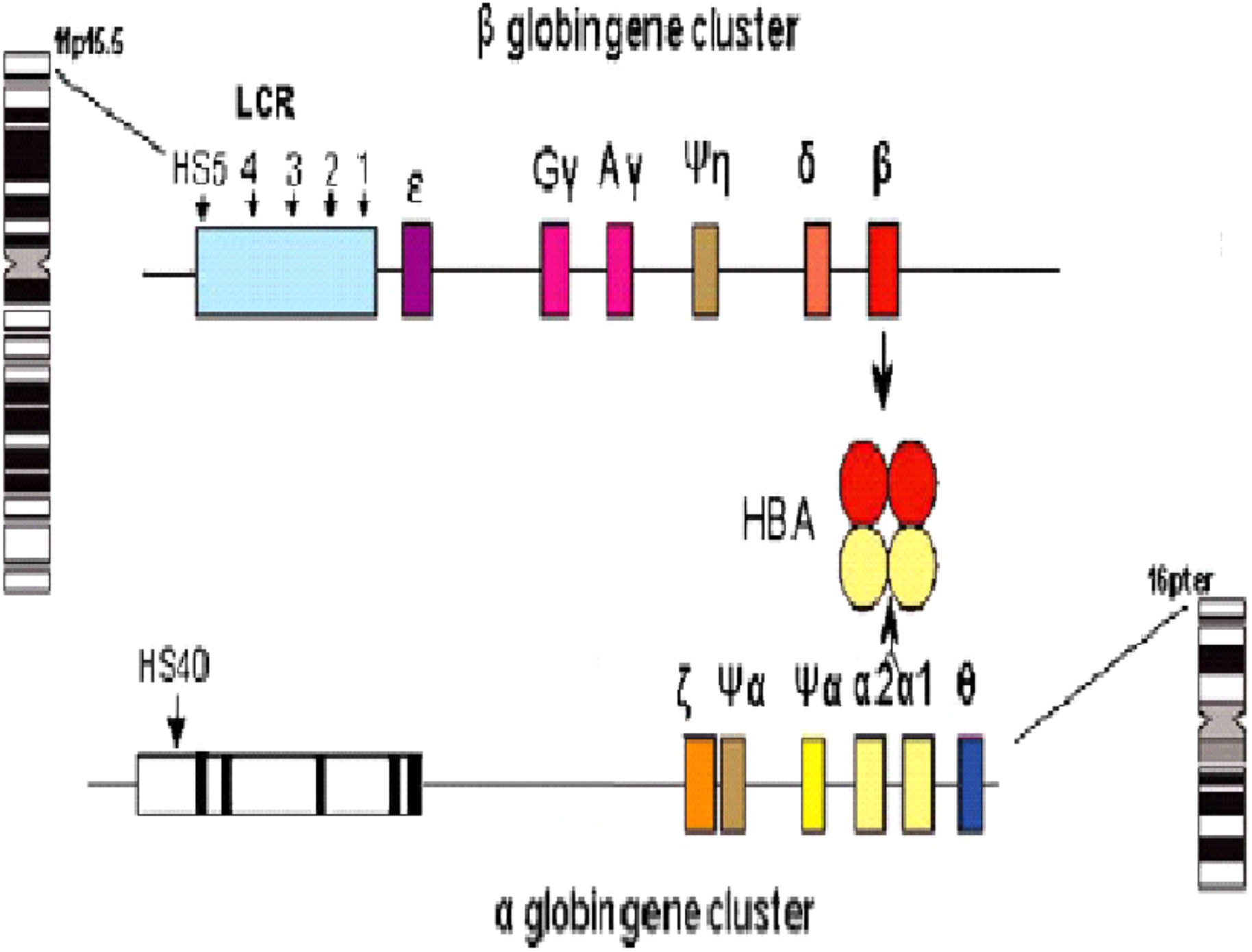
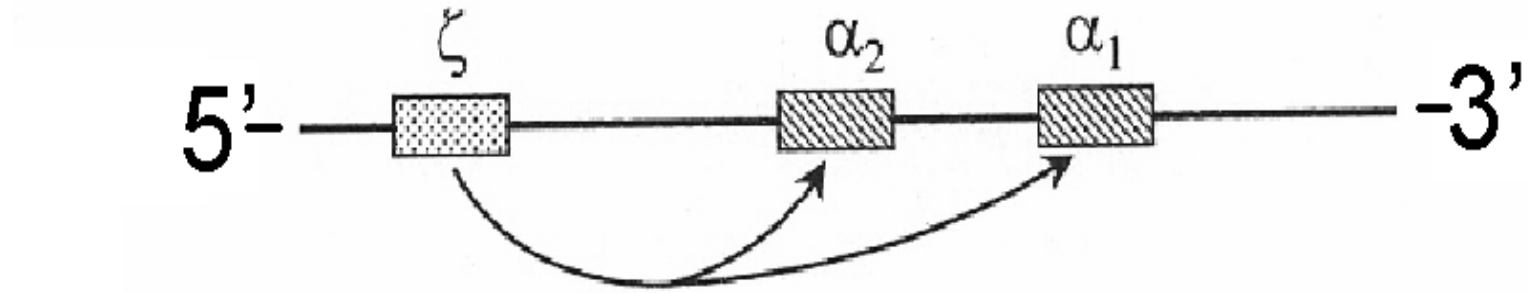
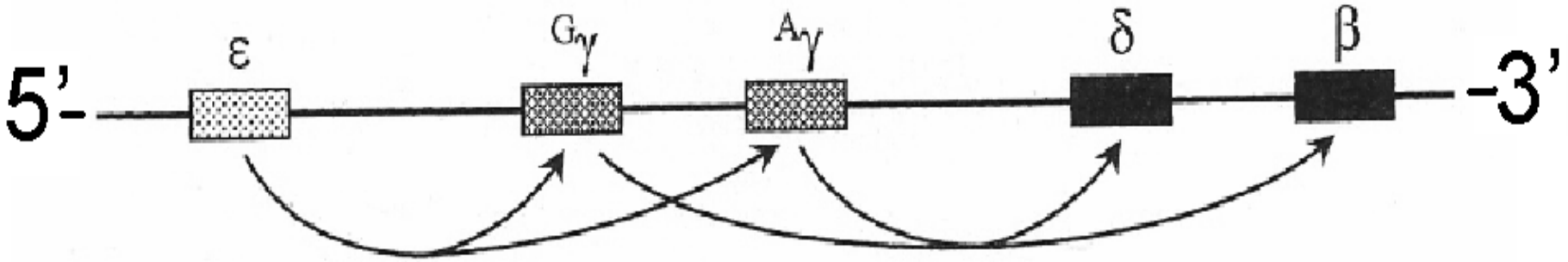
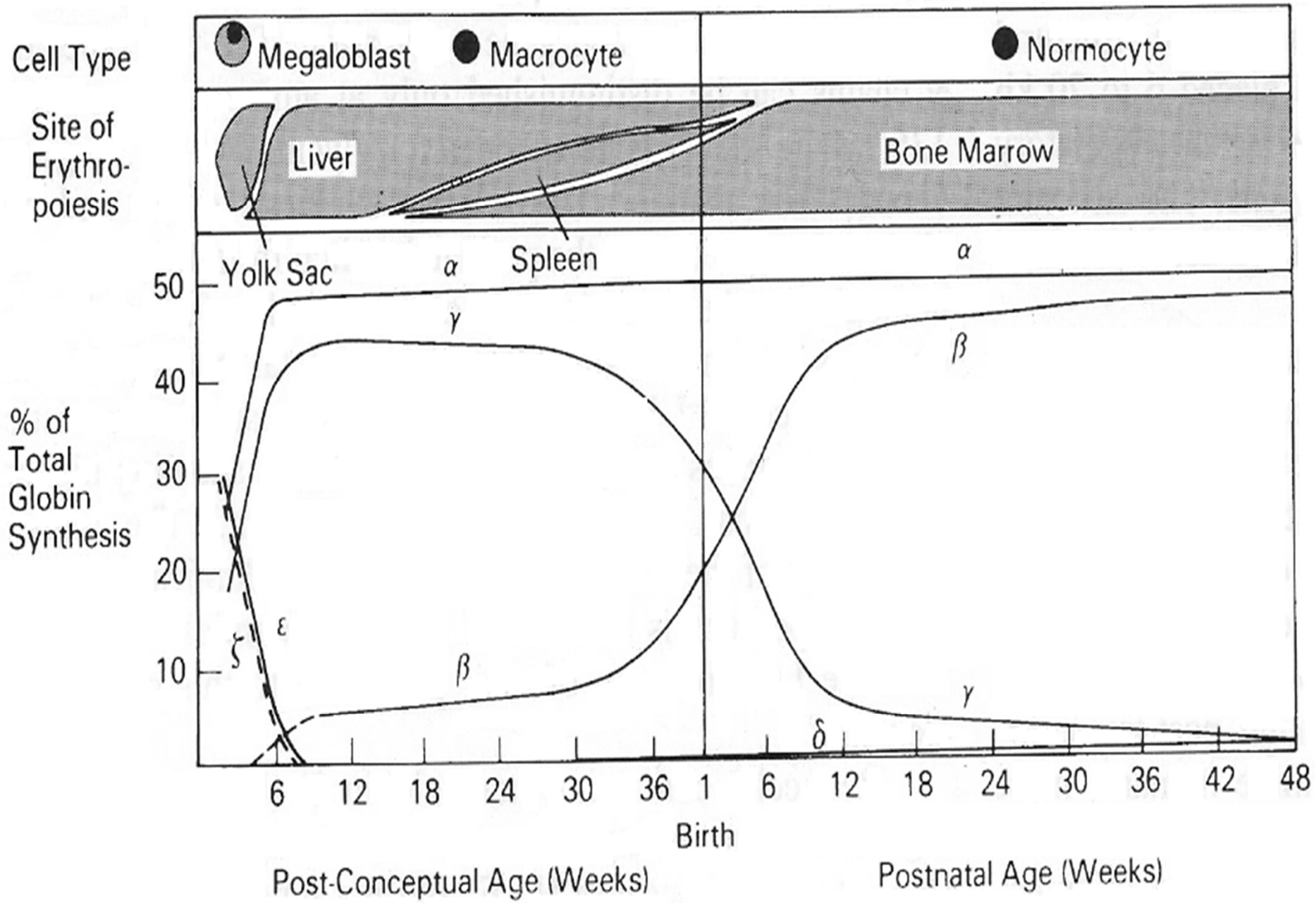
GLOBİN GEN REGÜLASYONU


β globin gene cluster
11p15.5
LCR
HS5 4 3 2 1
ε
Gγ Aγ
Ψη
δ
β
HBA
16pter
HS40
ζ Ψα Ψα α2α1 θ
α globin gene cluster





Embryonic	Fetal	Adult
	HbF	HbA
$\zeta_2\varepsilon_2$	$\alpha_2\gamma_2$	$\alpha_2\beta_2$
		HbA_2
		$\alpha_2\delta_2$


Cell Type
Megaloblast
Macrocyte
Normocyte
Site of Erythro-poiesis
Liver
Bone Marrow
Yolk Sac
Spleen
% of Total Globin Synthesis
50
40
30
20
10
α
γ
ε
ζ
β
α
β
γ
δ
6
12
18
24
30
36
1
6
12
18
24
30
36
42
48
Birth
Post-Conceptual Age (Weeks)
Postnatal Age (Weeks)

GLOBİN GENLERİN REGÜLASYONU

- Her bir globin genin dokuya ve gelişime spesifik ekspressiyonu regülatör dizilimdeki transkripsiyon faktörlerinin etkisi ile sağlanmaktadır.

- Globin gen regülasyonunu sağlayan faktörler ya doku ile sınırlıdırlar yada ekspressiyona bağlı olarak ubiquitos dağılım göstermektedir.

ERİTROİD TRANSKRİPSİYON FAKTÖRLERİ

GATA-1

- Transkripsiyon faktörü
- Organların uyumlu çalışmasında merkezi rol oynamakta,
- DNA'ya-bağlanan protein ailesindendir
- GATA diğer hematopoetik düzenlerde de bulunmaktadır

GATA protein ailesi

- İki zinc-finger ile karakterize
- DNA heliksinin major oluğunda nüleotidlere spesifik etkileşim göstermekte
- Eritroid hücre proliferasyonu ile hücre farklılaşması arasında kritik rolü bulunmakta

GATA-1

- Fetal eritroid farklılaşmasında gerekli
- GATA-1 eksikliğinde primitif ve olgun hücrelerin proeritroblast evrede matürasyonda eksiklik sonucunda hücreler apoptozise gitmektedirler
- Aşırı expressiyonunda proeritroblast hücrelerin proliferasyonunu stimüle etmekte ve farklılaşmasını inhibe etmekte

Diğer GATA proteinleri

- Diğer 5 GATA geni (GATA2-6) zinc finger protein ailesinde
- GATA-2 ve 3 hematopoetik sistemde ve diğer dokularda expresse edilmekte
- GATA-2'nin inaktivasyonu letal embriyonik fenotipi göstermekte
- GATA-2 ve 3 globin gen ekspressiyonunda direk olarak rol almamakta

FOG PROTEİNİ

- 9 zinc-finger içermekte
- Yolk sac kan adacıklarında ve fetal KC eritroid prekörsörlerinde GATA-1 ile yüksek düzeyde koekspresse olmakta
- GATA-1 ve FOG etkileşimi terminal eritroid farklılaşmasında önemlidir
- Eksikliğinde embriyo ağır anemiden letal olmakta

NUCLEAR FAKTÖR-ERİTROİD2 (NF-E2)

- Eritroid, mast ve megakaryositleri içeren bir çok hücrede ekspresse olmaktadır.

- β globin lokusunun HS-2 bölgesine (T/C)GCTGA(C/G)TCA(T/C) bağlanmakta

EKLF
(Eritroid Kruppel-Like Factor)

- Ekspressiyonu primitif ve olgunlaşmış eritroid hücreleri ile sınırlıdır
- Bilinen tek hedef geni β globin geni
- Yokluğunda β globin gen ekspressiyonu olmadığından fetal evrede letal olmakta
- Aşırı ekspressiyonunda sirkülasyondaki platelet sayısı azalmakta
  - Megakaryosit ve eritroid dengesinin saptanmasında rolü olabilir

EKLF-2

- Yetişkin β globin genin promotöründe CACC motifine bağlanmakta ve transkripsiyonel aktivatör olarak işlev görmekte
- HS3'te G'ce zengin dizilim ile etkileşime girerek LCR'nin aktivasyonunda rol alıyor olabilir

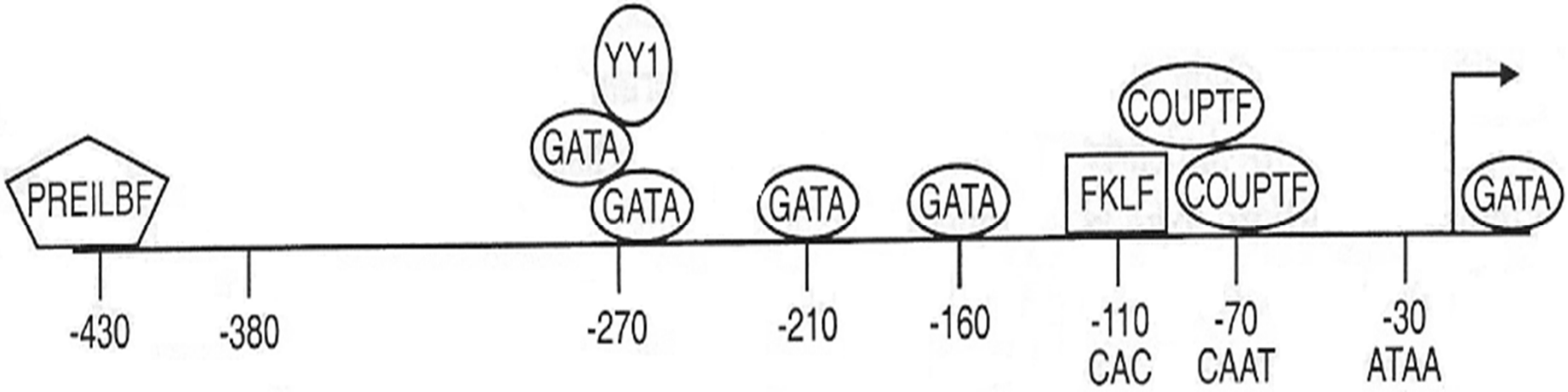
Diğer DNA'ya Bağlanan Proteinler

- Embriyonik ve fetal genlerin silencing de suppressor olarak rol oynayan faktörler
  - Ör; ubiquitous olarak ekspresse olan YY1 embriyonik ε globin genin silencer elementi
  - NF-E3 γ globin genin repsressor kompleksi
- DNA'ya direk olarak bağlanan faktörlere ek olarak birçok kofaktör eritroid spesifiteliğini sağlamakta

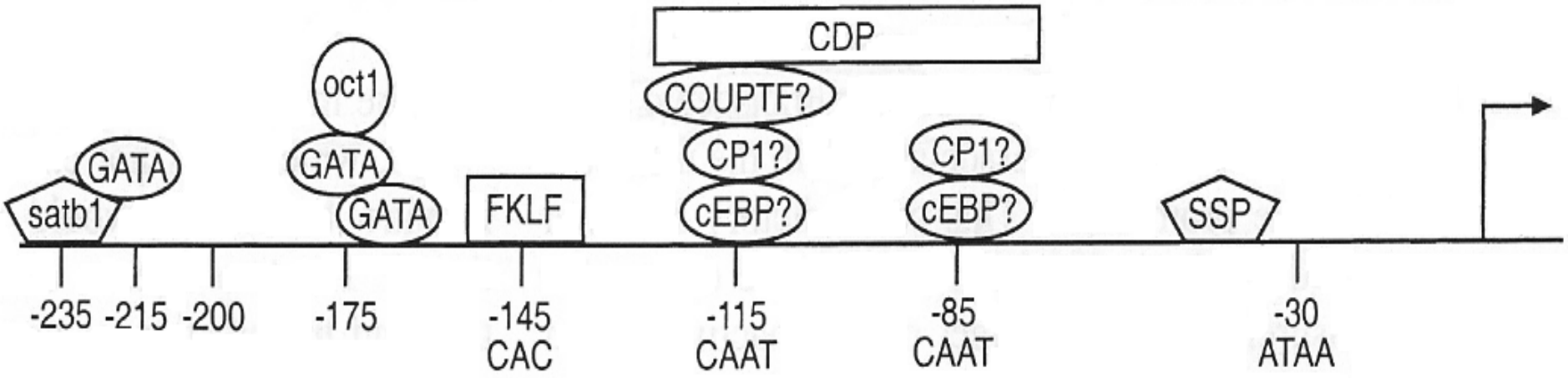
GLOBİN GENLERİN REGÜLATÖR ELEMENTLERİ

- Her bir globin genin birçok promotör, enhancer veya silencer regülatör elementi var
- Gen ekspressiyonun yüksek düzeyde yapılabilmesi için
  - β globin genin LCR
  - α lokusundaki HS-40 ile etkileşimde bulunmakta?

ε globin geni



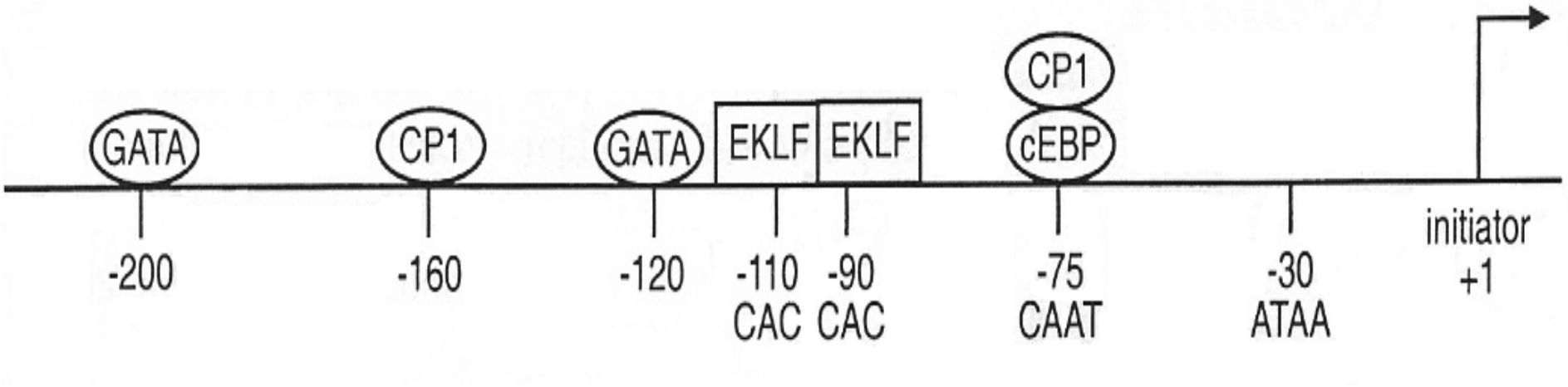
γ globin genleri



δ globin geni

- Normal olarak çok zayıf promotörü olduğundan çok düşük düzeyde ekspresse olmakta
- Çok düşük düzeyde ekspresse olmasının nedeni fonksiyonel CACCC kutusunun olmaması
- EKLF faktörü bağlanamamakta
- CACCC kutusunun insersiyonu ekspressiyonu 10 kat artırmaktadır

β globin geni



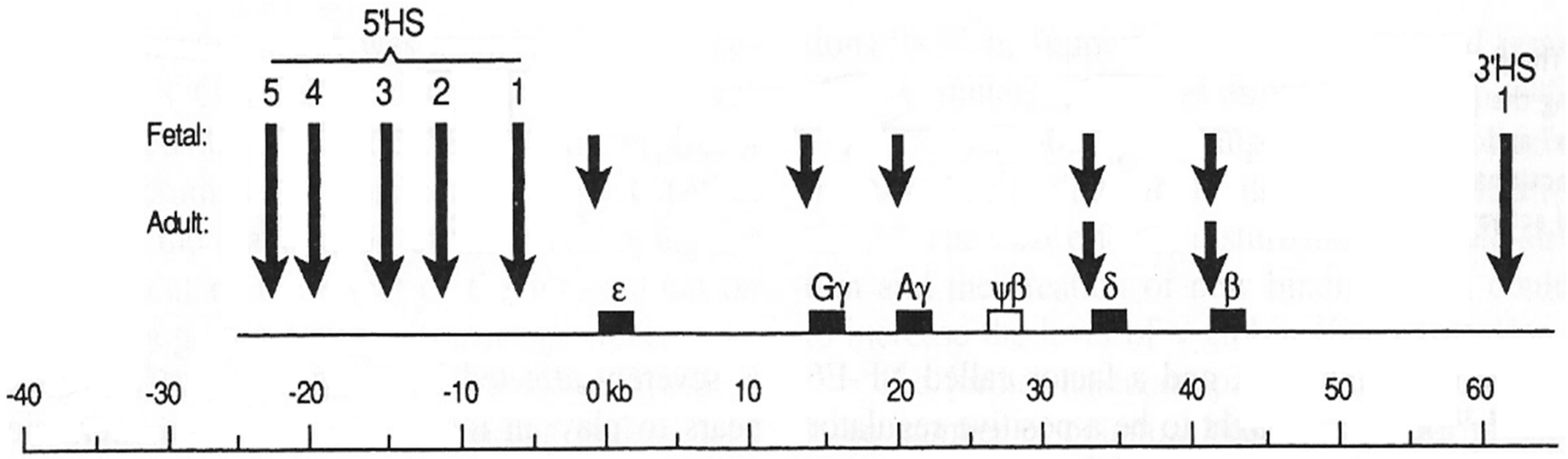
ζ globin geni

- TATA kutusu -28
- CAAT kutusu -66 → CP2
- CACCC kutusu -95 → Sp1-benzeri proteinler bağlanmakta
- ζ mRNA'nın sentezi olgunlaşmış eritropoez hücrelerinde devam etmekte ama mRNA unstabil, hemen yıkılmaktadır

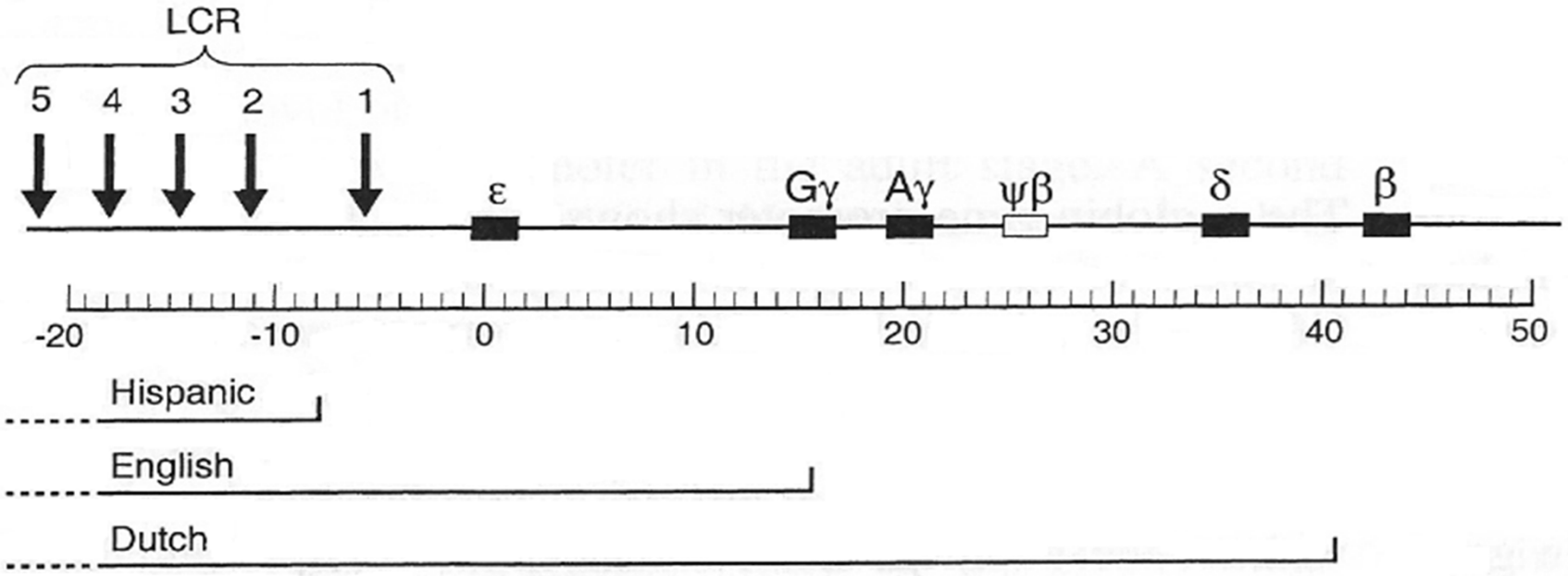
α globin geni

- α globin geni diğerlerinden farklı olarak bazı eksiklikleri vardır
  - CACCC kutusu yok ama gen içinde uzanan GC'ce zengin promotör alanı var
- β-promotörü gibi GATA-1 (-185) veya CP1(-90) ve CP2 gibi birkaç farklı protein α promotör bölgesine bağlanmaktadır

LCR


5'HS
5 4 3 2 1
3'HS
1
Fetal:
Adult:
ε
Gγ
Aγ
ψβ
δ
β
-40 -30 -20 -10 0 kb 10 20 30 40 50 60

$(\epsilon\gamma\delta\beta)^0$



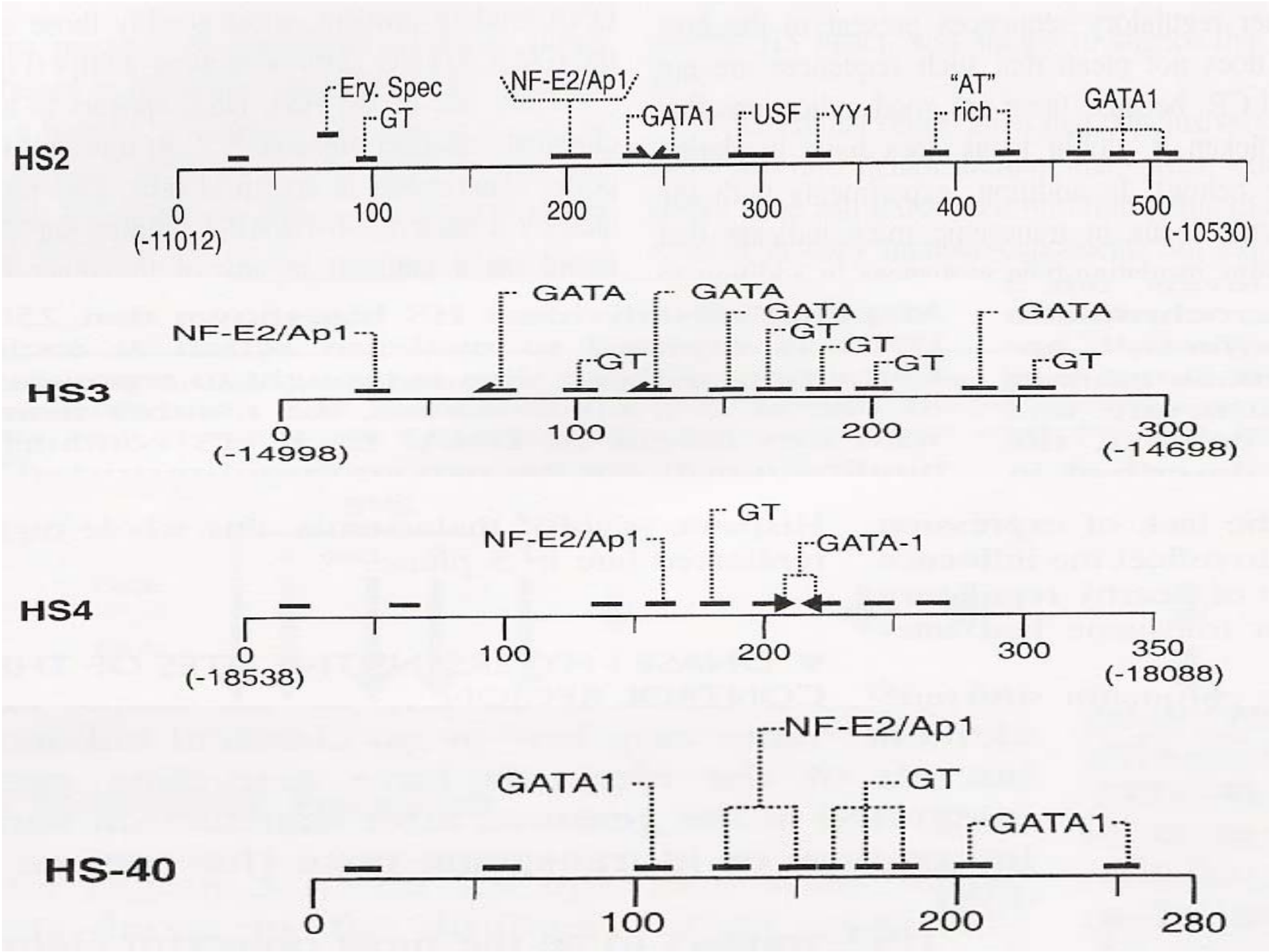
β LCR

- Erken DNA replikasyonunda eritroid-spesifikliği veya gelişime bağlı olarak ekspressiyonun sınırlandırılmasında

- Ama yüksek düzeyde spesifik ekspressiyonun sağlanmasında rolü olduğu açıktır.

β LCR

- Her HS bölgesi bir veya daha çok eritroid-sınırlı transkripsiyonel aktivatörler (GATA-1 ve NF-E2) ve diğer ubiquitous DNA-bağlayıcı proteinler için bağlanma bölgeleri vardır


HS2
Ery. Spec
GT
NF-E2/Ap1
GATA1
USF
YY1
"AT"
rich
GATA1
0
(-11012)
100
200
300
400
500
(-10530)
HS3
NF-E2/Ap1
GATA
GATA
GATA
GT
GT
GT
GT
GATA
GT
0
(-14998)
100
200
300
(-14698)
HS4
GT
NF-E2/Ap1
GATA-1
0
(-18538)
100
200
300
350
(-18088)
HS-40
NF-E2/Ap1
GATA1
GT
GATA1
0
100
200
280

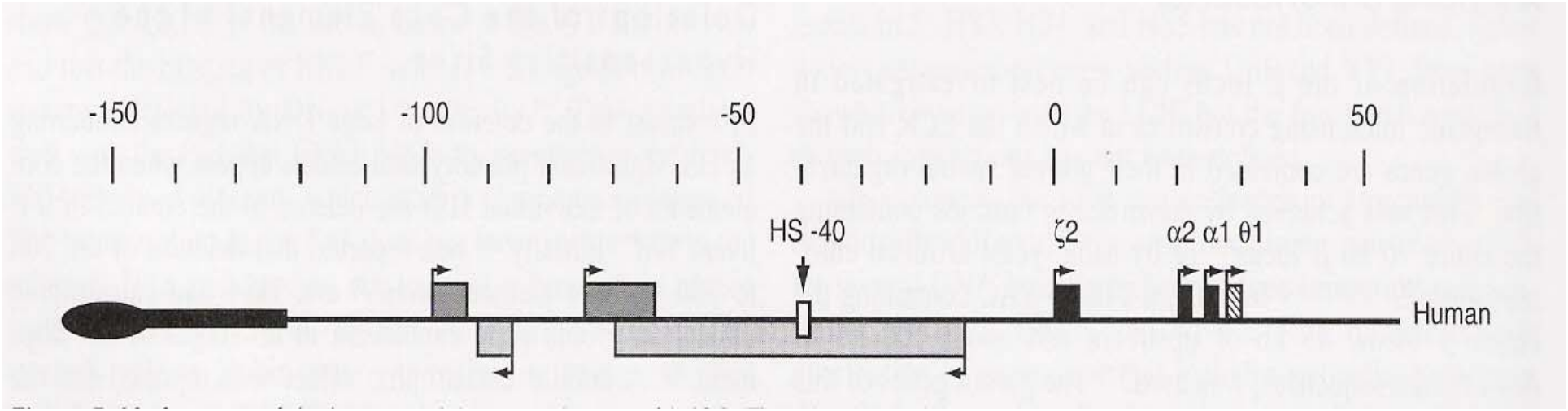
- Sonuç olarak
  - GATA proteinlerinin hem kromatin yapısında hem de gen aktivitesinde LCR fonksiyonu için kritik rolleri vardır
  - EKLF proteini kromatin yapısının modifikasyonunda
  - NF-E2 HS bölgelerin tümünde bulunmakta ve HS2'nin enhancer aktivitesi için gereklidir

- Bu üç transkripsiyon aktivatörü (GATA-1, EKLF ve NF-E2);
  - Eritroid hücre gelişimi
  - Hemoglobin sentezinin regülasyonunda direk olarak ilişkili olduğu bulunmuştur

- Diğer proteinler ubiquitous olarak ekspresse olduğu gibi LCR ile etkileşime girmektedir

α globin gen kümesinin regülatör elementleri

- GC bileşeni: %54
- Yüksek Alu tekrarlı sıralar: %26
- Bir çok CpG adası
- Kromatin açıklığı 300 kb
- Bir çok eritroidlere spesifik HS bölgesi
  - HS-40 bunlardan sadece biridir ve majör regülatör elementtir
  - HS-40 gen ekspressiyonunda eritroidlere spesifik güçlü enhancer olarak rol oynamaktadır


-150
-100
-50
0
50
HS -40
ζ2
α2 α1 θ1
Human

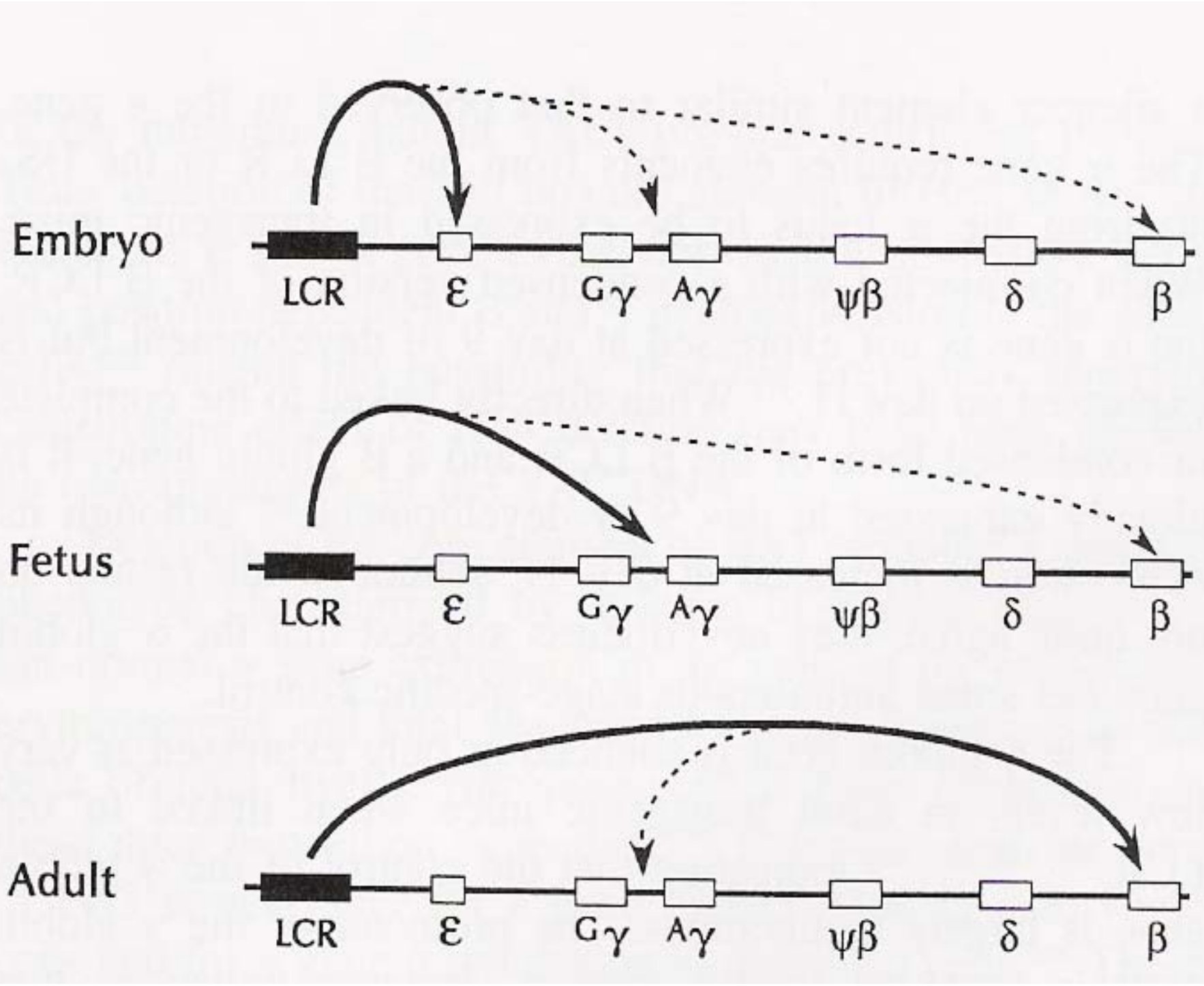
Globin gen silencing

- Gen özerk olarak sessiz olmaktadır ve ε geninden silencing elementin delesyonu sonucunda ekspressiyon olgun hücrelerde düşük düzeyde devam etmektedir
- Özerk silencing ζ geni içinde geçerlidir ama delesyon çalışmaları ε geninde gözlendiği gibi silencing elementin varlığı gösterilememiştir

Gen yarışması

- β globin genin sessiz olabilmesi için γ genin normal konfigrasyonda olması, β globin genin γ geni ile yarışarak regüle olduğunu göstermektedir

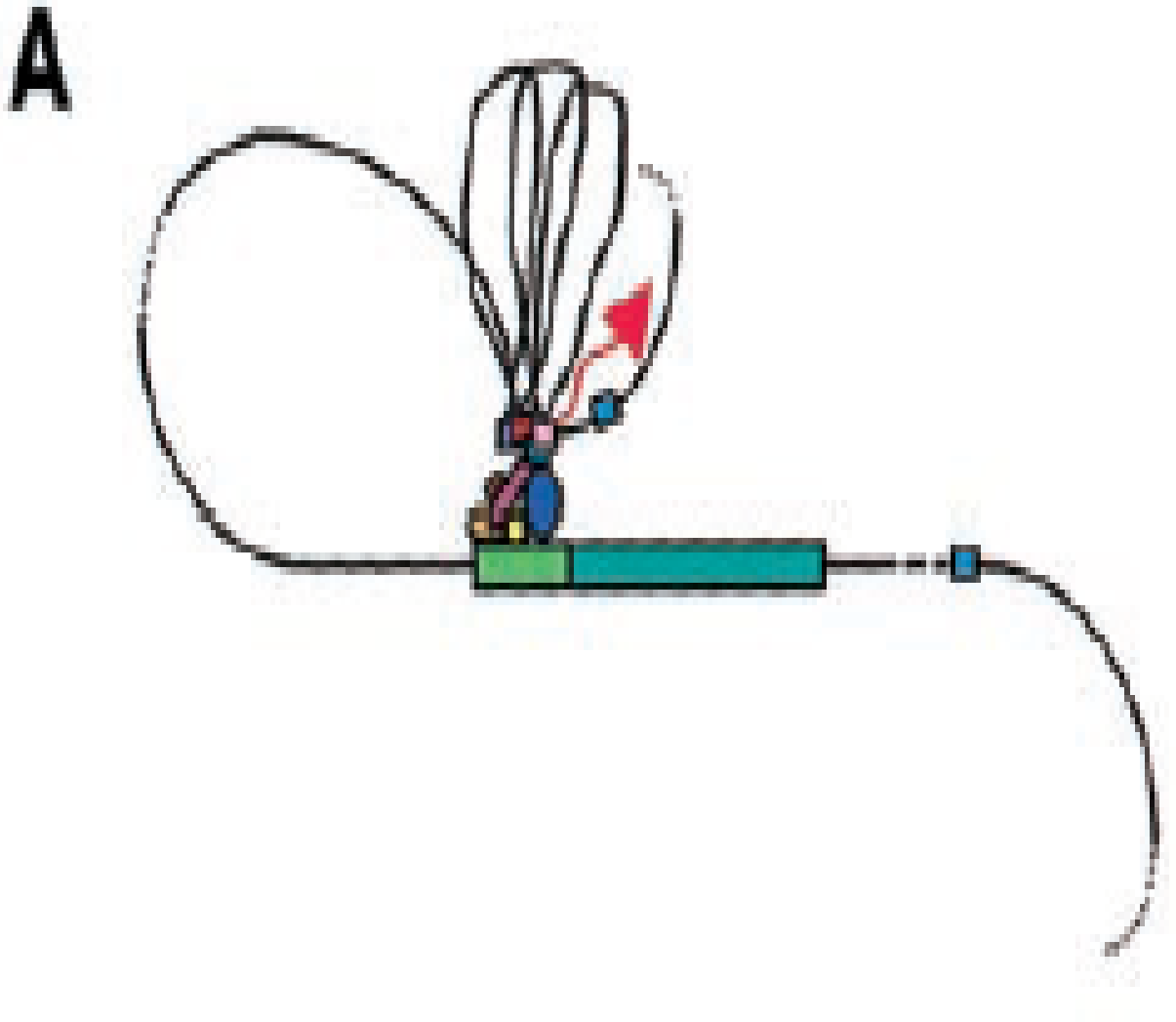
- Gen ekspresse olmak LCR için yarışmakta


Embryo
LCR
ε
Gγ
Aγ
ψβ
δ
β
Fetus
LCR
ε
Gγ
Aγ
ψβ
δ
β
Adult
LCR
ε
Gγ
Aγ
ψβ
δ
β

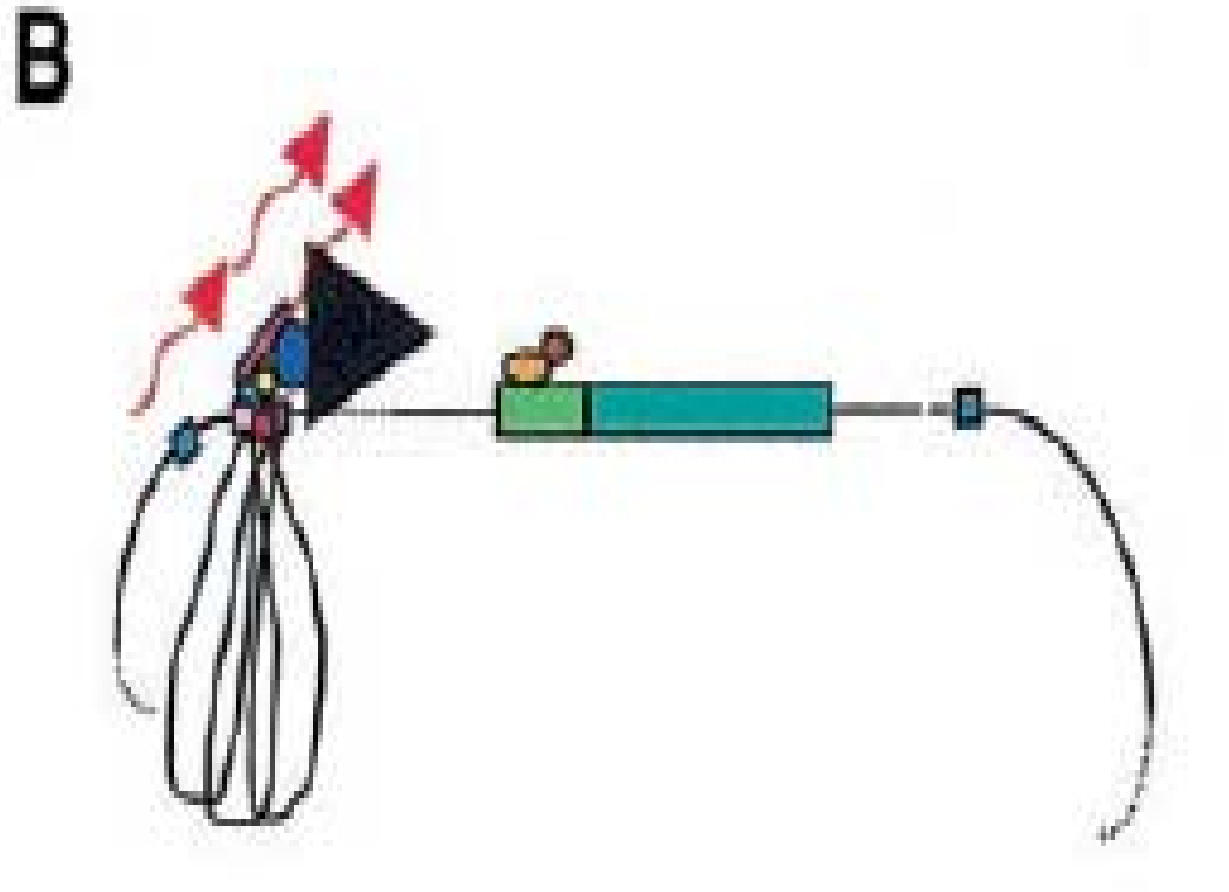
Regülasyon modelleri

- Regülasyonu açıklamak için;
  - Looping modeli
  - Tracking modeli
  - Facilitated tracking modeli
  - Linking model

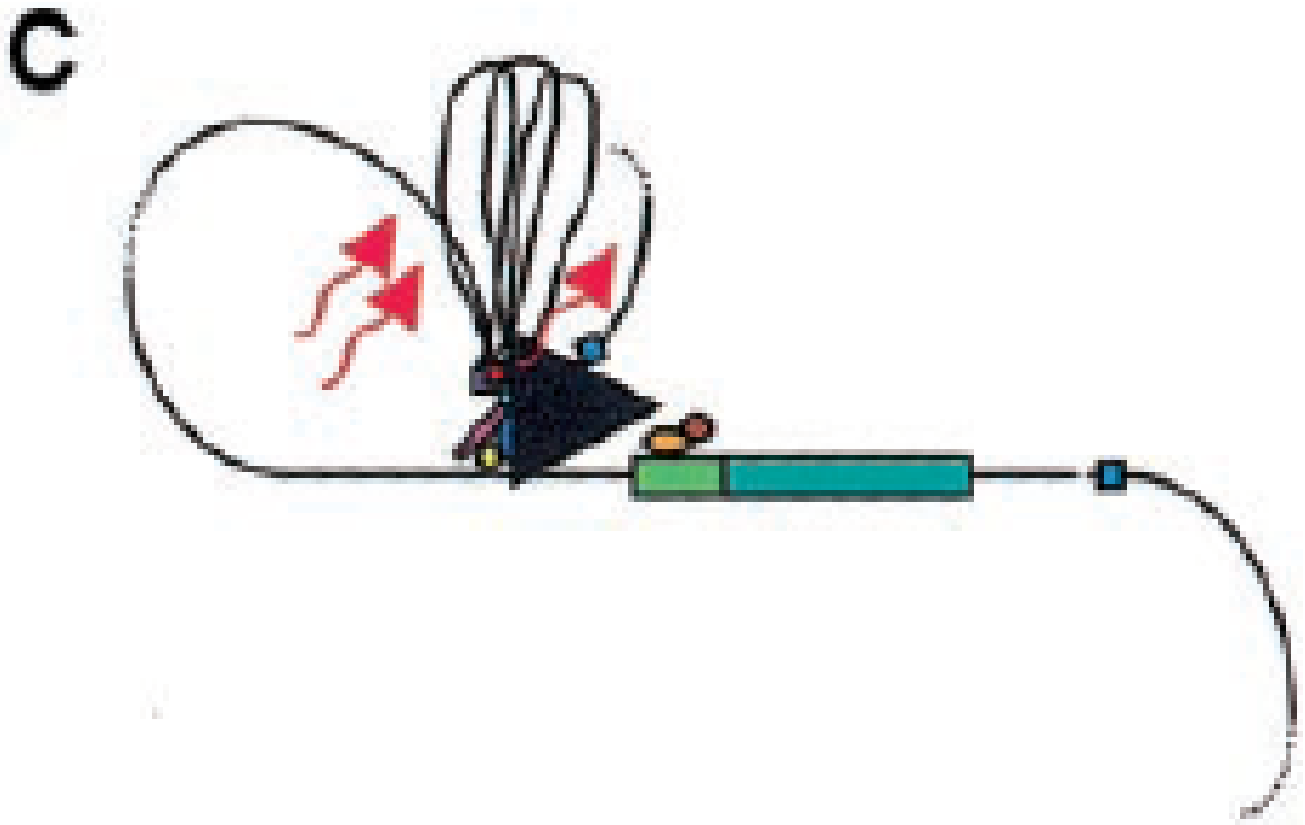
Looping model



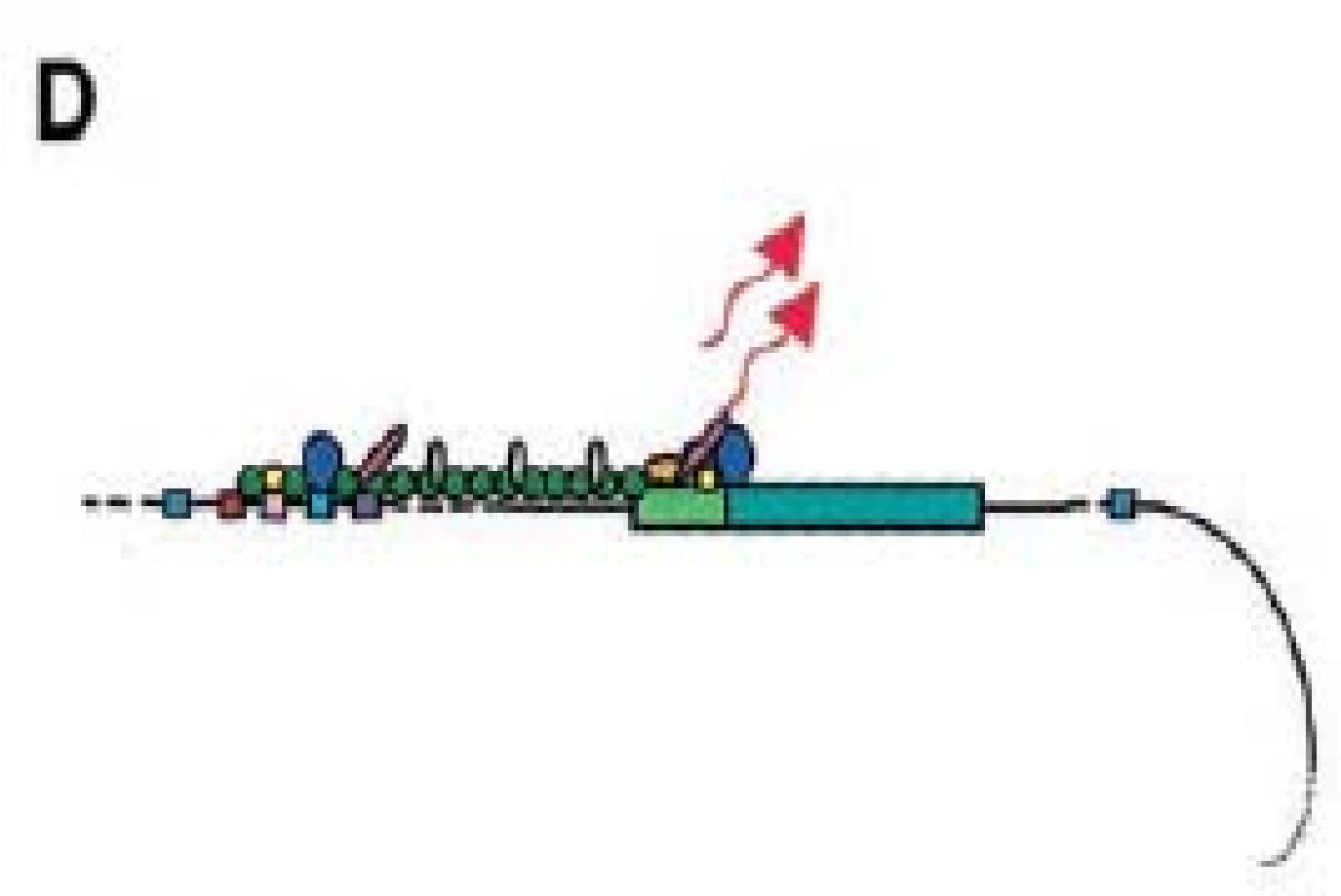
Tracking model



Facilitated tracking modeli



Linking model

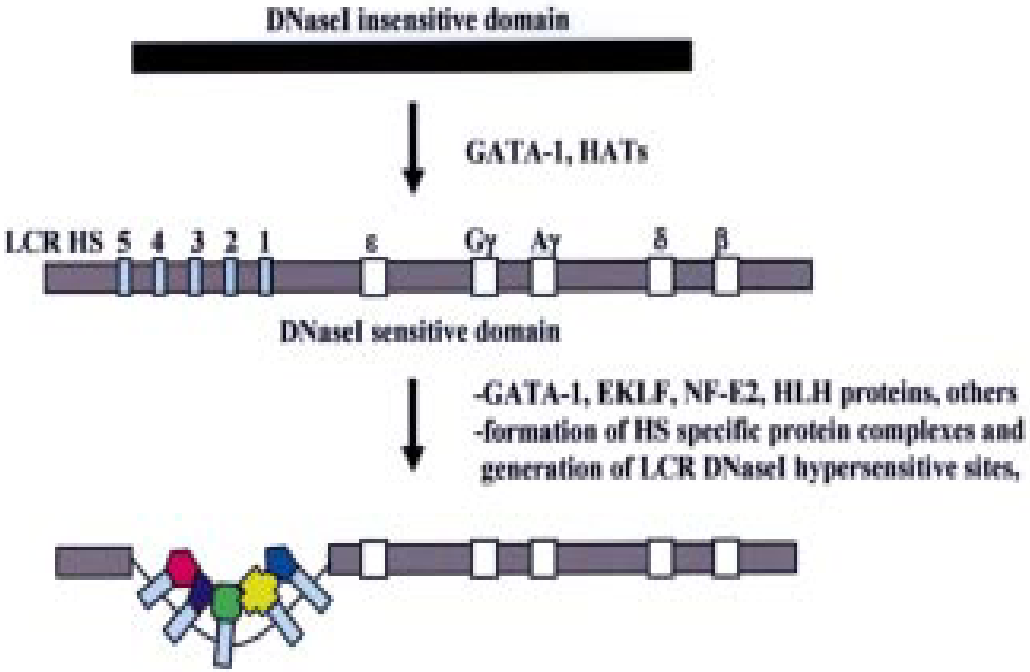


Human:

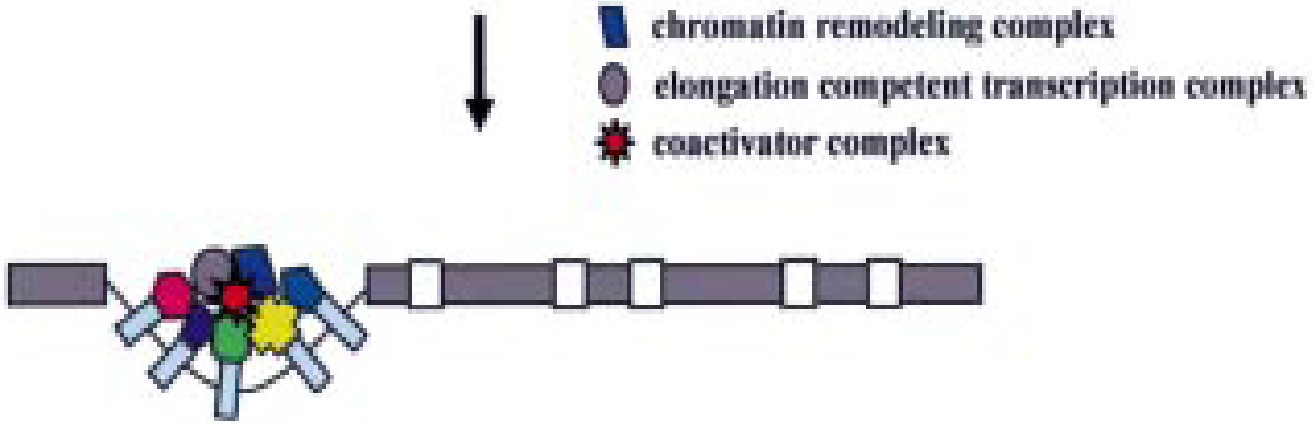
1. β-globin locus[1]
2. Adenosine deaminase gene, human[86]
3. Apolipoprotein E/C-1 gene locus, human[105]
4. T cell receptor α/δ locus, human[96]
5. CD2 gene, human[95]
6. S100 β gene, human[106]
7. Growth hormone gene, human[107]
8. Apolipoprotein B gene, human[108]
9. β myosin heavy chain gene, human[109]
10. MHC class I HLA-B7 gene, human[87]
11. Immunoglobulin heavy chain locus, human[110]
12. Immunoglobulin C alpha 1 & 2 genes, human[111]
13. Keratin 18 gene, human[112]
14. MHC class I HLA G gene, human[113]
15. Complement component C4A & B genes, human[114]
16. Red and green visual pigment genes, human[115]
17. CD4 gene, human[116]
18. α-lactalbumin, human[117]
19. Desmin gene, human[118]
20. CYP19 (aromatase) gene, human[119]
21. *c-fes* proto-oncogene, human[120]

teşekkürler…

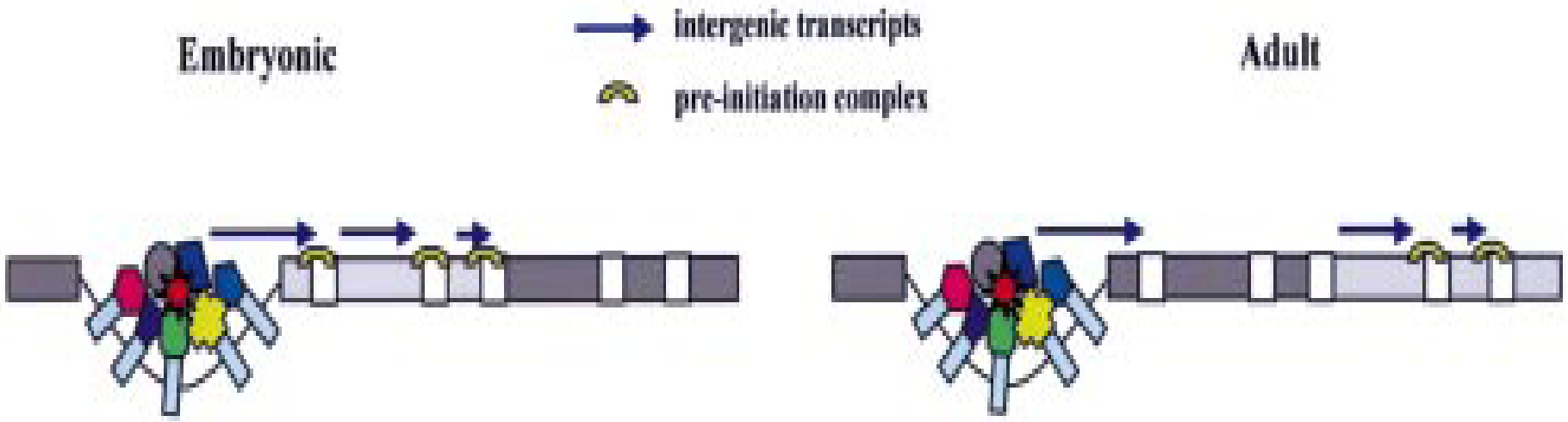
A. Generation of a highly accessible LCR holocomplex



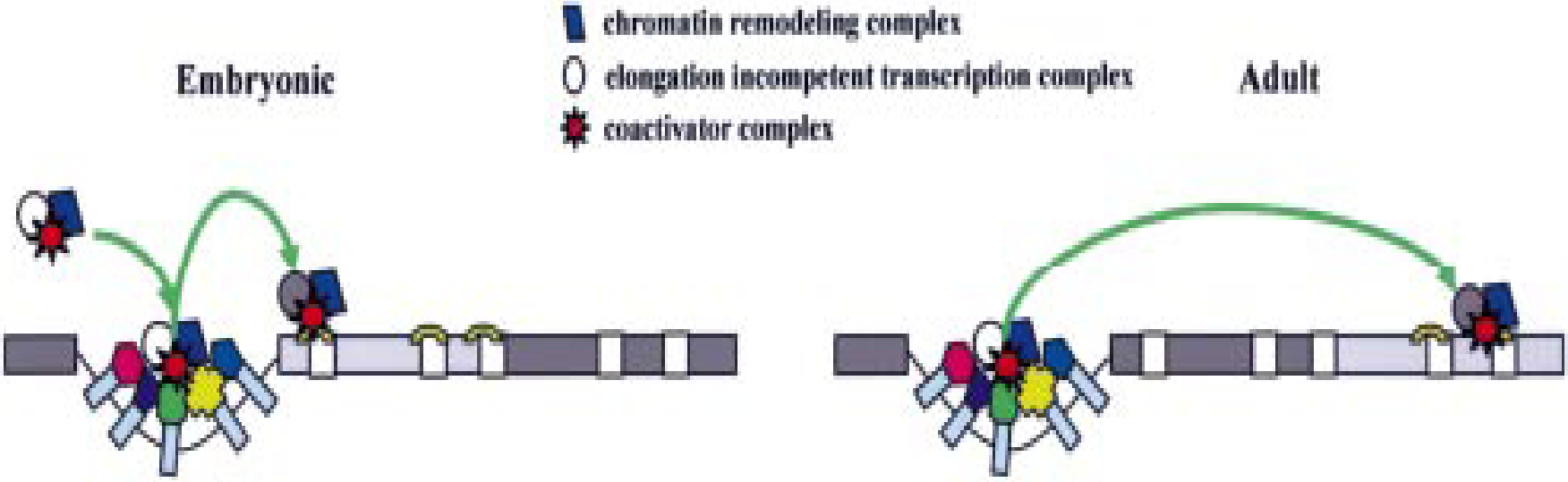
B. Recruitment of chromatin remodeling, coactivator, and transcription complexes



C. Establishment of chromatin domains permissive for transcription



D. Transfer of macromolecular protein complexes to individual globin gene promoters



Binary model

- Lokus aktive olduğu zaman, gen transkripsiyon faktörlerini bağlamakta ve her gen birbirinden bağımsız olarak hareket etmektedir
- Upstream genin transkripsiyonu bilinmeyen downstream değişimlere göre

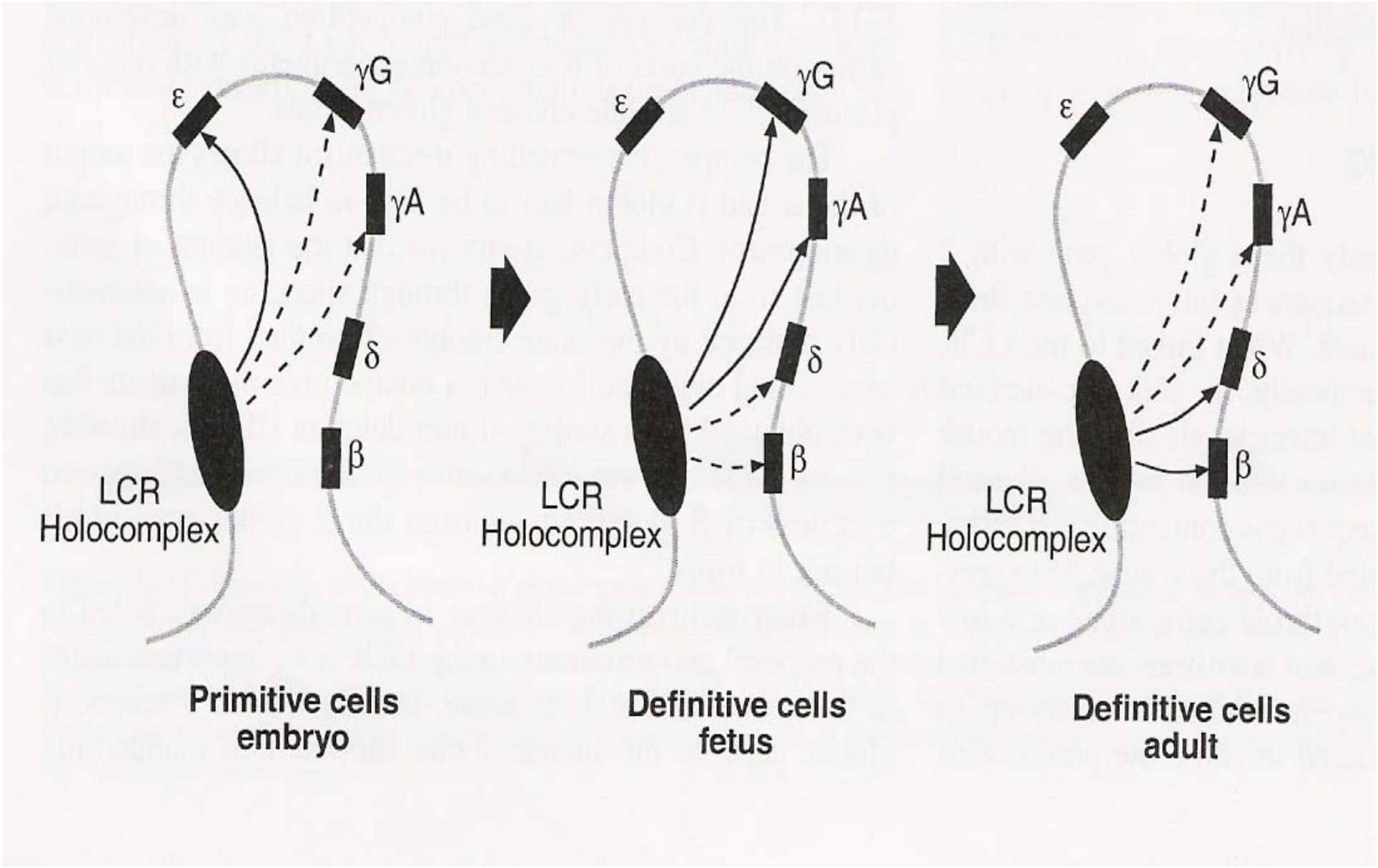
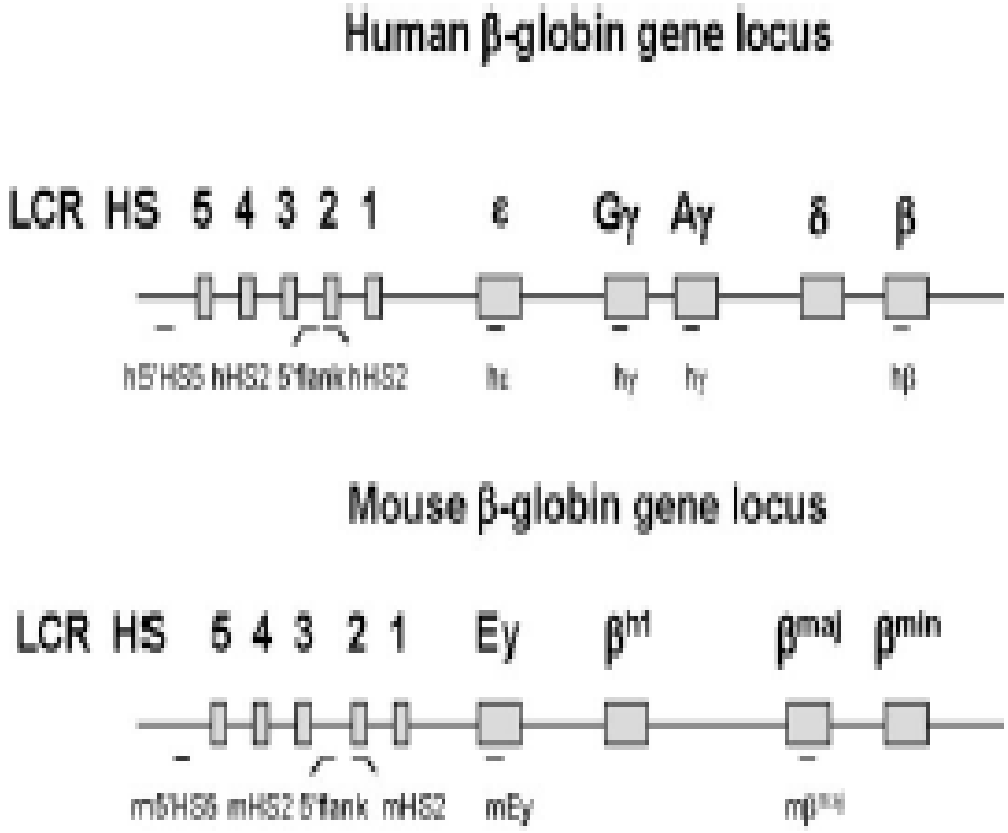
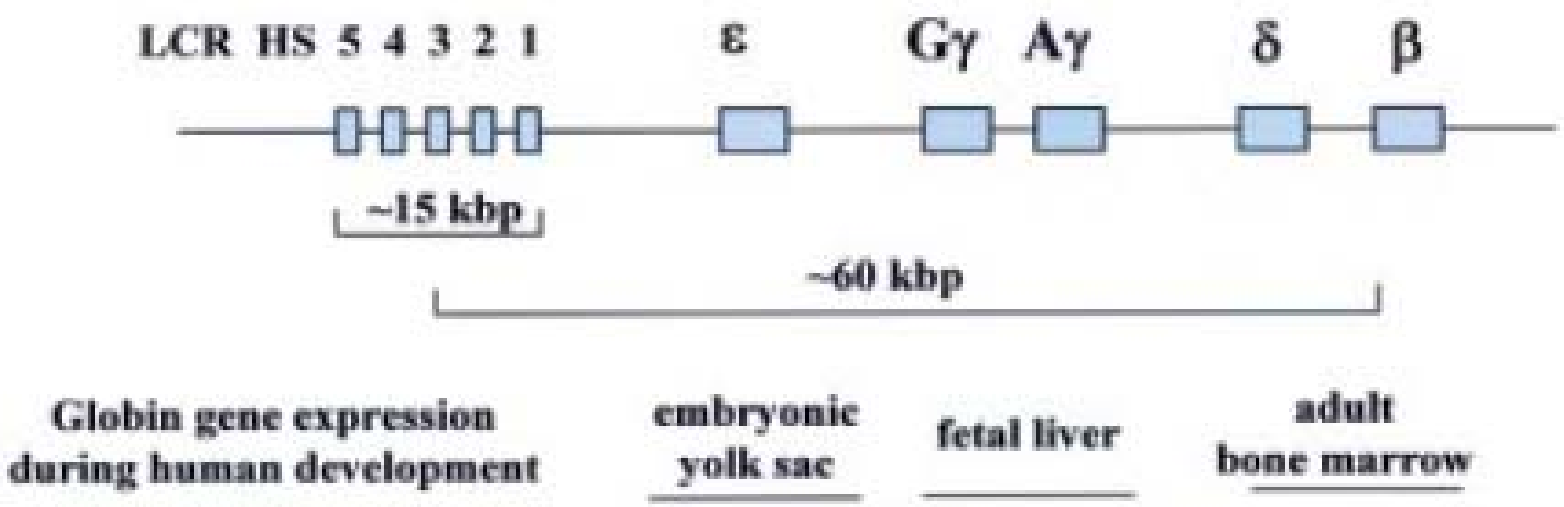

ε
γG
γA
δ
β
LCR
Holocomplex
Primitive cells
embryo
ε
γG
γA
δ
β
LCR
Holocomplex
Definitive cells
fetus
ε
γG
γA
δ
β
LCR
Holocomplex
Definitive cells
adult

FIG. 1. Diagrammatic representation of the human and murine β-globin gene locus. The human β-globin gene locus is located on chromosome 11 and consists of five developmentally regulated genes: the embryonic ε-globin gene, the two fetal γ-globin genes, and the adult δ- and β-globin genes. The LCR is located upstream of the ε-globin gene and is composed of at least five HS sites. The murine β-globin locus is located on chromosome 7 and contains four developmentally regulated genes: the Ey- and β^{h1}-globin genes, which are co-expressed in the embryonic stage and the β^{maj}- and β^{min}-globin genes, which are expressed during the fetal and adult stages of erythropoiesis. The overall organization of the LCR is highly conserved between the murine and the human β-globin locus (2). PCR fragments analyzed in the ChIP experiments or used as probes in the NRO assays are indicated as lines (*horizontal bars*) below the respective loci.




LCR HS 5 4 3 2 1
ε
Gγ
Aγ
δ
β
~15 kbp
~60 kbp
Globin gene expression
during human development
embryonic
yolk sac
fetal liver
adult
bone marrow

E42

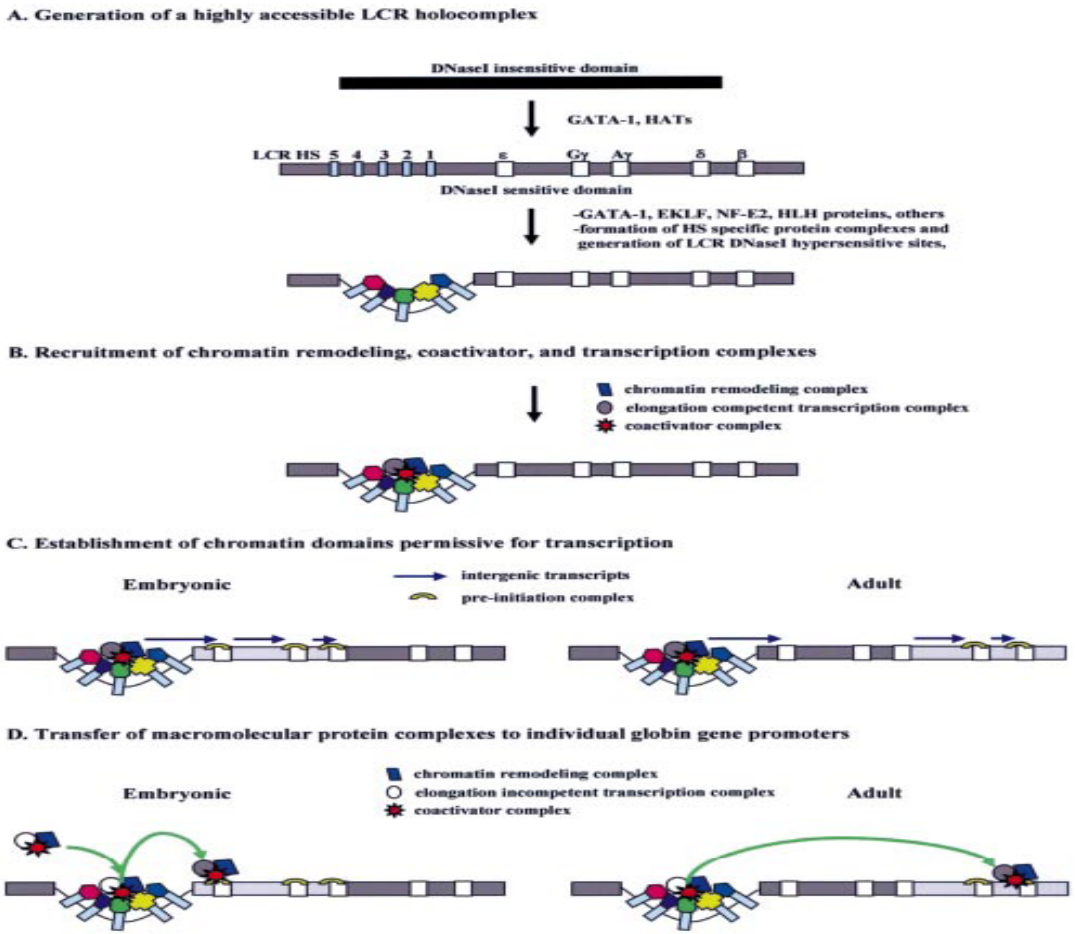
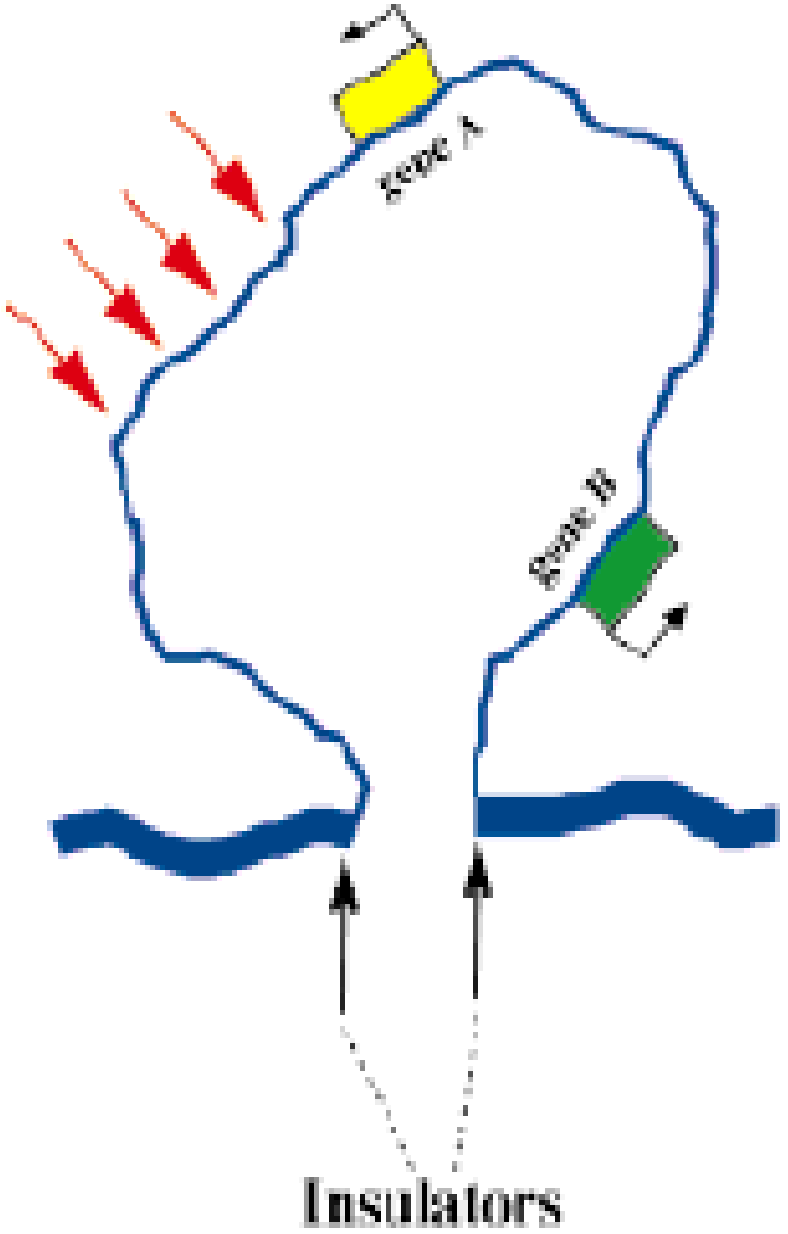


Fig. 2. Multistep model for human β-globin gene regulation. The model depicts four steps proposed to be involved in the regulation of chromatin structure and gene expression in the human β-globin locus. The model focuses on the regulation of the human globin locus in the context of transgenic mice, but it is assumed that the same principal mechanisms govern the correct expression of the β-globin genes during human development, except that the timing of expression of the genes is somewhat different (see Fig. 1). (A) Generation of a highly accessible LCR holocomplex. We propose that the initial events in activating the human globin gene locus during differentiation involves the partial unfolding of the chromatin structure into a DNase I-sensitive domain and the binding of protein complexes to the LCR HS sites. This will then generate the LCR holocomplex, the protein-mediated interaction of HS sites. (B) Recruitment of chromatin-remodeling, coactivator and transcription complexes. Once the LCR holocomplex is generated, the globin locus is relocated to an area of the nucleus enriched for macromolecular complexes involved in coactivation, chromatin remodeling (or modification of histone tails) and transcription. These complexes are recruited to the LCR, which provides a highly accessible platform for recruiting these activities. (C) Establishment of chromatin domains permissive for transcription. The macromolecular protein complexes recruited to the LCR will initially be used to establish chromatin domains that allow transcription of the genes. Specifically, we propose that the LCR recruits elongation-competent transcription complexes (or complexes that are rendered elongation competent at the LCR) that track along the DNA and modify the chromatin structure. This reorganization of the chromatin structure will render the promoters accessible for activating proteins and components of the preinitiation complex. Data published by Gribnau *et al*. [7] suggest that intergenic transcription and chromatin reorganization is stage-specific and restricted to the genes that are expressed either at the embryonic or adult stage. (D) Transfer of macromolecular protein complexes to individual globin gene promoters. Once active chromatin domains are established, the LCR recruits elongation-incompetent transcription complexes, which are transferred to the individual globin gene promoters present in the accessible chromatin domains. The polymerases are then rendered elongation-competent, possibly through phosphorylation of the C-terminal domain [88].

A

Strong domain
Model

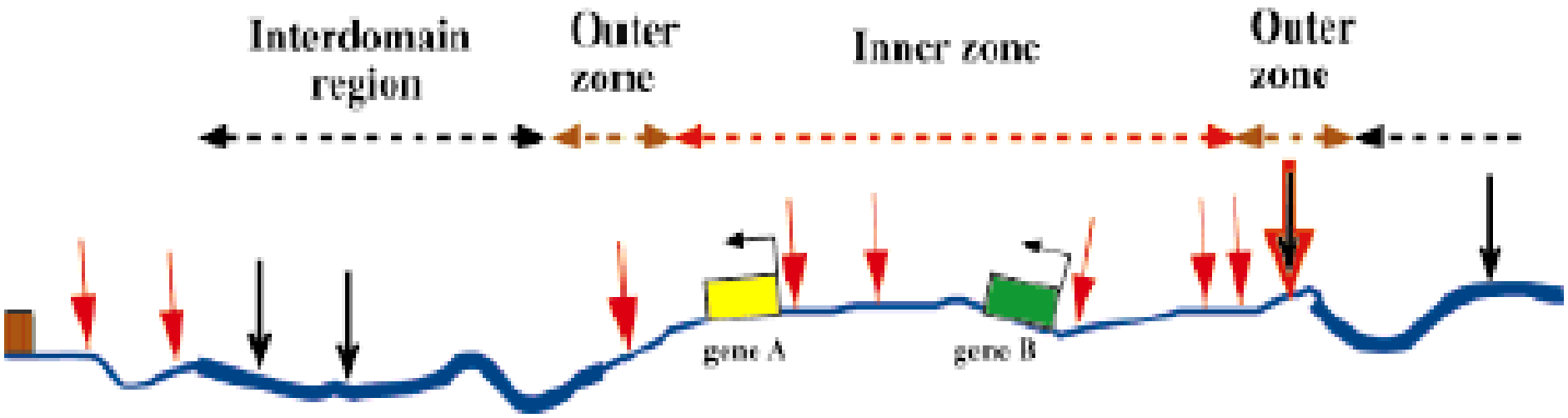


E207

B

Weak domain Model

E207



HS associated with enhancer, LCR or promoter

Insulator

Enhancer or promoter with insulator properties

Condensed chromatin

Open chromatin

E207

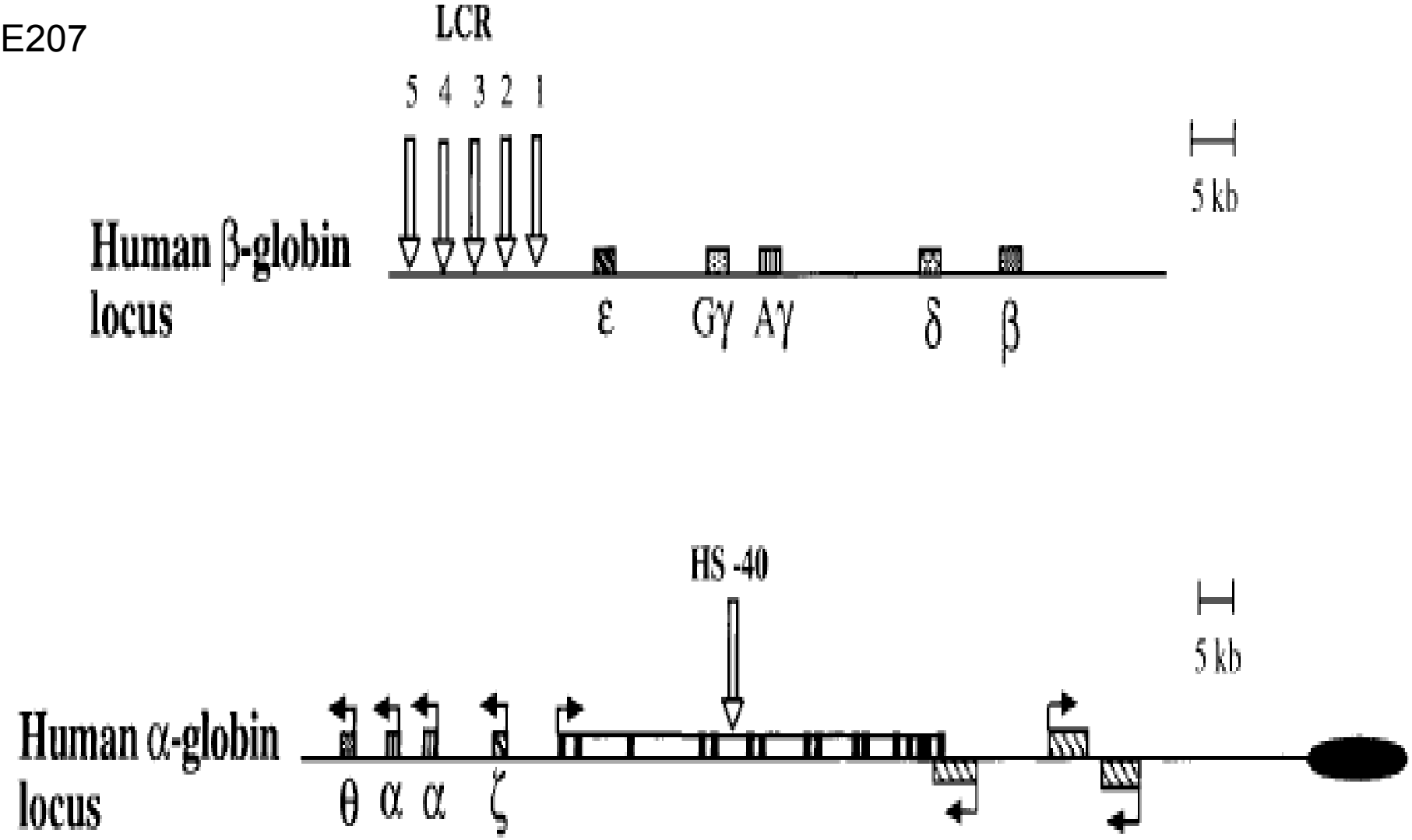


Figure 3. The human α- and β-globin loci are located on separate chromosomes (HS in both loci are represented by vertical arrows). The α-globin locus is located in a gene-dense region close to a telomere. The major HS (HS-40), which is required for expression of the α-globin genes, is located in an intron of a ubiquitously expressed housekeeping gene.[37]

E28

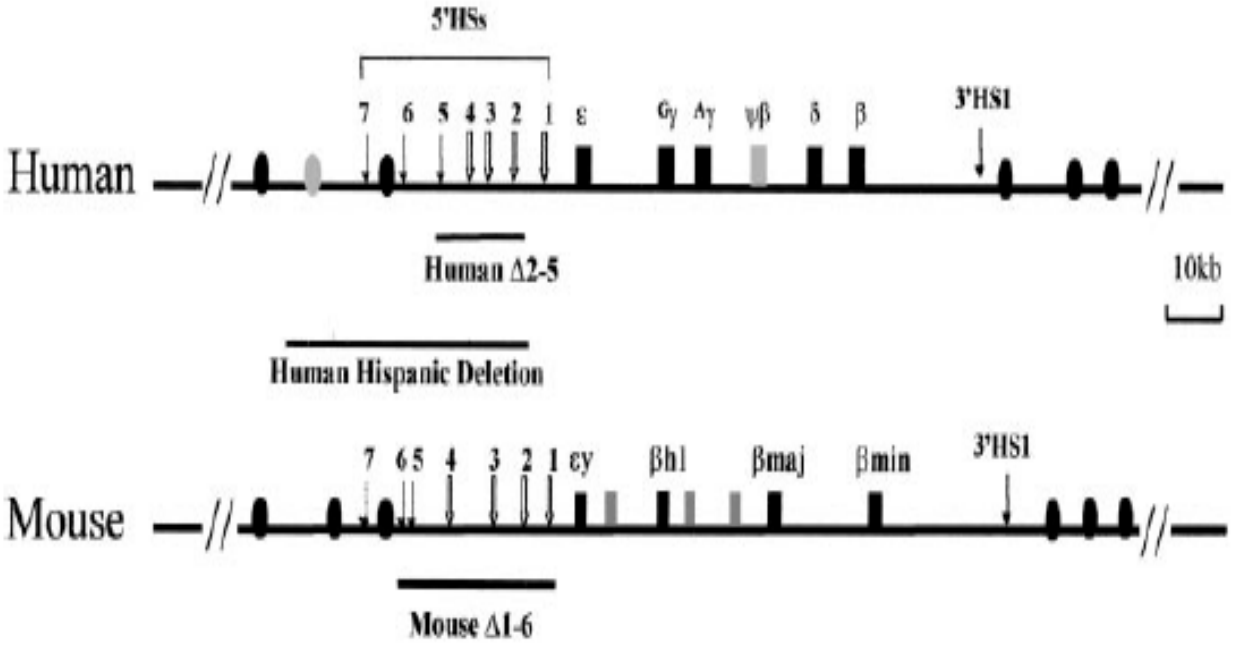


Figure 1. The human and mouse β-globin loci. The human locus consists of 5 functional genes, indicated as dark boxes, arrayed in their order of developmental expression, 5′-ε-^{G}γ-^{A}γ-δ-β-3′. There are 2 developmental switches in globin chain synthesis coincident with changes in site and type of erythropoiesis. During primitive erythropoiesis, the ε-globin gene is expressed in the embryonic yolk sac. The first switch occurs at approximately 8 weeks' gestation; the ε-globin gene is silenced and the ^{G}γ- and ^{A}γ-globin genes are expressed during definitive erythropoiesis in the fetal liver. The second switch occurs shortly after birth; the γ-globin genes are silenced and the β-globin gene and, to a lesser extent, the δ-globin gene are activated in the bone marrow. The HSs 5′HS1 through 5′HS7 are located −6, −11, −15, −18, −22, −28, and −35 kb relative to the ε-globin gene, respectively, and are indicated by arrows. 5′HS1 through 5′HS4 are erythroid specific, but 5′HS5 through 5′HS 7 are not. Another HS (3′HS1) is located 20 kb downstream of the β-globin gene; 3′HS1 is found only in erythroid cells. Boxes represent globin genes and ovals represent olfactor receptor genes; filled ones represent the productive genes and shaded ones the pseudogenes. The lines below the diagram of the locus indicate deletions of the LCR discussed in "In vivo function of LCRs." The Hispanic deletion, which causes $(\gamma\delta\beta)^0$ thalassemia in humans, extends an additional 20 kb 5′ of the LCR 5′HS5.

Figure 2. Models of LCR function. A globin gene is denoted by a green rectangular box with the promoter region indicated in a lighter green. Transcription factors are shown as colored ovals and circles. The 4 erythroid-specific hypersensitive site cores (HSs) are indicated by small red boxes. Blue boxes are the positions of 5′HS5 and 3′HS1, representing likely insulator elements. The flanking DNA sequences of the HSs are depicted as loops between the HS cores. Transcripts are denoted by wavy arrows. (A) Looping model. Transcription factors bind to the LCR HSs and the gene promoter. The LCR directly interacts with the gene promoter by looping out the intervening DNA, thus forming an active transcription complex at the gene promoter. (B) Tracking model. Sequence-specific transcription factors bind to the LCR, forming a complex that tracks down the DNA sequence, as depicted by the large black arrowhead, until encountering transcription factors bound to the appropriate gene promoter, initiating high-level gene expression. (C) Facilitated tracking model. Aspects of both looping and tracking models are combined. Sequence-specific transcription factors bind the LCR; looping then occurs to deliver the bound transcription factors proximal to the gene promoter, followed by tracking, until they encounter transcription factors bound to the appropriate gene promoter. (D) Linking model. Sequential binding of transcription factors along the DNA directs changes in chromatin conformation and defines the transcriptional domain. The transcription factors are linked to one another from the LCR to the gene promoter by non–DNA-binding proteins and chromatin modifiers (shown as small colored circles).

E125

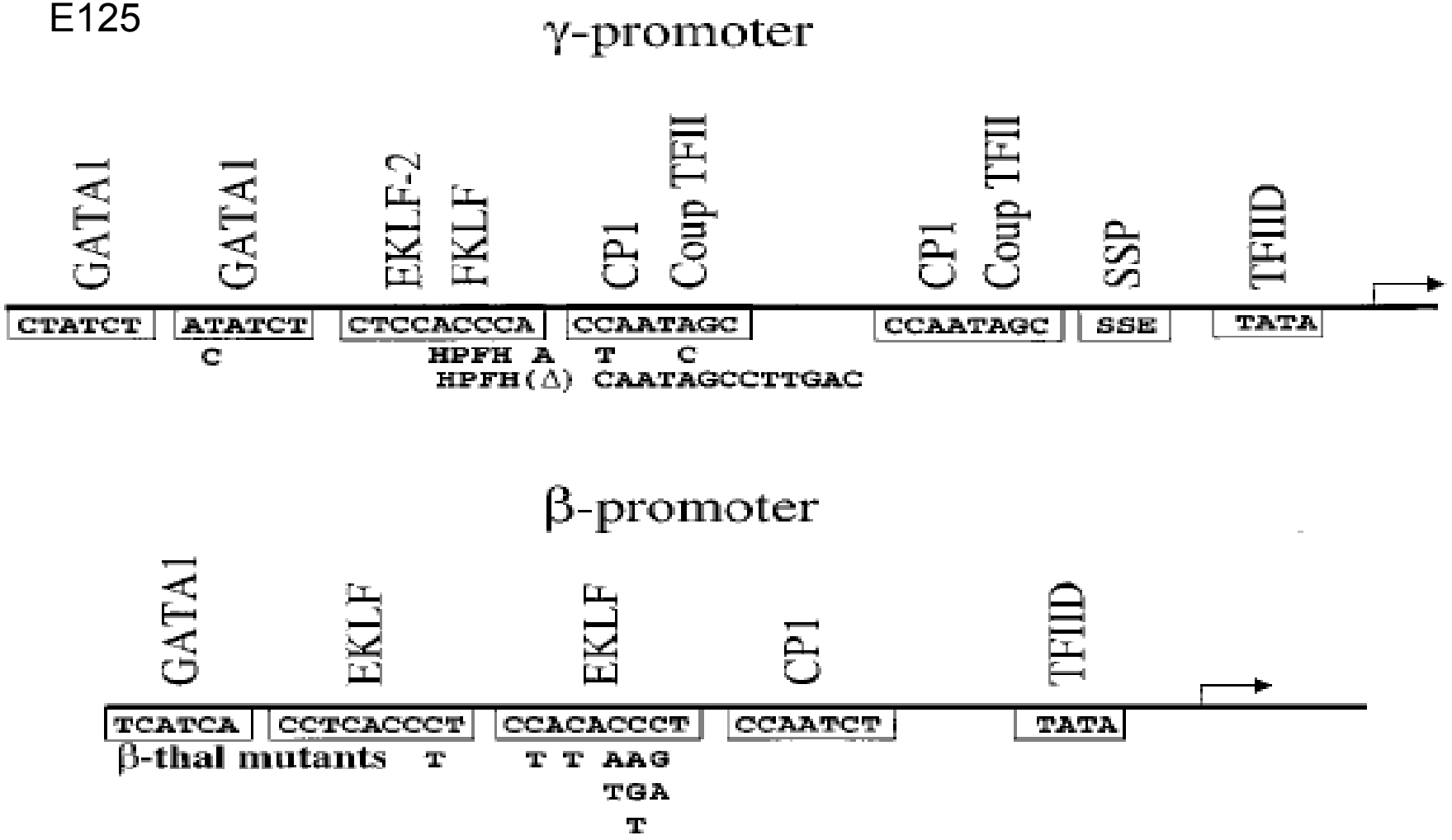


Figure 2. Organization of the γ- and β-globin promoters. The sequences of the most relevant transcription factors binding sites are boxed. Above each site are indicated the factors that are more likely to bind, and below the HPFH or thalassemia point are mutants likely to interfere with transcription factor binding. Note clustering of the HPFH mutants in the γ-globin distal CAAT site and of the thalassemia mutants in the proximal β-globin CACCC box.

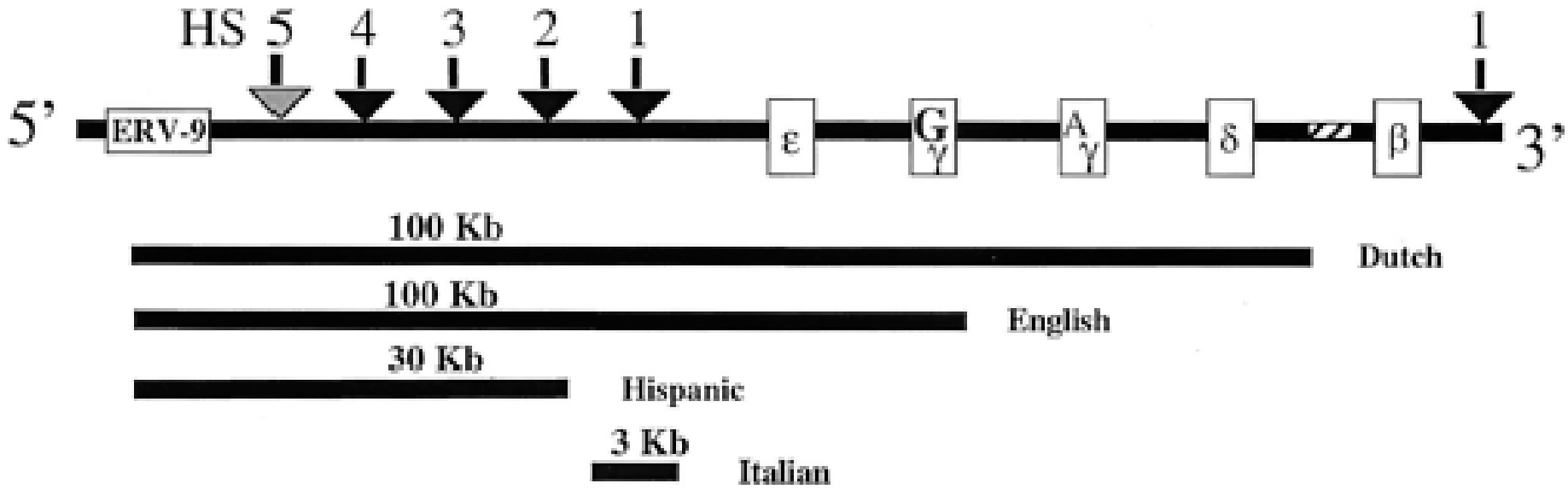


Figure 1. Organization of the β-globin gene cluster. Genes are indicated by *open boxes*, hypersensitive sites (*HS*) by *vertical arrows*, and the origin of replication by a *hatched box*. The position of the long-terminal repeat of the retrotransposon ERV-9 is indicated upstream of the LCR. The four lower *lines* represent the extent of the thalassemia deletions of the LCR.

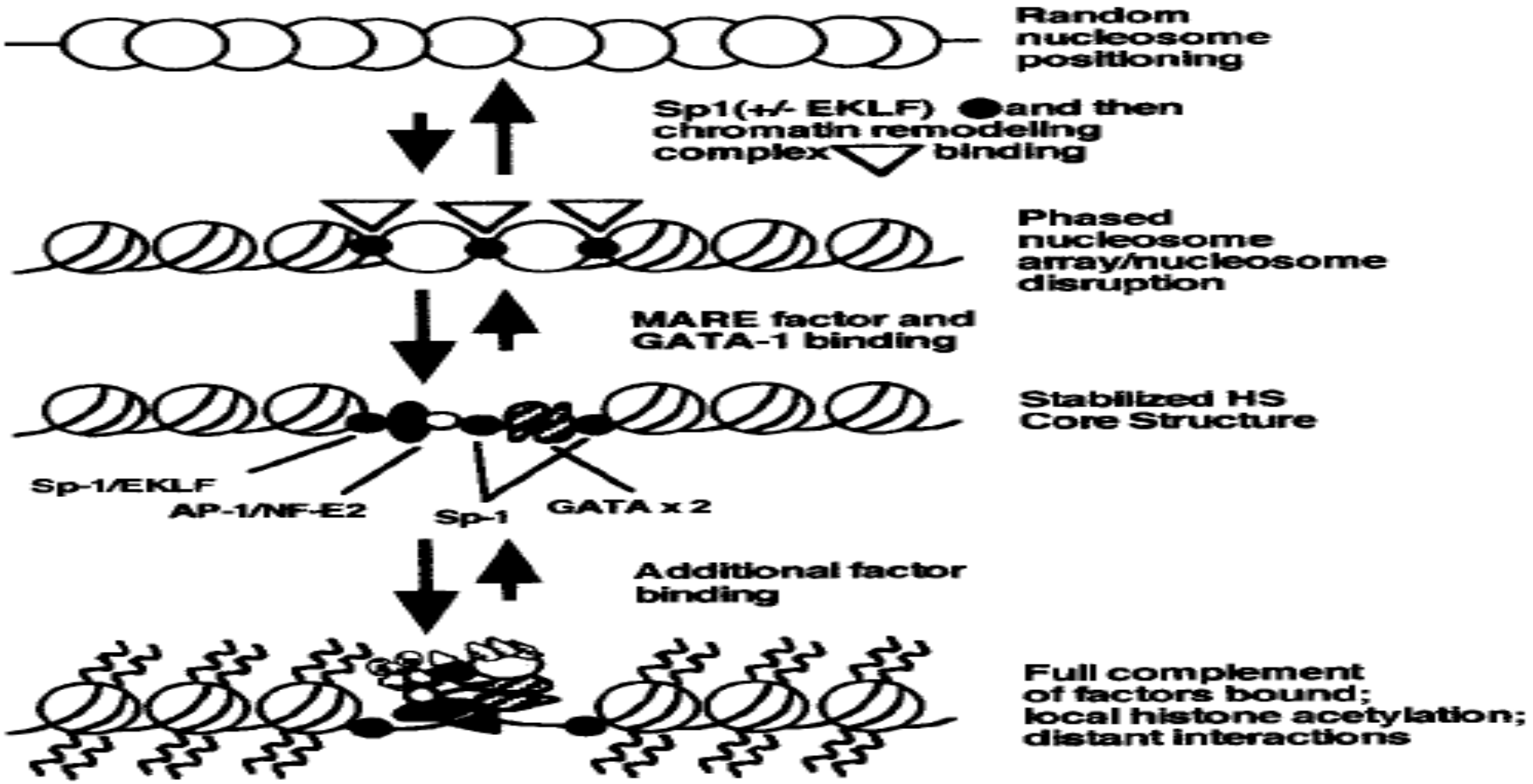


FIG. 10. **Model of β-globin LCR DNase I HS formation.** Prior to factor binding the nucleosomes in the regions of the LCR HSs were randomly positioned. Step 1, Sp1 and/or EKLF binding to sites within the HS core region results in the formation of a phased nucleosomal array and a weak DNase HS. Nucleosome remodeling complexes are proposed to be involved in this step. Step 2, this positioning allows the binding of additional factors, including GATA-1 and NF-E2 (or related factors) to the core region, resulting in the stabilization of the nucleosome-free region, as indicated by the formation of a strong DNase I HS. Step 3, the full complement of factors that comprise the HS core are assembled, resulting in a functional HS core element. Localized histone acetylation mediated by secondary core proteins is also likely to be an important part of the core structure. *S*, Sp1 binding site; *M*, maf-related factor or maf response element (*MARE*) binding site; *GG*, tandem GATA binding sites.

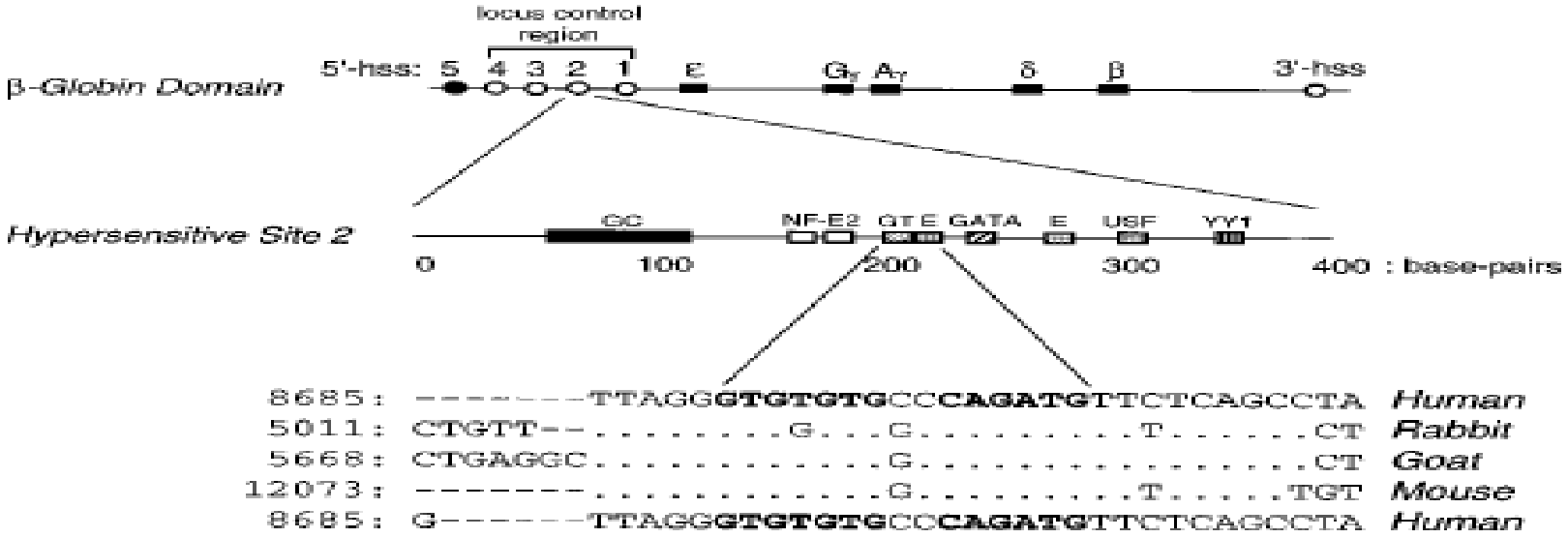


FIG. 1. **A conserved sequence within HS2 of the human β-globin locus control region.** The *diagram* at the *top* shows the β-globin locus on chromosome 11. Transcription factor binding sites within HS2 are shown in the *middle*. *GC*, a GC-rich region that binds multiple proteins; NF-E2, GATA-1, USF, and YY1, known transcription factor binding sites; *E*, E boxes; *GT*, repetitive GT residues. Sequences from the human, rabbit, goat, and mouse β-globin locus have been aligned to reveal conserved sites within the HS2 core. The *boldface sequences* depict two conserved recognition sites for known transcription factors: a GT repeat and an E box.

E99

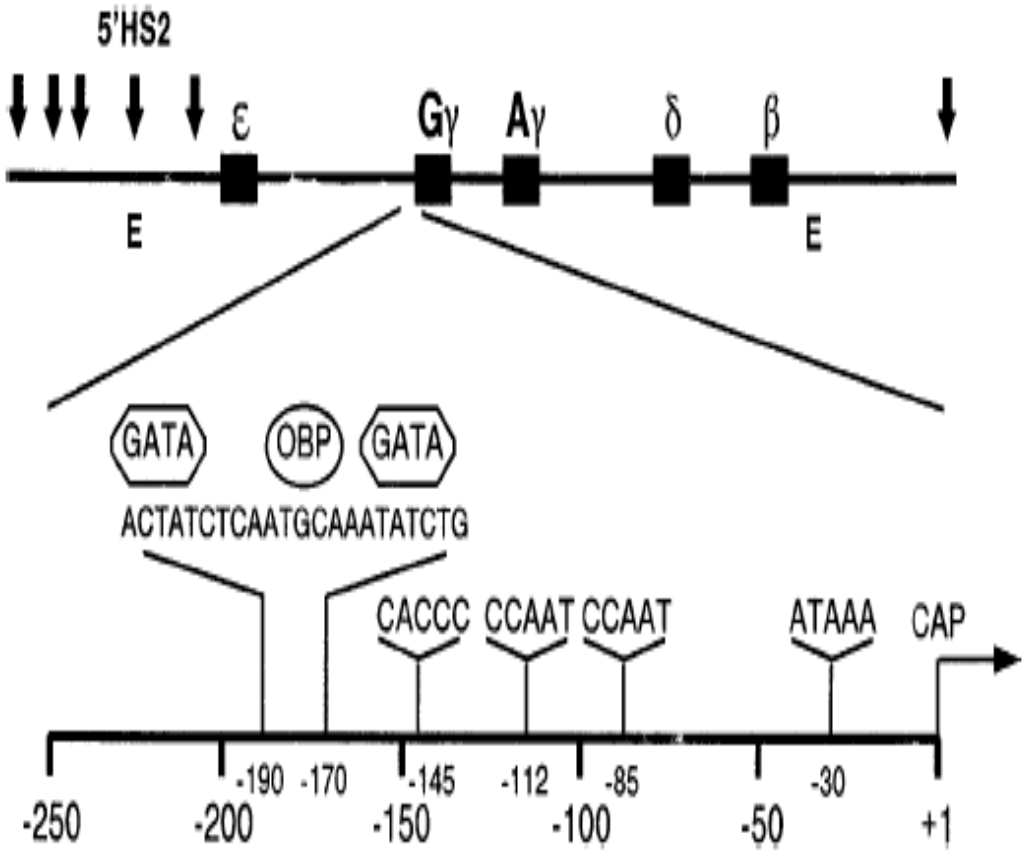
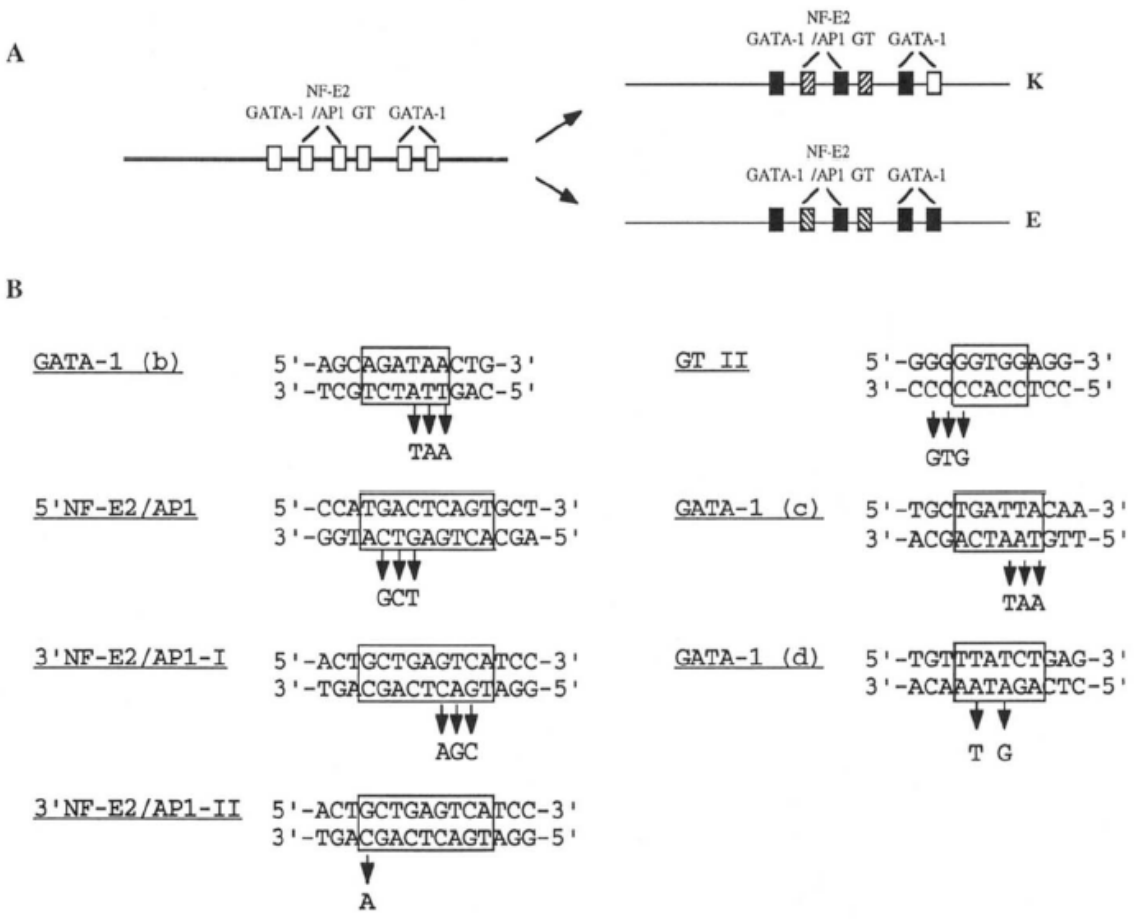


Fig 1. Diagram of the human β-globin gene domain showing the arrangement of the five β-globin genes, the position of the six identified DNase I hypersensitive (HS) sites (arrows), and the location of the 5′HS2 and 3′β enhancer elements (E). The expanded view of the G γ gene promoter region shows the recognition sites for the DNA-binding proteins GATA-1 and OBP.

E113

FIG. 1. *A*, nuclear factor binding motifs of HS-40. The motifs that bind nuclear factors *in vitro* are indicated by the *blank boxes* on the *left map*. Those motifs that bind nuclear factors *in vivo* in K562 cells (*K*) and human adult erythroblasts (*E*) are indicated by the *filled* or *hatched boxes*. Genomic footprints are similar for motifs with *filled boxes*, while those of the *hatched motifs* are different between the two types of erythroid cells. The three GATA-1 motifs shown are GATA-1(b), GATA-1(c), and GATA-1(d), respectively, as arranged on HS-40 from *left* to *right*. The two NF-E2/AP1 motifs are 5′NF-E2/AP1 and 3′NF-E2/AP1, also arranged from *left* to *right*. The single GT motif is termed GT II throughout the text. The maps are constructed from the data of references (Jarman *et al.*, 1991; Strauss *et al.*, 1992; Zhang *et al.*, 1993). *B*, mutants of HS-40 motifs. The mutations of the nuclear factor binding motifs of HS-40 were generated by site-directed mutagenesis. For each motif and its flanking DNA, the wild type sequence is shown in full, with the central binding site of nuclear factor(s) *boxed*. The mutated base or bases are indicated by *downward arrows*. Note that two different mutants of the 3′NF-E2/AP1 motif, I and II, have been generated and analyzed (see text for details).



E121

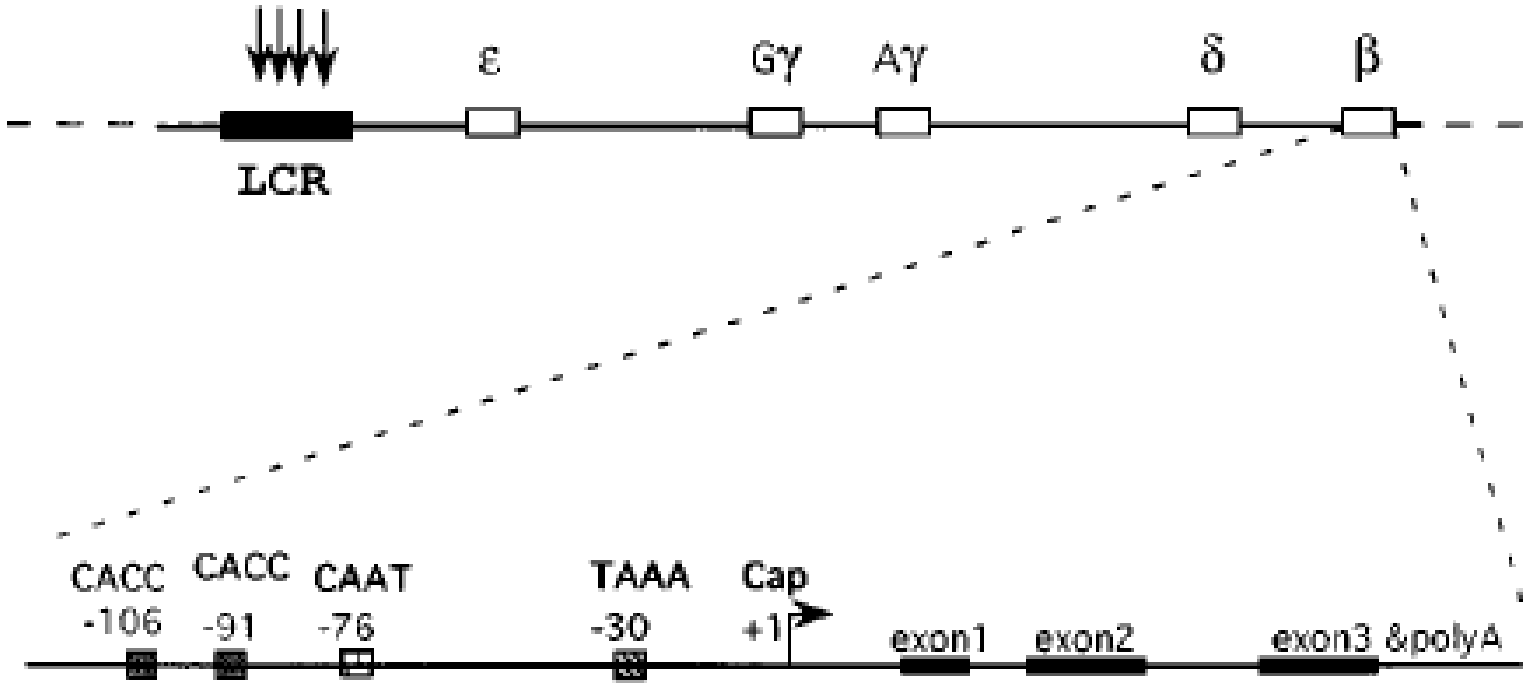
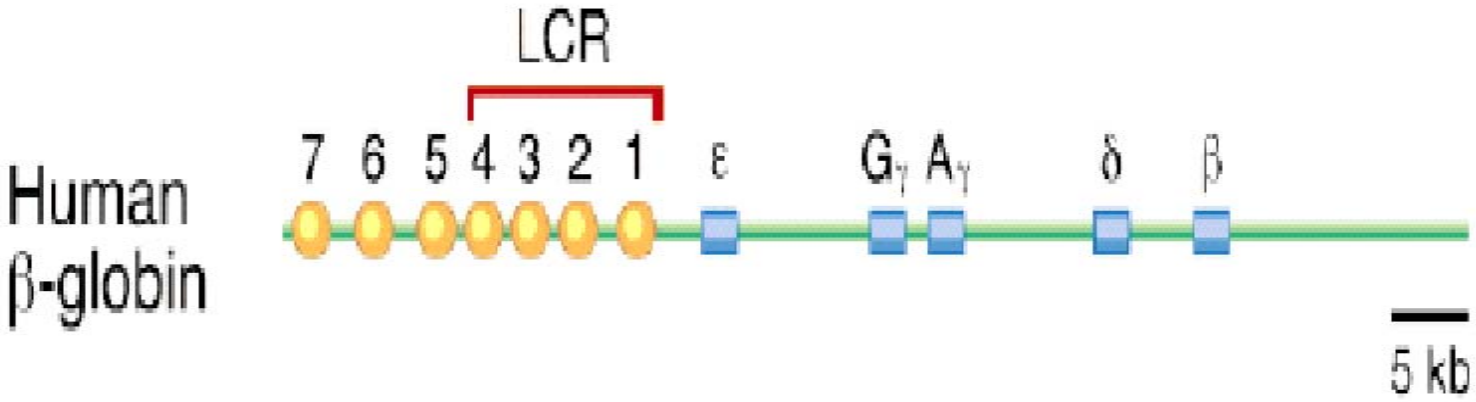


FIG. 1. **Structure of the human β-globin gene complex.** The promoter and the rest of the β-globin gene are not drawn to scale relative to each other.


LCR
7 6 5 4 3 2 1
ε
Gγ Aγ
δ
β
Human
β-globin
5 kb

E207

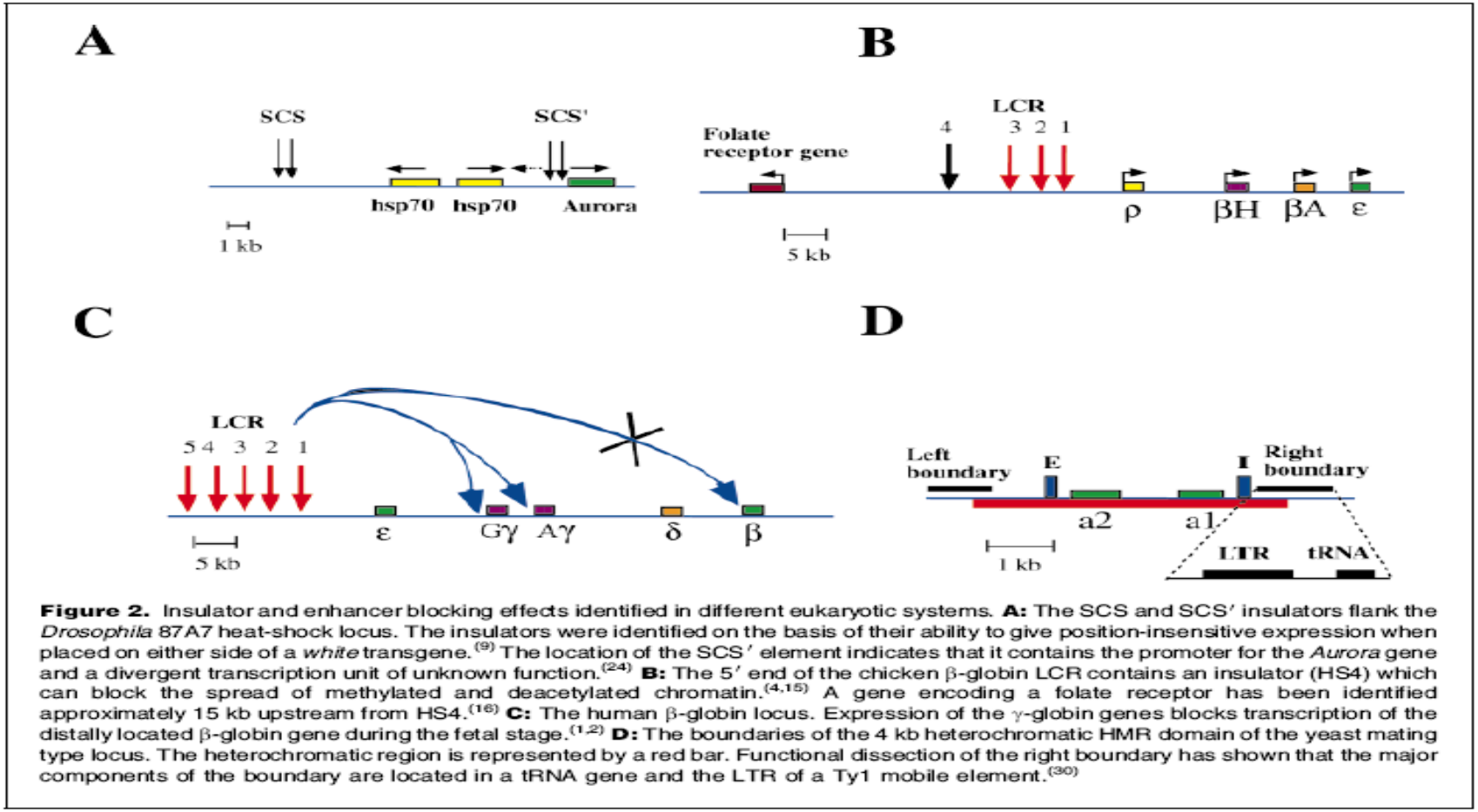


Figure 2. Insulator and enhancer blocking effects identified in different eukaryotic systems. **A:** The SCS and SCS′ insulators flank the *Drosophila* 87A7 heat-shock locus. The insulators were identified on the basis of their ability to give position-insensitive expression when placed on either side of a *white* transgene.[9] The location of the SCS′ element indicates that it contains the promoter for the *Aurora* gene and a divergent transcription unit of unknown function.[24] **B:** The 5′ end of the chicken β-globin LCR contains an insulator (HS4) which can block the spread of methylated and deacetylated chromatin.[4,15] A gene encoding a folate receptor has been identified approximately 15 kb upstream from HS4.[16] **C:** The human β-globin locus. Expression of the γ-globin genes blocks transcription of the distally located β-globin gene during the fetal stage.[1,2] **D:** The boundaries of the 4 kb heterochromatic HMR domain of the yeast mating type locus. The heterochromatic region is represented by a red bar. Functional dissection of the right boundary has shown that the major components of the boundary are located in a tRNA gene and the LTR of a Ty1 mobile element.[30]

E207

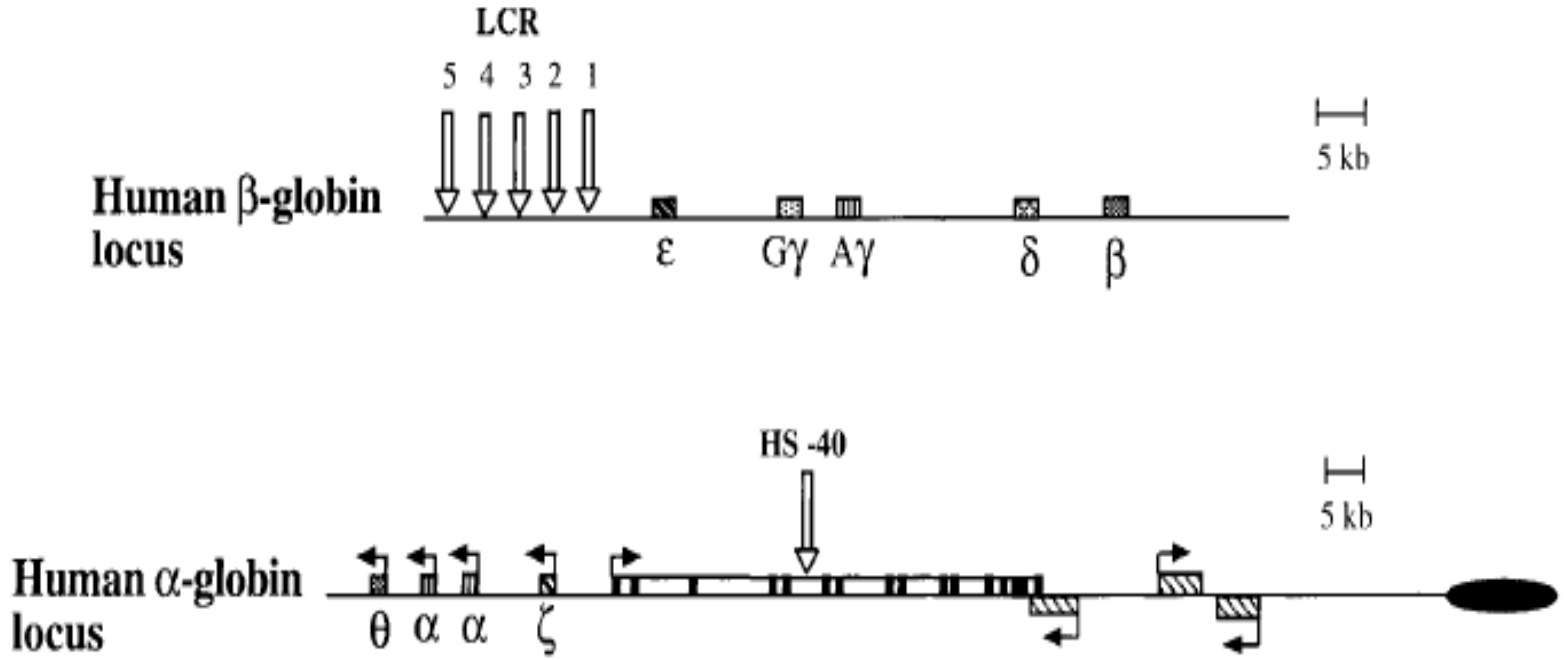


Figure 3. The human α- and β-globin loci are located on separate chromosomes (HS in both loci are represented by vertical arrows). The α-globin locus is located in a gene-dense region close to a telomere. The major HS (HS-40), which is required for expression of the α-globin genes, is located in an intron of a ubiquitously expressed housekeeping gene.[37]

E207

Functional definitions used to classify cis-acting sequences that affect transcription*

Promoter	Specifies site of transcriptional initiation in *in vitro* and *in vivo* assays. Promoters bind the basal transcriptional machinery and additional potentiating factors. Protein coding genes always have their promoters located immediately upstream from the site of transcriptional initiation.
Enhancer	Increases transcription from a linked promoter. Effect is orientation independent and shows some flexibility with respect to distance. The definition does not specify the level of transcriptional enhancement or the assay used to measure it, although enhancers were originally defined in transient expression assays in cultured cells.
Locus control region (LCR)	Dominant activating sequences that confer position independent and copy dependent expression on a linked gene in transgenic mice. Capable of activating expression at single copy. Level of expression per copy should be equivalent to that of the gene in its normal location.
Insulator	Blocks the action of an enhancer on a promoter when placed between them. Insulates transgenes from positive and negative position effects.
DNase I hypersensitive site (HS)	Short region of sequence (200–300 bp) which has an increased sensitivity to digestion with DNase I compared to surrounding chromatin. HS are generally detected by digesting isolated nuclei with increasing amounts of DNase I. Hypersensitivity is caused by binding of factors to clustered binding sites resulting in displacement of a nucleosome.

*Note: The operational nature of these definitions means that they are dependent on the parameters of the assays used to formulate them. This makes it problematic to directly compare different definitions.

E198

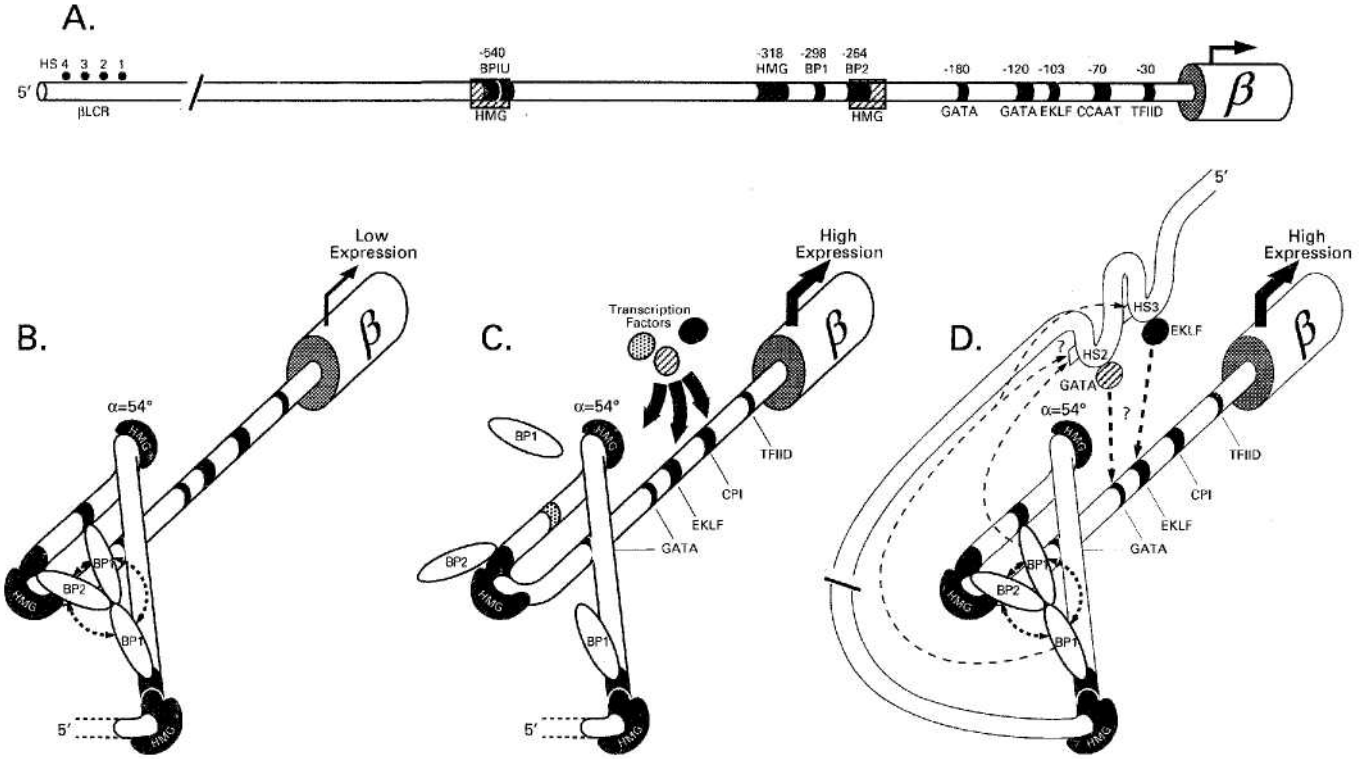
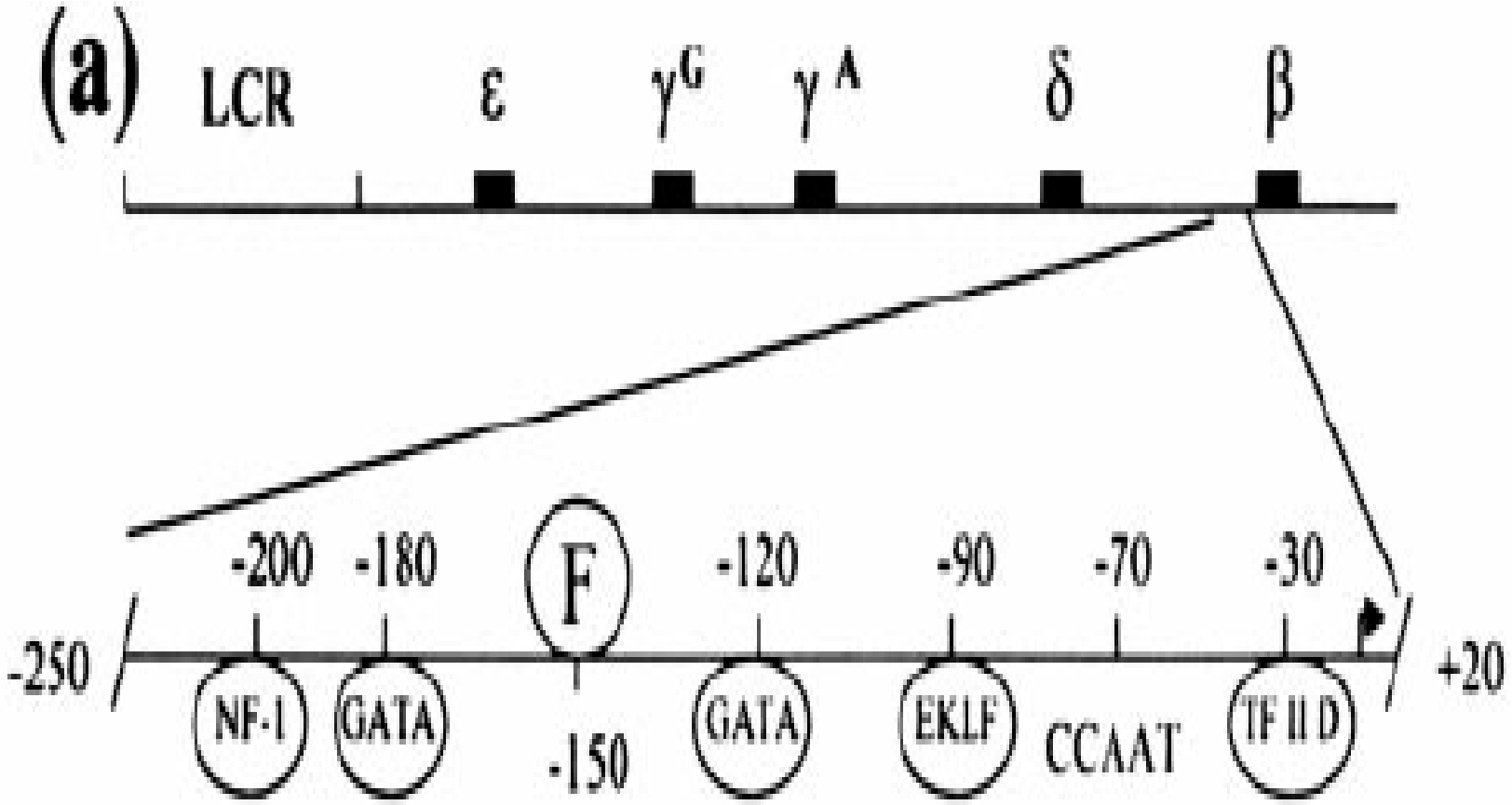


Figure 5. Combinatorial model of adult β-globin gene repression. (**A**) Wild-type β-globin gene showing known transcription factor DNA-binding sites. (**B**) Model shows interaction between these DNA binding sites with HMG1+2, BP1 and BP2 factors. Collectively, they provide DNA–protein and protein–protein interactions needed to form a stable silencing complex. (**C**) Modification of the model in which disruption of the stable DNA–protein, protein–protein silencing complex (seen here by *in vitro* mutagenesis of functional assays by the DNA mutation of BP1 binding) activates the β-globin gene in K562 environment. (**D**) Silencing disruption in erythroid cells via βLCR interaction with the β-promoter, leading to gene activation.

E183


(a)
LCR
ε
γ G
γ A
δ
β
-200
-180
F
-120
-90
-70
-30
-250
+20
NF-1
GATA
-150
GATA
EKLF
CCAAT
TF II D

E159

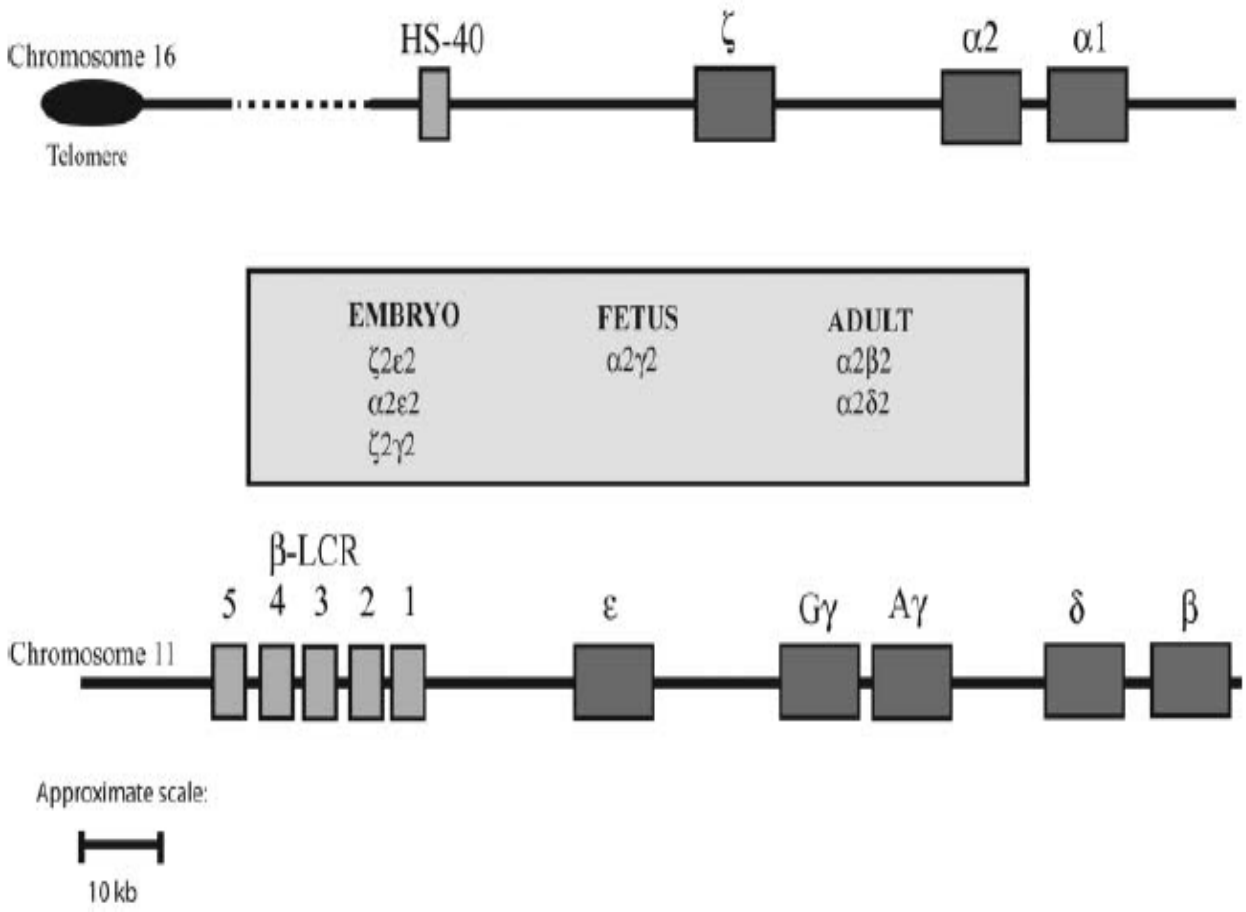


Figure 1. Schematic representation of the human α (top) and β (bottom) globin gene clusters.

Pseudogenes are not represented. In order to make them easier to visualize, the globin-like genes are larger than scale.

E156

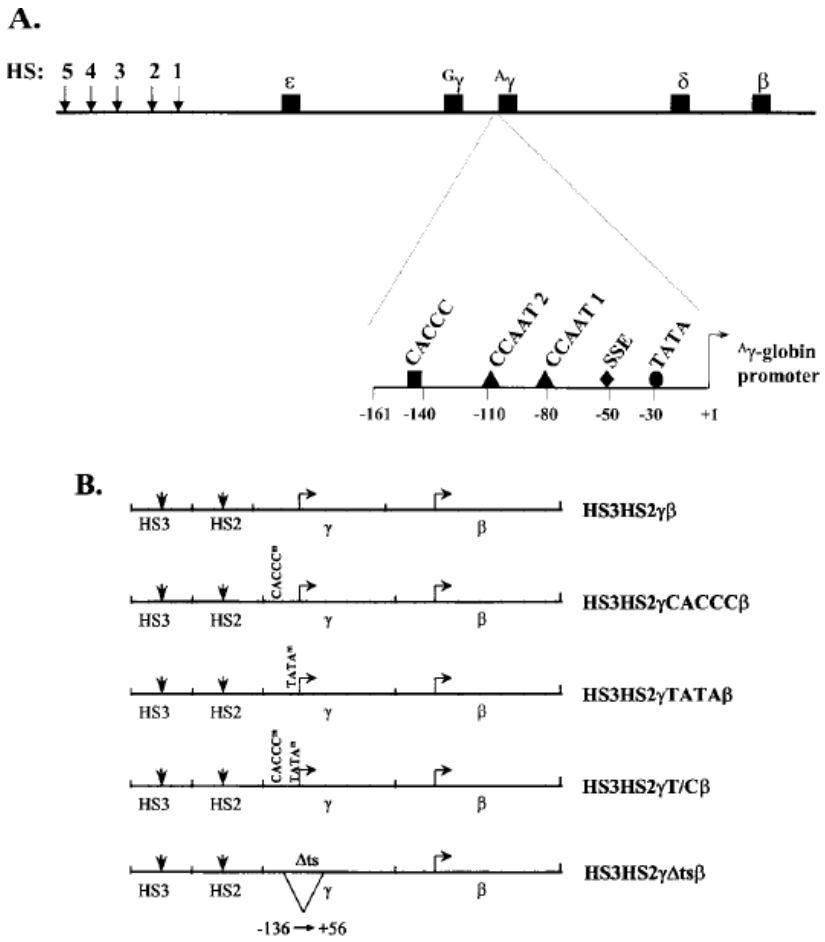


FIG. 1. *A*, the human β-globin locus and regulatory elements in the ^{A}γ-globin promoter. The *arrows* represent the five DNase I super-hypersensitive (*HS*) sites in the LCR. Below the locus is a diagram indicating the locations of the CACCC, the two CCAAT, the stage-selector element, and the TATA elements downstream of −161 in the γ-globin promoter. The CACCC and/or TATA elements in the ^{A}γ-globin promoter have been mutated in the constructs described below. *B*, the DNA constructs tested in stable transfection assays and transgenic mice. All of the constructs contain a 1.9-kb fragment containing HS3 from the LCR, a 1.9-kb fragment containing HS2, a 3.3-kb human γ-globin gene, and a 4.5-kb human β-globin gene. The HS3HS2γβ construct has the wild-type γ-globin gene, and the other constructs contain the indicated mutations in the γ-globin promoter. HS3HS2γTATAβ has base substitutions in the TATA box at −30, HS3HS2γ-CACCCβ has base substitutions in the CACCC element at −140, and HS3HS2γT/Cβ has both the TATA and the CACCC mutations. HS3HS2γΔtsβ has a deletion from −136 to +56 in the γ-globin promoter.

E137

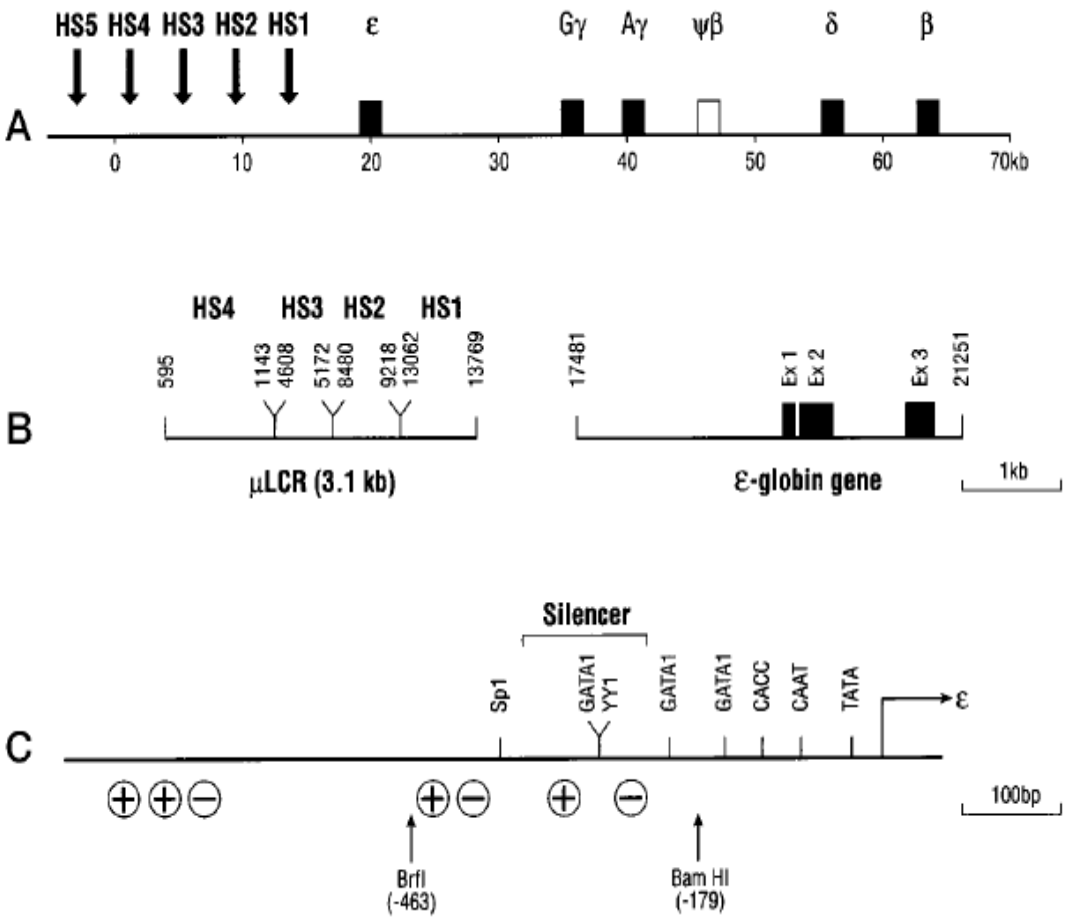


FIG. 1. **ε-gene constructs used for production of transgenic mice.** *A*, schematic representation of the human β-globin locus. Numbers correspond to GenBank™ coordinates (Humhbb). The *thick vertical arrows* indicate the DNase I hypersensitive sites of the LCR. *Filled boxes* show the five transcribed globin genes, whereas the *open box* marks the position of the pseudo β gene. *B*, the *left line* shows the μLCR, a 3.1-kb truncated version of the LCR. The *numbers* above the line indicate the 5′ and 3′ ends of each HS fragment. The *right line* is a 3.7-kb *Eco*RI fragment encompassing the ε-globin gene spanning from −2025 to +1745 relative to the cap site. *C*, ε-globin gene promoter. Shown are the location of the conserved boxes and the binding motifs for various proteins. + and − correspond to the sites of positive and negative elements identified by transient transfection assays (43). The truncated positions of the two μLCRε constructs used in this study are indicated by *arrows* (*Bam*HI and *Brf*I).

E100

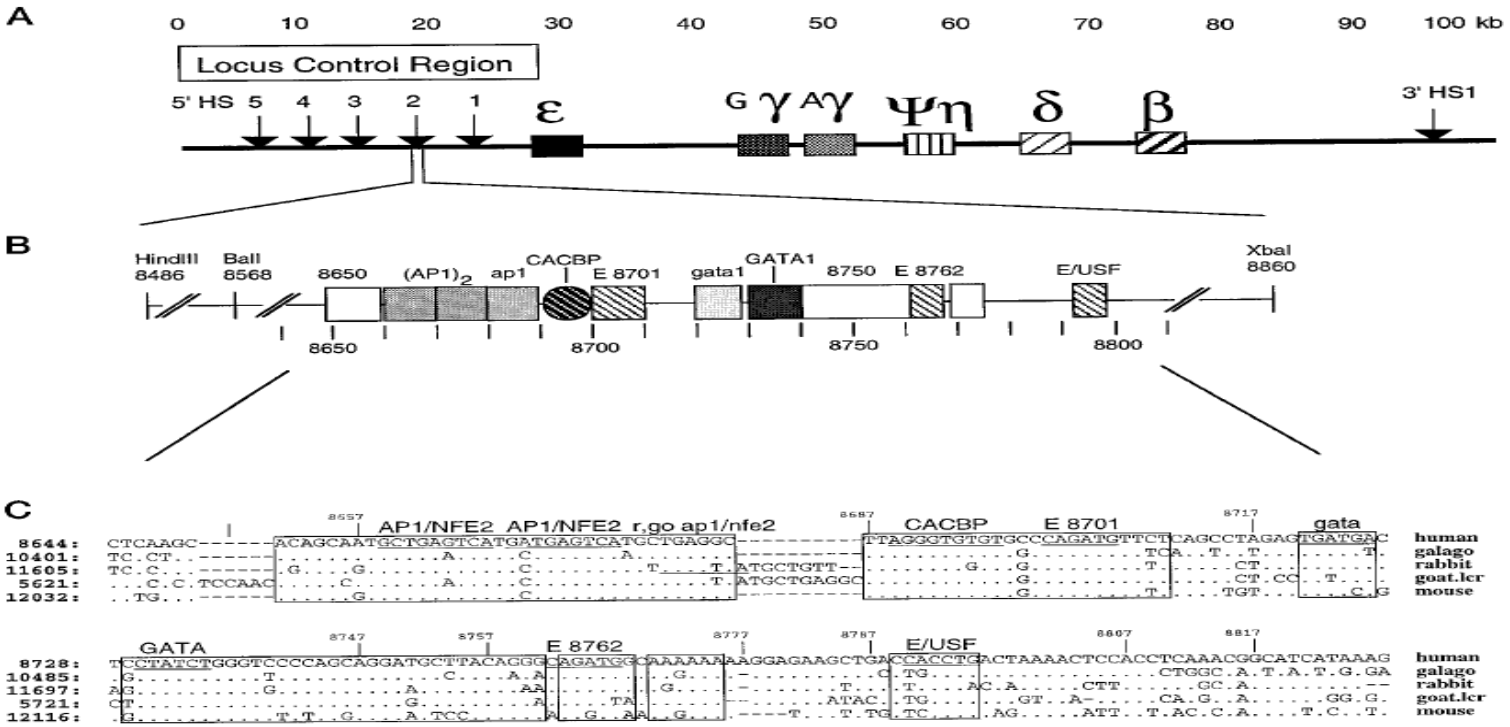


FIG. 1. **Important segments in HS2 of the β-LCR.** ***A***, a map of the human β-globin gene cluster, showing the β-LCR and developmentally stable DNase HSs as *arrows*. ***B***, a diagram of notable sites within the core of HS2. Binding sites for proteins are labeled and given a distinctive fill, and *open boxes* are conserved sequences with no currently identified binding protein. Nucleotide positions refer to GenBank file HUMHBB. ***C***, a simultaneous alignment of DNA sequences of HS2 from human, galago, rabbit, goat, and mouse. ***Boxes*** are drawn around runs of consecutive conserved columns that are not contained in a longer such run. An alignment column is called "conserved" if it is contained in a run of 6 consecutive columns that have a "model row" of length 6, such that each row (of length 6) has at most one mismatch with the model row.[3] This criterion was developed to find blocks that are likely to represent protein-binding sites while allowing one mismatch per species. Binding sites for proteins are *underlined* and labeled.

E3

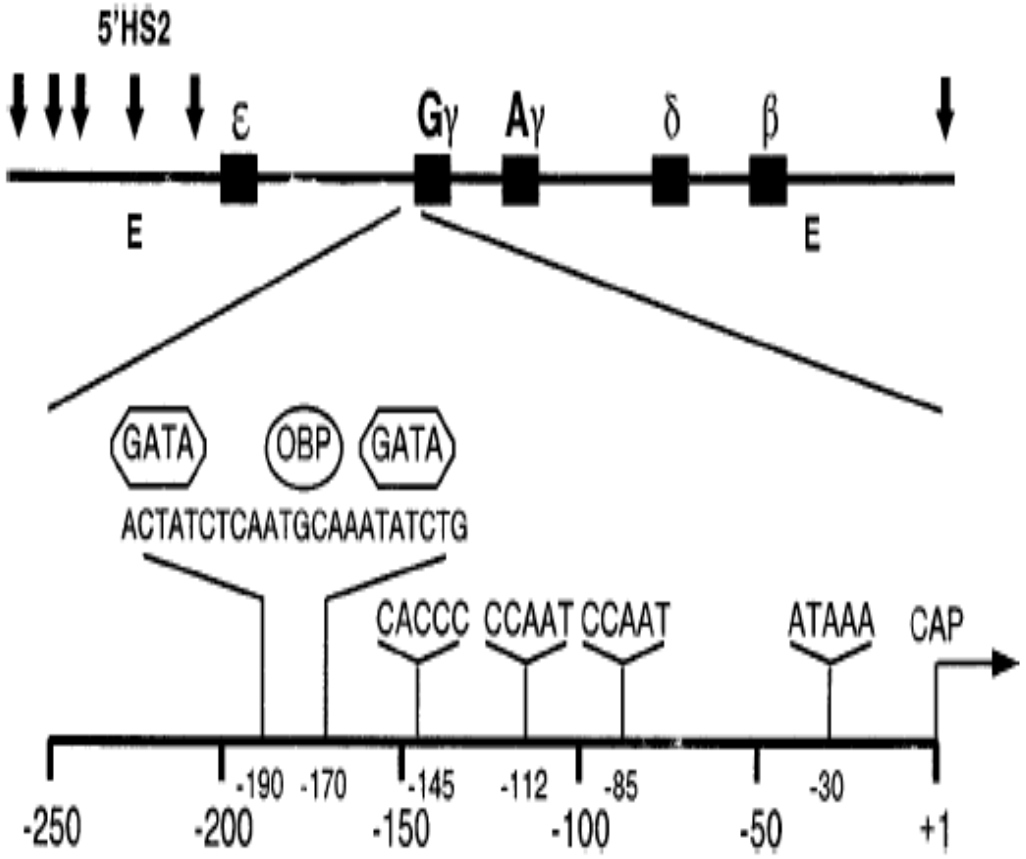


Fig 1. Diagram of the human β-globin gene domain showing the arrangement of the five β-globin genes, the position of the six identified DNase I hypersensitive (HS) sites (arrows), and the location of the 5′HS2 and 3′β enhancer elements (E). The expanded view of the G γ gene promoter region shows the recognition sites for the DNA-binding proteins GATA-1 and OBP.

E21

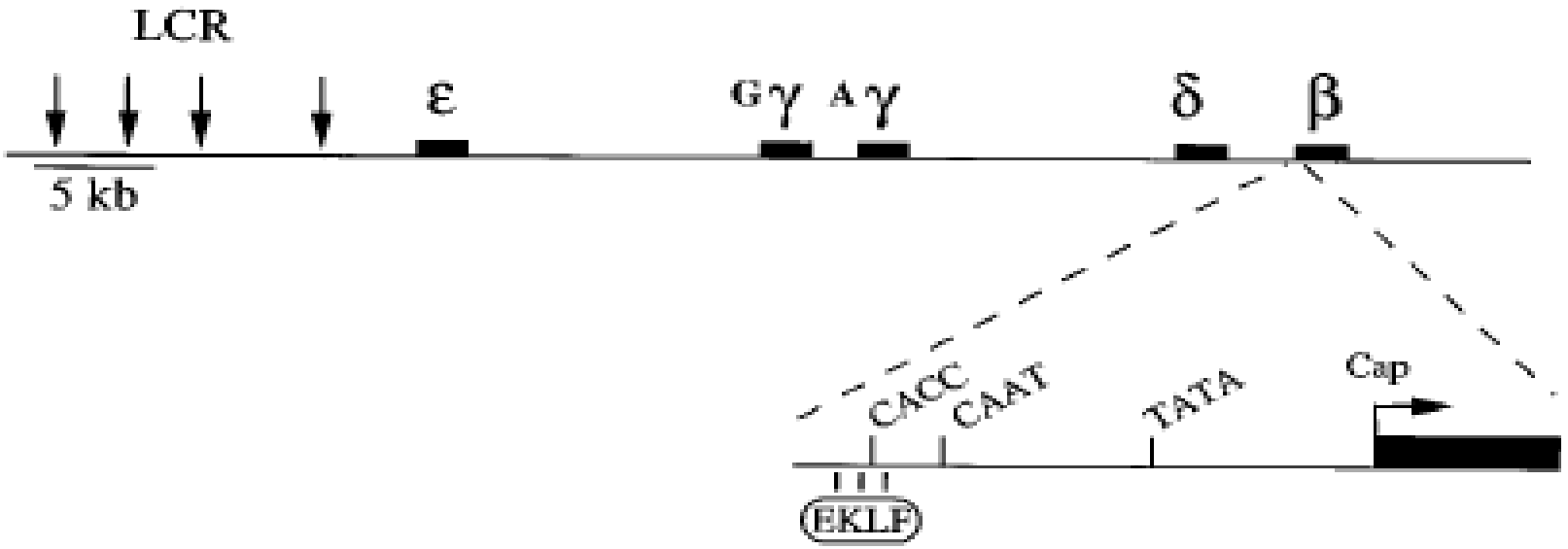


Fig 1. Structure of the human β-like globin locus and the β-globin gene minimal promoter. The figure also shows EKLF that interacts with the CACC motif element of the promoter.

E24

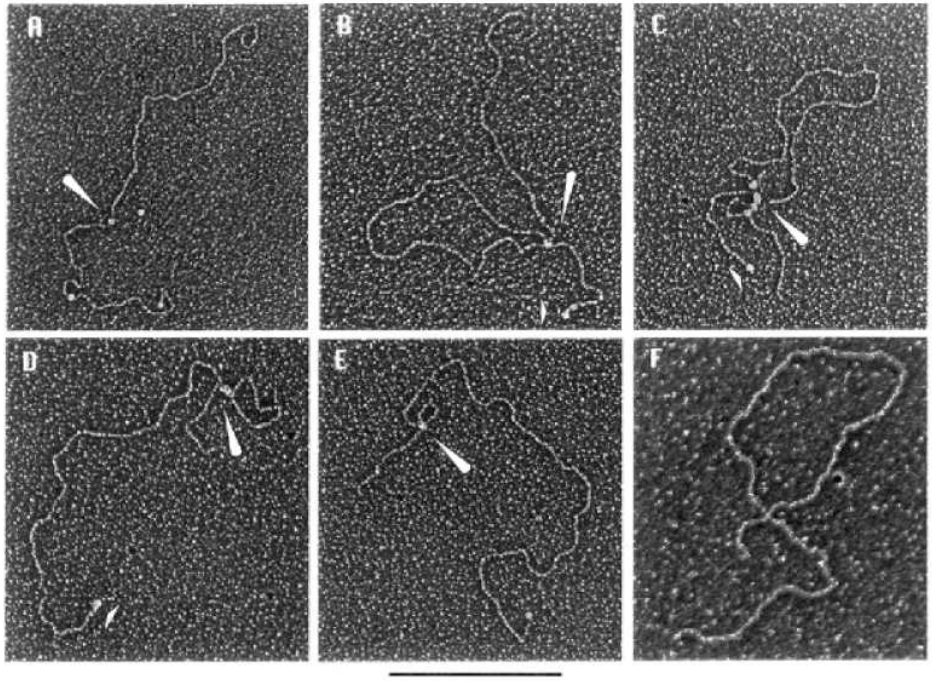


Figure 5. HMG I/Y forms DNA loops by cooperative binding and self-association. Electron micrographs of HMG I-mediated loops in streptavidin–biotin end-labeled β-globin DNA. (**A–E**) A sampling of the types of molecules seen at HMG I concentration of 15 ng for 50 ng DNA. Enhancer-containing β-globin DNA–HMG I complexes are shown in (A)–(C) and enhancerless β-globin DNA–HMG I complexes are shown in (D) and (E). (**F**) Demonstration of a protein-stabilized large loop in enhancer-containing β-globin DNA using the classic surface method employing a denatured film of cytochrome c protein. Arrows in (A)–(E) represent rHMG I particles bound to DNA and the short arrow in (D) represents the streptavidin–biotin tag. The size bar indicates 1 kb.

E85

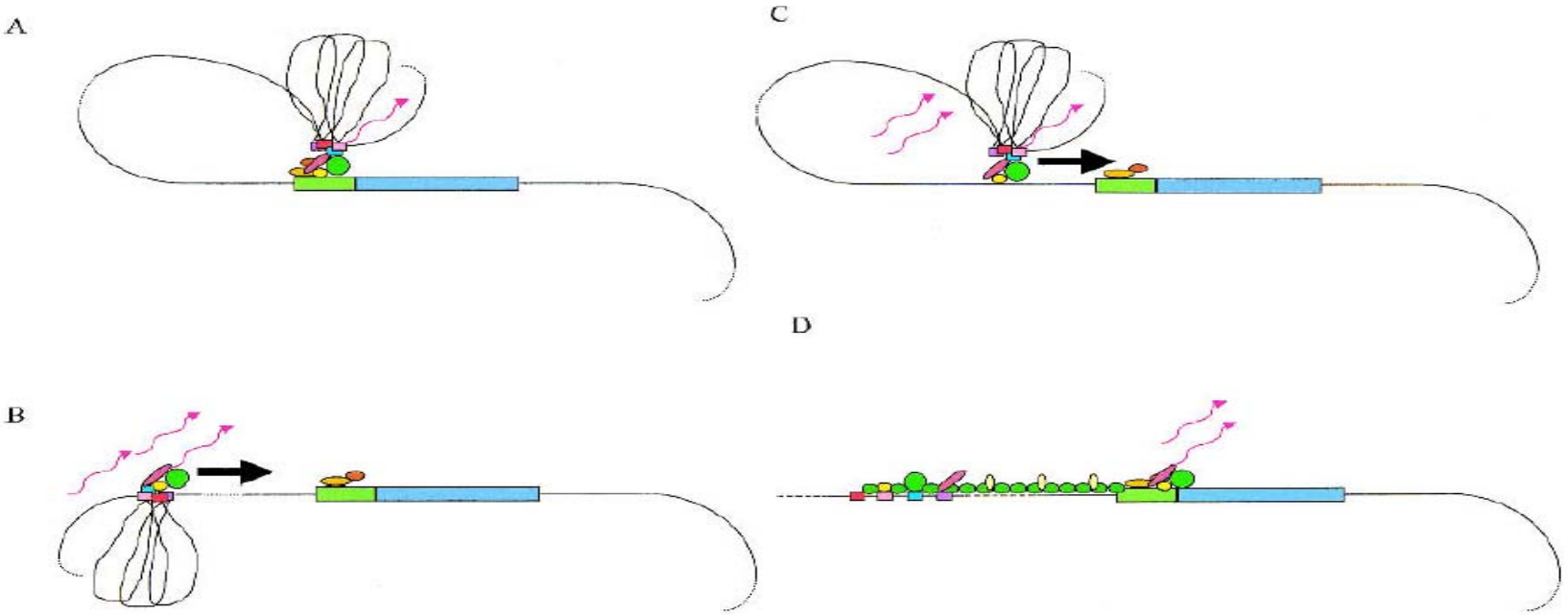


Figure 2. Models of LCR function. A globin gene is denoted as a green rectangular box with the promoter region indicated in a lighter green. Transcription factors are shown as colored ovals and circles. The four erythroid-specific hypersensitive site cores (HSs) are indicated by small colored boxes. The flanking DNA sequences of the HSs are depicted as loops between the HS cores. Transcripts are denoted by wavy arrows. (A) Looping model. Transcription factors bind to the LCR HSs and the gene promoter. The LCR directly interacts with the gene promoter by looping out the intervening DNA, thus forming an active transcription complex at the gene promoter. (B) Tracking model. Sequence-specific transcription factors bind to the LCR forming a complex that tracks down the DNA sequence until encountering transcription factors bound to the appropriate gene promoter, initiating high-level gene expression. (C) Facilitated tracking model. Aspects of both looping and tracking models are combined. Sequence-specific transcription factors bind the LCR, looping then occurs to deliver the bound transcription factors proximal to the gene promoter, followed by tracking until they encounter transcription factors bound to the appropriate gene promoter. (D) Linking model. Sequential binding of transcription factors along the DNA directs changes in chromatin conformation and defines the transcriptional domain. The transcription factors are linked to one another from the LCR to the gene promoter by non–DNA-binding proteins and chromatin modifiers (shown as colored ovals and circles).

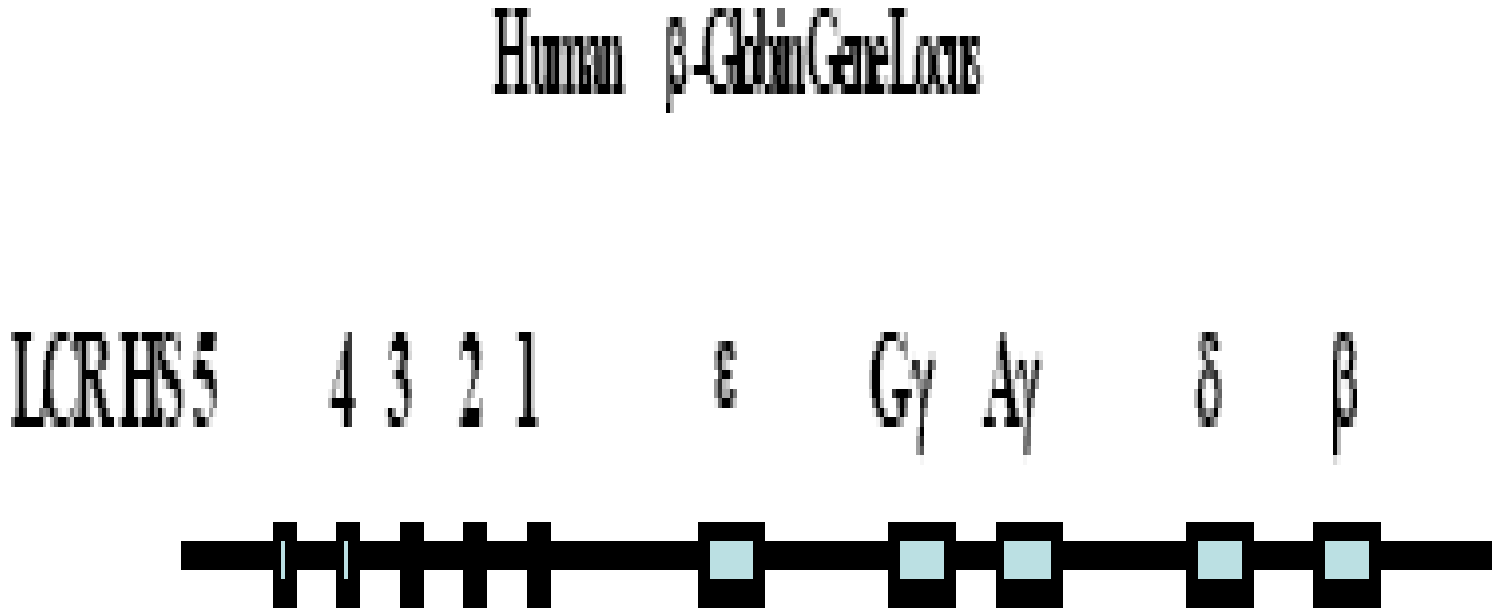
E85

Table III. Erythroid Transcription Factors

Protein	Protein family	DNA/protein interactions	Expression pattern	[a]KO/ Overexpression	Modulation	References
DRED (direct-repeat-erythroid-definitive protein)		Binds the direct repeat element in ε-globin promoter	MEL cells			(98)
EKLF (erythroid Krüppel-like factor)	Krüppel-like zinc finger	Binds to CACCC box element in β-globin promoter, LCR 5′HS3; interacts with GATA-1	Erythroid cell lineages throughout erythropoiesis	KO lethal by 14 dpc due to β-thalassemia	GATA-1 activates expression; acetylated *in vitro* by CPB an α p300; phosphorylated by CKII	(85,87,88,93, 95,124)
FKLF, FKLF-2 (fetal Krüppel-like factor)	EKLF/Ap1; TIEG (FLKF), BTEB1 (FKLF-2) subfamilies; Krüppel-like zinc finger	Binds to CACCC box of γ-globin promoter	Predominantly in erythroid cells, MEL cells	Overexpression of both activates γ- and ε-globin; FKLF-2 activates other erythroid promoters such as GATA-1 and glycophorin B		(96,97)
FOG (Friend-of-GATA)	Zinc finger	Interacts with GATA-1	Coexpressed with GATA-1 during embryonic development in erythroid and megakaryocytic cells	KO lethal during embryonic development due to failure of megakaryopoiesis and arrested erythropoiesis		(83,84)
GATA-1	GATA family	Binds (T/A)GATA(A/G) DNA sequence; interacts with FOG	Vertebrate erythroid, megakaryocyte and mast cell lineages	KO lethal at 10.5–11.5 dpc; overexpression inhibits terminal differentiation of erythroid cell lines	GATA-1/GATA-2 binding competition; acetylation by p300 and CBP causes conformational change and stimulates GATA-1 transcriptional activity *in vitro;* perturbs nucleosomes	(75,76,126, 144-149)
GATA-2	GATA family	GATA-element	Early erythroid cells, mast cells, megakaryocytes, pluripotent hematopoietic stem cells	KO causes lethality *in utero* due to anemia resulting from an early hematopoietic defect in all hematopoietic cell lineages; overexpression blocks hematopoiesis	GATA-1/GATA-2 binding competition; phosphorylated via the MAP kinase pathway	(81,82,150)
NF-E2	Basic leucine-zipper family of transcription factors; small subunits part of the Maf protein family	GCTGA(G/C)TCA (Maf-recognition element or MARE); this sequence includes the core AP-1-binding motif; large p45-NF-E2 subunit interacts with small p18 Maf subunits	p45 and Mafs expressed in the hematopoietic system; in addition, Mafs widely expressed during embryogenesis	p45 KO leads to fatal hemorrhage due to failure of megakaryocyte differentiation (lack of platelets); MafG KO lethal due to impaired megakaryopoiesis; MafK KO has no phenotype. MafG::MafK KO lethal; MafK overexpression catalyzes terminal erythroid differentiation	Phosphorylated via the MAP kinase pathway; acetylated by CBP/p300; chromatin remodeling at 5′HS2 and ε-globin promoter	(64,65,114, 151-157)
COUP-TFII (NF-E3)	Nuclear orphan receptor	Binds to the direct repeat elements in ε- and γ-globin promoters	Erythroid cells; peak of expression coincides with γ- to β-globin switch	KO lethal at 10.5 dpc		(99)
SSP (stage-selector protein)		Binds the stage selector element (SSE) in ε- and γ-globin promoters and LCR 5′HS2 and 3; contains NF-E4	Hematopoietic cells of fetal liver, cord blood, bond marrow; K562 and HEL cell lines	Overexpression in K562 cells induces γ-globin gene expression		(158–161)
YY1 (NF-E1)	GLI-Krüppel-like family	Developmental repressor of the human ε-globin gene together with GATA-1	Ubiquitous			(16,158)

[a] KO: knockout mutation

E93



E99

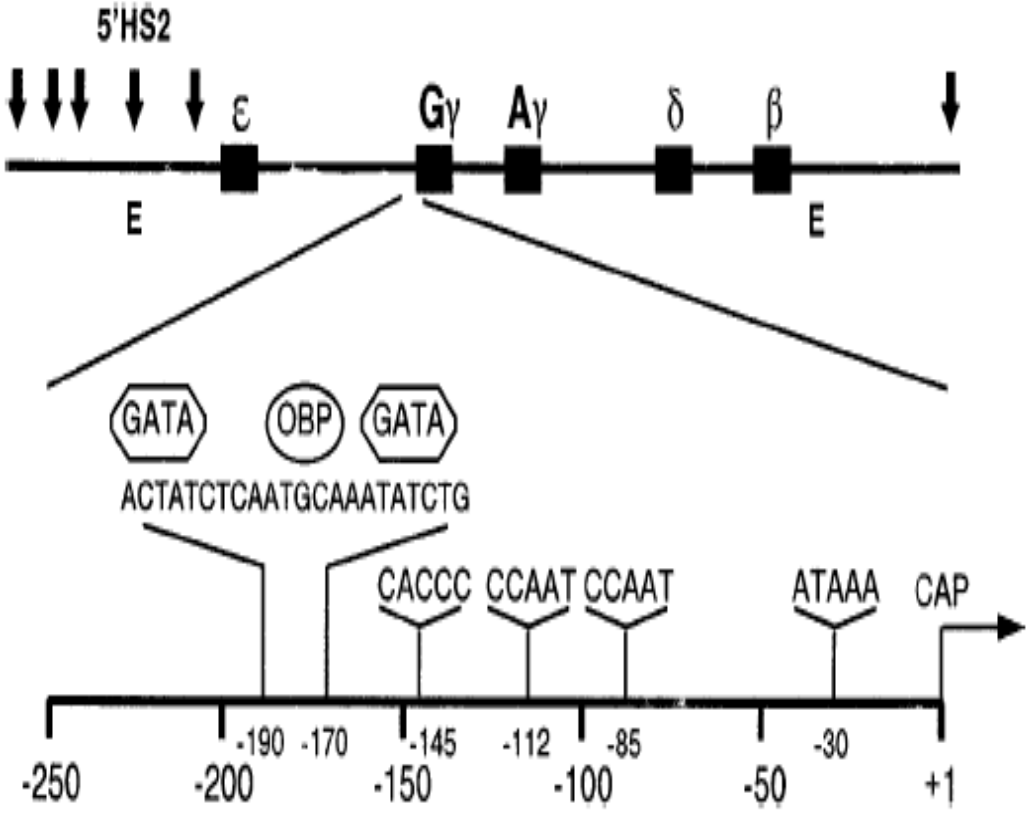


Fig 1. Diagram of the human β-globin gene domain showing the arrangement of the five β-globin genes, the position of the six identified DNase I hypersensitive (HS) sites (arrows), and the location of the 5'HS2 and 3'β enhancer elements (E). The expanded view of the G γ gene promoter region shows the recognition sites for the DNA-binding proteins GATA-1 and OBP.

E100

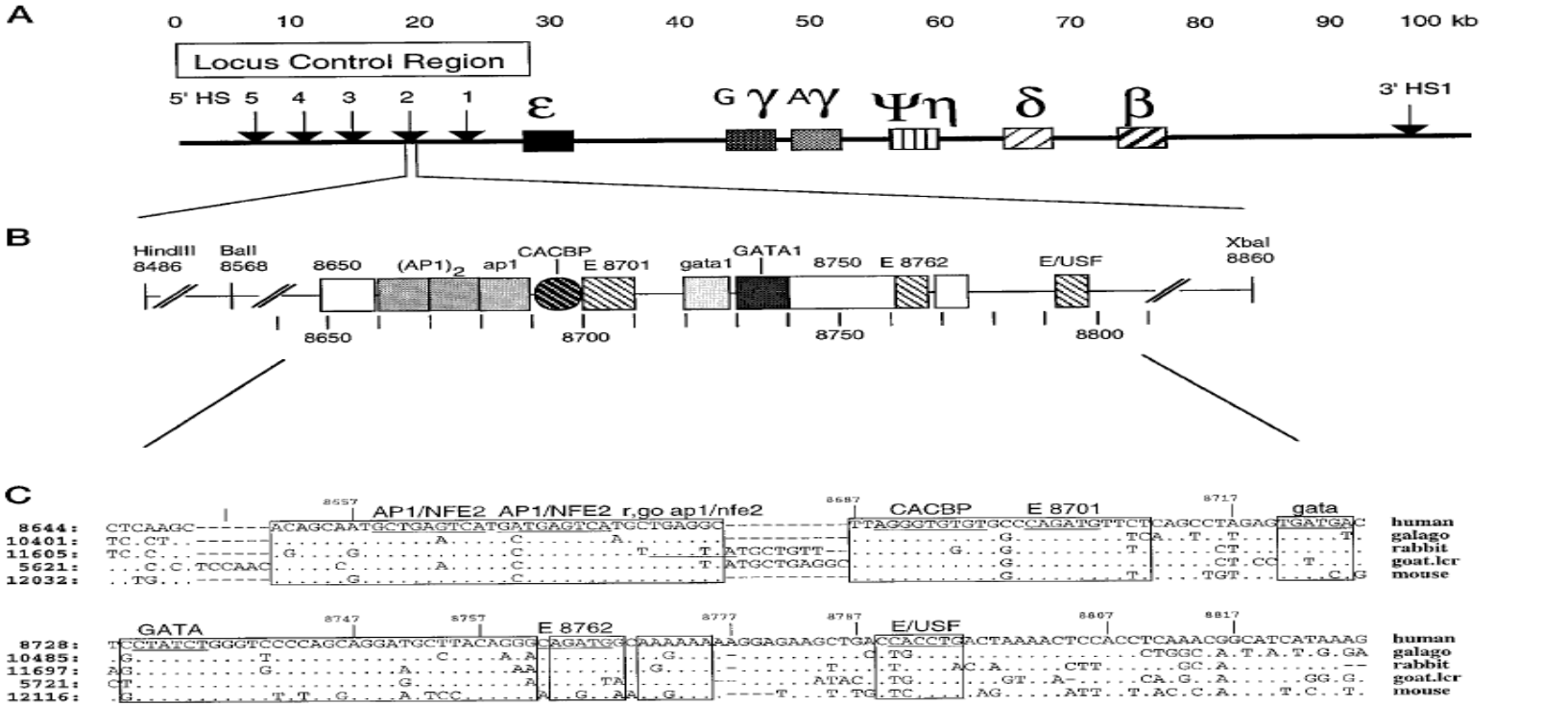


FIG. 1. **Important segments in HS2 of the β-LCR.** *A*, a map of the human β-globin gene cluster, showing the β-LCR and developmentally stable DNase HSs as *arrows*. *B*, a diagram of notable sites within the core of HS2. Binding sites for proteins are labeled and given a distinctive fill, and *open boxes* are conserved sequences with no currently identified binding protein. Nucleotide positions refer to GenBank file HUMHBB. *C*, a simultaneous alignment of DNA sequences of HS2 from human, galago, rabbit, goat, and mouse. *Boxes* are drawn around runs of consecutive conserved columns that are not contained in a longer such run. An alignment column is called "conserved" if it is contained in a run of 6 consecutive columns that have a "model row" of length 6, such that each row (of length 6) has at most one mismatch with the model row.[3] This criterion was developed to find blocks that are likely to represent protein-binding sites while allowing one mismatch per species. Binding sites for proteins are *underlined* and labeled.