DNA-protein interaction at erythroid important regulatory elements of MEL cells by *in vivo* footprinting

JI Xinjun, LIU Depei, XU Dongdong, LI Lei, WANG Jing & LIANG Zhiquan

National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, Faculty of Basic Medical Sciences, Peking Union Medical College, Beijing 100005, China (e-mail:guaji@yahoo.com)

Abstract Using ligation-mediated PCR method to study the status of DNA-protein interaction at hypersensitive site 2 of locus control Region and β^{maj} promoter of MEL cell line before and after induction, MEL cell has been cultured and induced to differentiation by Hemin and DMSO, then the live cells have been treated with dimethyl sulfate. Ligation mediated PCR has been carried out following the chemical cleavage. The results demonstrate that before and after induction, the status of DNA-protein interaction at both hypersensitive site 2 and β^{maj} promoter change significantly, indicating that distal regulatory elements (locus control region, hypersensitive sites) as well as proximal regulatory elements (promoter, enhancer) of *β-globin* gene cluster participate in the

regulation of developmental specificity.

Keywords: *β-globin* gene, HS, LM-PCR, *in vivo* footprinting, MEL cell line.

Gene expression is regulated by the complex interaction of *cis*-acting elements and *trans*-acting factors. The correct expression of *β-globin* gene cluster mainly depends on two kinds of regulatory elements: the locus control region (LCR) which is located far upstream of this cluster[1] and proximal regulatory elements such as promoter and enhancer. LCR is composed of 4 DNase I hypersensitive sites (HS) and individual HS activity of LCR is defined to its core sequence. The HS core sequences of the LCR and the promoter of individual *globin* genes have quite a few binding sites for erythroid-specific as well as ubiquitous proteins. The involvement of specific DNA-protein interaction in the formation of HSs, in the assembly of basal transcription apparatus and in the contact between LCR and downstream gene promoter is suggested by a large body of experimental evidence[2]. So the study of DNA-protein interaction is important and significant to elucidate the mechanism for LCR action and *globin* gene switching.

In this note, we used ligation-mediated PCR (LM-PCR) and *in vivo* footprinting to study DNA-protein interaction at *β-globin* gene promoter and HS2 of LCR of MEL cells. The results demonstrated that before and after induction, the status of DNA-protein interaction at both hypersensitive site 2 and β^{maj} promoter changed significantly, and indicated that distal regulatory elements (locus control region, hypersensitive sites) as well as proximal regulatory elements (promoter, enhancer) of *β-globin* gene cluster participate in the regulation of developmental specificity.

1 Materials and methods

(i) Cell culture. MEL cells were grown in DMEM medium containing 10% fetal bovine serum supplemented with penicillin (100 units/mL) and streptomycin (100 units/mL) at 37℃ with 5%CO_2. Hemin (100 μmol/L) and DMSO (1.5%) were used to induce cell differentiation for 3 d.

(ii) Preparation of cell suspension. Collect MEL cells by centrifuge at room temperature. Discard the supernatant and resuspend MEL cells in 1 mL DMEM medium without serum, adjust cell concentration to 5×10^7/mL.

(iii) *In vivo* methylation. The suspension cells were pre-warmed at 37℃ for 10 min, then added with 10 μL 10% dimethyl sulfate (DMS), mixed gently and incubated at 37℃ for 2 min. The cells were immediately transferred to an ice-cold 49-mL PBS aliquot, mixed by gentle inversion, and centrifuged 5 min at 1 500×*g*, 4℃. Discard the supernatant. Repeat the PBS washing twice. Discard the supernatant again, then add 4 mL lysis solution to the tube, digest overnight[3].

(iv) Extraction of DNA. Extraction of genomic DNA of *in vivo* or *in vitro* methylation was performed by routine methods[3]. Finally the DNA samples were resuspended in TE buffer.

(v) *In vitro* methylation. *In vitro* methylation of control protein-free DNA was performed according to the method of Mueller and Wold[3].

(vi) Base-specific DNA cleavage. Guanine-specific piperidine cleavage of *in vivo* and *in vitro* methylated DNA was done according to the method of Mueller and Wold[3].

(vii) LM-PCR. LM-PCR genomic footprinting was performed essentially as described by Garrity and Wold[4], the primers used were the same as Reddy's[5].

(viii) Labeling reaction and sequencing PAGE. Finally, the PCR products were used as substrate in 2 cycles of labeling reaction according to Garrity and Wold[4]. The labeled samples were subjected to 8 mol/L urea-6%PAGE sequencing gels at 35 W for 3 h. Gels were dried and exposed to X-ray film at −70℃ for 2—3 d.

(ix) Result statistics. After sequencing PAGE, the qualified gels were exposed to phosphate screen and scanned by PhosphorImager (Molecular Dynamics). The percent protection or hypersensitivity was determined by comparison of individual G residues reactivity to the reactivity of corresponding G residue on the naked DNA lane. The reactivity of *in vitro* methylation control G residues was defined as 100%. G residues that were protected or 10%—40% hypersensitive were

classified as weak hypersensitivity or weak protection. G residues having 40% or more reactivity were considered as strong hypersensitivity or strong protection. In order to verify the reproducibility of *in vivo* footprints, every set of reaction, sequencing PAGE and PhosphorImager scanning was repeated at least twice under similar conditions.

2 Results

We used LM-PCR and *in vivo* footprinting to study the status of DNA-protein interaction at both HS2 of LCR and β^{maj} promoter. The structure of mouse *β-globin* gene cluster, the arrays of *cis*-acting elements at HS2 and β^{maj} promoter were presented in fig. 1.

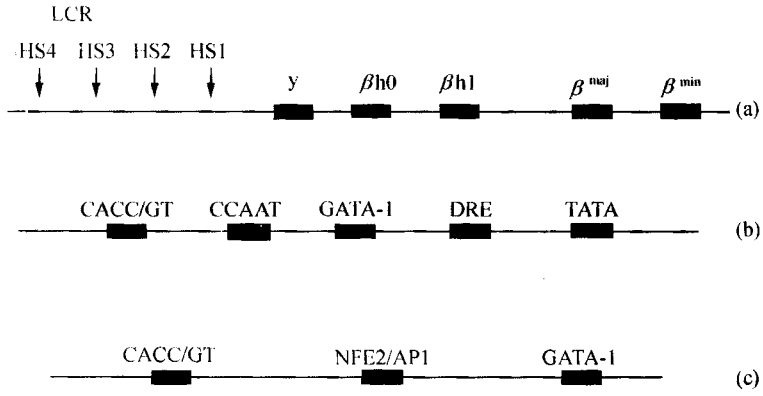

Fig. 1. The structure of mouse *β-globin* gene cluster, HS2 and promoter. (a) Mouse *β-globin* gene cluster; (b) the sequence of promoter and protein binding sites; (c) the sequence of HS2 and protein binding sites.

(i) *In vivo* footprints of HS2. The mouse HS2, just like human's, consists of CACC/GT, NFE2/AP1 and GATA elements, which are the binding sites of Sp1(or EKLF, BKLF/TEF2)[6,7], NFE2 (or AP1, Nrf2, LCRF1/Nrf1)[8] and GATA-1[9], respectively. The interaction of these *cis*-acting elements and relevant *trans*-acting factor is required for LCR formation and LCR function. In order to facilitate the description, "G" combining with numbers up (coding strand) or down (non-coding strand) was used to refer to individual "G" residue.

CACC/GT element. The sequence of this element and flanking region are as follows:

```
                                12   13 14 15
TCTCTAC TCCCCAC CCTGTGGGTGTGTTCA
AGAGATGAGGGGTGGGACACCCACACAAGT
 1  2   3  4 5 6 7  8 9 10                 11
```

Results of *in vivo* footprints of the coding strand of this element showed that G residues have the same reactivity before and after induction, that is to say, G14 and G15 were protected (fig. 2(a)). On the other hand, *in vivo* footprinting pattern of the non-coding strand exhibited some differences between these two groups: G6 was hypersensitive, while G4, G5, G7, G8, G9, G10 and G11 were protected after induction. Before induction, however, G1, G2, G6, G7 and G10 were hypersensitive, while G8 and G11 were protected (fig. 2(b)).

The result described above showed that G1, G2, G7 and G10 were hypersensitive before induction. However, G7 and G10 were protected with G4, G5 and G10 being additionally protected after induction. The results indicated that DNA-protein interaction at this element was different.

NFE2/AP1 element. DNA sequence corresponding to this element is as follows:

```
9      10  11 12  13 14     15  16 17     18  19
GCACAGCAGTGCTGAGTCATGCTGAGTCATGCTG
CGTGTCGTCACGACTCAGTACGACTCAGTACGAC
 1  2  3      4       5     6       7   8
```

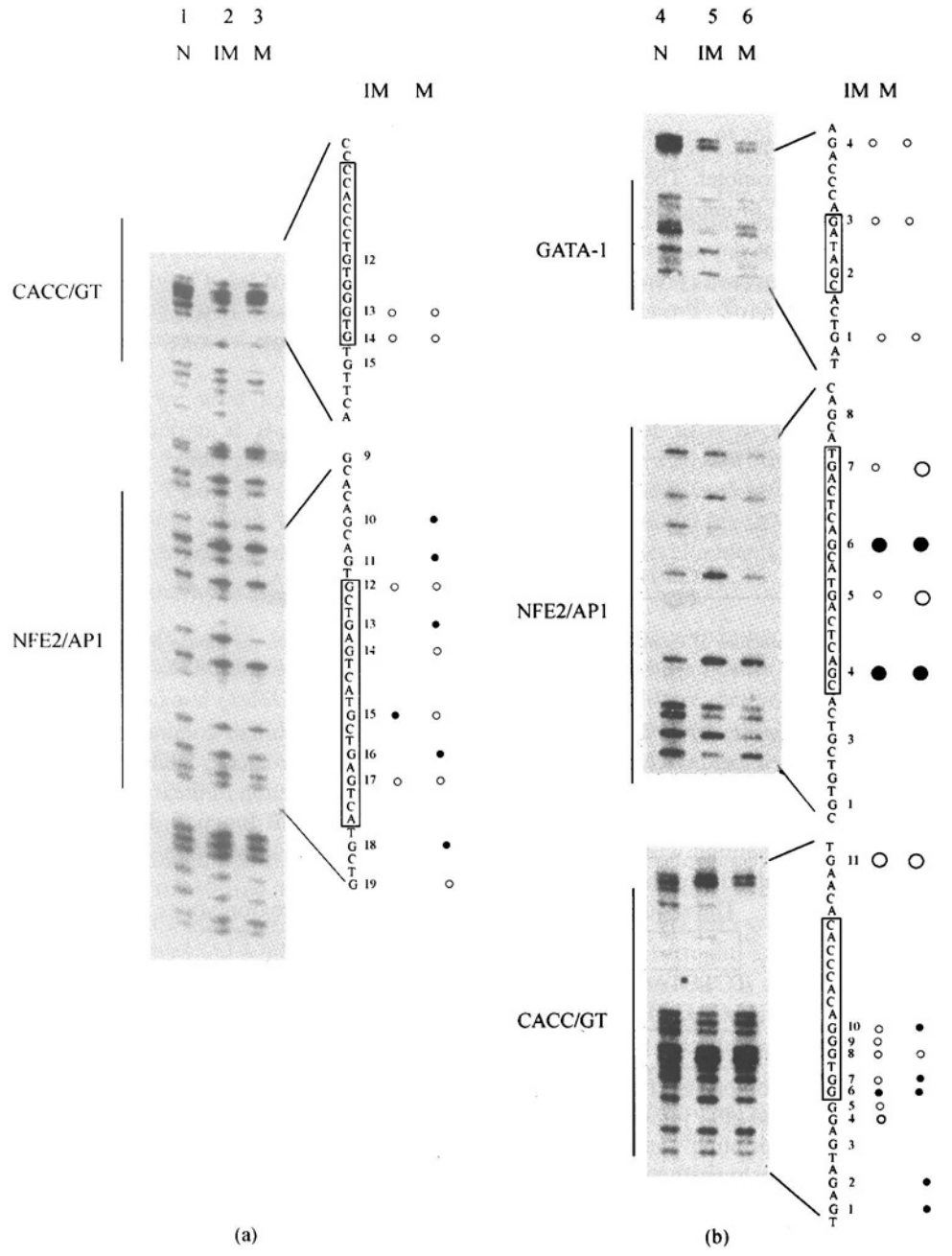

Fig. 2. Autoradiographs of *in vivo* footprints of LCR-HS2 of *β-globin* gene cluster of MEL cell. (a) Coding strand; (b) Non-coding strand. The vertical lines on the left represent different motifs, the corresponding DNA sequence is on the right. Open and solid circles denote protection and hypersensitivity, respectively. Sizes of the circles represent the relative extents of G reactivity. Lanes 1 and 4, Control sample of *in vitro* methylation (N); lanes 2 and 5, induced MEL cell sample of *in vivo* methylation (IM); lanes 3 and 6, uninduced MEL cell sample of *in vivo* methylation (M).

The results of *in vivo* footprints at this site indicated that DNA-protein interaction mode at this element had changed in these two groups. Results of coding strand showed that, before induction, G10, G11, G13 G16 and G18 were hypersensitive, while G12, G14, G15, G17 and G19 were protected.

However, after induction, only G12 and G17 were protected with G15 being hypersensitive (fig. 2(a)). Results of non-coding strand showed that the reactivities of G residues were basically the same in the two groups, that is to say, G5 and G7 were protected and G4 and G6 were strongly hypersensitive (fig. 2(b)).

The above result demonstrated that although G12 and G17 were protected before and after induction, additional 5 G residues were hypersensitive before induction. More attention should be given to G15: it was protected before induction; however, it was hypersensitive after induction. The result inferred that DNA-protein interaction at this element changed in the two groups of cells.

GATA element. The following sequence is for GATA element:

TAGTCACGATAGACCCAGA (non-coding strand)

1	2	3	4

The reactivity of G residues at this element was basically the same; that is to say, G1, G3 and G4 were protected in these two groups (fig. 2(b)). This result demonstrated that the mode of DNA-protein interaction at this element did not change before and after induction.

(ii) *In vivo* footprints of promoter. The mouse β^{maj} promoter consists of five *cis*-acting elements: CACC/GT, CCAAT, GATA, DRE, and TATA. These elements are necessary for its appropriate transcription activation. CACC/GT, CCAAT, GATA, and TATA are recognized by Sp1(or EKLF, BKLF/TEF2)[6,7], NFE3[10] or other factors, GATA-1[9] TFIID[11], respectively. The two DRE elements are very crucial to the induction of *globin* gene expression, but the factor which recognizes and binds to this element is not yet identified.

CACC/GT element. The following is the sequence for this element:

1	2					3	4

TAGAGCCACACCCTGGT

A TCTCGGTGTGGGACCA

5	6	7	8	9	10

The reactivity of G residues of MEL cells at this element was different before and after induction. The result of coding strand showed that G3 was hypersensitive and G2, G4 were protected before induction; G3 and G4 were hypersensitive after induction (fig. 3(a)). The results of non-coding strand showed the pattern of *in vivo* footprints before and after induction was basically the same; that is to say, G5, G6, G9 and G10 were hypersensitive (fig. 3(b)).

The result at this element demonstrated that although the reactivity of G residues from the non-coding strand was the same, the result of coding strand indicated that the reactivity was different before and after induction; that is to say, G2 was additionally protected before induction, G4 is also protected before induction, however, G4 was hypersensitive after induction.

CCAAT element. The DNA sequence of this element and its flanking region are as follows:

1	2	3			4

AAGGGCCAATCTGC

TTCCCGGTTAGACG

5	6		7

The result of the coding strand showed that G1, G2, G3 and G4 were hypersensitive before induction, while only G1 was hypersensitive with G2, G3, G4 having no reactivity after induction (fig.3(a)). The *in vivo* footprints of non-coding strand showed that G5 and G7 were protected before induction, while G6 was additionally protected (fig. 3(b)). The above result indicated that different complex formed at this element before and after induction.

GATA-DRE elements.

1	2		3	4	5	6	7		8	9	10		11	12	13		14

CAGGATAGAGAGGGCAGGAGCCAGGGCAGAGCAT

GTCCTATC TCT CCCGTCCTCGG TCCC G TCTCGTA

15				16			17	18		19		20

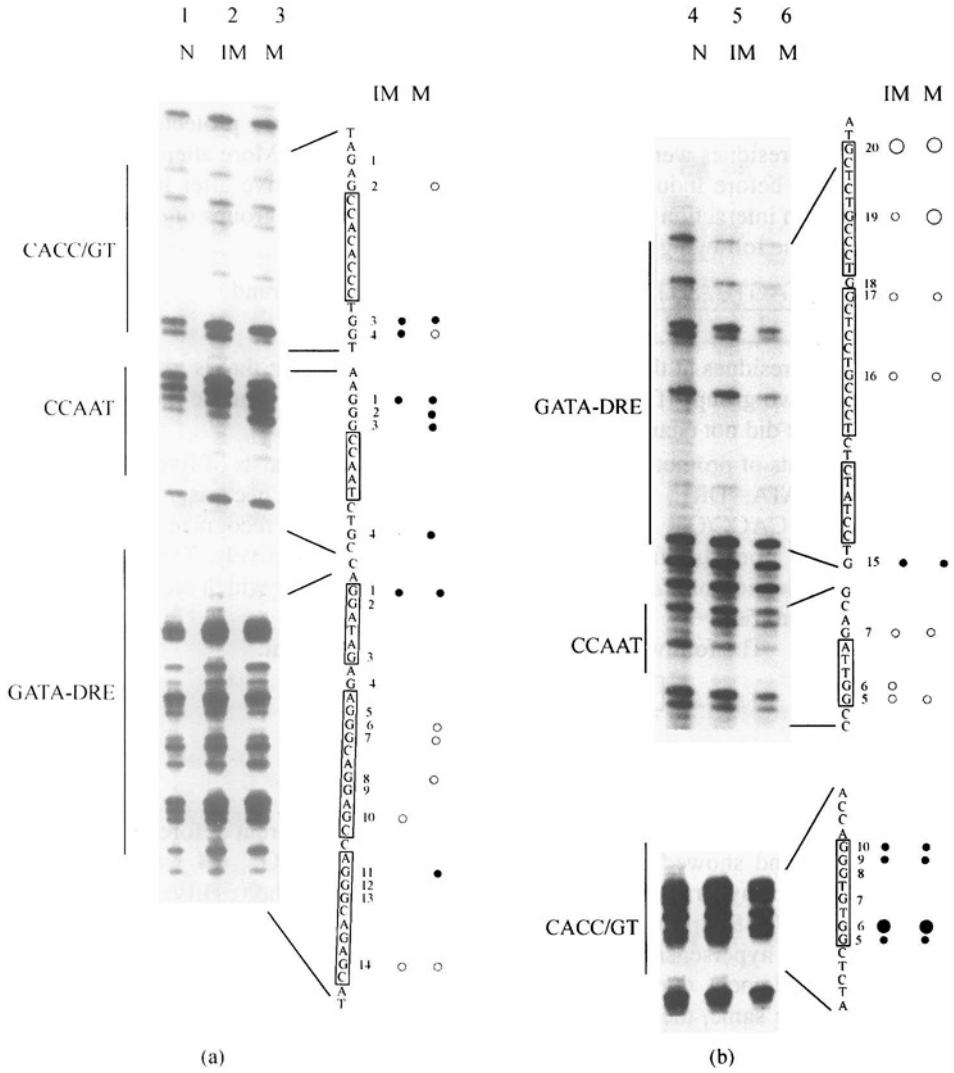

Fig. 3. Autoradiographs of *in vivo* footprints of *β-globin* gene promoter of MEL cells. (a) Coding strand; (b) non-coding strand. The vertical lines on the left represent different motifs, the corresponding DNA sequence is on the right. Open and solid circles denote protection and hypersensitivity, respectively. Sizes of the circles represent the relative extents of G reactivity. Lanes 1 and 4, Control sample of *in vitro* methylation (N); lanes 2 and 5, induced MEL cell sample of *in vivo* methylation (IM); lanes 3 and 6, uninduced MEL cell sample of *in vivo* methylation (M).

The result of *in vivo* footprints showed that G6, G7, G8 and G14 were protected and G1 and G11 were hypersensitive before induction, while only G1 was hypersensitive and G10, G14 were protected after induction (fig. 3(a)). The results of non-coding strand exhibited that the pattern of *in vivo* footprints was basically the same in the two groups; that is to say, G16, G17, G19 and G20 were protected and G15 was hypersensitive (fig. 3(b)).

The reactivity of G residues at this element was different in these two groups, and this is mainly manifested from the coding strand. Except G1 was hypersensitive and G14 was protected in the two groups, G11 was hypersensitive and G6, G7, G8 were protected before induction, while only G10 was additionally protected after induction.

3 Discussion

The study of LCR has greatly deepened the insight into the regulation of *globin* gene expression.

At present, it is generally agreed that individual HS element interacts with one another to form a functional unit (holocomplex) which flip-flops via looping among *globin* genes[12]. At present, however, there is no universal agreement as to the role of LCR in the regulation of the *globin* gene switching during ontogenesis. According to one hypothesis, tissue- and stage-specificity may be conferred only by gene-proximal sequence (promoter and enhancer) of individual genes. The other hypothesis thinks that the switching process is determined by the combinatory interaction of gene-proximal regulatory elements and LCR[13].

In present study, we used *in vivo* footprinting and LM-PCR to study the DNA-protein interaction mode at HS2 of LCR and β^{maj} promoter of MEL cells. MEL cells are arrested at a later stage of erythroid development. Upon induction with DMSO, they are subjected to termination differentiation. The promoter region is crucial to the activation of β^{maj} *globin* gene, at the same time, the transcription of *β-globin* gene cluster is regulated by the upstream LCR[12]. The result described here showed that not only the DNA- protein interaction pattern at CACC/GT, CCAAT and DRE elements of promoter, but also the interaction pattern at CACC/GT and NFE2/AP1 elements of HS2 changed before and after induction. This is different from Reddy's[5].

The expression of *β-globin* gene in transgenic mice with LCR-*β-globin* gene construct loses developmental specificity, individual HSs within the β-LCR contributes preferentially to the developmental regulation of specific *globin* genes[14]. The disruption of HS3 in YAC-transgenic mice differentially influences the expression of *γ-globin* gene during development, it means that the interaction mode between LCR (or individual HSs) and downstream genes is different at different developmental stages[15]. Human HS3 does not exhibit *in vivo* footprints in K562 cells[16,17], it is footprinted in HU11 cells only after being induced with DMSO, so it seems that HS3 is developmentally regulated[18]. Although human HS2 has *in vivo* footprints throughout ontogenesis (K562 cells and human erythroblast)[16,18—20], the status of *in vivo* footprints in K562 cells is different before and after induction[16]. The *in vivo* footprints of CACC/GT element of HSs in HU11 cell line which expresses *γ globin* gene are different from that in HU11 cell line which expresses *β globin* gene[18].

Based on the result described in this note and previous researches in this field, we put forward the following theory: the individual HSs of LCR form a holocomplex and interact with downstream gene promoters; however, it is possible that different holocomplexes interact with different gene promoter regions, that is to say, LCR may form more than one kind of holocomplex. At different developmental stages, the difference of variety and quantity of *trans*-acting factors will result in the form of different complexes. Alternatively, LCR may indeed form only one kind of holocomplex; however, when it interacts with individual gene promoter, the individual HS of LCR has different action manner. Because the holcomplex is very large, it is impossible to interact with the promoter region by the whole LCR holocomplex, the feasibility is that only part of the LCR holocomplex (HS) is suitable to interact with one downstream gene promoter, the suitable part of the LCR (HS) may be different at different developmental stages. It can be concluded that the DNA-protein interaction could be affected at any circumstances (mentioned above). At last, the change of *trans*-acting factor environment during development will necessarily result in the change of DNA-protein interaction at promoter. When LCR holocomplex interacts with downstream gene via a looping mediated by protein-protein interaction, because of the alternate influence between the two DNA-protein complexes, we can infer that the status of DNA-protein interaction at any regulatory elements will also change.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant No. 39470390) and the Science Foundation for the Chinese Outstanding Youth (Grant No. 3925006) and the Key Programs of the National Natural Science Foundation of China (Grant No. 39893320).

References

1. Grosveld, F., van Assendelft, G. B., Greaves, D. R. et al., Position-independent, high-level expression of the human *β-globin* gene in transgenic mice, Cell, 1987, 51: 975.
2. Fraser, P., Pruzina, S., Antoniou, M. et al., Each hypersensitive site of the human *β-globin* locus control region confers a different developmental pattern of expression on the globins, Genes & Dev., 1993, 7: 106.
3. Frederick, M. A., Short Protocol in Molecular Biology, Chapter15.5, New York: John Wiley & Sons, Inc., 1995.

4. Garrity, D. A., Wold, B. J., Effects of different DNA polymerases in ligation-mediated PCR: Enhanced genomic sequencing and *in vivo* footprinting, Proc. Natl. Acad. Sci. USA, 1992, 89: 1021.
5. Reddy, P. M. S., Shen, C. K. J., Erythroid differentiation of mouse erythroleukemia cells results in reorganization of protein-DNA complexes in the mouse β^{maj} *globin* gene promoter but not its distal enhancer, Mol. Cell. Biol., 1993, 13: 1093.
6. Asano, H., Stamatoyannopoulos, G., Activation of *β-globin* promoter by erythroid kruppel-like factor, Mol. Cell. Biol., 1998, 18: 102.
7. Kadonaga, J. T., Carner, K. R., Masiara, F. R. et al., Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain, Cell, 1987, 51: 1079.
8. Andrews, N. C., Erdjument-Bromage, H., Davidson, M. B. et al., Erythroid transcription factor NFE2 is a haemetopoietic-specific basic-leucine zipper protein, Nature, 1993, 362: 722.
9. Trainor, C. D., Omichinski, J. G., Vandergon, T. L. et al., A palindromic regulatory site within vertebrate GATA-1 promoters requires both zinc fingers of the GATA-1 DNA binding domain for high-affinity interaction, Mol. Cell. Biol., 1996, 16: 2238.
10. Ronchi, A. E., Bottardi, S., Mazzucchelli, C. et al., Differential binding of the NFE3 and CP1/NFY transcription factors to the human gamma- and epsilon-globin CCAAT boxes, J. Biol. Chem., 1995, 270: 21934.
11. Evans, T., Felsenfeld, G., Reitman, M., Control of globin gene transcription, Annu. Rev. Cell. Biol., 1990, 6: 95.
12. Wijgerde, M., Grosveld, F., Fraser, P., Transcription complex stability and chromatin dynamics *in vivo*, Nature, 1995, 377: 209.
13. Baron, M. H., Transcriptional control of globin gene switching during vertebrate development, BBA, 1997, 1351: 51.
14. Wood ,W. G., The complexities of beta globin gene regulation, Trends Genet, 1996, 12: 204.
15. Navas, P. A., Peterson, K. R., Li, Q. et al., Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion, Mol. Cell Biol., 1998, 18: 4188.
16. Ikuta, T., Kan, Y. W., *In vivo* protein-DNA interactions at the beta-globin gene locus, Proc. Natl. Acad. Sci. USA, 1991, 88: 10188.
17. Strauss, E. C., Orkin, S. H., *In vivo* protein-DNA interactions at hypersensitive site 3 of the human beta-globin locus control region, Proc. Natl. Acad. Sci. USA, 1992, 89: 5809.
18. Ikuta, T., Papayannopoulou, T., Stamatoyannopoulos, G. et al., Globin gene switching, *in vivo* protein-DNA interactions of the human beta-globin locus in erythroid cells expressing the fetal or the adult globin gene program, J. Biol. Chem., 1996, 271: 14082.
19. Reddy, P. M. S., Stamatoyannopoulos, G., Papayannopoulou, T. et al., Genomic footprinting and sequencing of human beta-globin locus, Tissue specificity and cell line artifact, J. Biol. Chem., 1994, 269: 8287.
20. Reddy, P. M. S., Shen, C. K. J., Protein-DNA interactions *in vivo* of an erythroid-specific, human beta-globin locus enhancer, Proc. Natl. Acad. Sci. USA, 1991, 88: 8676.

(Received September 6, 1999)