

CONTROL OF GLOBIN GENE TRANSCRIPTION¹

Todd Evans, Gary Felsenfeld, and Marc Reitman

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892

KEY WORDS: transcription factors, dominant control regions, locus activating regions, developmental gene regulation

CONTENTS

INTRODUCTION	95
CIS ELEMENTS AND GLOBIN GENE REGULATION	96
<i>Globin Cluster Organization</i>	96
<i>Local Control Elements</i>	97
REGULATION OF GLOBIN EXPRESSION BY TRANS-ACTING CELLULAR FACTORS	102
<i>Cellular Factors Influence Globin Expression</i>	102
<i>DNA-Binding Proteins</i>	104
CHROMATIN STRUCTURE OF THE GLOBIN GENES	110
<i>Matrix Attachment Regions (MARs)</i>	110
<i>Chromatin Structure and Globin Gene Expression</i>	111

INTRODUCTION

The globin gene families have long served as a paradigm for the study of developmental regulation of gene expression. Some of the earliest work on gene sequence organization, control of transcription, and chromatin structure in relationship to gene activation was carried out in this system (see reviews by Collins & Weissman 1984; Bunn & Forget 1986; Nienhuis & Maniatis 1987; Stamatoyannopoulos & Nienhuis 1987; Weatherall et al 1989).

There is now new information about globin promoters and enhancers, the local *cis*-acting DNA regulatory elements. Much progress has also

¹ The US Government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

been made in the identification and characterization of the specific DNA-binding factors that interact with these two classes of regulatory elements. Most recently distant sequences have been identified that may exert their regulatory effect by structural changes in large chromatin domains. It is not possible to provide a comprehensive account of these advances in a review of this size. We have, therefore, summarized some essential findings from past investigations and emphasized results of recent studies that we believe illustrate the important conclusions. There is every reason to assume that the principles governing control of the globin genes are applicable to other gene families as well.

CIS ELEMENTS AND GLOBIN GENE REGULATION

Globin Cluster Organization

We begin by considering the chromosomal organization of the α and β globin loci (see Bunn & Forget 1986). Within every globin cluster studied, the direction of transcription in each transcription unit is the same. It is not known whether this is important for regulation or is simply a remnant of locus evolution by gene duplication. In most cases the genes at the 5' end of the cluster are expressed earliest in development; those at the 3' end are expressed latest. For example, the human β locus (Figure 1) is arranged in the order embryonic (ϵ), fetal (γ^G , γ^A), adult (δ , β) and the human α locus in the order embryonic (ζ), fetal/adult ($\alpha 2$, $\alpha 1$). Exceptions to this order include the chicken β locus (Figure 1), where the early embryonic genes (ρ , ϵ) flank the late embryonic (β^H) and late embryonic/adult (β^A) genes, and the goat β locus, in which the cluster is a tandem triplication. The significance of gene order for expression is unknown. An unusual transcription unit, oriented in the opposite direction from the globin genes, is found between β^H and β^A genes in the chicken β cluster. This unit may ensure termination of β^H transcripts so that no interference with initiation at the β^A promoter occurs (Pribyl & Martinson 1988).

The embryonic/adult regulatory mechanisms of the β cluster may differ between chickens and humans. At the time of divergence of Aves and Mammalia, only a single β gene existed. After divergence independent duplications occurred (Goodman et al 1987). Therefore there is no reason to expect that the β cluster gene choice regulatory mechanisms in the chicken should be related to those in mouse or man. In contrast, the embryonic/adult duplication in the α locus occurred before the Aves/Mammalia divergence (Goodman et al 1987), and the mechanisms governing α embryonic/adult gene choice are expected to be conserved.

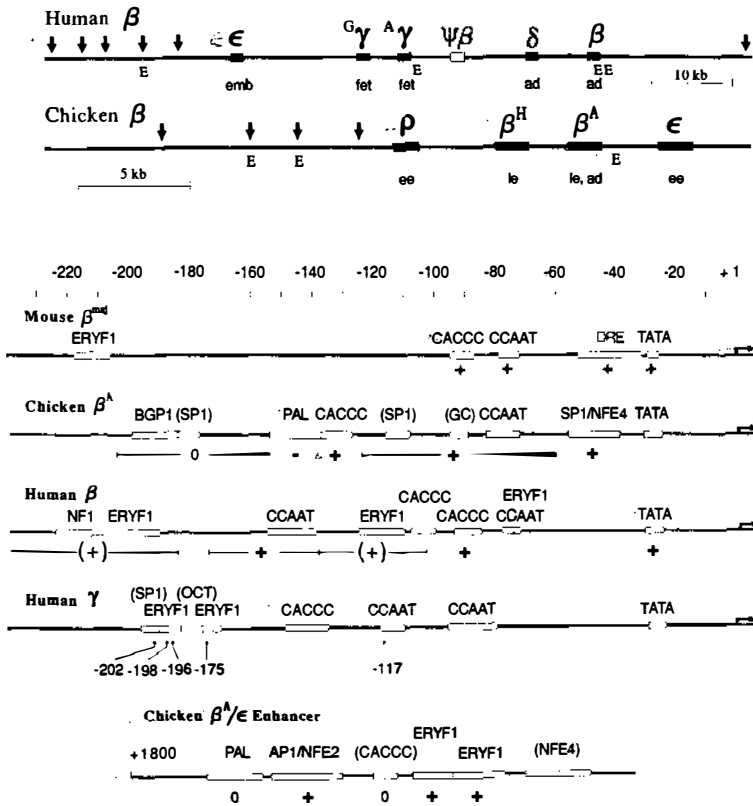


Figure 1 Organization of some globin clusters, promoters, and an enhancer. The cluster maps depict the position of the genes (*solid boxes*), distant super-hypersensitive sites (*arrows*), enhancers (E), and stage of expression (emb, embryo; fet, fetus; ad, adult; ee, early embryo; and le, late embryo). The promoter maps show the sequence motifs/factor binding sites and their functional contribution (–, 0, or +). Parentheses indicate some uncertainty in functional significance of the site. Indicated under the consensus human γ promoter are the positions of point mutations associated with HPFH. The scale for the promoters, in bp from the cap site, is shown. The enhancer is drawn to the same scale, with base +1800 relative to the β^A cap site indicated. For further details see the text and Table 1.

Local Control Elements

From mutation studies a great deal is known about the anatomy of promoters (regulatory regions near the mRNA-cap site) and enhancers (more distant regulatory regions) in the globin gene families. Those regions of DNA whose regulatory function has been assayed will be discussed herein. Many other potential regulatory elements have been recognized by

sequence similarity to known functional sites, or by protein binding studies, but they will be ignored in this section unless there is direct evidence for their function. We do not discuss the DNA sequence elements involved in RNA metabolism (RNA splicing, polyadenylation, transcription termination, stability). Note, however, that these processes are important and mutations that affect them result in production of low levels of globin protein (reviewed in Weatherall et al 1989).

Most of the evidence for *cis*-element function in transcription comes from studying the expression of cloned globin genes transfected into cultured cells. These experiments vary widely (for example, in the cell-type, transfection assay, and properties of the DNA vectors), making it difficult to derive general conclusions from them. Nevertheless, it is possible to conclude that globin promoters alone do not confer absolute erythroid specificity. Thus globin genes can be expressed in non-erythroid cells, although a viral enhancer is often required.

Some expression experiments give results that appear to mimic the *in vivo* situation. For example, in transient COS cell expression assays, the human δ gene is expressed at $\sim 2\%$ of the level of the human β gene. In this case, the $\delta : \beta$ ratio during transient expression appears to be a measure of the relative effectiveness of the two regulatory regions *in vivo* (Humphries et al 1982; note that the β construct includes a weak enhancer, whereas the δ construct is not known to contain any). Similar expression experiments using hybrid genes have allowed enhancers to be localized. After accounting for the enhancers' action, the data suggest that the remaining regulation usually occurs within ~ 200 bp of the cap site (i.e. at the promoter). In the absence of an enhancer, some globin promoters are expressed in a partially tissue- and stage-specific manner (e.g. Rutherford & Nienhuis 1987), which suggests that isolated promoters contain some, but not all, of the tissue and stage regulatory functions.

Globin gene transcription can initiate, at low frequency, upstream of the predominant mRNA cap site (Allan et al 1983; Ley & Nienhuis 1983; Lois et al 1990; Broders et al 1990 and references therein). Globin transcription has also been reported to occur in non-erythroid cells (e.g. Humphries et al 1976; Groudine & Weintraub 1980), often with overrepresentation of the upstream initiation sites (e.g. Allan et al 1983; Lois et al 1990). These observations are intriguing, but it is not clear that they have functional significance.

PROMOTERS A typical β globin promoter has three positive acting sequence elements: a TATA box at about position -30 (relative to the mRNA cap site), a CCAAT box near -70 to -90 , and a CACCC motif

at variable positions, often near -95 to -120 . The CACCC sequence was originally identified as a common element of adult β -globin promoters (Dierks et al 1983). None of these motifs is restricted to globin genes, and proteins that bind these motifs have been identified in both erythroid and non-erythroid cells. Some specific promoters, each of which contains additional regulatory elements, are analyzed below. Supplementary data are presented in Figure 1 and Table 1.

SPECIFIC β -GLOBIN PROMOTERS The mouse β^{maj} globin promoter has been examined in exquisite detail by Myers et al (1986), who produced 130 point mutants between $+24$ and -101 . The CACCC, CCAAT, and TATA regions are all needed for maximal activity in both non-erythroid and erythroid cells (Cowie & Myers 1988). In contrast, the -32 to -53 region containing a direct repeat (DRE) is important only in erythroid cells induced to differentiate (Stuve & Myes 1990).

For the chicken β^A -globin gene, a combination of in vitro transcription (Emerson et al 1989) and transient expression (Jackson et al 1989; Gallarda et al 1989) studies have defined four functional regions in addition to the TATA box. Protein binding studies have identified the *trans*-acting factors associated with these sites (discussed below).

In the human β promoter, the importance in vivo of the CACCC and TATA elements is demonstrated by the existence of patients with point mutations in these motifs who show reduced β -globin expression (reviewed in Weatherall et al 1989). Eight protein binding regions in the β -globin promoter are known; functional studies demonstrated that the -150 region is needed for maximal activity and suggest that either the -200 or the -120 region is also required (Antoniou et al 1988; deBoer et al 1988). Berg et al (1989) have detected silencer elements (negatively-acting enhancers) upstream of the promoter, which can function in either K562 or HeLa cells. Cao et al (1989) have also found a silencer upstream of the human ϵ promoter.

CHICKEN β CLUSTER ENHANCERS Three enhancers have been found in the chicken β cluster (Figure 1). The most thoroughly characterized is equidistant from the β^A and ϵ promoters (Hesse et al 1986; Choi & Engel 1986); the other two are found upstream of the ρ gene (~ 6.4 and ~ 4.2 kb, respectively; M. Reitman, unpublished results). The β^A/ϵ enhancer is conserved between chickens and ducks in both sequence and enhancer activity (Kretsovali et al 1987). This enhancer is erythroid-specific and mediates a 50- to 100-fold increase in transcription (Hesse et al 1986). The smallest DNA fragment with full enhancer activity contains four discrete protein-binding regions, two of which confer the positive activity (Emerson

Table 1 Sites near globin genes that have been assayed for function

Region	Position ^a	Motif/ factor	Function ^b	Assay ^c	Mutation method ^d
Chick β^A promoter	-39 to -58	SP1/NFE4	+	trans ¹ ; TE, RBC ²	5' del
	-58 to -122	multiple	+	trans ¹ ; TE, RBC ²	5' del
	-122 to -142	CACCC	+	trans ¹ ; TE, RBC ²	5' del
	-142 to -165	PAL	-	trans ¹ ; TE, RBC ²	5' del
	-165 to -214	BGP1	0	trans ¹ ; TE, RBC ²	5' del
Mouse β^{maj} promoter	-26 to -30	TATA	+	TE, HeLa ³ ; TE, MEL ⁴	point
	-32 to -53	DRE	+	TE, MEL ⁵	point
			0	TE, HeLa ⁵	
	-72 to -77	CCAAT	+	TE, HeLa ³ ; TE, MEL ⁴	point
	-87 to -95	CACCC	+	TE, HeLa ³ ; TE, MEL ⁴	point
Human β promoter	-113 to -600		0/+	TE, HeLa ⁶	5' del
	-28 to -31	TATA	+	thal ⁷	point
	-87, -88	CACCC	+	thal ⁷	point
	-184 to -138		+	SE, MEL ⁸	5' del
	-338 to -233		-	TE, K562 ⁹	5' del
Human ϵ promoter	-610 to -490		-	TE, K562 ⁹	5' del
	-177 to -392		-	TE, HeLa, K562 ¹⁰	5' del
Human α promoter	-55 to -87	CCAAT	+	TE, COS ¹¹	5' del
	-87 to -220		0/+	TE, COS ¹¹	5' del
	-220 to -525		0/-	TE, COS ¹¹	5' del
	-110 to -450		+	TE, HeLa, K562 ¹²	5' del
Chicken β enhancer	1617 to 1814		0	TE, RBC ¹³	del
	1825 to 1843	PAL	0	TE, RBC ¹⁴	cluster
	1847 to 1869	AP-1/NFE2	+	TE, RBC ¹⁴	cluster
	1879 to 1887	CACCC(?)	0/+	TE, RBC ¹⁴	cluster
	1892 to 1904	ERYF1	+	TE, RBC ¹⁴	cluster
	1905 to 1922	ERYF1	+	TE, RBC ¹⁴	cluster
	1939 to 2095		0/-	TE, RBC ¹³	del
	1877 to 2180		+	TE, HD3 ¹⁵	del

^a Relative to the cap site.

^b Function is indicated using the following symbols, which depict the effect mediated by the indicated element on transcription: +, definite increase; 0/+, weak or questionable increase; 0, no effect; 0/-, weak or questionable decrease; -, definite decrease.

^c Abbreviations used are: trans, in vitro transcription; TE, transient expression; SE, stable expression; thal, mutations in patients with β -thalassemia; RBC, embryonic chicken erythroid cells. References are:

¹Emerson et al 1989; ²Jackson et al 1989; ³Myers et al 1986; ⁴Cowie & Myers 1988; ⁵Myers et al 1989; ⁶Charnay et al 1985; ⁷reviewed in Weatherall et al 1989; ⁸Antoniou et al 1988; ⁹Berg et al 1989; ¹⁰Cao et al 1989; ¹¹Mellon et al 1981; ¹²Whitelaw et al 1989; ¹³Emerson et al 1987; ¹⁴Reitman & Felsenfeld 1988; ¹⁵Choi & Engel 1986.

^d Mutation method abbreviations: point, point mutations; 5' del, successive 5' deletions towards the cap site; cluster, clustered point mutations.

et al 1987, Reitman & Felsenfeld 1988). This enhancer can act on both the β^A and ε genes (Choi & Engel 1988; Nickol & Felsenfeld 1988); whether it also controls ρ and β^H is unknown. Choi & Engel (1988) have proposed that the switching between the ε and β^A promoters is governed by the relative strength of the physical interaction between each promoter and the β^A/ε enhancer. Such a competition model is supported by an experiment in which two enhancers are placed on the same plasmid; both promoters are then stimulated. Gallarda et al (1989) have suggested that in definitive cells (cells expressing adult-stage globins) the enhancer recognizes a stage selector element located at -44 to -55 bp in the β^A promoter.

HUMAN β CLUSTER ENHANCERS Four regions with enhancer activity have been discovered in this locus (Figure 1). In addition, the second intron of the γ gene may contain a fifth region (Donovan-Peluso et al 1987). The strongest known enhancer (up to 300-fold) in the human β locus is located 10.2–11.0 kb upstream of the ε gene (Tuan et al 1989) and is associated with the dominant control region (DCR). It and DCRs are discussed later in this review.

Earlier experiments, using constructs that did not include DCRs, revealed that both the human β and γ genes are developmentally regulated (human β like the mouse adult gene and γ like the mouse embryonic gene) in transgenic mice (summarized in Trudel & Costantini 1987). During the course of dissecting this regulation, two enhancers were found in the neighborhood of the human β -globin gene: one within the gene and another 550–900 bp downstream (Behringer et al 1987; Trudel & Costantini 1987; Kollias et al 1987). While the boundaries of the intragenic element have not been defined, the second intron (Collins et al 1990) and the third exon (Behringer et al 1987; Antoniou et al 1988) clearly have functional roles. The downstream enhancer increases expression about tenfold (Trudel & Constantini 1987) and contains multiple binding sites (Wall et al 1988) for an erythroid-specific protein (Eryf1²). When the intragenic and downstream enhancers were used together to drive a promiscuous promoter (SV40 early) in transgenic mice, the transgene was expressed in both a tissue- (erythroid) and stage- (adult) specific manner (Magram et al 1989).

A survey of 22 kb in the human β locus from 5' of ε to 3' of γ revealed an enhancer located 400–1150 bp 3' of the γ polyadenylation site. This enhancer maps to a tissue-specific DNase I hypersensitive site, but is not tissue specific in transient expression assays. It stimulates transcription 6- to 23-fold (Bodine & Ley 1987).

² An erythroid-specific protein was identified independently by several groups; it has been referred to as Eryf1 (Evans et al 1988b), NF-E1 (Wall et al 1988), or GF-1 (Martin et al 1989). For brevity and clarity, we will refer to it as Eryf1 throughout this text.

α -GLOBIN REGULATION To date no detailed functional analysis of any α -globin promoter has been performed. The only known α enhancer is located at the 3' end of the chicken α cluster; an ~ 35 -fold increase in expression is mediated by a 439 bp fragment containing three Eryf1 binding sites (Knezetic & Felsenfeld 1989). At present no human or mouse α cluster enhancers have been found (see Whitelaw et al 1989; Ryan et al 1989b; Hanscombe et al 1989). Strong evidence that such an enhancer-like element must exist is provided by a patient with α -thalassemia. This patient's DNA has a 62 kb deletion that ends 10 kb upstream of apparently normal $\alpha 2$ and $\alpha 1$ genes, yet no α -globin is made (Nicholls et al 1987).

A puzzling observation is that the human α genes, when carried on a replicating plasmid and assayed by transient expression in COS cells, do not require an enhancer for expression (Mellon et al 1981; Humphries et al 1982). When assayed on a non-replicating plasmid, however, an enhancer is required (Whitelaw et al 1989). In contrast, the mouse α (and human β and δ) genes are enhancer-dependent, even on replicating plasmids (Humphries et al 1982). This is thought to reflect plasmid copy number rather than a fundamentally replication-dependent mechanism (Whitelaw et al 1989). In contrast, Enver et al (1988) have proposed that replication per se is required for *Xenopus* β -globin expression in HeLa cells. The relevance of these observations to the situation in vivo remains to be established.

REGULATION OF GLOBIN EXPRESSION BY TRANS-ACTING CELLULAR FACTORS

Early work employing chromosome transfer, cell fusion, and transfection of cloned genes revealed the presence of cellular factors that mediate globin gene regulation. Recently, much of the emphasis has been on the identification and fractionation of nuclear DNA-binding proteins that recognize *cis*-linked regulatory elements. These factors can be grouped into three categories: (a) general RNA polymerase II promoter factors; (b) erythroid-specific factors that may potentiate expression of globin genes regardless of stage; (c) factors proposed to be involved in regulation of globin switching, that is, developmental stage-specific factors.

Cellular Factors Influence Globin Expression

ACTIVATION OF GLOBIN GENES FOLLOWING CELL FUSION Some of the first evidence that diffusible factors regulate globin expression came from chromosome transfer experiments performed by somatic cell fusion (summarized in Baron & Maniatis 1986). The formation of transient hetero-

karyons between two erythroid cells—mouse MEL cells and human K562 cells—resulted in the reciprocal activation of the human adult β and the mouse embryonic ϵ genes (Baron & Maniatis 1986). Cells with no history of globin expression were also fused to erythroid cells. Fusion of mouse cells to the embryonic-like human K562 cells resulted in activation of mouse embryonic globin genes, while fusion of human cells to the adult-like mouse MEL cells resulted in activation of human adult globin genes. There was one exception; the human ζ gene was never activated. Wrighton & Grosveld (1988) introduced a marker gene, under the control of various globin promoters, into non-erythroid cells and followed this gene as a sensitive reporter of gene activation in subsequent fusion experiments. The adult β promoter was only activated by fusion to MEL cells, whereas the γ promoter was activated by fusion to either MEL or K562 cells. Fusion to non-erythroid cells failed to activate either promoter.

These data not only directly demonstrate that diffusible cellular factors are involved in activation of the globin locus, but also implicate such factors in directing developmental switching. This must be reconciled with the results of Papayannopoulou et al (1986) who transferred human chromosome 11 (which contains the human β globin locus) from fetal erythroblasts into hybrid cells by fusion to the adult-like mouse MEL cells. In the adult-like MEL environment, the human γ gene is expressed initially. During time in culture, however, human globin transcription switches from the γ gene to the β gene. Thus, hemoglobin switching can, at least in these circumstances, occur within a single cell lineage (as opposed to replacement of a lineage expressing fetal globin by a distinct cell lineage expressing adult globin; see Ingram 1972; Weatherall et al 1976; Beaupain 1985). The existence of a developmental clock-like mechanism is supported by fusion of MEL cells to later stage human fetal erythroblasts; these hybrids switch from γ to β gene transcription after a shorter time in culture than do the hybrids derived from early fetal cells. Switching may be controlled by a developmental clock acting in *cis*, or by a *trans*-acting mechanism that is linked to human chromosome 11.

TRANS-ACTIVATION IN VITRO AND IN TRANSGENIC MICE In experiments in vitro designed to detect erythroid factors that influence globin gene expression, Bazett-Jones et al (1985) found that pre-incubation of globin gene templates in K562 cell extracts gave up to a 30-fold increase of transcription in a HeLa whole cell extract. The effect was specific for extracts from K562 cells that had been induced by hemin; although non-globin genes were not affected by the incubation, both γ and β gene transcription was enhanced, even though the latter gene is not expressed in K562 cells.

More recently Emerson et al (1989) have prepared transcriptionally active extracts from embryonic chick erythroid cells. The extracts maintain some developmental specificity since transcription of the adult β^A gene is most efficient in extracts from the adult-like definitive cell lineage. This system is particularly promising for the analysis of developmental changes in the activities of factors that interact with globin promoter sequences.

Experiments with transgenic mice provide the most direct evidence for *trans*-acting factors that regulate the expression of globin genes. Human globins undergo two developmental switches (embryonic \rightarrow fetal \rightarrow adult), but mice undergo only one (embryonic \rightarrow fetal/adult). Costantini and co-workers (Chada et al 1986) introduced individual human γ or β globin genes into mice to determine which developmental program they followed. The transgenic human β gene was expressed in both fetal and adult mouse erythroid cells; the human γ gene was expressed only in mouse embryonic erythroid cells. Under the conditions of these experiments, the *trans*-factors of the mouse cells appear to be dominant in determining cell-type and developmental regulation of the transgene. These effects are modified, however, by the action of the dominant control region (see DCRs, below).

DNA-Binding Proteins

UBIQUITOUS FACTORS BIND TO CONSERVED PROMOTER ELEMENTS By analogy to other eukaryotic genes that also contain a conserved ATA sequence around position -30 (TATA box), the globin TATA box is presumed to bind the pol II transcription factor TFIID (Parker & Topol 1984; Sawadogo & Roeder 1985; Buratowski et al 1988; Cavallini et al 1988). At this time there is no evidence for or against cell-type variants of this factor within a given species.

The literature on CCAAT binding factors (CBFs) is rather confusing because of the presence of at least five distinct proteins that recognize this element with varying affinities (Chodosh et al 1988; Dorn et al 1987; Graves et al 1986; Jones et al 1987; Santoro et al 1988). Analysis of the factors that interact with globin gene CCAAT elements provides some evidence as to which CBFs interact preferentially with globin promoters *in vitro*. For example, it seems likely that the ubiquitous CP1 binds with highest affinity to CCAAT elements in the promoters of the human α , β , and γ genes (Chodosh et al 1988; deBoer et al 1988; Gumucio et al 1988; Superti-Furga et al 1988), while Yu et al (1990) have suggested that CP2 preferably interacts with the human ζ promoter.

Two lines of evidence indicate that CP1 may be a globin transcriptional activator: first, an upstream CP1 binding site is necessary (in addition to a binding site for the erythroid-specific factor Eryf1) for regulated expression from the human β promoter in MEL cells (deBoer et al 1988). Second, a

mutation that increases the similarity of a CCAAT sequence to the consensus binding site for CP1 (Chodosh et al 1988) creates a stronger binding site in the human γ promoter (Gumucio et al 1988), which correlates with increased expression of this gene (see — 117 HPFH, below).

CTF/NF1 from HeLa cells has been shown to bind to a CCAAT box and stimulate transcription *in vitro* from the human α promoter (Jones et al 1987). In contrast, in the chicken β^A promoter, CTF/NF1 or a related factor interacts not with a CCAAT element, but with an NF1 binding site (see Pal protein, below); this binding prevents interaction of a positive transcription factor with an adjacent site.

CTF/NF1 or a related protein is also capable of inhibiting the binding of a positive transcription factor to an adjacent site in the promoters of the chicken β^H (Plumb et al 1986) and human β genes (deBoer et al 1988). At least in the latter case, the CTF/NF1 binding activities are the same as, or related to, the vertebrate homologue of the sea urchin CCAAT displacement protein (CDP; Barberis et al 1987). CDP also binds to the duplicated CCAAT elements of the human γ promoter and inhibits binding of CP1 (Superti-Furga et al 1988). A similar situation may exist in the mouse α promoter (Barnhart et al 1988). Although only the case of the chicken β^A promoter has been confirmed by *in vivo* footprinting (Jackson et al 1989), we might speculate that displacement of positive transcription factors by proteins related to CTF/NF1 is a general globin gene regulatory mechanism.

Although the CACCC box is an important part of β globin promoters, the sequence is also found in a number of non-erythroid regulatory regions (Schule et al 1988; transcription factors. For example, several groups have shown that the pol II transcriptional activator Sp1 (Kadonaga et al 1986), derived from either erythroid or non-erythroid cells, interacts with the CACCC box (Lewis et al 1988; Barnhart et al 1988; Emerson et al 1989; Jackson et al 1989). This may explain the functional importance of the CACCC box for expression in non-erythroid cells (Myers et al 1986).

On the other hand, proteins that may be more tissue-restricted have also been identified as binding to the CACCC box. For example, a cell-type specific binding activity (TEF-2; Xiao et al 1987; Nomiyama et al 1987; Davidson et al 1988) to a CACCC motif in the SV40 enhancer is distinct from Sp1. Most of the evidence supports the notion that erythroid cells contain two or more CACCC binding proteins, probably some or all of which are related to Sp1 and TEF2, and at least one of which is usually abundant in erythroid cells (Mantovani et al 1988; Emerson et al 1989; Jackson et al 1989; Catala et al 1989). The CACCC binding proteins are transcriptional activators; addition of DNA site affinity-enriched factors

to depleted *in vitro* transcription reactions restores basal promoter activity of the chicken β^A gene (Emerson et al 1989). Furthermore, the CACCC element is bound by protein specifically in nuclei of erythrocytes that are transcribing this gene (Jackson et al 1989).

The same CACCC factor may also bind to sequences in the human 3' (Giglioni et al 1989) and chicken β^A/ϵ (Emerson et al 1987) enhancers, but there is no evidence that these CACCC enhancer sites mediate transcriptional activity (Reitman & Felsenfeld 1988).

Other binding sites for Sp1, besides the CACCC element, have been identified in a number of globin promoters. The α promoters of chicken (Kemper et al 1987), mouse (Barnhart et al 1988), and human (Whitelaw et al 1989) all contain sequences that resemble Sp1 consensus binding sites (Briggs et al 1986). The chicken β^A promoter has been shown to bind human or avian Sp1 (Evans et al 1988a; Lewis et al 1988; Emerson et al 1989) at a site critical for transcriptional activity (Jackson et al 1989; Emerson et al 1989). Thus Sp1 may be important for globin transcription as it is for many other eukaryotic genes (Kadonaga et al 1986). This regulation, however, might involve erythroid-specific or developmental stage-specific variants or modifications of Sp1 (see NF-E4, below). Yu et al (1990) have shown that Sp1 binds to the human ζ promoter, but may be displaced *in vitro* by erythroid-specific factors that bind to overlapping sites. Furthermore, aberrant binding of Sp1 has been proposed to mediate certain forms of HPFH (see below).

A distinct protein that may be related to Sp1 binds to a stretch of dG residues in the chicken β^A promoter (Emerson et al 1985; Plumb et al 1986). The protein, designated BGP1, was purified by Lewis et al (1988); their data indicate that both BGP1 and Sp1 require zinc for binding and cross-react with the same antibody. It has been speculated that binding of BGP1 to the upstream promoter may be involved in generating the chromatin structure associated with active transcription of the β^A gene in definitive erythrocytes; no data yet supports this idea. Similar binding activities to the poly(dG) region have since been found in non-erythroid and primitive erythroid cells (Jackson et al 1989; Emerson et al 1989), although it is not known if they are identical to BGP1.

ERYTHROID-SPECIFIC FACTORS The erythroid-specific factor Eryf1 (erythroid factor 1; Evans et al 1988b) was identified independently in several laboratories; it is also referred to as GF-1 (Martin et al 1989) or NF-E1 (Wall et al 1988). It has become apparent that Eryf1 binding sites are present in regulatory regions (promoters and/or enhancers) for all chicken globin genes (Evans et al 1988b; Knezetic & Felsenfeld 1989) and probably for all human and mouse globin genes as well (Galson & Hous-

man 1988; Wall et al 1988; deBoer et al 1988; Catala et al 1989; Barnhart et al 1989; Perkins et al 1989). The DNA sequence to which Eryf1 binds includes the conserved motif WGATWR (W = A or T; R = A or G).

The cDNAs encoding chicken (Evans & Felsenfeld 1989), mouse (Tsai et al 1989), and human (Trainor et al 1990; Zon et al 1990) Eryf1 have been cloned. Murine and human Eryf1 are highly conserved proteins; 86% of the 413 amino acids in the two proteins are identical. The chicken protein is smaller (304 amino acids) and has apparently undergone several deletions relative to the mammalian factors (Trainor et al 1990). All three proteins are highly conserved in a cysteine rich, central region containing two repeats of the form CXNCX₄TXLWRRX₃GXCNAC. Although these repeats are similar in cysteine arrangement to zinc fingers of the class found in the DNA-binding domain of the steroid hormone receptor family (Evans 1988; Berg 1989), Eryf1 does not otherwise resemble these receptor proteins. The Eryf1 fingers are actually much more similar to a single finger present in the regulatory protein areA from *Aspergillus nidulans* (Arst et al 1989), which indicates that the motif is perhaps of ancient and distinct origin. The protein sequence in this region of Eryf1 has evolved at a remarkably low rate; other regions of the protein may be derived from a distinct repeat and appear to be evolving at a much higher rate (Trainor et al 1990).

Expression of Eryf1 is regulated at the RNA level; no Eryf1 RNA is detectable in a variety of non-erythroid cells (Evans & Felsenfeld 1989; Tsai et al 1989). No major differences in the abundance of Eryf1 RNA or in the amount of Eryf1 binding activity have been noted in different erythroid developmental lineages, or as a result of chemical induction of erythroid cell differentiation in culture.

Binding sites for Eryf1 are present in control regions of erythroid genes encoding proteins other than the globins. These include the erythroid-specific promoters of the human genes encoding uroporphobilinogen deaminase (Plumb et al 1989; Mignotte et al 1989b), carbonic anhydrase I (Brady et al 1989), and the erythropoietin receptor (S. Orkin, personal communication), as well as the chicken histone H5 enhancer (Rousseau et al 1989). These observations indicate that Eryf1 expression may be an important factor in establishment of erythroid lineages. Because Eryf1 has been detected in the developmentally related megakaryocyte lineage, however, Eryf1 expression alone may not be sufficient for erythroid determination (Martin et al 1990; Romeo et al 1990).

An erythroid-specific factor distinct from Eryf1 has been described and designated NF-E2. This factor recognizes the same DNA consensus sequence (TGACTCA) as AP-1 (*jun*; Lee et al 1987; Angel et al 1988; Bohmann et al 1987), but appears to be distinct; Mignotte et al (1989a)

studied a point mutation that abolished binding of NF-E2 without inhibiting binding of the ubiquitous factor AP-1. The NF-E2 binding motif was originally identified as a promoter element of the porphobilinogen deaminase gene (Mignotte et al 1989b). Perhaps it is the same protein that binds to an AP-1 site within a positive activating element of the chicken β^A/ϵ enhancer (Reitman & Felsenfeld 1988); binding sites for AP-1 or NF-E2 are also found within the human β DCR at -11 kb (P. Ney & A. Nienhuis, personal communication).

STAGE-SPECIFIC FACTORS A protein designated the Pal factor binds within the nuclease hypersensitive domain of the chicken β^A promoter. This protein recognizes a *palindromic* sequence related to the binding site for CTF/NF1; it seems likely that it is also related to the TGGCA protein isolated from chicken liver (Rupp & Sippel 1987; Rupp et al 1988). Sequence data obtained from purified peptides confirm that Pal is a member of the NF1 family (C. Lewis & G. Felsenfeld, unpublished results).

Binding activity of the Pal factor increases two orders of magnitude during the course of erythroid cell development, so that it is most abundant in late embryonic or adult circulating erythrocytes in which the globin gene is transcriptionally inactive; the Pal factor site is progressively bound in nuclei as development proceeds (Jackson et al 1989). Occupancy of the Pal site correlates with vacancy of the adjacent CACCC site *in vivo*, thus indicating that Pal may be a transcriptional repressor that acts to shut down adult globin expression late in differentiation by preventing binding of (or by displacing) the positive acting CACCC factor. Promoter constructions in which the Pal binding site is deleted show increased transcription levels (Jackson et al 1989; Emerson et al 1989). Furthermore, highly enriched preparations of Pal protein act to repress *in vitro* transcription of β globin templates unless the Pal site is mutated (Emerson et al 1989). The β^H gene, which is turned off in circulating embryonic erythrocytes prior to the β^A gene, contains an even stronger Pal factor promoter binding site (Plumb et al 1986). It has been suggested that Pal might, therefore, direct the coordinate repression of chicken β genes (Jackson et al 1989). As mentioned above, displacement of a positive activator by a Pal-related protein may be a general globin gene regulatory mechanism.

Although Pal factor activity is most abundant in transcriptionally inactive adult circulating erythrocytes, it is not yet known whether Pal activity accumulates during definitive cell maturation. Perhaps there are specific negative regulatory mechanisms that function for the β^A gene, superimposed on the general suppression of transcription that attends erythroid cell maturation.

Choi & Engel (1988) identified a region of the β^A promoter (the stage-selector element) that is required in transient transfections of definitive erythroid cells for both the expression of the adult gene and the suppression of a linked embryonic ε gene. Although they did not analyze proteins from erythrocytes that express the adult gene, Gallarda et al (1989) identified a specific binding activity to this region and to the enhancer. This activity (designated NF-E4) is present in mature adult erythrocytes, but not in primitive erythrocytes that express the ε gene. Because the β^A enhancer is shared between the β^A and ε genes (Nickol & Felsenfeld 1988; Choi & Engel 1988), NF-E4 might augment a productive interaction between the enhancer and the adult-stage promoter. The same promoter site has previously been shown to bind strongly to Sp1 from both primitive and definitive cells (Lewis et al 1988; Jackson et al 1989; Emerson et al 1989). It is not clear how Sp1 and NF-E4 are related.

FACTORS IMPLICATED IN MEDIATING THE HPFH SYNDROME Hereditary persistence of fetal hemoglobin (HPFH) is a syndrome characterized by the inappropriate expression of human γ genes in adults (reviewed in Wood 1989). In a few cases, linkage analysis has shown that the HPFH mutation is not linked to the β locus (Gianni et al 1983; Jeffreys et al 1986). Characteristic of these cases is the overexpression of both $^G\gamma$ and $^A\gamma$; the genetic lesion could be due either to a defective DNA-binding transcription factor, or to an altered protein involved in maturation of erythroid cells. The recent mapping of the human Eryf1 (GF-1) gene to the X-chromosome (Zon et al 1990) may be relevant in explaining an X-linked form of HPFH (Miyoshi et al 1988).

Very strong evidence that the cellular environment provides diffusible factors that help mediate HPFH was provided by Papayannopoulou et al (1988), who transferred human lymphoid chromosomes into MEL cells by cell fusion. When lymphocytes containing a normal chromosome are fused, hybrids support expression of only the adult human β gene. In contrast the γ genes on chromosomes containing the HPFH mutations become activated after fusion with MEL cells. The cell type from which the HPFH chromosome was derived does not affect the activation, which occurs at the transcriptional level.

Numerous cases have been identified in which the only difference between a HPFH mutant and a normal γ gene is a point mutation in the promoter (see Figure 1). In these cases, only the γ gene that is mutated is abnormally expressed. The failure of these HPFH genes to be turned off efficiently could result from the abnormal binding of a positive transcription factor to the mutated site. Alternatively, the normal binding of a negative regulator could be blocked by the mutation (see Gumucio et al 1988; Wood 1989).

The HPFH mutation at -117 in the distal CCAAT results in a better match to a consensus CPl binding site; increased binding (about twofold) of this ubiquitous factor to the mutated site has been reported (Gumucio et al 1988). Superti-Furga et al (1988) proposed a different mechanism involving decreased binding of a factor that normally represses the γ gene. They find that the -117 mutation inhibits binding by an erythroid-specific activity designated NF-E3. This interpretation is supported by the discovery of an HPFH promoter that is deleted for the entire CCAAT element and thus binds no CBF (Mantovani et al 1989).

Interpretation of another kind of HPFH, associated with a point mutation at -175, is complicated by the binding at this site of multiple factors (Eryf1 and the ubiquitous octamer-binding protein OTF1). The mutation does not affect the binding affinity of Eryf1, but greatly inhibits binding of OTF1 (Martin et al 1989). Nevertheless, the effect on transcription appears to be mediated by Eryf1 and not OTF1; other mutations that inhibit OTF1 binding have no effect on transcription (Martin et al 1989; Nicolis et al 1989). Gumucio et al (1990) have noted that the OTF1 binding site is not conserved among primates, while the Eryf1 binding site is conserved. Martin et al (1989) have proposed that the -175 mutation alters the binding of Eryf1 in some way, perhaps changing normal protein-protein contacts needed for gene activation. It is not clear, however, why certain in vitro generated mutations that completely abolish Eryf1 binding have no effect on transcription.

Another class of HPFH mutations map to a G/C rich region around position -200 of the γ promoter. Mutations at -198 or -202 are found to result in increased binding in vitro by Spl (Ronchi et al 1989; Sykes & Kaufman 1990). As is true for all of the HPFH point mutations, interpretation of these correlations is difficult because of the lack of an appropriate model system. The mutated promoters, when transfected into cultured cells, generally do not accurately reproduce the high level of γ expression observed in the HPFH syndrome (see also, Charnay & Henry 1986; Stoeckert et al 1987; Gumucio et al 1988; Rixon & Gelinis 1988; Mantovani et al 1989; Nicolis et al 1989; Lloyd et al 1989; Martin et al 1989; Ulrich & Ley 1990).

CHROMATIN STRUCTURE OF THE GLOBIN GENES

Matrix Attachment Regions (MARs)

MARs are defined as regions of DNA that preferentially associate with detergent-extracted nuclear remnants (which are called nuclear matrices

or scaffolds). While unproven, the rationale behind the MAR assay is the detection of physiologic sites of DNA attachment to the nuclear matrix. In this view, MARs should delineate the ends of DNA loops. Their relationship to DCRs, or other elements that regulate gene expression, is uncertain.

In the chicken β cluster, MAR activity is found in two regions, one including the ρ gene and flanking regions, and the other the β^A/ϵ enhancer and flanking DNA (Bennet et al 1989). The human β cluster has at least eight MARs in the 90 kb mapped, ranging from ~ 20 kb upstream of ϵ to ~ 18 kb 3' of the β polyadenylation site (Jarman & Higgs 1988). In contrast, no MARs were detected in 140 kb encompassing the α locus. Unfortunately, no simple model is suggested by these data. For example, if MARs do represent loop boundaries, then the average loop size in the globin cluster would only be ~ 10 kb. This is possible, but is not in agreement with physical estimates for the average genome loop size. It is not yet clear whether all MARs have the same structure, or the same function.

Chromatin Structure and Globin Gene Expression

NUCLEOSOMES IN THE GLOBIN CLUSTER Genes that are competent for transcription are packaged as chromatin, although the structure of the chromatin of active genes (active chromatin) may involve partial modification or disruption of the packaging. Analysis of nucleosome phasing along the mouse β^{maj} gene shows that, aside from a disruption between positions -200 and $+500$ when the gene is active, the nucleosome positions are quite similar on the transcriptionally active or inactive gene (Benezra et al 1986). The transcriptionally active chicken β^A gene is also covered with nucleosomes, although in this case there is some evidence for changes in local structure (Sun et al 1986).

The folding of globin sequences into higher order chromatin structures has also been studied. It is possible to isolate fragments of 30-nm thick chromatin fiber from erythroid nuclei by digestion with either micrococcal nuclease (Fisher & Felsenfeld 1986), or the restriction enzyme EcoRI (Kimura et al 1983; Caplan et al 1987). The extent of compaction of the isolated chromatin fiber can be deduced from its sedimentation properties in sucrose gradients. Fragments containing the β^A gene obtained from nine day or adult erythroid cells are only slightly less compacted than bulk chromatin fragments containing DNA of the same size. The β^A gene in these erythroid cells thus exists in a considerably compacted state that may reflect the presence of either a partially disrupted 30-nm fiber (Fisher & Felsenfeld 1986) or a fully folded fiber interrupted by nucleosome-free regions (Caplan et al 1987).

Nucleosome-free regions exist not only in the neighborhood of the chicken β^A gene, but throughout the β cluster: they give rise to the sites, hypersensitive to nuclease digestion, that mark the presence of active DNA regulatory elements. Much of the early work (Stalder et al 1980; McGhee et al 1981) focused on local hypersensitive sites near individual globin genes in chicken erythroid cells. Similar behavior is seen in all species, and the phenomenon of hypersensitivity is quite general (reviewed in Gross & Garrard 1988). The presence of hypersensitivity is often correlated with the potential for expression of the nearby genes; subsequent investigation has identified many of the globin gene hypersensitive regions as the sites of promoters or enhancers.

SUPER-HYPERSENSITIVE SITES The discovery (Tuan et al 1985; Forrester et al 1986) in the human β cluster of a quite different class of hypersensitive sites, located far from the individual genes, has led to important new insights concerning developmental regulatory mechanisms. These have been named major or super-hypersensitive sites (SH sites) because they appear earlier in the course of DNase I digestion than do the local (minor) hypersensitive sites described above. Tuan et al (1985) identified five SH sites in the region 6–25 kb 5' of the human ϵ gene, and a single site about 20 kb 3' of the β gene (Figure 1; see below for further discussion). The sites thus demarcate a region of about 90 kb that contains within it the entire family of β -like genes. Most of the SH sites are erythroid cell-specific, but unlike the stage-specific local hypersensitive sites, they are found in a variety of erythroid cells, regardless of whether the principal expressed gene is embryonic, fetal, or adult globin. The chicken β cluster has four upstream SH sites, one of which is constitutively expressed; the rest are erythroid-specific (Reitman & Felsenfeld 1990).

The location of the SH sites relative to the globin genes, taken together with their developmental program of appearance, led to the suggestion (Tuan et al 1985) that these sites might define the boundaries of a large domain of active chromatin and might also be involved in generating the domain's conformation. Evidence for the existence of such domains comes from earlier studies of the chicken β (Stalder et al 1980) and ovalbumin (Lawson et al 1982) gene families. In erythroblasts and oviduct, respectively, these domains exhibit diffuse low-level DNase I sensitivity both within the gene clusters and extending many kilobases beyond them in each direction. Evidence that the SH sites might serve as locus activation regions also came from experiments (Forrester et al 1987) in which human non-erythroid cells were fused with MEL cells; after fusion the human β -globin cluster displayed all of the upstream and downstream SH sites characteristic of transcriptionally active tissue.

DOMINANT CONTROL (LOCUS ACTIVATION) REGIONS The properties of the human globin SH sites suggested (Tuan et al 1985; Forrester et al 1986) that they might function as permissive regulatory elements which, through their effect on chromatin structure, act on all of the genes in the β cluster. Strong evidence for the existence of regulatory elements far upstream in the β locus had already been provided by Kioussis et al (1983), who analyzed an extensive upstream deletion locus in a γ , β -thalassemic patient. They showed that the β gene itself contained no point mutations and that it could be transcribed in HeLa cells to give an mRNA indistinguishable from that found in normal reticulocytes. Moreover, they demonstrated that the DNA brought by the deletion into proximity with the β gene is packaged in chromatin that is insensitive to DNase I, and that it is methylated in a way consistent with transcriptional inactivity. They concluded that the deletion provided "a position effect similar to those found in *Drosophila*." This could have arisen either through the action of a distant negative element brought close to the gene, or by removal of one or more positive *cis*-acting elements.

Recent results entirely substantiate the latter point of view and show that the distant control elements are associated with DNA in or near the SH sites. In order to examine the properties of the sites, Grosveld et al (1987) constructed a minilocus linking ~21 kb of DNA containing all of the upstream SH sites, the human β gene with all of its known local regulatory elements, and a 12 kb region containing the downstream SH site. Transgenic mice carrying this minilocus-express human β -globin in a tissue-specific manner, but with two novel properties: (a) With the exception of individuals in which the DNA has undergone deletion or rearrangement, every transgenic mouse expresses the human β gene. (b) High levels of human β mRNA and protein are produced, with levels per introduced gene copy comparable to the levels found for the endogenous mouse globin genes; the amount of human β -globin is thus roughly proportional to the number of gene copies.

This is quite different from the results obtained in experiments with transgenic mice in which only the local regulatory sites accompany the introduced β gene. In typical cases (see Ryan et al 1989a), most of the transgenic animals (70%) do not express detectable amounts of the transgene; when they do, the levels of expression per copy are typically less than 1% that of the endogenous gene, and there is no simple relationship between gene copy number and expression levels. The large number of failures to express has usually been ascribed to position effects: a DNA sequence inserted at an arbitrary position in the genome is likely to find itself surrounded by compact heterochromatin. It is clear that the presence of DNA from the distant regions near the SH sites not only liberates

associated genes from the position effect, but also assures high levels of expression. It is therefore reasonable to call these locus activation (Forrester et al 1987) or dominant control (Grosveld et al 1987) regions (LARs or DCRs).

A number of obvious questions must now be asked: How many sequence elements contribute to DCR activity? Is the sequence that confers position independence identical to that which results in high expression levels, or are they separable? Are these activities directly related to the generation of the extended domain of low-level DNase I sensitivity? Does the DCR also contain sites of attachment to the nuclear matrix? Does the DCR confer both tissue- and developmental-stage-specificity, and can it operate on any gene?

DOMAINS WITHIN THE DCR Work addressing some of these questions has begun in several laboratories. Considerable effort has been concentrated on the dissection of the DCR. The five upstream SH sites of the human β -locus are situated (Tuan et al 1985) about 6, 11, 14.5, 17.5, and 21 kb upstream (5') of the ϵ gene.³ Of these, the sites at -6, -11, and -17.5 kb are erythroid-specific. The site at -21 kb is present both in erythroid and at least one non-erythroid cell line and might belong to an adjacent gene cluster; the site at -14.5 kb is present in HEL cells (a human erythroleukemia line), but not in K562 cells (Tuan et al 1987). In addition to the three upstream erythroid-specific sites, a fourth such site is found downstream, 20 kb 3' of the β gene.

Not all of the sites are necessary for the action of the DCR. Omission of the downstream site has little effect on the production of β mRNA in transgenic mice when compared to a construction containing all of the sites (a mean of 108% of the amount of mouse β message per human gene copy, with a range of 16–200% with the downstream site vs 52%, with a range of 20–84% without it; Ryan et al 1989a; see also Talbot et al 1989). Further deletions that retain only the -6 and -11 kb sites, or the -11 kb site alone, result in transgenic mice that still express human β mRNA at a level 40–50% that of the mouse β mRNA per gene copy. Other laboratories have also identified the -11 kb site as sufficient to produce high levels of DCR activity and have defined the active element more precisely. A DNA segment only 882 bp long, containing the -11 kb site, is sufficient to give expression in transgenic mice at a level of 30–300% of mouse β expression (Curtin et al 1989). A patient with β -thalassemia, whose DNA contains a deletion encompassing the four SH sites from 11–

³ These are numbered I–V by Tuan et al (1985); Grosveld et al (1987) considered only the first four, and numbered them 4–1. We will refer to them by their approximate distance from the ϵ gene, as given in the first reference.

21 kb upstream of the ϵ gene, provides evidence that the site at -6 kb, together with the single downstream site, is not sufficient for DCR activity (Driscoll et al 1989).

Quite similar conclusions have been derived from studies with stably transformed cell lines. MEL cells transformed with DNA containing the entire upstream DCR, coupled to the human β gene, support high levels of globin mRNA production, again in quantities roughly proportional to gene copy number (Forrester et al 1989; van Assendelft et al 1989). In contrast, transformed cells carrying the gene not coupled to the DCR are an order of magnitude less efficient in expression. Constructions that contain only the upstream sites at -6 and -11 kb function at about 30% of the efficiency of the full DCR when transformed into MEL cells, consistent with the transgenic results (Forrester et al 1989). Furthermore, constructions containing only the pair of sites at -14.5 and -17.5 kb also function effectively (40% efficiency) to activate human globin gene expression. These results suggest that DCR activity is shared in an additive fashion among DNA sequences near two or three of the upstream SH sites.

FUNCTION OF THE DCR The human β minilocus has been fused (van Assendelft et al 1989) either to the human β gene (see above) or the Thy-1 gene in constructions also carrying a selectable marker, tk-neo. When MEL cells are stably transformed with this DNA and induced, high and copy number-dependent levels of expression are observed for either β or Thy-1, and in both cases for tk-neo. Elevated levels of human β expression are also obtained upon transformation of the human γ -expressing cell line, K562, but the DCR has no effect when the non-erythroid mouse L cell line is transformed. It appears that the β -domain's DCR can confer tissue specificity on a variety of genes (provided that they possess appropriate local regulatory elements) whether or not they are normally erythroid-specific, but that in these experiments the DCR alone is not sufficient to confer stage specificity of expression.

It had not been possible to obtain α expression in transgenic mice carrying either individual α genes or larger portions of the human α cluster. The ability of the β DCR to activate other genes has been employed to overcome this obstacle. Introduction of the human β DCR, or a DNA segment carrying the -6 and -11 kb sites, is sufficient to produce high levels of α message (Ryan et al 1989a; Hanscombe et al 1989). This provides a practical solution for those who wish to assemble a complete human hemoglobin in the mouse, which is particularly useful for the development of systems for studying human hemoglobinopathies (Behringer et al 1989; Greaves et al 1990; Ryan et al 1990).

As noted above, the human β DCR is not sufficient to confer developmental stage specificity in stably transformed cell lines. The same behavior has been observed in transgenic mice (Enver et al 1989). Mice carrying the human fetal globin gene, γ^A (with its local regulatory elements), fused to DNA containing the four upstream SH sites, express γ^A in an erythroid-specific manner, but in both primitive and definitive cells. In earlier transgenic mouse studies in which the same gene was introduced without the DCR, the γ^A gene was expressed in a stage-specific manner (i.e. in primitive cells only), although at a lower level (Chada et al 1986; Kollias et al 1986). There are a number of possible explanations for this behavior. It was suggested (Enver et al 1989) that in constructions such as these, which bring the DCR much closer than normal to the gene, the DCR might act as a powerful stage-independent enhancer (see below); alternatively, the deleted sequences that normally separate the DCR from the genes might contain negative control elements.

A different and attractive model is based on the idea that the local control elements (the promoters and enhancers of the individual genes) compete with each other for interaction with the DCR, and that stage-specific expression depends upon the presence of the full complement of genes. In this view, γ^A is expressed in definitive cells in the above experiments because it is no longer in competition with other genes in the cluster. The model is supported by recent results in transgenic mice (Enver et al 1990; Behringer et al 1990) showing that stage-specific regulation is restored when constructions carrying both β - and γ -globin genes are used, although it must be kept in mind that these experiments do not exclude possible effects of proximity.

The competition model has also been proposed to account for altered relative levels of γ and β gene expression in certain HPFH mutants (see for example Wood 1989); direct evidence for the operation of such a mechanism in the chicken ϵ to β^A switch has been provided by Choi & Engel (1988). Competition models of this kind usually invoke physical contact between the proteins of the DCR and the promoter or enhancer, accompanied by looping out of the intervening DNA. In assessing such proposals, a distinction can be made between reversible and irreversible models. In a reversible model we suppose (for example) that the β and γ promoters are both accessible to compete for interaction with the DCR within the expressing cell and that the observed preferential interaction reflects higher binding affinity or frequency of interaction with one of the two promoters. Such a model requires that all of the necessary *trans*-acting factors for both promoters still be present. In irreversible models, the competition event has occurred earlier (e.g. during cell division), perhaps inducing a chromatin conformation that locks in the ultimate pattern of

expression. (Of course “irreversibly” formed structures might be unlocked by events like cell fusion.) The fact that cells at different developmental stages have different chromatin structures at the inactive promoters (hypersensitive domains are absent) suggests that any early competition events must have been followed by further changes. It must be kept in mind that in most cases cells expressing different globin genes belong to different lineages.

THE DCR AS AN ENHANCER The properties of the β DCR in some ways resemble those of a distant and powerful enhancer, although the DCR clearly raises transcription levels in a way that the local β family enhancers and promoters do not. Various portions of the β DCR have been tested for enhancer function in transient expression experiments in K562 cells. In these experiments (Tuan et al 1989) the DCR element to be tested is positioned upstream of an ϵ promoter fused to the CAT gene. Levels of expression are raised up to 300-fold by introduction of a fragment 0.8 kb long that spans the -11 kb SH site (site II). This is the same fragment shown by Curtin et al (1989, see above) to function as a fully competent DCR in transgenic mice. In earlier experiments (Tuan et al 1987), fragments containing this site gave 20–30-fold stimulation of a reporter gene, but none of the other upstream DCR elements had any effect. [Note that two of the other elements, at -14.5 and -17.5 kb, are together sufficient to give appreciable DCR activity in stably transformed cells (Forrester et al 1989)].

Whether or not the DCR functions primarily as a super-enhancer, interacting with local regulatory elements, it has a dominant effect on chromatin structure as well, insulating the genes in its domain against the presumed suppressive effects of surrounding heterochromatin. Perhaps this effect arises directly from the presence of unusually large SH sites, which may create defects in the cooperative packing of nucleosomes that propagate outward from the sites over large distances. It is, however, equally possible that some elements of the DCR function specifically as nuclear matrix attachment points and that mechanisms associated with the matrix induce the unfolding. It should be noted (Jarman & Higgs 1988) that MARs have been identified in the human β locus, some of them near the DCR, but not near the -11 kb SH site that, when present alone, displays a large fraction of the activity of the complete DCR. In any case, the discovery of dominant control (locus activation) regions provides our first information about the coupling of chromatin structure to gene expression and how this coupling is exploited to control globin gene activation and developmental switching.

Our understanding of promoters, enhancers, and dominant control

regions is conditioned by the assay systems used to study them. Although transient expression studies (and transcription *in vitro*) can reveal certain aspects of regulatory mechanisms, they are not likely to detect effects that depend upon long-range chromatin structure. The full effects of DCRs may be displayed only in stably transformed cells or transgenic mice.

The properties of the DCR suggest (Grosveld et al 1987) a hierarchical model of control in which tissue specificity is determined *in vivo* primarily by the DCR, and stage specificity is conferred by the individual enhancers and promoters. In this view, the role of the DCR is at least permissive: it makes the local promoters and enhancers accessible to the binding of *trans*-acting factors. Although the DCRs are a great distance away in terms of DNA sequence, the packaging of DNA in chromatin fibers, and its probable attachment to the nucleoskeleton, open the possibility for intimate contact of all the regulatory elements within a functioning complex. New methods must now be devised to permit study of those complexes.

ACKNOWLEDGMENTS

We thank members of the Laboratory of Molecular Biology and Dr. A. W. Nienhuis for helpful comments on the text. M. R. is a Lucille P. Markey Scholar, and this work was supported in part by a grant from the Lucille P. Markey Charitable Trust.

Literature Cited

- Allan, M., Lanyon, W. G., Paul, J. 1983. Multiple origins of transcription in the 4.5 kb upstream of the epsilon-globin gene. *Cell* 35: 187-97
- Angel, P., Allegretto, E. A., Okino, S. T., Hattori, K., Boyle, W. J., et al. 1988. Oncogene *jun* encodes a sequence-specific *trans*-activator similar to AP-1. *Nature* 332: 166-71
- Antoniou, M., deBoer, E., Habets, G., Grosveld, F. 1988. The human beta-globin gene contains multiple regulatory regions: Identification of one promoter and two downstream enhancers. *EMBO J.* 7: 377-84
- Arst, H. N., Kudla, B., Martinez-Rossi, N., Caddix, M. X., Sibley, S., Davies, R. W. 1989. *Aspergillus* and mouse share a new class of 'zinc finger' protein. *Trends Genet.* 5: 291
- Barberis, A., Superti-Furga, G., Busslinger, M. 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. *Cell* 50: 347-59
- Barnhart, K. M., Kim, C. G., Banerji, S. S., Sheffery, M. 1988. Identification and characterization of multiple erythroid cell proteins that interact with the promoter of the murine alpha-globin gene. *Mol. Cell. Biol.* 8: 3215-26
- Barnhart, K. M., Kim, C. G., Sheffery, M. 1989. Purification and characterization of an erythroid cell-specific factor that binds the murine alpha- and beta-globin genes. *Mol. Cell. Biol.* 9: 2606-14
- Baron, M. H., Maniatis, T. 1986. Rapid reprogramming of globin gene expression in transient heterokaryons. *Cell* 46: 591-602
- Bazett-Jones, D. P., Yeckel, M., Gottesfeld, J. M. 1985. Nuclear extracts from globin-synthesizing cells enhance globin transcription *in vitro*. *Nature* 317: 824-28
- Beaupain, D. 1985. Line-restricted hemoglobin synthesis in chick embryonic erythrocytes. *Cell Diff.* 16: 101-7

- Behringer, R. R., Hammer, R. E., Brinster, R. L., Palmiter, R. D., Townes, T. M. 1987. Two 3' sequences direct adult erythroid-specific expression of human beta-globin genes in transgenic mice. *Proc. Natl. Acad. Sci. USA* 84: 7056-60
- Behringer, R. R., Ryan, T. M., Palmiter, R. D., Brinster, R. L., Townes, T. M. 1990. Human gamma-to beta-globin gene switching in transgenic mice. *Genes Dev.* 4: 380-89
- Behringer, R. R., Ryan, T. M., Reilly, M. P., Asakura, T., Palmiter, R. D., et al. 1989. Synthesis of functional human hemoglobin in transgenic mice. *Science* 245: 971-73
- Benezra, R., Cantor, C. R., Axel, R. 1986. Nucleosomes are phased along the mouse beta-major globin gene in erythroid and nonerythroid cells. *Cell* 44: 697-704
- Bennett, J. P., Brotherton, T. W., Ginder, G. D. 1989. Nuclear matrix attachment regions of the chicken beta-type globin gene cluster. In *Hemoglobin Switching Part B: Cellular and Molecular Mechanisms*, ed. G. Stamatoyannopoulos, A. W. Nienhuis. pp. 47-55. New York: Liss.
- Berg, J. M. 1989. DNA binding specificity of steroid receptors. *Cell* 57: 1065-68
- Berg, P. E., Williams, D. M., Qian, R.-L., Cohen, R. B., Cao, S.-X., et al. 1989. A common protein binds to two silencers 5' to the human beta-globin gene. *Nucleic Acids Res.* 17: 8833-51
- Bodine, D. M., Ley, T. J. 1987. An enhancer element lies 3' to the human γ -globin gene. *EMBO J.* 6: 2997-3004
- Bohmann, D., Bos, T. J., Admon, A., Nishimura, T., Vogt, P. K., Tjian, R. 1987. Human proto-oncogene *c-jun* encodes a DNA binding protein with structural and functional properties of transcription factor AP-1. *Science* 238: 1386-92
- Brady, H. J. M., Sowden, J. C., Edwards, M., Lowe, N., Butterworth, P. H. W. 1989. Multiple GF-1 binding sites flank the erythroid specific transcription unit of the human carbonic anhydrase I gene. *FEBS Lett.* 257: 451-56
- Briggs, M. R., Kadonaga, J. T., Bell, S. P., Tjian, R. 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* 234: 47-52
- Broders, F., Zahraoui, A., Scherrer, K. 1990. The chicken alpha-globin gene domain is transcribed into a 17-kilobase polycistronic RNA. *Proc. Natl. Acad. Sci. USA* 87: 503-7
- Bunn, H. F., Forget, B. G. 1986. *Hemoglobin: Molecular, Genetic and Clinical Aspects*. Philadelphia: Saunders. 690 pp.
- Buratowski, S., Hahn, S., Sharp, P. A., Guarente, L. 1988. Function of a yeast TATA element-binding protein in a mammalian transcription system. *Nature* 334: 37-42
- Cao, S.-X., Gutman, P. D., Dave, H. P. G., Schechter, A. N. 1989. Identification of a transcriptional silencer in the 5'-flanking region of the human epsilon-globin gene. *Proc. Natl. Acad. Sci. USA* 86: 5306-9
- Caplan, A., Kimura, T., Gould, H., Allan, J. 1987. Perturbation of chromatin structure in the region of the adult beta-globin gene in chicken erythrocyte chromatin. *J. Mol. Biol.* 193: 57-70
- Catala, F., deBoer, E., Habets, G., Grosveld, F. 1989. Nuclear protein factors and erythroid transcription of the human γ -globin gene. *Nucleic Acids Res.* 17: 3811-27
- Cavallini, B., Huet, J., Plassat, J.-L., Sentenac, A., Egly, J.-M., Chambon, P. 1988. A yeast activity can substitute for the HeLa cell TATA box factor. *Nature* 334: 77-80
- Chada, K., Magram, J., Costantini, F. 1986. An embryonic pattern of expression of a human fetal globin gene in transgenic mice. *Nature* 319: 685-89
- Charnay, P., Henry, L. 1986. Regulated expression of cloned human fetal γ -globin genes introduced into murine erythroleukemia cells. *Eur. J. Biochem.* 159: 475-78
- Charnay, P., Mellon, P., Maniatis, T. 1985. Linker scanning mutagenesis of the 5'-flanking region of the mouse beta-major-globin gene: sequence requirements for transcription in erythroid and nonerythroid cells. *Mol. Cell. Biol.* 5: 1498-1511
- Chodosh, L. A., Baldwin, A. S., Carthew, R. W., Sharp, P. A. 1988. Human CCAAT-binding proteins have heterologous subunits. *Cell* 55: 11-24
- Choi, O.-R., Engel, J. D. 1986. A 3' enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene. *Nature* 323: 731-34
- Choi, O.-R. B., Engel, J. D. 1988. Developmental regulation of beta-globin gene switching. *Cell* 55: 17-26
- Collins, F. S., Weissman, S. M. 1984. The molecular genetics of human hemoglobin. *Prog. Nucleic Acid Res. Mol. Biol.* 31: 315-462
- Collis, P., Antoniou, M., Grosveld, F. 1990. Definition of the minimal requirements within the human beta-globin gene and the dominant control region for high level expression. *EMBO J.* 9: 233-40
- Cowie, A., Myers, R. M. 1988. DNA sequences involved in transcriptional regu-

- lation of the mouse beta-globin promoter in murine erythroleukemia cells. *Mol. Cell. Biol.* 8: 3122-28
- Curtin, P. T., Liu, D., Liu, W., Chang, J. C., Kan, Y. W. 1989. Human beta-globin gene expression in transgenic mice is enhanced by a distant DNase I hypersensitive site. *Proc. Natl. Acad. Sci. USA* 86: 7082-86
- Davidson, I., Xiao, J. H., Rosales, R., Staub, A., Chambon, P. 1988. The HeLa cell protein TEF-1 binds specifically and cooperatively to two SV40 enhancer motifs of unrelated sequence. *Cell* 54: 931-42
- deBoer, E., Antoniou, M., Mignotte, V., Wall, L., Grosveld, F. 1988. The human beta-globin promoter; nuclear protein factors and erythroid specific induction of transcription. *EMBO J.* 7: 4203-12
- Dierks, P., van Ooyen, A., Cochran, M. D., Dobkin, C., Reiser, J., Weissmann, C. 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. *Cell* 32: 695-706
- Donovan-Peluso, M., Acuto, S., Swanson, M., Dobkin, C., Bank, A. 1987. Expression of human gamma-globin genes in human erythroleukemia (K562) cells. *J. Biol. Chem.* 262: 17051-57
- Dorn, A., Bollekens, J., Staub, A., Benoist, C., Mathis, D. 1987. A multiplicity of CCAAT box-binding proteins. *Cell* 50: 863-72
- Driscoll, M. C., Dobkin, C. S., Alter, B. P. 1989. Gamma/delta/beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. *Proc. Natl. Acad. Sci. USA* 86: 7470-74
- Emerson, B. M., Lewis, C. D., Felsenfeld, G. 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult beta-globin gene: nature of the binding domain. *Cell* 41: 21-30
- Emerson, B. M., Nickol, J. M., Jackson, P. D., Felsenfeld, G. 1987. Analysis of the tissue-specific enhancer at the 3' end of the chicken adult beta-globin gene. *Proc. Natl. Acad. Sci. USA* 84: 4786-90
- Emerson, B. M., Nickol, J. M., Fong, T. C. 1989. Erythroid-specific activation and derepression of the chick beta-globin promoter in vitro. *Cell* 57: 1189-1200
- Enver, T., Brewer, A. C., Patient, R. K. 1988. Role for DNA replication in beta-globin gene activation. *Mol. Cell. Biol.* 8: 1301-8
- Enver, T., Ebens, A. J., Forrester, W. C., Stamatoyannopoulos, S. 1989. The human beta-globin locus activation region alters the developmental fate of a human fetal globin gene in transgenic mice. *Proc. Natl. Acad. Sci. USA* 86: 7033-37
- Enver, T., Raich, N., Ebens, A. J., Papayannopoulou, T., Costantini, F., Stamatoyannopoulos, G. 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* 344: 309-13
- Evans, R. M. 1988. The steroid and thyroid hormone receptor superfamily. *Science* 240: 889-95
- Evans, T., DeChiara, T., Efstratiadis, A. 1988a. A promoter of the rat insulin-like growth factor II gene consists of minimal control elements. *J. Mol. Biol.* 199: 61-81
- Evans, T., Felsenfeld, G. 1989. The erythroid-specific transcription factor Eryf1: A new finger protein. *Cell* 58: 877-85
- Evans, T., Reitman, M., Felsenfeld, G. 1988b. An erythrocyte-specific DNA-binding factor recognizes a regulatory sequence common to all chicken globin genes. *Proc. Natl. Acad. Sci. USA* 85: 5976-80
- Fisher, E. A., Felsenfeld, G. 1986. A comparison of the folding of beta-globin and ovalbumin gene-containing chromatin isolated from chicken oviduct and erythrocytes. *Biochem.* 25: 8010-16
- Forrester, W. C., Thompson, C., Elder, J. T., Groudine, M. 1986. A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proc. Natl. Acad. Sci. USA* 83: 1359-63
- Forrester, W. C., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G., Groudine, M. 1987. Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res.* 15: 10159-77
- Forrester, W. C., Novak, U., Gelinas, R., Groudine, M. 1989. Molecular analysis of the human beta-globin locus activation region. *Proc. Natl. Acad. Sci. USA* 86: 5439-43
- Gallarda, J. L., Foley, K. P., Yang, Z., Engel, J. D. 1989. The beta-globin stage selector element factor is erythroid-specific promoter/enhancer binding protein NF-E4. *Genes Dev.* 3: 1845-59
- Galson, D. L., Housman, D. E. 1988. Detection of two tissue-specific DNA-binding proteins with affinity for sites in the mouse beta-globin intervening sequence 2. *Mol. Cell. Biol.* 8: 381-92
- Gianni, A. M., Bregni, M., Cappellini, M. D., Fiorelli, G., Taramelli, R., et al. 1983. A gene controlling fetal hemoglobin expression in adults is not linked to the non-alpha-globin cluster. *EMBO J.* 2: 921-25
- Giglioli, B., Comi, P., Ronchi, A., Man-

- tovani, R., Ottolenghi, S. 1989. The same nuclear proteins bind the proximal CACCC box of the human beta-globin promoter and a similar sequence in the enhancer. *Biochem. Biophys. Res. Comm.* 164: 149-55
- Goodman, M., Czelusniak, J., Koop, B. F., Tagle, D. A., Slightom, J. L. 1987. Globins: A case study in molecular phylogeny. *Cold Spring Harbor Symp. Quant. Biol.* 52: 875-90
- Graves, B. J., Johnson, P. F., McKnight, S. L. 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. *Cell* 44: 565-76
- Gross, D. S., Garrard, W. T. 1988. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* 57: 159-97
- Grosveld, F., van Assendelft, G. B., Greaves, D. R., Kollias, G. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51: 975-85
- Groudine, M., Weintraub, H. 1980. Activation of cellular genes by avian RNA tumor viruses. *Proc. Natl. Acad. Sci. USA* 77: 5351-54
- Gumucio, D. L., Lockwood, W. K., Weber, J. L., Saulino, A. M., Delgrosso, K., et al. 1990. The -175 T → C mutation increases promoter strength in erythroid cells; Correlation with evolutionary conservation of binding sites for two trans-acting factors. *Blood* 75: 756-61
- Gumucio, D. L., Rood, K. L., Gray, T. A., Riordan, M. F., Sartor, C. I., Collins, F. S. 1988. Nuclear proteins that bind the human gamma-globin gene promoter: alterations in binding produced by point mutations associated with hereditary persistence of fetal hemoglobin. *Mol. Cell. Biol.* 8: 5310-22
- Hanscombe, O., Vidal, M., Kaeda, J., Luzatto, L., Greaves, D. R., Grosveld, F. 1989. High-level, erythroid-specific expression of the human alpha-globin gene in transgenic mice and the production of human hemoglobin in murine erythrocytes. *Genes Dev.* 3: 1572-81
- Hesse, J. E., Nickol, J. M., Lieber, M. R., Felsenfeld, G. 1986. Regulated gene expression in transfected primary chicken erythrocytes. *Proc. Natl. Acad. Sci. USA* 83: 4312-16
- Humphries, R. K., Ley, T., Turner, P., Moulton, A. D., Nienhuis, A. W. 1982. Differences in human alpha-, beta- and delta-globin gene expression in monkey kidney cells. *Cell* 30: 173-83
- Humphries, S., Windass, J., Williamson, R. 1976. Mouse globin gene expression in erythroid and non-erythroid tissues. *Cell* 7: 267-77
- Ingram, V. M. 1972. Embryonic red cell formation. *Nature* 239: 338-39
- Jackson, P. D., Evans, T., Nickol, J. M., Felsenfeld, G. 1989. Developmental modulation of protein binding to beta-globin gene regulatory sites within chicken erythrocyte nuclei. *Genes Dev.* 3: 1860-73
- Jarman, A. P., Higgs, D. R. 1988. Nuclear scaffold attachment sites in the human globin gene complexes. *EMBO J.* 7: 3337-44
- Jeffreys, A. J., Wilson, V., Thein, S. L., Weatherall, D. J., Ponder, B. A. J. 1986. DNA "fingerprints" and segregation analysis of multiple markers in human pedigrees. *Am. J. Hum. Genet.* 39: 11-24
- Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T. J., Tjian, R. 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* 48: 79-89
- Kadonaga, J. T., Jones, K. A., Tjian, R. 1986. Promoter-specific activation of RNA polymerase II transcription by Spl. *Trends Biochem.* 11: 20-23
- Kemper, B., Jackson, P. D., Felsenfeld, G. 1987. Protein-binding sites within the 5' DNase I-hypersensitive region of the chicken alpha^H-globin gene. *Mol. Cell. Biol.* 7: 2059-69
- Kimura, T., Mills, F. C., Allan, J., Gould, H. 1983. Selective unfolding of erythroid chromatin in the region of the active beta-globin gene. *Nature* 306: 709-12
- Kioussis, D., Vanin, E., de Lange, T., Flavell, R. A., Grosveld, F. G. 1983. Beta-globin gene inactivation by DNA translocation in gamma, beta thalassemia. *Nature* 306: 662-66
- Knezetic, J. A., Felsenfeld, G. 1989. Identification and characterization of a chicken alpha-globin enhancer. *Mol. Cell. Biol.* 9: 893-901
- Kollias, G., Wrighton, N., Hurst, J., Grosveld, F. 1986. Regulated expression of human alpha-gamma-, beta-, and hybrid gamma/beta-globin genes in transgenic mice: manipulation of the developmental expression patterns. *Cell* 46: 89-94
- Kollias, G., Hurst, J., deBoer, E., Grosveld, F. 1987. The human beta-globin gene contains a downstream developmental specific enhancer. *Nucleic Acids Res.* 15: 5739-47
- Kretsovali, A., Muller, M. M., Weber, F., Marcaud, L., Farache, G., et al. 1987. A transcriptional enhancer located between adult beta-globin and embryonic epsilon-globin genes in chicken and duck. *Gene* 58: 167-75
- Lawson, G. M., Knoll, B. J., March, C. J., Woo, S. L. C., Tsai, M., 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: 1501-7
- Lee, W., Mitchell, P., Tjian, R. 1987. Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. *Cell* 49: 741-52
- Lewis, C. D., Clark, S. P., Felsenfeld, G., Gould, H. 1988. An erythrocyte-specific protein that binds to the poly (dG) region of the chicken beta-globin gene promoter. *Genes Dev.* 2: 863-73
- Ley, T. J., Nienhuis, A. W. 1983. A weak upstream promoter gives rise to long human beta-globin RNA molecules. *Biochem. Biophys. Res. Comm.* 112: 1041-48
- Lloyd, J. A., Lee, R. F., Lingrel, J. B. 1989. Mutations in two regions upstream of the γ -globin gene canonical promoter affect gene expression. *Nucleic Acids Res.* 17: 4339-52
- Lois, R., Freeman, L., Villeponteau, B., Martinson, H. G. 1990. Active beta-globin gene transcription occurs in methylated, DNase I-resistant chromatin of non-erythroid chicken cells. *Mol. Cell. Biol.* 10: 16-27
- Magram, J., Niederreither, K., Costantini, F. 1989. Beta-globin enhancers target expression of a heterologous gene to erythroid tissues of transgenic mice. *Mol. Cell. Biol.* 9: 4581-84
- Mantovani, R., Margaretti, N., Nicolis, S., Giglioni, B., Comi, P., et al. 1988. An erythroid specific nuclear factor binding to the proximal CACCC box of the beta-globin gene promoter. *Nucleic Acids Res.* 16: 4299-313
- Mantovani, R., Superti-Furga, G., Gilman, J., Ottolenghi, S. 1989. The deletion of the distal CCAAT box region of the γ -globin gene in black HPFH abolishes the binding of the erythroid specific protein NFE3 and of the CCAAT displacement protein. *Nucleic Acids Res.* 17: 6681-91
- Martin, D. I. K., Tsai, S.-F., Orkin, S. H. 1989. Increased gamma-globin expression in a nondeletion HPFH mediated by an erythroid-specific DNA-binding factor. *Nature* 338: 435-38
- Martin, D. I. K., Zon, L. I., Mutter, G., Orkin, S. H. 1990. Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages. *Nature* 344: 444-47
- McGhee, J. D., Wood, W. I., Dolan, M., Engel, J. D., Felsenfeld, G. 1981. A 200 base pair region at the 5' end of the chicken adult beta-globin gene is accessible to nuclease digestion. *Cell* 27: 45-55
- Mellon, P., Parker, V., Gluzman, Y., Maniatis, T. 1981. Identification of DNA sequences required for transcription of the human alpha-globin gene in a new SV40 host-vector system. *Cell* 27: 279-88
- Mignotte, V., Elequet, J. F., Raich, N., Romeo, P.-H. 1989a. *Cis*- and *trans*-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. *Proc. Natl. Acad. Sci. USA* 86: 6548-52
- Mignotte, V., Wall, L., deBoer, E., Grosveld, F., Romeo, P.-H. 1989b. Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res.* 17: 37-54
- Miyoshi, K., Kaneto, Y., Kawai, H., Ohchi, H., Niki, S., et al. 1988. X-linked dominant control of F-cells in normal adult life: Characterization of the Swiss type as hereditary persistence of fetal hemoglobin regulated dominantly by gene(s) on X chromosome. *Blood* 72: 1854-60
- Myers, R. M., Tilly, K., Maniatis, T. 1986. Fine structure genetic analysis of a beta-globin promoter. *Science* 232: 613-18
- Myers, R. M., Cowie, A., Stuve, L., Hartzog, G., Gaensler, K. 1989. Genetic and biochemical analysis of the mouse beta-major globin promoter. In *Hemoglobin Switching Part A: Transcriptional Regulation*, ed. G. Stamatoyannopoulos, A. W. Nienhuis, pp. 117-127. New York: Liss
- Nicholls, R. D., Fischel-Ghodsian, N., Higgs, D. R. 1987. Recombination at the human alpha-globin gene cluster: sequence features and topological constraints. *Cell* 49: 369-78
- Nickol, J. M., Felsenfeld, G. 1988. Bidirectional control of the chicken beta- and epsilon-globin genes by a shared enhancer. *Proc. Natl. Acad. Sci. USA* 85: 2548-52
- Nicolis, S., Ronchi, A., Margaretti, N., Mantovani, R., Giglioni, B., Ottolenghi, S. 1989. Increased erythroid-specific expression of a mutated HPFH gamma-globin promoter requires the erythroid factor NF-E1. *Nucleic Acids Res.* 17: 5509-16
- Nienhuis, A. W., Maniatis, T. 1987. Structure and expression of globin genes in erythroid cells. In *The Molecular Basis of Blood Diseases*, ed. G. Stamatoyannopoulos, A. W. Nienhuis, P. Leder, P. W. Majerus, pp. 28-65. Philadelphia: Saunders
- Nomiyama, H., Fromental, C., Xiao, J. H., Chambon, P. 1987. Cell-specific activity of the constituent elements of the simian virus 40 enhancer. *Proc. Natl. Acad. Sci. USA* 84: 7881-85
- Papayannopoulou, T., Brice, M., Stamatoyannopoulos, G. 1986. Analysis of human hemoglobin switching in MEL \times

- human fetal erythroid cell hybrids. *Cell* 46: 469-76
- Papayannopoulou, T., Enver, T., Takegawa, S., Anagnou, N. P., Stamatoyannopoulos, G. 1988. Activation of developmentally mutated human globin genes by cell fusion. *Science* 242: 1056-58
- Parker, C. S., Topol, J. 1984. A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an hsp 70 gene. *Cell* 37: 273-83
- Perkins, N. D., Nicolas, R. H., Plumb, M. A., Goodwin, G. H. 1989. The purification of an erythroid protein which binds to enhancer and promoter elements of haemoglobin genes. *Nucleic Acids Res.* 17: 1299-1314
- Plumb, M., Frampton, J., Wainwright, H., Walker, M., Macleod, K., et al. 1989. GATAAG; a *cis*-control region binding an erythroid-specific nuclear factor with a role in globin and non-globin gene expression. *Nucleic Acids Res.* 17: 73-91
- Plumb, M. A., Lobanenko, V. V., Nicolas, R. H., Wright, C. A., Zavou, S., Goodwin, G. H. 1986. Characterisation of chicken erythroid nuclear proteins which bind to the nuclease hypersensitive regions upstream of the β^A - and β^H -globin genes. *Nucleic Acids Res.* 14: 7675-93
- Pribyl, T. M., Martinson, H. G. 1988. Transcription termination at the chicken β^H -globin gene. *Mol. Cell. Biol.* 8: 5369-77
- Reitman, M., Felsenfeld, G. 1988. Mutational analysis of the chicken β -globin enhancer reveals two positive-acting domains. *Proc. Natl. Acad. Sci. USA* 85: 6267-71
- Reitman, M., Felsenfeld, G. 1990. Developmental regulation of topoisomerase II sites and DNase I-hypersensitive sites in the chicken β -globin locus. *Mol. Cell. Biol.* 10: 2774-86
- Rixon, M. W., Gelinas, R. E. 1988. A fetal globin gene mutation in γ non-deletion hereditary persistence of fetal hemoglobin increases promoter strength in a nonerythroid cell. *Mol. Cell. Biol.* 8: 713-21
- Romeo, P.-H., Prandini, M.-H., Joulin, V., Mignotte, V., Prenant, M., et al. 1990. Megakaryocytic and erythrocytic lineages share specific transcription factors. *Nature* 344: 447-49
- Ronchi, A., Nicolis, S., Santoro, C., Ottolenghi, S. 1989. Increased Sp1 binding mediates erythroid-specific overexpression of a mutated (HPFH) γ -globin promoter. *Nucleic Acids Res.* 17: 10231-41
- Rousseau, S., Renaud, J., Ruiz-Carrillo, A. 1989. Basal expression of the histone H5 gene is controlled by positive and negative *cis*-acting sequences. *Nucleic Acids Res.* 17: 7495-511
- Rupp, R. A. W., Nicolas, R. H., Borgmeyer, U., Lobanenko, V. V., Plumb, M. A., et al. 1988. TGGCA protein is present in erythroid nuclei and binds within the nuclease-hypersensitive sites 5' of the chicken β^H - and β^A -globin genes. *Eur. J. Biochem.* 177: 505-11
- Rupp, R. A. W., Sippel, A. E. 1987. Chicken liver TGGCA protein purified by preparative mobility shift electrophoresis (PMSE) shows a 36.8 to 29.8 kd micro-heterogeneity. *Nucleic Acids Res.* 15: 9707-26
- Rutherford, T., Nienhuis, A. W. 1987. Human globin gene promoter sequences are sufficient for specific expression of a hybrid gene transfected into tissue culture cells. *Mol. Cell. Biol.* 7: 398-402
- Ryan, T. M., Behringer, R. R., Townes, T. M., Palmiter, R. D., Brinster, R. L. 1989a. High-level erythroid expression of human α -globin genes in transgenic mice. *Proc. Natl. Acad. Sci. USA* 86: 37-41
- Ryan, T. M., Behringer, R. R., Martin, N. C., Townes, T. M., Palmiter, R. D., Brinster, R. L. 1989b. A single erythroid-specific DNase I super-hypersensitive site activates high levels of human β -globin gene expression in transgenic mice. *Genes Dev.* 3: 314-23
- Ryan, T. M., Townes, T. M., Reilly, M. P., Asakura, T., Palmiter, R. D., et al. 1990. Human sickle hemoglobin in transgenic mice. *Science* 247: 566-68
- Santoro, C., Mermod, N., Andrews, P. C., Tjian, R. 1988. A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. *Nature* 334: 218-24
- Sawadogo, M., Roeder, R. G. 1985. Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell* 43: 165-75
- Schule, R., Muller, M., Otsuka-Murakami, H., Renkawitz, R. 1988. Cooperativity of the glucocorticoid receptor and the CACCC-box binding factor. *Nature* 332: 87-90
- Stalder, W. E., Larsen, A., Engel, J. D., Dolan, M., Groudine, M., Weintraub, H. 1980. Tissue-specific DNA cleavage in the globin chromatin domain introduced by DNase I. *Cell* 20: 451-60
- Stamatoyannopoulos, G., Nienhuis, A. W. 1987. Hemoglobin Switching. In *The Molecular Basis of Blood Diseases*, ed. G. Stamatoyannopoulos, A. W. Nienhuis, P. Leder, P. W. Majerus, pp. 66-105. Philadelphia: Saunders
- Stoeckert, C. J., Metherall, J. E., Yama-

- kawa, M., Weissman, S. M., Forget, B. G. 1987. Expression of the affected γ gamma-globin gene associated with Greek non-deletion hereditary persistence of fetal hemoglobin. *Mol. Cell. Biol.* 7: 2999–3003
- Stuve, L. L., Myers, R. M. 1990. A directly repeated sequence in the beta-globin promoter regulates transcription in murine erythroleukemia cells. *Mol. Cell. Biol.* 10: 972–81
- Sun, Y. L., Xu, Y. Z., Bellard, M., Chambon, P. 1986. Digestion of the chicken beta-globin gene chromatin with micrococcal nuclease reveals the presence of an altered nucleosomal array characterized by an atypical ladder of DNA fragments. *EMBO J.* 5: 293–300
- Superti-Furga, G., Barberis, A., Schaffner, G., Busslinger, M. 1988. The -117 mutation in Greek HPFH affects the binding of three nuclear factors to the CCAAT region of the gamma-globin gene. *EMBO J.* 7: 3099–3107
- Sykes, K., Kaufman, R. 1990. A naturally occurring gamma globin gene mutation enhances SP1 binding activity. *Mol. Cell. Biol.* 10: 95–102
- Talbot, D., Collis, P., Antoniou, M., Vidal, M., Grosveld, F., Greaves, D. R. 1989. A dominant control region from the human beta-globin locus conferring integration site-independent gene expression. *Nature* 338: 352–55
- Trainor, C. D., Evans, T., Felsenfeld, G., Boguski, M. S. 1990. Structure and evolution of a human erythroid transcription factor. *Nature* 343: 92–96
- Trudel, M., Costantini, F. 1987. A 3' enhancer contributes to the stage-specific expression of the human beta-globin gene. *Genes Dev.* 1: 954–61
- Tsai, S.-F., Martin, D. I. K., Zon, L. I., D'Andrea, A. D., Wong, G. G., Orkin, S. H. 1989. Cloning of cDNA for the major DNA-binding protein of the erythroid lineage through expression in mammalian cells. *Nature* 339: 446–51
- Tuan, D., Abeliovich, A., Lee-Oldham, M., Lee, D. 1987. Identification of regulatory elements of human beta-like globin genes. In *Developmental Control of Globin Gene Expression*, ed. G. Stamatoyannopoulos, A. W. Nienhuis, pp. 211–20. New York: Liss
- Tuan, D., Solomon, W., Li, Q., London, I. M. 1985. The "beta-like-globin" gene domain in human erythroid cells. *Proc. Natl. Acad. Sci. USA* 82: 6384–88
- Tuan, D. Y. H., Solomon, W. B., London, I. M., Lee, D. P. 1989. An erythroid-specific, developmental-stage-independent enhancer far upstream of the human "beta-like globin" genes. *Proc. Natl. Acad. Sci. USA* 86: 2554–58
- Ulrich, M. J., Ley, T. J. 1990. Function of normal and mutated gamma-globin gene promoters in electroporated K562 erythroleukemia cells. *Blood* 75: 990–99
- van Assendelft, G. B., Hanscombe, O., Grosveld, F., Greaves, D. R. 1989. The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 56: 969–77
- Wall, L., deBoer, E., Grosveld, F. 1988. The human beta-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein. *Genes Dev.* 2: 1089–1100
- Weatherall, D. J., Clegg, J. B., Wood, W. G. 1976. A model for the persistence or reactivation of fetal haemoglobin production. *Lancet* 2: 660–63
- Weatherall, D. J., Clegg, J. B., Higgs, D. R., Wood, W. G. 1989. The Hemoglobinopathies. In *The Metabolic Basis of Inherited Disease*, ed. C. R. Scriver, A. L. Beaudet, W. S. Sly, D. Valle, pp. 2281–339. New York: McGraw-Hill, 6th ed.
- Whitelaw, E., Hogben, P., Hanscombe, O., Proudfoot, N. J. 1989. Transcriptional promiscuity of the human alpha-globin gene. *Mol. Cell. Biol.* 9: 241–51
- Wood, W. G. 1989. HbF Production in Adult Life. In *Hemoglobin Switching Part B: Cellular and Molecular Mechanisms*, ed. G. Stamatoyannopoulos, A. W. Nienhuis, pp. 251–67. New York: Liss
- Wrighton, N., Grosveld, F. 1988. A novel in vivo transcription assay demonstrates the presence of globin-inducing *trans*-acting factors in uninduced murine erythroleukemia cells. *Mol. Cell. Biol.* 8: 130–37
- Xiao, J.-H., Davidson, I., Macchi, M., Rosales, R., Vigneron, M., et al. 1987. In vitro binding of several cell-specific and ubiquitous nuclear proteins to the GT-1 motif of the SV40 enhancer. *Genes Dev.* 1: 794–807
- Yu, C.-Y., Chen, J., Lin, L.-I., Tam, M., Shen, C.-K. J. 1990. Cell type-specific protein-DNA interactions in the human zeta-globin upstream promoter region: Displacement of Sp1 by the erythroid cell-specific factor NF-E1. *Mol. Cell. Biol.* 10: 282–94
- Zon, L. I., Tsai, S.-F., Burgess, S., Matsu-daira, P., Bruns, G. A. P., Orkin, S. H. 1990. The major human erythroid DNA-binding protein (GF-1): Primary sequence and localization of the gene to the X chromosome. *Proc. Natl. Acad. Sci. USA* 87: 668–72