ISSN 0006-2979, Biochemistry (Moscow), 2007, Vol. 72, No. 5, pp. 467-470. © Pleiades Publishing, Ltd., 2007.
Original Russian Text © S. V. Razin, E. S. Ioudinkova, 2007, published in Biokhimiya, 2007, Vol. 72, No. 5, pp. 581-585.

MINI-REVIEW

# Mechanisms Controlling Activation of the α-Globin Gene Domain in Chicken Erythroid Cells

**S. V. Razin[1,2]* and E. S. Ioudinkova[1]**

[1]*Institute of Gene Biology, Russian Academy of Sciences, ul. Vavilova 34/5, 119334 Moscow, Russia; E-mail: sergey.v.razin@usa.net*
[2]*Biological Faculty, Lomonosov Moscow State University, 119992 Moscow, Russia*

Received January 25, 2007
Revision received January 30, 2007

**Abstract**—In this review we consider the organization of the chicken α-globin gene domain and mechanisms regulating the activity of this tissue-specific gene domain located in a potentially active (characterized by an increased sensitivity to nucleases) chromatin configuration in cells of all lineages. Both regulatory mechanisms ensuring repression of α-globin genes in non-erythroid cells and mechanisms responsible for activation of transcription of these genes during erythroid cell differentiation are discussed. The analysis of the structure–function organization of the chicken α-globin gene domain presented in this review is based mainly on the authors' own results obtained over the last 20 years. On discussing the hypotheses explaining the mechanisms controlling the functional activity of chicken α-globin gene domain, data obtained in studies of α-globin gene domains of other vertebrates are also analyzed.

**DOI**: 10.1134/S000629790705001X

*Key words*: chromatin, genomic domains, α-globin genes, locus control region, insulators

A genome region containing a gene (or a group of genes) together with a complex of regulatory elements controlling functioning of the gene(s) is called a genomic domain. Genomic domains can be of open and closed types.

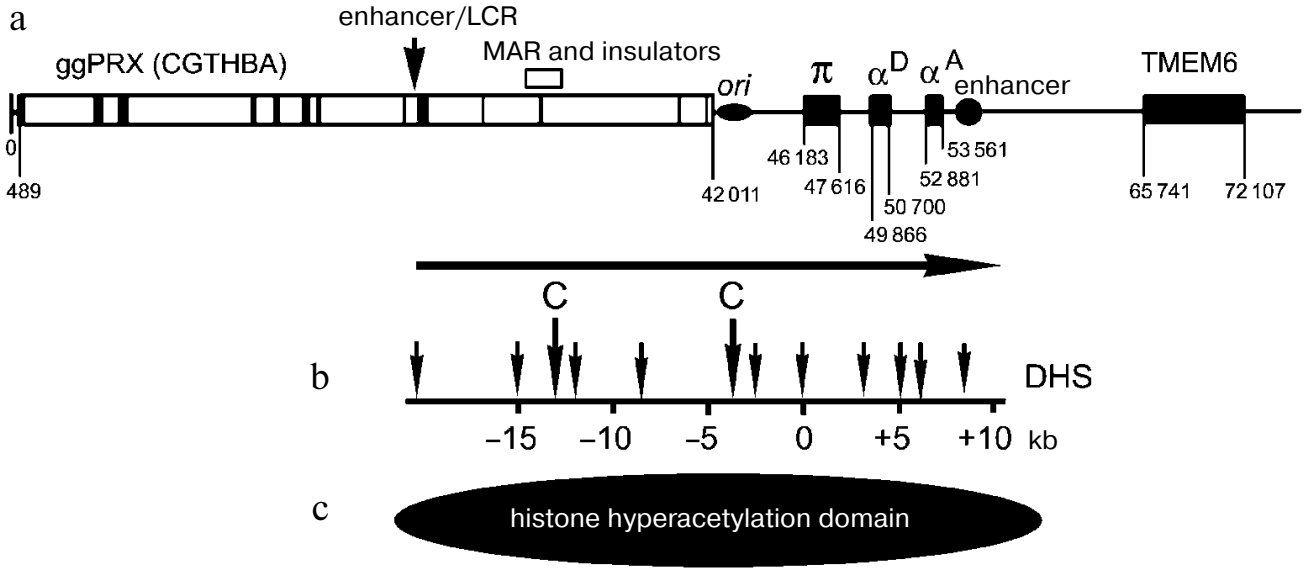
The open type domains include tissue-specific gene domains with an increased sensitivity to DNase I in cells with the active corresponding genes. Such domains are exemplified by domains of β-globin genes of vertebrates and the chicken ovalbumin gene domain. Boundaries of a closed type domains coincide with those of a region with increased sensitivity to DNase I. The open type domains are located in the gene-rich genomic regions, have no distinct boundaries, and are often overlapped with domains of other genes. Moreover, the open type domains containing tissue-specific genes are characterized by increased sensitivity to DNase I in cells of all lineages (also including cells without expression of the domain residing tissue-specific genes). Specific features of organization of the different type genomic domains have been discussed in detail in our preceding reviews [1-3].

*Abbreviations*: MAR) matrix-associated region; kb) thousand base pairs.

* To whom correspondence should be addressed.

The chicken α-globin gene domain is an open type domain. The figure presents a scheme of the organization of this domain. The α-globin gene cluster proper is ~8 kb in length. Immediately before this cluster, the housekeeping gene ggPRX with yet unknown functions is located [4]. The major regulatory element controlling the functional activity of the α-globin gene domain is located in the fifth intron of this gene [5]. Between this regulatory element and the globin genes proper, the MAR-element (Matrix Associated Region is a DNA fragment capable of preferential binding *in vitro* with isolated nuclear matrix) [6], which acts as an insulator, is located [7]. Insulators are special genomic elements that limit the region of enhancer effect and prevent propagation of various signals along a chromatin fibril [3, 8, 9].

A reasonable question arises—is it possible to at least approximately determine the length and boundaries of the α-globin gene domain if this domain has no differentiated sensitivity to DNase I in erythroid and non-erythroid cells? This question may well be answered in the affirmative. There are two groups of experimental data that allow us to say that the functional domain of α-globin genes is significantly larger than the gene cluster itself. First of all, analysis of distribution of erythroid-specific regions of hypersensitivity to DNase I allows us to




Organization of the chicken α-globin gene domain. a) The domain scheme presenting locations of the genes, enhancers, and insulators. The distances in the scheme are given in base pairs according to the nucleotide sequence deposited in the GeneBank database (No. AY016020). The horizontal black arrow under the scheme indicates a genome-wide transcript [11]. The origin of the replication region (*ori*) corresponds to data presented in works [34, 35]. The gene TMEM6 is an open reading frame presumably encoding a transmembrane protein [4]. b) Distribution of regions hypersensitive to DNase I (DHS) in the long genomic stretch containing the α-globin gene cluster. The short arrows indicate erythroid-specific regions of hypersensitivity, the long arrows with the letter "C" indicate permanent regions of hypersensitivity. c) Position of the histone hyperacetylation domain arising in pre-erythroblasts

approximately map boundaries of the erythroid-specific genomic domain containing the chicken α-globin gene cluster [9] (figure, panel (b)). Moreover, the same genomic domain in erythroid cells can also be identified on analyzing the site-specific acetylation of histones [10] (figure, panel (c)). Both experimental approaches suggest that the chicken α-globin gene domain begins at the distance of ~21 kb before the embryonal α-globin gene π (i.e., approximately at the location of the major regulatory element of the α-globin gene cluster) and terminates several kilobase pairs after the gene $\alpha^A$, which is the last gene of the α-globin gene cluster. It has long been known that not only globin genes are transcribed in chicken erythroid cells but also a rather long 5′-terminal region of the domain [11, 12]. The situation seems to be rather complicated for analysis because this region contains a housekeeping gene (figure, panel (a)), which is transcribed in direction opposite to that of the α-globin gene transcription. However, working with K. Scherrer, we demonstrated that this region is also transcribed in the same direction as that of the globin genes [13]. It has been also shown that, first, the transcript synthesized in the direction of the α-globin gene transcription is erythroid-specific; second, the transcriptional unit boundaries virtually coincide with those of the α-globin gene domain, and the transcript is continuous and, hence, called a domain-wide transcript ([13] and references therein); third, the domain-wide transcript origin is really coincident with the location of the major regulatory element of the α-globin gene domain [14], synthesis of the domain-wide transcript begins in pre-erythroblasts, and the transcription product (the domain-wide transcript) remains in the cell nucleus [15] where it contributes to formation of the nuclear matrix [16]. All these observations suggest that the α-globin gene domain in pre-erythroblasts should be activated according to Travers' hypothesis [17]. According to this hypothesis, the elongating complex of RNA polymerase II can act as a vehicle which, moving along the genomic domain, can "carry" histone acetylase, chromatin remodeling complexes, and other enzymes necessary for activation of the chromatin domain. By now, Travers' hypothesis has been directly confirmed experimentally [18, 19]. The major regulatory element of the chicken α-globin gene domain located in the fifth intron of gene ggPRX (and similar regulatory elements of the α-globin gene domains of other vertebrates [5, 20, 21]) includes a powerful erythroid-specific enhancer operating with involvement of the protein factor NF-E2 and low-molecular-weight proteins of the Maf group [5, 22]. Our experimental findings suggest that, in addition to the enhancer, this regulatory element of the chicken α-globin gene domain also includes a domain-wide transcript promoter [14]. It seems that the enhancer is activated even in immature erythroblasts due to an increase in the concentration of erythroid-specific transcriptional factors. This triggers the domain-wide tran-

script synthesis coupled with changes in the chromatin configuration of the α-globin gene domain, which, in particular, is manifested by changes in the acetylation type of histones throughout the whole domain [10]. It should be noted that a similar mechanism is realized during the activation of the "adult" subdomain of human β-globin gene domain. A long transcript containing globin genes (δ and β) functioning in the adult organism and intergene stretches was also identified there. Deletion of this promoter transcript resulted in inactivation of the subdomain containing the δ and β genes, although the promoters of these genes and the locus-controlling region were unaffected [23].

Returning to the chicken α-globin gene domain, we would like to emphasize that a distant enhancer cannot interact with globin genes in pre-erythroblasts because of the presence between it and the promoters of a group of CTCF-dependent insulators [7]. These insulators are inactivated in mature erythroblasts, which is manifested by the absence of CTCF bound with the insulators [7]. The nature of the insulator inactivation is not quite clear. Up to now only one class of CTCF-dependent conditional (i.e. switching on or off in response to certain signals) insulators has been described, with the activity regulated through methylation of the CTCF binding site [3]. In the case of insulators located in the 5′-terminal region of the α-globin gene domain, this mechanism is unable to function because of absence of CpG dinucleotides in the CTCF recognition sites necessary for functioning of these insulators. It should be emphasized that the presence of 11 zinc fingers allows CTCF to recognize different genomic sequences, which cannot be represented as some even rather degenerated consensus. Most probably, in mature erythroblasts CTCF is displaced with involvement of erythroid-specific transcriptional factors competing for the overlapping binding sites. This hypothesis is indirectly confirmed by location of the insulators under discussion within the limits of the erythroid-specific regions of hypersensitivity to DNase I [7]. In any case, insulators separating the distant enhancer and globin gene cluster are inactivated in mature erythroblasts, and this ensures the possibility of direct interaction of this enhancer with the globin gene promoters.

The chicken α-globin gene domain, similarly to the α- and β-globin gene domains of other vertebrates, is characterized by switching of the gene expression during the development. During chicken embryogenesis, the gene π located in the beginning of the α-globin gene cluster is preferentially expressed. The preferential expression of this gene may be due to its closest location to the distant enhancer element. Consequently, from the pure statistical viewpoint, the complexing of the distant enhancer and this gene promoter is the most probable. Here, the situation is quite similar to the situation with the human β-globin gene domain, where the embryonal gene ε is located most closely to the locus control region [24]. On the termination of embryogenesis, the gene π promoter is inactivated by methylation, which stops the expression of this gene [25]. In mature chicken erythroblasts, two α-globin genes ($\alpha^A$ and $\alpha^D$) are expressed. The expression of these genes is controlled by both the distant enhancer located in the 5′-terminal region of the domain and a powerful erythroid-specific enhancer located after the gene $\alpha^A$ [26]. The activity of this enhancer crucially depends on the presence of the erythroid-specific transcriptional factor GATA-1 [26]. Correspondingly, this activity can be regulated through changes in the expression level of this transcriptional factor. Moreover, a silencer is located adjacent to the enhancer [27]. In what manner the correlative activities of the enhancer and silencer are regulated during the differentiation of erythroid cells remains unclear. It has been recently shown [28] that, in addition to the GATA-1 factor, the factor YY1 physically interacting with GATA-1 also binds with this enhancer under certain conditions. The same authors have also shown that in pre-erythroblasts this enhancer activity is suppressed on the level of packing into chromatin [28]. The distant enhancer and the enhancer adjacent to the $\alpha^A$ gene are functioning with involvement of different transcriptional factors (NF-E2 and GATA-1, respectively), and this suggests a possibility of differential activation of these enhancers through changes in the correlative concentrations of NF-E2 and GATA-1.

As noted above, the chicken α-globin gene domain, similarly to α-globin gene domains of other vertebrates, is characterized by overlapping with a housekeeping gene (the –14 gene in human [29], the ggPRX gene in chicken [4]). Under conditions of α-globin gene domain activation accompanied, in particular, by arising of a hyperacetylated chromatin domain, which also contains promoters of the above-mentioned genes [10], the expression of these genes can increase. It is difficult to predict the consequences of this, because functions of the housekeeping genes overlapping with the α-globin gene cluster are still undetermined. However, it is unlikely that the increase in the expression level of these genes is necessary for differentiating erythroid cells. In fact, we have found that in the chicken α-globin gene domain the gene ggPRX promoter is controlled by the CTCF-dependent conditional silencer, which is activated in erythroid cells [30].

In conclusion, note that organization and operation principles of the chicken α-globin gene domain are substantially similar to those of the α-globin gene domains of other vertebrates. Additional information concerning this problem can be found in some recently published papers [31-33].

This work was supported by the Russian Foundation for Basic Research (project No. 05-04-48005).

## REFERENCES


1. Ioudinkova, E. S., and Razin, S. V. (2002) *Mol. Biol. (Moscow)*, **36**, 947-955 (Russ.).
2. Ioudinkova, E. S., and Razin, S. V. (2003) *Genetika*, **39**, 182-186.
3. Razin, S. V., Farrell, C. M., and Recillas-Targa, F. (2003) *Int. Rev. Cytol.*, **226**, 63-125.
4. Sjakste, N., Iarovaia, O. V., Razin, S. V., Linares-Cruz, G., Sjakste, T., Le Gac, V., Zhao, Z., and Scherrer, K. (2000) *Mol. Gen. Genet.*, **262**, 1012-1021.
5. Flint, J., Tufarelli, C., Peden, J., Clark, K., Daniels, R. J., Hardison, R., Miller, W., Philipsen, S., Tan-Un, K. C., McMorrow, T., Frampton, J., Alter, B. P., Frischauf, A. M., and Higgs, D. R. (2001) *Hum. Mol. Genet.*, **10**, 371-382.
6. Ioudinkova, E. S., and Razin, S. V. (1999) *Mol. Biol. (Moscow)*, **33**, 603-604 (Russ.).
7. Valadez-Graham, V., Razin, S. V., and Recillas-Targa, F. (2004) *Nucleic Acids Res.*, **32**, 1354-1362.
8. Kuhn, E. J., and Geyer, P. K. (2003) *Curr. Opin. Cell. Biol.*, **15**, 259-265.
9. Razin, S. V., de Moura Gallo, C. V., and Scherrer, K. (1994) *Mol. Gen. Genet.*, **242**, 649-652.
10. Anguita, E., Johnson, C. A., Wood, W. G., Turner, B. M., and Higgs, D. R. (2001) *Proc. Natl. Acad. Sci. USA*, **98**, 12114-12119.
11. Imaizumi, M.-T., Diggelmann, H., and Scherrer, K. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 1122-1126.
12. Broders, F., Zahraoui, A., and Scherrer, K. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 503-507.
13. Razin, S. V., Rynditch, A., Borunova, V., Ioudinkova, E., Smalko, V., and Scherrer, K. (2004) *J. Cell. Biochem.*, **92**, 445-457.
14. Borunova, V. V., Iarovaia, O. V., Vassetzky, Y. S., and Razin, S. V. (2005) *FEBS Lett.*, **579**, 4746-4750.
15. Iarovaia, O. V., Razin, S. V., Linares-Cruz, G., Sjakste, N., and Scherrer, K. (2001) *Exp. Cell Res.*, **270**, 159-165.
16. Ioudinkova, E., Razin, S. V., Borunova, V., de Conto, F., Rynditch, A., and Scherrer, K. (2005) *J. Cell. Biochem.*, **94**, 529-539.
17. Travers, A. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 13634-13637.
18. Sims, R. J., Belotserkovskaya, R., and Reinberg, D. (2004) *Genes Dev.*, **18**, 2437-2468.
19. Gerber, M., and Shilatifard, A. (2003) *J. Biol. Chem.*, **278**, 26303-26306.
20. Higgs, D. R., Wood, W. G., Jarman, A. P., Sharpe, J., Lida, J., Pretorius, I.-M., and Ayyub, H. (1990) *Genes Dev.*, **4**, 1588-1601.
21. Gourdon, G., Sharpe, J. A., Higgs, D. R., and Wood, W. G. (1995) *Blood*, **86**, 766-775.
22. Wen, S. C., Roder, K., Hu, K. Y., Rombel, I., Gavva, N. R., Daftari, P., Kuo, Y. Y., Wang, C., and Shen, C. K. (2000) *Mol. Cell. Biol.*, **20**, 1993-2003.
23. Gribnau, G., Diderich, K., Pruzina, S., Calzolari, R., and Frazer, P. (2000) *Mol. Cell. Biol.*, **5**, 377-386.
24. Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and Grosveld, F. (1991) *Genes Dev.*, **5**, 1387-1394.
25. Singal, R., van Wert, J. M., and Ferdinand, L., Jr. (2002) *Blood*, **100**, 4217-4222.
26. Knezetic, J., and Felsenfeld, G. (1989) *Mol. Cell. Biol.*, **9**, 893-901.
27. Recillas Targa, F., de Moura Gallo, C. V., Huesca, M., Scherrer, K., and Marcaud, L. (1993) *Gene*, **129**, 229-237.
28. Rincon-Arano, H., Valadez-Graham, V., Guerrero, G., Escamilla-Del-Arenal, M., and Recillas-Targa, F. (2005) *J. Mol. Biol.*, **349**, 961-975.
29. Vyas, P., Vickers, M. A., Picketts, D. J., and Higgs, D. R. (1995) *Genomics*, **29**, 679-689.
30. Klochkov, D., Rincon-Arano, H., Ioudinkova, E. S., Valadez-Graham, V., Gavrilov, A., Recillas-Targa, F., and Razin, S. V. (2006) *Mol. Cell. Biol.*, **26**, 1589-1597.
31. Anguita, E., Hughes, J., Heyworth, C., Blobel, G. A., Wood, W. G., and Higgs, D. R. (2004) *EMBO J.*, **23**, 2841-2852.
32. Higgs, D. R., Garrick, D., Anguita, E., de Gobbi, M., Hughes, J., Muers, M., Vernimmen, D., Lower, K., Law, M., Argentaro, A., Deville, M. A., and Gibbons, R. (2005) *Ann. N. Y. Acad. Sci.*, **1054**, 92-102.
33. Higgs, D. R., Vernimmen, D., de Gobbi, M., Anguita, E., Hughes, J., Buckle, V., Iborra, F., Garrick, D., and Wood, W. G. (2006) *Biochem. Soc. Symp.*, Vol. 73, Portland Press, Oxford, pp. 11-22.
34. Razin, S. V., Kekelidze, M. G., Lukanidin, E. M., Scherrer, K., and Georgiev, G. P. (1986) *Nucleic Acids Res.*, **14**, 8189-8207.
35. Verbovaia, L., and Razin, S. V. (1995) *Gene*, **166**, 255-259.