Volume 15 Number 9 1987 Nucleic Acids Research

Changes in minor transcripts from the $\alpha 1$ and β^{maj} globin and glutathione peroxidase genes during erythropoiesis

Jonathan Frampton, David Conkie, Ian Chambers, Wendy McBain, Mike Dexter* and Paul Harrison

The Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD and *Paterson Laboratories, Wilmslow Road, Manchester, UK

Received February 9, 1987; Revised and Accepted April 16, 1987

ABSTRACT

We have analysed the transcriptional regulation of the murine $\alpha 1$ and β^{maj} globin genes and the glutathione peroxidase (GSHPx) gene, which are all highly expressed during erythropoiesis. The levels of minor RNAs compared to the major message were monitored throughout differentiation within the erythroid lineage. For each gene, upstream transcripts arise from distinct clusters of sites which are regulated differently during differentiation: some occur only during early erythropoiesis, some occur early and persist to the terminal stages, while others accumulate later and roughly in parallel with the main RNA transcript. In addition, opposite strand transcripts from the GSHPx gene were found in increasing amounts during later stages of erythropoiesis. The initiation sites for specific subsets of these minor transcripts lie close to sequences known to be involved in globin gene regulation (i.e. the TATA, CAAT and the CACCCT boxes) or other conserved sequences; others lie close to developmentally regulated DNase I hypersensitive sites around the globin and GSHPx genes.

INTRODUCTION

Murine erythropoiesis is a differentiation process which lends itself to analysis of the molecular mechanisms involved in the co-expression of genes. Much work has centered on control elements around the globin genes especially with regard to β globin gene transcriptional activity (1), inducibility of the β globin gene in erythroleukaemic cells (2-6), and minimum sequence requirements for erythroid-specific expression of β globin genes in transgenic mice (7,8). However, little attention has been directed towards erythroid-specific control elements in α globin genes or non-globin genes.

The best characterised mouse red blood cell non-globin proteins for which cloned genes are available are the carbonic anhydrase II (CAII) isoenzyme (P.Curtiss, personal communication; (9)) and glutathione peroxidase (GSHPx) (10). The GSHPx mRNA is interesting in that it is highly expressed in red cells and in a limited number of other tissues (notably liver and kidney) in contrast to its low level in other cell types (10,11).

© IRL Press Limited, Oxford, England.

The upstream flanking sequence of the gene exhibits features characteristic of many constitutive genes (10). However, the GSHPx gene's pattern of expression poses an interesting question in terms of how it comes to be co-expressed with various groups of other genes in different tissues.

Transcription occuring from sites other than the normal cap position has been documented for a number of genes (12-15) including the globin genes (16-21). Generally the transcripts arise from upstream positions, are of the same sense as the major transcript, and are polyadenylated. However, there are examples of minor transcripts occuring from the opposite-sense strand (14,15), and lacking a polyadenylic acid tail (22).

Often the ratios of minor RNAs to the major transcript from a particular gene are found to vary either in a developmental-specific manner (13,23), or in response to stimuli (14). Studies of sequences close to such minor transcription start sites, particularly from the human ε globin gene (24,25), have suggested a regulatory role for these regions. Moreover, there is a close correlation between the positions of ε globin gene upstream transcripts and tissue-specific nuclease hypersensitive sites which are thought to be involved in transcriptional activation (26). A further close correlation between minor transcripts and regulatory regions is seen in some viral promoters where minor RNA start sites occur within functionally defined transcriptional enhancer sequences (12,13).

In this report, the GSHPx gene is compared to the α1 and β^{maj} globin genes in terms of changes in minor transcripts throughout erythroid differentiation from haemopoietic stem cells through to terminally differentiated reticulocytes. The pattern of expression of the various minor transcripts implies that they are involved in, or are reflecting, developmentally specific events.

MATERIALS AND METHODS

Cell lines and tissues

Where possible all cells were derived from Balb/c mice in order to minimise the possibility of small regions of non homology between RNAs and the Balb/c DNA probes and primers. Balb/c Friend cells (clone FSD 2/3, a gift from W. Ostertag, Hamburg (unpublished)) were grown in Eagles MEM with double the concentrations of amino acids and vitamins and with 10% foetal calf serum. Induction of differentiation was achieved by the addition of 5mM HMBA to exponentially growing cells set up at $2x10^5$/ml, cells being collected 0-2 d later. DBA/2 Friend cells (clone 707B10/1) and the mouse

fibroblast line (Balb/C-N) were grown as described previously (10). The haemopoietic stem cell line DeCl.15 (27) was grown at 33° in Fischer's medium supplemented with 20% horse serum and 10% WEHI-conditioned medium. The pregranulocytic cell line, AD3 (28), was maintained in the same medium at 37°. The neuroblastoma line N18 was grown in 707B10/1 Friend cell medium (see above). Reticulocytes were prepared as described previously (10). Kidneys and livers were dissected from Balb/c mice, quickly cooled on ice, and RNA extracted immediately.

Isolation of committed erythroid progenitor cells (CFU-Es)

The basis of the method (29) involves thiamphenicol inhibition of erythropoiesis at the BFU-E stage causing anaemia and high serum erythropoietin levels. Upon drug removal, a wave of erythropoiesis occurs with accumulation of CFU-Es in the spleen; these are then purified by centrifugal elutriation and banding on a Percoll gradient. Female Balb/c mice (12-14w, 23-25g) were made anaemic by s.c. implanation of 400 mg thiamphenicol (Sigma) in 0.5 ml water contained in a dialysis bag. Mice were bled on d1 and 2 and the thiamphenicol removed at d4. On d8, the spleens (300-450 mg) were removed and mechanically disrupted to produce a single-cell suspension in alpha modified Eagle's medium. This was then injected into a centrifugal elutriator rotating at 2000 rpm with a flow rate of 17 ml/min of MEM/2% foetal calf serum. Elutriated cells were then collected in 200 ml at a flow rate of 30 ml/min at 20°. After pelleting and layering onto 60% Percoll (Pharmacia), discrete bands of cells were formed after 30 min centrifugation at 13000 rpm in an SS34 Sorval rotor. Cells banding at a density of 1.065-1.070 g/ml (typically, $1.5\text{-}2\times10^8$ cells from 10 spleens) were shown to comprise 80-85% CFU-E as determined by the standard CFU-E assay (30).

Molecular Biology Techniques

Recombinant DNA manipulations were carried out as described in Maniatis et al (31). Sequence analysis was performed using an IBM PC-AT microcomputer in conjunction with the "Micro-Genie" software (Beckman).

Preparation of total and polyA$^+$ RNAs, end-labelling of DNA, S1 nuclease protection, and primer extension on RNA templates were carried out as described previously (10).

RESULTS

Minor upstream transcripts from α1 and β^{maj} globin genes in reticulocytes

Mapping and quantitation of transcripts was routinely achieved by the

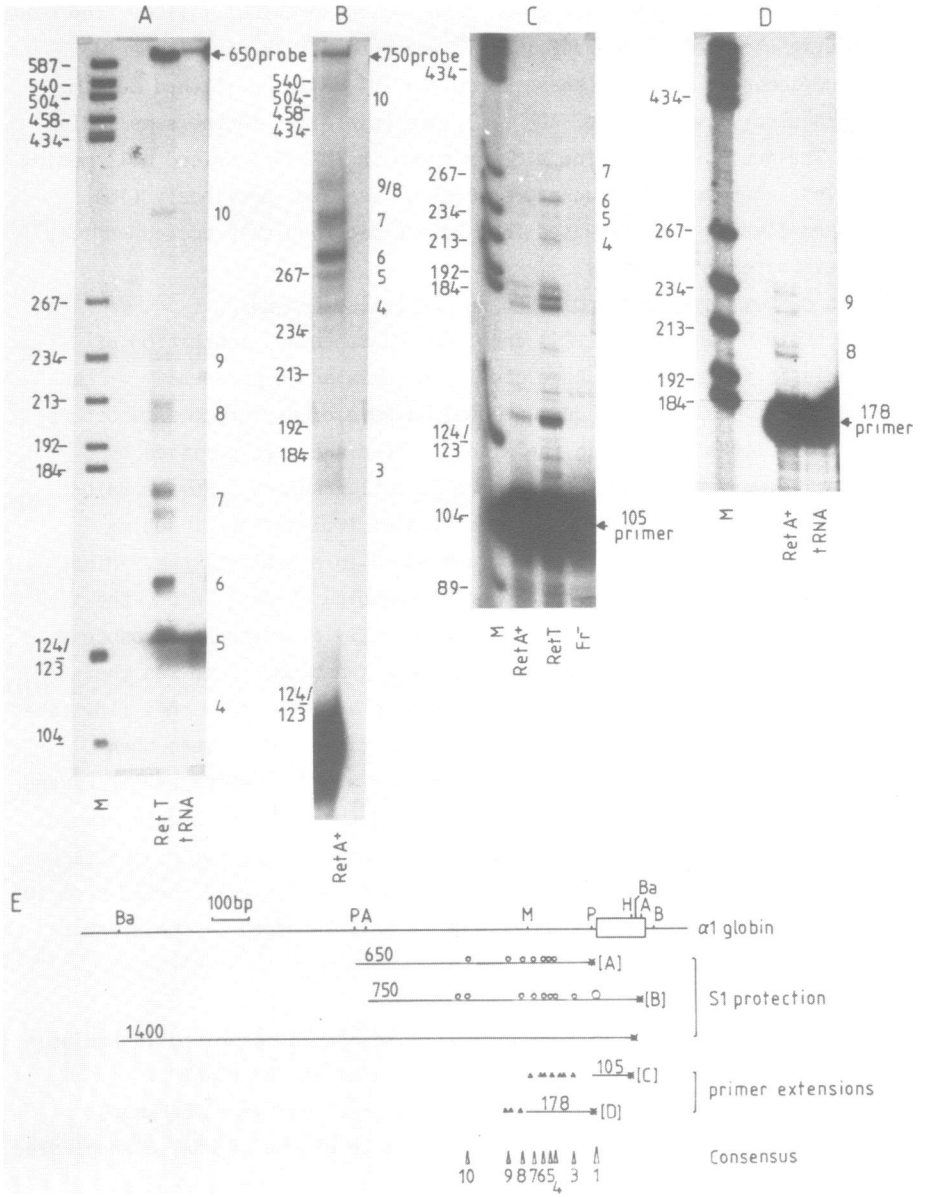

Fig.1. Mapping of α1 globin upstream transcripts in reticulocyte RNA S1 protection experiments (A and B) and primer extensions (C and D) were carried out as described in Materials and Methods. Hybridisations (in 10 µl) were at 52° for 16h with the total RNA concentration adjusted to 5 mg/ml with yeast RNA. S1 digestion was with 2000U S1 (Boehringer) at 37°

for 1h, initial experiments using a range of temperature from 25° to 42° having established that these conditions were optimum. In the summary diagram (E), open circles indicate RNA 5' termini deduced by S1 protection (the large circle being the major transcript) and small triangles indicate RNA 5' termini deduced by primer extension. The open box represents the first exon. A consensus for the pattern of transcripts deduced by both methods is shown at the bottom. Restriction sites used in generating primers and probes are: A, AluI; B, BamHI; Ba, BalI; H, HaeIII; M, MspI, and P, PstI. Primers and probes were derived from subclones of the α1 globin genomic clone (41) using restriction sites derived from the DNA sequence (41) and our own mapping data. RNAs used were: tRNA, yeast tRNA; RetT. 50 µg total reticulocyte RNA; Ret A^+, 1 µg reticulocyte poly$^+$ RNA; and Fr^-, 50 µg uninduced Friend cell (FSD2/3) total RNA. Markers (M) were pBR322 cut with HaeIII.

S1 nuclease protection method; in addition the majority of these sites were verified by primer extension. Molar excess of single-stranded end-labelled DNA was used allowing coding strand definition and quantitation. For reasons of simplicity bands different by only a few nucleotides are treated as the same transcript.

For the α globin gene, besides the major transcript (α1) eight predominant minor transcripts (referred to subsequently as α3-10 respectively) were identified by the S1 nuclease protection method using a 750nt AluI-AluI probe, a 650nt PstI-PstI probe, or a 1400nt BalI-BalI probe (Figure 1A,B, summarised in 1E): these map between -70nt and -360nt relative to the cap site, the majority (transcripts α4-9) being clustered between -125nt and -245nt. Confirmation of the distribution of the upstream start sites was attempted by primer extension from the 105nt PstI-HaeIII or the 178nt MspI-PstI primers (Figure 1C,D, summarised in 1E); however, particularly in the case of primer extensions on α globin RNAs it was noted that bands predicted from the results of S1 protection only represented a subset of those generated by primer extension. These additional bands, which represented a majority of the extension products using the 105nt PstI-HaeIII primer (extension products of 130nt and a cluster centered on 184nt - see Figure 1C), probably represent termination of reverse transcriptase caused by secondary structures in the RNA.

All of the predominant α globin upstream transcripts appear to be polyadenylated (Figure 1B and data not shown). An additional transcript mapping at -35nt (α2) at earlier stages of differentiation will be discussed below (Figure 4). Densitometric scanning gives estimates of the steady-state levels of the minor transcripts relative to the cap site from 1% for α7 to less than 0.01% (Table I). Preliminary S1 mapping data using other probes

Table I : Summary of α1 and βmaj upstream transcripts and variations during erythropoesis

Globin transcript (position relative to major start) [a]	Level of transcript as percentage of total transcripts (ratio of steady-state levels of minor transcript between indicated RNA and reticulocyte RNA) [b]							
	Reticulocyte	DeCl.15	AD3	Fr⁻ (707B10/1)	Fr⁻ (FSD2/3)	Fr+1	Fr+2	CFU-E
α1(0)	96.7	85(<0.001)	75(<0.001)	98(0.05)	98(0.1)	ND(0.4)	99(1.5)	98
α2(-35)	trace	10	20	2	1	ND	~0.3	0.2
α3(-70)	0.2	-	-	-	-	ND	trace	trace
α4(-125)	0.07	-	-	-	-	ND	~0.01(<0.1)	-
α5(-140)	0.3	-	-	-	-	ND	0.1(0.3)	-
α6(-150)	0.9	-	-	-	-	ND	0.02(0.03)	-
α7(-180)	1.0	-	-	-	-	ND	0.1(0.1)	-
α8(-220)/α9(-245)	0.7	-	-	-	-	ND	0.1(0.1)	-
α10(-360)	0.2	(5(0.02)	~5(~0.02)	ND	1(0.3)	ND(1)	0.2(1.5)	1.8(1)
β1(0)	93.7	100(0.001)	59(0.001)	83(0.007)	72(0.02)	53(0.015)	97.9(0.9)	98.9(0.3)
β2(-50)	1.0	-	13(0.02)	1.3(0.01)	5.8(0.15)	11.3(0.3)	0.8(0.7)	0.18(0.05)
β3(-60)	0.5	-	6(0.02)	2.5(0.04)	2.9(0.15)	5.7(0.3)	0.2(0.35)	0.09(0.05)
β4(-75)	0.5	-	15(0.05)	9.5(0.15)	9.6(0.5)	15.1(0.8)	0.5(0.8)	0.26(0.15)
β5(-90)	1.0	-	6(0.01)	2.5(0.02)	3.8(0.1)	7.6(0.2)	0.3(0.25)	0.11(0.03)
β6(-110)	2.0	-	-	1.3(0.005)	3.1(0.04)	3.8(0.05)	0.2(0.1)	0.07(0.01)
β7(-140)	0.25	-	-	0.1(0.005)	0.3(0.03)	0.3(0.03	0.04(0.15)	0.02(0.025)
β8(-160)/β9(-165)	0.6	-	-	-	0.9(0.04)	1.4(0.06)	0.03(0.2)	0.13(0.06)
β10(-170)	0.2	-	-	-	0.8(0.1)	0.8(0.1)	0.03(0.15)	0.08(0.12)
β11(-180)	0.15	-	-	-	0.6(0.1)	0.6(0.1)	0.02(0.1)	0.05(0.1)
β12(-195 to -225)	0.1	-	3(0.05)	-	0.4(0.1)	0.4(0.1)	0.01(0.1)	0.04(0.1)
β13(-540)	<0.01	ND	ND	ND	ND	ND	ND	ND

a Transcripts are designated α1-10 (α1 globin) and β1-13 (βmaj globin)
b Determined by densitometric scanning and/or serial dilutions of reaction products on sequencing gels
ND Not determined
- Below the level of detection by scanning of autoradiograph

extending as far as the EcoRI site at -4Kb has suggested the existence of very low level non-polyadenylated transcripts arising from -730nt, -835nt, -1000nt, and -1060nt (data not shown).

Analogous results were obtained for the β^{maj} globin gene (Figure 2, Table I). Ten predominant minor transcripts (β2-11) were identified between -50nt and -180nt using the 239nt DdeI-DdeI and 690nt Sau3A-Sau3A overlapping probes in S1 protection experiments (Figure 2A,B, summarised in 2E). Two additional minor start sites of lower abundance were found between -195nt and -225nt (β12 cluster) and at -540nt (β13). The β5-11 and β13 transcripts have been confirmed by primer extension experiments (Figure 2C,D). All of these minor transcripts appear to be polyadenylated, and represent from 2% (β6) to less than 0.01% (β13) of the total transcripts. Additional minor transcripts have been identified upstream of β13 using a 1070nt EcoRI-HindIII fragment in S1 protection experiments but it has as yet proved difficult to assign their orientation.

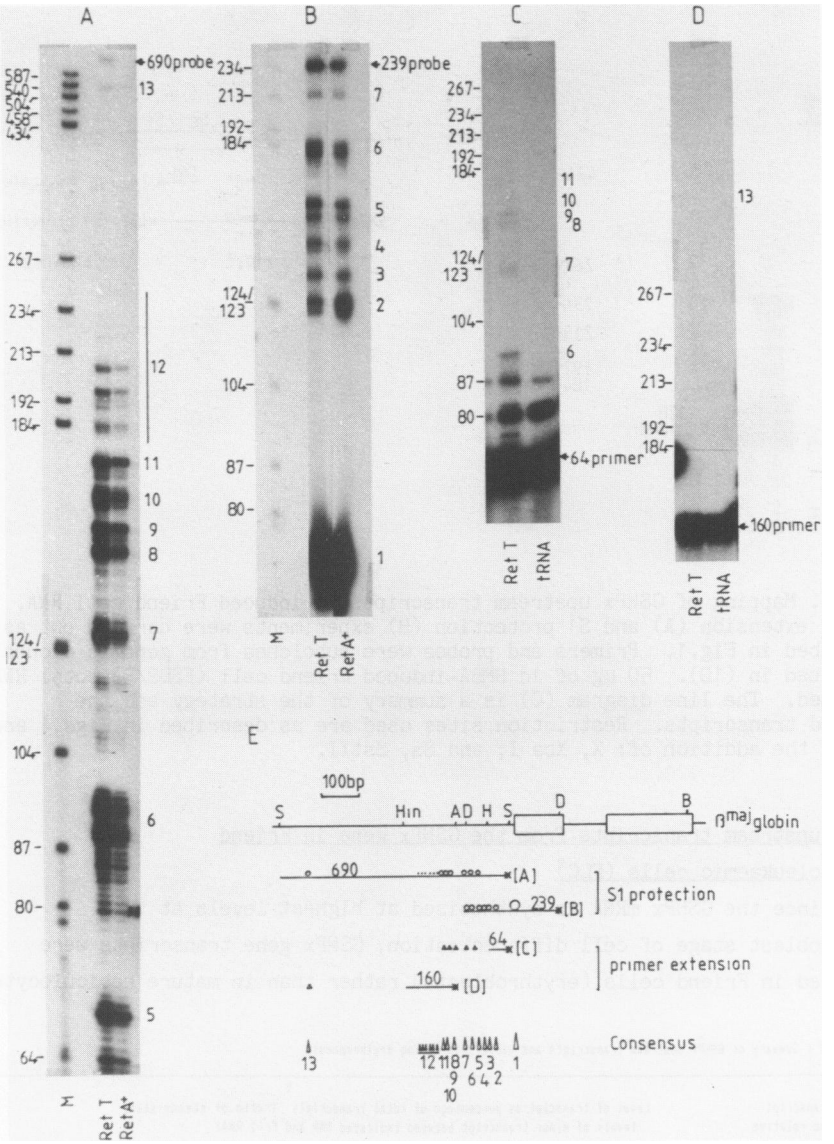

Fig. 2. Mapping of β^{maj} globin upstream transcripts in reticulocyte RNA S1 protection experiments (A and B) and primer extensions (C and D) were carried out as described in Fig. 1. The line diagram at the bottom summarises the strategy and the deduced transcription sites using the same notation as in Fig.1. Primers and probes were derived from subclones of the β^{maj} globin genomic clone CA11 (42) using restriction sites derived from the DNA sequence (43 and 44). Restriction sites used in generating fragments were; A, AluI; B, BamHI; D, DdeI; H, HaeIII; Hin, HindIII; and S, Sau3AI. The RNAs used were as in Fig. 1.

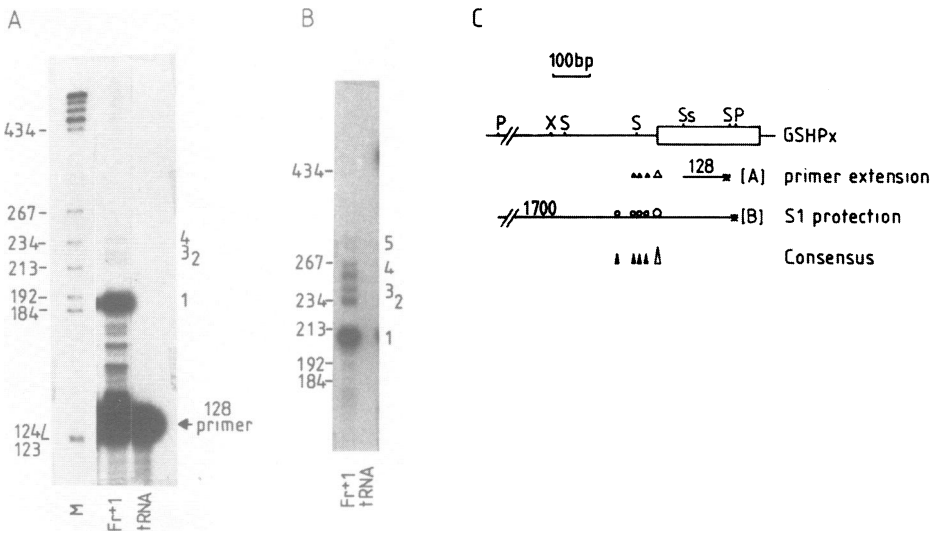

Fig. 3. Mapping of GSHPx upstream transcripts in induced Friend cell RNA. Primer extension (A) and S1 protection (B) experiments were carried out as described in Fig.1. Primers and probes were subcloned from genomic clones described in (10). 50 µg of 1d HMBA-induced Friend cell (FSD2/3) total RNA was used. The line diagram (C) is a summary of the strategy and the deduced transcripts. Restriction sites used are as described in Figs.1 and 2 with the addition of: X, Xba I; and Ss, SstII.

Minor upstream transcripts from the GSHPx gene in Friend erythroleukaemic cells (FLC)

Since the GSHPx mRNA is synthesised at highest levels at the erythroblast stage of cell differentiation, GSHPx gene transcripts were examined in Friend cells (erythroblasts) rather than in mature reticulocytes

Table II : Summary of GSHPx upstream transcripts and variations during erythropoesis

GSHPx transcript (position relative to major start)	Level of transcipt as percentage of total transcripts (ratio of steady-state levels of minor transcript between indicated RNA and Fr+1 RNA)[b]							
	Fr+1	DeCl.15	AD3	Fr⁻ (FSD2/3)	Fr+2	Reticulocyte	CFU-E	Kidney
G1(0)	96	95(0.16)	81(0.03)	97.7(0.6)	98(1.04)	98.8(0.92)	98(0.8)	99.2(0.9)
G2(-25)	1.0	1.5(0.25)	7(0.25)	0.68(0.4)	0.59(0.6)	0.28(0.25)	0.64(0.5)	0.23(0.2)
G3(-40)	0.7	1.1(0.25)	4.9(0.25)	0.48(0.4)	0.55(0.8)	0.23(0.3)	0.45(0.5)	0.24(0.3)
G4(-50)	1.0	2.5(0.4)	7(0.25)	0.68(0.4)	0.49(0.5)	0.22(0.2)	0.38(0.3)	0.23(0.2)
G5(-100)	0.5	ND	ND	0.42(0.5)	0.39(0.8)	0.45(0.8)	0.51(0.8)	0.06(0.1)
G6(-140)	<0.1	ND	ND	ND	ND	ND	ND	ND
G7(-190)	<0.1	ND	ND	ND	ND	ND	ND	ND
G8(-220)	<0.1	ND	ND	ND	ND	ND	ND	ND

See Table I for footnotes

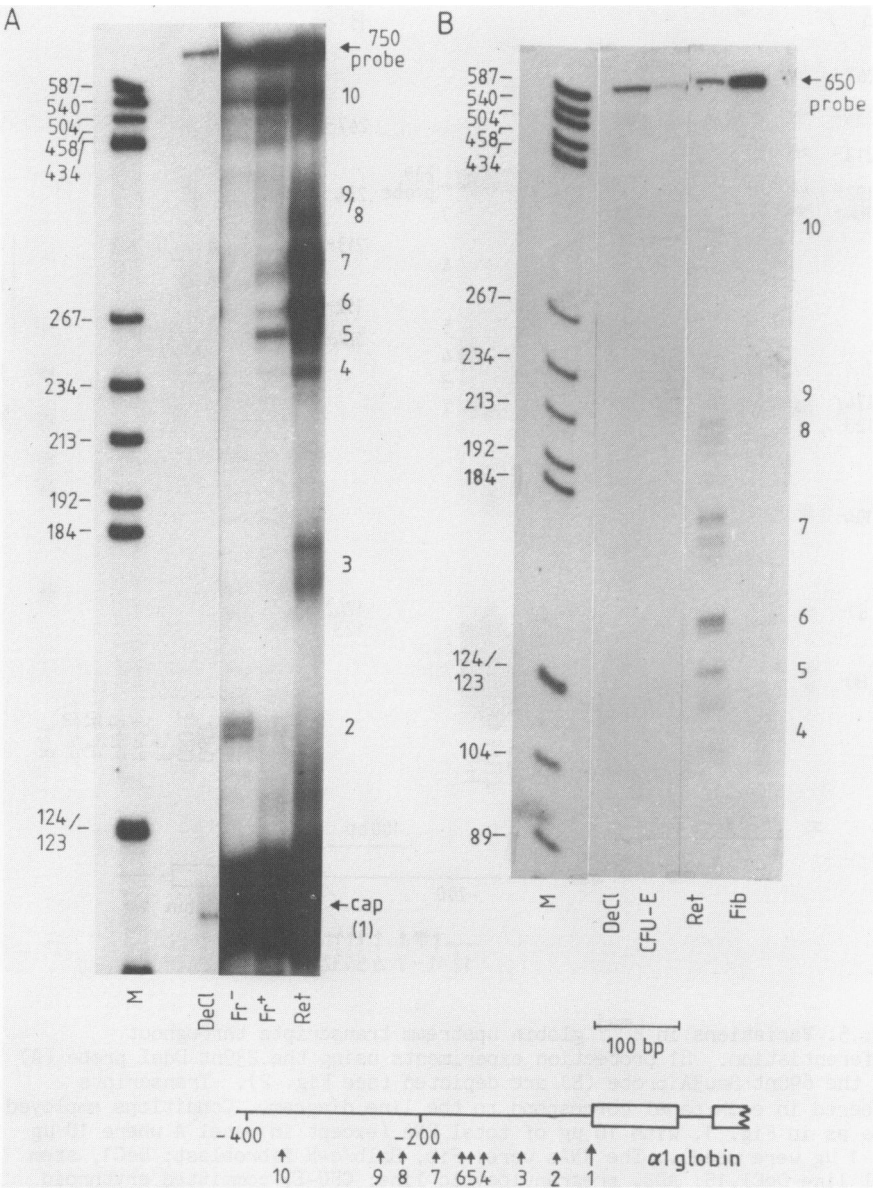

Fig. 4. Variations in α1 globin upstream transcripts throughout differentiation. S1 protection experiments using the 750nt AluI-AluI probe (A) and the 650nt PstI-PstI probe (B) are depicted (see Fig. 1). Transcripts numbered in each panel correspond to the line diagram. The RNAs used (50 µg total) were: DeC1, stem cell line DeC1.15; CFU-E, committed erythroid progenitor; Ret, reticulocyte; Fr^-, uninduced Friend cell 70710/1; and Fib, Balb/c-N cell line.

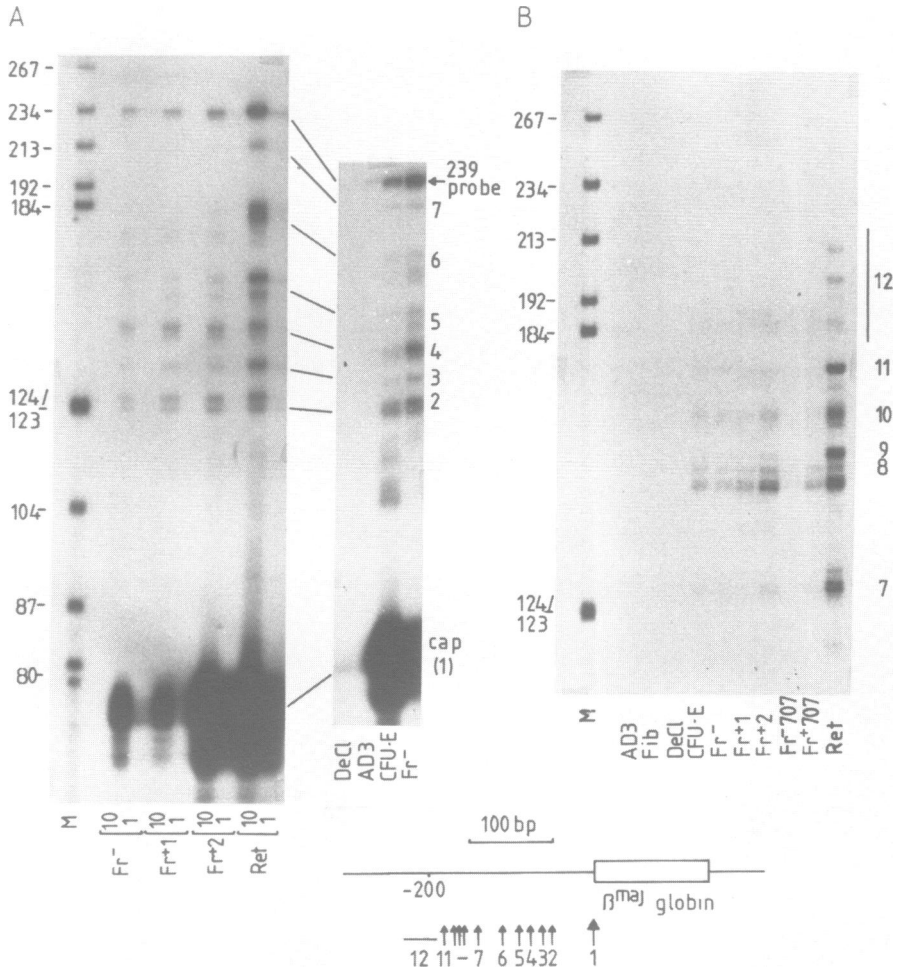

Fig. 5. Variations in β^{maj} globin upstream transcripts throughout differentiation. S1 protection experiments using the 239nt DdeI probe (A) and the 690nt Sau3A probe (B) are depicted (see Fig. 2). Transcripts numbered in each panel correspond to the line diagram. Conditions employed were as in Fig. 1, with 10 µg of total RNA (except in panel A where 10 µg and 1 µg were used). The RNAs were: Fib, Balb/c-N fibroblast; DeC1, stem cell line DeCl.15; AD3, pregranulocytic line; CFU-E, committed erythroid progenitors; Fr^{-}, FR^{+1}, FR^{+2}, Friend cells (FSD 2/3) zero, 1 or 2d after induction by HMBA; $Fr^{-}707$, $Fr^{+}707$; 0, 4d DMSO-induced Friend cell 707B10/1; Ret, reticulocyte. Note that the S1 protection using CFU-E RNA in panel A has been overdigested giving extra bands below the transcript 2 band.

by primer extension experiments using the 128nt SstII-Sau3A primer (Figure 3A) as well as by S1 protection experiments using the 1700nt PstI-PstI, 500nt XbaI-PstI, and 198nt Sau3A-Sau3A probes (Figure 3B, Figure 6 and data not shown). As with the globin genes, the minor transcripts are polyadenylated and appear clustered; however, in the case of the GSHPx gene the predominant minor transcripts (G2-4, representing individually approaching 1% of the total transcripts, Table II) occur close to the major start site (-25nt to -50nt).

Variations in α1 and β^{maj} upstream transcripts throughout differentiation

To elucidate the possible developmental significance of the upstream transcripts of the globin genes, their expression was monitored in haemopoietic stem cells (the DeCl cell lines which are capable of forming colonies comprising erythroid, myeloid and megakaryocytic cells in vitro in the presence of the growth factor IL3 (27,32)), in purified normal committed erythroid cells (CFU-Es) and in Friend cells at various stages of differentiation induced by HMBA or DMSO.

As others have also noted (eg (33)) the α globin cap site mRNA arises earlier in differentiation than the β globin cap site mRNA (cf Figures 4, 5, and Table I). There is evidence for low level cap site transcripts from both α and β globin genes in the stem cell line; this could be the result of spontaneous differentiation into the erythroid pathway due to low levels of growth factors in the serum. In terms of the level of cap site transcripts, the CFU-Es (although not visibly haemoglobinised) resemble slightly induced FLC (Figure 5, Table I). This may reflect the tendency of CFU-Es to differentiate rapidly after release of the differentiation block in vivo due to their exposure to very high erythropoeitin levels induced by the severe anaemia in the animals. Such a phenomenon has been noted before after treatment of early erythroblasts with high erythropoietin levels in vitro (34). The levels of globin cap site transcripts in the two Friend cell lines examined here, 707B10/1 and FSD 2/3, seem to imply that the latter has been arrested at a later point in erythroblast maturation (Table I).

The minor globin transcripts can be categorised in terms of their behaviour relative to the cap site transcript during erythropoeisis. The transcript most distal from the α globin promoter, α10, is present at low levels even in the stem cell line and then increases somewhat in late erythroid cells (CFU-E, FLC and reticulocytes - Figure 4; note, the α10 transcript in stem cells is not well reproduced in the photograph but it is

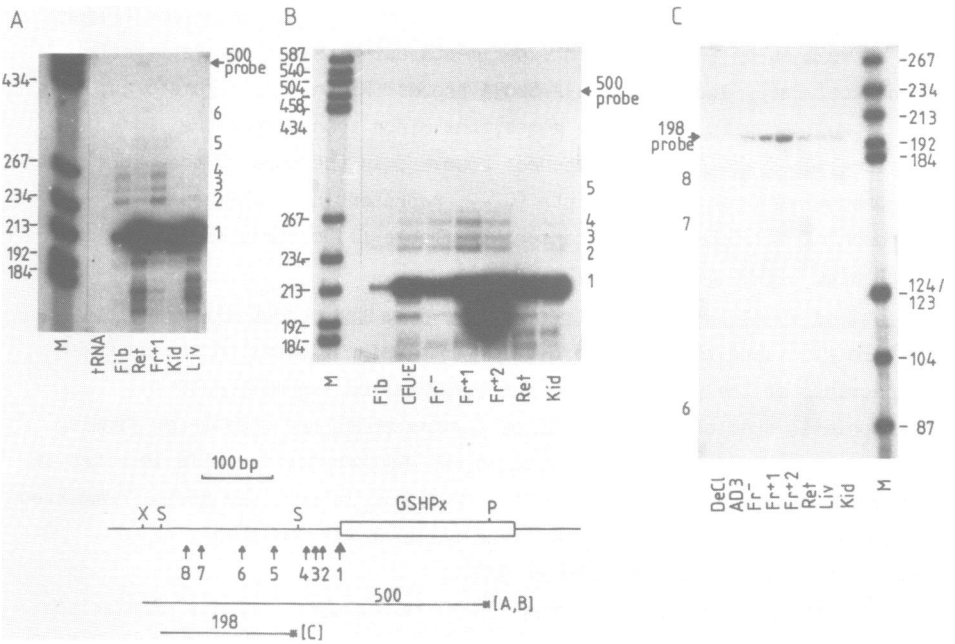

Fig. 6. Variations in GSHPx upstream transcripts throughout differentiation. Minor GSHPx gene transcripts (G2-5) were quantitated relative to the major RNA (G1) by S1 nuclease protection using a 500nt XbaI-PstI probe spanning from -285 to +215nt (A and B, see line diagram - notation as in Fig. 3). Very low level promoter distal transcripts (G6-8) were measured using a 198nt Sau3A-Sau3A probe from -243 to -45nt (C). Conditions employed were as in Fig.1 and the RNAs used (50 µg) as described in Fig. 5, also Balb/c kidney (kid) or liver (liv) RNAs.

evident on the original autoradiographs). The α2 transcript is also present in the stem cell lines; this increases considerably in uninduced FLC, but then declines dramatically in induced FLC and reticulocytes as the cap site transcript increases (Figure 4). The β2-5 group of transcripts arises somewhat later in uninduced FLC and then remains roughly constant or increases only slightly in later stages of differentiation: this is particularly obvious for transcript β4 (Figure 5 and TableI). Both the α2 and β2-5 transcripts are found in the pregranulocytic line AD3, but not in a lymphoma line, L5178Y, the macrophage cell line, WEHI3, or non-haemopoetic cells (Figure 5, and data not shown). Transcript cluster β12 is present at constant levels in erythroblasts (FLC) and is only seen to accumulate late in differentiation (reticulocytes); interestingly it is also present in the

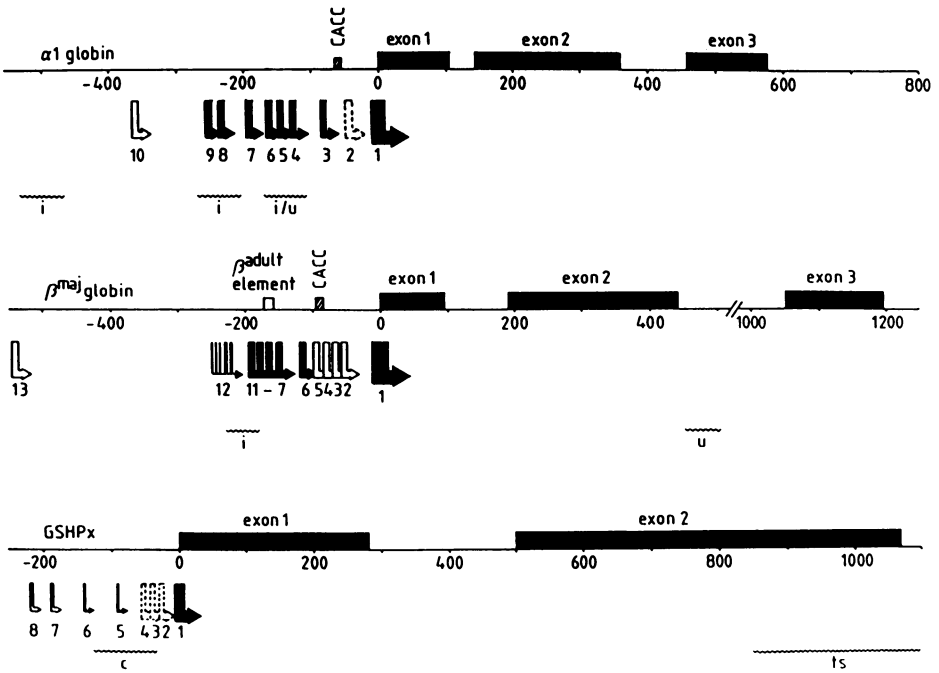

Fig. 7. Summary of α1 and β^{maj} globin and GSHPx gene upstream transcripts in relation to hypersensitive sites and consensus sequences. Each gene is represented as a solid line with exons denoted by filled boxes. Transcripts are represented by arrows (open - early and late expression; dotted - early expression, declines late; solid - induced like major RNA). DNase I hypersensitive sites are indicated below each gene by wavy lines (i - induced Friend cell; u - uninduced Friend cell; c - constitutive; ts - high expressing tissues). Data on hypersensitive sites from refs. 10,40. Location of consensus sequences are shown above each gene. β adult element represents the sequence TCCTAAGCCAGTG in β^{maj}globin which is highly conserved about 150 nt upstream of the adult mammalian β globin genes.

AD3 line. The remaining minor transcripts, α3-9 and β6-11 are undetectable in stem cells or other haemopoietic cells (data not shown) but increase relatively late in differentiation (induced FLC and reticulocytes) reaching only 5-30% of the reticulocyte level by the late erythroblast stage (induced FLC), compared to 90% for the cap site transcripts (Figures 4,5 and Table I).

To summarise, three groups of minor globin upstream transcripts can be loosely defined in terms of their behaviour during erythropoesis: some (α10, β2-5, β13) appear early and stay constant or increase slightly during differentiation; one (α2) appears early and then declines dramatically;

whereas most arise late in differentiation roughly in parallel with the cap site transcripts.

Variations in GSHPx upstream transcripts in erythroid and non-erythroid cells

GSHPx gene transcripts G2-4 can be grouped together since their levels relative to the major transcript increase in the early stages of FLC differentiation (c.f. 0, 1-2d induced FLC, Figure 6B) and then decline (c.f. CFU-E, 2d induced FLC and reticulocytes, Figure 6B). Furthermore, in low expressing cells (fibroblasts, neuroblastoma cells, the AD3 cell line, and stem cells), the relative levels of G2-4 to the major RNA are very high in contrast to the low ratio in terminally-differentiated high-expressing cells (reticulocyte, liver and kidney) (Figure 6A,B, Table II, and data not shown). Although the other GSHPx gene upstream transcripts (G5-8) occur in extremely low amounts, S1 protection experiments reveal that G6 seems to be exclusive to erythroid cells, being undetectable in fibroblast, liver or kidney (Figure 6A,B); in addition G6 appears to increase during erythroid cell differentiation (Figure 6C). The level of G7 also rises during erythroid differentiation; but unlike G6, it is also detected in kidney and to a lesser extent in liver (Figure 6C). In contrast, the level of G8 is more constant in erythroid and non-erythroid cells (Figure 6C).

DISCUSSION

How sets of genes come to be co-expressed at different times during cell differentiation is clearly an issue of major biological importance; but despite extensive work on the α and β globin genes for example, it is still not clear how this occurs. Indeed, what is known from transfection and other experiments indicates that these two genes behave differently in terms of their response to factors produced during erythroblast differentiation, to enhancers and to replication-associated events (35,36,3).

We describe here changes during erythroid cell differentiation (from stem cells to reticulocytes) in minor RNA transcripts from three genes all co-expressed at characteristically high levels in erythroid cells. Using S1 nuclease protection and primer extension to quantitate RNA steady-state levels (summarised in Tables I and II and Figure 7), we have shown that particular upstream regions of the α and β globin genes and the GSHPx gene become active in transcription very early in differentiation (generating transcripts α10, β2-5, β13); some regions are active early and show a peak of transcript accumulation before the terminal stages of differentiation

(α2, G2-4), whereas others become active more-or-less in parallel with the major (cap) transcript (α3-9, β6-11, G6), although within the latter category some transcripts are seen in advance of others (α4, α5, β8). In addition, the promoter region of the GSHPx gene was found to generate antisense transcripts, the major one of which rises during induced erythroid differentiation. No other study has been conducted making a comparison of minor transcripts from three genes throughout differentiation along the erythroid lineage: the only previous analyses of the mouse β^{maj} globin gene have mapped upstream transcripts that parallel the major mRNA (see reference 20 - defined by us as transcripts β6-12) but have failed to detect the early upstream transcripts.

At present we can only speculate about the possible functions of these different classes of upstream transcripts in activation of the erythroid genes. As already discussed earlier, upstream transcripts have been detected from many genes and in some cases, as here, the minor RNAs accumulate differentially compared to the major mRNA (13,37). Moreover, Allan **et al** (24) have shown that the ratio of the minor RNA originating at -200 nt to the major RNA from the ε globin gene can be affected by SV40 enhancer sequences in cis or by the adenovirus E1A gene in trans. Carlson and Ross (38) have also recently demonstrated a possible link between a decrease in upstream transcripts from around the -200 nt position in vitro and a point mutation at -202 nt in the human ϒ globin gene associated with hereditary persistence of foetal haemoglobin. Such studies suggest that upstream transcripts are mechanistically involved in transcriptional regulation or at least are reflecting molecular events at the promoter such as the binding of trans-acting factors.

Carlson and Ross (19,20) have suggested that RNA polymerase III is responsible, at least in vitro, for the generation of upstream β globin transcripts; however, their data only relates to transcripts that we designate β9-12. In contrast, Hess **et al** (21) have shown that human α globin transcripts are produced by a polII-dependent mechanism.

In the case of the β globin gene there is considerable evidence suggesting that the region from which upstream transcripts are generated is not essential for erythroid-specific regulation (3,4,6 and 8). Hence, appropriate regulation of β globin genes deleted in their 5' flanking sequences was observed in mice transgenic for the human gene deleted to -48nt (8) and in Friend cell transfectants containing the rabbit β globin gene deleted to -58nt (6) or the mouse β^{maj} globin gene truncated at -47nt

(4). However, it is possible that additional tissue-specific aspects of β globin gene regulation are conferred by sequences near minor RNA start sites. Indeed, human β globin sequences 5' to the translational initiation site have been shown to confer erythroid-specific expression (45) or inducibility (6) in FLC.

The hypothesis that the regions of the globin and GSHPx genes giving rise to minor transcripts may be involved somehow in regulation has been given support by correlations with the sites of developmentally-regulated DNaseI hypersensitive sites (Figure 7). Analysis of sequences around upstream RNA start sites of the globin and GSHPx genes revealed associations with known promoter control elements. Transcripts α2 and β2-4 arise from regions containing important promoter elements (TATA, CCAAT). Also, the α3 and β5 transcripts start near the functionally important globin promoter motif CACCCT (1). Interestingly, the inducible G6 minor RNA also arises from the same sequence. Transcript α10 at -360nt starts close to a SV40 core enhancer sequence that is also in the region of an erythroid-specific HSS (reference 33 - Figure 7). Moreover, comparison of the human β globin and mouse β^{maj} globin genes with respect to upstream minor RNAs (18, 19) reveals a close correspondence with our transcripts designated β8-11 (-160nt to -180nt). Moreover, this exactly correlates with a region that is highly conserved upstream of mammalian adult β globin genes (shown as β adult element in Figure 7) (39). As far as the GSHPx gene is concerned, most of the upstream transcripts originate from a GC-rich region containing potential Sp1-binding sites (10). The aim of our current work is to elucidate the roles of such sequences around upstream transcript start sites in globin and GSHPx gene regulation at different stages of differentiation using functional assays.

ACKNOWLEDGEMENTS

We thank Dr R Padgett (Department of Microbiology and Immunology, University of North Carolina, Chapel Hill) for providing clones containing the β globin genes. We would also like to thank Dr John Paul for critical reading of the manuscript.

This work has been supported by the Cancer Research Campaign. IC is supported by an MRC studentship and an ICI scholarship.

REFERENCES
1. Dierks, P., Van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and Weissman, C. (1983) Cell 32, 695-706.
2. Chao, M.V., Millan, P., Charnay, P., Maniatis, T. and Axel, R. (1983) Cell 32, 483-493.
3. Charnay, P., Treisman, R., Mellon, P., Chao, M., Axel, R. and Maniatis, T. (1984) Cell 38, 251-263.
4. Charnay, P., Mellon, P. and Maniatis, T. (1985) Mol. Cellular Biol. 5, 1498-1511.
5. Wright, S., deBoer, E., Grosveld, F.G. and Flavell, R.A. (1983) Nature 305, 333-336.
6. Wright, S., Rosenthal, A., Flavell, R. and Grosveld, F. (1984) Cell 38, 265-273.
7. Chada, K., Magram, J., Raphael, K., Radice, G., Lacy, E. and Constantini, F. (1985) Nature 314, 377-380.
8. Townes, T.M., Lingrel, J.B., Chen, H.Y., Brinster, R.L. and Palmiter, R.D. (1985) EMBO J. 4, 1715-1723.
9. Venta, P.J., Montgomery, J.C., Hewett-Emmett, D. and Tashian, R.E. (1985) Biochem. Biophys. Acta 826, 195-201.
10. Chambers, I., Frampton, J., Goldfarb, P., Affara, N., McBain, W., Harrison, P.R. (1986) EMBO J. 5, 1221-1227.
11. Affara, N., Fleming, J., Goldfarb, P.S., Black, E., Thiele, B. and Harrison, P.R. (1985) Nucleic Acids Res. 13, 5629-5744.
12. Contreras, R., Gheysen, D., Knowland, J., van de Voorde, A. and Fiers, W. (1982) Nature 300, 500-505.
13. Osborne, T.F. nd Berk, A.J. (1983) J. Virol. 45, 594-599.
14. Mitchell, P.J., Carothers, A.M., Han, J.H., Harding, J.D., Kas, E., Venolia, L. and Chasin, L.A. (1986) Mol. Cellular Biol. 6, 425-440.
15. Chang, D.D., Hixson, J.E. and Clayton, D.A. (1986) Mol. Cellular Biol. 6, 294-301.
16. Allan, M., Grindlay, G.J., Stephani, L. and Paul, J. (1982) Nucleic Acids Res. 10, 5133-5147.
17. Allan, M., Lanyon, W.G. and Paul, J. (1983) Cell 35, 187.
18. Grindlay, G.J., Lanyon, W.G., Allan, M. and Paul, J. (1984) Nucleic Acids Res. 12, 1811.
19. Carlson, D.P. and Ross, J. (1983) Cell 34, 857-864.
20. Carlson, D.P. and Ross, J. (1984) Proc. Natl. Acad. Sci. USA 81, 7782-7786.
21. Hess, J., Perez-Stable, C., Deisseroth, A. and Shen, C-K.J. (1985) Nucleic Acids Res. 13, 6059-6074.
22. Allan, M. and Paul, J. (1984) Nucleic Acids Res. 12, 1193-1200.
23. Allan, M., Montague, P., Grindlay, G.J., Sibbet, G., Donovan-Peluso, M., Bank, A. and Paul, J. (1985) Nucleic Acids Res. 13, 6125-6136.
24. Allan, M., Zhu, J-de, Montague, P. and Paul, J. (1984) Cell 38, 399-407.
25. Kollias, G., Sekeris, C.E. and Grosveld, F.G. (1985) Nucleic Acids Res. 13, 7993-8005.
26. Zhu, J-de, Allan, M. and Paul, J. (1984) Nucleic Acids Res. 12, 9191.
27. Spooncer, E., Boettiger, D. and Dexter, T.M. (1984) Nature 310, 228-230.
28. Dexter, T.M., Garland, J.M., Scott, D., Scolnick, E. and Metcalf, D. (1980) J. exp. Med. 152, 1036-1047.
29. Nijhof, W. and Wierenga, P.K. (1983) J. Cell Biol. 96, 386-392.

30. Iscove, N.N. and Sieber, F. (1975) Exp. Hematol. 3, 32-43
31. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in Molecular Cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
32. Spooncer, E., Heyworth, C.M., Dunn, A. and Dexter, T.M. (1986) Differentiation, in press.
33. Kaneda, T., Murate, T., Sheffery, M., Brown, K., Rifkind, R.A. and Marks, P.A. (1985) Proc. Natl. Acad. Sci. USA 82, 5020-5024.
34. Harrison, P.R., Conkie, D. and Paul, J. (1973) In: Brit. Soc. Dev. Biol. Symp. The Cell Cycle in Development and Differentiation (eds. M. Balls and F.S. Billett) Camb. Univ. Press pp341-364.
35. Humphries, R.K., Ley, T., Turner, P., Moulton, A.D. and Nienhuis, A.W. (1982) Cell 30, 173-183.
36. Treisman, R., Green, M.R. and Maniatis, T. (1983) Proc. Natl. Acad. Sci. USA 80, 7428-7432.
37. Ghosh, P. and Lebowitz, P. (1981) J. Virol. 40, 224-240.
38. Carlson, D.P. and Ross, J.R. (1986) Mol. Cellular Biol. 6, 3278-3282.
39. Mosconas, N., de Boer, E. and Flavell, R.A. (1982) Nucleic Acids Research 10, 2109.
40. Sheffery, M., Rifkind, R.A. and Marks, P.A. (1983) Proc. Natl. Acad. Sci. USA 80, 3349-3353.
41. Nishioka, Y. and Leder, P. (1979) Cell 18, 875-882.
42. Jahn, C.L., Hutchison III, C.A., Phillips, S.J., Weaver, S., Haigwood, N.L., Voliva, C.F. and Edgell, M.H. (1980) Cell 21, 159-168.
43. Gilmour, R.S., Spandidos, D.A., Vass, J.K., Gow, J.W. and Paul, J. (1984) EMBO J. 3, 1263-1272.
44. Konkel, D.A., Maizel, J.V. and Leder, P. (1979) Cell 18, 865-873.
45. Rutherford, T. and Nienhuis, A. (1987) Mol. Cellular Biol. 7, 398-402.