Regulation of Globin Gene Expression

by

Stephanie Claire Wright

PhD thesis 1985

Laboratory of Gene Structure and Expression
National Institute for Medical Research
Mill Hill, London

and

Dept. Biochemistry,
St. Mary's Hospital Medical School
Paddington, London

To my Parents

ABSTRACT

The human β-like globin gene cluster contains five genes, each of which is expressed in the erythroid cells at a specific stage of development. The ε-globin gene is expressed in the early embryonic yolk sac, the $^{A}\gamma$- and $^{G}\gamma$-globin genes are active in the erythroid cells of the fetal liver, and the δ- and β-globin genes are expressed in the bone marrow of the adult. In order to determine whether the individual genes of the β-like globin gene cluster are independently regulated, and to localize the DNA sequences which regulate β-globin gene activation, cloned human globin genes were introduced by DNA mediated transfer into cultured murine erythroleukemia (MEL) cells. MEL cells are committed adult erythroid cells which are arrested at the proerythroblast stage of differentiation. They may be chemically induced to complete erythroid maturation, whereupon the mouse adult globin genes become activated. The expression of foreign globin genes during differentiation in MEL cells therefore provides a useful system in which to study gene regulation in erythropoiesis.

By introducing cosmids containing different regions of the human β-like globin gene cluster into MEL cells, it was shown that expression of the foreign human adult (β) but not fetal (γ) or embryonic (ε) globin genes was inducible during chemically triggered erythroid differentiation. The human β-globin gene was inducible when introduced as part of the intact β-like globin gene cluster or as a gene fragment containing only 1.5 kb DNA flanking the gene to the 5' and 3' sides. The expression of the human γ-globin gene was simlarly unaffected by the presence of other regions of the human β-like globin gene cluster. The individual genes of the β-like globin gene cluster were therefore regulated differently and independently in MEL cells and the organization of the genes in a cluster did not influence their developmental-stage specific activation in this system; the specific activation in MEL cells of the human β-globin gene mimicked globin gene activation in adult type erythropoiesis.

In order to localize the DNA sequences which regulate β-globin gene activation, a series of rabbit β-globin gene mutants with deletions in the 5' flanking DNA of the gene were introduced into MEL cells. Not more than 58 bp DNA 5' to the transcription initiation site was required for induction of the foreign rabbit β-globin gene in MEL cells. In order to determine whether the regulatory DNA sequences resided 5' or 3' to the translation initiation codon of the β-globin gene, hybrid genes were constructed between the inducible human β-globin gene and either the noninducible human γ-globin gene or the noninducible mouse major histocompatibility $H\text{-}2K^{bm1}$ gene. By analysing the induction of these hybrid genes during differentiation in MEL cell transformants it was shown that the DNA sequences which regulate β-globin gene activation are located both 5' and 3' to the translation initiation site.

Acknowledgements

I would like to thank Dick Flavell, Frank Grosveld, Dimitris Kioussis, Arnon Rosenthal and Bob Williamson for many stimulating discussions. Frank Grosveld, Gerard Grosveld, Dimitris Kioussis and Arnon Rosenthal kindly provided clones as indicated in the text.

This work was supported by a research studentship from the British Medical Research Council.

Abbreviations

DNA	deoxyribonucleic acid
RNA	ribonucleic acid
mRNA	messenger RNA
dNTP	deoxynucleoside 5'-triphoshate
dATP	deoxyadenosine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphoshate
dCTP	deoxycytosine 5'-triphosphate
TTP	thymidine 5'-triphoshate
GTP	guanosine 5'-triphoshate
CTP	cytosine 5'-triphoshate
ATP	adenosine 5'-triphoshate
UTP	uridine 5'-triphoshate
ss	single-stranded
ds	double-stranded
bp	base-pair
kb	kilobasepair
nt	nucleotide
O.D.	optical density
rpm	revolutions per minute
BPV	bovine papilloma virus
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
DMSO	dimethylsulphoxide
DNaseI	deoxyribonucleaseI
EDTA	ethylenediaminetetra-acetic acid
FCS	fetal calf serum
HCl	hydrochloric acid
HEPES	N-2 hydroxyethylpiperazine N'-2-ethanesulfonic acid
HMBA	hexamethylenebisacetamide
HSV	Herpes Simplex Virus
IVS	intervening sequence
KAc	potassium acetate
KCl	potassium chloride

K_2HPO_4	dipotassium hydrogen orthophospate
KH_2PO_4	potassium dihydrogen orthophoshate
LTR	Long Terminal Repeat
MEL	murine erythroleukemia
$MgCl_2$	magnesium chloride
$MgSO_4$	magnesium sulphate
MMTV	mouse mammary tumour virus
MTX	metallothionen
NaAc	sodium acetate
Na_3 citrate	sodium citrate
NaCl	sodium chloride
Na_2HPO_4	disodium hydrogen orthophoshate
NaH_2PO_4	sodium dihydrogen orthophoshate
NaOH	sodium hydroxide
NH_4Ac	ammonium acetate
PCI	phenol:chloroform:isoamylalcohol (24:24:1,v:v:v)
PMSF	phenylmethylsulfonyl fluoride
RNase	ribonuclease
SDS	sodium dodecyl sulphate
TK	thymidine kinase
Tris	tris(hydroxymethyl)aminomethane

CONTENTS

Page No

CHAPTER 2 MATERIALS AND METHODS

Page No

Page No

FIGURES

Figure No — Page No

TABLES

Table No — Page No

CHAPTER 1

INTRODUCTION

Globin gene expression is an ideal system in which to study tissue specific and developmentally regulated gene expression. The globin genes encode the α-like and β-like protein subunits of the oxygen binding compound hemoglobin, which constitutes approximately 90% of the soluble protein in the circulating red blood cells of vertebrates. Due to the high levels of globin mRNA in the reticulocytes, the globin genes were among the first eukaryotic genes to be cloned, and are probably now the most well characterized genes with respect to structural organization, expression and evolution. ~~evolution~~. In addition, the analysis of a large number of naturally occurring mutant globin genes and the ability to study the transcription of *in vitro* mutated cloned genes in eukaryotic cell expression systems has led to the identification of the DNA sequences which regulate globin gene transcription and RNA splicing.

The globin genes become activated in the later stages of erythroid differentiation. The structure of both the α-like and β-like globin chains produced in the erythroid cell change at several stages during the development of the animal; these different developmental stage specific globin genes are found together in the respective α-like and β-like globin gene clusters. Probably the most important unanswered question in globin molecular biology is what is the mechanism that controls the activation and inactivation of the individual genes of the globin clusters at different stages of

development. Experiments described in this thesis approach this problem by studying the regulation of expression of cloned human globin genes introduced into a mouse erythroleukemia cell line which can be induced to undergo erythroid differentiation in culture.

Sections 1.1 and 1.2 of the introduction describe the regulation of erythroid differentiation in the normal animal and in the viral transformed murine erythroleukemia cell line. The developmental switches in globin gene expression are discussed in section 1.3. Section 1.4 describes the structure and transcription of the globin genes, and section 1.5 describes the methods for studying the transcription of cloned genes in eukaryotic cell expression systems.

1.1 HEMATOPOIETIC DIFFERENTIATION

i The hematopoietic stem cell

All types of blood cell are derived from a common pluripotent hematopoietic stem cell, which can either proliferate and self-renew, or differentiate along alternate pathways to give rise ultimately to the erythrocytes, platelets, granulocytes, macrophages and lymphocytes (Figure 1; see Till and McCulloch, 1980 for review). Differentiation takes place via a series of intermediate "committed" precurser cells which progressively lose the capacity to proliferate and become more restricted in the type of blood cell to which they may give rise.

The pathways of hematopoietic differentiation have been

characterized in the bone marrow, which is the major site of blood cell formation in all adult vertebrates and contains the pluripotent hematopoietic stem cells together with the different committed precursor cells. The stem cells represent only a minor subpopulation of the cells in hematopoietic tissue, and were originally functionally identified using the "spleen colony formation assay" devised by Till and Mc Culloch (1961). In these assays, bone marrow cells were injected into a lethally irradiated host mouse (i.e. one whose hematopoietic system had been destroyed) and then shown to colonize the spleen so as to produce nodules containing all types of blood cell. The clonal origin of the nodules was confirmed by injecting irradiated bone marrow cells which had chromosome abnormalities, and the founder cells of the clones were presumed to be the pluripotent hematopoietic stem cells. The proliferation and differentiation of hematopoietic stem cells may be studied in long term bone marrow cultures in which hematopoiesis occurs within a bone-marrow derived adherent multilayer consisting of endothelial, fibroblastic, fat, and macrophage cells (see Dexter, 1982 for review). Hematopoiesis in these cultures is influenced by the stromal cells both via direct cell-cell contact and by the action of secreted inhibitory and stimulatory molecules. Preferential differentiation of the hematopoietic cells along most of the different lineages may be achieved by using specific culture conditions; this is thought to operate via facilitation of the proliferation and differentiation of particular committed precursor cells rather than by causing the preferential committment of the pluripotent stem cells to a given lineage. The way in which the determination of the pluripotent stem

cells to one or other of the hematopoietic differentiation lineages is controlled remains largely unclear and is generally explained as a purely stochastic process (Korn et al., 1973).

ii Erythroid differentiation

The later stages of hematopoietic differentiation have been elucidated by analysing the colonies formed from bone marrow cells after their dispersal in culture. The use of specific culture conditions favours the proliferation and differentiation of a given type of hematopoietic precurser cell, thus giving rise to pure colonies of erythroblasts, macrophages, or one of the other blood cell types (see Metcalf, 1977 for review). These studies lead to the identification of specific hematopoietic precurser cells and to the discovery of the obligatory role of glycoprotein factors in erythroid differentiation. The formation of erythroblast clones in such assays required the presence of the glycoprotein hormone erythropoietin, and by dispersing the bone marrow cells in different concentrations of the hormone, a series of erythroid precurser cells with differing sensitivity to erythropoietin were discovered (Figure 1; Eliason et al., 1979; Iscove et al., 1974, 1980). Thus, the CFU-E (erythrocyte colony forming unit) is a late committed erythroid precurser cell; it is very sensitive to erythropoietin, has a limited proliferation capacity, and gives rise to erythroblasts after 6 cell divisions or less (Iscove et al., 1974, 1980). Dispersal of bone marrow cells in low concentrations of erythropoietin results in proliferation of these CFU-Es to yield erythroblast colonies containing less than 60 cells after 2 days. In contrast, the BFU-Es (erythrocyte burst

forming unit) are an earlier erythroid precurser cell; they have a higher proliferative capacity than CFU-Es and are less sensitive to erythropoietin (Axelrad et al., 1973). Culturing of dispersed bone marrow cells in high concentrations of erythropoietin leads to the proliferation and differentiation of these BFU-Es and gives rise colonies containing up to 5000 cells after 9 days. Such in vitro studies therefore explained the mechanism whereby erythrocyte production is stimulated in the animal by release of erythropoietin from the kidney under conditions of erythrocyte shortage. Proliferation of the BFU-Es is also controlled by the factor BPA (Johnson and Metcalf, 1977); during the transition from BFU-E to CFU-E, the erythroid precursers become progressively less sensitive to BPA and more sensitive to erythropoitein (Figure 1). Similar in vitro studies have defined a role for glycoprotein factors in the regulation of other hematopoietic differentiation lineages. For example, the GM-CSF factor, which has been recently cloned (Gough et al., 1984), promotes granulocyte and macrophage formation (Burgess et al., 1977). Factors which control different differentiation pathways do not compete for the same target cell; they act on separate committed precursor cells rather than on the pluripotent stem cells (Metcalf and Johnson, 1979).

Erythropoietin stimulates the differentiation of the early erythroid precurser cells through to the proerythroblast stage; further maturation of the proerythroblast to erythroblast results in a decrease in both proliferative capacity and sensitivity to erythropoietin, and during this transition many of the characteristic erythroid markers become apparent (Figure 1; section 1.2ii). The

globin genes become transcriptionally activated and levels of globin mRNA reach more than 25,000 molecules per cell by the stage of the orthochromatic erythroblast. Most of the polyA$^+$ RNA in the enucleate reticulocyte is globin mRNA, and hemoglobin is the major protein in the mature erythrocyte. The accumulation of globin mRNA results from an increase in both the transcription of the globin genes and in the relative stability of globin mRNA (discussed further in section 1.2ii). These terminal stages of erythroid maturation therefore constitute an ideal system in which to study tissue specific gene activation.

1.2 MURINE ERYTHROLEUKEMIA CELLS

Murine erythroleukemia (MEL) cells are Friend virus transformed erythroid cells which are arrested at the proerythroblast stage of differentiation (for review see Harrison, 1977). Cultured MEL cells may be induced to complete erythroid maturation by treatment with a variety of chemicals such as dimethylsulphoxide (DMSO) or hexamethylenebisacetamide (HMBA). This chemicallly triggered maturation closely resembles normal erythroid differentiation and results in an accumulation of mouse adult-type globin mRNA in the differentiated cell. Chemically induced differentiation of MEL cells is therefore a useful model system in which to study gene regulation during the later stages of erythroid differentiation. The more recent isolation of cell lines arrested at the BFU-E and CFU-E stages (Shibuya and Mak, 1983) will similarly enable analysis of these

earlier stages of erythroid differentiation.

i Isolation and characteristics of MEL cells

Friend erythroleukemia virus (FV-A) was first identified in cell free extracts of Ehrlich ascites carcinoma cells (Friend 1957). Passage of the extracts in certain strains of mice produced an anemia and accumulation of erythroid precursers in the spleen, resulting in erythroblastosis in the peripheral circulation and culmination of the disease in a fatal leukemia. Subsequently, variants of the virus were isolated which produced a leukemia in association with polycythemia (FV-P virus; Mirand et al., 1961).

Several distinct phases in the course of Friend virus induced erythroleukemia have been identified. The erythroid precursers proliferating in the spleen in the early stages of the disease continue to differentiate, show no obvious chromosome abnormalities, and are not malignant (Majumdar and Bikender, 1975; Matioli, 1973). Later, the proliferating cells become blocked at the proerythroblast stage of differentiation, and tumor cells showing chromosome abnormalities appear in the spleen, peripheral blood, bone marrow and liver. At this stage the erythroid cells are capable of unlimited growth outside the normal hematopoietic environment, and subcutaneous tumours may be formed from transplants of infected spleen (Friend and Haddad, 1960). Cell lines may be established either from fragments of late-stage infected spleen or from the transplanted subcutaneous tumors (Friend et al., 1966; Dube et al., 1973); growth of these cultured MEL cells is independent of the normally essential stromal cell environment (section 1.1i).

The mechanism whereby Friend virus causes proliferation of erythroid precursers has been elucidated both by studying the *in vitro* proliferation and differentiation of hematopoietic precursor cells taken from the spleens of virally infected mice, and by studying the *in vitro* infection of hematopoietic cell populations by Friend virus. Such analysis has shown that the FV-A and FV-P viruses cause erythroproliferation by different mechanisms; FV-P acts by allowing erythroid precursors (probably late BFU-Es and CFU-Es) to proliferate independently of erythropoietin (Liao and Axelrad, 1975; Hankins and Krantz, 1975), whereas FV-A results in abnormal growth regulation of the early BFU-E compartment, with the CFU-Es remaining under erythropoietin control (Fagg *et al*., 1980). The virus thus affects the committed erythroid cells rather than the stem cell per se; changes in other hematopoietic compartments merely result from indirect feedback responses due to changes in numbers of erythroid precursors, the decreased lifetime of mature blood cells because of the viral infection, and the immune response directed against virus infected cells. Tumour colonies arise as rare events in the proliferating cell population; these final erythroleukemic cells are thought to derive directly from the proliferative cells stimulated early in the course of the disease rather than independently.

The way in which Friend erythroleukemia viruses cause growth factor independent erythroproliferation is still unknown. Friend virus is an RNA tumor virus comprising a defective replication incompetent component and a helper replication competent one (Bernstein *et al*., 1977). The defective virus has been implicated directly in the proliferation and malignant transformation, although

the involvment of a specific viral gene has not been established. Several well defined mouse genes affect both the susceptibility to Friend leukemia virus and the types of hematopoietic modulation caused. Some of these genes have also been shown to influence normal hemopoiesis, e.g. the Fv-2 locus was originally discovered as affecting susceptibility to FV-A virus and has since been shown to control the release from the stromal cells of a factor which regulates BFU-E cycling. There is thus a close relationship between the control of normal- and Friend virus induced- erythropoiesis.

ii Differentiation of MEL cells

Cultured proerythroblast MEL cells can be induced into the terminal stages of erythroid maturation by treatment with chemicals such as dimethylsulphoxide (DMSO; Friend et al., 1971) and hexamethylenebis-acetamide (HMBA; Reuben et al., 1976). The process of MEL differentiation closely resembles the differentiation of normal proerythroblasts to orthochromatic erythroblasts as judged by the appearance of characteristic erythroid markers e.g. spectrin, glycophorin, heme synthetic enzymes, globin mRNA, and by morphological criteria, e.g. chromatin condensation (Ross et al., 1972; Nudel et al., 1977). Differentiation to the enucleate reticulocyte may be achieved by using improved culture conditions which stabilize the cells (Volloch and Housman, 1981). The isolation of MEL mutants that are inducible for early erythroid markers e.g. spectrin, but not for late characteristics, e.g. hemoglobin production and terminal differentiation, is enabling erythroid differentiation to be dissected into component processes (Eisen et

al., 1978).

The increase in the levels of globin mRNA observed upon differentiation of MEL cells is mainly a result of increased transcription of the globin genes (Orkin and Swerdlow, 1977), but is probably also due to an increase in the relative stability of globin mRNA. This change in relative mRNA stability is due to a decrease in stability of non-globin mRNAs upon differentiation; the half-life of globin mRNA itself (> 60 hours) does not change (Aviv et al., 1976; Volloch et al., 1981). Most MEL cell lines show an accumulation of adult but not embryonic type globins upon differentiation (Nudel et al., 1977); this is thought to reflect the "adult" nature of the erythroid environment in the MEL cells which were derived from an adult mouse. Interestingly, different inducers will preferentially activate one or other of the adult type mouse β-globin genes (see Figure 3). For example, DMSO induction results in preferential induction of the β-major versus the β-minor globin gene, whereas hemin induction results in preferential activation of the β-minor gene (Nudel et al., 1977; Curtis et al., 1980). The wide variety of chemical inducers now known do not share common chemical properties, and the isolation of MEL mutants by virtue of their resistance to certain inducers has shown that they do not all act by a common pathway (Harrison, 1977). Inducers have been categorized into two classes on the basis of whether they act via a mechanism involving polyamine synthesis, although the mechanism whereby they trigger the differentiation is unclear (Gazitt and Friend, 1980).

1.3 DEVELOPMENTAL SWITCHES IN GLOBIN GENE EXPRESSION

i Changes in site of erythropoiesis

Activation of the globin genes takes place in the later stages of erythropoiesis during the transition of the proerythroblast to erythroblast (section 1.1ii; Figure 1). Both the site of erythropoiesis and the specific types of globin chain produced in the erythroid cell change during development of the animal in all vertebrate species (see Wood and Weatherall, 1983 for review). Erythropoiesis in mammals and birds begins in the yolk sac derived erythrocytes and then passes to the developing fetal liver; this switch occurs at about 6 weeks gestation in man. Blood cell formation in the fetus also takes place to a lesser extent in the spleen and bone marrow. At 32-36 weeks gestation, erythropoiesis switches to the bone marrow, which remains the major site of blood cell formation throughout adult life. In mammals, these developmental changes are thought to result from migration of hematopoietic stem cells from one organ to another, rather than being due to *de novo* differentiation events in the developing tissues (Moore and Metcalf, 1970; Johnson and Moore, 1975). In birds, however, there is evidence for an intraembryonic origin of post yolk-sac hemopoiesis (Martin *et al.*, 1980; Dieterlen-Lievre and Martin, 1981). The mechanism by which the hematopoietic cell migration is triggerred and controlled remains unknown.

ii Developmental changes in globin structure

Hemoglobin proteins are tetramers comprising 2 α-like and 2 β-like subunits; the types of both α-like and β-like subunits produced in the erythroid cell change at several stages of development. In man, there are ~~two~~ one developmental switches in α-like chain synthesis and ~~2~~ 1 switches in β-like chain synthesis (Figure 2). The developmental stage specific types of β-like subunit are the ε (embryonic), γ (fetal), and δ and β (adult) chains; the α-like subunits are the ζ (embryonic) and α (adult) chains. In most other species, there are only two developmental types (embryonic and adult) of both α-like and β-like chains. The earliest embryonic hemoglobin in man has the structure $\zeta_2\varepsilon_2$ (Gale et al., 1979), and after about 6 weeks gestation, this is replaced by the fetal hemoglobin HbF ($\alpha_2\gamma_2$; Schroeder et al., 1968, 1972). During the embryonic to fetal transition period, $\alpha_2\varepsilon_2$ and $\zeta_2\gamma_2$ are also found. HbF is the predominant hemoglobin throughout fetal life and is replaced just before birth by HbA ($\alpha_2\beta_2$) and HbA_2 ($\alpha_2\delta_2$). Throughout adult life, 97% of hemoglobin is HbA, 2% is HbA_2, and 1% is HbF. In man, the α-globin chains are the products of 2 non-allelic genes, α1 and α2, whose protein products have the same amino-acid sequence; the α1/α2 ratio is 1.5-3:1 and does not change during development (Liebhaber and Kan, 1981; Orkin and Goff, 1981). The γ-globin chains are similarly the products of the non-allelic $^{A}\gamma$- and $^{G}\gamma$-globin genes, whose protein products differ by only 1 amino-acid; the $^{G}\gamma:{}^{A}\gamma$ ratio in the fetus is 3:1 (Nute et al., 1973). During development, the synthesis of α-like and β-like subunits has to be coordinated, since an excess of one type of chain

results in their precipitation in the erythrocyte; this is probably achieved by regulation at both the transcriptional and translational level.

All of the HbF found in adult man is confined to the F cells, which constitute 3% of mature erythroid cells (Wood et al, 1975). Increased levels of HbF in adult life may result from genetic defects in the β-globin gene cluster or from a response to a wide variety of conditions which are associated with acute erythroid expansion (Dover et al., 1979) or with a deficit in functioning stem cells or erythroid progenitors (Papayannopoulou et al., 1980). In these latter cases, an acute demand on the precurser population or a reduced population results in alteration of the normal proliferation and maturation pattern of red cells, which then synthesise HbF (Papayannopoulou et al., 1977; Stamatoyannopoulos and Papayannopoulou, 1978). In some diseases e.g. β-thalassemic homozygotes (section 1.4viii) an imbalance in the α/β chain ratio results in destruction of the normal blood cells due to precipitation of the excess globin chains; this may lead to an elevation in HbF levels due to the selective survival of F cells (Weatherall et al., 1976).

The mechanisms which regulate the developmental switches in globin gene expression have been studied 1) by analysing the structure of globin DNA and chromatin at different stages of development 2) by analysing the expression of cloned globin genes introduced into cultured erythroid cells (section 1.6) 3) by determining the nature of the lesion at the molecular level in genetic disorders of globin gene switching (e.g. the thalassemias and

HPFH, section 1.4viii) 4) by transplantation of fetal erythroid cells into an adult environment and vice versa and 5) by the clonal culture and differentation of erythroid precurser cells in vitro.

The fact that the switch from γ-globin to β-globin synthesis occurs simultaneously in all fetal hematopoietic tissues (spleen, bone marrow and liver) suggests that it is not controlled by the site of erythropoiesis (Wood and Weatherall, 1973); it has also been demonstrated that the switch is not influenced by severe endocrine perturbations (Wood et al., 1978). In order to distinguish between autonomous programming of hematopoietic cells and the influence of extracellular inducers, fetal hematopoietic cells were transplanted into irradiated adult sheep; after transplantation, the fetal cells switched to adult globin synthesis, although the time at which they did so depended on their gestational age, suggesting that the pattern of hemoglobin production was intrinsic to the hematopoietic cells but was modulated by extracellular influences of the adult environment (Zanjani et al., 1979; Bunch et al., 1981). In the converse experiment, adult cells transplanted into a fetus did not make fetal hemoglobins or fetal antigens, suggesting that they had already become irreversibly programmed and were not influenced by the fetal environment (Zanjani et al.,1982; Blanchet et al., 1982; section 1.4vi,vii).

Studies of the clonal culture of erythroid precurser cells in vitro have shown that the switch from fetal to adult globin gene expression occurs within single cells rather than via the selection of cells already programmed to make one or other type of globin. In these cultures, erythroid stem cells proliferate and differentiate to

form single-cell origin colonies derived from erythroid precursor cells at various stages of differentiation (usually CFU-Es and BFU-Es; section 1.1ii). The ratio of adult/fetal hemoglobin synthesis in these colonies reflected the developmental stage from which the precurser cells were taken, e.g. fetal erythroid progenitors differentiated in culture to produce colonies in which >95% of the globin synthesized was HbF. Cells taken from cord blood during the fetal-adult switch produced equal amounts of HbA and HbF (Kidoguchi et al., 1978, Stamatoyannopoulos et al., 1979). The HbA/HbF ratio in individual colonies derived from single precursor cells covered a wide range (Kidoguchi et al., 1978; Comi et al., 1980), indicating that the modulation of globin synthesis during the switch occurs within single precursor cells and not via selection of clones already programmed to make only HbA or HbF. This is consistent with the observation that HbF and HbA are found in the same cell in the cord blood during the switch. In vitro culture of erythroid precurser cells has also been used to isolate factors which might mediate globin switching, and an activity which accelerates the HbF to HbA transition in neonatal human BFU-Es has been identified in fetal sheep serum. This factor also effected the HbF to HbA switch in erythroid precurser cells from a patient with Hereditary Persistance of Fetal Hemoglobin patient (in this disease, γ-globin gene expression persists in vivo throughout adult life; section 1.4viii). This might indicate that the defect in this HPFH patient was one of interaction of the γ-globin gene with trans acting repressors, with this defect being overcome by high levels of the factor in the in vitro assay (Papyannopoulou et al., 1982, 1984).

1.4 GLOBIN GENE STRUCTURE AND EXPRESSION

The globin genes were among the first eukaryotic genes to be cloned and characterized. Early structural analysis of these genes showed firstly that they were interrupted by introns and secondly that the globin genes were organized in gene clusters, each of which contained several different developmental stage specific genes. More extensive characterization of the structure and expression of globin genes from a wide variety of species and from patients with inherited disorders in globin gene expression is now providing information about the evolution of these gene families and about the regulation of globin gene expression.

i Globin gene linkage

The first evidence suggesting that the human β-like globin genes were linked was obtained from structural analysis of hemoglobin variant proteins which comprised hybrid globin chains e.g. the Hb Lepore protein contained the N terminal sequences of the δ-globin chain joined to the C terminal sequences of the β-globin chain (Baglioni, 1962), and likewise Hb Kenya contained the N-terminal sequences of the $^{A}\gamma$-globin chain fused to the C-terminal segment of the β-chain (Huisman et al., 1972). Genetic studies of families with these variant proteins suggested that the $^{A}\gamma$, δ and β-globin genes were closely linked, and this was later confirmed by genomic blotting (Flavell et al., 1978; Little et al., 1979), and by the

isolation of clones containing overlapping regions of the β-like globin gene cluster from genomic libraries (Ramirez et al., 1979; Fritsch et al., 1979, 1980; Lawn et al., 1978). The human β-like globin gene cluster is organized 5'-ε-$^{G}\gamma$-$^{A}\gamma$-δ-β-3', with all the genes being transcribed from the same DNA strand (Figure 3). The genes are arranged along the chromosome in the order of their expression during development, and those genes which are expressed together at a given developmental stage (e.g. δ and β; $^{G}\gamma$ and $^{A}\gamma$) are found in pairs. A similar pattern of gene linkage is also found in the human α-like globin gene cluster (Figure 3; Lauer et al., 1980). The human α-like and β-like gene clusters have been shown by cell fusion studies to be located on chromosomes 16 and 11 respectively (Deisseroth et al., 1977, 1978). The human β-like gene cluster has more recently been assigned to the short arm of chromosome 11, lying between the parathyroid hormone and insulin loci (Antonarakis et al., 1983).

The globin genes of other species are similarly arranged in clusters, although the organization of the different developmental stage specific genes within the cluster varies (Figure 3). In amphibians, the α-like and β-like globin genes are syntenic, with the larval genes flanking the adult ones (Hosbach et al., 1983). In birds, the α-like and β-like clusters are no longer linked, but within the β-like cluster, the embryonic genes still flank the adult and neonatal ones (Engel and Dodgson, 1983; Dolan et al., 1981). In rodents and primates, the arrangment of the genes tends to match the order in which they are expressed during development, although the sizes of the cluster vary considerably among species e.g. the brown

lemur has only one β-like globin gene for each of the embryonic, fetal, and adult stages of development, and the entire cluster is contained within 25kb, as compared with the 50kb β-like globin gene cluster in man (Barrie et al., 1981). Globin gene organization in goats is quite different, with the gene cluster having arisen as a result of triplication of a basic group of four genes; the developmental patterns of globin production are, however, similar to those found in primates (Lingrel et al., 1983; Schon et al., 1981; Townes et al., 1984). Interestingly, genes which are structurally homologous between species may be expressed at different stages of development e.g. the rabbit β3 gene codes an embryonic globin chain, but is structurally homologous to the human γ-globin genes which are expressed in the post-embryonic stages (Hardison, 1981).

The intergenic regions of the globin gene clusters are large compared with the size of the genes themselves, and much of this intergenic DNA is repetitive (Shen and Maniatis, 1980). These repeat structures have been mapped in the human and rabbit loci; the position of some repeats have been conserved through evolution, while others are specific to only one species. The function of repeat sequences in the genome is unclear; previous data have described their genomic distribution and differential transcription during development (see Davidson and Britten, 1979 for review), and it has been suggested that they may play a role in controlling differential gene expression, DNA replication or chromatin structure. The presence of repeats within the globin cluster will, however, enable their study in relation to a well-defined set of developmentally regulated genes.

The organization of eukaryotic genes into clusters is now recognised as a widespread phenomenon, and there are many eukaryotic gene families in which several structurally and functionally related genes are closely linked e.g. the alphafetoprotein and albumin genes in mouse (Ingram et al., 1981), the Major Histocompatibility Gene Complex in mouse and man (Steinmetz et al., 1982). Such gene families are thought to have arisen from a single primordial gene via gene duplication, with divergence of the duplicated sequences yielding gene products of different function. The evolutionary relationships between the genes of the β-like globin cluster in man have been calculated by comparing the nucleotide sequence divergence between the coding sequences of the genes (Efstradiatis et al., 1980). The α- and β-globin genes are thought to have duplicated from an ancestral gene 500 million year ago i.e. at the time in evolution between the appearance of the amphibians and birds. Subsequently, the two gene clusters evolved independently by gene duplication, correction and divergence. More recently, the different developmental stage specific genes of the α- and β-like gene clusters evolved. Interestingly, the embryonic α- and β-like globin genes diverged from their adult counterparts at different times, with the α-cluster divergence occuring 400 million years ago, and that of the β-cluster being 200 million years ago (Proudfoot et al., 1982). Thus, although the temporal expression of the ζ- and ε-globin genes is coordinately controlled in development, their evolutionary histories are distinct, and the mechanisms which control gene switching within the two gene clusters may be different. The fetal-embryonic divergence within the β-globin gene cluster took place about 100

million years ago, i.e. at the time of mammalian radiation.

ii Globin gene fine structure

The development of techniques for genomic blotting and the subsequent comparison of restriction enzyme maps of the globin genomic and cDNA led to the discovery of introns within the mouse and rabbit globin genes (Jeffreys and Flavell, 1977; Tilghman et al., 1978). This was later confirmed by DNA sequence analysis of cloned chromosomal globin genes. Introns have since been found in a wide variety of eukaryotic genes, including plant, organelle, and viral DNA. Eukaryotic genes that are not split are relatively rare, with well documented examples of intronless genes including the interferon genes (Nagata et al., 1980) and most of the histone genes (Kedes, 1979; Hentschel and Birnstiel, 1981). Nearly all the globin genes from every species examined to date possess two introns; these introns are always found in the same relative position although they are variable in length (Figure 4). All the α-like and β-like globin genes have a small intron of 100-190 bp between codons 30 and 31 (Lawn et al., 1980). The β-like globin genes have a large intron of 573-930 bp between codons 104 and 105, whereas the α-like genes have a smaller large intron of 140-339 bp which lies between codons 99 and 100. Exceptionally long introns (886 and 239 bp) are found in the human α-like ζ-globin gene (Proudfoot et al., 1982). Similarly, the myoglobin gene, which encodes a monomeric globin protein responsible for oxygen binding in the muscle, has exceptionally long introns in both seal (4.8kb and 3.4kb; Blanchetot et al., 1983) and man; the myoglobin gene is thought to have diverged from the hemoglobins about

600-800 million years ago. The structure of the soybean leghemoglobin gene, which encodes a monomeric globin (Ostergaard-Jensen et al., 1981), is also different from that of other globin genes in that it has an "extra" third intron between codons 68 and 69. DNA sequence analysis has shown that the globin gene introns diverged much more rapidly than the coding sequences due to the occurence of base substitutions and insertions/deletions. A similar high rate of intron sequence divergence has been found in other gene families e.g. the ovalbumin genes in chicken (Breathnach and Chambon, 1981) and the vitellogenin genes in X.laevis (Wahli et al., 1980).

The evolution and function of introns has been the subject of much speculation. As in many systems, the positions of the introns within the globin genes correlate with discontinuities in the protein molecule that mark the boundaries of functional domains (Gilbert, 1978). Thus, the central exon of the globin gene encodes the region of the protein which makes most of the heme contacts, while the 3rd exon codes the region which makes most of the α-β subunit contacts (Craik et al., 1980). Protein sequence analysis of the globins suggested that the central domain might be constructed from two subdomains; the position of this boundary correlates with the position of the 3rd intron in the leghemoglobin gene. Similar exon-domain correlations are found in many systems e.g. the Ig subdomains are encoded on separate exons (Sakano et al., 1979). This has suggested that the intron-exon structure might have had an important role in evolution by allowing the DNA sequences coding protein subdomains to be joined to make proteins with novel properties. It should be noted that some of the exon-domain

correlations (e.g. the α-β contact residues specified by exons 1 and 3, and the Bohr effect residues specified by exon 3) relate only to the tetrameric hemoglobins, and must have evolved long after the basic exon-intron structure of the globin genes, which is also found in the monomeric globins, had been established (Blanchetot et al., 1983).

Globin gene introns may also have played a role in evolution by allowing gene correction mechanisms to prevent gene loss by mutational inactivation. The 5' two-thirds of the large intervening sequence of the $^{A}\gamma$-globin gene of the 2 allelic chromosomes show much less sequence homology than do the $^{A}\gamma$- and $^{G}\gamma$-globin genes from the same chromosome (Slightom et al., 1980), suggesting that the intron sequence conservation was due to intrachromosomal recombination which was possibly mediated via a $(TG)_n$ stretch in the intron. A similar phenomenon has also been observed for the two nonallelic α-globin genes in man.

A number of nonfunctional genes (pseudogenes) which are structurally related to a specific functional globin gene have been found within the globin gene clusters of many species (Figure 3). They were detected by genomic blot hybridization and had previously been undetected because of their lack of production of functional protein. Each of the pseudogenes shows 70-80% sequence homology with its corresponding normal gene, and may be defective for a number of different reasons. For example, the human pseudo α1 globin gene is homologous to the adult α-gene but has a 20 nucleotide deletion in the middle of the gene that puts the coding sequence out of phase (Proudfoot and Maniatis, 1980); pseudo ζ is homologous to the

embryonic ζ α-like globin gene but has a termination codon at aminoacid 6 (Proudfoot et al., 1982). Some pseudogenes have the properties of a "processed" gene, i.e. they have precisely lost their introns and have ~~the~~ acquired a characteristic polyA sequence (which is added post-transcriptionally to mRNAs) in the 3' flanking region of the gene. Thus the mouse α-globin pseudogene is intronless, has a poly A sequence in its 3' flanking region, and is found on a different chromosome from the mouse α-globin gene cluster (Nishioka et al., 1980; Vanin et al., 1980); "processed" genes have also been found in other gene families e.g. tubulin, immunoglobulin (Hollis et al., 1982; Wilde et al, 1982). It has been suggested that such processed genes arose by reverse transcription of an intronless and polyadenylated mRNA copy of the gene; this might occur via insertion of the gene into a retroviral genome, which replicates via an RNA intermediate, thus leading to the loss of introns from any gene it contained. The feasibility of the retroviral insertion model has been suggested by cloning the functional mouse α-globin gene into a retroviral vector and finding an intronless globin gene in the progeny virus (Shimotohno and Temin, 1982).

iii The globin gene transcription unit

The primary transcript of the globin gene was originally shown by pulse labelling of nuclear RNA transcripts to be colinear with the gene and to contain both introns (Curtis and Weissman, 1976). The 5' end of this earliest transcript coincided with the first capped nucleotide of the mature mRNA, suggesting that transcription initiation was at the position corresponding to this nucleotide. SI

nuclease mapping of RNA isolated from rabbit bone marrow has shown that the precursor RNAs are spliced via the stepwise removal of both introns, with poly A addition preceding RNA splicing (G.Grosveld et al., 1981a). The precise position of transcription termination in globin and other eukaryotic genes which yield polyadenylated transcripts has not been well characterized, although transcription of the mouse β-major globin gene has been shown to terminate at a specific site located 1400+/-100 bp downstream from the poly A addition site (Hofer et al., 1982).

iv The globin gene promoter

Comparison of the DNA sequences of the 5'flanking region of a large number of genes revealed nucleotide sequence homologies which were therefore presumed to play a role in transcription initiation. Thus, the ATA box, whose consensus sequence is GTATA(A/T)A(A/T)G is found 30 bp 5' to the mRNA cap site (Goldberg, 1979; Breathnach and Chambon, 1981) and the CAAT box, whose consensus sequence is GGCAAT is found 70-80 nucleotides 5' to the cap site (Benoist et al., 1980; Efstradiatis et al., 1980). Other conserved sequences which are specific to a given gene family have also been recognized e.g. conserved sequences in the 5' flanking DNA of the histone genes (Busslinger et al., 1979). The importance of these conserved sequences in the expression of globin and other genes has been confirmed by studying the expression of in vitro constructed mutants introduced into eukaryotic cell expression systems. Mutation of the ATA- or CAAT-box regions reduced the efficiency of correct transcription initiation of many genes including the rabbit β-globin

gene (G.Grosveld et al., 1981b, 1982a,b), and the Herpes Simplex Virus thymidine kinase gene (HSV tk; McKnight et al., 1981; McKnight, 1982). The ATA box was shown to direct the site at which transcription initiated, since deletion of DNA sequences between the ATA box and cap site lead to the use of downstream transcription initiation sites (G.Grosveld et al., 1981b, 1982a). Studies on the expression of in vitro constructed mutant genes also lead to the identification of another conserved element, the -90 box, which is located 80-100 bp 5' to the transcription initiation site and which has been shown to be essential for the transcription of the rabbit β-globin (G.Grosveld et al., 1982b) and HSV tk genes (McKnight et al., 1981). The 5' flanking region of the globin genes share other conserved sequence elements which are absent from non-globin genes; it has been suggested that these may be important in controlling specific expression of the globin genes in erythroid tissue (Moschonas et al., 1982).

Transcription of the ε-globin gene has recently been shown to initiate at several positions within the 4.5 kb region upstream of the normal major initiation site (Allan et al., 1982,1983). About 10-15% of ε-globin transcripts in purified erythroblasts from first trimester embryos initiate from these upstream sites, although their significance is unclear and no upstream initiation sites have been found in other globin genes. Multiple transcription initiation sites have been identified in several other genes; for example, transcription of the rat α-amylase gene initiates at different positions in the liver and salivary gland so as to generate alternately spliced transcripts (Young et al., 1981), and

transcription of the chicken skeletal muscle myosin L chain gene similarly initiates at two sites to form RNAs which are processed by different modes and which yield two distinct proteins (Nabeshima et al., 1984).

v Globin gene splicing and transcription termination

Compilation of the DNA sequences found at exon-intron boundaries of a number of genes revealed a sequence conservation, with the 5' exon-intron (donor) site having a consensus sequence 5'(C/A)AG*GT(G/A)AGT 3' and the 3'intron-exon boundary (acceptor) having the consensus sequence 5'(Py)6 N(T/C)AG*G 3' (Breathnach et al., 1978; Breathnach and Chambon, 1981; * represents the site of splicing). These sequences are thought to reflect the requirments for recognition, cleavage and ligation of pre-mRNA by the splicing enzymes. This recognition is thought to involve small nuclear ribonucleoproteins, and the 5' end of U1RNA, which is one of the most abundant snRNAs, shows sequence complementarity to the exon-intron and intron-exon boundaries (Lerner et al., 1980). The importance of these conserved boundary sequences has been analysed by studying the expression of in vitro constructed mutant globin genes after their introduction into eukaryotic cell expression systems (Wieringa et al., 1983). By mutating each nucleotide at the 5' splice site of the rabbit β-globin gene large intervening sequence, it was shown that all mutations reduced the efficiency of RNA splicing but that mutation of the first position of the intron (converting GT to AT) had the greatest effect. Mutation of the normal splice signals sometimes lead to the use of alternative splice sites within the

first exon which had sequence homology with the consensus splice site sequence. A number of thalassemic globin genes have been found to have point mutations at splice site junctions and undergo aberrant splicing both in the patient *in vivo* and upon transient expression of the cloned gene (Busslinger *et al.*, 1981; Treisman *et al.*, 1983; section 1.4viii). Comparison of three thalasemias with point mutations in the first intervening sequence 5' splice site has shown that the mutation with the greatest effect on RNA splicing was at the first position of the intron; the data obtained from expression of *in vitro* constructed mutant genes therefore agree with the observed effect of the naturally occurring thalassemic mutations (Treisman *et al.*, 1983).

Little is known about the control of transcription termination in eukaryotes. The formation of non-polyadenylated 3' ends of the histone gene transcripts requires the formation of a hairpin loop structure at the 3' end of the primary transcript (Birchmeier *et al.*, 1983), although no such requirment for polyadenylated transcripts has been shown. The recognition site for endonucleolytic cleavage of the precurser RNA and for polyA addition is thought to be the sequence AATAAA, which is found 20 bp ~~down~~up stream from the polyA addition site (Proudfoot and Brownlee, 1976).

vi Chromatin structure and globin gene expression

Early evidence suggesting that chromatin from actively transcribed genes had a different structure from that of non-transcribed regions originated from studies on the digestion of isolated nuclei with DNaseI. It was shown that the susceptibility of a given gene to

DNaseI digestion correlated with its transcriptional activity in the tissue from which the nuclei were isolated (Weintraub and Groudine, 1976). Thus, the globin gene was more sensitive to DNaseI digestion in nuclei isolated from chick erythrocytes as compared with chick oviduct, and the ovalbumin gene was more DNaseI sensitive in nuclei isolated from chick oviduct (Garel and Axel, 1976). The DNaseI sensitivity of transcribed genes is probably a general phenomenon and reflects the potential of a gene to be transcribed rather than its transcription level per se. Thus, the ovalbumin gene remains sensitive to DNaseI in the hormone withdrawn oviduct in which the gene is no longer being transcribed (Palmiter et al., 1977). Similarly, the mouse globin genes show the same sensitivity to DNaseI in MEL cells before and after the induction of their differentiation, i.e. the genes are already in a DNaseI sensitive state in the "committed" undifferentiated MEL cell (Miller et al., 1979; section 1.2).

"Active" chromatin is also characterized by the presence of DNaseI hypersensitive sites (for reviews see Weisbrod, 1982; Elgin, 1981). These sites are identified by mild digestion of chromatin with DNaseI and usually lie in the 5' flanking region of transcribed genes. For example, the chemically induced differentiation of cultured MEL cells (section 1.2) results in the appearance of a DNaseI hypersensitive site near the cap site of the globin gene, and the fact that mutant MEL cell lines which are resistant to induction do not form this hypersensitive site on treatment with inducing agents suggests that it may be relevant to gene activation (Sheffery et al., 1982, 1983). The DNaseI hypersensitive site of the chick

β-globin gene is one of the best characterized and extends from 60 to 260 bp 5' to the transcription initiation site. It is possible to excise this hypersensitive site as a 115 bp protein free DNA fragment by digestion of chromatin with MspI, suggesting that it is not associated with nucleosomes (McGhee et al., 1981). The appearance of this site has been correlated with globin gene activation in early development (Groudine and Weintraub, 1981). By transforming cultured chick embryo fibroblast cells with a virus which causes globin gene activation, it was shown that the hypersensitive site persists through many cell generations after the inducing agent is no longer present and the globin genes are no longer being transcribed. This indicated that the formation of hypersensitive sites might be an irreversible step in the committment to differentiation (Groudine and Weintraub, 1982). However, the DNaseI hypersensitive sites of the sea urchin histone genes change reversibly as these genes become activated and inactivated during early development (Bryan et al., 1983); these sites are found within the functionally important regions located 5' and 3' to the histone genes, and the differences in pattern of the appearance of the DNaseI hypersensitive sites in the histone and globin genes could reflect differences in the control of embryonic genes during early development and of genes in terminally differentiating tissues.

The DNaseI hypersensitivity of a given region of chromatin is probably due to a localized change in DNA conformation which is conferred by a specific DNA sequence. The single-stranded nature of the DNA in DNaseI hypersensitive sites is indicated by the fact that the single-strand specific SI nuclease cleaves the same

hypersensitive sites as does DNaseI in chick erythrocyte nuclei (Larsen and Weintraub, 1982). Several kinds of altered DNA conformation might result in an increased sensitivity of DNA to SI nuclease: the formation of cruciform structures at palindromic DNA sequences, the formation of left-handed Z-DNA at alternating pyrimidine-purine stretches, and the formation of melted loops in AT rich regions. The DNaseI hypersensitive site located 5' to the early region of SV40 has indeed been shown to exist in a Z-DNA conformation (Nordheim and Rich, 1983). All the above conformations relieve the torsional strain of supercoiling and are therefore stabilized when incorporated into supercoiled molecules. Several experiments indicate the importance of DNA supercoiling or torsional strain in the formation of DNaseI hypersensitive structures. For example, the DNaseI hypersensitive region of the chick globin gene has been cloned into pBR322 (Larsen and Weintraub, 1979) and was retained in the naked recombinant DNA molecule when it was in the supercoiled form, but was lost upon linearization of the plasmid. This implied firstly that the formation of the hypersensitive site was directed solely by DNA sequence of the cloned region and was not dependent on any specific protein interaction in the erythroid cell, and secondly suggested that the strain of supercoiling is required to induce the conformation change. Several other observations suggest that DNA topology may have a role in gene control. For example, the transcriptionally active fraction of SV40 minichromosomes extracted from infected monkey cells are torsionally stressed (Luchnik et al.,1982). The influence of DNA template topology on transcription efficiency has been studied more directly by analysing the expression

in Xenopus oocytes of the cloned HSV tk gene when introduced on supercoiled and linear plasmids (Harland et al., 1983). Transcription of the circular templates was >500 times more efficient than the linear ones, indicating the importance of DNA topology in determining the efficiency of gene transcription. Presumably the torsional strain which is provided by supercoiling and which is required both for the formation of hypersensitive sites in naked DNA and for efficient transcription on cloned DNA might be provided by other means in the cell, e.g. interaction with specific proteins. For example, interaction of the lac o and r proteins with the binding domain results in a 10-100 fold increase in the DNaseI sensitivity of the latter (Schmitz and Galas, 1979). Interestingly, the DNaseI hypersensitive sites in the 5' flanking DNA of the human c-myc gene coincide with regions which are thought to interact with specific regulatory proteins (Siebenlist et al., 1984)

vii Regulation of globin gene expression by DNA methylation

5-methyl cytosine (5meC) constitutes 2-7% of the total cytosine in mammalian DNA, and is the only modified base yet found in vertebrate DNA (Vanyushin et al., 1970). 90% of 5 me C is found within the dinucleotide sequence 5' CpG 3', and the symmetrical methylation of both DNA strands allows the methylation pattern to be retained throughout DNA replication by a methylase which acts on hemimethylated sites (Bird et al., 1978). The presence of 5meC within defined regions of DNA can be analysed by Southern blotting using restriction enzymes that recognize CpG containing DNA sequences (Roberts, 1978; McClelland, 1981); most of these enzymes will not

cleave DNA if the cytosine residues in their specific recognition site are methylated.

Data from many systems have shown an inverse correlation between the level of CpG methylation of a gene and its transcriptional activity. Measurement of the degree of methylation within 17 HpaII/MspI sites of the human β-like globin gene locus showed all the sites to be highly modified in sperm DNA, in which all of the genes are inactive, whereas there was a tissue specific demethylation around the transcribed genes in somatic cells. Thus, in adult bone marrow, the expressed β-gene was unmethylated as compared with the rest of the globin gene cluster, and in fetal liver, the expressed γ-globin genes were undermethylated (Van der Ploeg and Flavell, 1980). A more detailed analysis in other systems has shown that the demethylations which accompany gene activation usually take place at the 5' end of the gene as, for example, occurs in the activation of albumin (Ott et al., 1982), and adenovirus genes (Kruczek and Doerfler, 1982). An analysis of 15 mammalian genes has revealed that the 5' flanking sequences are enriched in CpG sequences relative to the rest of the gene (McClelland and Ivarie, 1982), and these combined data therefore suggested that a methylated gene cannot be expressed and that a low level of methylation, possibly specifically in the 5' flanking region of the gene, is required for gene activation. Other evidence suggesting that DNA methylation has a role in controlling gene expression derives from the use of the drug 5-azacytidine to induce demethylation of cytidine residues. 5-aza-C is a cytidine analog which cannot become methylated because it has a nitrogen atom at position 5; incorporation of 5 aza C into the

DNA of cultured cells leads to DNA demethylation (Jones and Taylor, 1981; Jones and Taylor, 1980). The azacytidine induced transcriptional activation of an endogenous retroviral genome in chick cells has been correlated with demethylation (Groudine et al., 1981), and similarly, the activation of the metallothionen gene in mouse thymoma cell lines by either 5 azaC or ultraviolet is associated with demethylation of the gene (Compere and Palmiter, 1981; Lieberman et al., 1983).

For a number of cellular genes, however, no clear correlation has been observed between the pattern of DNA methylation and gene expression in vivo. For example, during Xenopus development, the hormonally inducible vitellogenin genes of chick become expressed despite being highly methylated (Gerber-Huber et al., 1983). A direct approach to study the role of DNA methylation in gene regulation has been to methylate cloned DNA in vitro and determine the effect of this methylation on expression of the gene in a eukaryotic cell expression system (section 1.5ii). When in vitro methylated DNA is introduced into Xenopus oocytes or cultured mammalian cells, methylation at every CpG sequence is retained during DNA replication through hundreds of cell generations (Stein et al., 1982a; Wigler et al., 1981; Harland ,1982). Methylation of the cloned hamster aprt gene with HpaII methylase (which methylates cytosine residues in the sequence CCGG) inhibited its expression when assayed after introduction into L cells (Stein et al., 1982b); in vitro methylation of adenovirus genes inhibited their transcription in Xenopus ooctyes (Vardimon et al., 1982). Complete methylation of all CpG sequences of the human $^{A}\gamma$-globin gene inhibited its

expression in L cells, and by constructing molecules which were methylated only in specific regions of the gene, it was shown that methylation of only the CpGs within the region -760-+100 (relative to cap site) was sufficient to inhibit the transcription of an otherwise unmethylated gene; methylation of other regions of the gene had no effect on its expression (Busslinger et al., 1983). Not all genes, however, contain a large number of CpG dinucleotide sequences within this putative regulatory 5' flanking region, and other data suggest that the expression of certain genes is insensitive to inhibition by DNA methylation. Thus, totally methylated SV40 DNA and Xenopus rDNA are transcribed when introduced into L cells and Xenopus oocytes respectively (Graessman et al., 1983; Macleod and Bird, 1983). Some data suggest that hypermethylation of a gene may actually be a consequence of transcriptional inactivity rather than a cause of it. Thus, the infection of embryonal carinoma cells with retroviruses results in an immediate block in viral gene transcription, with de novo methylation of the viral genome taking place 8-15 days later (Gautsch and Wilson, 1983). This suggested that de novo methylation of the viral DNA may have been a consequence of its transcriptional inactivity, and that DNA methylation may be viewed as a means of stabilizing the structure of a gene in an active or inactive state, with its transcription level being primarily determined by other means, e.g. the binding of specific proteins to regulatory sequences, chromatin configuration etc. Methylation of CpG sequences is thought to favor the transition of B-form DNA to the left-handed Z form (Behe and Felsenfeld, 1981), although the significance of this observation to gene control is not clear.

viii The thalassemias

The thalassemias are a heterogenous group of genetic diseases which are characterized by a defect in globin chain production (for reviews, see Weatherall and Clegg, 1979, 1982). They provide a variety of naturally occuring models for studying both the regulation of globin gene transcription and the mechanisms which control globin gene switching during development. Thalassemia may result from a reduction in the synthesis of either an α-like or β-like chain of hemoglobin, and the resulting imbalance of α/βchain ratio leads to precipitation of the excess chain in the erythroid cell, thus leading to a reduction in red cell maturation and survival. The diseases are classified into α, β, δ, δβ and γδβ thalassemias according to the type of globin chain whose expression is affected; these types are further subdivided into the α^{o}, β^{o} thalassemias in which there is no globin chain synthesis from the defective allele, and the α^{+}, β^{+} types in which globin synthesis is merely reduced. In another group of related conditions, Hereditary Persistance of Fetal Hemoglobin (HPFH), γ-globin gene expression and HbF production persists into adult life.

The nature of the lesions in thalassemia are diverse at the molecular level. These defects have been characterized by genomic blotting, DNA sequence analysis of the cloned thalassemic genes, and by studying the level and structure of the transcripts produced from these defective genes both in the erythroid cells in vivo and after introduction of the cloned thalassemic gene into a eukaryotic cell expression system (section 1.5ii).

All of the α^{o} thalassemias characterized so far are the

result of deletions within the α-globin gene loci (Pressley et al., 1980,a,b; Orkin and Michelson, 1980). Most of the α^+ thalassemias are also due to deletions (Orkin et al., 1979), although recently an α-thalassemic gene has been shown to have a point mutation in the polyadenylation signal, with this defect giving rise to long transcripts which extend past the normal signal (Higgs et al., 1983).

In contrast, many of the β^0 thalassemias result from stop codons, notably at codons 17 and 39 (Orkin and Goff, 1981; Chang and Kan, 1979; Moschonas et al., 1981, see accompanying manuscript). Many of the β^+ thalassemias are due to point mutations that either destroy a normal splice site sequence or create a new splice site sequence elsewhere in the gene; these result in inefficient splicing and/or the production of aberrently spliced nonfunctional transcripts (Busslinger et al., 1981; Treisman et al., 1983). Another β^+ thalassemia is due to a point mutation residing 87 nucleotides upstream of the cap site within the promoter, thus resulting in inefficient transcription initiation (Treisman et al., 1983). Analysis of the expression of such splice site and promoter mutations in naturally occuring thalassemic genes is thus providing information about the DNA sequences required for transcription initiation and RNA splicing which is complementary to that obtained from studying the expression of *in vitro* constructed mutant genes (section 1.4iv,v).

Many of the δβ and γδβ thalassemias and most of the HPFHs are characterized by deletions of varying size within the β-like globin gene locus; some of these are associated with an elevated expression of the γ-gene in adult life (Bernards et al., 1979; Ottolenghi et

al., 1979; van der Ploeg et al., 1980). Analysis of the positions of these deletions and the resulting phenotype in various diseases has suggested that the deletions not only result in an obvious loss of expression of the deleted region, but also that they act in cis over considerable distances to influence differential gene expression within the human globin gene cluster. Thus, it has been suggested that removal of both an alu repeat residing 5' to the δ gene and sequences 3' to the β gene results in the HPFH phenotype. Examination of the lesions in the non-deletion HPFHs and thalassemias will help pinpoint more accurately the regulatory regions involved in controlling globin gene switching. Interestingly, the deletion in a Dutch case of γβ-thalassemia removes the ε, γ and δ-globin genes, but leaves the β-globin gene intact (van der Ploeg et al., 1980). The β-globin gene is not, however, expressed in the erythroid cells of the patient and its inactivity is probably the result of a translocation which placed transcriptionally inactive chromatin adjacent to the β-globin gene; this resulted in acquisition of an inactive chromatin structure by the affected β-globin gene, leading to its transcriptional silencing (Kioussis et al., 1983; see chapter 5).

1.5 EUKARYOTIC CELL EXPRESSION SYSTEMS

The mechanisms which control gene expression may now be studied by the reintroduction of cloned DNA into a eukaryotic environment. Expression of the given gene may be analysed either in a cell free transcription extract, or after its introduction into Xenopus

oocytes, cultured mammalian cells or whole animals. The *in vitro* transcription systems have been useful in the isolation of specific factors which regulate transcription initiation and RNA splicing, whereas the introduction of cloned genes into cultured cells or Xenopus oocytes has given information on the DNA sequences which regulate transcription initiation, RNA splicing, and gene activation during differentiation and in response to environmental changes.

i *Cell free transcription extracts*

In most *in vitro* transcription systems, expression of cloned DNA takes place in a HeLa cell extract which may or may not be supplemented with partially purified RNA polymeraseII (Weil *et al.* 1979; Manley *et al.*, 1980). Extracts from other cell types e.g. Xenopus oocytes have also been used (Bogenhagen *et al.*, 1980; Sakonju *et al.*, 1980). *In vitro* transcription has been used mainly to study the DNA sequences required for transcription initiation and to assay specific purified cellular components for their role in transcription initiation and RNA splicing; correct transcription termination, polyadenylation and efficient RNA splicing in these systems is not well documented.

HeLa cell transcription extracts have been used to determine the DNA sequences which regulate transcription initiation. Analysis of various base substitution and deletion mutants has shown that the ATA box (section 1.4iv) is required for efficient initiation of transcription of the Adenovirus 2 major late genes (Corden *et al.*, 1980), the chick conalbumin gene (Wasylyk *et al.*, 1980) and the rabbit β-globin gene (G.Grosveld *et al.*, 1981b), and has shown that

transcription always initiates 25-30 bp downstream from the ATA box (section 1.4iii). However, DNA sequences residing 5' to the ATA box are not required for transcription initiation *in vitro* (Corden *et al.*, 1980; Hu and Manley., 1981; Wasylyk *et al.*, 1980; G.Grosveld *et al.*, 1981b). This contradicts data obtained for the expression of *in vitro* constructed mutant genes in whole cell expression systems (section 1.5ii) in which case DNA sequences upstream of the ATA box are essential for efficient transcription. Such differences between the systems might reflect the absence of important factors from the transcription extracts or the fact that transcription in such extracts probably takes place on naked DNA. It also seems unlikely that the *in vitro* systems will allow reconstitution of complex regulatory phenomena; for example, the globin genes are transcribed in cell free extracts from HeLa cells which do not express their endogenous globin genes (G.Grosveld *et al.*, 1981b). Some cases of regulation have, however, been observed e.g. regulation of the SV40 early promoter *in vitro* by T Ag (Hansen *et al.*, 1981), and the specific transcription of vaccinia genes in extracts derived from viral infected cells (Puckett and Moss, 1983).

In vitro transcription extracts have also been used for the identification of factors which regulate transcription initiation. For example, the Xenopus transcription factor TFIII, which is required for transcription of the Xenopus 5S RNA genes, was isolated from Xenopus oocytes and was found to bind to the same internal region of the Xenopus 5S RNA gene as had previously been shown to be essential for its transcription (Engelke *et al.*, 1980). HeLa and Drosophila cells extracts have been fractionated to yield factors

which bind to the promoter region of genes transcribed by RNA polymeraseII (Dynan and Tjian, 1983; Parker and Topol, 1984; Davison et al., 1983). For example, fractionation of HeLa cell extracts has revealed the presence of a promoter specific transcription factor which activates the SV40 early promoter; the factor binds to the 21 bp repeat region which is located about 70-100bp upstream from the transcription initiation site and which is required for SV40 early gene activation in vivo (Dynan and Tjian, 1983). Hela cell extracts have also been fractionated to yield factors which bind to the ATA box of the conalbumin and Adenovirus late genes; the factors are necessary for correct transcription of these genes in HeLa extracts and bind to the ATA box in the absence of RNA polII (Davison et al., 1983).

The standard in vitro transcription extracts splice only poorly (Padgett et al., 1983), although nuclear extracts do show splicing activity when supplemented with various post nuclear fractions (Hernandez and keller, 1983). It is therefore possible to synthesis a precurser mRNA in the standard in vitro transcription system, and then study its splicing in these supplemented nuclear extracts. Although these systems are not efficient enough to test the splicing of in vitro constructed mutants, they have been useful in defining the role of snRNAs in RNA splicing (Hernandez and Keller, 1983; Padgett et al., 1983; section 1.4v) and in identifying intermediate RNA products of splicing reactions (Grabowski et al., 1984).

ii Expression of foreign DNA in cultured mammalian cells

Analysis of the expression of cloned genes after their introduction into eukaryotic cells has enabled the DNA sequences which are required for transcription initiation, RNA splicing, and gene activation to be studied in detail. In addition, the effect of DNA methylation and DNA conformation on transcriptional activity may be monitored in some of these systems. The methods for the introduction of cloned genes into eukaryotic cells include the microinjection of DNA into Xenopus oocytes (Kressman et al., 1977; Gurdon and Brown, 1978), and the introduction of DNA into cultured mammalian cells by calcium phosphate transformation (Graham and van der Eb, 1973; Wigler et al., 1977), dextran transformation (Sompayrac and Danna, 1981) microinjection (Capecchi 1980) protoplast fusion (Schaffner, 1980; Rassoulzadegan et al., 1982), or electric shock treatment (Neumann et al., 1982). In addition, the introduction of foreign DNA into the germline of Drosophila and mice promises to be a useful approach for studying tissue specific and developmental stage specific control of gene expression.

a Transient expression systems

Two main techniques are available for studying the expression of a foreign gene after its introduction into mammalian cells: the transient expression and the stable clone systems. The transient expression systems have been used to identify the nucleotide sequences which control transcription initiation and RNA splicing (section 1.4iv,v) and to identify transcriptional enhancer elements

(section 1.5iii). Cloned DNA is introduced into the cells by one of the above methods and its expression in the total cell population is analysed 24-48 hours later either by RNA hybridization analysis or by the detection of protein. Transient expression may be carried out in most types of eukaryotic cell line for which efficient introduction of foreign DNA can be obtained; under optimal conditions, 30-40% of the transfected cell population expresses the foreign DNA. Efficient expression of the foreign gene usually requires the presence of viral enhancer sequences (section 1.5iii) on the introduced vector. Most assays for promoter and splice site sequence requirements are carried out by cloning the gene of interest in a vector containing the SV40 enhancer, mutating the gene *in vitro*, and analysing its expression after introduction into HeLa cells; this approach was used to identify the promoter elements of the rabbit β-globin gene (G.Grosveld *et al.*, 1982a,b). Transient expression of cloned genes may also be studied after their introduction into COS cells (Gluzman, 1981); these are SV40 transformed monkey kidney cells which express SV40 T antigen and contain the permissivity factors required for SV40 DNA replication. Vectors containing the SV40 region required in cis for replication (i.e. the origin region) replicate to 100,000 copies per cell within a few days of their introduction into COS cells. Efficient expression of introduced DNA in COS cells requires the presence of an enhancer element; this system has been used to analyse the DNA sequences which control the initiation of human α-globin gene transcription (Mellon *et al.*, 1981). The Xenopus oocyte system may also be classified as a transient expression assay, since expression of the foreign DNA is assayed 24h after its injection into the cells.

Not all introduced genes are efficiently transcribed in Xenopus oocytes; globin genes are only poorly expressed, with the transcripts mainly initiating from sites upstream from the in vivo cap site. The cloned HSV tk gene is, however, efficiently transcribed in Xenopus oocytes, and the system has been used to identify the promoter elements of this gene (McKnight et al., 1981; Mc Knight, 1982). The Xenopus oocyte system has also been used to analyse the role of DNA template topology in gene transcription (section 1.4vi).

b Stable eukaryotic cell clones

DNA sequences required for transcription initiation and RNA splicing have also been analysed by studying the expression of cloned genes after their stable introduction into the genome of cultured eukaryotic cells (Dierks et al., 1981,1983; Treisman et al.,1983). The influence of DNA methylation on gene expression has also been studied in stable eukaryotic cell clones (section 1.4vii). Since only 1-1000 per 10^6 cells incorporate exogenous DNA in a stably heritable manner, a variety of selection systems are used to isolate such stable transformants. For example, tk^- and $aprt^-$ cell lines may be transformed to tk^+ or $aprt^+$ phenotypes by transfection with cloned tk or aprt genes; such transformed cells are selected by culturing the cells in aminopterin or asaserine which kill nontransformed tk^- or $aprt^-$ cells. Dominant selection systems may be used to transform any cell type i.e. one which does not necessarily have a mutant phenotype. For example, all mammalian cell are killed by appropriate concentrations of the antibiotic G418, and transformation with the cloned bacterial AGPT

gene renders cells resistant to the drug by inactivating it (Jimenez and Davies, 1980; Colbere-Garapin et al., 1981). Any nonselectable gene of interest may be introduced into the cells either by ligating it to the appropriate selectable gene or by cotransferring the selectable and nonselectable gene on separate molecules (Hsiung et al., 1980; Wigler et al., 1979). Expression of foreign DNA in a stable eukaryotic cell clone does not require the presence of a viral enhancer and transforming DNA usually integrates into the host chromosome and acquires a nucleosome conformation (Camerini-Otero and Zasloff, 1980). Expression of the foreign DNA is usually studied in clones of transformed cells or in mixed populations of clones by RNA hybridization analysis. Sometimes expression of the introduced DNA may be detected by its ability to confer a specific phenotype on the cells. For example, the transformation of mouse L cells with the cloned H-2Kb gene resulted in the expression of the H-2Kb protein on the cell surface, enabling its detection with allospecific anti-H-2Kb cytotoxic T cells (Mellor et al., 1982); by introducing recombinants constucted in vitro between the H-2Kb and another MHC gene, the region of the H-2Kb gene involved in cytotoxic T cell recognition was identified (Allen et al., 1984). Transfection of eukaryotic cells with total DNA or genomic libraries has been used to isolate genes which confer a specific phenotype on the cell e.g the hamster aprt and human tk genes have been isolated by transformation of $aprt^-$ and tk^- cells with total DNA (Lowy et al., 1980; Lin et al., 1983), and after transfection of mouse L cells with total human DNA, cells expressing HLA antigens have been selected by Fluorescence Activated Cell Sorting (FACS; Kavathas and

Herzenberg, 1983).

One of the major problems of studying the expression of cloned DNA in stable eukaryotic cell transformants is that both the number of copies of integrated foreign DNA and the level of expression of the foreign gene are very variable between individual transformants. The latter phenomenon presumably reflects an influence of host DNA at the site of integration, and complicates the interpretation of experiments in which a quantitative effect of a mutation on gene expression is analysed. The introduction of the foreign gene on a BPV vector obviates this problem (Di Maio et al., 1982; Law et al., 1981), since foreign BPV DNA remains in the cell as approximately 30 extrachromosomal elements per cell, and individual clones express the foreign gene at a similar level. The BPV vectors are therefore useful for analysing the effect of mutations on expression of a foreign gene in stable cell transformants (Treisman et al., 1983). Transformed clones containing BPV DNA may be selected by morphology or by including a dominant selection marker on the vector BPV construct.

iii Transcriptional enhancers

Transcriptional enhancer elements may be defined as segments of DNA which dramatically increase the level of transcription from a nearby promoter. Transcriptional enhancers were discovered in SV40 virus, with the SV40 enhancer being identified as a 200 bp segment of DNA which contained the 72 bp repeat region and which was essential for early gene transcription (Benoist and Chambon, 1981; Gruss et al., 1981; Fromm and Berg, 1982). Ligation of this enhancer to the rabbit

β-globin gene increased the level of expression of the latter by >100 fold in a transient expression assay (Banerji et al., 1981). Most genes have been shown to be efficiently expressed in transient expression assays only if they are linked to a viral enhancer; a notable exception is the α-globin gene, which is expressed in an enhancer independent manner in both COS and HeLa cells (Mellon et al., 1981; Treisman et al., 1983). In general, the enhancer may be positioned in either orientation and placed either 5' or 3' to the heterologous gene (Banerji et al., 1981; de Villiers and Schaffner, 1981; Moreau et al., 1981). Enhancers also act when located several kilobasepairs away from the gene, although specific DNA sequences located between the enhancer and promoter (e.g. another promoter) may interfere with the enhancement (Wasylyk et al., 1983). Enhancer elements have now been found in a wide variety of viruses, including SV40 (Banerji et al., 1981), Polyoma (de Villiers and Schaffner 1981, BKV (Rosenthal et al., 1983,), BPV (Lusky et al., 1983), adenovirus (Hearing and Shenk, 1983) and retroviral LTRs (Laimins et al., 1982; Levinson et al., 1982). Viral enhancers often comprise tandem repeats of 50-100 nucleotides, and many have a DNA sequence which has the potential of forming Z DNA (Nordheim and Rich, 1983). Enhancers usually show some species and tissue specificity in their action; for example, the polyoma enhancer is more efficient than the SV40 counterpart in mouse 3T6 fibroblasts, whereas the SV40 enhancer is more efficient in monkey kidney CV-1 cells (de Villiers et al., 1982). The mechanism of transcriptional enhancement is unclear, although a popular model is that they act as a "super entry site" for some component of the transcription machinery. Interestingly, the

protein product of the Adenovirus EIA gene may activate a heterologous gene when provided *in trans* in a transient expression assay; this E1A mediated gene activation does not require the presence of a viral enhancer *in cis* on the construct, and the only promoter elements that are required for efficient transcription of an E1A activated gene are those including and located 3' to the ATA box (Green *et al.*, 1983; Imperiale *et al.*, 1983).

The only well defined eukaryotic enhancer is the tissue specific immunoglobulin gene enhancer, which is located in an intron upstream of the constant region gene and which probably serves to activate the variable region promoter of the functionally rearranged immunoglobulin gene in the lymphocyte. Ligation of the immunoglobulin enhancer to cloned rabbit β-globin and SV40 T antigen genes increased the expression of these heterologous genes when introduced into lymphoid derived myeloid cells but not in other cell types (Queen and Baltimore, 1983; Banerji *et al.*, 1983; Gillies *et al.*, 1983). It will be of obvious interest to determine whether enhancers have a more widespread role in the regulation of eukaryotic gene expression, and whether all genes which are expressed in a tissue specific manner possess an appropriate enhancer element which acts only in certain cell types.

iv Regulated expression of cloned genes in cultured eukaryotic cells

Many genes which are physiologically expressed only in specific differentiated tissues are found to be efficiently transcribed when introduced as cloned DNA into heterologous tissue culture cells which do not express their equivalent endogenous gene. For example, the

rabbit β-globin gene is transcribed in stable mouse L cell clones (Dierks et al., 1981) and in the HeLa cell transient expression assay (G.Grosveld et al., 1982a,b). This implied that gene transfer systems, although ideal for studying DNA sequences which control transcription per se, were not suitable for analysing the sequences which regulate gene activation as occurs during cell differentiation or in response to environmental changes. In some systems, however, it has been possible to regulate the expression of a cloned gene after its introduction into a eukaryotic cell, and by studying the expression of in vitro constructed mutant genes it is possible to identify the regulatory sequences which control gene activation as opposed to those required for transcription per se. Such studies have shown that the information required for activation of a gene is an intrinsic property of its DNA sequences and is not dependent on its localization in a specific chromosomal domain or its acquisition of a particular chromatin conformation pattern or DNA modification in the host cell. Most of the regulatory elements identified so far are located within the 5' flanking DNA of the gene.

A number of genes whose expression is regulated by glucocorticoid hormones in vivo also show hormonal inducibility when introduced as cloned DNA into mouse fibroblast L cells which contain glucocorticoid receptors. Thus, hormonal regulation of the cloned human growth hormone gene (Robins et al., 1982), MMTV proviral DNA (Hynes et al., 1981; Ucker et al., 1981; Huang et al., 1981), and the rat α2 globulin gene (Kurtz, 1981) was observed when they were stably introduced into mouse L cells. The 5' flanking region of the growth hormone gene conferred inducibility upon the normally noninducible

Herpes virus thymidine kinase (tk) gene or the dihydrofolatereductase (DHFR) gene, suggesting that the growth hormone gene regulatory sequences resided within its 5' flanking DNA. Similarly, the 5' flanking region of MMTV will confer inducibility upon the HSV tk gene when placed at 430 (Hynes et al., 1983) or 120 (Chandler et al., 1983) bp upstream from the cap site. The presence of regulatory DNA sequences within the structural gene was not, however, rigorously tested, and the observation that purified glucocorticoid receptor binds not only to the 5' flanking region but also downstream from the transcription initiation site of the MMTV LTR suggests that these latter sequences may also be important in control (Payvar et al., 1983). Similarly, both the chicken lysozyme gene (Renkawitz et al., 1982) and the ovalbumin gene (Dean et al., 1983) are inducible by estrogen in a transient expression assay in primary chick oviduct cells, and the 5' region of these genes was shown to confer inducibility upon the normally noninducible chicken β-globin and SV40 T antigen genes (Dean et al.,1983). The 5' flanking region of the interferon gene has been shown to mediate its induction by virus or poly rIrC after transfer to L cells (Ragg and Weissmann, 1983; Weidle and Weissman, 1983; Zinn et al., 1983). The human metallothionen II-A (MTX) gene, which is physiologially inducible by both glucocorticoids and heavy metals in vivo, is similarly regulated by both agents when introduced into rat fibroblasts and has been shown by deletion mapping to possess two regulatory elements; the hormone responsive element is a 20 bp sequence which is centred at 250-260 bp upstream from the cap site and which binds purified GH-receptor complex in vitro. The metal ion responsive element

comprises 2 boxes of 14 nucleotides which are centred at 45 bp and 145 bp upstream from the cap site (Karin et al., 1984a). Both of these elements can activate heterologous promoters located at least 600 bp away, suggesting that the enhancement of transcription by glucocorticoids and heavy metals is mediated via the activation of enhancer like elements (Karin et al., 1984b). Both the interferon and MTX genes show regulated expression when introduced into eukaryotic cells on a BPV vector, and thus remaining as extrachromosomal elements (Zinn et al., 1982; Pavlakis and Hamer, 1983).

Cloned Drosophila heat shock genes retain their ability to be induced by high temperature after transfer into L cells (Corces et al., 1981), COS cells (Pelham 1982) or Xenopus oocytes (Bienz and Pelham, 1982). Deletion mapping has shown that DNA sequences residing 47-67bp upstream from the cap site are essential for heat inducible transcription, and a synthetic version of this region rendered a normally noninducible gene heat sensitive when placed at the varying distances from the cap site (Pelham and Bienz, 1982). This implied that heat shock transcription is controlled by an upstream element that acts as a transcriptional activator at high temperature but is inert at low temperature.

The introduction of cloned genes into cells which may be induced to differentiate in culture provides a system in which to study the regulation of gene activation during cell differentiation. For example, cultured embryonal carcinoma cells, which were derived from germ cell tumours and behave as a subset of early blastocyst cells, may be induced to differentiate in vitro along several

different pathways of embryogenesis (for review see Martin, 1980). Each type of differentiation is accompanied by the activation of specific sets of genes e.g. differentiation to visceral endoderm results in activation of the alphafetoprotein gene, whereas differentiation to either visceral or parietal endoderm leads to activation of the genes of the major histocompatibility complex. By introducing a cloned alphafetoprotein gene into cultured embryonal carcinoma cells, Scott et al. (1984) have shown that the foreign gene becomes specifically activated upon differentiation of the cells to parietal endoderm. Similarly, Rosenthal et al. (1984; see accompanying manuscript) have demonstrated the transcriptional activation of a foreign MHC gene in embryonal carcinoma cells upon differentiation to either parietal or visceral endoderm; experiments are in progress to define the regulatory elements which control the activation of the MHC genes during early embryonic development. The activation of foreign globin genes during erythroid differentiation in cultured murine erythroleukemia cells serves as a useful system in which to study globin gene activation during erythropoiesis (section 1.6).

v Introduction of cloned genes into whole animals

The mechanisms that control gene activation at specific stages in development and in specific tissues may be most rigorously studied by the introduction of cloned DNA into the germline of an animal; such foreign DNA is transmitted to the progeny and its expression may be studied in the various tissues of the animal throughout development.

The original experiments demonstrating the transfer of foreign

DNA into whole animals involved the introduction of viral DNA into the mouse by either infection of the preimplantation embryo, or by microinjection of DNA into the blastocyst cavity of the embryo (Jaenisch and Mintz, 1974; Gordon et al., 1980). The developing mice were mosaics; they contained foreign DNA in only some of their tissues and there was only limited success in introducing foreign DNA into the germline. Using this method, different mouse substrains were generated which carried Moloney Leukemia Virus (MoLV) integrated into different sites of the host chromosome; integration into each site resulted in a different temporal and tissue specific pattern of viral expression, suggesting that expression of the foreign DNA was influenced by the host DNA at the site of chromosomal integration (Harbers et al., 1981). Another method of introducing foreign DNA into mice by gene transfer at the blastocyst stage has been to introduce cultured embryonalcarcinoma cells (section 1.5iv) into the blastocyst; these transferred malignant cells then behave as normal embryonic cells, and participate in development to produce mosaic mice in which some of the tissues derived from the introduced teratocarcinoma cells (Mintz and Ilmensee, 1975). Transfection of the teratocarcinoma cells with foreign genes prior to their introduction into the animal enables the introduction of foreign DNA whose state in the teratocarcinoma cell can be characterized before transfer (Pellicer et al., 1980). There are, however, only a few reports of the introduction of injected teratocarcinoma cells into the germline (Bradley et al., 1984).

Many of the problems associated with the blastocyst injection technique have been overcome by microinjection of the DNA directly

into the pronucleus of the fertilised one-cell egg. 2-40% of the developing eggs retain the foreign DNA, and the resultant mice are usually nonmosaic and contain the introduced DNA in their germ-line. The foreign DNA is transmitted to the progeny mice and its expression in the developing mice can be studied to determine if it shows correct tissue specificity and developmental and environmental regulation.

Several foreign genes have now been introduced into the mouse germ line by pronucleus injection. These include the chicken transferrin gene (McKnight et al., 1983), rabbit β-globin gene (Costantini and Lacy, 1981; Lacy et al., 1983), immunoglobulin gene (Brinster et al., 1983) and hybrid genes containing the metallothionen gene promoter fused to HSV tk or growth hormone structural gene sequences (Palmiter et al., 1982a,b; Brinster et al., 1981). In all cases, germline integration of foreign DNA and transmission to progeny was reported, although success in obtaining correct tissue specificity of expression has been variable. The first successful experiments were those involving transfer of the metallothionen hybrid genes. The metallothionens are cysteine-rich polypeptides which bind heavy metals and give resistance to heavy metal toxicity; the metallothionen gene is expressed in most tissues, but mainly in liver, and is induced by heavy metals and glucocorticoids. Mice containing the metallothionen-growth hormone hybrid gene showed high levels of fusion mRNA in the liver, with the accumulation of this mRNA being comparable to that of the endogenous metallothionen gene (Palmiter et al., 1982b). Growth hormone levels in the transgenic mice were 100-800x higher than in normal mice,

resulting in animals of twice the normal weight. Transfer of this metallothionen-growth hormone hybrid gene into "little" mice that were genetically deficient in growth hormone resulted in the production of transgenic mice that were the size of normal nonmutant mice (Brinster et al., 1984).

Other foreign genes which have been introduced by the pronucleus injection technique into the mouse germ line also show correct tissue specificity in their expression. For example, the chicken transferrin gene is preferentially expressed in the liver of transgenic mice and the introduction of a functionally rearranged immunoglobulin gene into mice resulted in its correct tissue specific expression in the splenic B lymphocytes. The fact that this foreign immunoglobulin gene was not expressed in other tissues although it was present in a functionally rearranged form indicated that rearrangement was not the sole requirement for tissue specific expression of the immunoglobulin genes in the B cells. The immunoglobulin protein was correctly translated, glycosylated, and secreted in these transgenic mice.

Initial experiments involving the introduction of cloned globin genes into the germline of mice were disappointing. A foreign rabbit β-globin gene was not expressed in a correct tissue specific manner, and was transcribed specifically in nonerythroid tissues, e.g. skeletal muscle and testis, in some animals (Costantini and Lacy, 1981; Lacy et al., 1983). More recently, the introduction of other foreign globin genes into the mouse germline has resulted in their correct tissue specific expression in the erythroid tissues of some transgenic mice (Costantini et al., personal communication)

although the system is probably not yet efficient enough to enable analysis of the regulatory elements which control the tissue-specific and developmental stage-specific activation of the globin genes.

Another system for introduction of foreign DNA into the germline involves the introduction of cloned DNA into the Drosophila germline by P element mediated transformation (Spradling and Rubin, 1982; Rubin and Spradling, 1982). Correct tissue specific expression of cloned xanthine dehydrogenase, alcohol dehydrogenase and dopamine decarboxylase genes has been reported, and the system is now being used to analyse more complex phenomena such as dosage compensation in Drosophila (Spradling and Rubin, 1983; Goldberg et al., 1983; Scholnick et al., 1983)

1.6 AIMS OF THESIS

The introduction of cloned genes into eukaryotic cell lines which may be induced to differentiate in culture provides a useful system in which to study the regulation of gene expression during cell differentiation (section 1.5iv). The aims of this thesis were:-

1. To determine whether the expression of foreign globin genes was regulated during erythroid differentiation in cultured murine erythroleukemia cells (section 1.2) and hence whether this system is suitable for analysing globin gene activation during erythropoiesis

(Chapter 3).

2. To determine the interrelationships between the expression of the individual embryonic, fetal and adult globin genes of the human β-like globin gene cluster (Chapter 3), since analysis of genetic diseases of globin gene expression had previously suggested that the developmental stage specific expression of the individual genes was influenced by long range cis interactions within the gene cluster.

3. To determine whether the transcriptional inactivity of the β-globin gene in vivo in a γβ-thalassemic patient was due to a mutation in the regulatory DNA sequences required for globin gene activation (Chapter 5).

4. To localize the DNA sequences which regulate the activation of the β-globin gene during erythroid differentiation (Chapters 6,7,8).

Figure 1

Hematopoietic differentiation

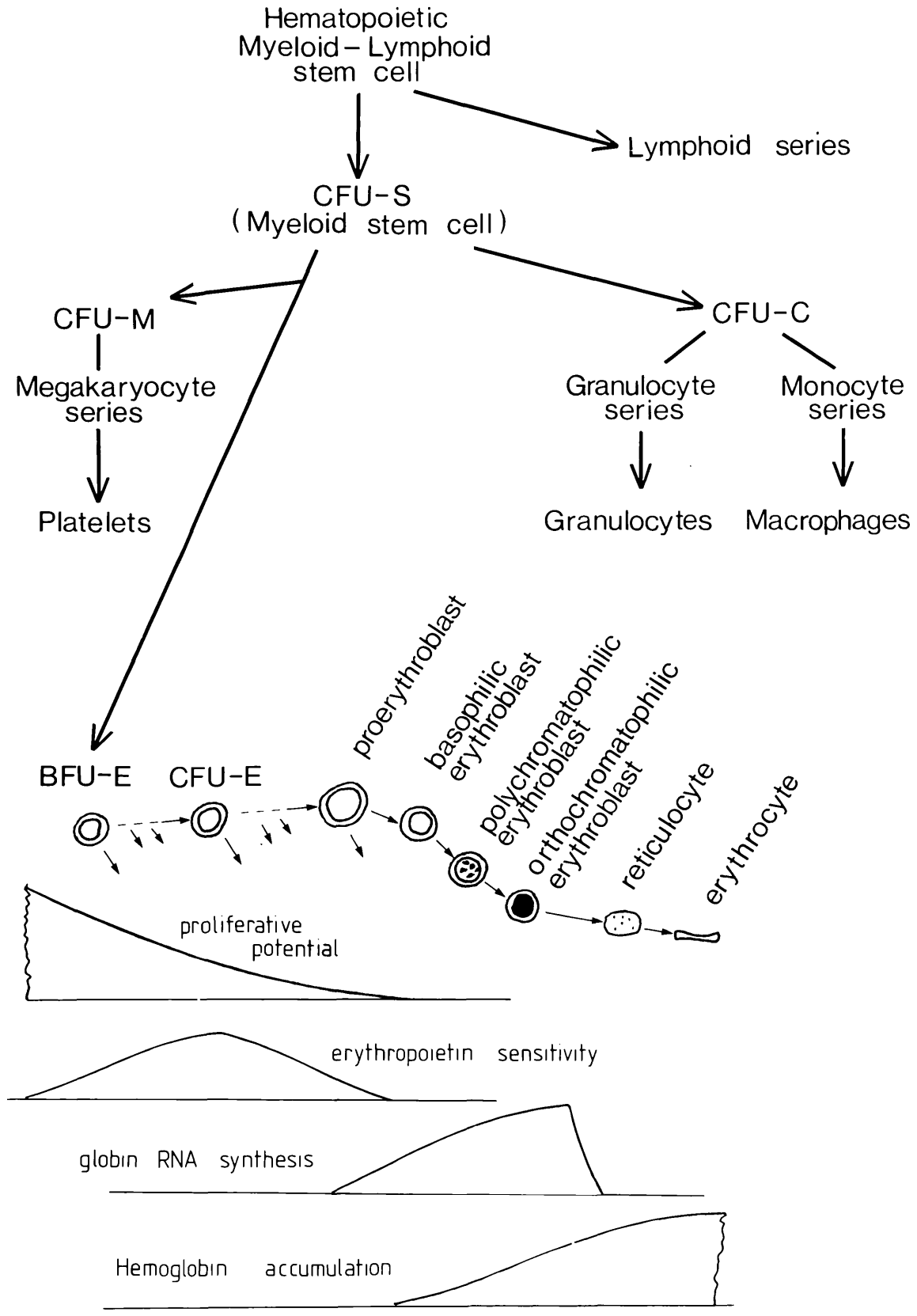
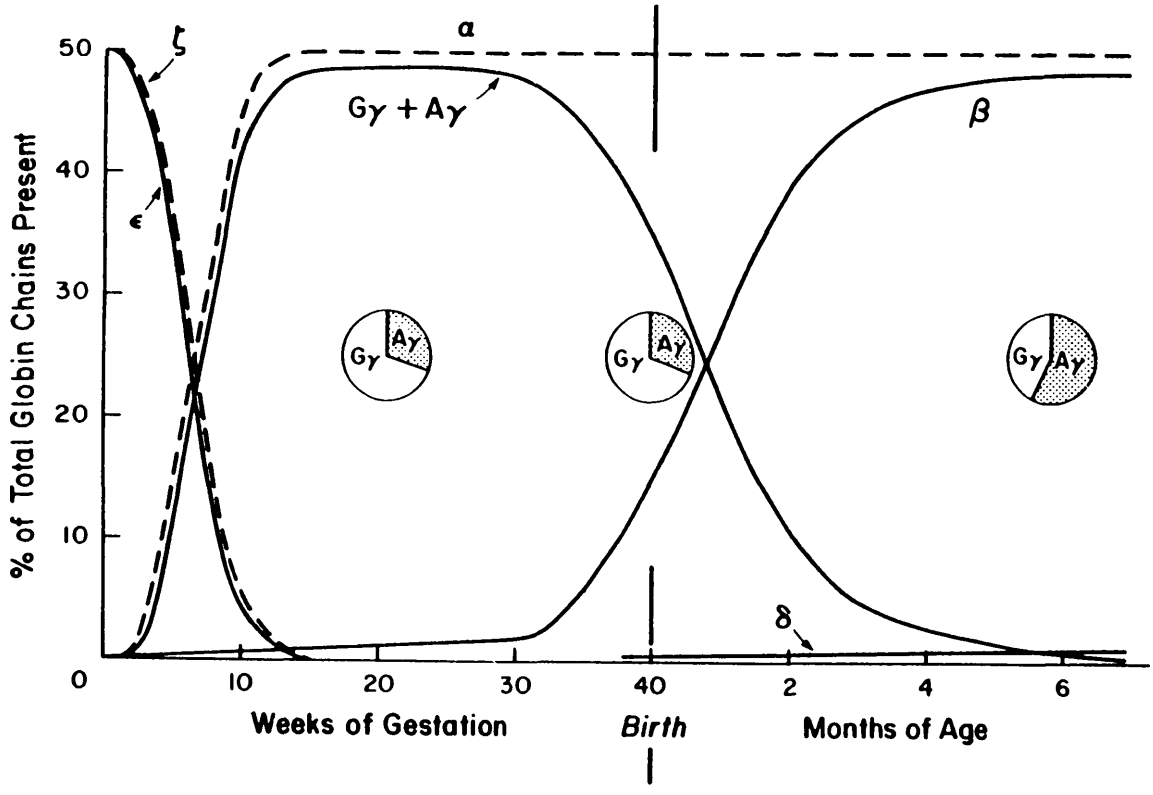

Hematopoietic
Myeloid – Lymphoid
stem cell
Lymphoid series
CFU-S
(Myeloid stem cell)
CFU-M
Megakaryocyte
series
Platelets
CFU-C
Granulocyte
series
Monocyte
series
Granulocytes
Macrophages
BFU-E
CFU-E
proerythroblast
basophilic
erythroblast
polychromatophilic
erythroblast
orthochromatophilic
erythroblast
reticulocyte
erythrocyte
proliferative
potential
erythropoietin sensitivity
globin RNA synthesis
Hemoglobin accumulation

Figure 2

Developmental switches in globin gene expression

A. Diagrammatic representation of globin gene expression during human development

B. Changes in hemoglobin composition during development



Stage of development	Hemoglobin	Chain composition	Proportion %
Embryo	Portland	$\zeta_2\gamma_2$	?
	Gower I	$\zeta_2\varepsilon_2$	?
	Gower II	$\alpha_2\varepsilon_2$	?
Fetus	F	$\alpha_2\gamma_2$	90
	A	$\alpha_2\beta_2$	5-10
Adult	A	$\alpha_2\beta_2$	97
	A2	$\alpha_2\delta_2$	2.5
	F	$\alpha_2\gamma_2$	0.5

Figure 3

Globin gene organization in different species

References:

Goat	Schon *et al.*, 1981
Rabbit	Lacy *et al.*, 1979
	Hardison *et al.*, 1979
Human	Fritsch *et al.*, 1980
	Lauer *et al.*, 1980
Lemur	Barrie *et al.*, 1981
Mouse	Jahn *et al.*, 1980
	Popp *et al.*, 1981
Chicken	Dolan *et al.*, 1981
	Engel *et al.*, 1980
Frog	Hosbach *et al.*, 1983
	Lee *et al.*, 1983

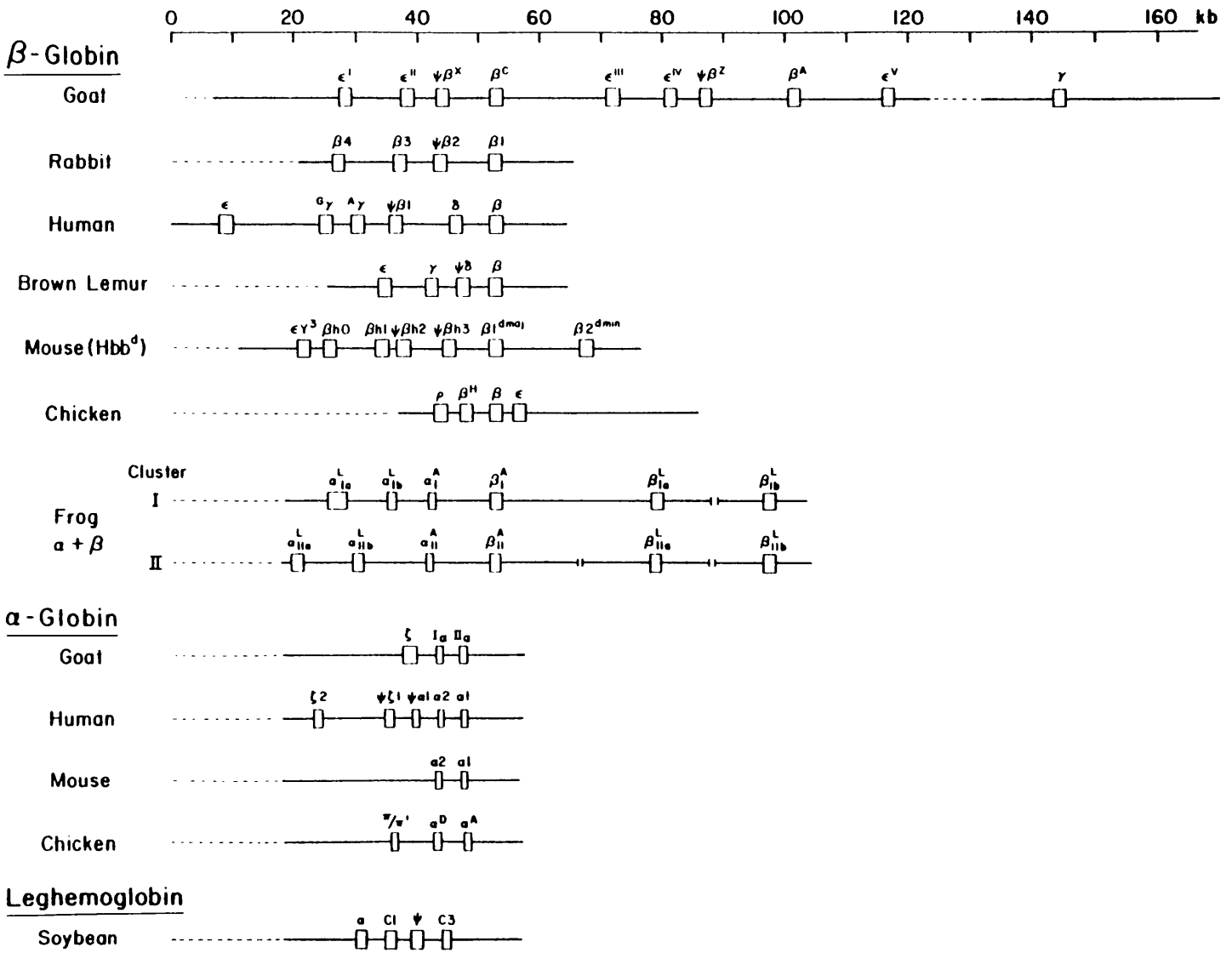

0
20
40
60
80
100
120
140
160 kb
β-Globin
Goat
Rabbit
Human
Brown Lemur
Mouse(Hbb^d)
Chicken
Frog
α + β
Cluster
I
II
α-Globin
Goat
Human
Mouse
Chicken
Leghemoglobin
Soybean

Figure 4

Globin gene structure and transcription

A. Structure of the human α- and β-globin genes

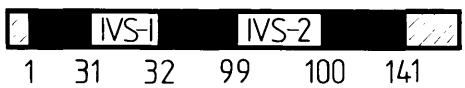
Hatched boxes: 5' and 3' untranslated regions
Open boxes : intron sequences
Solid boxes : coding regions

Codon numbers at the exon boundaries are indicated

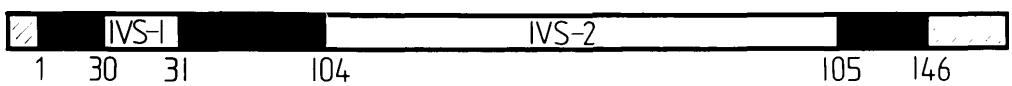
B. Rabbit β-globin gene promoter

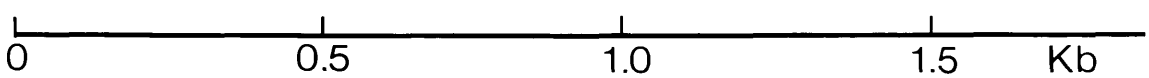
Nucleotide sequence of the 5' flanking region of the rabbit β-globin gene, showing positions of the -90 region (G.Grosveld et al., 1982b), CAAT box and ATA box (G.Grosveld et al., 1981b, 1982a).

Human α globin gene

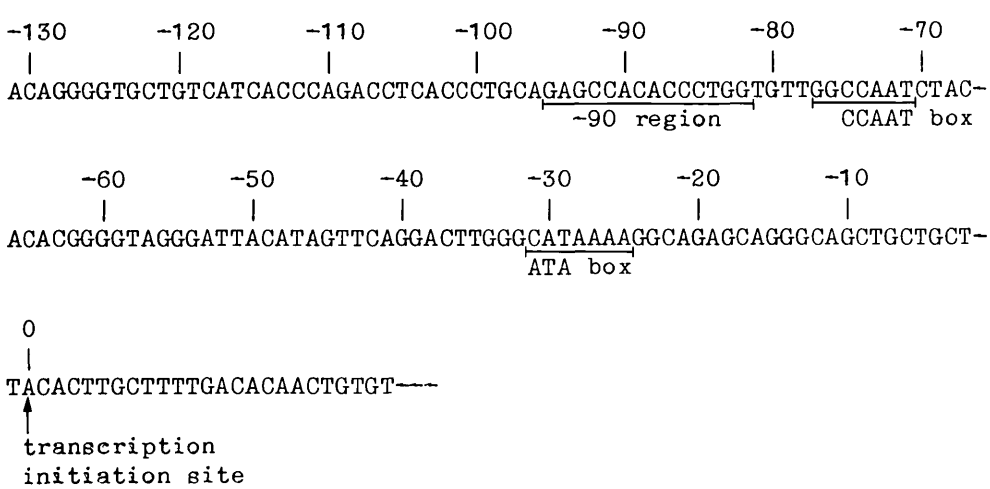


Human β globin gene





RABBIT β-GLOBIN GENE PROMOTER



CHAPTER 2

MATERIALS AND METHODS

2.1 Chemicals

All chemicals unless otherwise stated were AnalaR grade and purchased from B.D.H. Chemicals Ltd. Nucleotides were purchased from Sigma Chemical Company. Solutions were prepared using double distilled water; they were sterilized by autoclaving at 15 lb/in^2 for 30 minutes or by filtration through 0.20μm Millipore filters.

2.2 Glassware and plasticware

Disposable plastic 1.5ml snapcap tubes and plastic tips for automatic Gilson pipettors were purchased from Eppendorf and Jencons Ltd. respectively and were sterilized by autoclaving at 15 lb/in^2 for 30 minutes. Sterile plastic disposable tubes and pipettes were purchased from Sterilin Ltd. Pyrex glassware was sterilized by baking at 200°C for 5 hours.

2.3 Radioactivity

γ-^{32}P adenosine 5'-triphosphate (specific activity 5000 Ci/mmol), α-^{32}P deoxyribonucleoside 5'-triphosphates (specific activity 3000 Ci/mmol) and α-^{32}P uridine 5'-triphosphate (specific activity 3000 Ci/mmol) were purchased in stabilized aqueous solution

from Amersham International PLC. They were added directly to reactions without drying down.

2.4 Autoradiography

^{32}-P labelled samples were detected in acrylamide gels and nitrocellulose filters by autoradiography using preflashed Fuji RX NIF X-ray film with intensifying screens at -70°C (Swanstrom and Shank, 1978; Laskey and Mills, 1977).

2.5 Spectrophotometric determination of nucleic acids

Concentrations of nucleic acid solutions were determined from optical density (O.D.) readings at 260nm and 280nm. An O.D.260 reading of 1.0 corresponded to concentrations of 50μg/ml and 40μg/ml for solutions of double stranded DNA and RNA respectively. Pure preparations of DNA and RNA had O.D.260/280 ratios of 1.8 and 2.0 respectively. Small amounts of DNA and RNA (i.e. <2μg) were quantitated from the intensity of ethidium bromide flourescence in agarose gels (section 2.18).

2.6 Bacterial strains and growth

Plasmids and cosmids were propogated in the rec A$^-$ E.coli strains HB101 (RecA13; Boyer and Rouland Dussoix, 1969) or ED8767 (Murray et al., 1977).

Bacteria were grown in Luria-Bertani (LB) medium which

consisted of 10g/l bactotryptone, 5g/l yeast extract, 5g/l NaCl. Medium was supplemented with 50μg/ml ampicillin or 15μg/ml tetracycline when bacteria harboured plasmids containing the appropriate drug resistance gene. Liquid bacterial cultures were grown on rotary platforms (250-300rpm) at 37°C. Agar plates contained 12g bactoagar per liter in LB medium.

Bacterial stocks were frozen at -70°C at a density of 10^7-10^8 cells/ml in Hogness Modified Freezing Medium (HMFM). HMFM consisted of 4.4% glycerol, 3.6mM K_2HPO_4, 2mM Na_3 citrate, 1.3mM KH_2PO_4, 1mM $MgSO_4$.

Bactotryptone, yeast extract, and bactoagar were purchased from Difco Laboratories. Ampicillin and tetracycline were purchased from Sigma Chemical Company.

2.7 Preparation of competent HB101

The method used was a modification of those described by Mandel and Higa, 1973 and by Cohen et al., 1973, in which the uptake of DNA by bacteria was enhanced by their treatment with calcium chloride.

100μl of a logarithmic phase (O.D.600nm = 0.2-1.0) bacterial culture was innoculated into 20ml LB medium and grown until O.D.600 = 0.6. 1ml of culture was then diluted into 1l LB medium and grown until O.D.600= 0.1. The bacteria were left on ice for 20 minutes, pelleted by centrifugation at 4000g for 10 minutes at 4°C, and resuspended in 240ml 0.1M $CaCl_2$. They were left on ice for a further 20 minutes and pelleted as before. The competent cells were then resuspended in 6ml 14% glycerol, 86mM $CaCl_2$, transferred in

100µl aliquots into 1.5ml Eppendorf snapcap tubes, frozen in liquid nitrogen and stored at -70°C.

2.8 Transformation of competent HB101 by plasmid DNA

1.0ng of supercoiled plasmid DNA or 100-1000ng DNA in a ligation reaction was used for transformation.

50µl frozen competent cells (section 2.7) were thawed on ice. 1-10µl DNA and 100µl TCM buffer (TCM = 10mM tris, 10mM $CaCl_2$, 10mM $MgCl_2$) were added and the mixture left on ice for 30 minutes. The bacteria were subjected to heat shock by transfer to a 42°C waterbath for 90 seconds. 1.2ml LB medium was then added and the cells incubated at 37°C for 45 minutes. Bacteria were pelleted by centrifugation for 20 seconds in an Eppendorf microfuge, resuspended in 100µl LB medium, spread on an agar plate containing the appropriate antibiotic and incubated at 37°C. Colonies of transformed bacteria appeared after 5-14 hours. They were either picked, innoculated into 10ml LB medium and a miniculture grown (section 2.15), or they were transferred to a nitrocellulose filter for direct screening by hybridization (section 2.25).

2.9 Eukaryotic cell culture

The thymidine kinase negative mouse fibroblast L cell line (LD1 tk^-) was a gift of R. Axel. Thymidine kinase negative mouse erythroleukemia cells (F412B2 tk^- line) were obtained from D. Spandidos.

Cell cultures were grown at 37°C, 10%CO_2 in αMEM medium supplemented with 5-10% fetal calf serum (FCS). Cell lines transformed by plasmid DNA to a tk^+ phenotype were maintained in 15μg/ml hypoxanthine, 0.5μg/ml aminopterin, 5μg/ml thymidine (HAT medium). Cell lines transformed to G418 resistance were maintained in 0.5-1.0 mg/ml G418.

Cell stocks were frozen in liquid nitrogen at a density of 10^6-10^7 cells/ml in αMEM supplemented with 20% FCS and 10% DMSO.

Sterile plasticware for tissue culture was purchased from Sterilin Ltd. or from Flow Laboratories. αMEM medium, FCS and G418 were obtained from Gibco Europe. Hypoxanthine, aminopterin and thymidine were purchased from Sigma Chemical Company.

2.10 Differentiation of murine erythroleukemia cells

Murine erythroleukemia cells were induced to differentiate by culturing in hexamethylenebisacetamide (HMBA) as described by Reuben et al., 1976. Cells were seeded at a density of 10^5/ml in αMEM+10%FCS into 90mM petridishes. 24 hours later, the medium was changed and supplemented with 3mM HMBA. The cells were optimally differentiated 3-4 days later.

HMBA was a gift of P. Marks.

2.11 Stable transformation of eukaryotic cells by cloned DNA

Cells were transformed using a modification of the method described

by Wigler et al., 1979. For transformation of L cells, 100ng DNA carrying a selectable gene was added to each petridish which contained 1-5x10^6 cells; for transformation of MEL cells, 1-5µg selectable DNA was added. The total amount of added DNA was made up to 40µg with salmon sperm carrier DNA and/or nonselectable DNA to be introduced into the cells by cotransformation (Hsiung et al., 1980).

18-36 hours before addition of transforming DNA, 1-2x10^6 cells were seeded in αMEM+10% FCS into a 90 mM petridish. DNA to be introduced into the cells was precipitated with ethanol, dissolved in 450µl 1mM Tris, 0.1mM EDTA pH7.9, and 50µl 2.5M $CaCl_2$ added. 500µl 2xHBS (2xHBS = 280mM NaCl, 50mM HEPES, 1.5mM Na_2HPO_4 pH 7.12) was added dropwise, resulting in the formation of a fine white precipitate. This was left at room temperature for 10-30 minutes and then put on the cells. The medium of the cells (αMEM+10% FCS) was renewed 12-24 hours later, and after a further 24 hours the cells were diluted so as to maintain a density of 10^5 cells/ml into medium supplemented with HAT or G418 (concentrations as in section 2.9). Cells usually started dying after 2 days culturing in selection medium, and colonies of transformed cells appeared after a further 7-14 days. The transformation efficiency of L cells was usually 1000 clones/µg DNA/10^6 cells; the efficiency of MEL cell transformation was 1-50 clones/µg DNA/10^6 cells. Clones were transferred into 10mm microwell dishes, and maintained throughout subsequent passaging in selection medium.

2.12 Phenol/Chloroform extraction of nucleic acids

Protein was removed from DNA and RNA solutions by extraction with water-saturated phenol:chloroform:isoamylalcohol (24:24:1,v:v:v) containing 0.1% hydroxyquinoline (PCI). Samples were mixed with an equal volume of PCI until an emulsion formed, and the phases separated by centrifugation at 1600g for 10minutes. When isolating high molecular weight genomic DNA, the DNA solution and PCI were mixed by gentle swirling and large bore pipettes were used to transfer the aqueous phase. Traces of phenol or chloroform were removed from nucleic acid solutions by extraction with water-saturated ether.

2.13 Ethanol precipitation of nucleic acids

RNA and DNA were precipitated from solution by adjusting the salt concentration of the nucleic acid solution to 0.25M in NaAc, addition of 2.5 volumes of ethanol, and cooling at $-20^{o}C$ for 30 minutes. Nucleic acid was recovered by centrifugation at 12,000g for 10 minutes at $0^{o}C$. High molecular weight genomic DNA was precipitated by gently mixing the DNA solution with 2.5 volumes of ethanol and lifting out the fibrous DNA precipitate with a sealed pasteur pipette.

2.14 Large scale preparation of plasmid DNA from bacterial cultures

Plasmid DNA was prepared using a modification of the alkaline

lysis method of Birnboim and Doly, 1979, followed by further purification of the DNA on caesium chloride gradients (Radloff et al., 1967).

Bacterial cells of a 1 litre ~~logarithmic~~ saturated phase culture were pelleted by centrifugation at 4000g for 20 minutes, resuspended in 50ml 50mM glucose, 10mM EDTA pH8, 25mM Tris pH8, and lysed by mixing for 5 minutes with 100ml 0.2M NaOH, 1% SDS. Chromosomal DNA was then precipitated by addition of 50ml 3M KAc pH4.8, and pelleted by centrifugation at 4000g for 10 minutes. The supernatant (containing plasmid DNA) was filtered through muslin, and the plasmid DNA precipitated by addition of 0.6 volumes of isopropanol. The DNA was pelleted by centrifugation and dissolved in 9ml 100mM tris, 1mM EDTA pH8. 10g caesium chloride and 1ml of a 5mg/ml solution of ethidium bromide were added. The solution was transferred to a Beckman "quickseal" ultracentrifuge tube and the caesium chloride gradient centrifuged in a Beckman L8-55 ultracentrifuge for 16-30 hours at 55K. After equilibration, two bands were seen in the gradient; the lower one contained the plasmid DNA and was removed through the side of the tube using a syringe. This DNA was then diluted 3 times with water, precipitated with ethanol, pelleted, and dissolved in 500µl 100mM Tris, 0.1mM EDTA pH8. The solution was extracted twice with PCI and the DNA precipitated with ethanol and pelleted. The DNA was dissolved in 200µl of 100mM NaCl, 50mM tris pH7.5, 5mM EDTA, treated with 100µg/ml RNase for 30 minutes at 37°C, and the solution then extracted with PCI. The plasmid DNA was finally precipitated with ethanol, pelleted by centrifugation, and dissolved in 1ml 10mM tris pH 7.5. 1l of bacterial culture usually yielded 1mg plasmid DNA.

RNase A (from Bovine pancreas) was purchased from Sigma Chemical Company.

2.15 Preparation of plasmid DNA from bacterial minicultures

10ml bacterial minicultures were innoculated from single colonies of a bacterial transformation (section 2.8). Bacteria were pelleted by centrifugation at 4000g for 30 minutes, resuspended in 200µl 50mM glucose, 10mM EDTA, 25mM tris pH8, transferred to a 1.5ml snapcap Eppendorf tube, and lysed by addition of 400µl 0.2M NaOH,1% SDS. 200µl 3M KAc pH4.8 was added, and the precipitated chromosomal DNA pelleted by immediate centrifugation at 4000g for 10 minutes. 0.6 volumes of isopropanol was added to the supernatant, and the precipitated plasmid DNA pelleted. The DNA was dissolved in 200µl 100mM NaCl, 50mM tris pH7.5, 5mM EDTA, treated with 100µg/ml RNase for 30 minutes at 37^{o}C, and the sample then extracted with PCI. The DNA was finally precipitated with ethanol, pelleted by centrifugation and dissolved in 100µl 10mM tris pH 7.5.

10ml bacterial cultures yielded 10-20µg plasmid DNA which was subsequently analysed by restriction enzyme digestion and Southern blot hybridization for the presence of the desired construct.

2.16 Isolation of high molecular weight genomic DNA from tissue culture cells

The procedure was modified from that described by Jeffreys and Flavell, 1977. To avoid contamination of genomic DNA preparations by

plasmid DNA, freshly autoclaved solutions and disposable sterile plasticware were used throughout.

1-5x10^7 cells were pelleted by centrifugation at 1600g for 10 minutes, resuspended in 5ml 100mM NaCl, 50mM tris pH7.5, 5mM EDTA and lysed by addition of an equal volume of 4% SDS. The solution was treated with 100μg/ml proteinase K at 37°C for 4-16 hours, extracted twice with PCI, and the DNA precipitated with ethanol. The DNA was then dissolved in 5ml 10mM NaCl, 5mM tris pH7.5, 0.5mM EDTA and digested with 100μg/ml RNase for 1 hour at 37°C. SDS was added to a final concentration of 0.5%, the solution treated with 100μg/ml proteinase K for a further 2 hours at 37°C, and then extracted with PCI. The DNA was finally precipitated with ethanol and dissolved in 1-2ml 10mM tris pH7.5. 10^7cells usually yielded 200μg DNA.

Proteinase K was purchased from Boehringer Mannheim.

2.17 Restriction enzyme digestion of DNA

Restriction endonucleases were purchased from New England Biolabs or from Boehringer Mannheim. DNA was digested using the conditions recommended by the manufacturer for each enzyme.

In order to monitor digestion of genomic DNA, one tenth of the reaction mixture was removed before incubation and mixed with 0.5 μg λ-DNA. The λ-control and main genomic-digests were then incubated in parallel, and digestion of the genomic DNA was monitored by checking completion of digestion of the λ-DNA.

2.18 Agarose gel electrophoresis of DNA

DNA fragments of 0.1-30.0kb were separated by electrophoresis through 0.2-2% horizontal slab agarose gels (Aaij and Borst, 1972). Electrophoresis buffer (TBE) consisted of 0.089M tris borate, 0.089M boric acid, 0.002M EDTA and contained 0.5 μg/ml ethidium bromide. Gels were prepared, loaded and electrophoresed as described in Maniatis et al., 1982. DNA was visualized by ultraviolet (300nm) illumination (Sharp et al., 1973).

Restriction enzyme digests of λ (Daniells and Blattner, 1982), ØX174 (Sanger et al., 1977,1978), and pBR322 (Sutcliffe, 1978) DNAs were used as DNA fragment size markers (Helling et al., 1974).

Agarose (Type 1, low EEO) was purchased from Sigma Chemical Company.

2.19 Recovery of DNA from agarose gels

Agarose gels were made with "low melting agarose" when the electrophoresed DNA fragments were to be recovered. This agarose melts at 65°C, i.e. below the melting temperature of DNA. DNA was recovered from melted agarose by a modification of the method described by Weislander et al., 1979.

DNA was electrophoresed through 0.5-2.0% low melting agarose gels as described in Maniatis et al., 1982 and the appropriate DNA band excised under long wavelength ultraviolet illumination (300-360 nm). The gel slice was melted at 65°C and extracted three times using phenol saturated with 0.3M NaAc pH4.8. The DNA solution was

then extracted once with PCI and finally precipitated with ethanol.

Low melting agarose was purchased from FMC Corporation.

2.20 Polyacrylamide gel electrophoresis of DNA

^{32}P-labelled DNA fragments which were products of S1 nuclease mapping (section 2.27) and DNA sequencing experiments were analysed by electrophoresis through denaturing polyacrylamide-urea gels (Maniatis et al., 1975).

5-20% polyacrylamide gels of dimensions 30cm x 40cm x 0.3mm were poured, loaded and electrophoresed in TBE buffer (section 2.18) as described in Maniatis et al., 1982. Radioactive DNA fragments were detected by autoradiography (section 2.4). ^{32}P-labelled restriction enzyme digests of pBR322 and ØX174 were used as DNA fragment size markers.

2.21 Strand separation of DNA fragments on polyacrylamide gels

^{32}P-labelled DNA fragments of <2kb were strand-separated by electrophoresis through non-denaturing polyacrylamide gels (Szalay et al., 1977). Strand-separated DNA fragments were used as probes in S1 nuclease mapping experiments (section 2.27).

5-10% polyacrylamide gels of dimensions 30cm x 40cm x 0.5mm were poured, loaded and electrophoresed in 0.5xTBE (section 2.18) buffer as described in Maniatis et al., 1982. Separated DNA strands were located in the gel by autoradiography (section 2.4).

2.22 Recovery of DNA from polyacrylamide gels

^{32}P-labelled DNA was located in the gel by autoradiography and the appropriate DNA band excised as described in Maniatis et al., 1982. The DNA was eluted by leaving the gel slice to stand for 4-24 hours at room temperature in 500µl 0.5M NH_4Ac, 0.5% SDS, 1mM EDTA (Maxam and Gilbert, 1977). The gel slice was then removed from the buffer and discarded, and the remaining solution extracted with PCI. The eluted DNA was then precipitated with ethanol.

2.23 Southern blotting of DNA from agarose gels onto nitrocellulose filters

Restriction enzyme digests of cloned or genomic DNA were transferred from agarose gels to nitrocellulose filters using modifications of the methods described by Southern, 1975, 1980 and by Jeffreys and Flavell, 1977.

DNA digests were electrophoresed through 0.5-1.5% horizontal agarose gels as described in section 2.18 . Samples contained 0.2-1.0µg cloned DNA or 10-20µg genomic DNA. The gel was transferred to a glass tray and shaken in 0.25M HCl for 20 minutes at room temperature; this resulted in partial hydrolysis of the DNA by acid depurination and hence more efficient transfer of large DNA fragments. The DNA was then denatured by shaking the gel in 2-3 volumes of 1.5M NaCl, 0.5M NaOH for 1 hour at room temperature, and the gel neutralized by shaking for a further hour in 0.5M tris, 3.0M

NaCl pH 7.0. DNA was then transferred from the gel to a 0.1μm nitrocellulose filter using blotting apparatus as described in Maniatis et al., 1982. After transfer, nitrocellulose filters were baked for 2 hours at 85°C.

Nitrocellulose filters for Southern blotting were purchased from Schleicher and Schull.

2.24 Hybridization of Southern blot filters

Solutions:

20xSSC	3M NaCl, 0.3M NaCitrate pH6.15
100xDenhardt's	20g/l Ficoll 400 20g/l polyvinylpyrollidone (PVP) 20g/l bovine serum albumin (BSA)
Prehybridization	3xSSC 10xDenhardt's 0.1% SDS 50μg/ml denatured salmon sperm DNA
Hybridization	Prehybridization solution supplemented with 10% dextran sulphate
Wash solutions	1st 3xSSC, 0.1% SDS 2nd 1xSSC, 0.1% SDS 3rd 0.3xSSC, 0.1% SDS 4th 0.1xSSC, 0.1% SDS

Baked nitrocellulose filters from Southern blots were cut into strips of 15cmx3cm and wetted in 3xSSC. They were prehybridized by shaking in 50-100ml of prehybridization mixture for 2-3 hours at 65°C and

then transferred individually into 20-30 ml preheated (65°C) hybridization solution which contained 100-300 ng nick-translated DNA probe (specific activity >10^8cpm/μg DNA; section 2.33). Hybridization was for 12-24 hours at 65°C in plastic boxes. Filters were then washed to the desired stringency by shaking sequentially in the above wash solutions; each wash was for 30 minutes at 65°C in 500 ml solution. Filters were then rinsed in 3xSSC at room temperature and air-dried.

PVP-40, BSA (fraction V), and salmon sperm DNA were purchased from Sigma Chemical company. Dextran sulphate and Ficoll 400 were purchased from Pharmacia.

2.25 Colony hybridization

When transformed bacterial colonies (section 2.8) were to be screened by hybridization (Grunstein and Hogness, 1975), the transformation mixture was plated directly onto nitrocellulose filters on agar plates, and replica filters were prepared by filter-to-filter contact (Hanahan and Meselson, 1980). The procedures for plating and replica-plating using nitrocellulose filters, and for transfer of bacterial colonies from an agar plate to a nitrocellulose filter were as described in Maniatis et al., 1982.

Bacterial colonies were lysed on the nitrocellulose by placing the filter for 5 minutes on a sheet of Whatmann 3MM paper saturated with 0.5M NaOH. The filter was then floated for 5 minutes on 1M tris pH7.5 and then for a further 5 minutes on 0.5M tris, 1.5M NaCl. Cell debris were removed from the filter by wiping thoroughly with soft

Kleenex medical tissue soaked in 3xSSC, 0.1% SDS. The filter was rinsed in 3xSSC, baked for 2 hours at 85°C and hybridized to a nicktranslated probe as described in section 2.24.

2.26 Isolation of RNA from tissue culture cells

RNA was prepared using guanidinium chloride to simultaneously disrupt the cells and to inactivate nucleases (Chirgwin et al., 1979; Cox, 1968).

$1-5x10^7$ cells were pelleted by centrifugation for 10 minutes at 2000g, resuspended in 50–100µl 100mM NaCl, 50mM tris pH7.5, 5mM EDTA and lysed by addition of 500–1000µl 6M guanidinium chloride. The lysate was left on ice for 1–24 hours, extracted vigorously with PCI, and the phases separated by centifugation for 30 minutes at 1200g. The nucleic acids were precipitated with ethanol, pelleted by centrifugation, resuspended in 200µl 10mM tris pH7.5, 10mM $MgCl_2$, 50mM NaCl and treated with 1µg RNase free DNaseI for 10 minutes at 37°C. The solution was then extracted with PCI and the RNA precipitated with ethanol. 10^7 cells usually yielded 300–400µg RNA.

DNaseI was purchased from Sigma Chemical company. RNase was removed from commercial DNaseI by chromatography on agarose 5'-(4-aminophenylphosphoryl) uridine 2'(3')phosphate (Miles) as described by Maxwell et al., 1977.

2.27 SI nuclease mapping of RNA

Specific mRNAs were quantitated and mapped by SI nuclease analysis (Berk and Sharp, 1977) using end-labelled DNA probes (Weaver and Weissman, 1979; Favoloro et al., 1980).

^{32}P end-labelled double-stranded or single-stranded DNA probe (specific activity $>10^7$cpm/μg) was precipitated with ethanol and dissolved at a concentration of 0.2-1.0ng/μl in 40mM PIPES pH6.4, 1mM EDTA, 0.4M NaCl, 80% formamide. 10μl probe was added to 10-50μg dried RNA pellet in a 1.5ml snapcap Eppendorf tube, the RNA dissolved thoroughly by pipetting up and down, and the nucleic acids denatured by heating at 85°C for 10 minutes. The tube was then quickly transferred to a water bath at the desired hybridization temperature. Optimal hybridization temperatures were chosen so as to minimise the formation of DNA-DNA hybrids while allowing RNA-DNA hybrids to form (Casey and Davidson, 1977), and were determined empirically for each probe. The hybridization temperatures when using double-stranded and single-stranded probes were usually 52°C and 37°C respectively. Samples were hybridized for 4-24 hours and then digested for 2 hours at 20°C with 3000 units SI nuclease in 300μl 200mM NaCl, 30mM NaAc pH4.8, 2mM $ZnSO_4$. Samples were then extracted with PCI and the nucleic acids precipitated with ethanol. The products were electrophoresed on 5-15% denaturing polyacrylamide-urea gels (section 2.20).

SI nuclease was purchased from Boehringer Mannheim.

2.28 Ligation of DNA fragments

DNA molecules with compatible cohesive termini or with blunt ends were ligated using T4 DNA ligase (Weiss et al., 1968). Self-ligation of DNA fragments containing antibiotic resistance genes was prevented by their dephosphorylation prior to inclusion in a ligation reaction (section 2.29).

Ligations were carried out in 10-30µl 50mM tris pH7.4, 10mM $MgCl_2$, 10mM dithiothreitol, 1mM spermidine, 1mM ATP, 100µg/ml BSA. Reactions contained 100-200ng of each DNA fragment to be ligated and 10 units T4 DNA ligase. Ligation of fragments with cohesive ends was for 0.5-4 hours at 17°C; ligation of fragments with blunt ends was for 12 hours at 4°C. Ligation mixtures were added directly, without purification, to bacteria for transformation.

T4 DNA ligase was purchased from New England Biolabs.

2.29 Dephosphorylation of DNA

The terminal 5' phosphates were removed from DNA fragments either prior to labelling the 5' end with T4 polynucleotide kinase, or in order to prevent self-ligation. Lyophilized Calf Intestinal Phosphatase (Chaconas and van de Sande, 1980) was purchased from Boehringer Mannheim and was subsequently stored at a concentration of 500units/ml in 10mM tris pH9.5, 1mM $MgCl_2$, 1mM $ZnSO_4$, 50% glycerol.

0.1-5µg DNA was dephosphorylated by treating with 0.5 units phosphatase in 10-50µl 10mM tris pH9.5, 1mM spermidine, 0.1mM EDTA

for 1 hour at 37°C. Phosphatase was inactivated by incubating the reaction for a further 1 hour at 65°.

2.30 Labelling the 5' ends of DNA using T4 polynucleotide kinase

T4 polynucleotide kinase (Richardson et al., 1971) was used to label the 5' termini of DNA fragments to be used as probes in SI nuclease mapping (section 2.27).

0.1-5μg dephosphorylated DNA was end-labelled by treating with 10 units T4 polynucleotide kinase in 10-50μl 50mM tris pH9.5, 10mM $MgCl_2$, 5mM dithiothreitol, 1mM spermidine, 0.1mM EDTA, 20-100μCi γ^{32}P-ATP for 30 minutes at 37°C. Labelled DNA was separated from unincorporated γ^{32}P-ATP by chromatography on Sephadex G50 (section 2.34).

T4 polynucleotide kinase was purchased from P.L.Biochemicals Inc.

2.31 Filling in of 3' recessed termini of DNA fragments

3' recessed termini created by digestion of DNA with certain restriction endonucleases were filled in by treatment with the Klenow fragment of E.coli DNA polymeraseI (Jacobsen et al., 1974; Klenow et al., 1970) or with reverse transcriptase (Verma et al., 1977). These reactions were used to create blunt-ended molecules for ligation or to end-label DNA probes at the 3'end by filling in with α^{32}P-dNTPs.

a) using Klenow polymerase

0.1-1.0μg DNA was treated with 1 unit Klenow polymerase in 10-50μl 50mM tris pH7.8, 5mM $MgCl_2$, 10mM mercaptoethanol, 500μM dNTPs at 20°C for 1 hour. 20-100μCi α^{32}P-dNTP was included in 3' end-labelling reactions using Klenow polymerase.

Klenow polymerase was purchased from New England Biolabs.

b) using reverse transcriptase

0.1-1.0μg DNA was treated with 1 unit reverse transcriptase in 10-50μl 50mM tris pH8.3, 6mM $MgCl_2$, 40mM KCl, 100μg/ml BSA, 1mM dNTPs at 41°C for 1 hour. 20-100 μCi α^{32}dNTP was included in 3' end-labelling reactions using reverse transcriptase.

Reverse transcriptase was purchased from Anglian Biotechnology Ltd.

2.32 Removing 3' protruding termini of DNA fragments

3' protruding termini created by digestion of DNA with certain restriction endonucleases were removed by using the 3'-5' exonuclease activity of Klenow polymerase or by using SI nuclease. This created blunt-ended molecules for use in ligations.

a) using Klenow polymerase

0.1–1.0μg DNA was treated with 5 units Klenow polymerase in 10–50μl 50mM tris pH7.8, 5mM $MgCl_2$, 10mM mercaptoethanol at 20°C for 1 hour; dNTPs were then added to a final concentration of 500μM and incubated the reaction incubated for a further hour at 37°C.

b) using SI nuclease

1–10μg DNA was treated with 1000 units SI nuclease (Boehringer) in 100μl 200mM NaCl, 30mM NaAc pH4.8, 2mM $ZnSO_4$ at 20°C for 1 hour. DNA was then "repaired" using Klenow polymerase or reverse transcriptase as described in section 2.31.

2.33 Nick translation of DNA

DNA fragments were labelled with ^{32}P to a specific activity of $1–10 \times 10^8$ cpm/μg by nick translation using E.coli DNA polymerase I (Kelly et al., 1970; Rigby et al., 1977).

50–500ng DNA was labelled in a 16μl reaction containing 20μM dGTP, 20μM dCTP, 50mM tris pH7.8, 5mM $MgCl_2$, 10mM mercaptoethanol, 20μCi $\alpha^{32}P$-dATP, 20μCi $\alpha^{32}P$-dTTP, 2 units E.coli DNA polymerase I; the reaction was for 1–2 hours at 16°C and sometimes included an empirically determined amount of DNaseI (usually 0.5ng/ml). Under these reaction conditions, approximately 30% of labelled nucleotides were incorporated into the DNA, and labelled DNA strands were 400–800 nucleotides long. Nick-translated DNA was separated from unincorporated nucleotide by chromatography on

Sephadex G50.

E.coli DNA polymerase I was purchased from New England Biolabs.

2.34 Chromatography of DNA through Sephadex G50

After radioactive-labelling of DNA fragments using α^{32}P-dNTP or γ^{32}P-ATP (sections 2.30, 2.31, 2.33), DNA was separated from unincorporated nucleotide by chromatography through 2ml Sephadex G50 columns as described in Maniatis et al., 1982.

2.35 Northern blot hybridization

RNA samples were electrophoresed through formaldehyde gels, blotted onto nitrocellulose filters, and the filters hybridized to nick-translated probe using the method of Seed and Goldberg exactly as described in Maniatis et al., 1982.

2.36 Isolation of nuclei from tissue culture cells and in vitro nuclear transcription

Solutions Buffer A:

60mM	KCl
15mM	NaCl
0.15mM	spermine
0.50mM	spermidine
15mM	Na HEPES pH7.5

All sucrose solutions were made in Buffer A.

Nuclear storage buffer:

20mM	tris pH7.5
75mM	NaCl
0.50mM	EDTA

0.85mM DTT
0.12mM PMSF
50% glycerol

10^7-10^8 cells were pelleted by centrifugation at 1600g for 10 minutes, homogenized in 2 ml 0.3M sucrose using 10 strokes of a Dounce homogenizer, and centrifuged through 6 ml 30% sucrose for 10 minutes at 2.5K, 0^oC using a Sorvall HB4 rotor. Nuclei were resuspended in 2M sucrose and pelleted through this same solution by centrifugation at 30K for 1 hour, 4^oC using a Beckman SW50.1 rotor. The nuclei were then washed in nuclear storage buffer, pelleted by centrifugation for 5 minutes at 5K, 0^oC using a Sorvall HB4 rotor, and resuspended in nuclear storage buffer at a density of 10^6 nuclei/10µl. Nuclei were stored at -70^oC.

Nuclear transcription labelling reactions contained 10^7 nuclei in 100mM tris pH7.9, 50mM NaCl, 0.4mM EDTA, 0.1mM PMSF, 1.2mM DTT, 50mM NH_4SO_4, 4mM $MnCl_2$, 1mM GTP, 1mM ATP, 1mM CTP, 29% glycerol, 10mM creatine phosphate, 130µg/ml creatine phosphokinase, 300-500µCi ^{32}P-UTP in a 300µl reaction volume. Labelling was for 20 minutes at 26^oC.

2.37 Hybridization of labelled nuclear transcripts to immobilized DNA fragments on nitrocellulose filters

DNA fragments were spotted onto 0.45µm nitrocellulose filters using a Schleicher and Schull manifold apparatus as described in Kafatos et al., 1979. The filter was rinsed with water, placed onto the apparatus, and the dots rinsed through with 1M NH_4Ac. The DNA

sample was denatured by making it 0.3-0.4M in NaOH and leaving at room temperature for 10 minutes. The sample was then diluted with a equal volume of 2M NH_4Ac, and loaded onto the apparatus. Samples contained 1-2μg DNA fragment and were loaded in 20-50μl volume. After loading, the dots were rinsed with a few drops of 1M NH_4Ac. The filter was then rinsed thoroughly with 4xSSC, air dried and baked for 2 hours at 80°C. Individual nitrocellulose filter dots were cut out and hybridized to labelled nuclear transcripts as described in Groudine et al., 1984. Filters were prehybridized for 2 hours at 41°C. They were then hybridized at 41°C for 48 hours in a 100-400μl hybridization mixture which contained labelled nuclear RNA from 10^7 cells (approximately $2\text{-}10\text{x}10^7$ cpm). Hybridization was in 5ml polypropylene scintillation vial inserts. After hybridization, filters were washed 3 times in 0.3xSSC, 0.1%SDS at 54°C; each wash was for 15 minutes.

Prehybridization:	50% deionized formamide
	5xSSC
	5xDenhardt's solution (section 2.24)
	50mM sodium phosphate pH7
	0.1% SDS
Hybridization:	50% deionized formamide
	5xSSC
	1x Denhardt's solution (section 2.24)
	20mM sodium phosphate pH7
	0.1% SDS
	10% dextran sulphate

CHAPTER 3

REGULATED EXPRESSION OF THE HUMAN β-GLOBIN GENE FAMILY IN MURINE ERYTHROLEUKEMIA CELLS

3.1 Introduction

Murine erythroleukemia (MEL) cells are Friend virus transformed erythroid cells which are arrested at the proerythroblast stage of differentiation (section 1.2). Cultured MEL cells may be induced to complete erythroid maturation by treatment with a variety of chemicals such as dimethylsulphoxide (DMSO) or hexamethylene-bisacetamide (HMBA). This chemically triggered maturation closely resembles normal erythroid differentiation and results in an accumulation of adult-type mouse globin mRNA in the differentiated cell. A potential method for studying globin gene activation in erythropoiesis is therefore to analyse the expression of cloned globin genes after their introduction into MEL cells. Such experiments may ultimately give information about the DNA sequences which regulate β-globin gene activation and the mechanisms which control developmental switches in globin gene expression (sections 1.3, 1.5iv, 1.6).

In order to assess the feasibility of this approach, cosmids containing defined regions of the human β-like globin gene cluster were introduced into MEL cells, and expression of the foreign globin genes analysed during chemically induced erythroid differentiation. It was thus possible to determine:-

1. If expression of an introduced human globin gene was inducible upon differentiation of the host heterologous MEL cells.
2. If expression of the foreign human embryonic (ε), fetal (γ) and adult (β) globin genes behaved similarly during MEL differentiation.
3. If the expression of an introduced human globin gene was influenced by the presence of other genes and/or regions of the human β-like globin gene cluster i.e. if the genes were expressed in an independent manner.

3.2 Introduction of human globin gene cosmids into MEL cells

Cosmids containing specific regions of the human β-globin gene cluster (shown in Figure 5) were introduced into thymidine kinase negative (tk^-) MEL cells by calcium phosphate transformation and selection in HAT medium. Some of these cosmids were large recombinants (>40kb) containing several genes of the β-globin cluster (e.g. COS HG28TK, COS HG25); others contained only one gene (e.g. β-pRT, γ-pRT) together with several kilobasepairs of its flanking DNA. By analysing the expression of these cosmids during chemically induced differentiation of the host MEL cells it would be possible to determine:-

1. If the human β, γ, and ε-globin genes showed a similar pattern of expression during MEL differentiation when they were introduced together on cosmids containing 40kb regions of the gene cluster (COS HG28TK, COS HG25).

2. If the expression of the foreign human β-globin gene was influenced by the presence of other regions of the gene cluster (COS HG28TK, COS HG29TK, COS HGD17, COS HG6N, β-pRT), and likewise if expression of the human γ-globin gene was affected by the presence of other β-like globin genes (COS HG25, COS HG28TK, γ-pRT).

Before transformation, cosmid DNA was linearized within the vector sequences by cleavage with PvuI. Since linearized transforming DNA integrates into the host cell chromosome preferentially at its site of cleavage, linearization within vector sequences would reduce the probability of random integration and disruption of the globin gene region.

Stable tk^+ transformants were obtained at a frequency of 1-10 clones per 10^6 cells per μg DNA. Individual colonies were picked, cell cultures expanded, and DNA isolated and analysed for the presence of foreign human globin genes in the host MEL cells (section 3.3). 10-15 MEL transformants containing each cosmid were induced to differentiate by culturing for 3 days in the presence of 3mM hexamethylenebisacetamide (HMBA). RNA was prepared from differentiated and non-differentiated cells and expression of the endogenous mouse and foreign human globin genes analysed by SI nuclease mapping (sections 3.4-3.6).

3.3 Southern blot analysis of MEL transformants

The presence of human globin gene sequences in the tk^+ MEL transformants was shown by Southern blot hybridization. DNA from MEL transformants was digested with EcoRI, HindIII or BclI, transferred to a nitrocellulose filter, and hybridized to human globin gene probes. The relative intensities of the hybridizing bands were determined by densitometer tracing. The number of copies of the introduced human globin gene cosmid per cell was approximated by comparison of the intensity of the hybridizing bands in MEL transformant DNA with those in human genomic DNA and in known amounts of cloned human globin DNA. Figure 6 shows a Southern blot hybridization in which human γ-globin gene sequences were detected in HindIII digested DNA from MEL transformants containing cosmid COS HG25. The 3.0kb HindIII human $^{A}\gamma$-globin gene probe hybridized to both the 9.7kb $^{G}\gamma$-globin and the 3.0kb $^{A}\gamma$-globin gene HindIII fragments, and the intensity of the hybridizing bands showed the MEL transformants to contain 1-15 copies of the foreign human globin cosmid per cell. Approximately 2% of MEL clones did not contain cosmid DNA, i.e. they were tk^+ revertants of the tk^- cell line.

The hybridization of MEL transformant DNA to the human globin gene probes was shown not to be due to the presence of contaminating plasmid DNA in the samples by digestion with restriction enzymes which are sensitive to the methylation pattern acquired by plasmid DNA due to its propogation in E.coli. The enzyme BclI will not cleave plasmid DNA due to methylation by a bacterial methylase at the

sequence $G^{m}ATC$. It will, however, digest eukaryotic DNA and plasmid DNA which has replicated in eukaryotic cells and hence lost the prokaryotic methylation pattern. BclI will not, therefore, cleave cosmid COS HG28TK, whereas digestion of human placental DNA yields bands of 9.8 kb (β) and 4.8 kb (δ) which hybridize to a human β-globin gene probe (Figure 7). Figure 7 shows the detection of human β- and δ- globin gene sequences in EcoRI- and BclI-digested DNA from MEL transformants containing cosmid COS HG28TK. The presence of the 4.8 kb hybridizing band in BclI-digested DNA from the MEL clones shows them to be true human globin gene transformants; the 9.8 kb BclI band is not seen in the MEL transformant digests because the globin-vector junction of COS HG28TK lies within this fragment (Figure 5).

Hybridization of Southern blots to the human globin gene probes showed the presence in some MEL transformant DNA digests of "extra" hybridizing bands which were absent from cloned human globin or human genomic DNA digests. This indicated breakages within the globin gene fragments of the introduced cosmid, and could suggest integration of the human globin cosmid into the host cell chromosome or the formation of concatemers of transforming DNA. The introduced human globin cosmids were shown by Southern blot hybridizations to be non-rearranged in most of the MEL transformants.

3.4 Transcription of the endogenous mouse β-major globin gene in MEL transformants

MEL transformants were induced to differentiate by culturing for 3

days in the presence of 3mM HMBA. RNA was prepared from induced and non-induced cells, and levels of mouse β-major globin transcripts were measured by SI nuclease analysis using a probe which detected the 3' end of the transcript. Figure 9 shows SI nuclease analysis of mouse β-major globin transcripts in MEL transformants; hybridization of correctly terminated mouse β-major globin mRNA to the 390 nt probe yielded a 240 nt fragment which was protected from SI nuclease digestion. Levels of mouse β-major globin transcript were quantitated by densitometer tracing comparison of the intensity of the 240bp SI nuclease protected band in hybridizations of MEL transformant RNA and of standard amounts of mouse reticulocyte polyA$^+$ RNA (it was assumed that 1 cell contains 30 pg RNA and that 25% of mouse reticulocyte poly A$^+$ RNA is mouse β-major globin mRNA; G.Grosveld et al., 1982b). 64 out of 66 MEL transformants analysed showed a 100-200 fold increase in the level of mouse β-major globin mRNA upon differentiation. These levels corresponded to <100 copies of mouse β-major globin mRNA per noninduced cell and to >10,000 copies per induced cell. Six of the transformants which showed induction also had a high level of mouse β-major globin mRNA before HMBA treatment (>1000copies per cell; Table 1) suggesting that these cells may have undergone partial spontaneous differentiation in absence of the chemical inducer.

3.5 Transcription of introduced human adult (β) globin genes in MEL transformants

Human β-globin transcripts in MEL transformants were quantitated by

SI nuclease analysis using probes which detected the 5' and 3' ends of the mRNA. Under the hybridization and SI nuclease digestion conditions used, the human β-globin probes did not detect the endogenous mouse β-major globin transcripts; any heterologous hybrids which may have formed between mouse globin mRNA and the end-labelled human gene probe were cleaved by SI nuclease in the nonhomologous regions, and hence no discrete SI nuclease protected fragment was observed. Also, high levels of mouse β-globin mRNA in induced MEL cells did not, by competition for human β-globin gene probe, reduce the amount of hybridization of human β-globin mRNA to the probe, since all hybridizations were in vast DNA excess.

Of 40 MEL transformants that contained a foreign human β-globin gene, 30 showed a 4-200 fold increase in the level of human β-globin mRNA transcripts upon differentiation as measured by SI nuclease analysis of the 3' end of the mRNA (Table 1). Inductions of human β-globin mRNA in MEL transformants containing COS HG28TK, COS HG29TK, COS HGD17, COS HG6N and β-pRT are shown in Figures 10 and 11. Hybridization of correctly terminated human globin transcripts to the 700 nt EcoRI-MspI probe yielded a 212 nt fragment which was protected from SI nuclease digestion. Figures 9 and 8 show the corresponding induction of endogenous mouse β-globin mRNA in the same MEL transformants and Southern blot hybridizations of the clones. Both the percentage of MEL transformants showing inducibility and the levels of induced human β-globin mRNA were similar for each of the human β-globin gene cosmids; thus, cosmid β-pRT, which contains only 1.5kb DNA flanking the β-globin gene to the 5' and 3' sides, was as inducible as COS HG28TK which contained sequences spanning from the

fetal through to the adult region of the human β-like globin gene cluster. Of the ten transformants that showed no induction of human β-globin mRNA, two were noninducible for endogenous mouse β-globin mRNA (one COS HG6N and one β-pRT; Table 1), and four showed a high level of mouse β-globin mRNA before addition of HMBA (two COS HG6N and two COS HG28TK; Table 1). The level of induction of human β-globin mRNA did not correlate with the number of copies of the human β-globin gene in the transformant. Thus, of the inductions shown in Figure 10, that of the COS HG29TK clone was greater than that of the COS HGD17 clone, although the latter contained more copies of the human β-globin gene (Figure 8).

In order to determine whether the human β-globin transcripts in MEL transformants were correctly initiated, the 5' end of these mRNAs were mapped by SI nuclease analysis using a 330 nt HinfI probe which covered the transcription initiation site of the β-globin gene; hybridization of correctly initiated human β-globin mRNA to this probe yielded a 65 nt fragment which was protected from SI nuclease digestion (Figure 12). Human β-globin transcripts in all MEL transformants analysed were found to be correctly initiated and levels of mRNA as measured by 5' end SI nuclease mapping corresponded to those determined by analysis of the 3' end of the transcript. Figure 12 shows SI nuclease analysis of the 5' ends of human β-globin transcripts in three MEL transformants containing cosmid β-pRT; the 3' end SI nuclease analysis of the corresponding transformants is shown in Figure 11.

Human β-globin mRNA in MEL transformants was quantitated as described in section 3.4 using mRNA from a patient with sickle cell

anemia as a control. Non induced levels corresponded to 10-50 copies of human β-globin mRNA per cell; induced levels corresponded to 120-2500 copies mRNA per cell. For example, the COS HG29TK clone shown in Figure 10 shows a 50-fold induction of human β-globin mRNA and has 1500 copies mRNA per induced cell.

3.6 Transcription of introduced human fetal (γ) and embryonic (ε) globin genes in MEL transformants

Levels of human globin mRNA in MEL transformants containing the complete human γ-globin (COS HG25, COS HG28TK and γ-pRT) and ε-globin (COS HG25) genes were measured by SI nuclease analysis using probes which detected the 5' and 3' ends of the mRNA. Of 31 transformants containing the human γ-globin gene, 28 showed the same or a smaller amount of γ-globin mRNA after differentiation. Three transformants showed a <8 fold inducibility of γ-globin mRNA. Figure 13 shows SI nuclease analysis of human γ-globin mRNA in MEL transformants containing cosmids COS HG28TK and COS HG25, which contained the fetal-adult and embryonic-fetal regions of the gene cluster respectively; hybridization of correctly terminated γ-globin mRNA to the 600 nt EcoRI probe yielded a 168 nt fragment which was protected from SI nuclease digestion. SI nuclease analysis of γ-globin transcripts in MEL transformants containing cosmid γ-pRT is illustrated in Figure 14, in which the highest induction of γ-globin mRNA ever observed is shown. Human γ-globin mRNA was quantitated as described in section 3.4 by using RNA from a $\beta^{o}/\delta\beta^{o}$ thalassemic patient as a standard (in this patient, levels of

γ-globin mRNA were elevated; see section 1.4viii). Levels of human γ-globin mRNA in MEL transformants corresponded to 50-500 copies per cell; they were similar in MEL transformants containing each of the γ-globin gene cosmids (COS HG25, COS HG28TK, γ-pRT) and did not correlate with the number of copies of the human γ-globin gene in the clone.

All MEL transformants which contained the complete ε-globin gene (COS HG25) showed the same or reduced levels of ε-globin mRNA after differentiation (data not shown).

3.7 Induction of human β-globin mRNA in MEL transformants is due to transcriptional activation

The accumulation of mouse globin mRNA in differentiating MEL cells results from both an increase in the rate of globin gene transcription and an increase in the relative stability of globin mRNA (section 1.2). Since all inductions of the foreign human β-globin gene in MEL transformants were determined by SI nuclease analysis, which measures the steady state level of a specific transcript, it was necessary to determine whether the observed increase in human β-globin mRNA levels was due to an increase in rate of transcription of the foreign human β-globin gene. The relative rate of transcription of the foreign human β-globin gene was measured before and after differentiation using a nuclear transcription run-off assay (Groudine et al., 1981; Hofer, 1982). In this method, α^{32}P-UTP is incorporated into newly synthesised RNA in isolated nuclei, and labelled nuclear transcripts then hybridized to specific

specific DNA probes which are usually immobilized as dots on nitrocellulose filters. Since the DNA probes are present in vast excess over the specific globin gene transcripts, the amount of hybridizing labelled nuclear RNA is proportional to the rate of transcription of the gene. Using such a procedure, it has been shown that the rate of transcription of the mouse β-major globin gene in MEL cells increases upon differentiation (Hofer et al., 1982).

The relative rate of transcription of the foreign human β-globin gene was measured in an MEL transformant in which the induction and expression level of the foreign human β-globin gene was equal to that of the endogenous mouse β-major globin gene as determined by SI nuclease analysis (Figure 15, clone β-pRT-a) Labelled nuclear transcripts from this MEL transformant were hybridized to mouse and human globin DNA probes. The probes were isolated from the second intervening sequence of corresponding β-globin genes, since this is the region of their greatest sequence divergence. Hybridization of labelled nuclear transcripts from non-transformed MEL nuclei before and after differentiation is shown in Figure 15. Differentiation of the MEL cells resulted in an increase in hybridization to the mouse IVS2 probe, indicating an increase in the rate of transcription of the mouse β-major globin gene. In contrast, no increase in hybridization to the human IVS2 probe was observed, showing that the labelled mouse globin mRNA did not cross-hybridize with the human β-globin gene probe. Labelled nuclear transcripts from the human β-globin gene MEL transformant showed an increase in hybridization to both the mouse and human β-globin gene probes upon differentiation, indicating an increase in

rate of transcription of both the endogenous mouse and foreign human β-globin genes (Figure 15). The increased levels of human β-globin mRNA observed upon differentiation of the MEL transformants as measured by SI nuclease analysis therefore resulted at least in part from an increase in rate of transcription of the gene. As an additional control, the relative rate of transcription of the mouse and human globin genes was measured in a transformant which showed only a low induction of the endogenous mouse gene and no induction of the foreign human β-globin gene as measured by SI nuclease analysis (Figure 15, clone β-pRT-b). Upon differentiation of this β-pRT-b transformant, there was only a slight increase in hybridization to the mouse probe and no increase in hybridization to the human globin gene probe. As a control for background hybridization, labelled nuclear transcripts were also hybridized to an insulin gene probe, since this gene is not transcribed in either differentiated or nondifferentiated MEL cells (Figure 40).

3.8 Discussion

By introducing into MEL cells cosmid DNA containing defined regions of the human β-like globin gene cluster, it was shown that expression of the adult(β) but not fetal(γ) or embryonic(ε) gene was inducible during chemically triggered erythroid differentiation. Regulated expression of the human β-globin gene was observed when it was introduced either as part of the intact globin gene cluster or as an individual gene fragment containing only 1.5kb DNA flanking the gene to the 5' side; expression of the human γ-globin gene was similarly

unaffected by the presence of other regions of the gene cluster. Induction of the human β-globin gene resulted at least in part from an increase in the rate of transcription of the gene upon differentiation. The developmental stage specific human β-like globin genes thus behaved differently and independently, and the specific activation of the human adult β-globin gene mimicked globin gene activation during adult erythropoiesis.

Both the endogenous mouse and foreign human β-globin genes were inducible during chemically triggered erythroid differentiation in MEL cells. Expression of the human β-globin gene was regulated appropriately in the heterologous environment of the differentiating mouse cell even though it was introduced as naked DNA and probably did not acquire the particular chromatin constitution and DNA modification pattern obtained by the endogenous mouse globin gene by virtue of the passage of the MEL cell through the earlier stages of erythroid differentiation (see also Chapter 9). The inducibility of the human β-globin gene was thus an inherent property conferred by its DNA sequence, and the DNA sequences which regulate the specific activation of globin genes during erythroid differentiation are conserved between mouse and man.

The level of induction of the introduced human β-globin gene was usually about an order of magnitude lower than that of the endogenous mouse globin gene. This phenomenon may be explained by the fact that the human β-globin gene is activated less efficiently than the mouse counterpart in the heterologous murine erythroleukemia cells, or by the fact that expression of the introduced human globin gene may have been influenced by its site of integration into the host cell

chromosome. The chromatin configuration and expression level of the host cell chromosome at the site of integration of the foreign gene might influence both the acquired chromatin constitution and transcriptional activity of the latter. Several lines of evidence suggest that the level of transcription of a gene may be influenced by its chromosomal position. For example, Kioussis et al. (1983) have shown that the transcriptional inactivity of a human β-globin gene in a case of γβ thalassemia is due to a translocation which places the β-globin gene next to a region of chromosome which is in an "inactive" chromatin configuration (see also section 1.4vi and Chapter 5). Similarly, the integration of retroviral genomes into different regions of the host chromosome probably accounts for their different patterns of expression during development (Harbers et al., 1981; section 1.5v). The integration of the foreign human β-globin gene into different regions of the chromosome in MEL transformants might account for its variable expression and induction level, and differences in chromatin constitution between the endogenous mouse and foreign human β-globin genes may explain the lower induction level of the latter.

The expression level of the foreign human globin genes in MEL cells may also have been influenced by their introduction as cosmid recombinants containing the selectable HSV-tk gene. By introducing a HSV-tk-globin recombinant into tk^- L cells and selecting for tk^+ transformants, for tk^- revertants of these transformants and then for tk^+ rerevertants, Roginski et al. (1983) have shown that expression of the globin genes was activated and inactivated concomitantly with the activation and inactivation of the adjacent tk

gene during the selection processes. Selection for expression of the tk gene in the globin gene cosmids described here may explain why the level of expression of the foreign human β-globin gene in the nondifferentiated MEL transformant is higher than that of the endogenous mouse gene; it may similarly account for the low level of constitutive expression of the noninducible γ- and ε-globin genes (see below).

The induction in MEL cells of the foreign human globin genes was specific for the adult(β) versus embryonic(ε) or fetal(γ) genes when these genes were introduced either together as a gene cluster or separately on individual gene fragments. This differential induction specificity is also observed in the expression of the endogenous mouse adult versus embryonic globin genes, and presumably reflects the fact that the MEL cells are erythroid cells derived from an adult mouse. The low level of transcription of the introduced γ- and ε-globin genes both before and after differentiation was possibly due to the fact that these genes were adjoined to the selectable HSV-tk gene (see above). Differential induction of introduced human β-like globin genes in MEL cells was also observed by Willing et al. (1979) who introduced chromosome 11 of a human fibroblast cell line into MEL cells by cell fusion and observed specific inducibility of the human β- but not γ-globin genes. In these experiments it was not clear whether the chromatin constitution of the introduced DNA and/or the presence of an intact globin gene cluster influenced either the activation of the foreign human β-gene or the adult versus fetal induction specificity. For example, the transcriptional inactivity of the human β- and γ-globin genes in the noninduced MEL cell may have

been dependent on their introduction in an inactive fibroblastic chromatin configuration, with the specific induction of the β-gene reflecting specific chromatin changes upon differentiation. Data presented in this chapter confirm that the activation of foreign human β-like globin genes in MEL cells is specific for the adult versus fetal gene, and suggest that this specificity is conferred by the DNA sequences of the individual globin genes. The mechanisms by which such DNA sequences regulate gene activation are discussed in sections 9.2 and 9.3.

The induction in MEL cells of the introduced human β-globin gene was independent of the presence of the rest of the gene cluster. The foreign human β-gene was inducible when introduced as part of the intact gene cluster (COS HG28TK) or as an individual gene fragment (β-pRT). Similarly, the expression level of the foreign human γ-globin gene in MEL cells was unaffected by the presence of other regions of the cluster. Thus, although comparisons of the expression levels of the β- and γ-globin genes in various deletion thalassemia and HPFH patients had suggested that both the intergenic region between the $^{A}\gamma$- and δ-globin genes and the region 3' to the β-gene were important in regulating globin gene activation and inactivation during development (section 1.4viii), these regions did not affect expression of the human β-like globin genes when introduced into MEL cells. The position of the individual human β-like globin genes within the gene cluster did not influence either the activation of the adult β-globin gene or the lack of activation of the fetal γ-globin genes in the adult erythroid MEL environment, suggesting that the cluster organization of the human β-like globin genes may

not be relevant to globin gene activation and switching during development *in vivo*.

The demonstration of the specific inducibility in MEL cells of an introduced human β-globin gene suggests that the system will be valuable in studying the mechanisms which regulate β-globin gene activation during adult-type erythroid differentiation; it will also be of use in determining the lesion in certain types of thalassemia and HPFH (section 1.4viii, Ch. 5, Ch.9). The following chapters describe the localization of DNA sequences which regulate β-globin gene activation (Chs. 6, 7, 8) and an analysis of the expression of foreign non-globin genes in MEL cells (Chs. 4, 6).

	MOUSE β-MAJOR GLOBIN mRNA			HUMAN γ-GLOBIN mRNA		HUMAN β-GLOBIN mRNA	
COSMID	NO. OF CLONES TESTED	NO. SHOWING INDUCTION	NO. WITH HIGH LEVELS BEFORE DIFFERENTIATION	NO. OF CLONES TESTED	NO. SHOWING INDUCTION	NO. OF CLONES TESTED	NO. SHOWING INDUCTION
COS HG25	9	9	2	9	0	NA	NA
COS HGD17	8	8	0	NA	NA	8	6
COS HG28TK	10	10	2	10	2	10	7
COS HG29TK	9	9	0	9*	6*	4	4
COS HG6N	8	7	2	NA	NA	8	5
β-pRT	10	9	0	NA	NA	10	8
γ-pRT	12	12	0	12	1	NA	NA
TOTAL	66	64	6	31	3	40	30

*Transcription of the γ-globin gene in MEL transformants containing COS HG29TK initiates at a vector promoter (see Chapter 4) and is not included in the totals columns.

TABLE 1 INDUCTION OF HUMAN AND MOUSE GLOBIN mRNAs IN MEL TRANSFORMANTS CONTAINING HUMAN GLOBIN GENE COSMIDS

Figure 5

Structure of human globin cosmids introduced into MEL cells

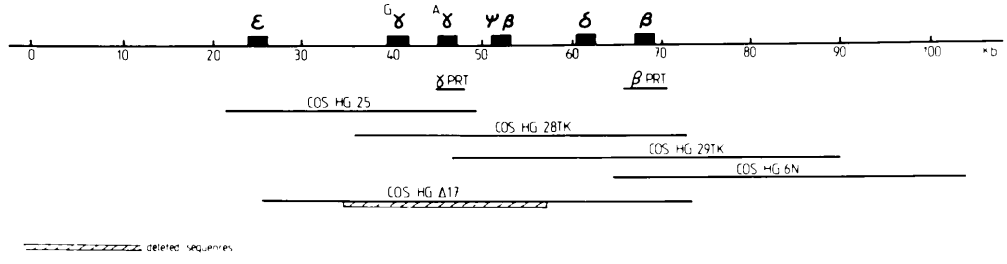
Upper diagram

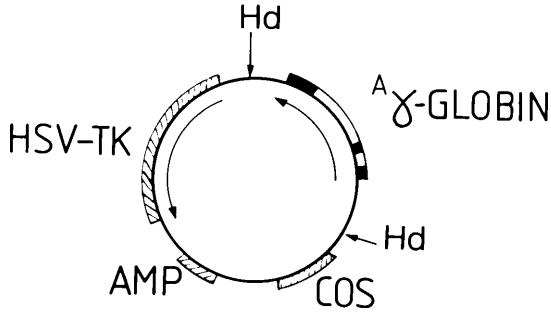
Cosmid recombinants contained the regions of the human β-like globin cluster indicated; vector sequences are not shown. Cosmids COS HG25, COS HG6N, COS HG28TK, COS HG29TK,and COS HGD17 were gifts of F.Grosveld (F.Grosveld et al., 1982,1983). COS HG25 and COS HG6N were isolated directly from a library constructed in the tk-vector pOPF. COS HG28TK and COS HG29TK were constructed from previously isolated cosmids and contained the HSV tk vector, pRT. COS HGD17 was constructed by packaging and transduction of a ligation mixture containing a ClaI-BamHI fragment from cosmid pRT, a BamHI-XmaI fragment containing the 3' side of the ε-globin gene, and an XmaI-PvuI fragment containing the $^{G}\gamma$, $^{A}\gamma$, δ and β globin genes together with part of the cosmid vector pJB8 (Ish Horowicz and Burke, 1981). Ligation yielded a DNA concatemer which deleted during the in vitro packaging to yield deletion cosmids including COS HGD17.

Lower diagram: Detailed structure of β-pRT-1 and γ-pRT-A

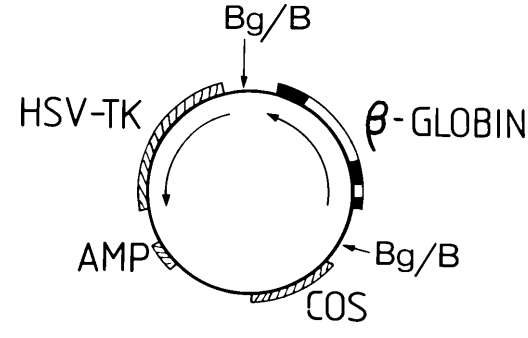
β-pRT-1 was constructed by inserting the 4.7 kb human β-globin BglII fragment into the BamHI site of pRT. γ-pRT-A was constructed by inserting the 3.0 kb $^{A}\gamma$-globin HdIII fragment into the HdIII site of pRT. β-pRT-1 and γ-pRT-A are referred to as β-pRT and γ-pRT in the text of Ch. 3. Arrows indicate the direction of transcription of the globin and HSV-tk genes.

Hd = HdIII restriction sites
Bg/B = sites of cloning of BglII fragment into BamHI site (both sites destroyed).


ε
Gγ
Aγ
ψβ
δ
β
0
10
20
30
40
50
60
70
80
90
100
γ PRT
β PRT
COS HG 25
COS HG 28TK
COS HG 29TK
COS HG 6N
COS HG Δ17


Hd
Aγ-GLOBIN
HSV-TK
Hd
AMP
COS

γ-pRT-A


Bg/B
HSV-TK
β-GLOBIN
AMP
Bg/B
COS

β-pRT-1

Figure 6

Southern blot analysis of MEL transformants containing the human globin cosmid COS HG25

DNA was digested with HdIII, electrophoresed on a 0.7% agarose gel, transferred to a nitrocellulose filter by Southern blotting and the filter hybridized to nick translated probe. The hybridization probe was the 3.0kb HindIII fragment isolated from the $^{A}\gamma$-globin gene. This probe hybridized to both $^{A}\gamma$- and $^{G}\gamma$-globin gene sequences.

Lanes 1-16 contain 10μg HdIII-digested DNA from individual MEL transformants.

Lane 17 contains 10μg HdIII-digested human placental DNA.

Lane C contains 10pg HdIII digested cloned COS HG25 DNA mixed with 10μg HdIII-digested untransformed MEL cell DNA.

HdIII-digested untransformed MEL cell DNA did not show any hybridizing bands (not shown).

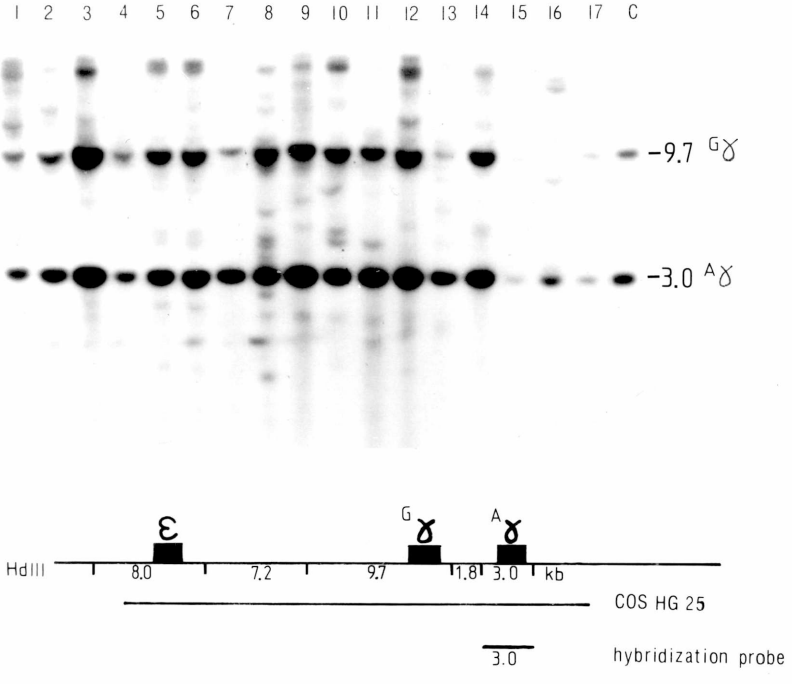

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 C
-9.7 Gγ
-3.0 Aγ
ε
Gγ
Aγ
HdIII
8.0
7.2
9.7
1.8
3.0
kb
COS HG 25
3.0
hybridization probe

Figure 7

Southern blot analysis of MEL transformants containing the human globin cosmid COS HG28TK

Upper Figure

DNA was digested with EcoRI or BclI, electrophoresed on a 0.7% agarose gel, transferred to a nitrocellulose filter by Southern blotting, and the filter hybridized to nick-translated probe. The hybridization probe was the 4.4kb PstI fragment isolated from the human β-globin gene. This probe hybridized to both human δ- and β-globin gene sequences.

Lane M contains 10μg EcoRI- or BclI-digested DNA from untransformed MEL cells.

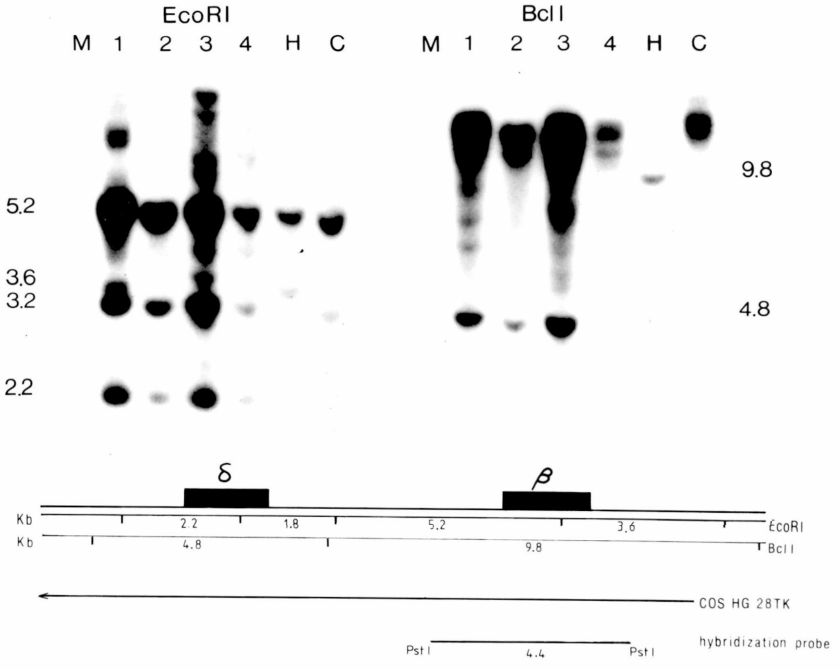
Lanes 1-4 contain 10μg digested DNA from individual MEL transformants.

Lane H contains 10μg digested human placental DNA.

Lane C contains 10pg digested cloned COS HG28TK DNA mixed with 10μg digested nontransformed MEL cell DNA.

Lower Figure

The filter containing the EcoRI digested samples above was rehybridized to the nick translated 1.6kb HindIII fragment isolated from the human $^{G}\gamma$-globin gene probe. This probe hybridized to both $^{A}\gamma$ and $^{G}\gamma$-globin gene sequences.


EcoRI
M 1 2 3 4 H C
BclI
M 1 2 3 4 H C
5.2
3.6
3.2
2.2
9.8
4.8
δ
β
Kb
2.2
1.8
5.2
3.6
EcoRI
Kb
4.8
9.8
BclI
COS HG 28TK
PstI
4.4
PstI
hybridization probe

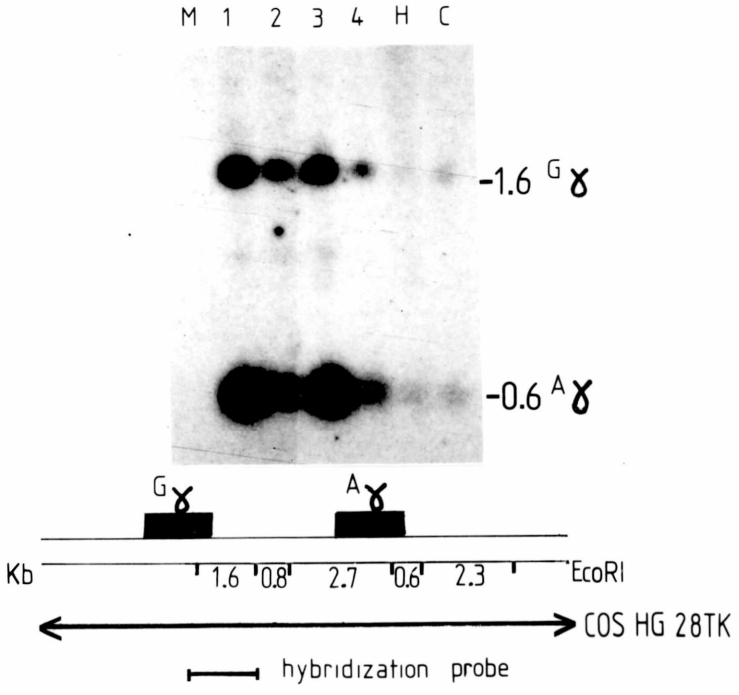

M 1 2 3 4 H C
-1.6 Gγ
-0.6 Aγ
Gγ
Aγ
Kb
1.6
0.8
2.7
0.6
2.3
EcoRI
COS HG 28TK
hybridization probe

Figure 8

Southern blot analysis of MEL transformants

DNA was digested with EcoRI, electrophoresed on a 0.7% agarose gel, transferred to a nitrocellulose filter by Southern blotting, and the filter hybridized to nick translated probe. The hybridization probe was the 4.4kb PstI fragment isolated from the human β-globin gene. This probe hybridized to both human δ- and β-globin gene sequences.

Lane 1 contains 10μg EcoRI-digested untransformed MEL cell DNA

Lanes 2-7 contain EcoRI-digested DNA from individual MEL transformants; these transformants correspond to those shown in Figures 9 and 10.

Lane 8 contains 10μg EcoRI-digested human placental DNA.

Lanes 9-12 contain 10pg EcoRI-digested cloned cosmid DNA mixed with 10μg digested untransformed MEL cell DNA.

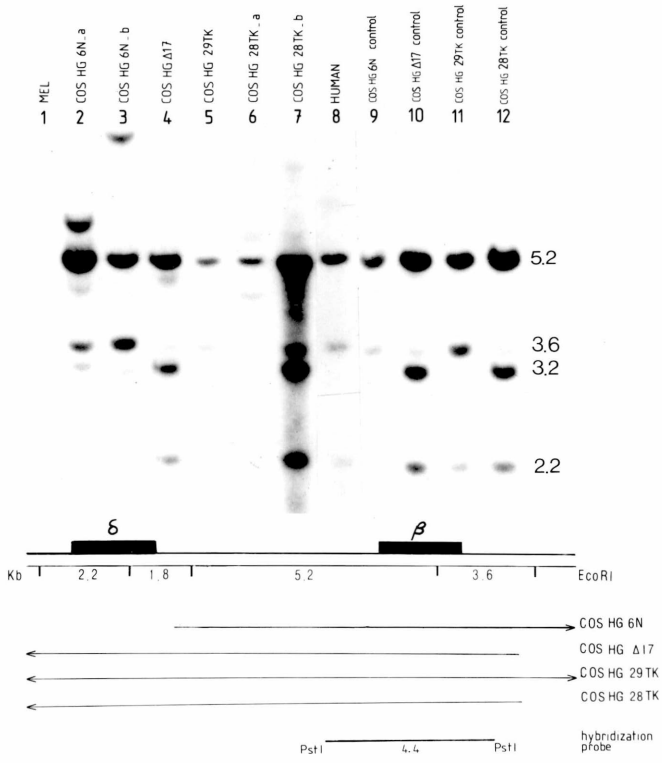

MEL
COS HG 6N. a
COS HG 6N. b
COS HG Δ17
COS HG 29TK
COS HG 28TK. a
COS HG 28TK. b
HUMAN
COS HG 6N control
COS HG Δ17 control
COS HG 29TK control
COS HG 28TK control
1 2 3 4 5 6 7 8 9 10 11 12
5.2
3.6
3.2
2.2
δ
β
Kb
2.2
1.8
5.2
3.6
EcoRI
COS HG 6N
COS HG Δ17
COS HG 29TK
COS HG 28TK
PstI
4.4
PstI
hybridization probe

Figure 9

SI nuclease analysis of mouse β-major globin gene transcripts in MEL transformants

Individual MEL transformants containing the cosmids indicated were grown for 3 days in the presence(+) or absence(-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method, and 1μg samples of total RNA were analysed by SI nuclease mapping using the mouse β-major globin gene 3' probe illustrated. The probe was end-labelled by reverse transcription to a specific activity of $>10^7$ cpm/μg. Hybridization was at 52°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control polyA$^+$RNA was obtained from mouse reticulocytes.

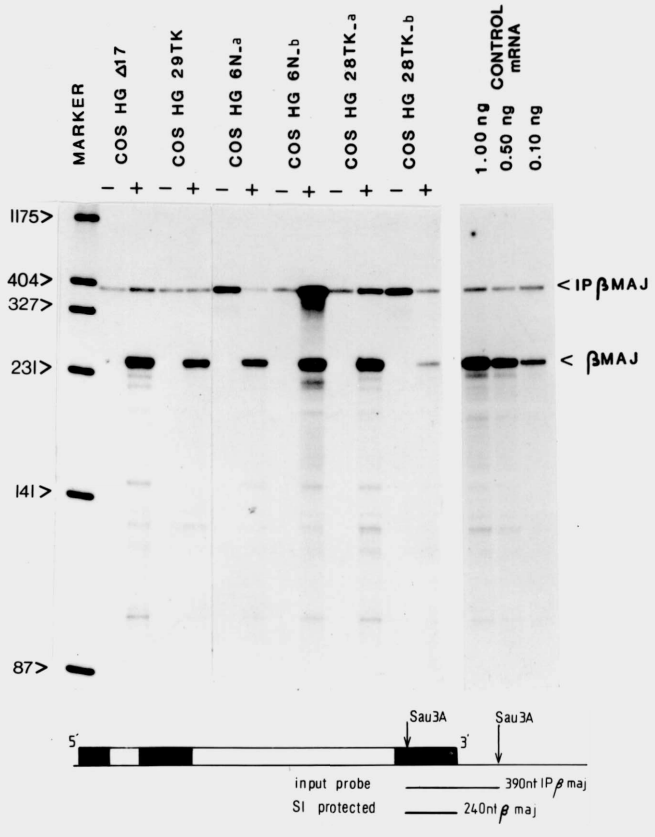

MARKER
COS HG Δ17
COS HG 29TK
COS HG 6N_a
COS HG 6N_b
COS HG 28TK_a
COS HG 28TK_b
CONTROL mRNA
1.00 ng
0.50 ng
0.10 ng
− + − + − + − + − + − +
1175>
404>
327>
231>
141>
87>
< IPβMAJ
< βMAJ
Sau3A
Sau3A
5'
3'
input probe
390nt IPβ maj
SI protected
240nt β maj

Figure 10

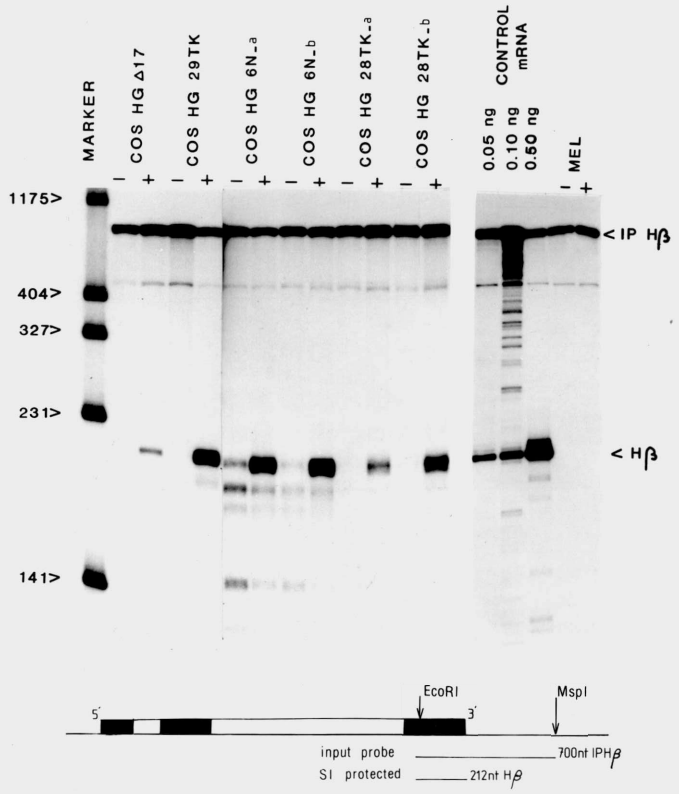
SI nuclease analysis of human β-globin gene transcripts in MEL transformants containing COS HG28TK, COS HG29TK, COS HGD17 and COS HG6N

Individual MEL transformants containing the cosmids indicated and corresponding to those shown in Figures 8 and 9 were grown for 3 days in the presence(+) or absence(-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 10μg samples of total RNA were analysed by SI nuclease mapping using the human β-globin gene 3' probe illustrated. The probe was end labeled by reverse transcription to a specific activity of $>10^7$ cpm/μg. Hybridization was at 52°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control polyA$^+$RNA was obtained from a patient with sickle cell anaemia.

Figure 11

SI nuclease analysis of human β-globin transcripts in MEL transformants containing β-pRT

Individual MEL transformants containing cosmid β-pRT were analysed as in Figure 10.


MARKER
COS HG Δ17
COS HG 29TK
COS HG 6N-a
COS HG 6N-b
COS HG 28TK-a
COS HG 28TK-b
CONTROL mRNA
0.05 ng
0.10 ng
0.50 ng
MEL
1175>
404>
327>
231>
141>
< IP Hβ
< Hβ
EcoRI
MspI
5'
3'
input probe
700nt IPHβ
SI protected
212nt Hβ

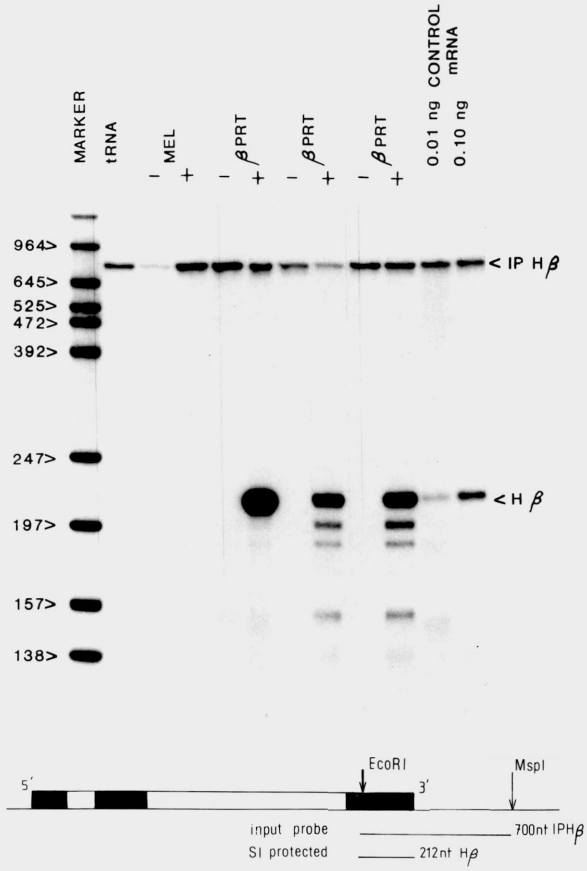

MARKER
tRNA
MEL
βPRT
βPRT
βPRT
CONTROL mRNA
0.01 ng
0.10 ng
964>
645>
525>
472>
392>
247>
197>
157>
138>
< IP Hβ
< Hβ
EcoRI
MspI
5'
3'
input probe
700nt IPHβ
SI protected
212nt Hβ

Figure 12

SI nuclease analysis of the 5' end of human β-globin gene transcripts in MEL transformants

Individual MEL transformants containing β-pRT and corresponding to those shown in Figure 11 were grown for 3 days in the presence(+) or absence(-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 100μg samples were analysed by SI nuclease mapping using the human β-globin gene 5' probe illustrated. The probe was end-labelled using T4 polynucleotide kinase to a specific activity of 10^7cpm/μg and then strand separated. Hybridization was at 37°C in a 40μl volume; SI nuclease digestion was for 1hour at 20°C followed by 1 hour at 40°C. Products were electrophoresed on a 7M urea 10% acrylamide gel. Control polyA$^+$RNA was obtained from a patient with sickle cell anaemia.

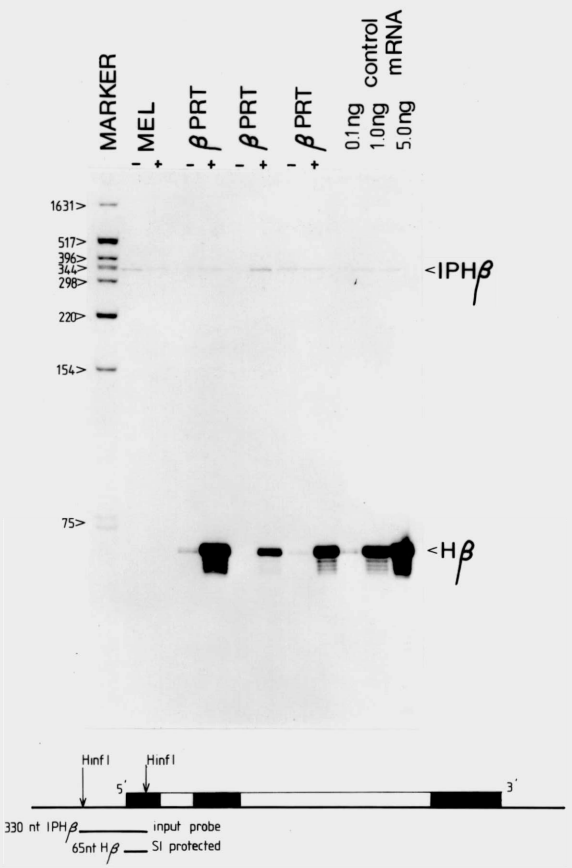

MARKER
MEL
- +
- β PRT
+
- β PRT
+
- β PRT
+
0.1ng control mRNA
1.0ng
5.0ng
1631>
517>
396>
344>
298>
220>
154>
75>
<IPHβ
<Hβ
Hinf I
Hinf I
5'
3'
330 nt IPHβ
input probe
65nt Hβ
SI protected

Figure 13

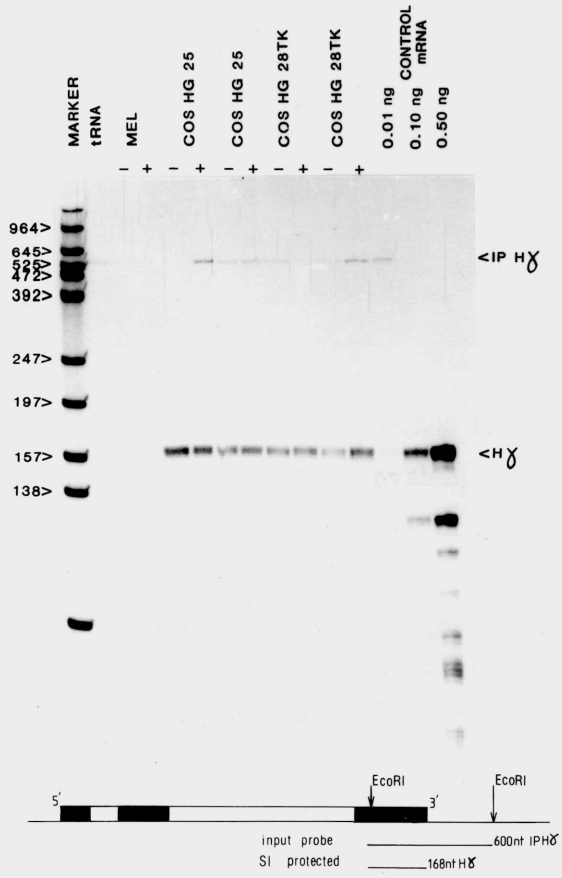
SI nuclease analysis of human γ-globin gene transcripts in MEL transformants containing COS HG25 and COS HG28TK

Individual MEL transformants containing the cosmids indicated were grown for 3 days in the presence(+) or absence(-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 10μg samples of total RNA were analysed by SI nuclease mapping using the human γ-globin gene 3' probe illustrated. The probe was end-labelled by reverse transcription to a specific activity of $>10^7$ cpm/μg. Hybridization was at 52°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control polyA$^+$RNA was obtained from patient with $\beta^o/\delta\beta^o$ thalassemia.

Figure 14

SI nuclease analysis of human γ-globin transcripts in MEL transformants containing γ-pRT

Individual MEL transformants containing cosmid γ-pRT were analysed as described in Figure 13. Each hybridization sample contained 30μg total RNA.


MARKER tRNA
MEL
COS HG 25
COS HG 25
COS HG 28TK
COS HG 28TK
0.01 ng
0.10 ng
0.50 ng
CONTROL mRNA
964>
645>
525>
472>
392>
247>
197>
157>
138>
<IP Hγ
<Hγ
EcoRI
EcoRI
5′
3′
input probe
600nt IPHγ
SI protected
168nt Hγ

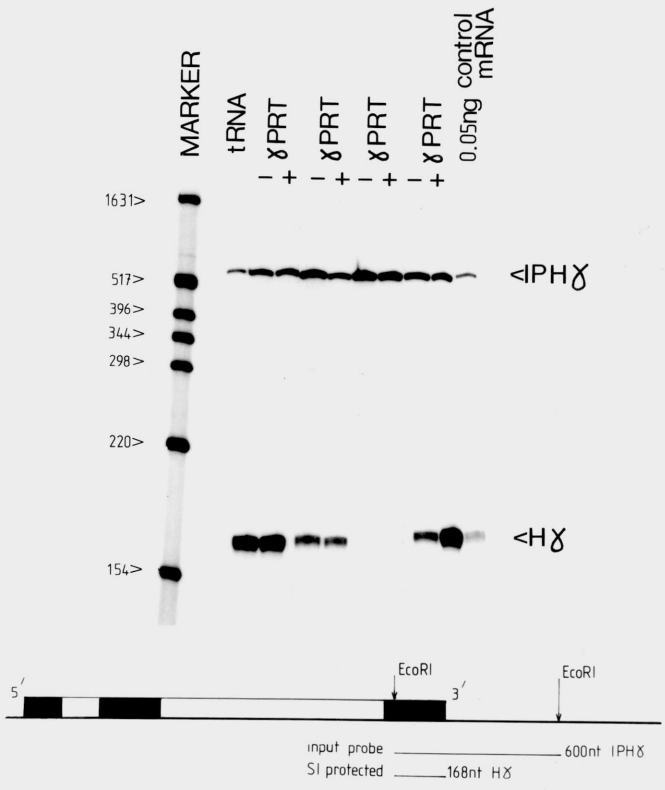
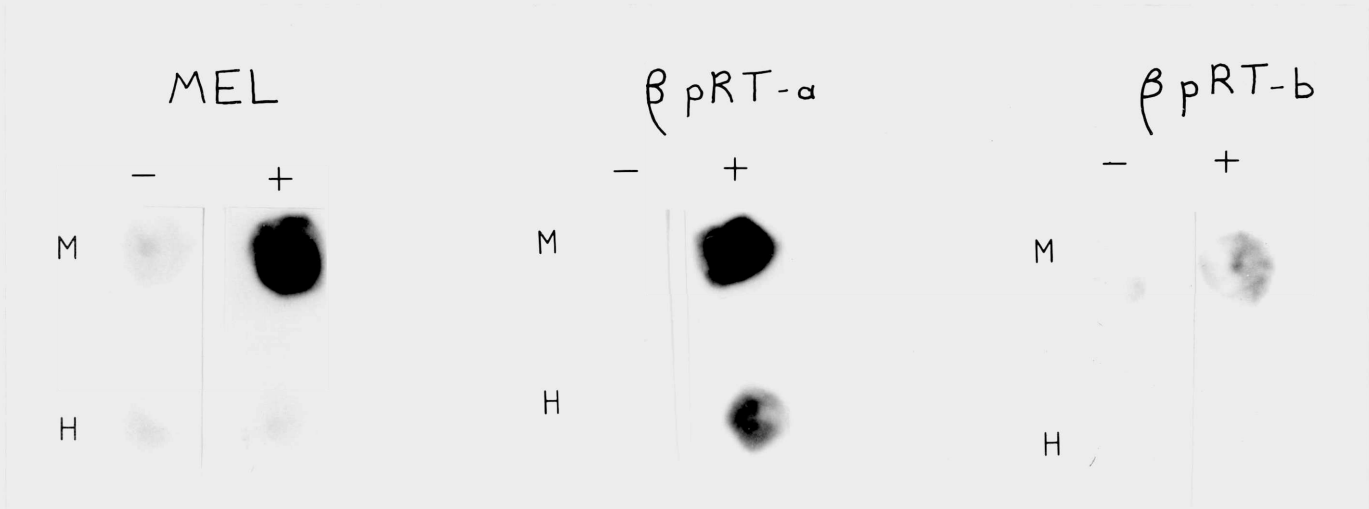

MARKER
tRNA
γPRT
γPRT
γPRT
γPRT
0.05ng control mRNA
1631>
517>
396>
344>
298>
220>
154>
<IPHγ
<Hγ
EcoRI
EcoRI
5′
3′
input probe
600nt IPHγ
SI protected
168nt Hγ

Figure 15

Dot hybridisation of labelled nuclear transcripts from MEL transformants containing human β-globin gene cosmid β-pRT

Nontransformed MEL cells and 2 β-pRT MEL transformants were grown for 3 days in the presence (+) or absence (-) of 3mM HMBA and nuclei isolated. Labelled nuclear transcripts were synthesized and hybridized to human β-globin (H) and mouse β-major globin (M) gene probes which were immobilized as dots on nitrocellulose filters. The human β-globin gene probe was the EcoRI-BamHI IVS2 fragment; the mouse β-major globin probe was the HdIII-PstI IVS2 fragment. Each hybridisation contained labelled RNA from 10^7 cells.


MEL
βpRT-a
βpRT-b
−
+
M
H

CHAPTER 4

REGULATED TRANSCRIPTION FROM A HERPES SIMPLEX VIRUS PROMOTER IN MURINE ERYTHROLEUKEMIA CELLS

4.1 Introduction

This chapter describes the identification and characterization of a Herpes Simplex Virus (HSV) transcript which initiated in the 5' flanking region of the HSV tk gene in the cosmid vector pRT and which was inducible during differentiation in MEL transformants. The transcript was detected in MEL transformants containing the human globin cosmids in which this HSV tk region of the pRT vector was adjoined directly to the 3' end of a human globin gene (COS HG29TK, COS HGD17). The induction of foreign genes in MEL cells was not, therefore, specific for globin, and DNA sequence comparison of the different inducible globin and non-globin genes may lead to the identification of sequences required for gene activation and to the characterization of specific classes of genes whose expression in MEL cells is regulated in a differentiation-dependent manner.

4.2 Identification of inducible human globin gene transcripts in MEL transformants containing truncated human γ- and ε-globin genes

Cosmids COS HG29TK and COS HGD17 contained the complete human β-globin gene together with the 3'end only of the human $^{A}\gamma$- and

ε-globin genes respectively (Figure 16). Cosmid COS HG29TK contained the 3'end of the human $^{A}\gamma$-globin gene joined at the BamHI site of its 2nd exon to HSV sequences of the vector pRT (Figure 16); the promoter and 5' end of the $^{A}\gamma$-globin gene were therefore absent from the construct. COS HGD17 contained the 3' end of the ε-globin gene joined in a likewise manner to HSV vector sequences (Figure 16). The introduction of COS HG29TK and COS HGD17 into MEL cells and analysis of transformants for expression of the human β-globin gene was described in sections 3.2, 3.3 and 3.5.

RNA from 9 MEL transformants containing COS HG29TK was also analysed by SI nuclease mapping using a probe which detected specifically the 3' end of human γ-globin mRNA (section 3.6). Surprisingly, 6 transformants showed a 10-30 fold induction of transcripts which hybridised to this probe (Figure 17). Since the normal human $^{A}\gamma$-globin gene promoter was absent, these inducible transcripts must have initiated either randomly or specifically within vector or host MEL cell DNA sequences and ended at the human $^{A}\gamma$-globin gene termination site. Similarly, MEL transformants containing COS HGD17 showed an induction of RNAs which were detected by SI nuclease analysis using a probe which hybridised to the 3' end of human ε-globin mRNA (Figure 18).

4.3 Northern blot analysis of the inducible hybrid globin transcripts

In order to determine the size and nature of the inducible human $^{A}\gamma$- and ε-globin gene transcripts detected in MEL transformants containing COS HG29TK and COS HGD17, RNA from these transformants was

analysed by Northern blot hybridization.

RNA was isolated from nine COS HG29TK MEL transformants both before and after differentiation, electrophoresed through a formaldehyde gel and transferred to a nitrocellulose filter by Northern blotting. These filters were hybridized to a human A_γ-globin gene and a vector DNA probe. Both probes detected an inducible RNA of approximately 700bp which did not comigrate with endogenous mouse globin mRNA. Similarly, Northern blot analysis of MEL transformants containing COS HGD17 showed the presence of a 700bp inducible transcript which hybridized to both ε-globin and to vector DNA probes. It therefore seemed probable that the inducible RNA was a hybrid vector-globin transcript. If it is assumed that the only RNA splicing occuring during maturation of this transcript is removal of the globin gene second intervening sequence, the RNA length observed on the Northern blot hybridization places the transcription initiation site close to that of the HSV tk gene of the vector, with transcription of the hybrid vector-globin mRNA and of tk mRNA being in the opposite direction.

4.4 Identification of the transcription initiation site of the inducible hybrid vector-globin RNA

The transcription initiation site of the inducible vector-globin hybrid RNA as indicated by Northern blot analysis was confirmed by SI nuclease mapping using a 510bp XmaI-BglII probe which covered the predicted 5' end of the transcript (Figure 17 and 18). Hybridization of the inducible RNA to this probe yielded a 280-300bp fragment which

was protected from SI nuclease digestion, thus mapping the 5' end of the hybrid transcript ("a-tk transcript") to be approximately 160bp away from the HSV tk gene cap site (Figure 19). Such an SI nuclease analysis does not, however, exclude the possibility that a small segment of RNA not detected by the probe is spliced onto the presumed 5' end of this inducible transcript.

Cosmids COS HG28TK, COS HG25, and COS HG6N also contained the region of HSV vector DNA in which the inducible hybrid transcript of COS HG29TK and COS HGD17 initiated. In the former cosmids, however, this HSV region was not linked to any known coding sequences (Figure 5), and no inducible transcript initiating from it was detectable by SI nuclease analysis using the probe which hybridized to the 5' end of the hybrid transcript of COS HG29TK and COS HGD17. This suggests that either an RNA splice site or a transcription terminator are required to yield the stable inducible transcript.

4.5 Discussion

It was surprising to find a transcript which initiated in the 5' flanking region of the HSV tk gene and which was inducible in MEL cells. Examination of the nucleotide sequence of this region of HSV DNA (McKnight, 1980) shows the presence of ATA and CAAT box promoter sequences (section 1.4iv, Figure 19) at the appropriate distances 5' to the initiation site of the inducible transcript. This indicates that this region of HSV DNA may be a "true" promoter and that transcription from it may occur physiologically in the viral life cycle, i.e. the inducible RNA observed in MEL cells was not merely

an aberrant transcription artifact of the system. A transcript initiating precisely at this site has also been observed in experiments showing *in vitro* transcription of HSV DNA fragments (Sullivan-Reed and Summers, 1983). Similarly, Wagner *et al*. (1983) have recently reported the identification of a late viral transcript initiating within this region; this viral RNA may therefore be that which was detected in the MEL transformants. Such a postulate may be tested by introducing total HSV DNA into MEL cells and determining whether a complete HSV transcript which initiates in the same region as the inducible hybrid HSV-globin transcript is present. The expression of viral genes is controlled by proteins produced both by the host cell and by the virus itself, and the finding of a HSV transcript which is regulated by proteins in induced MEL cells is not inconceivable. If this "anti-tk" promoter is indeed used naturally in the Herpes virus life cycle, it is possible that it normally responds to a trans-acting factor (produced either by the host cell or the virus itself) similar to the presumptive factor(s) which is present in induced but not in uninduced MEL cells and which activates expression of the cellular globin genes during the later stages of erythroid differentiation. Identification of the proteins which regulate expression of this Herpes inducible gene could be of use, by virtue of their possible homology with the putative MEL cell factor, in isolation of factors involved in regulating gene activity during erythroid differentiation. Also, the factors which physiologically regulate the expression of the HSV transcript might influence the expression of cloned globin genes in eukaryotic cell expression systems; this could be studied by analysing the expression of globin

genes introduced into viral -infected or -transformed cells. Such factors might act by a similar mechanism by whereby the Adenovirus immediate-early E1A protein, which is responsible for activating late viral gene expression *in vivo*, activates the expression of a foreign cloned globin gene when provided *in trans* in a transient expression assay (Green *et al.*, 1983; Imperiale *et al.*, 1983; section 1.5iii)

The identification of the inducible Herpes transcript showed that the activation of foreign genes in MEL cells was not specific for globin. Figure 20 shows a DNA sequence comparison of the 5' flanking region of the different globin and non-globin genes which have been shown to be inducible in MEL cells. All such genes share a block of homologous sequences 5' to the ATA box. The possibility that these sequences regulate gene activation in MEL cells was therefore tested by studying the expression in MEL cells of introduced mutant globin genes which have deletions in their 5' flanking region (Chapter 6).

Figure 16

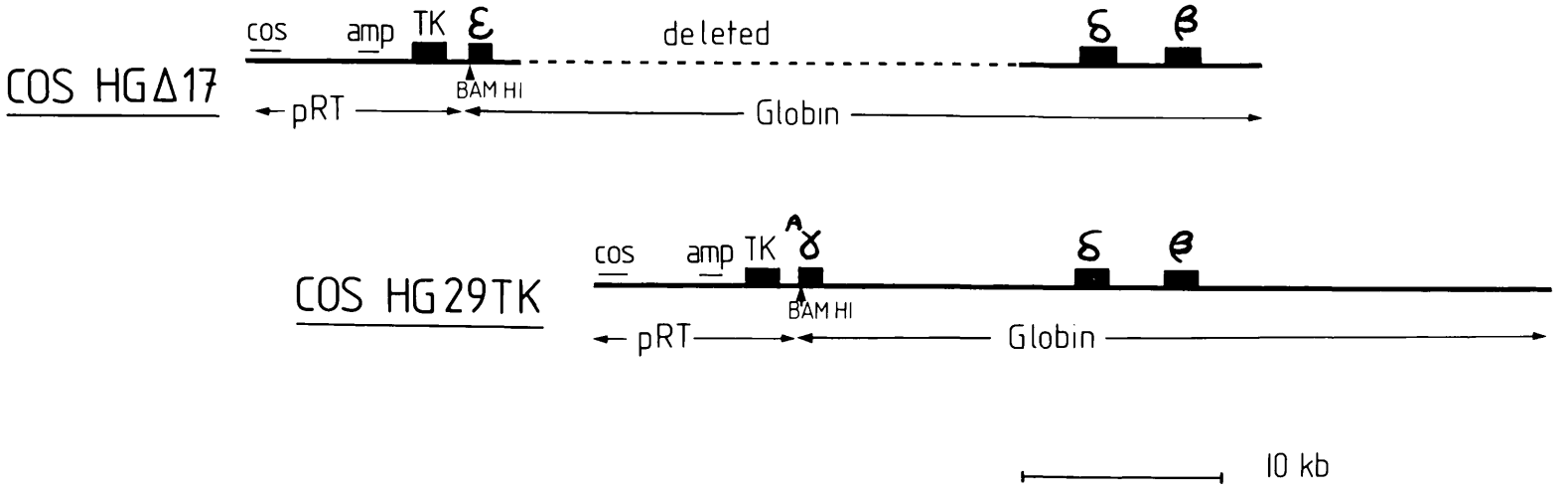
Structure of the globin-HSV junction region in cosmids COS HG29TK and COS HGD17

Upper Figure

Cosmids COS HGD17 and COS HG 29TK contained regions of the human β-globin cluster cloned into the HSV-tk vector pRT as indicated (see also Figure 5).

Lower Figure

Detailed structure of the vector-globin junction in cosmids COS HG29TK and COS HGD17. At the globin-vector junction, COS HG29TK contained the 3' end of the human $^{A}\gamma$-globin gene joined at the BamHI site of its 2nd exon to HSV sequences of pRT; COS HGD17 contained the 3' end of the human ε-globin gene joined in a likewise manner at this position. The direction of transcription of the HSV tk gene is indicated.


COS HGΔ17
cos
amp
TK
ε
deleted
δ
β
BAM HI
pRT
Globin
COS HG29TK
cos
amp
TK
Aγ
δ
β
BAM HI
pRT
Globin
10 kb

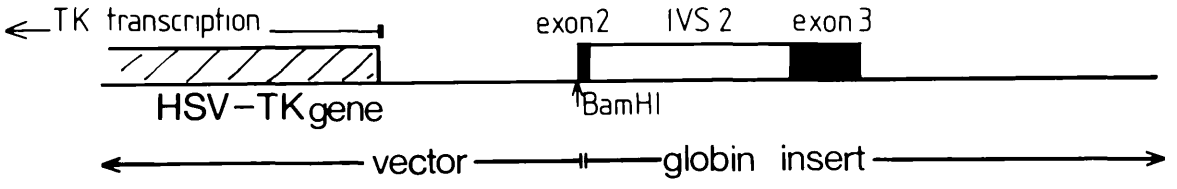
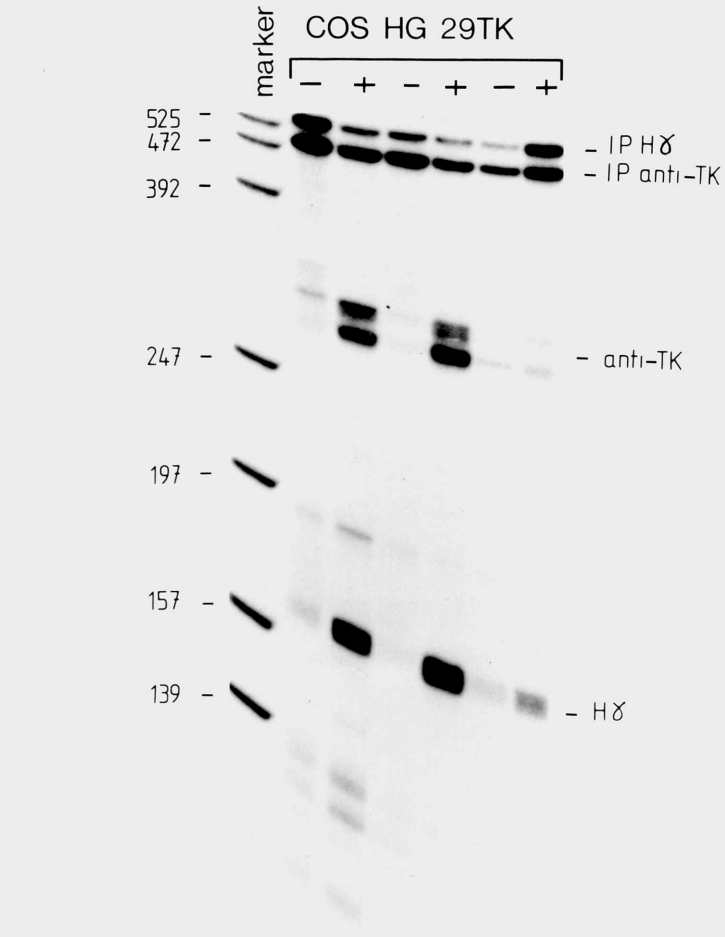

TK transcription
exon2
IVS 2
exon 3
HSV-TK gene
BamHI
vector
globin insert

Figure 17

SI nuclease analysis of hybrid vector-globin transcripts in MEL transformants containing COS HG29TK

MEL transformants containing COS HG29TK were grown for 3 days in the presence (+) or absence (-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 10μg samples of total RNA were analysed by SI nuclease mapping using the probes illustrated. The 600nt $^{A}\gamma$-globin gene probe mapped the 3' end of the hybrid transcript; it was labelled by reverse transcription to a specific activity of > 10^7 cpm/μg. The 510nt XmaI-BglII probe mapped the 5' end; it was labelled using T4 polynucleotide kinase to a specific activity of 10^7 cpm/μg. The 5' end of the inducible transcript initiating in HSV sequences is referred to as "anti-tk" RNA. Both probes were used simultaneously in the SI hybridization. Hybridization was at 52°C and SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel.


marker
COS HG 29TK
– + – + – +
525
472
392
– IP Hγ
– IP anti-TK
247
– anti-TK
197
157
139
– Hγ

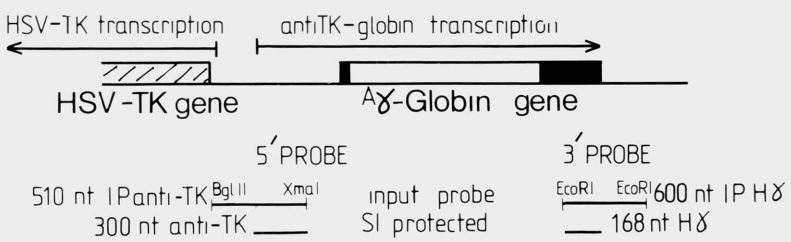

HSV-TK transcription
antiTK-globin transcription
HSV-TK gene
Aγ-Globin gene
5′ PROBE
3′ PROBE
510 nt IPanti-TK
Bgl II
Xma I
input probe
EcoRI
EcoRI
600 nt IP Hγ
300 nt anti-TK
S1 protected
168 nt Hγ

Figure 18

SI nuclease analysis of hybrid vector-globin transcripts in MEL transformants containing COS HGD17

MEL transformants containing COS HGD17 were grown for 3 days in the presence (+) or absence (-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 10μg samples of total RNA were analysed by SI nuclease mapping using the probes illustrated. The 190 nt ε-globin gene probe mapped the 3' end of the hybrid transcript; it was labelled by reverse transcription to a specific activity of $>10^7$ cpm/μg. The 510nt XmaI-BglII probe mapped the 5' end; it was labelled using T4 polynucleotide kinase to a specific activity of 10^7 cpm/μg. The 5' end of the inducible transcript initiating in HSV sequences is referred to as "anti-tk" RNA. Both probes were used simultaneously in the SI hybridization. Hybridization was at 52^oC and SI nuclease digestion was for 2 hours at 20^oC. Products were electrophoresed on a 7M urea 7% acrylamide gel.

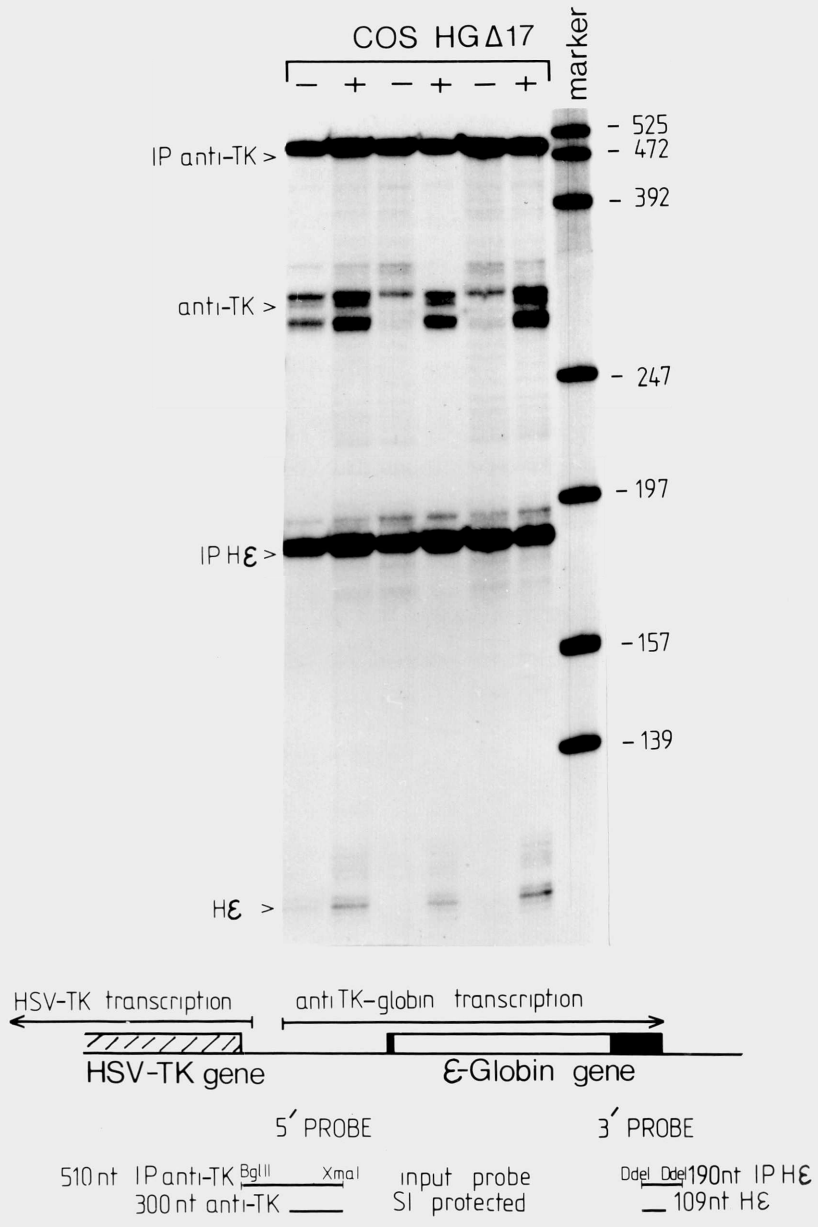

COS HGΔ17
marker
– + – + – +
IP anti-TK >
anti-TK >
IP Hε >
Hε >
525
472
392
247
197
157
139
HSV-TK transcription
anti TK-globin transcription
HSV-TK gene
ε-Globin gene
5′ PROBE
3′ PROBE
510 nt IP anti-TK
300 nt anti-TK
BglII
XmaI
input probe
SI protected
DdeI DdeI 190nt IP Hε
109nt Hε

Figure 19

Nucleotide sequnce of HSV DNA at the transcription initiation site of the inducible hybrid vector-globin transcript

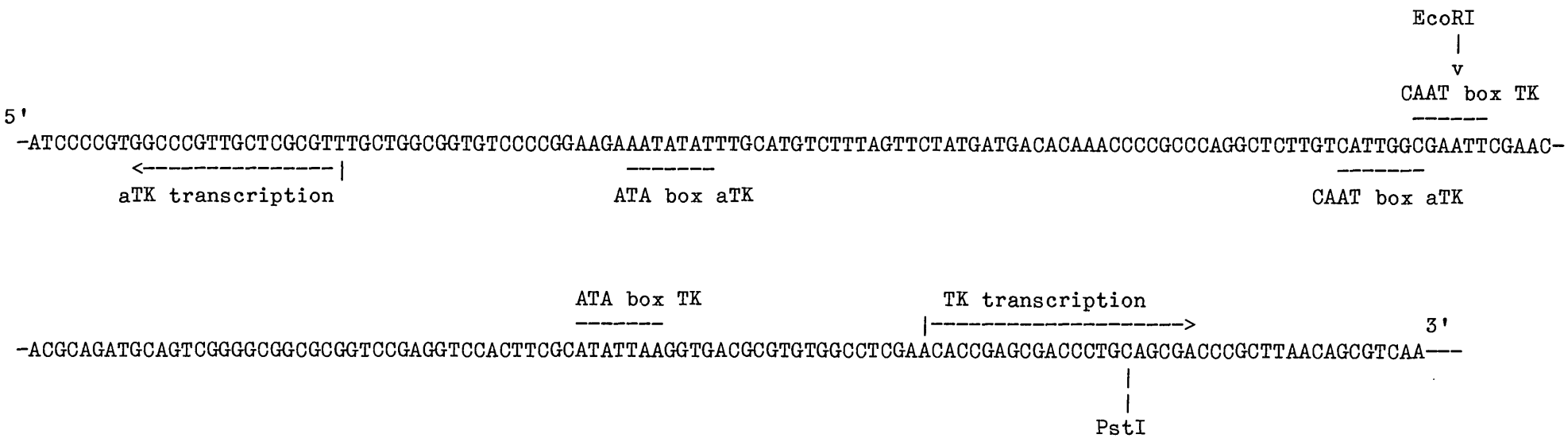
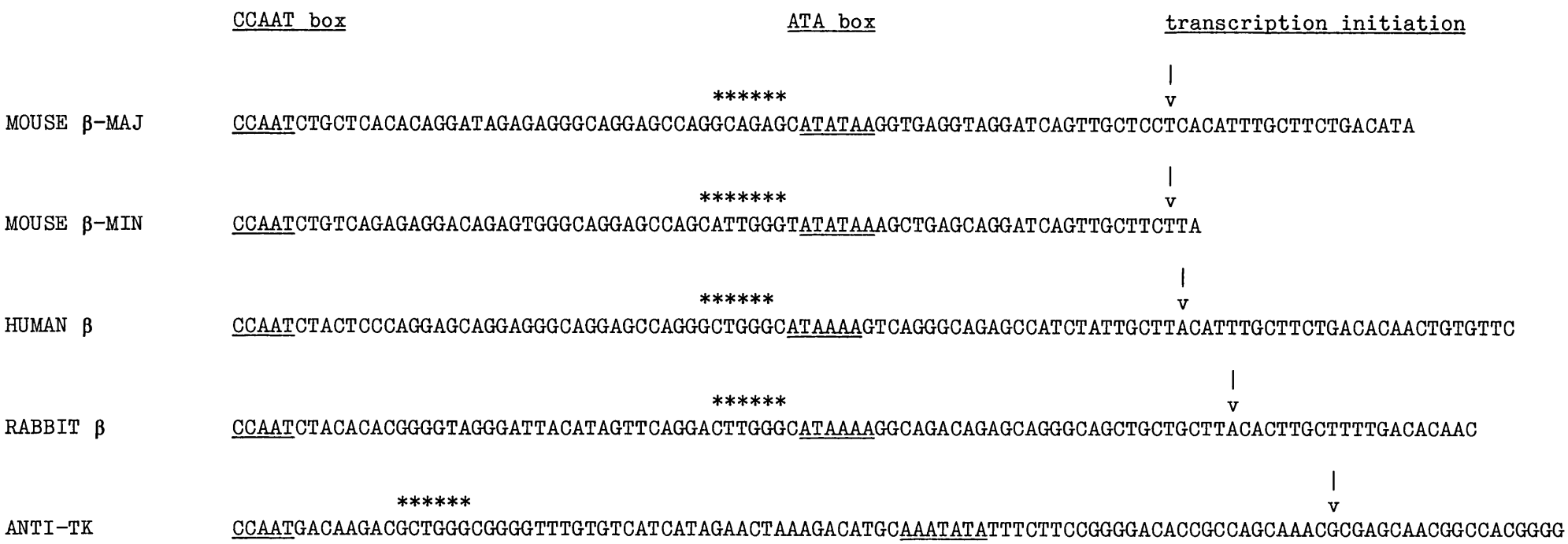

EcoRI
CAAT box TK
5'
-ATCCCCGTGGCCCGTTGCTCGCGTTTGCTGGCGGTGTCCCCGGAAGAAATATATTTGCATGTCTTTAGTTCTATGATGACACAAACCCCGCCCAGGCTCTTGTCATTGGCGAATTCGAAC-
aTK transcription
ATA box aTK
CAAT box aTK
ATA box TK
TK transcription
3'
-ACGCAGATGCAGTCGGGGCGGCGCGGTCCGAGGTCCACTTCGCATATTAAGGTGACGCGTGTGGCCTCGAACACCGAGCGACCCTGCAGCGACCCGCTTAACAGCGTCAA---
PstI

Figure 20

Nucleotide sequence of the promoter region of genes showing inducibility in MEL cells



**** indicates region of homology
Positions of CCAAT box, ATA box and transcription initiation are shown

CHAPTER 5

EXPRESSION OF A CLONED β-GLOBIN GENE FROM A γβ-THALASSEMIC PATIENT IN MURINE ERYTHROLEUKEMIA CELLS

5.1 Introduction

The expression of cloned human β-globin genes in MEL cells offers a useful system in which to determine whether the mutant phenotype of a thalassemic globin gene is due to a defect in the regulatory DNA sequences which control its activation during erythroid differentiation. γβ-thalassemia is characterized by a lack of expression of both the γ- and β-globin genes of the mutant allele (section 1.4viii). The disease is fatal in the homozygote form, and in heterozygotes results in a severe anemia in the newborn and a mild β-thalassemia in later life. All three cases of γβ-thalassemia examined so far are the result of large deletions within the β-globin gene cluster. Interestingly, in one form of Dutch γβ-thalassemia, the deletion removes only the ε, γ and δ-globin genes and leaves the β-globin gene together with several kilobasepairs of its 5' and 3' flanking DNA intact (van der Ploeg et al., 1980). In order to determine the cause of the inactivity of this apparently intact β-globin gene, Kioussis et al. (1983) cloned the β-like globin gene cluster from the normal and mutant chromosomes of the Dutch thalassemic patient. Analysis of the DNA sequence of the cloned thalassemic β-globin gene and of its expression in a HeLa cell transient expression assay (section 1.5ii) showed that its phenotype

was not caused by a mutation which created a stop codon or one which affected RNA splicing, transcription initiation or transcription termination. In order to determine whether the thalassemic phenotype was due to a mutation in the DNA sequences required for transcriptional activation as opposed to those required for transcription per se, the expression of the cloned γβ-thalassemic β-globin gene was analysed here during differentiation in MEL cell transformants.

5.2 Structure of the γβ-thalassemic globin gene locus and analysis of of its expression in MEL cells

The human β-like globin gene cluster of the normal and mutant chromosomes of the γβ-thalassemic patient were cloned in the HSV tk vector pOPF by Kioussis et al. (1983). The structure of the mutant β-globin locus is shown in Figure 21. The endpoint of the deletion, which removes the ε, γ and δ-globin genes, lies approximately 2 kb 5' to the β-globin gene; the γβ5 cosmid clone indicated contained appromixately 11 kb DNA 5' to the β-gene (of which 9 kb was translocated DNA) and 25 kb DNA to the 3' side.

The cosmid γβ5 was introduced into tk^- MEL cells by calcium phosphate transformation. 10 MEL transformants were isolated and expression of the foreign thalassemic β-globin gene analysed as described in section 3.5. 6 of these clones showed a >10 fold induction of human β-globin mRNA upon differentiation, as measured by SI nuclease analysis using a probe which detected the 3' end of human β-globin mRNA (Figure 22). The inducibility of the β-globin gene of

the γβ5 clone indicated that its thalassemic phenotype was not caused by a mutation in the DNA sequences which regulate β-globin gene activation during the later stages of erythroid differentiation.

5.3 Discussion

The thalassemic phenotype of the β-globin gene in a Dutch case of γβ thalassemia has previously been shown not to be the result of mutations generating stop codons, RNA splicing defects, or defects in transcription initiation or termination. The inducibility of the gene in MEL transformants further indicates that a mutation in the DNA sequences which regulate β-globin gene activation is not involved. This suggests that there is no defect in the β-globin gene *per se* and indicates that in the thalassemia, the translocated DNA which becomes juxtaposed to the β-globin gene is exerting a "negative effect" on the expression of the latter. The thalassemic β-globin gene was introduced into MEL cells as a cosmid clone containing 9kb of translocated DNA. This indicates that, unless sequences >9kb away from the gene were involved, the DNA sequence *per se* of the translocated DNA was not sufficient to mediate the suppression of the adjacent β-globin gene. Kioussis *et al*. (1983) have indeed shown that the chromatin constitution rather than the DNA sequence *per se* of the translocated DNA plays a role in the suppression of the adjacent β-globin gene. It was found that the transposed DNA is normally found in a region of the chromosome which is in a DNaseI insensitive and hypermethylated state, indicating that it is transcriptionally inactive (section 1.4vi,vii). The β-globin gene on

the translocated chromosome acquires this inactive chromatin configuration, whereas the allelic β-globin gene on the normal chromosome remains in a DNaseI sensitive and hypomethylated state and is transcriptionally active. The inactivity of the β-gene in the Dutch γβ-thalassemia is therefore probably due to the transposition of inactive chromatin. It would be of obvious interest to determine the chromatin structure of the introduced γβ5 clone in MEL transformants. The inducibility of the thalassemic β-globin gene in MEL cells suggests that the inactive chromatin constitution of the translocated DNA as found in vivo is not restored when it is introduced into MEL cells as naked DNA; this is in agreement with most evidence which suggests that DNA introduced into eukaryotic cells acquires an active chromatin structure even when the endogenous counterpart gene is inactive (discussed further in Chapter 9).

Figure 21

Structure of cosmid clone γβ5 from a γβ thalassemic patient

Cosmid γβ5 was isolated by D. Kioussis (Kioussis et al., 1983) from a cosmid library constructed from blood DNA of a γβ thalassemic patient in the vector pOPF (Grosveld et al., 1982). The cosmid contained the β-globin gene from the mutant chromosome together with 11 kb DNA 5' to the β-gene and 25 kb DNA to the 3' side. The translocation breakpoint on the mutant chromosome resided about 2 kb 5' to the transcription initiation site of the β-globin gene. Positions of EcoRI restriction sites are indicated.

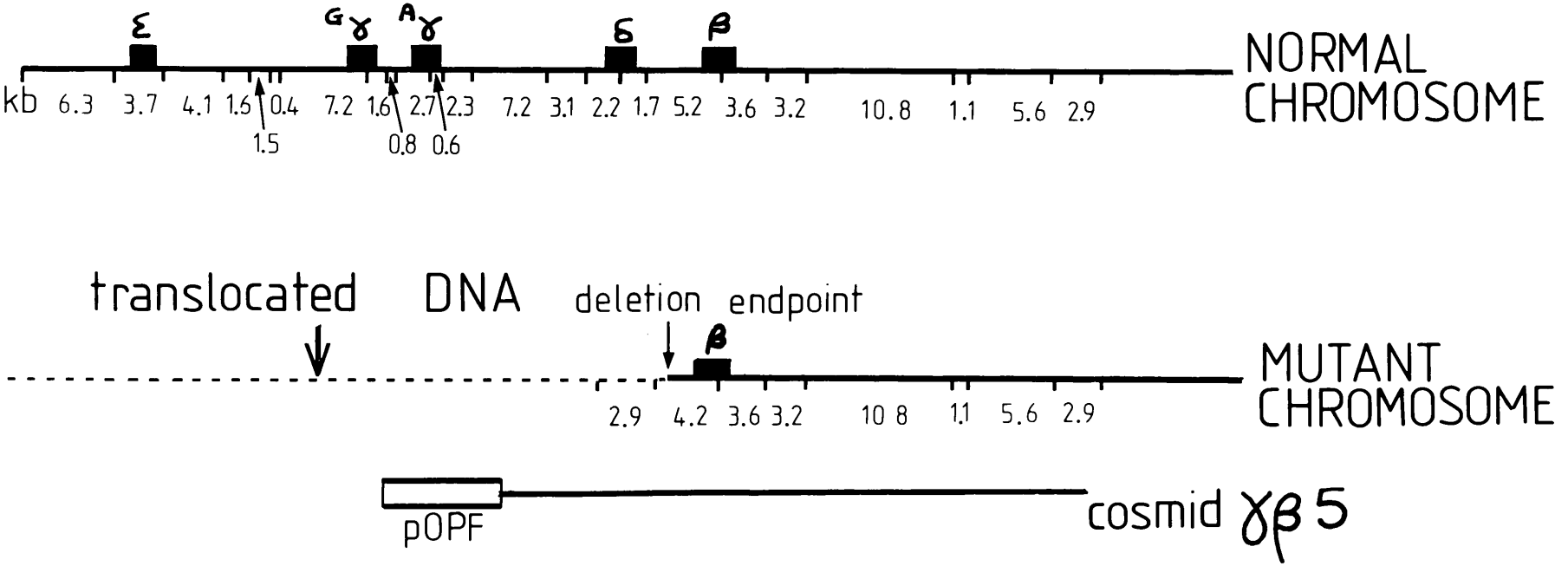
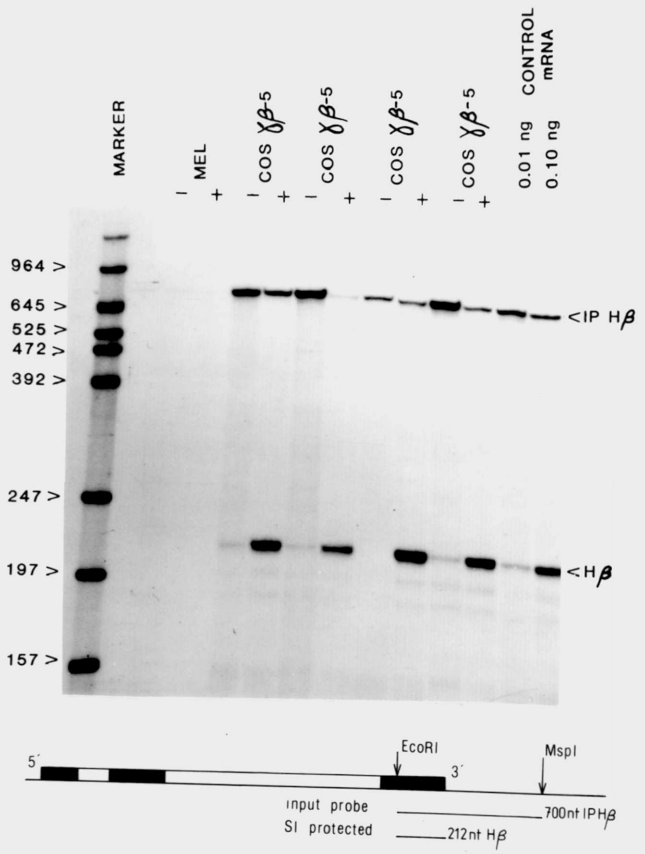

ε
Gγ
Aγ
δ
β
kb 6.3 3.7 4.1 1.6 0.4 7.2 1.6 2.7 2.3 7.2 3.1 2.2 1.7 5.2 3.6 3.2 10.8 1.1 5.6 2.9
1.5
0.8 0.6
NORMAL CHROMOSOME
translocated DNA
deletion endpoint
β
2.9 4.2 3.6 3.2 10 8 1.1 5.6 2.9
MUTANT CHROMOSOME
pOPF
cosmid γβ5

Figure 22

SI nuclease analysis of human β-globin transcripts in MEL transformants containing cosmid γβ5

MEL transformants containing cosmid γβ5 were grown for 3 days in the presence (+) or absence (-) of 3mM HMBA. RNA was prepared and 10μg samples analysed by SI nuclease mapping using the 3' human β-globin gene probe illustrated and the procedure as described in Figure 10. Each hybridization contained 10μg total RNA.


MARKER
MEL
COS γβ-5
COS γβ-5
COS γβ-5
COS γβ-5
0.01 ng CONTROL mRNA
0.10 ng mRNA
964 >
645 >
525 >
472 >
392 >
247 >
197 >
157 >
<IP Hβ
<Hβ
5'
EcoRI
3'
MspI
Input probe
700nt IPHβ
SI protected
212nt Hβ

CHAPTER 6

THE INDUCTION IN MEL CELLS OF A FOREIGN RABBIT β-GLOBIN GENE REQUIRES ONLY 58BP DNA 5' TO THE TRANSCRIPTION INITIATION SITE

6.1 Introduction

By introducing human globin gene cosmids into MEL cells, it was shown in Chapter 3 that expression of a foreign human β-gene was inducible during chemically triggered erythroid differentiation and that the induction did not require the presence of more than 1.5 kb DNA 5' to the transcription initiation site of the globin gene. Since the regulatory DNA sequences of many genes have been shown to reside within their 5' flanking DNA (see section 1.5iv), a series of rabbit β-globin gene mutants with deletions within this 5' flanking region was introduced into MEL cells in order to make a preliminary localization of the DNA sequences which are required for globin gene induction.

6.2 Transcription of an introduced rabbit β-globin gene in MEL cells

In order to determine whether it would be feasible to study the expression and induction of a series of rabbit β-globin deletion mutants (G.Grosveld et al., 1981b, 1982a,b) in MEL cells, the inducibility of an intact foreign rabbit β-globin gene was firstly ascertained. The rabbit β-globin gene clone RβpBSV, which contained

425 bp DNA 5' to the transcription initiation site (Figure 23) was introduced into MEL cells by cotransfection with the human β-globin cosmid β-pRT. The latter clone contained the human β-globin gene 4.7 kb BglII fragment cloned into the HSV-tk vector pRT (Figure 5), and thus served as an "induction" control against which expression of the rabbit β-globin gene and its deletion mutants could be measured.

5 MEL clones transfected with both the rabbit and human β-globin genes were induced to differentiate and levels of human β-globin mRNA measured by SI nuclease mapping using a 700 nt EcoRI-MspI probe which detected the 3' end of the transcript. 4 out of the 5 clones showed a 5-80 fold induction of human β-globin mRNA upon differentiation. Rabbit β-globin mRNA levels in these 4 transformants were measured by SI nuclease mapping using a probe which detected the 5' end of the transcript. All 4 clones showed a 10->40 fold induction of rabbit β-globin mRNA. Figure 25 shows the induction of rabbit β-globin mRNA in 2 of these clones; hybridization of the 221nt BstNI probe to correctly initiated rabbit β-globin transcripts yielded a 139nt fragment which was protected from SI nuclease digestion. The observed induction of the rabbit β-globin gene in these transformants suggested that it would be feasible to study the expression of the globin deletion mutant genes in this system.

6.3 Effect of SV40 DNA sequences on MEL cell transformation efficiency and expression of introduced genes

The rabbit β-globin gene clones had been constructed in vectors

containing SV40 DNA (G.Grosveld et al., 1982a,b). Since SV40 sequences have been reported to affect both the efficiency of transformation of some eukaryotic cell lines and the level of expression of foreign genes in eukaryotic cells (Luciw et al., 1983), it was necessary to determine the effect of the SV40 sequences in the rabbit β-globin gene constructs on both of these parameters. The SV40 sequences of the rabbit β-globin gene clone (Figure 23) were therefore removed by cleavage of the DNA with HindIII, and the transformation experiment described in section 6.2 repeated. The presence of SV40 DNA sequences in the rabbit β-globin gene clone was shown to have no effect on either the efficiency of transformation of the MEL cells, or on the inducibility and expression level of the introduced human and rabbit β-globin genes.

The expression of the intact SV40 DNA sequences in the MEL transformants described in section 6.2 was also analysed. RNA from 3 of the these MEL transformants was analysed by SI nuclease mapping using a probe which detected the 5' end of SV40 early gene transcripts. Hybridization of the 645nt AvaII probe (Figure 29) to correctly initiated SV40 early transcripts yielded fragments of 115-140 nt which were protected from SI nuclease digestion. The control mRNA used in this SI nuclease analysis was obtained from CV1 cells which were lytically infected with SV40, and the appearance of many SI nuclease protected fragments indicates the characteristic multiple initiation sites of SV40 early transcription. All 3 MEL transformants showed inducibility of SV40 early transcripts (Figure 29); interestingly, only one or two of the multiple transcription initiation sites was observed in the SV40 transformed MEL cells.

6.4 Structure of rabbit β-globin gene mutants

Rabbit β-globin gene deletion mutants were a gift of G. Grosveld; their structure is shown in Figure 24. Figure 24 also shows the relative transcription level of each mutant in the HeLa cell transient expression assay (section 1.5ii); this gives a measure of the effect of each deletion on the efficiency of transcription *per se* of the rabbit β-globin gene. The -100 mutant contained all of the promoter sequences required for transcription *per se*, but had no extra 5' flanking DNA; analysis of the inducibility of the -425 and -100 mutants would determine whether the regulatory sequences resided within or outside the promoter boundary. The other mutants had deletions within the conserved sequence boxes required for transcription *per se*; the D76-93 mutant lacked the conserved -90 region, and the -58 mutant, which had only 58 bp DNA 5' to the transcription initiation site, lacked both the conserved -90 region and CAAT box. Deletion mutations extending into the ATA box resulted in too great a reduction in transcription efficiency for the evaluation of inducibility in MEL cells to be possible (sections 6.5 and 6.6).

6.5 Transcription of rabbit β-globin gene mutants in stable mouse L cell transformants

The transcription efficiency of the rabbit β-globin gene mutants was originally determined using the HeLa cell transient expression system

(G.Grosveld et al., 1982a,b). In this assay, transcripts from a foreign gene represent a 10-100 fold greater fraction of total cellular RNA than when the foreign gene is stably introduced into a eukaryotic cell (section 1.5ii). The expression of the rabbit β-globin gene mutants was therefore firstly studied in stable mouse L cell transformants in order to assess the feasibility of detecting their induction in stable MEL cell clones i.e. to ensure that any lack of transcription which might be observed in MEL transformants was not merely the result of the transcription level of mutant gene being below that of detection. It was also of interest to determine whether a given mutation had the same effect on transcription efficiency per se in both the HeLa cell transient expression system and in stable eukaryotic cell clones.

The rabbit β-globin gene mutants (Figure 24) had been constructed in a vector which lacked a selectable gene for eukaryotic cell transformation and were introduced into tk^- L cells by cotransfection with the vector pRT (F.Grosveld et al., 1982, see Figure 5) which contained the HSV tk gene. A mutant containing only 35 bp DNA 5' to the transcription initiation site of the rabbit β-globin gene was also introduced into L cells in the same manner. Levels of rabbit β-globin mRNA in individual L cell transformants were measured by SI nuclease analysis using a probe which detected the 3' end of the transcript. Figure 28 shows SI nuclease analysis of rabbit β-globin transcripts in transformants containing each of the mutants. L cell transformants containing deletion mutants D76-93 and -58 showed a 10-50 fold lower level of rabbit β-globin transcripts than did transformants containing the intact gene (-425

and -100 mutants). The level of transcription of any given mutant varied by 5-20 fold between individual transformants, and did not correlate with the number of copies of the rabbit β-globin gene in the cell. No rabbit β-globin transcripts were detectable in transformants containing deletion mutants which lacked the ATA box (-35 mutant). For any given mutant, the fraction of total cellular RNA represented by rabbit β-globin transcripts was 10-30 fold lower in the stable L cell transformants than when measured in the HeLa cell transient expression system. This implied that it would not be feasible to test deletion mutants extending into the ATA box for their inducibility in MEL cells since the absence of transcripts from such mutants in differentiated MEL cells would most likely result from the effect of the deletion mutation on transcription of the gene per se rather than being due to a possible deletion of regulatory sequences required for gene activation. Rabbit β-globin transcripts in L cell transformants were quantitated as in section 3.4 using rabbit bone marrow polyA$^+$ RNA as a control. Levels of transcripts in L cell transformants containing the intact rabbit β-globin gene corresponded to 800-5000 copies mRNA per cell.

6.6 Transcription of rabbit β-globin gene mutants in MEL cells

The rabbit β-globin gene mutants (Figure 24) were introduced into MEL cells by cotransfection with the human globin cosmid β-PRT, and 10-15 MEL clones transfected with each mutant were induced to differentiate by culturing for 3 days in the presence of 3mM HMBA. Levels of mouse, human, and rabbit β- globin mRNA were measured by SI

nuclease analysis using probes which detected the 5' or 3' ends of the specific transcripts. Only transformants showing inducibility of both endogenous mouse β-major globin mRNA and foreign human β-globin mRNA were analysed for expression of the rabbit β-globin gene mutant.

SI nuclease mapping of rabbit β-globin gene transcripts in MEL transformants containing each of the mutants is shown in Figures 25 and 26. Figures 27 and 28 show the induction of endogenous mouse β-major globin mRNA in the same MEL transformants and the level of transription of each mutant in stable L cell transformants.

6 out of the 8 transformants which contained the rabbit β-globin gene -100 mutant and which showed inducibility of human β-globin mRNA also showed induction of rabbit β-globin mRNA upon differentiation, thus indicating that DNA sequences 5' to the promoter boundary are not essential for activation of the rabbit β-globin gene. 4 out of 7 transformants containing the D76-93 deletion mutant and 3 out of 6 transformants containing the -58 mutant showed a 5-50 fold induction of correctly initiated rabbit β-globin transcripts. The absolute levels of rabbit β-globin mRNA varied extensively between individual transformants containing the same mutant. The level of transcription of the D76-93 and -58 mutant genes was, however, generally lower than that of the -425 and -100 mutants, presumably due to the effect of these promoter mutations on the efficiency of transcription per se. Thus, removal of the -90 and CAAT box regions of the promoter sequences resulted in a reduction in efficiency of transcription but did not affect activation of the rabbit β-globin gene. It should be noted that the effect of the D76-93 and -58 deletion mutations on expression of the rabbit

β-globin gene was not as great in the differentiated MEL cells as it was as measured in the HeLa cell transient expression assay (G.Grosveld et al., 1982a,b). Rabbit β-globin mRNA levels in MEL transformants were quantitated as described in section 6.5; levels in induced cells containing the intact rabbit β-globin gene (-425 mutant) corresponded to 500-3000 copies mRNA per cell. These data therefore indicate that not more than 58 bp DNA 5' to the transcription initiation site is required for induction of a transfected rabbit β-globin gene in MEL cells.

6.7 Discussion

It was shown that the expression in MEL cells of an introduced rabbit β-globin gene was inducible upon chemically triggered differentiation. Thus, as was observed for the human β-globin gene, the expression of the rabbit β-globin gene was regulated appropriately in the heterologous mouse erythroid cell environment. Induction of the rabbit β-globin gene did not require the presence of >58 bp DNA 5' to transcription initiation site; this implies that the DNA sequences which regulate β-globin gene activation during erythroid differentiation reside either within the 3' region of the promoter or within the structural gene and/or its 3' flanking region. The presence of regulatory DNA sequences within the close vicinity of an inducible gene has been previously described in other systems. For example, the hormonal induction of a cloned metallothionen gene in Rat2 cells is mediated by elements residing 237-268bp upstream from the cap site (Karin et al., 1984a; section 1.5iv). This and most

other cases of gene induction differ from that of globin gene activation in that their regulatory element has been located 5' to the DNA sequences which are required for transcription per se; the DNA sequences which regulate globin gene induction were shown in this chapter to reside downstream from the CAAT box (discussed in Chapter 9). Because both the level of induction and transcription of the rabbit β-globin gene mutants was very variable between individual MEL transformants, it was not possible to compare either of these parameters between the different mutants, and hence the presence of regulatory DNA sequences upstream from the -58 position which have a quantitative effect on the induction cannot be excluded. Since deletion of further sequences of the promoter greatly reduced the efficiency of transcription (section 6.5), the DNA sequences which regulate β-globin gene activation were further localized by the construction of hybrid genes between inducible and non-inducible genes (Chapters 7 and 8).

Figure 23

Structure of rabbit β-globin gene clone RβpBSV

RβpBSV was a gift of G. Grosveld. It contained the 2.2 kb rabbit β-globin gene BglII partial fragment inserted into the BamHI site of vector pBSV (G.Grosveld et al., 1982a).

B = BamHI restriction site
H = HindIII " "
E = EcoRI " "
Bg= BglII " "

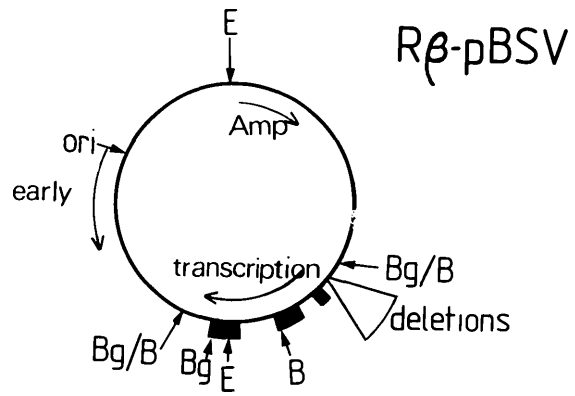
Bg/B = site of cloning of BglII fragment into BamHI site (both sites destroyed).

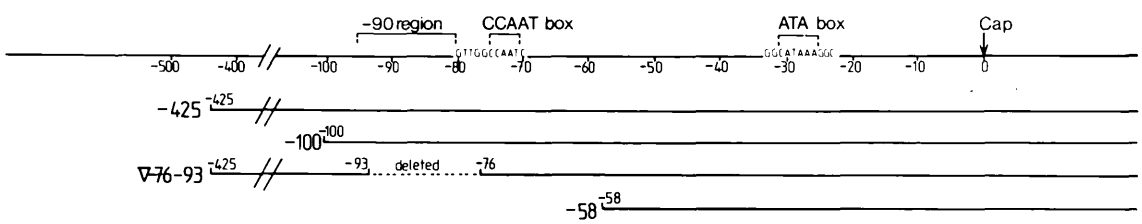
Figure 24

Structure of rabbit β-globin gene deletion mutants

The construction of the rabbit β-globin gene mutants is described in G.Grosveld et al., 1982a,b. The -425 construct corresponds to the RβpBSV clone above; the mutants had deletions within these 425 bp of 5' flanking DNA as illustrated. The relative transcription level of the mutants as measured in the HeLa cell transient expression system is indicated.

Mutants were introduced into tk^- L cells by cotransfection with the cosmid tk-vector pRT; they were introduced into MEL cells by cotransfection with the human β-globin gene cosmid β-pRT (Figure 5).





	Transcription Level (HeLa cells)
-425	1.0
-100	1.0
∇-76-93	0.02
-58	0.12

Figure 25

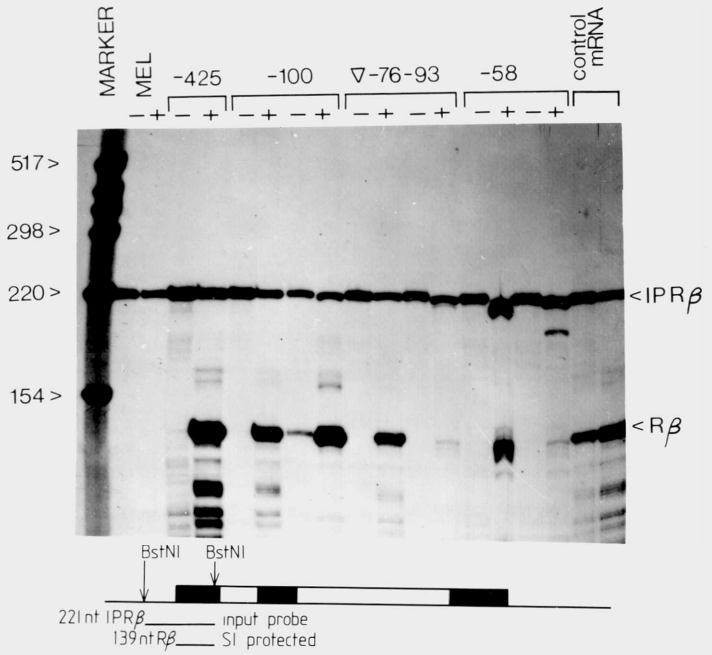
SI nuclease analysis of the 5' end of rabbit β-globin transcripts in MEL transformants

Individual MEL transformants containing the rabbit β-globin gene mutants indicated and corresponding to those shown in Figure 26 were grown for 3 days in the presence (+) or absence (-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 100μg samples were analysed by SI nuclease mapping using the rabbit β-globin gene 5' probe illustrated. The probe was end-labelled using polynucleotide kinase to a specific activity of $>10^7$ cpm/μgDNA and strand separated.
Hybridization was at 37°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control polyA$^+$ RNA was obtained from rabbit bone marrow.

Figure 26

SI nuclease analysis of the 3' end of rabbit β-globin transcripts in MEL transformants

The 3' end of rabbit β-globin transcripts in MEL transformants corresponding to those shown in Figure 25 was analysed using the probe illustrated and the method described in Figure 28.
Each sample contained 10μg total RNA.


MARKER
MEL
-425
-100
∇-76-93
-58
control mRNA
517 >
298 >
220 >
154 >
<IPRβ
<Rβ
BstNI
BstNI
221 nt IPRβ
139 nt Rβ
input probe
SI protected

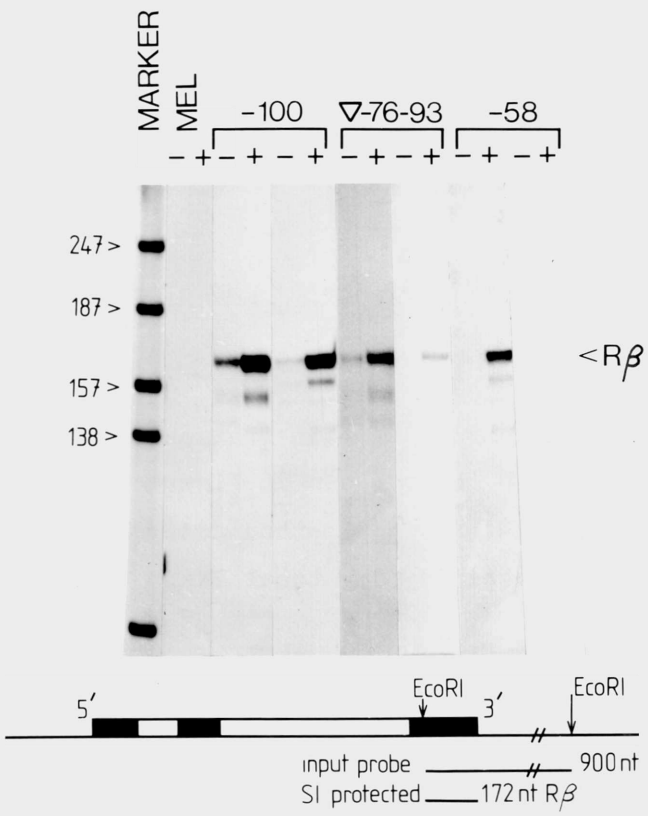

MARKER
MEL
-100
∇-76-93
-58
247 >
187 >
157 >
138 >
<Rβ
5′
EcoRI
3′
EcoRI
input probe
900 nt
SI protected
172 nt Rβ

Figure 27

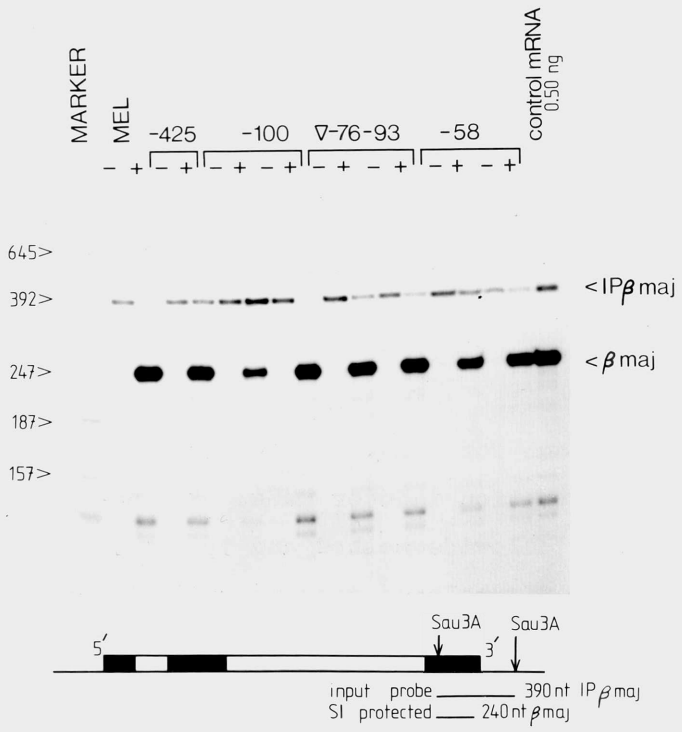
SI nuclease analysis of mouse β-major globin transcripts in MEL transformants containing rabbit β-globin gene mutants

Individual MEL transformants containing the rabbit β-globin gene mutants indicated (corresponding to those shown in Figures 25 and 26) were grown for 3 days in the presence (+) or absence (-) of the diferentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 1μg samples of total RNA were analysed by SI nuclease mapping using the mouse β-major globin gene 3' probe illustrated. The probe was end labelled by reverse transcription to a specific activity of $>10^7$ cpm/μgDNA. Hybridization was at 52°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control polyA$^+$ RNA was obtained from mouse reticulocytes.

Figure 28

SI nuclease analysis of rabbit β-globin transcripts in L cell transformants

L cell transformants contained the rabbit β-globin gene mutants indicated. 10μg samples of total RNA from individual transformants were analysed by SI nuclease mapping using the rabbit β-globin gene 3' probe illustrated. The probe was end labelled by reverse transcription to a specific activity of $>10^7$ cpm/μgDNA. Hybridization was at 52°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed ot a 7M urea 7% acrylamide gel. Control poly A$^+$ RNA was obtained from rabbit bone marrow.


MARKER
MEL
-425
-100
∇-76-93
-58
control mRNA
0.50 ng
645>
392>
247>
187>
157>
<IPβ maj
<β maj
Sau3A
Sau3A
5′
3′
input probe 390 nt IPβ maj
SI protected 240 nt β maj

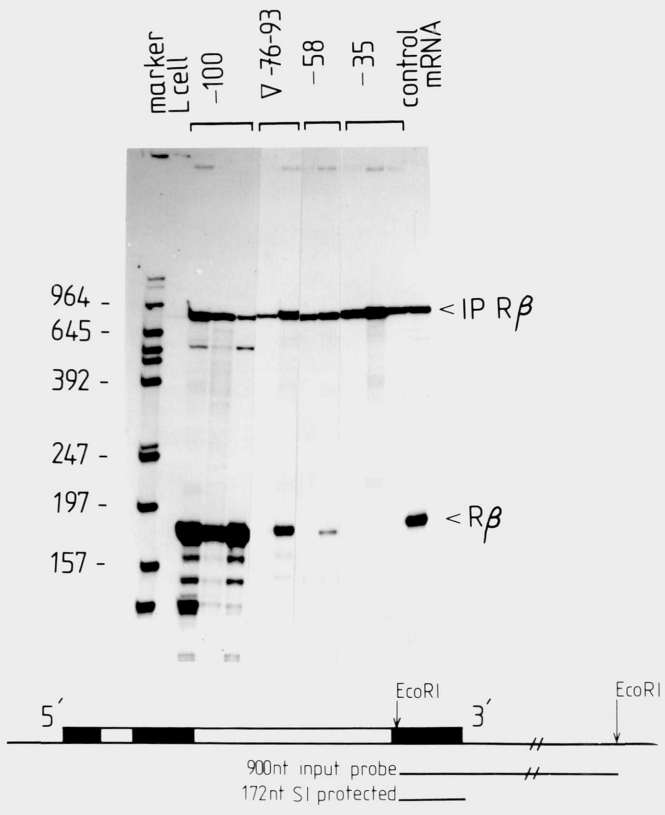
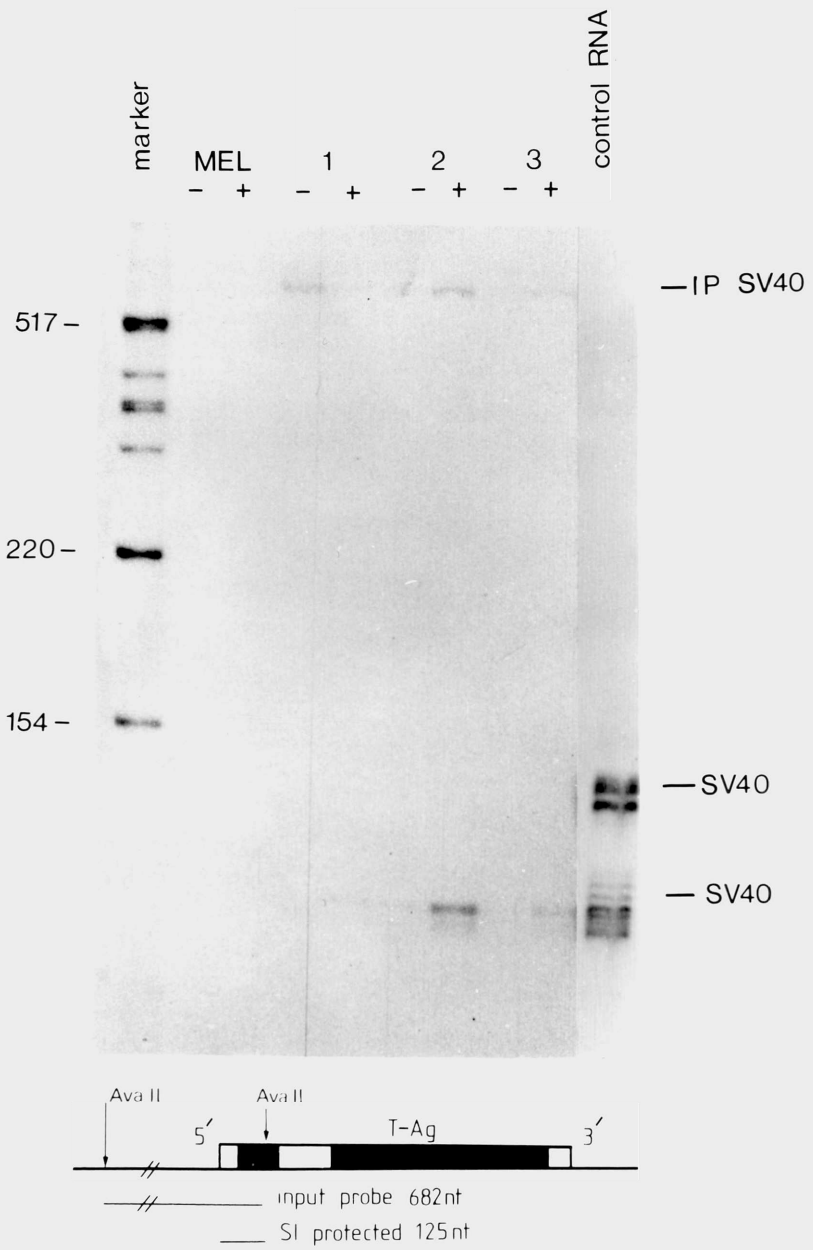

marker
L cell
-100
∇ -76-93
-58
-35
control mRNA
964 -
645 -
392 -
247 -
197 -
157 -
< IP Rβ
< Rβ
EcoRI
EcoRI
5′
3′
900nt input probe
172nt SI protected

Figure 29

Expression of SV40 early region transcripts in MEL transformants containing RβpBSV

3 MEL transformants containing RβpBSV were grown for 3 days in the presence (+) or absence (-) of 3mM HMBA. RNA was prepared and 10 μg samples analysed by SI nuclease mapping using the 645 nt AvaII probe which detects the 5' end of SV40 early gene transcripts. The probe was end-labelled using polynucleotide kinase. Hybridization was at 37°C; SI nuclease digestion was for 2 hours at 20°C. Control RNA (a gift of D. Kalderon) was from CVI cells which had been lytically infected with SV40.


marker
MEL
1
2
3
control RNA
– + – + – + – +
517 –
220 –
154 –
IP SV40
SV40
SV40
Ava II
Ava II
5′
T-Ag
3′
input probe 682nt
SI protected 125nt

CHAPTER 7

INDUCTION OF FOREIGN HYBRID GENES CONTAINING THE HUMAN β-GLOBIN 5' FLANKING REGION IN MEL CELLS

7.1 Introduction

Data presented in Chapter 6 showed that not more than 58 bp DNA 5' to the transcription initiation site were required for induction of a transfected rabbit β-globin gene in MEL cells. In order to determine whether the β-globin gene regulatory sequences resided within its 5' flanking DNA, hybrid genes were constructed between the inducible human β-globin gene and either the noninducible human γ-globin gene or the noninducible mouse major histocompatibility H-2K^{bm1} gene. The hybrids described in this chapter contained the 5' flanking region of the human β-globin gene joined at the translation initiation site to the complementary region of the γ-globin or H-2K^{bm1} genes; the converse constructs are described in Chapter 8.

7.2 Transcription of an introduced mouse major histocompatibility H-2K^{bm1} gene is not inducible in MEL transformants

In order to determine whether a mouse major histocompatibility (MHC) gene would be suitable for use as the "noninducing" moiety in the construction of hybrid genes, the plasmid pTME10, which contains the 10kb H-2K^{bm1} EcoRI fragment cloned in the cosmid vector pTM (Rosenthal et al., 1984, accompanying manuscript; Figure 30) was

introduced into MEL cells by cotransfection with the HSV tk vector pRT. Levels of H-2K^{bm1} transcript were measured in 12 MEL transformants by SI nuclease analysis using a 600 nt probe which was isolated from the H-2K^{b} gene and which covered the 3rd exon-3rd intron boundary. Because the nucleotide sequences of the H-2K^{bm1} and H2K^{b} genes differ in the 3rd exon (Weiss et al., 1983), hybridization of the H-2K^{b} probe to H-2K^{bm1} transcripts yielded a 140 nt fragment that was protected from digestion by SI nuclease (Figure 31). None of the 12 MEL transformants analysed showed inducibility of H-2K^{bm1} transcripts (Figure 31); the H-2K^{bm1} gene was therefore suitable for use as the noninducing moiety in the construction of hybrid globin genes.

7.3 Construction of 5'β-3'γ and 5'β-3'H-2K^{bm1} hybrid genes

In order to determine whether human β-globin gene sequences located 5' to the translation initiation site can confer inducibility upon a noninducible gene, they were adjoined to the complementary 3' region of ~~the~~ both the noninducible human γ-globin and H-2K^{bm1} genes.

The 5'β-3'γ hybrid gene was constructed at the NcoI site of both the human β- and γ-globin genes; this site lies exactly at the translation initiation site, which is approximately 40 bp downstream from the transcription initiation site. The hybrid gene contained approximately 1.5 kb of the 5' flanking region of the β-globin gene and 1kb DNA 3' to the γ-globin gene (Figure 32).

The 5'β-3'H-2K^{bm1} construct was made at the Nco1 site of the β-globin gene and at the NruI site of the H-2K^{bm1} gene; the

latter site lies between the transcription initiation site and translation initiation site. The hybrid gene contained approximately 1.5 kb of the 5' flanking region of the human β-globin gene and 1 kb DNA 3' to the H-2K^{bm1} gene (Figure 33).

7.4 Transcription of 5'β-3'γ and 5'β-3'H-2K^{bm1} hybrid genes in MEL transformants

The 5'β-3'γ hybrid gene clone contained the HSV tk gene and was introduced into MEL cells by calcium phosphate transformation and selection in HAT medium. 10 transformants were induced to differentiate, and transcripts from the introduced hybrid gene were quantitated by SI nuclease analysis using a probe which detected the 3' end of human γ-globin mRNA. Five transformants showed a 5-50 fold induction of the 5'β-3'γ hybrid transcript upon differentiation (Figure 34).

The 5'β-3'H-2K^{bm1} hybrid gene was similarly introduced into MEL cells by cotransfection with the HSV tk vector pRT. 2 out of 3 transformants showed a >10 fold inducibility of hybrid gene transcript as measured by SI nuclease analysis using the H-2K^{b} gene exon-intron probe described in section 7.2 (Figure 35).

7.5 Discussion

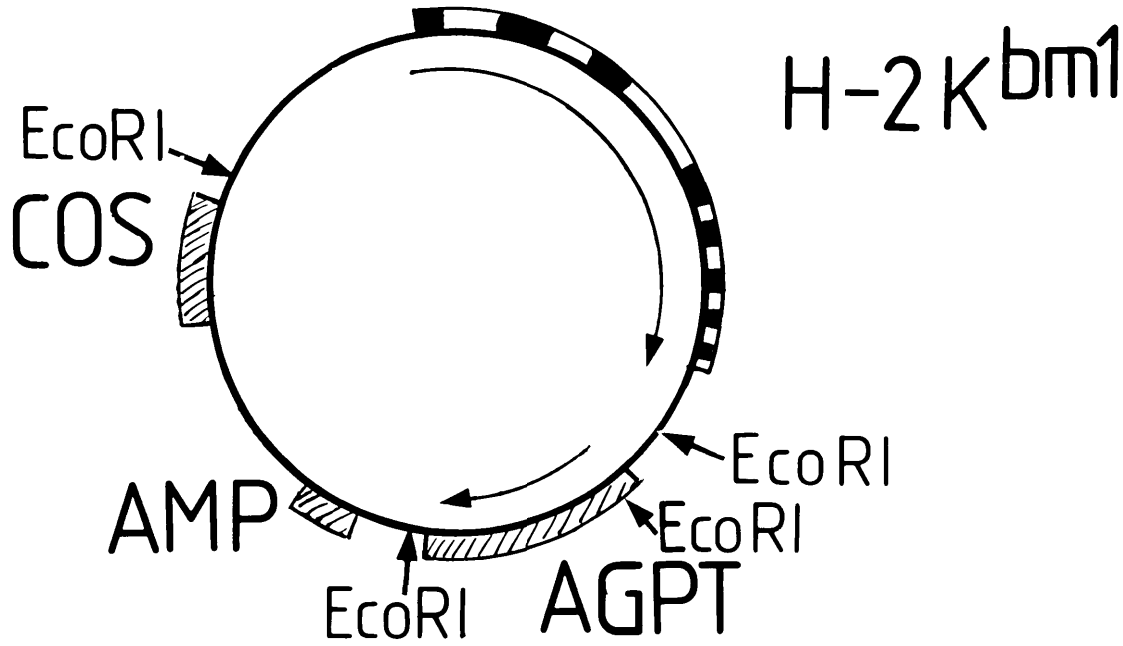
The 5' flanking DNA of the human β-globin gene conferred inducibility in MEL cells upon both the noninducible human γ-globin and mouse H-2K^{bm1} genes (5'β-3'γ and 5'β-3'H-2K^{bm1}). DNA sequences 5'

to the translation initiation site therefore regulate activation of the β-globin gene. Because the level of induction was so variable between transformants, it was not possible to determine whether the induction of the 5'β-3'γ and 5'β-3'H-2K^{bm1} genes was as great as that of the intact β-globin gene (Chapter 3), and hence whether all the regulatory DNA sequences were contained within this 5' region. The data presented in this chapter are similar to those shown for other systems in which the 5' flanking region of an inducible gene will confer regulation upon a normally noninducible gene (discussed in sections 1.5iv, 6.7, 9.3). The presence of β-globin regulatory DNA sequences 3' to the translation initiation site was studied in Chapter 8.

Figure 30

Structure of H2-K^{bm1} gene clone, pTME10

The H2-K^{bm1} clone was a gift of A. Rosenthal (Rosenthal et al., 1982, accompanying manuscript); it contained the 10 kb H2-K^{bm1} gene EcoRI fragment inserted into an EcoRI site of the vector pTM (F. Grosveld et al., 1982)



pTME10

Figure 31

SI nuclease analysis of H2-K^{bm1} transcripts in MEL transformants containing pTME10

Individual MEL transformants containing pTME10 (Figure 30) were grown for 3 days in the presence (+) or absence (-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 10μg samples of total RNA were analysed by SI nuclease analysis using the H2-K^b gene probe illustrated. Since the nucleotide sequence of the H-$2K^b$ and H-2^{bm1} genes differ in the 3rd exon at the position shown, hybridization of the H-$2K^b$ probe to H-$2K^{bm1}$ transcripts yields a 140 nt fragment which is protected from SI nuclease digestion. The probe was end-labelled by reverse transcription to a specific activity of $>10^7$ cpm/μg. Hybridization was at 60°C; SI nuclease digestion was for 1 hour at 20°C followed by 1 hour at 40°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control RNA was obtained from the liver of a H2-K^{bm1} haplotype mouse.

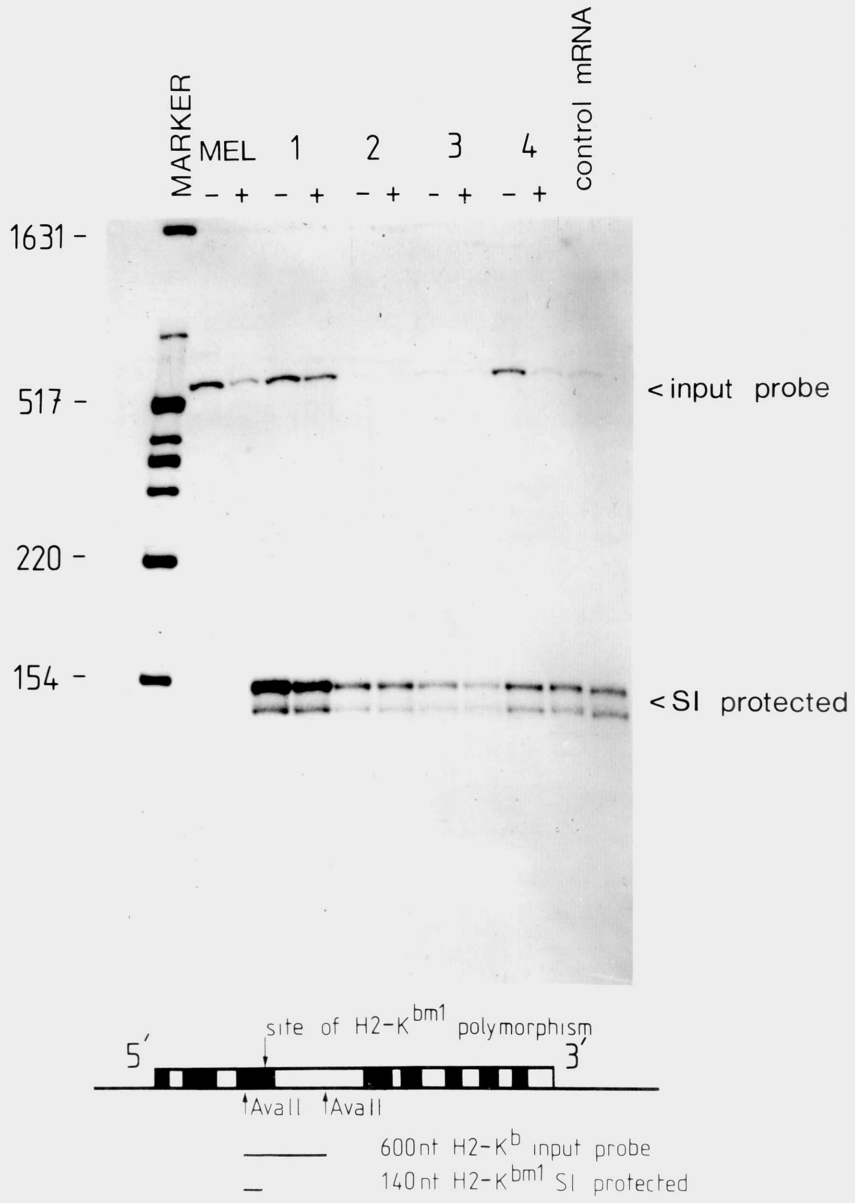

MARKER
MEL
1
2
3
4
control mRNA
- + - + - + - + - +
1631 -
517 -
220 -
154 -
< input probe
< SI protected
site of H2-K bm1 polymorphism
5'
3'
AvaII
AvaII
600nt H2-K b input probe
140nt H2-K bm1 SI protected

Figure 32

Construction of the 5'β-3'γ hybrid human globin gene

The 5'β-3'γ hybrid gene was constructed from globin gene subclones γ-pRT-A and β-pRT-1 (Figure 5).

The 3.8 kb NcoI fragment containing the 3' region of the human γ-globin gene was isolated from γ-pRT and the 6.0 kb NcoI fragment containing the 5' region of the human β-globin gene isolated from β-pRT. The β- and γ-globin gene NcoI fragments were ligated and the ligation mixture transfected into E. coli HB101. Minipreparations of DNA from transformants were screened by digestion with EcoRI for the desired construct, which contained a 5'β-3'γ hybrid gene lacking a 240bp NcoI fragment from the 5' end of the γ-globin gene moieity. This construction intermediate was partially digested with NcoI, and the linear form isolated on a 0.5% low melting agarose gel. The 240 bp NcoI fragment containing the 5' coding region of the human γ-globin gene was isolated from γ-prt and ligated to the NcoI linear fragment. The ligation mixture was transfected into E.coli HB101 and minipreparations of DNA from transformants were screened by digestion with EcoRI and with BamHI+BglI for the desired construct.

N = NcoI restriction site
H = HindIII " "

Bg/B = site of cloning of BglII fragment into BamHI site (both sites destroyed).

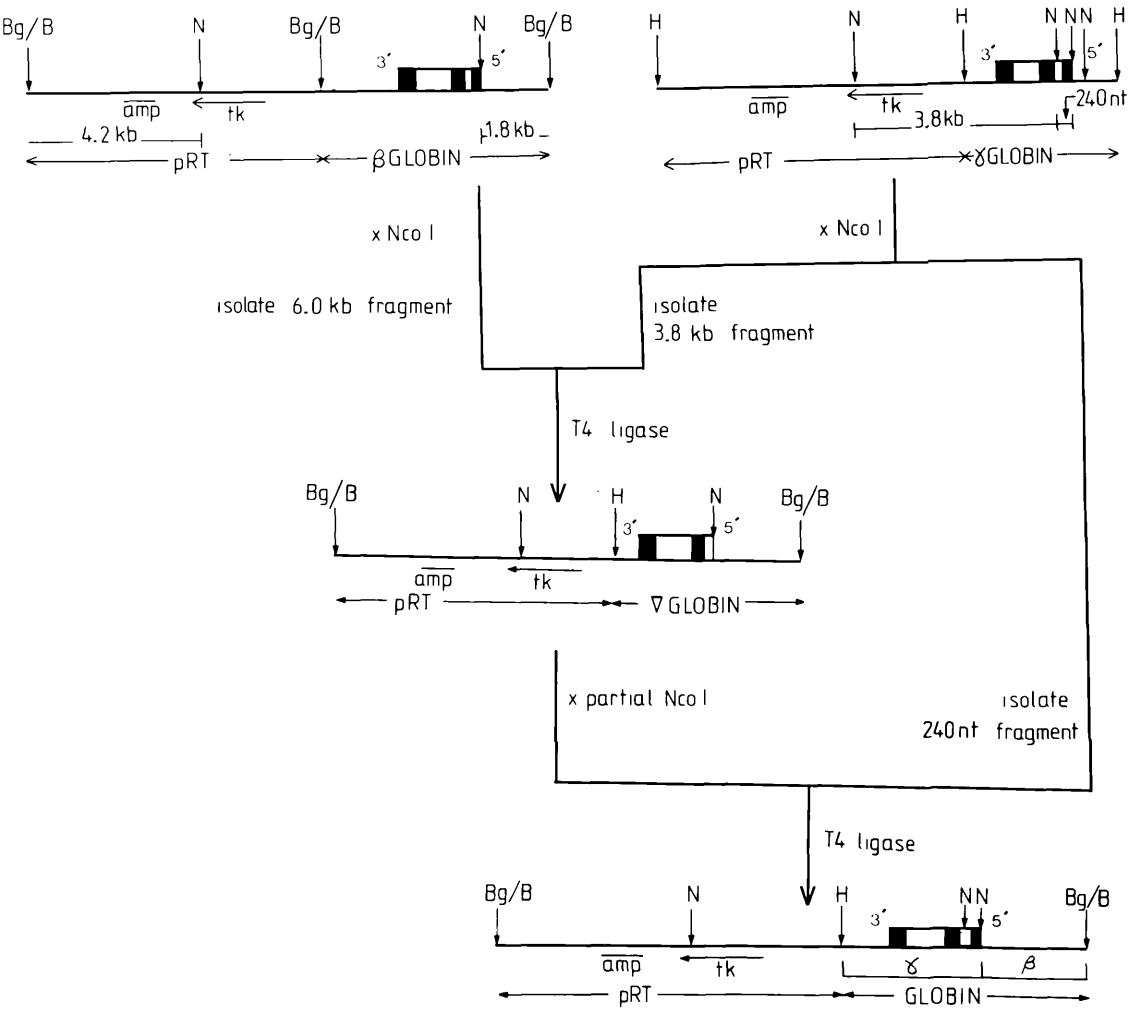

Bg/B
N
Bg/B
N
Bg/B
3'
5'
amp
tk
4.2 kb
1.8 kb
pRT
βGLOBIN
H
N
H
N NN
H
3'
5'
amp
tk
240nt
3.8kb
pRT
γGLOBIN
x Nco I
x Nco I
isolate 6.0 kb fragment
isolate
3.8 kb fragment
T4 ligase
Bg/B
N
H
N
Bg/B
3'
5'
amp
tk
pRT
∇GLOBIN
x partial Nco I
isolate
240nt fragment
T4 ligase
Bg/B
N
H
NN
Bg/B
3'
5'
amp
tk
γ
β
pRT
GLOBIN

Figure 33

Construction of the 5'β-3'H2-K^{bm1} hybrid gene

The 5'β-3'H-$2K^{bm1}$ hybrid gene was constructed from the human β-globin gene subclone β-pRT-1 (Figure 5) and the H-$2K^{bm1}$ gene subclone pTME10 (Figure 30).

The human β-globin gene clone was digested with NcoI and the 3' recessed termini at the NcoI sites were made blunt by treatment with reverse transcriptase in the presence of 1mM each of dATP, dGTP, dCTP, TTP. The sample was then digested with PvuI and the 5.0 kb fragment isolated from a 0.5% low melting agarose gel.

The H2-K^{bm1} clone was digested with NruI (which yields blunt-ended molecules) and then with PvuI. The 8.5 kb NruI-PvuI fragment was isolated from a 0.5% low melting agarose gel.

The 8.5 kb H2-K^{bm1} and 5.0 kb β-globin fragments were ligated, the ligation mixture transfected into E.coli HB101 and minipreparations of DNA from the transformants analysed by digestion with EcoRI for the desired construct.

P, E, N, Nr represent PvuI, EcoRI, NcoI and NruI restriction sites.

Bg/B represents cloning of BglII fragment into BamHI site (both sites destroyed).

Nr/N represents ligation of blunted NcoI site with NruI site (both sites destroyed).

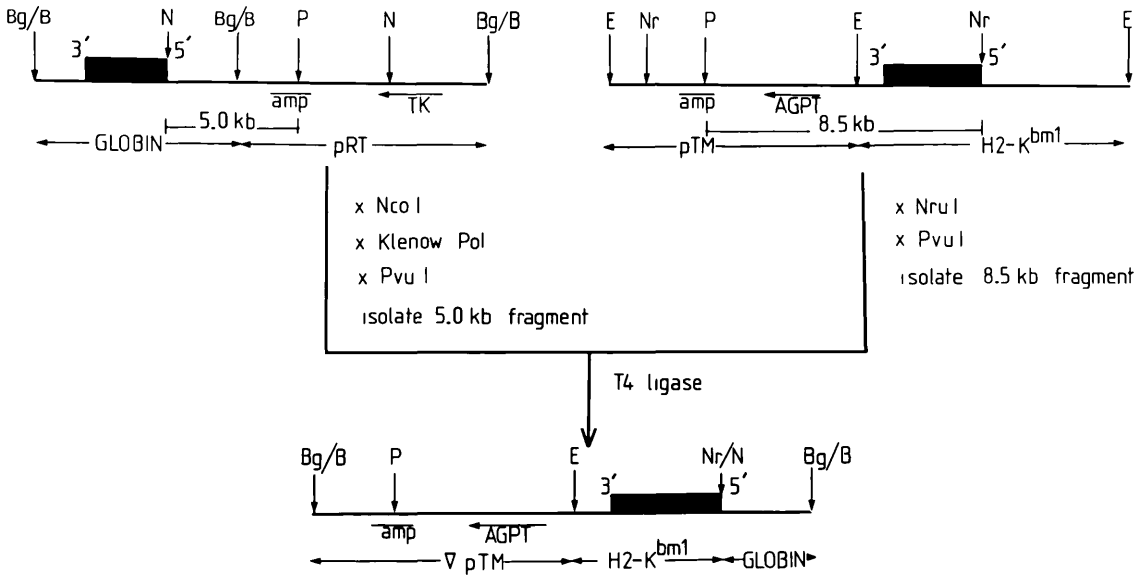

Bg/B
N
Bg/B
P
N
Bg/B
3'
5'
amp
TK
5.0 kb
GLOBIN
pRT
E
Nr
P
E
Nr
E
3'
5'
amp
AGPT
8.5 kb
pTM
H2-K bm1
x Nco I
x Klenow Pol
x Pvu I
isolate 5.0 kb fragment
x Nru I
x Pvu I
isolate 8.5 kb fragment
T4 ligase
Bg/B
P
E
Nr/N
Bg/B
3'
5'
amp
AGPT
pTM
H2-K bm1
GLOBIN

Figure 34

SI nuclease analysis of 5'β-3'γ hybrid gene transcripts in MEL transformants

Individual MEL transformants containing the 5'β-3'γ globin gene were grown for 3 days in the presence (+) or absence (-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 10μg samples of total RNA were analysed by SI nuclease mapping using the human γ-globin gene 3' probe illustrated. The probe was end-labelled by reverse transcription to a specific activity of $>10^7$ cpm/μg. Hybridization was at 52°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control polyA$^+$ RNA was obtained from a patient with $\beta^o/\delta\beta^o$ thalassemia.

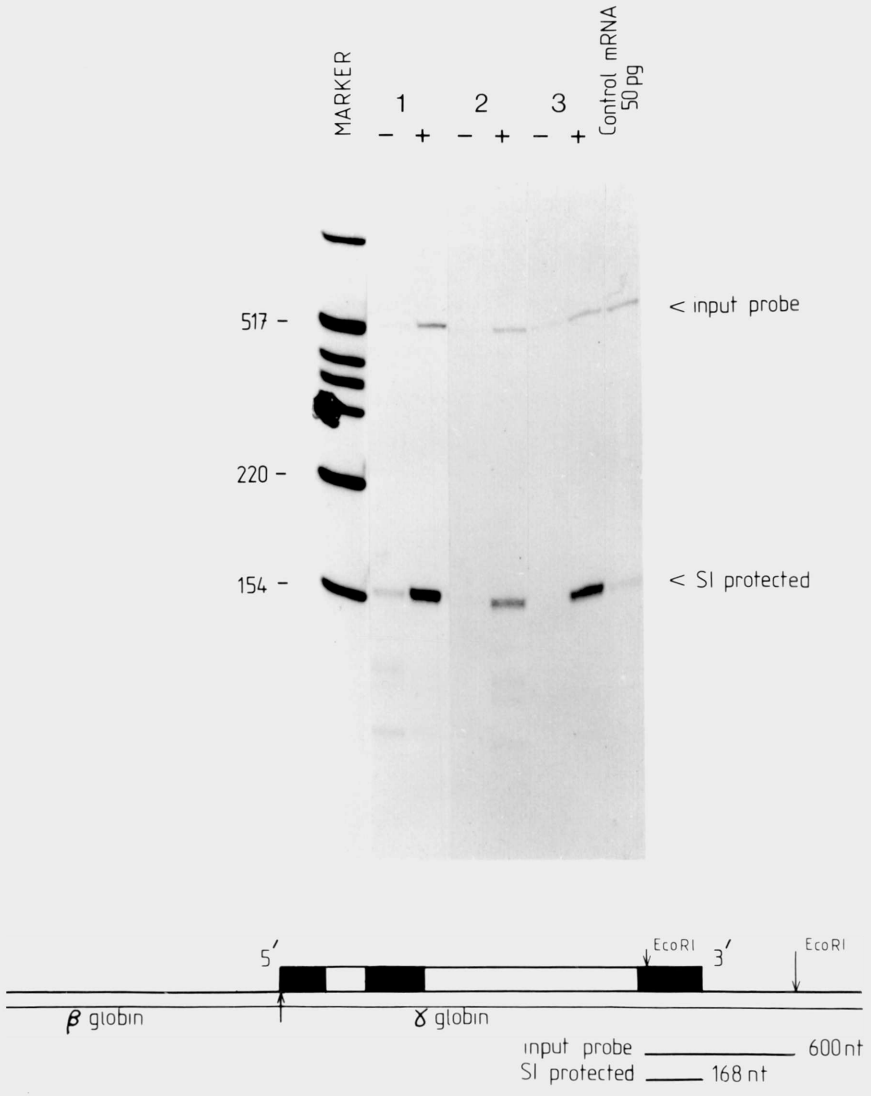
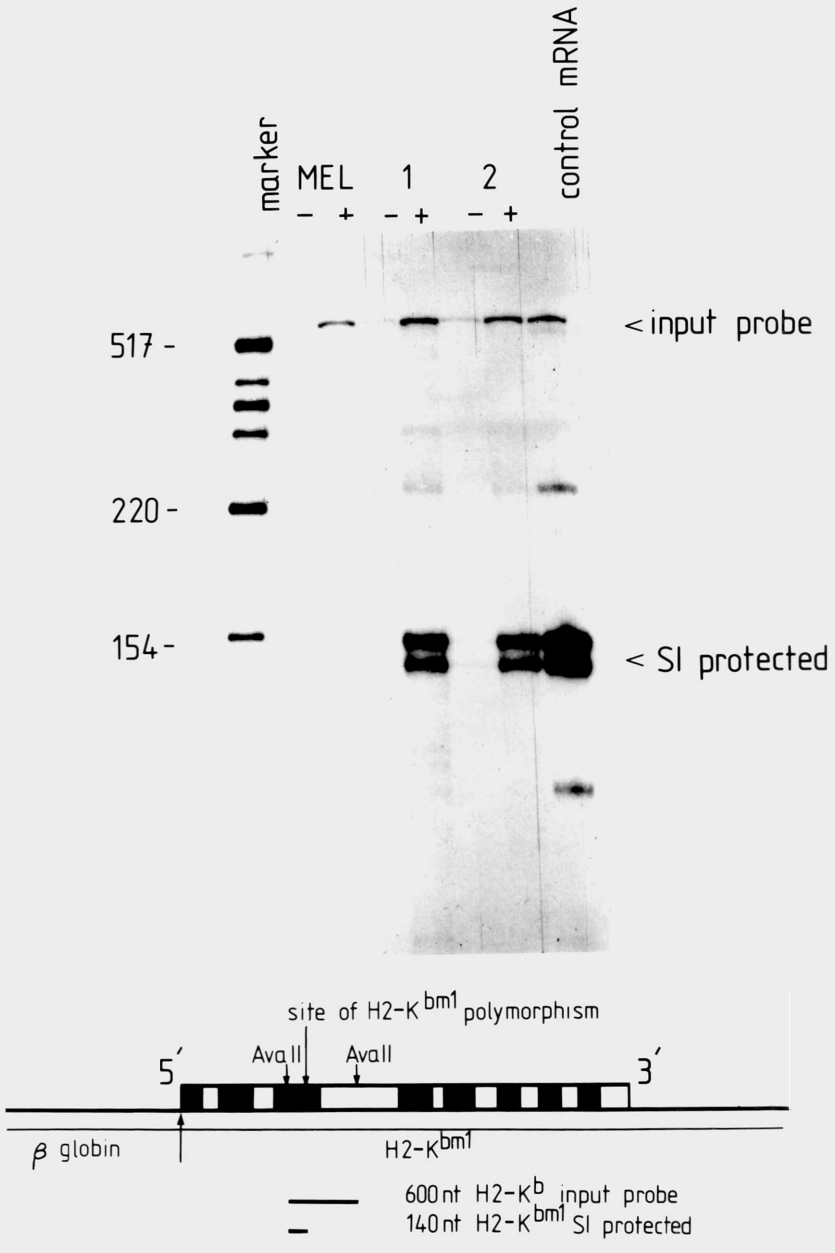

MARKER
1
2
3
Control mRNA
50 pg
– + – + – +
517 –
220 –
154 –
< input probe
< SI protected
5′
EcoRI
3′
EcoRI
β globin
δ globin
input probe 600 nt
SI protected 168 nt

Figure 35

SI nuclease analysis of 5'β-3'H-2K^{bm1} hybrid gene transcripts in MEL transformants

Individual MEL transformants containing the 5'β-3'H-2K^{bm1} hybrid gene were analysed as described in Figure 31. Each hybridization contained 80μg total RNA.


marker
MEL
1
2
control mRNA
– + – + – +
517 -
220 -
154 -
< input probe
< SI protected
site of H2-K bm1 polymorphism
AvaII
AvaII
5′
3′
β globin
H2-K bm1
600 nt H2-K b input probe
140 nt H2-K bm1 SI protected

CHAPTER 8

INDUCTION IN MEL CELLS OF FOREIGN HYBRID GENES CONTAINING HUMAN β-GLOBIN GENE SEQUENCES 3' TO THE TRANSLATION INITIATION SITE

8.1 Introduction

By analysing the expression in MEL cells of hybrid genes containing the human β-globin gene 5' flanking region it was shown in Chapter 7 that these DNA sequences can confer inducibility upon the noninducible human γ-globin and mouse H-2K^{bm1} genes. The possibility that some regulatory DNA sequences also resided within the β-globin gene itself was investigated here by analysing the expression of hybrid genes of the converse construction, i.e. ones in which the 5' flanking region of the noninducible human γ-globin and mouse H-2K^{bm1} genes were adjoined to the complementary structural gene region of the inducible human β-globin gene.

8.2 Construction of 5'γ-3'β and 5'H-2K^{bm1}-3'β hybrid genes

Hybrid genes were constructed as shown in Figures 36 and 37. The 5'γ-3'β hybrid gene was constructed at the NcoI site at the translation initiation codon of both human β- and γ-globin genes. The hybrid gene contained 200 bp of the 5' flanking region of the γ-globin gene and 1.5 kb of the 3' flanking region of the human β-globin gene (Figure 36).

The 5'H-2K^{bm1}-3'β hybrid gene was constructed at the NcoI

site of the β-globin gene and the KpnI site of the H-2K^{bm1} gene; the latter site lies between the transcription initiation and translation initiation sites. The hybrid gene contained approximately 4 kb DNA from the 5' flanking region of the H-2K^{bm1} gene and 1.5 kb DNA from the 3' flanking region of the β-globin gene (Figure 37).

8.3 Expression of 5'γ-3'β and 5'H-2K^{bm1}-3'β hybrid genes in MEL transformants

5'γ-3'β and 5'H-2K^{bm1}-3'β hybrid genes were introduced into MEL cells by cotransfection with the cosmid pRT. Upon differentiation, four out of eight clones transfected with the 5'γ-3'β gene showed a 5->100 fold induction of hybrid gene transcript as measured by SI nuclease analysis using a probe which detected the 3' end of human β-globin mRNA (Figure 38). Four out of five clones transfected with the 5' H-2K^{bm1}-3'β hybrid gene showed a 5->50 fold induction of hybrid gene transcript as measured by the same SI nuclease analysis (Figure 39).

The inducibility of both the 5'γ-3'β and 5'H-2K^{bm1}-3'β hybrid genes suggested that some of the regulatory DNA sequences required for β-globin gene activation resided 3' to the translation initiation site.

8.4 Induction of 5'γ-3'β and 5'H-2K^{bm1}-3'β hybrid genes is due to transcriptional activation

The induction in MEL cells of introduced hybrid genes containing human β-globin structural gene sequences (5'γ-3'β and 5'H-2K^{bm1}-3'β) may be due to either an increase in relative stability of hybrid gene transript or to an increase in rate of transcription of the genes upon differentiation (see section 3.7). The relative rate of transcription of these hybrid genes in MEL transformants was therefore measured as described in section 3.7 using a nuclear transcription runoff assay.

Labelled nuclear transcripts were synthesized in two MEL transformants which showed a high induction of 5'γ-3'β and 5'H-2K^{bm1}-3'β hybrid gene transcript as measured by SI nuclease analysis. Nuclear transcripts from induced and noninduced cells were hybridized to mouse and human β- globin gene probes which were immobilized as dots on nitrocellulose filters (Figure 40). Differentiation of the MEL cells containing either 5'H-2K^{bm1}-3'β or 5'γ-3'β hybrid genes resulted in a >10 fold increase in hybridization to both the mouse and human globin probes, indicating an increase in the rate of transcription of both the endogenous mouse globin and introduced hybrid globin genes. Thus, the increase in level of hybrid 5'γ-3'β and 5'H-2K^{bm1}-3'β gene transcripts following MEL differentiation resulted at least in part from an increase in the rate of transcription of these genes.

8.5 Discussion

The presence of regulatory DNA sequences within the 5' flanking region of the human β-globin gene has been previously shown by analysing the expression in MEL cells of 5'β-3'γ and 5'β-3'H-2K^{bm1} hybrid genes. Surprisingly, hybrid genes containing human β-globin gene sequences 3' to the translation initiation site (5'H-2K^{bm1}-3'β and 5'γ-3'β) were also inducible in MEL cells, with the induction resulting at least in part from an increase in the rate of transcription of these genes upon MEL differentiation. DNA sequences which regulate β-globin gene activation during erythroid differentiation therefore reside both 5' and 3' to the translation initiation codon. The inducibility of both the 5'β-3'H-2K^{bm1} and 5'H-2K^{bm1}-3'β hybrid genes further indicates that these regulatory elements may act independently. Because the level of induction of hybrid gene transcripts was variable between transformants, the quantitative contribution of each regulatory region cannot be assessed. The significance of these data is discussed further in the following chapter (section 9.3).

Figure 36

Construction of the 5'γ-3'β hybrid human globin gene

The 5'γ-3'β hybrid gene was constructed from globin gene subclones γ-pRT-A and β-pRT-2 (Figure 5).

The 190 bp NcoI fragment containing the human γ-globin gene 5' flanking region was isolated from γ-pRT and the 6.7 kb NcoI fragment containing the 3' region of the human β-globin gene isolated from β-pRT. The β- and γ-globin gene NcoI fragments were ligated and the ligation mixture transfected into E.coli HB101. Minipreparations of DNA from transformants were screened by digestion with NcoI, EcoRI, and BalI+BamHI for the desired construct.

N = NcoI restriction site
H = HindIII " "

Bg/B = site of cloning of BglII fragment into BamHI site (both sites destroyed).

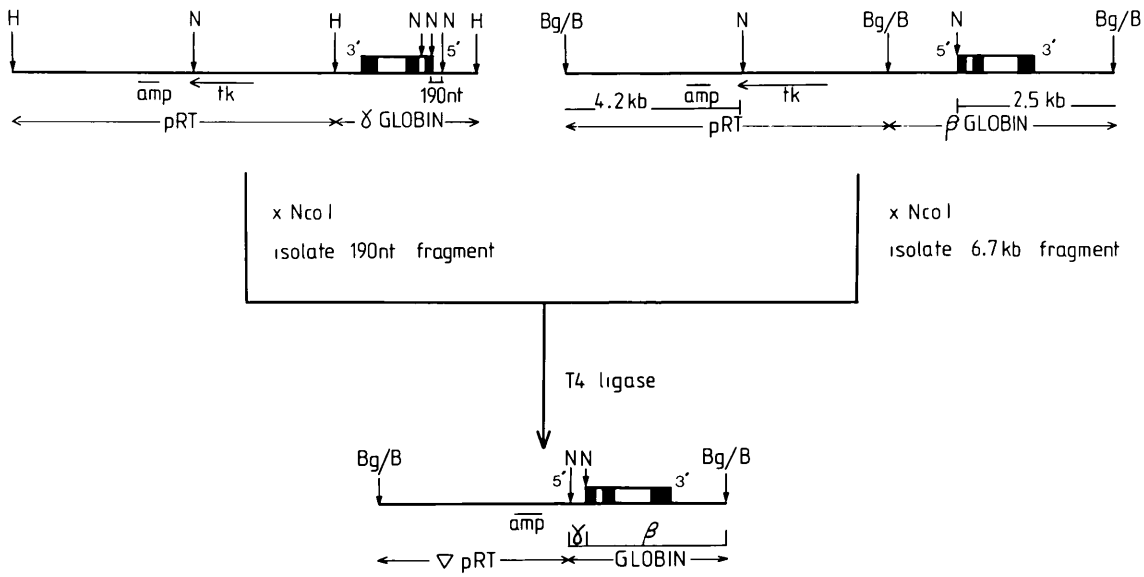

H
N
H
N N N
H
3'
5'
amp
tk
190nt
pRT
γ GLOBIN
Bg/B
N
Bg/B
N
Bg/B
5'
3'
4.2 kb
amp
tk
2.5 kb
pRT
β GLOBIN
x Nco I
isolate 190nt fragment
x Nco I
isolate 6.7 kb fragment
T4 ligase
Bg/B
NN
Bg/B
5'
3'
amp
γ
β
▽ pRT
GLOBIN

Figure 37

Construction of the 5'H2-K^{bm1}-3'β-globin hybrid gene

The 5'H-$2K^{bm1}$-3'β hybrid gene was constructed from the human β-globin gene clone HβpBSV, which contained the human β-globin gene 4.7 kb BglII fragment inserted into the BamHI site of the vector pBSV, and from the H2-K^{bm1} clone illustrated. The latter clone contained the 8 kb HindIII-EcoRI fragment of the H2-K^{bm1} gene inserted into a BPV vector and was a gift of A. Rosenthal.

The human β-globin gene clone was digested with NcoI and the 3' recessed termini at the NcoI sites were made blunt by treatment with reverse transcriptase in the presence of 1mM of each of dATP, dGTP, dCTP, TTP. The sample was then digested with HindIII and the 7.5 kb NcoI-HindIII fragment isolated from a 0.5% low melting agarose gel.

The H2-K^{bm1} clone was digested with KpnI and the 3' protruding termini at the KpnI sites removed by treatment with Klenow polymerase in the presence of 500μM dGTP at 20°C for 2 hours. The sample was then digested with HindIII and the 2 kb HindIII-KpnI fragment isolated from a 0.7% low melting agarose gel.

The 2 kb HindIII-KpnI H2-K^{bm1} and the 7.5 kb HindIII-NcoI β-globin fragments were ligated, the ligation mixture transfected into E.coli HB101, and minipreparations of DNA from the transformants analysed by digestion with EcoRI and with NruI+BamHI for the desired construct.

N, K, Hd, E represent NcoI, KpnI, HindIII and EcoRI restriction sites.
K/N represents ligation of blunted KpnI and NcoI sites (both sites destroyed).
Bg/B represents cloning of BglII fragment into BamHI site (both sites destroyed).

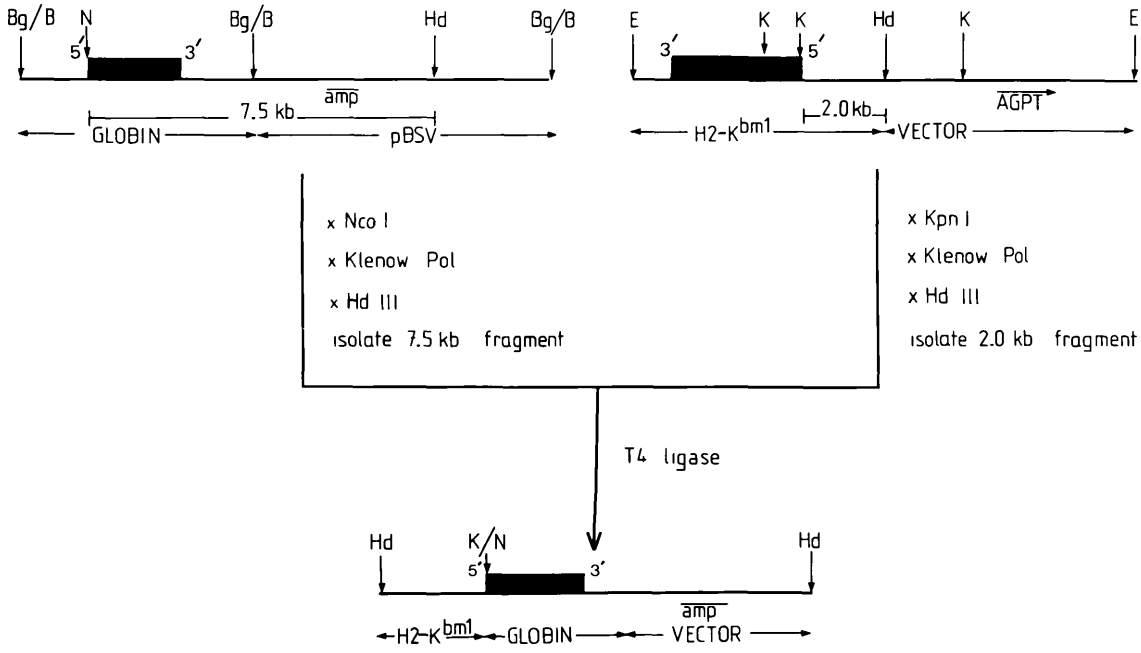

Bg/B
N
Bg/B
Hd
Bg/B
5′
3′
amp
7.5 kb
GLOBIN
pBSV
E
K
K
Hd
K
E
3′
5′
AGPT
2.0kb
H2-K^bm1
VECTOR
x Nco I
x Klenow Pol
x Hd III
isolate 7.5 kb fragment
x Kpn I
x Klenow Pol
x Hd III
isolate 2.0 kb fragment
T4 ligase
Hd
K/N
Hd
5′
3′
amp
H2-K^bm1
GLOBIN
VECTOR

Figure 38

SI nuclease analysis of 5'γ-3'β hybrid gene transcripts in MEL transformants

Individual MEL transformants containing the 5'γ-3'β hybrid human globin gene were grown for 3 days in the presence (+) or absence (-) of the differentiation inducer HMBA. RNA was prepared by the guanidinium chloride method and 10μg samples were analysed by SI nuclease mapping using the human β-globin gene 3' probe illustrated. The probe was end-labelled by reverse transcription to a specific activity of $>10^7$ cpm/μg. Hybridization was at 52°C; SI nuclease digestion was for 2 hours at 20°C. Products were electrophoresed on a 7M urea 7% acrylamide gel. Control $polyA^+$ RNA was obtained from a patient with sickle cell anaemia.

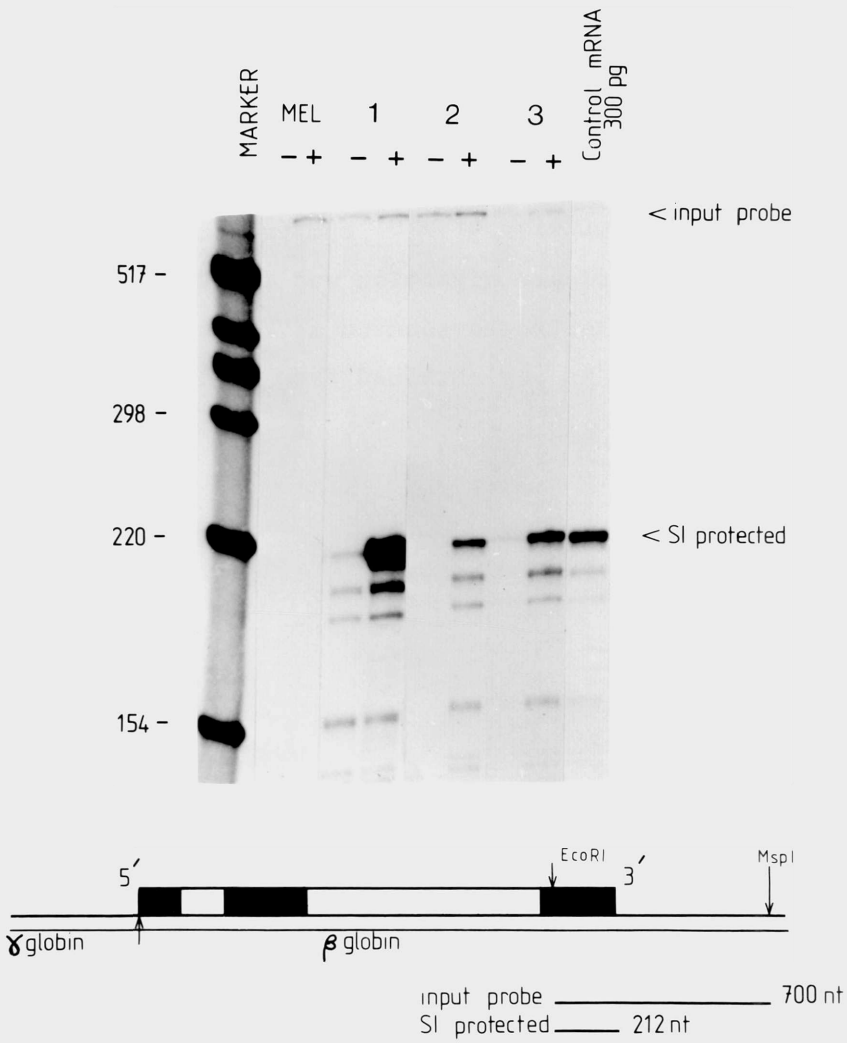

MARKER
MEL
1
2
3
Control mRNA 300 pg
− + − + − + − +
< input probe
517 –
298 –
220 –
154 –
< SI protected
5′
EcoRI
3′
MspI
γ globin
β globin
input probe 700 nt
SI protected 212 nt

Figure 39

SI nuclease analysis of 5'H-2-3'β transcripts in MEL transformants

Individual MEL transformants containing the 5'H-2K^{bm1}-3'β hybrid gene were grown for 3 days in the presence (+) or absence (−) of 3mM HMBA. RNA was prepared and analysed using the 3'human β-globin gene probe illustrated and the procedure described in Figure 38. Each hybridization contained 10μg total RNA.

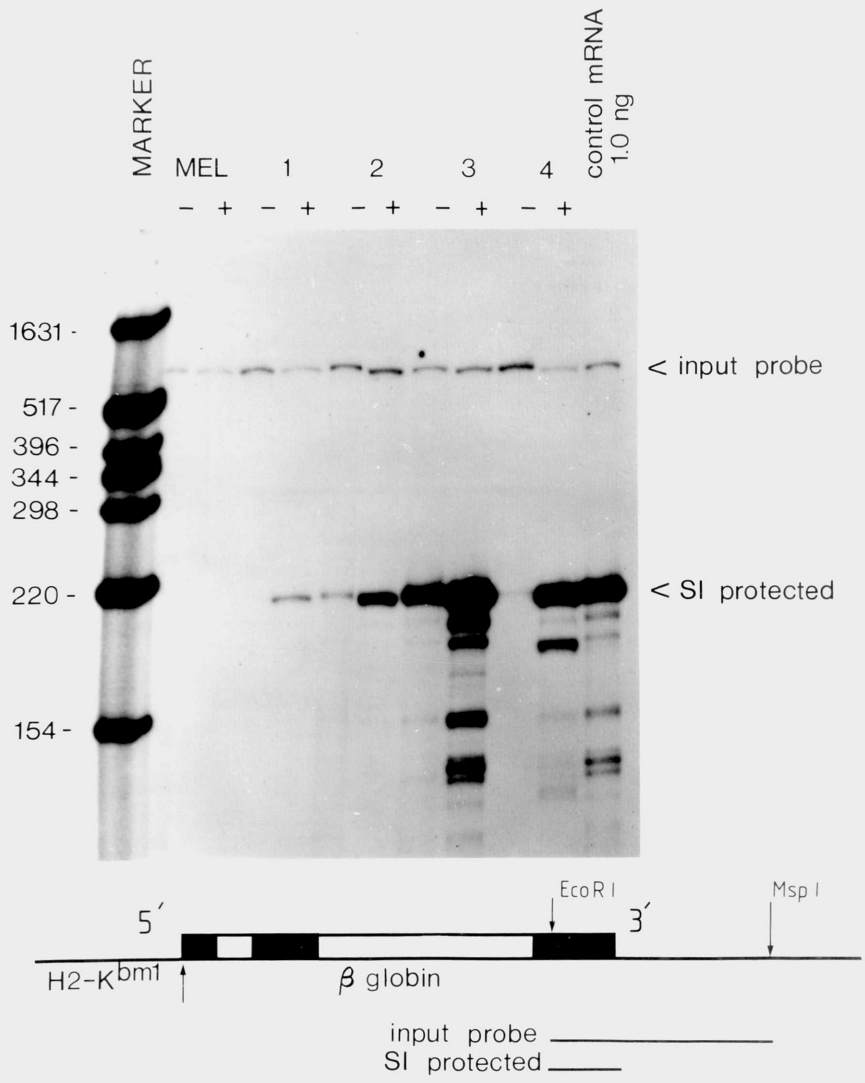
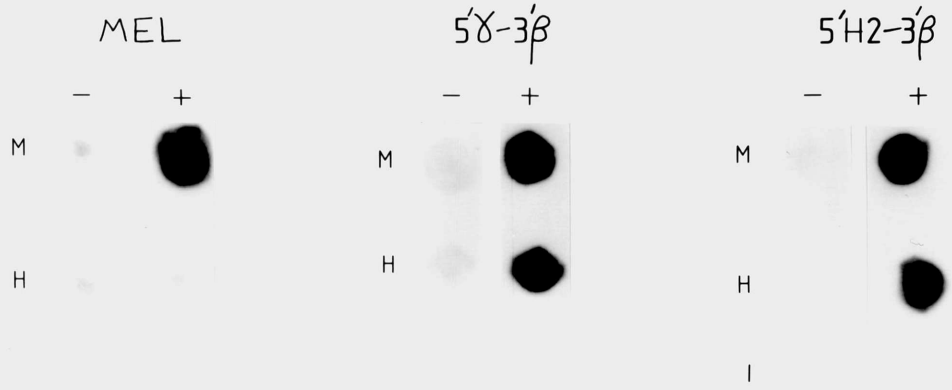

MARKER
MEL
1
2
3
4
control mRNA
1.0 ng
− + − + − + − + − +
1631 -
517 -
396 -
344 -
298 -
220 -
154 -
< input probe
< SI protected
EcoR I
Msp I
5′
3′
H2-K^bm1
β globin
input probe
SI protected

Figure 40

Dot hybridization of labelled nuclear transcripts from MEL transformants containing 5'γ-3'β and 5'H-2K^{bm1}-3'β hybrid genes

2 MEL transformants containing 5'γ-3'β and 5'H2-3'β hybrid genes were grown for 3 days in the presence (+) or absence (-) of 3mM HMBA and nuclei isolated. Labelled nuclear transcripts were synthesized and hybridized to human β-globin (H) and mouse β-major (M) globin probes which were immobilized as dots on nitrocellulose filters. The human β-globin probe was the BamHI-EcoRI IVS2 fragment; the mouse β-globin probe was the HdIII-PstI IVS2 fragment. Nuclear transcripts from the 5' H-2K^{bm1}-3'β transformant were also hybridized to a 1.8 kb BamHI fragment isolated from the human insulin gene probe (I) as a control for background hybridization (Bell et al., 1982).


MEL
5′γ-3′β
5′H2-3′β
−
+
M
H
I

CHAPTER 9

DISCUSSION

9.1 Regulated expression of foreign β-globin genes in MEL cells

Foreign human and rabbit β-globin genes were both inducible during chemically triggered differentiation in MEL cells. The inducibility of the foreign β-globin gene was an inherent property conferred by its DNA sequence and retained after its probable integration into a variety of chromosomal locations and in spite of an acquired DNA methylation and chromatin configuration which was different from that of the endogenous mouse β-globin gene. Several lines of evidence suggest that DNA introduced into eukaryotic cells acquires an active chromatin configuration. For example, both the rat α2 globulin and human growth hormone genes are transcriptionally active and inducible by glucocorticoids when transferred into mouse L cells in which the endogenous counterpart genes remain inactive (Robins et al., 1982; Kurtz, 1981). Similarly, cloned human β-globin genes introduced into HeLa cells in a transient expression assay may be transcriptionally activated in trans by viral immediate-early gene products, whereas the endogenous globin genes in the same cells remain inactive (Green et al., 1983; Imperiale et al., 1983). In these cases, the transcriptional inactivity of the endogenous genes probably reflects their inactive chromatin configuration, since the expression of the exogenous gene indicates that all the factors necessary for their transcription per se are present. The "active"

chromatin configuration acquired by the foreign DNA is possibly influenced by the selection for expression of the HSV tk gene on the vector. Other evidence suggests that an active chromatin configuration is necessary for β-globin gene activation in vivo. For example, the transcriptional inactivity of a β-globin gene in a patient with γβ-thalassemia is probably due to a chromosomal translocation which places the β-globin gene adjacent to a region of chromatin which is normally in an "inactive" configuration. The β-globin gene on the translocation chromosome acquires a similar inactive chromatin configuration and becomes transcriptionally inactive (Kioussis et al., 1983), whereas the β-globin gene on the allelic chromosome in the same cell remains in an "active" chromatin configuration and is transcribed. However, the fact that foreign genes introduced into eukaryotic cells are often not expressed at a "maximal" level until induced by differentiation of the cells, application of hormones or heat shock etc. implies that an "active" chromatin configuration is not by itself sufficient to ensure a high rate of gene transcription. For example, the activation of the endogenous mouse β-major globin gene in MEL cells is associated with the appearance of new DNaseI hypersensitive sites in its 5'flanking DNA; cloned mouse globin genes introduced into MEL cells have been shown to possess the 5' hypersensitive site both before and after chemical induction, even though they are expressed at a high level only after differentiation (Charnay et al., 1984). Other evidence suggesting that chromatin configuration alone does not determine the level of expression of the globin gene in erythroid cells in vivo derives from studies in which the fetal and embryonic globin genes in

adult erythroid cells were activated by treatment of the cells with agents that induce DMA demethylation (e.g. 5azacytidine) or ones that induce chromatin conformation changes (e.g. sodium butyrate; Ginder et al., 1984). The levels of expression of the activated embryonic genes in the adult cell are, however, 50 fold lower than the level of expression of the embryonic gene in an embryonic erythroid cell or adult gene in an adult cell. These low levels of expression of the activated embryonic gene in the adult environment are in agreement with the low constitutive level of expression of the noninducible γ-globin and $H\text{-}2K^{bm1}$ genes, which are probably expressed in MEL transformants, albeit at a low level, due to their acquisition of an active chromatin configuration. These foreign γ-globin and $H\text{-}2K^{bm1}$ genes are not however inducible because they do not respond to the specific adult-type erythroid factors which mediate induction of the adult globin genes.

The high level of transcription of the β-globin gene in induced MEL cells may be mediated by the appearance of specific "activator" factors in the differentiated cell and/or the removal of repressors which are present in the nondifferentiated cell. Charnay et al. (1984) have suggested that a specific repression of globin gene transcription in the nondifferentiated MEL cell may indeed occur. It was found that a cloned human α-globin gene was expressed at a high level in nondifferentiated MEL cell transformants and was not further inducible upon differentiation of the cells with HMBA; the expression level of the foreign α-globin gene was about 100 times greater than that of a foreign human β-globin gene in the same (noninduced) cells. In contrast, when the human α-globin gene was introduced into MEL

cells on an intact chromosome, it was expressed at a low level before differentiation and was inducible on subsequent treatment of the cells with HMBA, suggesting that the incorrect regulation of the foreign cloned human α-globin gene was due to a lack of its repression in the nondifferentiated MEL cell. The fact that the human αbut not β-globin gene is expressed in an enhancer independent manner when introduced into HeLa or COS cell transient expression assays (Mellon et al., 1981; Humphries et al. 1982; Treisman et al., 1983b; section 1.5iii) may further indicate that expression of the human αand β-globin gene families is regulated by fundamentally different mechanisms.

9.2 Individual human β-like globin genes are independently expressed in MEL cells

The human β-globin gene was inducible in MEL cells when introduced either as a short DNA fragment or as cosmids containing 40 kb segments of the β-globin gene cluster; the human γ-globin gene was noninducible when introduced either as a gene fragment or as part of the intact gene cluster. The adult and fetal genes were therefore regulated independently and differently, and the induction was specific for the adult type globin gene in the case of both endogenous mouse and foreign human globin genes; this presumably reflected the fact that the MEL cells were derived from an adult mouse and either lacked the appropriate factors required for fetal and embryonic gene activation or possessed factors responsible for their repression.

Studies of deletion thalassemia and HPFH patients had

previously suggested that DNA sequences residing 3' to the β-globin gene and in the γ-δ intergenic region were involved in repressing γ-globin gene expression in an adult erythroid environment (section 1.4viii); however,the presence or absence of these regions had no effect on γ-globin gene expression in MEL cells. This may imply that the structural organization of the globin gene cluster is not relevant to the activation and inactivation of the individual fetal and adult genes during development. More information on the mechanism of γ-globin gene repression in an adult erythroid environment will come from studying the expression in MEL cells of mutated γ-globin genes and cloned γ-globin genes from HPFH patients. (in the latter disease, expression of the fetal globin genes is not repressed in the adult erythroid environment in vivo, section 1.4viii). Since the level of γ-globin gene expression in vivo in an adult erythroid cell is probably not solely determined by its DNA modification pattern and chromatin structure (see above) the adult and fetal globin genes may differ in the DNA sequences which regulate their developmental stage specific activation; such differences could exist in the sequences which bind a putative activator of adult globin gene expression in the differentiated adult erythroid cell. It is also possible that the γ-globin gene has specific DNA sequences responsible for binding factors which repress its expression in the environment of the differentiating adult erythroid cell.

9.3 DNA sequences which regulate β-globin gene activation

The inducibility in MEL cells of both 5'β-3'H-2K^{bm1} and 5'H-2K^{bm1}-3'β hybrid globin genes suggested that the DNA sequences which regulate β-globin gene activation reside both 5' and 3' to the translation initiation site and that these regulatory elements may act independently.

Most of the characterized sequence elements which regulate the transcription of a gene are located in its 5' flanking region. Thus, the promoter region of all eukaryotic genes transcribed by RNApolII resides within the 100 bp DNA immediately 5' to the transcription initiation site. In addition, DNA seqences within the 5' flanking DNA of a gene play a role in regulating gene activation in addition to being required for transcription per se (see section 1.5iv) Thus, the region 47-66 bp upstream of the cap site of the Drosophila heat shock protein genes mediates their induction by high temperature (Pelham and Bienz, 1982). Similarly, DNA sequences residing within 117 bp of the transcription initiation site of the interferon gene regulate its induction by virus or poly I poly C (Ragg and Weissman, 1983; Weidle and Weissman, 1983). In other cases, regulatory sequences have been located in the 5' flanking region of the gene, although their presence elsewhere in the gene has not been rigorously investigated. For example, by analysing the expression of hybrid genes containing the 5' flanking region of a glucocorticoid responsive gene ligated to the complementary region of a nonresponsive gene, it was shown that the 5' flanking DNA of the growth hormone, MMTV, and metallothionen genes contained the glucocorticoid responsive element

(Robins et al., 1982; Huang et al., 1981; Karin et al., 1984a,b). This was further substantiated by the demonstration of binding of the glucocorticoid hormone-receptor complex to this 5' region of the MMTV LTR region (Payvar et al., 1983) and metallothionen gene (Karin et al., 1984a). However, the finding of glucocorticoid binding downstream from the transcription initiation site in the MMTV LTR region (Payvar et al., 1983) and of a sequence of exact homology to the gluocorticoid receptor binding site within the first intron of the growth hormone gene raises the possibility that regulatory elements are also located downstream from the transcription initiation site. From analysis of the rabbit β-globin deletion mutants (Chapter 6) and hybrid genes (Chapter 7) it was not possible to precisely localize the 5' regulatory region of the β-globin gene, since all the rabbit β-globin gene deletion mutants contained the β-globin structural sequences which themselves possessed inducing elements (Chapter 8). The localization of this 5' regulatory region is at present being determined by the construction of 5'β-3'H-2K^{bm1} hybrid genes in which the 5'-β region is mutated.

The 5' regulatory sequences of the β-globin gene might mediate its induction either via the binding of activator molecules in the differentiated cell, as, for example, the binding of glucocorticoid receptor activates hormone responsive genes, or by the binding of repressor molecules in the nondifferentiated cell, with these repressors being released upon differentiation. The fact that constructs containing the 5' flanking region of the human β-globin gene (βpRT, 5'β-3'H-2K^{bm1}) showed a lower level of expression in the noninduced MEL cell than did constructs containing a 5'

$H\text{-}2K^{bm1}$ region ($H\text{-}2K^{bm1}$, 5'H2-3'β) may indicate that a repression of the β-globin gene in the nondifferentiated MEL cell as mediated by its 5' flanking region does indeed occur; the major histocompatibility $H\text{-}2K^{bm1}$ gene, which is physiologically expressed at a low level both before and after differentiation, does not show such repression. Differentiation of the MEL cells might result in loss of repression on this 5'β region and hence induction of the β-globin gene.

It was surprising to find that DNA sequences 3' to the translation initiation site of the human β-globin gene contain a regulatory element. The fact that constructs containing this 3' β region (5'$H\text{-}2K^{bm1}$-3'β and β-pRT) are expressed at a higher level after MEL differentiation than are ones which contain a 3' $H\text{-}2K^{bm1}$ region (5'β-3'$H\text{-}2K^{bm1}$ and pTME10) may suggest that the β-globin gene 3' regulatory element acts by binding activators of globin gene expression in the differentiated cell. There are several well defined systems in which gene transcription is influenced by DNA sequences residing 3' to the transcription initiation site. The Xenopus 5S RNA genes contain an internal promoter region which lies 50-80 bp downstream from the transcription initiation site, which is required for transcription initiation, and which binds the factor TFIIIA that is essential for 5S gene transcription (Bogenhagen et al,. 1980; Sakonju et al., 1980; Engelke et al., 1980). DNA sequences located downstream from the cap site may also regulate the transcription of genes transcribed by polII; thus viral enhancer elements activate transcription of heterologous genes in a transient epression assay when located at

varying distances either 5' or 3' to the gene (Banerji et al., 1981; section 1.5iii). The finding of a tissue specific enhancer element downstream from the transcription initiation site of the immunoglobulin genes suggests that such elements may play a physiological role in the regulation of eukaryotic gene expression. The immunoglobulin enhancer is located in an intron upstream from the constant region gene and probably serves to activate the variable region promoter of the functionally rearranged Ig gene in the lymphocyte. The enhancer will activate the expression of heterologous genes, e.g. β-globin, when ligated to them and introduced into T cells in both a transient expression assay and in stable clones; the enhancer does not function in other cell types. (Gillies et al., 1983; Queen and Baltimore, 1983; Banerji et al., 1983). The 3' regulatory region of the β-globin gene may similarly contain a tissue specific enhancer element which is active in differentiated but not in nondifferentiated MEL cells and which serves to specifically activate the globin gene. Recent evidence suggests that enhancer elements interact with specific cellular factors, with the tissue and host specificity of the enhancer being mediated by its ability to bind only the factors present in a specific cell type (Scholer and Gruss, 1984). Thus, both the immunoglobulin and putative globin enhancers could mediate tissue specific gene activation by their ability to bind only lymphocyte or erythroid factors. The precise location of the human β-globin 3'-regulatory element is at present being determined by studying the expression in MEL cells of gene hybrids comprising the mouse $H\text{-}2K^{bm1}$ gene into which short fragments (400 bp) derived from

different regions of the human β-globin gene have been inserted. If the regulatory element is precisely localized it will then be possible to ligate in various positions and orientations to a heterologous gene and determine whether it functions as a true erythroid cell specific enhancer. A final understanding of the mechanisms by which these regulatory elements of the β-globin gene mediate its transcriptional activation will require the identification of the erythroid specific cellular factors with which they interact.

REFERENCES

Aaij, C. and Borst, P. (1972) Biochem. Biophys. Acta. 269, 192

Allan, M., Grindlay, G. J., Stefani, L. and Paul, J. (1982) Nucl. Acids Res. 10, 5133

Allan, M., Lanyon, W. G. and Paul, J. (1983) Cell 35, 187

Allen,H., Wraith, D., Pala, P., Askonas, B. and Flavell, R. A. (1984) Nature 309, 279

Antoniarkis, S. E., Philips, J., Mallonee, R., Kazazian, H., Fearon, E., Waber, P., Kronenberg, H., Ullrich, A., and Meyer, D. A. (1983) Proc. Natl. Acad. Sci. U.S.A. 60, 6615

Aviv, H., Voloch, Z., Bastos, R. and Levy, S. (1976) Cell 8, 495

Axelrad, A. A., McLeod, D., Shreeve, M. and Heath, D. S. (1973) In: Proc. Second. Int. Workshop on Hemopoiesis in Culture (Ed. Robinson, W. A.)

Baglioni, C. (1962) Proc. Natl. Acad. Sci. U.S.A. 48, 1880

Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell 27, 299

Banerji, J., Olson, L. and Schaffner, W. (1983) Cell 33, 729

Barrie, P. A., Jeffries, A. J. and Scott, A. F. (1981) J. Mol. Biol. 149, 319

Behe, M. and Felsenfeld, G. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 1619

Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucl. Acids. Res. 8, 127

Benoist, C. and Chambon, P. (1981) Nature 290, 304

Berk, A. J. and Sharp, P. A. (1977) Cell 12, 721

Bernards, R., Kooter, J. M. and Flavell, R. A. (1979) Gene 6, 265

Bernstein, A., Mak, T. and Stephenson, J. (1977) Cell 12, 287

Bienz, M. and Pelham, H. (1982) EMBO J. 1, 1583

Birchmeier, C., Folk, W. and Birnstiel, M. L. (1983) Cell 35, 433

Bird, A. (1978) J. Mol. Biol. 118, 49

Bird, A. and Southern, E. M. (1978) J. Mol. Biol. 118, 27

Birnboim, H. C. and Doly, J. (1979) Nucl. Acids Res. 7, 1513.

Blanchet, J. P., Fleishman, R. A. and Mintz, B. (1982) Dev. Genet. 3, 197

Blanchetot, A., Wilson, V., Wood, D. and Jeffreys, A. J. (1983) Nature 301, 732

Bogenhagen, D. F., Sakonju, S. and Brown, D. D. (1980) Cell 19, 27

Boyer, H. W. and Rouland-Dussoix, D. (1969) J. Mol. Biol. 41, 459

Bradley, A., Evans, M., Kaufman, M. and Robertson, E. (1984) Nature 309, 255

Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 4853

Breathnach, R., Mantei, N. and Chambon, P. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 740

Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349

Brinster, R. L., Chen, H. Y., Trumbauer, M., Senear, A. W., Warren, R., and Palmiter, R. D. (1981) Cell 27, 223

Brinster, R. L., Ritchie, K. A., Hammer, R. E., O'Brien, R. L., Arp, B. and Storb, U. (1983) Nature 306, 332

Brinster, R. L. (1984) In: Phil. Soc. Trans. Roy. Soc. London B in press

Bryan, P. N., Olah, J. and Birnstiel, M. L. (1983) Cell 33, 843

Bunch, C., Wood, W. G., Weatherall, D. J. Br. J. Hematol. 49, 325

Burgess, A. W., Camakaris, J. and Metcalf, D. (1977) J. Biol. Chem. 252, 1998

Busslinger, M., Portmann, R. and Bernstiel, M. L. (1979) Nucl. Acids. Res. 6, 2997

Busslinger, M., Moschonas, N. and Flavell, R. A. (1981) Cell 27, 289

Busslinger, M., Hurst, J. and Flavell, R. A. (1983) Cell 34, 197

Camerini-Otero, R. D. and Zasloff, M. A. (1980) Proc. Natl, Acad. Sci. U.S.A. 77, 5079

Capecchi, M. R. (1980) Cell 22, 479

Casey, J. and Davidson, N. (1977) Nucl. Acids. Res. 4, 1539

Chaconas, G. and van de Sande, J. (1980) Methods Enzymol. 65, 75

Chandler, V., Maler, B. and Yamamoto, K. (1983) Cell 33, 489

Chang, L. J-A and Kan, Y. W. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 2886

Charnay, P., Treisman, R., Mellon, P., Chao, M., Axel, R. and Maniatis, T. (1984) Cell 38, 251

Chirgwin, J. M., Przybyla, A., MacDonald, R. and Rutter, W. (1979) Biochemistry 18, 5294

Cohen, S. N., Chang, A. C. Y. and Hsu, L. (1973) Proc. Natl. Acad. Sci. U.S.A. 69, 2110.

Colbere-Garapin, F., Horodniceanu, F., Kourilsky P. and Garapin, A-C (1981) J. Mol. Biol. 150, 1

Compere, S. J. and Palmiter, R. D. (1981) Cell 25, 233

Comi, P., Giglioni, B., Ottolenghi, S., Gianni, A., A., Polli, E., Barba, P, Covelli, H., Migliacci, G., Candorelli, M. and Peschle, C. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 362

Corces, V., Pellicer, A., Axel., R. and Meselson, M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 7038

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, L. and Chambon, P. (1980) Science 209, 1406

Costantini, F. and Lacy, E. (1981) Nature 294, 92

Cox, R. A. (1968) Methods Enzymol. 12B, 120

Craik, C. S., Buchman, S. R. and Beychok, S. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 1384

Curtis, P. J. and Weissman, C. (1976) J. Mol. Biol. 106, 1061

Curtis, P. J., Finnigan, A. C. and Rovera, G. (1980) J. Biol. Chem. 255, 8971

Daniels and Blattner, F. (1982) Virology 117, 81

Davidson, E. and Britten, R. (1979) Science 204, 1052

Davison, B. L., Egly, J-M., Mulvihill, E. R. and Chambon, P. (1983) Nature 301, 680

Dean, D. C., Knoll. B. J., Riser, M. E. and O'Malley, B. W. (1983) Nature 305, 551

Deisseroth, A., Nienhuis, A., Turner, P., Velez, R., Anderson, W., Ruddle, F., Lawrence, J., Creagan, R. and Kucherlapati, R. (1977) Cell 12, 205

Deisseroth, A., Nienhuis, A., Lawrence, J., Giles, R., Turner, P and Ruddle, F. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 1456

de Villiers, J. and Schaffner, W. (1981) Nucl. Acids. Res. 9, 6251

de Villiers, J., Olson, L., Tyndall, C. and Schaffner, W. (1982) Nucl. Acids. Res. 10, 7965

Dexter, T. M. (1982) J. Cell Physiol. Suppl. 1, 87

Dierks, P., van Ooyen, A., Mantei, N. and Weissmann, C. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 1411

Dierks, P., Van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J. and Weissmann (1983) Cell 32, 695

Dieterlen-lievre, F. and Martin, C. (1981) Dev. Biol. 88, 180

DiMaio, D., Treisman, R. and Maniatis, T. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 4030

Dodgson, J. B., Strommer, J. and Engel, J. D. (1979) Cell 17, 879

Dolan, M., Sugarman, B., Dodgson, J. and Engel, J. (1981) Cell 24, 669

Dover, G. J., Boyer, S. H. and Zinkham, W. H. (1979) J. Clin. Invest, 62, 173

Dube, S. K., Gaedicke, G., Kluge, N., Weimann, B., Melderis, H., Stenheider, G., Crozier, F., Beckmann, H. and Ostertag, W. (1973) In: Differentiation of and control of malignancy of tumor cells, Eds. Nakaharo, W. et.al. Univ. Tokyo Press.

Dynan, W. S. and Tjian, R. (1983) Cell 35, 79

Efstatiadis, A., Posakony, J., Maniatis, T and Lawn, R. M. (1980) Cell 21, 653

Eisen, H., Keppel-Ballivet, F., Georgopoulos, C., Sassa, S., Granick, J., Pragnell, I. and Ostertag, W. In: Differentiation of normal and neoplastic hematopoietic cells. Cold Spring Harbour

Conferences on Cell Proliferation, Vol 5. (1978)

Eliason, J. F., Testa, N. G. and Dexter, T. M. (1979) Nature 281, 382

Elgin, S. C. R. (1981) Cell 27, 413

Engel, J. D. and Dodgson, J. B. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 1392

Engelke, D. R., Ng, S. Y., Shastry, B. S. and Roeder, R. G. (1980) Cell 19, 717

Fagg, B., Vehmeyer, K., Ostertag, W., Jasmin, C. and Klein, B. (1980) In: In vivo and in vitro erythropoiesis: the Friend system p 163 Elsevier/North Holland ed. Rossi, G. B.

Favaloro, J., Treisman, R. and Kamen, R. (1980) Methods Enzymol. 65, 718.

Flavell, R. A., Kooter, J., deBoer, E., Little, P . and Williamson, R. (1978) Cell 15, 25

Friend, C. (1957) J. Exp. Med. 105, 307

Friend, C. and Haddad, J. R. (1960) J. Nat. Can. Inst. 25, 1279

Friend, C., Patuleia, M. C. and De Harven, E. (1966) Nat. Can. Inst. Monogr. 22, 505

Friend, C., Scher, W., Holland, J. G. and Sato, T. (1971) Proc. Natl. Acad. Sci. U.S.A. 68, 378

Fritsch, E. F., Lawn, R. M. and Maniatis, T. (1979) Nature 279, 599

Fritsch, E. F., Lawn, R. M. and Maniatis, T. (1980) Cell 19, 959

Fromm, M. and Berg, P. (1982) J. Mol. Appl. Gen. 1, 457

Gale, R., Clegg, J. and Huens, E. (1979) Nature 280, 162

Garel, A. and Axel, R. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 3966

Gautsch, J. W. and Wilson, M. C. (1983) Nature 301, 32

Gazitt, Y. and Friend, C. (1980) In: In vivo and in vitro erythropoieis: The Friend Cell System, Ed. Rossi, G. B. Elsevier, North Holland, Amsterdam.

Gerber-Huber, S., May, F. E. B., Westley, B. R., Felber, B. K., Hosbach, H. A., Andres, A. C. and Ryffel, G. U. (1983) Cell 33, 43

Gilbert, W. (1978) Nature 271, 501

Gillies, S. D., Morrison, S. L., Oi, V. T. and Tonegawa, S. (1983) Cell 33, 717

Ginder, G. D., Whitters, M. J. and Pohlman, J. K. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 3954

Gluzman, Y. (1981) Cell 23, 175

Goldberg, M. L. (1979) PhD thesis Stanford University

Goldberg, D. A., Posakony, J. W. and Maniatis, T. (1983) Cell 34, 59

Gordon, J. W., Scangos, G., Plotkin, D., Barbosa, J. and Ruddle, F. H. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 7380

Gough, N., Gough, J., Metcalf, D., Kelso, A., Grail, D., Nicola, N., Burgess, A. W. and Dunn, A.R. (1984) Nature 309, 763

Grabowski, P. J., Padgett, R. A. and Sharp, P. A. (1984) Cell 37, 415

Graham, F. and van der Eb, L. (1973) Virology 53, 536

Graessman, M., Graessmann, A., Wagner, H., Werner, E. and Simon, D. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 6470

Green, M. R., Treisman, R. and Maniatis, T. (1983) Cell 35, 137

Grosveld, F. G., Dahl, H-H, de Boer, E. and Flavell, R. A. (1981) Gene 13, 227

Grosveld, F. G., Lund, T., Murray, E., Dahl, H-H and Flavell, R. A. (1982) Nucl. Acids. Res. 10, 6715

Grosveld, G. C., Koster, A. and Flavell, R. A. (1981a) Cell 23, 573

Grosveld, G. C., Shewmaker, C. K., Jat, P. and Flavell, R.A. (1981b) Cell 25, 215

Grosveld, G. C., de Boer, E., Shewmaker, C. K. and Flavell, R. A. (1982a) Nature 295, 120

Grosveld, G. C., Rosenthal, A. and Flavell, R. A. (1982b) Nucl. Acids. Res. 10, 4951

Groudine M., Eisenman, R. and Weintraub, H. (1981) Nature 292, 311

Groudine, M., Peretz, M. and Weintraub, H. (1981b) Mol. Cell. Biol. 1, 281

Groudine, M. and Weintraub, (1982) Cell 30, 131

Groudine, M. and Casimir, C. (1984) Nucl. Acids. Res. 12, 1427

Grunstein, M. and Hogness, D. (1975) Proc. Natl. Acad. Sci., 72, 3961.

Gruss, P., Dhar, R. and Khoury, G. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 943

Gurdon, J. B. and Brown, D. D. (1978) Dev. Biol. 67, 346

Gusella, J. and Housman, D. (1976) Cell 8, 263

Hanahan, D. and Meselson, M. (1980) Gene, 10, 63.

Hankins, W. D. and Krantz, S. B. (1975) Nature 253, 731

Hansen, V., Tenen, D. G., Livingstone, D. M. and Sharp, P. A. (1981) Cell 27, 603

Harbers, K., Jahner, D. and Jaenisch, R. (1981) Nature 293, 540

Hardison, R. C., Butler, E. T., Lacy, E., Maniatis, T., Rosenthal, N. and Efstratiadis, A. (1979) Cell 18, 1285

Hardison, R. C. (1981) J. Biol. Chem. 256, 117

Harland, R. (1982) Proc. Nat. Acad. Sci. U.S.A. 79, 2323

Harland, R. M., Weintraub, H. and McKnight, S. L. (1983) Nature 302, 38

Harrison, P. R. (1977) In: Int. Rev. Biochem., 15, 227 University Park Press, Baltimore.

Hearing, P and Shenk, T. (1983) Cell 33, 695

Helling, R. B., Goodman, H. M. and Boyer, H. W. (1974) J. Virol., 14, 1235

Hentschel, C. and Bernstiel, M. L. (1981) Cell 25, 301

Higgs, D. R., Goodbourn, S., Lamb, J., Clegg, J. B., Weatherall, D. J. and Proudfoot, N. J. (1983) Nature 306, 398

Hofer, E., Hofer-Warbinek, R.and Darnell, J. E. Jr. (1982) Cell 29, 887

Hollis, G. F., Hieter, P. A., McBride, O., Swan, D. and Leder, P. (1982) Nature 296, 321

Hosbach, H. A., Wyler, T. and Weber, R. (1983) Cell 32, 45

Huang, A., Ostrowski, M., Berard, D. and Hager, G. L. (1981) Cell 27, 245

Huisman, T. H. J., Wrightston, R., Wilson, J., Schroeder, W., Kendall, A. (1972) Arch. Bioch. Biophys. 153, 850

Humphries, R. K., Ley, T., Turner, P., Moulton, A., and Nienhuis, A. W. (1982) Cell 30, 173

Hsuing, N., Warrick, H., deRie, J. R., Tuan, D., Forger, B., Skoultchi, A. and Kucherlapati, R. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 4852

Hu, S. L. and Manley, S. L. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 820

Hynes, N. E., Kennedy, N., Rahmsdorf, U. and Groner, B. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 2038

Hynes, N. E., van Ooyen, A., Kennedy, N., Herrlich, P., Ponta, H. and Groner, B. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 3637

Imperiale, M. J., Feldman, L. T. and Nevins, J. R. (1983) Cell 35, 127

Ingram, R., Scott, R. M. and Tilghman, S. M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 4694

Iscove, N. N., Sieber, F. and Winterhalter, K. H. (1974) J. Cell. Physiol. 83, 309

Iscove, N. N., Guilbert, L. J. and Weyman, C. (1980) Exp. Cell Res. 126, 121

Ish-Horowicz, D. and Burke, J. F. (1981) Nucl. Acids. Res. 9, 2989

Jacobsen, H., Klenow, H., and Ovargaard-Hansen, K. (1974) Eur. J. Biochem., 45, 623

Jaenisch, R. and Mintz, B. (1974) Proc. Natl. Acad. Sci. U.S.A. 71, 1250

Jahn, C. L., Hutchinson, C., Phillips, S., Weaver, N., Haigwood, C., Voliva, F. and Edgell, M. H. (1980) Cell 21, 159

Jeffreys, A. J. and Flavell, R. A. (1977) Cell 12, 1079

Jimenez and Davies (1980) Nature 287, 869

Johnson, G. R. and Moore, M. A. S. (1975) Nature 258, 726

Johnson, G. R. and Metcalf, D. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3879

Jones, P. A. and Taylor, S. M. (1980) Cell 20, 85

Jones, P. A. and Taylor, S. M. (1981) Nucl. Acids. Res. 9, 2933

Kafatos, F. C., Jones, C. W. and Efstratiadis, A. (1979) Nucl. Acids. Res. 7, 1541

Karin, M., Haslinger, A., Holtgrove, H., Richards, R., Krauter, P., Westphal, H. and Beato, M. (1984a) Nature 308, 513

Karin, M., Haslinger, A., Holtgrove, H., Cathala, G., Slater, E. and Baxter, J. D. (1984b) Cell 36, 371

Kavathas, P. and Herzenberg, L. A. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 524

Kedes, L. H. (1979) Ann. Rev. Bioch. 48, 837

Kelly, R. G., Cozzarelli, N., Deutscher, M. P., Lehman, I. and Korberg, A. (1970) J. Biol. Chem. 245, 39

Kidoguchi, K., Ogawa, M., Karam, J., McNiel, J. and Fitch, M. (1978) Hemoglobin 2, 503

Kioussis, D., Vanin, E. F., de Lange, T., Flavell, R. A. and Grosveld, F. G. (1983) Nature 38, 662

Klenow, H. and Henningsen, I. (1970) Proc. Natl. Acad. Sci., 65, 168.

Korn, A. P., Henkelman, R., Ottensmeyer, F. and Till, J. E. (1973) Exp. Hematol. 1, 362

Kressmann, A., Clarkson, S. G., Telford, J. and Birnstiel, M. (1978) Cold. Spring. Harbour. Quant. Biol. 42, 1077

Kruczek, I. and Doerfler, W. (1982) EMBO J. 1, 409

Kurtz, D. T. (1981) Nature 291, 629

Lacy, E., Hardison, R. C., Quon, D. and Maniatis, T. (1979) Cell 18, 1273

Lacy, E., Roberts, S., Evans. E., Burtenshaw, M. and Costantini, F. (1983) Cell 34, 343

Laimins, L A., Khoury, G., Gorman, C., Howard, B. and Gruss, P. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 6453

Larsen, A. and Weintraub, H. (1982) Cell 29, 609

Laskey, R. A. and Mills, A. D. (1977) FEBS Lett. 82, 314.

Lauer, J., Shen, C-K. J. and Maniatis, T. (1980) Cell 20, 119

Law, M-F., Lowy, D. R., Dvoretzky, I. and Howley, P. M. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 2727

Lawn, R. M., Fritsch, E., Parker, R. C., Blake, G. and Maniatis, T. (1978) Cell 15, 1157

Lawn, R. M., Efstatiadis, A., O'Connell, C. and Maniatis, T. (1980) Cell 21, 647

Lerner, M. R., Boyle, J., Mount, S. M., Wolin, S. L. and Steitz, J. A. (1980) Nature 283, 220

Levinson, B., Khoury, G., Van de Woude, G., and Gruss, P. (1982) Nature 295, 568

Liao, S. K. and Axelrad, A. A. (1975) Int. J. Cancer 15, 467

Lieberman, M. W., Beach, L. R. and Palmiter, R. D. (1983) Cell 35, 207

Lin, P-F., Zhao, S-H., and Ruddle, F. H. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 6528

Lingrel, J. B. et. al. (1983) in: Hemoglobin switching Eds, Stamatoyannopoulos, G. and Nienhuis, A. W. A. R. Liss, N.Y.

Little, P., Flavell, R. A., Kooter, J. M., Annison, G. and Williamson, R. (1979) Nature 278, 227

Lowy, I., Pellicer, A., Jackson, J., Sim, G-K., Silverstein, S. and Axel, R. (1980) Cell 22, 817

Luchnik, A., Bakeyev, V., Zbarsky, I. and Georgiev, G. (1982) EMBO J. 1, 1353

Luciw, P. A., Bishop, J. M., Varmus, H. E. and Cappechi, M. R. (1983) Cell 33, 705

Lusky, M., Berg, L., Weiher, H. and Botchan, M. (1983) Molec. and Cell. Biol. 3, 1108

MacLeod, D. and Bird, A. (1983) Nature 306, 200

Majumdar, S. K. and Bikender, J. D. (1975) J. Nat. Can. Inst. 54, 503

Mandel, M. and Higa, A. (1970) J. Mol. Biol., 53, 154.

Maniatis, T., Jeffrey, A. and Kleid, D. G. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 1184

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory

Manley, J. L., Fire, A., Cano, A., Sharp, P. A. and Gefter, M. L. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 3855

Manley, J. L. and Hu, S-L. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 820

Marks, P. A. and Rifkind, R. A. (1978) Ann. Rev. Bioch. 47, 419

Martin, C., Beaupain, D., Dieter-Lievre, F. (1980) Dev. Biol. 75, 303

Martin, G. R. (1980) Science 209, 768

Matioli, G. (1973) J. Reticuloendothel. Soc. 14, 380

Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci., 74, 560.

Maxam, A. M. and Gilbert, W. (1980) Methods Enzymol. 65, 499 Eds. Grossman, L. and Moldave, K. Acad. Press. N. Y.

Maxwell, I. H., Maxwell, F. and Hahn, W. E. (1977) Nucl. Acids. Res. 4, 241.

McClelland, M. (1981) Nucl. Acids. Res. 9, 5859

McClelland, M. and Ivarie, R. (1982) Nucl. Acids. Res. 10, 7865

McGhee, J. D., Wood, W. I., Dolan, M., Engel, J. and Felsenfeld, G. (1981) Cell 27, 45

McKnight, G. S., Hammer, R., Kuenzel, E. and Brinster, R. L. (1983) Cell 34, 335

McKnight, S. L. (1980) Nucl. Acids. Res. 8, 5949

McKnight, S. L., Gavis, E. R., Kinsbury, R. and Axel, R. (1981) Cell 25, 385

McKnight, S. L. (1982) Cell 31, 355

Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981) Cell 27, 279

Mellor, A. L., Golden, L., Weiss, E., Bullman, H., Hurst, J., Simpson, E., James, R., Townsend, A., Taylor, P., Schmidt, W., Ferulga, J., Leben, L., Santamaria, M., Atfield, G., Festenstein, H. and Flavell, R. A. (1982) Nature 298, 529

Metcalf, D. (1977) Hemopoietic colonies in vitro: Cloning of normal and leukemic cells. Springer Verlag, Berlin.

Metcalf, D. and Johnson, G. R. (1979) J. Cell. Physiol. 99, 159

Metcalf, D., Johnson, G. R. and Mandel, T. E. (1979) J. Cell. Physiol. 98, 401

Miller, D. M., Turner, P., Nienhuis, A. W., Axelrod, D. E. and Gopalakrishnan, T. (1979) Cell 14, 511

Mintz, B. and Illmensee, K. (1975) Proc. Natl. Acad. Sci. U.S.A. 72, 3585

Mirand, E. A., Prentice, T. C. and Hoffmann, J. G. (1961) Proc. Soc. Exp. Biol. Med. 106, 423

Moore, M. A. S. and Metcalf, D. (1970) Br. J. Hematol. 18, 279

Moreau, P., Hen, H., Wasylyk, B., Everett, R., Gaub, M. and Chambon, P. (1981) Nucl. Acids. Res. 9, 6047

Moschonas, N., deBoer, E., Grosveld, F. G., Dahl, H-H. M., Wright, S., Shewmaker, C. K. and Flavell, R. A. (1981) Nucl. Acids. Res. 9, 4391

Moschonas, N., deBoer, E. and Flavell, R. A. (1982) Nucl. Acids. Res. 10, 2109

Murray, N., Brammar, W. and Murray, K. (1977) Mol. Gen. Genetics 150, 53

Nabeshima, Y., Kuriyama, Y., Muramatsu, M., and Ogata, K. (1984) Nature 308, 333

Nagata, S., Mantei, N. and Weissmann, C. (1980) Nature 287, 401

Neumann, E., Schaefer-Ridder, M., Wang, Y., and Hofschneider, P. H. (1982) EMBO J. 1, 841

Nishioka, Y., Leder, A. and Leder, P. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2806

Nordheim, A. and Rich, A. (1983) Nature 303, 674

Nudel, U., Salmon, J., Fibach, E., Terada, M., Rifkind, R., Marks, P. and Bank, A. (1977) Cell 12, 463

Nute, P. E., Pataryas, H. A. and Stamatoyannopoulos, G. (1973) Am. J. Human. Genet. 25, 271

Orkin, S. H., Old,J., Lazarus, H., Altay, C., Gurgey, A., Weatherall, D. J. and Nathan, D. (1979) Cell 17, 33

Orkin, S. H. and Michelson (1980) Nature 286, 538

Orkin, S. H. and Goff, S. C. (1981) J. Biol. Chem. 256, 9782

Orkin, S. H. and Swerdlow, P. S. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 2475

Ostergaard Jensen, E., Paludan, K., Hyldrij-Nielson, J., Jorgensen, P and Marcker, K. et. al. (1981) Nature 291, 677

Ott, M-O., Sperling, L., Cassio, D., Levilliers, J., Sala-Trepat, J. and Weiss, M. L. (1982) Cell 30, 825

Ottolenghi, S., Giglioni, B., Come, P., Giaanni, A., Polli, E., Acquaye, C., Oldhom, J. H. and Masera, G. (1979) Nature 278, 654

Padgett, R. A., Mount, S. M., Steitz, J. A. and Sharp, P. A. (1983) Cell 35, 101

Palmiter, R., Mulvihill, E., McKnight, S., Sencar, A. (1977) Cold Spring Harbour Symp. Quant. Biol. 42, 639

Palmiter, R. D., Chen, H. Y. and Brinster, R. L. (1982a) Cell 29 701

Palmiter, R. D., Brinster, R., Hammer, R., Trumbauer, M., Rosenfeld, M., Birnberg, N. C. and Evans, R. M. (1982b) Nature 300, 611

Papayannopoulou, T., Brice, M. and Stamatoyannopoulos, G. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 2033

Papayannopoulou, T., Brice,M., and Stamatoyannopoulos, G. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 2933

Papayannopoulou, T., Kalmantis, T. and Stamatoyannopoulos, G. (1979) Proc. Natl. Acad. Sci. 76, 6420

Papayannopoulou, T., Vichinsky, E., and Stamatoyannopoulos, G. (1980) Br. J. Hematol. 44, 535

Papayannopoulou, T. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 6579

Papayannopoulou, T., Tatsis, B., Kurachi, S., Nakamoto, B. and Stamatoyannopoulos, G. (1984) Nature 309, 71

Parker, C. S. and Topol, J. (1984) Cell 37, 273

Pavlakis, G. N. and Hamer, D. H. (1983) Proc. Natl. Acad. Sce. U.S.A. 80, 397

Payvar, F., DeFranco, D., Firestone, G., Edgar, B., Wrange, O., Okret, S., Gustafsson, J-A. and Yamamoto, K. R. (1983) Cell 35, 381

Pelham, H. (1982) Cell 30, 517

Pelham, H. and Bienz, M. (1982) EMBO J. 1, 1473

Pellicer, A., Wagner, E., Kareh, A., Dewey, M., Reuser, A., Silversteinn, S., Axel, R. and Mintz, B. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2098

Popp, R. A., Lalley, P. A., Whitney, J. and Anderson, W. F. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 6362

Pressley, L., Higgs, D. R., Aldridge, B., Metaxatou, A., Clegg, J. B. and Weatherall, D. J. (1980) Nucl. Acids. Res. 8, 4889

Pressley, L., Higgs, D. R., Clegg, J. B. and Weatherall, D. J. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 3586

Proudfoot, N. J. and Brownlee, G. G. (1976) Nature 263, 211

Proudfoot, N. J. and Maniatis, T. (1980) Cell 21, 537

Proudfoot, N. J., Gil, A. and Maniatis, T. (1982) Cell 31, 553

Puckett, C. and Moss, B. (1983) Cell 35, 441

Queen, C. and Baltimore, D. (1983) Cell 33, 741

Radloff, R., Bauer, W. and Vinograd, J. (1967) Proc. Natl. Acad. Sci. U.S.A. 57, 1514

Ragg, H. and Weissman, C. (1983) Nature 303, 439

Ramirez, F., Burns, A., Mears, J. G., Spence, S., Starkman, D. and Bank, A. (1979) Nucl. Acids. Res. 7, 1147

Rassoulzadegan, M., Binetruy, B. and Cuzin, F. (1982) Nature 295, 257

Renkawitz, R., Beug, H., Graf, T., Matthia, P., Grez, M. and Schutz, G. (1982) Cell 31, 167

Reuben, R., Wife, R., Breslow, R., Rifkind, R. and Marks, P. A. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 862

Richardson, C. C. (1971) Proc. Natl. Acad. Sci. 2, 815

Rigby, P. W. J., Diekmann, C., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237

Roberts, R. J. (1978) Gene 4, 183

Robins, D. M., Ripley, S., Henderson, A. S. and Axel. R. (1981) Cell 23, 29

Robins, D. M., Paek,I, Seeburg, P. and Axel, R. (1982) Cell 29, 623

Roginski, R. S., Skoultchi, A. I., Henthorn, P., Smithies, O., Hsiung, N. and Kucherlapati, R. (1983) Cell 35, 149

Rosenthal, A., Wright, S., Cedar, H., Flavell, R. A. and Grosveld, F. G. (1982) Nature 310, 415

Rosenthal, N., Kress, M., Gruss, P. and Khoury, G. (1983) Science 222, 749

Ross, J., Ikawa, Y. and Leder, P. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 3620

Rubin, G. M. and Spradling, A. C. (1982) Science 218, 348

Sakano, H., Rogers, J. H., Huppi, K., Brack, C., Traunecker, A., Maki, R. and Tonegawa, S. (1979) Nature 277, 627

Sakonju, S., Bogenhagen, D. F. and Brown, D. D. (1980) Cell 19, 13

Sanger, F., Air, G., Barrell, B., Brown, N. L., Coulson, A., Fiddes, J., Hutchison, C. A. III, Slocombe, P. and Smith, M. (1977) Nature 265, 687

Sanger, F., Coulson, A., Friedmann, T., Air, G., Barrell, B., Brown, N., Fiddes, J., Hutchison, C. A. III, Slocombe, P. and Smith, M. (1978) J. Mol. Biol. 125, 225

Schaffner, W. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 2163

Schibler, U., Hagenbuchle, O., Wellauer, P. K. and Pittet, A. C. (1983) Cell 33, 501

Schmitz, A. and Galas, D. (1979) Nucl. Acids. Res. 1, 11

Scholer, H. and Gruss, P. (1984) Cell 36, 403

Scholnick, S. B., Morgan, B. A. and Hirsh, J. (1983) Cell 34, 37

Schon, E. A., Wernke, S. M. and Lingrel, J. B. (1982) J. Biol. Chem. 257, 6825

Schroeder, W. A., Huisman, T., Shelton, J. R., Shelton, J. B., Kleihauer, E., Dozy, A. and Robertson, B. (1968) Proc. Natl. Acad. Sci. U.S.A. 60, 537

Schroeder, W. A., Shelton, J. R., Shelton, J. B., Apell, G., Huisman,. T. and Bouver, N. G. (1972) Nature New Biol. 240, 272

Scott, R. W., Vogt, T., Croke, M. and Tilghman, S. M. (1984) Nature 310, 562

Sharp, P. A., Sugden, B. and Sambrook, J. (1973) Biochemistry 12, 3055

Sheffery, M., Rifkind, R. A. and Marks, P. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 1180

Sheffery, M., Rifkind, R. A. and Marks, P. A. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 3349

Shen, C-K. J. and Maniatis, T. (1980) Cell 19, 379

Shibuya, T. and Mak, T. W. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 3721

Shimotohno, K. and Temin, H. M. (1982) Nature 299, 265

Siebenlist, U., Hennighausen, L., Battey, J. and Leder, P. (1984) Cell 37, 381

Slightom, J., Blechl, A. and Smithies, O. (1980) Cell 21, 627

Sompayrac, P. and Danna, K. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 7575

Southern, E. (1975) J. Mol. Biol. 98, 503.

Southern, E. (1980) Methods Enzymol. 69, 152.

Spradling, A. C. and Rubin, G. M. (1982) Science 218, 341

Spradling, A. C. and Rubin, G. M. (1983) Cell 34, 47

Stamatoyannopoulos, G., and Papayannopoulou,T. (1978) In: Cellular and Molecular Regulation in Hemoglobin Switching, Eds. Stamatoyannopoulos, G. and Nienhuis, A. W. Grune and Stratton, N.Y.

Stamatoyannopoulos, G., Roseblum, B., Papayannopoulou, T. and Brice, M. (1979) Blood 54, 440

Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A. and Cedar, H. (1982a) Proc. Natl. Acad. Sci. U.S.A. 79, 61

Stein, R., Razin, A. and Cedar, H. (1982b) Proc. Natl. Acad. Sci. U.S.A. 79, 3418

Steinmetz, M., Winoto, A., Minard, K., and Hood, L. (1982) Cell 28, 489

Sullivan-Read, G. and Summers, W. C. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 5215

Sutcliffe, J. G. (1978) Nucl. Acids Res. 5, 2721

Swanstrom, R. and Shank, P. R. (1978) Anal. Biochem. 86, 184

Szaley, A. A., Grohmann, K. and Sinsheimer, R. L. (1977) Nucl. Acids Res. 4, 1569

Talkington, C. A., Nishioka, Y. and Leder, P. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 7132

Talkington, C. A. and Leder, P. (1982) Nature 298, 192

Tilghman, S. M., Tiemeier, D., Seidman, J., Peterlin, B., Sullivan, M., Maizel, J. and Leder, P. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 725

Till, J. E. and McCulloch, E. A. (1961) Radiat. Res. 14, 213

Till, J. E. and McCulloch, E. A. (1980) Biophys. Biochem. Acta. 605, 431

Townes, T., Shapiro, S., Wernke, S. and Lingrel, J. (1984) J. Biol. Chem. 259, 1896

Treisman, R., Orkin, S. H. and Maniatis, T. (1983) Nature 302, 591

Treisman, R., Green., M. and Maniatis, T. (1983b) Proc. Natl. Acad. Sci. U.S.A. 80, 7428

Ucker, D. S., Ross, S. R. and Yamamoto, K. R. (1981) Cell 27, 257

van der Ploeg, L. H. T. and Flavell, R. A. (1980) Cell 19, 947

van der Ploeg, L. H. T., Konings, A., Oort, M., Roos, D., Bernini, L.and Flavell, R. A. (1980) Nature 283, 637

Vanin, E. F., Goldberg, G. I., Tucker, P. W. and Smithies, O. (1980) Nature 286, 222

Vanyushin, B. F., Tkacheva, S. G. and Belozerdky, A. N. (1970) Nature 225, 948

Vardimon, L., Kressmann, A., Cedar, H., Maechler, M. and Doerfler, W. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 1073

Verma, I. M. (1977) Biophys. Biochem. Acta. 473, 1

Volloch, V. and Housman, D. (1981) Cell 23, 509

Wagner, E. K. (1983) in: Advances in Viral Oncology, Vol 3 Ed, Klein, G. Raven Press N.Y.

Wahl, G. M., Stern, M. and Stark, G. R. (1979) Proc. Natl. Acad. Sci., 76, 3683.

Wahli, W., Dawid, I., Wyler, T., Weber, R. and Ryffel, G. (1980) Cell 20, 107

Wasylyk, B., Kedinger, C., Corder, J., Brisan, O. and Chambon, P. (1980) Nature 285, 388

Wasylyk, B., Wasylyk, C., Augereau, P. and Chambon, P. (1983) Cell 32, 503

Weatherall, D. J., Clegg, J. B. and Wood, W. G. (1976) Lancet ii 660

Weatherall, D. J. and Clegg, J. B. (1979) Cell 16, 467

Weatherall, D. J. and Clegg, J. B. (1982) Cell 29, 7

Weaver, R. F. and Weissman, C. (1979) Nucl. Acids. Res., 6, 1175.

Weidle, U. and Weissmann, C. (1983) Nature 303, 442

Weil, P. A., Luse, D. S., Segall, J. and Roeder, R. G. (1979) Cell 18, 469

Weintraub, H. and Groudine, M. (1976) Science 93, 848

Weintraub, H., Larsen, A. and Groudine, M. (1981) Cell 24, 333

Weisbrod, S. (1982) Nature 297, 289

Weislander, L. (1979) Anal. Biochem. 98, 305

Weiss, B., Jacquemin-Sablon, A., Live, T. R., Fareed, G. and Richardson, C. (1968) J. Biol. Chem. 243, 4543

Weiss, E. H., Mellor, A., Golden, L., Fahrner, K., Simpson, E., Hurst, J. and Flavell, R. A. (1983) Nature 301, 671

Wieringa, B., Meyer, F., Reiser, J. and Weissman, C. (1983) Nature 301, 38

Wigler, M., Silverstein, S., Lee, L., Pellicer, A., Cheng, Y. and Axel, R. (1977) Cell 11, 223

Wigler, M., Sweet, R., Sim, G., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. and Axel, R. (1979) Cell 24, 33

Wigler, M., Levy, D. and Perucho, M. (1981) Cell 24, 33

Wilde, C. D., Crowther, C. E., Cripe, T. P., Lee, M. G. and Cowan, N. J. (1982) Nature 297, 83

Willing, M. C., Nienhuis, A. W. and French Anderson, W. (1979) Nature 277, 534

Wood, W. G. and Weatherall, D. J. (1973) Nature 244, 162

Wood, W. G. and Weatherall, D. J. (1983) Bioch. J. 215, 1

Wood, W. G., Stamatoyannopoulos, G., Lim, G. and Nute, P. E. (1975) Blood 46, 671

Wood, W. G., Nash, J., Weatherall, D. J., Robinson, J. and Harrison, F. A. (1978) In: Cellular and Molecular Regulation of Hemoglobin Switching, Eds. Stamatoyannopoulos, G. and Nienhuis, A. W., Grune and Stratton, N. Y.

Wood, W. G. and Bunch, C. (1983) in: Hemoglobin switching Eds. Stamatoyannopoulos, G. and Nienhuis, A. W. A. R. Liss, N. Y.

Young, R. A., Hagenbuchle, O. and Schibler, U. (1981) Cell 23, 451

Zanjani, E. D., McGlave, P., Bhakthavathsalan, A. and Stamatoyanopoulos, G. (1979) Nature 280, 495

Zanjani, E. D., Lim, G., McGlave, P., Clapp, J., Mann, L., Norwood, T. and Stamatoyanopoulos, G. (1982) Nature 295, 244

Zinn, K., Mellon, P., Ptashne, M. and Maniatis, T. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 4897

Zinn, K., DiMaio, D. and Maniatis, T. (1983) Cell 34, 865

Reprinted from Nature, Vol. 305, No. 5932, pp. 333-336, 22 September 1983

Regulated expression of the human β-globin gene family in murine erythroleukaemia cells

S. Wright, E. deBoer, F. G. Grosveld* & R. A. Flavell†

Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Chemically induced differentiation of cultured murine erythroleukaemia (MEL) cells results in a several hundred-fold increase in transcription of the adult mouse globin genes[1-3] and thus serves as a model for gene activation during erythropoiesis. One approach to study gene regulation in this system has been to analyse the expression of foreign globin genes introduced into MEL cells[4-8]. By introducing cosmid DNA containing the human adult(β), fetal(γ) and embryonic(ε)-globin genes, we have shown here that expression of the β, but not the γ or ε genes, is regulated during MEL differentiation. Regulated expression of the human β-globin gene was observed when it was introduced either as part of the intact globin gene cluster or as an individual gene with 1.5 kilobases (kb) of 5′ flanking DNA. Transcription from a herpes simplex virus (HSV) promoter adjacent to the thymidine kinase (*tk*) gene is also inducible in MEL cells.

Cosmids containing regions of the human β-globin gene cluster (see Fig. 1) were introduced into thymidine kinase-negative (tk^-) MEL cells[6] by calcium phosphate transformation[9]. Before transformation, DNA was linearized within the vector sequences by cleavage with *Pvu*I. Stable tk^+ transformants were obtained at a frequency of 1–10 clones per 10^6 cells per μg DNA. Transformants were shown by Southern blot analysis[10] to contain 1–15 copies of non-rearranged introduced DNA per cell (data not shown); 10–15 transformants containing each cosmid were induced into erythroid differentiation by culturing for 3 days in the presence of 3 mM hexamethylene bisacetamide (HMBA). Levels of human and mouse globin transcripts were quantitated by S_1 nuclease analysis[11,12] using probes which mapped the 5′ and 3′ ends of the mRNAs.

Endogenous mouse β-major globin mRNA levels in transformants as measured by S_1 nuclease analysis of the 3′ end of the mRNA are shown in Table 1 and Fig. 2*a*; 52 of 54 clones tested showed a >100-fold increase in mouse β-globin mRNA levels on differentiation, corresponding to >10,000 copies of mRNA per induced cell. Six of the transformants which showed induction had a high level of mouse β-major globin mRNA before HMBA treatment (>1,000 copies per cell).

Of 40 MEL transformants that contained a foreign human β-globin gene, 30 showed a 4–100-fold increase in the level of human β-globin mRNA on differentiation as measured by S_1 nuclease analysis of the 3′ end of the mRNA (Fig. 2*b*, *c*; Table 1). Of the 10 transformants that showed no induction, two were non-inducible for endogenous mouse β-globin mRNA (one CosHG6N and one β-pRT) and four showed a high level of mouse globin mRNA before addition of the inducing agent (two CosHG6N and two CosHG28TK). Correct initiation of the human globin transcripts was shown by S_1 nuclease mapping of the 5′ end (Fig. 2*d*). Induced human β-globin mRNA levels corresponded to 120–2,500 copies per cell, and induction levels did not correlate with the copy number of human β-globin genes present in the transformant.

Of 19 transformants containing the complete human γ-globin genes (CosHG25 and CosHG28TK), 17 showed the same or a smaller amount of γ-globin mRNA after differentiation (~50 copies per cell; Fig. 2*e*, Table 1). Two transformants showed a less than sixfold inducibility of γ-globin mRNA. Similarly, transformants which contained the complete ε-globin gene (CosHG25) showed the same or reduced levels of ε-globin mRNA after differentiation (data not shown).

CosHG29TK and CosHGΔ17 contain the 3′ ends of the A_γ- and ε-globin genes adjoined to vector sequences as shown in Figs 1 and 3. Surprisingly, transformants containing these cosmids showed a 10–30-fold inducibility of RNAs which were detected by S_1 nuclease analysis using probes which hybridized to the 3′ ends of A_γ- and ε-globin mRNAs. Northern blot analysis[13] showed that the inducible RNA was 700 base pairs (bp) long and that it hybridized with both vector and globin DNA probes (not shown). If we assume the only RNA splicing in this vector–globin hybrid mRNA to be removal of the globin large intervening sequence, the observed RNA length places the transcription initiation site close to that of the HSV *tk* gene[14], but on the opposite DNA strand. This was confirmed by S_1 nuclease mapping using a 510-bp *Xma*I–*Bgl*II probe which covers the predicted 5′ end and gives a 280–300-bp protected band. The 5′ end of this 'anti-tk' (*a-tk*) transcript thus maps ~160 bp away from the *tk* gene cap site (see Fig. 3). A transcript initiating at this site has been previously reported from *in vitro*

* To whom correspondence should be addressed.
† Present address: Biogen Research Corporation, 14 Cambridge Center, Cambridge, Massachusetts 02142, USA.

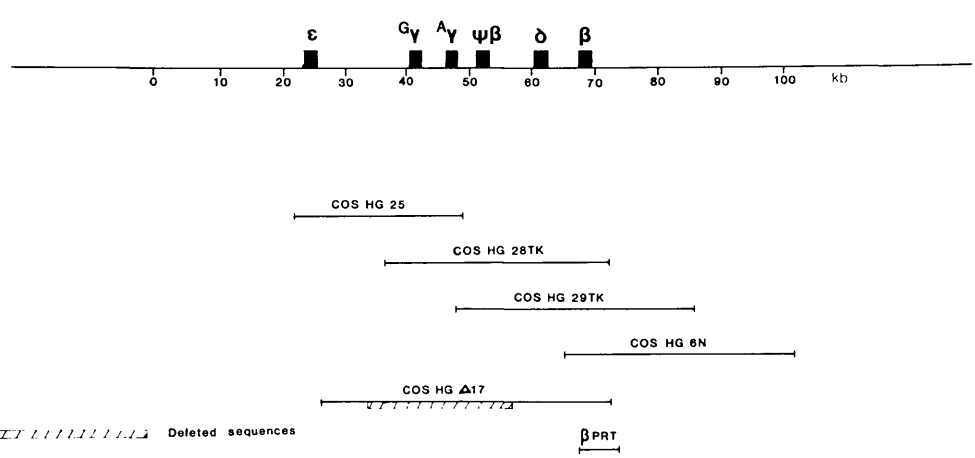


Fig. 1 Structure of human globin gene cosmids introduced into MEL cells. Regions of the human globin cluster as indicated were cloned in cosmid vectors containing the HSV *tk* gene[17,18] (vector sequences not shown). CosHG25 and CosHG6N were isolated directly from a tk vector library. CosHG28TK and CosHG29TK were constructed from previously isolated cosmids, CosHG28 and CosHG29 (ref. 18) by the exchange of vector sequences containing the *tk* gene. CosHGΔ17 was constructed by packaging and transduction of a ligation mixture containing a *Cla*I–*Bam*HI fragment from cosmid pRT, a *Bam*HI–*Xma*I fragment containing the 3′ side of the ε-gene, and *Xma*I–*Pvu*I fragment containing the Gγ-, Aγ-, δ- and β-globin genes and part of the cosmid vector pJB8 (ref. 19). Ligation yielded a DNA concatemer which deleted during the *in vitro* packaging to give deletion cosmids including CosHGΔ17. β-pRT is a subclone of the 4.7-kb β-globin *Bga*I fragment inserted in the *Bam*HI site of vector pRT; the *tk* and β-globin genes in this construction are transcribed in the same direction. Cosmid DNA was prepared by the method of Birnboim and Doly[20] followed by purification on CsCl gradients. *Pvu*I-linearized cosmid DNA was introduced into tk^- MEL cells[6] by calcium phosphate transformation[8] and selection in hypoxanthine–aminopterin–thymidine medium.

Table 1 Induction of human and mouse globin mRNAs in cosmid-containing MEL transformants

Cosmid	Mouse β-major globin mRNA			Human γ-globin mRNA			Human β-globin mRNA		
	No. of clones tested	No. showing induction	No. with high non-induced levels	No. of clones tested	No. showing induction	No. showing no induction	No. of clones tested	No. showing induction	No. showing no induction
CosHG25	9	9	2	9	0	9	—	—	—
CosHGΔ17	8	8	0	—	—	—	8	6	2
CosHG28TK	10	10	2	10	2	8	10	7	3
CosHG29TK	9	9	0	9*	6*	3*	4	4	0
CosHG6N	8	7	2	—	—	—	8	5	3
β-pRT	10	9	0	—	—	—	10	8	2
Total	54	52	6	19	2	17	40	30	10

* Transcription of the γ-globin gene in MEL clones containing CosHG29TK initiates at a vector promoter (see text) and is not included in the totals column.

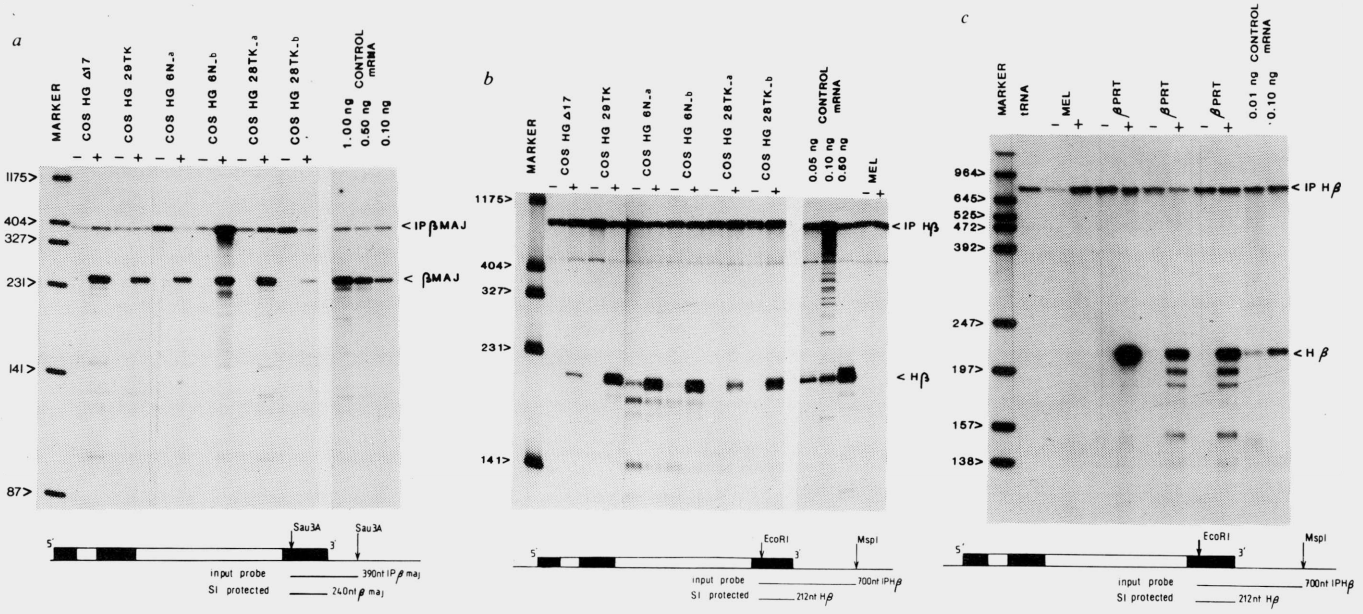


Fig. 2 *a*, S_1 nuclease analysis of mouse β-major globin mRNA in MEL transformants. Control RNA was obtained from mouse reticulocytes. *b*, MEL transformants containing the cosmids indicated and corresponding to those in *a* were analysed using the 3′ human β-globin gene probe illustrated and the procedure described for *a* below; 10 μg of total RNA were used in each hybridization. Control mRNA was obtained from a patient with sickle cell anaemia. *c*, MEL transformants containing cosmid β-pRT were analysed as described for *b*. *d*, S_1 nuclease analysis of the 5′ end of human β-globin transcripts in MEL transformants. *e*, S_1 nuclease analysis of γ-globin transcripts in MEL transformants. Control mRNA was obtained from a $\beta^0/\delta\beta^0$ thalassaemic patient[21].

Methods: *a*, MEL transformants containing the cosmids indicated were grown for 3 days in the presence (+) or absence (−) of 3 mM HMBA. RNA was prepared by the LiCl–urea method[22]. The mouse β-major globin gene probe illustrated (IPβmaj) was end-labelled by reverse transcription and hybridized to 1 μg total RNA at 52 °C for 12–16 h in 10 μl of 80% formamide, 40 mM PIPES *p*H 6.4, 1 mM EDTA, 400 mM NaCl. The mixture was digested with 3,000 units of S_1 nuclease (Boehringer) in 300 μl of 300 mM NaAc *p*H 4.8, 200 mM NaCl, 2 mM $ZnSO_4$ for 2 h at 20 °C. S_1-protected DNA was ethanol-precipitated and electrophoresed on a 7 M urea, 7% acrylamide gel. *d*, MEL transformants containing cosmid β-pRT and corresponding to those used in *c* were analysed. The procedure was as for *a* except that the 5′ human β-globin probe was end-labelled by kinase, strand-separated, and hybridized to 120 μg total RNA in a 50 μl hybridization volume. *e*, RNA from MEL transformants containing the cosmids indicated was analysed using the γ-globin gene probe illustrated and the procedure described for *a*. 10 μg of total RNA were used in each hybridization.

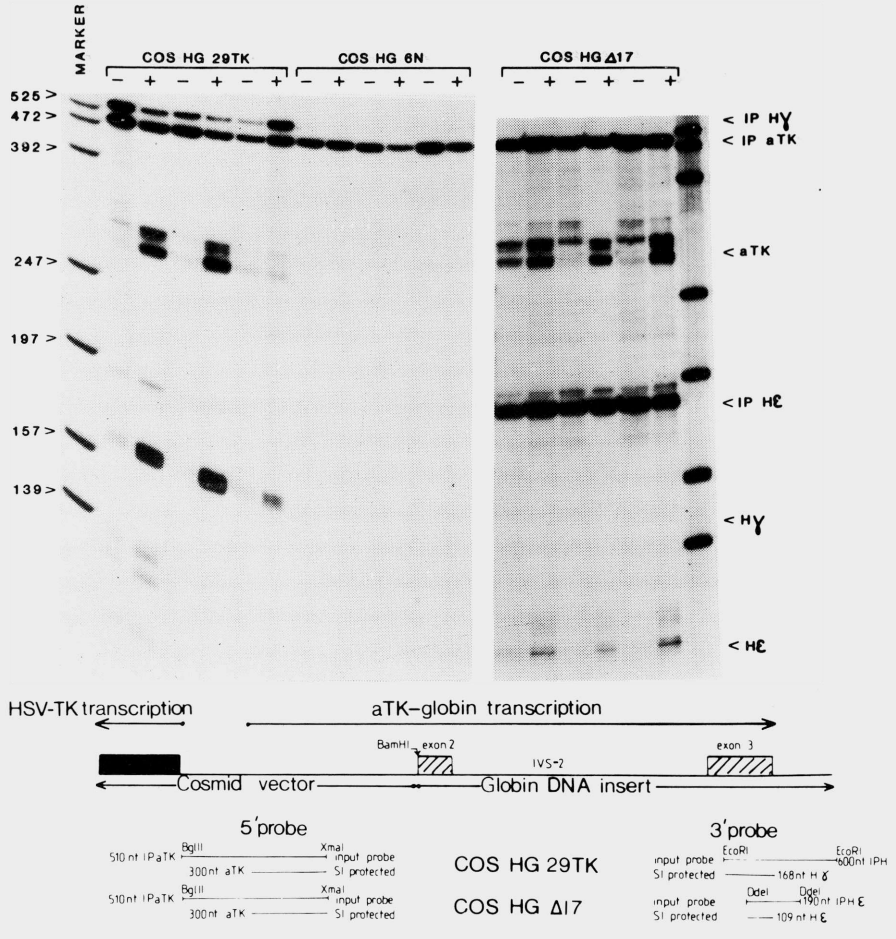


Fig. 3 S_1 nuclease analysis of anti-*tk* transcripts in MEL transformants. MEL transformants containing cosHG29TK, cosHG6N and cosHGΔ17 were grown for 3 days in the presence (+) or absence (−) of 3 mM HMBA. CosHG29TK and cosHGΔ17 contain the 3′ ends of the Aγ- and ε-globin genes adjoined to vector sequences near the HSV *tk* gene as shown. The 5′ ends of *a-tk* transcripts initiating near the *tk* gene were mapped with a 510 bp *Xma*I–*Bgl*II probe (IP aTK) using the procedure described for Fig. 2*a*. This probe covers the expected 5′ end of the *a-tk* transcript and gives an S_1 nuclease-protected fragment of 280–300 nucleotides (aTK). RNA from cosHG29TK and cosHGΔ17 transformants was simultaneously hybridized to 3′ γ-globin and 3′ ε-globin gene probes, respectively.

transcription of HSV DNA fragments[15]. ATA and CAAT box sequences required for transcription initiation are located at the appropriate distances 5′ to the *a-tk* transcript. When this promoter is not linked to known coding sequences (in CosHG25, CosHG28TK and CosHG6N), no transcripts originate from it (Fig. 3, centre panel), suggesting that either an RNA splice or transcription terminator are required to yield a stable transcript.

We have shown independent regulation in the expression of individual human ε-, γ- and β-globin genes when introduced into MEL cells as a gene cluster, in accordance with previous work by Willing *et al.*[4] and Pyati *et al.*[5] who used human chromosome 11–MEL cell hybrids. Expression of human adult, but not fetal or embryonic globin genes, is regulated during MEL cell differentiation. This, therefore, mimics the specific activation of adult-type globin genes during erythropoiesis in man. Regulated expression of the human β-globin gene is independent of the presence of the rest of the globin gene cluster and is observed when it is introduced with only 1.5 kb of 5′ flanking DNA (βpRT; Fig. 2*c*). We have not directly excluded the possibility that the increased globin mRNA levels found after MEL cell differentiation are due to changes in mRNA stability as opposed to changes in transcription rate, although strong arguments suggest that this is not the case. First, the human ε- and γ-globin mRNAs are constitutively produced in MEL cells and in some exceptional clones even β-globin mRNA is constitutive. This suggests that ε- and γ-globin mRNAs are stable in uninduced MEL cells and therefore makes it unlikely that β-globin mRNA is unstable. Note that γ- and β-globin mRNAs stably coexist in human erythroid cells, for example, in heterozygotes for hereditary persistence of fetal haemoglobin[16]. Moreover, the 3′ regions of the same ε- and γ-globin mRNAs become inducible when linked to the HSV promoter, which suggests that the induction is not related to the stability of the mRNA.

The level of induction of the introduced gene is about one order of magnitude lower than that observed for the endogenous mouse β-major globin gene. This lower level probably cannot be explained by the fact that the human β-globin gene and mouse MEL cells represent a heterologous system; Chao *et al.*[8] have recently shown that the transcription of a mouse–human hybrid gene is also inducible in this system. As the level of hybrid mouse–human mRNA derived from the mouse promoter is about the same as that derived from the human promoter (present results and ref. 8) a heterologous effect is unlikely. It is, therefore, probable that the lower level of induction of the foreign compared with the endogenous globin genes reflects differences in chromosomal location and chromatin structure of the introduced gene. Such differences would also explain the variable level of induction of the human β-globin gene between different transformants.

We were surprised to find an inducible transcript which initiated in the 5′ flanking region of the HSV *tk* gene. If this viral promoter is used naturally in the herpes life cycle, it is possible that it responds to a viral or cellular *trans*-acting factor similar to that present in induced, but not in uninduced MEL cells.

Received 13 April; accepted 25 July 1983.

1. Friend, C., Scher, W., Holland, J. G. & Sato, T. *Proc. natn. Acad. Sci. U.S.A.* **68,** 378–382 (1971).
2. Reuben, R., Wife, R., Breslow, R., Rifkind, R. & Marks, P. *Proc. natn. Acad. Sci. U.S.A.* **73,** 862–866 (1976).
3. Ross, J., Ikawa, Y. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69,** 3620–3623 (1972).
4. Willing, M. C., Niehus, A. W. & French Anderson, W. *Nature* **277,** 534–538 (1979).
5. Pyati, J., Kucherlapati, R. S. & Skoultchi, A. J. *Proc. natn. Acad. Sci. U.S.A.* **77,** 3435–3439 (1980).
6. Spandidos, D. & Paul, J. *EMBO J.* **1,** 15–20 (1982).
7. Flavell, R. A. *et al.* in *UCLA Symp. molec. cell. Biol.* **26,** 17–33 (1982).
8. Chao, M. V., Mellon, P., Charnay, P., Maniatis, T. & Axel, R. *Cell* **32,** 483–493 (1983).
9. Wigler, M. *et al. Cell* **16,** 777–785 (1979).
10. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
11. Berk, A. J. & Sharp, P. A. *Cell* **12,** 721–732 (1977).

12. Weaver, R. F. & Weissmann, C. *Nucleic Acids Res.* **6,** 1175–1192 (1979).
13. Seed, B. in *Genetic Engineering* Vol. 4 (eds Setlow, J. K. & Hollaender, A.) (Academic, New York, 1982).
14. McKnight, J. *Nucleic Acids Res.* **8,** 5949–5964 (1980).
15. Sullivan Read, G. & Summers, W. *Proc. natn. Acad. Sci. U.S.A.* **79,** 5212–5219 (1982).
16. Weatherall, D. J. & Clegg, J. A. *Cell* **29,** 7–9 (1982).
17. Grosveld, F. G., Dahl, H. H. M., deBoer, E. & Flavell, R. A. *Gene* **13,** 227–237 (1981).
18. Grosveld, F. G. *et al. Nucleic Acids Res.* **10,** 6715–6732 (1982).
19. Ish-Horowicz, D. & Burke, J. F. *Nucleic Acids Res.* **9,** 2989–2998 (1981).
20. Birnboim, H. C. & Doly, J. *Nucleic Acids Res.* **7,** 1513–1523 (1979).
21. Moschonas, N. *et al. Nucleic Acids Res.* **9,** 4391–4401 (1981).
22. Auffray, C. & Rougeon, F. *Eur. J. Biochem.* **107,** 303–314 (1980).

Printed in Great Britain by Macmillan Production Ltd., Basingstoke, Hampshire

Cell, Vol. 38, 265–273, August 1984, Copyright © 1984 by MIT 0092-8674/84/080265-09 $02.00/0

DNA Sequences Required for Regulated Expression of β-Globin Genes in Murine Erythroleukemia Cells

Stephanie Wright,* Arnon Rosenthal,† Richard Flavell,‡ and Frank Grosveld*
*Laboratory for Gene Structure and Expression
National Institute for Medical Research
The Ridgeway, Mill Hill
London NW7 1AA, England
†Department of Molecular Biology
The Hebrew University
Hadassah Medical School
Jerusalem, Israel
‡Biogen Research Corporation
14 Cambridge Center
Cambridge, Massachusetts 02142

Summary

We introduced into MEL cells rabbit β-globin gene deletion mutants and two sets of hybrid genes constructed from the inducible human β-globin gene and noninducible human γ-globin gene or the murine H-2K^{bm1} class I MHC gene. S1 nuclease analysis of gene transcripts before and after MEL differentiation showed that induction of the rabbit β-globin gene did not require more than 58 bp of DNA 5′ to the transcription initiation site. Hybrid genes were constructed with human β-globin DNA sequences from either 5′ or 3′ of the translation initiation site linked to the complementary parts of the γ or H2K^{bm1} genes. Both types of constructs were inducible during MEL differentiation. The relative rates of transcription of the 5′γ-3′β and 5′H2-3′β hybrid genes show that induction of the hybrid gene transcripts results at least in part from transcriptional activation of the genes. We suggest that DNA sequences that regulate β-globin gene transcription during MEL differentiation are located both 5′ and 3′ to the translation initiation site.

Introduction

Murine erythroleukemia (MEL) cells are Friend virus transformed erythroid cells that are arrested at the proerythroblast stage of differentiation (for review, see Marks and Rifkind, 1978). Cultured MEL cells may be induced to complete erythroid maturation by treatment with a variety of chemicals such as dimethylsulphoxide (DMSO) or hexamethylenebisacetamide (HMBA) (Friend et al., 1971; Reuben et al., 1976; Marks and Rifkind, 1978). This chemically triggered maturation closely resembles normal erythroid differentiation and results in an accumulation of mouse adult globin mRNA in the differentiated cell (Ross et al., 1972; Nudel et al., 1977; Orkin and Swerdlow, 1977). The accumulation of globin transcripts results from both transcriptional activation of the mouse globin genes and an increase in the relative stability of globin mRNA. One approach in studying the DNA sequences that regulate globin gene activation during erythroid differentiation is to analyze the expression of cloned globin genes introduced into MEL cells (Spandidos and Paul, 1982; Wright et al., 1983; Chao et al., 1983).

We have previously introduced genes from the human β-like globin gene cluster containing the embryonic (ε), fetal (γ), and adult (β) genes into MEL cells (Wright et al., 1983). Transcription of the foreign adult β-globin gene was induced following MEL differentiation, whereas expression of the fetal (γ) and embryonic (ε) genes was not regulated. This specific induction of the human adult β-globin gene therefore mimicked globin gene activation during adult type erythropoiesis and showed that the information required for regulation of the human β-globin gene resided within the introduced DNA sequences.

In the experiments described here, we have localized the DNA sequences that regulate the activation of an introduced β-globin gene in MEL cells. The results suggest that the DNA sequences required for regulation of β-globin gene expression during erythroid differentiation are located both 5′ and 3′ to the translation initiation codon.

Results

Regulated Expression of Rabbit β-Globin Gene Deletion Mutants in MEL Cells

In order to determine whether specific DNA sequences within the 5′ flanking region of the β-globin gene regulate their activation during erythroid differentiation, rabbit β-globin gene deletion mutants as shown in Figure 1 were introduced into MEL cells. We analyzed the expression of a rabbit β-globin gene with 425 bp of DNA 5′ to the transcription initiation site and three mutants containing truncated 5′ ends. The first of these had 100 bp of DNA 5′ to the transcription initiation site and thus contained all the DNA sequences required for transcription (Grosveld et al., 1982a, 1982b; Dierks et al., 1981). This −100 mutant therefore enabled us to determine whether the regulatory DNA sequences required for induction reside within or outside the promoter boundary. The other mutants had deletions within the conserved sequence boxes required for transcription per se; the ∇-76-93 mutant thus lacked the conserved −90 region, and the −58 mutant, which had only 58 bp 5′ to the transcription initiation site, lacked both the conserved −90 region and CAAT box. The rabbit β-globin gene mutants had been constructed in vectors which contained SV40 DNA sequences (Grosveld et al., 1982a, 1982b). These sequences were removed prior to transformation by digestion with Hind III. The resultant rabbit β-globin fragments were transfected into tk$^-$ MEL cells with a Pvu I linearized human β-globin cosmid, β-pRT. This cosmid contains the 4.7 kb human β-globin gene Bgl II fragment cloned into the Bam HI site of the HSV-tk cosmid pRT. The expression of the cotransferred human β-globin gene served as a control for induction of the introduced DNA since previous experiments have shown this human β-gene to be inducible in 80% of MEL transformants (Wright et al., 1983). Stable tk$^+$ transformants were obtained at frequencies of 5–20 clones per micro-

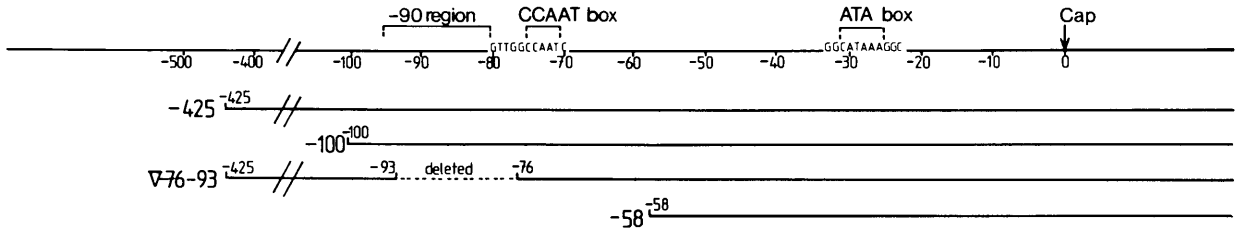


Figure 1. Structure of Rabbit β-Globin Gene Deletion Mutants

Deletions within the 5′ flanking region of the rabbit β-globin gene were as shown. The construction of these mutants was described in Grosveld et al., 1982.

gram β-pRT per 10^6 cells. Seven to ten transformants containing each of the mutants were induced to differentiate by culturing for three days in the presence of 3 mM HMBA, and levels of mouse β-major-, human β-, and rabbit β-globin transcripts were measured by S1 nuclease analysis using probes that detected the 5′ or 3′ ends of these mRNAs. The MEL cell transformants that showed inducibility for the human and mouse β-globin genes were subsequently analyzed for expression of the foreign rabbit β-globin gene by S1 nuclease mapping (Berk and Sharp, 1977; Weaver and Weissmann, 1979) (Figure 2). Hybridization of correctly terminated rabbit β-globin gene transcripts to the 900 nucleotide 3′ labeled Eco RI probe yielded a 172 nucleotide fragment which was protected from S1 nuclease digestion. Similarly, hybridization of correctly initiated rabbit β-globin gene transcripts to the 221 nucleotide 5′ labeled Bst NI probe yielded a 139 nucleotide S1 nuclease protected fragment. Figure 2 also shows the induction of endogenous mouse β-globin mRNA in the same MEL transformants as measured by S1 nuclease analysis of the 3′ end of the mRNA. All S1 nuclease analyses were carried out using an excess of DNA in the hybridization reaction, and the intensity of the S1 nuclease protected fragments was, therefore, proportional to the amount of globin mRNA present. Six out of six MEL transformants containing the −425 rabbit β-globin gene showed a 10- to more than 40-fold increase in rabbit β-globin mRNA levels upon differentiation. Induced mRNA levels corresponded to 500–3000 copies rabbit β-globin mRNA per cell. A similar inducibility was observed for the −100 rabbit β-globin gene mutant (Figure 2), thus indicating that DNA sequences 5′ to the promoter are not essential for activation of the gene. Four out of seven transformants containing the ∇-76-93 mutant and three out of six transformants containing the −58 mutant showed inducibility of correctly initiated rabbit β-globin gene transcripts. The absolute levels of rabbit β-globin mRNA varied extensively between different transformants containing the ∇-76-93 and −58 mutants (Figure 2, the first and second examples of both the ∇-76-93 and the −58 mutants). Consequently, a large number of transformants would have to be analyzed to obtain an accurate representative mRNA level for each mutant. Nevertheless, the average level from the small numbers of transformants analyzed is lower in the ∇-76-93 and −58 clones than in the −425 and −100 rabbit β-globin transformants. This is presumably due to the effect of these promoter deletions on the efficiency of transcription per se (Dierks et al., 1981; Grosveld et al., 1982a, 1982b). Thus removal of DNA sequences that are required for transcription per se resulted in a reduction in the efficiency of transcription, but did not affect gene activation. Deletion mutations which removed part of the conserved ATA box resulted in so great a reduction in transcription of the gene that evaluation of their inducibility in MEL cells was not feasible (not shown). These data therefore indicate that not more than 58 bp of DNA 5′ to the transcription initiation site is required for induction of the transfected rabbit β-globin gene in MEL cells.

Regulated Expression in MEL Cells of Gene Hybrids Containing a 5′ Human β-Globin Region

Expression of the rabbit β-globin gene deletion mutants in MEL transformants indicated that the DNA sequences required for induction reside either within the 3′ region of the promoter or downstream of the transcription initiation site. In order to distinguish between these possibilities, we constructed hybrid genes between the inducible human β-globin gene and either the noninducible human γ-globin gene or the noninducible mouse H-2K^{bm1} gene (Weiss et al., 1983).

We have previously shown that the human γ-globin gene is not induced in MEL cells when it is introduced as part of a large cosmid recombinant. This was confirmed by analyzing in MEL cells the expression of the human γ-globin clone γ-pRT, which contained the Aγ-globin 3 kb Hind III fragment inserted in the tk vector pRT. Only one out of 12 γ-pRT transformants showed inducibility of human γ-globin mRNA (Figure 3a).

In order to determine whether the mouse H-2K^{bm1} gene is inducible, pTME10, which contains the H-2K^{bm1} gene on a 10 kb Eco RI fragment, was introduced into MEL cells by cotransfection with the HSV-tk vector pRT. Levels of H-2K^{bm1} transcript in 12 MEL transformants were measured by S1 nuclease analysis using a 600 nucleotide probe that spanned the junction between the third exon and third intron of the H2-K^{b} gene. Because the H2-K^{b} and H2-K^{bm1} genes differ in the third exon (Weiss et al., 1983), hybridization of this probe to H-2K^{bm1} transcripts yields a 140 nucleotide fragment that is protected from digestion by S1 nuclease (Rosenthal et al., submitted). None of these 12 MEL transformants showed inducibility of H-2K^{bm1} transcripts (Figure 3b).

In order to determine whether human β-globin DNA

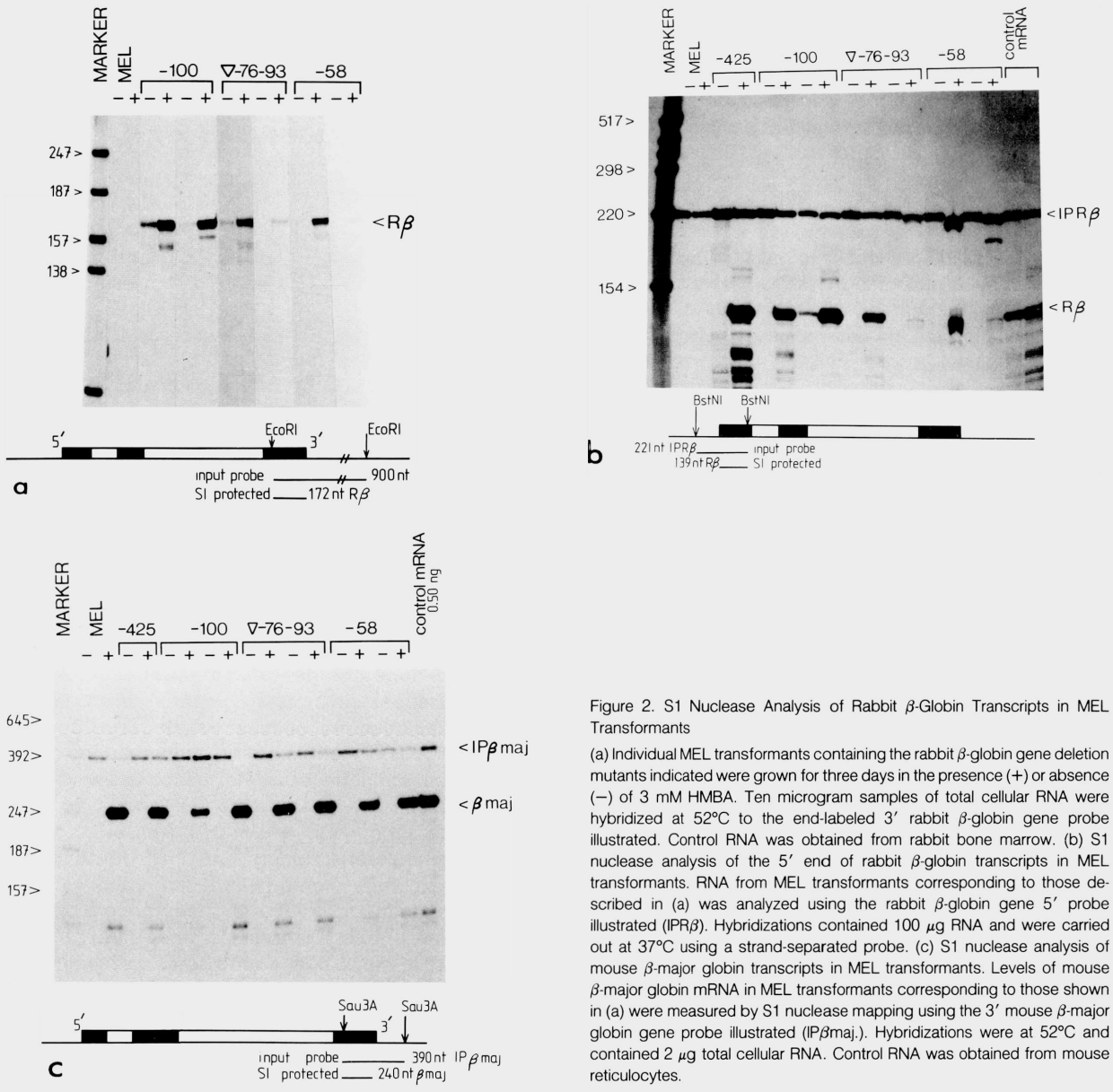


Figure 2. S1 Nuclease Analysis of Rabbit β-Globin Transcripts in MEL Transformants

(a) Individual MEL transformants containing the rabbit β-globin gene deletion mutants indicated were grown for three days in the presence (+) or absence (−) of 3 mM HMBA. Ten microgram samples of total cellular RNA were hybridized at 52°C to the end-labeled 3′ rabbit β-globin gene probe illustrated. Control RNA was obtained from rabbit bone marrow. (b) S1 nuclease analysis of the 5′ end of rabbit β-globin transcripts in MEL transformants. RNA from MEL transformants corresponding to those described in (a) was analyzed using the rabbit β-globin gene 5′ probe illustrated (IPRβ). Hybridizations contained 100 μg RNA and were carried out at 37°C using a strand-separated probe. (c) S1 nuclease analysis of mouse β-major globin transcripts in MEL transformants. Levels of mouse β-major globin mRNA in MEL transformants corresponding to those shown in (a) were measured by S1 nuclease mapping using the 3′ mouse β-major globin gene probe illustrated (IPβmaj.). Hybridizations were at 52°C and contained 2 μg total cellular RNA. Control RNA was obtained from mouse reticulocytes.

sequences 5′ to the translation initiation site can confer inducibility upon a noninducible gene, they were adjoined to the complementary 3′ region of the noninducible human γ-globin and mouse MHC H-2K^{bm1} genes. The construction of these hybrid genes is shown in Figure 4, using the Nco I site at the translation initiation site of the β- and γ-genes.

The hybrid gene containing the 5′ region of the human β-globin gene joined to the 3′ region of the γ-globin gene (5′β–3′γ) was introduced into tk$^-$ MEL cells. Ten transformants were induced into differentiation and transcripts from the introduced hybrid gene were quantitated by S1 nuclease analysis using a probe that detected the 3′ end of human γ-globin mRNA. Five transformants showed a 5- to 50-fold induction of the 5′β–3′γ hybrid transcript upon differentiation (Figure 5a).

In another set of transformations, the 5′β–3′H2 hybrid gene was introduced into tk$^-$ MEL cells by cotransfection with the HSV-tk vector pRT. Two out of three transformants containing this hybrid gene showed an induction of hybrid gene transcript as measured by S1 nuclease analysis using the H-2K^b gene exon–intron probe (Figure 5b).

Regulated Expression in MEL Cells of Gene Hybrids Containing the Human β-Globin Gene 3′ Region

We have also analyzed the expression in MEL cells of hybrid genes in which only DNA sequences 3′ to the translation initiation codon were derived from the human β-globin gene. The construction of these hybrid genes (5′-H-2K^{bm1}–3′β and 5′γ–3′β) is shown in Figure 6.

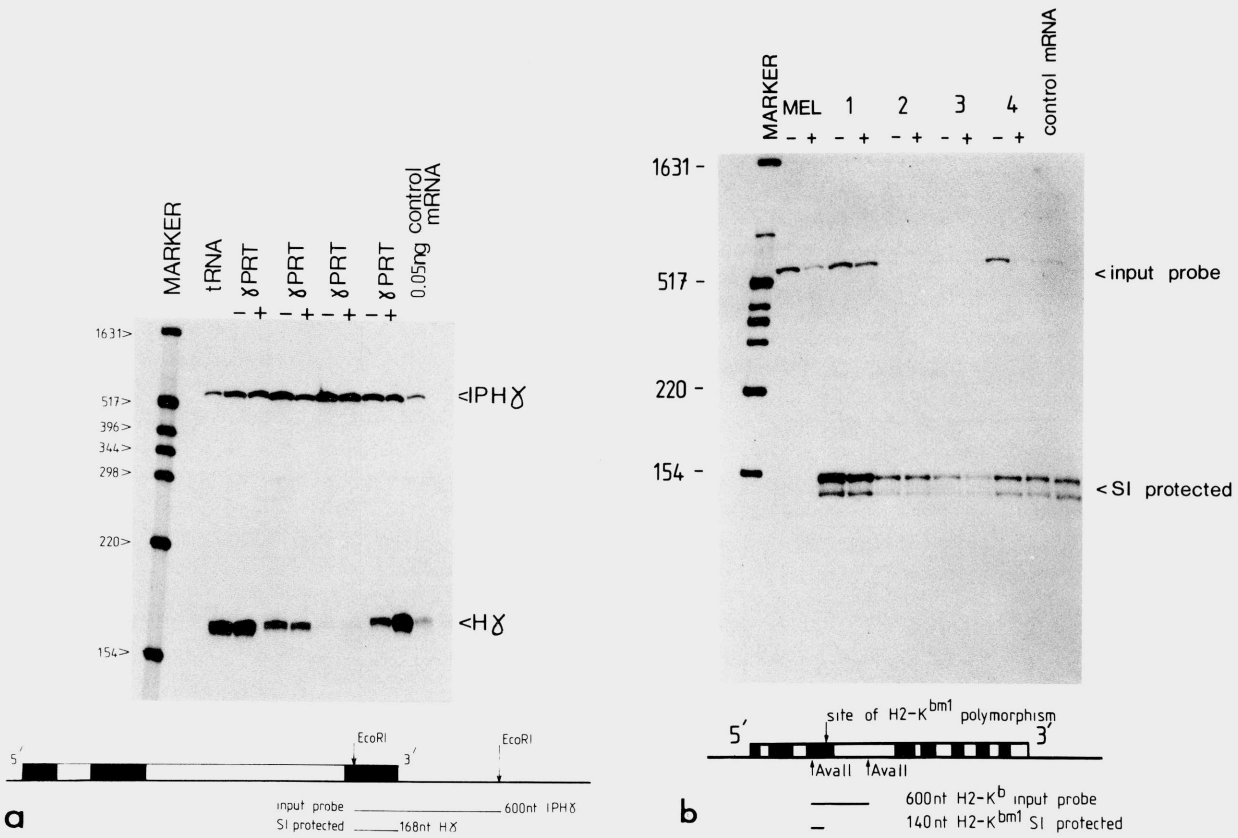


Figure 3. S1 Nuclease Analysis of Human γ-Globin Gene Transcripts in MEL Transformants Containing γ-pRT

(a) Thirty microgram samples of RNA from noninduced (−) and induced (+) MEL transformants containing γ-pRT were analyzed by S1 nuclease mapping using the 3′γ-globin gene probe illustrated (IPHγ). Hybridization was at 52°C. Control RNA was obtained from a patient with $\beta^0/\delta\beta^0$ thalassaemia. (b) S1 nuclease analysis of H-2K^{bm1} transcripts in MEL transformants containing the H-2K^{bm1} cosmid pTME10. pTME10 contained the 10 kb H-2K^{bm1} gene Eco RI fragment cloned in the Eco RI site of the pTM cosmid vector. Ten microgram samples of total cellular RNA from noninduced (−) and induced (+) MEL transformants containing pTME10 were analyzed by S1 nuclease mapping using the H-2K^b gene exon–intron probe illustrated. The DNA sequence of the H-2K^b and H-2K^{bm1} genes differ in the third exon at the position indicated. Hybridization of H-2K^{bm1} transcripts to the H-2K^b gene probe thus yields a 140 nucleotide fragment which is protected from S1 nuclease digestion. Hybridization was at 60°C and S1 digestion was for 1 hr at 20°C followed by 1 hr at 40°C. Control RNA was obtained from the liver of an H-2K^{bm1} mouse.

The 5′γ-3′β and 5′H2-3′β hybrid genes were introduced into tk$^-$ MEL cells by cotransfection with pRT. Upon differentiation, four out of eight transformants containing the 5′γ–3′β gene showed a 5- to more than 100-fold induction of hybrid gene transcript as measured by S1 nuclease analysis using a probe which detected the 3′ end of human β-globin mRNA (Figure 7a). Four out of five transformants containing the 5′H-2K^{bm1}-3′β hybrid gene showed a similar induction of hybrid gene transcript (Figure 7b).

The accumulation of mouse globin mRNA in differentiating MEL cells results from both an increase in the rate of globin gene transcription and an increase in the relative stability of globin mRNA. Similarly, the induction of introduced hybrid genes containing human β-globin structural gene sequences (5′γ-3′β and 5′H-2K^{bm1}-3′β) may be due to either an increase in the relative stability of the hybrid gene transcript or to an increase in the rate of transcription of these genes. In order to distinguish between these two possibilities, we measured the relative rate of transcription of these hybrid genes before and after differentiation using a nuclear run-off transcription assay (Groudine et al., 1981; Hofer et al., 1982).

Nuclei were isolated before and after differentiation from two MEL transformants containing the 5′H2-3′β and 5′γ-3′β hybrid genes. In these two transformants, the induction and expression level of the foreign hybrid gene was approximately equal to that of the endogenous mouse β-major globin gene, as measured by S1 nuclease. Isolated nuclei were incubated in a nuclear transcription reaction containing ^{32}P-UTP, which was therefore incorporated into newly synthesized RNA. RNA was extracted and hybridized to specific globin gene DNA probes that were immobilized as dots on nitrocellulose filters. The probes were isolated from the second IVS of the mouse and human β-globin genes for two reasons; first, maximal sequence divergence between the mouse and human genes is located in this region. Second, these probes hybridize only to precursor mRNA. DNA probes were present in vast excess over the specific globin transcripts and conse-

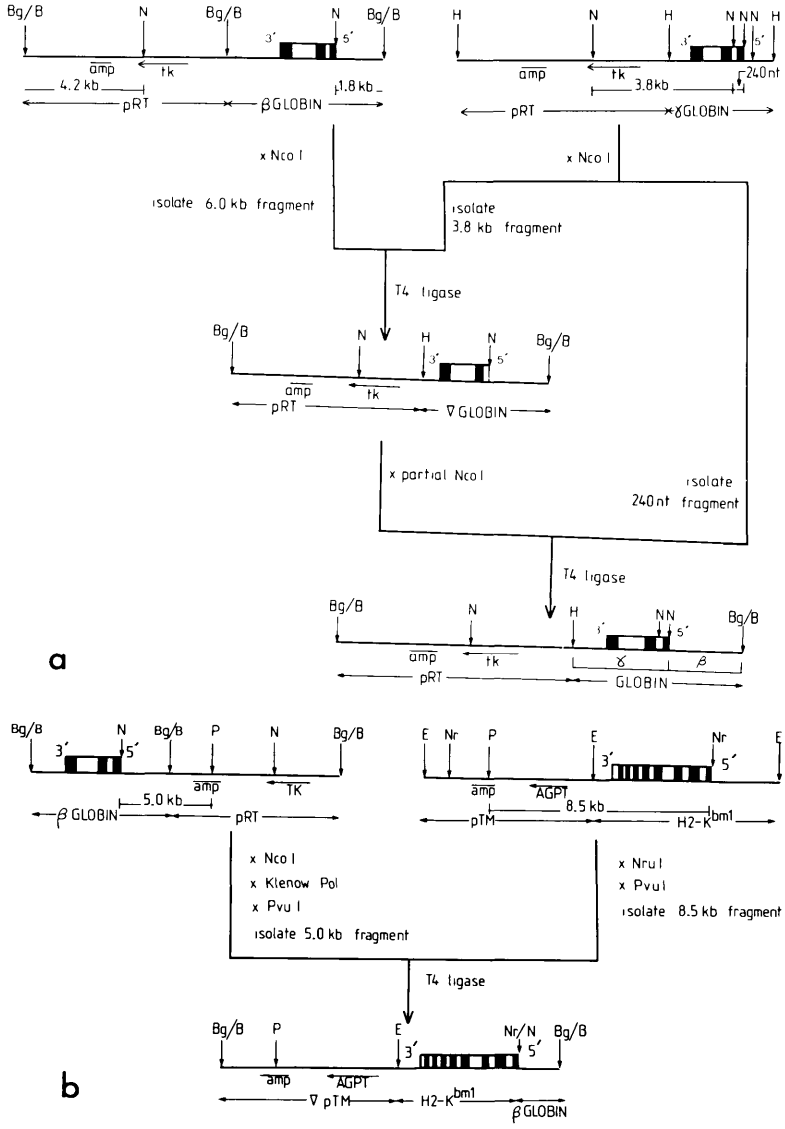


Figure 4. Construction of 5′β-3′γ and 5′β-3′H-2K^{bm1} Hybrid Genes

Hybrid genes were constructed from the globin and H-2K^{bm1} gene subclones as described in the diagrams.

N, P, E, Nr, and Bg/B refer to Nco I, Pvu I, Eco RI, Nru I, and joined Bgl II–Bam HI sites, respectively.

quently, the amount of hybridizing labeled nuclear RNA was proportional to the rate of transcription of the globin gene. All hybridizations contained RNA isolated from an equal number of nuclei, so that differences in the absolute rate of gene transcription could be measured (see Experimental Procedures).

Hybridization of labeled nuclear transcripts from non-transformed MEL nuclei before and after differentiation is shown in Figure 8. Differentiation of the MEL cells resulted in an increase in hybridization to the mouse IVS2 probe, indicating an increase in the rate of transcription of the mouse β-major globin gene. In contrast, no increase in hybridization to the human IVS2 probe was observed, showing that the labeled mouse globin mRNA did not cross-hybridize with the human β-globin gene probe. Labeled nuclear transcripts from MEL transformants containing the 5′γ-3′β and 5′H2-K^{bm1}-3′β hybrid genes showed an increase in hybridization to both mouse and human probes upon differentiation, indicating an increase in the rate of transcription of both the endogenous mouse globin and introduced hybrid globin genes. Thus the increase in the level of hybrid 5′γ-3′β and 5′H2-K^{bm1}-3′β gene transcripts following MEL differentiation resulted at least in part from an increase in the rate of transcription of these genes.

Discussion

By introducing into MEL cells a series of rabbit β-globin gene deletion mutants, we have shown that no more than 58 bp of DNA 5′ to the transcription initiation site are required for induction of the β-globin gene during erythroid differentiation. We also analyzed the expression in MEL cells of hybrid genes in which DNA sequences 5′ to the translation initiation codon were derived from the human β-globin gene (5′β-3′γ and 5′β-3′H-2K^{bm1}). Expression of these hybrid genes was induced during differentiation,

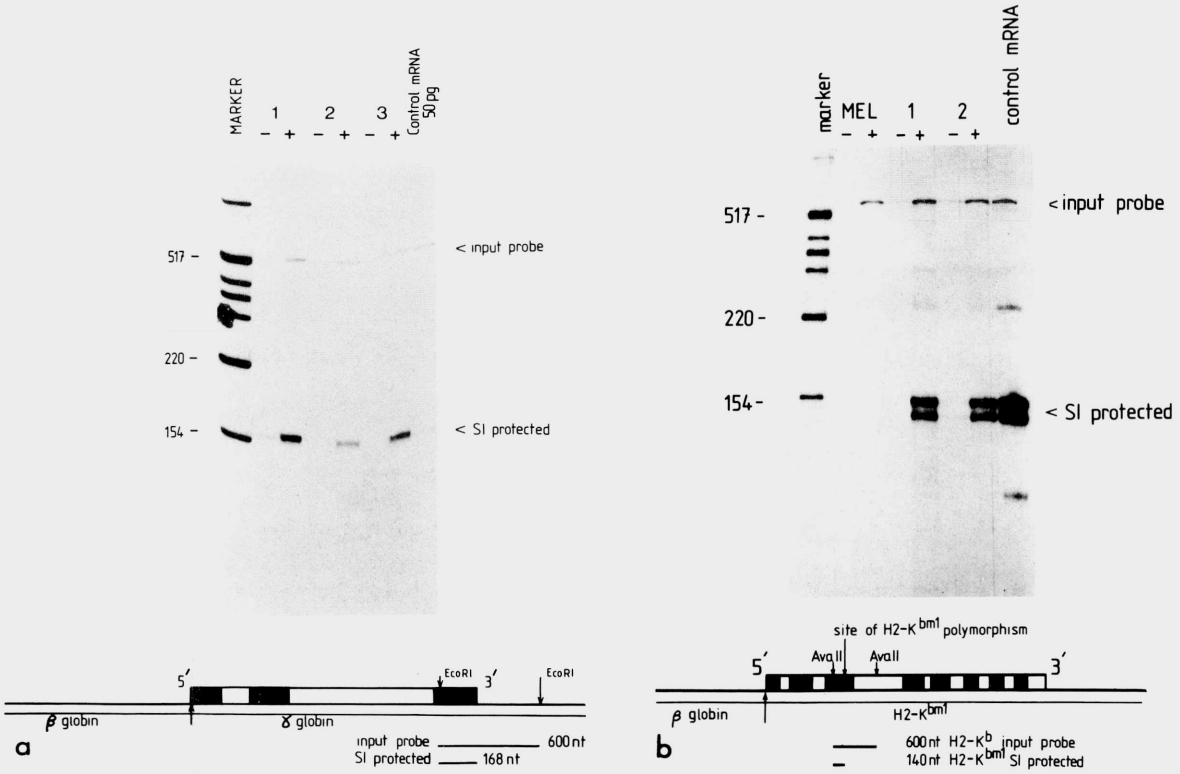


Figure 5. (a) S1 nuclease mapping of 5′β-3′γ hybrid gene transcripts in MEL transformants. Ten microgram samples of total cellular RNA from noninduced (−) and induced (+) MEL transformants were analyzed by S1 nuclease mapping using the 3′ human γ-globin gene probe illustrated. Hybridization was at 52°C. Control RNA was from a patient with $\beta^0/\delta\beta^0$ thalassaemia. (b) S1 nuclease mapping of 5′β-3′H-2K^{bm1} transcripts in MEL transformants. Eighty microgram samples of RNA from noninduced (−) and induced (+) MEL transformants were analyzed by S1 nuclease analysis using the H-2K^b gene probe as described in Figure 4.

showing that the human β-globin 5′ flanking DNA sequences can confer inducibility on the noninducible human γ-globin and mouse H-2K^{bm1} genes. In addition, we analyzed the expression of hybrid genes in which only DNA sequences 3′ to the translation initiation codon were derived from the human β-globin gene (5′γ-3′β and 5′H-2K^{bm1}-3′β). Interestingly, we observed an increased level of transcripts from these hybrid genes upon MEL differentiation. Measurement of the relative rate of transcription of these hybrid genes by nuclear transcription run-off analysis showed that these increased levels of hybrid transcript resulted at least in part from transcriptional activation of the genes upon differentiation. We therefore suggest that DNA sequences required for activation of the β-globin gene during erythropoiesis are found both 5′ and 3′ to the translation initiation codon.

Several cloned genes have been shown to be expressed in a regulated manner after introduction into an appropriate eucaryotic cell line e. g., the growth hormone (Robins et al., 1982) and α-2 globulin (Kurtz, 1981) synthesis are inducible by glucocorticoids after the introduction of these genes into L cells that contain receptors for this hormone, and the Drosophila heat shock genes are inducible by high temperature after introduction into a variety of cell types (Pelham and Bienz, 1982). Approaches to localize the regulatory DNA sequences that confer inducibility on the foreign gene have involved studying the expression of in vitro constructed deletion mutants and hybrid genes. In all cases studied thus far, the foreign gene is expressed in a regulated manner after transfer into cells as a short DNA fragment containing only several hundred base pairs of DNA 5′ to the gene. In some systems, most notably in the induction of the interferon (Ragg and Weissmann, 1983; Weidle and Weissmann, 1983; Zinn et al., 1983) and heat shock genes (Pelham and Bienz, 1982), a distinct region within the 5′ flanking region of the gene has been shown to confer inducibility on an otherwise noninducible gene. In other cases, however, the presence of regulatory DNA sequences 3′ to the cap site of the gene has not always been rigorously investigated. There are several well-defined examples of gene transcription being influenced by DNA sequences located 3′ to the cap site. These include the internal promoter region of the Xenopus 5S RNA gene (Bogenhagen et al., 1980), which is required for transcription per se, and the tissue-specific enhancer of the immunoglobulin gene (Queen and Baltimore, 1983; Banerji et al., 1983), which is located in an intron upstream of the constant region gene and which probably serves to activate the variable region promoter of the functionally rearranged immunoglobulin gene in the lymphocyte. In addition, the expression of a foreign gene in a eucaryotic cell may be increased by a viral enhancer sequence when the

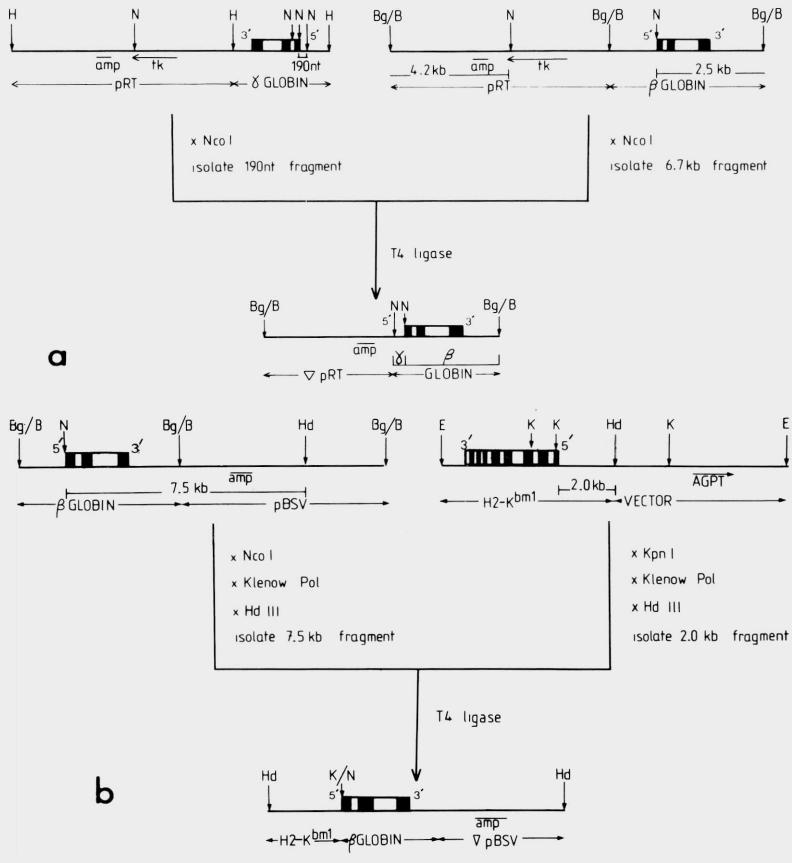


Figure 6. Construction of 5′γ-3′β and 5′H-2K^{bm1}-3′β Hybrid Genes

Hybrid genes were constructed from globin and H-2K^{bm1} gene subclones as described in the diagram. N, P, E, K, and Bgl/B refer to Nco I, Pvu I, Eco RI, Kpn I, and joined Bgl II–Bam HI sites, respectively.

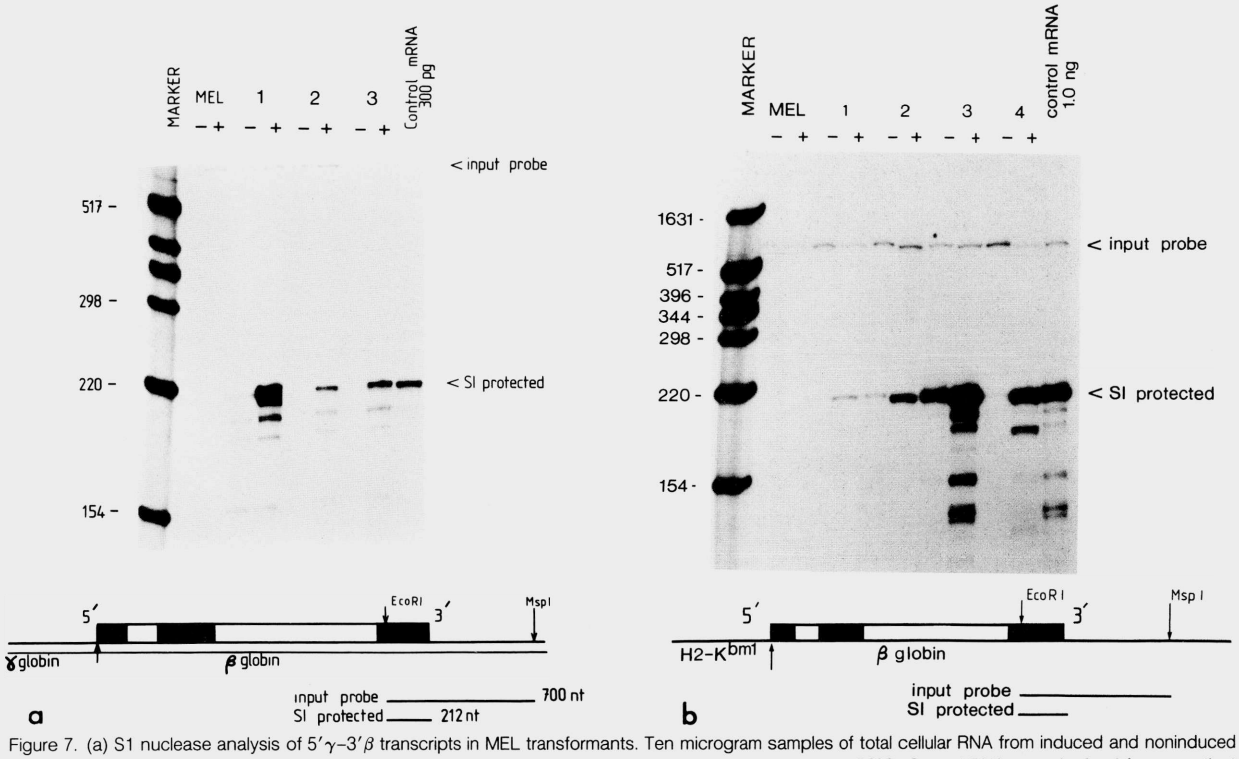


Figure 7. (a) S1 nuclease analysis of 5′γ–3′β transcripts in MEL transformants. Ten microgram samples of total cellular RNA from induced and noninduced (−) MEL transformants were analyzed using the 3′ human β-globin probe illustrated. Hybridization was at 52°C. Control RNA was obtained from a patient with sickle cell anemia. (b) S1 nuclease analysis of 5′H2K^{bm1}-3′β transcripts in MEL transformants. Ten microgram samples of total cellular RNA from MEL transformants containing the 5′H2K^{bm1}-3′β hybrid gene were analyzed as in (a).

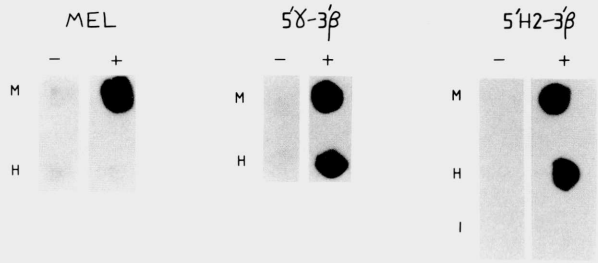


Figure 8. Dot Hybridization of Labeled Nuclear Transcripts from MEL Transformants

Nuclei were isolated from induced (+) and noninduced (−) individual MEL 5′γ-3′β (Figure 7a, clone 1) and 5′H2-3′β (Figure 7b, clone 4) transformants. RNA synthesized in these isolated nuclei was labeled for 20 min with ^{32}P-UTP as described in Experimental Procedures. Nuclear transcripts were hybridized to human β-globin (H), mouse β-major globin (M), and human insulin (I) gene probes as indicated. The human β-globin probe was the Eco RI–Bam HI IVS-2 fragment. The mouse β-major globin probe was the Hind III–Pst I IVS-2 fragment. The human insulin gene probe was a 1.8 kb Bam HI fragment described by Bell et al., 1982.

letter is placed either 5′ or 3′ to the gene (Banerji et al., 1981).

The activation of a variety of glucocorticoid-responsive genes is mediated by the binding of the hormone–receptor complex to the gene. Specific binding of the glucocorticoid receptor to MTV DNA in vitro occurs at sites both within and upstream of the transcribed LTR region (Payvar et al., 1983). This implies that DNA sequences both within and upstream of the gene may regulate its activation. The results described in this paper may similarly indicate that DNA sequences both within and upstream of the globin gene may play a role in their activation during erythroid differentiation. Whether the activation is accomplished by the binding of a positive stimulatory factor(s) or the removal of a negative repressive factor(s) supplied in *trans* by the MEL cell during differentiation is not clear from these results. The low levels of expression obtained with the 5′ β-globin constructs, 5′β-3′H2 (Figure 7) suggest that the β-globin promoter might be very weak or repressed in undifferentiated MEL cells. Although these results cannot be quantitated reliably, it is tempting to suggest that a negative regulation would operate on the 5′ end and a positive regulation on the 3′ end of the β-globin gene. The induction process would only operate after the β-globin chromatin had been activated during erythroid differentiation in vivo (Groudine et al., 1981; Groudine et al., 1983). In one form of γβ-thalassaemia (van der Ploeg et al., 1980; Kioussis et al., 1983), a large deletion has positioned an unknown locus next to the β-globin gene resulting in an inactive chromatin structure and a silencing of the gene. When this β-globin gene is cloned and introduced into MEL cells, it is expressed and induced normally (Wright, unpublished observation). Because this abnormal β-globin locus can respond normally to MEL cell differentiation, and because the heterozygous patient also contains a normal and expressed β-globin allele, it suggests that factors that act in *trans* could only act in vivo after the activation of globin chromatin structure.

Experimental Procedures

Transformation and Induction of MEL Cells

tk$^-$ MEL cells (Spandidos and Paul, 1982) were maintained in α-MEM medium supplemented with 10% FCS. Cells were transformed as described (Wigler et al., 1979) by adding 1 μg tk plasmid (pRT1 or βpRT) plus 20 μg nonselectable DNA (rabbit β-globin mutant or hybrid gene) per 10^6 cells. tk$^+$ transformants were selected in α-MEM supplemented with 10% FCS, 15 μg/ml hypoxanthine, 0.5 μg/ml aminopterin, 5 μg/ml thymidine.

MEL transformants were induced into differentiation by culturing for 3 days in the presence of 3 mM HMBA.

Purification of RNA from MEL Transformants

10^7 cells were lysed with 1-ml 6 M guanidinium chloride and phenol extracted. Nucleic acid was ethanol precipitated, dissolved in 10 mM Tris (pH 7.5), 10 mM $MgCl_2$, 50 mM NaCl, and treated with RNAase-free DNase. RNA was finally phenol extracted and ethanol precipitated.

S1 Nuclease Analysis

Levels of globin mRNA in MEL transformants were measured by S1 nuclease analysis (Berk and Sharp, 1977; Weaver and Weissmann, 1979) using end-labeled DNA probes. Probes were end labeled by T4 polynucleotide kinase or reverse transcriptase to a specific activity of $>10^7$ cpm/μg. Labeled probe was hybridized to 10–30 μg samples of total cellular RNA in 10 μl 40 mM PIPES (pH 6.4), 400 mM NaCl, 1 mM EDTA, and 80% recrystallized formamide for 8–12 hr. Hybridization temperatures were as indicated in the figure legends. After hybridization, samples were digested for 2 hr with 3000 U S1 nuclease in 300 μl 200 mM NaCl, 30 mM NaAc (pH 4.8), 2 mM $ZnSO_4$; digestion was at 20°C unless otherwise stated. In some instances this resulted in an overdigestion of the RNA–DNA hybrids (e. g., Figure 7b). S1 protected DNA was ethanol precipitated and electrophoresed on a 7 M urea 7% acrylamide gel. The amount of protected probe was quantitated by scanning the autoradiographs.

Transcription in Isolated Nuclei

Transcription in isolated nuclei from differentiated and nondifferentiated MEL transformants was carried out as described in Groudine et al. (1984). Each transcription reaction contained 5×10^7 nuclei and 300 μCi ^{32}P-UTP (3000 Ci/mmole).

Purified DNA fragments were bound to a 0.45 μm nitrocellulose filter as described by Kafatos et al. (1979). Two micrograms of fragment was loaded per dot. Individual dots were hybridized to labeled nuclear transcripts in a 400 μl volume as described in Groudine et al. (1984). Each hybridization contained nuclear RNA from 10^7 cells, which corresponded to $5–10 \times 10^7$ cpm for the uninduced cells and $2–5 \times 10^7$ cpm for the induced cells. After hybridization, the filters were washed to a stringency of 0.3× SSC.

Acknowledgments

We are grateful to Cora O'Carroll for typing this manuscript. A. R. was supported by a fellowship from the "Friends of the Hebrew University." This work was supported by the Medical Research Council.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 8, 1984; revised June 6, 1984

References

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.

Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell *33*, 729–740.

Bell, G. I., Karam, J. H., and Rutter, W. J. (1981). Polymorphic DNA region adjacent to the 5′ end of the human insulin gene. Proc. Nat. Acad. Sci. USA *78*, 5759–5763.

Berk, A. J., and Sharp, P. A. (1977). Sizing and mapping of early adenovirus

mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell *12,* 721–732.

Bogenhagen, D. F., Sakonju, S., and Brown, D. D. (1980). A control region in the center of the 5S RNA gene directs specific initiation of transcription II. The 3′ border of the region. Cell *19,* 27–35.

Chao, M. V., Mellon, P., Charnay, P., Maniatis, T., and Axel, R. (1983). The regulated expression of β-globin genes introduced into mouse erythroleukemia cells. Cell *32,* 483–493.

Dierks, P., van Ooyen, A., Mantei, N., and Weissmann, C. (1981). DNA sequences preceding the rabbit β-globin gene are required for formation in mouse L cells of β-globin RNA with correct 5′ terminus. Proc. Nat. Acad. Sci. USA *78,* 1411–1415.

Friend, C., Scher, W., Holland, J. G., and Sato, T. (1971). Hemoglobin synthesis in murine virus induced leukemic cells *in vitro*: stimulation of erythroid differentiation by dimethyl sulfoxide. Proc. Nat. Acad. Sci. USA *68,* 378–382.

Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M., and Flavell, R. A. (1982). The construction of cosmid libraries which can be used to transform eukaryotic cells. Nucl. Acids Res. *10,* 6715–6732.

Grosveld, G. C., deBoer, E., Shewmaker, C. K., and Flavell, R. A. (1982a). DNA sequences necessary for the transcription of the rabbit β-globin gene *in vivo*. Nature *295,* 120–126.

Grosveld, G. C., Rosenthal, A., and Flavell, R. A. (1982b). Sequence requirements for the transcription of the rabbit β-globin gene *in vivo*: the −80 region. Nucl. Acids Res. *10,* 4951–4971.

Groudine, M., and Casimir, C. (1984). Post-transcriptional regulation of the chicken thymidine kinase gene. Nucl. Acids Res. *12,* 1427–1446.

Groudine, M., Peretz, M., and Weintraub, H. (1981). Transcriptional regulation of hemoglobin switching in chicken embryos. Mol. Cell. Biol. *1,* 281–288.

Hofer, E., Hofer-Warbinek, R., and Darnell, J. E., Jr. (1982). Globin RNA transcription: a possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. Cell *29,* 887–893.

Kafatos, F. C., Jones, C. W., and Efstratiadis, A. (1979). Determination of nuleic acid sequence homologies and relative concentrations by a dot hybridization procedure. Nucl. Acids Res. *7,* 1541–1551.

Kioussis, D., Vanin, E., deLange, T., Flavell, R., and Grosveld, F. G. (1983). β-globin gene inactivation by DNA translocation in γβ-thalassaemia. Nature *38,* 662–666.

Kurts, D. T. (1981). Hormonal inducibility of rat $\alpha_{2\mu}$ globulin genes in transfected mouse cells. Nature *291,* 629–631.

Marks, P. A., and Rifkind, R. A. (1978). Erythroleukemic differentiation. Ann. Rev. Biochem. *47,* 419–426.

Nudel, U., Salmon, J., Fibach, E., Terada, M., Rifkind, R., Marks, P. A., and Bank, A. (1977). Accumulation of α- and β-globin messenger RNAs in mouse erythroleukemia cells. Cell *12,* 463–469.

Orkin, S. H., and Swerdlow, P. S. (1977). Globin RNA synthesis *in vitro* by isolated erythroleukemia cell nuclei: direct evidence for increased transcription during erythroid differentiation. Proc. Nat. Acad. Sci. USA *74,* 2475–2479.

Payvar, F., DeFranco, D., Firestone, G. L., Edgar, B., Wrange, O., Okret, S., Gustafsson, J.-A., and Yamamoto, K. R. (1983). Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. Cell *35,* 381–392.

Pelham, H. R. B., and Bienz, M. (1982). A synthetic heat shock promoter element confers heat-inducibility on the herpes simplex virus thymidine kinase gene. EMBO J. *1,* 1473.

Queen, C., and Baltimore, D. (1983). Immunoglobulin gene transcription is activated by downstream sequence elements. Cell *33,* 741–748.

Ragg, H., and Weissmann, C. (1983). Not more than 117 base pairs of 5′-flanking sequence are required for inducible expression of a human IFN-α gene. Nature *303,* 439–442.

Reuben, R. C., Wife, R. L., Breslow, R., Rifkind, R. A., and Marks, P. A. (1976). A new group of potent inducers of differentiation in murine erythroleukemia cells. Proc. Nat. Acad. Sci. USA *73,* 862–866.

Robins, D. M., Paek, I., Seeburg, P. H., and Axel, R. (1982). Regulated expression of human growth hormone genes in mouse cells. Cell *29,* 623–631.

Ross, J., Itawa, Y., and Leder, P. (1972). Globin messenger RNA induction during erythroid differentiation of cultured leukemia cells. Proc. Nat. Acad. Sci. USA *69,* 3620–3623.

Spandidos, D. A., and Paul, J. (1982). Transfer of human globin genes to erythroleukemic mouse cells. EMBO J. *1,* 15–20.

van der Ploeg, L. H. T., Konings, A., Oort, M., Roos, D., Bernini, L. F., and Flavell, R. A. (1980). γ-β-thalassaemia: deletion of the γ- and δ-genes influences β-globin gene expression in man. Nature *283,* 637–642.

Weaver, R. F., and Weissmann, C. (1979). Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β-globin mRNA precursor and 10S β-globin mRNA have identical map coordinates. Nucl. Acids Res. *6,* 1175–1192.

Weidle, U., and Weissmann, C. (1983). The 5′-flanking region of a human IFN-α gene mediates viral induction of transcription. Nature *303,* 442–446.

Weiss, E. H., Mellor, A., Golden, L., Fahrner, K., Simpson, E., Hurst, J., and Flavell, R. A. (1983). Structure of a mutant H-2 gene: generation of polymorphism in H-2 genes by gene conversion-like events. Nature *301,* 671–674.

Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979). Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell *16,* 777–785.

Wright, S., deBoer, E., Grosveld, F. G., and Flavell, R. A. (1983). Regulated expression of the human β-globin gene family in murine erythroleukemic cell hybrids. Nature *305,* 333–336.

Zinn, K., DiMaio, D., and Maniatis, T. (1983). Identification of two distinct regulatory regions adjacent to the human β-interferon gene. Cell *34,* 865–879.

Reprinted from Nature, Vol. 310, No. 5976, pp. 415-418, 2 August 1984

Regulated expression of an introduced MHC *H–2K*^{bm1} gene in murine embryonal carcinoma cells

Arnon Rosenthal*†, Stephanie Wright*, Howard Cedar†, Richard Flavell*‡ & Frank Grosveld*

* Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK
† Department of Molecular Biology, The Hebrew University, Hadassah Medical School, Jerusalem, Israel
‡ Biogen Research Corporation, 14 Cambridge Center, Cambridge, Massachusetts 02142, USA

The transplantation antigens H–2K, H–2D and H–2L are developmentally regulated[1–3], highly polymorphic[4] cell surface proteins encoded by the major histocompatibility gene complex (MHC)[5–7]. First detectable on the early embryo[2,3], they are subsequently expressed on most somatic cells of the adult mouse in association with the protein β_2-microglobulin (β_2M; ref. 5). Cultured F9 embryonal carcinoma (EC) cells can be induced to differentiate along alternative pathways to form either parietal[8] or visceral[9] extra-embryonic endoderm, each concomitant with a change in morphology and pattern of gene expression. Previous reports have demonstrated an increased level of transplantation antigens in differentiated F9 EC cells[10–14], but the cell types expressing them were not defined. Here we show that the level of MHC H–2K^b and β_2M transcripts is increased during both pathways of this differentiation. Expression of a foreign MHC *H–2K*^{bm1} gene was found to be regulated in a similar manner when the gene was introduced into EC cells. In contrast, an introduced rabbit β-globin gene was not regulated but expressed constitutively.

F9 EC cells were differentiated into parietal endoderm by culturing as a monolayer with 10^{-7} M retinoic acid for 14 days[8], or with 10^{-7} M retinoic acid plus 10^{-4} M N^6,O^2-dibutyryladenosine 3′,5′-cyclic monophosphate (dbcAMP) plus 10^{-4} μM 3-isobutylmethylxanthine for 5 days, conditions in which 90% of the cells are differentiated[15,16]. For differentiation to visceral endoderm, cells were grown for 6 days as small floating aggregates in bacteriological Petri dishes with 5×10^{-8} M retinoic acid[9]. MHC H–2K^b and β_2M mRNA transcripts were quantitated by S_1 nuclease analysis using an exon–intron and a cDNA probe respectively, which can detect ~10 pg of specific mRNA after densitometric scanning.

Using a 161-nucleotide *Hpa*II fragment spanning part of the β_2M cDNA and adjacent plasmid sequences as a probe for S_1 nuclease protection, no β_2M transcripts were detected in undifferentiated F9 cells (Fig. 1). Differentiation to parietal or visceral endoderm resulted in at least a 20-fold increase in the level of β_2M transcripts, as demonstrated by the presence of 110- and 80-nucleotide fragments resistant to digestion by S_1 nuclease (Fig. 1). This is consistent with the high level of this transcript observed in visceral and parietal endoderm isolated from the extra-embryonic tissues of a 13.5-day-old mouse. The shorter than full-length 80-nucleotide fragment may be due to either a small difference between the cDNA probe and the transcript or cleavage by S_1 nuclease of an A+T-rich region located 85 base pairs (bp) from the labelled *Hpa*II site[17].

S_1 nuclease analysis using a 600-nucleotide *Ava*II exon III–intron probe (Fig. 2*b*) was used to determine the level of H–2K^b mRNA. As judged by the presence of a 230-nucleotide protected fragment, this transcript is produced in very small quantities in undifferentiated F9 cells. After differentiation either to visceral endoderm using retinoic acid or to parietal endoderm using retinoic acid alone or in combination with dbcAMP (Fig. 2*b*), the level of H–2K mRNA increased 5–10-fold as determined by densitometer scanning.

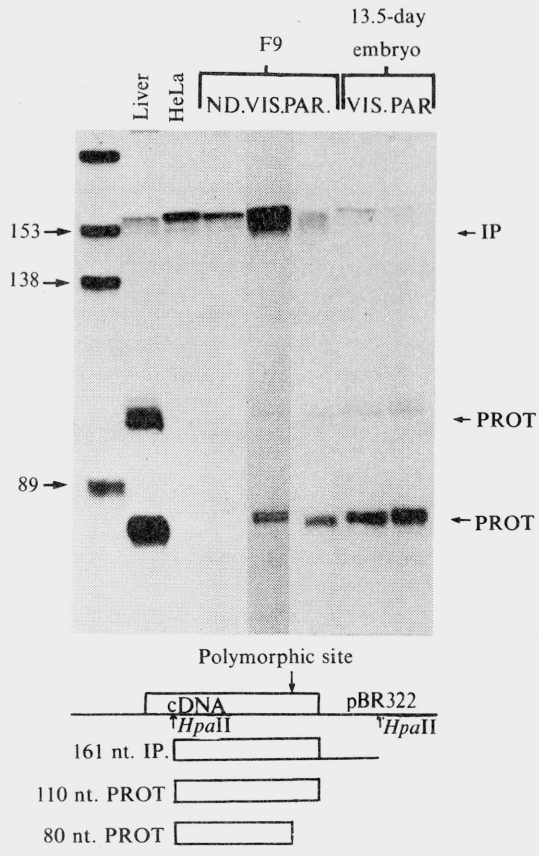


Fig. 1 S_1 nuclease analysis of β_2M mRNA in F9 EC cells before (ND) and after differentiation to parietal (PAR) or visceral (VIS) endoderm. Control mRNA was obtained from visceral and parietal extra-embryonic endoderm of a 13.5-day-old mouse embryo, and from adult mouse liver. Lane 1 contains a radioacivity labelled ΦX × *Rsa* marker.
Methods: A cDNA probe for β_2M (IP) (illustrated at bottom of figure) was end-labelled by reverse transcription, strand-separated and hybridized with total RNA at 37 °C for 12 h in 10 μl of 80% formamide, 40 mM PIPES *p*H 6.4, 1 mM EDTA, 400 mM NaCl. The mixture was digested with 3,000 U of S_1 nuclease (Boehringer) in 300 μl of 300 mM NaAc *p*H 4.8, 200 mM NaCl, 2 mM $ZnSO_4$ at 20 °C for 2 h. S_1-protected DNA (Prot.) was ethanol-precipitated, denatured and electrophoresed on a 7M urea/7% acrylamide gel. F9 lanes: ND, 15 μg of undifferentiated F9 cell RNA; VIS, 3 μg of visceral endoderm F9 cell RNA (overexposed fivefold); PAR, 15 μg of parietal endoderm F9 cell RNA. Embryo lanes: 5 μg of visceral or parietal endoderm (from B. Hogan).

Analysis of RNA samples from other undifferentiated and differentiated F9 cells (given by B. Hogan)[9,15,16], showed levels of H–2K^b mRNAs identical to ours (not shown). Low levels of H–2K^b mRNA in undifferentiated EC cells have been reported previously[18,19] and resemble the results described for laminin. Laminin mRNA is also transcribed at low levels in undifferentiated F9 cells and its transcription is increased following differentiation[15]. The fact that β_2M RNA is not detectable in undifferentiated F9 cells may indicate that the *H–2K* and β_2M genes are regulated differently. Note that despite the low levels of H–2K^b RNA synthesis, H–2K^b antigen might not be detected[10,20] on the cell surface in the absence of β_2M.

Experiments using cell fusion or chromosome transfer between somatic cells and EC cells have shown that the somatic *H–2K* gene is either repressed[21] or remains active in the hybrid cells[19,22,23]. As these experiments were inconclusive, we have used DNA-mediated gene transfer to study the control of H–2K expression.

A 10-kilobase (kb) fragment containing a variant *H–2K*^{b} gene called *H–2K*^{bm1} was introduced into F9 cells[24]. *H–2K*^{bm1} was

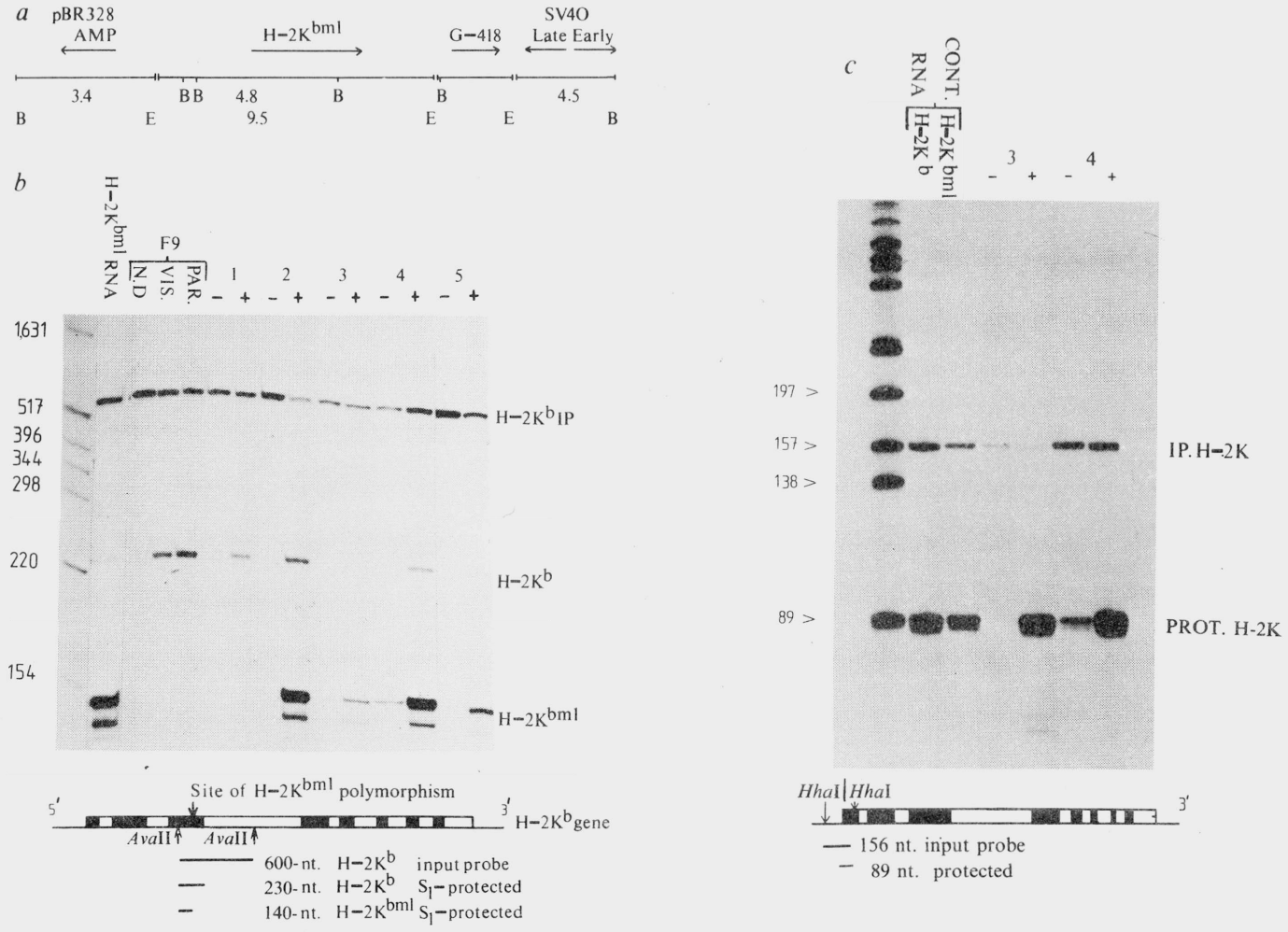


Fig. 2 *a*, Structure of the plasmid introduced into F9 cells. Transcripts are illustrated by arrows in the 5′ to 3′ direction. E and B, *Eco*RI and *Bam*HI sites, respectively. *b*, S_1 nuclease analysis of H–2K^b and H–2K^{bm1}, RNA from untransformed (F9) and from *H–2K^{bml}*-transformed (1–5 EC cells) before (−) and after (+) differentiation. Lane labelled H–2K^{bm1} RNA, 10 μg of RNA from a transient expression of the *H–2K^{bml}* gene in HeLa cells[31]. F9 lanes: 10 μg of RNA as in Fig. 1. Lanes 1–5, 10 μg of transformed F9 cell RNA containing, respectively, 1–3, 5–8, 3–5, 10–12 and 3–5 integrated copies of the *H–2K^{bml}* gene. *c*, S_1 nuclease analysis of the 5′ end of H–2K mRNA in two F9 transformants, before (−) and after (+) differentiation.

Methods: *b*, A 600-nucleotide intron–exon probe that allowed a distinction to be made between H–2K^{bm1} and H–2K^b transcripts (see *a*) was end-labelled with reverse transcriptase, strand-separated and hybridized as in Fig. 1, except that S_1 digestion was for 1 h at 20 °C followed by 1 h at 40 °C. Lane 4 contains a pBR322 × *Hin*f marker. *c*, The 5′ H–2K^b probe was end-labelled with T4 polynucleotide kinase, strand-separated and hybridized as described in Fig. 1.

covalently linked to Simian virus 40 (SV40) DNA and to the transposon Tn*5*-derived aminoglycosyl phosphotransferase II (AGPT) as a selective marker (Fig. 2*a*). Using the calcium phosphate transformation method[25], 1–10 clones per 5×10^5 cells per μg DNA were obtained. Southern blotting showed that 12 transformants contained 1–15 copies of the introduced DNA per cell (data not shown).

Taking advantage of the fact that the introduced *H–2K^{bml}* gene and the endogenous *H–2K^b* gene differ in the third exon sequence[24], we prepared a 600-bp *H–2K^b Ava*II probe from this region which allowed us to distinguish between the transcripts from the two genes by S_1 digestion analysis (Fig. 2*b*). Because partial S_1 nuclease digestion of the *H–2K^b Ava*II probe protected by *H–2K^{bml}* mRNA yields the same fragment as that protected by the *H–2K^b* transcript, we confirmed the results using the corresponding probe isolated from the *H–2K^{bml}* gene (not shown).

Out of eight clones that contained intact copies of the *H–2K^{bml}* introduced DNA, seven clones showed a 6–15-fold increase in the level of *H2–K^{bml}* mRNA on differentiation to parietal endoderm (Fig. 2*b*, clones 2–5). Only one clone expressed the gene in a constitutive manner (not shown). Four clones were shown to contain an incomplete or rearranged copy of the introduced DNA and on differentiation, only the endogeneous *H–2K^b* transcripts could be detected (Fig. 2*b*, clone 1). Note that the level of the endogenous *H–2K^b* mRNA is 5–10-fold lower than the level found in liver and spleen of *H–2K^b* mice. In contrast, the exogenous *H–2K^{bml}* gene is expressed at higher levels. This effect may be due either to a higher copy number (see legend to Fig. 2*b*) or to selective integration sites.

The increased levels of *H–2K* mRNA observed on differentiation of EC cells could be a result of transcriptional activation or specific stabilization of *H–2K* mRNA; we cannot yet distinguish between these possibilities.

All of the above S_1 analyses were done using a probe which spans one of the exon–intron junctions within the gene. To determine whether these transcripts were properly initiated, we used a probe that covered the 5′ end of the *H–2K^b* and *H–2K^{bml}* genes, which are identical in this region. Although this probe does not distinguish between the mRNA produced from the endogeneous and exogenous genes, it showed only the correct 5′ end expected for H–2K^b and H–2K^{bm1} mRNAs, therefore we conclude that transcription of the exogenous gene is correctly

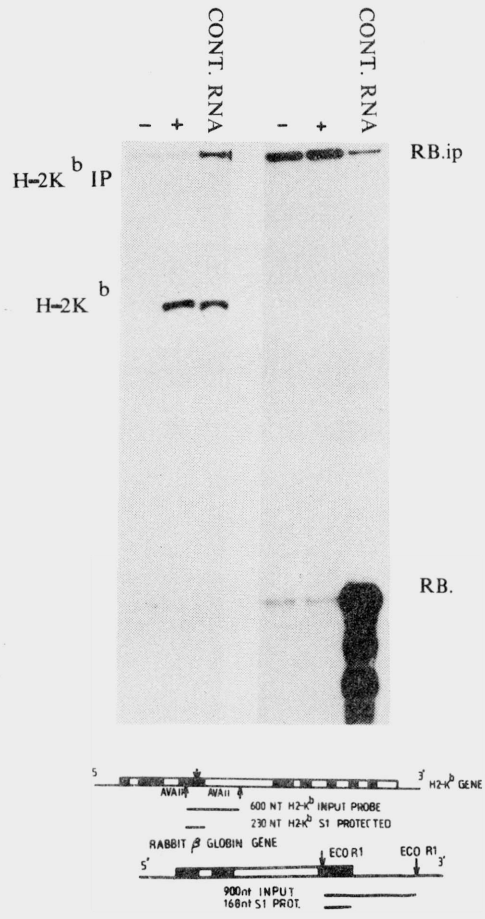


Fig. 3 S_1 nuclease analysis of H–2K^b and rabbit β-globin mRNA before (−) and after (+) differentiation. Control mRNA for H–2K^b was obtained from nontransformed F9 cells after differentiation to parietal endoderm. Control mRNA for rabbit β-globin was obtained from rabbit bone marrow. The 3′ rabbit β-globin probe (RBip) was end-labelled using reverse transcriptase and hybridized as described in Fig. 1 legend.

initiated (Fig. 2*c*).

In another set of transfection experiments, the rabbit β-globin gene, covalently linked to the same SV40 vector, was introduced into F9 EC cells. S_1 nuclease analysis of four different transformants showed a low level of β-globin mRNA which did not increase following differentiation, while the endogenous H–2K^b mRNA level in the same transformants increased (Fig. 3).

These results suggest that the DNA sequences involved in this induction process must be present on the 10-kb fragment containing the *H–2K^{bm1}* gene used in this experiment and that these sequences are specific for the *H–2K* gene. It is highly unlikely that the control sequences would be tissue-specific enhancer elements, as first observed for immunoglobulin genes[26–28]; instead, they may control the activation of the *H–2K* gene at a certain time in the development programme and its subsequent expression in all somatic cells. Induction of this sort may be mediated either by the release of gene suppression of by gene activation and our data are compatible with both of these models.

The fact that visceral-like endoderm and parietal-like endoderm F9 cells, which differ morphologically and functionally, both express H–2K and β_2M may imply that the F9 EC cells originate from embryonal cells predetermined[29] to express transplantation antigens. Thus, although F9 EC cells can be differentiated along alternative pathways[8,9,30], with respect to the expression of MHC genes they may not be different from other, already committed cell systems where increased gene expression can be induced chemically.

This work was supported by the BMRC (UK). A.R. was supported by The Michael and Anna Wix Trust through the Friends of the Hebrew University. We thank D. Barlow and B. Hogan for gifts of RNA and for technical advice, and R. Axel for his hospitality and assistance in the initial transformations.

Received 14 February; accepted 21 May 1984.

1. Strickland, S., Smith, K. K. & Maratti, K. K. *Cell* **21,** 347–355 (1980).
2. Ostrand-Rosenberg, S. *Curr. Topics dev. Biol.* **14,** 147–168 (1980).
3. Webb, C. G. *et al. J. exp. Med.* **146,** 923–932 (1977).
4. Klein, J. *Science* **203,** 516–521 (1977).
5. Ploegh, H. L., Orr, H. T. & Strominger, J. L. *Cell* **24,** 287–299 (1981).
6. Hood, L., Steimmetz, M. & Goodenow, R. *Cell* **28,** 685–687 (1982).
7. Klein, J. *Biology of the Mouse Major Histocompatibility Complex* (Springer, Berlin, 1975).
8. Strickland, S. & Mahdavi, V. *Cell* **15,** 393–403 (1980).
9. Hogan, B. L. & Taylor, A. *Nature* **291,** 235–237 (1981).
10. Nicolas, J. F. *et al. Ann. Microbiol.* **A126,** 3–22 (1975).
11. Croce, C. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **78,** 5754–5748 (1981).
12. Knowles, B. B. *et al. Nature* **288,** 615–617 (1980).
13. Jacob, F. *Immunol. Dev.* **33,** 3–32 (1977).
14. Tanaha, K., Appella, E. & Jay, G. *Cell* **35,** 457–465 (1983).
15. Cooper, A. R., Taylor, A. & Hogan, B. L. M. *Devl Biol.* **99,** 510–516 (1983).
16. Kurkinen, M. *et al. Nucleic Acids Res.* **11,** 6200–6209 (1983).
17. Parnes, J. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78,** 2253–2257 (1981).
18. Sherman, M. I. & Miller, R. A. *Devl Biol.* **63,** 27–34 (1978).
19. Benham, P. J., Quintero, M. A. & Goodfellow, P. M. *EMBO. J.* **2,** 1963–1968 (1983).
20. Artzt, K. & Jacob, F. *Transplantation* **17,** 632–625 (1975).
21. Forejt, J. *Cold Spring Harb. Symp. quant. Biol.* **48,** 541–552 (1983).
22. Andrews, D. W. & Goodfellow, P. N. *Somat. Cell Genet.* **6,** 271–284 (1980).
23. Correani, A. & Croce, C. M. *J. cell. Physiol.* **105,** 73–99 (1980).
24. Weiss, E. H. *et al. Nature* **301,** 671–674 (1983).
25. Wigler, G. R. *et al. Cell* **16,** 777–785 (1979).
26. Queen, C. & Baltimore, D. *Cell* **33,** 741–748 (1983).
27. Gillies, S. D. *et al. Cell* **33,** 719–728 (1983).
28. Banerji, J. *et al. Cell* **33,** 729–740 (1983).
29. Murphy, D. *et al. Cell* **35,** 865–871 (1983).
30. Liesi, P. *et al. Nature* **306,** 265–267 (1983).
31. Grosveld, G. C. *et al. Nature* **295,** 120–126 (1982).

Printed in Great Britain by Macmillan Production Ltd., Basingstoke, Hampshire

Phil. Trans. R. Soc. Lond. B **000**, 000–000 (1984)
Printed in Great Britain

DNA sequences required for regulated expression of β-globin genes in murine erythroleukaemia cells

By S. Wright[1], E. deBoer[1], A. Rosenthal[2], R. A. Flavell[3] and F. Grosveld[1]

[1] *Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.*

[2] *Department of Molecular Biology, The Hebrew University, Hadassah Medical School, 91 010 Jerusalem, Israel*

[3] *Biogen Research Corporation, 14 Cambridge Center, Cambridge, Massachusetts 02142, U.S.A.*

We have introduced into murine erythroleukaemia (MEL) cells a series of human globin gene cosmids and two sets of hybrid genes constructed from the human β-globin gene and the human γ-globin or murine H-2K^{bm1} genes. S1-nuclease analysis of the mRNA products from these genes before and after MEL cell differentiation showed that the human β-globin gene, but not the human ε- or γ-globin or H-2K^{bm1} genes, is induced specifically. Hybrid genes containing human β-globin DNA sequences from either 5′ or 3′ side of the translation initiation site were both inducible. Measurement of the relative rate of transcription showed this induction to be the result of transcriptional activation. We therefore suggest that DNA sequences which regulate β-globin gene expression during MEL differentiation are located both 5′ and 3′ to the translation initiation site.

Introduction

The human β-like globin genes are differentially expressed during the course of development. The five active genes which are located in a cluster on the short arm of chromosome 11 are expressed at different times and in different erythroid tissues. The embryonic ε-globin gene is expresed in the yolk sac, the foetal ^{G}γ- and ^{A}γ-globin genes in the foetal liver, and the adult δ- and β-globin gene in adult bone marrow (for review, see Maniatis *et al.* 1981). To study this tissue and stage specific expression, we have used DNA mediated gene transfer into cultured murine erythroleukaemia (MEL) cells. Chemically induced differentiation of these cells results in an increase in transcription by a factor of several hundred of the adult mouse globin genes (Friend *et al.* 1971; Ross *et al.* 1972; Reuben *et al.* 1976; Marks & Rifkind 1978). These cells thus serve as a model for gene activation during erythropoiesis and can be used to study the expression of foreign β-globin genes, as first shown by Willing *et al.* (1979) and Pyati *et al.* (1980). They used human chromosome 11-MEL cell hybrids to show that expression of the human β-globin, but not the human ε- and γ-globin genes, is regulated in such MEL cell hybrids. MEL cells were subsequently used to study the expression of foreign globin genes introduced by DNA mediated gene transfer (Spandidos & Paul 1982; Wright *et al.* 1983; Chao *et al.* 1983). We have used a series of cosmid DNA containing the human ε-, γ- and β-globin genes to show that the expression of the β-globin, but not the ε- and γ-globin genes, is regulated during MEL cell differentiation. These were followed by a series of recombinant genes with parts of the human β-globin gene to localize the sequence involved in the regulated expression.

REGULATED EXPRESSION OF THE HUMAN β-GLOBIN GENE

A series of recombinant cosmids containing different regions of the human β-globin locus (figure 1) was introduced into thymidine kinase negative (tk^-) MEL cells (Spandidos & Paul 1982). Before the calcium phosphate transformation (Wigler *et al.* 1979), the DNA was linearized within the vector sequences with PvuI. Between one and twenty stable tk^+ transformants per 10^6 cells per microgram DNA were obtained after selection in HAT medium.

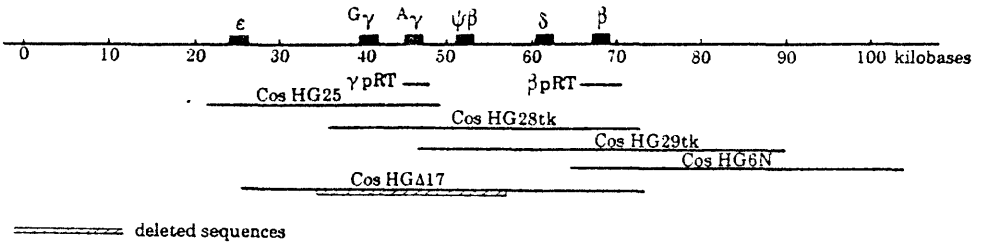


FIGURE 1. Structure of human globin gene cosmids. The human globin cluster was cloned in cosmid vectors containing the HSV *tk* gene (Grosveld *et al.* 1981; Grosveld *et al.* 1982). Cosmid DNA was prepared by the method of Birnboim & Doly 1979). PvuI linearized cosmid DNA was introduced into tk^- MEL cells (Spandidos & Paul 1982) by calcium phosphate transformation (Wigler *et al.* 1979) and selection in hypoxanthine-aminopterin-thymidine (HAT) medium. Reprinted by permission from *Nature*, vol. 305, no. 5932, pp. 333–336.

Each transformant was shown to contain 1–15 copies of non-rearranged exogenous DNA by Southern blot analyses (not shown). Ten different transformants from each cosmid transformation were induced to differentiate by culturing for three days in the presence of 3 mM hexamethylene bisacetimide (HMBA). RNA was prepared from each culture (Auffrey & Rougeon 1980) before and after differentiation. The levels of human and mouse globin RNA transcripts were quantitated by S1 nuclease protection (Berk & Sharp 1977; Weaver & Weissmann 1979) by using probes from the 5′ and 3′ ends of the genes.

Of the 54 clones tested, 52 showed more than a 100-fold increase in the levels of mouse β-major globin mRNA as measured by S1 nuclease protection analysis with a probe corresponding to the 3′ end of the gene (figure 2*a*). This corresponds to more than 20000 copies of mRNA per induced cell. Six of the transformants that showed induction had high levels (over 1000 copies per cell) of mouse β-major globin mRNA before induction.

Of 30 transformants that contained a foreign human β-globin gene, 23 showed an increase in the level of human β-globin mRNA after induction (average 40- to 50-fold), as shown in figure 2*b* with a probe corresponding to the 3′ end of the human β-globin gene. The induced levels of human β-globin mRNA corresponded to 150–2000 copies per cell. Of the seven transformants that did not induce, one also showed no induction for the endogenous mouse globin gene, and four showed high levels of mRNA before induction by HMBA. The levels of human β-globin mRNA either before or after induction did not correspond to the number of β-globin genes present in the transformants. Seventeen of nineteen transformants and all of nine transformants containing the complete human γ-globin gene (figure 2*c*), or the human ε-globin gene (not shown) showed the same or a smaller amount of γ- or ε-globin mRNA respectively. Two γ-globin transformants showed a low inducibility of γ-globin mRNA.

These data with cosmid recombinants strongly suggest that the large flanking sequences are

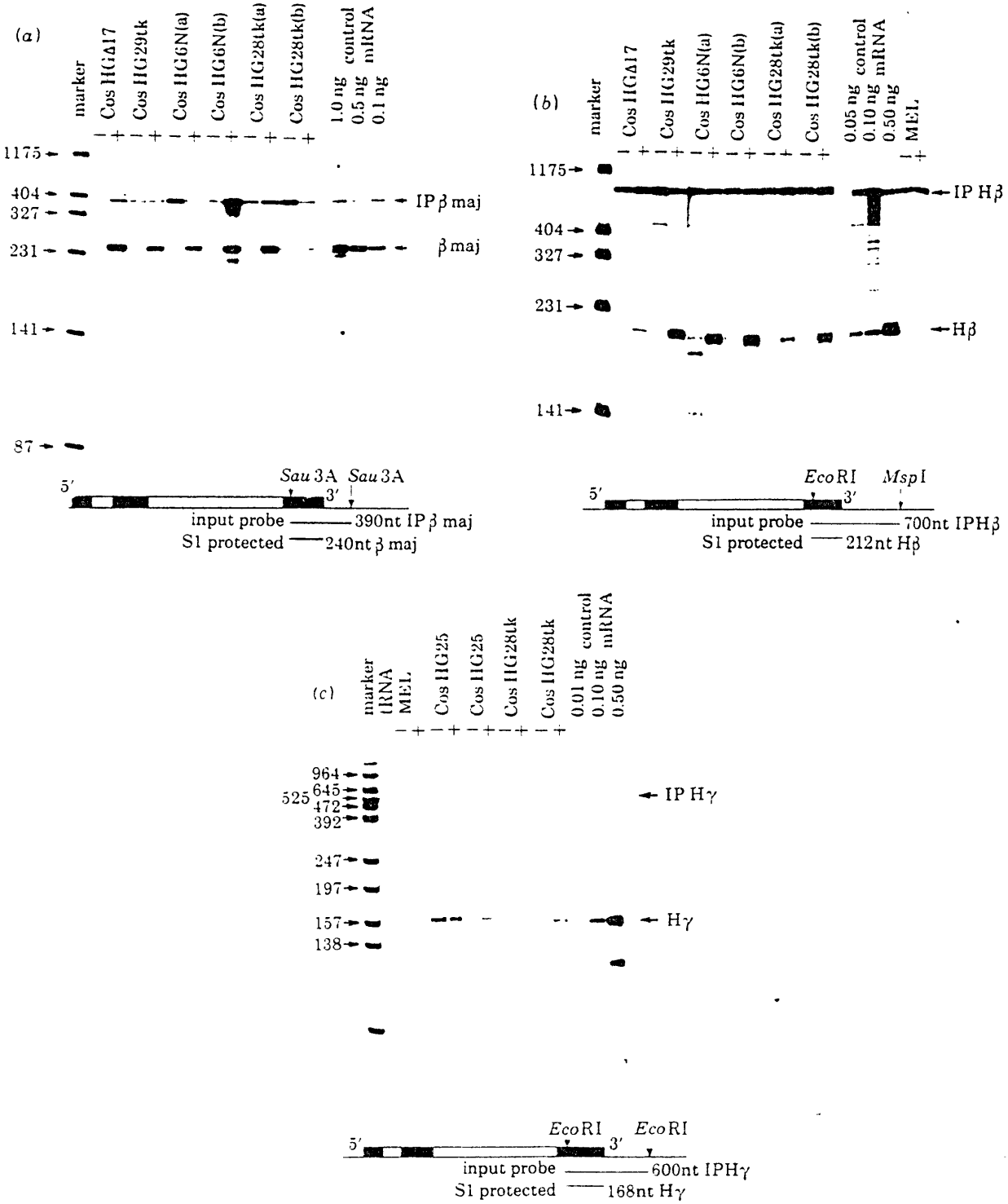


FIGURE 2. (*a*) S1 nuclease analysis of mouse β-major globin mRNA in MEL cell transformants. Control RNA was obtained from mouse reticulocytes. (*b*) S1 nuclease analysis of human β-globin cosmid containing MEL transformants (figure 1). Control RNA was obtained from a patient with sickle cell anaemia. (*c*) S1 nuclease analysis of human γ-globin cosmid containing MEL transformants (figure 1). Control RNA was obtained from a patient with $\beta^0/\delta\beta^0$ thalassaemia (Moschonas *et al.* 1981). MEL transformants were grown for three days in the absence (−) or presence (+) of HMBA, and RNA was prepared Auffrey & Rougeon 1980). The S1 nuclease analysis was as described by Wright *et al.* 1983). Reprinted by permission from *Nature*, vol. 305, no. 5932, pp. 333–336.

not required for the observed regulated expression of the β-globin gene. Regulation is obtained for both Cos HG28tk and Cos HG6N, which only contain approximately 2 kilobases of 3′ flanking DNA and 5 kilobases of 5′ flanking DNA sequences. To confirm these results a second series of MEL cells was selected after transformation with βpRT, γpRT and pTME10. These plasmids contain the human β-globin gene, the human γ-globin gene and the murine major histocompatibility gene H-2K^{bm1} respectively (Weiss *et al.* 1983) and short flanking sequences at the 5′- and 3′- ends of the genes (figure 3).

Eight out of nine transformants with the human β-globin gene (*βpRT*) showed the same increased levels of β-globin mRNA (figure 3*a*, *b*) as the transformants described above, when analysed with 5′- or 3′- and S1 nuclease protected probes. In contrast, only one of twelve γ-globin gene transformants and none of twelve H-2K^{bm1} gene transformants showed any increase in the respective mRNA levels (figure 3*c*, *d*). These results indicate therefore that only the β-globin gene expression is regulated in MEL-cells and that the regulatory DNA sequences appear to be located within a BglII fragment of 4.7 kilobases spanning the gene.

Localization of the sequences required for regulated expression

To determine whether the human β-globin DNA sequences that confer inducibility upon the gene are located in the 5′ promoter region, two hybrid genes were constructed (figure 4). The 5′ end of the human β-globin gene (up to the translation initiation site) was adjoined to the complementary 3′ region of either the human γ-globin or mouse H-2K^{bm1} genes, both of which are non inducible (see above).

The hybrid genes were introduced into tk$^-$ MEL cells and individual transformants isolated. Five out of ten 5′β–3′γ-globin transformants showed a 5- to 50-fold induction of the mRNA level as measured with a probe that detects the 3′ end of the γ-globin gene (figure 5*a*). In the second set, two out of three 5′β–3′H-2K hybrid transformants showed a similar induction, as measured with a probe specific for the H-2K mRNA (figure 5*b*).

A second set of hybrid genes was constructed to determine whether any inducible sequences were located to the 3′ site of the initiation codon of the β-globin gene. In these cases, the complementary constructs to those described above were made, i.e. the γ-globin or the H-2K 5′ end of the genes were adjoined to the body of the β-globin gene. The construction of these 5′H-2K–3′β and 5′γ–3′β hybrid genes is shown in figure 6. Both sets of genes were introduced into tk$^-$ MEL cells by co-transfer with *pRT*. S1 nuclease protection, with a 3′β-globin specific probe showed that four out of eight 5′γ–3′β and four out of five 5′H-2K–3′β hybrid gene mRNA levels increased at least five to more than 100-fold after induction (figure 7*a*, *b*). It has been shown that the accumulation of mouse globin mRNA in differentiating MEL cells is not only the result of an increase in the rate of transcription, but also of an increase in the relative stability of globin mRNA. It is therefore possible that the increase in the mRNA levels of the 5′γ–3′β and the 5′H-2K–3′β hybrid genes may be due to an increase in the stability of the mRNA rather than to an increase in the role of transcription. To distinguish between these possibilities, we measured the relative rate of transcription of these genes before and after differentiation using *in vitro* nuclear run-off transcription (Groudine *et al.* 1981; Hofer *et al.* 1982). Nuclei were isolated from MEL cells containing the hybrid genes, both before and after differentiation. RNA was extracted from these nuclei after a nuclear transcription reaction in the presence of ^{32}P-UTP and hybridized to globin probes spotted on to nitrocellulose filters. Three DNA probes were

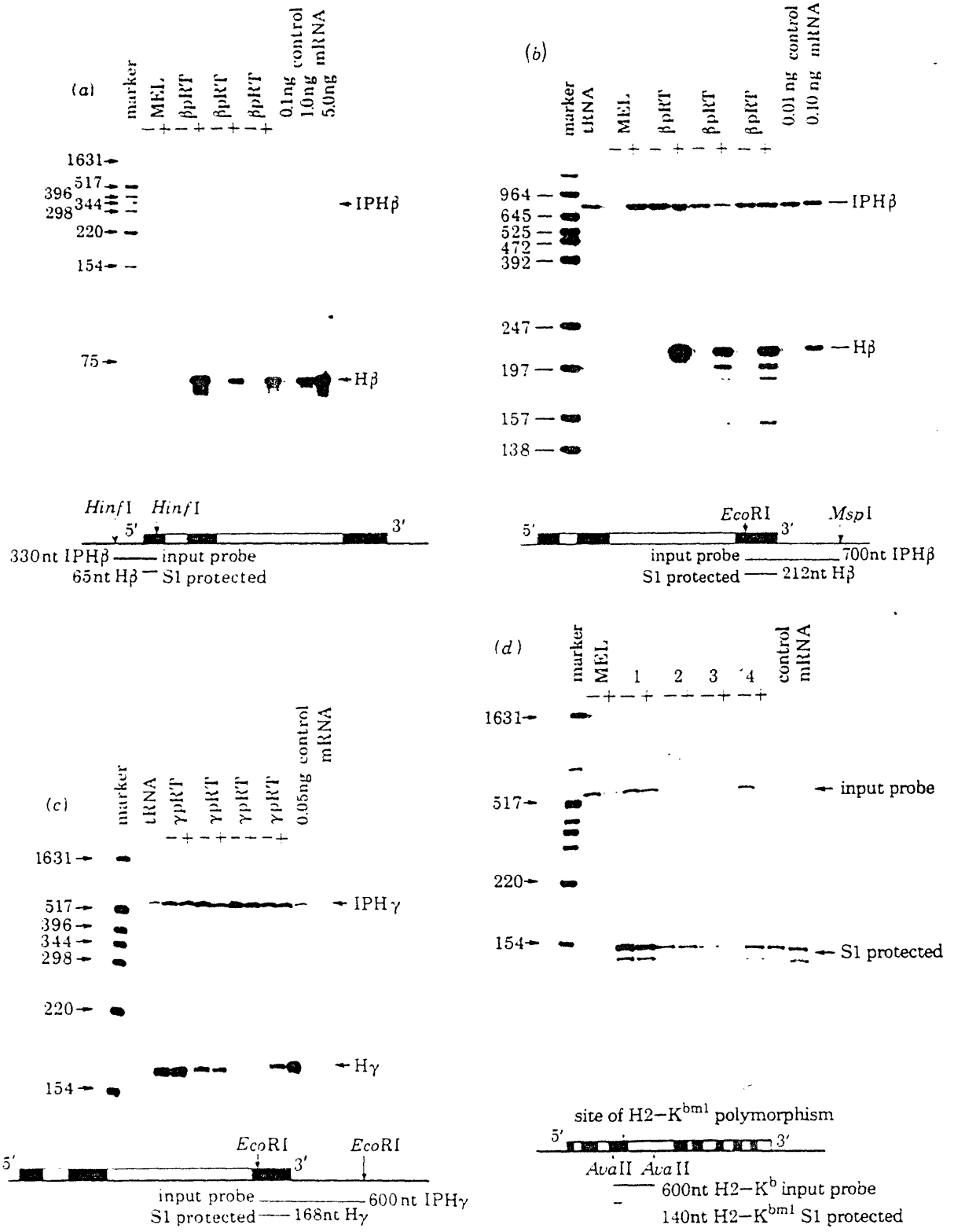


FIGURE 3. *(a)*, *(b)* S1 nuclease analysis of MEL transformants containing β *pRT* with a 5′ S1 nuclease probe and a 3′ S1 nuclease probe, respectively. Controls and methods as in figure 2. *(c)*, *(d)* S1 nuclease analysis of MEL cell transformants containing γ *pRT* or the H-2K^{bm1} gene respectively. Control RNAs were as in figure 2 and mouse spleen mRNA respectively. The methods were as described in figure 2. Reprinted by permission from *Nature*, vol. 305, no. 5932, pp. 333–336. Copyright © 1983 Macmillan Journals Ltd.

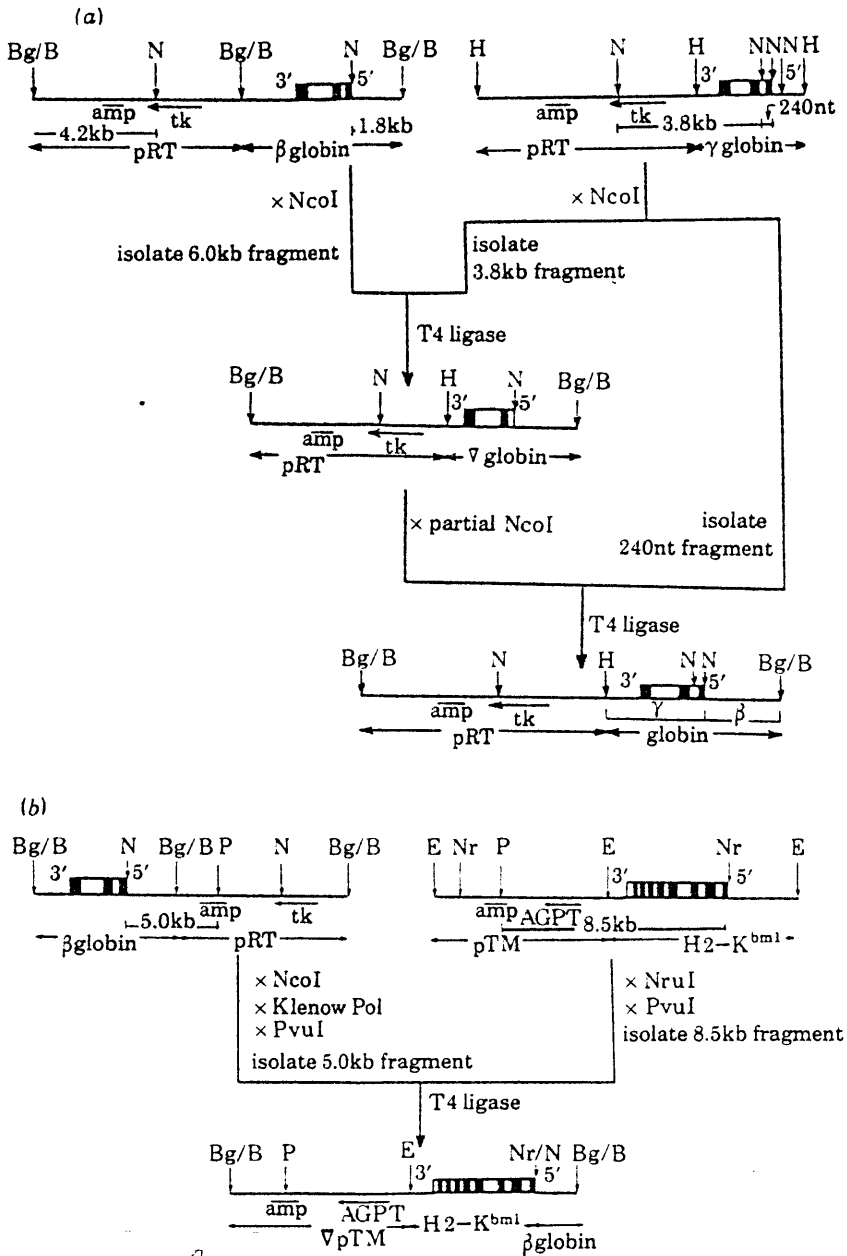


FIGURE 4. Construction of (*a*) 5′β–3′γ and (*b*) 5′β–3′H-2K^{bm1} hybrid genes. The hybrid genes were constructed from the globin and H-2K^{bm1} genes as described in the diagrams. N, P, E, Nr refer to *Nco*I, *Pvu*I, *Eco*RI and *Nru*I sites respectively.

used, an insulin probe (Bell *et al.* 1982) as a negative control and two globin probes from the large IVS of the mouse and human β-globin genes respectively. These probes therefore measure the amount of precursor mRNA in an area of the greatest sequence divergence between the mouse and human globin genes. The hybridization contained RNA from an equal number of nuclei and a vast excess of filter bound DNA probe, so that differences in the absolute rate of gene transcription could be measured.

Differentiation of control untransformed MEL cells showed an increase in hybridization to

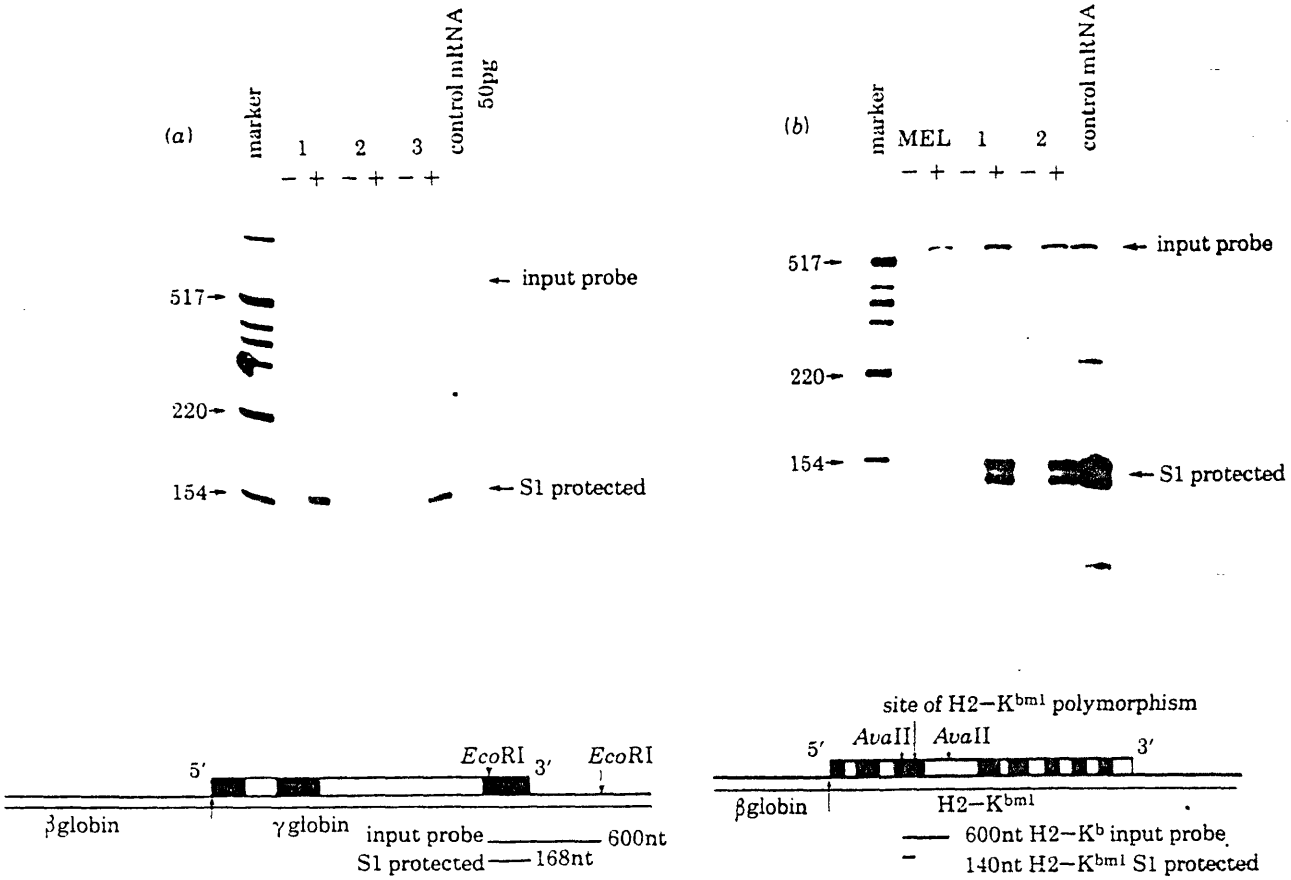


FIGURE 5. (*a*) S1 nuclease analysis of MEL cell transformants containing 5′β–3′γ hybrid globin genes. (*b*) S1 nuclease analysis of MEL cell transformants containing 5′β–3′H-2K^{bm1} globin genes. Control RNAs and methods were as described in figures 2 and 3.

the mouse probe, but not the human probe after differentiation (figure 8). This indicated an increase in the rate of transcription of the mouse β-globin mRNA, which did not cross hybridize with the human probe. In contrast, MEL cells containing the 5′γ–3′β or 5′H-2K–3′β hybrid genes showed an increased hybridization to both the mouse and human probes. Thus, the increase in the level of 3′β hybrid gene transcripts resulted (at least in part) from an increase in the rate of transcription following MEL differentiation.

DISCUSSION

We have shown independent regulation in the expression of individual human ε-, γ- and β-globin genes when introduced into MEL cells as a gene cluster. In accordance with the results obtained with human chromosome 11-MEL cell hybrids (Willing *et al.* 1979; Pyati *et al.* 1980), the expression of human adult, but not foetal or embryonic genes, is regulated during MEL cell differentiation. In our case, all the genes are expressed constitutively, probably because the genes are linked to a selective marker. The marker gene (*tk*) in turn will be integrated in an active part of the MEL cell genome and be expressed constitutively to grow under selective conditions. Different integration sites and chromatin structures of the introduced genes in

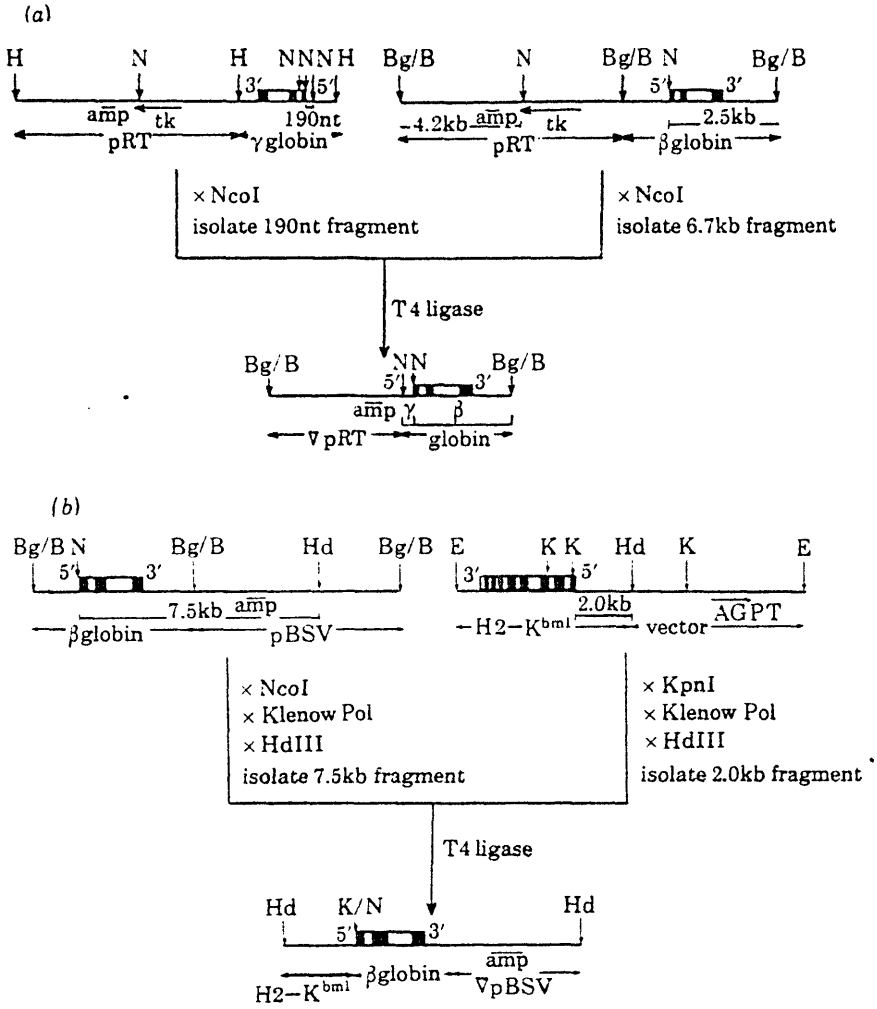


FIGURE 6. Construction of (*a*) 5′γ–3′β and (*b*) 5′H-2K^{bm1}–3′β genes. The genes were constructed from globin and H-2K^{bm1} plasmids as described in the diagram. N, R, E, K refer to *Nco*I, *Pvu*I, *Eco*RI and *Kpn*I sites respectively.

separate transformants, rather than any heterologous effects would also explain the variable levels of expression and induction of the introduced genes (Wright *et al.* 1983; Chao *et al.* 1983). To define the DNA region of the β-globin gene involved in induction, we constructed two types of hybrid genes: one set containing the upstream sequences from the initiation codon of the β-globin gene linked to the downstream complementary part of two non inducible genes (5′β–3′γ and 5′β–3′H-2K). The mRNA levels of these hybrid genes was increased after MEL cell differentiation, which suggests that the 5′- flanking upstream region of the β-globin gene contains sequences that confer inducibility of transcription.

A second set of hybrid genes contained the upstream region from the non inducible γ-globin and H-2K genes linked to the part of the β-globin gene downstream from the initiation codon (5′γ–3′β and 5′H-2K–3′β). Surprisingly, we also observed an increased level of transcripts from these hybrid genes after MEL cell differentiation. Nuclear run-off analysis showed this increase to be (at least partly) due to an increase in the rate of transcription. We therefore suggest that DNA sequences required for the activation of the β-globin gene during erythropoiesis are also

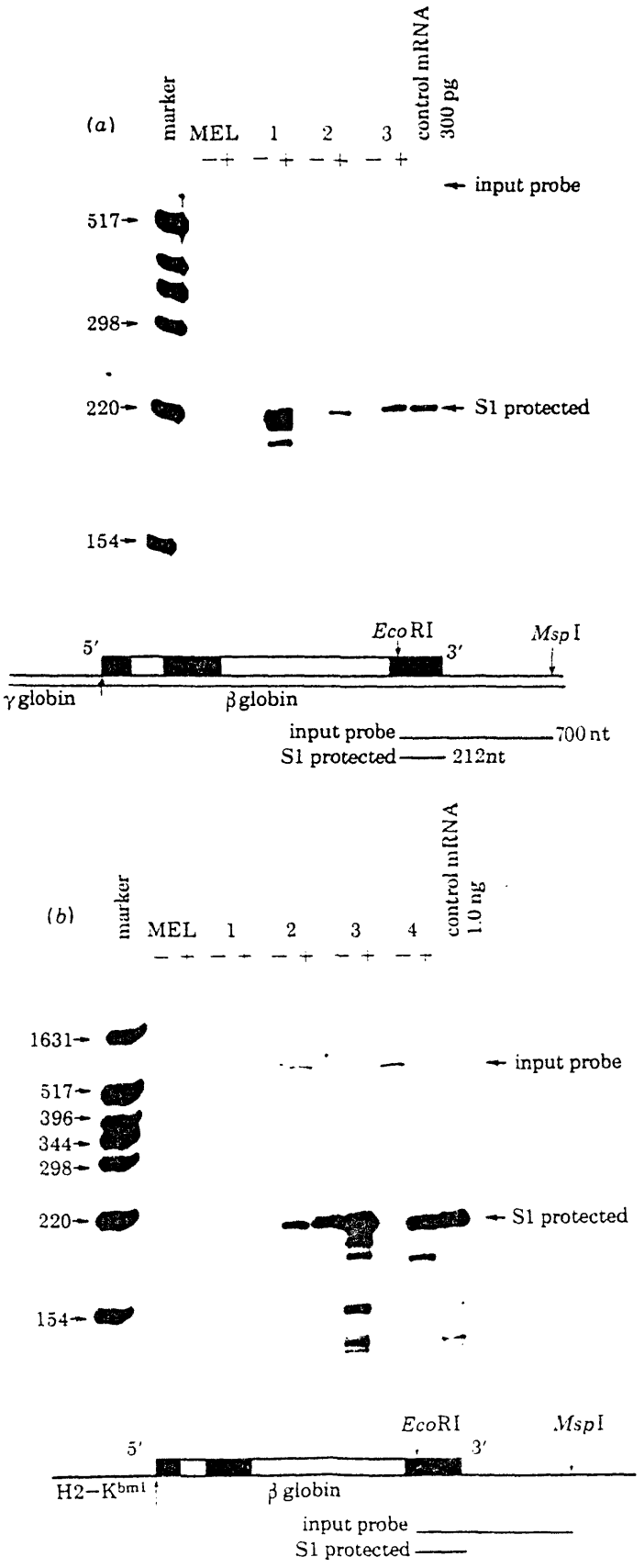


FIGURE 7. (*a*), (*b*) S1 nuclease analysis of the 5′γ–3′β and 5′H-2K^{bm1}–3′β hybrid genes respectively. Control RNA and methods were as described in the previous figures.

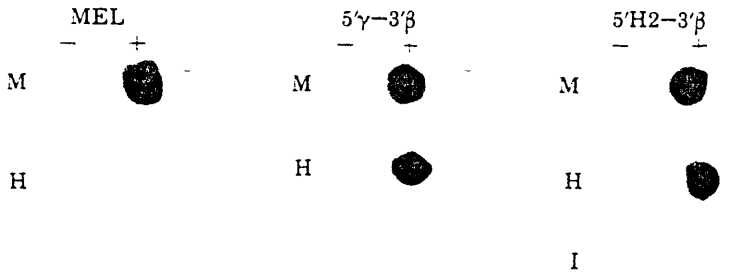


FIGURE 8. Nuclear run-off transcription of MEL cell transformants containing hybrid globin genes. Nuclei were isolated from non-induced (−) and induced (+) MEL cell transformants containing a 5′γ–3′β or a 5′H-2K^{bm1}-3′β hybrid gene. RNA was labelled with ^{32}P-UTP (300 μCi, 3000 Ci/mmol per 5×10^7 nuclei) as described by Groudine *et al.* (1984). Two μg of purified DNA fragments were bound to a 0.45 μm nitrocellulose filter, as described by Kafatos *et al* (1979), and hybridized with labelled nuclear transcript as described by Groudine *et al.* (1984). Each hybridization contained nuclear RNA from 10^7 cells which corresponded to 5×10^7 to 10×10^7 cpm for the non-induced cells and 2×10^7 to 5×10^7 for the induced cells. After hybridization, the filters were washed to a stringency of 0.3 XSSC. The human β-globin (H), mouse β-major globin (M) and human insulin (I) fragment probes were an *Eco*RI–*Bam*HI IVS fragment, a *Hin*III–*Pst*I IVS fragment and a 1.8 kilobases *Bam*HI fragment (Bell *et al.* 1982), respectively.

located at the 3′ side of the translation initiation codon. This would also explain the fact that the rabbit β-globin gene is still induced upon differentiation when large regions of the upstream sequences have been deleted (Wright, unpublished observations).

There are several other examples of gene transcription influenced by DNA sequences located downstream of the camp site, e.g. the internal promoter region of the *Xenopus* 5*S* RNA gene and the tissue specific enhancer of the immunoglobulin gene, which is located in an intron. In addition, it has been shown that the expression of foreign genes in eukaryotic cells can be increased by a rival enhancer when it is placed either at the 5′- or 3′- side of the gene.

Specific binding of the glucocorticoid receptor to MTV DNA *in vitro* occurs at sites both within and upstream of the transcribed region. This suggests that DNA sequences both within and upstream of the gene may regulate this glucocorticoid responsive gene. The results described in this paper may similarly indicate that DNA sequences both within and upstream of the globin gene may play a role in their activation during erythroid differentiation. Whether the activation is accomplished by the binding of a positive stimulatory factor(s), or the removal of a negative repressive factor(s) supplied *in trans* by the MEL cell during differentiation, is not clear from these results. The low levels of expression obtained with the 5′β-globin constructs 5′β–3′H-2K (figure 7) suggest that the β-globin promoter might be very weak or repressed in undifferentiated MEL-cells, although these results cannot be quantitated reliably. Nevertheless, it is tempting to suggest that both a negative regulation would operate on the 5′ end and a positive regulation on the 3′ end of the β-globin gene.

The evidence to date suggests that the induction *in trans* could only operate *in vivo* after the β-globin chromatin structure has been activated during erythroid differentiation. Changes in globin chromatin structure have been shown to occur with globin gene activation in chicken (Groudine *et al.* 1981) and humans (Groudine *et al.* 1983). Similarly, in MEL cells, new DNaseI hypersensitive sites appear in the mouse β major globin region after induction of MEL cell differentiation (Hofer *et al.* 1982; Sheffery *et al.* 1982). Although not proven, this suggests a causal relationship between globin gene activation and changes in chromatin structure. This relationship has been used to explain the phenotype of one form of γβ-thalassaemia (van der

Ploeg *et al.* 1979; Kouissis *et al.* 1983). In these patients a large deletion has removed the normal upstream sequences and positioned an unknown locus next to the β-globin gene. When this β-globin gene is cloned and introduced into MEL cells, it is expressed and induced normally (Wright, unpublished observations). Consequently, because this abnormal β-globin locus can respond normally to MEL cell differentiation, and because the heterozygous patient also contains a normal and expressed β-globin allele, it suggests that factors which act *in trans* could only act *in vivo* after the activation of the globin chromatin structure.

This work was supported by the Medical Research Council and A.R. was supported by a fellowship from the 'Friends of the Hebrew University'.

References

Auffrey, C. & Rougeon, F. 1980 Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. *Eur. J. Biochem.* **107**, 303–314.

Bell, G. I., Karam, J. M. & Rutter, W. J. 1981 Polymorphic DNA region adjacent to the 5′ end of the human insulin gene. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5759–5763.

Berk, A. J. & Sharp, P. A. 1977 Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell* **12**, 721–732.

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

Chao, M. V., Mellon, P., Charnay, P., Maniatis, T. & Axel, R. 1983 The regulated expression of β-globin genes introduced into mouse erythroleukemia cells. *Cell* **32**, 483–493.

Friend, C., Scher, W., Holland, J. G. & Sato, T. 1971 Hemoglobin synthesis in murine virus induced leukemic cells *in vitro*: stimulation of erythroid differentiation by dimethyl sulfoxide. *Proc. natn. Acad. Sci. U.S.A.* **68**, 378–382.

Grosveld, F. G., Dahl, H. H. M., deBoer, E. & Flavell, R. A. 1981 Isolation of β-globin-related genes from a human cosmid library. *Gene* **13**, 227–237.

Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M. & Flavell, R. A. 1982 The construction of cosmid libraries which can be used to transform eukaryotic cells. *Nucl. Acids Res.* **10**, 6715–6732.

Groudine, M., Kohwi-Shigematsu, T., Gelinas, R., Stamatogannopoulos, G. & Papayannopoulou, T. 1983 Human fetal to adult hemoglobin switching, changes in chromatin structure of the *β-globin* locus. *Proc. natn. Acad. Sci. U.S.A.* **80**, 7551–7555.

Groudine, M., Peretz, M. & Weintraub, H. 1981 Transcriptional regulation of hemoglobin switching in chicken embryos. *Molec. Cell. Biol.* **1**, 281–288.

Hofer, E., Hofer-Warbinek, R. & Darnell, J. E. Jr 1982 Globin RNA transcription: a possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. *Cell* **29**, 887–893.

Kafatos, F. C., Jones, C. W. & Efstratiadis, A. 1979 Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. *Nucl. Acids Res.* **7**, 1541–1581.

Kioussis, D., Vanin, E., deLange, T., Flavell, R. & Grosveld, F. G. 1983 β-globin gene inactivation by DNA translocation in γβ-thalassaemia. *Nature, Lond.* **308**, 662–666.

Maniatis, T., Fritsch, E. F., Lauer, J. & Lawn, R. M. 1981 Molecular genetics of human hemoglobins. *A. Rev. Genet.* **14**, 145–178.

Marks, P. A. & Rifkind, R. A. 1978 Erythroleukemic differentiation. *A. Rev. Biochem.* **47**, 419–426.

Moschonas, N., deBoer, E., Grosveld, F. G., Dahl, H. H. M., Wright, S., Shewmaker, C. K. & Flavell, R. A. 1981 Structure and expression of a cloned β^0 thalassaemic globin gene. *Nucl. Acids Res.* **9**, 4391–4401.

Pyati, J., Kucherlapati, R. S. & Skoultchi, A. I. 1980 Activation of human β-globin genes from non erythroid cells by fusion with murine erythroleukemia cells. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3435–3439.

Reuben, R. C., Wife, R. L., Breslow, R., Rifkind, R. A. & Marks, P. A. 1976 A new group of potent inducers of differentiation in murine erythroleukemia cells. *Proc. natn. Acad. Sci. U.S.A.* **73**, 862–866.

Ross, J., Itawa, Y. & Leder, P. 1972 Globin messenger RNA induction during erythroid differentiation of cultured leukemia cells. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3620–3623.

Sheffery, M., Rifkind, R. A. & Marks, P. A. 1982 Murine leukaemia cells differentiation: DNaseI hypersensitivity and DNA methylation near the globin genes. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1180–1184.

Spandidos, D. A. & Paul, J. 1982 Transfer of human globin genes to erythroleukemic mouse cells. *EMBO J.* **1**, 15–20.

van der Ploeg, L. H. T., Konings, A., Oort, M., Roos, D., Bernini, L. F. & Flavell, R. A. 1980 γ-β-thalassaemia: deletion of the γ- and δ-genes influences β-globin gene expression in man. *Nature, Lond.* **283**, 637–642.

Weaver, R. F. & Weissmann, C. 1979 Mapping of RNA by a modification of the Berk–Sharp procedure: the 5′ termini of 15*S* β-globin mRNA precursor and 10*S* β-globin mRNA have identical map coordinates. *Nucl. Acids Res.* **6**, 1175–1192.

Weiss, E. H., Mellor, A., Golden, L., Fahrner, K., Simpson, E., Hurst, J. & Flavell, R. A. 1983 Structure of a mutant H-2 gene: Generation of polymorphism in H-2 genes by gene conversion-like events. *Nature, Lond.* **301**, 671–674.

Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. & Axel, R. 1979 Transformation of mammalian cells with genes from procaryotes and eucaryotes. *Cell* **16**, 777–785.

Willing, M. C., Nienhuis, A. W. & Anderson, W. F. 1979 Selective activation of human β- but not γ-globin gene in human fibroblast X mouse erythroleukaemia cell hybrids. *Nature, Lond.* **277**, 534–538.

Wright, S., deBoer, E., Grosveld, F. G. & Flavell, R. A. 1983 Regulated expression of the human β-globin gene family in murine erythroleukemic cell hybrids. *Nature, Lond.* **305**, 333–336.

STRUCTURE AND EXPRESSION OF GLOBIN AND H-2 GENES

R. A. Flavell, F. G. Grosveld, G. C. Grosveld, S. Wright, M. Busslinger, E. deBoer, D. Kioussis, A. L. Mellor, L.Golden, E. Weiss, J. Hurst, H. Bud, H. Bullman, E. Simpson,[1] R. James,[2] A. R. M. Townsend,[3] P. M. Taylor,[3] W. Schmidt,[4] J. Ferluga[4], L. Leben,[4] M. Santamaria,[4] G. Atfield,[4] and H. Festenstein [4]

Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA

ABSTRACT The use of DNA-mediated gene transfer for the study of the rabbit β-globin promotor, the molecular basis of β^+thalassaemia and the specific induction of human β-globin expression in mouse erythroid cells is described. In addition, we describe the isolation of the class I mouse MHC genes from the H-2^b haplotype. We further show that an H-$2K^b$ gene when introduced into mouse L cells renders these cells a target for both allospecific and H-2 restricted, virus-specific cytotoxic T cell killing.

INTRODUCTION

The globin system has served us well as a model for the ays in which genes are organized on the chromosome and xpressed in the cells of higher eukaryotes. In this article e consider three aspects of globin gene transcription and NA processing, all studied using the techniques of DNA- ediated gene transfer. In the second part of the paper, we escribe the application of these techniques to the analysis f the functioning of cell surface antigens of the major istocompatibility complex (H-2) of the mouse.

[1]Clinical Research Centre, Northwick Park Hospital, Watford Road, Harrow, Middlesex HA1 3UJ.

[2]ICRF Tumour Immunology Unit, University College, Gower Street, London WC1E 6BT.

[3]Division of Immunology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA.

[4]Dept. of Immunology, The London Hospital Medical School, University of London, Turner Street, London E1 2AD.

THE RABBIT β-GLOBIN PROMOTOR

For a number of years we have studied the DNA sequences required for the transcription, in vivo and in vitro, of the rabbit β-globin gene. DNA sequence comparisons made by others, have pointed to two sequences, present to the 5' side of genes transcribed by RNA polymerase II called the "ATA" box ($^{T}_{C}ATA^{T}_{A}AG$; refs. 1, 2, 3) and the "CAAT" box ($GG^{T}_{C}CAATCT$ (3). Subsequent to their discovery, considerable effort has been expended to test the requirement of these sequences for transcription of cellular and viral genes. Thus, deletion of the ATA box causes a loss of specificity in the choice of the site of initiation of RNA synthesis (4, 5, 6, 7) and usually (4, 5) but not always (5, 6) a strong down mutation. Deletion of the CAAT box region is usually a strong down mutation (8, 5, 6, 7, 9), but these studies did not determine the necessary DNA sequences within this region. In the past nine months we have attempted to delineate these DNA sequences, within the -80 region which are required for transcription in vivo. The approach used is shown in Fig. 1. A variety of deletion mutants have been generated and then analysed in the SV40 HeLa cell transient expression assay system (10) as described (7). Basically deletions emanating around the CCAAT sequence at position -75 or further upstream at -96 have been analyzed. The mutant DNAs were introduced into HeLa cells as globin-SV40 -pBR328 plasmid recombinants. The transcription, which occurs from the rabbit β-globin gene promotor requires the presence of the SV40 enhancer sequence (72bp repeat); the plasmid used by us also contains the early region of SV40 so that T antigen is also expressed in these cells (about 30% of the cells are T antigen positive two days after transformation).

To quantitate our data, we cotransformed the HeLa cells with mutant rabbit β-globin genes and the 'wild type' human β-globin gene; transcripts from both genes are analyzed simultaneously in S_1 mapping experiments and the human β-globin gene serves as an internal control. A summary of these experiments performed with a number of deletion mutants is shown in Fig. 2.

The following conclusions can be drawn:

1. Most, if not all, of the entire GGCCAATCT sequence is required for efficient transcription in vivo. Deletion of the GG, or AATCT residues are down mutations.
2. Deletion of a region to the 5' side of the CCAAT sequence has an even stronger effect. Deletion -76 to -93 has an essentially intact "CAAT" box yet it is a 50-fold down mutation. Although this upstream region is less precisely defined, it is likely that the sequences from -83 to -94 are involved. This region is characterized by many C residues and is therefore

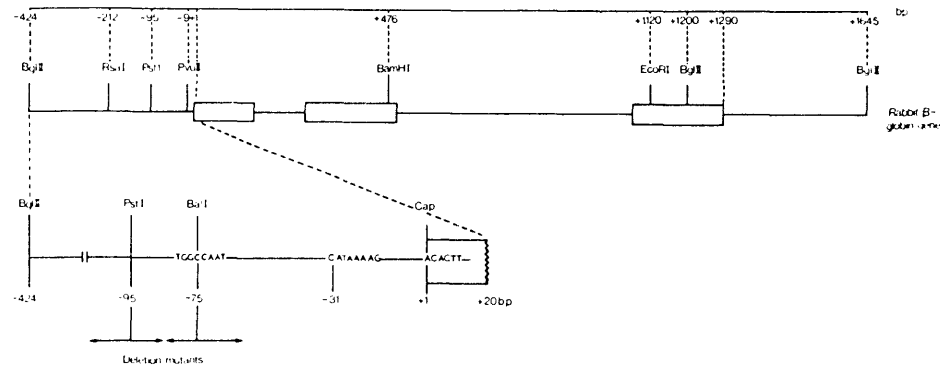


FIGURE 1. General structure of the rabbit β-globin gene and the strategy used to generate deletion mutants. The rabbit β-globin gene is shown together with a number of relevant restriction enzyme cleavage sites. The bottom half of the figure shows an expanded view of the promotor region.

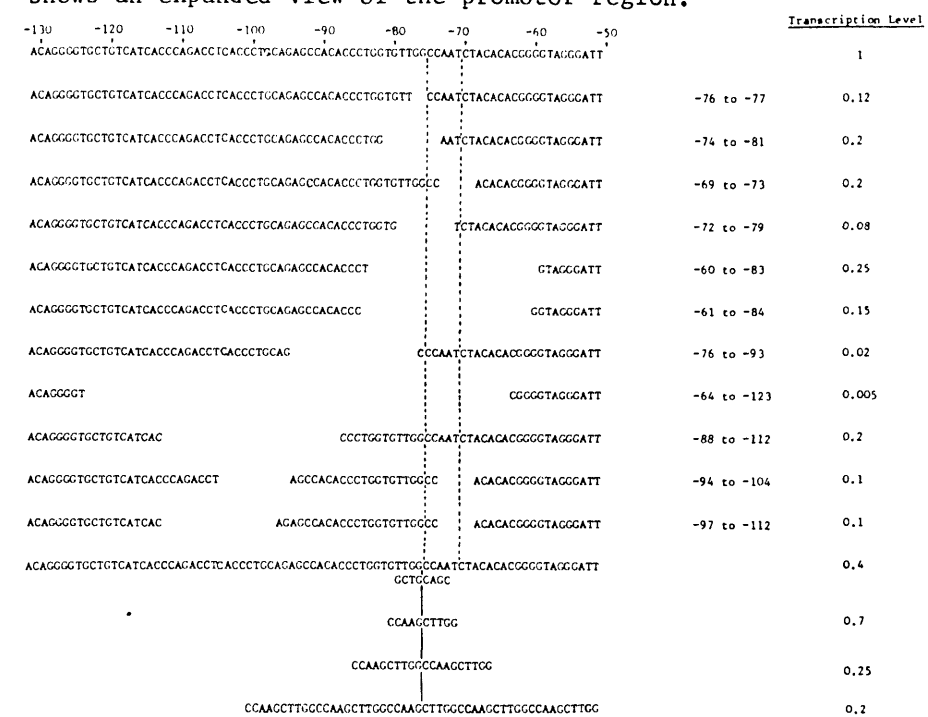


FIGURE 2. The transcription efficiencies of deletion and insertion mutants in the -80 region of the rabbit β-globin gene. The sequence in the -80 region of each mutant template is presented, and the gaps show the DNA residues deleted. Transcription level was determined by scanning the autoradiograps

similar to a region of the TK gene, mapping from -95 to -105, which is required for transcription of that gene in mouse L cells (6).

The rabbit β-globin gene promotor can therefore be seen conceptually as a tripartite structure, the ATA box, CAAT box and the -90 region. In our transient expression system the enhancer sequence of SV40 is required for expression. Why this is, and whether there is a cellular counterpart for enhancing globin gene expression, is not clear.

PHENOTYPIC ASSAY OF INHERITED DEFECTS IN HUMAN GENES

Structural studies of cloned genes from patients with inherited diseases have in many cases provided a clear explanation for the genetic defect. Thus, many such diseases result from a single base substitution which generates a stop codon; as a result the protein is not formed (e.g. 11, 12,13). Yet it is likely that genetic lesions exist for which the molecular defect is not so easy to determine, either because the mutations detected by structural analysis are not clearly interpretable, or because a large number of polymorphic differences are present which mask the true mutation that causes the disease. To solve this, it is essential to correlate structure with function.

β-Thalassaemia is an excellent model system for this type of study. It is by far the commonest form of thalassaemia in the β-related globin gene family and it is generally divided into two classes. In the first, β^+-thalassaemia, a low but detectable level of β-globin protein is found; this low level of β-globin protein reflects a low level of β-globin mRNA. It seems likely, a priori, that this disease would be the result of a defect in the transcription or processing of β-globin mRNA precursors. In β^0thalassaemia no β-globin protein is detected. In this case, therefore, a number of causes of the genetic lesion can be envisaged. In some cases, globin mRNA is readily detectable (14), whereas in others, β-globin mRNA could not be detected using cDNA hybridization techniques (e.g. see reference 15). It is likely, therefore, that there are several different molecular forms of β^0thalassaemia.

The molecular defect in β^+-thalassaemia has been studied by structural analysis of cloned thalassaemic β-globin genes (16,17). Interestingly, both genes show a single base substitution when compared with the normal β-globin gene -- TTGG -> TTAG at a site 21 nucleotides to the 5' side of the 3' splice junction of the small intervening sequence of the β-globin gene. Among the possible phenotypes that could be

postulated to explain this result, the most interesting is that this generates a new splice acceptor site in the small intervening sequence. We have recently shown this using the SV40-HeLa cell system (18). S_1 mapping of the β-globin RNA produced in HeLa cells from a β^+thalassaemic β-globin gene shows that about 95% of the β-globin mRNA produced is spliced to this abnormal splice site (not shown). To establish the nature of the aberrant RNA produced, we have used a 5'-labelled DNA fragment as a primer to synthesize, using reverse transcriptase, a cDNA complementary to this β^+-globin RNA. This cDNA has been purified and its sequence determined (Fig. 3). As a control, cDNA synthesized from globin mRNA from a patient with sickle cell haemoglobin was also sequenced. It can be seen that the abnormal splicing of the small intron generates a β-globin mRNA with an additional 19 nucleotides which derive from the small intervening sequence; since this is not a multiple of 3, this insertion results in the alteration of the translational reading frame and generates an in-phase stop codon 5 amino acids after the splice.

From the above, it should be clear that a combination of structural and functional analyses will make it possible in the future to elucidate the molecular nature of defects in a large number of human diseases. In the examples provided here, we show that defects at the level of protein synthesis and RNA processing can be elucidated using these systems. Since the globin mRNA produced in these cells is spliced, polyadenylated and exported to the cytoplasm (18), it is likely that defects at the level of RNA transport can also be studied. It is to be anticipated that the combination of structural and functional analyses in the next few years will bring some interesting discoveries in the field of the molecular biology of human diseases.

SPECIFIC EXPRESSION OF THE HUMAN β-GLOBIN GENE IN ERYTHROID CELLS

In the above mentioned experiments, globin gene expression has been analysed in cells (such as HeLa cells) which do not express their endogenous globin genes. Although this permissive behaviour is useful, this system is not suited to a study of those factors which determine the specific expression of globin genes in erythroid cells. To examine this, we have introduced various segments of the human β-related globin gene locus as cosmid recombinants into an erythroid cell line, the well known Friend cell (murine erythroleukaemia cells or MEL) and then determined the level of human β-globin mRNA in the transformed MEL cells before and after the

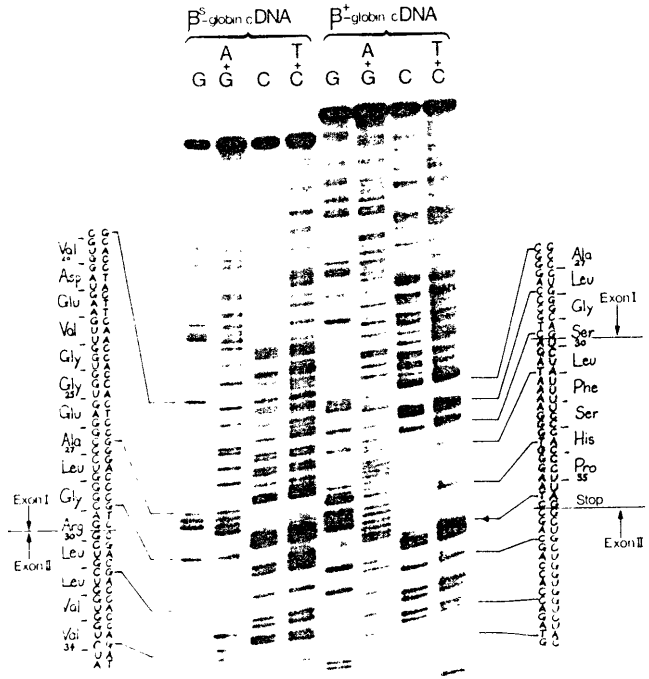


FIGURE 3. cDNA sequence over the junction of Exon I and Exon II of β^+-globin mRNA.

We synthesized cDNA by hybridizing an end-labelled, 87 nucleotide CvnI-Cvn I DNA primer to β^+-globin mRNA obtained from transformed HeLa cells and by subsequent primer extension with reverse transcriptase. Full-length cDNA synthesized up to the 5' end of β-globin mRNA was sequenced, and then separated on a 10% polyacrylamide sequencing gel. β^S-globin mRNA isolated from reticulocytes of a patient with sickle cell anaemia was used in the control experiment. The cDNA sequences and the complementary mRNA sequences, together with the deduced amino acid sequences, are shown to the right and left. Amino acids are numbered starting at the N terminus of the β-globin protein. Shaded area: the additional 19 nucleotides of the β^+-globin mRNA that are inserted between Exon I and Exon II of the normal β-globin mRNA.

induction of erythroid differentiation. The DNA segments that we have introduced into these cells are shown in Fig. 4. In the majority (15 of 24 clones examined) of clones that contain the human β-globin gene, the level of β-globin mRNA is stimulated by from three to greater than ten-fold by inducers of erythroid differentiation (Fig. 5 shows an example). This result is found for all the relevant cosmid clones; since the cosmid clones used have from 2.5kb to about 35kb of the DNA sequences flanking the β-globin gene on the 5' side and the same amount of 3'-flanking DNA, it follows that DNA sequences far from the human β-globin gene are not required for the induction phenomenon. Of those clones that contain the γ-globin genes, the majority are not inducible; instead, constitutive synthesis is usually seen or a decrease in the level of γ-globin mRNA after induction (see the MEL clone in,e.g.,Fig. 6 which is inducible for β-globin mRNA but not γ-globin mRNA).

It is likely that the induction of human β-globin mRNA is regulated at the level of transcription, although we have not specifically addressed this question. In all clones examined, the mouse α-globin mRNA induction was monitored and was greater than ten fold in all cases. Secondly, in clones where the β- and γ-globin genes (such as Fig. 5) were present γ-globin mRNA, but not β-globin mRNA, was present before induction; upon induction β-globin mRNA is found. It is therefore unlikely that the human β-globin mRNA, but not the γ-globin mRNA is being synthesized but selectively degraded before induction.

MOLECULAR ANALYSIS OF A COMPLEX GENE LOCUS; THE MAJOR HISTOCOMPATIBILITY COMPLEX

The major histocompatibility complex (MHC) plays an important role in the regulation of the immune response in vertebrates. There are three classes of MHC proteins, the classical transplantation antigens (class I: in man HLA-A B and C; in mouse the H-2 K L and D proteins), immune response-associated antigens (class II: in man HLA DR; in mouse Ia antigens) and the complement products (class III). The genes for the three classes of antigens are closely linked on chromosome 17 of the mouse. Adjacent to the H-2 complex in mouse is the TL complex which contains several genes encoding lymphoid differentiation antigens, the Qa and Tla antigens. See Fig. 6 for a schematic genetic map of the MHC complex in mouse.

Both the H-2 antigens (K, D, L,and R) and the Qa and Tla are integral membrane proteins of 40-45000 molecular weight found associated with β_2-microglobulin. It is,

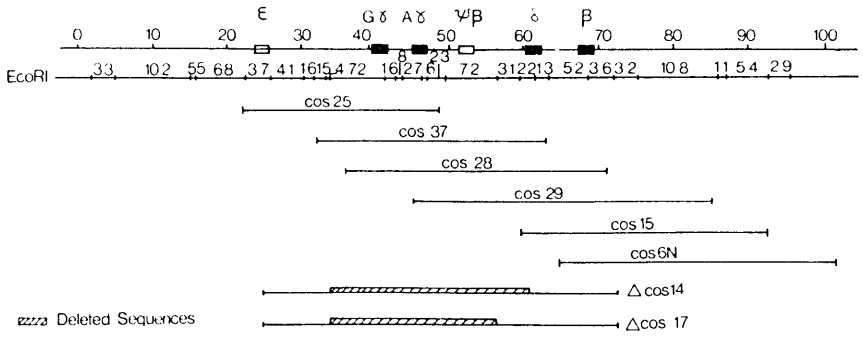


FIGURE 4. Cosmid clones used to transform MEL cells. A map of the human β-globin is shown; the coordinates of the cosmids are indicated by the solid lines.

FIGURE 5. S_1-mapping of RNA from MEL cells before (-) and after (+) induction of erythroid differentiation. 3'-labelled probes were used to measure simultaneously γ- and β-globin mRNA. The position of the relevant bands is indicated. The bands seen in the upper part of the gel are input fragments.

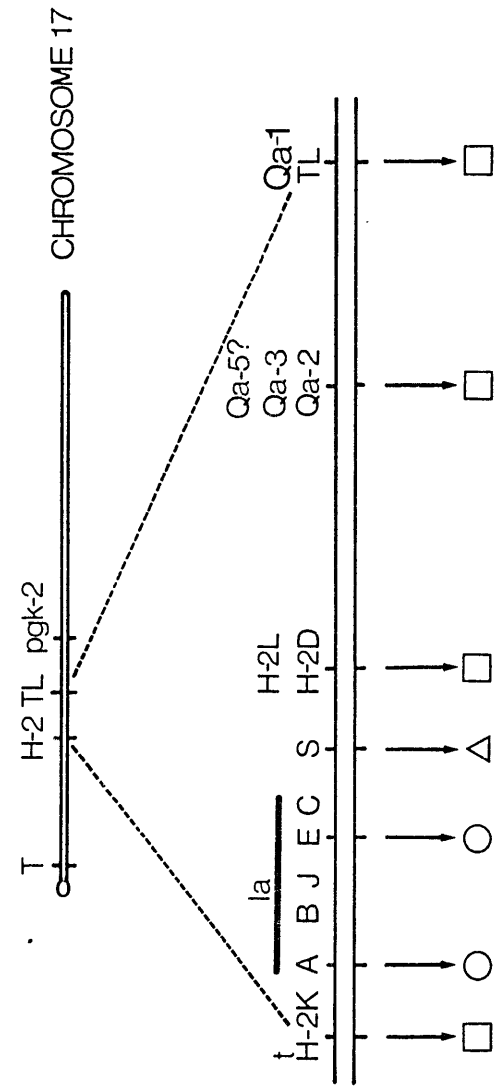


FIGURE 6. Schematic genetic map of the H-2 and TL complexes of chromosome 17 of the mouse. The relative order of the loci listed vertically is not known. The distance from H-2K to H-2D is about 0.3 centimorgans (cm) and from H-2D to Tla is about 1 cm.

□ Class I antigens; ○ Class II antigens;

△ Class III antigens.

One of the most intriguing aspects of the H-2 antigens is their extreme polymorphism. Allelic products from different inbred mouse strains differ in most cases as shown by both serological analyses and protein sequencing. In fact, about 50 alleles at each of the H-2K and H-2D loci have been observed (for a review see 21). H-2 haplotypes are defined serologically and consist of a particular set of alleles found at each H-2 locus (e.g. the C57 black/10, or B10, mouse is $H\text{-}2^b$ at each H-2 locus, the AKR strain $H\text{-}2^k$, DBA/2 $H\text{-}2^d$ and so on). Some inbred mouse strains appear to be natural recombinants between other strains (e.g. the A mouse strain with the $H\text{-}2^a$ haplotype is $H\text{-}2K^k$ but $H\text{-}2D^d$). It is likely that the analysis of the gene clusters by recombinant DNA technology will provide some clues to the nature of the polymorphic differences.

A number of groups have isolated cDNA clones which encode proteins whose sequences are homologous to known transplantation antigens (see e.g., 22 and 23). Southern blotting analysis of genomic DNA probed with these cDNAs shows many cross-hybridizing DNA fragments which reinforces the notion that the MHC is a large multi-gene family. We have isolated a large number of mouse genomic clones which contain sequences complementary to H-2 cDNA from a cosmid library (24) using spleen DNA from the B10 mouse (A. Mellor, L. Golden and H. Bullman, unpublished). Analysis of these clones by standard restriction enzyme digestion techniques establishes that many of these clones fall into a number of gene clusters (Table I); in total about 17 different H-2 like genes have been isolated in these cosmids.

Which H-2 genes are contained within these clusters? The number of independent H-2-like DNA clones exceeds that expected from the 4-5 known H-2 genes. This suggests that other H-2 related sequences are present in the mouse genome. A first step to identifying these genes is to determine where they are located on the MHC genetic map. To do this we have made use of polymorphic differences at restriction sites that exist between different mouse strains in the DNA sequences flanking a given H-2 gene. This is done by cutting DNA from, say, B10 ($H\text{-}2^b$), AKR ($H\text{-}2^k$) and BALB/c ($H\text{-}2^d$) mice with a number of restriction enzymes and blotting this onto nitrocellulose filters. The filters are then probed with a fragment from a given cosmid in order to detect polymorphic restriction site differences between the parental inbred strains. For example, using a probe from a clone in cluster 4 in PstI digests of these mouse DNAs shows a 2.5kb band in B10 DNA, but a 3.8kb band in AKR DNA. Recombinant mice have been produced which contain a given segment of the MHC map from one haplotype (say AKR; $H\text{-}2^k$) and a second segment from a second

mouse (say C57 black/6 or B6; $H\text{-}2^b$). The origin of genes which have been recombined in these mice has been shown by serological analysis of the recombinants. Thus, in the recombinant mouse B6K2 the T1a and Qa1 loci are derived from the AKR ($H\text{-}2^k$) mouse, whereas the H-2 region comes from the B6 mouse. Blotting PstI-cut DNA from the B6K2 DNA with the same probe shows that the AKR pattern (3.8kb band) and not the B6 pattern is obtained. It follows, therefore, that gene cluster 4 is localized in the AKR segment of the B6K2 mouse; that is, in the T1a region.

We have used the same approach to localize DNA segments in the H-2K region, the H-2D region and the Qa2, 3 region. Thus, class I, H-2 genes, or pseudogenes are found spread over chromosome 17 between the H-2K and TL genetic loci; a region of 1.3 cM and perhaps as much as 10^6-10^7 base pairs of DNA. This gene cluster might then be two orders of magnitude larger than the β-globin cluster (Table I).

EXPRESSION OF AN $H\text{-}2K^b$ GENE IN MOUSE L CELLS

Although cosmid clones can be mapped to a given locus by the mapping procedures described above, to establish the identity of any given gene, a more detailed analysis is required. We have, therefore, introduced H-2 cosmids into mouse L cells to ask if these cells which contain a given cosmid express a known H-2 gene. We have transformed L cells with cosmid DNA as a calcium phosphate coprecipitate and then selected for stable TK positive transformants. Since the cosmid vector used for most of our experiments, pOPF1 (F. Grosveld and T. Lund, unpublished) contains the Herpes Simplex Virus (HSV) thymidine kinase (tk) gene, transformants are readily obtained. Transformants were screened for the binding of monoclonal antibodies which react with the H-2K or H-2D molecules. L cells transformed with the cosmid H8 bind the $H\text{-}2K^b$ monoclonal specifically, while none of this group binds the $H\text{-}2D^b$ monoclonal (not shown). Cosmid clones H8, H24 and H25 are all found in cluster 1 and map at the H-2K locus. This cluster (Fig. 7) contains two genes; since both H24 and H25 contain gene A and do not express $H\text{-}2K^b$ we deduce that gene B is the $H\text{-}2K^b$ gene. Sequence data confirm this conclusion (E. Weiss, R. Zakut, unpublished). L cells containing the B gene of cluster I (called gene IB here) bind a variety of other alloantisera directed against the $H\text{-}2K^b$ molecule and immuno precipitations with anti $H\text{-}2K^b$ antibodies reveal the expected 45K protein. We are therefore confident that the cells express the $H\text{-}2K^b$ protein.

TABLE I

Summary of H-2^b gene clusters isolated from B10 mouse library

Cluster	map location	number of genes	number of cosmids
1	H-2K	2	7
2	H-2D	1	4
3	Qa 2/3	3	25
4	Tla	4	39
5	?	3	11

H-2 cosmids were mapped using standard restriction enzyme digestion and blot-hybridization with probes for the 5' and 3' regions of H-2 cDNA, mouse genomic DNA and the vector used for the construction of the library, nine cosmids have not yet been arranged into clusters. Of these one maps to Qa 2/3 and one appears to map outside the entire locus.

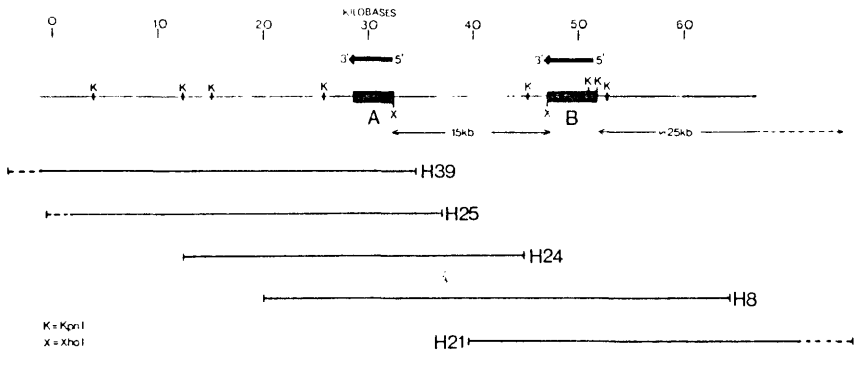


FIGURE 7. A restriction enzyme map of the region cloned in the H8 cosmid cluster. The top line shows the extent of the cloned region and the approximate positions of XhoI (X) and KpnI (K) restriction sites (see top scale calibrated in kilobases, kb). Class I gene regions (A and B) are shown as thick lines and their 5' to 3' orientation is indicated by the thick arrows over the line. The approximate extent of each cosmid is shown below the line beside the name assigned to each cosmid. Dotted lines at the end of cosmids indicate that the end of the cosmid has not been determined precisely.

CYTOTOXIC T CELL MEDIATED LYSIS OF TRANSFORMED CELLS

The presence of the H-$2K^b$ molecule on a cell renders these cells as targets for allogeneic cytotoxic T-cell (T_C-cell) mediated killing by T_C cells raised against H-2^b antigens (25). A number of indirect arguments suggest that the presence on the cell membrane of a given H-2 polypeptide, such as the H-$2K^b$ molecule, is sufficient to generate a target for T_C-cells. For example, this process can be blocked by monoclonal antibodies directed against a single H-2 polypeptide (26). To test this directly, and also to further characterize the H-$2K^b$ polypeptide detected by serological assays on the surface of L-cells transformed by cosmid H8, we generated T_C cells directed against H-2^b or H-2^k antigens and tested their ability to kill various target cells (not shown). Whereas anti-H-2^k T_C cells show specific killing of all L-cell lines tested, anti-H-2^b T_C-cells show specific killing of only control B10 target cells and two L-cell clones (LH8-1 and LH8-2) transformed by cosmid H8. Untransformed L-cells (Ltk^-), a clone transformed by cosmid H25 (LH25-1) and other clones transformed by cosmids H24 and H39 (data not shown) are not seen as targets for killing by anti-H-2^b T_C-cells. This data confirms that the H-$2K^b$ molecule detected by serological assays in cosmid H8 transformed cells, can act as a target for anti-H-2^b T_C-cell killing.

We performed a similar experiment using B10.D2 (H-2^d) T_C-cells generated by immunisation with target spleen cells from the recombinant mouse B10.A (5R) (K^b,D^d). Such T_C-cells kill cells expressing the H-$2K^b$, but not the H-$2D^b$ molecule. These T_C-cells also kill L-cells transformed with cosmid H8, but not L-cells transformed with cosmid H39 (not shown).

CYTOTOXIC T CELL-MEDIATED LYSIS OF CELLS INFECTED WITH INFLUENZA VIRUS

T_C-cell-mediated lysis of cells expressing viral antigens on their surface can only take place when the viral antigen is presented to T_C-cells in association with an appropriate H-2 antigen; in this way H-2 antigens act as a restriction element for the cytotoxic T-cell response (25, 27). Thus, we tested the ability of the H-$2K^b$ molecule expressed on the surface of L-cells transformed by cosmid H8 to act as a restriction element for T_C-cell-mediated killing of influenza virus-infected cells.

Influenza-specific T_C clones restricted to one H-2 region can be selected and grown in the presence of T-cell growth factor (28). L-cells (H-2^k) and LH8-1 were infected

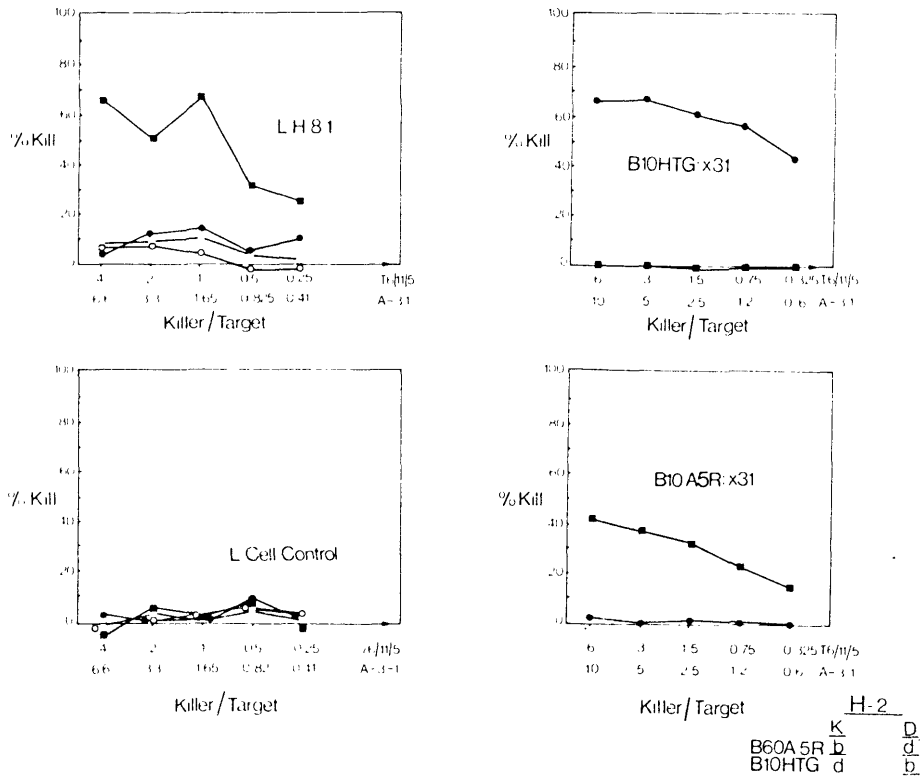


FIGURE 8. Cytotoxic T-cell-mediated killing of transformed L-cells infected with influenza virus.

Two influenza specific T_C-clones, T6/11/5 (H-2K^b restricted; square symbols) or A3.1 (H-2D^b restricted; circular symbols) were incubated with influenza virus (strain X31) infected (closed symbols) or non-infected (open symbols) LH8-1 or Ltk$^-$ (L-cell control) target cells. The percent specific lysis estimated by ^{51}chromium release (29) at each killer (T_C-cell) to target (LH8.1 or Ltk$^-$) ratio was calculated using the formula $\frac{(E-C)}{(M-C)} \times 100\%$. Where E = cpm from target cells incubated with killer cells, C = cpm from target cells incubated with medium alone and M = cpm from target cells lysed with 5% triton. Controls using influenza virus infected macrophage* targets from B10HTG mice (H-2K^d, D^b) or B10A5R (H-2K^b, D^d) demonstrate the H-2K^b and H-2D^b restricted killing of T_C-clones T6/11/5 and A3.1 respectively. The isolation and characterization of these T_C-clones will be described in detail elsewhere (Townsend and Taylor, manuscript in preparation).

*

for two influenza specific H-2^b T_C clones, one restricted to H-2K^b (T6/11/5) and the other (A/3.1) to H-2D^b molecules (A. Townsend and P. M. Taylor, unpublished data). Figure 8 shows that X31 infected, untransformed L-cells (H-2^k) are unable to act as targets for killing by either the H-2K^b or H-2D^b-restricted T_C-cell clones. LH8-1 cells infected with X31, however, act as targets for killing by the H-2K^b restricted T_C cell clone, but not the H-2D^b restricted clone. This shows that the H-2K^b molecule expressed on the surface of LH8-1 cells is able to act as a restriction element for T_C-cell killing of X31 virus infected cells. We have recently carried out the same experiments with L cells transformed with an H-2D^b cosmid. Again, using an H-2D^b-restricted, influenza virus specific T_C cell clone, we have obtained specific lysis. It is clear that this will be a general phenomenon.

In summary, it is clear that the presence of the H-2K^b gene in mouse L cells is sufficient to evoke essentially all the immunological properties that have been associated with this molecule by a number of more conventional approaches. The fact that these processes can be studied by DNA mediated gene transfer will make it possible to carry out a detailed molecular analysis of the determinants recognised by antibodies and by the T cell receptor. Specifically, it should now be possible to transform L-cells with hybrid H-2 genes, constructed in vitro, which contain DNA segments from, say, the H-2K^b and H-2D^b genes and to ask whether T_C-cell clones, either allospecific or H-2 restricted, can kill such target cells. In this way, we hope to map the determinants recognized by specific monoclonal antibodies and by the T-cell receptor.

ACKNOWLEDGMENTS

M.B, L.G. and E. W. were the recipients of fellowships from the Swiss National Fonds, the Helen Hay Whitney Foundation and the Deutsche Forschungsgemeinschaft respectively. Work in the laboratories of R.A.F., E.S. and A.R.M.T: was supported by the British Medical Research Council, work in the laboratory of H.F. by the Cancer Research Campaign and work by R.J. by the Imperial Cancer Research Fund. Cora O'Carroll is gratefully thanked for preparing this manuscript.

REFERENCES

1. Goldberg, M. (1979). Sequence analysis of Drosophila histone genes. Ph.D. thesis, Stanford University, Stanford, California, U.S.A.
2. Proudfoot, N. J. (1979). Nature 279, 376.
3. Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980). Nucl. Acids Res. 8, 127.
4. Grosschedl, R. and Birnstiel, M. L. (1980). Proc. Natl. Acad. Sci. U.S.A. 77, 1432.
5. Benoist, C. and Chambon, P. (1981). Nature 290, 304.
6. McKnight, S. L., Gavis, E. R. and Kingsbury, R. (1981). Cell 25, 385.
7. Grosveld, G. C., deBoer, E., Shewmaker, C. K. and Flavell, R. A. (1982). Nature 295, 120.
8. Dierks, P., van Ooyen, A., Mantei, M. and Weissmann, C. (1981). Proc. Natl. Acad. Sci. U.S.A. 78, 1411.
9. Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981). Cell 27, 279.
10. Banerji, J., Rusconi, S. and Schaffner, W. (1981). Cell 27, 299.
11. Chang, J. C. and Kan, Y. W. (1979) Proc. Natl. Acad. Sci. U.S.A. 76, 2886.
12. Trecartin, R. F., Liebhaber, S. A., Chang, J. C., Lee, K. Y., Kan, Y. W., Furbetta, M., Angius, A. and Cao, A. (1981). J. Clin. Invest. 918, 1012.
13. Moschonas, N., deBoer, E., Grosveld, F. G., Dahl, H. H.M., Wright, S., Shewmaker, C. K. and Flavell, R. A. (1981). Nucl. Acids Res. 9, 4391.
14. Temple, G. F., Chang, J. C. and Kan, Y. W. (1977). Proc. Natl. Acad. Sci. U.S.A. 74, 3047.
15. Ottolenghi, S., Lanyon, W. G., Williamson, R., Weatherall, D. J., Clegg, J. B. and Pitcher, C. S. (1975). Proc. Natl. Acad. Sci. U.S.A. 72, 2294.
16. Spritz, R. A., Jagadeeswaran, P., Choudary, P. V., Biro, P. A., Elder, J. T., deRiel, J. K., Manley, J. L., Gefter, M. L., Forget, B. G. and Weissman, S. M. (1981). Proc. Natl. Acad. Sci. U.S.A. 78, 2455.
17. Westaway, D. and Williamson, R. (1981). Nucl. Acids Res. 9, 1777.
18. Busslinger, M., Moschonas, N. and Flavell, R. A. (1981). Cell 27, 289.
19. Michaelson, J., Flaherty, L., Vitetta, E. and Poulik, M. (1977). J. Exp. Med. 145, 1066.
20. Stanton, T. H. and Hood, L. (1980). Immunogenetics 11, 309.
21. Klein, J. (1979). Science 203, 516.

22. Kvist, S., Bregerere, F., Rask, L., Cami, B., Garoff, H., Daniel, F., Wiman, K., Larhammar, D., Abastado, J. P., Gachelin, G., Peterson, P. A., Dobberstein, B. and Kourilsky, P. (1981). Proc. Natl. Acad. Sci. 78, 2772.
23. Steinmetz, M., Frelinger, J. G., Fisher, D., Hunkapiller, T., Pereira, D., Weissman, S. M., Uehara, H., Nathenson, S. and Hood, L. (1981) Cell 24, 125.
24. Grosveld, F. G., Dahl, H. H. M., deBoer, E. and Flavell, R. A. (1981). Gene 13, 227.
25. Doherty, P. C. and Zinkernagel, R. M. (1975). J. Exp. Med. 141, 502.
26. Weyand, C., Hammerling, G. J. and Goronzy, J. (1981). Nature 292, 627.
27. Zinkernagel, R. M. and Doherty, P. C. (1979). Adv. Immunol. 27, 51.
28. Lin, Y-L. and Askonas, B. A. (1981). J. exp. Med., 154, 225.
29. Sanderson, A. R. (1964). Nature 204, 250.

DNA Methylation and Globin Gene Expression

F. G. GROSVELD, D. KIOUSSIS, M. BUSSLINGER,[1] S. WRIGHT, J. HURST, E. VANIN,[2] AND R. A. FLAVELL[3]

Laboratory of Gene Structure and Expression
National Institute for Medical Research
London, England

INTRODUCTION

The regulation of gene expression in higher eukaryotes is likely to occur in several steps. Initiation of transcription is believed to be at least one of the major steps in gene control. Although transcriptional regulation is at present poorly understood, the interaction of the 5′ flanking region of a given gene with regulatory proteins is thought to play an important role in this process (Davison *et al.*, 1). Furthermore, the chromatin structure of transcribed genes differs from that of inactive genes, and as a consequence, changes in the chromatin configuration apparently precede transcriptional activation of a gene (for review, see Weisbrod, 2). Over the last years, the modulation of DNA methylation has been implicated to be yet another mechanism of gene regulation (for review, see Razin and Riggs, 3). Detailed studies of the tissue specific pattern of DNase I sensitivity and DNA methylation showed that actively transcribed genes were DNase I sensitive and undermethylated, whereas inactive genes were DNase I insensitive and usually, but not always, methylated (Weintraub and Groudine, 4; Stalder *et al.*, 5; McGhee and Ginder, 6; Mandel and Chambon, 7; van der Ploeg and Flavell, 8). This suggested that DNase I sensitivity is a necessary condition for gene expression,

[1] Present address: Institut fur Molekularbiologie II der Universität Zürich, CH-8093 Zürich, Switzerland.

[2] Present address: Department of Biochemistry, Ohio State University, 484 N-12th Street, Columbus, Ohio 43210.

[3] Present address: Biogen Research Corporation, 14 Cambridge Center, Cambridge, Massachusetts 02142.

TRANSFER AND EXPRESSION OF EUKARYOTIC GENES

ISBN 0-12-284650-8

whereas a low level of DNA methylation might be necessary, but not a sufficient condition for gene expression in higher eukaryotes. In this article, we describe the relationship between DNase sensitivity and DNA methylation and gene expression *in vivo* for the β-globin genes in a case of Dutch γβ-thalassemia (van der Ploeg *et al.*, 9), which has an expressed normal globin and a nonexpressed rearranged globin locus. We report here that the unexpressed β-globin gene from the mutant chromosome, which is not expressed in cells that actively express the β-globin gene from their unaffected chromosome, is DNase I resistant and hypermethylated. Second, we address the issue as to which C residues must be methylated with respect to the human γ-globin gene in order to inhibit globin gene expression.

We report here that methyl C residues from position −760 in the 5′ flanking region to position +100 of the γ-globin gene prevent transcription.

RESULTS

SEQUENCE ANALYSIS AND TRANSCRIPTION OF THE γβ-THALASSEMIA GLOBIN GENES

In a Dutch case of γβ-thalassemia (van der Ploeg *et al.*, 9), one of the β-globin alleles is normal and expressed, whereas the second allele is not expressed. The nonexpressing locus contains the entire β-globin gene and its normal 5′ and 3′ flanking sequences, whereas the γ- and δ-globin genes are deleted (Fig. 1). Several possible explanations have been given for the inactivity of the β-globin gene in the abnormal locus; the thalassemia might be the result of a double mutation, i.e., the large deletion might be irrelevant, but a second mutation, near or in the β-globin gene, may render it inactive. Alternatively, the disease might be the result of an altered chromatin structure, either because regulatory sequences lying far from the β-globin gene are deleted or because the β-globin has been juxtaposed to a chromatin structure that is incompatible with gene expression in erythropoietic cells.

To distinguish between these possibilities, the β-globin genes from a Dutch γβ-thalassemic patient were isolated from a cosmid library (Grosveld *et al.*, 10). Three clones contained the mutant 4.2 kb *Eco*RI fragment, whereas a fourth clone contained the normal 5.2 kb *Eco*RI fragment, which contains the 5′ end of the β-globin gene (Fig. 1). The *Bgl*II fragments, containing the β-globin gene from two of the cosmid recombinants (one normal, one mutant), were subcloned for DNA sequence analysis. The

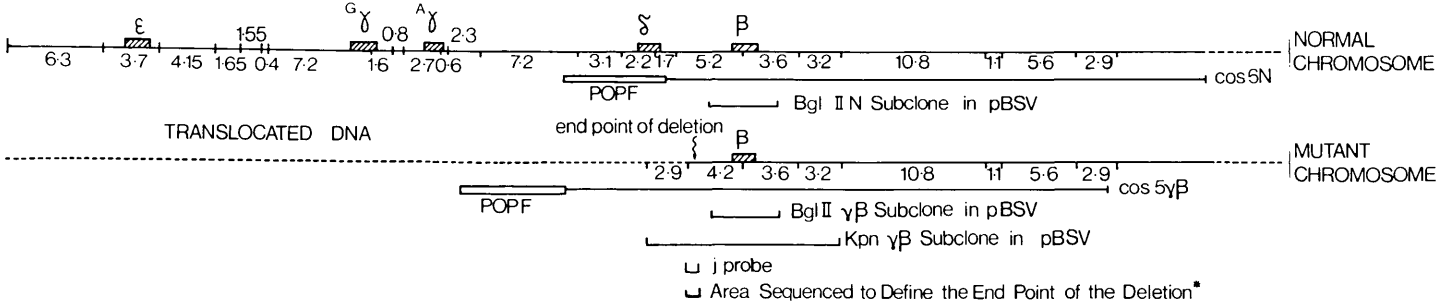
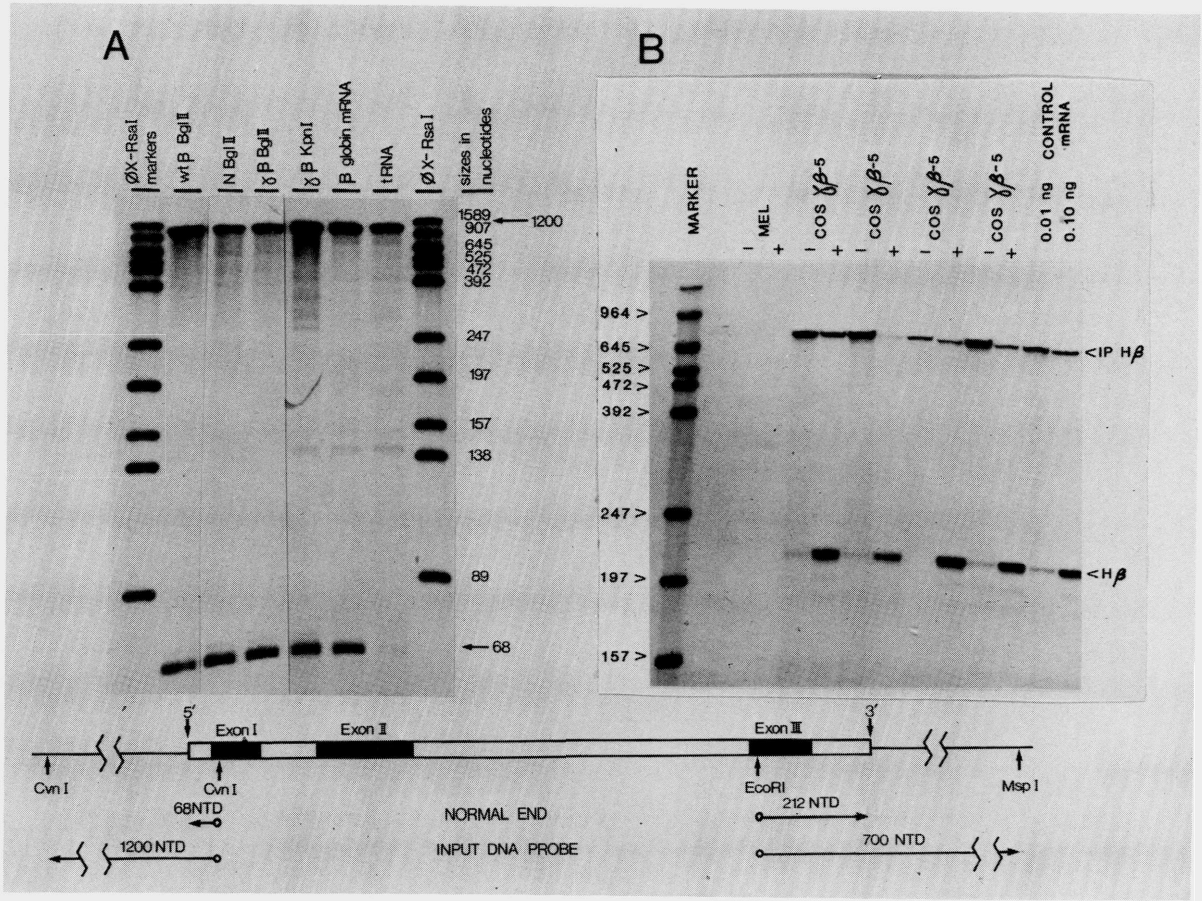


Fig. 1. Physical map of the normal and $\gamma\beta$-thalassemic β-globin region. Each map shows the position of the genes in relation to the *Eco*RI restriction map. The relevant clones (cos 6N and cos $5\gamma\beta$) were isolated from a cosmid library constructed from blood DNA of a $\gamma\beta$-thalassemic patient and the cosmid vector pOPF. The brackets underneath indicate the fragments that were subcloned in the expression vector pBSV and the location of the junction probe (j). The arrow indicates the deletion end-point in the affected locus.


A
ØX-RsaI markers
wt β BglII
N BglII
δβ BglII
δβ KpnI
β globin mRNA
tRNA
ØX-RsaI
sizes in nucleotides
1589
907
645
525
472
392
247
197
157
138
89
68
1200
B
MARKER
MEL
COS δβ-5
CONTROL mRNA
0.01 ng
0.10 ng
964
645
525
472
392
247
197
157
<IP Hβ
<Hβ
5′
Exon I
Exon II
Exon III
3′
Cvn I
Cvn I
EcoRI
Msp I
68NTD
1200 NTD
NORMAL END
INPUT DNA PROBE
212 NTD
700 NTD

three exons, the promoter regions, and all the exon–intron boundaries of each of the genes were sequenced by the method of Maxam and Gilbert (11). Both genes showed a sequence identical to that of a normal β-globin gene in these regions, i.e., from a position 100 bp upstream of the cap site through the three coding segments to a position 50 bp downstream from the poly(A) addition site (data not shown). Consequently, these data exclude the possibility that a point mutation in the coding sequence is responsible for the thalassemic phenotype. It leaves the possibility that the gene is transcriptionally nonfunctional. To test this possibility, subclones of the β-globin genes were made in the transient expression vector pBSV (Grosveld *et al.*, 12) and introduced into HeLa cells and the complete cosmids were introduced into MEL cells by $CaPO_4$ precipitation (Wigler *et al.*, 13). The RNA from the HeLa cells and from the MEL cells (before and after globin induction with HMBA) were analyzed for the presence of β-globin mRNA using S1 nuclease mapping (Weaver and Weissmann, 14). The following pBSV recombinants were constructed (Fig. 1):

1. A *Bgl*II fragment from both the normal and the mutant chromosome that contains the β-globin gene and about 3 kb of flanking sequences was cloned in the *Bam*HI site of pBSV.
2. A 14 kb *Kpn*I fragment from the mutant chromosome was cloned into the *Kpn*I site of pBSV. This fragment contains additional flanking DNA in both 5′ and 3′ direction, including 2.5 kb of the sequences that have been juxtaposed to the 5′ side of the β-globin gene by the deletion. Insertion into the *Bam*HI or *Kpn*I site of pBSV does not affect the effi-

---

Fig. 2. (A) Transient expression of the normal and γβ-thalassemic β-globin gene in HeLa cells. The *Bgl*II and *Kpn*I restriction fragments indicated in Fig. 1 were subcloned in the expression vector pBSV by standard procedures. 15 μg DNA from each of these subclones were mixed with 25 μg of salmon sperm DNA and used to transform half confluent 100 mm dishes of HeLa cells by the calcium phosphate method. After 16 hr the medium was changed; the cells were grown for 36 hr and collected; the RNA was isolated by the LiCl-urea method. The β-globin present in the HeLa RNA was detected by S1 mapping and end-labeled DNA probes. The lower part of the figure shows the 5′ labeled 1200 nucleotide probe, which contains 68 nucleotides from the 5′ end of the gene that are protected against S1 nuclease digestion of the RNA–DNA hybrids. The panel shows the 5′ end analysis of the subclones indicated in Fig. 1, plus two controls, a β-globin subclone from normal DNA and *in vivo* produced mRNA from reticulocytes. The labeled marker in ∅X174 × *Rsa*I, and the numbers are the length of the marker fragments in nucleotides. (B) S1 nuclease analysis of human β-globin mRNA in MEL cell transformants. MEL cell transformants containing the cosmid cos HG γβ5 were grown in the presence (+) or absence (−) of 3 mm HMBA. The human β-globin mRNA was detected by S1 mapping essentially as described in Fig. 2A. The probe was a 3′ β-globin fragment illustrated in the bottom half of the panel.

ciency of transcription. A fourth plasmid containing the *Bgl*II β-globin fragment from a normal individual was used as a control.

Figure 2A shows the protected fragments after S1 nuclease digestion, when a 5′ end labeled 1200 bp *Cvn*I probe from the 5′ end of the β-gene is used. In all cases, the same amount of a 68 bp fragment is protected which represents the 5′ end of mature β-globin mRNA. Similar results are found when each of the splice junctions and the 3′ end of the mRNA were analyzed. Again, no differences were found between mature β-globin mRNA and the RNA of each of the transformants (data not shown). Fig. 2B shows the protected fragments after S1 nuclease digestion when a 3′ end labeled 700 bp *Eco*RI/*Msp*I probe from the 3′ end of the β-globin gene is used, for the analysis for the RNA from stable MEL cell transformants before and after induction. The levels of induction of the transfected mutant β-globin gene (average 50-fold) are the same as the induction levels of a normal β-globin gene in MEL cells (not shown, Flavell *et al.*, 15). These results show that by these methods both the normal and the mutant gene from the patient are inducible and transcribed equally to give an mRNA indistinguishable from β-globin mRNA from reticulocytes. In addition, the levels of mRNA are the same for the β-globin genes from the "mutant" and normal chromosomes. Consequently, the transcription and sequence data indicate that the phenotype observed in the patient cannot be caused by a promoter defect, a "splice" mutation or a mutation in the coding sequence.

DNASE I SENSITIVITY AND METHYLATION OF THE γβ DNA SEQUENCE

Because transcriptional activity of a region of the genome has been associated with an increased sensitivity to DNase I (Weintraub and Groudine, 4) and a hypomethylation of the DNA (McGhee and Ginder, 6; Mandel and Chambon, 7; van der Ploeg and Flavell, 8), these assays were applied to the rearranged β-globin locus. We tested the susceptibility to DNase I of fetal liver and fetal brain chromatin from a normal fetus and of fetal liver chromatin of a γβ-thalassemic fetus.

As expected, in the nuclei of normal fetal liver samples the γ-globin genes (expressed genes) are very sensitive to DNase I. In contrast, 80% of the junction sequences (j probe; see Fig. 1) in the normal liver are still present when > 80% of the γ-globin sequences have been digested. In normal fetal brain tissue, both sequences are resistant (not shown). When a γβ-thalassemic fetal liver sample is treated with DNase I, the normal locus shows a "sensitive" pattern for the normal β- and δ-globin genes, whereas the mutant locus β-globin gene is resistant (Fig. 3, the 4.0 kb and 2.2 kb bands versus the 4.2 kb band). This suggests that the "inactive"

configuration of the junction sequences in normal erythropoietic tissue is transferred to the β-globin gene in the mutant locus.

In addition to the DNase I sensitivity, the methylation patterns of the mutant and normal locus were analyzed, using the restriction enzymes *Hpa*II and *Msp*I. *Hpa*II will not cleave the sequence C^{m}CGG or ^{m}C^{m}CGG, whereas *Msp*I will not cleave the sequence ^{m}CCGG or ^{m}C^{m}CGG. Because most of the methylated C residues (^{m}C) occur in the dinucleotide ^{m}CG, the difference in the cleavage pattern of these two enzymes provides a measure for the methylation of a particular region of the DNA. Southern blots of the DNA digests were hybridized to the same probes as in the DNase I experiments. Hybridization with a γ-globin probe shows the expected hypomethylation pattern for the fetal liver and the hypermethylation pattern for fetal brain (van der Ploeg and Flavell, 8). This is consistent with the fact that active globin genes are hypomethylated in erythroid tissues. Figure 3 shows that hybridization of the β-probe to the fetal liver DNA shows that the *Hpa*II sites at the 3′ side of the β-globin gene (there are no sites directly 5′) are hypomethylated by the presence of a 20 kb *Hpa*II fragment. In the brain, all of this signal has shifted to high molecular weight DNA. Hybridization to the fetal liver DNA of the patient shows two complete digest bands with *Msp*I, an 11.0 kb band for the normal locus and a 4.7 kb band for the mutant locus (see maps in Fig. 3). A partial digest band of 5.3 kb is visible, which is probabily caused by the failure of *Msp*I to cut certain C^{m}CGG sites (15a). The *Hpa*II digest shows the normal 20 kb band and a high molecular weight signal, indicating that the 3′ β-globin sites in the normal locus are hypomethylated. Hybridization with the junction probe shows the expected 0.6 kb *Msp*I digest band and a 1.1 kb partial digest band (for the same reasons as described above). The *Hpa*II digest of all the tissues examined fail to detect any low molecular weight band, indicating that the *Msp*I sites adjacent to the β-globin gene in the γβ-thalassemia locus are methylated. This is confirmed by double digest experiments (data not shown).

We conclude that the mutant locus has become hypermethylated, which, as with the DNase I sensitivity assays, indicates that the mutant locus is present in a transcriptionally inactive state. These observations postulate a cis influence of sequences far from the β-globin gene. Whether the effect is exerted in this patient by the removal of regulating sequences or by the addition of actively suppressing sequences upstream from the β-globin gene is presently unclear. Either possibility could block the normal progression of globin gene expression during normal erythropoiesis or alter the ability of this chromosomal region to be expressed. Either way the net result is a position effect similar to those found in *Drosophila*.

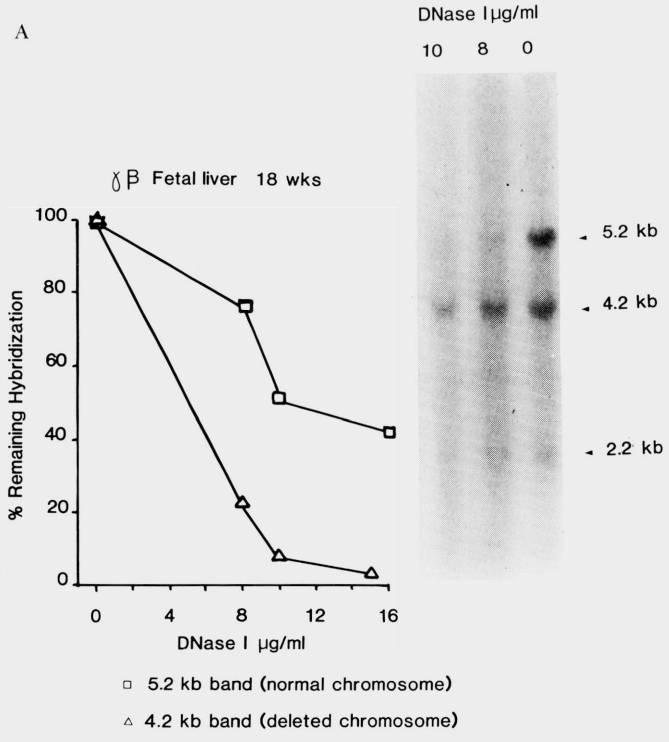


Fig. 3. (A) The DNase I sensitivity of the normal and the mutant β-globin genes. Nuclei and DNA samples were isolated from an 18-week fetal liver of a $\gamma\beta$-thalassemic patient and treated essentially as described. The panel on the left shows the plot obtained from the complete blot, including lanes not shown in the panel on the right, which is the autoradiograph of three lanes obtained with the 5.2 kb β-globin probe (Fig. 1) and increasing concentration of DNase I. The numbers indicate the fragment length. (B) Methylation of the translocated and normal $\gamma\beta$-thalassemia locus. DNA was isolated from 18-weeks fetal liver of a normal and a $\gamma\beta$-thalassemic patient and brain of a normal 18 weeks fetus. 10 μg of DNA was digested with *Msp*I or *Hpa*II, electrophoresed through a 0.5% agarose gel and blotted. The filters were hybridized to a 0.6 kb 3′ γ-globin fragment, a 5.2 kb 5′ β-globin fragment, and a 0.8 kb junction probe. The numbers indicate the fragment size of a λ-X-*Hin*dIII marker. The control lanes contained 10 μg of mouse DNA mixed in with 1 ng of cosmid DNA.

IN VITRO GENE METHYLATION

To test the direct relationship between methylation and gene expression, we used an *in vitro* methylation technique developed by Stein *et al.* (16). They showed that hemimethylated DNA can be synthesized *in vitro* using 5-methyldeoxycytidine (5-methyl dCTP) and the three other standard deoxynucleotide triphosphates.

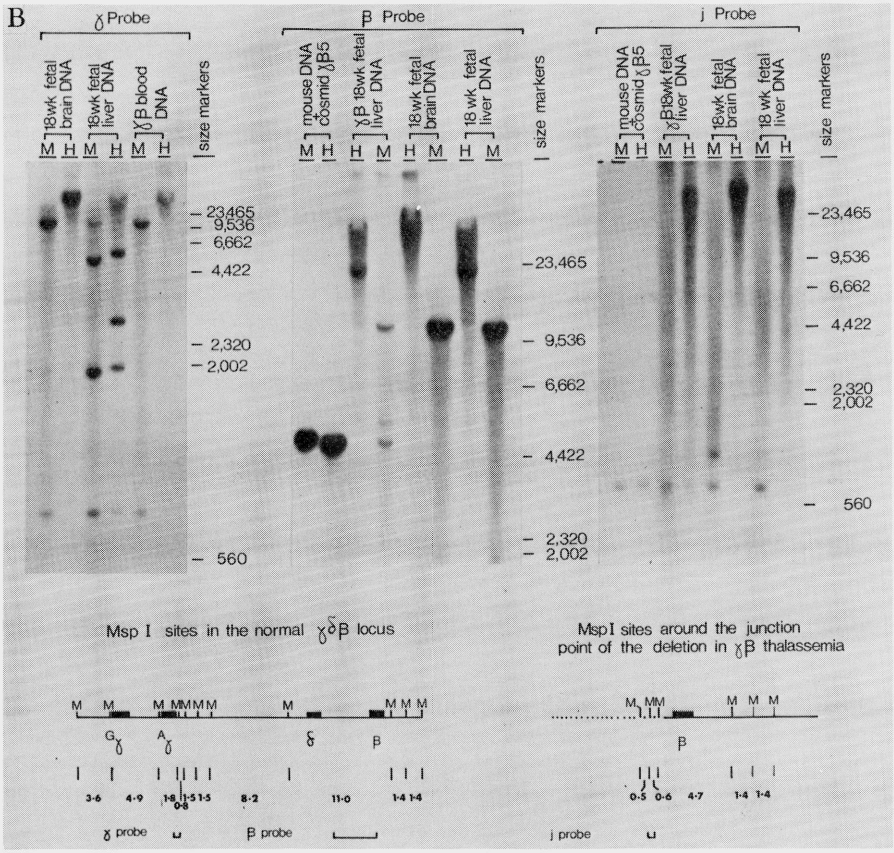


Fig. 3. *Continued*

The resulting DNA molecules contain in the newly synthesized DNA strand only methyl-C residues, while the template strand remains unmodified. When introduced into mouse L cells, the template strand of this hemimethylated DNA is methylated by cellular methyltransferases at most or all of their recognition sites, that is, predominantly at CpG residues (Stein *et al.*, 17). The methylation pattern, which is thereby established, is inherited from one cell generation to the next and can be studied in the DNA of stable transformants by restriction and Southern blot analysis.

To study the effect of DNA methylation on globin gene expression, we have modified this approach as shown in Fig. 4. The human γ-globin gene and some 1340 bp of its 5′ flanking region have been cloned in both orientations into the replicative form (RF) of M13mp8 (Messing and

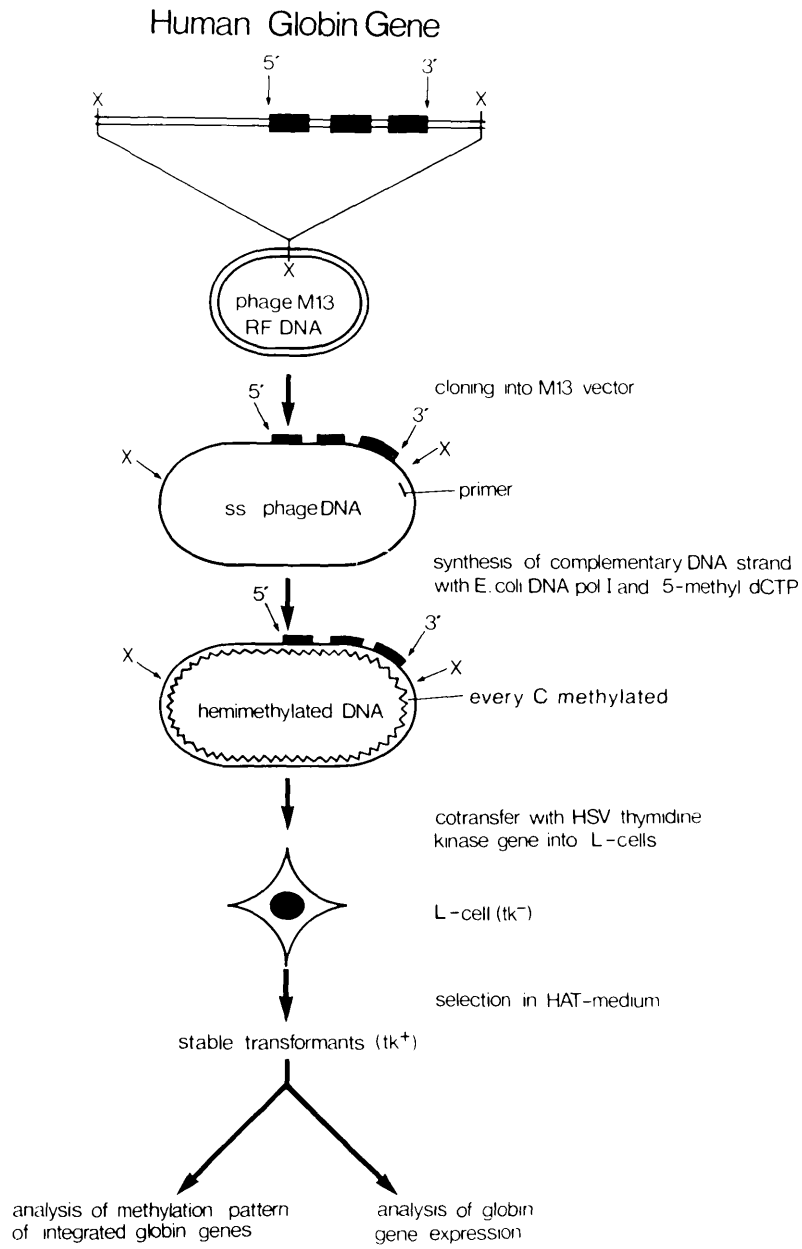


Fig. 4. Procedure used to study the effect of DNA methylation on globin gene expression. The restriction enzyme used for M13 cloning is denoted by X, i.e., *Hin*dIII for the human γ-globin gene and *Pst*I for the β-globin gene. The plus strand of the recombinant

Vieira, 18) to give recombinants Mγ1 and Mγ2. Single-stranded DNA of Mγ1 and Mγ2 was converted into double-stranded DNA by *E. coli* DNA polymerase I, dATP, dGTP, dTTP, and 5-methyl dCTP instead of normal dCTP. This hemimethylated DNA was introduced into mouse L cells by co-transformation with the Herpes Simplex Virus (HSV) thymidine kinase *(tk)* gene as a selective marker. The transformed *tk*-positive cells were then analyzed by Southern blotting for the inheritance of methyl-C residues in the integrated γ-globin gene and by S1 mapping for the presence of globin gene transcripts. In order to have an internal positive control for globin gene transcription in these cells, we have introduced the unmethylated human β-globin gene, also cloned on M13 (Mβ1) together with the methylated γ-globin gene into the same L cells.

This *in vitro* methylation technique has the following advantage over direct *in vitro* methylation with bacterial or eukaryotic methylases. First, it allows us to methylate all and not only a few of the CpG residues in the γ-globin gene. Second, and more importantly, segmental methylation patterns can be generated by using DNA primers of different length and map position for the synthesis of hemimethylated DNA by the large fragment of *E. coli* DNA polymerase I (Klenow polymerase). Since mC residues are only introduced into newly synthesized DNA, the primer region stays unmethylated on both DNA strands and remains so even in the L cell, which has little or no *de novo* methylase activity (Pollack *et al.*, 19).

METHYLATION IN THE 5′ REGION OF THE β-GLOBIN GENE SUPPRESSES TRANSCRIPTION

We have synthesized a number of different hemimethylated M13-γ-globin molecules with different primers (Fig. 5) in the following manner:

1. Using a short M13-specific DNA primer (SP.6), we first synthesized Mγ1 DNA, which was totally hemimethylated in both the globin and M13 sequences. Transformation of L cells with this DNA lead to the isolation of four different cell clones M1-M4.
2. Using the *Hin*dIII insert of the γ-globin gene as a primer (A, Fig. 5),

---

phages Mγ1 and Mβ1 contains the anticoding DNA strand of the γ- and β-globin gene, whereas phage Mγ2 contains the coding strand of the γ-globin gene. Hemimethylated DNA was synthesized and introduced into mouse L cells as described by Stein *et al.* (16). Transformed *tk*-positive cells were analyzed by Southern blotting for the inheritance of DNA methylation in the integrated globin gene and by S1 mapping for the presence of globin gene transcripts.

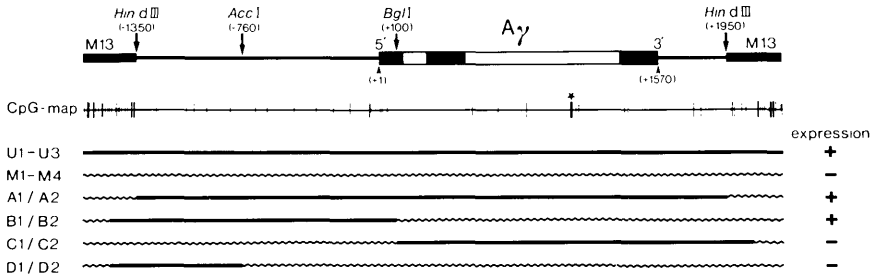


Fig. 5. Summary of the various methylation patterns and their influence on γ-globin gene expression. The data for the CpG map are taken from Slightom *et al.* (20, from position −56 to +1410), from Shen *et al.* (21, remaining γ-globin gene DNA sequences), and from van Wezenbeek *et al.* (22, Ml3 DNA sequences). Thick bars indicate multiple CpG sites. The asterisk denotes the position of the DNA sequence consisting of a run of 5 CpG dinucleotides. Unmodified DNA seqences are shown as straight lines, whereas methylated DNA regions are indicated as wavy lines. The abbreviations for the cell clones analyzed are shown to the right.

an Mγ1 DNA was synthesized that is only hemimethylated in the M13 sequences. Two cell clones A1 and A2 were isolated for further analysis.

3. A *Hin*dIII/*Bgl*II DNA fragment (primer B) was used to synthesize an Mγ1 DNA; this is only hemimethylated downstream from position +100 in the structural gene and the M13 sequences. Two cell clones B1 and B2 were isolated for further analysis.

4. The reciprocal experiment to number 2 was done by using a *Bgl*I fragment as a primer (C, Fig. 5). This results in a molecule which has the 5′ flanking sequences from −1350 to −100 and the M13 vector hemimethylated, but leaves the structural gene unmethylated. Two clones C1 and C2 were analyzed.

5. Finally, a *Hin*dIII–*Acc*I fragment was used, which results in a completely methylated Mγ1 except the sequences between −1350 (*Hin*dIII) and −760 (*Acc*I). Two clones D1 and D2 were analyzed.

As a control three cell clones (U1-U3) were isolated that were transfected with completely unmethylated Aγ1. Southern blot analysis of all the clones showed that inheritance of the methyl groups is almost 100% at the four *Hha*I and two *Hpa*II sites of the γ-globin genes as methylated DNA is resistant to *Hha*I and *Hpa*II cleavage. RNA from all of these cell clones was assayed for the presence of γ- and β-globin mRNA by S1 mapping using 5′ or 3′ specific probes (Figs. 6 and 7). All of the isolated cell clones expressed the control β-globin gene (Figs. 6 and 7; data not shown). The fact that the cell clones U1–U3 and A1, A2, but not M1–M4 express the

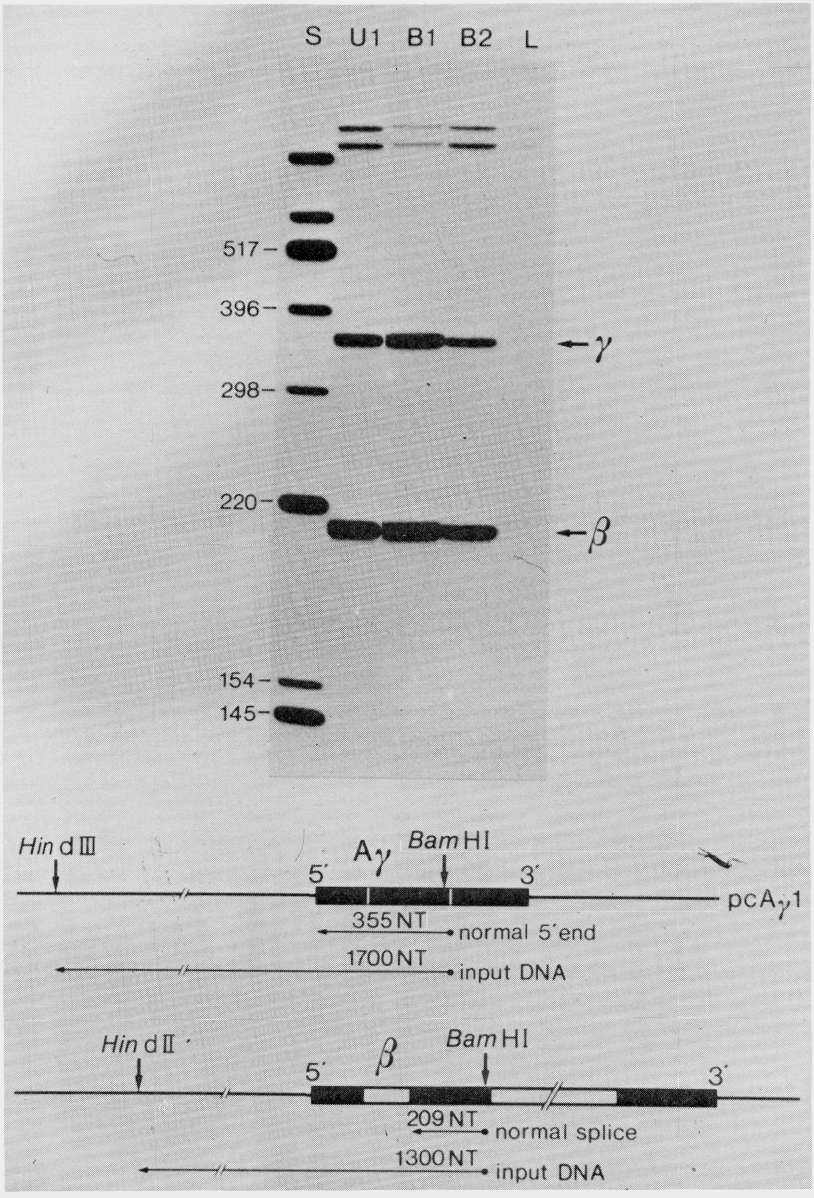


Fig. 6. Globin gene expression. Total RNA from cell clones U1, B1, and B2 was analyzed by S1 mapping. RNA from L cells was used as a control (land L) and labeled pAT153 DNA digested with *Hin*fI as size marker (lane S).

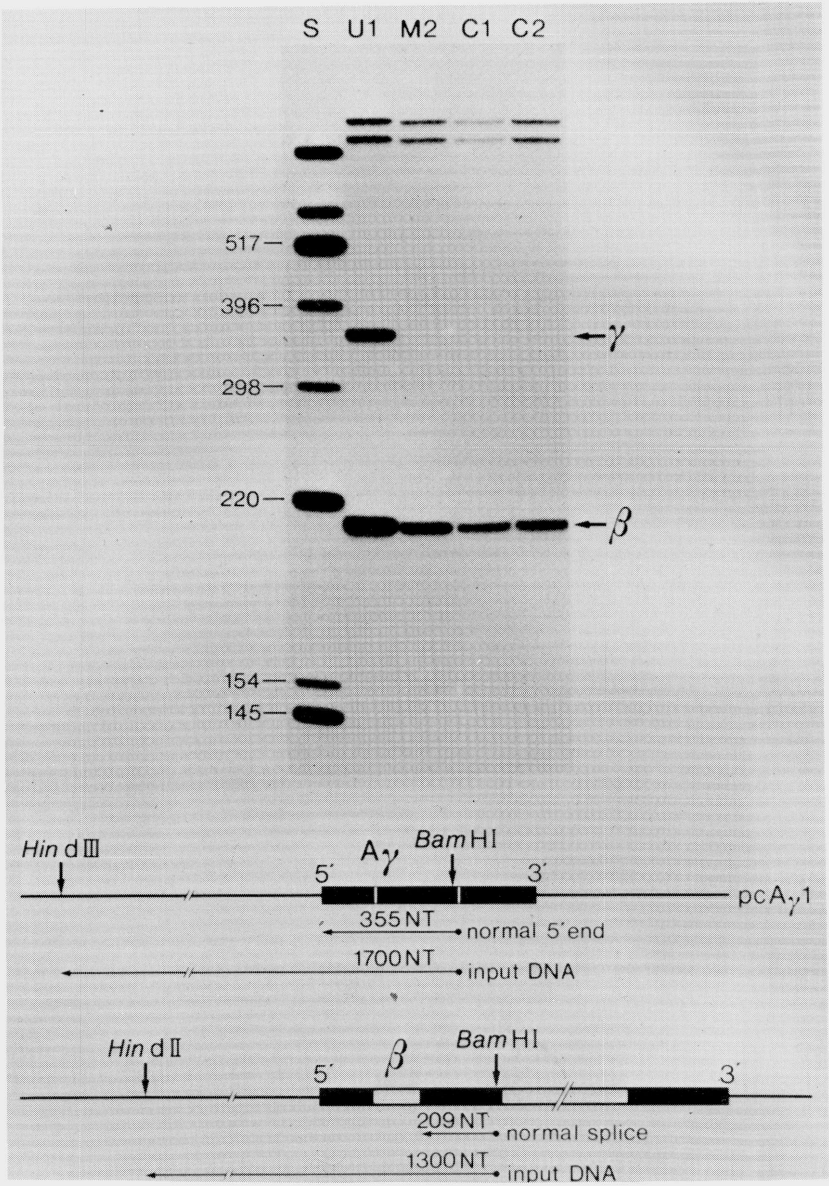


Fig. 7. Globin gene expression. Total RNA from cell clones U1 (containing unmethylated γ-globin genes), M2 (containing fully methylated γ-globin genes), C1, and C2 was analyzed by S1 mapping with the two DNA probes shown in the bottom part as arrows below the map of the respective globin gene. Labeled pAT153 DNA digested with *Hinf* I ws used as a size marker (lane S).

γ-globin gene (Fig. 5; data not shown) shows that methylation of the CpG residues (M1–M4) of the human γ-globin DNA and its flanking DNA sequences is sufficient to suppress globin gene transcription in L cells. The fact that A1 and A3 express the γ-globin gene excludes the possibility that the mere presence of many methyl-C residues anywhere would be sufficient for suppression but that the effect is localized within the γ-globin *Hin*dIII insert.

The analysis of the cell clones B1, B2 (Fig. 6, positive for γ-mRNA) and C1, C2 (Fig. 7, negative for γ-mRNA) shows that the suppressive effect of DNA methylation is localized in the 5′ region of the gene and that transcription can proceed through a methylated γ-globin gene. Cell clones D1, D2 negative γ-mRNA; data not shown) excludes that the sequences far upstream from the promoter are important, and we conclude from all experiments together (Fig. 5) that the methyl-C residues in the DNA sequences from position −760 to +100 of the γ-globin gene prevent globin transcription, whereas methyl-C in the DNA sequences further upstream (−1350 to −760) or in the structural gene and 3′ coding sequences (+100 to +1950) have no effect on globin gene expression. There are 11 CpG dinucleotides in the DNA sequence from −760 to +100 (Fig. 5), and some or all of these may be involved in the suppression of transcription. The effect of the methyl-C residues could be to alter DNA–protein interactions within the γ-globin gene promoter. However, we have not established this yet and further experiments will be necessary to determine precisely which methyl-CpG sites in the 5′ region of the γ-globin gene suppress transcription *in vivo.*

SUMMARY

We have studied the effect of chromatin structure on human globin gene expression. First, we have shown that the β-globin gene present on the chromosome that carries the γγδ-globin gene deletion in a Dutch βδ-thalassemic patient is identical to the normal β-globin gene with respect to DNA sequence and its transcription in HeLa cells and Friend cells. However, studies of DNase I sensitivity and the extent of DNA methylation shows that the affected β-globin gene is in an inactive configuration *in vivo.* Second, we have used an *in vitro* methylation technique to study the expression of the γ-globin gene in mouse L cells. The data indicate that DNA methylation in the 5′ region of the gene might play a direct role in the regulation of expression.

ACKNOWLEDGMENTS

We are grateful to F. Wilson for excellent technical assistance and to Professor M. Birnstiel for providing the opportunity to finish this work in his laboratory. This work was supported by the British Medical Research Council. M. B. was the recipient of a fellowship from Schweizerischer Nationalfonds; D. K. was supported by a long-term EMBO fellowship and E. V. by NIH Grant GM 20069 to O. Smithies.

REFERENCES

1. Davison, B. L., Egly, J. -M., Mulwihill, E. R., and Chambon, P. (1983) *Nature (London)* **301,** 608.
2. Weisbrod, S. (1982) *Nature (London)* **297,** 289.
3. Razin, A., and Riggs, A. D. (1980) *Science* **210,** 604.
4. Weintraub, H., and Groudine, M. (1976) *Science* **93,** 848.
5. Stalder, J., Groudine, M., Dodgson, J. B., Engel, J. D., and Weintraub, H. (1980) *Cell* **19,** 973.
6. McGhee, J. D., and Ginder, G. D. (1979) *Nature (London)* **280,** 419.
7. Mandel, J. L., and Chambon, P. (1979) *Nucleic Acids Res.* **7,** 2081.
8. van der Ploeg, L. H. T., and Flavell, R. A. (1980) *Cell* **19,** 947.
9. van der Ploeg, L. H. T., Konings, A., Oort, M., Roos, D., Bernini, L. F., and Flavell, R. A. (1980) *Nature (London)* **283,** 637.
10. Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M., and Flavell, R. A. (1982) *Nucleic Acids Res.* **10,** 6715.
11. Maxam, A. M., and Gilbert, W. (1980) *In* "Methods in Enzymology" (L. Grosman and K. Moldave, eds.), vol. 65, Part 1, p. 499. Academic Press, New York.
12. Grosveld, G. C., deBoer, E., Shewmaker, C. K., and Flavell, R. A. (1982) *Nature (London)* **295,** 120.
13. Wigler, M., Sweet, R., Sim, G. K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S., and Axel, R. (1979) *Cell* **16,** 777.
14. Weaver, R. F., and Weissmann, C. (1979) *Nucleic Acids Res.* **6,** 1175.
15. Flavell, R. A., Grosveld, F. G., Grosveld, G. C., Wright, S., Busslinger, M., deBoer, E., Kioussis, D., Mellor, A. L., Golden, L., Weiss, E., Hurst, J., Bud, H., Bullman, H., Simpson, E., James, R., Townsend, A. R. M., Taylor, P. M., Schmidt, W., Ferluga, J., Leben, L., Santamaria, M., Atfield, G., and Festenstein, H. (1982). *UCLA Symp. Mol. Cell. Biol.* **26,** 17.

15a. Busslinger, M., deBoer, E., Wright, S., Grosveld, F. G., and Flavell, R. (1983) *Nucleic Acids Res.* **11,** 355.

16. Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A., and Cedar, H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79,** 61.
17. Stein, R., Razin, A., and Cedar, H. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79,** 3418.
18. Messing, J., and Vierira, J. (1982) *Gene* **19,** 269.
19. Pollack, V., Stein, R., Razin, A., and Cedar, H. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77,** 6463.
20. Slightom, J. L., Blechl, A. E., and Smithies, O. (1980) *Cell* **21,** 627.
21. Shen, S.-H., Slightom, J. L., and Smithies, O. (1981) *Cell* **26,** 191.
22. van Wezenbeek, P. M. G. F., Hulsebos, T. J. M., and Schoenmakers, J. G. G. (1980) *Gene* **11,** 129.

The sequence GGC^{m}CGG is resistant to *Msp*I cleavage

M.Busslinger*, E.deBoer, S.Wright, F.G.Grosveld and R.A.Flavell+

Laboratory of Gene Structure and Expression, NIMR, Mill Hill, London NW7 1AA, UK, *Institut für Molekularbiologie II, University of Zürich, Zürich, Switzerland and +Biogen Research Corp., 241 Binney Street, Cambridge, MA 02142, USA

Received 22 March 1983; Accepted 11 May 1983

ABSTRACT

MspI essentially fails to cut the sequence GGC^{m}CGG at enzyme concentrations which give total digestion of CCGG, C^{m}CGG and GGCCGG sites. This result explains why certain sites in mammalian DNA are resistant to both MspI and HpaII and shows that this results from an idiosynchracy of MspI rather than a novel form of DNA methylation at this site in mammalian cells.

INTRODUCTION

The analysis of DNA methylation in eukaryotic cells has been greatly aided by the use of restriction enzymes which are sensitive to the presence of 5 methyl Cytosine (^{m}C) residues in their recognition site. Thus, Bird and Southern (1) showed that HhaI (GCGC) and HpaII (CCGG) would not cut Xenopus rDNA if the internal C residue was methylated and that some, but not all of these sites were methylated in genomic rDNA. The failure of these enzymes to cut at a given site indicated the presence of ^{m}C in that sequence or the absence of that site in the DNA because of a polymorphic difference in primary DNA sequence. It became easier to discriminate between these two possibilities for CCGG sites when Waalwijk and Flavell (2) showed that MspI would cut at CCGG, irrespective of ^{m}C at the internal C residue; MspI could, therefore, be used to show the presence of that site and HpaII to diagnose the presence of ^{m}C residues.

An apparent exception to this rule was fortuitously observed in a study of the methylation of the human β-related globin genes (3). They showed that MspI failed to cut at two CCGG sites present in the 5' flanking regions of the Gγ- and Aγ-globin genes, respectively. Since this site was cut after cloning the γ-globin DNA in a phage vector (4), this effect had to be explained by modification of the DNA in human cells. Furthermore, since it was shown (5) that MspI could not cut the sequence ^{m}CCGG, it was suggested that the sites in the human γ-globin genes were modified at the external C residue (3, 4). Others have also noted such sites (8). An alternative explanation would be,

however, that those exceptional MspI sites are a subset of the CCGG sites which are uncut when methylated at the internal C residue as a result of some unique property of that site. We show here that these unique sites have the consensus sequence GGCCGG and that when the internal C residue is methylated (GGC^{m}CGG) MspI cuts this site with great difficulty. That MspI does not cut at these sites is therefore an interesting enzymological artefact and does not necessarily indicate $^{m}C^{m}CGG$ sites in mammalian DNA.

MATERIALS AND METHODS

Enzymes and reagents

Restriction enzymes, E. coli DNA polymeraseI, HpaII methylase and T4 DNA ligase were purchased from N.E.Bio Labs. α-^{32}P-dATP (3000Ci/mMol) and α-^{32}P-dCTP (300 Ci/mMol) was from the Radiochemical Centre, Amersham, 5-methyl dCTP from P-L Biochemicals, dATP, dCTP, dTTP from Boehringer Mannheim and dextransulfate from Pharmacia.

M13 cloning and in vitro DNA methylation

The human Aγ-globin gene was isolated on a 3300bp long HindIII fragment from cosmid clone HG25 (9) and ligated into the HindIII site of the replicative form of phage M13mp8 (10). The human α1-globin gene was inserted into M13mp8 as a 1500bp long PstI fragment isolated from DNA clone pRBα1 (11). These ligated DNAs were used to transfect E.coli JM103 and single stranded DNA of the recombinant phages named Mγ1 and Mα1, respectively, was isolated according to Messing and Vieira (10). These phage DNAs were used as templates for the in vitro synthesis of hemimethylated DNA essentially as described by Stein et al. (6) (Fig. 1). An M13-specific DNA primer (SP16) was heat-denatured and added to a reaction mixture (100μl) containing 2μg of single stranded phage DNA, 66mM Tris-HCl (pH7.5), 6.6mM MgCl2, 10mM dithiothreitol and 50μM each of rATP, dATP, dGTP, dTTP and 5 methyl dCTP (5-methyldeoxycytidine-5'-triphosphate). The complementary strand was synthesized at 30^{o}C for 1hr. with E.coli DNA polymeraseI (25 units) in the presence of T4 DNA ligase (400 units). The DNA was then extracted with phenol and separated on a 0.8% low-melting agarose gel in the presence of EtBr (0.5μl/ml). The covalently closed circular DNA molecules were isolated from the agarose by phenol extraction and ethanol precipitation and used directly for transformation of L-cells. Methylation of DNA with HpaII methylase was performed according to the instructions of the manufacturer (N.E.Biolabs) and methylation with Bacillus phage methylase essential as in Jentsch et al.(5). Under the conditions used (1hr. incubation) only the outer C residue of CCGG

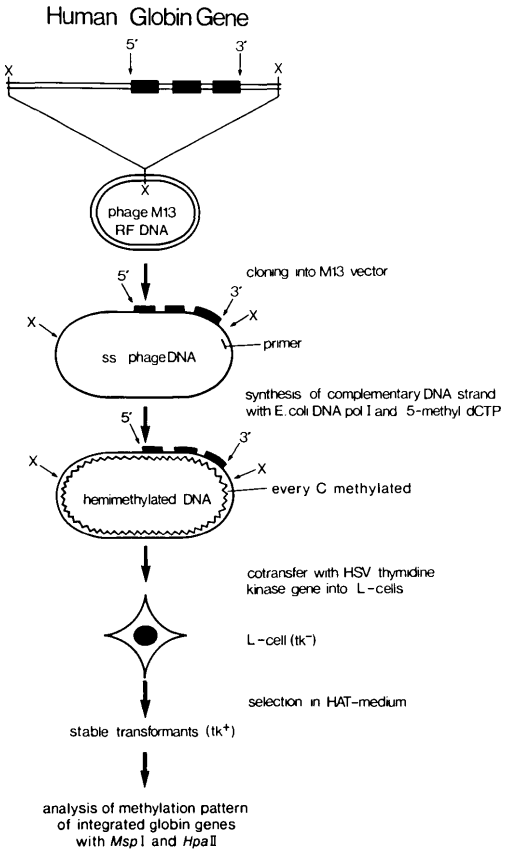


Fig. 1. Procedure used to study the inheritance of DNA methylation.
X denotes the restriction enzyme used for M13 cloning, i.e. PstI for the α1-globin gene and HindIII for the Aγ-globin gene; the resulting phages are named Mα1 and Mγ1, respectively. For experimental details see Methods section.

is methylated. This will be described in detail elsewhere (U. Gunthert et al., in preparation).

Cell Culture and Transformation

Ltk⁻ mouse cells obtained from R. Axel were maintained in Dulbecco's modified Eagle's (DME) medium supplemented with 5% new born calf serum. These cells were transformed with the plasmid pTKLM176 which contains the gene of herpes simplex virus type I inserted into pBR322 (T. Lund and A. Mellor, unpublished data). 40ng of pTKLM16 DNA, together with 100-200ng of hemimethylated Mα1 DNA and 20μg of salmon sperm DNA were added as calcium phosphate coprecipitate to each petri dish containing 10^6 L-cells. In some experiments

unmethylated replicative form DNA of Mα1 or Mγ1 were used as cotransforming DNAs instead of hemimethylated DNA. Transformants (tk^+) were selected in DME medium supplemented with 5% new born calf serum and hypoxanthine, aminopterin and thymidine (HAT). Individual colonies were picked and grown into mass cultures.

Restriction and Hybridization Analysis

DNA of transformed cells was isolated as described by Stein et al.,(6). 10-50μg DNA was digested with a 5 to 10 fold excess of restriction endonuclease in the buffer recommended by the suppliers. After addition of the enzyme, a sample (0.5μg) of the DNA digest was removed and incubated with 0.5μg of phage λDNA, which allowed us to monitor the digestion. DNA fragments were electrophoresed on 1% or 1.5% agarose gels in Loening E buffer and then blotted onto nitrocellulose filters. The filter strips were hybridized overnight at 65°C to ^{32}P-labelled Aγ-globin, α1-globin or M13 DNA probes in 2 X SSC, 10% dextran sulfate, 0.1% SDS, and 10 X Denhardt's solution containing 20μg/ml of sheared mouse liver DNA. Post-hybridization washes were carried out at 65°C in 1 X SSC, 0.1% SDS (M13 DNA probes) or in 0.1 X SSC, 0.1% SDS (globin DNA probes) for 1-2 hrs. All DNA probes were labelled by nick-translation with α-^{32}P-dATP or α-^{32}P-dCTP to a specific activity of 10^8 cpm/μg. Complete digestion with the methylation-sensitive enzymes MspI and HpaII was controlled by adding a small amount of λDNA as internal control to the digestion mixture. The restriction pattern of this λDNA was made visible by hybridization of nick-translated λDNA to the blots previously hybridized with globin of M13 DNA probes.

RESULTS AND DISCUSSION

MspI does not cut certain CCGG sites in methylated human globin genes

To determine whether the MspI resistant cleavage sites have a specific DNA sequence, we carried out a search for such sites in the human α- and γ-globin genes and in the M13 viral cloning vector. The two globin genes were cloned into M13 and the single stranded DNA of the recombinant phages (Mα1 and Mγ1) was used as template for DNA synthesis using DNA polymeraseI, a short oligonucleotide as primer and as substrates dATP, TTP, dGTP and d^mCTP instead of normal dCTP. In this way, a duplex DNA methylated in one strand is produced. Stein et al. (6) showed that the unmethylated template strand of this hemimethylated DNA is methylated in vivo when this DNA is introduced into L cells by DNA mediated gene transfer. The methylated residues are predominantly or exclusively at ^mCpG sequences. The methylation pattern

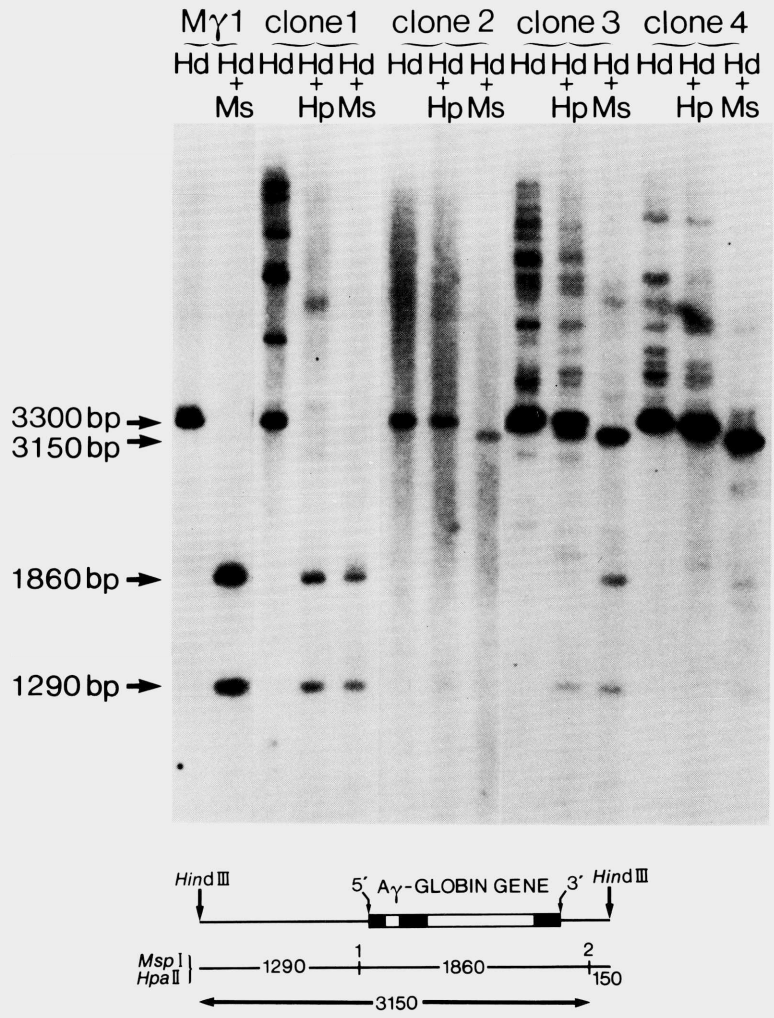


Fig. 2. MspI does not cut the 5' CCGG site of the Aγ-globin gene.
High molecular weight DNA of L-cells transformed with unmethylated (clone 1) or hemimethylated (clones 2, 3, 4) Mγ1 DNA was digested with the restriction enzymes indicated: HindIII (Hd), HpaII (Hp) and MspI (Ms). DNA fragments were electrophoresed on a 1% agarose gel, transferred to nitrocellulose and hybridized to the nick-translated 3300bp HindIII fragment of the Aγ-globin gene. Mγ1 denotes the M13 clone containing the γ-globin HindIII insert and was used as a size marker. The relevant restriction map of the γ-globin gene is shown in the bottom part with the sizes of the corresponding restriction fragments indicated.

thereby established is inherited faithfully from one cell generation to the next and can therefore be studied in transformed cell lines. DNA from L cells (clone 2, 3, 4) transfected with hemimethylated Mγ1 DNA was cleaved with HindIII (to release the human DNA insert from the vector(, HindIII plus HpaII or HindIII plus MspI (Fig. 2). As expected, the two CCGG sites present on this fragment are uncut by HpaII, showing that both sites are methylated at the internal C residues of the CCGG site. MspI cleaves the CCGG site present on the 3' side of the γ-globin gene, but fails to cut the site to the 5' side of the gene, resulting in a DNA fragment of 3150bp (Fig. 2). Since the same site is uncut (3) in human DNA from most tissues (Fig. 2), it follows that the 'inheritance' of this type of modification in this transformation system is similar to that of other ^{m}CpG's and that we can use this method to find other such sites. When unmethylated γ-globin DNA was used in the transformations (clone 1), both CCGG sites were cut with MspI and HpaII.

We next checked the cloned human α-globin gene (Mα1) for CCGG sites uncut by MspI. One such site was detected (number 16 in Fig. 3). DNA from transformed L cells was cut with PstI (to release the 1.5kb α-globin DNA insert) and either HpaII or MspI. Digestion with HpaII shows essentially no cleavage of the PstI fragment in DNA from cells (clones 6, 7, 8) transformed with hemimethylated α-globin genes, although total cleavage of this fragment was obtained with DNA for cell lines transformed with unmethylated α-globin DNA (clone 5). All 16 CCGG sites therefore retain the internal ^{m}C residue. Since there are so many CCGG sites in the PstI fragment cloned in Mα1, cleavage with MspI yields small fragments, most of which run off the gel showing in Fig. 3. However, two MspI fragments can be seen of 490bp and 460bp, respectively. The latter fragment results from cleavage at site 16. This site is partially resistant to MspI digestion resulting in the larger 490bp fragment. Digestion of DNA from unmethylated α-globin DNA (either cloned DNA or DNA from transfected cells) gives only the 460bp fragment. That the difference between methylated and unmethylated DNA is real is shown by mixing cloned α-globin DNA with the DNA from the L cells transfected with hemimethylated α-globin genes. Here two poorly resolved bands are found instead of a single band (Fig. 3).

M13 DNA also contains MspI-resistant CCGG Sites

Finally, we screened M13 vector DNA for CCGG sites which are resistant to cleavage by MspI. Two such sites are found at residues 2552 and 7007, respectively. To do this we purified the 818bp DNA fragment which flanks the

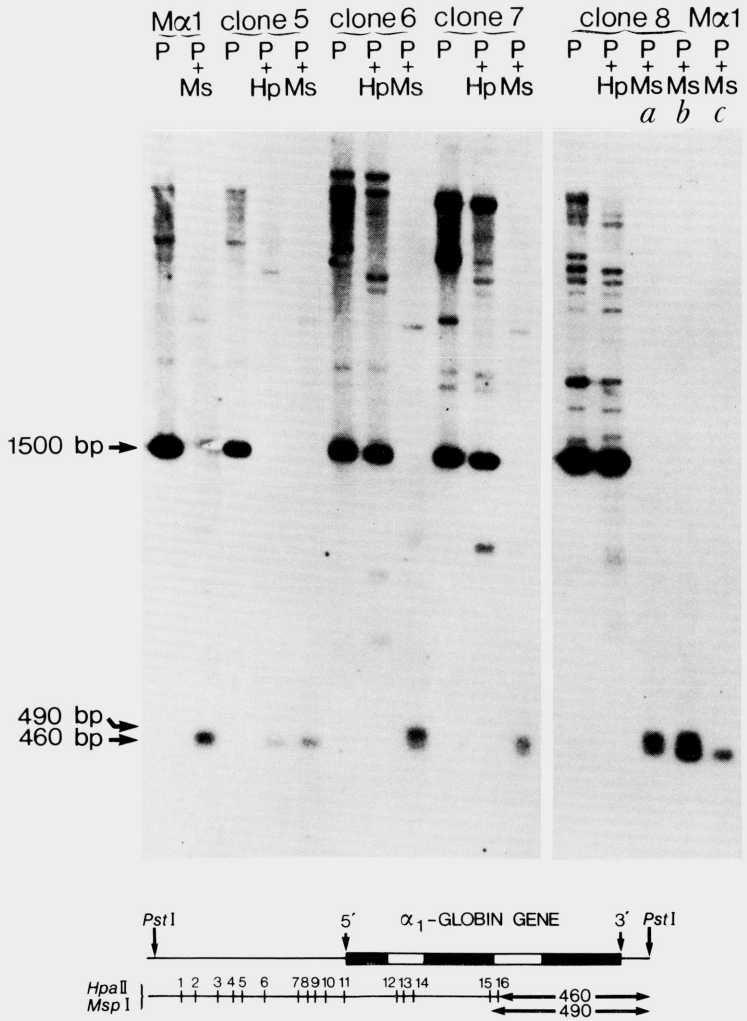


Fig. 3. Cleavage in the human α1-globin gene with MspI.

High molecular weight DNA of L-cells transformed with unmethylated (clone 5) or hemimethylated (clone 6, 7, 8) Mα1 DNA was digested with the restriction enzymes indicated: PstI (P), HpaII (Hp) and MspI (Ms). DNA fragments were separated on 1.5% agarose, transferred to nitrocellulose and hybridized to the nick-translated 1500bp long PstI fragment of the α-globin gene. Mα1, the M13 recombinant containing the α-globin PstI insert, was used as a size marker. The HpaII restriction map of the α-globin gene, shown in the bottom part, was determined by partial restriction mapping of the cloned PstI fragment. The HpaII sites are numbered from left to right with the size of the largest restriction fragment indicated. In one case (clone 8) we added MspI digested cloned Mα1 to the cut L cell DNA (in lane b) to show the resolution of the two bands.

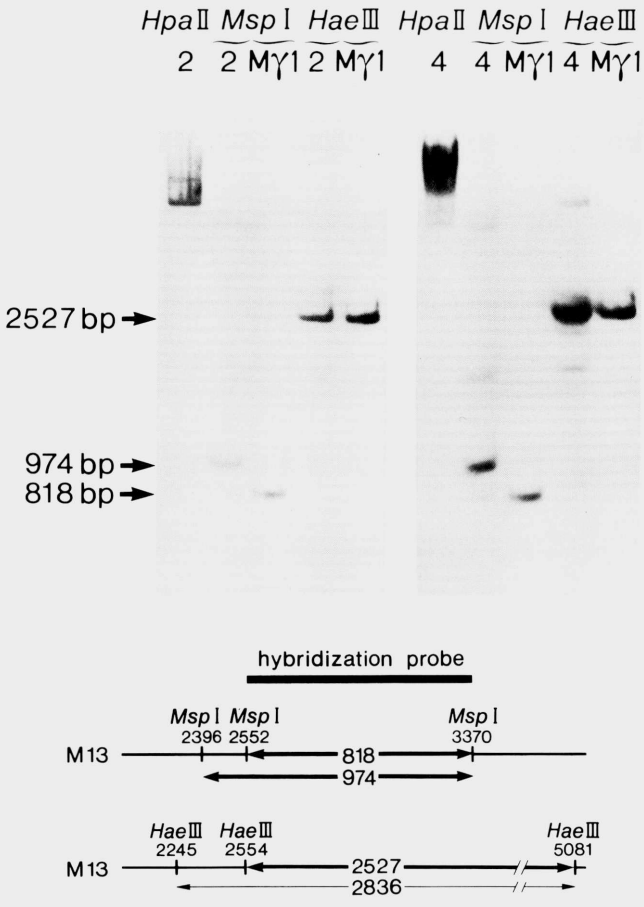


Fig. 4. Failure of MspI to cut a CCGG site of phage M13.
DNA from L cells transformed with hemimethylated Mγ1 (clones 2 and 4 of Fig. 2) was cut with HpaII, MspI or HaeIII, electrophoresed on an agarose gel and hybridized to the 818bp MspI fragment shown in the scheme.

site at residue 2552 from M13 RF DNA and used this as a probe in our Southern blots of DNA from L cells transformed with hemimethylated Mγ1 DNA (clones 2 and 4). The 818bp fragment is however not seen upon MspI cleavage. Instead, a fragment of about 970bp is seen which results from the failure of MspI to cut at site 2552 (Fig. 4). We have performed the similar experiments with the MspI site at residue 7007 with identical results (not shown).

We have aligned the DNA sequences of the four MspI-resistant sites in Fig. 5 (13, 14, 15). It can be seen that they have in common the sequence

*Msp*I #1 ($^{A}\gamma$) TGAGGCCAGG**GGCCGG**CGGCTGGCTA
*Msp*I #16 (α_1) GTGAGCGGCG**GGCCGG**GAGCGATCTG
*Msp*I 2552(M13) CCATTAGCAA**GGCCGG**AAACGTCACC
*Msp*I 7007(M13) GGGTGAGAAA**GGCCGG**AGACAGTCAA

Fig. 5. Alignment of the MspI sequences uncut by MspI in 'methylated' DNA.

(Pu)GGCCGG. Since the consensus sequence of the site contains overlapping HaeIII(GGCC) sites, this permitted a direct test of the methylation status of the two C residues of this sequence. If both C residues are methylated, then the site will be resistant to cleavage by HaeIII (blocked by GG^{m}CC) as well as MspI (blocked by ^{m}CCGG). This GG↓CCGG site at position 2554 of M13 is however cut by HaeIII in the DNA from the L cells transformed with hemimethylated DNA despite the fact that it is resistant to MspI; a 2557bp HaeIII fragment is seen rather than the 2836bp fragment which would be expected if the site were uncut (see Fig. 4).

Modification with HpaII Methylase Blocks MspI Cleavage at (Pu)GGCCGG

Since the four sites have a common DNA sequence and because of the HaeIII results, we were concerned that the failure of MspI to cut at this site could derive simply from the presence of the sequence GGC^{m}CGG, that is, a site with only a single ^{m}C residue. We therefore used HpaII methylase to methylate the sequence of Mγ1 and then digested the DNA with MspI, HpaII and HaeIII. MspI cleaves the unmethylated DNA much more readily than the methylated DNA. In addition, MspI only partially cuts at the MspI site at the 5' side of the γ-globin gene at enzyme amounts that give complete cleavage at other sites; at extremely high enzyme doses (200X excess) about 50% cleavage is found. Similar results are found for the M13 GGCCGG sites. In contrast, MspI cuts the CCGG sites 3' to the γ-globin gene and the remaining CCGG sites in the M13 vector to completion. As expected, HpaII fails to cut all CCGG sites (Fig. 6a). This result suggests that either MspI is blocked by the single ^{m}C residue, or the HpaII methylase modifies both C residues.

To exclude the latter possibility we cleaved methylated DNA with HaeIII and asked whether the GGC^{m}CGG site of M13 was cut by HaeIII. If the sequence was GG^{m}C^{m}CGG, HaeIII cannot cut this site (12). HaeIII, however, cleaves the *in vitro* methylated DNA to completion (Fig. 6b). To establish beyond doubt that the sequence $\begin{matrix} GG^{m}CCG\ G \\ CC\ GGC^{m}C \end{matrix}$ is not cut by HaeIII we used DNA methylase of Bacillus phage SPR19 which gives this modification pattern on short incubation times (U. Gunthert, unpublished). This methylated sequence is not cut by HaeIII (not shown).

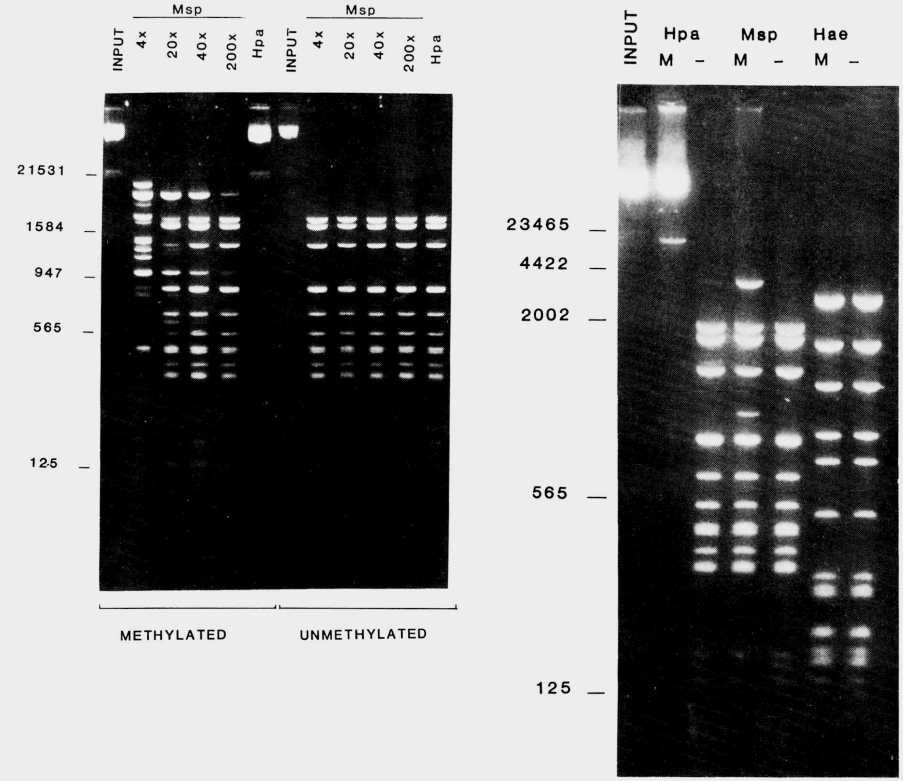


Fig. 6. Cleavage of Mγ1 DNA methylated *in vitro* with HpaII methylase.
Panel a, the left hand five lanes show the MspI digestion of methylated Mγ1 DNA up to 200 units/μg of DNA. Lane 6 shows the HpaII digest of methylated DNA. The right hand five lanes show the same MspI digests of unmethylated Mγ1 DNA. The last lane shows the HpaII digest of unmethylated DNA.
Panel b, shows the input DNA followed by HpaII, MspI and HaeIII digests of methylated (M) and unmethylated (-) Mγ1 DNA.

The results show that MspI fails to cut DNA at the sequence GGC^{m}CGG. It is theoretically possible that HpaII methylase generates the sequence $\begin{matrix} GGC^{m}C\ G\ G \\ CCG\ G^{m}C^{m}C \end{matrix}$ at this specific site even though the only modification detected up to now with the enzyme is $\begin{matrix} C^{m}C\ GG \\ G\ G^{m}CC \end{matrix}$. We consider this highly unlikely. The methylated strand of the M13 DNA sequence introduced into the animal cells in all our experiments was $\begin{matrix} GG^{m}C^{m}CGG \\ CC\ G\ GCC \end{matrix}$ and our data show that the external ^{m}C residue is lost after passaging through the cells ($\rightarrow \begin{matrix} GGC^{m}C\ GG \\ CCG\ G^{m}CC \end{matrix}$): it is implausible that this loss is accompanied by a transfer of the methyl group to the external C residue of the other DNA strand to give $\begin{matrix} GGC^{m}C\ C\ C \\ CCG\ G^{m}C^{m}C \end{matrix}$. Moreover, the fact

that we have observed the same phenomenon with both eukaryotic and bacterial methylase makes this explanation still more unlikely.

It is not clear why MspI has such difficulty cleaving this site when the internal C residue is methylated. ^{m}C residues have been implicated in the transition of DNA from the B to Z configuration (7) and it is possible that the site GGCCGG forms a non-B configuration (under MspI incubation conditions) when the internal C residue is methylated. Whatever the explanation, this phenomenon is clearly of some practical importance in the study of DNA methylation since such sites cannot be characterized with MspI. HaeIII and HpaII are, however, diagnostic for the presence of methyl group's sequence GGCCGG.

H. Cedar and his colleagues have independently examined this phenomenon (see this issue of NAR) and have also concluded that the internal ^{m}C residue prevents cleavage of these sites by MspI.

ACKNOWLEDGEMENTS

We are indebted to Dr. U. Gunthert for performing the experiment with the Bacillus phage DNA methylation. M. Busslinger was the recipient of a fellowship from the Schweizerischer National Fonds. We are grateful to H. Cedar for helpful discussions, to M. L. Birnstiel for providing the opportunity to M.B. to finish this work in his laboratory, and to F. Ochsenbein for graphical work. This work was supported by the British Medical Research Council.

REFERENCES

1. Bird, A. and Southern, E. M. (1978) J. Mol. Biol., 118, 27-48.
2. Waalwijk, C. and Flavell, R. A. (1978) Nucl. Acids Res., 5, 3231-3236.
3. van der Ploeg, L. H. T. and Flavell, R. A. (1980) Cell, 19, 947-958.
4. van der Ploeg, L. H. T., Groffen, J. and Flavell, R. A. (1980). Nucl. Acids Res., 8, 4563-4574.
5. Jentsch, S., Gunthert, U. and Trautner, T. A. (1981) Nucl. Acids Res., 9, 2753-2759.
6. Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A. and Cedar, H. (1982) Proc. Natl. Acad. Sci. U.S.A., 79, 61-65.
7. Behe, N. and Felsenfeld, G. (1981) Proc. Natl. Acad. Sci. U.S.A., 78, 1619-1623.
8. Sanders, Haigh, L., Blanchard Owens, B., Hellewell, S. and Ingram, V. M. (1982) Proc. Natl. Acad. Sci. U.S.A., 79, 5332-5336.
9. Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M. and Flavell, R. A. (1982) Nucl. Acids Res., 10, 6715-6732.
10. Messing, J. and Vieira, J. (1982) Gene, 19, 269-276.
11. Lauer, J., Shen, C. K. and Maniatis, T. (1980) Cell, 20, 119-130.
12. Mann, N. B. and Smith, H. O. (1977) Nucl. Acids Res., 4, 4211-4221.
13. Slightom, J. L., Blechl, A. E. and Smithies, O. (1980) Cell, 21, 627-638.
14. Michelson, A. M. and Orkin, S. H. (1980) Cell, 22, 371-377.
15. van Weezenbeck, P. N. T. F., Hulsebos, T. J. M. and Schoenmakers, J. G. G. (1980) Gene, 11, 129-148.

Volume 9 Number 17 1981 Nucleic Acids Research

Structure and expression of a cloned β^0thalassaemic globin gene

N.Moschonas, E.de Boer, F.G.Grosveld, H.H.M.Dahl, S.Wright, C.K.Shewmaker and R.A.Flavell

Laboratory of Gene Structure and Expression, National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

Received 2 July 1981

ABSTRACT

We have cloned the single β-globin gene from an Italian patient who is a double heterozygote for $\beta^0/\delta\beta^0$ thalassaemia. RNA isolated from nucleated red cells from this patient can be translated in vitro to give detectable levels of Aγ- Gγ and α-globin, but no β-globin. S_1-mapping transcription studies show that β-globin mRNA is present at about 1-3 % of the level of α- and γ-globin mRNA. In addition, the expression of this gene has been studied by reversed genetics. SV40-plasmid-β^0-globin gene recombinants have been transfected into Hela cells and analysed for β-globin mRNA. In contrast to the results obtained with mRNA isolated directly from the blood of this patient, in the transfected Hela cells the same level of β-globin mRNA is seen for the β^0 thalassaemic globin gene and for a normal β-globin gene.

To elucidate the nature of the lesion, the entire DNA sequence of the β-globin gene of this patient has been determined. The sequence shows that this gene contains a termination codon at position 39 (CAG -> UAG). Otherwise, there is a remarkable conservation of the entire DNA sequence.

INTRODUCTION

The hereditary disease β-thalassaemia is characterized by a reduced level of β-globin chains; since the α-globin chains are produced at the normal level, an imbalance in the synthesis of the α and β-chains ensues. β-thalassaemia can be divided into two types, β^0thalassaemia, where no β-globin chains are produced, and β^+thalassaemia, where low levels of structurally normal β-globin chains can be found (1-3).

Restriction enzyme mapping data have shown that the gross structure of the δ β-globin gene region is unaltered in most cases of β-thalassaemia (4, 5), although in one form (4, 5) the 3' terminal exon of the β-globin gene is deleted. In another form of β^0thalassaemia, Kan and his colleagues (6) have shown that globin mRNA is present at an approximately seven-fold reduced level, relative to α-globin mRNA and that the mRNA contains a translational stop codon at position 17 (AAG -> UAG) (7). Other forms of β^0-thalassaemia

have been characterized with levels of β-globin mRNA that could not be detected by cDNA titration studies.

To determine the nature of the lesion in this type of β^{o}-thalassaemia, we have cloned the β-globin gene from a patient who has been studied previously (8). We have analyzed the globin mRNA directly from the patient as well as analysing the expression of the β-globin gene using SV40 recombinants. DNA sequence studies show that this patient too, has a translational stop codon, this time at position 39 (CAG -> UAG).

MATERIALS AND METHODS

Cloning of the β-globin gene from a patient with $\beta^{o}/\delta\beta^{o}$thalassaemia

DNA from the $\beta^{o}/\delta\beta^{o}$ thalassaemia patient was isolated from blood as previously described (19). 10.8 mg of the DNA was then digested to completion with restriction endonuclease Hind III and fractionated by electrophoresis on a 0.5% preparative agarose gel. Fractions were phenol extracted, ethanol precipitated and resuspended in 10 mM Tris/HCl, pH 7.5. Fractions containing the β-globin gene fragment were determined by Southern blot analysis of aliquots of the fractions using nick translated plasmid HβG-DNA as probe (19). Positive hybridization was obtained with DNA from fractions containing DNA in the 7-9 kb size range.

Phage λ Ch21A was grown and DNA isolated as published (e.g. 20). 100 μg λ DNA (in 100 mM Tris/HCl, 10 mM $MgCl_2$, pH 8.0) was annealed at 42°C for 1 hr. followed by ligation at 37°C for 3 hrs in a buffer containing 100 mM Tris/HCl, 10 mM $MgCl_2$, 10 mM DTT, 1mM ATP, 0.02% gelatin and 1000 U T4 DNA ligase. The λ Ch21A DNA was then digested to completion with HindIII, treated with bacterial alkaline phosphatase, phenolextracted, ethanol precipitated and resuspended in 10 mM Tris/HCl, pH 7.5.

Human DNA fragments (containing the β-globin gene) were ligated to the vector DNA in the ratio 1:3, packaged *in vitro* as described by Hohn and Murray (21) and plated on bacterial strain DP50 sup F (20). 200,000 plaques were screened according to the method of Benton and Davis (22). One plaque was found to be positive on rescreening. The human β-globin Hind III DNA fragment was then subcloned in the Hind III site plasmid pBR322.

Detection of globin mRNA by the S_1 nuclease mapping

RNA from blood cells from the patient or from Hela cells was extracted as described (13) and hybridized to DNA probes as described previously (14). The probe for the α-globin gene product was a 5' labelled HindIII fragment (from a Sst I fragment of the α-globin gene subcloned in pBR322)which should

give a 179 nucleotide product corresponding to the fragment from the HindIII site in the second exon to the splice acceptor of the second exon. The probe was labelled as described (14). The γ-globin probe was a 3' labelled EcoRI fragment from the γ globin gene. This fragment was recovered from a cloned γδβ globin gene cosmid (cosHG28; see ref. 15). This probe should generate a 167 NT long S_1 nuclease resistant fragment which extends from the EcoRI site in the 3rd exon to the 3' end of the γ globin gene. The β-globin probe was a 3' labelled EcoRI-MspI double digest fragment labelled at the EcoRI site. Hybridization to β-globin mRNA generates a 212 NT S_1 nuclease resistant fragment. The 3' labelled probes were prepared as described previously (13) using α-^{32}P labelled dATP and TTP(2000/3000 Ci/mMol, Amersham).

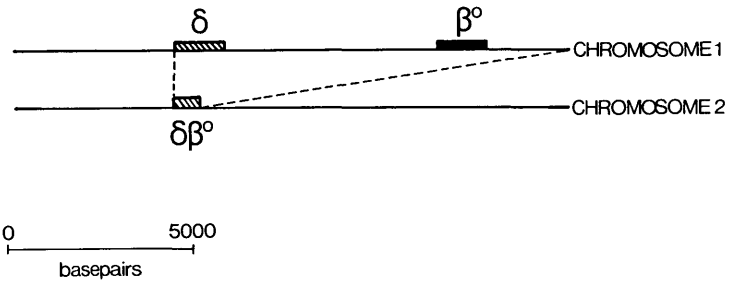
Transformation of Hela cells with β-globin gene-SV40 recombinants

A BamHI-EcoRI fragment of SV40 was subcloned in pBR328 (a generous gift of G.C. Grosveld). The 4.7kb BglII fragment, which contains the β-globin gene was isolated from a normal (15) or a β^othalassaemic individual (this report) and subcloned in the BamHI site of the SV40-pBR328 recombinant. 15 μg of the recombinant DNA was applied to Hela cells (grown in α-MEM + 10% newborn calf serum) as a calcium phosphate coprecipitate as described (15). After removal of the excess DNA-precipitate after about 16 h, the cells were grown for 48h, RNA isolated (13) and the level of β-globin mRNA determined by S_1 mapping. Under these conditions the rabbit β-globin gene generates about 10^4 β-globin mRNA molecules per cell.

RESULTS

Expression of the β^othalassaemic globin gene in vivo

Most studies of β^othalassaemia have been performed on "homozygous" patients. Though technically simple, this approach suffers from the disadvantage that if each respective allele exhibits a different form of β^o thalassaemia, then it is not usually possible to distinguish these in in vivo studies of gene expression. The general consensus is that β-thalassaemia is a heterogeneous disease which may be a manifestation of several different types of defect (see e.g. refs. 1-3). For this reason, we chose to study the β^o-globin gene of a patient doubly heterozygous for $\beta^o/\delta\beta^o$ thalassaemia. Since in δβothalassaemia the β-globin gene is entirely deleted (9,10,11),all studies of β-globin gene expression must therefore only refer to the single remaining β-globin gene in this patient (see Fig. 1).



Schematic Genotype of the Thalassaemic Patient

Fig. 1. Schematic genotype of the patients at the δ and β-globin locus. chr. = chromosome. δβ° refers to the partially deleted δβ° thalassaemic globin gene and β° to the thalassaemic β-globin gene. Blotting experiments described previously have shown that this patient has the classical Southern Italian form of δβ°-thalassaemia depicted here (9).

We first isolated total RNA from this patient and translated this in vitro using a wheat germ extract to establish that no translatable β-globin mRNA is present. The translation product was analysed on Triton-urea gels (12), since these separate both Gγ-, Aγ-, β-and α-globins; while Gγ, Aγ and α-globins were abundantly produced, no β-globin could be detected (not shown).

The RNA preparation was then assayed for α-, β-, γ- and δ-globin mRNA by hybridizing the preparation with 3' end-labelled probes for the β-γ- and δ-globin genes and a 5' probe for the α-globin gene as described in Materials and Methods. The hybrids were treated with S_1 nuclease and the S_1-resistant products analysed on polyacrylamide-urea 'sequencing' gels to determine the size of the product. Previous studies on the globin mRNA of this patient showed that the level of β-globin mRNA was below the detection level of the cDNA titration method used although α-globin mRNA was shown to be present (7). Using S_1-mapping, we found high levels of α- and γ-globin mRNA (Fig. 2) and a low level of S_1 resistant hybrid was formed with a δ-globin gene probe (not shown). Low but significant levels of β-globin mRNA are, however, also detectable in the RNA of this patient with this method. Both the 5' (not shown) and 3' ends (Fig. 2) of this mRNA are the same as those of normal β-globin mRNA, since the S_1 nuclease resistant products have the same size as those obtained with normal human β-globin mRNA. The extensive sequence divergence of the δ- and β-globin genes in the 3' untranslated region precludes the detection of heterologous δ-β hybrids with the β-globin

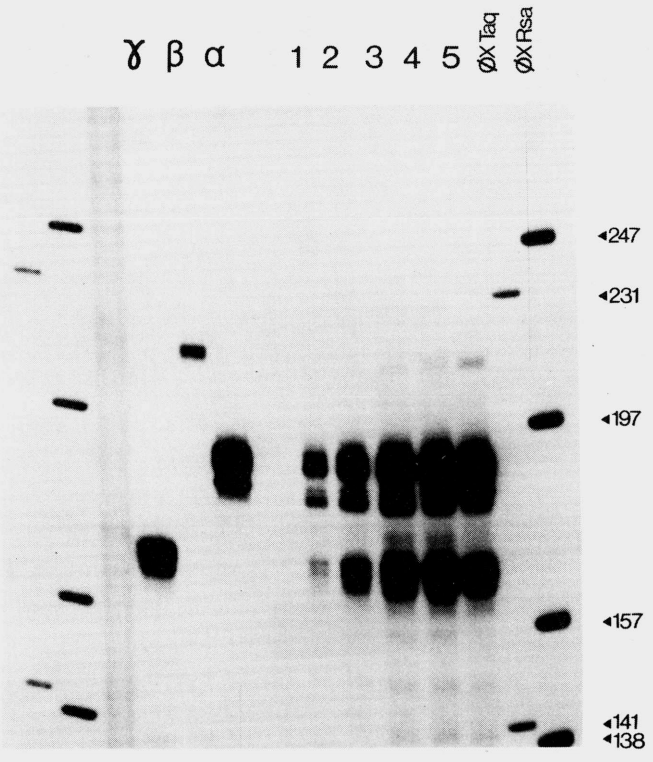


Fig. 2. Determination of the α, β and γ-globin mRNA levels in whole blood RNA from this patient. The RNA levels were determined by hybridization to 3' end labelled probes for the β (2×10^4 cpm/p mol 3' ends) and γ-globin genes (1.72×10^4 cpm/p mol 3' ends) and to a 5' labelled probe for the α-globin gene (1.5×10^4 cpm/p mol 5' ends). On the left side of the figure the probes have been hybridized separately to an excess of total blood RNA in a volume of 10 µl. On the right side the three probes were mixed and hybridized to 4ng, 20ng, 100ng, 400ng and 2µg respectively (1 to 5), of total RNA. The markers are 5' labelled fragments of ØX 174 DNA cut with TaqI or RsaI.

gene probe using this method.

Cloning of the thalassaemic β^0-globin gene and determination of the β-globin gene DNA sequence

A variety of mechanisms could be envisaged to explain the low level of β-globin mRNA and the absence of β-globin in the red cell precursors of this patient. To determine the molecular basis of the lesion in this case, we cloned a 7.5kb HindIII fragment which contains the β-globin gene using as vector bacteriophage λ charon 21A (see Materials and Methods). Standard

restriction mapping studies (F.G.Grosveld and H.H.M. Dahl, unpublished) confirmed that no detectable deletions of the DNA in the vicinity of the β-globin gene have occurred.

The primary sequence of the normal β-globin gene has been determined by Lawn et al. (16). We therefore determined the sequence of the same DNA region of the β^{o}-thalassaemic globin gene. Fig. 3 shows the sequencing strategy and the differences between the normal and thalassaemic gene sequence. Out of 2043 nucleotides, only 3 nucleotide differences were found between the normal and thalassaemic gene. The first difference is in the second exon of the β-globin gene and results in a translational termination codon at residue 39 (CAG -> UAG). This is obviously the causal agent of β-thalassaemia in this case. The other two differences (TT -> AA) reside in the 3' extragenic regions, 214 nucleotides downstream from the 3' end of the β-globin gene (defined as those sequences that encode β-globin mRNA). Remarkably, the entire β-globin gene sequence, including the entire sequence of the large intron, is otherwise identical to that of the normal β-globin gene.

We have also cloned the same HindIII fragment from another Italian β^{o}thalassaemic patient and determined part of the DNA sequence of the β-globin gene. The only difference observed here is the same transition at the codon for amino acid residue 39 (CAG -> UAG). The 5' extragenic region, the 5' and 3' intron-exon junctions and the first and second exons show no differences from the DNA sequence of the normal β-globin gene.

Expression of the β^{o}thalassaemic globin gene on an SV40 recombinant

The facts that the transcribed portion of the β^{o}-thalassaemic globin

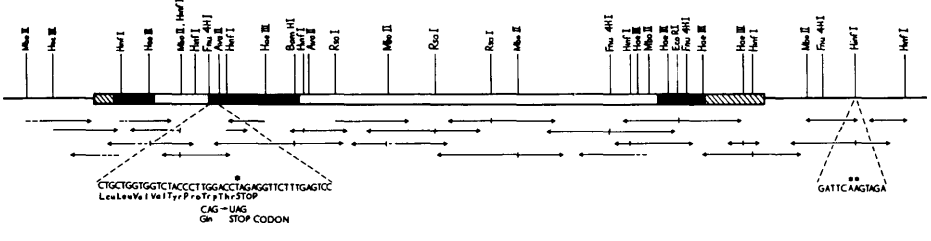


Fig. 3. The sequencing strategy for the β^{o}thalassaemic globin gene and the sequence differences observed. The β-globin gene is indicated as protein coding (filled boxes) 5' or 3' untranslated sequences (hatched boxes) or intron (open boxes). The TT -> AA difference in the 3' extragenic regions may not be real. The sequence of this region of the normal β-globin gene requires reconfirmation (T. Maniatis, A. Efstratiadis and N.Proudfoot, personal communication).

gene differs from the normal gene by only a single base substitution and that the 5' flanking DNA sequences are identical for both genes (see Discussion), suggest that the translational stop codon causes the low level of β-globin mRNA in the thalassaemic. Alternatively, it could be argued that a second mutation exists in the vicinity of the thalassaemic β-globin gene which causes, for example, a reduction in the level of transcription.

To examine this possibility we linked the β^{o}thalassaemic globin gene to an SV40-pBR328 plasmid vector which we have recently used to study the expression of the rabbit β-globin gene (G. C. Grosveld, E. de Boer, C. K. Shewmaker and R. A. Flavell, in preparation). The 4.7kb BglII fragment (which contains the human β-globin gene together with about 2 kb of 5' flanking DNA sequences and about 1.2 kb of 3' flanking DNA sequences), was introduced into the BamHI site of the SV40 vector by cloning in E.coli as described in Materials and Methods. This inserts the β-globin gene segment between the early region of SV40 and the tetracycline gene of pBR328. Both the normal and β^{o}thalassaemic globin genes were inserted into the SV40 vector and plasmids containing the human β-globin genes in both orientations (i.e. with the 5' end of the β-globin gene adjacent to the SV40 and pBR328 sequences respectively) were obtained.

The SV40 β-globin gene recombinants were then introduced into Hela cells as a calcium phosphate co-precipitate (see Materials and Methods). Under these conditions, W. Schaffner (personal communication) has shown that the rabbit β-globin gene inserted into the SV40 KpnI site of an SV40-pBR322 plasmid, is expressed specifically from its own promotor and that translatable β-globin mRNA is produced in these cells. To detect the human β-globin gene transcripts in the Hela cells, a 3' labelled 700 bp EcoRI +MspI fragment, which spans the 3' end of the β-globin gene was used as a hybridization probe in S_1 mapping experiments as described above and in Materials and Methods. Fig. 4 shows that the expected DNA fragment is also found when RNA from the Hela cells that have been transfected with human β-globin gene is hybridized with the 3' labelled fragment and digested with S_1 nuclease. Moreover, approximately the same level of β-globin mRNA is produced from the normal or thalassaemic β-globin gene. No protected fragment is seen for control Hela cells (not shown) or Hela cells transformed with the SV40 pBR328 vectors (Fig. 4). It is therefore unlikely that the low level of β-globin mRNA in the thalassaemic patient is due to a strongly reduced transcription level in vivo.

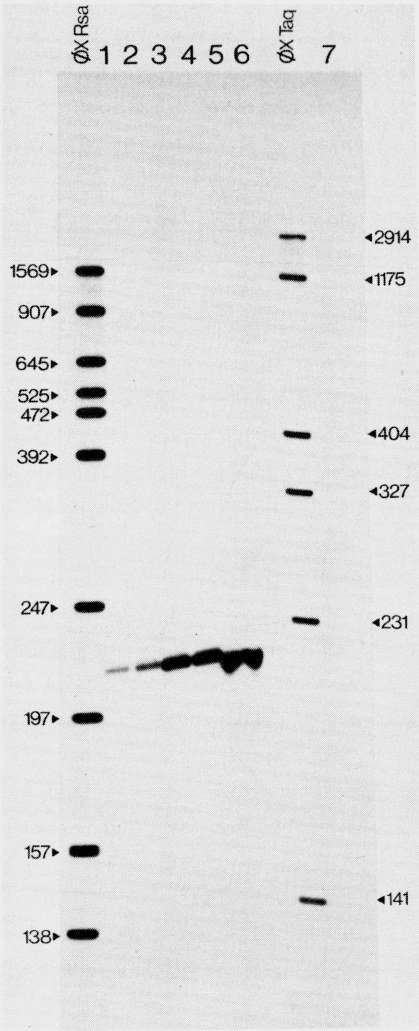


Fig. 4. β-globin mRNA levels produced from either the thalassaemic or normal β-globin gene as SV40 recombinants in Hela cells. The level of β-globin mRNA was determined by S_1 nuclease mapping as described in Materials and Methods using as probe a 3' labelled 0.7kb EcoRI + MspI fragment of the β-globin gene. The hybrids were analysed on an 80% polyacrylamide urea sequencing gel. As markers ^{32}P 5' labelled ØX DNA x RsaI and ØX x TaqI were used. Hybridization and S_1 nuclease treatment was performed as described in Materials and Methods. The RNA samples hybridized were 5µg 'normal' (1) or 'thalassaemic' (2) 20 µg 'normal' (3) or 'thalassaemic' (4) or 50µg of 'normal' (5) or 'thalassaemic' (6). As a control, RNA from Hela cells transformed with the SV40-pBR328 vector (7) was used. The number refers to the molecular weights of the marker fragments.

DISCUSSION

In this study we have employed a combination of structural and functional analyses to elucidate the nature of the lesion in a case of β^{o} thalassaemia. The DNA sequence of the thalassaemic β-globin gene in these two Italian patients, shows that the primary lesion is a translational termination codon at residue 39; Kan and his colleagues (6) previously showed a termination codon at residue 17 in a Chinese form of β^{o} thalassaemia and more recently, they have demonstrated an identical lesion to the one described here (CAG -> UAG) in another Italian β^{o} thalassaemic patient (25). It seems possible, therefore, that this will turn out to be a common lesion in β^{o} thalassaemia. It should be noted that while earlier studies suggested that these patients did not produce β-globin mRNA (e.g. 11), these current studies show this RNA to be present. Whether β-thalassaemic cases with zero levels of β-globin mRNA exist remains to be seen.

Interestingly, in the patient we have characterized in most detail, the level of β-globin mRNA is only about 1-3% of the normal level. It follows that β-globin mRNA level is about 50 fold lower than the normal value. Y. W. Kan and his colleagues have made a similar observation for their patient (25). Since the entire β-globin gene only differs at one nucleotide position, it would follow that the β-globin mRNA produced only differs at a single residue, namely in the translational stop codon and that this is the cause of the low level of β-globin mRNA. The alternative explanation is, of course, that an additional mutation has occurred, for example, in the β-globin gene promotor which results in a reduced rate of RNA synthesis. We do not consider this likely. The 5' extragenic nucleotide sequence of the β^{o} thalassaemic globin gene is identical to that of the normal β-globin gene up to residue -340 (this report and N.M., E.de B. and R.A.F., unpublished data). This is beyond the region considered to contain the β-globin gene promotor (for discussion, see 18). More important, the β-thalassaemic gene generates the same level of β-globin mRNA as is found for the normal β-globin gene, when these genes are transcribed in Hela cells.

At this stage it seems likely that the low level of β-globin mRNA in the erythroblasts of this patient may be a result of a reduced stability of this defective β-globin mRNA, as was previously proposed for the Chinese β^{o} thalassaemia (6). Such instability might result from the nuclease susceptibility of mRNAs which are not protected by ribosomes because protein synthesis terminated prematurely on these mutant mRNAs. Since this difference

is not seen in relatively short-term (48 h) in vivo experiments in Hela cells, we believe that it results from a difference between the Hela cell and in vivo nucleated red cell cytoplasm that must at present remain the subject of speculation.

Finally, it is noteworthy that exceedingly few polymorphic differences exist between the DNA sequence of this β^{o}thalassaemic β-globin gene and the normal β-globin gene sequence. In fact, of the total of 1600 nucleotides present in the transcribed segments of each of four β-globin genes (one β^{o} thalassaemic (this report), two identical β^{+}thalassaemics (23,24)and one normal gene (17)), only one nucleotide difference is seen per abnormal gene and this difference appears to be the causal agent of the disease in each case - no polymorphic differences have been seen. This differs markedly from the human foetal globin genes where two allelic Aγ globin genes differed by 15 base substitutions and three deletions of 18, 4 and 4 nucleotides respectively.

ACKNOWLEDGEMENTS

F.G.G., H.H.M.D., E.deB. and R.A.F. are grateful to J. Collins for his hospitality in Braunschweig in the earliest stages of this work, to T. Maniatis for a gift of the α-globin genomic subclone, and to R. Harvey and A. Smith for help with the in vitro translation of globin mRNA. F.G.G. and R.A.F. gratefully acknowledge an EMBO short term fellowship to fund the original visit to Braunschweig. H.H.M.D. and N.M. were recipients of EMBO long term fellowships. F.G.G. and C.K.S. were recipients of a Royal Society European Programme fellowship and a Jane Coffin Childs fellowship respectively. This work was supported by the British Medical Research Council.

REFERENCES

1. Weatherall, D. J. and Clegg, J. B. (1972) in: The Thalassaemis Syndromes. Blackwell Scientific Publications, Oxford (second edition).
2. Forget, B. G. (1978). Trends Biochem. Sci. 3, 86-89.
3. Bank, A., Mears, J. G. and Ramirez, F. (1980). Science 207, 486-493.
4. Flavell, R. A., Bernards, R., Kooter, J. M., de Boer, E., Little, P.F.R., Annison, G. and Williamson, R. (1979). Nucl. Acids Res. 6, 2749-2460.
5. Orkin, S. H., Old, J. M., Weatherall, D. J. and Nathan, D. G. (1979). Proc. Natl. Acad. Sci. U.S.A. 76, 2400-2404.
6. Temple, G. F., Chang, J. C. and Kan, Y. W. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3047-3051.
7. Chang, J. C. and Kan, Y. W. (1979). Proc. Natl. Acad. Sci. U.S.A. 76, 2886-2889.

8. Ottolenghi, S., Lanyon, W. G., Williamson, R., Weatherall, D. J. Clegg, J. B. and Pitcher, C. S. (1975). Proc. Natl. Acad. Sci. U.S.A., 72, 2294-2299.
9. Bernards, R., Kooter, J. M. and Flavell, R. A. (1979). Gene 6, 265-280.
10. Fritsch, E. F., Lawn, R. M. and Maniatis, T. (1979) Nature, 279, 598-603.
11. Ottolenghi, S., Giglioni, B., Comi, P., Gianni, A. M., Polli, E., Acquaye, C. T. A., Oldham, J. H. and Masera, G. (1979) Nature, 278, 654-657.
12. Alter, B. P., Goff, S. C., Efremor, G. D., Gravely, M. A. and Huisman, T. H. (1980). Brit. J. Haematol. 44, 527-534.
13. Auffrey, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
14. Grosveld, G. C., Koster, A. and Flavell, R. A. (1981) Cell 23, 573-584.
15. Grosveld, F. G., Dahl, H. H. M., de Boer, E. and Flavell, R. A. (1981) Gene 13, 227-237.
16. Wigler, M., Silverstein, S., Lee, L-S., Pellicer, A., Cheng, Y. and Axel, R. (1977) Cell 11, 223-232.
17. Lawn, R. M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980) Cell 21, 647-651.
18. Efstratiadis, A., Posakony, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C. and Proudfoot, N. J. (1980) Cell 21, 653-668.
19. Flavell, R. A., Kooter, J. M., deBoer, E., Little, P. F. R. and Williamson, R. (1978) Cell 15, 25-41.
20. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. and Efstratiadis, A. (1978) Cell 15, 687-701.
21. Hohn, B. and Murray, K. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 3259-3263.
22. Benton, W. and Davis, R. (1977) Science 196, 180-182.
23. Westaway, D. and Williamson, R. (1981) Nucl. Acids Res., 9, 1777-1788.
24. Spritz, R. A., Jagadeeswaran, P., Choudary, P. V., Biro, P. A., Elder, J. T., deRiel, J. K., Manley, J. L., Gefter, M. L., Forget, B.G. and Weissman, S. M. (1981) Proc. Natl. Acad. Sci. U.S.A. (in press).
25. Trecartin, R. F., Liebhaber, S. A., Chang, J. C., Lee, K. Y., Kan, Y. W., Furbetta, M., Angius, A. and Cao, A. (1981). J. Clin. Invest. (in press).