The EMBO Journal Vol.18 No.14 pp.4035–4048, 1999

An insulator element and condensed chromatin region separate the chicken β-globin locus from an independently regulated erythroid-specific folate receptor gene

Marie-Noëlle Prioleau[1], Pascale Nony[2], Melanie Simpson and Gary Felsenfeld[3]

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0540, USA

[1]Present address: Laboratoire de Génétique Moléculaire, CNRS URA 1302, 46 rue d'Ulm, 75230 Paris Cedex 05, France
[2]Present address: Laboratoire de Thérapie Génique, Bâtiment Jean Monnet-Hôtel Dieu 30, boulevard Jean Monnet, 44035 Nantes Cedex 1, France

[3]Corresponding author
e-mail: gary.felsenfeld@nih.gov

We have identified a folate receptor gene upstream of the chicken β-globin locus and separated from it by a 16 kbp region of silent chromatin. We find that this receptor is expressed only at a stage of erythroid differentiation (CFU-E) preceding the activation of β-globin genes, consistent with the role of folate receptors in proliferation. This discovery raises the question of how these two loci are regulated during erythropoiesis. Our data suggest that the folate receptor gene and the β-globin locus are regulated independently. We show that a 3.3 kbp DNA region upstream of the folate receptor gene is sufficient to induce strong expression of a transgene in CFU-E stage cells. We also find that the region between the β-globin locus and the folate receptor gene is fully methylated and condensed at this stage of differentiation. Its 3′ boundary coincides with the 5′ β-globin insulator. We speculate that the 5′ β-globin boundary element might be important for the proper regulation of two adjacent domains activated at two different stages during differentiation.

Keywords: β-globin locus/boundary/condensed chromatin/domain/methylation

Introduction

Chromosomal position effects in mammals have been observed frequently in transgenic animals and in many diseases caused by natural mutations (Ton *et al*., 1991; Joos *et al*., 1992; Tommerup *et al*., 1993; Foster *et al*., 1994; Wagner *et al*., 1994; Fantes *et al*., 1995). These observations demonstrate that in addition to its role as a structural scaffold important for compaction, chromatin plays a major role in the utilization of genetic information. One attractive hypothesis is that chromosomes are subdivided into independent functional domains, which are either permissive or repressive toward gene expression. This model is reinforced by the observation that transcriptionally active genes are embedded in DNase I-sensitive domains, extending many kilobases beyond the genes (Stalder *et al*., 1980; Lawson *et al*., 1982; Jantzen *et al*., 1986; Stratling *et al*., 1986; Levy-Wilson and Fortier, 1989).

For a domain to be a truly independent unit, it should contain all of the regulatory elements necessary for the correct expression of the genes located within; some well characterized domains do maintain a higher level of tissue- and stage-specific expression when integrated into the genome than constructs containing only proximal *cis*-regulatory elements (Greaves *et al*., 1989; Bonifer *et al*., 1994; Mason *et al*., 1995; Dillon *et al*., 1997). DNA sequences that induce efficient gene expression independently of the site of integration have been identified in a number of different domains. Usually they are marked by a series of DNase I hypersensitive sites. Such clusters of regulatory elements required to induce copy-number-dependent expression have been designated as locus control regions (LCRs), and have been most thoroughly investigated in the human β-globin gene locus (Forrester *et al*., 1987, 1989; Grosveld *et al*., 1987; Festenstein *et al*., 1996).

The β-globin LCR, also studied in mouse and chicken (Moon and Ley, 1990; Hug *et al*., 1992; Mason *et al*., 1995), can activate genes located as far as 50 kbp away, and is responsible for the proper expression of the β-globin genes (Epner *et al*., 1998; Reik *et al*., 1998). This raises the problem of how activation can be confined to a particular set of genes within the domain without affecting adjacent genes. It has been proposed that DNA sequences at the boundaries of an active chromatin domain may help in their establishment and maintenance (Geyer, 1997). The chicken β-globin locus spans 33 kbp (Figure 1A), with boundaries identified by transitions to DNase I inaccessibility and low levels of histone acetylation (Hebbes *et al*., 1994). At the 5′ boundary, a DNA sequence element marked by a constitutive DNase I hypersensitive site (HS4) has been shown to possess the properties of an insulator; it blocks enhancer action, but only when it lies between the enhancer and a promoter (Chung *et al*., 1993).

We analyzed the region upstream of the well defined 5′ boundary of the chicken β-globin locus, and found a region of condensed chromatin extending without interruption 16 kbp upstream of the β-globin 5′ boundary. Immediately beyond this we discovered a novel locus that is activated during erythropoiesis. This locus encodes a folate receptor and has a different program of expression from the globin genes. The chicken folate receptor is expressed in erythroid precursor cells in the developing embryo and is repressed during terminal differentiation. We identify a strong regulatory element upstream of the folate receptor which by itself directs expression of a linked gene, fairly independently of the chromosomal site of integration in an erythroid cell line. We also show that at the time when the folate receptor gene is activated, the region between

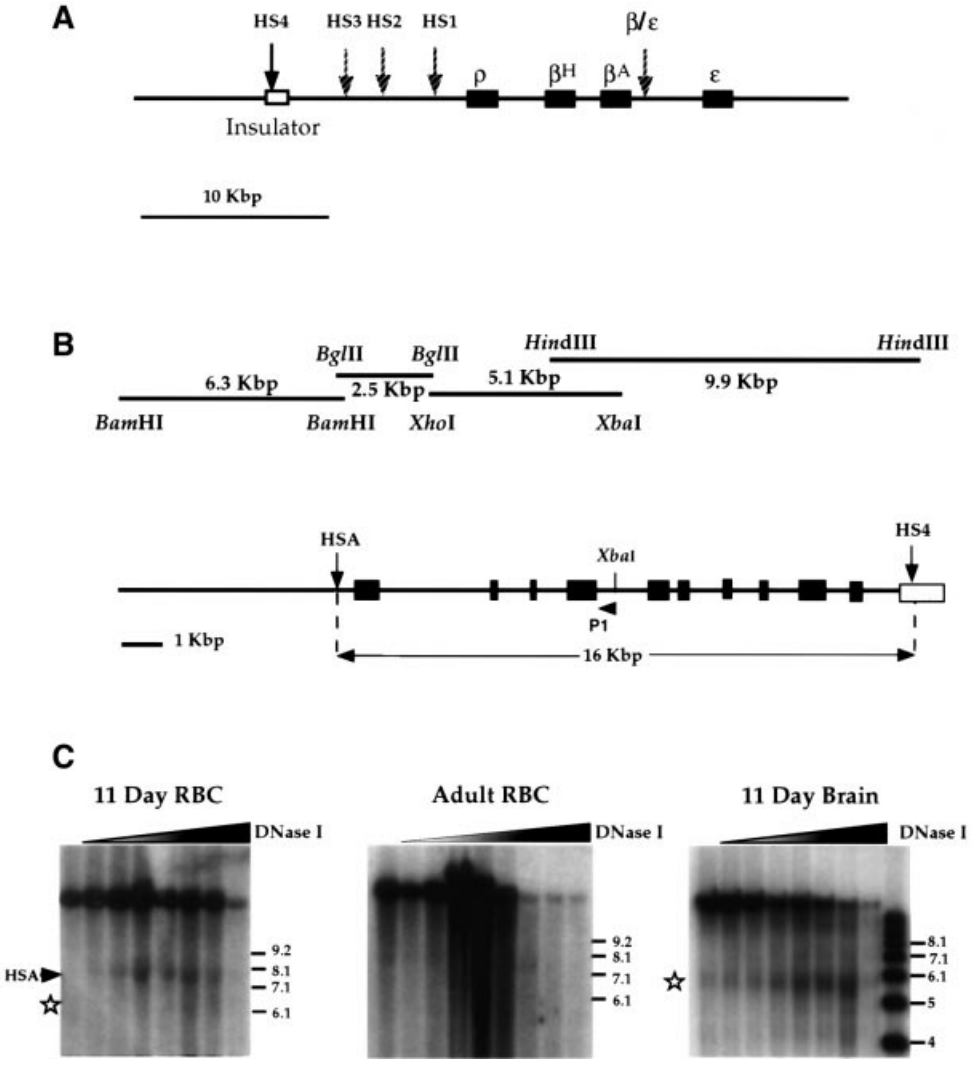


Fig. 1. A tissue-specific hypersensitive site (HSA) is formed 16 kbp upstream of the chicken β-globin insulator. (**A**) Map of the chicken β-globin domain. The constitutive 5′ HS4 is shown with a closed arrow, erythroid-specific hypersensitive sites are shown with stippled arrows. Erythroid-specific hypersensitive sites constitute the chicken β-globin LCR. Open box represents the 1.2 kbp insulator. The four β-globin genes are indicated by solid boxes. (**B**) The four clones isolated from partial libraries are represented at the top of the figure. The probe (P1) and the *Xba*I site used for the analysis of DHSs in different cells in (C) are represented. Black boxes represent repeated sequences identified between HS4 and HSA. (**C**) DNase I hypersensitive site mapping of the DNA region upstream of the chicken β-globin insulator. DNA from DNase I-treated nuclei was digested with *Xba*I and probed with probe P1. DNA size standards are indicated on the right in kbp. The hypersensitive band corresponding to HSA appears around 7.9 kbp. Stars indicate weak hypersensitive sites. DHSs detected in Day 11 brain do not correspond to DHSs observed in 11 Day RBC.

the β-globin locus and the folate receptor locus is organized in a micrococcal nuclease-resistant chromatin structure, and the DNA of this region is fully methylated. Our results indicate that the β-globin and the folate receptor domains are regulated independently despite their proximity. We propose that the 5′ β-globin insulator and the condensed region are important in assuring the proper expression of these two loci by establishing two units of independent gene activity.

Results

Identification of an active erythroid-specific locus 16 kbp upstream of the β-globin locus

It has been demonstrated previously that for at least 4 kbp upstream of the chicken β-globin locus, chromatin is DNase I-resistant and its histones are hypoacetylated (Hebbes *et al*., 1994). This study defined the beginning of a new condensed chromatin region. The first step in understanding how adjacent loci are organized on the chromosome was the identification of the 5′ boundary of this condensed region. We searched for local perturbation in chromatin structure that might indicate the presence of active loci by mapping DNase I hypersensitive sites (DHS). The proximal DHS (designated HSA) was found 16 kbp upstream of the insulator in 11-day embryonic erythrocytes, or red blood cells (RBCs) (Figure 1B and C, 11 Day RBC), which are transcriptionally active and expressing the β^A and β^H globin genes (Hesse *et al*., 1986; Mason *et al*., 1995). To determine whether HSA is developmentally regulated, DHSs were mapped in other cells. In contrast with all the sites observed within the β-globin locus, HSA is not present in terminally differentiated adult erythrocytes (Figure 1C, Adult RBC) or

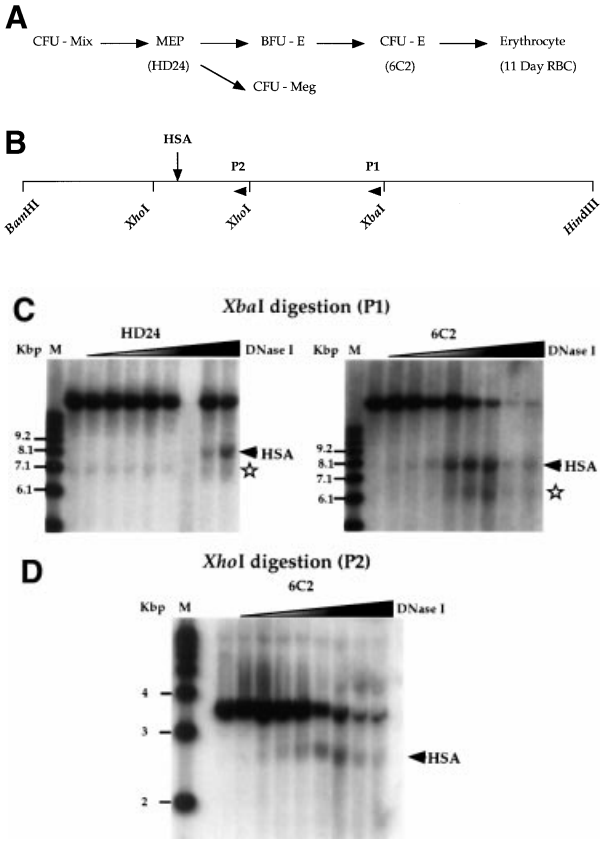


Fig. 2. HSA is formed precociously during erythroid differentiation. (**A**) The diagram represents different stages of erythropoiesis from the cell lines HD24 (Metz and Graf, 1991) and 6C2 (Beug *et al*., 1982) corresponding to the MEP and the CFU-E stages, respectively. (**B**) The map indicates the restriction sites and probes used to analyze HSA. The *Xba*I site and probe P1 used in (C) are indicated. The 3692 bp *Xho*I parental fragment and probe P2 used in (D) are also shown. The distance between *Xho*I and *Xba*I is 5150 bp. (**C**) DNA from DNase I-treated nuclei (HD24 or 6C2) were digested with *Xba*I and probed with probe P1. Stars indicate weak DHSs. (**D**) Precise localization of HSA. DNA from DNase I-treated 6C2 nuclei used in (C) was digested with *Xho*I and probed with probe P2. The hypersensitive band corresponding to HSA appears at ~2740 bp.

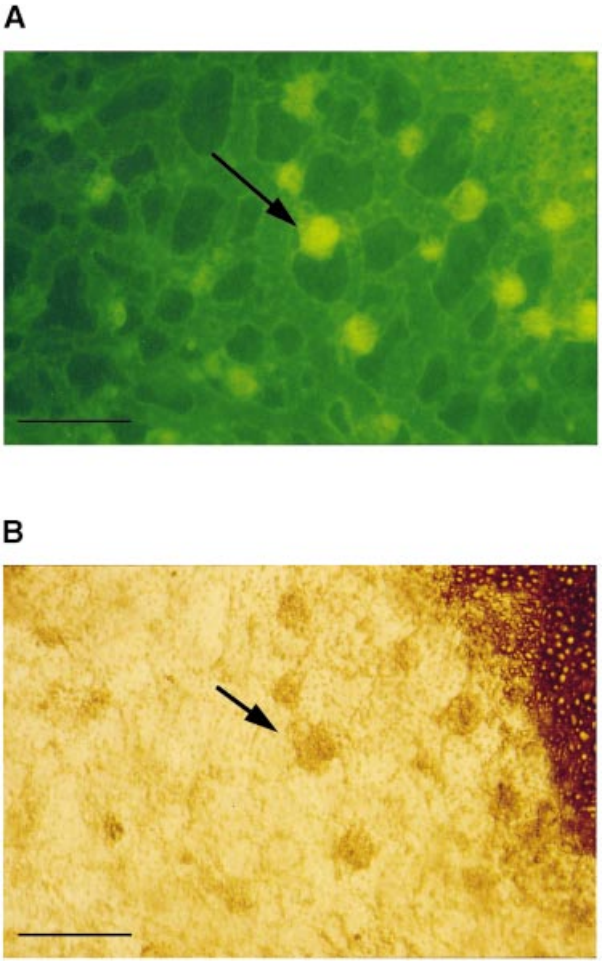


Fig. 3. Folate receptor mRNA expression in the blood islands of a chick embryo. (**A**) Whole-mount, *in situ* hybridization of folate receptor mRNA expression in the area pellucida of a 48 h embryo. Expression in the blood islands is indicated (typical example marked by arrow). Scale bar represents 0.1 mm. (**B**) The same embryo counter stained with Giemsa to distinguish the blood islands. Note co-localization with folate receptor mRNA expression (arrow) seen in (A).

embryonic brain (Figure 1C, 11 Day Brain). Thus, in contrast to HS4, the hypersensitive site that marks the 5′ boundary of the β-globin locus, HSA is not a constitutive DHS.

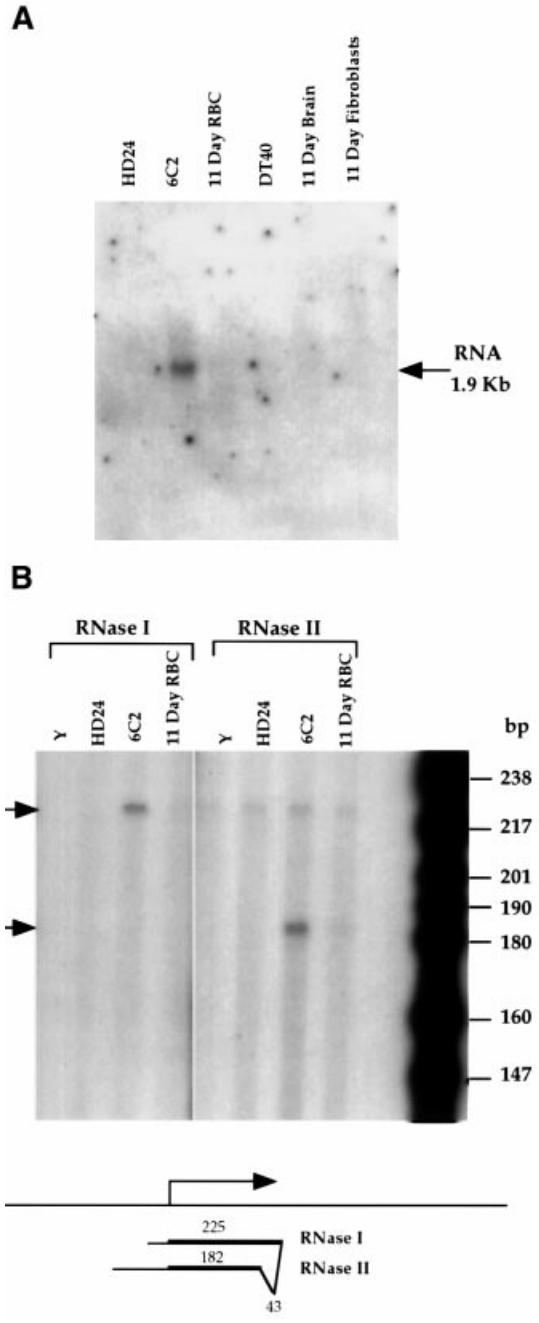
The regulation of HSA was also analyzed in two virally transformed chicken erythroid cell lines: HD24 cells are arrested at the multipotent erythroid-myeloid precursor cell (MEP) stage; 6C2 cells are arrested at the colony forming unit erythrocyte stage cell line (CFU-E) (Figure 2A). These cell lines are model systems that have been exploited in numerous studies to examine chromatin structure and the regulation of erythroid-specific genes at early developmental stages (see Discussion). HSA is present in both of these cell lines (Figure 2C). The location of HSA was determined more precisely by utilizing a different probe and restriction digest combination (Figure 2D). These observations define a 16 kbp region which is bordered in erythrocytes by two hypersensitive sites: HSA and the insulator, HS4 (Figure 1B). The 5′ end of this region is delimited during erythropoiesis by HSA. We thought that its presence could signal the beginning of a new active domain, suggesting that it could be a regulatory element for a gene important during erythroid differentiation. We sequenced a cloned 6.3 kbp fragment that contains HSA near its 3′ end (Figure 1B, *Bam*HI clone), and found that it has considerable identity (see below) to known human genes of the folate receptor family (Shen *et al*., 1994). It is very likely, therefore, that these sequences encode a chicken folate receptor. The orientation of this novel coding region is opposite to that of the β-globin genes.

To confirm erythroid-specific expression of this folate receptor gene, we performed wholemount *in situ* hybridization on 1- to 3-day-old chick embryos. The first erythrocyte precursors are detectable in the blastoderm after ~24 h (BFU-E stage), later developing into colonies of rapidly proliferating cells, known as blood islands, that precede terminal differentiation (Sabin, 1920; Lucas and Jamroz, 1961; Samarut *et al*., 1979). Expression of folate receptor mRNA was restricted to yolk sac blood islands, clearly recognizable by their characteristic cobblestone appearance when stained with Giemsa (Sabin, 1920) (Figure 3). No significant mRNA expression was detected in either

pre-blood island blastoderm or circulating embryonic red cells.

Using Northern blot analysis, we found that transcripts of this gene are detectable only in 6C2 cells (Figure 4A), consistent with the observation in embryos and erythrocytes that the expression of the folate receptor gene is tissue- and stage-specific. In order to characterize this coding region fully, we identified the site of initiation of transcription by RNase mapping. We used two different overlapping probes to analyze transcripts in the erythroid lines HD24 and 6C2 and in erythrocytes (Figure 4A, 11 Day RBC), loading a quantity of RNA for each sample corresponding to the same number of nuclei (see Materials and methods). RNase mapping with the two probes identified the same site for transcriptional initiation and also confirmed that expression of this gene is only detectable in 6C2 cells (Figure 4B), in contrast to the presence of HSA in both cell lines. The alignment of the chicken DNA sequence with the three human known isoforms of the folate receptor and the use of the GT–AG rule allowed a prediction of the exon–intron junctions (Figure 4C). These junctions were confirmed by reverse transcriptase polymerase chain reaction (RT–PCR) on the total RNA of 6C2 cells (data not shown). We found a polyadenylation signal downstream of the stop codon of the coding sequence. The sum of the lengths of the four exons leads to a predicted 1.9 kb transcript, which is the size observed in the Northern blot analysis (Figure 4A and C). The protein is composed of three regions: a conserved central region, and variable N- and C-terminal regions. The identity found in the conserved region between the different human isoforms and the chicken gene is between 65 and 70% (Figure 5A). One of the human isoforms corresponds to a truncated form of the folate receptor (FRγ), which is not bound to the cellular membrane. The other two isoforms (FRα and FRβ) possess a signal for attachment to a glycosylphosphatidylinositol (GPI) membrane anchor at their C-termini (Shen *et al.*, 1994). As can be seen in Figure 5B, the C-terminal region of the chicken folate receptor contains such an attachment signal sequence, suggesting that it is anchored to the extracellular membrane.

In summary, we have identified a gene 16 kbp upstream of the 5′ boundary of the β-globin locus that is expressed at a specific stage of erythroid development. Terminal differentiation is accompanied by the activation of globin



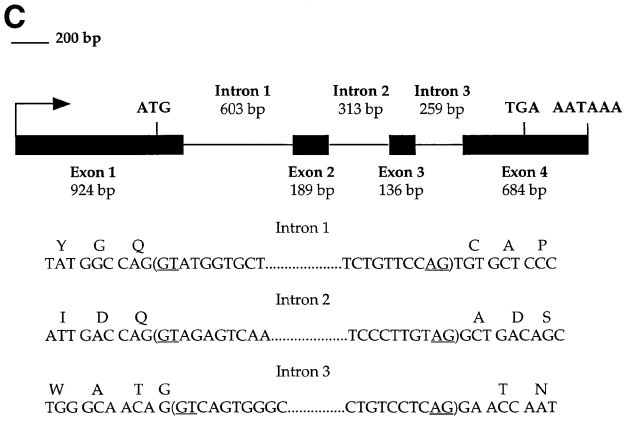


Fig. 4. An active erythroid locus is located 16 kbp upstream of the β-globin locus. (**A**) Expression of the folate receptor gene in different cell lines and tissues. Thirty micrograms of total RNA was analyzed by Northern blot hybridization with a probe covering the first exon of the chicken folate receptor gene. (**B**) Determination of the transcriptional initiation site of the folate receptor. An RNase protection experiment was performed using two overlapping riboprobes. These probes were hybridized with total RNA extracted from HD24, 6C2 or 11 Day RBC. For each lane, the amount of RNA corresponds to the same number of nuclei (see Materials and methods). Y corresponds to the control with yeast RNA. The RNase I probe protects 225 bp of the first exon, and the RNase II probe 182 bp of this exon. The difference in the size of the protected fragments reflects the difference of the 3′ extension of the two probes. (**C**) Organization of the chicken folate receptor gene. The gene is composed of four exons (solid boxes). A consensus polyadenylation signal was found downstream of the STOP codon. Sequences of the exon/intron boundaries are represented at the bottom of the figure. The chicken folate receptor transcript DDBJ/EMBL/GenBank accession No. is AF160960.

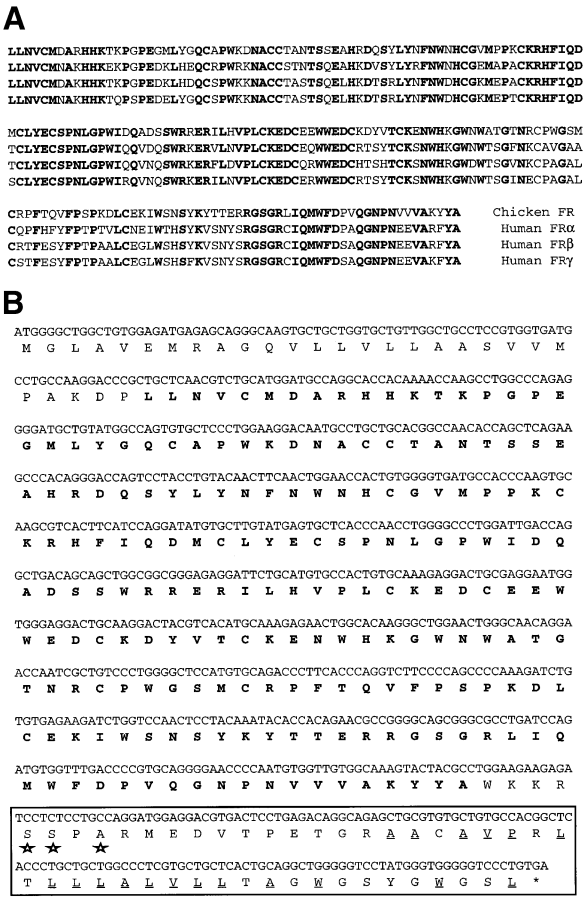


Fig. 5. The chicken folate receptor identified is a membrane-bound isoform. (**A**) Amino acid alignment of the conserved region of the chicken folate receptor with human FRα, FRβ and FRγ isoforms. Bold letters indicate amino acids conserved among all FR isoforms. (**B**) Total amino acid sequence of the chicken folate receptor. Bold letters indicate the amino acids contained in the central conserved region. The signal for attachment to a glycosyl phosphatidyl inositol (GPI) membrane anchor is surrounded by an open box. This signal has been characterized in a few model proteins and is composed of a moderately hydrophobic C-terminal region of 10–20 amino acids, a site of proteolysis/GPI modification, and a spacer of 8–12 amino acids that separates the modification site from the hydrophobic sequence (Yan and Ratnam, 1995). Stars indicate possible sites for proteolysis/GPI modification. Amino acids G, A, D, S and N are efficient sites of proteolysis, and C and V are low efficiency sites of proteolysis. Underlined amino acids are C-terminal hydrophobic amino acids.

genes with the repression of this folate receptor gene. The proximity of two loci and their activation in the same developmental pathway but at two different stages of differentiation suggests that complex controls of gene regulation are employed which are not yet understood.

A 3.3 kbp region upstream of the folate receptor gene contains regulatory elements sufficient to induce strong expression in 6C2 cells

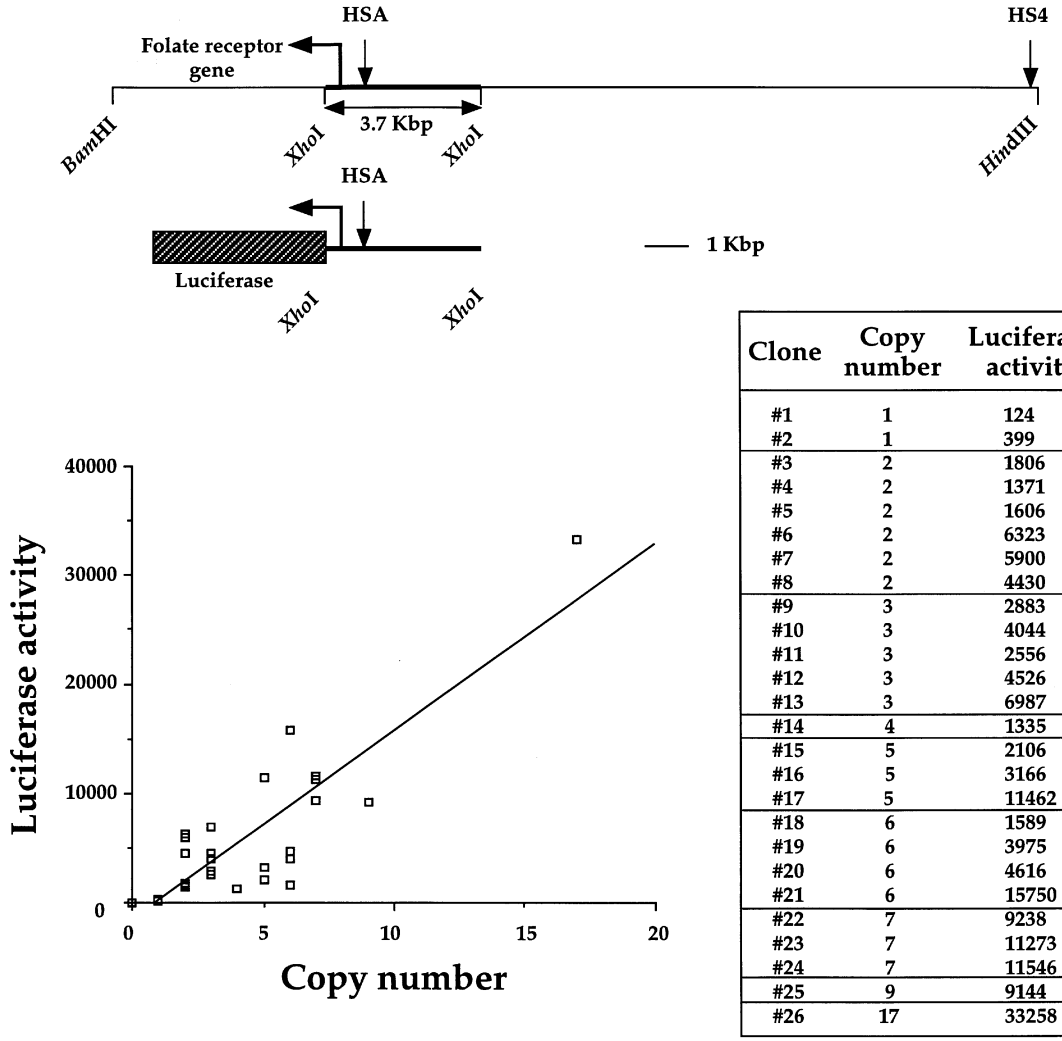
The proper developmental regulation of the chicken β-globin locus (33 kbp) necessitates a complex combination of upstream regulatory elements marked by upstream hypersensitive sites (HS1-3) and the β/ε enhancer in addition to globin gene promoters (Mason *et al*., 1995). This model implicates regulatory elements that function over long distances (e.g. the distance between HS3 and the ε promoter is 20 kbp; see Figure 1A). The identification of a gene located 16 kbp upstream of the β-globin domain and activated just before β-globin expression during erythropoiesis raises the question of whether the folate receptor shares regulatory elements with β-globin genes. It is equally possible that the insulator and condensed chromatin lying between these regions may generate a physical shield, causing the two loci to be regulated independently. The identification of potent regulatory elements located proximally to the folate receptor gene would suggest that these, and not the β-globin elements, regulate folate receptor expression. We searched for additional regulatory elements downstream of the folate receptor gene, but no DHSs were identified in 6C2 cells within 5.5 kbp downstream of the polyadenylation signal (data not shown). Thus, the only detectable hypersensitive site near the folate receptor gene is HSA.

We decided, therefore, to fuse only the upstream region of the gene, which contains HSA, to a luciferase reporter gene. The fusion contained 3.3 kbp upstream of the folate receptor transcriptional start site and 0.4 kbp downstream (Figure 6). The linearized plasmid containing the promoter and the luciferase reporter gene was stably transfected into 6C2 cells. The number of integrated copies and the luciferase activity were determined for each resulting independent clone (see Materials and methods; Figure 6). All of the clones analyzed expressed high levels of luciferase activity, suggesting that most of the regulatory elements required for transcription of the folate receptor gene were included in the construction.

We asked whether the level of expression obtained was directly related to the number of copies integrated into the genome of 6C2 cells. A good correlation between gene copy number and luciferase expression is observed. The regression line produced is shown in Figure 6 and has a correlation coefficient of 0.85, which is statistically significant at a confidence level of >99%. This compares favorably with similar assays for copy-number-dependence of expression in hematopoietic cell lines (e.g. see May and Enver, 1995). The copy-number-dependent expression induced by this sequence suggests that it contains powerful regulatory elements, although it should be pointed out that in all stable transformation experiments of this kind, the construct may have preferred sites of integration within the genome that support greater activity. Furthermore, we obviously may have excluded additional regulatory elements that may be located much further downstream of the gene. With these reservations, the results support the conclusion that this upstream region could be sufficient to direct proper expression of the folate receptor *in situ*.

A 16 kbp region located between the β-globin insulator and the folate receptor is condensed in 6C2 cells

Overexpression of the folate receptor has been observed in several human tumor cells (Campbell *et al*., 1991; Coney *et al*., 1991; Weitman *et al*., 1992; Ross *et al*., 1994). It has been proposed that the folate receptor plays a role in cell proliferation. We find (Figure 3) that a folate receptor gene is expressed at a stage of cell proliferation which precedes the final differentiation and expression of globin genes during erythropoiesis in chicken embryos. The role of the folate receptor in proliferation suggests that proper repression of its gene may be crucial for



Clone	Copy number	Luciferase activity
#1	1	124
#2	1	399
#3	2	1806
#4	2	1371
#5	2	1606
#6	2	6323
#7	2	5900
#8	2	4430
#9	3	2883
#10	3	4044
#11	3	2556
#12	3	4526
#13	3	6987
#14	4	1335
#15	5	2106
#16	5	3166
#17	5	11462
#18	6	1589
#19	6	3975
#20	6	4616
#21	6	15750
#22	7	9238
#23	7	11273
#24	7	11546
#25	9	9144
#26	17	33258

Fig. 6. A 3.3 kbp region upstream of the folate receptor gene is sufficient to induce strong copy-number-dependent expression at the CFU-E stage. A 3.7 kbp *Xho*I fragment containing HSA and the initiation site of the folate receptor gene was fused to a luciferase reporter gene. The construct was stably integrated into 6C2 cells and 26 individual clones were analyzed for transgene copy number and expression. The gene copy number and the luciferase expression (Materials and methods) for each clone are presented in the table. A scatter plot of luciferase activity versus gene copy number and the regression line are also shown.

differentiation. Because the establishment of a condensed region between the β-globin locus and the folate receptor gene could be necessary for the repression of the folate receptor gene, we investigated the structural properties of the region between these two loci.

The whole region between HS4 and HSA was cloned (Figure 1B). DNA sequencing revealed that 30% of the DNA contained in the 9.9 kbp *Hin*dIII clone is repeated sequence, principally the CR1 repeat commonly found in the chicken genome (Stumph *et al.*, 1981). We did not identify any known open reading frames. Previous studies have demonstrated that the 3′ end of this 16 kbp region possesses the characteristics of repressed chromatin; the 4 kbp located immediately upstream of the insulator (HS4) contain hypoacetylated histones and the chromatin in this region is more resistant to DNase I digestion than the chromatin contained in the β-globin locus in immature erythrocytes (Hebbes *et al.*, 1994). The absence of DHSs and the high concentration of repeated sequence in this 16 kbp region suggests that a repressed chromatin structure might extend further upstream, thus defining a 16 kbp condensed domain at this stage of differentiation.

Further structural analysis of this domain was carried out in 6C2 cells. The nuclease sensitivity of regions flanking HSA, inside and downstream of the folate receptor gene, and toward the center of the 16 kbp region were examined using micrococcal nuclease (MNase). Genomic DNA was purified after MNase digestion of 6C2 nuclei, blotted and hybridized successively with four different probes (probes MN1, MN2, MN3 and P1; Figure 7A). Inspection of the digestion patterns shows that the DNA surrounding the folate receptor gene (probes MN1, MN2 and MN3) is more sensitive than the DNA located towards the center of the 16 kbp region (probe P1) (Figure 7B). Sensitivity is greatest within the active folate receptor gene (probe MN2): at the last point of digestion, the amount of DNA detected as mononucleosomes is higher than the DNA, which migrates at ~1.3 kb. In contrast, for the same level of digestion with probe P1, DNA detected as mononucleosomes is as abundant as 1.3 kbp DNA. The

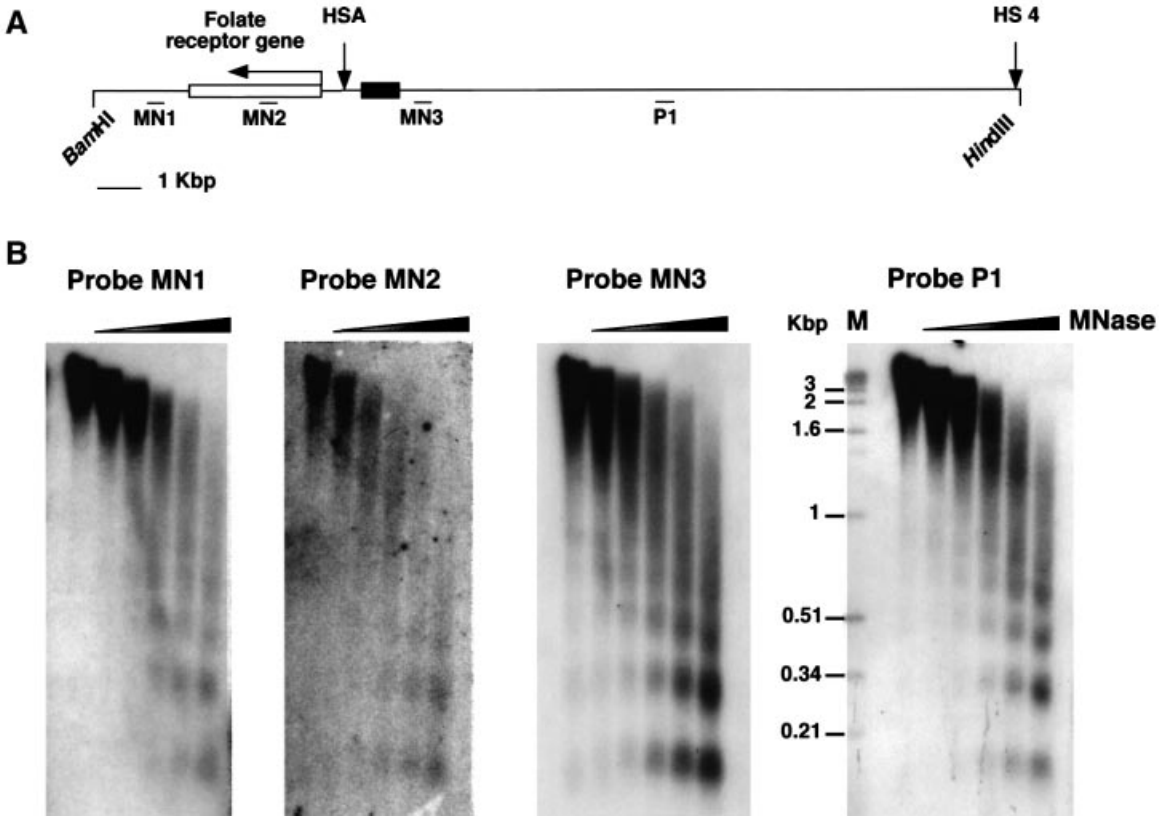


Fig. 7. The region between HS4 and HSA is condensed at the CFU-E stage. (**A**) The different probes (MN1, MN2, MN3 and P1) used for the analysis of micrococcal nuclease digestion are represented. The closed box between MN2 and MN3 represents a repeated sequence. (**B**) 6C2 nuclei were digested with increasing amounts of micrococcal nuclease. After purification, genomic DNA was analyzed by Southern blot hybridization. Two samples of the same digestion were run on the same gel and transferred together, then hybridized successively with probes MN1, P1, MN3 and MN2.

digestion profiles observed with probe MN1 also show that the DNA located downstream of the folate receptor gene is more resistant than that located inside and upstream of HSA (probes MN2 and MN3).

These results suggest that in cells arrested at the CFU-E stage, even in the absence of activation of the globin genes, the active locus containing the folate receptor gene is separated from the inactive β-globin domain by a region condensed into MNase-resistant chromatin. The observation of condensed chromatin between the folate receptor locus and the β-globin locus in these cells is another indication that the β-globin LCR is not implicated in the expression of the folate receptor gene.

We also found that HS3 is already formed in 6C2 cells, and that HS1 and HS2 are not detectable in contrast with erythrocytes, where HS1, HS2 and HS3 which define the LCR are present (Figure 8). The DHS over the β/ε enhancer is also present in 6C2 cells, but it is weaker than in erythrocytes (Boyes and Felsenfeld, 1996). These data reinforce the hypothesis that the β-globin LCR is not involved in the activation of the folate receptor gene in 6C2 cells.

Establishment of a fully unmethylated domain

Another important transition which is also found at the 3′ boundary of the condensed region (5′ boundary of the β-globin locus) is a progressive increase in DNA methylation as one proceeds into the condensed region in the 5′ direction (Chung *et al.*, 1997). We examined the methylation of sites at the 5′ end of this region, near the folate receptor. Methylation of the first *Hpa*II site, located 250 bp downstream of HSA (site 1, small arrow numbered 1 in Figure 9A), was analyzed in different tissue and cell lines by *Hpa*II–*Msp*I restriction analysis. In all the erythroid tissue and cell lines tested this site is cleaved by both enzymes, which shows that it is unmethylated (Figure 9B, lanes 2, 3, 5, 6, 8, 9, 11 and 12). In contrast to the disappearence of HSA, this unmethylated state is maintained in terminally differentiated erythrocytes (Figure 9B, Adult RBC). The site is unmethylated in all erythroid cells; in 11-day brain and in a lymphoid cell line (DT40) it is methylated (Figure 9B, lanes 14 and 23). Methylation of the next *Hpa*II site (site 2, small arrow numbered 2 in Figure 9A), located 1.2 kbp downstream of HSA, was also analyzed. This site is strongly methylated in HD24 (Figure 9C, lane 2), unmethylated in 6C2 (Figure 9C, lane 5) and methylated in 11-day and adult erythrocytes (Figure 9C, lanes 8 and 11). These results indicate that absence of methylation at this site is precisely correlated with folate receptor expression. In contrast, the *Hpa*II site 1 located proximally to HSA is unmethylated in all erythroid cells tested, but not in non-erythroid cell lines [even in DT40, where HSA is detectable (data not shown)]. Since the DNA is unmethylated at this site in erythrocytes (adult RBCs), where HSA does not exist, it is clear that there is no direct correlation between DNA methylation and DHS formation at this locus.

It is known that high levels of methyl-CpG correlate with transcriptional inactivity and nuclease resistance in chromosomes (Antequera *et al.*, 1989). Our previous

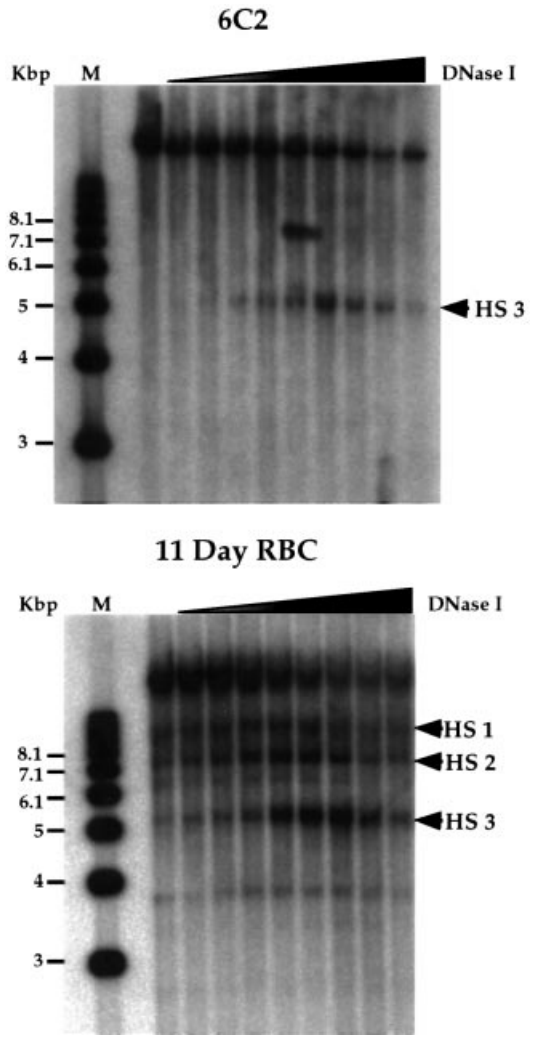


Fig. 8. Detection of DNase I hypersensitive sites. DNA from DNase I-treated nuclei (6C2 cells or 11 Day RBC) was digested with *Kpn*I and probed with probe P6. The hypersensitive bands corresponding to HS1, HS2 and HS3 are indicated.

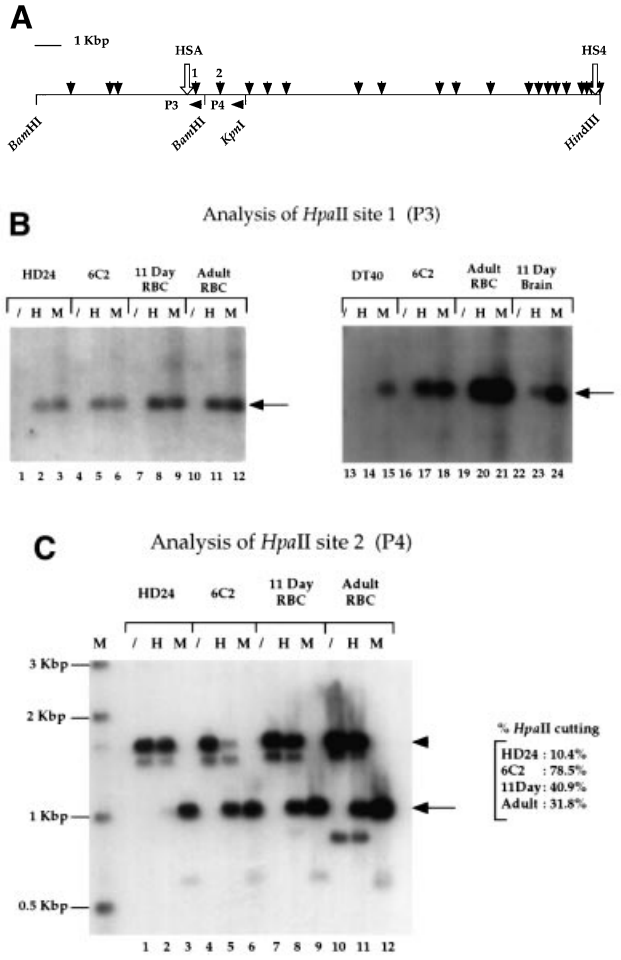


Fig. 9. Transition in DNA methylation close to HSA is tissue- and stage-specific. (**A**) The map represents restriction fragments and probes used to analyze the DNA methylation close to HSA. The parental *Bam*HI fragment detected with probe P3 is 6538 bp. The parental *Bam*HI–*Kpn*I fragment detected with P4 is 1574 bp. *Hpa*II sites are represented by small arrows. *Hpa*II sites analyzed in (B) and (C) are numbered 1 and 2, respectively. (**B**) The *Hpa*II site next to HSA (site1) is unmethylated only in erythroid cells. Genomic DNA from different cell lines or tissues was digested with *Bam*HI (/), *Bam*HI + *Hpa*II (H) or *Bam*HI + *Msp*I (M). Digested samples were detected with probe P3. Arrows mark the product of *Hpa*II digestion close to HSA. (**C**) The *Hpa*II site located further downstream of HSA (site 2) is unmethylated only in cells which express the folate receptor gene. Genomic DNA from different stages of erythroid differentiation was digested with *Bam*HI + *Kpn*I (/), *Bam*HI + *Kpn*I + *Hpa*II (H) or *Bam*HI + *Kpn*I + *Msp*I (M). Digested samples were detected with probe P4. Long arrow represents the product of *Hpa*II digestion, the small arrow shows the uncut parental fragment. The percentage of *Hpa*II cutting was quantified in each case by phosphoimager.

results indicated that at the CFU-E stage (6C2 cells), the domain between HS4 and HSA is condensed. DNA methylation around the folate receptor has a tissue- and stage-specific pattern, suggesting that there is precise control of methylation at this locus during erythropoiesis. As DNA methylation could be a mechanism for establishment or maintenance of condensed chromatin, DNA methylation was also analyzed along the domain between HS4 and HSA. *Hpa*II sensitivity was measured in both directions from the center of the 16 kbp region. Over this region, a distinct characteristic pattern of DNA methylation is displayed in 6C2 cells. The DNA is completely methylated over this entire region, yet becomes unmethylated close to HS4 and HSA (Figure 10, lanes 3 and 10). In contrast, in both immature erythrocytes (11-day RBCs) and mature, adult erythrocytes (Figure 10, lanes 5 and 12), only a fraction of the sites is methylated along the domain. Partial *Hpa*II digestion is observed at every site (Figure 10, lanes 4 and 11), but the boundaries are maintained; *Hpa*II sites close to HSA and HS4 are still fully unmethylated. As a control, the DNA methylation of this region was also analyzed in the embryonic chick brain. Two main features were observed. As in erythrocytes, only a fraction of *Hpa*II sites is methylated along the 16 kbp domain, and *Hpa*II sites close to the insulator are fully unmethylated (Figure 10, lane 14). In contrast to the erythroid cells, however, the DNA close to HSA is methylated, as shown by the presence of fragments derived from *Hpa*II cuts beyond HSA (Figure 10, lane 7). This result is in agreement with the results presented in Figure 9.

This study in 6C2 cells defines a methylated segment of DNA delineated by two unmethylated boundaries. In order to look at the accessibility of these boundaries in the nucleus, we tried to excise this piece of chromatin from 6C2 nuclei with a simple *Hpa*II restriction digest.

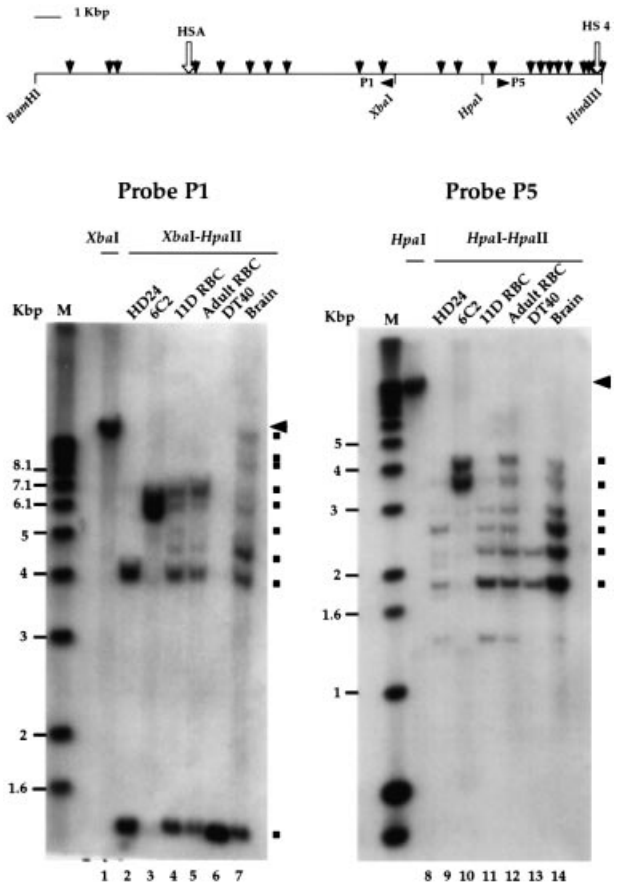


Fig. 10. Establishment of a fully methylated region between the folate receptor gene and the β-globin locus at the CFU-E stage. The map shows the position of *Xba*I and *Hpa*I restriction sites and probes P1 and P5 used to look for *Hpa*II cutting. Small arrows represent *Hpa*II sites. Analysis of *Hpa*II methylation toward the folate receptor gene is shown on the left panel. Genomic DNA from different tissues or cell lines was digested with *Xba*I + *Hpa*II and detected with probe P1. The small arrow indicates the parental band. Dots show the position of *Hpa*II sites. Analysis of *Hpa*II methylation toward the insulator is shown on the right panel. Genomic DNA from different tissues or cell lines was digested with *Hpa*I + *Hpa*II and detected with probe P5. The small arrow indicates the parental band.

The prediction was that the unmethylated *Hpa*II site close to HSA should be accessible, because at this stage this DHS is already formed but we did not know whether the chromatin close to the insulator was accessible in these cells. A discrete DNA fragment is indeed excised from 6C2 nuclei by *Hpa*II digestion (Figure 11, lane 1). The cutting of chromatin occurs at both ends very close to HS4 and HSA (Figure 11, lanes 2 and 5). This result suggests that a fully methylated domain is surrounded by an active locus at its 5′ boundary and that its 3′ boundary is also already established. We found that HS3 and HS4 are already formed; HS1 and HS2 are not detectable (Figure 8). The DHS over the β/ε enhancer is also present in 6C2 cells but it is weaker than in immature erythrocytes (Boyes and Felsenfeld, 1996). All these results suggest that the β-globin domain may already be pre-set at the CFU-E stage, reinforcing the idea that the 5′ β-globin insulator and the condensed region that we have identified might be crucial for the proper independent regulation of the folate receptor and the β-globin locus.

Discussion

Expression of the folate receptor at a proliferative stage

The purpose of this study was to define new chromatin regions upstream of the well characterized chicken β-globin domain and its 5′ insulator. We hoped that this analysis would provide some insight into the role of the boundary element in gene regulation. In the course of this study, we found a 16 kbp silent region upstream of the globin locus, and beyond that a novel gene that codes for a folate receptor and is expressed during erythroid differentiation. We first explored the pattern of expression of this gene in 1- to 3-day-old developing chick embryos, and found that it was confined to the blood islands (Figure 3); precursor cells at this stage do not express β-globin genes (Groudine and Weintraub, 1981). Folate receptor gene expression was not observed, however, in circulating red cells isolated from later stage embryos or adults, which do express β-globin genes. This pattern of expression is consistent with the observation that a dietary deficiency in folate induces megaloblastic anemia, characterized by a prevalence of large, early stage hematopoietic precursor cells (Koury *et al*., 1997). Moreover, it is known that folate receptors are attached to the membrane of human RBC and it has been shown that addition of antibodies to placental folate receptor increases the proliferation of BFU-E and CFU-E in culture (Antony *et al*., 1987, 1991). Our observations provide the first evidence that the expression of a folate receptor gene is regulated precisely during erythroid differentiation. This observation also raises the question of how two proximal loci, folate receptor and β-globin, activated at two different stages of erythroid differentiation, are regulated.

In order to study chromatin structure and DNA methylation patterns in cells that might express the folate receptor gene, we made use of two virally transformed avian erythroid precursor cell lines. These lines (6C2 and HD24) result specifically from transformation of BFU-E stage chicken pre-erythroid cells. Detailed analysis of surface markers has determined that 6C2 cells and other avian erythroblastosis virus transformed lines are arrested at the CFU-E stage of development (Beug *et al*., 1982), whereas the HD24 line is blocked at an earlier multipotent erythroid-myeloid precursor stage (Metz and Graf, 1991). We therefore analyzed methylation and chromatin structural patterns in these lines with the expectation that they would reflect the pattern of gene expression we had observed in embryonic erythroid development. For this reason, we concentrated our attention on 6C2 cells, and found that these CFU-E stage cells alone expressed the folate receptor. Consistent with these results, expression was also detected in the rapidly proliferating cells of the yolk sac blood islands, but not in differentiated erythrocytes. Like the 6C2 cells, yolk sac CFU-E cells do not express β-globin genes (Groudine and Weintraub, 1981). The HD24 line (arrested at MEP) expresses neither of these genes, and experiments are underway to determine whether embryonic blood precursor cells follow this same pattern of folate receptor and globin gene expression *in vivo*.

Establishment of a condensed region between two erythroid loci

It has been demonstrated previously that for at least 4 kbp upstream of the chicken β-globin locus, chromatin in 11-

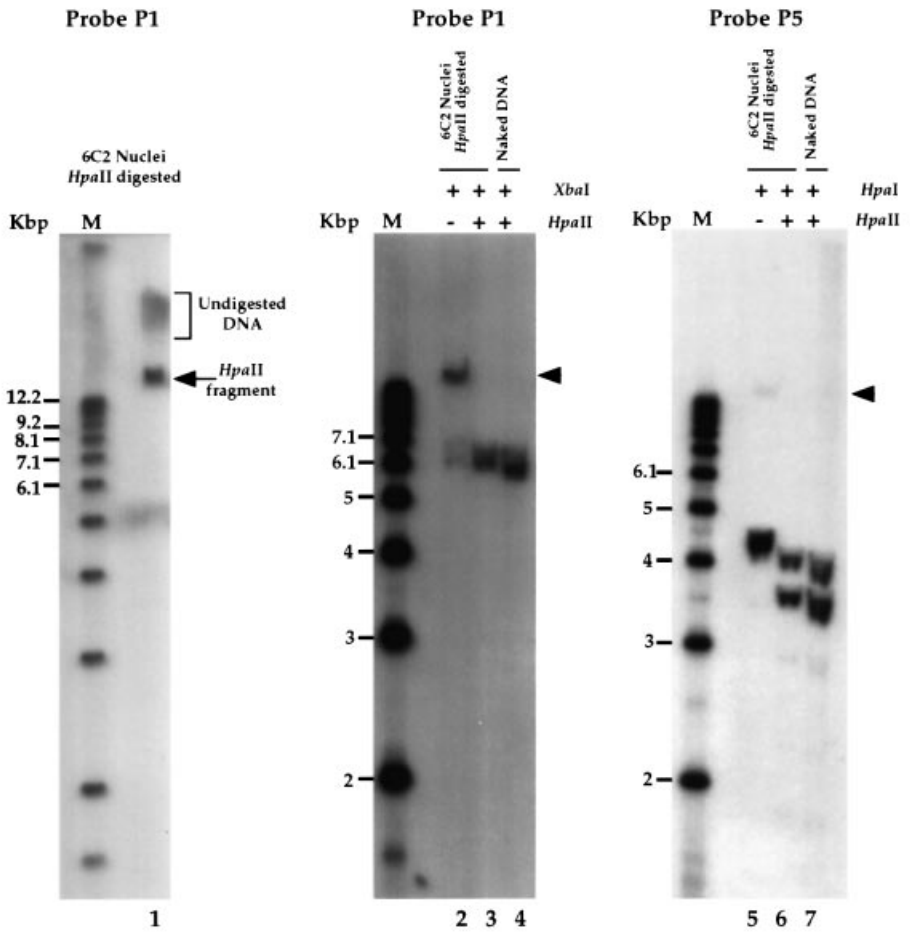


Fig. 11. The unmethylated boundaries of the domain are accessible to nuclease at the CFU-E stage. In the left panel, 6C2 nuclei were digested with *Hpa*II (see Materials and methods). Purified DNA was detected with probe P1 (Figure 10). *Hpa*II cutting on 6C2 nuclei was analyzed close to HSA in the center panel. DNA from *Hpa*II-treated nuclei was further digested with *Xba*I or *Xba*I + *Hpa*II and then detected with probe P1. The small arrow represents the parental band. The right panel shows that *Hpa*II sites close to HS4 are accessible in 6C2 nuclei. DNA from *Hpa*II-treated nuclei was digested further with *Hpa*I or *Hpa*I + *Hpa*II and then detected with probe P5 (Figure 10).

day erythrocytes is DNase I-resistant and packaged with hypoacetylated histones (Hebbes *et al*., 1994) (Figure 12). The establishment of a condensed region between the β-globin locus and the folate receptor could be necessary for the proper repression of the folate receptor gene and/or the absence of β-globin transcription at the CFU-E stage. We find that in 6C2 cells, the region between the folate receptor locus and the β-globin locus possesses properties of condensed chromatin (Figure 12). The DNA along this 16 kbp region is fully methylated, with boundaries that are unmethylated and accessible to nuclease digestion. In contrast, we find that this region is only partially methylated in erythrocytes, although it is known that at least 4 kbp of its 3′ part remains compacted (Figure 12). This suggests that when the folate receptor gene is active, the condensed DNA is maintained fully methylated, but that in differentiated cells in which it is inactive, DNA is randomly demethylated (note that adult RBCs are transcriptionally dead). It is not clear whether DNA methylation is a cause or effect of gene activation. It is possible, however, that DNA methylation along this domain is an important epigenetic mechanism for the maintenance of a condensed chromatin structure, especially given the known conection between DNA methylation and histone deacetylation in some systems (Jones *et al*., 1998; Nan *et al*., 1998). Once the condensed structure is established, perhaps the DNA can be randomly demethylated in quiescent cells without affecting the compaction of the DNA.

Recent results support this hypothesis. A mechanistic bridge between DNA methylation and histone deacetylation has been established in mammalian cells; it has been shown that at least one methyl-CpG-binding protein (the mouse protein MeCP2) associates with a corepressor complex that includes histone deacetylases (Jones *et al*., 1998; Nan *et al*., 1998). Such indirect recruitment of histone deacetylases by methyl-CpG could be an efficient mechanism for re-establishment of repressive chromatin structure after the passage of the replication fork. Our observations suggest that at the stage where the folate gene is expressed the entire region between the β-globin domain and the folate receptor may contain hypoacetylated histones. We have already shown that in 6C2 cells, the boundaries of this region are both accessible to nucleases and unmethylated. Whether these boundaries also mark a transition in histone acetylation, from hypoacetylated to hyperacetylated histones, is currently being investigated. The presence of a condensed chromatin structure close to

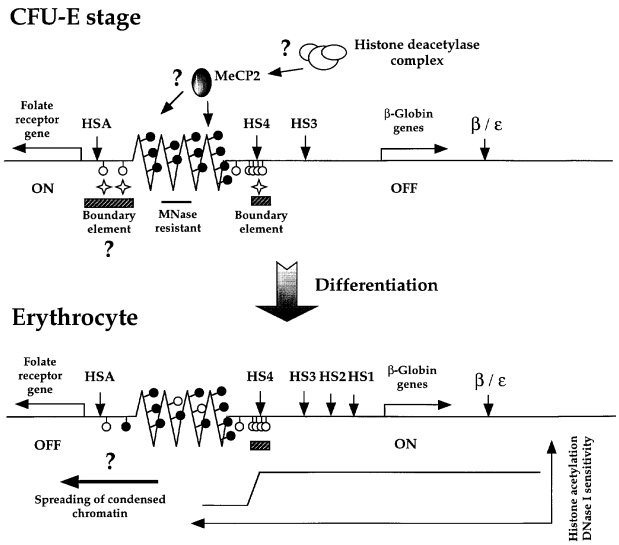


Fig. 12. Schematic representation of structural and functional modifications occurring during differentiation from CFU-E cells to erythrocytes on the β-globin and the folate receptor loci and the condensed region in between. At the CFU-E stage *Hpa*II sites are fully methylated in the region between the β-globin locus and the folate receptor locus, represented as closed circles. Unmethylated *Hpa*II sites are represented as open circles. Those unmethylated sites close to HSA and HS4 are accessible to *Hpa*II in 6C2 nuclei, and are indicated by stars. Question marks indicate hypotheses, which are formulated in the Discussion.

active loci raises the question of how compacted chromatin can be restricted to a precise region. We suggest that the insulator element previously identified at the 5′ end of the β-globin locus may participate in establishing such a structural boundary at the 3′ end of the condensed region. This element, contained in a 1.2 kbp fragment which includes HS4, has been shown to function both to block activation of an enhancer by a promoter when placed between them, and in *Drosophila* and 6C2 cells to confer position independence of expression on a weakly expressed reporter (Chung *et al*., 1993; Pikaart *et al*., 1998).

The 5′ boundary of the condensed region is not yet as well defined as the 3′ boundary. However, we have identified a region upstream of the folate receptor gene sufficient to induce strong expression of a transgene in 6C2 cells, which shows a fairly linear dependence on copy number. This observation suggests that the region may contain regulatory elements sufficient to prevent chromosomal position effects. Further characterization of this region will tell us whether such a function is provided by an LCR-like activity alone, or whether an insulating activity is also present. It is possible that repression of transcription of the folate receptor gene in erythrocytes involves the spreading of the condensed domain toward the folate receptor (Figure 12).

Boundaries and the role of the β-globin insulator

The organization of the folate receptor and β-globin gene cluster, and their programs of expression, suggests that there might be an important role for the β-globin insulator in establishing and maintaining separate regulatory compartments. Recent studies in our laboratory (Pikaart *et al*., 1998) have shown that a stably integrated reporter gene can be protected from both positive and negative position effects if it is flanked on both sides by two copies of the 1.2 kbp β-globin insulator element. This is consistent with a related role for the insulator element in its natural location at the 5′ end of the β-globin locus, in protecting the folate receptor and β-globin gene cluster from cross-interaction. An analogous organization of two loci has been described in human T cells. T cell receptor α and δ gene segments are organized within a single genetic locus but are differentially regulated during T cell development. An enhancer-blocking element has been identified between these segments within the locus and it has been proposed that this insulator functions as a boundary to separate the TCRα/δ locus into distinct regulatory domains (Zhong and Krangel, 1997).

Insulator elements also play a role in *Drosophila* development. Parasegment-specific expression of the genes that comprise the Bithorax complex (BX-C) is controlled by a *cis*-regulatory region that spreads over 300 kbp of DNA. This large region is subdivided into nine parasegment-specific *cis*-regulatory subregions (for a review see Duncan, 1987). Mutations in the *Mcp* and *Fab*-7 elements, which lie between particular pairs of parasegment *cis*-regulatory subregions, affect control of expression in BX-C, and consequently the identities of abdominal segments. This has been shown to arise from inactivation of an insulator function present in *Mcp* and *Fab*-7 (Gyurkovics *et al*., 1990; Galloni *et al*., 1993; Karch *et al*., 1994; Mihaly *et al*., 1997). *Mcp* and *Fab*-7 thus represent a distinct class of boundary elements; instead of separating adjacent domains that contain separate structural genes, they delimit *cis*-regulatory domains.

The folate receptor and the β-globin gene cluster are two loci that have different programs of expression during erythroid development. The folate receptor is expressed at an earlier proliferative stage than the globin genes, which are expressed during terminal differentiation. Throughout differentiation, both these loci display hypersensitive sites at their regulatory elements. In 6C2 cells arrested at the CFU-E stage, we found both HSA at the folate receptor, and HS4 and HS3 in the globin LCR; it is also known that the hypersensitive site at the β/ε enhancer is present (Boyes and Felsenfeld, 1996). Furthermore, HSA is still detectable in immature erythrocytes (Figure 12). These two loci are separated by 16 kbp of condensed chromatin and by the insulator sequence. The distance between the loci is within the range over which LCR elements are capable of activating transcription. We propose that the insulator may be important for the independent regulation of the two flanking loci, serving to prevent inappropriate activation of genes on one side of the 'barrier' by transactivating elements on the other. This may be an activity separate from, but related to the boundary-forming activity discussed above.

The proximity of two independently regulated erythroid-specific loci separated by an insulator and a relatively short segment of condensed chromatin will be valuable for investigating the interplay of chromatin structure and gene expression.

Materials and methods

Cloning and sequencing of the region upstream of the β-globin locus

Cloning of the upstream region was carried out by 'chromosome walking'. Four partial libraries were made successively. For the construction of

each library, Southern blot analysis was first used to determine an adequate restriction enzyme and the size of the genomic fragment after digestion. Genomic DNA was digested with appropriate restriction enzymes and then purified on an agarose gel. Fragments within ±500 bp of the size of the genomic fragment to be cloned were purified. After purification, partial libraries were generated by ligation in pBlueScript SK (Stratagene). For each library, 30 000 clones were screened. Isolation of the right clone was confirmed by DNA sequencing of the 3′ end of the clone and also by restriction digestion. Each clone was sent for DNA sequencing to Biotech Research laboratories (Rockville, MD). The sequence of the coding region was confirmed using the Thermo Sequenase system (Amersham).

DNase I, micrococcal and restriction endonuclease digestions in nuclei

For micrococcal nuclease digestion, 10^8 cells were washed twice in PBS, then resuspended in 1 ml of lysis buffer (10 mM Tris pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.4% NP-40). Nuclei were pelleted and resuspended in the same buffer plus 1 mM $CaCl_2$. Five microliters were mixed with 200 µl of NaCl 2 M/urea 5 M. The A_{260} was adjusted to 0.5. Six aliquots of 300 µl were digested with 0, 3, 6, 15, 30 and 60 U of MNase (Worthington) for 10 min at 37°C. The reaction was stopped with 60 µl of a 60 mM EDTA, 3% SDS. Samples were treated with 10 µg of RNase A at 37°C for 30 min, 100 µg of proteinase K was then added and samples were incubated at 55°C for 1 h. Samples were phenol–chloroform extracted three times, and chloroform extracted and then ethanol precipitated.

For *Hpa*II digestion, 10^8 cells were washed twice in PBS, then resuspended in 1 ml of lysis buffer 10 mM Tris pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.2% Triton X-100, 0.5 mM EGTA, 1 mM DTT and PMSF 0.4 mM, aprotinin 10 µg/ml. The sample was washed once again in the lysis buffer and then washed twice in the same buffer without Triton. Nuclei were pelleted and resuspend in 500 µl of a buffer (10 mM Tris pH 7.9, 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT), and 10 000 U of concentrated *Hpa*II (New England Biolabs) were added. The sample was incubated 1 h at 37°C. The reaction was stopped by addition of EDTA to a final concentration of 15 mM. DNA was extracted and precipitated as described above.

For DNase I digestion, ~10^8 cells were washed twice in cold PBS. Cells were lysed in 5 ml of buffer A (10 mM Tris–HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.2% Triton X-100, 1 mM DTT and 0.5 mM EGTA) plus PMSF 0.4 mM, aprotinin 10 µg/ml. Nuclei were pelleted for 5 min at 1000 *g* and washed once in buffer A without Triton. The pellet was resuspended in 1 ml of buffer B (10 mM Tris–HCl pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 1 mM $CaCl_2$) at room temperature, and 100 µl of the nuclei preparation were used for each DNase I (Worthington) digestion. The final DNase I concentrations were 0, 0.8, 1.6, 2.4, 3.2, 4, 4.8 and 5.6 U/ml. Digestions were performed at room temperature for 5 min and stopped by addition of 20 µl of buffer C (10 mM Tris pH 8, 60 mM EDTA and 3% SDS). Samples were treated with RNase A and proteinase K as described above. DNA was extracted and precipitated as described above.

Southern blotting and hybridization

Ten micrograms of each DNA sample was digested with the restriction endonucleases indicated in the particular experiment. Following electrophoresis through agarose gel, the gel was treated for 15 min in 0.25 M HCl and then incubated in denaturing buffer (3 M NaCl, 0.4 NaOH) twice, for 30 min each. After 15 min incubation in transfer buffer (3 M NaCl, 8 mM NaOH), the gel was transferred to a Genescreen plus membrane (DuPont) in 3 M NaCl, 8 mM NaOH for 90 min using a TurboBlotter system (Schleicher and Schuell).

Hybridizations were performed using 50 ng of probes labeled by random priming. Filters were prehybridized for 30 min and hybridized for 90 min at 68°C using Quick Hyb solution (Stratagene). Following hybridization, filters were washed as follows: once in 2× SSC, 0.1% SDS at room temperature, once in 2× SSC, 0.1% SDS at 68°C for 30 min and finally once in 0.1× SSC, 0.1% SDS at 68°C for 30 min.

In situ hybridization

Antisense riboprobes were prepared using digoxygenin-UTP (Boehringer Mannheim) and T7 RNA polymerase (Stratagene) from linearized RNase I plasmid. Sense riboprobes were prepared with the insert cloned in the opposite direction from the same vector (pCR2.1; Invitrogen). Fertilized eggs were incubated until the desired developmental stage. The whole blastoderm was harvested and fixed flat in 2 cm tissue culture plates in 4% paraformaldehyde overnight at 4°C. Embryos were washed in PBS then dehydrated gradually on ice in increasing concentrations of methanol to 100%, then rehydrated into PBS–0.1% Triton X-100 (PBT).

Prehybrization was for 2 h at 65°C in hybridization buffer (50% formamide, 5× SSC pH 6.0, 100 µg/ml yeast tRNA, 100 µg/ml salmon sperm DNA, 50 µg/ml heparin) followed by hybridization overnight at 65°C with riboprobe. Embryos were washed extensively first in 2× SSC at 65°C then at 2× SSC at room temperature in PBT, then blocked in 5% sheep's sera (Gibco-BRL) for 1 h, followed by a 2 h antibody incubation at RT (FITC-anti-Dig IgG; Boehringer Mannheim). PBT washes were monitored until background staining was negligible. Photographs were taken on an Olympus SZ12 microscope using Kodak 1600 ASA color film. After photography, embryos were counter stained with Giemsa (Sigma Accustain).

Construction

pGL3-FR4 was constructed by coinserting two *Xho*I–*Bam*HI fragments into pGL3-basic linearized by *Xho*I (Promega). The *Xho*I–*Bam*HI fragment linked to the luciferase reporter gene is located in the 3′ end of the 6.3 kbp *Bam*HI clone. The other *Xho*I–*Bam*HI fragment was excised from the 2.5 kbp *Bgl*II clone.

Cell lines

HD24 cells are chicken multipotent erythroid-myeloid cells transformed by the E26 virus. They cannot differentiate as efficiently as primary transformants and express some markers for early erythroid progenitors. They were grown in blastoderm media (Metz and Graf, 1991). 6C2 are CFU-E stage erythroid precursor cells, obtained by transformation of bone marrow with wild-type avian erythroblastosis virus (AEV). 6C2 cells were grown in alpha modified Eagle's medium (MEM) supplemented with 10% fetal bovine serum (FBS), 2% chicken serum, 1 mM HEPES, 50 mM β-mercaptoethanol and a standard complement of antibiotics. DT40 cells, purchased through American Type Culture Collection (ATCC, Rockville, MD), were grown in DMEM supplemented with 50 mM β-mercaptoethanol, 2 mM glutamine, 10% FBS, 5% chicken serum, 10% tryptose phosphate broth and antibiotics. All the cells were maintained at 37°C in 5% CO_2.

Cell culture and transfections

6C2 cells were grown, transfected and selected as described previously (Boyes and Felsenfeld, 1996). The test construct pGL3-FR4 was linearized by either *Nhe*I or *Sal*I. *Nhe*I cuts the vector just upstream of the 5′ end of the folate receptor upstream region; *Sal*I cuts 3′ of the SV40 late poly(A) signal of the luciferase reporter gene. Both linearizations gave the same results. Individual colonies were picked after 2–3 weeks and expanded in medium containing 1000 U/ml hygromycin (Calbiochem), an amount that kills any non-transfected cells. After cell growth for 4–6 days in 3 ml of fresh selective medium, half of the culture was spun down and resuspended in 500 µl fetal calf serum plus 10% DMSO and stored in liquid nitrogen. The rest of the culture was spun down and the pellet was treated for DNA extraction.

Determination of the copy number and transcriptional activity of 6C2-transfected colonies

The cell pellet of each individual colony was resuspended in 250 µl of 400 mM NaCl, 200 mM Tris pH 8.5 and lysed by adding 250 µl of 0.4% SDS, 10 mM EDTA. The sample was incubated for 30 min at 37°C with 100 µg/ml RNase A and then with 100 µg/ml Proteinase K for 1 h at 55°C. Samples were extracted twice with phenol–chloroform and once with chloroform, and the DNA was precipitated with isopropanol. The DNA pellet was resuspended in 50 µl 10 mM Tris–HCl pH 7.5, 1 mM EDTA (TE). Ten microliters of the DNA sample was digested by *Hpa*I and *Kpn*I. Digested samples were analyzed by Southern blotting with probe RNase I. This probe detects a 4.9 kbp fragment for the transgene and a 7.5 kbp fragment for the endogenous folate receptor gene after *Hpa*I and *Kpn*I digestion. The ratio between the transgene copy number and the endogenous gene was quantified by Phosphorimage analysis. A ratio of 0.5 was identified as a single copy integration. This was confirmed by a restriction digestion which cuts once in the construct integrated. Clones stored in liquid nitrogen were cultured for 4–6 days in fresh medium containing hygromycin. For luciferase assays, cells from 1 ml of culture were harvested, washed twice in PBS and resuspended in 100 µl of Reporter Lysis buffer (Promega). The supernatants were assayed for luciferase activity using the Promega luciferase assay system. Luciferase activity was normalized to the protein concentration in each samples. Protein concentration was determined with Bio-Rad protein assay kit.

RNA preparations and analysis

Total RNA was prepared using RNA STAT-60 solution (TEL-TEST) according to the manufacturer's instructions and was treated with RNase-free DNase I (Boehringer Mannheim). RNA integrity of each preparation was checked on ethidium bromide stained 1% agarose/MOPS-formaldehyde gels. RNA was prepared from HD24, 6C2 and DT40 chicken cell lines, circulating red blood cells of 11-day embryos, chicken fibroblasts prepared by trypsin digestion of 11-day embryo, brain (cerebral hemisphere) of 11-day embryo. Fertilized White Leghorn chicken eggs were obtained from Truslow Farms, Chestertown, MD. The amount of RNA extracted per nucleus from 6C2 cells, HD24 and 11-day RBCs was determined. Before extraction, cells were counted with a hemacytometer. RNA (4.7 pg) per nucleus, 7.2 pg of RNA per nucleus, and 0.6 pg per nucleus were extracted respectively from HD24 cells, 6C2 cells and 11-day RBCs. For the RNase protection assay, 19.6, 30 and 2.5 μg of RNA from HD24, 6C2 and 11-day RBCs, respectively, were hybridized.

RNA was analyzed by RNase protection assay using the RPA II kit (Ambion). Radiolabeled antisense transcripts were generated using T7 RNA polymerase (Stratagene) and [^{32}P]CTP. RNA probes were acrylamide gel-purified prior to hybridization. Hybridizations were performed overnight at 45°C. RNA was also analyzed by Northern blot. Thirty micrograms were loaded in each lane. The blot was hybridized with a 1.2 kbp *Xho*I–*Bgl*II fragment excised from 6.3 kbp *Bam*HI clone.

Probes

P1: 404 bp, was generated by PCR amplification from the 9.9 kbp *Hin*dIII clone with oligonucleotides GATGGCTTCCCCAGCAGCAA and AATGGTCTTTGCCTGGGTCA.

P2: 566 bp, *Kpn*I–*Xho*I restriction fragment excised from the 2.5 kbp *Bgl*II clone.

P3: 263 bp, *Bgl*II–*Bam*HI restriction fragment excised from the 2.5 kbp *Bgl*II clone.

P4: 689 bp, *Eco*RV–*Kpn*I restriction fragment excised from the 2.5 kbp *Bgl*II clone.

P5: 394 bp, *Nco*I–*Bam*HI restriction fragment excised from pCBG D (Hebbes *et al*., 1994).

P6: 1062 bp, *Sac*II–*Kpn*I restriction fragment excised from pCBG C (Reitman and Felsenfeld, 1990).

MN1: 378 bp, was generated by PCR amplification from the 6.3 kbp *Bam*HI clone with oligos CAGCTTCTGGAAATAACTGA and CTCTCCTGCACAGCTGTGGT.

MN2: 394 bp, was generated by PCR amplification from the 6.3 kbp *Bam*HI clone with oligos CATGTCAGCATCCCACAAAG and CAGTGGTGTGGGTAATTGGA.

MN3: 377 bp, was generated by PCR amplification from the 2.5 kbp *Bgl*II clone with oligos CTGCTGATCACTGAACTGTG and TATGCTGGGAACAGCTATAC.

RNase I: 267 bp, was generated by PCR amplification from the 6.3 kbp *Bam*HI clone with oligos ACCGTGCCTTCTCCTCCTCT and CTGTGTTGCTTTAATGCTGTCA. The PCR fragment was then cloned in pCR2.1 (Invitrogen) using the TA Cloning kit. Orientation of the insert was determined by restriction digestion.

RNase II: 295 bp, was generated by PCR amplification from the 6.3 kbp *Bam*HI clone with oligos GCCTCCTGACACAGCAGAGC and GCAGGTTTGAGGCAAGTTAG. The PCR fragment was then cloned in pCR2.1 (Invitrogen) using the TA Cloning kit. Orientation of the insert was determined by restriction digestion.

Acknowledgements

We are grateful to Olivier Hyrien, Mike Litt, Felix Recillas-Targa and Adam West for critical reading of this manuscript, and members of the Felsenfeld laboratory for helpful discussions. M.-N.P. was supported by a long-term EMBO fellowship and a postdoctoral fellowship from the Human Frontiers Science Project Organization.

References

Antequera,F., Macleod,D. and Bird,A.P. (1989) Specific protection of methylated CpGs in mammalian nuclei. *Cell*, **58**, 509–517.

Antony,A.C., Bruno,E., Briddell,R.A., Brandt,J.E., Verma,R.S. and Hoffman,R. (1987) Effect of perturbation of specific folate receptors during *in vitro* erythropoiesis. *J. Clin. Invest.*, **80**, 1618–1623.

Antony,A.C., Briddell,R.A., Brandt,J.E., Straneva,J.E., Verma,R.S., Miller,M.E., Kalasinski,L.A. and Hoffman,R. (1991) Megaloblastic hematopoiesis *in vitro*. Interaction of anti-folate receptor antibodies with hematopoietic progenitor cells leads to a proliferative response independent of megaloblastic changes. *J. Clin. Invest.*, **87**, 313–325.

Beug,H., Doederlein,G., Freudenstein,C. and Graf,T. (1982) Erythroblast cell lines transformed by a temperature-sensitive mutant of avian erythroblastosis virus: a model system to study erythroid differentiation *in vitro*. *J. Cell. Physiol. Suppl.*, **1**, 195–207.

Bonifer,C., Yannoutsos,N., Kruger,G., Grosveld,F. and Sippel,A.E. (1994) Dissection of the locus control function located on the chicken lysozyme gene domain in transgenic mice. *Nucleic Acids Res.*, **22**, 4202–4210.

Boyes,J. and Felsenfeld,G. (1996) Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. *EMBO J.*, **15**, 2496–2507.

Campbell,I.G., Jones,T.A., Foulkes,W.D. and Trowsdale,J. (1991) Folate-binding protein is a marker for ovarian cancer. *Cancer Res.*, **51**, 5329–5338.

Chung,J.H., Whiteley,M. and Felsenfeld,G. (1993) A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell*, **74**, 505–514.

Chung,J.H., Bell,A.C. and Felsenfeld,G. (1997) Characterization of the chicken beta-globin insulator. *Proc. Natl Acad. Sci. USA*, **94**, 575–580.

Coney,L.R., Tomassetti,A., Carayannopoulos,L., Frasca,V., Kamen,B.A., Colnaghi,M.I. and Zurawski,V.R.,Jr (1991) Cloning of a tumor-associated antigen: MOv18 and MOv19 antibodies recognize a folate-binding protein. *Cancer Res.*, **51**, 6125–6132.

Dillon,N., Trimborn,T., Strouboulis,J., Fraser,P. and Grosveld,F. (1997) The effect of distance on long-range chromatin interactions. *Mol. Cell*, **1**, 131–139.

Duncan,I. (1987) The bithorax complex. *Annu. Rev. Genet.*, **21**, 285–319.

Epner,E. *et al.* (1998) The β-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse β-globin locus. *Mol. Cell*, **2**, 447–455.

Fantes,J. *et al.* (1995) Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum. Mol. Genet.*, **4**, 415–422.

Festenstein,R., Tolaini,M., Corbella,P., Mamalaki,C., Parrington,J., Fox,M., Miliou,A., Jones,M. and Kioussis,D. (1996) Locus control region function and heterochromatin-induced position effect variegation. *Science*, **271**, 1123–1125.

Forrester,W.C., Takegawa,S., Papayannopoulou,T., Stamatoyannopoulos,G. and Groudine,M. (1987) Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res.*, **15**, 10159–10177.

Forrester,W.C., Novak,U., Gelinas,R. and Groudine,M. (1989) Molecular analysis of the human β-globin locus activation region. *Proc. Natl Acad. Sci. USA*, **86**, 5439–5443.

Foster,J.W. *et al.* (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature*, **372**, 525–530.

Galloni,M., Gyurkovics,H., Schedl,P. and Karch,F. (1993) The bluetail transposon: evidence for independent *cis*-regulatory domains and domain boundaries in the bithorax complex. *EMBO J.*, **12**, 1087–1097.

Geyer,P.K. (1997) The role of insulator elements in defining domains of gene expression. *Curr. Opin. Genet. Dev.*, **7**, 242–248.

Greaves,D.R., Wilson,F.D., Lang,G. and Kioussis,D. (1989) Human CD2 3′-flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. *Cell*, **56**, 979–986.

Grosveld,F., van Assendelft,G.B., Greaves,D.R. and Kollias,G. (1987) Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell*, **51**, 975–985.

Groudine,M. and Weintraub,H. (1981) Activation of the globin genes during chicken development. *Cell*, **24**, 393–401.

Gyurkovics,H., Gausz,J., Kummer,J. and Karch,F. (1990) A new homeotic mutation in the *Drosophila* bithorax complex removes a boundary separating two domains of regulation. *EMBO J.*, **9**, 2579–2585.

Hebbes,T.R., Clayton,A.L., Thorne,A.W. and Crane-Robinson,C. (1994) Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken β-globin chromosomal domain. *EMBO J.*, **13**, 1823–1830.

Hesse,J.E., Nickol,J.M., Lieber,M.R. and Felsenfeld,G. (1986) Regulated gene expression in transfected primary chicken erythrocytes. *Proc. Natl Acad. Sci. USA*, **83**, 4312–4326.

Hug,B.A., Moon,A.M. and Ley,T.J. (1992) Structure and function of the murine β-globin locus control region 5′ HS-3. *Nucleic Acids Res.*, **20**, 5771–5778.

Jantzen,K., Fritton,H.P. and Igo-Kemenes,T. (1986) The DNase I sensitive domain of the chicken lysozyme gene spans 24 kb. *Nucleic Acids Res.*, **14**, 6085–6099.

Jones,P.L., Veenstra,G.J., Wade,P.A., Vermaak,D., Kass,S.U., Landsberger,N., Strouboulis,J. and Wolffe,A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genet.*, **19**, 187–191.

Joos,S., Falk,M.H., Lichter,P., Haluska,F.G., Henglein,B., Lenoir,G.M. and Bornkamm,G.W. (1992) Variable breakpoints in Burkitt lymphoma cells with chromosomal t(8;14) translocation separate c-myc and the IgH locus up to several hundred kb. *Hum. Mol. Genet.*, **1**, 625–632.

Karch,F., Galloni,M., Sipos,L., Gausz,J., Gyurkovics,H. and Schedl,P. (1994) Mcp and Fab-7: molecular analysis of putative boundaries of *cis*-regulatory domains in the bithorax complex of *Drosophila melanogaster*. *Nucleic Acids Res.*, **22**, 3138–3146.

Koury,M.J., Horne,D.W., Brown,Z.A., Pietenpol,J.A., Blount,B.C., Ames,B.N., Hard,R. and Koury,S.T. (1997) Apoptosis of late-stage erythroblasts in megaloblastic anemia: association with DNA damage and macrocyte production. *Blood*, **89**, 4617–4623.

Lawson,G.M., Knoll,B.J., March,C.J., Woo,S.L., Tsai,M.J. and O'Malley,B.W. (1982) Definition of 5′ and 3′ structural boundaries of the chromatin domain containing the ovalbumin multigene family. *J. Biol. Chem.*, **257**, 1501–1507.

Levy-Wilson,B. and Fortier,C. (1989) The limits of the DNase I-sensitive domain of the human apolipoprotein B gene coincide with the locations of chromosomal anchorage loops and define the 5′ and 3′ boundaries of the gene. *J. Biol. Chem.*, **264**, 21196–21204.

Lucas,A.E. and Jamroz,C. (1961) *Atlas of Avian Hematology.* United States Department of Agriculture.

Mason,M.M., Lee,E., Westphal,H. and Reitman,M. (1995) Expression of the chicken β-globin gene cluster in mice: correct developmental expression and distributed control. *Mol. Cell. Biol.*, **15**, 407–414.

May,G. and Enver,T. (1995) Targeting gene expression to haemopoietic stem cells: a chromatin-dependent upstream element mediates cell type-specific expression of the stem cell antigen CD34. *EMBO J.*, **14**, 564–574.

Metz,T. and Graf,T. (1991) v-myb and v-ets transform chicken erythroid cells and cooperate both *in trans* and *in cis* to induce distinct differentiation phenotypes. *Genes Dev.*, **5**, 369–380.

Mihaly,J., Hogga,I., Gausz,J., Gyurkovics,H. and Karch,F. (1997) *In situ* dissection of the Fab-7 region of the bithorax complex into a chromatin domain boundary and a Polycomb-response element. *Development*, **124**, 1809–1820.

Moon,A.M. and Ley,T.J. (1990) Conservation of the primary structure, organization and function of the human and mouse β-globin locus-activating regions. *Proc. Natl Acad. Sci. USA*, **87**, 7693–7697.

Nan,X., Ng,H.H., Johnson,C.A., Laherty,C.D., Turner,B.M., Eisenman, R.N. and Bird,A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, **393**, 386–389.

Pikaart,M.J., Recillas-Targa,F. and Felsenfeld,G. (1998) Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.*, **12**, 2852–2861.

Reik,A., Telling,A., Zitnik,G., Cimbora,D., Epner,E. and Groudine,M. (1998) The locus control region is necessary for gene expression in the human β-globin locus but not the maintenance of an open chromatin structure in erythroid cells. *Mol. Cell. Biol.*, **18**, 5992–6000.

Reitman,M. and Felsenfeld,G. (1990) Developmental regulation of topoisomerase II sites in the chicken β-globin locus. *Mol. Cell. Biol.*, **10**, 2774–2786.

Ross,J.F., Chaudhuri,P.K. and Ratnam,M. (1994) Differential regulation of folate receptor isoforms in normal and malignant tissues *in vivo* and in established cell lines. Physiologic and clinical implications. *Cancer*, **73**, 2432–2443.

Sabin,F.R. (1920) Studies on the origin of blood-vessels and of red blood-corpuscles as seen in the living blstoderm of chicks during the second day of incubation. *Carnegie Contrib. Embryol.*, **9**, 213–262.

Samarut,J., Jurdic,P. and Nigon,V. (1979) Production of erythropoietic colony-forming units and erythrocytes during chick embryo development: an attempt at modelization of chick embryo erythropoiesis. *J. Embryol. Exp. Morph.*, **50**, 1–20.

Shen,F., Ross,J.F., Wang,X. and Ratnam,M. (1994) Identification of a novel folate receptor, a truncated receptor and receptor type β in hematopoietic cells: cDNA cloning, expression, immunoreactivity and tissue specificity. *Biochemistry*, **33**, 1209–1215.

Stalder,J., Larsen,A., Engel,J.D., Dolan,M., Groudine,M. and Weintraub,H. (1980) Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNAase I. *Cell*, **20**, 451–460.

Stratling,W.H., Dolle,A. and Sippel,A.E. (1986) Chromatin structure of the chicken lysozyme gene domain as determined by chromatin fractionation and micrococcal nuclease digestion. *Biochemistry*, **25**, 495–502.

Stumph,W.E., Kristo,P., Tsai,M.J. and O'Malley,B.W. (1981) A chicken middle-repetitive DNA sequence which shares homology with mammalian ubiquitous repeats. *Nucleic Acids Res.*, **9**, 5383–5397.

Tommerup,N. *et al*. (1993) Assignment of an autosomal sex reversal locus (SRA1) and campomelic dysplasia (CMPD1) to 17q24.3-q25.1. *Nature Genet.*, **4**, 170–174.

Ton,C.C. *et al*. (1991) Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell*, **67**, 1059–1074.

Wagner,T. *et al*. (1994) Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell*, **79**, 1111–1120.

Weitman,S.D., Lark,R.H., Coney,L.R., Fort,D.W., Frasca,V., Zurawski, V.R.,Jr and Kamen,B.A. (1992) Distribution of the folate receptor GP38 in normal and malignant cell lines and tissues. *Cancer Res.*, **52**, 3396–3401.

Yan,W. and Ratnam,M. (1995) Preferred sites of glycosylphosphatidyl-inositol modification in folate receptors and constraints in the primary structure of the hydrophobic portion of the signal. *Biochemistry*, **34**, 14594–14600.

Zhong,X.P. and Krangel,M.S. (1997) An enhancer-blocking element between α and δ gene segments within the human T cell receptor α/δ locus. *Proc. Natl Acad. Sci. USA*, **94**, 5219–5224.

Received October 9, 1998; revised April 8, 1999; accepted June 1, 1999