24-48 h. RNA was then pelleted by centrifugation (16 000g, 15 min, $4^{o}C$), dissolved in 10ml 8M urea and reprecipitated by the addition of 10ml 4M LiCl and incubation at $4^{o}C$ for 24 h. RNA was

again pelleted by centrifugation (16 000g, 15 min, 4°C), resuspended in 7.5ml 0.1% w/v SDS; 1mM EDTA; 5mM tris-HCl, pH7.5 and extracted thoroughly against 7.5ml phenol-chloroform. Phase separation was achieved by centrifugation (16 000g, 10 min, 20°C) after which the RNA was precipitated (section 2.9.2.) and stored under ethanol at -20°C.

iii. Selection of poly A^+ RNA.

This was performed as outlined by Favaloro et al (1980). That is, oligo(dT)-cellulose from which the "fines" had been removed in elution buffer (0.1% w/v SDS; 1mM EDTA; 5 mM tris-HCl, pH7.5) was poured to form a 1ml column in a sterile 10ml BioRad plastic disposable column. The column was then washed in binding buffer (0.1% w/v sarkosyl; 0.5M NaCl; 10mM tris-HCl, pH7.5) and RNA dissolved in binding buffer was applied to the column at an amount not exceeding 25% of the binding capacity of the column. The flow-through was collected and re-applied to the column, twice. The column was then washed once with 2ml binding buffer and then thrice with 2ml midwash buffer (0.1% w/v sarkosyl; 0.1M NaCl; 10mM tris-Hcl, pH7.5). The poly A^+ RNA was then eluted from the column by the application of three successive 1ml aliquots of elution buffer, precipitated (section 2.9.2.) and stored under ethanol at -20°C.

2.4.2. Eukaryotic genomic DNA

Nuclei, prepared by NP-40 lysis of cells as described in section 2.3.1. ii, were resuspended in 2.5ml 140mM NaCl; 1.5mM $MgCl_2$; 10mM tris-HCl, pH7.4 and lysed by the addition of 12.5ml

1% w/v SDS; 20mM EDTA from a pipette at an even, forceful rate, in order to give a homogeneous solution. Proteinase K was then added to 0.5mg/ml and the solution gently agitated at 37°C overnight. The mixture was then extracted successively once against phenol, twice against phenol-chloroform and once against chloroform by thorough mixing and the phases separated by centrifugation (16 000g, 10 min, 20°C). Nucleic acid was then precipitated (section 2.9.2.) at -20°C overnight. After pelleting by centrifugation (16 000g, 10 min, 4°C), nucleic acid was dissolved in 0.1x SSC (1xSSC is 0.15M NaCl; 0.015M sodium citrate, pH7.0) and RNA hydrolysed by the addition of a mixture of RNAse A and RNAse T_1 (which had previously been boiled for 10 min to destroy any contaminating DNAse) to give final concentrations of 20µg/ml and 40 units/ml respectively and incubated at 37°C for 30 min. DNA was then extracted with phenol, phenol-chloroform, chloroform and ethanol precipitated as detailed above.

2.4.3. Prokaryotic plasmid DNA.

i. Small scale preparation.

In preliminary examinations, plasmid DNA was prepared by the "zippy prep" method of Holmes and Quigley (1981). A single bacterial colony containing the plasmid of interest was inoculated into 1.5ml L-broth supplemented with 100µg/ml ampicillin and incubated at 37°C overnight in an orbital shaker. 0.5ml of culture was saved and, if desired, was diluted with 0.5ml glycerol and stored at -20°C. Bacteria were pelleted from the remaining culture by a 15 sec microfugation, resuspended in 8% w/v sucrose;

5% v/v Triton X-100; 50mM EDTA; 50mM tris-HCl, pH 8.0; 0.6mg/ml lysozyme and placed at 100°C for 45 sec. Chromosomal DNA was pelleted by a 5 min microfugation, the supernatant removed to a fresh tube and phenol-chloroform extracted. Plasmid DNA was then precipitated (section 2.9.2.), resuspended in T.E. buffer (1mM EDTA; 10mM tris-HCl, pH7.5) and stored at -20°C.

ii. Bulk preparation

This was prepared essentially as detailed by Birnboim and Doly (1979). Bacteria containing the plasmid of interest were taken from storage (section 2.4.3.i.) and 20µl inoculated into 5ml of L-broth supplemented with 100µg/ml ampicillin and grown overnight at 37°C in an orbital shaker. The culture was then transferred to an additional 500ml of growth medium and incubated overnight at 37°C in an orbital shaker. Bacteria were then pelleted by centrifugation (4 000g, 10 min, 4°C), resuspended in 10ml of lysis solution (50mM glucose; 25mM tris-HCl, pH8.0; 10mM EDTA; 5mg/ml lysozyme) and left at ambient temperature for 5 min to allow bacterial lysis. 20ml of freshly prepared 0.2M NaOH; 1% w/v SDS was then added, mixed gently and the whole left on ice for 10 min. The solution was then neutralised by addition of 15ml 5M potassium acetate, pH4.8, thorough mixing and incubation of 0°C for 10 min. Precipitated genomic DNA was then removed by centrifugation (16 000g, 10 min, 4°C) and plasmid DNA precipitated from the supernatant by addition of 0.6 volumes of iso-propanol and incubation at ambient temperature for 15 min. After centrifugation (16 000g, 10 min, 20°C) DNA was rinsed in ethanol, dried and dissolved in 6ml T.E. buffer. Caesium chloride (8g) was

then added and dissolved at 37°C, 700µl ethidium bromide (3mg/ml) added and the refractive index adjusted to 1.394-1.396 with T.E. buffer. The mixture was then transferred to a 10ml tube and centrifuged at 45 000rpm in a 10x10 fixed angle rotor at 20°C for 36 h using a Damon IEC/B-60 ultracentrifuge. The lower DNA band containing plasmid was then collected using a hypodermic needle and ethidium bromide immediately removed by four successive extractions with iso-propanol saturated with aqueous caesium chloride . Dialysis against several changes of 10mM tris-HCl, pH 7.4; 1mM EDTA; 10mM NaCl removed caesium chloride from the DNA, which was then precipitated (section 2.9.2.), resuspended in T.E. buffer and stored at -20°C.

2.4.4. Preparation of nuclei and digestion with DNAse I

Two separate methods were used, the first being a standard technique, the second an adaption of this specifically designed to yield intact nuclei from sources know to have a high endogenous nuclease content, such as liver. For uniformity, the latter method was used on all tissues examined, the former reserved for cell cultures.

i. From cell cultures

All manipulations were performed at temperatures between 0 and 4°C unless otherwise stated.

Cells were pelleted from culture medium by centrifugation (2 000g, 15 min, 4°C) and washed twice in PBS. Nuclei were prepared by resuspending cells in reticulocyte standard buffer (RSB - 10mM NaCl; 3mM $MgCl_2$; 10mM tris-HCl, pH7.4) supplemented with 0.25M

sucrose and 0.2% v/v NP-40 and incubating at 0°C for 10 min (Weintraub and Groudine, 1976). Nuclei were then pelleted by centrifugation (4 000g, 10 min, 4°C) and resuspended in RSB + 0.25M sucrose but lacking detergent. Microscopic examination showed lysis of plasma membranes to be complete. Nuclei were then washed three times in RSB + 0.25M sucrose and resuspended at a nucleic acid concentration of 1mg/ml.

Aliquots of nuclei were then digested by adding DNAse I in the concentration range 0-10 units/mg DNA for 2 min at 37°C after first preincubating the nuclei at 37°C for 4 min. I found it necessary to add the DNAse I from a stock solution at 12 500 units/ml prepared by dissolution of DNAse I in 0.2mM $CaCl_2$; 10mM tris-HCl, pH 7.5, since simply dissolving DNAseI in water did not give digested nuclei. Digestions were stopped and, at the same time, nuclei were lysed by the addition of 5 volumes of 1% w/v SDS; 20mM EDTA. Genomic DNA was then prepared as detailed in 2.4.2.

ii. From tissues

Liver, kidney and brain tissue dissected from BALB/c mice were used immediately; all steps were performed between 0 and 4°C. Nuclei were prepared essentially as outlined in Becker *et al* (1984) and digested with DNAse I essentially as detailed by Fritton *et al* (1983). Tissue was suspended in homogenisation buffer [60mM KCl; 15mM NaCl; 15mM tris-HCl, pH7.5; 0.5mM spermidine; 0.15mM spermine (buffer A) containing 0.3M sucrose; 2mM EDTA; 0.5mM EGTA and 1mM phenylmethylsulphonylfluoride (PMSF) (freshly added from a 100mM *iso*-propanolic solution)] by 20

strokes in a Dounce homogenizer. The homogenate was filtered through gauze and cells pelleted by centrifugation (2 000g, 15 min, 4°C). Nuclei were then released by a 10 min incubation in homogenization buffer containing 0.2% v/v Triton X-100, after which plasma membrane lysis was judged complete by microscopic examination. Nuclei were then washed thrice in homogenization buffer and resuspended in buffer A plus 1mM PMSF; 0.2mM EGTA at a nucleic acid concentration of 0.5mg/ml.

0.5ml aliquots of nuclei were then digested by DNAse I in the concentration range 0-2 units/µg DNA by the addition of 5µl of the appropriate dilution of enzyme (prepared by dissolution in 0.2mM $CaCl_2$; 10mM tris-HCl, pH7.5; 0.5M $MgCl_2$) and incubation at 0°C for 10 min. (Dilution from the stock was such that each aliquot of nuclei was resuspended at 5mM $MgCl_2$, thereby allowing nuclease action to proceed.) Nuclei digestions were then stopped by nuclear lysis and genomic DNA prepared as outlined in section 2.4.4.i.

2.5. Manipulation of DNA

2.5.1. Digestion of DNA with restriction endonucleases.

Plasmid DNA was digested with 2-5 units enzyme/µg for 3 h under conditions specified by the manufacturer. When necessary reactions were terminated by phenol-chloroform extraction and ethanol precipitation (section 2.9.2.).

Genomic DNA was digested with 10-20 units enzyme/µg for 18 h and a further 10-20 units enzyme/µg was then added and the

incubation continued for a further 3 h. In most cases, digests were phenol-chloroform extracted and ethanol precipitated (section 2.9.2.) before proceeding to the next step.

2.5.2. Purification of DNA restriction enzyme fragments

Depending upon the size of the DNA fragment of interest and the size of contaminating DNA fragments electrophoresis was carried out either through low melting point agarose or polyacrylamide gels. Following ethidium bromide staining, the appropriate region of the gel containing the DNA of interest was excised and the DNA purified by one of the following methods.

i. From low melting point agarose.

The agarose slice containing DNA was diluted in approximately twice its own volume of T.E. buffer and heated at 70°C until the agarose had melted. An equal volume of ice cold phenol was then added and mixed thoroughly with the DNA/agarose solution by vortexing. Following a 2 min microcentrifugation the aqueous phase was removed and re-extracted with phenol until no interphase was visible after a 2 min microcentrifugation. The volume of the DNA solution was reduced by iso-butanol extraction before ethanol precipitation of DNA (section 2.9.2.)

ii. From polyacrylamide.

For DNA fragments smaller than about 400 b.p. the gel slice was soaked in 0.5M ammonium acetate; 1mM EDTA at 37°C overnight by which time most of the DNA had eluted from the gel.

For DNA fragments larger than about 400 b.p. the gel slice

was placed in a dialysis tube in TAE buffer and the dialysis tube placed in an electrophoresis tank containing TAE buffer. Quantitative elution of DNA was achieved by passing a current of 200 mA through the buffer for 1 h. Since, during this period, the DNA had moved to the side of the dialysis tube a reversal of the current was applied for 1 min when the DNA was then present in the buffer.

DNA was ethanol precipitated (section 2.9.2.) and resuspended in TE buffer.

2.5.3. Removal of 5' terminal phosphate groups.

In order to reduce the number of bacterial colonies containing non-recombinant plasmid molecules during sub-cloning experiments the 5' phosphate groups of the linearised vector DNA were removed by treatment with calf intestinal phosphatase (C.I.P.). These phosphate groups were also removed from DNA fragments as a prelude to 5' end labelling with ^{32}P. 20 min before the end of a restriction enzyme digestion which would produce the DNA termini containing the unwanted phosphate groups, C.I.P. was added (1 unit/µg DNA) and the incubation continued. The reaction was stopped by the addition of EDTA to 20mM and SDS to 0.5% w/v. Incubation at 65^{o}C for 10 min destroyed the phosphatase which was then removed by phenol-chloroform extraction and the DNA ethanol precipitated (section 2.9.2.).

2.5.4. Conversion of 5' single stranded regions of duplex DNA to double stranded DNA.

DNA fragments produced by digestion with particular restriction enzymes cannot always be readily cloned because of the sequence of the 5' single stranded termini (5' overhang). One way round this problem is to insert nucleotides complementary to those of the 5' overhang using deoxynucleotide triphosphates (dNTPs) and the Klenow fragment of E.coli DNA polymerase I. 2µg of digested DNA was incubated at 15°C for 1 h in 50µl of 50mM tris-HCl, pH7.8; 5mM Mg Cl_2; 1mM dithiothreitol; 20µM each dNTP and 5 units Klenow fragment. The reaction was stopped by phenol-chloroform extraction and DNA ethanol precipitated (section 2.9.2.).

2.5.5. Addition of short synthetic double stranded DNA to linear DNA.

In order to clone a DNA fragment into a plasmid such that the inserted DNA is at a particular position in the plasmid with respect to other plasmid sequences it is sometimes necessary to change the restriction enzyme sites at the end of the insert DNA. This can be done by adding short synthetic double stranded DNA (molecular linkers) containing a recognition site for a restriction enzyme to a blunt ended DNA fragment. 2µg of blunt ended DNA was ligated with a 50-fold excess of phosphorylated molecular linkers as described in section 2.5.7. below. Ligation was terminated by heating at 65°C for 10 min, DNA ethanol precipitated (section 2.9.2.) and then digested with a 50-fold excess of the appropriate restriction enzyme for 7 h to remove

multiple copies of the linker molecules. DNA was then purified from digested linkers by low melting point agarose gel electrophoresis (section 2.5.2.i).

2.5.6. Labelling DNA with radioisotopes.

i. Nick-translation of DNA.

Purified DNA fragments were uniformly labelled by nick-translation (Rigby et al, 1977) using the Amersham kit (N.5000). This contains the necessary enzymes and buffers to which DNA, radioactively labelled [^{32}P] dCTP and water were added in the following quantities. 150ng DNA was added to 3µl 5x buffer solution and this was made up to 10.5µl with water. 3µl of α-[^{32}P]-dCTP (>3000 Ci/m mole) were then added and finally 1.5µl of the enzyme solution. This was then incubated at 11^{o}C for 1-2 h and the reaction stopped by addition of 5µl 100mM EDTA. Labelled DNA was purified from unincorporated label by gel filtration over a sephadex G50 column in 0.1xSSC, 0.1% w/v SDS. The specific activity of the labelled DNA (which, routinely was between 1-5x10^{8}cpm/µg) was determined by measuring the radioactivity of an aliquot of the DNA solution in a scintillation counter. Just prior to use the DNA was denatured by boiling for 10 min and then immediately cooling on ice.

ii. 5' end labelling DNA.

This was executed essentially as described by Maxam and Gilbert (1980).

Purified calf intestinal phosphatase treated DNA (<20 pmoles 5' ends) in 11µl of 1mM spermidine; 0.1mM EDTA; 20mM tris-HCl,

pH9.5 was heated at 70°C for 2 min, plunged on ice and left for 5 min before being transferred to another tube containing 5µCi dried γ-[^{32}P] ATP (>5000 Ci/mmol), 6µl H_2O and 2µl 100mM $MgCl_2$; 50mM dithiothreitol; 50% v/v glycerol; 500mM tris-HCl, pH9.5 20 units (2µl) of T4 polynucleotide kinase were then added and the reactants incubated at 37°C for 2 h. The reaction was then stopped and DNA precipitated by addition of 80µl 2.5M ammonium acetate, 1µl (10µg) yeast tRNA, 275µl ethanol and incubation at -70°C for 10 mins.

2.5.7. Ligation of DNA fragments into vectors.

Ligations were performed in 66mM tris-HC1 pH7.4, 1mM spermidine, 10mM $MgCl_2$, 4mM DTT and 1mM ATP overnight at 14°C. For cohesive ended ligation 10 units of T4 DNA ligase was added per microgram of vector DNA whilst 20-100 times as much enzyme was employed in ligating blunt ends. The optimum molar ratio of vector to insert was determined in each situation by trial ligations, but was generally in the range of 1:1 to 5:1 (in blunt end ligations a greater proportion of insert was used to counter the more probable event of vector religation). Reaction volumes were such that the concentration of 5' ends was 0.1-1.0µM.

2.5.8. Transfection of E.coli cells with DNA.

This was carried out essentially as described by Mandel and Higa (1970).

Fresh overnight cultures were diluted 1 in 100 with 100ml L-broth (1% w/v bacto-tryptone; 0.5% w/v yeast extract; 1% w/v

NaCl) and grown to an OD_{600} of 0.4 to 0.6. Cells were pelleted by centrifugation (1 000g, 5 min, 4°C) resuspended in 50ml precooled 50mM $CaCl_2$, incubated at 0°C for 30 min, pelleted and resuspended in 10ml 50mM $CaCl_2$. Competent cells were stored for up to 2d at 4°C.

Ligation mixtures containing up to 1µg DNA were added to 200µl competent cells, incubated at 0°C for 30 min and then heat shocked at 42°C for 2 min. 1ml of L-broth was then added and the culture incubated at 37°C for 1 h to allow expression of plasmid encoded functions before spreading onto 90mm plates containing L-broth supplemented with 1.5% w/v agar and 100µg/ml ampicillin (25µg/ml of the chromogenic substrate for β-galactosidase, 5-bromo-4-chloro-3-indolyl-ß-galactoside [X-gal] was also included in the growth medium when JM83 was the host bacterium).

2.6. Sequencing of nucleic acids.

2.6.1. DNA.

The methods and principles involved in the subcloning and sequencing reactions are clearly outlined in the M13 cloning and sequencing handbook proved by Amersham.

Briefly, DNA fragments were cloned into each of the M13 vectors mp10 and mp11 (Norrander et al, 1983), single stranded DNA prepared and sequenced by the dideoxy chain termination method (Sanger et al., 1977) in the presence of α-[^{35}S] dATP (Biggen et al., 1983). Reaction products were then denatured and electrophoresed through 6% polyacrylamide/7M urea/TBE gels (section 2.3.4.).

2.6.2. RNA.

RNA was sequenced by a modification of the method of Hamlyn et al (1978) using reverse transcription from a 5' end labelled primer in the presence of dideoxynucleotidetriphoshates (ddNTPs). The primer used was an Msp I-Pst I fragment and because Msp I gives a 5' overhang and Pst I a 3' overhang the two strands are of different lengths. The purified, phosphatased DNA fragment was 5' end labelled as described (2.5.6.ii.) and the two strands separated by denaturation and electrophoresis through a 15% polyacrylamide/7M urea/TBE gel. After recovery, 5ng of primer was mixed with 100µg of reticulocyte poly A^+ RNA in 7.5µl water and heated to 70°C for 1 min followed by rapid cooling on ice. KCl was then added to 400mM and the primer-RNA mix annealed at 20°C for 2 h. The mix was then split into five equivalent aliquots, four of which were made 60µM with respect to one of each of the four ddNTPs. The volume of each aliquot was then adjusted to 50µl containing 100mM tris-HCl (pH8.3 at 42°C)/10mM $MgCl_2$/140mM KCl/20mM ß-mercaptoethanol/400µM each dNTP/5 units reverse transciptase and incubated at 42°C for 1 h. The reaction was then stopped and RNA hydrolysed by addition of EDTA to 10mM and NaOH to 0.2M and incubation at 42°C for 40 min. After neutralisation with HCl, the cDNA was phenol extracted, ethanol precipitated (section 2.9.2), denatured and run on an 8% polyacrylamide/7M urea/TBE gel (section 2.3.4.).

2.6.3. Computer analysis of sequences.

Sequence analysis was performed using an IBM PC-AT microcomputer in conjunction with the "MicroGenie" software (Beckmann).

2.7. DNA-DNA and DNA-RNA hybridisation techniques.

2.7.1. Southern blotting.

DNA fragments separated by agarose gel electrophoresis (2.3.1.) were transferred to nitrocellulose paper for hybridisation to DNA probes by the method of Southern (1975). Following photography, the gel was rinsed twice in 0.5M NaOH; 1.5M NaCl for 15 min and then twice in 3M NaCl; 0.5M tris-HCl, pH7.5 for 30 min. The gel was then transferred to a raised platform covered with two sheets of Whatman 3MM paper soaked in 20xSSC (1xSSC is 0.15M NaCl; 0.015M sodium citrate, pH7.0) the ends of which extended below the platform into a reservoir of 20xSSC. Care was taken to ensure that there were absolutely no visible bubbles between the platform and the gel. The complete assembly around the gel was then sealed off from the space above using parafilm to ensure that all movement of liquid occurred through the gel. A sheet of nitrocellulose (previously soaked in 2xSSC) was then placed in direct contact with the gel, without any intervening bubbles. Two sheets of Whatman No.1 filter paper (pre-soaked in 2xSSC) were then placed directly atop the nitrocellulose, with no intervening bubbles. Finally, the contents of a box of tissues were placed above the filter paper and weighted down with a 0.5kg mass. The complete assembly was

left overnight to allow transfer of DNA. The nitrocellulose was then retrieved and the position of the gel wells marked. Completion of transfer of DNA was monitored by re-staining the gel in ethidium bromide (section 2.3.1.). The nitrocellulose filter was rinsed briefly in 2xSSC, air dried, baked at 80°C for 2-3 h and then stored at room temperature until required.

2.7.2. Northern blotting.

RNA separated in denaturing agarose gels (section 2.3.2.) was transferred to nitrocellulose as described in section 2.7.1. except that no pretreatment of the gels was performed before assembly into the transfer apparatus.

2.7.3. Hybridisation of nucleic acids immobilised on nitrocellulose filters.

Nitrocellulose filters were first wetted in 2xSSC and then transferred to heat sealable plastic bags. Prehybridisation buffer (50% v/v formamide; 5xSSC; 0.5% w/v Marvel; 0.1 mg/ml denatured salmon sperm DNA; 0.1% w/v SDS) was then added, the bag sealed and incubated with shaking overnight at 42°C. Prehybridisation buffer was then removed and replaced with fresh buffer (for hybridisations to Southern blots, this was supplemented with 10% w/v Dextran sulphate) to which had been added denatured radioactively labelled probe (prepared as detailed in 2.5.6.i) such that the probe concentration was 10-20ng/ml (concentrations at the higher end of this range were preferred for genomic Southerns). Bags were then sealed and incubation at 42°C

continued for a further 18-24 h. Filters were then washed, routinely twice in 2xSSC; 0.1% w/v SDS, room temperature for 15 min and then twice in 0.1xSSC; 0.1% w/v SDS, 60°C for 30 min. Filters were then air dried, wrapped in thin plastic and exposed to Kodak XAR-5 or Kodak X-S film at -70°C with intensifying screens.

2.7.4. S1 nuclease protection assay.

This was performed essentially as outlined by Weaver and Weissmann (1979).

An excess of end labelled single stranded probe was heated at 100°C in 10µl 80% v/v formamide for 2 min and plunged on ice. RNA was precipitated, dried and the total amount of RNA made up to 50µg with yeast tRNA. 1µl 4M NaCl; 10mM EDTA; 0.4M PIPES [Piperazine-N,N'-bis (2-ethanesulphonic acid)] pH6.4 was added to the RNA, dried and the whole resuspended in the 10µl of probe/formamide mix and incubated at 52°C overnight. Hybridised RNA: DNA was then diluted into 250µl 250mM NaCl; 1mM Zn SO_4; 40mM sodium acetate, pH4.4. 2000 units of S1 nuclease were then added and the mixture incubated at 37°C for 1h. Hybrids were then precipitated, denatured and electrophoresed through a 6% polyacrylamide/7M urea/TBE gel (section 2.3.4.).

2.8. Transient expression of DNA introduced into eukaryotic cells in culture.

2.8.1. DNA mediated gene transfer into eukaryotic cells.

DNA was transferred into cells by a modification of the DEAE-dextran method of McCutchan and Pagano (1968).

Exponentially growing cells were set up in multiple 9cm petri dishes at $2x10^6$ cells per dish and grown overnight. Medium was then removed and cells washed very gently once in 10ml serum free medium. 9ml serum free medium was then added to each dish. 100µl of a freshly prepared stock of serum free medium containing 10µg DNA was then added, followed immediately by 1ml 2.5mg/ml DEAE-dextran, pH7.3 and the cells incubated at 37°C for 1h. Medium was then removed, the cells rinsed once very gently in serum free medium, refed and incubated at 37°C for 48h.

2.8.2. Harvesting transfected cells; preparation of soluble protein extracts

After short term expression of introduced DNA, cells were harvested essentially as described (Gorman et al., 1982).

Cells were washed very gently once in PBS, then 1.5ml 150mM NaCl; 1mM EDTA; 40mM tris-HCl, pH7.4 added and left at room temperature for 5 min. The cells were then scraped off the dish surface using a disposable spatula, transferred to an eppendorf centrifuge tube and pelleted by a 2 min microfugation at 4°C. The cell pellet was then resuspended in 100µl 250mM tris-HCl, pH7.8 and mixed thoroughly by vortexing. Cells were disrupted by three freeze-thaw cycles, alternately at -70°C for 5 min and 37°C for 5

min. After pelleting debris by a 2 min microfugation at 4°C, the supernatant was removed to a fresh tube and stored at -20°C.

2.8.3. Determination of protein concentration of extracts.

This was determined by the method of Bradford (1976) using prepared reagents purchased from BioRad.

Briefly, the extinction at 595nm caused by complex formation between Coomassie blue and protein from an aliquot (2-5µl) of extract was measured and compared against that due to complex formation between Coomassie blue and a range of concentrations of bovine serum albumin prepared at the same time.

2.8.4. Measurement of chloramphenicol acetyltransferase activity

Aliquots of cell extracts prepared as detailed in section 2.8.2. were taken and diluted to 89µl with 250mM tris-HCl, pH7.8 so that the same amount of protein (as determined in section 2.8.3.) was taken from each of the extracts prepared from a given cell line in a given experiment. These were then heated at 65°C for 10 mins to denature proteins other than CAT which utilise acetyl CoA as a substrate (Fordis et al, 1986). Assays were then performed by the addition of 11µl of a pre-mix consisting of 1 volume ^{14}C-chloramphenicol and 10 volumes 25mM acetyl CoA to each sample and incubation at 37°C for 60 min. Reactions were terminated by the addition of 300µl ethyl acetate and vortexing. Phase separation was by microfugation (30 sec) after which the organic phase was removed to a fresh tube and dried down on a

Savant speedivac concentrator (15 min).

The mixture of unacetylated and acetylated chloramphenicol was then dissolved in 30µl ethyl acetate, transferred to t.l.c. plates and chromatography performed using 5% v/v methanol, 95% v/v chloroform as the mobile phase. T.l.c. plates were then exposed to Kodak XAR-5 at room temperature. Quantitation was achieved by separately counting the radioactivity in the unacetylated and acetylated forms for each sample in a scintillation counter.

2.9. Miscellaneous methods.

2.9.1. Determination of nucleic acid concentrations.

i. Spectrophotometric determination.

Nucleic acid concentrations were measured by determining the absorbance of an aqueous solution of the acid at 260nm using the convention that an absorbance of 1 unit was equivalent to a double stranded DNA concentration of 50µg/ml and an RNA concentration of 40µg/ml.

ii. Gel electrophoretic determination.

Following purification of DNA fragments, the recovery was estimated by agarose gel electrophoretic analysis (section 2.3.1.) of an aliquot of the DNA solution alongside doubling dilutions of a stock of linearised pUC12 of known concentration. This analysis has the advantage of revealing any contamination by other DNA sequences in the DNA fragment preparation.

2.9.2. Concentration of nucleic acids.

Solutions containing nucleic acids were adjusted to 0.3M with respect to monovalent cations using a 3M sodium acetate pH5.0 stock. Precipitation was achieved by addition of 2.5 volumes ethanol followed by chilling at -20°C overnight or at -70°C for 30 min. The precipitates were collected by spinning at 0°C for 10 min in a microfuge or for 10 to 30 min at 12 000g. The pellet was then washed with 70% v/v ethanol and dried.

2.9.3. Preparation of solutions for organic extraction of nucleic acid solutions

i. Phenol.

Water-saturated phenol was made 0.1% w/v with respect to 8-hydroxyquinoline, extracted once with 1M tris-HCl, pH8.0 and then twice with 0.1M tris-HCl, pH8.0; 0.2% v/v ß-mercaptoethanol.

ii. Chloroform.

The solution referred to in these methods as chloroform is in fact a mixture of chloroform and iso-amyl alcohol (24:1 v/v).

iii. Phenol-chloroform.

This is a 1:1 v/v mixture of the phenol and chloroform solutions referred to above.

Chapter 3

Results

The discussion in Chapter 1 has provided an outline of the cell biology of the murine erythropoietic system and information concerning the current state of knowledge of eukaryotic transcriptional control.

The eventual aim of the analysis of GSHPx gene expression is to elucidate the mechanisms underlying the different levels of expression of the GSHPx gene in diverse cell types. In particular it is of interest to determine how the gene is expressed at a high level in erythroid cells and whether its expression is co-ordinated with that of other "erythroid" functions.

It was therefore decided that the analysis should begin with a structural characterization of the ep19 (GSHPx) gene. Following this, determination of the possible mechanisms leading to varying cytoplasmic levels of ep19 (GSHPx) mRNA was performed by analysis of the RNA from different cell types. Regions of the DNA surrounding the GSHPx gene which were identified as potentially important to its regulation either by sequence homology to other transcriptional regulatory elements or by virtue of their position immediately 5' to the major transcription initiation site were tested for their effect upon transcription by short-term expression analysis.

Since different transcriptional states of genes are reflected in altered chromatin structures around the gene of interest (section 1.3.3) the DNAse I hypersensitivity of the GSHPx domain in different cell types which expressed the GSHPx gene at high or low levels was examined. In this way it was hoped that information regarding the mechanisms by which the GSHPx gene was

expressed in different cell types could be obtained and that this might pinpoint regions of the DNA surrounding the GSHPx gene which might be binding sites for trans-regulatory proteins. The identification of such putative trans-acting factor binding sites would be an important step towards understanding whether transcription of the GSHPx gene was co-ordinated with transcription of other genes.

3.1. Characterisation of the ep19 gene.

3.1.1. Restriction mapping λR68A.

The three Eco RI fragments of the genomic DNA recombinant (λR68A) encoding the RBC 19kD polypeptide (ep19) had been subcloned into pAT153 and a limited restriction map of the clone was already known (Goldfarb et al, 1983). However, to provide a more detailed map and also to produce subcloned probes, the four Bam HI fragments of the 12kb Eco RI fragment and the three Xba I fragments of the 2.2kb Eco RI fragment were subcloned into pUC plasmids. Using these subclones the restriction map shown in Figure 3.I.1 was obtained.

3.1.2. ep19 gene sequence : identification of ep19 as glutathione peroxidase (GSHPx).

To further characterize the ep19 gene the two cDNA recombinants pFC5 and pFA6 and the Xba I - Eco RI 0.8kb and the Eco RI - Eco RI 0.7kb fragments of λR68A known to contain the 5' end (P. Goldfarb, personal communication) and at least part of the ep19 gene (Goldfarb et al, 1983) were subcloned into M13 and sequenced according to the strategy shown in Figure 3.I.2. Screening the nucleic acids data base with the sequence so obtained did not reveal any significant homologies. However, when translations of the ep19 sequence were used to screen the protein data base a significant homology (86%) with the complete amino acid sequence of bovine erythrocyte glutathione peroxidase (GSHPx) was obtained (Gunzler et al, 1984; Figure 3.I.3.). The sequence of a fragment of rat liver GSHPx (Condell and Tappel, 1982) is

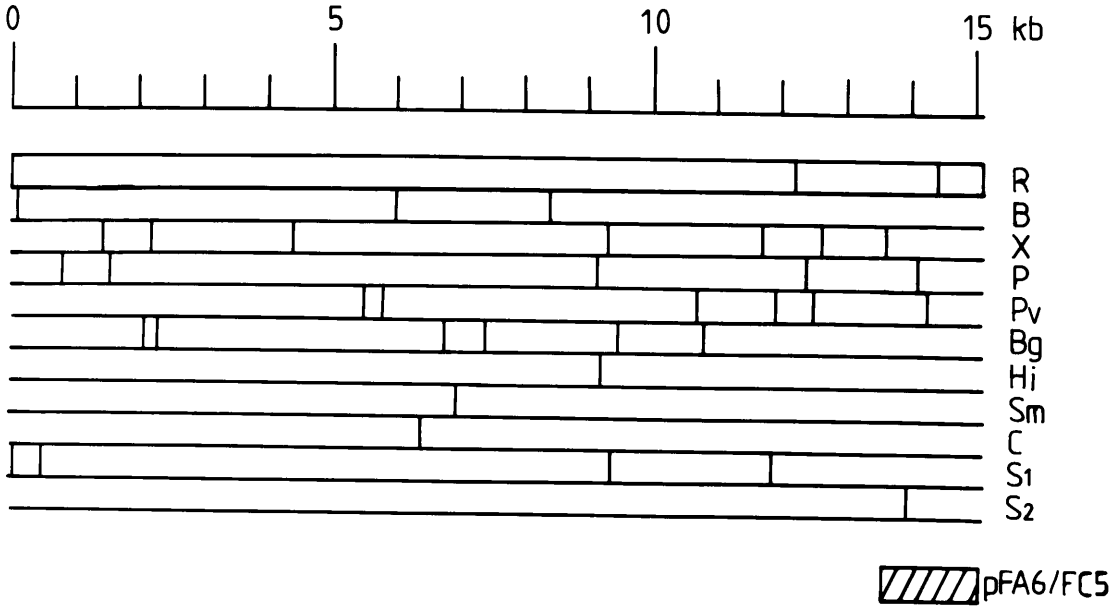


Figure 3.I.1. : Restriction map of λ R68A.

The limits of the cloned DNA are defined by two Eco RI sites created by linker addition to a DNA fragment produced by partial Hae III digestion of BALB/c genomic DNA. Vertical lines in each row indicate positions where the DNA is restricted by the corresponding enzyme. Only enzymes which were tested against all the λR68A subclones are shown. No restriction sites for Sal I, Xho I or Kpn I were detected (note: within the 5.9kb Bam HI fragment there are four Xba I sites although only three are shown. The fourth lies about 70 b.p. from one of the others, as yet unidentified). The striped bar indicates the region of hybridization to the cDNAs pFC5 and pFA6 (Goldfarb et al., 1983), both of which are derived from a mRNA encoding a polypeptide of Mr 19kD (Affara et al., 1983).

The abbreviations used for restriction enzymes in this Figure and throughout this thesis are listed below.

A,	Acc I	Av,	Ava I	B,	Bam HI	Bc,	Bcl I	Bg,	Bgl II
Bs,	Bst EII	C,	Cla I	H,	Hae III	Hi,	Hind III	Hp,	Hpa II
K,	Kpn I	M,	Msp I	Na,	Nar I	Nc,	Nco I	P,	Pst I
Pv_1,	Pvu I	Pv,	Pvu II	R,	Eco RI	R_2,	Eco RII	Rs,	Rsa I
S,	Sau 3 A1	Sl,	Sal I	Sm,	Sma I	Sp,	Sph I	St,	Stu I
S_1,	Sst I	S_2,	Sst II	T,	Taq I	X,	Xba I	Xh,	Xho II
Xm,	Xmn I								

An unbound copy of this Figure is enclosed within this thesis.

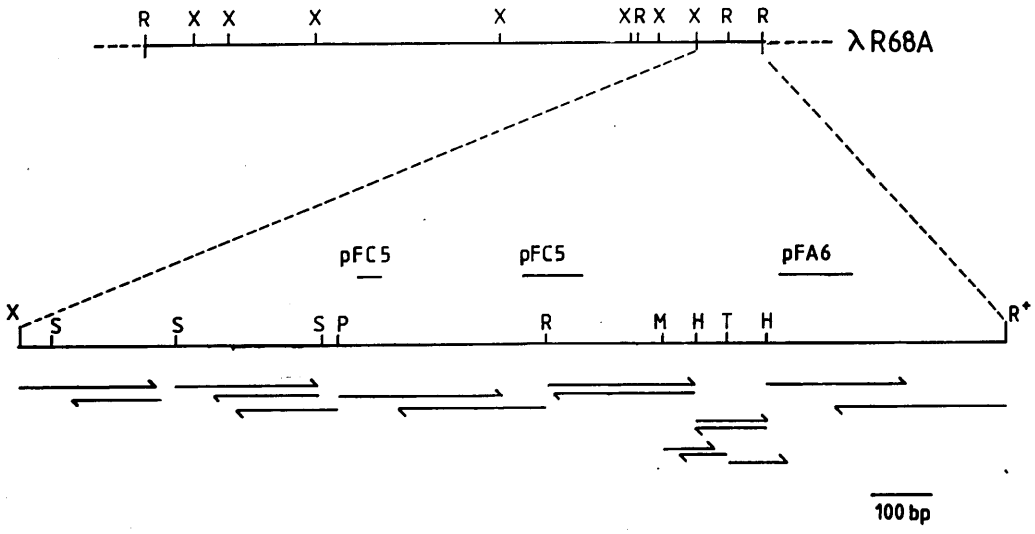


Figure 3.I.2. Restriction map of the ep19 gene and sequencing strategy.

Top: map of λR68A, the genomic DNA recombinant containing the gene (Goldfarb et al., 1983). Middle: enlarged map of the gene showing restriction sites relevant to the sequencing strategy. Regions homologous to cloned cDNAs are shown by overlining. Bottom: sequencing strategy. Arrows indicate the extend and direction of sequencing of sublcones. Abbreviations: As for Figure 3.I.1.

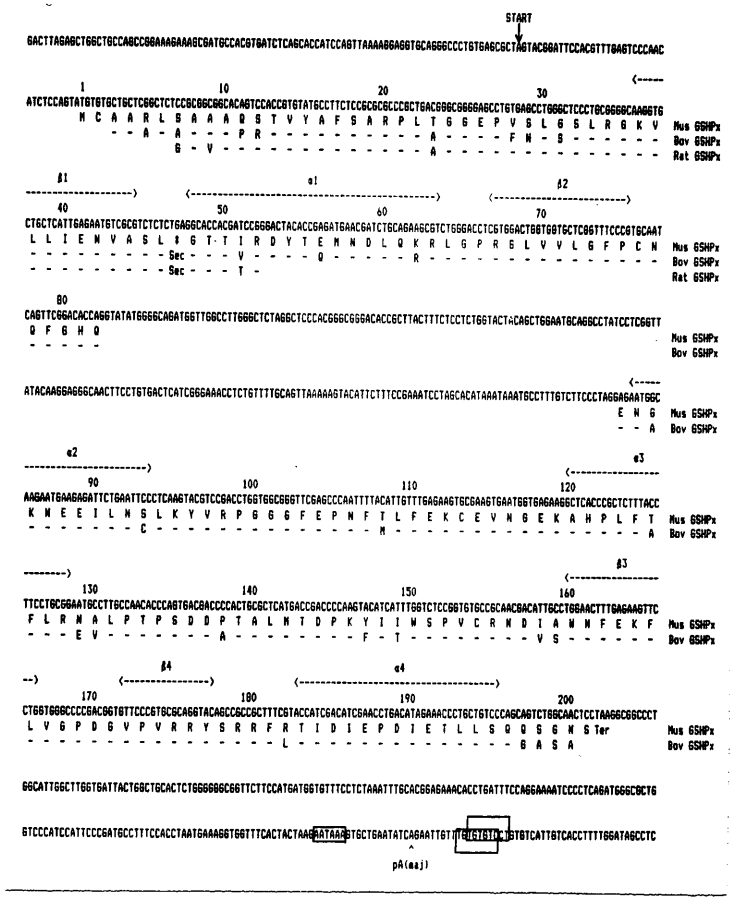


Figure 3.I.3. Sequence of the mouse GSHPx gene compared to the known sequences of GSHPx proteins.

The nucleotide and deduced amino acid sequences of mouse GSHPx are aligned with those known for the bovine and rat enzymes. The position of the in-frame TGA codon is marked by an asterisk and occurs at the same position as selenocysteine (Sec) in the protein. Numbering of amino acids begins at the initiator methionine. The major sites of transcription initiation and polyadenylation are marked START and pA (maj) respectively. The latter is preceded by a consensus polyadenylation signal (Proudfoot and Brownlee, 1976) and two overlapping sequences which match the consensus sequence (PyGTGTTPyPy), shown to be important for efficient polyadenylation (McLauchlan et al, 1985); these sequences are boxed. Arrows above the sequence define the limits of elements of secondary structure determined for the bovine enzyme (Epp et al, 1983). The original nomenclature of the elements has been retained; α, α-helix; β, β-strand; numbering increases from N- to C- termini.

also known and within this region the rat enzyme is more homologous (91%) to the deduced mouse GSHPx sequence than is the bovine enzyme (83%). These homologies increase to 95% and 89% respectively if the highly variable N and C termini (corresponding to the first 11 and last 7 amino acids of the murine sequence) are ignored. These extreme segments of the protein are not considered to have any structural or catalytic role (Gunzler et al, 1984; Chaudiere et al, 1984); indeed they were not detected in the GSHPx crystal lattice during X-ray diffraction (Epp et al, 1983). Within this central portion of the protein, twenty amino acid substitution have occurred between the murine and bovine sequences. However, neither of the residues postulated as forming hydrogen bonds to the selenocysteine selenolate (glutamine 82 and tryptophan 160) have been altered, nor have the arginines thought to form salt bridges to the carboxylate groups of GSH (residues 52 and 179) (Epp et al, 1983). The arginine at position 52 is also conserved in rat GSHPx.

This sequence comparison also reveals the presence of a single intron within the protein coding region of the GSHPx gene: this is corroborated by an interruption in one of the GSHPx cDNA sequences (pFC5: see Figure 3.I.2), when compared to the sequence of the gene. Moreover, the intron boundary sequences conform to the consensus splice site sequences (Padgett et al, 1986).

3.1.3. The selenocysteine in the active site of GSHPx is encoded by UGA.

The most surprising point to emerge from this analysis was that the selenocysteine in the active site of the enzyme appeared to be encoded by one of the usual "stop" codons, UGA. The accuracy of the sequence of the genomic recombinant in this region was first confirmed by sequencing both strands of the DNA. However, the possibility remained that a mutation had occurred during isolation of the genomic recombinant to produce the UGA codon. In this case selenocysteine would be encoded, not by UGA but by another codon, perhaps one of the normal cysteine codons (UGU, UGC). Unfortunately, this could not be excluded from the sequences of the two GSHPx cDNAs since neither is derived from this region of the mRNA.

To clarify this point, the sequence of the mRNA in the relevant region was determined by reverse transcription under dideoxy sequencing conditions using a short single-stranded 23-nucleotide (nt) primer (the Msp I - Pst I fragment 18-40nt downstream from the putative TGA codon; Figure 3.I.3). The result (Figure 3.I.4) confirms that the GSHPx mRNA sequence encoding the selenocysteine residue is in fact UGA. There is no evidence for any further bands in the primer extension sequencing ladder other than very minor components present in the control lane. In contrast, the termination codon used in the GSHPx mRNA is UAA (Figure 3.I.3).

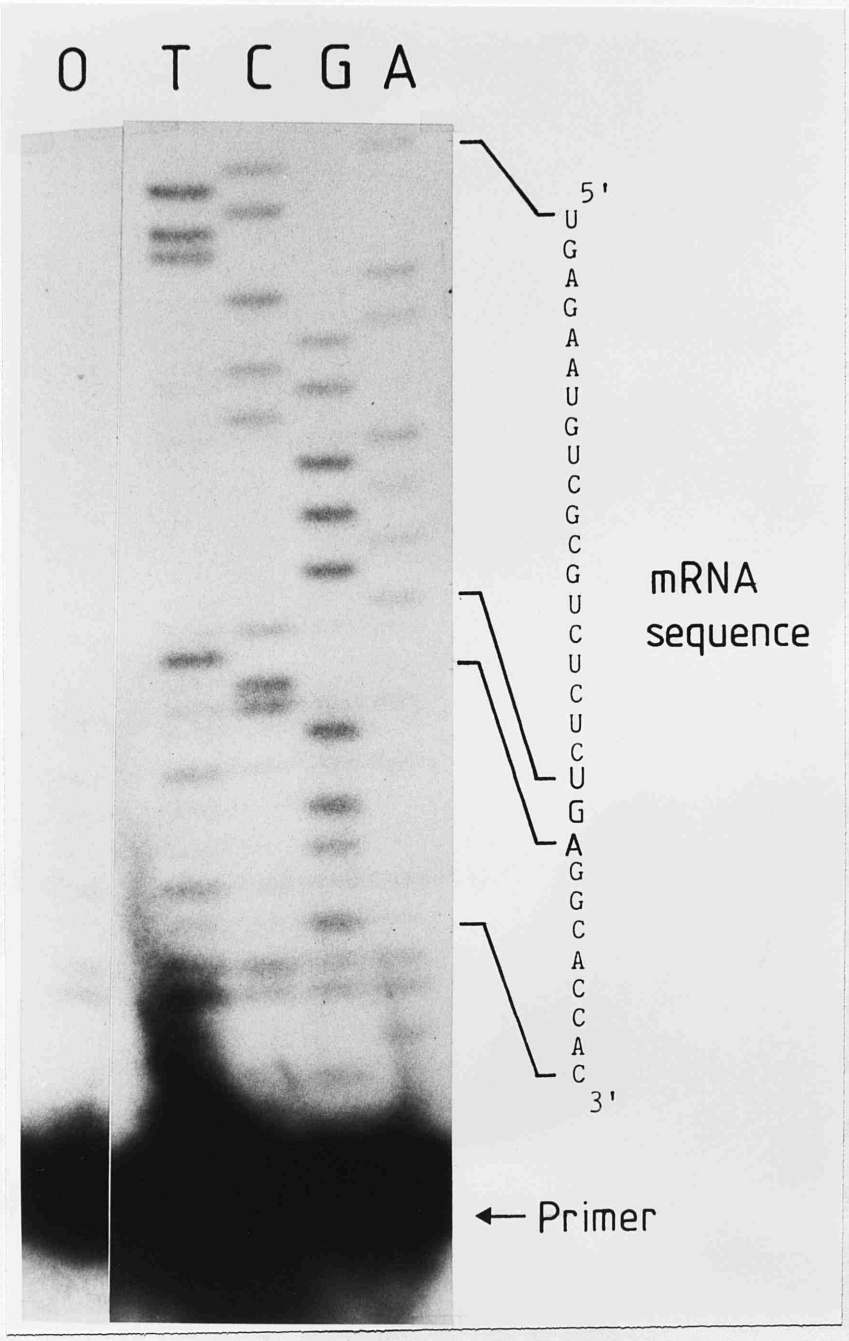


Figure 3.I.4. Dideoxy sequencing of GSHPx mRNA around the selenocysteine codon.

The reactions were performed as described in Chapter 2. The primer is the 5' end labelled 23 b.p. strand of the Msp I - Pst I fragment described in the text. Lanes T, C, G, A and O are the products of reactions containing ddTTP, ddCTP, ddGTP, ddATP and no ddNTPs, respectively. The sequence shown at the side is that deduced for the mRNA. The position of the UGA codon is indicated.

3.1.4. Potential regulatory sequences around the GSHPx gene.

The GSHPx gene sequence from -290 to +1234nt was examined for the presence of potential transcriptional regulatory motifs (Figure 3.I.5). A sequence with good homology to the "ATA" sequence (consensus TATARAR; Corden et al, 1980) shown by in vitro mutagenesis to be important in determining the efficiency and accuracy of transcription initiation (Breathnach and Chambon, 1981) is present between -31 and -28nt (TTAAAA). In contrast, no matches to the "CAAT" sequence (consensus GGPyCAATCT; Benoist et al, 1980) exist upstream of the GSHPx major transcription initiation site.

The upstream region of the gene, from -260 to +1nt is G-C rich (61%) with the region from -200 to -70nt being particularly so (67%). Similar G-C richness is seen in the promoter regions of a number of genes expressed in a wide variety of cell types, the so-called "housekeeping" genes such as those encoding 3-hydroxy-3 methylglutaryl CoA reductase (Reynolds et al, 1984) hypoxanthine phosphoribosyltransferase (Melton et al, 1984) and adenosine deaminase (Valerio et al, 1985). This region of the GSHPx gene also resembles the promoters of such genes by its possession of five sequences having 8/10nt homology to the decanucleotide consensus sequence for binding of the transcription factor Sp1 (Kadonaga et al, 1986). In addition to these, there are another seven such sequences downstream of the major transcription initiation site showing at least the same homology to this consensus (Figure 3.I.5). Of all of these sequences, five contain the Sp1 hexanucleotide core sequences : these homologies are

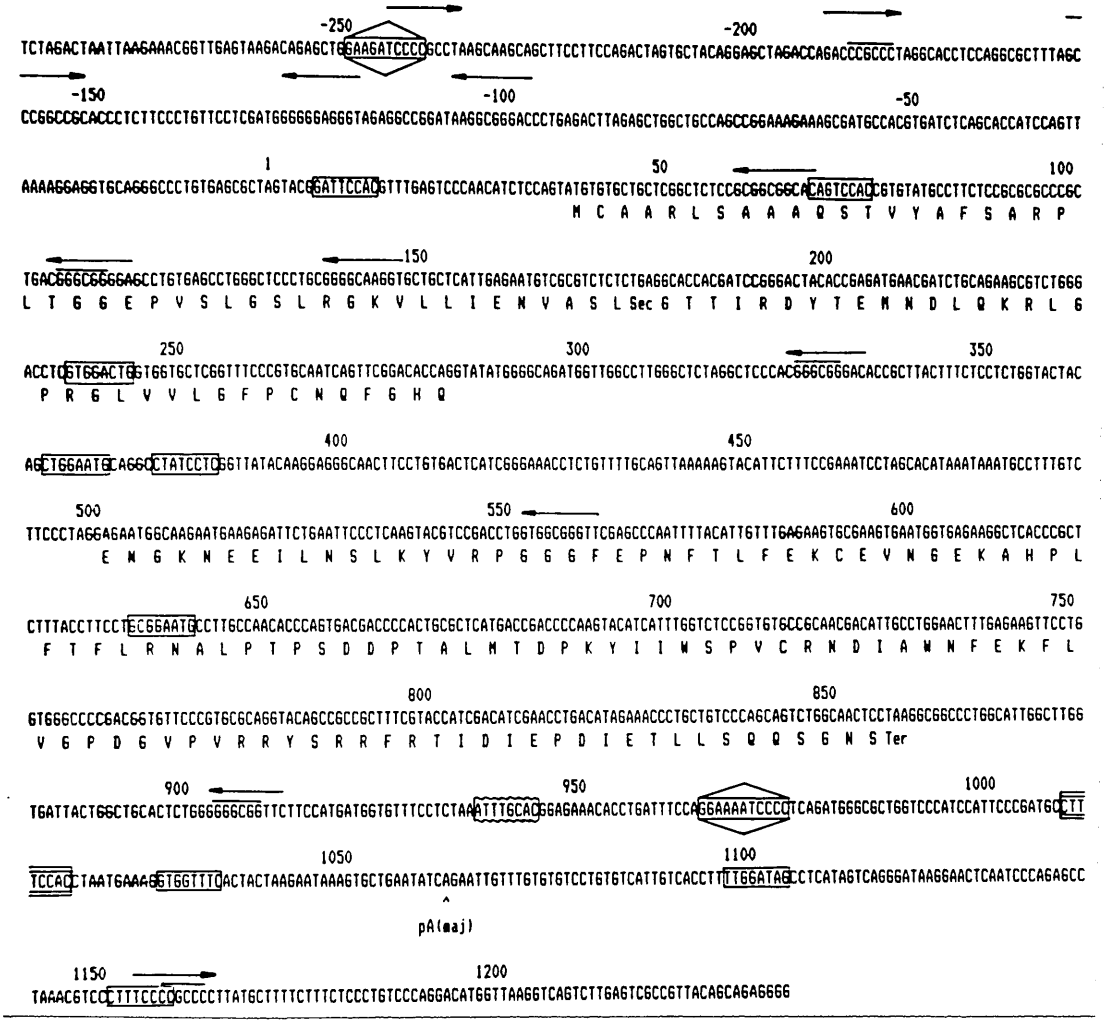


Figure 3.I.5 Sequence of the GSHPx gene : potential regulatory elements.

The nucleotide sequence of the GSHPx gene is shown. To aid orientation the deduced amino acid sequence and the position of polyadenylation of the mRNA are also shown. Numbering of nucleotides begins at the major position of transcription initiation and increases in the 5' to 3' direction of the coding strand. Sequences with 8-9/10 nt homology to the Sp1 binding site consensus sequence are overlined with arrows (←——— = WGGGCGGPuPuPy) and those which also possess the hexanucleotide core (GGGCGG) have that sequence overlined. Sequences with at least 7/8 nt homology to the SV40 core enchancer sequence (GTGGRRRG) are boxed: the copy showing perfect homology is double boxed. Sequences with 9-10/11nt homology to the NFk-B binding site consensus sequence (GGGGACTTTCC) are boxed by open triangles while the sequence with 7/8nt homology to the immunoglobulin octamer sequence (ATTTGCAT) is boxed by squiggly lines.

summarised in Table 3.I.1.

Interestingly, the similarity to this class of genes can be taken further since the regions 24nt downstream of the "ATA" sequences of several of these genes can be fitted to a consensus sequence (Martini et al, 1986). This consensus was arrived at after examining the sequences of eleven housekeeping genes. Eight of these possessed sequences which showed some homology to the "ATA" box sequence 20-30nt upstream from their main transcription initiation site and it was by comparison of these genes that the consensus was deduced. However, the significance of this observation is dubious since the homology to the consensus was quite strained in some cases and appears to rely mostly upon GC content. Moreover the number of sequences examined was limited. Nevertheless, the homology in all cases was better than 64% and for GSHPx is 71% (Table 3.I.2).

There are also a number of sequences around the GSHPx gene which show a 7/8nt homology to the SV40 core enhancer sequence (Weiher et al, 1983) as well as one sequence at +1009nt which is a perfect match to this consensus (Table 3.I.1). Interestingly, this sequence and the related inverted sequence at +1027nt are centred within a 250nt sequence flanked by two oppositely orientated Sp1 hexanucleotide core sequences (Figure 3.I.5). A sequence with 7/8nt homology to the conserved octanucleotide motif (ATTTGCAT) found 70nt upstream of the transcription initiation site in immunoglobulin genes (Parslow et al, 1984) and also within the immunoglobulin μ enhancer is also located within this 250nt sequence (at +936nt). This sequence would appear to have a more

Table 3.I.1. Consensus sequence homologies within the GSHPx gene.

Sp1 decanucleotide consensus sequence[o] (WGGGCGGPuPuPy) homologies.

Position	Sequence	% homology	SP1 hexanucleotide core
-245	TCCCCGCCTA	80	No
-191	GACCCGCCCT	80	Yes
-161	GCCCGGCCGC	80	No
-128	GGGGGAGGGT	80	No
-107	AAGGCGGGAC	80	No
+59	GCGGCGGCAC	80	No
+105	CGGGCGGGGA	80	Yes
+138	GGGGCAAGGT	80	No
+325	CGGGCGGGAC	90	Yes
+552	GTGGCGGGTT	80	No
+903	GGGGCGGTTC	80	Yes
+1154	TCCCCGCCCC	90	Yes

SV40 core consensus sequence* (GTGGRRRG) homologies.

Position	Sequence	% homology
+7	GATTCCAC	88
+48	CAGTCCAC	88
+237	GTGGACTG	88
+364	CTGGAATG	88
+377	CTATCCTC	88
+634	GCGGAATG	88
+1009	CTTTCCAC	100
+1097	TTGGATAG	88
+1151	CTTTCCCC	88

Octamer sequence[+] (ATTTGCAT) homology.

+936	ATTTGCAC	88

NF-kB binding site[x] (GGGGACTTTCC) homologies.

-249	GGAAGATCCCC	82
+964	GGAAAATCCCC	91

[o]Kadonaga et al, 1986; [*]Weiher et al, 1983; [+]Parslow et al, 1984; [x]Sen and Baltimore, 1986

The positions at which the homologies occur in the GSHPx gene are indicated according to the numbering used in Figure 3.I.5. Nucleotides not matched to the relevant consensus sequence are underscored. For sequences homologous to the Sp1 binding site, those also possessing a perfect match to the Sp1 hexanucleotide core sequence (GGGCGG; Gidoni et al,1984) are indicated.

Table 3.I.2. Homology between a housekeeping gene consensus sequence and the GSHPx gene sequence downstream from the "ATA" box.

Consensus[+]	TATAAAWGCGGCCGCCGCGGCWCGGPyG-SCS
	* ***** ** ** * *** * ** * ***
GSHPx	TTAAAAGGAGGT-GCAG-GGCCCTGT GAGCG

% agreement 24 b.p. downstream from "ATA" sequence = 71%

"ATA" homologies are underscored, matched nucleotides are indicated by *

[+]The consensus sequence (Martini et al, 1986) was deduced by comparison of sequences from the following eight genes; human dihydrofolate reductase (Chen et al, 1984); chicken thymidine kinase (Kwoh and Engler, 1984); human Cu/Zn superoxide dismutase (Levanon et al, 1985); chicken glyceraldehyde-3-phosphate dehydrogenase (Stone et al, 1985); chicken triosephosphate isomerase (Straus and Gilbert, 1985); human adenosine deaminase (Valerio et al, 1985); chicken hepatic 5-amino levulinate synthase (Maguire et al, 1986) human glucose-6-phosphate dehydrogenase (Martini et al, 1986).

general function than just to operate on B-cell specific genes since it is also found within the promoters of ubiquitously expressed genes (Sive et al, 1986) as well as within the SV40 enhancer (Bohmann et al, 1987). In fact, recent work has shown that there are at least two distinct factors which can interact with this sequence; one is B-cell specific whereas the other does not seem to have any cell specificity (Singh et al, 1986; Staudt et al, 1986; Bohmann et al, 1987). It is intriguing however that the octamer homologous sequence present in the GSHPx gene is located only 28nt upstream from a sequence having 10/11nt homology to the NF-kB binding sequence present in the immunoglobulin kappa enhancer (Sen and Baltimore, 1986a). This sequence binds a factor thought originally to be B-cell specific (Sen and Baltimore, 1986a) but since shown to be present in several other cell types in an inactive form which can be activated by post-translational modification (Sen and Baltimore, 1986b). More recently, this factor has been shown to interact with sequences in the human immunodeficiency virus long terminal repeat following activation of a T-cell leukaemia line, leading to a 50 fold elevation of expression of the linked gene (Nabel and Baltimore, 1987).

Palindromic sequences have been known for some time to bind regulatory proteins in prokaryotes (reviewed by Pabo and Sauer, 1984). More recently, palindromes present in the adenovirus major late promoter (Carthew et al, 1985) and the immunoglobulin enhancers (Church et al, 1985; Ephrussi et al, 1985; Sen and Baltimore, 1986a) have been shown to be binding sites for nuclear factors. The GSHPx gene sequence was therefore scanned for such

motifs. One such dyad symmetry, which has a 5/6nt homology to itself on either side of a 2nt spacer is centered at +953nt, that is, between the sequences with homology to the octamer motif and the NF-kB binding sequence. This is curious, since the palindrome is of the CA/TG type found in the immuoglobulin enhancers which are binding site for protein *in vivo* (Church *et al*, 1985; Ephrussi *et al*, 1985). Other palindromes were also found (Table 3.I.3). One of these, which is present close to the other sequence homologous to the NF-kB binding sequence is centred at -228nt and extends 8nt on either side of a central A with only 1 mismatch. The remaining dyad symmetries detected, including a perfect palindrome covering 10nt centred at -13nt, are listed in Table 3.I.3.

3.1.5. Relationship of sequences within the GSHPx gene to other sequences in the mouse genome

By utilising the technique of southern blot hybridisation, relationships between sequences within and upstream of the GSHPx gene and other sequences in the mouse genome were established.

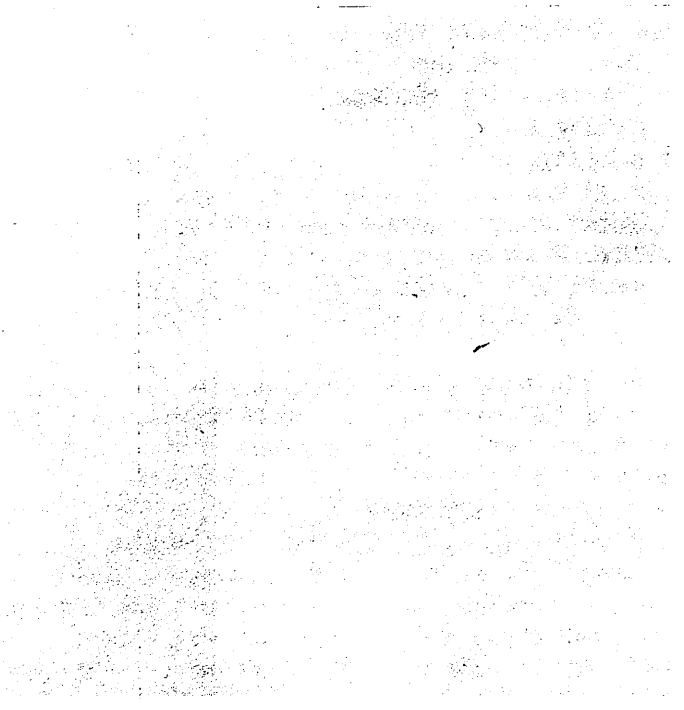
DNA, purified from the Friend cell line M707T was digested with various restriction enzymes, electrophoresed through agarose and transferred to a nitrocellulose filter. This was then hybridised under normal conditions of stringency (50% v/v formamide; 5xSSC; 42°C) to the 706bp *Eco* RI, exon 2 probe, but was washed at reduced stringency (0.5xSSC; 37°C) (Figure 3.I.6). This shows that there are no sequences in the mouse genome closely related to GSHPx exon 2 other than the GSHPx gene itself. Bands

Table 3.I.3. Palindromic sequences around the GSHPx gene.

Position[+]	Palindrome[*]	Spacer length	Homology on either side of spacer (nts)
-228	AAGCAAGCaGCTTCCTT	1	7/8
-13	CAGGGCCCTG	0	5/5
+953	GAAA-CAccTGATTTC	2	5/6
+1028	TAATGAAAggtggTTTCACTA	5	7/8
+1189	CCCTGTCCcaGGACATGG	2	7/8

[+]Indicates the position of the palindromes central nucleotide according to the numbering of Figure 3.I.4.

[*]Spacers are shown in lower case, mismatched nucleotides are underlined and gaps introduced into the sequence are indicated by dashes.



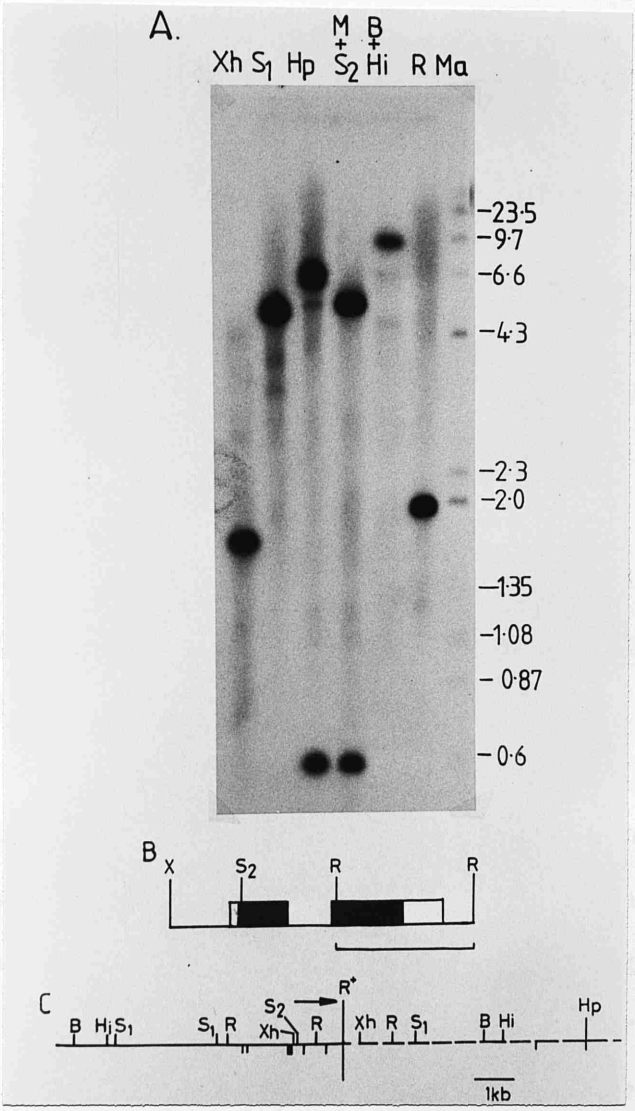


Figure 3.I.6. Detection of sequences related to the GSHPx gene.

DNA purified from the Friend cell line M707T was digested with the various restricition enzymes indicated and electrophoresed through a 1% w/v agarose gel cast and run in TBE buffer, together with end labelled Hind III digested λDNA Mr markers (Ma). Following transfer to a nitrocellulose filter the DNA was hybridised overnight at normal stringency to the 706 b.p. Eco RI probe. The filter was then washed twice 2xSSC, 0.1% w/v SDS at room temperature for 15 min and twice in 0.5xSSC, 0.1% w/v SDS at 37°C for 30 min. After air-drying, the filter was then exposed at -70°C overnight to Kodak XAR-5 film (A).

A line diagram (B) showing the GSHPx gene and the origin of the probe is shown in the middle. Below this (C) is a restriction map of the GSHPx locus, showing only the restriction sites pertinent to this experiment. The solid line corresponds to cloned DNA, the dashed line to uncloned DNA: these are separated by a vertical line indicating the synthetic Eco RI site (R^+) used to clone the GSHPx genomic DNA. Vertical lines above the horizontal indicate the restriction sites shown. Those below the line show recognition sites for Msp I (the thick bar is three Msp I sites). The extent of the GSHPx gene (from transcription initiation site to polyadenylation site) and the direction of transcription are indicated by the arrow above the map.

additional to the main GSHPx bands which can be seen in Figure 3.I.6 are of dubious significance since such bands are not detectable in all the lanes. Any relationship which these bands may have to the GSHPx gene will require a more extensive analysis, perhaps by hybridisation at reduced stringency (e.g. 30-40% v/v formamide).

3.1.6. Highly repetitive sequences located upstream of the GSHPx gene.

To identify DNA fragments from around the GSHPx gene containing highly repetitive DNA sequences, DNA from the upstream subclones (see Section 3.1.1) was digested with a variety of different restriction enzymes, electrophoresed through agarose and transferred to nitrocellulose filters. These filters were then hybridised to sheared mouse genomic DNA. This type of analysis should only identify fragments of DNA containing copies of the most repetitive sequences in the mouse genome, that is, sequences represented by about $5x10^4$ - $3x10^5$ copies per genome such as those described by Bennet et al (1984). In addition, similar filters were hybridised to a probe containing a mouse B1 element, a repetitive sequence represented by approximately $0.5\text{-}1.0x10^5$ copies per genome (Krayev et al, 1980). The B1 containing probe used is a 2.8kb Bgl II fragment from the 3' end of the mouse β^{maj} globin gene, known to contain a single B1 element (Coggins et al, 1982). The analyses are presented below beginning with the furthest upstream clone and working towards the GSHPx gene. Figure 3.I.7A shows that highly repeated sequences within the

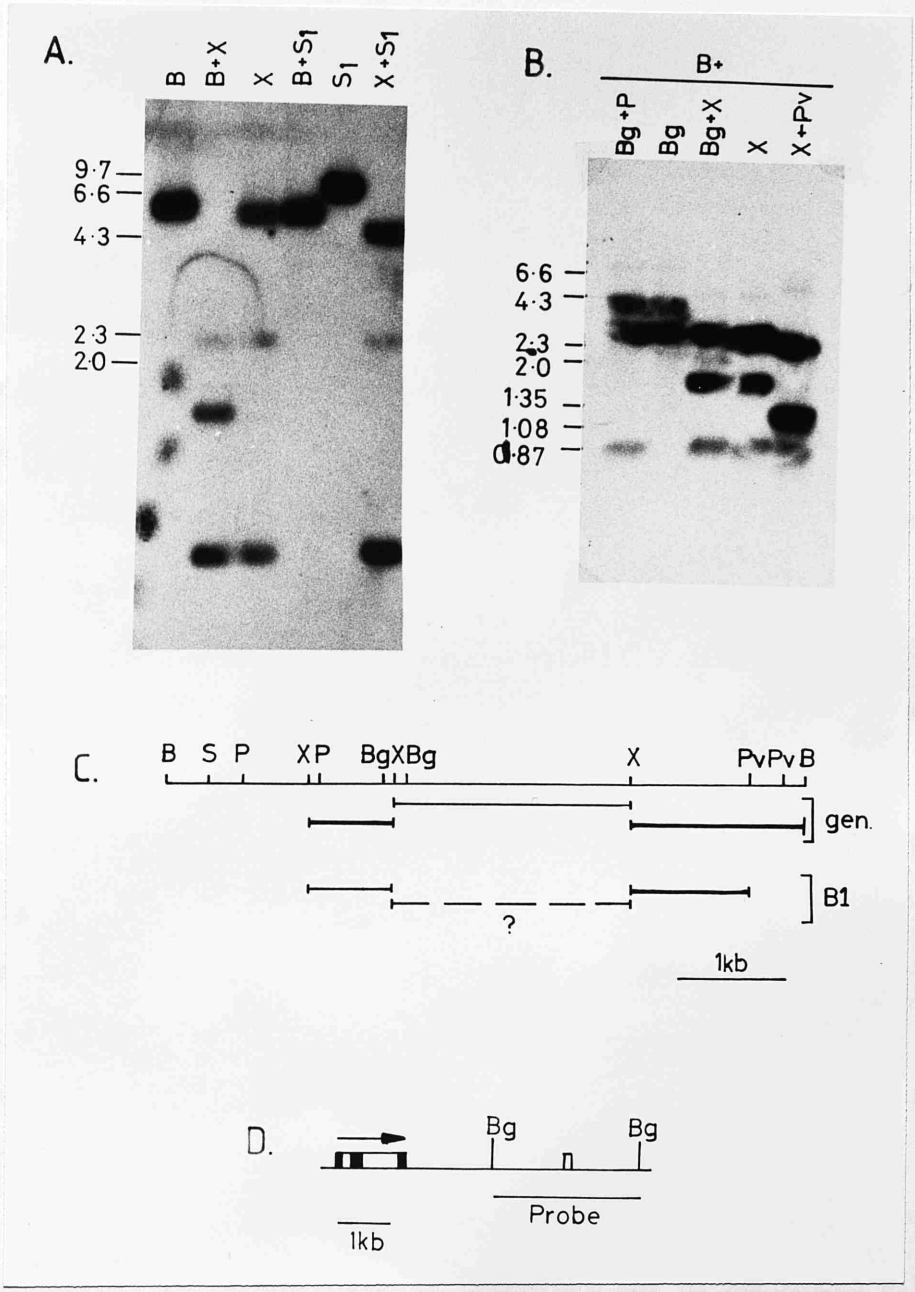


Figure 3.I.7. Analysis of repetitive sequences within the 5.9kb Bam HI fragment upstream from the GSHPx gene.

Purified DNA from the pUC8 subclone of the 5.9kb fragment was digested with the various restriction enzymes indicated and 1µg of each digest electrophoresed through 1% w/v agarose gels, cast and run in TAE buffer. Following transfer to nitrocellulose, DNA was hybridised to, A, sheared genomic DNA prepared from 707B10/1 cells; B, a B1 containing probe from the 3' end of the mouse β^{maj} globin gene. The position of the Mr markers (Hind III cut λ DNA and Hinf I cut ϕ x174 DNA) are shown. C, A restriction map of the 5.9kb fragment is shown with the regions of hybridisation to the probes indicated. Differences in intensity of hybridisation are indicated by the thickness of the corresponding line. The dashed line with the associated question mark reflects uncertainty regarding hybridisation of this region to the probe. D, the origin of the B1 probe. The β^{maj} globin gene is shown boxed (closed segments, exons; open segments, introns) and its direction of transcription indicated by the arrow. The Bgl II fragment containing the B1 sequence (open box) is indicated.

5.9kb Bam HI fragment (see Figure 3.1.1 for the relationship of this DNA to the GSHPx gene) are located within three sub-fragments produced by Xba I and Bam HI digestion. Furthermore, the relative intensities of hybridisation to each of these bands suggests that either the repetitive sequences present in the 1.6 and 0.8kb sub-fragments are homologous to a more highly repetitive element in the mouse genome than those in the 2.2kb sub-fragment or that there are a larger number of equally repetitive sequences present in the 1.6 and 0.8kb sub-fragments than in the 2.2kb sub-fragment.

At least some of this hybridisation to genomic DNA is due to the presence of B1 sequences within the 5.9kb fragment (Figure 3.I.7B). This analysis is complicated by the fact that the 5.9kb fragment was separated from the vector but was not purified away from it before being digested with other restric tion enzymes and electrophoresed. This has resulted in a relatively strong signal on the autoradiograph due to hybridisation between the vector and the small amount of pUC sequences contaminating the probe. Consequently, it is not clear if there are any B1 sequences within the 2.2kb Xba I fragment which runs close to these vector sequences. Nevertheless, it is evident that B1 elements are present in the 0.8kb Xba I fragment and the 1.6kb Xba I - Bam HI fragment. Indeed B1 sequences can be further localised to the 1.1kb Xba I - Pvu II part of the 1.6kb Xba I - Bam HI fragment, although this of course does not mean that all of the hybridisation of genomic DNA to the 1.6kb fragment need necessarily be accounted for by sequences within the 1.1kb fragment. Moreover, it should be stressed that the presence of a

B1 element within a DNA fragment which hybridises to genomic DNA does not rule out the possibility of that fragment containing other repetitive sequences, unless the fragment is close to the size of a B1 element (130bp; Krayev et al, 1980).

The difference in intensity of hybridisation of the B1 probe to the 0.8kb and 1.6kb fragments is interesting. This could be caused either by the presence of a greater number of B1 sequences in the larger fragment or by a greater homology between the probe and sequences in the larger fragment than between the probe and sequences in the smaller fragment. However, the latter possibility seems unlikely since it has been reported that B1 sequences differ from one another by less than 10% (Kramerov et al, 1985).

Figure 3.I.8 shows a similar analysis of the 2.4kb Bam HI fragment (see Figure 3.1.1 for the relationship of this DNA to the GSHPx gene). Repetitive sequences here can be deduced to lie on the Cla I - Bgl II 0.33kb, the Sma I - Bgl II 0.46kb and the Bgl II - Bam HI 1.0kb sub-fragments. However, it is not possible to compare these fragments in terms of the number or degree of repetition of the repetitive sequences they contain, by virtue of the differing intensities of hybridising bands, since, on this autoradiograph, the intensity of a given band is somewhat variable from lane to lane for some unknown reason.

This is not the case for the identical blot hybridized to the B1 probe where a strong hybridisation can be seen to the Bgl II - Bam HI fragment and a weaker hybridisation to the Cla I - Bgl II fragment. As argued above, this probably reflects the presence of

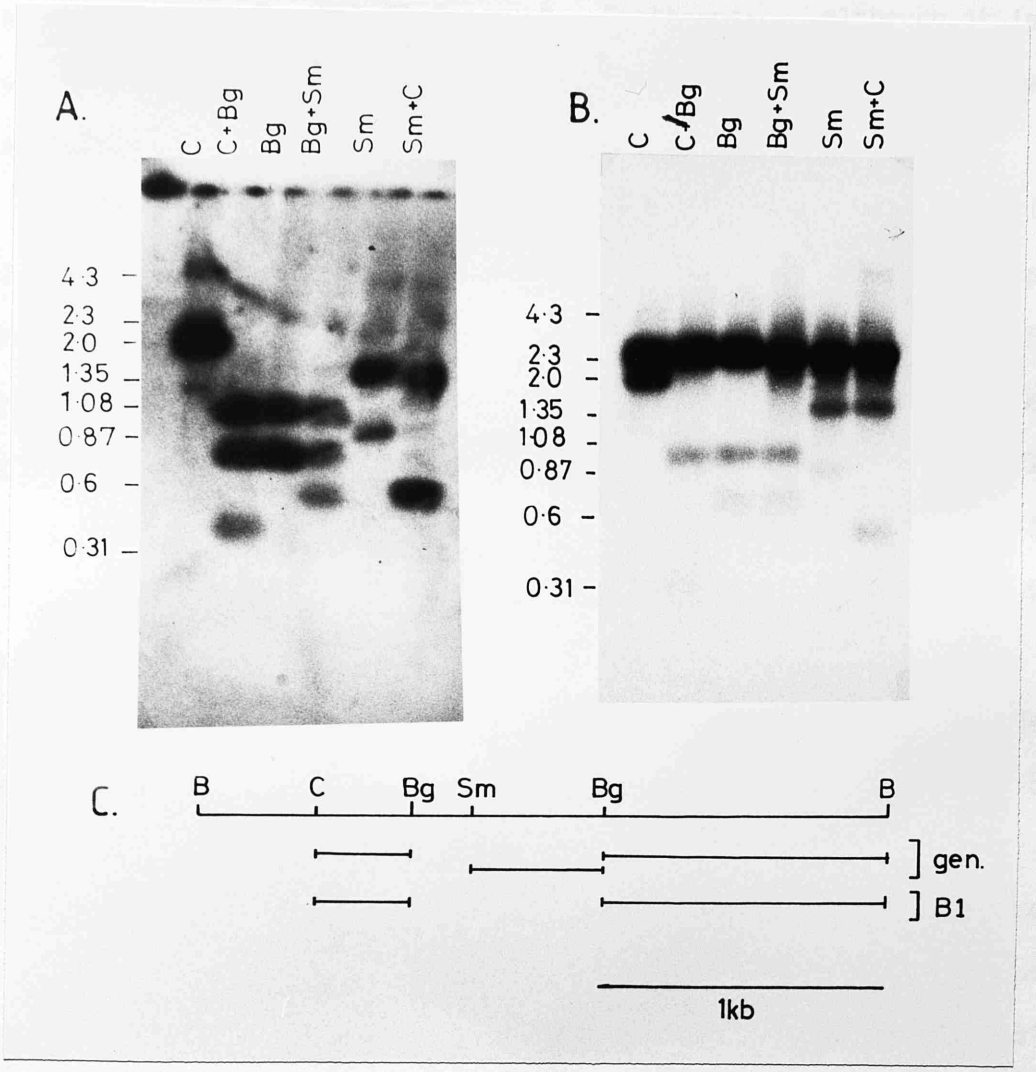


Figure 3.I.8. Analysis of repetitive sequences within the 2.4kb Bam HI fragment upstream of the GSHPx gene.

Purified DNA from the pUC8 subclone of the 2.4kb fragment was digested with Bam HI and then with the additional enzymes indicated and 1µg of each digest electrophoresed through a 1% w/v agarose gel, cast and run in TAE buffer. Following transfer to nitrocellulose, DNA was hybridised to A, sheared genomic DNA prepared from 707B10/1 cells; B, the B1 containing probe described in Figure 3.I.7 legend. The positions of the Mr markers (Hind III cut λDNA and Hinf I cut ϕx174 DNA) are shown. C. A restriction map of the 2.4kb fragment is shown with the regions of hybridisation to the probes indicated.

a greater number of B1 elements in the Bgl II - Bam HI fragment than in the Cla I - Bgl II fragment. Furthermore, although it is unclear whether these two fragments contain any non-B1 repetitive sequences, the Sma I - Bgl II fragment obviously does.

Similar analyses were conducted on the 3.8kb Bam HI - Eco RI fragment (Figure 3.I.9; Figure 3.1.1 shows the relationship between this DNA and the GSHPx gene). Although the intensity of hybridisation in each of the lanes varies, three separate fragments hybridising to genomic DNA can be discerned; the Bam HI - Xba I 0.9kb, the Xba I - Pvu II 1.35kb and the Pvu II - Xba I 1.0kb sub-fragments. Of these only the Xba I - Pvu II fragment contains B1 sequences.

A previous analysis of the 2.2kb Eco RI fragment containing the 5' end of the GSHPx gene had located B1 sequences within 0.3 to 1.2kb upstream of the transcription initiation site (P. Goldfarb, unpublished data). To locate these sequences more precisely, the 2.2kb Eco RI fragment was digested with various enzymes and a southern blot prepared and hybridised to the B1 containing probe (Figure 3.I.10). This shows hybridising signals to be located predominantly within the Acc I - Xba I 150bp fragment. Weaker hybridising bands of higher Mr are not visible on the ethidium bromide stained gel and so probably correspond to small amounts of undigested plasmid DNA. Since the B1 sequence is 130bp long (Krayev et al, 1980) the obvious explanation is that there is one B1 sequence within the 2.2kb fragment and that the major portion of this lies within the Acc I - Xba I fragment located approximately 1050bp to 1200bp upstream from the major

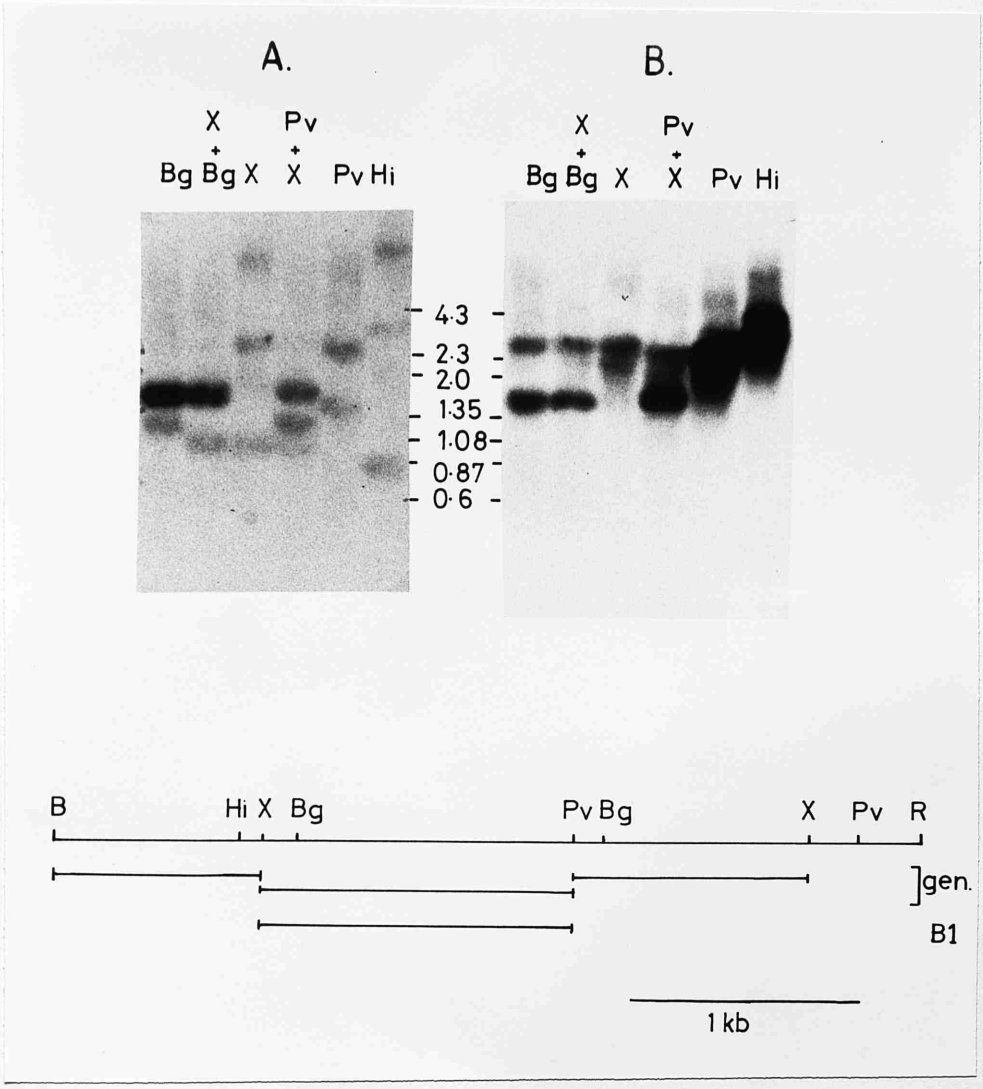


Figure 3.I.9. Analysis of repetitive sequences within the 3.8kb Bam HI - Eco RI fragment upstream of the GSHPx gene.

Purified DNA from the pUC8 subclone of the 3.8kb fragment was digested with Bam HI and Eco RI and then with the additional enzymes indicated. 1µg of each digest was then electrophoresed through a 1% w/v agarose gel, cast and run in TAE buffer. Following transfer to nitrocellulose, DNA was hybridised to A, sheared genomic DNA prepared for 707B10/1 cells; B, the B1 containing probe described in Figure 3.I.7 legend. The positions of the Mr markers (Hind III cut λDNA and Hinf I cut ϕx174 DNA) are shown. C. A restriction map of the 3.8kb fragment is shown with the regions of hybridisation to the probes indicated.

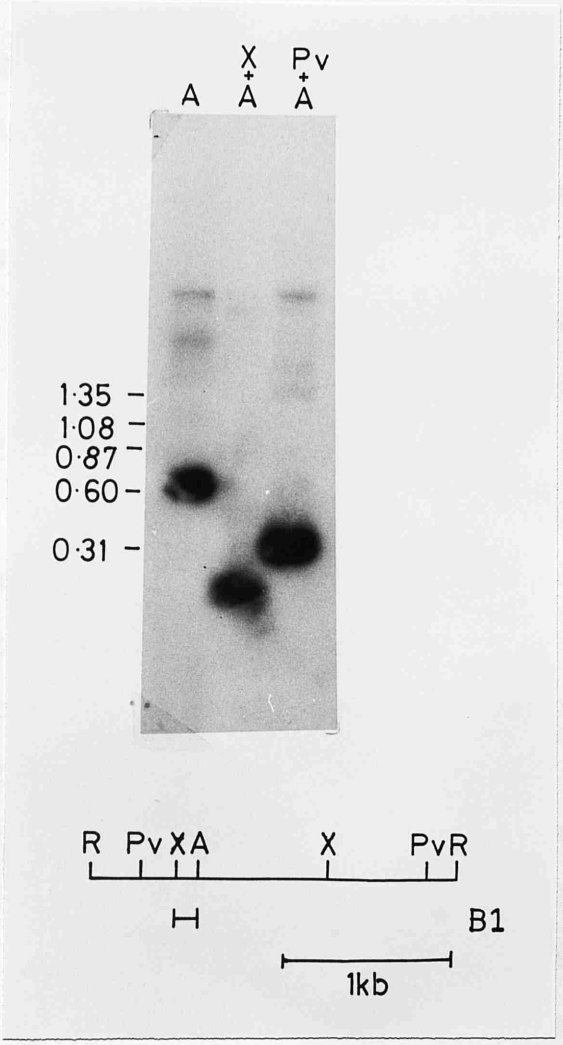


Figure 3.I.10. Analysis of B1 homologous sequences within the 2.2kb Eco RI fragment also containing part of the GSHPx gene.

The 2.2kb Eco RI fragment was gel purified from the pAT153 vector in which it was cloned, digested with restriction enzymes and 1µg DNA from each digest electrophoresed through a 1% w/v agarose gel cast and run in TBE buffer. Following transfer to nitrocellulose, DNA was hybridised to the B1 containing probe described in Figure 3.I.7 legend. The filter was exposed to Kodak XAR-5 at -70°C overnight with an intensifying screen. The positions of Mr markers (Hinf I cut ϕx174 DNA) are shown. A restriction map of the 2.2kb fragment showing the region of hybridisation to the probe is sketched below.

transcription initiation site of the GSHPx gene.

3.1.7. A lowly reiterated sequence located 1.2 - 1.7kb upstream from the major GSHPx cap site.

Southern blot experiments had shown that the Eco RI - Xba I fragment located 1.2 to 1.7kb upstream of the major cap site hybridised to seven Eco RI bands present in DNA prepared from both Friend cells and from normal BALB/c mouse tissues, under normal conditions of stringency (data not shown). This pattern of hybridisation suggests that the probe is homologous to lowly reiterated sequences such as may exist in related endogenous retroviruses or members of a gene family. As restriction sites in retroviral long terminal repeats are present at each end of the integrated retrovirus, digestion of DNA with an enzyme for which there is a site in the long terminal repeat but not in the rest of the retrovirus would produce similar sized bands for each retrovirus in the genome. DNA was therefore digested with various restriction enzymes and southern blots of the digests analysed by hybridisation to the Eco RI - Xba I probe (Figure 3.I.11A). This shows the presence of 6-8 bands of various sizes in each digest. Thus it was not possible to prove a retroviral origin for these sequences. Nevertheless it is interesting that all the family members detected are present on Pvu II fragments of less than 2kb.

By increasing the stringency of washing it was possible to establish the degree of relatedness of the family members to the probe member. As expected some family members are more homologous to the probe than others, since some of the bands are washed off

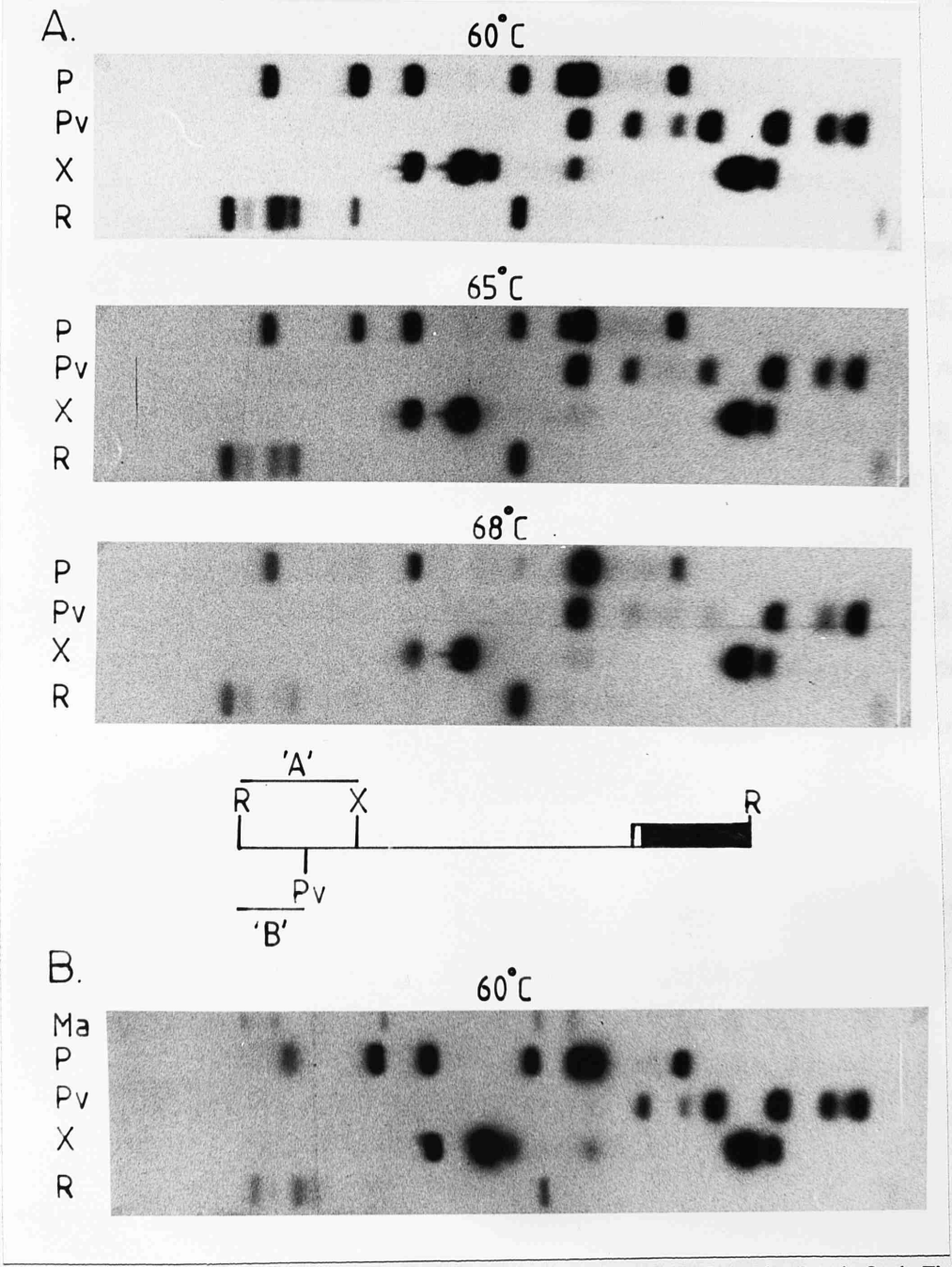


Figure 3.I.11. A lowly reiterated sequence located 1.2-1.7kb upstream from the GSHPx gene.

DNA purified from the Friend cell line M707T was digested with the indicated enzymes and 20µg electrophoresed alongside Hind III digested λDNA Mr markers (Ma) through a 0.8% w/v agarose gel, cast and run in TAE buffer. Following transfer to nitrocellulose, DNA was hybridised to either the 500 b.p. Eco RI - Xba I probe (panel A) or to the 300 b.p. Eco RI-Pvu II probe (panel B) at normal stringency overnight. Filters were then washed as normal (final wash, 0.1xSSC, 0.1% SDS, 60°C, 30 mins, twice) and exposed to Kodak XAR-5 film at -70°C for 16hrs. The filter used in panel A was then rewashed twice for 30 mins in 0.1xSSC, 0.1% SDS at 65°C and exposed to Kodak XAR-5 film at -70°C for 42hrs. This filter was then washed twice for 30 mins in 0.1xSSC, 0.1% SDS at 68°C and exposed to Kodak XAR-5 film for 65hrs. The line diagram shows a restriction map outlining the origin of the probes used in panels 'A' and 'B' and their relationship to the GSHPx gene.

the blot or decreased in intensity before others. This also suggests that the number of bands detected following a 60°C wash is a minimum estimate of the number of members of this family since it is possible that other members exist which are not sufficiently homologous to the probe to remain hybridised during a 60°C wash.

To determine whether the repetitive sequence was located 5' or 3' to the internal Pvu II site of the probe, the Eco RI - Xba I fragment was digested with Pvu II and the 5' 300bp sub-fragment gel purified and used to probe an identical blot. The result (Figure 3.I.11B) shows that all the bands seen using the Eco RI - Xba I probe are still detected at the same relative intensity, except the Pvu II fragment corresponding to the sequences which had been removed from the probe. Therefore, the repetitive sequences which define this family lie, at least partially, 5' to the Pvu II site.

3.1.8. Summary of the relationships of sequences at the GSHPx locus to other sequences in the mouse genome.

The hybridisation analyses presented in the previous sections showed that the DNA containing the GSHPx gene is essentially unique within the genome. Furthermore, the DNA located within the 14kb upstream of the GSHPx gene contains at least six, and probably more, copies of the B1 sequence of highly reiterated elements. Moreover, at least three non-B1 highly reiterated elements are also contained within this DNA. Consequently, there are only a few short DNA fragments within the 14kb upstream of the

GSHPx gene which contain potentially unique sequences. Closer examination of one such fragment shows it not to be unique but in fact to contain sequences which are repeated a relatively small number of times within the genome. It is at present unclear what function this lowly reiterated sequence may possess. It is also unclear what role, if any, the other repetitive sequences may have in relation to GSHPx gene expression.

3.2. Expression of GSHPx mRNA.

3.2.1. Tissue distribution.

The sequence data shown in Figure 3.I.3 predict that transcription of the GSHPx gene from the major transcription initiation site to the polyadenylation site would produce a polyadenylated mRNA precursor (pre-mRNA) of 1320-1360nt (assuming the usual poly A tail length of 260-300nt, Brawerman, 1981). Removal of the intron by splicing would then generate a mature mRNA of 1105-1145nt.

To examine the tissue distribution and size of GSHPx RNAs in different tissues total cellular RNA from several tissues was denatured and electrophoresed through agarose under denaturing conditions. Following transfer to nitrocellulose, GSHPx related RNAs were revealed by hybridisation to the 706bp Eco RI exon 2 probe (Figure 3.II.1). The main band detected in this analysis is approximately 1.0 - 1.15kb in length in all tissues and this is close to that anticipated for a poly A^+ spliced GSHPx mRNA. The origin of the faster migrating band present in liver, kidney and reticulocytes is uncertain; it may possibly be spliced poly A^-

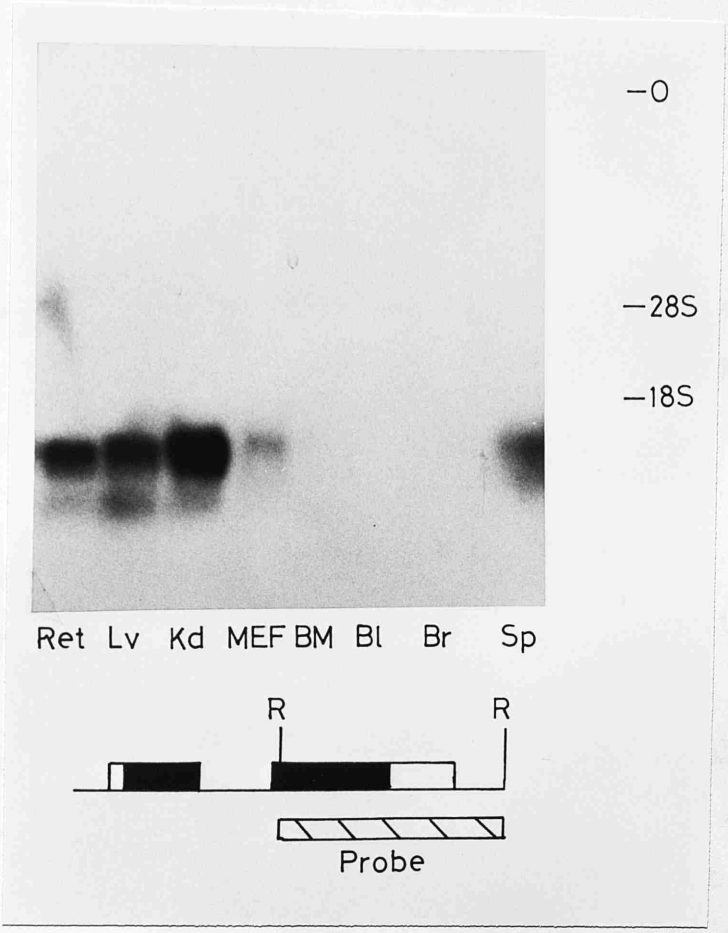


Figure 3.II.1. Tissue distribution of GSHPx mRNA.

10μg of total cellular RNA prepared from various tissues was denatured and electrophoresed through a 1.1% w/v agarose gel containing 2.2M formaldehyde, cast and run in MOPS buffer until the bromophenol blue dye had travelled 8cm. Following transfer to nitrocellulose, GSHPx related RNAs were visualised by hybridisation to the 706 b.p. Eco RI probe. After washing, the filter was exposed to Kodak XAR-5 film for 40hrs at -70°C with intensifying screens. Lines at the side of the autoradiograph mark the positions of the origin (O) and of 18 and 28S rRNAs. RNAs in each lane are from Ret, reticulocytes; Lv, liver; Kd, kidney; MEF, primary mouse embryo fibroblasts; BM, bone marrow; Bl, whole blood; Br, brain; Sp, spleen. The line diagram outlines the origin of the probe.

GSHPx mRNA, although its size (estimated at approximately 660nt) is less than expected (845nt). Such poly A^- mRNAs have been detected in Friend cells, where it seems (at least in the case of actin and tubulin mRNAs) that the poly A^- mRNAs are as stable as their poly A^+ counterparts (Krowczynska et al, 1985). Figure 3.II.1 also shows that the steady state level of GSHPx mRNA sized RNA varies between different tissues being considerably higher (20 - 50 fold by densitometric scanning) in liver, kidney and reticulocytes than in other tissues such as brain.

3.2.2. Expression in haemopoietic cells.

The steady state level of GSHPx mRNA was known to differ by less than a factor of two between uninduced Friend cells, induced Friend cells, foetal liver and reticulocytes (Affara et al, 1983; 1985). However, it was not known if this GSHPx mRNA level is established in haemopoietic stem cells, nor whether it changes during differentiation into other haemopoietic cell lineages.

To address this question, total cellular RNAs prepared from a cloned haemopoietic stem cell line (DeCl.15) and from a pregranulocytic cell line (AD.3) were analysed by the same technique used for total cellular tissue RNAs in the previous sub-section (Figure 3.II.2). This shows the GSHPx mRNA level to be lower in the cloned stem cell line than in committed maturing erythroid cells. However, the GSHPx mRNA level is higher in DeCl.15 than in either AD.3 or a neuroblastoma cell line (N18). It is unclear if this represents a decrease in the GSHPx mRNA level as the multipotent stem cell becomes committed to the

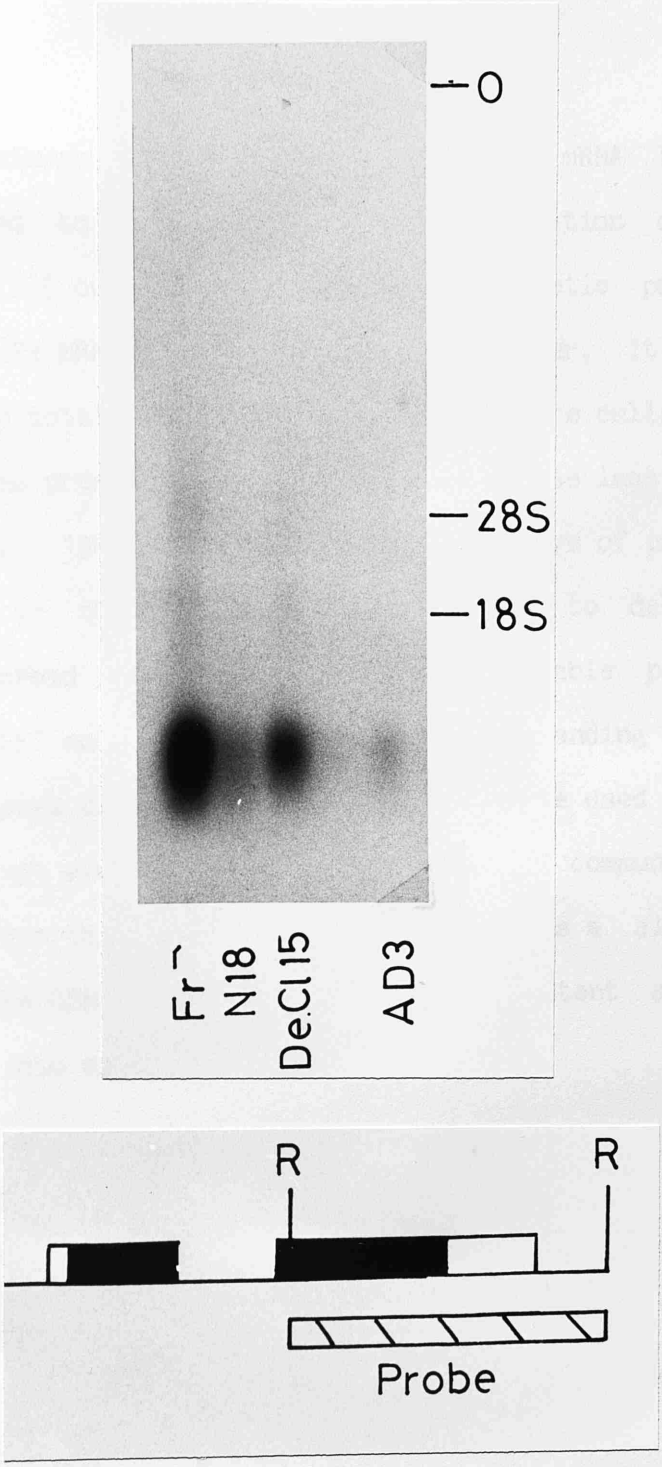


Figure 3.II.2. GSHPx mRNA level during haemopoiesis.

5µg of total cellular RNA prepared from various cell lines was denatured and electrophoresed through a 1.1% w/v agarose gel containing 2.2M formaldehyde, cast and run in MOPS buffer until the bromophenol blue had travelled 10cm. Following transfer to nitrocellulose, GSHPx related RNAs were visualised by hybridisation to the 706 b.p. Eco RI probe. Lines at the side of the autoradiograph mark the positions of the 18 and 28S rRNAs. RNAs in each lane are from the cell lines indicated. The line diagram shows the origin of the probe.

granulocytic pathway or whether the higher GSHPx mRNA level in DeCl.15 compared to AD.3 is caused by a subpopulation of cells within the DeCl.15 culture maturing down haemopoietic pathway(s) in which the GSHPx mRNA level is increased. However, it is worth noting that the total number of recognisable mature cells of all myeloid lineages present in the DeCl.15 culture was less than 1% (Spooncer et al, 1986). Moreover, colony assays of progenitor cells present in the DeCl.15 culture failed to detect any committed erythroid precursors, although a variable proportion (1-10%) of GM-CFC and CFC-Mix were detected, depending upon the particular passage examined; data for the passage used in this experiment are not available (M. Dexter, personal communication). What is clear from this experiment is that there is a significant increase in the GSHPx mRNA level as the multipotent stem cell differentiates into an erythroblast.

3.2.3. Expression in murine cell lines.

Established cell lines which express steady state levels of GSHPx mRNA comparable to those in the tissues from which they originated are likely to be useful tools with which to examine the possible role of cis-acting DNA sequences elements in producing the various different GSHPx mRNA levels characteristic of different tissues. Therefore, in an attempt to identify such cell lines, an analysis of the steady state level of GSHPx mRNA in several murine cell lines was performed by the same technique as used in the previous sub-sections (Figure 3.II.3). This shows that the GSHPx mRNA levels in a liver cell line (BW1-J) and in a

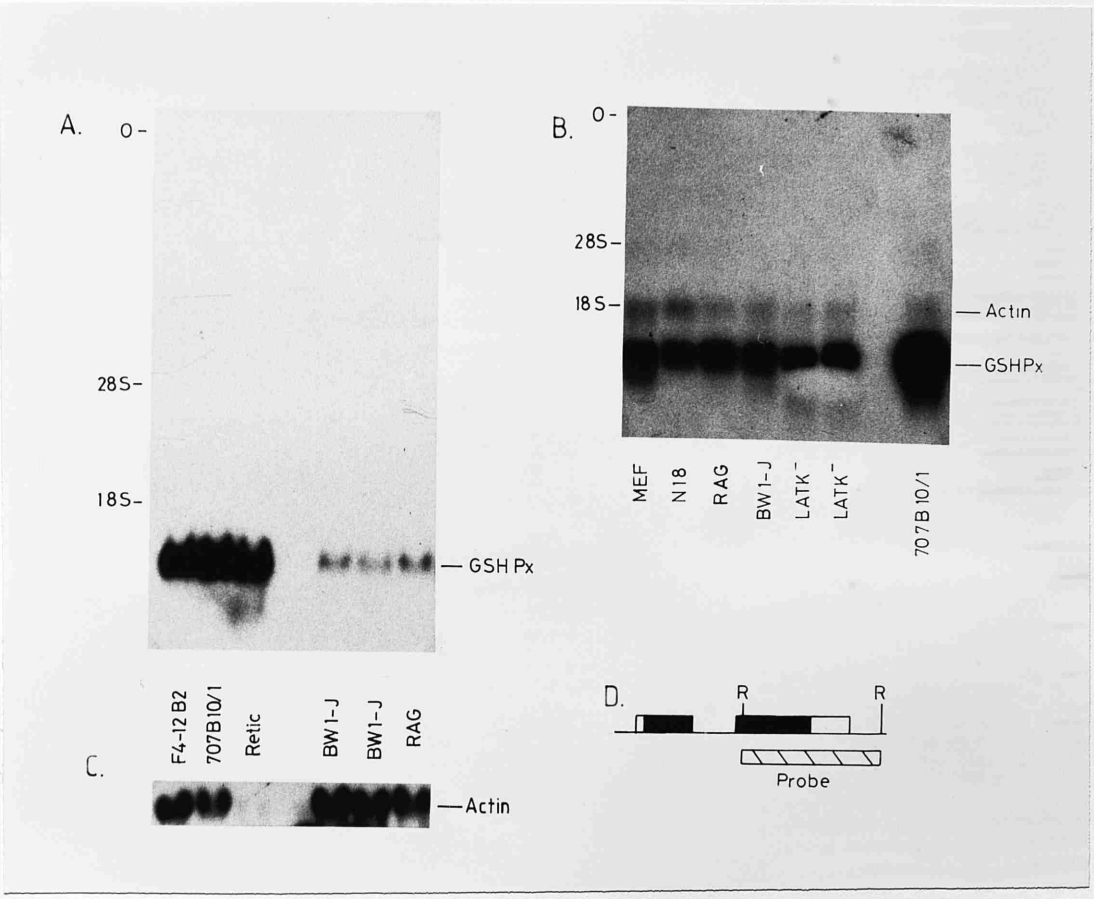


Figure 3.II.3. GSHPx mRNA levels in various murine cell lines.

20µg total cytoplasmic RNA prepared from various cell lines was denatured and electrophoresed through a 1.1% w/v agarose gel, containing 2.2M formaldehyde, cast and run in MOPS buffer until the bromophenol blue had run 15cm (panel A) or 9cm (panel B). Following transfer to nitrocellulose, RNA was hybridised to the 706 b.p. Eco RI exon 2 probe shown in D (panel A) or the the rat skeletal muscle actin cDNA (panel C) or to a mixture of both (panel B) under normal stringency conditions. After washing under normal stringency conditions the filters were exposed to film (panel A; Kodak X-S; panel B, Kodak XAR-5) at -70°C with intensifying screens for 20hrs (A and C) or 70hrs (B). Lines at the sides of the autoradiographs mark the positions of the origins and of 18 and 28S rRNAs and identify GSHPx and actin mRNAs. RNA in the two lanes marked BW1-J in panel A are from separate preparations, as is the RNA from the two lanes marked LATK⁻ in panel B.

kidney cell line (RAG) are comparable to other low expressing cell lines rather than being the high levels characteristic of liver and kidney *in vivo*. The GSHPx protein of normal adult rats has been localised immunohistochemically, predominantly within the proximal convoluted tubules of the kidney and exclusively within the hepatocytes of the liver (Mizuiri *et al*, 1986; Yoshimura *et al*, 1980). Interestingly, in the latter case the protein appeared to be more abundant in hepatocytes around the portal areas than in hepatocytes around the central vein. In addition, BW1-J cells have been shown to express significant quantities only of neo-natal products (Cassio and Weiss, 1979) and in this respect it may be relevant that the level of GSHPx mRNA in neo-natal hepatocytes remains undetermined. It is therefore possible that RAG and BW1-J cell lines are derived from low GSHPx expressing cells within the tissues of origin. Alternatively, the decreased level of GSHPx mRNA may simply reflect the loss of normal functions during adaption of these cells to tissue culture. Indeed, in a study of several different hepatoma cell lines, including BW1-J, each was found to have a decreased level of expression of all eight liver specific functions tested, compared to normal adult liver and this was shown to be caused by a decreased rate of transcription of the corresponding genes (Clayton *et al*, 1985). Whatever the reasons for the low levels of GSHPx mRNA in RAG and BW1-J cells they are clearly not appropriate cells to use to attempt to understand the mechanisms leading to an elevated level of GSHPx mRNA in liver and kidney. Figure 3.II.3 also shows that the adherent Friend cell line F4-12B2 has

approximately the same GSHPx mRNA level as the suspension Friend cell line used in previous experiments (707B10/1) and that L cells ($LATK^-$) have approximately the same GSHPx mRNA level as a number of low expressing cell lines. Thus it would appear that these two cell lines are phenotypically appropriate cells in which to study the basis of an elevated GSHPx mRNA level in erythroid cells compared to low expressing cells. This is a useful practical observation since short-term DNA expression experiments were planned to elucidate the basis of the elevated GSHPx mRNA level in erythroid cells. F4-12B2 cells have been shown to give rise to stable DNA transfectants at a level 100-fold greater than most suspension Friend cells (Spandidos and Paul, 1982); $LATK^-$ cells also readily take up DNA in these kind of experiments in contrast to most suspension cells such as N18.

3.2.4. Mechanisms responsible for variations in GSHPx mRNA levels.

The preceding experiments show that the steady state level of GSHPx RNA varies between different cell types. These differences could be due to differences in the rate of gene transcription or RNA stability between the different cell types (reviewed by Nevins, 1982; Darnell, 1982; Brawerman, 1981). Alternatively, a block in either the processing of the primary transcript or the nucleo-cytoplasmic transport of the mRNA could result in a decreased cytoplasmic level of mRNA in one cell type relative to another (Padgett *et al*, 1986; Fulton *et al*, 1985).

Prior analyses had shown the steady state level of GSHPx mRNA

to be 20-50 fold higher in the cytoplasm of reticulocytes, liver and kidney than in the cytoplasm of other tissues such as brain (Affara et al, 1983; 1985). As the steady state levels of total cellular GSHPx mRNA-sized RNAs vary similarly between these different tissues (Figure 3.II.1) a block in the nucleo-cytoplasmic transport of the mRNA is not the cause of the reduced levels of GSHPx mRNA in the cytoplasm of low expressing tissue. The GSHPx pre-mRNA (which should run 3mm behind the GSHPx mRNA in Figure 3.II.1) is not detectable on the autoradiograph shown in Figure 3.II.1, although it may be present in the tissues expressing higher levels of GSHPx mRNA but be unresolved from the main mRNA band. To assess the contribution due to a block in precursor splicing, the relative levels of GSHPx pre-mRNA and mRNA were compared in a low and in a high expressing cell line (N18-TG2 and 707B10/1). An intron probe corresponding to nucleotides +305 to +480 of Figure 3.I.5 was first prepared. To eliminate any contaminating exon sequences, this probe was subcloned into pUC12, the insert repurified and used to analyse a northern blot containing 5µg poly A^+ RNA from uninduced 707B10/1 cell nuclei. However, no hybridisation was detected (data not shown) probably due to the low level of GSHPx pre-mRNA (and also possibly fragmentation of the probe during nick-translation). An alternative strategy was therefore adopted. This involved preparation of DNA fragments extending from the Xba I site at -286bp (see Figure 3.I.5) to either the Pst I site within exon 1 (located at +215bp) or the Pvu II site within the intron (located at +363bp). The 5' termini of these DNA fragments were then

labelled with ^{32}P, the two strands of each of these DNA fragments separated and hybridised to cellular RNA from neuroblastoma cells (N18-TG2) or Friend cells (707B10/1). Following digestion of unhybridised single stranded regions with S1 nuclease, a measure of the level of spliced and unspliced RNAs can be obtained by denaturation of hybrid molecules and electrophoresis through denaturing gels. Figure 3.II.4 shows that this approach detected the pre-RNA in both 707B10/1 and N18-TG2 cells. Furthermore, densitometric comparison of the levels of pre-mRNA in the two cell lines shows that these vary in the same way as the major mRNA transcripts, thereby suggesting that there is no difference in the control of processing of the pre-mRNA in the two cell lines. However, calculations showed the S1 protection experiment with 707B10/1 RNA to have been conducted at only a 2-fold probe excess, a situation which does not yield precisely quantitative results. Nevertheless, it is clear that there is no massive control at the level of pre-mRNA splicing and it would therefore appear that the different levels of GSHPx mRNA in these two cell lines is due either to differences in mRNA stability or to differences in the rate of GSHPx gene transcription in the two cell lines.

3.2.5. Variation in the use of different transcription initiation sites in different cell lines.

Mapping of the GSHPx transcription initiation sites in several cell lines by primer extension and S1 nuclease protection had detected several other discrete sites between -220nt and +1nt in addition to the major transcription initiation site (located at

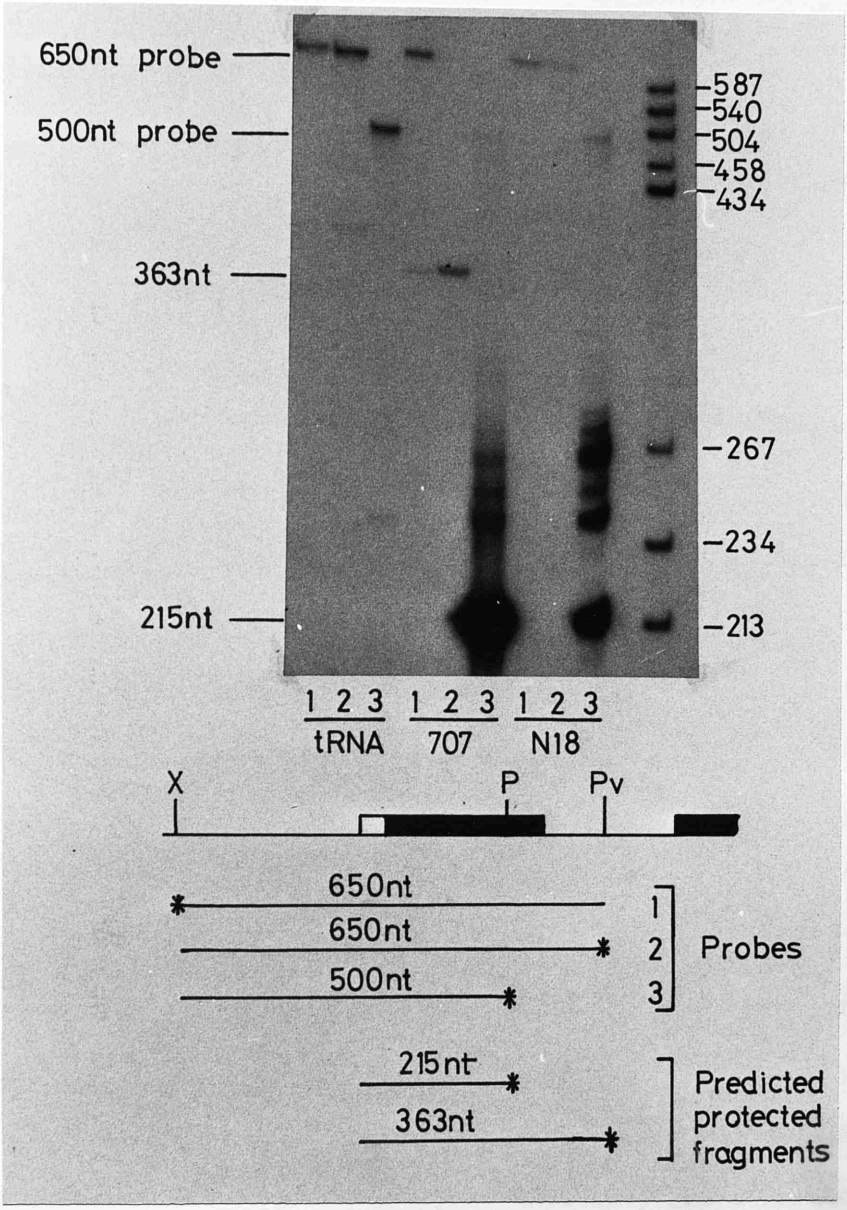


Figure 3.II.4. S1 nuclease protection assay of GSHPx mRNA and pre-mRNA in high and low expressing cell lines.

200µg each of the 650 b.p. Xba I - Pvu II fragment and the 500 b.p. Xba I - Pst I fragment were 5' end-labelled, denatured and the strands separated on a polyacrylamide gel as described in Chapter 2. Single stranded probes (1-3) were then recovered from the gel and 25ng of each hybridised to 5µg poly A^+ RNA from 707B10/1 or N18 plus 45µg yeast tRNA. After hybridisation and S1 digestion hybrid molecules were denatured and electrophoresed through a denaturing 6% polyacrylamide/7M urea/TBE gel until the bromophenol blue had run 35cm. After drying the gel was exposed to Kodak XAR-5 film at -70°C for 60 h. Top: Autoradiograph showing the positions and sizes (in nucleotides) of the Hae III cut pBR322 Mr markers and indicating the sizes of the probes and the protected fragments. Bottom: Line diagram showing the origin of the single stranded probes (asterisks mark the ends at which each strand is labelled) and relating these to the lanes on the autoradiograph containing the products of the reactions between each of these probes and tRNA (tRNA), 707B10/1 RNA (707^-) and N18 RNA (N18). The sizes of the protected fragments predicted after hybridisation to mRNA or pre mRNA are also indicated.

+1nt, Figure 3.I.5; Frampton et al, 1987). In uninduced Friend cells 97% of the GSHPx transcripts originate at +1nt with each of the upstream initiation sites accounting for less than 1% of the total GSHPx transcripts. In induced Friend cells as well as reticulocytes, liver and kidney this bias in favour of +1nt transcription initiation is even more pronounced. A comparison of the transcript profile of a high expressing cell line (707B10/1) and a low expressing cell line (N18-TG2) is presented in Figure 3.II.4. This shows not only that the absolute level of GSHPx transcripts is much lower in N18-TG2 than in 707B10/1 but also that the proportion of upstream transcripts is much higher in N18-TG2 than in 707B10/1. This is a similar observation to that noted for a fibroblast cell line (Frampton et al, 1987) except that in the case of N18-TG2 the proportion of upstream transcripts is even higher.

What might cause these differences in transcript profile in different cells? One possibility is that they are a reflection of different transcriptional states of the GSHPx gene. Alternatively, they may be due at least partly to differences in transcript stability in the two cell types. (However, it is difficult to see how stability differences alone might account for the complex differences seen in transcript profile in the two cell types). The simplest way of distinguishing between these alternatives is to measure the polymerase densities on the GSHPx gene by in vitro RNA elongation ("run-on" transcription) in isolated nuclei prepared from the different cell types (Derman et al, 1981).

3.3. Search for cis-control regions regulating transcription of the GSHPx gene.

The position of the 5' ends of the GSHPx transcripts in several different cells have been determined (Chambers et al, 1986; Frampton et al, 1987; section 3.2.5.). In all cases the majority of transcripts originate from the major transcription initiation site (+1nt, Figure 3.I.5). It thus seems probable that the sequences located upstream from the major transcription initiation site may function as a RNA polymerase II promoter in mouse cells. The apparently ubiquitous presence of a DNAse I hypersensitive (DH) site in this region (Affara et al, 1985; see also below) lends weight to this idea since it is often found that the promoter regions of RNA polymerase II genes are DH sites in the cells in which these genes are expressed (McGhee et al, 1981; Wu, 1980; Wu and Gilbert, 1981; Saragosti et al, 1982; Sweet et al, 1982). To test the validity of this assumption, a DNA fragment extending 650nt upstream from the major transcription initiation site was cloned into a plasmid immediately upstream from the bacterial chloramphenicol acetyltransferase (CAT) gene (Gorman et al, 1982) and this was then used to transfect mouse cells. Since eukaryotic cells do not possess any CAT enzyme, any such activity detectable in the transfected cell population must originate from the introduced DNA. Moreover, if transfections are conducted using two plasmids which differ only in the presence of a given DNA fragment, then any differences in the level of CAT enzyme activity detected in the transfected cell populations must be due to that DNA fragment.

This basic approach was extended to address the question of whether the potential transcriptional regulatory sequences at the 3' end of the gene (see section 3.1.4) could function to alter the transcription of a CAT gene linked in *cis* to them.

3.3.1. CAT plasmids used as vectors and controls.

The starting plasmids for these experiments was pLW2 (a gift from J.B. Clements, see Figure 3.III.1), the derivation of which has been described (Gaffney *et al*, 1985). Briefly, this plasmid contains the CAT gene under transcriptional control of the herpes simplex virus type 2 immediate early (HSV-2 IE) gene -4/-5 promoter and having the terminator fragment from the 3' end of the HSV-2 IE gene -5 at its 3' end. This terminator fragment contains an AATAAA sequence, which forms an essential part of the recognition sequences for 3' end processing ; downstream from this is a sequence homologous to the consensus sequence PyGTGTTPyPy which is found 3' to the majority of polyadenylation sites and is important in directing efficient 3' end formation (McLauchlan *et al*, 1985; a recent comprehensive review of 3' end formation is given in Birnsteil *et al*, 1985).

To construct a CAT plasmid lacking the viral promoter, whilst retaining the sequences necessary for mRNA 3' end formation, the CAT coding sequences and the terminator sequences of pLW2 were excised using *Bam* HI and *Xba* I and cloned into pUC12 to give plasmid p22 (a gift from Jas Lang, see Figure 3.III.2A).

This plasmid was used to generate another plasmid containing a "minimal" SV40 promoter which would give a low level of CAT

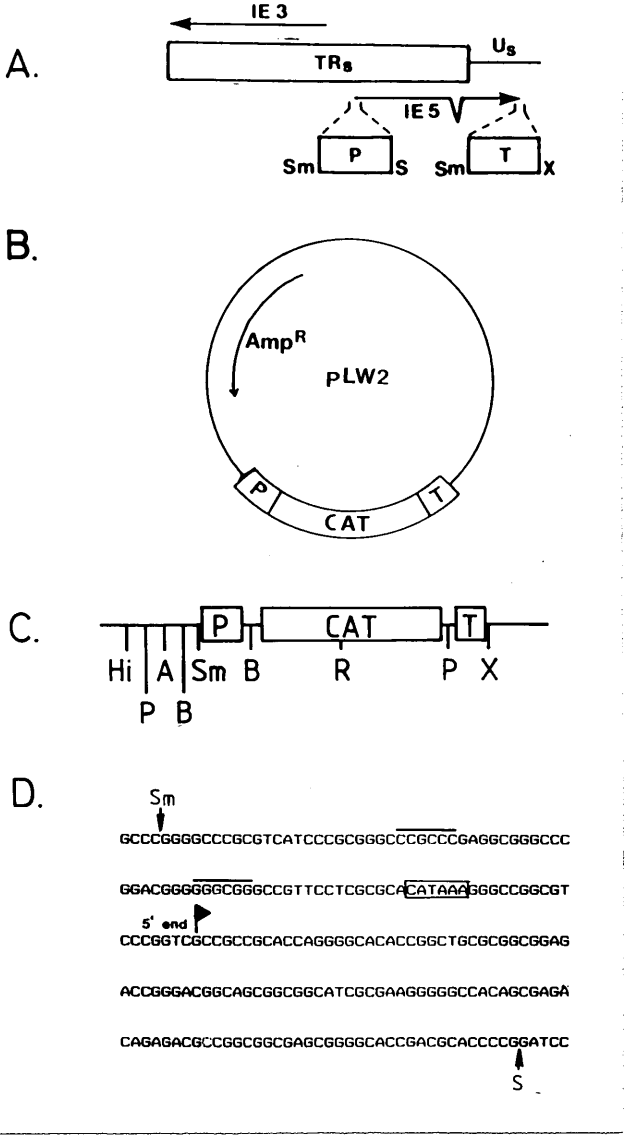


Figure 3.III.1. Derivation of pLW2; a HSV immediate early gene promoter -CAT fusion plasmid.

A. Partial map of the HSV-2 genome around the immediate early-3 (IE3) and the immediate early-5 (IE5) genes. Arrows indicate the extent of the IE genes and their direction of transcription. TR_s, short terminal repeat; U_s, short unique region; P, IE-5 promoter; T, IE-5 3' end processing signals. The restriction sites used for excision of the promoter and terminator fragments are shown. B. Map of pLW2. The 210 b.p. promoter fragment and the 100 b.p. terminator fragments from the HSV-2 IE-5 gene have been inserted around the CAT gene in the same orientations as in the IE-5 gene. The remaining plasmid sequences are pUC derived (Gaffney et al, 1985). C. Expanded map of the CAT transcription unit of pLW2 showing the relative positions of some useful restriction sites. D. Sequence of the 210 b.p. IE-5 promoter. The restriction sites used in excision of the fragment are indicated by arrows. Sp1 binding site hexanucleotide sequences are overlined and the 5' end of the transcription unit and the direction of transcription are indicated by the flag.

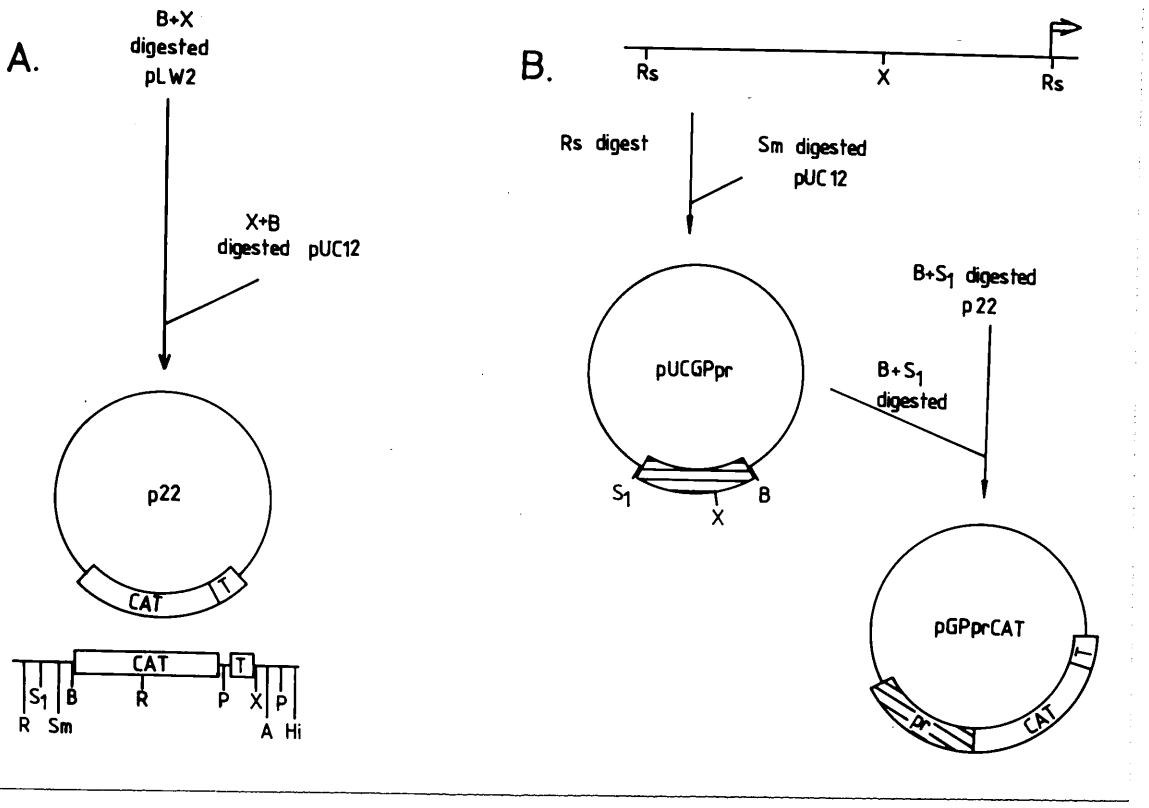


Figure 3.III.2. Construction of pGPprCAT; a CAT plasmid under control of the putative GSHPx promoter.

A. Construction of the promoter negative CAT plasmid p22 from pLW2. A map of the restriction site around the CAT transcription unit is shown. B. Construction of pGPprCAT from the GSHPx gene and p22. The line diagram at the top represents the GSHPx 5' flanking sequence upstream from the major transcription initiation site (indicated by the arrow showing the direction of transcription). The 650 b.p. Rsa I fragment from the 5' end of the GSHPx gene was first cloned into pUC12 and its orientation with respect to the polylinker deduced by cleavage at the asymmetrically disposed Xba I site. The plasmid having the desired orientation (pUCGPpr) was then digested with Bam HI and Sst I and the GSHPx 5' flanking fragment recloned into p22 in the same orientation with respect to the CAT gene as it is with respect to the GSHPx gene in the mouse genome.

activity and which would show little cell-specificity, pSVCATO. This was done by cloning the Sph I - Hind III fragment from the SV40 early promoter into the Sma I site of p22, thus placing the CAT gene under transcriptional control of the "enhancer-less" SV40 early promoter (Jon Frampton, personal communication; Figure 3.III.3). This plasmid is essentially equivalent to pA10CAT2 constructed by Laimins et al (1982), which gives a low level of CAT activity in all cells examined.

3.3.2. Construction of pGPprCAT : a plasmid containing the CAT gene under transcriptional control of the putative GSHPx gene promoter.

To produce a CAT gene plasmid under the transcriptional control of the putative GSHPx gene promoter, the Rsa I fragment extending from approximately -650nt to +3nt of the GSHPx gene was prepared. Since this fragment has blunt ends it was subcloned into Sma I cut pUC12 before being recloned into p22 (Figure 3.III.2B). Subclones putatively containing the GSHPx promoter were digested with Bam HI and Sst I to identify plasmids carrying a single copy of the Rsa I fragment (Figure 3.III.4). These plasmids were also digested with Xba I separately in order to orient the insert with respect to the polylinker (Figure 3.III.4). This is possible (see Figure 3.III.2B) since Xba I cuts the insert once asymmetrically and cuts the vector once in the polylinker; therefore plasmids having the polylinker Xba I site at the end of the insert normally closest to the GSHPx gene will release a 300bp fragment upon Xba I digestion (clones 2, 5 and 6, Figure 3.III.4). These clones

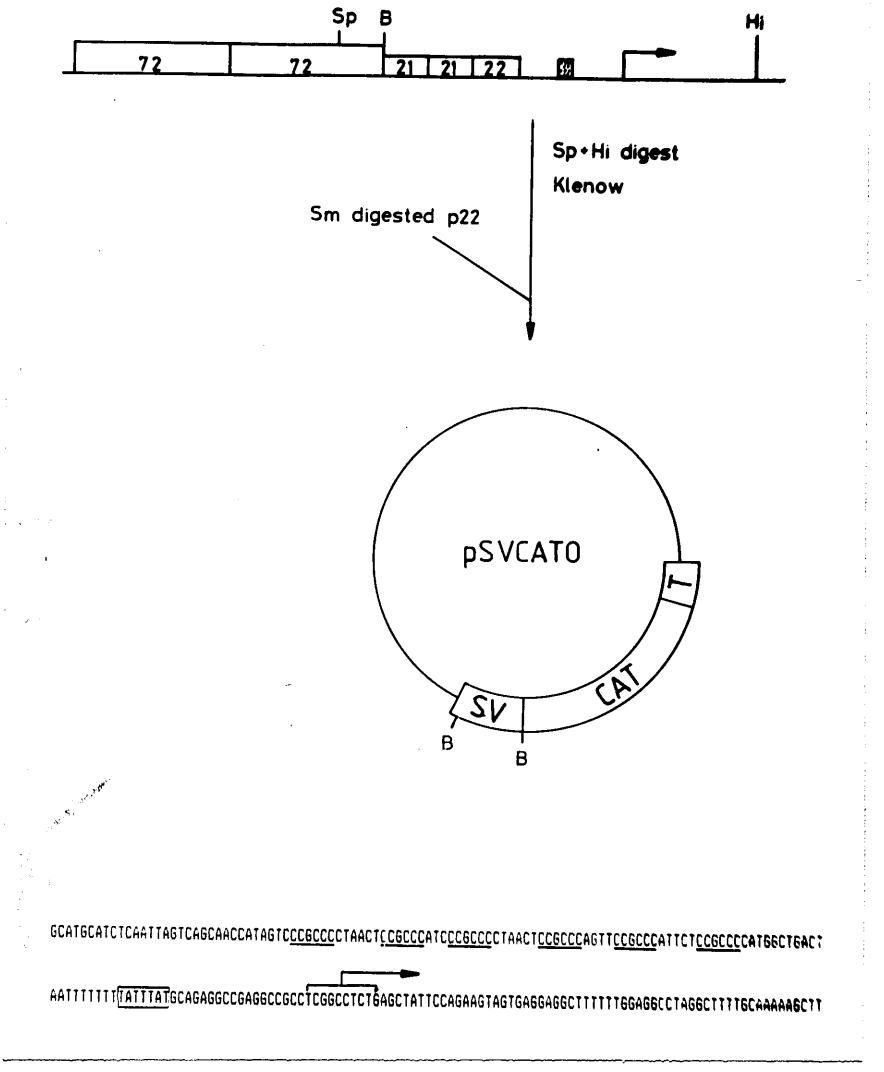


Figure 3.III.3. Construction of pSVCATO; a CAT plasmid under control of the SV40 early promoter.

The line diagram at the top shows the region of the SV40 genome surrounding the early promoter. The arrow indicates the position of the early-early start sites and the direction of transcription. The boxes upstream of this represent the "ATA" box (stipled), the two perfect and one imperfect 21 b.p. repeats containing the Sp1 binding sites, and the two 72 b.p. repeats of the SV40 enhancer. The Sph I-Hind III 200 b.p. fragment was isolated from the SV40 genome, the ends repaired using the Klenow fragment of E.coli DNA polymerase I, and cloned into Sma I digested p22. The orientation of the inserts was ascertained by Bam HI digestion and clones isolated (pSVCATO) with the 200 b.p. fragment in the same orientation with respect to the CAT gene as it is in the SV40 genome with respect to the early transcription unit. The sequence of the 200 b.p. fragment showing the positions of early-early transcription initiation (arrow), "ATA" box (boxed) and Sp1 hexanucleotide core sequences within the 21 b.p. repeats (underlined).

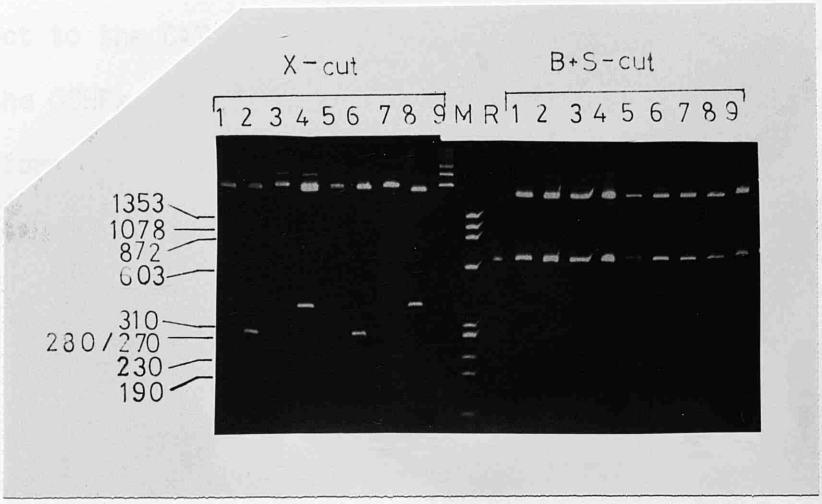


Figure 3.III.4. Identification of pUC subclones containing the 650 b.p. Rsa I GSHPx gene promoter fragment.

Following transfection of E.coli JM83 with the products of a ligation reaction of Sma I cut pUC12 and the 650 b.p. Rsa I fragment from the GSHPx gene, plasmid DNA was prepared by the rapid method from individual white colonies growing on agar containing 5-bromo-4-chloro-3-indolyl-β-galactoside (X-gal) and ampicillin (Amp). DNA was the digested with the indicated enzymes and electrophoresed through a 6% polyacrylamide gel. The sizes of the Mr markers (ϕx174 DNA Hinf I digest) are indicated at the left hand side.

(designated pUCGPpr) were the ones of interest and were taken for further manipulation to produce pGPprCAT. These clones having the opposite orientation were designated pUCGPpr(-) and were stored.

The Bam HI site in pUCGPpr lies between the polylinker Xba I site and the Sma I site used for cloning the Rsa I fragment and the Sst I site lies on the opposite side of the insert (Figure 3.III.2B). Therefore, the cloned putative GSHPx gene promoter (GPpr) fragment can be excised by digestion with Bam HI plus Sst I and recloned into Bam HI - Sst I digested p22. The resultant plasmid will have the 650bp fragment in the same orientation with respect to the CAT gene as it is in the mouse genome with respect to the GSHPx gene (Figure 3.III.2B). These manipulations were therefore performed and plasmid DNA isolated from the putative subclones. Since p22 contains two Eco RI sites, one within the polylinker upstream from the Sst I site and another approximately 260bp downstream within the CAT gene (Figure 3.III.2A), digestion of subclones carrying a single copy of the insert with Eco RI should release a fragment of approximately 910bp. This indeed occurred for most of the plasmids examined (Figure 3.III.5) although one clone (No.1) released a much larger fragment: this was probably caused by cloning of two copies of the GPpr fragment into the one vector molecule.

To check that the other plasmids were as expected, two (Nos. 3 and 4) were analysed by digestion with Bam HI and Sst I to release the GPpr fragment and by digestion with Bam HI and Xba I to check the orientation of the insert. This latter digest should also release the CAT gene plus terminator fragment. The results

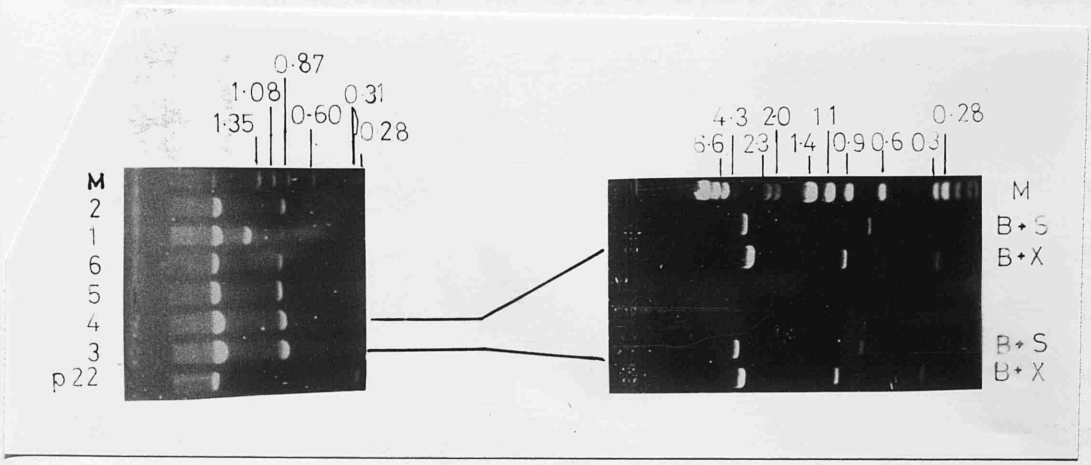


Figure 3.III.5. Identification of pGPprCAT.

Following transfection of E.coli HB101 with the products of a ligation reaction of BamHI, Sst I cut p22 with the 650 b.p. Bam HI, Sst I fragment from pUCGPpr, plasmid DNA was prepared by the rapid method from individual colonies growing on agar containing Amp. DNA was then digested with either Eco RI (left hand panel) or other indicated enzymes and the products display on 1% w/v agarose gels. The sizes of the Mr markers (ϕx174 DNA, Hinf I digest; λDNA, Hind III digest) are indicated.

of this analysis (Figure 3.III.5) show that clones 3 and 4 are in fact of the desired design and so these plasmids (designated pGPprCAT clones 3 and 4) were carried forward for transient expression analysis and also for further genetic manipulation.

3.3.3. Construction of CAT plasmids designed to analyse the transcriptional effect of the 706 b.p. Eco RI fragment from the 3' end of the GSHPx gene.

As discussed in section 3.1.4, there are several motifs at the 3' end of the GSHPx gene, close homologues of which have been shown to be important in regulating transcription in other contexts. In addition, the data of Affara *et al* (1985) showed that in cells having a high GSHPx mRNA level a region of the chromatin downstream from the transcription initiaion site was hypersensitive to DNAse I. Moreover, at the resolution obtained in their analysis it was possible that this DH site lay within the region of the chromatin corresponding to the 706bp *Eco* RI fragment. This seemed intriguing since both viral and cellular enhancer elements correspond to DH sites in the cells in which they act (Parslow and Granner, 1982; Jongstra *et al*, 1984; Zaret and Yamamoto, 1984). Consequently, to investigate whether sequences at the 3' end of the GSHPx gene had similar transcriptional enhancing properties, the 706bp *Eco* RI fragment containing most of exon 2 and all the available cloned 3' flanking sequences (Figure 3.I.3) was prepared, the 5' overhangs repaired with the Klenow fragment of *E.coli* DNA polymerase I and *Hind* III linkers added before cloning this fragment into the *Hind* III sites

upstream of the CAT gene in pLW2 or downstream of the CAT gene in pGPprCAT (Figure 3.III.6).

i. Characterisation of pGP0.7LW2 constructs.

Pst I cuts 3' to the Hind III site and Pvu II cuts 5' to the Hind III site in pLW2; therefore putative pGP0.7LW2 subclones were examined by digestion with Pst I plus Pvu II to ensure that they contained only one insert (not shown). The orientation of the insert was then determined by digesting at the asymmetrically disposed Msp I site within the insert and at the Pst I site situated on the CAT gene proximal side of the Hind III cloning site (see Figure 3.III.7 for the positions of the relevant restricition sites). When the insert is in the opposite orientation, with respect to the direction of transcription of the CAT gene, as it is in the mouse genome with respect to the direction of transcription of the GSHPx gene, the Msp I - Pst I fragment released should be approximately 190bp. In two subclones (Nos. 11 and 16, Figure 3.III.7) this is indeed the case and these subclones were designated pGP0.7-LW2 clones 11 and 16.

In the opposite orientation a Msp I - Pst I fragment of 540bp would be expected; however in clone 14 the fragment observed runs at 610bp (Figure 3.III.7). This might mean that the fragment which has been cloned is not the 706bp fragment or that some additional DNA has been added onto it. However, since the 527bp Msp I - Eco RI sub-fragment of the original 706bp Eco RI fragment runs under non-denaturing conditions at an apparent molecular weight greater than expected (data not shown) but under denaturing conditions runs at the expected position for a band of 527nt

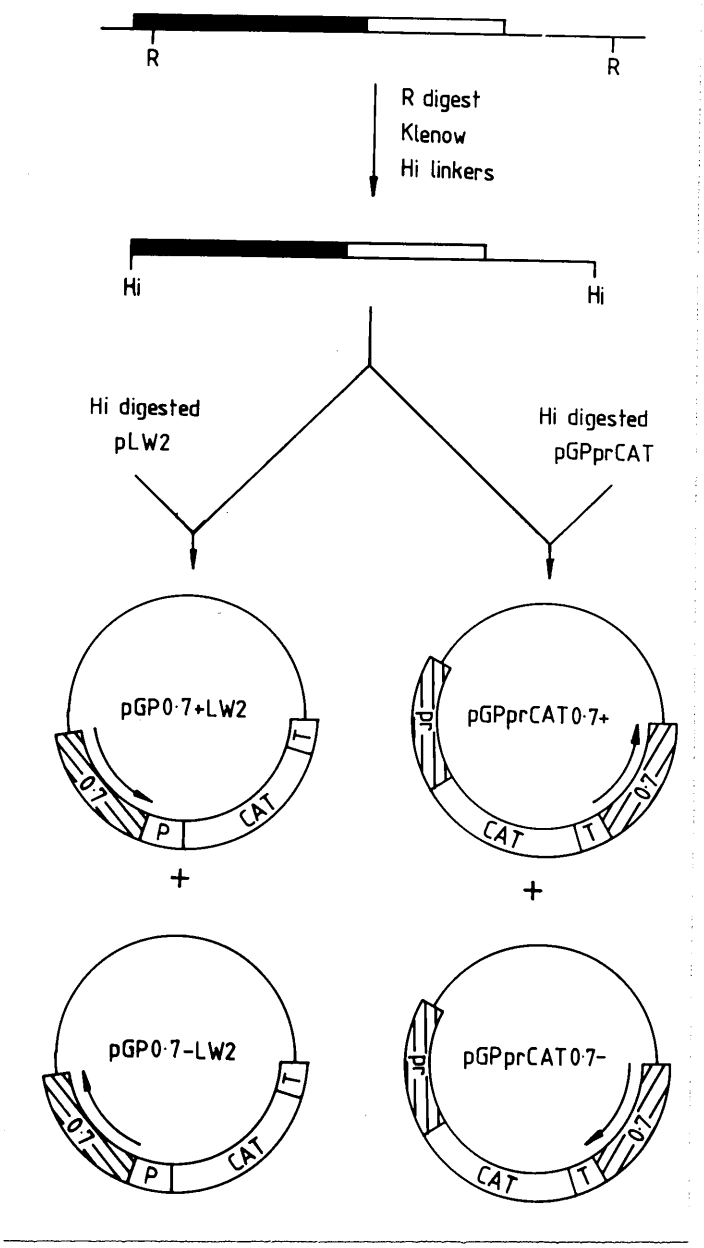
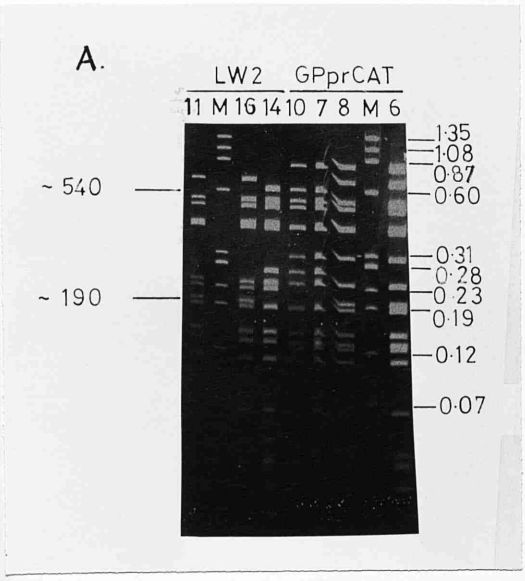
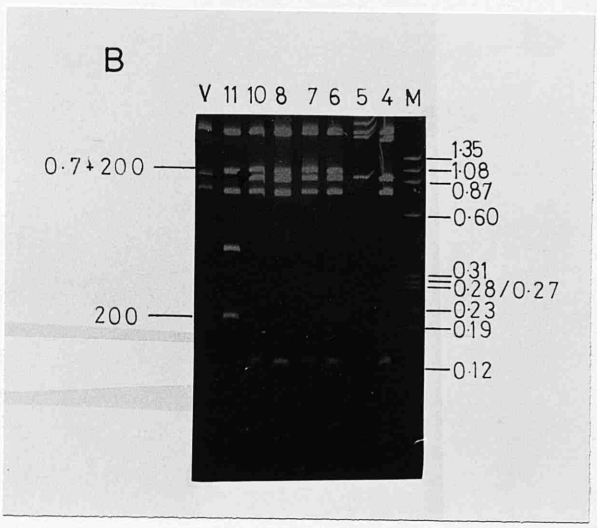
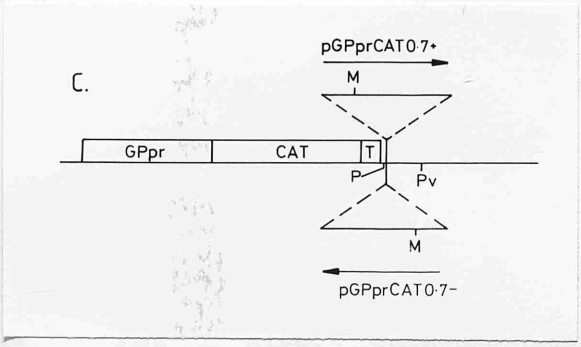


Figure 3.III.6. Construction of CAT plasmids containing the 706 b.p. Eco RI fragment from the 3' end of the GSHPx gene.

The scale diagram at the top represents the 3' part of the GSHPx gene showing the second exon of the gene (box) with protein coding region (filled segment) and untranslated region (open segment) indicated. The 706 b.p. fragment was digested with Eco RI, gel purified and its ends repaired by the Klenow fragment of E.coli DNA polymerase I. Hind III linkers were then added and the fragment cloned in both orientations into the Hind III sites of either pLW2 (Figure 3.III.1) or pGPprCAT (Figure 3.III.2B).







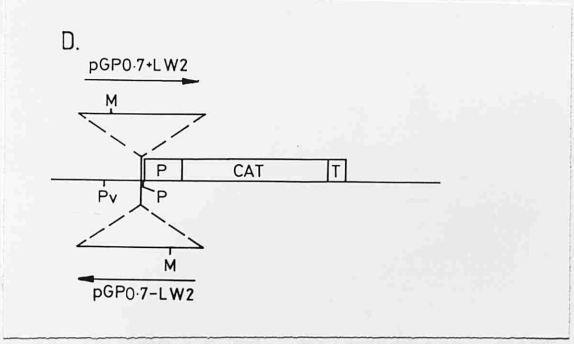


Figure 3.III.7. Identification of pGP0.7LW2 and pGPprCAT0.7 constructs.

Following transfection of E.coli HB101 with the products of a ligation reaction of Hind III cut pLW2 with the Hind III linkered 706 b.p. fragment from the 3' end of the GSHPx gene (panel A, LW2) or the products of a ligation reaction of Hind III cut pGPprCAT with the Hind III linkered 706 b.p. fragment (panel A, GPprCAT; panel B), plasmid DNA was prepared by the rapid method from individual colonies growing on agar containing Amp. A. DNA was digested with Pst I and Msp I and electrophoresed through 6% polyacrylamide. The line below LW2 indicates the clones from an experiment using pLW2 as the vector, that below GPprCAT those from an experiment using pGPprCAT as the vector. M; Hinf I cut ϕx174 DNA Mr marker fragments, sizes indicated at the side of the figure. The positions of the Pst I - Msp I fragments of interest are indicated. B. DNA was digested with Pst I and Pvu II and electrophoresed through 6% polyacrylamide. M; Hinf I cut ϕx174 DNA Mr marker fragments, sizes indicated at the side of the figure. V; Pst I and Pvu II digested pGPprCAT. The Pst I - Pvu II fragments of interest are indicated. Below are shown the restriction sites around the CAT transcription unit relevant to the diagnosis of the orientation of pGPprCAT0.7 (C) and pGP0.7LW2 (D).

(Chambers *et al* 1986), it is clear that the fragment is correct but that it migrates anomalously in neutral gels. Consequently, clone 14 was designated pGP0.7+LW2.

ii. Characterisation of pGPprCAT0.7 constructs.

Pst I cuts 5' to the *Hind* III site and *Pvu* II cuts 3' to the *Hind* III site in pGPprCAT, therefore putative pGPprCAT0.7 subclones were examined by digestion with *Pst* I plus *Pvu* II to ensure that they contained only one insert (Figure 3.III.7). In fact, compared to a similar digest of pGPprCAT (lane V) clones 6,7,8 and 10 have lost a 200bp fragment and gained a fragment of approximately 900bp. These subclones were therefore examined further in order to determine the orientation of their inserts by the same method used for pGP0.7LW2 constructs: that is, digestion at the internal *Msp* I site in combination with digestion at the 5' *Pst* I site. When the insert is in the same orientation with respect to the direction of transcription of the CAT gene, as it is in the mouse genome with respect to the direction of transcription of the GSHPx gene, the *Msp* I - *Pst* I fragment released will be approximately 190bp. This is indeed the case for subclones 6 and 8 (Figure 3.III.7), which were therefore designated GPprCAT0.7+. When the insert is in the opposite orientation, the anomalously migrating 540bp fragment should appear, which indeed it does when clones 7 and 10 are digested with *Msp* I and *Pst* I (Figure 3.III.7). These clones were therefore designated pGPprCAT0.7-.

3.3.4. The 650 bp Rsa I fragment from the 5' end of the GSHPx gene functions as a promoter in mouse cells with no apparent cell preference.

In order to examine the ability of the putative GSHPx gene promoter to function as such, 10µg of each of the plasmids p22 (Figure 3.III.2A), pGPprCAT (Figure 3.III.2B) and pSVCAT0 (Figure 3.III.3) were transfected, in duplicate, into mouse cells as detailed in Chapter 2. As the promoters of a number of genes have been shown to be responsible, at least in part, for the cell specificity of expression of the associated gene (Ott *et al*, 1984; Edlund *et al*, 1985; Gopal *et al*, 1985; Grosschedl and Baltimore, 1985; Mason *et al*, 1985; Picard and Schaffner, 1985) the above plasmids were transfected into two cell lines, one having a high level of GSHPx mRNA (F412-B2) and the other a low level of GSHPx mRNA ($LATK^-$), in an attempt to determine whether the putative promoter has a role in directing these different levels of expression. Cells were harvested 48 hours after transfection and soluble protein extracts prepared. An equivalent amount of protein from each sample was then assayed for CAT activity and the mixture of acetylated and unacetylated chloramphenicol derivatives separated by thin layer chromatography (Figure 3.III.8). This shows that in $LATK^-$ cells the presence of the putative GSHPx promoter caused a significant increase in the transcription of the CAT gene above the background level caused by prokaryotic sequences. Moreover, the GSHPx promoter was approximately equivalent in activity to the SV40 early promoter (quantitative estimates of promoter activities are presented in Table 3.III.1).

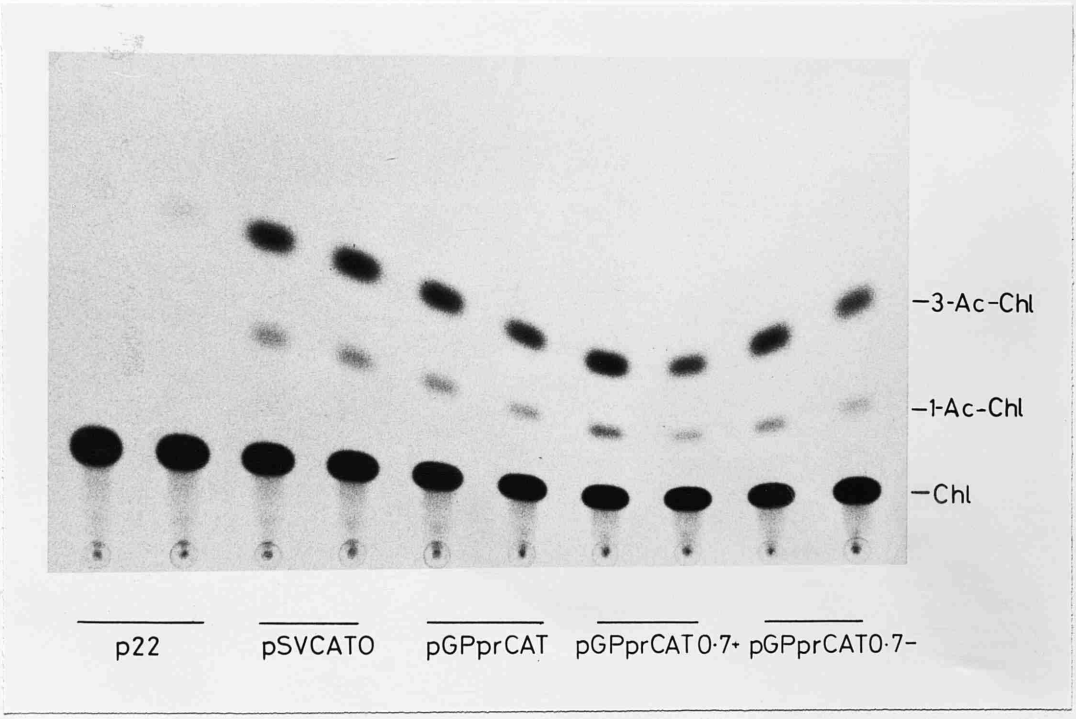


Figure 3.III.8. CAT activities directed by GSHPx promoter based CAT constructs during transient expression in LATK⁻ cells.

Extracts from duplicate transfections were prepared and assayed for CAT activity as described in Chapter 2. The mixtures of 1-acetyl chloramphenicol (1-Ac-Chl), 3-acetyl chloramphenicol (3-Ac-Chl) and unacetylated chloramphenicol (Chl) were separated by t.l.c. and the plates exposed to Kodak XAR-5 film at room temperature for 20h.

Table 3.III.1. GSHPx promoter function during transient expression in two cell lines.

Plasmid Transfected	Cell Line		
	LATK$^-$	F412B2	
	Experiment No.1	Experiment No 1	Experiment No 2
p22	0.3	0.3	0.1
pSVCATO	5.5	0.4	0.2
pGPprCAT	4.2	0.4	0.2
pLW2	16	11	10
pGPprCAT / pSVCATO	0.8	1.0	1.0

Transient expression of plasmids and CAT assays were performed as detailed in Chapter 2. Values presented are percentage conversion of chloramphenicol to acetylated chloramphenicol derivatives after assaying extracts at 37^{o}C for 1h and are the average of determinations from duplicate flasks. Within each experiment equivalent amounts of protein have been assayed from each sample. The ratio pGPprCAT/pSVCATO is a measure of the cell specificity of pGPprCAT since pSVCATO shows no cell specificity (Laimins et al, 1982) (the background level due to p22 was not subtracted from the values of pGPprCAT or pSVCATO obtained in Friend cells before calculating this ratio).

The results for Friend cells (also summarised in Table 3.III.1) are less clear cut since the levels of CAT activity are very close to background (the Friend cells did however take up the DNA as evidenced by the CAT activity directed by pLW2). Nevertheless since the ratios of activities of pGPprCAT and pSVCATO are the same in both cell types it seems that the GSHPx promoter does not show any preference in its function between LATK$^-$ and F4-12B2 cells.

3.3.5. Examination of the transcriptional influence of sequences at the 3' end of the GSHPx gene.

As discussed previously there exist at the 3' end of the GSHPx gene several sequences which could potentially affect the level of GSHPx gene transcription (Section 3.1.4). Moreover, the data of Affara *et al* (1985) suggest that these sequences might be coincident with those over which a DH site is formed in cells expressing high levels of GSHPx mRNA. Therefore, in order to test if these sequences did in fact possess any ability to influence transcription from a promoter linked in *cis*, the plasmids pGPO.7+LW2, pGPO.7-LW2, pGPprCATO.7+ and pGPprCATO.7- (Figure 3.III.6) were transfected into cell lines expressing GSHPx mRNA at a low or a high level (LATK$^-$ and F4-12B2). Cells were harvested 48 hours later and soluble protein extracts prepared. Aliquots from each sample containing equivalent amount of protein were then assayed for CAT activity (Table 3.III.2). The results of this analysis show first of all that the 706bp fragment does not affect the transcription from either of the two

Table 3.III.2. Effect of the 706 bp Eco RI fragment upon transcription of a CAT gene linked in cis during transient expression in two cell lines.

Cell line	Plasmid transfected	Experiment No. 1	2	3
F412-B2	p22	0.5	0.3	0.1
	pGPprCAT	0.7	0.4	0.2
	pGPprCAT0.7+	0.9(1.3)	0.6(1.5)	0.2(1.0)
	pGPprCAT0.7-	0.8(1.1)	0.4(1.0)	0.3(1.5)
	pLW2		11	10
	pGP0.7+LW2		21	10 (1.0)
	pGP0.7-LW2		19 (1.7)	23 (2.3)
LATK$^-$	p22	0.3		0.3
	pGPprCAT	2.4		4.2
	pGPprCAT0.7+	3.5(1.5)		4.5(1.1)
	pGPprCAT0.7-	5.7(2.6)		3.8(0.9)
	pLW2		23*	16
	pGP0.7+LW2		34* (1.5)	13 (0.8)
	pGP0.7-LW2		27* (1.2)	23 (1.4)

Transient expression of plasmids and CAT assays were performed as detailed in Chapter 2. Values presented are percentage conversion of chloramphenicol to acetylated chloramphenicol derivatives after assaying extracts at 37^{o}C for 1h and are the average of determinations from duplicate flasks except * where only one flask of cells was transfected. For each cell line within a given experiment the same amount of protein has been assayed from each sample. Figures in brackets represent the ratios between the indicated plasmid and the relevant plasmid lacking the 706 bp fragment (for LATK$^-$ these ratios were calculated after first subtracting the background activity from the promoterless p22).

promoters by more than a factor of two when linked in *cis* to them in either orientation in LATK$^-$ cells. Likewise, in Friend cells the 706bp fragment does not affect transcription from the HSV-2 IE -4/-5 gene promoter when linked in *cis*, in either orientation. If there are any effects in these cases they are marginal and will require a greater number of experiments to be performed before they can be quantitated with any degree of confidence. The analysis of the effect of the 706bp fragment upon transcription when linked in *cis* to the GSHPx promoter, in Friend cells is more difficult to assess since all the CAT activities are close to background. Nevertheless, the 706bp fragment does not appear to have any greater enhancing effect upon transcription from the GSHPx promoter.

These experiments show a distinct lack of any large modulating influence of the 706bp *Eco* RI fragment upon the homologous or a heterologous promoter linked in *cis*. It is worth noting however, that this does not preclude that such a function may be mediated by these sequences in another unidentified cell type or under particular environmental conditions.

3.4. Chromatin structure of the GSHPx gene in different tissues.

As was argued in Chapter I, sites of transcriptional regulatory significance are often located in regions of chromatin which are hypersensitive to DNAse I digestion compared to bulk chromatin. Therefore, identification of DNase I hypersensitive (DH) sites close to the GSHPx gene might be expected to pinpoint

regions involved in the control of transcription of the gene.

The chromatin structure of the GSHPx gene in erythroid (induced Friend) cells and in two non-erythroid (neuroblastoma and T-lymphoma) cell lines had previously been examined (Affara et al, 1985). This showed the presence of a DH site at the 5' end of the gene in all cells, regardless of the GSHPx mRNA level. In addition, a DH site at the 3' end of the gene was detected only in induced Friend cells. Since induced Friend cells have a GSHPx mRNA level approximately 20x higher than either of the two non-erythroid cell lines, it seemed possible that formation of this 3' DH region was required for increased transcription of the gene. This idea is supported by the observation that cell hybrids formed between neuroblastoma and Friend cells not only have a GSHPx mRNA level equal to that of the neuroblastoma parent, but do not possess the 3' DH site.

3.4.1. Comparison of the GSHPx gene chromatin structure in induced Friend and other cells.

To determine if this 3' DH site was always associated with an increased steady state level of GSHPx mRNA or if it was a peculiarity of induced Friend cells, the chromatin structure of the GSHPx gene was examined in cells expressing a high level of GSHPx mRNA (liver, kidney, induced and uninduced Friend cells) and cells expressing a low level of GSHPx mRNA (brain).

To address this question, nuclei were prepared from the different sources, aliquots digested with a range of DNAse I concentrations and DNA purified as described in Chapter 2.

Following the strategy of Affara et al (1985) for detecting both 5' and 3' DH sites, these samples were digested with Hind III, electrophoresed through agarose, transferred to nitrocellulose and hybridised to the 706bp Eco RI exon 2 probe as described in Chapter 2. Although this approach does not locate the precise positions of DH sites, it is a useful diagnostic procedure to allow sequences within approximately 5kb upstream and downstream of the transcription initiation site to be scanned simultaneously. Therefore, any gross differences between the various cell types/tissues will be apparent using this approach.

Figure 3.IV.1 shows the results of such an experiment. In uninduced Friend cells the parental band decreases in intensity with increasing DNAse I concentration and as this happens 3-4 bands of lower Mr appear or increase in intensity. Some of these bands are present in the DNA prepared from non-DNAse I-digested nuclei and these are probably the result of partial degradation of nuclei from dead or fragile differentiating cells present in the cultures, since they do not occur in similarly treated nuclei from tissues (compare with 0 DNAse I - treated nuclei from kidney in this figure and also from brain and liver - below).

The bands of lower Mr present in uninduced Friend cells have been labelled I-IV in Figure 3.IV.1. The band II/III present in the 0.5 units DNAse I/mg DNA lane has been so labelled since it is broader than the corresponding band in the 0 units DNAse I/mg DNA lane. It therefore seemed likely that it was composed of two DH sites although these were not readily resolved by this analysis.

As the parental band is approximately 9.7kb and the position

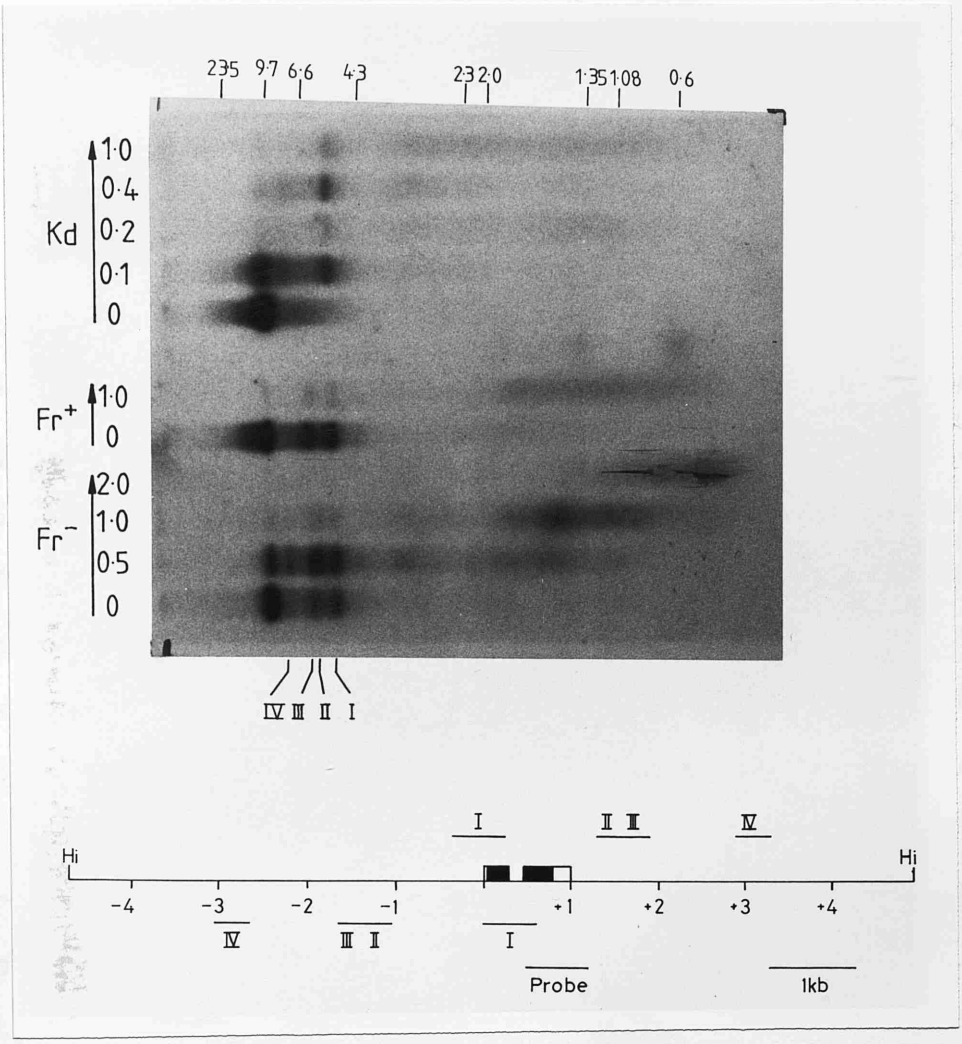


Figure 3.IV.1. Chromatin structure of the GSHPx locus in erythroid and kidney cells.

Aliquots of nuclei from uninduced (Fr^-) or induced (Fr^+) Friend cells or from kidney (Kd) were digested with the various concentrations of DNAse I indicated (units/mg DNA, Fr; units/µg DNA, Kd). Different concentrations of DNAse I are required in digestion of nuclei from tissues as opposed to cell cultures since the composition of the digestion buffers and the temperature of digestion are different in the two cases (see Chapter 2 for details). Following purification, DNA was digested with Hind III and 20µg electrophoresed through a 1% w/v agarose gel cast and run in TAE buffer. Following transfer of DNA to nitrocellulose, GSHPx homologous sequences were revealed by hybridisation to the 706 b.p. Eco RI probe. Top: Autoradiograph. The lines along the top of the autoradiograph indicate the positions of the Hind III cut λDNA Mr markers, whose sizes (in kb) are indicated. Bottom: Line diagram showing the GSHPx locus. The positions of the two exons are indicated by the boxed regions, filled-in areas corresponding to protein coding segments. Arabic numerals show the distance (in kb) from the major transcription initiation site. The probe origin is indicated. Bars associated with Roman numerals above or below the main line indicate the two alternative possibilities for the positions of DH sites I - IV discussed in the text.

of the Hind III site upstream of the GSHPx gene has been mapped (Figure 3.I.1), the downstream Hind III site could be calculated to be approximately 5kb from the major transcription initiation site. In addition, since the sub-bands produced by DNAse I cleavage at DH sites must span at least part of the probe sequences and have as one of their ends a Hind III site, possible positions for the DH sites can be calculated. Sub-band I, which is 4.7-5.3kb in length, is produced by a DH site just upstream from the major transcription initiation site, since the other possible position for a DH site generating a similarly sized band (centered within the intron) overlaps the probe sequences by only 50bp and is therefore unlikely to give a strong hybridising signal. Moreover, if the DH site were in this second position, then another band 4.4 to 5.0kb in size would be visible, which it is not.

The remaining bands visible in DNA from uninduced Friend cell nuclei could be produced by DNAse I cleavage at the positions upstream or downstream indicated on the line diagram. However, there are lower Mr diffuse bands of weaker intensity visible in lanes containing DNA from nuclei treated with higher concentrations of DNAse I. These are the lanes in which DNA fragments produced by DNAse I cleavage at two DH sites would be expected. Such bands would only be detected in a blot of the kind shown in Figure 3.IV.1 if they spanned at least part of the probe. Therefore, the other DH sites present in uninduced Friend cells are located 3' to the gene. This conclusion is supported by the sizes of the bands (3.4kb, cleavage at I and IV; 1.9kb, cleavage

at I and III; 1.6kb, cleavage at I and II). Further studies (see below) confirm this assignment.

In HMBA induced Friend cells only DH sites I and III are visible. This on its own is not enough to allow the conclusion that site II and IV are absent from induced Friend cells since the nuclei were treated with only a single concentration of DNAse I. However, the fact that the internal cleavage fragments at 3.4kb and 1.6kb visible at a higher DNAse I concentration in uninduced cells are not as apparent in induced cells argues against the presence of sites II and IV in induced Friend cells. These results for Friend cells agree with those obtained previously (Affara et al, 1985 and further unpublished data of Fleming and Harrison).

The analysis of kidney nuclei clearly shows the existence of DH sites I and IV, whereas sites II and III are absent over a broad range of DNAse I concentrations (Figure 3.IV.1). This is essentially the situation seen in liver (Figure 3.IV.2) except that DH IV is stronger relative to DH I in liver than in kidney or uninduced Friend cells. This indicates a significant difference in the chromatin structure of the GSHPx gene between these two tissues and erythroid cells. Moreover, a similar analysis of brain nuclei provided convincing evidence for the existence of DH site I alone (Figure 3.IV.2), the other band visible in this autoradiograph being irreproducible.

Therefore, three different states can be discerned for the GSHPx gene chromatin. In the first, represented by brain, where the GSHPx mRNA level is low, only the cap site DH site can be

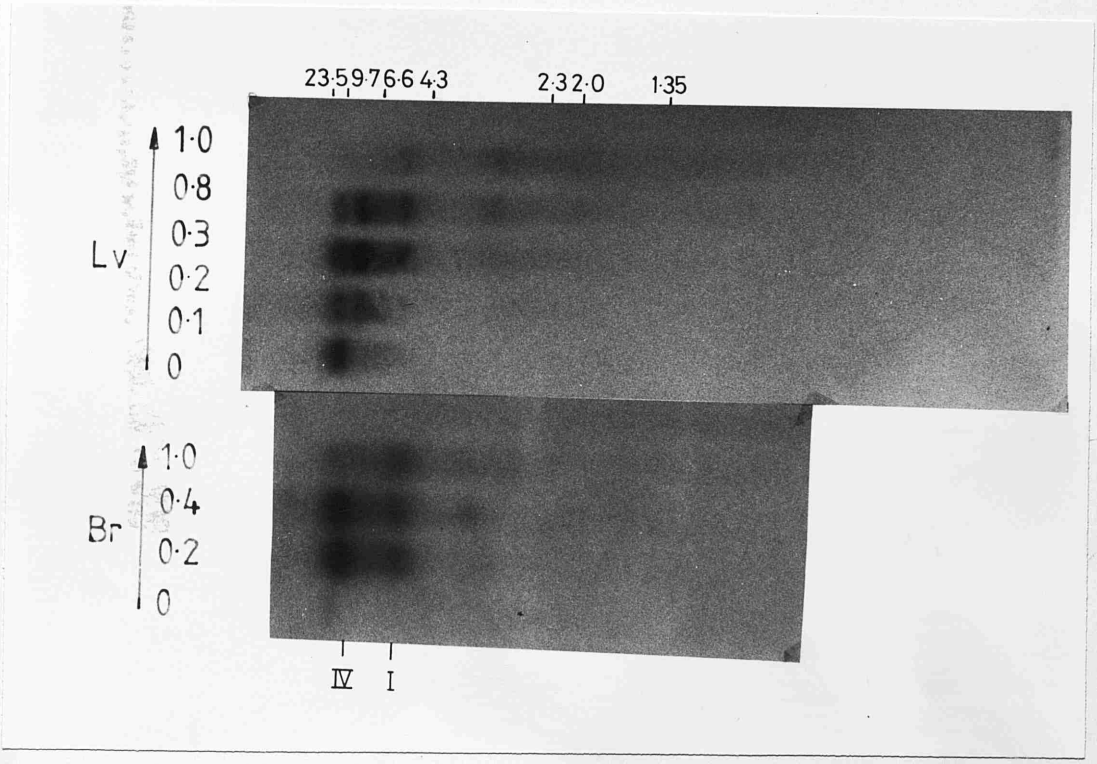


Figure 3.IV.2. Chromatin structure of the GSHPx locus in liver and brain.

Aliquots of nuclei from liver (Lv) or brain (Br) were digested with the various DNAse I concentrations indicated (units/µg DNA). Following purification DNA was digested with Hind III and 20µg electrophoresed through a 1% w/v agarose gel cast and run in TAE buffer. Following transfer of DNA to nitrocellulose, GSHPx homologous sequences were revealed by hybridisation to the 706 b.p. Eco RI probe. Roman numerals refer to DH sites in the same possible positions indicated in Figure 3.IV.1.

seen. The second, characterized by liver and kidney, in which the GSHPx mRNA level is approximately 20-fold higher than in brain, the cap site DH site is again present, but, in addition, DH site IV is also detectable. The third class, the erythroid cells, where the GSHPx mRNA level is similar to liver and kidney, these two DH sites are again present but are now joined by two others, DH sites II and III. Moreover, at least one of these DH sites disappears as the Friend cells are induced to undergo erythroid maturation. This analysis was therefore very informative. It was still, necessary, however, to determine the precise positions of the DH sites in the different tissues/cell types. To do this the indirect end-labelling technique developed by Nedospasov and Georgiev (1980) was adopted.

3.4.2. Determination of the position of DH I by indirect end-labelling.

The above experiments showed the presence of DH I in all tissues/cell types examined although the exact position of this DH site could not be assessed accurately. To determine this, DNA, purified as before from DNAse I treated nuclei, was restricted with Eco RI, electrophoresed through agarose, transferred to nitrocellulose and hybridised to the 469bp Sst II - Eco RI probe. Figure 3.IV.3 shows the presence of a broad region of DNAse I hypersensitivity extending 300bp upstream from the major transcription initiation site in uninduced Friend cells. This figure also shows the absence of other bands which would have been detected were other DH sites present in the 2.2 kb Eco RI

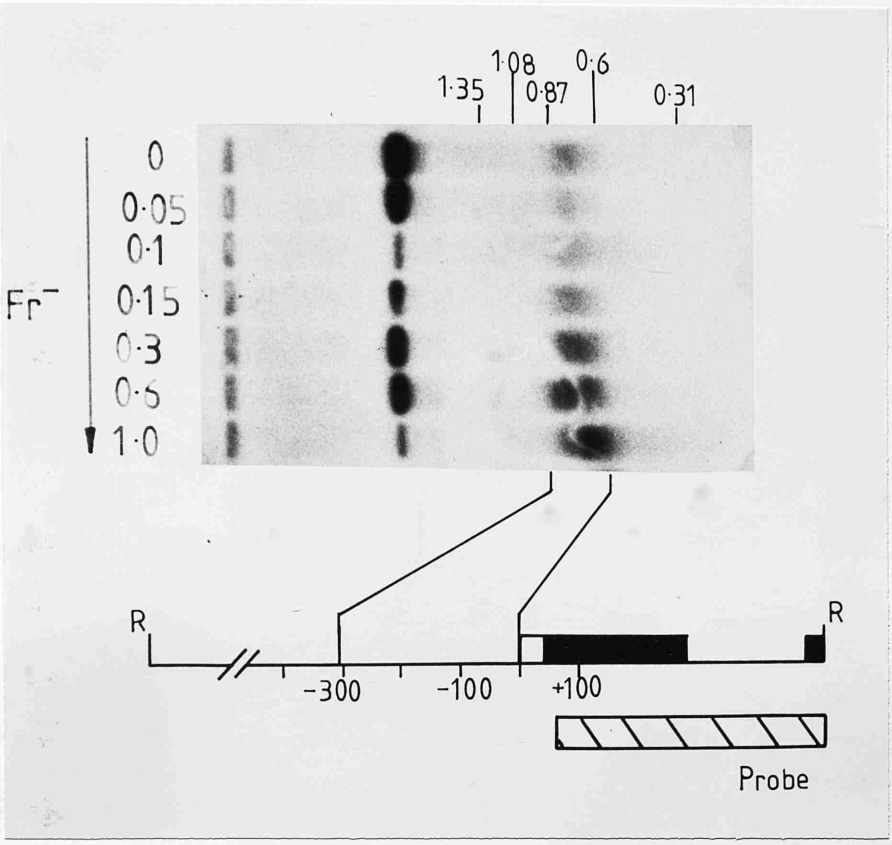


Figure 3.IV.3. Chromatin structure at the 5' end of the GSHPx gene in uninduced Friend cells.

Aliquots of nuclei from uninduced Friend cells were treated with the various DNAse I concentrations indicated (units/mg DNA). Following purification, DNA was digested with Eco RI and 20µg electrophoresed through a 2% w/v agarose gel cast and run in TBE buffer until the Orange G dye had run approximately 10cm. Following transfer of DNA to nitrocellulose GSHPx homologous sequences were revealed by hybridisation to the 469 b.p. Sst II - Eco RI probe. Top: Autoradiograph. The lines along the top of the autoradiograph indicate the positions of the Hinf I cut ϕx174 DNA Mr markers, whose sizes (in kb) are indicated. Bottom: Line diagram of the 2.2kb Eco RI fragment from the 5' end of the GSHPx gene. The positions of the exons and the origin of the probe are indicated. Numerals below the line diagram mark the distance (in b.p.) from the major transcription initiation site. The extent of DHI is indicated by the lines connecting the autoradiograph and line diagram.

fragment.

To ascertain if there were any differences in the structure of DH I in the different tissues/cell types, a similar analysis of DNA purified from these sources was performed using long 2% w/v agarose gels, to obtain higher resolution. Figure 3.IV.4 shows the presence of a broad DH site in uninduced Friend cells extending from -310bp to +10bp. Because of the diffuse nature of this hypersensitive region it is difficult to be confident of the exact position of its boundaries. Indeed estimates of these vary from experiment to experiment depending upon the amount of DNA loaded per lane, the efficiency of transfer of DNA to the membrane and the specific activity of the probe. However, within any one experiment the results from different tissues/cell types should be directly comparable. Thus DH I appears tighter in brain and kidney than it does in uninduced Friend cells (Figure 3.IV.4). However, since this could be due to the DH site being weaker in brain than uninduced Friend cells or the fact that less DNA appears to have been loaded in the kidney panel than was estimated, it remains likely that the boundaries of DH I are roughly the same in all three tissues. Moreover, it is noteworthy that in each of the tissues examined DH I appears as a doublet with the "cold spot" being located in the same position in kidney and brain but further upstream in uninduced Friend cells. This difference may be more apparent than real since the Friend cell nuclei were at least partly digested by an endogenous nuclease and microheterogeneity in the exact position of nuclease hypersensitive sites has previously been observed using different

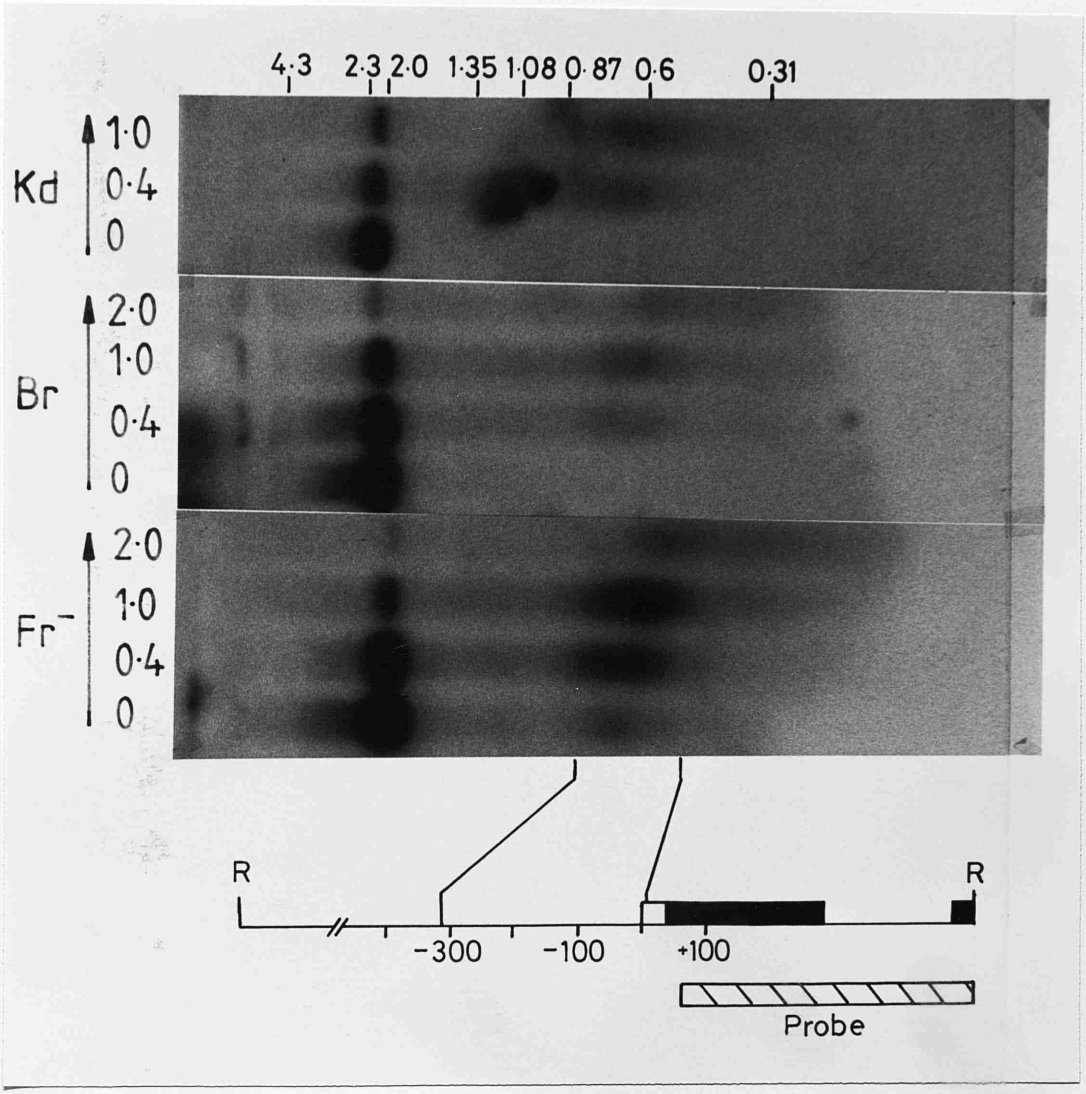


Figure 3.IV.4. Chromatin structure at the 5' end of the GSHPx gene in kidney, brain and uninduced Friend cells.

Aliquots of nuclei were digested with the various DNAse I concentrations indicated (units/mg DNA, Fr^-; units/µg DNA, Br, Kd). Following purification, DNA was analysed as described in Figure 3.IV.3, except that electrophoresis was allowed to continue until the Orange G dye had run approximately 17 cm. Top: Autoradiograph. The lines along the top of the autoradiograph indicate the positions of the Hind III cut λDNA and Hinf I cut ϕx174 DNA Mr markers, whose sizes (in kb) are indicated. Bottom: Line diagram of the 2.2kb Eco RI fragment from the 5' end of the GSHPx gene. The various features of this diagram are described in Figure 3.IV.3 legend. The extent of the DH site in Fr^- is indicated by the lines connecting the autoradiograph and the line diagram.

nucleases (Eissenberg et al, 1985).

A similar analysis was performed on induced Friend cell nuclei (Figure 3.IV.5). In this experiment DH I appears to extend from -220bp to -25bp in uninduced Friend cells. The difference between this result and that shown in Figure 3.IV.4 is probably caused by less DNA having been used in this experiment. Nevertheless, the important point to note is that the same diffuse DH site is present in induced Friend cells as in uninduced Friend cells.

Liver and kidney nuclei were compared by the same procedure (Figure 3.IV.6). Although the hybridising signals here are weak, the same broad DH site is evident in kidney as was seen in Figure 3.IV.4. In addition, a similar diffuse DH site is present in liver. From these experiments it seems that DH I is present as a diffuse region of hypersensitivity similarly located in each of these tissues/cell types.

3.4.3. Determination of DH sites II - IV by indirect end-labelling.

In order to analyse the chromatin structure over a region extending at least 9 kb downstream from the major transcription initiation site, DNA, purified after digestion of uninduced Friend cell nuclei with DNAse I, was digested with Sst II. Following electrophoresis through agarose and transfer to nitrocellulose, DNA was hybridised to the 706bp Eco RI probe. This confirms the presence of DH sites II and IV (Figure 3.IV.1) as broad regions of hypersensitivity extending from +1.25kb to +1.45kb (II) and

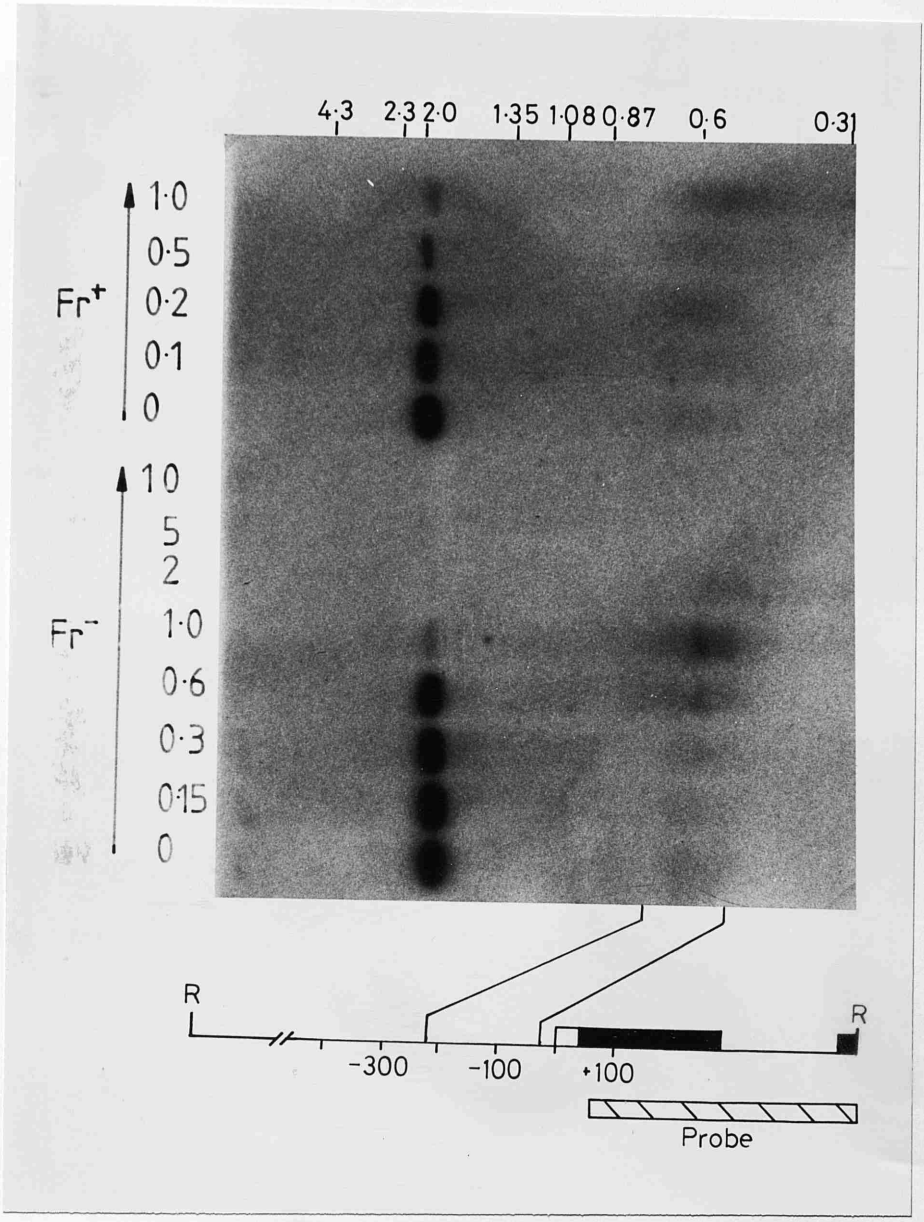


Figure 3.IV.5. Chromatin structure at the 5' end of the GSHPx gene before and after induction of erythroid maturation.

Aliquots of nuclei were digested with the various DNAse I concentrations indicated (units/mg DNA). Following purification DNA was analysed as described in Figure 3.IV.3, except that only 10µg of DNA was loaded per lane. Top: Autoradiograph. The lines along the top of the autoradiograph indicate the positions of the Hind III cut λDNA and Hinf I cut ϕx174 DNA Mr markers, whose sizes (in kb) are indicated. Bottom: Line diagram of the 2.2kb Eco RI fragment from the 5' end of the GSHPx gene. The various features of this diagram are described in Figure 3.IV.3 legend. The extent of the DH site is indicated by the lines connecting the autoradiograph and line diagram.

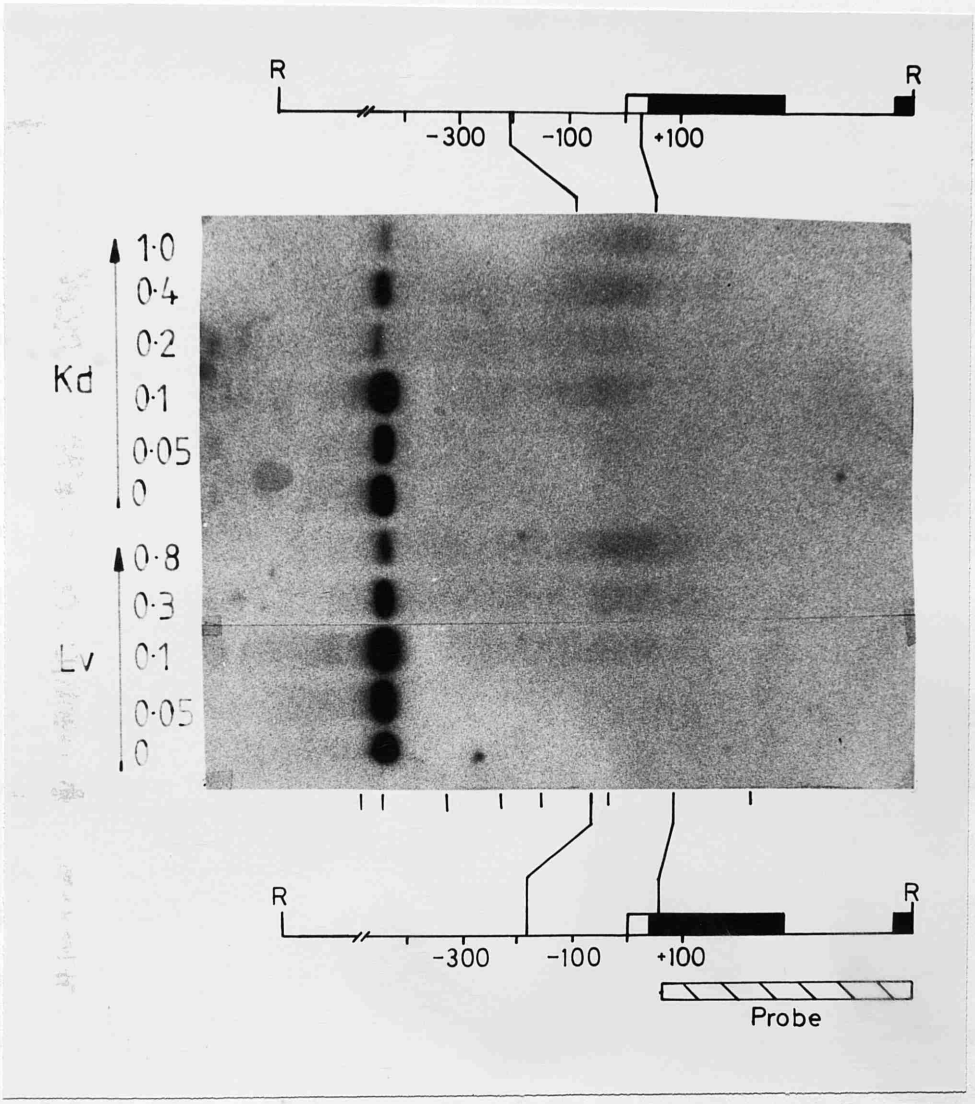


Figure 3.IV.6. Chromatin structure at the 5' end of the GSHPx gene in liver and kidney cells.

Aliquots of nuclei were digested with the various DNAse I concentrations indicated (units/µg DNA). Following purification DNA was analysed as described in Figure 3.IV.3. Autoradiograph: The lines along the bottom of the autoradiograph indicate the positions of Mr markers whose sizes are (from left to right) 2.3, 2.0, 1.35, 1.08, 0.87, 0.6 and 0.31kb. Line Diagram: The various features of these line diagrams are described in Figure 3.IV.3 legend. The extent of the DH sites in both tissues are indicated by connecting lines between the autoradiograph and the line diagrams.

+3.45kb to +3.65kb (IV) (Figure 3.IV.7). Moreover, DH site III was resolved into a doublet of DH sites centred at +1.7kb (III_1) and +1.85kb (III_2). In addition, several minor bands are visible in Figure 3.IV.7 which may correspond to weak DH sites at +0.75 kb, +1.05kb, +2.75kb and +4.85kb.

Applying a similar analysis to induced Friend cell nuclei confirms the absence of any DH sites other than DH III, which as with uninduced Friend cells, was resolved into a doublet of DH sites (Figure 3.IV.8). The doublet nature of this DH site is not obvious in Figure 3.IV.8 although it is apparent on the original autoradiograph.

Liver nuclei were analyzed similarly and the presence of DH site IV was thus confirmed (Figure 3.IV.9). Although the positions of DH site IV in liver and uninduced Friend cells overlap (as determined by indirect end-labelling), the liver site is centred slightly upstream from the Friend cell site. This latter point, together with the fact that the peak of DNAse I hypersensitivity is towards the 5' edge of DH site IV in liver, might mean that, in fact, the chromatin structure in this region is different in these two cell types. Although the background hybridisation on the autoradiograph shown in Figure 3.IV.9 is higher than usual, there is evidence for other weak DH sites besides DH IV. These may be genuine, but this would would require confirmation.

A closer examination of the more proximal DH sites was performed by digesting purified DNA from DNAse I treated uninduced Friend cell nuclei with Eco RI [which cuts the genomic DNA 1.4kb

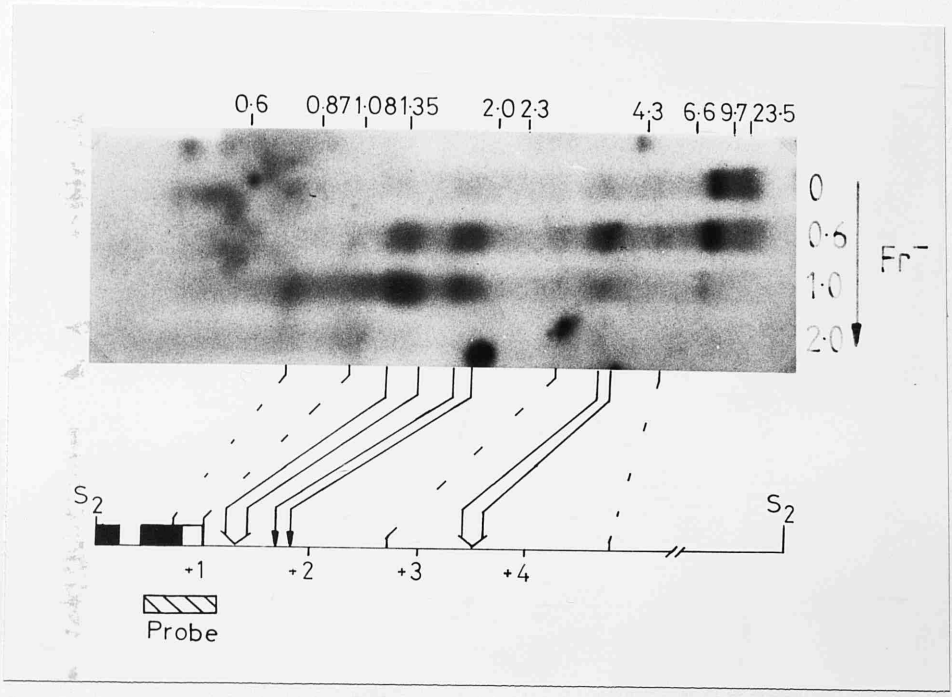


Figure 3.IV.7. Chromatin structure at the 3' end of the GSHPx gene in uninduced Friend cells.

Aliquots of nuclei were digested with the various DNAse I concentrations indicated (units/mg DNA). Following purification, DNA was restricted with Sst II and 20µg electrophoresed through a 1% w/v agarose gel cast and run in TAE buffer until the Orange G dye had run approximately 17 cm. Following transfer of DNA to nitrocellulose GSHPx homologous sequences were revealed by hybridisation to the 706 b.p. Eco RI probe. Top: Autoradiograph. The lines along the top of the autoradiograph mark the positions of the Hind III cut λ and Hinf I cut ϕx174 DNA Mr markers, whose sizes (in kb) are shown. Bottom: Line diagram of the Sst II fragment extending downstream from the Sst II site at +59 b.p. The positions of the exons and the origin of the probe are indicated. Numerals below the line diagram mark the distance (in kb) from the major transcription initiation site. Arrowheaded lines mark the positions of the corresponding DH sites, while broken lines mark the position of weak putative DH sites.

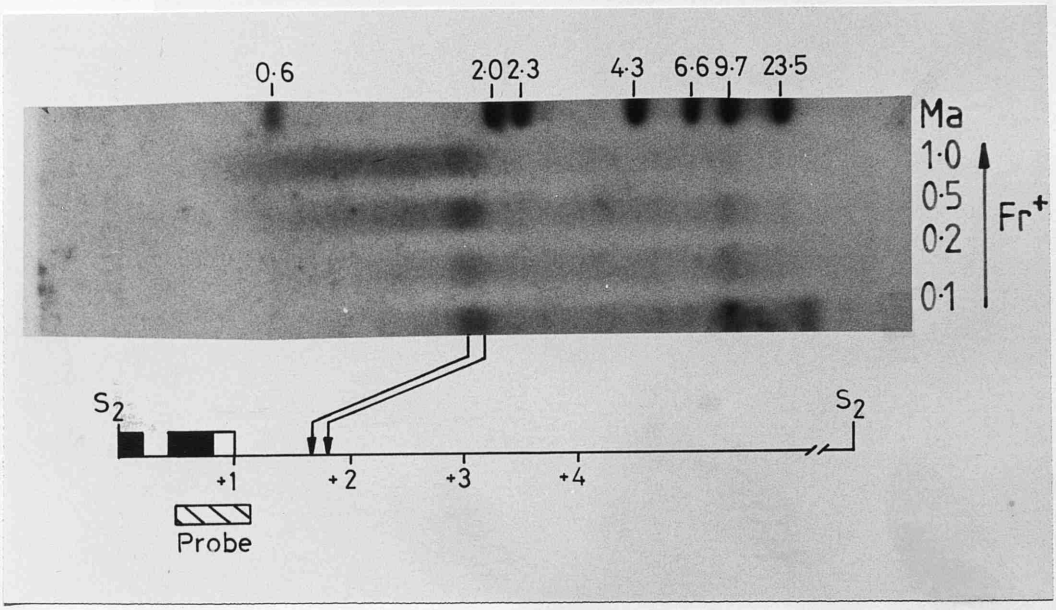


Figure 3.IV.8. Chromatin structure at the 3' end of the GSHPx gene in induced Friend cells.

Aliquots of nuclei were digested with the various DNAse I concentrations indicated (units/mg DNA). Following purification DNA was analysed as described in Figure 3.IV.7. Top: Autoradiograph. The lines along the top of the autoradiograph mark the positions of the Hind III cut λDNA Mr markers (Ma), whose sizes (in kb) are shown. Bottom: Line diagram of the Sst II fragment extending downstream from the Sst II site at +59 b.p. The various features of this line diagram are described in Figure 3.IV.7 legend. The double arrow head marks the positions of the doublet DH site III.

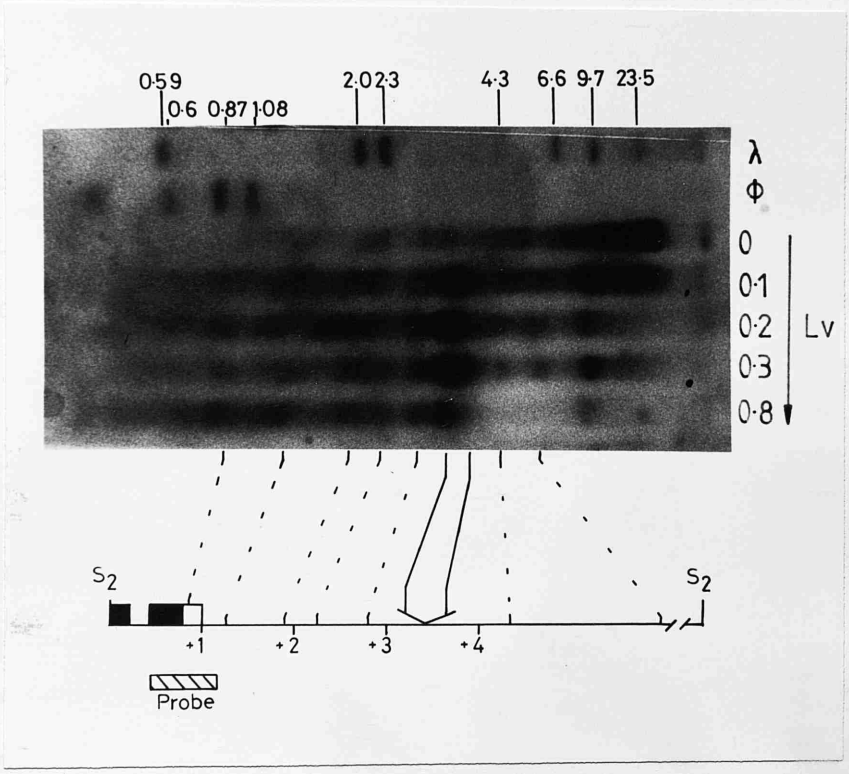


Figure 3.IV.9. Chromatin structure at the 3' end of the GSHPx gene in liver cells.

Aliquots of nuclei were digested with the various concentrations of DNAse I indicated (units/μg DNA). Following purification DNA was analysed as described in Figure 3.IV.7. Top: Autoradiograph. The lines along the top of the autoradiograph mark the positons of the Hind III cut λ and Hinf I cut ϕx174 DNA Mr markers, whose sizes (in kb) are shown. Bottom: Line diagram of the Sst II fragment extending downstream from the Sst II site at +59 b.p. The various features of this line diagram, including the meaning of the arrowheaded lines and broken lines, are as described in Figure 3.IV.7 legend.

downstream from the poly (A) site] and indirectly end-labelling with the 706bp Eco RI fragment from λR68A. This is possible since the 3' Eco RI site of the 706bp fragment was artificially created during construction of the BALB/c genomic library from which λR68A was isolated; therefore no such Eco RI site exists in the mouse genome. The result (Figure 3.IV.10) shows that the broad DH site II present in uninduced Friend cells and previously mapped at +1.25kb to +1.45kb (Figure 3.IV.7) is in fact made up of a doublet of DH sites centred at +1.28kb (II_1) and +1.37kb(II_2). In addition the positions of the other doublet bands III_1 and III_2 at +1.71kb and +1.85kb were confirmed. As anticipated, none of these bands was detected in brain.

3.4.4. Summary of DH sites present around the GSHPx gene in cells from different tissues.

A diagrammatic summary of the results obtained in this study is presented in Figure 3.IV.11. A broad region of DNAse I hypersensitivity (DH I) is present upstream from the major transcription initiation site in all cells examined. Two doublet DH sites (II and III) have been detected as major DH sites so far only in erythroid cells and one of these (DH II) disappears upon HMBA induced erythroid maturation of Friend cells. It is possible that these DH sites are present as weak DNAse I cleavage sites in liver and kidney, but these tissues were not examined closely enough to resolve this. Another DH site present in uninduced Friend cells (DH IV) which is lost upon terminal maturation is also present in liver cells and probably also kidney cells (the

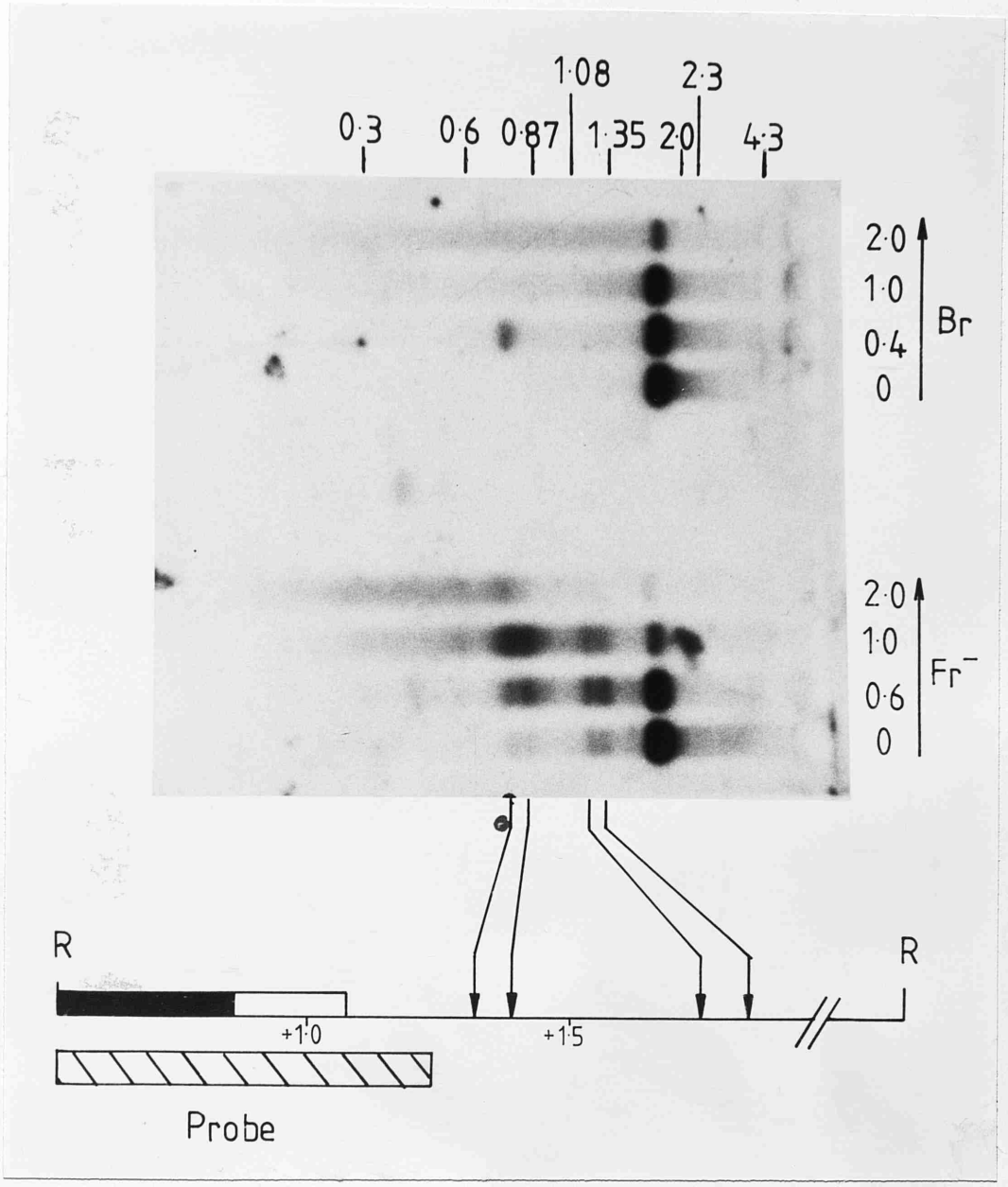


Figure 3.IV.10. Fine mapping of DH sites II and III.

Aliquots of nuclei were digested with the various concentrations of DNAse I indicated (units/mg DNA, Fr^-; units/µg DNA, Br). Following purification, DNA was digested with Eco RI and 20µg electrophoresed through a 2% w/v agarose gel cast and run in TBE buffer until the Orange G dye had run approximately 17cm. Following transfer of DNA to nitrocellulose GSHPx homologous sequences were revealed by hybridisation to the 706 b.p. Eco RI probe. Top: Autoradiograph. The lines along the top of the autoradiograph mark the positions of the Hind III cut λ and Hinf I cut ϕx174 DNA Mr markers, whose sizes (in kb) are shown. Bottom: Line diagram of the 1.9 kb Eco RI fragment. The position of GSHPx exon 2 sequences and the origin of the probe are indicated. Numbers below the line mark the distance (in kb) from the major transcription initation site. Arrowheads mark the positions of the DH sites.

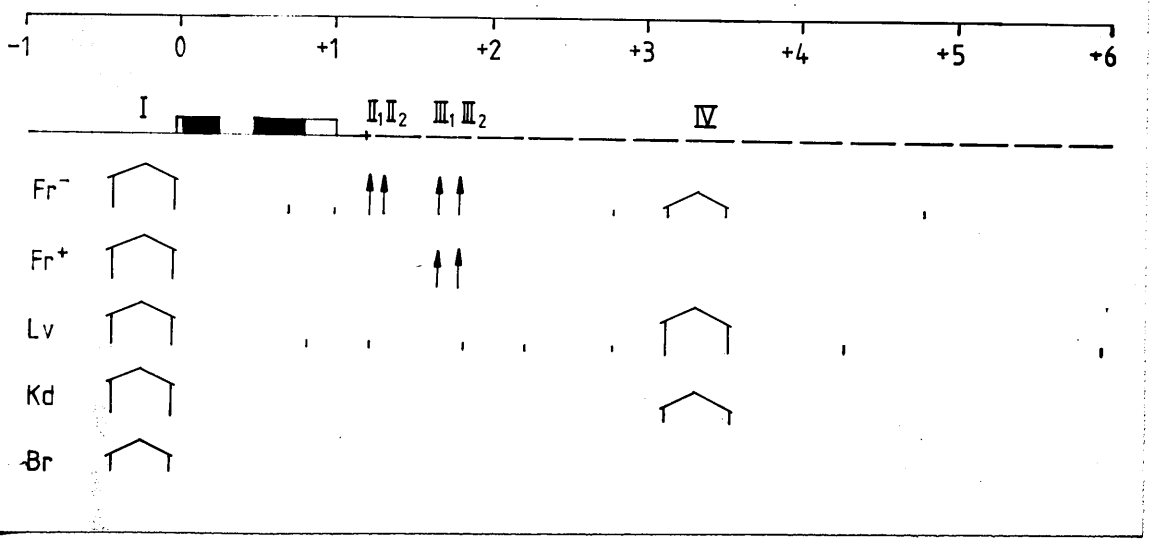


Figure 3.IV.11. Summary of the chromatin structure around the GSHPx gene in different cell types.

The line diagram represents the DNA at the GSHPx locus. The two boxes represent the exons, filled in regions corresponding to protein coding segments. A solid line represents the cloned DNA with uncloned DNA indicated by the broken line. The scale above the line diagram marks the distance (in kb) from the major transcription initiation site. The arrows below the line indicate the presence of particular DH sites within a given tissue/cell. Variations in the size of the arrow reflect differences in the intensity of the DH site within different tissues/cells. The different DH sites are marked by Roman numerals. Small lines indicate the possible presence of additional weak DH sites.

position of DH IV in kidney has not been determined by indirect end-labelling and so it is still possible that this site is located at the alternative upstream position shown in Figure 3.IV.1). This DH site appears stronger in liver cells than in kidney cells or uninduced Friend cells. In kidney this probably reflects the fact that a smaller proportion of cells express high levels of GSHPx in kidney than in liver (Yoshimura et al, 1980; Mizuiri et al, 1986). The difference between liver and uninduced Friend cells may be explained by the presence of DH sites II and III in Friend cells causing cleavage and therefore reduction in intensity of the hybridising bands seen as DHIV. In addition, some minor bands were detected using filters prepared from Sst II digested DNA. In each case these bands have only been detected once; moreover, the background on the blots on which they were detected was high. Consequently, their presence would require confirmation by independent digests before any great significance could be attached to them.

Assuming that the downstream DH sites (II, III and IV) are related to GSHPx expression and not to expression of an, as yet unidentified, adjacent transcription unit, it would appear that the GSHPx gene is regulated differently in erythroid as opposed to liver and kidney cells. Although this study has not provided any evidence to support the notion, it is still possible that the GSHPx gene is regulated differently in high GSHPx expressing kidney and liver cells. In this respect, it is noteworthy that an enhancer element (present within a tissue-specific DH site) responsible for constitutive expression of the chicken lysozyme

gene in macrophages is located 6.1 kb upstream of the gene's transcription initiation site (Theisen et al, 1986). This is further upstream than the extent of the analysis of the GSHPx chromatin presented here and it may therefore be worth re-examining the filters used in Figures 3.IV.1 and 3.IV.2 with a probe from the 5'side of the upstream Hind III site.

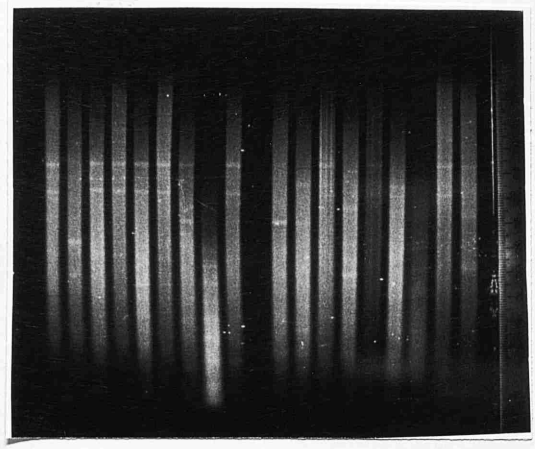
Thus far, DHIV is the site best correlated with an increased GSHPx mRNA level, although this correlation is not absolute since DHIV is absent from induced Friend cells. This absence may be related to the fact that the nucleus is becoming less active as maturation proceeds and in this context it would be interesting to know when, during erythroid maturation, DH sites II and IV disappear and if transcription of the GSHPx gene then continues at the same rate.

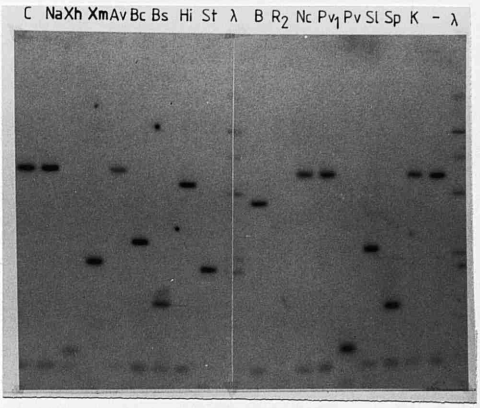
3.4.5. Further restriction mapping of the GSHPx locus as a preliminary to cloning of the 3' flanking DNA.

Before the DNA covering DH sites II, III and IV can be analysed further it must first be cloned. Previous attemps to obtain clones which overlapped the 3' end of λR68A by rescreening the library from which λR68A was isolated were unsuccessful (J. O'Prey and P. Harrison, personal communication).

A method has recently been described for cloning genomic DNA in plasmids following fragment enrichment on an agarose gel (Nicholls et al, 1985). What is required for this is that a fragment generated by digestion of genomic DNA with one or two restriction enzymes is (1) not digested by several other enzymes

(a total of at least five) (2) sufficiently large that electrophoresis of the total digest resolves it from >95% of the genome (3) not so large that it will not be ligated efficiently into a plasmid vector. A partial map of the uncloned DNA was available (Affara et al, 1985) but this was not detailed enough to allow cloning of the flanking DNA via fragment enrichment. The available data showed that there is an Msp I site in λR68A 526bp from its 3' end and that the next Msp I site is 5.2kb downstream from this in the 3' flanking DNA. Furthermore, the 5.2kb Msp I fragment is not digested by Sst II. Therefore, mouse genomic DNA was digested with Msp I and Sst II and then with a variety of other restriction enzymes before electrophoresis through agarose, transfer to nitrocellulose and hybridisation to the 706bp Eco RI probe (Figures 3.IV.12). It was hoped that the results of this analysis would provide additional information concerning restriction enzyme sites around the GSHPx gene which would allow a successful cloning attempt of DNA covering DH sites II, III and IV to be made. Thus, any of the enzymes used in Figure 3.IV.12 which can be judged by visual inspection of the ethidium bromide stained gel to have digested the genomic DNA but which have left the 5.2 kb Msp I fragment intact should be useful in this type of exercise. Figure 3.IV.12 shows that the majority of enzymes tested reduced the size of the Msp I fragment, but six enzymes did not (Cla I, Nar I, Ava I, Nco I, Pvu I and Kpn I). Of these six, three had definitely digested the genomic DNA (Cla I, Nco I and Kpn I), the others had either not digested the DNA, or if they had, were unlikely to give a useful enrichment, since they did not





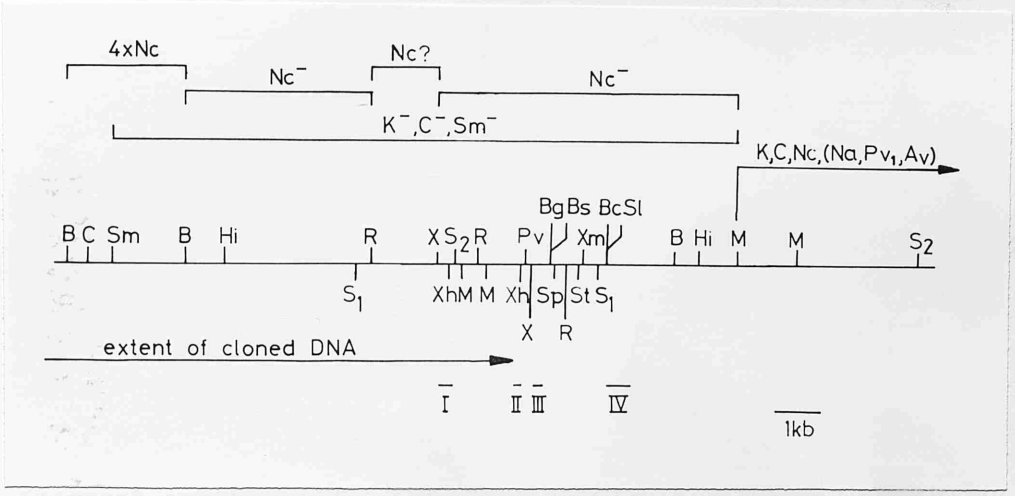


Figure 3.IV.12. Restriction mapping at the 3' end of the GSHPx gene.

Genomic DNA purified from M707T cells was digested with Sst II and Msp I. 20µg aliquots of the resultant digested DNA were then digested with a third enzyme and electrophoresed through a 0.8% w/v agarose gel cast and run in TAE buffer. Following transfer to nitrocellulose, GSHPx 3' sequences were revealed by hybridisation to the 706 b.p. Eco RI probe. Top left: ethidium bromide stained gel. Top right: overnight exposure (-70°C with intensifying screen) of the washed filter to Kodak XAR-5 film. Note: the ethidium bromide stained gel was photographed upside down, therefore the order of the lanes shown is opposite to that for the autoradiograph. Bottom: Restriction map of the GSHPx locus showing the positons of the relevant enzymes as well as the absence of particular restriction enzymes sites within some DNA fragments (restriction enzyme⁻). The extent of the cloned λR68A DNA and the positions of DH sites I, II, III & IV are indicated.

significantly reduce the average size of the genomic DNA. In addition since DH site IV is located 5' to the downstream Bam HI site, both this enzyme and Hind III may be used to enrich for DNA covering DH sites II, III and IV. The additonal information to be gleaned from this analysis allowed a more detailed map of the uncloned DNA to be constructed (Figure 3.IV.12).

DNA was next restricted with Msp I, Sst II, Cla I, Nco I and Kpn I and electrophoresed through agarose. The 5.2kb fragment was however, judged not to be sufficiently resolved from the bulk of the DNA for a representative library to be constructed, (data not shown; Nicholls et al, 1985), a situation which was not improved by further digestion of DNA with Hind III and Bam HI.

There are three possible directions from which to proceed from here. Genomic DNA could be digested with Bam HI and Hind III, the 9.3 kb GSHPx gene fragment purified, digested with Sst II, Msp I, Kpn I, Cla I and Nco I and the 3.9kb Msp I - Bam HI fragment purified and cloned. This method has the disadvantage of requiring two purification steps from agarose and so losses are likely to be high. An alternative would be to screen for additional enzymes which would not cut the Msp I - Bam HI fragment but would give a useful enrichment compared to the bulk of DNA; this has the disadvantage of requiring additional analysis before being begun. The third possibility, and the one which I would choose, is to clone the 9.3 kb Hind III - Bam HI fragment. Kpn I and Cla I do not cut the 5.2 kb Msp I fragment; in addition it is known that neither of these enzymes cuts λR68A 3' to the Hind III site (Figure 3.1.1) and therefore these could be used for fragment

enrichment. Furthermore, Sma I may also be useful here, since this enzyme has only one site in λR68A located 5' to the Hind III site and as the Msp I recognition site is a subset of that for Sma I it does not have a restriction site 5' to the downstream Msp I site. Nco I might also be useful here but further analysis is required to determine this (Figure 3.IV.12).

Chapter 4

Discussion

The discovery that the ep19 gene encodes glutathione peroxidase was a major milestone in the development of this thesis. By drawing upon the literature surrounding GSHPx, it allows for a re-evaluation of the significance of the tissue distribution of ep19 mRNA as well as indicating additional cell types where the level of GSHPx mRNA may be elevated. Moreover, it has suggested potential aspects of the regulation of GSHPx gene expression which have yet to be investigated. Before moving on to discuss these topics in relation to the results obtained in this study, it will be useful to review the implications of the discovery that selenocysteine is encoded by UGA.

4.1 For selenocysteine read UGA.

4.1.1. GSHPx selenocysteine is encoded by UGA.

Numerous studies have shown that selenium-deficient animals or cultured cells have decreased levels of GSHPx activity (reviewed by Neve et al, 1985), which are reflected in decreased levels of GSHPx protein (Takahashi; et al, 1986b; Takahashi and Cohen, 1986). Indeed, the livers of rats fed a selenium-free diet contain no protein recognisable by antibodies directed against GSHPx (Yoshida et al, 1982). Furthermore, during selenium repletion recovery of GSHPx activity requires new protein synthesis (Perona et al, 1978; Sunde and Hoekstra, 1980b; Cohen et al, 1985b; Hornsby et al, 1985; Speier et al, 1985). Taken together, these facts indicate that either selenium is required for translation of the GSHPx mRNA or selenium is required to stabilise a hypothetical apo-GSHPx, which in the absence of selenium is rapidly degraded.

More recent evidence has shown selenocysteine to be incorporated into protein during translation (Hawkes and Tappel, 1983) and has identified a selenocysteyl-tRNA ($tRNA^{Sec}$) in rat liver (Hawkes et al, 1982). However, neither the sequence nor the codon recognition of $tRNA^{Sec}$ have been determined. Consequently, the demonstration that UGA is present in GSHPx mRNA in a position colinear with that of selenocysteine in GSHPx provides the first evidence for a selenocysteine codon distinct from other "sense" codons. Furthermore, the recent report of an in-frame TGA codon within the gene encoding the selenopeptide subunit of E.coli formate dehydrogenase (fdhF) (Zinoni et al,

1986) suggests that the appearance of UGA as a selenocysteine codon was an early event in the evolution of the genetic code. Prior demonstrations that E.coli (Young and Kaiser, 1975) or rabbit (Wilhelmsen et al, 1985) cysteineyl-tRNA can be aminoacylated with selenocysteine at selenocysteine concentrations in excess of 40μM are almost certainly physiologically irrelevant, since the estimated maximum physiological concentration of selenocysteine is 200nM (Wilhelmsen et al, 1985).

4.1.2. Translation of selenopeptide mRNAs during selenium deficiency.

What happens to a ribosome in a selenium deprived eukaryotic cell when it reaches a selenocysteine codon is presently unclear. However, it is notable that in E.coli, the expression of β-galactosidase sequences located 3' to the UGA codon of fdhF mRNA occurs at 10% of the normal level when the cells are grown in selenium deficient medium (unpublished data of Bock and Zinoni, cited in Zinoni et al, 1986). Since this effect has been shown not to be due to a transcriptional stimulation by selenium, this means that in the absence of selenocysteyl-tRNA, the ribosome either stalls indefinitely during translation of a selenopeptide mRNA or terminates translation at the UGA codon. Assuming that a similar situation pertains in a eukaryotic cell and that the ribosome terminates translation at the UGA codon of GSHPx then the N-terminal GSHPx sequences would be released and could then either be degraded or be used to signal cellular

selenocysteine starvation. However, it is perhaps unlikely that the same peptide could have two such diverse functions. It is known that reinitiation of translation following termination can occur in eukaryotic cells provided an AUG initiator is present within 50-90nt upstream or downstream of the termination codon (Liu et al, 1984; Peabody et al, 1986; Peabody and Berg, 1986). Therefore, it is possible that the open reading frame extending from +161nt to +250nt (see Figure 4.1) could be translated by this mechanism to yield a selenium starvation signal peptide. This is a rather speculative suggestion; however it has been demonstrated that some tissues, particularly testes and adrenals, preferentially retain selenium at the expense of other tissues during selenium deprivation (Behne et al, 1982; Behne and Hofer-Bosse, 1984). Thus it is clear that some mechanism exists preferentially to route scarce selenium to these sites. The selenium in these two tissues is involved in reproduction and this may provide the rationale for the sequestration of the element within these sites : at least part of the adrenal gland selenium is in the form of GSHPx which is thought to have a role in the biosynthesis of steroids (Murakoshi et al, 1981), while the testes' selenium is incorporated into a selenopeptide which forms part of the sheath surrounding the midpiece mitochondria of spermatozoa and without which proper flagellar motility is not achieved (Behne et al, 1982; Sunde, 1984).

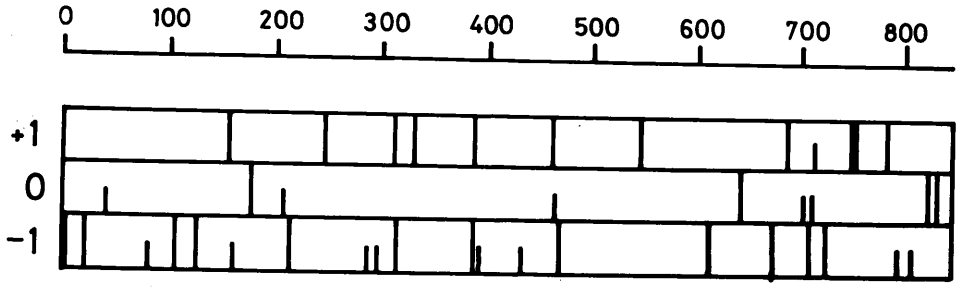


Figure 4.1. Positions of "termination" and methionine codons in all three reading frames of GSHPx mRNA. The sequence of the GSHPx mRNA between the major cap site and the major polyadenylation site has been translated in all three reading frames; 0, GSHPx frame; -1, +1, reading frames beginning 1nt before and 1nt after the GSHPx frame. The positions of "termination" codons are indicated by full vertical lines and those of methionine codons by half vertical lines in the relevant row. The scale at the top of the figure is in nt.

4.1.3. Coding potential of "termination" codons in other systems.

There have been previous reports of termination codons being used to insert specific amino acids. For example, the prokaryote Mycoplasma capricolum uses UGA to encode tryptophan (Yamao et al, 1985). In addition ciliated protozoans possess an unusual genetic code in which all genes examined to date use UGA as the termination codon and in which UAA and UAG both specify glutamine insertion rather than termination (reviewed in Fox, 1985). These organisms possess three glutamyl tRNAs one recognizing the normal glutamine codons CAA and CAG, one recognizing UAA and UAG, and the other just UAG (Hanyu et al, 1986). The fact that these latter two tRNAs are more homologous to one another than either is to the more normal glutamyl tRNA has been used to argue that the two deviant $tRNA^{Gln}$s arose by gene duplication and divergence from the archetypal ciliate $tRNA^{Gln}$ gene early in ciliate evolution but after the ciliates branched off from the primitive eukaryotic ancestor (Hanyu et al, 1986).

The above examples refer to all-or-none functions of termination codons within a cell or a subcellular organelle. This is quite different from the situation with respect to selenocysteine incorporation where the same codon is utilised under different circumstances within the same cellular compartment to mean "stop" or "insert selenocysteine".

4.1.4. The context effect for selenocysteine incorporation.

Obviously, some other aspect(s) of the mRNA structure/sequence must be involved in distinguishing between a "stop" or an "insert selenocysteine" function for a particular UGA codon. Clues as to what these might be may be gleaned from ideas developed in an attempt to explain the action of experimentally induced prokaryotic suppressor tRNAs. Suppressor tRNAs can be of two types : nonsense suppressors insert an amino acid in response to a termination codon; whereas missense suppressors insert an incorrect amino acid in response to a sense codon. The suppression afforded by a particular tRNA is only very rarely complete, being dependent upon the suppressible codon : this is known as the context effect (Bossi, 1983). Suppressor tRNAs vary in their responsiveness to the effects of context; some are relatively efficient regardless of context, whereas others, which are often very weak, can have their activity altered dramatically by a change of context (Bossi and Roth, 1980; Ayer and Yarus, 1986). These context effects have been shown to be due primarily to the nucleotides immediately 3' to the suppressible codon in the case of nonsense suppression (Bossi, 1983; Miller and Albertini, 1983); whereas for missense suppression the immediate 5' nucleotides are of prime importance (Murgola et al, 1984).

How might these observations be explained? Several hypothesis have been advanced. The original suggestion (Bossi and Roth, 1980) that a fourth base pair formed between the suppressor tRNA and the mRNA caused the context effect has not

been upheld experimentally (Ayer and Yarus, 1986). The alternative suggestion (Miller and Albertini, 1983) that a hairpin loop structure formed next to the suppressible codon has a role in context remains a possibility, although if this is so it must be an infrequently used mechanism since such potential structures are only rarely observed (Miller and Albertini, 1983). Interestingly, Zinoni *et al* (1986) noted the presence of a potential hairpin loop structure around the UGA codon of fdhF mRNA. If this is involved in the recognition of the UGA as a selenocysteine codon then the context effect for selenocysteine incorporation must be different in *E.coli* and mice since no such potential structure exists around the UGA codon of GSHPx mRNA.

The remaining context effect hypotheses all stress the importance of tRNA:tRNA interactions occurring at the ribosome. All elongator tRNAs have uridine at position 33 (i.e. the nucleotide immediately 5' to the anticodon; see Figure 4.2). This fact has been used to postulate that hydrogen bonding between this base and the variable nucleotide 37 of the tRNA binding to the adjacent codon 3' to the suppressible codon [i.e. occupying the acyl (A-) site] could account for the context effect for nonsense suppression (Bossi, 1983; Figure 4.2). However, this idea presupposes that the suppressor tRNA can compete efficiently with termination factor(s) for binding to the A-site, is translocated to the peptidyl (P-) site and will dissociate from the ribosome in the absence of such hydrogen bonding to the tRNA binding to the A-site (Bossi, 1983; Menninger, 1976).

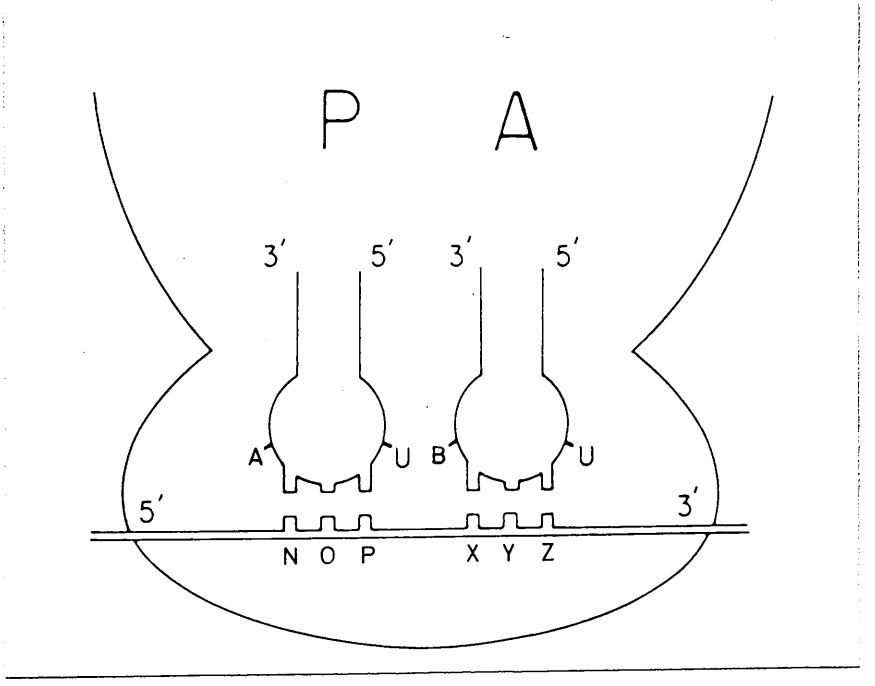


Figure 4.2. Representation of a ribosome translating the hypothetical mRNA sequence NOPXYZ. tRNAs occupying the peptidyl (P-) and acyl (A-) sites are shown together with their invariant U at nt 33 and their variable (A or B) nucleotides at position 37.

An adaptation of this hypothesis has been advanced to account for the context effect of nucleotides 5' to the suppressible codon in missense suppression (Murgola et al, 1984). In this case the variable nucleotide (nt37 of the tRNA at the P-site) cannot hydrogen bond to the suppressor tRNA, but it may stabilize the adjacent double helix formed between the codon and the anticodon via base stacking (Carrier and Buckingham, 1984). This extended structure may then interact more efficiently with the incoming suppressor tRNA depending upon its anticodon sequence. In support of this idea is the fact that the identity of nucleotide 37 of the tRNA at the P-site is correlated with the degree of suppression at a given context (Murgola et al, 1984).

How might these considerations help in identifying the sequences involved in the context effect for selenocysteine incorporation? It is first of all necessary to assume that since the selenocysteine codon became fixed early in evolution, the context effect for selenocysteine is the same in E.coli and mice. The two known selenopeptide mRNA sequences can then be compared in the hope that the nucleotides involved in the selenocysteine context effect will be revealed by their commonality between the two mRNAs (Table 4.1). The most obvious similarity between these two sequences is the presence of the dinucleotide UC immediately preceding the UGA codon. However, since this dinucleotide is also present in one of twenty-three rodent sequences which definitely use UGA as a termination codon (β-lutropin, Chin et al, 1983; see Table 4.1), it would seem that UC alone is not

Table 4.1. Comparison of sequences around the selenocysteine codons of fdhF and GSHPx mRNAs with the sequences around UGA termination codons of rodent genes.

A.

fdhF	GCU	CGU	GUC	UGA	CAC	GGC	CCA
	**	* *	**	***	*	*	*
GSHPx	GCG	UCU	CUC	UGA	GGC	ACC	ACG

B.

Mouse genes							
ϵ-Casein	AGC	CAG	G*U	***	GUG	GAA	AAU
IFβ	UUC	CAA	AA*	***	AGA	CCU	GUC
Kallikrein	AAA	AA*	GC*	***	GUG	UCA	CAU
MT I	UGC	UG*	GC*	***	UGU	GA*	GAA
MT II	UGC	UG*	GC*	***	AGG	GGG	G*G
NGF β	AGA	AGA	GG*	***	CUU	GC*	UGC
WAP	A*C	CU*	CAG	***	GCA	GC*	GGC
Rat genes							
At	AAU	GUG	G*G	***	UG*	CU*	CUG
Alp E	AGA	AA*	CAA	***	UCA	UC*	CUC
Cp A	CAC	CCC	UA*	***	AAG	CAG	C*C
Cyt P450-Pb	UCA	GC*	CGG	***	UC*	GG*	UGA
γ-Casein	UAC	AUG	GCU	***	GUA	GAA	AAU
γ-Cryst	*AU	UUC	UAU	***	AAU	AUU	UUU
Igδ	C	UU*	CC*	***	GA*	CUU	U*U
β-Lutropin	CUC	CUC	U**	***	UG*	CCA	C*C
MT-I	UGC	UG*	GC*	***	AGU	GA*	GAA
MBP	G*A	AGA	CG*	***	GAG	CCU	C*C
Insulin I	UAC	UGC	AA*	***	GU*	CA*	CAC
PPR	AAA	CUC	TG*	***	UA*	GGA	G*U
PSBP	*AU	AUG	AA*	***	UUU	UU*	CUG
SVSP IV	*UA	CUC	AA*	***	UAG	UG*	AUC
V-N	CGG	GUC	UA*	***	GC*	AU*	G*C
WAP	U*C	UU*	CAG	***	GAA	GC*	UGC

A. The sequences of fdhF and GSHPx mRNAs around the selenocysteine codons are aligned. Nucleotides common to both sequences are indicated by asterisks.

B. The DNA sequences around the UGA termination codons of rodent genes are presented. Nucleotides common to both fdhF and GSHPx as well as the mRNA indicated are represented by asterisks. Abbreviations: IFβ, β-interferon; MT, metallothionien; NGF, Nerve growth factor; WAP, Whey acidic protein; At, angiotensinogen; Alp, Apolipoprotein; Cp, Carboxypeptidase; Cyt P450-Pb, phenobarbitol induced cytochrome P450; Cryst, Crystallin; MBP, myelin basic protein; PPR, preprorelaxin; PSBP, prostatic steroid binding protein; SVSP IV, seminal vesicle secretory protein IV; V-N, Vassopressin-neurophysin.

sufficient to account for the putative selenocysteine context effect.

Bearing in mind the previously mentioned correlation between the identity of nucleotide 37 of tRNAs binding to codons adjacent to the suppressible codon and the context effect for a particular suppressor tRNA (Bossi, 1983; Murgola, 1984), it may be fruitful to compare the sequences of the tRNAs which recognize the codons 5' and 3' to the UGA in fdhF, GSHPx and β-lutropin mRNAs. However at present this is not possible since the relevant rodent tRNA sequences are not available (Sprinzl et al, 1987). Thus the identity of the nucleotides in the selenopeptide mRNAs responsible for the selenocysteine context effect remain elusive.

Identification of the nucleotides relevant to the selenocysteine context effect could be achieved experimentally, for example, by cloning versions of the DNA surrounding the TGA of the GSHPx cDNA carrying point mutations into an SP6 vector and assaying the corresponding mRNA transcripts for their ability to incorporate radiolabelled selenium into protein. This would not only advance understanding of selenium biochemistry but may also provide some additional information regarding the incompletely understood process of translational termination (Caskey, 1979; Moldave, 1985).

4.1.5. Other eukaryotic suppressor tRNAs and the origin of selenocysteyl-tRNA.

Selenocysteyl-tRNA is by no means the only naturally occuring tRNA which recognizes a stop codon within the same cellular compartment in which it is used to signal translational termination. By analysis of the proteins produced by readthrough of the termination codons of ß-globin (UGA) and α-globin (UAG) mRNAs in rabbit reticulocytes, the existence of tRNAs capable of suppressing termination by each of these codons was inferred (Geller and Rich, 1980). It should be stressed that the degree of suppression afforded at these sites was far from complete, underlining once more the importance of context in codon fuction. One of the tRNAs which suppressed UGA termination was identified as a $tRNA^{Trp}$. However, this observation is of dubious significance since this tRNA may be solely of mitochondrial origin (Barrel *et al*, 1979). Perhaps of greater importance were the readthrough products of α-globin and those of ß-globin produced by insertion of amino acids other than tryptophan. Several different proteins were detected in each case, implying the existence of several different suppressor tRNAs specific for either UGA or UAG. This is consistent with more recent findings which demonstrate the existence of UAG suppressor tyrosinyl-tRNAs in fruit flies and tobacco and of UGA suppressor seryl-tRNAs in cows, chickens, rabbits, frogs and man (Hatfield, 1985). However, it is difficult to see why the cell should have evolved a mechanism for inserting, in response to particular stop codons, amino acids for which there already exist adequate pathways of

insertion. Moreover, no genes have yet been identified which contain an in-frame stop codon in a position colinear with the position of a serine or a tyrosine in the corresponding protein. Since suppressor seryl-tRNAs can be enzymatically converted to phosphoseryl-tRNAs (Hatfield et al, 1982), presumably a similar reaction could occur for tyrosinyl-tRNAs; thus the possibility emerges that all naturally occurring suppressor tRNAs might encode unusual amino acids.

This serine phosphorylation reaction is interesting in relation to the suggestion that selenocysteine is formed by substitution of the phosphate group of phosphoserine by selenite (Hawkes and Tappel, 1983). A similar reaction with phosphoseryl tRNA seems an attractive alternative particularly in view of some older evidence that inorganic selenium is more efficiently incorporated into GSHPx than selenocystine (Sunde and Hoekstra, 1980a). The recent demonstration that serine provides the carbon skeleton of the selenocysteine of GSHPx in rat liver (Sunde and Evanson, 1987) also supports this idea.

An alternative suggestion that the UGA in fdhF mRNA (and therefore presumably also GSHPx mRNA) is recognized by a $tRNA^{Trp}$ after replacement of the indole group of tryptophan by selenide (Zinoni et al, 1986) seems unlikely since such a tRNA has not been found in mammalian cytosol, with the exception of reticulocytes (Geller and Rich, 1980) in which the mitochondria have been lysed.

Resolution of these questions regarding the pathway of formation of the aminocylated form of $tRNA^{Sec}$ requires a more

extensive analysis of $tRNA^{Sec}$, which will only be possible once it has been purified.

4.2. Structure of the GSHPx gene.

4.2.1. GSHPx mRNA is the product of a simple transcription unit.

The data presented concerning the expression and size of the GSHPx mRNA indicate that, within the limits of the experiments, the main mRNA detected in all cells is the same size. Furthermore, since the major sites of transcription initiation and polyadenylation are the same in several tissues, this suggests that the same intron sequences are spliced out in all tissues. Therefore, the GSHPx gene is representative of a simple transcription unit as defined by Nevins (1983) [another example of which would be β-globin]. So, differing levels of GSHPx mRNA in different cell types must result from the varying transcriptional activity of the same promoter or varying post-transcriptional effects upon the same transcript. The GSHPx gene thus differs from some complex transcription units (Leff *et al*, 1986) such as α-amylase (Schibler *et al*, 1983) whose transcripts accumulate to different levels in various cell types by transcription from alternative promoters of differing strengths.

4.2.2. Various GSHPx mRNA 5' ends are represented to differing degrees in RNA from diverse cell types.

Although the different levels of GSHPx mRNA do not result from a switch in major promoter usage in different cell types, a variety of minor GSHPx transcripts originating upstream of the major GSHPx gene promoter are represented to differing extents in RNA from different cell types. Thus, by using a combination of S1 nuclease protection assays and primer extension, transcription initiation sites were mapped 25, 40 and 50nt upstream from the major cap site, as well as very minor sites at -100, -140, -190 and -220nt (Frampton et al, 1987). The data in Figure 3.II.4 corroborate and extend these findings ; one of the major points to emerge from the combined analyses is that the ratio of upstream transcripts to major cap transcripts is relatively high in cells expressing the GSHPx mRNA at low levels and, conversely, low in cells expressing the GSHPx mRNA at high levels.

A similar heterogeneity in the 5' ends of transcripts from several other genes has been noted (Contreras et al, 1982; Osborne and Berk, 1983; Allan et al, 1983; Carlson and Ross, 1983; 1984; Reynolds et al, 1984; Hess et al, 1985; Masters and Attardi, 1985; Chang et al, 1986; Mitchell et al, 1986). Particularly interesting among these is the human ϵ-globin gene, which is normally only expressed in embryonic erythroid cells and has 5' termini spread in discrete clusters over a 4.5kb region extending upstream from the major cap site (Allan et al, 1983). In non-erythroid cell lines a low level of ϵ-globin RNA is detectable and in these cases the 5' termini are located

predominantly at a site 200nt upstream of the major cap site with none arising from the major cap site. When plasmids containing the ϵ-globin gene are introduced into non-erythroid cells, transcripts are likewise initiated predominantly from the -200nt site. Co-transfection of a plasmid containing the adenovirus E1A gene together with the ϵ-globin plasmid does not affect the total level of ϵ-globin transcripts but causes a re-direction of transcription initiation from the -200nt site to the major cap site. Additional experiments showed that, whereas the -200 cap site was unresponsive to the effects of an SV40 enhancer linked in *cis* to the ϵ-globin gene, the level of transcripts originating from the major cap site was increased 200-fold (Allan *et al*, 1984). It is therefore not difficult to imagine how similar mechanisms could give rise to the differences in GSHPx transcript profile in cells expressing high and low levels of GSHPx mRNA.

4.2.3. GSHPx gene structure vis-a-vis protein structure : an evolutionary speculation.

The structure of the GSHPx protein has been described in detail (Epp *et al*, 1983). The enzyme is a homotetramer with its active sites being present on the surface of the molecule as flat depressions. The selenocysteine residue forms part of the active site and is located at the N-terminal end of the first α-helix in the primary structure, which is sandwiched between two parallel β-strands (Figure 3.I.3). Such $\beta\alpha\beta$ structures favour the occurrence of a binding site at the C-terminal end of the

parallel ß-strands (Branden, 1980) which is where the GSHPx active site in fact lies (Epp et al, 1983). Moreover, the alignment of peptide dipoles parallel to the helical axis gives rise to a macrodipole, whose magnitude at the helical N-terminus is approximately half a positive unit charge (Hol et al, 1978; see also Hol, 1985 for a review). This dipole should stabilize the selenocysteine selenolate, increasing its nucleophilicity.

When considered in relation to the protein structure, the position of the intron within the GSHPx gene is intriguing since it splits the $\beta_1\alpha_1\beta_2$ motif from the rest of the protein (Figure 3.I.3). Since all the characterized selenoproteins which contain selenocysteine are redox enzymes, the supersecondary structural element present in exon 1 (ßSecαß) could be a modular component in the assembly of potent redox enzymes which has spread as such during evolution by exon shuffling (Gilbert, 1978; Doolittle, 1978; Blake, 1978). The sequences of other eukaryotic selenoproteins and their corresponding genes are required before the validity of this hypothesis can be assessed.

4.3. Expression of GSHPx mRNA.

4.3.1. Tissue distribution of GSHPx.

The GSHPx tissue distribution has been assayed enzymatically in humans (Carmagnol et al, 1983), rats (Hawkes et al, 1985) and in a variety of species including mice (Tappel et al, 1982). These studies show that although the activity of GSHPx in a given tissue varies between species, the highest level is always in liver with kidney or RBCs having the next highest level; other

tissues have a considerably lower GSHPx activity. The fact that the GSHPx activity in mouse liver is approximately three fold higher than in kidney seems at odds with the tissue distribution of GSHPx mRNA, which shows the mRNA level to be three fold higher in kidney than in liver (Figure 3.II.1). This may indicate that GSHPx mRNA is differentially controlled at the translational level in these two tissues. On the other hand, it may reflect a more serious problem associated with determinations of GSHPx levels.

For example, a high GSHPx activity, determined enzymatically, has been reported in human platelets (Guidi et al, 1984; Ramos-Martinez, 1979) where the enzyme is thought to have a role in arachidonic acid metabolism (Needleman et al, 1986; Lands et al, 1984). However, a high GSHPx activity was not seen in rat platelets either enzymatically (Levander et al, 1983) or by radioimmunoassay (Baret et al, 1983). More importantly, the latter study also failed to detect a high GSHPx protein level in human platelets. These discrepancies between human platelet GSHPx levels determined enzymatically (Guidi et al, 1984; Ramos-Martinez, 1979) and by radioimmunoassay (Baret et al, 1983) could be due to at least one other enzyme, peroxidation-inhibiting protein (PIP), which exhibits glutathione peroxidase activity towards hydroperoxides within membranes (Ursini et al, 1982; Fujii et al, 1984), something of which GSHPx is incapable (Grossman and Wendel, 1983; Sevanian et al, 1983; Yasuda and Fujita, 1977). In addition, the fact that the PIP activity is higher in porcine liver than kidney (Ursini et al,

1983) may at least partly explain the differences between the GSHPx enzyme activities in murine liver and kidney and the GSHPx mRNA levels in these tissues. It may be necessary to perform radioimmunoassays to determine the true levels of GSHPx protein in different tissues.

Some immunological studies on the distribution of GSHPx protein within tissues have already been conducted. Immunohistochemical analysis of rat liver (Yoshimura et al, 1980), kidney (Mizuiri et al, 1986) and adrenals (Murakoshi et al, 1981) has shown each of these tissues to be composed of cells that contain widely differing amounts of GSHPx protein. For example, in liver, GSHPx protein was only detectable in hepatocytes. However, those around the portal areas (where blood carrying newly absorbed nutrients from the gut first penetrates the hepatic lobule) had a higher GSHPx level than those around the central vein (Yoshimura et al, 1980). Analysis of the corresponding mouse tissues via in situ hybridisation of RNA to GSHPx probes should reveal whether these differences in protein level are a reflection of varying steady state levels of GSHPx mRNA within the same cells.

4.3.2. Alterations in GSHPx levels during haemopoiesis.

The changes in GSHPx mRNA level during haemopoiesis are intriguing. Although the pregranulocytic cell line AD-3 has a low level of GSHPx mRNA, rat granulocytes and lymphocytes have been shown by radioimmunoassay to have a high GSHPx protein level comparable to RBCs (Baret et al, 1983). It would be interesting,

therefore, to discover whether the GSHPx mRNA level increases during murine granulo- or lymphopoiesis and if so whether the mechanism(s) involved are the same as those operating during erythropoiesis. If the mechanism is the same then it would be useful to investigate whether the biochemical machinery responsible was present prior to commitment of haemopoietic progenitor cells and how this relates to proposed schemes of haemopoiesis involving sequential restriction in lineage potentials (Nicola and Johnson, 1982). However, in murine cell lines derived from a T-lymphoma (L5178Y) and a B-myeloma (P1BU1), the GSHPx mRNA level is not elevated (Affara et al, 1985; and further unpublished data) although it has been shown to increase when a human myeloid cell line (HL-60; Tsiftsoglou and Robinson, 1985) is induced to differentiate into granulocytes but not into macrophages (I.C. and S. Graham, unpublished observation).

Since the GSHPx mRNA level increases following the commitment and maturation of a haemopoietic stem cell down the erythroid lineage, it seems highly probable that the GSHPx protein level also increases during erythropoiesis. However, it is unclear when this rise might occur. It may parallel the rise in GSHPx mRNA or it could lag behind by several cell divisions, as has been demonstrated for the erythroid lipoxygenase (Thiele et al, 1979; 1982). In this context it is noteworthy that a 12-14S mRNA exists in the untranslated ribonucleoprotein particles (RNPs) in duck erythroblasts (Vincent et al, 1977; 1983). Since this mRNA is translated in vitro to give a 22kD polypeptide, it is not inconceivable that it is duck GSHPx mRNA.

If indeed there is this behavioural similarity between GSHPx mRNA and erythroid lipoxygenase mRNA, then this could be reflected in the possession by both mRNAs of a similar sequence in their untranslated regions responsible for their initial targetting to the untranslated RNP fraction.

4.3.3. Possible induction of GSHPx by peroxides and selenium.

Rats exposed to low levels of ozone (0.2 ppm) for eight days have an increased GSHPx activity in their lungs (Chow et al, 1974). Moreover, patients with glucose-6-phosphate dehydrogenase deficiency (who consequently have a reduced capacity to regenerate GSH-see Figure 1.3) also have an increased RBC GSHPx activity (Beutler, 1977) as do Down's syndrome patients (who have an additional Cu/Zn superoxide dismutase gene) (Sinet et al, 1975; Frischer et al, 1981). In none of these cases is it clear at what level this induction occurs. However, an increased GSHPx activity was seen following the generation of activated oxygen in human RBCs by acetylphenyl hydrazine treatment (Perona et al, 1978). This was taken as evidence of allosteric activation of the enzyme, but this does not mean that induction cannot also occur at other levels. It may therefore be fruitful to determine whether the GSHPx mRNA level in cells expressing high or low levels of GSHPx can be regulated by peroxide treatment.

In this respect it would be interesting to know whether the higher level of GSHPx protein seen in hepatocytes around the portal areas (Yoshimura et al, 1980) was constitutive or was due to induction by hydroperoxides in the portal blood. Perhaps the

differential staining of hepatocytes with anti-GSHPx antibodies could be abolished by prior extensive liver perfusion with peroxide free medium. The related question of whether the hepatocytes around the central vein are inducible for GSHPx protein by peroxide could be addressed by liver perfusion in the reverse direction with peroxide supplemented medium.

Interestingly, the enteric bacteria E.coli and S.typhimurium both contain an enzyme activity analogous to GSHPx (alkyl hydroperoxide reductase) which is inducible by cumene hydroperoxide. The corresponding gene is part of a larger network of genes whose products are inducible by hydrogen peroxide, via the action of the positively acting product of the oxyR gene (Christman et al, 1985; Morgan et al, 1986). Points of note here are that: (1) cadmium induces most of the oxyR stress proteins (Morgan et al, 1986); selenium has been shown to protect mammals from cadmium toxicity (Neve et al, 1985); (2) bacterial alkyl hydroperoxide reductase is composed of two subunits, one is a 52kD FAD containing polypeptide, the other is a polypeptide of similar size (22kD) to GSHPx, which has not yet been further characterized.

If GSHPx is indeed inducible by peroxide then it should be possible to obtain regulatory mutants which constitutively express the peroxide resistance by selecting for cells which can grow in concentrations of peroxide which kill normal cells. GSHPx activity has already been shown to correlate with tolerance of bovine adrenocortical cells to hydroperoxides when the cells are cultured in varying states of selenium deficiency (Hornbsy et

al, 1985). (A useful practical observation to emerge from this work is that some batches of commercially available serum contain little or no bioavailable selenium.) It would obviously be interesting to identify which other enzyme activities were elevated in such potential regulatory mutants. Another question which may be of interest concerns whether such potential regulatory mutants had an altered susceptibility to tumour promoters, since increased intracellular levels of active oxygen species are thought to be involved in tumour promotion (Cerutti, 1985).

This touches the broader issue of whether GSHPx is involved in mediating the anticarcinogenic effects of selenium (Ames, 1983). It is known that treatment of cells with selenium prior to insult with tumour promoters or radiation increases the activity of GSHPx and decreases the tumour promoting activity of these agents (Perchellet et al, 1986; Borek et al, 1986). However, other work suggests that the anticarcinogenic effects of selenium are not mediated by GSHPx (Ip, 1985) since the in vivo levels of selenium but not GSHPx in certain tissues are inversely correlated with tumour yield in some model systems. Therefore a causal relationship between GSHPx activity and selenium anticarcinogenesis remains to be established. This question could perhaps be approached using controlled diets fed to mice transgenic for an antisense GSHPx gene under control of an inducible promoter.

4.4. Control of GSHPx gene expression.

GSHPx mRNA is present in all cell types examined to date, generally at a low basal level, but in some cell types this level is elevated 20-50 fold. In this respect, the GSHPx gene resembles certain other "housekeeping" genes whose levels of expression differ in different cell types e.g. chicken 5-aminolevulinate synthase (Maguire et al, 1986), glyceraldehyde-3-phosphate dehydrogenase (Stone et al, 1985), adenosine deaminase (Valerio et al, 1985), Cu/Zn superoxide dismutase (Sherman et al, 1983). The present results (section 3.2) show that the different cytoplasmic GSHPx mRNA levels do not appear to be controlled at the level of nucleo-cytoplasmic transport. Neither is there any apparent difference, at least between neuroblastoma and Friend cells, in the ratio of GSHPx pre-mRNA to mature mRNA. Presumably, therefore the differences in GSHPx mRNA level in various cell types reflect either transcriptional or mRNA stability differences between the cells. At present it is not possible to distinguish between these alternatives definitively. However, indirect support for a mechanism involving alternative transcriptional states of the GSHPx gene is provided by the different chromatin structures of the GSHPx gene in cells expressing high or low levels of GSHPx mRNA. One aspect of the GSHPx gene chromatin structure which does not appear to change in cells expressing high or low levels of GSHPx mRNA is the DH site (I), present at the 5' end of the gene.

4.4.1. DNA sequences 5' to the GSHPx gene.

The DNA underlying this apparently invariant DH site contains a G-C rich (61%) segment extending from -260 to +1nt, within which there are three Hpa II sites. These may be examples of the Hpa II tiny fragments suggested by Bird et al (1985) to be unmethylated in all cells and to be located at the 5' ends of ubiquitously expressed genes. Indeed the Hpa II sites at the 5' end of the GSHPx gene which have been examined are unmethylated in Friend cells (Figure 3.I.6) although this has not been checked in other cell types. This region of the GSHPx gene also resembles the promoters of "housekeeping" genes inasmuch as it possesses five sequences which show at least 8/10nt match to the Sp1-binding decanucleotide consensus sequence (Kadonaga et al, 1986). Although only one of these possesses the exact hexanucleotide core sequence, they may all bind Sp1 : for instance the human immunodeficiency virus long terminal repeat promoter contains three Sp1 binding sites, two of which do not conform to the core sequence (Jones et al, 1986). Clearly, this issue can only be resolved by binding of Sp1 to this region. However, in this context it should be noted that Sp1 has thus far not been isolated from mouse cells, although it is present in hamster, as well as human cells (Briggs et al, 1986). Therefore, if these sequences do bind Sp1 it will then be necessary to determine whether such binding stimulates transcription of the GSHPx gene in vitro and whether point mutations within the putative Sp1 binding sites have a negative effect upon transcription of a linked gene following transfection into murine

cells. These same considerations apply to the potential Sp1 binding sites located within the gene and at its 3' end.

4.4.2. Comparison of the SV40 early and GSHPx promoters.

That the region upstream from the major transcription initiation site in the GSHPx gene does in fact function as a promoter of transcription has been shown by introduction of a CAT gene under control of the putative GSHPx promoter into murine cell lines. The fact that the level of transcription directed by the GSHPx promoter is similar to that directed by the SV40 early promoter is interesting. Although these two promoters differ in the A-T richness of sequences surrounding their "ATA" sequences and in the distance between the potential/actual Sp1 binding sites and the "ATA" sequences, they do have common features. For instance, neither promoter contains a sequence exactly homologous to the consensus "ATA" sequence (Corden et al, 1980), although both contain sequences closely related to this consensus located approximately 30nt upstream from the major transcription initiation site. Moreover, the sequences of both genes downstream from the "ATA" sequence show a similar degree of homology (approximately 70%) to the consensus sequence for "housekeeping" gene promoters (Table 3.I.2). Perhaps the most interesting similarity between the two genes is the different transcription initiation sites used by both promoters under different physiological conditions. Early after SV40 infection, transcription initiates from a cluster of sites located 20-30nt downstream for the "ATA" box, referred to as the early-early

start sites (Wasylyk et al, 1983b). Later in infection T-antigen accumulates and by binding to DNA surrounding the early-early start sites, switches transcription to a cluster of sites 30-40nt further upstream referred to as the late-early start sites (Hansen et al, 1981; Buchman et al, 1984).

The Sp1 binding sites in the SV40 early promoter are required for efficient transcription from the early-early sites (Fromm and Berg, 1982; Hansen and Sharp, 1983; Vigneron et al, 1984, Baty et al, 1984). However, the accuracy of transcription initiation at these sites is governed by the "ATA" sequence; a double point mutation within it causes a large decrease in early-early start site transcription (Wasylyk et al, 1983b). This is accompanied by a large increase in the transcription from the late-early and other minor start sites, suggesting that late in infection the switch to late-early start sites is mediated by a lack of binding of an "ATA" box factor required for initiation of transcription from the early-early start sites. It is tempting to speculate that in cells expressing high levels of GSHPx mRNA, where transcription is predominantly from the major transcription initiation site, a productive interaction between an "ATA" box factor and the "ATA" box is allowed; conversely such an interaction might be precluded in cells expressing a low level of GSHPx mRNA. It would be interesting to see what effect mutation of the GSHPx "ATA" sequence had upon the transcription profile in cells expressing a high GSHPx mRNA level. It may also be fruitful to investigate the effect of deleting the G-C rich region upon transcription of the GSHPx gene.

4.4.3. Sequence homologies at the 3' end of the GSHPx gene and their relevance to gene regulation.

Analysis of the GSHPx gene for other sequence homologies focused attention upon a region of 250nt at the 3' end of the gene flanked by two inverted putative Sp1 binding sites and containing centred within it two inverted copies of the SV40 core enhancer sequence as well as sequences homologous to the Ig octamer sequence and the NF-kB binding site (section 3.1.4). However, constructs designed to test the transcriptional influence of the DNA fragment containing these homologies upon transcription from homologous or heterologous promoters linked in *cis*, failed to reveal any function for this DNA in either Friend cells or fibroblasts. However, this does not mean that this DNA (or some part of it) might not be part of a larger domain which can enhance transcription in either of these cell types, particularly in view of its proximity to sequences forming the putatively erythroid-specific DH sites. It is notable in this context that the SV40 enhancer has been shown to consist of three separate domains each of which alone exhibits no enhancer function; however different domains in combination or single domains in tandem duplication do act as enhancers (Zenke *et al*, 1986; Herr and Clarke, 1986). Perhaps duplication of the 706bp fragment from the 3' end of the GSHPx gene might cause the appearance of enhancer activity.

It is also worth stressing that sequences within the GSHPx 706bp fragment might act as an enhancer in different cell types or under different environmental conditions. For instance, as

discussed previously (section 4.3.2) lymphocytes have been reported to contain a high GSHPx protein level (Baret *et al*, 1983). This is intriguing in view of the fact that the 706bp fragment of the GSHPx gene contains an Ig octamer homology (Parslow *et al*, 1984) and an NF-kB homology (Sen and Baltimore, 1986a) separated by 20bp within which is located a palindromic sequence of the CA/TG type which resembles palindromes which bind B-cell proteins *in vivo* (Church *et al*, 1985; Ephrussi *et al*, 1985) and *in vitro* (Sen and Baltimore , 1986a). Assuming that these sequences are still homologous enough to the consensus to allow binding by the corresponding B-cell *trans*-factors, then this region of the GSHPx gene might be expected to function as an enhancer in B-cells. However, preliminary attempts to introduce this DNA into MPC11 cells (a mouse B-myeloma cell line) were unsuccessful. Moreover, the level of GSHPx mRNA in P1BU1 cells (a B-myeloma cell line) was found to be low, comparable to that in neuroblastoma cells (Affara, unpublished data).

An alternative speculation is that the GSHPx gene 3' fragment acts as an enhancer in response to elevated levels of peroxide in the environment. The fact that NF-kB can be activated in non B-cells post-translationally may be relevant in this respect (Sen and Baltimore, 1986b).

4.4.4 Chromatin structure of the GSHPx gene : implications for gene expression.

The DH site at the 5' end of the GSHPx gene is broad (200-300bp) in all cells examined : no apparent differences in the structure of the DH site in different cell types was revealed by close examination on long gels. Similar broad DH sites have been observed at the 5' ends of the HSV-1 thymidine kinase gene (Sweet et al, 1982), the SV40 genes (Saragosti et al, 1982) and the rat preproinsulin II gene (Wu and Gilbert, 1981) : these appear to reflect the binding of several different proteins (McKnight and Tjian, 1986; Eissenberg et al, 1985). It is therefore possible that several proteins bind to the GSHPx promoter and that these are the same in different cell types.

The presence of a DH site (IV), approximately 3.5kb downstream from the cap site has been detected in all cells expressing a high level of GSHPx mRNA. However, these cells do not all have the same GSHPx chromatin structure, additional DH sites being present in Friend cells compared to liver and kidney. This suggests that the GSHPx gene is regulated differently in these three cell types/tissues with the downstream DH sites reflecting the activity of regulatory elements in erythroid cells alone (DH II and III) or in liver, kidney and erythroid cells (DH IV). The latter regulatory element may also be active to differing extents in each of these cell types/tissues. Alternatively the analysis may have overlooked some difference in the chromatin structure between these cell types/tissues outwith the DNA already examined (5kb upstream and 10kb downstream from

the GSHPx gene cap site). However, as noted earlier, the possiblity of post-transcriptional controls being involved in some of these cases cannot at present be excluded.

It is interesting that the uninduced Friend cell DH sites exist as two doublets downstream of the GSHPx gene. By analogy with other similar doublet DH sites (Costlow et al, 1985; Wu, 1984; Plumb et al, 1986) this suggests that they are binding sites for trans-factors, although the extent of the "cold spot" between the doublet bands in each case (approximately 40-50bp) seems too large to be explained by the binding of a single protein.

Two of the Friend cell DH sites (II and IV) disappear following HMBA induced terminal maturation. It would be interesting to know at what stage during Friend cell differentiation these DH sites disappear, whether they disappear together or sequentially and whether their disappearance occurs during the commitment process or during the subsequent maturation. This question could be addressed following induction of differentiation in the presence of inhibitors of commitment or maturation such as cordycepin or dexamethasone (see section 1.2.2 for references). Possibly relevant here is the situation in lymphoma x Friend cell hybrids which can be induced to accumulate globin chains by DMSO but do not differentiate terminally, unless haemin is also added (Harrison et al, 1977). The results presented by Affara et al (1985) seem to show that these hybrid cells, treated with DMSO, retain the GSHPx chromatin structure characteristic of uninduced Friend cells. It may therefore be

informative to determine whether loss of these DH sites occurs when the DMSO-treated hybrids are allowed to differentiate by the additon of haemin. This might allow further dissection of the molecular events attending the alterations in GSHPx chromatin structure during differentiation.

A related question concerns the GSHPx gene chromatin structure in haemopoietic stem cells and how this changes during erythroid versus non-erythroid differentiation. It may be possible to address this question using cloned haemopoietic stem cell lines, such as the DeCl.15 cells (section 1.1; Spooncer *et al*, 1984; 1986) if these can be made to differentiate selectively into different myeloid lineages under different *in vitro* conditions.

Recent experiments have provided evidence for the existence of *trans*-dominant repressor molecules which can interact with in the tyrosine amino transferase (Killary and Fournier, 1984) and albumin genes (Petit *et al*, 1986) causing their repression in non-hepatic cells. While characterisation of the controls operating upon the GSHPx gene is not as advanced as either of these examples, evidence does exist for *trans*-dominant molecules present in neuroblastoma cells which are able to impose a neuroblastoma-like chromatin structure upon the GSHPx genes in the Friend cell derived chromosomes of neuroblastoma-Friend cell hybrids and concomitantly cause a reduction in the steady state GSHPx mRNA level in such hybrids to a level comparable to that of the neuroblastoma parent (Affara *et al*, 1985). Since exposure of nuclei from one cell population to the cytoplasm from another

cell population confers upon such nuclei a DNAse I sensitivity more characteristic of the nuclei from the cells from which the cytoplasm was derived (Prentice et al, 1983) it may be worth investigating the ability of neuroblastoma cytoplasmic factors to alter the GSHPx chromatin structure of Friend cell chromosomes when mixed with Friend cell nuclei. If this were successful then it would be possible to isolate the responsible factors by fractionation of neuroblastoma cytoplasm. Alternatively the trans dominant repressor(s) from neuroblastoma cells could be isolated by the methods alluded to above (Killary and Fournier, 1984; Petit et al, 1986).

Finally, it is to be noted that the two doublet DH sites detected in Friend cells lie outwith the cloned DNA and could not therefore be tested for their transcriptional influence upon a promoter linked in cis by transfection studies. These experiments must await the cloning of the relevant DNA, which, hopefully, should shortly be in hand using the strategy outlined in section 3.4.5.

Chapter 5

References

Affara, N., Goldfarb, P.S., Yang, Q.-S. & Harrison, P.R. (1983) Patterns of expression of erythroblast non-globin mRNAs. Nucl. Acids Res. 11, 931-945.

Affara, N., Fleming, J., Goldfarb, P.S., Black, E., Thiele, B. & Harrison P.R. (1985) Analysis of chromatin changes associated with the expression of globin and non-globin gnees in cell hybrids between erythroid and other cells. Nucl. Acids Res. 13, 5629-5644.

Allan, M., Lanyon, G. W. & Paul, J. (1983) Multiple origins of transcription in the 4.5kb upstream of the ϵ-globin gene. Cell 35, 187-197.

Allan, M., Zhu, J.-D., Montague, P. & Paul, J. (1984) Differential response of multiple ϵ-globin cap sites to *cis* and *trans*-acting controls. Cell 38, 399-407.

Allen, T.D. & Dexter, T.M. (1983) Long term bone marrow cultures : an ultrastructural review. Scan. Elect. Mic. 4, 1851-1866.

Ames, B.N. (1983) Dietary carcinogens and anticarcinogens Science 221, 1256-1264.

Ayer, D. & Yarus, M. (1986) The context effect does not require a fourth base pair. Science 231, 393-395.

Baret, A., Courtiere, A., Rouquette, A. & Michelson, A.M. (1983) Determination of glutathione peroxidase in circulating blood cells : use of a radioimmunoassay. Mol. Physiol. 4, 313-322.

Barrel, B.G., Bankier, A.T. & Drouin, J. (1979) A different genetic code in human mitochondria. Nature 282, 189-194.

Baty, D., Barrera-Saldana, H.A., Everett, R.D., Vigneron, M. & Chambon, P. (1984) Mutational dissection of the 21bp repeat region of the SV40 early promoter reveals that it contains overlapping elements of the early-early and late-early promoters. Nucl. Acids Res. 12, 915-932.

Becker, P., Renkawitz, R. & Schutz, G.(1984) Tissue-specific DNAse I hypersensitive sites in the 5'- flanking sequences of the tryptophan oxygenase and the tyrosine aminotransferase genes. EMBO J. 3, 2015-2020.

Becker, P., Gloss, B., Schmid, W., Strahle, U. & Schutz, G. (1986) *In vivo* protein-DNA interactions in a glucocorticoid response element require the presence of the hormone. Nature 324, 686-688.

Behne, D., Hofer, T., von Berswordt-Wallrabe, R. & Elger, W. (1982) Selenium in the testis of the rat : studies on its regulation and its importance for the organism. J. Nutr. 112, 1682-1687.

Behne, D. & Hofer-Bosse, T. (1984) Effects of a low selenium status on the distribution and retention of selenium in the rat. J. Nutr. 114, 1289-1296.

Benezra, R., Cantor, C.R. & Axel, R. (1986) Nucleosomes are phased along the mouse β-major globin gene in erythroid and non-erythroid cells. Cell 44, 697-704.

Bennet, K.L., Hill, R.E., Pietras, D.F., Woodworth-Gutai, M., Kane-Haas, C., Houston, J.M., Heath, J.K. & Hastie, N.D. (1984) Most highly repeated dispersed DNA families in the mouse genome. Mol. Cell. Biol. 4, 1561-1571.

Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980) The ovalbumin gene-sequence of putative control regions. Nucl. Acids Res. 8, 127-142.

Bentley, D.L. & Groudine, M. (1986) A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. Nature 321, 702-706.

Berg, P.E., Popovic, Z. & Anderson, W.F. (1984) Promoter dependance of enhancer function. Mol. Cell. Biol. 4, 1664-1668.

Bestor, T.H., Hellewell, S.B. & Ingram, V.M. (1984) Differentiation of two mouse cell lines is associated with hypomethylation of their genomes. Mol. Cell. Biol. 4, 1800-1806.

Beutler, E. (1977) Glucose-6-phosphate deficiency and red cell glutathione peroxidase. Blood 49, 467-469.

Biggen, M.D., Gibson, T.J. & Hong, G.F. (1983) Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination Proc. Natl. Acad. Sci. U.S.A. 80, 3963-3965.

Bird, A.P. (1984) DNA methylation - how important in gene control? Nature 307, 503-504.

Bird, A., Taggart, M., Frommer, M., Miller, O.J. & MacLeod, D. (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell 40, 91-99.

Birnboim, H.C. & Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7, 1513-1523.

Birnstiel, M.L., Busslinger, M. & Strub, K. (1985) Transcription termination and 3' processing : the end is in site! Cell 41, 349-359.

Bishop, J.M. (1986) Oncogenes as hormone receptors. Nature 321, 112-113.

Blake, C.C.F. (1978) Do genes-in-pieces imply proteins-in-pieces? Nature 273, 267.

Boettiger, D., Anderson, S. and Dexter, T.M. (1984) Effect of *src* infection in long-term marrow cultures: Increased self renewal of haemopoietic progenitor cells without leukaemia. Cell 36, 763-773.

Bohmann, D., Keller, W., Dale, T., Scholer, H.R., Tebb, G. & Mattaj, I.W. (1987) A transcription factor which binds to the enhancers of SV40, immunoglobulin heavy chain and U2 snRNA genes. Nature 325, 268-272.

Borek, C., Ong, A., Mason, H., Donahue, L. & Biaglow, J.E. (1986) Selenium and vitamin E inhibit radiogenic and chemically induced transformation *in vitro* via different mechanisms. Proc. Natl. Acad. Sci. U.S.A. 83, 1490-1494.

Bossi, L. (1983) Context effects : translation of UAG codon by suppressor tRNA is affected by the sequence following UAG in the message. J. Mol. Biol. 164, 73-87.

Bossi, L. & Roth, J.R. (1980) The influence of codon context on genetic code translation. Nature 286, 123-127.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72 248-254.

Branden, C.-I. (1980) Relation between structure and function of α/β-proteins. Quart. Rev. Biophys. 13, 317-338.

Brawerman, G. (1981) The role of the poly (A) sequence in mammalian messenger RNA. CRC Crit. Rev. Biochem. 10, 1-38.

Breathnach, R. & Chambon, P. (1981) Organization and expression of eukaryotic split genes coding for proteins. Ann. Rev. Biochem. 50 349-383.

Bridges, K., Levenson, R., Housman, D. & Cantley, L. (1981) Calcium regulates the commitment of murine erythroleukaemia cells to terminal erythroid differentiation. J. Cell Biol. 90, 542-544.

Briggs, M.R., Kadonaga, J.T., Bell, S.P. & Tjian, R. (1986) Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science 234, 47-52.

Brown, D.D. (1984) The role of stable complexes that repress and activate eukaryotic genes. Cell 37, 359-365.

Bryan, P.N. & Folk, W.R. (1986) Enhancer sequences responsible for DNAse I hypersensitivity in polyomavirus chromatin. Mol. Cell. Biol. 6, 2249-2252.

Buchman, A.R., Fromm, M. & Berg, P. (1984). Complex regulation of simian virus 40 early-region transcription from different overlapping promoters. Mol. Cell. Biol. 4, 1900-1914.

Burch, J.B.E. & Evans, M.I. (1986) Chromatin structural transitions and the phenomenon of vitellogenin gene memory in chickens. Mol. Cell. Biol. 6, 1886-1893.

Busslinger, M., de Boer, E., Wright, S., Grosveld, F.G. & Flavell, R.A. (1983) The sequence GGC^{m}CGG is resistant to Msp I cleavage. Nucl. Acids Res. 11, 3559-3569.

Carlson, D.P. & Ross, J. (1983) Human β-globin promoter and coding sequences transcribed by RNA polymerase III. Cell 34, 857-864.

Carlson, D.P. & Ross, J. (1984) α-Amanitin-insensitive transcription of mouse β^{maj}-globin 5'-flanking and structural gene sequences correlates with mRNA expression. Proc. Natl. Acad. Sci. U.S.A. 81, 7782-7786.

Carmagnol, F., Sinet, P.M. & Jerome, H. (1983) Selenium dependent and non-selenium dependent glutathione peroxidases in human tissue extracts. Biochim. Biophys. Acta. 759, 49-57.

Carrier, M.J. & Buckingham, R.H. (1984) An effect of codon context on the mistranslation of UGU codons in vitro. J. Mol. Biol. 175, 29-38.

Carthew, R.W., Chodosh, L.A. & Sharp, P.A. (1985) An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major later promoter. Cell, 43, 439-448.

Caskey, C. Th. (1979) Peptide chain termination. T.I.B.S. 4, 234-237.

Cassio, D. & Weiss M.C. (1979) Expression of foetal and neonatal hepatic functions by mouse hepatoma-rat hepatoma hybrids. Somatic Cell Genetics 5, 719-738.

Cerutti, P.A. (1985) Prooxidant states and tumour promotion. Science 227, 375-381.

Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W. & Harrison, P.R. (1986) The structure of the mouse glutathione peroxidase gene : the selenocysteine in the active site is encoded by the termination codon, TGA. EMBO J. 5, 1221-1227.

Chang, D.D., Hixson, J.E. & Clayton, D.A. (1986) Minor transcription initiation events indicate that both human mitochondrial promoters function bidirectionally. Mol. Cell. Biol. 6, 294-301.

Charnay, P., Mellon, P. & Maniatis, T. (1985) Linker scanning mutagenesis of the 5'-flanking region of the mouse ß-major-globin gene : sequence requirements for transcription in erythroid and non-erythroid cells. Mol. Cell. Biol. 5, 1498-1511.

Charron, J. & Drouin, J. (1986) Glucocorticoid inhibition of transcription from episomal proopiomelanocortin gene promoter. Proc. Natl. Acad. Sci. U.S.A. 83, 8903-8907.

Chaudiere, J., Wilhelmsen, E.C. & Tappel, A.L. (1984) Mechanism of selenium-glutathione peroxidase and its inhibition by mercaptocarboxylic acids and other mercaptans. J. Biol. Chem. 259, 1043-1050.

Chen, M.J., Shimada, T., Moulton, A.D., Cline, A., Humphries, K., Maizel, J. & Nienhuis, A.W. (1984) The functional human dihydrofolate reductase gene. J. Biol. Chem. 259, 3933-3943.

Chen, T.R. (1977) In situ detection of mycoplasma contamination in cell cultures by fluorescent Hoechst 33258 stain. Exp. Cell Res. 104, 255-262.

Chin, A.C. & Fournier, R.E.K. (1987) A genetic analysis of extinction : trans-regulation of 16 liver-specific genes in hepatoma-fibroblast hybrid cells. Proc. Natl. Acad. Sci. U.S.A. 84, 1614-1618.

Chin, W.W., Godine, J.E., Klein, D.R., Chang, A.S., Tan, L.K. & Habener, J.F. (1983) Nucleotide sequence of the cDNA encoding the precursor of the ß subunit of rat lutropin. Proc. Natl. Acad. Sci. U.S.A. 80, 4649-4653.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. & Rutter, W. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochem. 18, 5294-5299.

Chiu, D., Lubin, B. & Shohet, S.B. (1982) Peroxidative reactions in red cell biology. In "Free radicals in biology, Vol. V". Pryor, W.A. (ed) N.Y. Academic Press pp.115-160.

Chiu, C.-P. & Blau, H.M. (1985) 5-azacytidine permits gene activation in a previously non-inducible cell type. Cell 40, 417-424.

Chow, C.K., Dillard, C.J. & Tappel, A.L. (1974) Glutathione peroxidase system and lysozyme in rats exposed to ozone or nitrogen dioxide. Environ. Res. 7, 311-319.

Christman, M.F., Morgan, R.W., Jacobson, F.S. & Ames, B.N. (1985) Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in Salmonella typhimurium. Cell 41, 753-762.

Church, G.M., Ephrussi, A., Gilbert, W. & Tonegawa, S. (1985) Cell-type-specific contacts to immunoglobulin enhancers in nuclei. Nature 313, 798-801.

Clayton, D.F., Weiss, M. & Darnell, J.E. (1985) Liver-specific RNA metabolism in hepatoma cells : variations in transcription rates and mRNA levels. Mol. Cell. Biol. 5, 2633-2641.

Coggins, L.W., Vass, J.K., Stinson, M.A., Lanyon W.G. & Paul, J. (1982) A B1 repetitive sequence near the mouse β-major globin gene. Gene 17, 113-116.

Cohen, D.R., Hapel, A.J. & Young, I.G. (1986) Cloning and expression of the rat interleukin-3 gene. Nucl. Acids Res. 14, 3641-3658.

Cohen, H.J., Chovaniec, M.E., Mistretta, D. & Baker, S.S. (1985b) Selenium repletion and glutathione peroxidase differential effects on plasma and red blood cell enzyme activity. Am. J. Clin. Nutr. 41, 735-747.

Cohen, R.B. & Sheffery, M. (1985a) Nucleosome disruption precedes transcription and is largely limited to the transcribed domain of globin genes in murine erythroleukaemia cells. J. Mol. Biol. 182, 109-129.

Colantuoni, V., Pirozzi, A., Blance, C. & Cortese, R. (1987) Negative control of liver-specific gene expression : cloned human retinol-binding protein gene is repressed in HeLa cells. EMBO J. 6, 631-636.

Colin, Y., Rahuel, C., London, J., Romeo P.-H., d'Auriol, L., Galibert, F. & Cartron, J.-P. (1986) Isolation of cDNA clones and complete amino acid sequence of human erythrocyte glycophorin C. J. Biol. Chem. 261, 229-233.

Condell, R.A. & Tappel, A.L. (1982) Amino acid sequence around the active site selenocysteine of rat liver glutathione peroxidase. Biochim. Biophys. Acta, 709, 304-309.

Conkie, D., Kleiman, L., Harrison, P.R. & Paul, J. (1975). Increase in the accummulation of globin mRNA in immature erythroblasts in response to erythroprotein *in vivo* or *in vitro*. Exp. Cell Res. 93, 315-324.

Contreras, R., Gheysen, D., Knowland, J., van der Voorde, A. & Fiers, W. (1982) Evidence for the direct involvement of DNA replication origin in synthesis of late SV40 RNA. Nature 300, 500-505.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. & Chambon, P. (1980) Promoter sequences of eukaryotic protein coding genes. Science 209, 1406-1414.

Costlow, N.A., Simon, J.A. & Lis, J.T. (1985) A hypersensitive site in *hsp 70* chromatin requires adjacent not internal DNA sequences. Nature 313, 147-149.

Courey, A.J., Plon, S.E. & Wang, J.C. (1986) The use of psoralen-modified DNA to probe the mechanism of enhancer action. Cell 45, 567-574.

Crabtree, G.R. & Kant, J.A. (1982) Organization of the rat γ-fibrinogen gene : alternative mRNA splice patterns produce the γA and γB (γ') chains of fibrinogen. Cell 31, 159-166.

Darnell, J.E. (1982) Variety in the level of gene control in eukaryotic cells. Nature 297, 365-371.

Davidson, E.H. (1976) Gene activity in early development. 2nd edition, N.Y. Acad. Press.

Davidson, E.H., Hough-Evans, B.R., Britten, R.J. (1982) Molecular biology of the sea urchin embryo. Science 217, 17-26.

Davison, B.L., Egly, J.-M., Mulvihill, E.R. & Chambon, P. (1983) Formation of stable preinitiation complexes between eukaryotic class B transcription factors and promoter sequences. Nature 301, 680-686.

Demuth, D.R., Showe, L.C., Ballantine, M., Palumbo, A., Fraser, P.J., Cioe, L., Rovera, G. & Curtis, P.J. (1986) Cloning and structural characterization of a human non-erythroid bands 3-like protein. EMBO J. 5, 1205-1214.

Derman, E., Kravter, K., Walling, L., Weinberger, C., Ray, M. & Darnell, J.E., (1981) Transcriptional control in the production of liver-specific mRNAs. Cell 23, 731-739.

Dexter, T.M. & Lajtha, L.G. (1974) Proliferation of haemopoietic stem cells *in vitro*. Br. J. Haematol. 28, 525-530.

Dexter, T.M. & Shadduck, R.K. (1980) The regulation of haemopoiesis in long-term bone marrow cultures I : Role of L-cell CSF. J. Cell Physiol. 102, 279-286.

Dexter, T.M., Garland, J., Scott, D., Scolnick, E. & Metcalf, D. (1980) Growth of factor-dependent haemopoietic precursor cell lines. J. Exp. Med. 152, 1036-1047.

Dexter, T.M., Testa, N.G., Allen, T.D., Rutherford, T. & Scolnick, E. (1981) Molecular and cell biologic aspects of erythropoiesis in long-term bone marrow cultures. Blood 58, 699-707.

Dexter, T.M., Spooncer, E., Schofield, R., Lord, B.I. and Simmons, P. (1984) Haemopoietic stem cells and the problem of self-renewal. Blood Cells 10, 315-339.

Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. & Weissmann, C. (1983) Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit ß-globin gene in mouse 3T6 cells. Cell 32, 695-706.

Doerfler, W. (1983) DNA methylation and gene activity. Ann. Rev. Biochem. 52, 93-124.

Doolittle, W.F. (1978) Genes in pieces : were they ever together? Nature 272, 581-582.

Dynan, W.S. & Tjian, R. (1985) Control of eukaryotic mRNA synthesis by sequence-specific DNA-binding proteins. Nature 316, 774-778.

Dyson, P.J., Cook, P.R., Searle, S. & Wyke, J. (1985) The chromatin structure of Rous sarcoma proviruses is changed by factors that act in *trans* in cell hybrids. EMBO J. 4, 413-420.

Earnshaw, W.C., Halligan, B., Cooke, C.A., Heck, M.M.S. & Liu, L.F. (1985) Topoisomerase II is a structural component of mitotic chromosome scaffolds. J. Cell. Biol. 100, 1706-1715.

Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) Two mRNAs can be produced from a single immunoglobulin µ gene by alternative RNA processing pathways. Cell 20, 313-319.

Eaves, C.J., Humphries, R.K. & Eaves, A.C. (1979) *In vitro* characterization of erythroid precursor cells and the erythropoietic differentiation process. In "Cellular and molecular regulation of haemoglobin switching", Stamatoyannopoulos, G. & Neinhuis, A.W. (eds.) New York : Grune and Stratton pp. 251-278.

Edlund, T., Walker, M.D., Barr, P.J. & Rutter, W.J. (1985) Cell specific expression of the rat insulin gene : evidence for role of two distinct 5' flanking elements. Science 230, 912-916.

Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M. O'Connell, C., Spiritz, R.A., DeRiel, J.K., Forget, B.G., Slighton, L., Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C. & Proudfoot, N.J. (1980) The structure and evolution of the human ß-globin gene family. Cell 21, 653-668.

Eick, D. & Bornkamm, G.W. (1986) Transcriptional arrest within the first exon is a fast control mechanism in c-myc gene expression. Nucl. Acids Res. 14, 8331-8346.

Eisen, H. (1978) Terminal erythroid differentiation in normal and leukaemic mouse cells. Blood Cells 4, 177-188.

Eisen, H., Keppel-Ballivet, F., Georgopoulos, C.P., Sassa, S., Granick, J., Pragnell, I. & Ostertag, W. (1978) Biochemical and genetic analysis of erythroid differentiation in Friend virus transformed murine erythroleukaemia cells. Cold Spring Harbor conferences on cell proliferation, Vol. 5, pp.277-294.

Eissenberg, J.C., Cartwright, I.L., Thomas, G.H. & Elgin, S.C.R. (1985) Selected topics in chromatin structure. Ann. Rev. Genet. 19, 485-536.

Eliason, J.F., Testa, N.G. & Dexter, T.M. (1979) Erythropoietin-stimulated erythropoiesis in long-term bone marrow culture. Nature 281, 382-384.

Ephrussi, A., Church, G.M., Tonegawa, S. & Gilbert, W. (1985) B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. Science 227, 134-140.

Epp, O., Ladenstein, R. & Wendel, A. (1983) The refined structure of the selenoenzyme glutathione peroxidase at 0.2nm resolution. Eur. J. Biochem. 133, 51-69.

Favaloro, J.M., Treisman, R.H. & Kamen, R.I. (1980) Transcription maps of polyoma virus-specific RNA : analysis by two dimensional nuclease S1 gel mapping. Methods Enzymol. 65, 718-749.

Falkner, F.G. & Zachau, H.G. (1984) Correct transcription of an immunoglobulin k gene requires an upstream fragment containing conserved sequence elements. Nature 310, 71-74.

Falkner, F.G., Mocikat, R. & Zachau, H.G. (1986) Sequences closely related to an immunoglobulin gene promoter/enhancer element occur also upstream of other eukaryotic and of prokaryotic genes. Nucl. Acids. Res. 14, 8819-8827.

Fibach, E., Gambari, R., Shaw, P.A., Maniatis, G., Reuben, R.C., Sassa, S., Rifkind, R.A. and Marks, P.A. (1979) Tumour-promoter-mediated inhibition of cell differentiation : suppression of the expression of erythroid functions in murine erythroleukaemia cells. Proc. Natl. Acad. Sci. U.S.A. 76, 1906-1910.

Fisher, E.F. & Caruthers, M.H. (1979) Studies on gene control region XII. The functional significance of a lac operator constitutive mutation. Nucl. Acids Res. 7 401-416.

Flohe, L. (1982) Glutathione peroxidase brought into focus. In "Free radicals in biology. Vol. V." Pryor, W.A. (ed.) N.Y. Acad. Press pp.223-254.

Flohe, L., Gunzler, W.A. & Shock, H.H. (1973) Glutathione peroxidase ; a selenoenzyme. FEBS letts. 32 132-134.

Fordis, C.M., Nelson, N., McCormick, M., Padmanabhan, R., Howard, B. & Schechter, A.N. (1986) The 5'-flanking sequences of human globin genes contribute to tissue specific expression. Biochem. Biophys. Res. Commun. 134, 128-133.

Forstrom, J.W., Zakowski, J.J. & Tappel, A.L. (1978) Identification of the catalytic site of rat liver glutathione peroxidase as selenocysteine. Biochem. 17, 2639-2644.

Foster, J., Stafford, J. & Queen, C. (1985) An immunoglobulin promoter displays cell-type specificity independently of the enhancer. Nature 315, 423-425.

Fox, T.D. (1985) Diverged genetic codes in protozoans and a bacterium. Nature 314, 132-133.

Frampton, J., Conkie, D., Chambers, I., McBain, W., Dexter, M. & Harrison, P.R. (1987) Changes in minor transcripts from the $\alpha 1$ and β^{maj} globin and glutathione peroxidase genes during erythropoiesis. Nucl. Acids Res. 15, 3671-3688.

Friend, C., Scher, W., Holland, J.G. & Sato, T., (1971) Haemoglobin synthesis in murine virus-induced leukaemic cells in vitro : stimulation of erythroid differentiation by dimethylsulphoxide. Proc. Natl. Acad. Sci. U.S.A. 68, 378-382.

Frischer, H., Chu, L.K., Ahmad, T., Justice, P. & Smith, G.F. (1981) Superoxide dismutase and glutathione peroxidase abnormalities in erythrocytes and lymphoid cells in Down syndrome. Prog. Clin. Biol. Res. 55, 269-283.

Fritton, H.P., Sippel, A.E. & Igo-Kemenes, T. (1983) Nuclease-hypersensitive sites in the chromatin domain of the chicken lysozyme gene. Nucl. Acids Res. 11, 3467-3485.

Fritton, H.P., Igo-Kemenes, T., Nowock, J., Strech-Jurk, U., Theisen, M. & Sippel, A.E. (1984) Alternative sets of DNAse I hypersensitive sites characterize the various functional states of the chicken lysozyme gene. Nature 311, 163-165.

Fromm, M. & Berg, P. (1982) Deletion mapping of DNA regions required for SV40 early region promoter function in vivo. J. Mol. Appl. Genet. 1, 457-481.

Fromm, M. & Berg, P. (1983) SV40 early-region and late-region promoter functions are enhanced by the 72-base pair repeat inserted at distant locations and inverted orientations. Mol. Cell. Biol. 3, 991-999.

Fujii, S., Dale, G.L. & Beutler, E. (1984) Glutathione-dependent protection against oxidative damage of the human red cell membrane. Blood 63, 1096-1101.

Fulton, R., Birnie, G.D. & Knowler, J.T. (1985) Post-transcriptional regulation of rat liver gene expression by glucocorticoids. Nucl. Acids Res. 13, 6467-6482.

Gaffney, D.F., McLauchlan, J., Whitton, J.L. & Clements, J.B. (1985) A modular system for the assay of transcription regulatory signals : the sequence TAATGARAT is required for herpes simplex virus immediate early gene activation. Nucl. Acids Res. 13 7847-7863.

Garcia, J.V., Bich-Thuy, L.T., Stafford, J. & Queen, C. (1986) Synergism between immunoglobulin enhancers and promoters. Nature 322, 383-385.

Gazitt, Y. & Friend, C. (1980) The possible role of polyamine biosynthetic enzymes in the induction of differentiation in Friend erythroleukaemic cells. in "In vivo and in vitro erythropoiesis : the Friend system". Rossi, G.B. (ed.) Elsevier/North Holland, Amsterdam pp.209-218.

Geller, A. and Rich, A. (1980) A UGA termination suppression $tRNA^{Trp}$ active in rabbit reticulocytes. Nature 283, 41-46.

Gerster, T., Matthias, P., Thali, M., Jiricny, J. & Schaffner, W. (1987) Cell type-specificity elements of the immunoglobulin heavy chain gene enhancer. EMBO J. 6, 1323-1330.

Gidoni, D., Dynan, W.S. & Tjian, R. (1984) Multiple specific contacts between a mammalian transcription factor and its cognate promoters. Nature 312, 409-413.

Gidoni, D., Kadonaga, J.T., Barrera-Saldana, H., Takahashi, K., Chambon, P. & Tjian, R. (1985) Bidirectional SV40 transcription mediated by tandem Sp1 binding interactions. Science 230 511-517.

Gil, A. & Proudfoot, N.J. (1987) Position-dependant sequence elements downstream of AAUAAA are required for efficient rabbit ß-globin mRNA 3' end formaion. Cell 49, 399-406.

Gilbert, W. (1978) Why genes in pieces? Nature 271, 501.

Gilmour, D.S. & Lis, J.T. (1986) RNA polymerase II interacts with the promoter region of the noninduced hsp 70 gene in Drosophila melanogaster cells. Mol. Cell. Biol. 6, 3984-3989.

Godowski, P.J., Rusconi, S., Miesfeld, R. & Yamamoto, K.R. (1987) Glucocorticoid receptor mutants that are constitutive activators of transcriptional enhancement. Nature 325, 365-368.

Goldfarb, P.S., O'Prey, J., Affara, N., Yang, Q.-S. & Harrison, P.R. (1983) Isolation of non-globin genes expressed preferentially in mouse erythroid cells. Nucl. Acids Res. (1983) 11, 3517-3530.

Goodbourn, S., Burstein, H. & Maniatis, T. (1986) The human ß-interferon gene enhancer is under negative control. Cell 45, 601-610.

Gopal, T.V., Shimada, T., Baur, A.W. & Nienhuis, A.W. (1985) Contribution of promoter to tissue-specific expression of the mouse immunoglobulin kappa gene. Science 229, 1102-1104.

Gorman, C.M., Moffat, L.F. & Howard, B.H. (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. 2, 1044-1051.

Grandchamp, B., Romeo, P.-H., Dubart, A., Raich, N., Rosa, J., Nordmann, Y, & Goosens, M. (1984) Molecular cloning of a cDNA sequence complementary to porphobilinogen deaminase mRNA from rat. Proc. Natl. Acad. Sci. U.S.A. 81, 5036-5040.

Graves, B.J., Johnson, P.F. & McKnight, S.L. (1986) Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell 44, 565-576.

Green, S. & Chambon, P. (1986) A super family of potentially oncogenic hormone receptors. Nature 324, 615-617.

Gregory, C.J. & Henkelman, R.M. (1977). Relationships between early hemopoietic progenitor cells determined by correlation analysis of their numbers in individual spleen colonies. In "Experimental haematology today", Baum, S.J. & Ledney, G.D. (eds.) New York : Springer-Verlag pp.93-101.

Greiser-Wilke, I., Ostertag, W., Goldfarb, P., Lang, A., Furusawa, M. & Conscience, J.F. (1980) Inducibility of spleen focus-forming virus by BrdUrd is controlled by the differentiated state of the cell. Proc. Natl. Acad. Sci. U.S.A. 78, 2995-2999.

Grosschedl, R. & Baltimore, D. (1985) Cell-type specificity of immunoglobulin gene expression is regulation by at least three DNA sequence elements. Cell 41, 885-897.

Grossman, A. & Wendel, A. (1983) Non-reactivity of the selenoenzyme glutathione peroxidase with enzymatically hydroperoxidised phospholipids. Eur. J. Biochem. 135, 549-522.

Groudine, M. & Weintraub, H. (1982) Propagation of globin DNAse I-hypersensitive sites in absence of factors required for induction. A possible mechanism for determination. Cell 30, 131-139.

Guidi, G., Schiavon, R., Biasioli, A. & Perona, G. (1984) The enzyme glutathione peroxidase in arachidonic acid metabolism of human plalelets. J. Lab. Clin. Med. 104 (4), 574-582.

Gunzler, W.A., Steffens, G.J., Grossmann, A., Kim, S.M.A., Otting, F., Wendel, A. & Flohe, L. (1984) The amino acid sequence of bovine glutathione peroxidase. Hoppe-Seyler's Z. Physiol. Chem. 365, 195-212.

Gusella, J., Geller, R., Clarke, B., Weeks, V. & Housman, D. (1976) Commitment to erythroid differentiation by Friend erythroleukaemia cells : a stochastic analysis. Cell 9, 221-229.

Guyette, W.A., Matusik, R.J. & Rosen, J.M. (1979) Prolactin-mediated transcriptional and post-transcriptional control of casein gene expression. Cell 17, 1013-1023.

Hamlyn, P.H., Brownlee, G.G., Cheng, C.-C., Gait, M.J. & Milstein, C. (1978) Complete sequence of constant and 3' noncoding regions of an immunoglobulin mRNA using the dideoxynucleotide method of RNA sequencing. Cell 15, 1067-1075.

Hansen, U. & Sharp, P.A. (1983) Sequences controlling *in vitro* transcription of SV40 promoters. EMBO J. 2, 2293-2303.

Hansen, U., Tenen, D.G., Livingston, D.M. & Sharp, P. (1981) T antigen repression of SV40 early transcription from two promoters. Cell 27, 603-612.

Hanyu, N., Kuchino, Y., Nishimura, S. & Beier, H. (1986) Dramatic events in ciliate evolution : alteration of UAA and UAG termination codons to glutamine codons due to anticodon mutations in two Tetrahymena $tRNAs^{Gln}$. EMBO J. 6, 1307-1311.

Harrison, P.R. (1977) The biology of the Friend cell. In "Biochemistry of cell differentiation, II" Vol. 15. J. Paul (ed.) Baltimore, University Park Press. pp. 227-267.

Harrison, P.R. (1982) Regulation of differentiation in retrovirus-induced murine erythroleukaemias : A personal view from the fridge. Cancer Surveys 1, 231-277.

Harrison, P.R. (1984) Molecular analysis of erythropoiesis. Exp. Cell Res. 155, 321-344.

Harrison, P.R., Conkie, D., Affara, N. & Paul, J. (1974) In situ localization of globin messenger RNA formation I. During mouse foetal liver development. J. Cell Biol. 63, 402-413.

Harrison, P.R., Affara, N., McNab, A. & Paul, J. (1977) Erythroid differentiation in a Friend erythroleukaemia cell x lymphoma hybrid cell line is limited, possibly due to reduced haem levels. Exp. Cell Res. 109, 237-246.

Harrison, P.R., Rutherford, T., Conkie, D., Affara, N., Sommerville, J., Hissey, P. & Paul, J. (1978) Analysis of erythroid differentiation in Friend cells using noninducible variants. Cell 14, 61-70.

Hatfield, D., Diamond, A. & Dudock, B. (1982) Opal suppressor serine tRNAs from bovine liver form phosphoseryl-tRNA. Proc. Natl. Acad. Sci. U.S.A. 79, 6215-6219.

Hatfield, D. (1985) Suppression of termination codons in higher eukaryotes. T.I.B.S. 10, 201-204.

Hawkes, W.C., Lyons, D.E. & Tappel, A.L. (1982) Identification of selenocysteine-specific aminoacyl transfer RNA from rat liver. Biochim. Biophys. Acta 699, 183-191.

Hawkes, W.C. & Tappel, A.L. (1983) In vitro synthesis of glutathione peroxidase from selenite : translational incorporation of selenocysteine. Biochem. Biophys. Acta 739, 225-234.

Hawkes, W.C., Wilhelmsem E.C. & Tappel, A.L. (1985) Abundance and tissue distribution of selenocysteine-containing proteins in the rat. J. Inorg. Biochem. 23, 77-92.

Herbomel, P., Saragosti, S., Blangy, D. & Yaniv, M. (1981) Fine structure of the origin-proximal DNAse I hypersensitive region in wild type and EC mutant polyoma. Cell 25, 651-658.

Herr, W. & Clarke, J. (1986) The SV40 enhancer is composed of multiple functional elements that can compensate for one another. Cell 45, 461-470.

Hess, J., Perez-Stable, C., Deisseroth, A. & Shen, C.-K.J. (1985) Characterization of an unique RNA initiated immediately upstream from human α1 globin gene in vivo and in vitro : polymerase II dependence, tissue specificity and subcellular location. Nucl. Acids Res. 13, 6059-6074.

High, S. & Tanner, M.J.A. (1987) Human erythrocyte membrane sialoglycoprotein β. The cDNA sequence suggests the absence of a cleaved N-terminal signal sequence. Biochem. J. 243, 277-280.

Hol, W.G.J. (1985) The role of the α-helix dipole in protein function and structure. Prog. Biophys. Molec. Biol. 45, 149-195.

Hol, W.G.J., van Duijnen, P.T. & Berendsen, H.J.C. (1978) The α-helix dipole and the properties of proteins. Nature 273, 443-446.

Hollenberg, S.M., Giguere, V., Segui, P. & Evans, R.M. (1987) Colocalization of DNA-binding and transcriptional activation functions in the human glucocorticoid receptor. Cell 49, 39-46.

Holmes, D.S. & Quigley (1981) A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114, 193-197.

Hornsby, P.J., Pearson, D.W., Autor, A.P., Aldern, K.A. & Harris, S.E. (1985) Selenium deficiency in cultured adrenocortical cells : restoration of glutathione peroxidase and resistance to hydroperoxides on addition of selenium. J. Cell. Physiol. 123, 33-38.

Humphries, R.K., Eaves, A.C. & Eaves, C.J. (1981) Self-renewal of haemopoietic stem cells during mixed colony formation in vitro. Proc. Natl. Acad. Sci. U.S.A. 78, 3629-3633.

Hutchison, N. & Weintraub, H. (1985) Localization of DNAse I-sensitive sequences to specific regions of interphase nuclei. Cell 43, 471-482.

Imbra, R.J. & Karin, M. (1986) Phorbol ester induces the transcriptional stimulatory activity of the SV40 enhancer. Nature 323, 555-558.

Imbra, R.J. & Karin, M. (1987) Metallothionein gene expression is regulated by serum factors and activators of protein kinase C. Mol. Cell. Biol. 7, 1358-1363.

Ip, C. (1985) Selenium inhibition of chemical carcinogenesis. Fed. Proc. 44, 2573-2578.

Jackson, D.A. & Cook, P.R. (1985) Transcription occurs at the nucleoskeleton. EMBO J. 4, 919-925.

Jantzen, H.-M., Strahle, U., Gloss, B., Stewart, F., Schmid, W., Boshart, M., Miksicek, R. & Schutz, G. (1987) Cooperativity of glucocorticoid response elements located far upstream of the tyrosine aminotransferase gene. Cell 49, 29-38.

Johnson, G.R., Keller, G.M. & Nicola, N.A. (1982) Differentiation and "renewal" of multipotential cells in vitro. J. Cell Physiol. [Suppl.1] 23-30.

Jones, K.A. & Tjian, R. (1985) Sp1 binds to promoter sequences and activates HSV "immediate-early" gene transcription in vitro. Nature 317, 179-182.

Jones, K.A., Yamamoto, K.R. & Tjian, R. (1985) Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell 42 559-572.

Jones, K.A., Kadonaga, J.T.,, Luciw, P.A. & Tjian, R. (1986) Activation of the AIDS retrovirus promoter by the cellular transcription factor Sp1. Science 232, 755-759.

Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J. & Tjian, R. (1987) A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell 48, 79-89.

Jones, N.C. (1986) Negative regulation of enhancers. Nature 321, 202-203.

Jongstra, J., Reudelhuber, T.L., Oudet, P., Benoist, C., Chae, C.-B., Jeltsch, J.-M., Mathis, D.J. & Chambon, P. (1984) Induction of altered chromatin structures by simian virus 40 enhancer and promoter elements. Nature 307, 708-714.

Joulin, V., Peduzzi, J., Romeo, P.-H., Rosa, R., Valentin, C., Dubart, A., Lapeyre, B., Blouquit, Y., Garel, M.-C., Goosens, M., Rosa, J., Cohen-Solal, M. (1986) Molecular cloning and sequencing of the human erythrocyte 2, 3-bisphosphoglycerate mutase cDNA : revised amino acid sequence. EMBO J. 5, 2275-2283.

Kadesch, T. & Berg, P. (1986) Effects of the position of the simian virus 40 enhancer on expression of multiple transcription units in a single plasmid. Mol. Cell. Biol. 6, 2593-2601.

Kadonaga, J.T., Jones, K.A. & Tjian, R. (1986) Promoter-specific activation of RNA polymerase II transcription by Sp1. TIBS 11, 20-23.

Karin, M., Haslinger, A., Holtgreve, H., Richards, R.I., Krauter, P., Westphal, H.M. & Beato, M. (1984) Characterization of DNA sequences through which cadmium and glucocorticoid hormones induce human metallothionein IIA gene. Nature 308, 513-519.

Karin, M., Imbra, R.J., Heguy, A. & Wong, G. (1985) Interleukin 1 regulates human metallothionein gene expression. Mol. Cell. Biol. 5, 2866-2869.

Kasturi, K. & Harrison, P.R. (1985) The cell specificity and biosynthesis of mouse glycophorins studied with monoclonal antibodies. Exp. Cell Res. 157, 253-264

Keegan, L., Gill, G. & Ptashne, M. (1986) Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science 231, 699-704.

Keshet, E. & Cedar, H. (1983) Effect of CpG methylation on Msp I. Nucl. Acids Res. 11, 3571-3580.

Keshet, I., Lieman-Hurwitz, J. & Cedar, H. (1986) DNA methylation affects the formation of active chromatin. Cell 44, 535-543.

Khoury, G. & Gruss, P. (1983) Enhancer elements. Cell 33, 313-314.

Killary, A.M. & Fournier, R.E.K. (1984) A genetic analysis of extinction : Trans-dominant loci regulate expression of liver-specific traits in hepatoma hybrid cells. Cell 38, 523-534.

Kirk, M.M. & Kirk, D.L. (1985) Translational regulation of protein synthesis, in response to light, at a critical stage of Volvox development. Cell 41, 419-428.

Kit, S., Dubbs, D.R., Pierkarski, L.J. & Hsu, T.C. (1963) Deletion of thymidine kinase activity from L cells resistant to bromodeoxyuridine. Exp. Cell Res. 31, 297-312.

Klebe, R.J., Chen, T.-R. & Ruddle, F.H. (1970) Controlled production of proliferating somatic cell hybrids. J. Cell Biol. 45, 74-82.

Klein, S., Sablitzky, F. & Radbruch, A. (1984) Deletion of the IgH enhancer does not reduce immunoglobulin heavy chain production of a hybridoma IgD class switch variant. EMBO J. 3, 2473-2476.

Kmiec, E.B. & Worcel, A. (1985) The positive transcription factor of the 5S RNA gene induces a 5S DNA-specific gyration in Xenopus oocyte extracts. Cell 41, 945-953.

Kmiec, E.B., Razvi, F. & Worcel, A. (1986) The role of DNA-mediated transfer of TFIIIA in the concerted gyration and differential activation of the Xenopus 5S RNA genes. Cell 45, 209-218.

Konialis, C.P., Barlow, J.H. & Butterworth, P.H.W. (1985) Cloned cDNA for rabbit erythrocyte carbonic anhydrase I : a novel erythrocyte-specific probe to study development in erythroid tissues. Proc. Natl. Acad. Sci. U.S.A. 82, 663-667.

Konieczny, S.F. & Emerson, C.P. (1984) 5-azacytidine induction of stable mesodermal stem cell lineages from 10T1/2 cells : evidence for regulatory genes controlling determination. Cell 38, 791-800.

Kopito, R.R. & Lodish, H.F. (1985) Primary structure and transmembrane orientation of the murine anion exchange protein. Nature 316, 234-238.

Korner, A.M., Feinstein, S.I. & Altman, S. (1979) Transfer RNA-mediated suppression. In "Transfer RNA" Altman, S. Cambridge, Mass. MIT Press pp. 105-135.

Kovesdi, I., Reichel, R. & Nevins, J.R. (1986a) E1A transcription induction : enhanced binding of a factor to upstream promoter sequences. Science 231, 719-722.

Kovesdi, I., Reichel, R. & Nevins, J.R. (1986b) Identification of a cellular transcription factor involved in E1A trans-activation. Cell 45, 219-228.

Kramerov, D.A., Tillib, S.V., Ryskov, A.P. & Georgiev, G.P. (1985) Nucleotide sequence of small polyadenylated B2 RNA. Nucl. Acids Res. 13, 6423-6437.

Krantz, S.B. & Goldwasser, E. (1984) Specific binding of erythropoietin to spleen cells infected with the anaemia strain of Friend virus. Proc. Natl. Acad. Sci. U.S.A. 81, 7574-7578.

Krayev, A.S., Kramerov, D.A., Skryabin, K.G., Ryskov, A.P., Bayev, A.A. & Georgiev, G.P. (1980) The nucleotide sequence of the ubiquitous repetitive DNA sequence B1 complementary to the most abundant class of mouse fold-back RNA. Nucl. Acids Res. 8, 1201-1215.

Krowczynska, A., Yenofsky, R. & Brawerman, G. (1985) Regulation of messenger RNA stability in mouse erythroleukaemia cells. J. Mol. Biol. 181, 231-239.

Kwoh, T.J. & Engler, J.A. (1984) The nucleotide sequence of the chicken thymidine kinase gene and the relationship of its predicted polypeptide to that of the vaccinia virus thymidine kinase. Nucl. Acids Res. 12, 3959-3971.

Lacy, E. & Maniatis, T. (1980) The nucleotide sequence of a rabbit β-globin pseudogene. Cell 21, 545-553.

Laimins, L.A., Khoury, G., Gorman, C., Howard, B.H., & Gruss, P. (1982) Host specific activation of transcription by tandem repeats from Simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. 79, 6453-6457.

Lands, W.E.M., Kulmacz, R.J. & Marshall, P.J. (1984) Lipid peroxide actions in the regulation of prostaglandin biosynthesis. In "Free Radicals in Biology. Vol. VI" Pryor, W.A. (ed.) N.Y. Academic Press. pp. 39-61.

Laskey, R.A. & Mills, A.D. (1977) Enhanced autoradiographic detection of ^{32}P and ^{125}I using intensifying screens and hypersensitized film. FEBS Lett. 82, 314-316.

Lee, W., Haslinger, A., Karin, M. & Tjian, R. (1987) Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature 325, 368-372.

Leff, S.E., Rosenfeld, M.G. & Evans, R.M. (1986) Complex transcriptional units : diversity in gene expression by alternative RNA processing. Ann. Rev. Biochem. 55, 1091-1117.

Lehrach, H., Diamond, D., Wozney, J.M. & Boedtker, H. (1977) RNA molecular weight determination by gel electrophoresis under denaturing conditions, a critical re-examination. Biochemistry 16, 4743-4751.

Levander, O.A., de Loach, D.P., Morris, V.C. & Moser, P.B. (1983) Platelet glutathione peroxidase activity as an index of selenium status in rats. J. Nutr. 113, 55-63.

Levanon, D., Lieman-Hurwitz, J., Dafni, N., Bigderson, M., Sherman, L., Bernstein, Y., Laver-Rudich, Z., Danciger, E., Stein, O. & Groner, Y. (1985) Architecture and anatomy of the chromosomal locus in human chromosome 21 encoding the Cu/Zn superoxide dismutase. EMBO J. 4, 77-84

Levenson, R. & Housman, D. (1979a) Memory of MEL cells to a previous exposure to inducer. Cell 17, 485-490.

Levenson, R. & Housman, D. (1979b) Developmental program of murine erythroleukaemia cells. Effect of the inhibition of protein synthesis. J. Cell Biol. 82, 715-725.

Levenson, R., Kerner, J. & Housman, D. (1979) Synchronization of MEL cell commitment with cordycepin. Cell 18, 1073-1078.

Levenson, R., Housman, D. & Cantley, L. (1980) Amiloride inhibits erythroleukaemia cell differentiaiton : Evidence for a Ca^{2+} requirement for commitment. Proc. Natl. Acad. Sci. U.S.A. 77, 5948-5952.

Lindquist, S. (1986) The heat-shock response. Ann. Rev. Biochem. 55, 1151-1191.

Liu, D., Simonsen, C.C. & Levinson, A.D. (1984) Initiation of translation at internal AUG codons in mammalian cells. Nature 309, 82-85.

Lorch, Y., LaPointe, J.W. & Kornberg, R.D. (1987) Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. Cell 49, 203-210.

McCutchan, J.H. & Pagano, J.S. (1968) Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl Dextran. J. Natl. Cancer Inst. 41, 351-356.

McGhee, J.D., Wood, W.I., Dolan, M., Engel, J.D. & Felsenfeld, G. (1981) A 200bp region at the 5' end of the chicken adult β-globin gene is accessible to nuclease digestion. Cell 27, 45-55.

McKnight, S. & Tjian, R. (1986) Transcriptional selectivity of viral genes in mammalian cells. Cell 46, 795-805.

McLauchlan, J., Gaffney, D., Whitton, J.L. & Clements, J.B. (1985) The consensus sequence YGTGTTYY located downstream form the AAUAAA signal is required for efficient formation of mRNA 3' termini. Nucl. Acids Res. 13, 1347-1368.

Macguire, D.J., Day, A.R., Borthwick, I.A., Srivastava, G., Wigley, P.L., May, B.K. & Elliot, W.H. (1986) Nucleotide sequence of the chicken 5-aminolevulinate synthase gene. Nucl. Acids Res. 14, 1379-1391.

Mandel, M. & Higa, A.J. (1970) Calcium-dependant bacteriophage DNA interaction. J. Mol. Biol. 53, 159-162.

Magli, M.C., Iscove, N.N. & Odartchenko, N. (1982) The transient nature of early haemopoietic spleen colonies. Nature 295, 527-529.

Marks, P.A. & Rifkind, R.A. (1978) Erythroleukaemia differentiation. Ann. Rev. Biochem. 47, 419-448.

Martini, G., Toniolo, D., Vulliamy, T., Luzzatto, L., Dono, R., Viglietto, G., Paonessa, G., D'Urso, M. & Persico, M.G. (1986) Structural analysis of the X-linked gene encoding human glucose-6-phosphate dehydrogenase. EMBO J. 5, 1849-1855.

Mason, J.O., Williams, G.T. & Neuberger, M.S. (1985) Transcription cell type specificity is conferred by an immunoglobulin VH gene promoter that includes a functional consensus sequence. Cell 41, 479-487.

Masters, J.N. & Attardi, G. (1985) Discrete human dihydrofolate reductase gene transcripts present in polysomal RNA map with their 5' ends several hundred nucleotides upstream of the main mRNA start sites. Mol. Cell. Biol. 5, 493-500.

Maxam, A.M. & Gilbert, W. (1980) Sequencing end-labelled DNA with base specific chemical cleavages. Methods Enzymol. 65, 499-560.

Mellon, P., Parker, V., Gluzman, Y. & Maniatis, T. (1981) Identification of DNA sequences required for transcription of the human α1-globin gene in a new SV40 host-vector system. Cell 27, 279-288.

Melton, D.W., Konecki, D.S., Brennand, J. & Caskey, C.T. (1984) Structure, expression and mutation of the hypoxanthine phosphoribosyl transferase gene. Proc. Natl. Acad. Sci. U.S.A. 81, 2147-2151.

Menninger, J.R. (1976) Peptidyl transfer RNA dissociates during protein synthesis from ribosomes of Escherichia coli. J. Biol. Chem. 251, 3392-3398.

Messner, H.A., Izaguirra, C.A. & Jamal, N. (1982) Identification of T-lymphocytes in human mixed haemopoietic colonies. Blood 58, 402-405.

Metcalf, D. (1985) The granulocyte-macrophage colony stimulating factors. Cell 43, 5-6.

Metcalf, D. (1981) Haemopoietic colony stimulating factor. In "Tissue Growth Factors" Baserga, R. (ed.) Berlin, Springer-Verlag. pp. 343-384.

Miller, J.H. & Albertini, A.M. (1983) Effects of surrounding sequence on the suppression of nonsense codons. J. Mol. Biol. 164, 59-71.

Mills, G.C. (1957) Haemoglobin catabolism I. Glutathione peroxidase, an erythrocyte enzyme which protects haemoglobin from oxidative breakdown. J. Biol. Chem. 229, 189-197.

Minty, A.J. & Gros, F. (1980) Coding potential of non-polyadenylated mRNA in mouse Friend cells. J. Mol. Biol. 139, 61-83.

Mirkovitch, J., Mirault, M.-E. & Laemmli, U.K. (1984) Organization of the higher-order chromatin loop : Specific DNA attachment sites in nuclear scaffolds. Cell 39, 223-232.

Mitchell, P.J., Carothers, A.M., Han, J.H., Harding, J.D. Kas, E., Venolia, L. & Chasin, L.A. (1986) Multiple transcription start sites, DNAse I hypersensitive sites and an opposite strand exon in the 5' region of the CHO dhfr gene. Mol. Cell. Biol. 6, 425-440.

Mizuiri, S., Hirata, K., Izumi, S., Komatsu, N., Yoshimura, S. & Watanabe, K. (1986) Immunocytochemical study of glutathione peroxidase in normal rat kidney. Renal Physiol. 9, 249-256.

Mizutani, T. & Hashimoto, A. (1984) Purification and properties of suppressor seryl-tRNA:ATP phosphotransferase from bovine liver. FEBS Lett. 169, 319-322.

Mizutani, T. & Tachibana, Y. (1986) Possible incorporation of phosphoserine into globin readthrough protein via bovine opal suppressor phosphoseryl-tRNA. FEBS Lett. 207, 162-166.

Moldave, K. (1985) Eukaryotic protein synthesis. Ann. Rev. Biochem. 54, 1109-1149.

Moore, M. & Dexter, T.M. (1986) Growth and development in the haemopoietic system : the role of lymphokines and their possible therapeutic potential in disease and malignancy. Carcinog. 7, 509-516.

Morgan, R.W., Christman, M.F., Jacobson, F.S., Storz, G. & Ames, B.N. (1986) Hydrogen peroxide-inducible proteins in Salmonella typhimurium overlap with heat shock and other stress proteins. Proc. Natl. Acad. Sci. U.S.A. 83, 8059-8063.

Muglia, L. & Rothman-Denes, L.B. (1986) Cell type-specific negative regulatory element in the control region of the rat α-fetoprotein gene. Proc. Natl. Acad. Sci. U.S.A. 83, 7653-7657.

Murakoshi, M., Osamura, Y., Yoshimura, S., Izumi, S., Komatsu, N. & Watanabe, K. (1981) Biochemical and immunohistochemical studies of glutathione peroxidase in the rat adrenal cortex. Acta histochem. cytochem. 14, 571-580.

Murgola, E.J., Pagel, F.T. & Hijazi, K.A. (1984) Codon context effects in missense suppression. J. Mol. Biol. 175, 19-27.

Myers, R.M., Tilly, K. & Maniatis, T. (1986) Fine structure genetic analysis of a β-globin promoter. Science 232, 613-618.

Nabel, G. & Baltimore, D. (1987) An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature 326, 711-713.

Nedospasov, S.A. & Georgiev, G.P. (1980) Non-random cleavage of SV40 DNA in the compact minichromosome and free in solution by micrococcal nuclease. Biochem. Biophys. Res. Commun. 92, 532-539.

Needleman, P., Turk, J., Jakschik, B.A., Morrison, A.R. & Lefkowith, J.B. (1986) Arachidonic acid metabolism. Ann. Rev. Biochem. 55, 69-102.

Neve, J., Vertongen, F. & Molle, L. (1985) Selenium deficiency. Clinics in endocrinology and metabolism. 14, 629-656.

Nevins, J.R. (1983) The pathway of eukaryotic mRNA formation. Ann. Rev. Biochem. 52, 441-466.

Newsholme, E.A. & Start, C. (1973) Regulation in metabolism. Chicester : Wiley.

Nicholls, R.D., Hill, A.V.S., Clegg, J.B. & Higgs, D.R. (1985) Direct cloning of specific genomic DNA sequences in plasmid libraries following fragment enrichment. Nucl. Acids Res. 13, 7569-7578.

Nicola, N.A. & Johnson, G.R. (1982) The production of committed haemopoietic colony-forming cells from multipotential precorsors *in vitro*. Blood 60, 1019-1029.

Nir, U., Walker, M.D. & Rutter, W.J. (1986) Regulation of rat insulin I gene expression : evidence for negative regulation in nonpancreatic cells. Proc. Natl. Acad. Sci. U.S.A. 83, 3180-3184.

Nishizuka, Y. (1986) Studies and perspectives of protein kinase C. Science 233, 305-312.

Nomura, S., Yamagoe, S., Kamiya, T. & Oishi, M. (1986) An intracellular factor that induces erythroid differentiation in mouse erythroleukaemia (Friend) cells. Cell 44, 663-669.

Norrander, J., Kempe, T. & Messing, J. (1983) Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene 26, 101-106.

Ondek, B., Shepard, A. & Herr, W. (1987) Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. EMBO J. 6, 1017-1025.

Orkin, S.H., Swan, D. & Leder, P. (1975) Differential expression of α- and ß- globin genes during differentiation of cultured erythroleukaemia cells. J. Biol. Chem. 250, 8753-8760.

Osborne, T.F. & Berk, A.J. (1983) Far upstream initiation sites for adenovirus early region 1A transcription are utilized after the onset of viral DNA replication. J. Virol. 45, 594-599.

Ott, M.-O., Sperling, L., Herbomel, P., Yaniv, M. & Weiss, M. (1984) Tissue-specific expression is conferred by a sequence from the 5' end of the rat albumin gene. EMBO J. 3, 2505-2511.

Pabo, C.O. & Sauer, R.T. (1984) Protein-DNA recognition. Ann. Rev. Biochem. 53, 293-321.

Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. & Sharp, P.A. (1986) Splicing of messenger RNA precursors. Ann. Rev. Biochem. 55, 1119-1150.

Parker, C.S. & Topol, J. (1984) A Drosophila RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. Cell 36, 357-369.

Parslow, T.G. & Granner, D.K. (1982) Chromatin changes accompany immunoglobulin k gene activation : a potential control region within the gene. Nature 299, 449-451.

Parslow, T.G., Blair, D.L., Murphy, W.J. & Granner, D.K. (1984) Structure of the 5' ends of immunoglobulin genes : a novel conserved sequence. Proc. Natl. Acad. Sci. U.S.A. 81, 2650-2654.

Peabody, D.S. & Berg, P. (1986) Termination-reinitiation occurs in the translation of mammalian cell mRNAs. Mol. Cell. Biol. 6, 2695-2703.

Peabody, D.S., Subramani, S. & Berg, P. (1986) Effect of upstream reading frames on translation efficiency in simian virus 40 recombinants. Mol. Cell. Biol. 6, 2704-2711.

Perchellet, J.-P., Perchellet, E.M., Orten, D.K. & Schneider, B.A. (1986) Decreased ratio of reduced/oxidised glutathione in mouse epidermal cells treated with tumour promoters. Carcinog. 7, 503-506.

Perona, G., Guidi, G.C., Piga, A., Cellerino, R., Menna, R. & Zatti, M. (1978) In vivo and in vitro variations of human erythrocyte glutathione peroxidase activity as result of cells ageing, selenium availability and peroxide activation. Br. J. Haematology 39, 399-408.

Petit, C., Levilliers, J., Ott, M.-O. & Weiss, M.C. (1986) Tissue-specific expression of the rat albumin gene : Genetic control of its extinction in microcell hybrids. Proc. Natl. Acad. Sci. U.S.A. 83, 2561-2565.

Picard, D. & Schaffner, W. (1985) Cell type preference of immunoglobulin κ and λ gene promoters. EMBO J. 4, 2831-2838.

Plumb, M.A., Lobanenkov, V.V., Nicolas, R.H., Wright, C.A., Zavou, S. & Goodwin, G.H. (1986) Characterization of chicken erythroid nuclear proteins which bind to the nuclease hypersensitive regions upstream of the β^A- and β^H- globin genes. Nucl. Acids Res. 14, 7675-7693.

Prentice, D.A., Tobey, R.A. & Gurley, L.R. (1983) DNAse I and cellular factors that affect chromatin structure. Biochim. Biophys. Acta 741, 288-296.

Proudfoot, N.J. & Brownlee, G.G. (1976) 3' non-coding region sequences in eukaryotic messenger RNA. Nature 263, 211-214.

Ptashne, M. (1986) Gene regulation by proteins acting nearby and at a distance. Nature 322, 697-701.

Quan, F., Korneluk, R.G., Tropak, M.B. & Gravel, R.A. (1986) Isolation and characterization of the human catalase gene. Nucl. Acids Res. 14, 5321-5335.

Ramos-Martinez, J.I., Launay, J.M., Dreux, C. (1979) A sensitive fluorimetric microassay for the determination of glutathione peroxidase activity. Application to human blood platelets. Anal. Biochem. 98, 154-159.

Rapoport, S.M., Schewe, T., Wiesner, R., Halangk, W., Ludwig, P., Janicke-Hohne, M., Tannert, C., Hiebsch, C. & Klatt, D. (1979) The lipoxygenase of reticulocytes. Purification, characterization, and biological dynamics of the lipoxygenase; its identity with the respiratory inhibitors of the reticulocyte. Eur. J. Biochem. 96, 545-561.

Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Scarpa, S., Carotti, D. & Cantoni, G.L. (1986) Replacement of 5-methylcytosine by cytosine : A possible mechanism for transient DNA demethylation during differentiation. Proc. Natl. Acad. Sci. U.S.A. 83, 2827-2831.

Remmers, E.F., Yang, J.-Q. & Marcu, K.B. (1986) A negative transcriptional control element located upstream of the murine c-*myc* gene. EMBO J. 5, 899-904.

Reynolds, G.A., Basu, S.K., Osborne, T.F., Chin, D.J., Gil, G., Brown, M.S., Goldstein, J.L. & Luskey, K.L. (1984) HMG CoA reductase : A negatively regulated gene with unusual promoter and 5' untranslated regions. Cell 38 275-285.

Rich, I.N. & Kubanek, B. (1985) The central role of the macrophage in haemopoietic microenvironmental regulation. In "Hematopoietic Stem cell physiology" N.Y. Liss pp. 283-298.

Rigby, P.W.J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113, 237-251.

Rink, T.J. & Hallam, T.J. (1984) What turns platelets on? T.I.B.S. 9, 215-219.

Robbins, P.D., Rio, D.C. & Botchan, M.R. (1986) *trans* activation of the simian virus 40 enhancer. Mol. Cell. Biol. 6, 1283-1295.

Rogers, J., Early, P., Carter, C., Calame, K., Bond, M., Hood, L. & Wall, R. (1980) Two mRNAs with different 3' ends encode membrane-bound and secreted forms of immunoglobulin μ chain. Cell 20, 303-312.

Romeo, P.-H., Dubart, A., Grandchamp, B., de Verneuil, H., Rosa, J., Nordmann, Y. & Goossens, M. (1984) Isolation and identification of a cDNA clone coding for rat uroporphyrinogen decarboxylase. Proc. Natl. Acad. Sci. U.S.A. 81, 3346-3350.

Rotruck, J.T., Pope, A.L., Ganther, H.E., Swanson, A.B., Hafeman, D.G., Hoekstra, W.G. (1973) Selenium, biochemical role as a component of GSHPx. Science 179, 588-590.

Russell, E.S. (1979) Hereditary anaemias of the mouse : a review for geneticists. Adv. Genet. 20, 357-459.

Rutherford, T.R. & Harrison, P.R. (1979) Globin synthesis and erythroid differentiation in a Friend cell variant deficient in heme synthesis. Proc. Nucl. Acad. Sci. U.S.A. 76, 5660-5664.

Ryoji, M. & Worcel, A. (1984) Chromatin assembly in Xenopus oocytes : in vivo studies. Cell 37, 21-32.

Sander, M. & Hsieh, T.-S. (1985) Drosophila topoisomerase II double-strand DNA cleavage. Analysis of DNA sequence homology at the cleavage site. Nucl. Acids Res. 13, 1057-1072.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.

Santi, D.V., Garret, C.E. & Barr, P.J. (1983) On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogues. Cell 33, 9-10.

Santoro, M.G., Beneditto, A. & Jaffe, B.M. (1978) Hydrocortisone inhibits DMSO-induced differentiation of Friend erythroleukaemia cells. Biochem. Biophys. Res. Commun. 85, 1510-1517.

Saragosti, S., Moyne, G. & Yaniv, M. (1980) Absence of nucleosomes in a fraction of SV40 chromatin between the origin of replication and the region coding for the late leader RNA. Cell 20, 65-73.

Saragosti, S., Cereghini, S. & Yaniv, M. (1982) Fine structure of the regulatory region of simian virus 40 minichromosomes revealed by DNAse I digestion. J. Mol. Biol. 160, 133-146.

Sassa, S., Granick, J.L., Eisen, H. & Ostertag, W. (1978) Regulation of haem biosynthesis in mouse Friend virus-transformed cells in culture. In "In vitro aspects of erythropoiesis" Peschle, C., Gordon, A.S. & Mirand, E.A. (eds.) N.Y. Springer-Verlag pp.135-142.

Schevitz, R.W., Otwinowski, Z., Joachimiak, A., Lawson, C.L. & Sigler, P.B. (1985) The three-dimensional structure of trp repressor. Nature 317, 782-786.

Schibler, U., Hagenbuchle, O., Wellauer, P.K. & Pittet, A.C. (1983) Two promoters of different strengths control the transcription of the mouse alpha-amylase gene Amy-1a in the parotid gland and the liver. Cell 33, 501-508.

Schoenhaut, D.S. & Curtis, P.J. (1986) Nucleotide sequence of mouse 5-aminolevulinate synthase cDNA and expression of its gene in hepatic and erythroid tissues. Gene 48, 55-63.

Schofield, R. & Dexter, T.M. (1985) Studies on the self-renewal ability of CFU-S which have been serially transferred in long term culture or *in vivo*. Leuk. Res. 9, 305-313.

Scholer, H.R. & Gruss, P. (1984) Specific interaction between enhancer-containing molecules and cellular components. Cell 36, 403-411.

Scholer, H.R. & Gruss, P. (1985) Cell type-specific transcriptional enhancement *in vitro* requires the presence of *trans*-acting factors. EMBO J. 4, 3005-3013.

Sen, R. & Baltimore, D. (1986a) Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell 46, 705-716.

Sen, R. & Baltimore, D. (1986b) Inducibility of k immunoglobulin enhancer-binding protein NF-kB by a post-transcriptional mechanism.

Serfling, E., Jasin, M. & Schaffner, W. (1985) Enhancers and eukaryotic gene transcription. TIG 1, 224-230.

Sevanian, A., Muakkassah-Kelly S.F. & Montestruque, S., (1983) The influence of phospholipase A_2 and glutathione peroxidase on the elimination of membrane lipid peroxides. Arch. Biochem. Biophys. 223, 441-452.

Shadduck, R.K., Waheed, A., Greenberger, J.S. & Dexter, T.M. (1983) Production of colony stimulating factor in long-term bone marrow cultures. J. Cell. Physiol 114, 88-92.

Shatkin, A.J. (1976) Capping of eukaryotic mRNAs. Cell 9, 645-653.

Shaw, G. & Kamen, R. (1986) A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46, 659-667.

Sherman, L., Dafni, N., Lieman-Hurwitz, J. & Groner, Y. (1983) Nucleotide sequence and expression of human chromosome 21-encoded superoxide dismutase RNA. Proc. Natl. Acad. Sci. U.S.A. 80, 5465-5469.

Shermoen, A.W. & Beckendorf, S.W. (1982) A complex of interacting DNAse I - hypersensitive sites near the Drosophila glue protein Sgs-4. Cell 29, 601-607.

Siebert, P.D. & Fukuda, M. (1986a) Isolation and characterization of human glycophorin A cDNA clones by a sythetic oligonucleotide approach : nucleotide sequence and mRNA structure. Proc. Natl. Acad. Sci. U.S.A. 83, 1665-1669.

Siebert, P.D. & Fukuda, M. (1986b) Human glycophorin A and B are encoded by separate, single copy genes coordinately regulated by a tumour-promoting phorbol ester. J. Biol. Chem. 261, 12433-12436.

Siminovitch, L., McCulloch, E.A. & Till J.E. (1963) The distribution of colony forming cells among spleen colonies. J. Cell Comp. Physiol. 62, 327-336.

Sinet, P.M., Michelson, A.M., Bazin, A., Lejeune, J. & Jerome, H. (1975) Increase in glutathione peroxidase activity in erythrocytes from trisomy 21 subjects. Biochem. Biophys. Res. Comm. 67, 910-915.

Singer-Sam, J., Keith, D.H., Tani, K., Simmer, R.L., Shively, L., Lindsay, S., Yoshida, A. & Riggs, A.D. (1984) Sequence of the promoter region of the gene for human X-linked 3-phosphoglycerate kinase. Gene 32, 409-417.

Singh, H., Sen, R., Baltimore, D. & Sharp, P.A. (1986) A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature 319, 154-158.

Sive, H.L., Heintz, N. & Roeder, R.G. (1986) Multiple sequence elements are required for maximal *in vitro* transcription of a human histone H2B gene. Mol. Cell. Biol. 6, 3329-3340.

Southern E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Spandidos, D.A. & Paul, J. (1982) Transfer of human globin genes to erythroleukaemic mouse cells. EMBO J. 1, 15-20.

Speier, C., Baker, S.S. & Newburger, P.E. (1985) Relationships between *in vitro* selenium supply, glutathione peroxidase activity, and phagocytic function in the HL-60 human myeloid cell line. J. Biol. Chem. 260, 8951-8955.

Spooncer, E., Boettiger, D. & Dexter, T.M. (1984) Continuous *in vitro* generation of multipotential stem cell clones from *src*-infected cultures. Nature 310, 228-230.

Spooncer, E., Heyworth, C.M., Dunn, A. & Dexter, T.M. (1986) Self-renewal and differentiation of interleukin-3-dependant multipotent stem cells are modulated by stromal cells and serum factors. Differentiation 31, 111-118.

Sprinzl, M., Hartmann, T., Meissner, F., Moll, J. & Vorderwulbecke, T. (1987) Compilation of tRNA sequences and sequences of tRNA genes. Nucl. Acids Res. 15 (supplement) r53-r188.

Staudt, L.M., Singh, H., Sen, R., Wirth, T., Sharp, P. & Baltimore, D. (1986) A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature 323, 640-643.

Stein, R.W. & Ziff, E.B. (1987) Repression of insulin gene expression by adenovirus type 5 EIA proteins. Mol. Cell. Biol. 7, 1164-1170.

Steiner, C., Muller, M., Baniahmad, A. & Renkawitz, R. (1987) Lysozyme gene activity in chicken macrophages is controlled by positive and negative regulatory elements. Nucl. Acids Res. 15, 4163-4178.

Stone, E.M., Rothblum, K.N., Alevy, M.C., Kuo, T.M. & Schwartz, R.J. (1985) Complete sequence of the chicken glyceraldehyde-3-phosphate dehydrogenase gene. Proc. Natl. Acad. Sci. U.S.A. 82, 1628-1632.

Straus, D. & Gilbert, W. (1985) Genetic engineering in the Precambrian : structure of the chicken triosephosphate isomerase gene. Mol. Cell. Biol. 5, 3497-3506.

Struhl, K. (1986) Constitutive and inducible Saccharomyces cerevisiae promoters : evidence for two distinct molecular mechanisms. Mol. Cell. Biol. 6, 3847-3853.

Struhl, K. (1987) Promoters, activator proteins, and the mechanisms of transcription initiation in yeast. Cell 49, 295-297.

Sunde, R.A. (1984) The biochemistry of selenoproteins. J. Amer. Oil Chem. Soc. 61, 1891-1900.

Sunde, R.A. & Hoekstra, W.G. (1980a) Incorporation of selenium from selenite and selenocystine into glutathione peroxidase in the isolated perfused rat liver. Biochem. Biophys. Res. Commun. 93, 1181-1188.

Sunde, R.A. & Hoekstra, W.G. (1980b) Incorporation of selenium into liver glutathione peroxidase in the Se-adequate and Se-deficient rat. Proc. Soc. Exp. Biol. Med. 165, 291-297.

Sunde, R.A. & Evanson, J.K. (1987) Serine incorporation into the selenocysteine moiety of glutathione peroxidase. J. Biol. Chem. 262, 933-937.

Sweet, R.W., Chao, M.V. & Axel, R. (1982) The structure of the thymidine kinase gene promoter : nuclease hypersensitivity correlates with expression. Cell 31, 347-353.

Szpirer, C. & Szpirer, J. (1975) A mouse hepatoma cell line which secretes several serum proteins including albumin and α-fetoprotein. Differentiation 4, 85-91.

Takahashi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. & Chambon, P. (1986a) Requirement of stereospecific alignments for initiation from the simian virus 40 early promoter. Nature 319, 121-126.

Takahashi, K. & Cohen, H.J. (1986) Selenium-dependant glutathione peroxidase protein and activity : immunological investigations on cellular and plasma enzymes. Blood 68, 640-645.

Takahashi, K., Newburger, P.E. & Cohen, H.J. (1986b) Glutathione peroxidase protein : absence in selenium deficiency states and correlation with enzymatic activity. J. Clin. Invest. 77, 1402-1404.

Tappel, A.L. Hawkes, W.C., Wilhelmsen, E.C. & Motsenbocker, M.A. (1984) Selenocysteine-containing proteins and glutathione peroxidase. Methods Enzymol. 107, 602-619.

Tappel, M.E., Chaudiere, J. & Tappel, A.L. (1982) Glutathione peroxidase activities of animal tissues. Comp. Biochem. Physiol. 73B, 945-949.

Testa, N.G. & Dexter, T.M.(1977) Long term production of erythroid precursor cells (BFU) in bone marrow cultures. Differentiation 9, 193-195.

Theisen, M., Stief, A. & Sippel A.E. (1986) The lysozyme enhancer : cell-specific activation of the chicken lysozyme gene by a far-upstream DNA element. EMBO J. 5, 719-724.

Thiele, B.J., Belkner, J., Andree, H., Rapoport, T.A. & Rapoport, S.M. (1979) Synthesis of non-globin proteins in rabbit-erythroid cells : synthesis of a lipoxygenase in reticulocytes. Eur. J. Biochem. 96, 563-569.

Thiele, B.J., Andree, H., Hohne, M & Rapoport, S.M. (1982) Lipoxygenase mRNA in rabbit reticulocytes : its isolation, characterization and translational repression. Eur. J. Biochem. 129, 133-141.

Thiele, B.J., Fleming, J., Kasturi, K., O'Prey, J., Black, E., Chester, J., Rapoport, S.M. & Harrison, P.R. (1987) Cloning of a rabbit erythroid-cell specific lipoxygenase mRNA. Gene, in press.

Till, J.E. & McCulloch, E.A. (1961) A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Rad. Res. 14, 213-222.

Till, J.E., McCulloch, E.A. & Siminovitch, L. (1964) A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. Proc. Natl. Acad. Sci. U.S.A. 51, 29-36.

Till, J.E. & McCulloch, E.A. (1980) Haemopoietic stem cell differentiation. Biochim. Biophys. Acta 605, 431-459.

Trentin, J.J. (1970) Influence of haemopoietic organ stroma (haemopoietic inductive microenvironment) on stem cell differentiation. In "Regulation of Hematopoiesis, Vol. I." Gordon, A.S. (ed.) N. Y., Appleton-Century-Crofts pp.161-186.

Triesman, R. & Maniatis, T. (1985) Simian virus 40 enhancer increases number of RNA polymerase II molecules on linked DNA. Nature 315, 72-75.

Treisman, R., Green, M.R. & Maniatis, T. (1983) *cis* and *trans* activation of globin gene transcription in transient assays. Proc. Natl. Acad. Sci. U.S.A. 80, 7428-7432.

Tsiftsoglou, A.S. & Robinson, S.H. (1985) Differentiation of leukaemic cell lines : a review focusing on murine erythroleukaemia and human HL-60 cells. Int. J. Cell Cloning 3, 349-366.

Ursini, F., Maiorino, M., Valente, M., Ferri, L. & Gregolin, C. (1982) Purification from pig liver of a protein which protects liposomes and biomembranes from peroxidative degradation and exhibits glutathione peroxidase activity on phosphatidylcholine hydroperoxides. Biochim. Biophys. Acta. 710, 197-211.

Ursini, F., Maiorino, M. & Gregolin, C. (1983) A glutathione peroxidase active on phospholipid hydroperoxides protects phosphatidylcholine liposomes and biomembranes from peroxidation. In "Oxyradicals and their scavenging systems, Vol.II" Greenwald, R.A. & Cohen, F (eds.) Elsevier Science Publ. pp.224-229.

Valerio, D., Duyvesteyn, M.G.C., Dekker, B.M.M., Weeda, G., Berkvens, Th. M., van der Voorn, L., van Ormont, H. and van der Eb, A.J. (1985) Adenosine deaminase : characterization and expression of a gene with a remarkable promoter. EMBO J. 4, 437-443.

Venta, P.J., Montgomery, J.C., Hewett-Emmett, D. & Tashian, R.E. (1985) Comparison of the 5' regions of human and mouse carbonic anhydrase II genes and identification of possible regulatory elements. Biochim. Biopys. Acta 826, 195-201.

Vigneron, M., Barrera-Saldana, H.A., Baty, D., Everett, R.E. & Chambon, P. (1984) Effect of the 21bp repeat upstream element on *in vitro* transcription from the early and late SV40 promoters. EMBO J. 3, 2373-2382.

Villeponteau, B., Lundell, M. & Martinson, H. (1984) Torsional stress promotes the DNAse I sensitivity of active genes. Cell 39, 469-478.

Vincent, A. (1986) TFIIIA and homologous genes. The 'finger' proteins. Nucl. Acids Res. 14, 4385-4391.

Vincent, A., Civelli, O., Buri, J.-F. & Scherrer, K. (1977) Correlation of specific coding sequences with specific proteins associated in untranslated cytoplasmic messenger ribonucleoprotein complexes of duck erythroblasts. FEBS Letts. 77, 281-286.

Vincent, A., Goldenberg, S., Standart, N., Civelli, O., Imaizumi-Scherrer, T., Maundrell, K. & Scherrer, K. (1981) Potential role of mRNP proteins in cytoplasmic control of gene expression in duck erythroblasts. Molec. Biol. Rep. 7, 71-81.

Vincent, A., Akhayat, O., Goldenberg, S. & Scherrer, K. (1983) Differential repression of specific mRNA in erythroblast cytoplasm : a possible role for free mRNP proteins. EMBO J. 2, 1869-1876.

Volloch, V. & Housman, D. (1981) Stability of globin mRNA in terminally differentiating murine erythroleukaemia cells. Cell 23, 509-514.

Voss, S.D., Schlokat, U. & Gruss, P. (1986) The role of enhancers in the regulation of cell-type-specific transcriptional control. T.I.B.S. 11, 287-289.

Wabl, M.R. & Burrows, P.D. (1984) Expression of immunoglobulin heavy chain at a high level in the absence of a proposed immunoglobulin enhancer element in *cis*. Proc. Natl. Acad. Sci. U.S.A. 81, 2452-2455.

Wang, J.C. (1985) DNA Topoisomerases. Ann. Rev. Biochem. 54, 665-697.

Wang, X.-F. & Calame, K. (1986) SV40 enhancer-binding factors are required at the establishment but not the maintenance step of enhancer-dependent transcriptional activation. Cell 47, 241-247.

Wasylyk, B., Wasylyk, C., Augereau, P. & Chambon, P. (1983a) The SV40 72bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. Cell 32, 503-514.

Wasylyk, B., Wasylyk, C., Matthes, H., Wintzerith, M. & Chambon, P. (1983b) Transcription from the SV40 early-early and late-early overlapping promoters in the absence of DNA replication. EMBO J. 2, 1605-1611.

Weaver, R.F. & Weissmann, C. (1979) Mapping of RNA by a modification of the Berk-Sharp procedure : the 5' termini of 15S ß-globin mRNA precursor and mature 10S ß-globin mRNA have identical map co-ordinates. Nucl. Acids Res. 7, 1175-1193.

Weber, F. & Schaffner, W. (1985) Simian virus 40 enhancer increases RNA polymerase density within the linked gene. Nature 315, 75-77.

Weiher, H., Koenig, M. & Gruss, P. (1983) Multiple point mutations affecting the simian virus 40 enhancer. Science 219, 626-631.

Weintraub, H. (1985) Assembly and propagation of repressed and derepressed chromosomal state. Cell 42, 705-711.

Weintraub, H. & Groudine, M. (1976) Chromosomal subunits in active genes have an altered conformation. Science 193, 848-856.

Weintraub, H., Beug, H., Groudine, M. & Graf, T. (1982) Temperature-sensitive changes in the structure of globin chromatin in lines of red cell precursors transformed by ts-AEV. Cell 28, 931-940.

Weisbrod, S. (1982) Active chromatin. Nature 297, 289-295.

Weiss, L. (1984) The blood cells and haematopoietic tissues, 2nd edition. Elsevier, N.Y.

Wetmur, J.G., Bishop, D.F., Cantelmo, C. & Desnick, R.J. (1986) Human 5-aminolevulinate dehydratase : Nucleotide sequence of a full-length cDNA clone. Proc. Natl. Acad. Sci. U.S.A. 83, 7703-7707.

Whetton, A.D., Bazill, G.W. & Dexter, T.M. (1984) Haemopoietic cell growth factor mediates cell survival via its action on glucose transport. EMBO J. 3, 409-413.

Whetton A.D. & Dexter, T.M. (1986) Haemopoietic growth factors. TIBS 11, 207-211.

Wildeman, A.G., Sassone-Corsi, P., Grundstrom, T., Zenke, M. & Chambon, P. (1984) Stimulation of in vitro transcription from the SV40 early promoter by the enhancer involves a specific trans-acting factor. EMBO J. 3, 3129-3133.

Wilhelmsen, E.C., Hawkes, W.C. & Tappel, A.L. (1985) Substitution of selenocysteine for cysteine in a reticulocyte lysate protein synthesis system. Biological Trace Element Research 7, 141-151.

Willmann, T. & Beato, M. (1986) Steroid-free glucocorticoid receptor binds specifically to mouse mammary tumour virus DNA. Nature 324, 688-691.

Wolff, S.P., Garner, A. & Dean, R.T. (1986) Free radicals, lipids and protein degradation. T.I.B.S. 11, 27-31.

Wu, C. (1980) The 5' ends of Drosophila heat shock genes in chromatin are hypersensitive to DNAse I. Nature 286, 854-860.

Wu, C. (1984) Two protein-binding sites in chromatin implicated in the activation of heat-shock genes. Nature 309, 229-234.

Wu, C. & Gilbert, W. (1981) Tissue-specific exposure of chromatin structure at the 5' terminus of the rat preproinsulin II gene. Proc. Natl. Acad. Sci. U.S.A. 78, 1577-1580.

Wyke, J.A., Stoker, A.W., Searle, S., Spooncer, E., Simmons, P. & Dexter, T.M. (1986) Perturbed haemopoiesis and the generation of multipotential stem cell clones in src-infected bone marrow cultures is an indirect or transient effect of the oncogene. Mol. Cell. Biol. 6, 959-963.

Yamao, F., Muto, A., Kawauchi, Y., Iwami, M., Iwagami, S., Azumi, Y. & Osawa, S (1985) UGA is read as tryptophan in Mycoplasma capricolum. Proc. Natl. Acad. Sci. U.S.A. 82, 2306-2309.

Yamomoto, K. (1985) Steroid receptor regulated transcription of specific genes and gene networks. Ann. Rev. Genet. 19, 209-252.

Yang, L., Rowe, T.C., Nelson, E.M. & Liu, L.F. (1985) In vivo mapping of DNA topoisomerase II-specific cleavage sites on SV40 chromatin. Cell 41, 127-132.

Yasuda, M. & Fujita, T. (1977) Effect of lipid peroxidation on phospholipase A2 activity of rat liver mitochondria. Japan J. Pharmacol. 27, 429-435.

Yoshida, M., Iwami, K. & Yasumoto, K. (1982) Purification and immunochemical analysis of rat liver glutathione peroxidase. Agric. Biol. Chem. 46, 41-46.

Yoshimura, S., Komatsu, N. & Watanabe, K. (1980) Purification and immunohistochemical localization of rat liver glutathione peroxidase. Biochim. Biophys. Acta. 621, 130-137.

Young, P.A. & Kaiser, I.I. (1975) Aminoacylation of Escherichia coli cysteine tRNA by selenocysteine. Arch. Biochem. Biophys. 171, 483-489.

Zaret, K.S. & Yamamoto, K.R. (1984) Reversible and persistent changes in chromatin structure accompanying activation of a glucocorticoid-dependant enhancer element. Cell 38, 29-38.

Zaricznyj, C. & Macara, I.G. (1987) Lithium inhibits terminal differentiation of erythroleukaemia cells : evidence for a pre-commitment 'priming' event. Exp. Cell Res. 168, 402-410.

Zenke, M., Grundstrom, T., Matthes, H., Wintzerith, M., Schatz, C., Wildeman, A. & Chambon, P. (1986) Mutliple sequence motifs involved in SV40 enhancer function. EMBO J. 5, 387-397.

Zinn, K. & Maniatis, T. (1986) Detection of factors that interact with the human β-interferon regulatory region in vivo by DNAse I footprinting. Cell 45, 611-618.

Zinoni, F., Birkmann, A., Stadtman, T.C. & Bock, A. (1986) Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 83, 4650-4654.

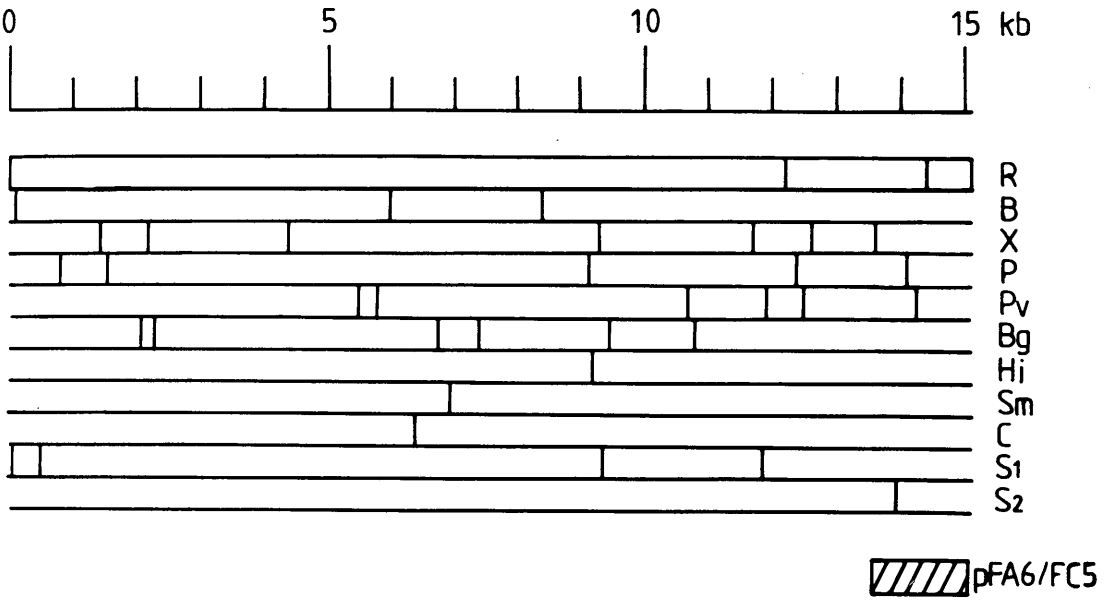


Figure 3.I.1. : Restriction map of λR68A.

The limits of the cloned DNA are defined by two Eco RI sites created by linker addition to a DNA fragment produced by partial Hae III digestion of BALB/c genomic DNA. Vertical lines in each row indicate positions where the DNA is restricted by the corresponding enzyme. Only enzymes which were tested against all the λR68A subclones are shown. No restriction sites for Sal I, Xho I or Kpn I were detected (note: within the 5.9kb Bam HI fragment there are four Xba I sites although only three are shown. The fourth lies about 70 b.p. from one of the others, as yet unidentified). The striped bar indicates the region of hybridization to the cDNAs pFC5 and pFA6 (Goldfarb et al., 1983), both of which are derived from a mRNA encoding a polypeptide of Mr 19kD (Affara et al., 1983).

The abbreviations used for restriction enzymes in this Figure and throughout this thesis are listed below.

A,	Acc I	Av,	Ava I	B,	Bam HI	Bc,	Bcl I	Bg,	Bgl II
Bs,	Bst EII	C,	Cla I	H,	Hae III	Hi,	Hind III	Hp,	Hpa II
K,	Kpn I	M,	Msp I	Na,	Nar I	Nc,	Nco I	P,	Pst I
Pv_1,	Pvu I	Pv,	Pvu II	R,	Eco RI	R_2,	Eco RII	Rs,	Rsa I
S,	Sau 3 A1	Sl,	Sal I	Sm,	Sma I	Sp,	Sph I	St,	Stu I
S_1,	Sst I	S_2,	Sst II	T,	Taq I	X,	Xba I	Xh,	Xho II
Xm,	Xmn I								

An unbound copy of this Figure is enclosed within this thesis.