Regulatory Elements in Vectors for Efficient Generation of Cell Lines Producing Target Proteins

O. Maksimenko, N. B. Gasanov, P. Georgiev*
Institute of Gene Biology, Russian Academy of Sciences, Vavilova str. 34/5, 119334, Moscow, Russia
*E-mail: georgiev_p@mail.ru
Received 14.04.2015

ABSTRACT To date, there has been an increasing number of drugs produced in mammalian cell cultures. In order to enhance the expression level and stability of target recombinant proteins in cell cultures, various regulatory elements with poorly studied mechanisms of action are used. In this review, we summarize and discuss the potential mechanisms of action of such regulatory elements.
KEYWORDS insulators, recombinant proteins, protein production in mammalian cells, UCOE, S/MAR, STAR.
ABBREVIATIONS AND TERMS RP – recombinant protein; CHO – Chinese hamster ovary cells; UTR – untranslated gene region; kbp – a kilobase pair, a unit of length of nucleic acids equal to 1,000 base pairs; S/MAR – DNA sequences corresponding to nuclear matrix-attachment regions; insulators – regulatory elements that block interaction between enhancer and promoter; UCOE –regulatory elements containing strong promoters of housekeeping genes; STAR – regulatory elements protecting from HP1-dependent repression; MSX –L-methionine sulphoximine; MTX – methotrexate.

INTRODUCTION

Therapeutic proteins are major structural and regulatory molecules essential for a normal functioning of the human body. Part of recombinant proteins that are small in size and do not require additional modifications are produced in the most economical bioreactor – *E.coli* cells. However, the production of proteins in bacteria is associated with a number of limitations: disrupted folding of some proteins, absence of crucial modifications, and the inability to produce larger molecules [1, 2]. Some of these limitations are avoided when using yeast cells, which can produce high-quality recombinant proteins at sufficiently low cost. However, many recombinant human proteins require specific modifications that can only be obtained in higher eukaryotic cells [3]. Therefore, nowadays there is an increasing number of drugs obtained from mammalian cell cultures grown in bioreactors, mostly Chinese hamster ovary (CHO) cells [4, 5]. CHO cells were first isolated in 1957 [6]. It soon became evident that these cells are ideal for biomass scaling during the generation of recombinant proteins at the bioreactor level, since they are undemanding vis-a-vis growth conditions. Several lines were obtained from the initial clone of CHO cells, among which the CHO-K1 line became the most commonly used [7]. Optimal conditions for providing high growth density of CHO cells in bioreactors were selected using this cell line, enabling a significant increase in the product (target protein) yield alongside a reduced chance of human virus transmission [1, 4, 5]. Yet, the main problem in recombinant protein production in cultured cells is the extremely low product cost; therefore, there is a constant effort to reduce the expenses on obtaining high-producing cell cultures and elevate the yield of the protein product by enhancing the expression level of the target protein, cell culture density, and decreasing cell death. One of such approaches is vector improvement for obtaining transgenes, which can significantly reduce expenses in the generation of producing cell cultures. This paper presents an overview of the regulatory elements used in vector constructs for the generation of transgenic lines.

VECTOR CONSTRUCTS FOR GENERATION OF TARGET PROTEIN-EXPRESSING CELL LINES

The most widely used method in industrial biotechnology is transfection with linearized plasmid DNA [3, 8], which allows one to obtain cell lines containing multiple copies of the expression vector that are usually integrated in one or, rarely, several genomic sites. The mechanism of vector construct integration into the genome has not yet been fully elucidated. Introduction of

linear DNA into the nucleus results in the activation of reparation systems that provide cross-linking of linear DNA ends through two main mechanisms: homologous recombination and ligation of nonhomologous DNA ends [9–11]. As a result, long linear DNA molecules are formed bearing several copies of the vector construct capable of integrating into the genome with some probability. Linear DNA can integrate a genomic region that already contains other copies of the same DNA through the mechanism of homologous recombination, which leads to an increase in the copy number of the construct integrated in a specific genomic site [12]. Thus, one genomic site can bear up to several hundreds of integrated copies of a vector construct.

In order to enhance the producing capacity of clones, selective increase in the number of construct copies, which results in target protein expression, is used in industrial biotechnology [8]. This is achieved by decreasing the functional activity of the reporter gene encoding enzyme dihydrofolate reductase (DHFR), which catalyzes conversion of dihydrofolate to tetrahydrofolate and is essential for the synthesis of glycine, purines, and thymidine acid [13, 14]. DHF-auxotrophic cell lines can grow in media that necessarily contain glycine, as well as a source of purines (hypoxanthine) and thymidine. The derivative line CHO-DG44 with mutations of both alleles of the dihydrofolate reductase gene was obtained by random mutagenesis of the CHO-K1line quite a long time ago [15, 16]. It enabled to use dihydrofolate reductase as a selection gene for the generation of producing cell lines. Furthermore, in order to selectively increase the number of expression vector copies, which usually correlates with an increase in the target protein level, methotrexate (MTX) is used, which is able to selectively inhibit dihydrofolate to tetrahydrofolate conversion [2]. Treatment of transfected cell lines with MTX results in the survival only of cells with a significantly elevated level of dihydrofolate reductase. In most cases, this is caused by an increase in the copy number of the *dhfr* gene included in the construct and, as a result, of the gene encoding the target protein. Another frequently used selectable marker is the gene encoding glutamine synthetase (GS). The CHO-K1 cell line, which contains mutated allele of *gs*, is used when working with this marker. In this case, *L*-methionine sulphoximine (MSX) is used as a selective agent facilitating the selection of the most effective clones [2].

Clone selection can be conducted using other reporter genes, as well. The most promising options are the ones that do not require selection of a mutant cell line. Fluorescent protein technology, which allows one to select cells with maximum expression of a target gene based on emission at a particular wavelength, can be considered [17, 18]. Such an approach can be used for the generation of stable cell lines, which, in contrast to mutant derivatives deficient in *dhfr* and *gs*, exhibit a higher proliferative potential and viability. Among the disadvantages of using fluorescent protein genes as selectable markers is the inability to amplify the copy number, which in most cases would lead to increased producing capacity of a cell clone.

Recently, the application of special robots capable of selecting individual cell clones with the most efficient expression of the target protein, which is identified using antibodies, has become widespread [4]. Alongside with other advantages, this technology enables to avoid using markers the expression of which does not always correlate with the target protein level.

Vector constructs have also been developed based on viruses, mobile elements, bacterial anti-phage protection system, and recombination systems in phages and yeasts [3, 19]. Application of such vectors in several cases enables single insertions of a target gene into a specific genomic region, which is commonly used in gene therapy when generating transgenic cell lines and animals in order to obtain model systems for the study of gene expression regulation processes [20].

The main challenge in obtaining producing cell lines containing multiple expression vector copies is heterochromatin formation of repeated sequences of vector constructs, which usually enhances upon cell proliferation. The main role in heterochromatin formation of a repeated DNA sequence is played by RNA interference and noncoding RNAs that can stimulate repressive chromatin zone formation at promoters, as well as methylation of CpG sites in promoter regions, which decreases the efficacy of transcriptional factor binding in such regions [21–23]. This can result in a significant decrease in the target protein level after the obtainment of highly producing cell lines for some period. Moreover, heterochromatin formation at repeated copies of a vector construct can negatively affect the activity of adjacent cellular genes, which often leads to a decreased viability of producing cell lines. Repression of transcription from integrated repeats of a vector construct is caused by the cellular response to the introduction of foreign information, the expression of which should be suppressed, into the genome. Thus, application of regulatory elements capable of supporting efficient performance of the target protein gene alongside isolating regulatory elements of a vector construct from genomic regulatory elements seems to be extremely important.

USE OF PROMOTERS AND ENHANCERS FOR GENERATION OF CELL LINES PRODUCING TARGET PROTEINS

Strong viral promoters, such as the cytomegalovirus (CMV) promoter and the early promoter of SV40,

as well as strong cellular promoters of housekeeping genes, such as β-actin and factor EF1α genes, are usually used for transgene expression [24].

Strong viral or cellular promoters contain a minimal promoter of approximately 100 bp, which serves as the transcription start site (TSS), and a strong enhancer located in close vicinity to the promoter. For instance, the most widely used CMV promoter bears a core region located between positions −62 and −1 bp from TSS and the enhancer (−544 through −63 bp) [25]. There are several motifs in the region of the minimal promoter that determine the association with various components of the core transcription complex (TFIID): TATA [26], INR [27], DPE [28], BRE [29], DCE [30], and MTE [31]. However, such regions are not essential, since most strong promoters do not contain these elements or belong to the class of GC-rich promoters. One can assume the existence of a practically unexplored group of so-called architectural proteins determining the capacity of the core promoter to recruit the TFIID complex [32]. Unfortunately, the promoter architecture has not been studied deeply enough to determine, based on the sequence of a minimal promoter, its capacity to effectively bind the core transcription factors necessary for the performance of a strong promoter.

One of the approaches in applying effective promoters in biotechnology is identification of strong promoters directly in the cell cultures that are further to be utilized for the generation of target protein-producing lines. Thus, total genome screening of the strongest promoters in CHO cells that are most commonly used as the expression system in mammalian cells has been performed [33–35]. As expected, the most effective promoters appeared to be the promoters of housekeeping genes, including some ribosomal genes. However, the pitfall of this approach is a significantly high chance that the promoters identified by a genome-wide analysis may perform effectively only when located in a certain genomic region (the position effect) or when containing a complicated regulatory region, which significantly decreases the attractiveness of using such promoters in expression systems for obtaining target proteins from transgenic CHO cell lines.

One of the solutions to the position effect is using long regulatory sequences of actively transcribed housekeeping genes located on both sides of the coding region of the gene. Thus, a high expression level of target proteins (6–35 times the level of expression from a standard CMV promoter) has been obtained for vectors bearing a 12 kbp regulatory region or a 4 kbp 3'-region of the Chinese hamster *EF1α* gene [36]. One of the problems in utilizing long regulatory sequences for target protein expression is the instability of large vectors and decreased efficiency in generating multicopy lines that predominantly bear full constructs capable of expressing a target protein. The perspective model is the construct that includes long regulatory DNA regions from terminal repeats of the Epstein-Barr virus, which provides an order increase in efficiency in obtaining stably transfected cells [37].

Artificial modification of promoters is another promising approach to enhance their activity. For instance, a strong CMV promoter has been demonstrated to undergo negative regulation resulting in methylation of GC regions at transcription factor (TF) binding sites comprising promoters and, as a consequence, inhibition of TF recruitment to the promoter. As a result, the activity of the CMV promoter is greatly decreased. Such a negative affect can be avoided by integrating between the enhancer and core elements of the CMV promoter a regulatory sequence that binds transcriptional factors suppressing the DNA methylation process [38]. An effective promoter consisting of two divergent core elements with a single CMV enhancer integrated between them has been developed based on two CMV promoters [39]. This bidirectional promoter is able to express two divergent genes with approximately the same efficiency, which plays an important role in the production of proteins consisting of two different subunits (e.g., monoclonal antibodies).

A novel means to increasing a target protein expression is artificial recruitment of effective transcription-associated complexes to the promoter [40, 41]. For instance, histone acetyltransferase p300 binds active enhancers and promoters and also participates in the stimulation of transcription [42]. Recruitment of p300 to promoters significantly enhances efficiency in generating stable cell clones with a high level of a reporter gene expression [43]. It is worth mentioning that opposite results have been obtained in similar experiments with the Brahma remodeling complex, which provides increased mobility of nucleosomes and is capable of positively/negatively regulating transcription depending on a particular gene.

In order to enhance reporter gene transcription and reduce transgene expression dependence on the surrounding chromatin, strong cellular enhancers are used [44, 45]. One of the most frequently used enhancers is LCR (Locus Control Region), which controls the expression of human β-globin locus genes [46]. The main disadvantage of using enhancers for elevating transgene expression is due to their specificity, i.e. the ability to function only in certain cell lines, which imposes certain restrictions on their use as a general regulatory element. The search for the enhancers most efficient in the cell lines that are used for the generation of proteins at an industrial scale seems to be a promising trend in this direction [47].

PROSPECTIVE USE OF INSULATORS FOR ENHANCING THE EFFICIENCY OF TARGET GENE EXPRESSION IN PRODUCING CELL LINES

In order to increase the efficiency and stability of target protein expression, known insulators are used [41, 48-51]. Insulators are regulatory elements that block interaction between the enhancer and promoter if interposed between them [52, 53]. In addition, insulators do not directly affect an enhancer's and promoter's activity, which means that the promoter can be activated by any other enhancer, and the enhancer, in its turn, is capable of activating any other promoter. In addition, some insulators can serve as boundary elements between transcriptionally active chromatin and heterochromatin. The best-studied examples are insulators of fruit fly Drosophila and vertebrates. Initially, it was assumed that insulators determine the borders of transcriptional domains within which gene expression does not depend on the negative effects of the surrounding genome [54, 55]. However, it has been later demonstrated that insulator proteins are considerably more flexibly integrated into the gene regulatory system [52, 53].

Recently, certain insulator proteins have been shown to participate in the organization of specific, long-range interactions between distal regions of chromatin [56–59]. Insulator proteins can support interactions between enhancers and promoters, boundaries of transcriptional domains which are usually located up to several hundreds of kbp away [34, 60–62]. The obtained results allowed one to refer the class of insulator proteins to chromatin architectural proteins [32, 53].

To date, insulator architectural proteins (IAP) of Drosophila remain poorly described, which is largely due to the ease of generating transgenic lines of flies. The study of insulator properties in Drosophila transgenic lines showed that each insulator binds several IAP, which in turn determine the specificity of long-range interactions [32, 53]. As a result, two identical insulators can provide sufficient specific ultra long-distance interactions between regulatory elements in Drosophila transgenic lines [63, 64], which allowed researcher to propose a model in which IAP associated with specific binding sites, comprising regulatory elements, create a code, which in turn determines how effective a long-range interaction established between these regulatory elements will be (*Fig. 1A*) [32].

Usually, transgenic Drosophila lines are obtained by injecting a vector that contains P element ends flanking the transgenic construct alongside the gene encoding transposase essential for construct integration into the genome [65]. Cases have been described of highly specific integration of the P element that includes an insulator or a strong promoter into a certain genomic region containing the cognate endogenous regulatory element [66, 67]. Such highly specific integration of the P element called *homing* can be explained by the recruitment of IAP to the regulatory element that comprises the P element upon introduction of the construct into an embryo. This results in specific interaction between the P element and the cognate endogenous regulatory element, with further integration of transposon into a certain region at the chromosome (*Fig. 1B*).

Genome-wide studies of IAP-binding site distribution unambiguously demonstrated that insulators are not fixed boundaries between transcriptional domains [62, 68]. Two mechanisms of enhancer activity suppression by insulators have been described in Drosophila transgenic lines, cell cultures, and *in vitro* [53]. The first

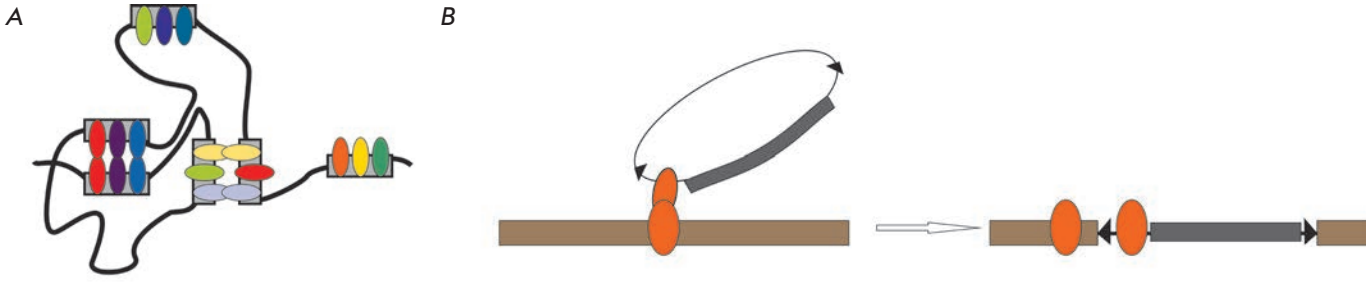

Fig. 1. *A* – Model of establishing specific long-range interactions between regulatory elements. Several IAP (insulator architectural proteins) bind with each element. As a result, two identical elements are capable of providing highly specific and efficient ultra long-range interactions between regulatory elements, but only weak contacts are formed upon partial overlap of IAP-binding sites in regulatory elements. Grey rectangles represent regulatory elements binding with IAP; colored ellipses depict combinations of IAP that specifically bind their own sites and interact with each other. *B* – Specific integration of a transgenic construct into a particular genomic region. The transgenic construct is presented as a circle. Black triangles correspond to P element end repeats; grey rectangle, –elements comprising a transgenic construct. Orange ellipses –regulatory elements that bind with IAP. Brown rectangle – a genome region

mechanism is based on the appearance of topological obstacles that block interactions between the enhancer and promoter. As a result of the formation of stable contacts between insulators, one of the interacting regulatory elements appears to be isolated in the independent chromatin loop. Such a mechanism of insulator action was found using artificial insulators that contained binding sites for proteins capable of effective interaction with each other and formation of stable chromatin loops [69, 70], and using transgenic lines of Drosophila [71]. This mechanism manifests itself effectively only in the case when insulators are immediately adjacent to the suppressed elements (enhancers and promoters). When the size of the chromatin loop formed by insulators is larger, such an enhancer blockage pathway is not utilized [72].The second, more common, mechanism is based on the establishment of direct contacts between the proteins associated with the insulator and enhancer-promoter complex. For example, it has been demonstrated in transgenic lines of Drosophila that insulators may directly interact with promoters and enhancers [71, 73]. In such a case, when the insulator is located between the enhancer and promoter, insulator proteins interfere with the establishment of proper contacts between the transcriptional complexes assembled at the enhancer and promoter, which leads to partial or complete inability of the enhancer to stimulate transcription from the promoter.

The mechanisms that determine the barrier function of insulators have been described in detail in Drosophila and mammals. In particular, it was found that IAP help to recruit the protein complexes responsible for nucleosome remodeling and modification to insulators [74-78], resulting in the formation of open chromatin zones. At the same time, some insulators can recruit the protein complexes directly involved in transcription stimulation [75]. Due to the formation of nucleosome-free DNA regions and recruitment of transcriptional complexes, insulators suppress the spread of repressive chromatin, which, nevertheless, does not exclude the possibility of direct interaction between insulators and silencers initiating heterochromatization.

To date, only a single DNA-binding insulator protein, CTCF, has been described in vertebrates [79, 80], which is probably due to the absence of convenient model systems for the study of insulators. CTCF is capable of supporting long-range interactions between distal areas of chromatin [60, 79–81]. Thus, CTCF is the first architectural protein characterized in a mammalian genome [32, 80].

Except for their key role in the formation of chromatin architecture, CTCF protein domains remain poorly studied and the mechanism of CTCF performance in maintaining long-range interactions has not been characterized yet. The main part of the protein consists of 11 C2H2-type zinc fingers (ZF), with only four of them (4th to 7th) being essential for the recognition of the core DNA motif [82]. The remaining zinc fingers seem to recognize the specific nucleotide sequences stabilizing CTCF association with DNA. The most logical suggestion is that a protein supporting long-range interactions is capable of effective di- and multimerization. Indeed, CTCF has been shown to be able to homodimerize; however, the domain responsible for this activity has not been identified yet [83]. Evidence has been obtained that the C-terminal domain of CTCF also interacts directly with its own zinc fingers [84]. However, such interaction cannot be highly specific, because CTCF zinc fingers bind many other transcription factors, as well: CHD8, Sin3A, and YB-1 [85-87].

The cohesin complex, which associates directly with CTCF [90], is suggested to play a significant role in the organization of long-range interactions [60, 80, 88, 89]. The cohesin complex is recruited to chromatin by CTCF and facilitates the formation of long-range interactions between CTCF genomic sites. This model is consistent with genome-wide studies demonstrating a high degree of colocalization of CTCF and cohesin subunits [91, 92]. However, a very slight decrease in binding of cohesins to chromatin has been shown in experiments on CTCF inactivation, suggesting the implication of other transcription factors in the recruitment of the cohesin complex at chromatin [92-94]. Moreover, inactivation of CTCF and cohesins leads to various disruptions of the chromatin architecture [95, 96], which can be attributed to independent functioning of these proteins.

The most well-studied vertebrate insulator, HS4, consisting of 1,200 bp and located at the 5'-end of the chick β-globin locus, is used in biotechnology (*Fig. 2*) [97, 98]. A core region of 250 bp has been found in this insulator, which exhibits the activity of the complete insulator and contains five fragments (FI, FII, FIII, FIV, FV), each of which has its own functional value. A site that binds CTCF, which is necessary and sufficient for the manifestation of the enhancer-blocking activity of HS4, has been identified in the FII region of the insulator [99]. Proteins USF1 and USF2, which bind as heterodimers to the FIV region, are responsible for boundary formation between active chromatin and heterochromatin [100]. USF has been shown to recruit the protein complexes responsible for modification of the histones associated with transcription stimulation [100, 101]. The protein BGP1/Vezf1, which possesses a DNA-binding domain consisting of zinc fingers, associates with other regions of the HS4 insulator (FI, FIII, FV) [102]. The protein BGP1/Vezf1 protects GC-rich regions of the insulator from methylation, which af-

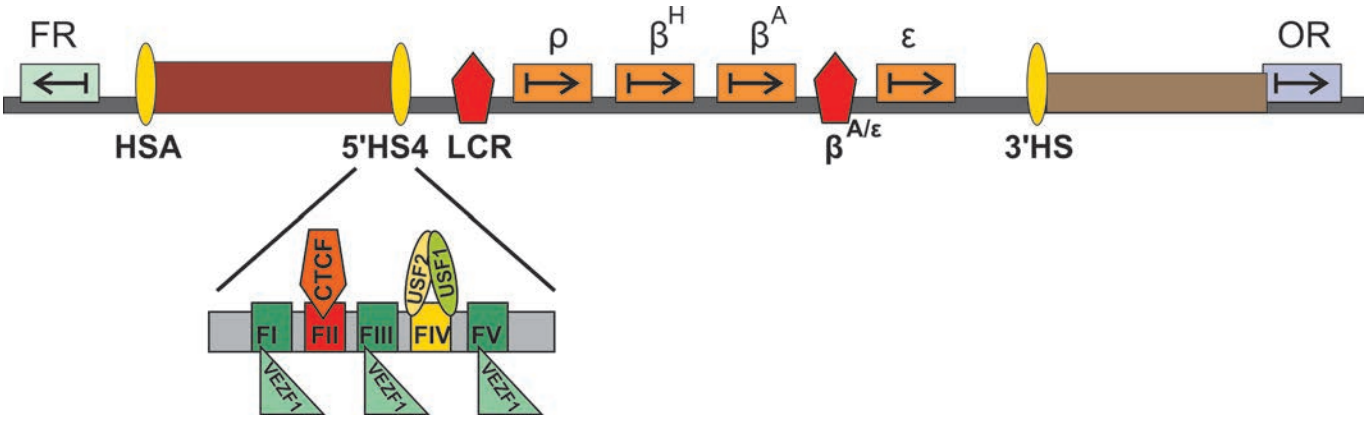

Fig.2. Schematic representation of β-globin locus and adjacent sequences. Designations ρ, β^H, β^A, ε correspond to the genes β-globin locus; FR – folate receptor gene; OR – olfactory receptor gene. Arrows indicate the direction of gene transcription. HSA, 5′HS4, 3′HS – insulators; LCR, $\beta^{A/\varepsilon}$ – enhancers of β-globin locus. Insulator HS4 is represented schematically in detail. FI, FIII, FV –binding sites for protein Vezf1/BGP1. FII and FIV – protein CTCF- and heterodimer USF1/2-binding sites, respectively

fects the recruitment of insulator proteins to DNA and, therefore, results in insulator inactivation. According to the existing model, BGP1/Vezf1 terminates weak transcription from the region of heterochromatin, which may play an important role in protection of the β-globin locus from the spread of inactive chromatin [103].

Since its discovery, insulator HS4 has been actively utilized for transgene expression in mammalian cell cultures [98]. Two complete copies of HS4 have been integrated into a vector for producing transgenic animals expressing a target protein in milk [104]. It was demonstrated that the insulator substantially enhances the expression of target proteins, but it has no significant affect on the specificity of transgene expression only in the mammary gland, and does not provide a direct correlation between the copy number of the construct and the level of target protein production [105].

Most effectively, HS4 insulator can be applied in vectors that for some reason have a limited size. Thus, a full-size 1.2 kbp insulator significantly reduces the efficiency of cellular transformation with lentiviral vectors (probably due to the limitations imposed on the size of the viral particle). Therefore, HS4-duplicated core element of 250 bp, which contains the binding sites of all the identified transcription factors required for the manifestation of insulator activity, is used in vectors of this class [98, 106]. Insulators are also successfully used for protecting reporter gene expression in vectors designed based on mobile elements [51] and retroviruses [107].

Despite the examples of successful use of a 1.2-kbp HS4 insulator or its core region [98, 108], abundant data have been obtained showing that HS4 does not have a positive effect on target gene expression. This can be explained by the fact that the cell cultures that were used in the experiments differed significantly in the set of transcription factors that bind with the HS4 insulator.

In conclusion, the following basic mechanisms for the protection of transgene expression using the HS4 insulator can be put forth: 1) formation of a nucleosome-free DNA region that can disrupt the linear spread of heterochromatin; and 2) recruitment of protein complexes that enhance nucleosome mobility, modify histones, stimulate transcription, protect CpG-sites from methylation, and terminate weak transcription. It has not been determined yet whether the HS4 insulator is capable of guiding construct integration into transcriptionally active chromatin zones and directly interacting with the target gene promoter for further transcription stimulation. Apparently, the main disadvantage of HS4 and other insulators is the dependence of their activity on the set of particular transcription factors expressed in the cell line.

ENHANCING TRANSGENE EXPRESSION USING A/T-RICH SEQUENCES ASSOCIATED WITH NUCLEAR MATRIX PROTEINS (S/MAR)

In order to enhance transformation efficiency and improve the stability of transgene expression, sequences of 300-5,000 bp, usually A/T-rich, which interact with a fraction of the nuclear matrix (S/MAR, scaffold matrix attachment region), as shown in experiments *in vitro*, have been widely used from the beginning of the 90s [108-110]. S/MAR regions possess a number of distinctive properties: they are A/T-rich regions, sensitive to DNase I, and potentially tend to form left-hand-

ed helix and triplex structures [111, 112]. It is assumed that it is the A/T-rich composition of these elements which leads to the destabilization of the double helix and ability of MAR to generate areas rich in various secondary structures [113, 114].

Based on the characteristics of the secondary structure, there is about an order of 50,000 elements predicted for the human genome that supposedly share the properties of S/MAR [115, 116]. A total of 1,500 regions that have the most relevant characteristics of S/MAR have been selected from this pool. Only several of them turned out to share a high level of homology with mouse orthologs, implying that the nucleotide sequence of S/MAR regions lacks distinctive, conserved elements.

The structure of S/MAR elements indicates that they might serve as recombination hotspots. Indeed, it was shown that disruption sites that occur due to inversions associated with human diseases are often localized in S/MAR-elements [117, 118], and integration of retroviruses into the genome occurs in close vicinity to S/MAR at a high frequency [119, 120]. According to some reports, S/MARs participate in the regulation of DNA replication [121-123]. It was found that S/MAR elements enhance transgene expression and reduce expression variability during the generation of stable cell lines [41, 124]. Expression of a gene surrounded by S/MAR elements has been experimentally established to be proportional to the gene copy number [125]. It is assumed that S/MAR elements can be functionally regarded as insulators that protect transgene expression from the positive/negative effects of the surrounding chromatin.

Initially, it was thought that lamina are a major component of nuclear matrix proteins [126]. Later, many other additional proteins, including transcription factors, were found in the nuclear matrix [127]. S/MAR often contain the binding sites of such transcription factors as SATB1, Fast1, CEBP, SAF-A, and SAF-B (proteins that preferentially bind to A/T-rich regions), NMP4 (matrix protein), CTCF and Hox family proteins [83, 116, 128-131]. Topoisomerase II also predominantly associates with A/T-rich regions within S/MAR [132-134]. Reduced density of nucleosome distribution in S/MAR elements and increased concentration of histone acetylation complexes is explained by the association of numerous transcription factors and ability of these elements to form secondary structures.

SATB1 is the most well-studied matrix protein involved in many biological processes, such as differentiation of T cells and epidermis [135–137]. SATB1 can be included in the class of architectural proteins capable of maintaining specific long-range interactions [135]. SATB1 forms homodimers and binds to A/T-rich sequences with two CUT domains and one C-terminal homeobox. Apart from participation in chromatin domain formation, SATB1 recruits ASF1 (ATP-dependent factor involved in chromatin organization) and the ISWI complex (enhances nucleosome mobility) [128, 138].

SAF-A, another matrix protein, includes a DNA-binding (SAF) and an RNA-binding (RGG) domain. It is interesting that Xist RNA, which regulates dosage compensation in mammals, also associates with the RGG domain. This interaction determines the localization of Xist RNA on the X chromosome [139]. According to the existing model [140], SAF-A recruits Xist RNA to a S/MAR element located in the region of initiation of heterochromatin formation on the X chromosome. Interaction between the proteins SAF-A and SATB1 further results in the formation of a loop between neighboring S/MAR complexes, which ultimately leads to the spread of Xist RNA on chromosome X and its subsequent inactivation.

According to the most commonly used model, S/MAR elements interact with the proteins of the nuclear skeleton (matrix proteins), resulting in the formation of chromatin loops where S/MAR serves as a core element [141]. Genes located within a chromatin loop formed by S/MAR are assumed to be protected from the negative influence of the surrounding chromatin [109]. Nevertheless, the structure of the nuclear matrix and role of S/MAR in the organization of the chromosome architecture still remain elusive. According to recent concepts, the matrix presents labile conglomerates of proteins, which transiently interact with S/MAR protein complexes comprising the chromosomes [141].

Despite the lack of understanding of the mechanisms underlying S/MAR action, abundant experimental data has been obtained demonstrating the effectiveness of using these elements for enhancing the expression of target proteins in mammalian cell cultures [142]. For example, S/MAR of the lysozyme gene from the chicken egg causes a 5- to 10-fold increase in the level of monoclonal antibody expression in CHO cells [109, 143]. Later, other, more effective mammalian S/MARs were characterized [144]. S/MARs have been successfully used to increase the expression level of erythropoietin, as well as human growth factor TGF-β receptor type II [108]. Furthermore, S/MAR appears to function both within viral vectors [145] and vectors designed based on transposable elements [146]. S/MAR effectively protects transgene expression from repression (barrier activity) and also supports a higher level of transcription from promoters within expression vectors (stimulatory activity) [144]. Several S/MARs are capable of increasing the efficacy of vector construct integration into chromosomes [147, 148]. Moreover, S/MAR proteins are able to provide integration of new additional

copies of the vector construct in a region already containing integrated construct copies. The best possible explanation for such a characteristic of S/MAR is that the interaction between architectural proteins (such as SAF-A and SATB1) provides a contact between S/MAR copies located in the genome and plasmid (analogue of the homing phenomenon in Drosophila mentioned above). Apparently, the increased recombination activity provided by S/MAR elements enhances the efficiency of transgene integration into specific regions of the genome. Thus, the positive impact of S/MAR on transgene expression might be largely determined by directed integration of a vector construct into normally transcriptionally active S/MAR-containing areas. This implication is consistent with the finding that many of the studied S/MARs have a positive influence on gene expression only when integrated into the genome [125, 144]. In order to increase the amplification rate of a construct in cell clones, S/MAR was combined with the mammalian replication initiation region [149, 150]. Treatment of primary transfectants with MTX enabled to achieve large-scale amplification of a construct in cell clones, which led to a stable multifold increase in the production of the target protein [150].

In conclusion, it can be stated that S/MAR present regulatory elements that are less studied than insulators. The most probable mechanisms of S/MAR action in enhancing transgene expression are 1) site-specific integration of S/MAR-bearing constructs into the regions of transcriptionally active chromatin and amplification of the copy number of the integrated construct, 2) association with S/MAR complex elements that guide transcriptionally active chromatin zones and thus suppress the spread of heterochromatin, and 3) immediate promoter activation by transcription factors that directly bind to S/MAR.

ENHANCING TRANSGENE EXPRESSION IN CELL CULTURES BY REGULATORY ELEMENTS BEARING A STRONG PROMOTER OF HOUSEKEEPING GENES

Between 2000 and 2002, a small company named CobraTherapeutics developed a technological platform based on regulatory elements isolated from housekeeping genes, which are actively transcribed at all stages of development and in all cells of an organism, for obtaining efficient cell lines producing recombinant proteins [45]. These regulatory elements were named ubiquitous chromatin opening elements (UCOE) since the promoter regions of the actively transcribed genes are characterized by a low density of nucleosomes, which is due to the presence of DNA-binding TF stimulating transcription. The best characterized UCOE are DNA regions that contain a pair of divergent gene promoters, *HNRPA2B1* and *CBX3* or *TBP* and *PSNB1*, which are actively transcribed in all cells of an organism [151]. The first experiments used large regulatory elements of 12-16 kbp, which significantly increased the percentage of transfected cells and provided high-level and stable transgene expression for a long cultivation period [151, 152]. Thus, UCOE causes a 16-fold increase in the efficiency of the CMV promoter, which is highly susceptible to inhibition by RNA interference and methylation of CpG regions [152, 153]. It was shown that UCOE can maintain a high expression level of a transgene integrated into pericentromeric heterochromatin. UCOE are also effective as part of lentiviral vectors [154-159]. It can be assumed that UCOE are bound by transcription factors that recruit the complexes preventing methylation of CpG repeats and forming chromatin areas with reduced nucleosome density in the promoter regions comprising lentiviral vectors [155, 160].

It should be noted that, unlike other regulatory elements such as LCR and enhancers, which exhibit pronounced cell specificity, promoters of housekeeping genes can function effectively in various cell lines. In experiments with various UCOE that were reduced in size in order to assess the possibility of using UCOE in expression vectors, more compact-size variants of UCOE (1.5 to 3 kbp) have been obtained. Such truncated elements completely retain their activity during the generation of high-producing cell lines [152].

UCOE actively participate in the process of transcriptional regulation, which implies the existence of direct interactions between the promoter regulatory elements responsible for the expression of the reporter gene located in the vector and the transcription factors associated with UCOE. Therefore, UCOE can effectively act only on certain promoters and the functional activity of these promoters does not manifest itself in cell lines [161]. Some studies have demonstrated that UCOE themselves can be used as promoters for providing stable expression of a reporter gene [162]. However, contribution in transgene expression of the transcription initiated from promoters comprising UCOE is ambiguous, and its role remains elusive. In particular, there are experimental data showing that UCOE do not always effectively enhance the expression of a target protein in CHO cells [162, 163]. Negative results obtained using UCOE can be explained by the fact that strong promoters comprising UCOE induce transcription that in some cases is capable of triggering RNA interference and/or recruiting repressive complexes to the promoter that transcribes the reporter gene.

It was also demonstrated that a combination of two strong promoters may in some cases facilitate the generation of stable cell lines and enhance transgene expression [164]. Analysis of various combinations of two of the promoters CMV, SV40, RPL32, EF1-α and

β-actin showed that only the *RPL32* promoter, integrated before any other of the studied promoters, can significantly increase the efficiency of stable cell clone selection. It is worth mentioning that the direction of the *RPL32* promoter should coincide with the direction of the promoter responsible for reporter gene expression, and core elements of the *RPL32* promoter in this case are essential components of the system for providing a stimulating effect.

In general, combining strong promoters is a promising way to improve the efficiency of generating cell clones producing a target protein. Strong promoters bearing a combination of enhancer and core promoter recruit protein complexes, which in turn support a transcriptionally active state of chromatin. According to the data of genome-wide studies, promoters act as effective boundaries that are able to protect genome areas against the spread of repressive chromatin regions [165, 166].Transcription factors that bind to properly matched promoter pairs can mutually reinforce each other's activities. Apparently, the use of some promoters can provide advantageous integration of a construct into certain chromosome areas with the highest levels of transcription. A more complete understanding of the mechanisms of transcription activation will further allow researchers to modify promoters in order to improve their performance when using them in expression systems.

ENHANCING THE TRANSGENE EXPRESSION LEVEL IN CELL CULTURES BY REGULATORY ELEMENTS PROTECTING FROM HP1-DEPENDENT REPRESSION

The Chromagenic company has developed a technological platform based on a test-system which effectively allows the identification of regulatory elements capable of suppressing the spread of heterochromatin areas [45]. The test-system is based on the recruitment of the HP1 protein, which is responsible for heterochromatin formation, to a plasmid using the DNA-binding domain of the Lex protein [167]. The chimeric protein HP1-Lex binds to Lex-specific sites on the plasmid and recruits other components of the heterochromatin complex, which launches the inactivation of the adjacent promoter. This results in repression of zeo^R, which is responsible for resistance to the antibiotic Zeocin, and the death of transfected cells when cultured in a selective medium with the addition of Zeocin. The screening aimed at detecting DNA fragments, integration of which between Lex binding sites and the zeo^R promoter protects the promoter from HP1-dependent repression, enabled to find a series of regulatory elements 500 to 2,000 bp in length called antirepressors (STAR). Other known regulatory elements such as the insulator HS4, MAR, and UCOE lack that ability. A comparative analysis of various regulatory elements [168] has shown that STAR elements are most effective when using them for generating high-producing CHO cell lines. However, the mechanism of STAR element action remains unexplored. There is still no evidence on what transcription factors bind with elements of this class and provide them with functional activity.

CONCLUSION

To date, no universal regulatory element with a clear mechanism of action has been found that can be effectively used in all types of vector constructs designed to generate cell lines producing various proteins at a high level. This is largely due to the complexity of the mechanisms that regulate promoter activity and also the absence of actual evidence on the original concepts of strict organization of genes with the same expression profile into transcriptional domains surrounded by special regulatory elements from a class of insulators or S/MAR. Clearly, some mechanisms must exist that suppress excessive transcription even for the strongest promoters. RNA interference is one of such mechanisms. It is possible that a detailed understanding of the mechanisms of transcription activation and suppression will lead to the development of artificial promoters that allow researchers to obtain stable high levels of target gene expression in transgenic systems. ●

The work was supported by the Russian Scientific Fund (grant № 14-24-00166).

REFERENCES

1. Demain A.L., Vaishnav P. // Biotech. Advances. 2009. V. 27. P. 297–306.
2. Durocher Y., Butler M. // Curr. Opin. Biotech. 2009. V. 20. P. 700–707.
3. Khan K.H. // Adv. Pharm. Bull. 2013. V. 3. P. 257–263.
4. Kim J.Y., Kim Y.G., Lee G.M. // Appl. Microbiol. Biotechnol. 2012. V. 93. P. 917–930.
5. Hacker D.L., De Jesus M., Wurm F.M. // Biotechnol. Adv. 2009. P. 1023–1027.
6. Tjio J.H., Puck T.T. // J. Exp. Med. 1958. V. 108. P. 259–271.
7. Kao F.T., Puck T.T. // Proc. Natl. Acad. Sci. USA. 1968. V. 60. P. 1275–1281.
8. Browne S.M., Al-Rubeai M. // Trends Biotech. 2007. V. 25. P. 425–432.
9. Folger K.R., Wong E.A., Wahl G., Capecchi M.R. // Mol. Cell. Biol. 1982. V. 2. P. 1372–1387.
10. Folger K.R., Thomas K., Capecchi M.R. // Mol. Cell Biol. 1985. V. 5. P. 59–69.
11. Robins D.M., Ripley S., Henderson A.S., Axel R. // Cell. 1981. V. 23. P. 29–39.
12. Thomas K.R., Folger K.R., Capecchi M.R. // Cell. 1986.

V. 44. P. 419–428.
13. Hayward B.E., Hussain A., Wilson R.H., Lyons A., Woodcock V., McIntosh B., Harris T.J. // Nucl. Acids Res. 1986. V. 14. P. 999–1008.
14. Wuest D.M., Harcum S.W., Lee K.H. // Biotechnol. Adv. 2012. V. 30. P. 629–638.
15. Urlaub G., Chasin L.A. // Proc. Natl. Acad. Sci. USA. 1980. V. 77. P. 4216–4220.
16. Urlaub G., Kas E., Carothers A.M., Chasin L.A. // Cell. 1983. V. 33. P. 405–412.
17. Inouye S., Tsuji F.I. // FEBS Lett. 1994. V. 341. P. 277–280.
18. Stepanenko O.V., Verkhusha V.V., Kuznetsova I.M., Uversky V.N., Turoverov K.K. // Curr. Protein Pept. Sci. 2008. V. 9. P. 338–369.
19. Niu J., Zhang B., Chen H. // Mol. Biotechnol. 2014. V. 56. P. 681–688.
20. Wijshake T., Baker D.J., van de Sluis B. // Biochim. Biophys. Acta. 2014. V. 1842. P. 1942–1950.
21. Castel S.E., Martienssen R.A. // Nat. Rev. Genet. 2013. V. 14. P. 100–112.
22. Fatica A., Bozzoni I. // Nat. Rev. Genet. 2014. V. 15. P. 7–21.
23. Morris K.V., Mattick J.S. // Nat. Rev. Genet. 2014. V. 15. P. 423–437.
24. Lai T., Yang Y., Ng S.K. // Pharmaceuticals (Basel). 2013. V. 6. P. 579–603.
25. Bradley A.J., Lurain N.S., Ghazal P., Trivedi U., Cunningham C., Baluchova K., et al. // J. Gen. Virol. 2009. V. 90. P. 2375–2380.
26. Goldberg M.L. Sequence Analysis of Drosophila Histone Genes: PhD thesis. Stanford University, 1979.
27. Corden J., Wasylyk B., Buchwalder A., Sassone-Corsi P., Kedinger C., Chambon P. // Science. 1980. V. 209. P. 1406–1414.
28. Burke T.W., Kadonaga J.T. // Genes Dev. 1996. V. 10. P. 711–724.
29. Lagrange T., Kapanidis A.N., Tang H., Reinberg D., Ebright R.H. // Genes Dev. 1998. V. 12. P. 34–44.
30. Lewis B.A., Kim T.K., Orkin S.H. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. P. 7172–7177.
31. Lim C.Y., Santoso B., Boulay T., Dong E., Ohler U., Kadonaga J.T. // Genes Dev. 2004. V. 18. P. 1606–1617.
32. Maksimenko O., Georgiev P. // Front. Genet. 2014. V. 5. P. 28.
33. Xu X., Nagarajan H., Lewis N.E., Pan S., Cai Z., Liu X., Chen W., Xie M., Wang W., Hammond S., et al. // Nat. Biotechnol. 2011. V. 29. P. 735–741.
34. Becker J., Hackl M., Rupp O., Jakobi T., Schneider J., Szczepanowski R., Bekel T., Borth N., Goesmann A., Grillari J., et al. // J. Biotechnol. 2011. V. 156. P. 227–235.
35. Jakobi T., Brinkrolf K., Tauch A., Noll T., Stoye J., Pühler A., Goesmann A. // J. Biotechnol. 2014. V. 190. P. 64–75.
36. Running Deer J., Allison D.S. // Biotechnol. Prog. 2004. V. 20. P. 880–889.
37. Orlova N.A., Kovnir S.V., Hodak J.A., Vorobiev I.I., Gabibov A.G., Skryabin K.G. // BMC Biotechnol. 2014. V. 14. P. 56.
38. Mariati, Koh E.Y., Yeo J.H., Ho S.C., Yang Y. // Bioengineered. 2014. V. 5. P. 340–345.
39. Andersen C.R., Nielsen L.S., Baer A., Tolstrup A.B., Weilguny D. // Mol. Biotechnol. 2011. V. 48. P. 128–137.
40. de Groote M.L., Verschure P.J., Rots M.G. // Nucl. Acids Res. 2012. V. 40. P. 10596–10613.
41. Kwaks T.H., Otte A.P. // Trends Biotechnol. 2006. V. 24. P. 137–142.
42. Holmqvist P.H., Mannervik M. // Transcription. 2013. V. 4. P. 18–23.
43. Kwaks T.H., Sewalt R.G., van Blokland R., Siersma T.J., Kasiem M., Kelder A., Otte A.P. // J. Biotechnol. 2005. V. 115. P. 35–46.
44. Moltó E., Fernández A., Montoliu L. // Brief Funct. Genomic Proteomic. 2009. V. 8. P. 283–296.
45. Palazzoli F., Bire S., Bigot Y., Bonnin-Rouleux F. // Nat. Biotechnol. 2011. V. 29. P. 593–597.
46. Kim A., Dean A. // Mol. Cells. 2012. V. 34. P. 1–5.
47. Vishwanathan N., Le H., Le T., Hu W.S. // Curr. Opin. Biotechnol. 2014. V. 30. P. 113–119.
48. Recillas-Targa F., Valadez-Graham V., Farrell C.M. // BioEssays. 2004. V. 26. P. 796–807.
49. Emery D.W. // Hum. Gene Ther. 2011. V. 22. P. 761–774.
50. Furlan-Magaril M., Rebollar E., Guerrero G., Fernández A., Moltó E., González-Buendía E., Cantero M., Montoliu L., Recillas-Targa F. // Nucl. Acids Res. 2011. V. 39. P. 89–103.
51. Bire S., Ley D., Casteret S., Mermod N., Bigot Y., Rouleux-Bonnin F. // PLoS One. 2013. V. 8. e82559.
52. Chetverina D., Aoki T., Erokhin M., Georgiev P., Schedl P. // BioEssays. 2014. V. 36. P. 163–172.
53. Kyrchanova O., Georgiev P. // FEBS Lett. 2014. V. 588. P. 8–14.
54. Labrador M., Corces V.G. // Cell. 2002. V. 111. P. 151–154.
55. Bell A.C., West A.G., Felsenfeld G. // Science. 2001. V. 291. P. 447–450.
56. Handoko L., Xu H., Li G., Ngan C.Y., Chew E., Schnapp M., Lee C.W., Ye C., Ping J.L., Mulawadi F., et al. // Nat. Genet. 2011. V. 43. P. 630–638.
57. Dixon J.R., Selvaraj S., Yue F., Kim A., Li Y., Shen Y., Hu M., Liu J.S., Ren B. // Nature. 2012. V. 485. P. 376–380.
58. Sanyal A., Lajoie B.R., Jain G., Dekker J. // Nature. 2012. V. 489. P. 109–113.
59. Sexton T., Yaffe E., Kenigsberg E., Bantignies F., Leblanc B., Hoichman M., Parrinello H., Tanay A., Cavalli G., et al. // Cell. 2012. V. 148. P. 458–472.
60. Holwerda S., de Laat W. // Front. Genet. 2012. V. 3. P. 217.
61. Nora E.P., Dekker J., Heard E. // BioEssays. 2013. V. 35. P. 818–828.
62. Gibcus J.H., Dekker J. // Mol. Cell. 2013. V. 49. P. 773–782.
63. Kravchenko E., Savitskaya E., Kravchuk O., Parshikov A., Georgiev P., Savitsky M. // Mol. Cell Biol. 2005. V. 25. P. 9283–9291.
64. Li H.B., Müller M., Bahechar I.A., Kyrchanova O., Ohno K., Georgiev P., PirrottaV. // Mol. Cell Biol. 2011. V. 31. P. 616–625.
65. Venken K.J., Bellen H.J. // Meth. Mol. Biol. 2012. V. 859. P. 203–228.
66. Fujioka M., Wu X., Jaynes J.B. // Development. 2009. V. 136. P. 3077–3087.
67. Fujioka M., Sun G., Jaynes J.B. // PLoS Genet. 2013. V. 9. e1003883.
68. Tanay A., Cavalli G. // Curr. Opin. Genet. Dev. 2013. V. 23. P. 197–203.
69. Ameres S.L., Drueppel L., Pfleiderer K., Schmidt A., Hillen W., Berens C. // EMBO J. 2005. V. 24. P. 358–367.
70. Bondarenko V.A., Jiang Y.I., Studitsky V.M. // EMBO J. 2003. V. 22. P. 4728–4737.
71. Kyrchanova O., Maksimenko O., Stakhov V., Ivlieva T., Parshikov A., Studitsky V.M., Georgiev P. // PLoS Genet. 2013. V. 9. e1003606.
72. Savitskaya E., Melnikova L., Kostuchenko M., Kravchenko E., Pomerantseva E., Boikova T., Chetverina D., Parshikov A., Zobacheva P., Gracheva E., et al. // Mol. Cell Biol. 2006. V. 26. P. 754–761.
73. Erokhin M., Davydova A., Kyrchanova O., Parshikov A.,

Georgiev P., Chetverina D. // Development. 2011. V. 138. P. 4097–4106.
74. Nakayama T., Shimojima T., Hirose S. // Development. 2012. V. 139. P. 4582–4590.
75. Ghirlando R., Giles K., Gowher H., Xiao T., Xu Z., Yao H., Felsenfeld G. // Biochim. Biophys. Acta. 2012. V. 1819. P. 644–651.
76. Vorobyeva N.E., Mazina M.U., Golovnin A.K., Kopytova D.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Georgiev P.G., Krasnov A.N. // Nucl. Acids Res. 2013. V. 41. P. 5717–5730.
77. Yajima M., Fairbrother W.G., Wessel G.M. // Development. 2012. V. 139. P. 3613–3622.
78. Li M., Belozerov V.E., Cai H.N. // Mol. Cell Biol. 2010. V. 30. P. 1067–1076.
79. Herold M., Bartkuhn M., Renkawitz R. // Development. 2012. V. 139. P. 1045–1057.
80. Merkenschlager M., Odom D.T. // Cell. 2013. V. 152. P. 1285–1297.
81. Chaumeil J., Skok J.A. // Curr. Opin. Immunol. 2012. V. 24. P. 153–159.
82. Nakahashi H., Kwon K.R., Resch W., Vian L., Dose M., Stavreva D., Hakim O., Pruett N., Nelson S., Yamane A., et al. // Cell Rep. 2013. V. 3. P. 1678–1689.
83. Yusufzai T.M., Felsenfeld G. // Proc. Natl. Acad. Sci. USA. 2004. V. 101. P. 8620–8624.
84. Pant V., Kurukuti S., Pugacheva E., Shamsuddin S., Mariano P., Renkawitz R., Klenova E., Lobanenkov V., Ohlsson R. // Mol. Cell Biol. 2004. V. 24. P. 3497–3504.
85. Chernukhin I.V., Shamsuddin S., Robinson A.F., Carne A.F., Paul A., El-Kady A.I., Lobanenkov V.V., Klenova E.M. // J. Biol. Chem. 2000. V. 275. P. 29915–29921.
86. Lutz M., Burke L.J., Barreto G., Goeman F., Greb H., Arnold R., Schultheiss H., Brehm A., Kouzarides T., Lobanenkov V., et al. // Nucl. Acids Res. 2000. V. 28. P. 1707–1713.
87. Ishihara K., Oshimura M., Nakao M. // Mol. Cell. 2006. V. 23. P. 733–742.
88. Mehta G.D., Kumar R., Srivastava S., Ghosh S.K. // FEBS Lett. 2013. V. 587. P. 2299–2312.
89. Lee B.K., Iyer V.R. // J. Biol. Chem. 2012. V. 287. P. 30906–30913.
90. Xiao T., Wallace J., Felsenfeld G. // Mol. Cell Biol. 2011. V. 31. P. 2174–2183.
91. Parelho V., Hadjur S., Spivakov M., Leleu M., Sauer S., Gregson H.C., Jarmuz A., Canzonetta C., Webster Z., Nesterova T., et al. // Cell. 2008. V. 132. P. 422–433.
92. Wendt K.S., Yoshida K., Itoh T., Bando M., Koch B., Schirghuber E., Tsutsumi S., Nagae G., Ishihara K., Mishiro T., et al. // Nature. 2008. V. 451. P. 796–801.
93. Hadjur S., Williams L.M., Ryan N.K., Cobb B.S., Sexton T., Fraser P., Fisher A.G., Merkenschlager M. // Nature. 2009. V. 460. P. 410–413.
94. Nativio R., Wendt K.S., Ito Y., Huddleston J.E., Uribe-Lewis S., Woodfine K., Krueger C., Reik W., Peters J.M., Murrell A. // PLoS Genet. 2009. V. 5. e1000739.
95. Zuin J., Dixon J.R., van der Reijden M.I., Ye Z., Kolovos P., Brouwer R.W., van de Corput M.P., van de Werken H.J., Knoch T.A., van IJcken W.F., et al. // Proc. Natl. Acad. Sci. USA. 2014. V. 111. P. 996–1001.
96. Gosalia N., Neems D., Kerschner J.L., Kosak S.T., Harris A. // Nucl. Acids Res. 2014. V. 42. P. 9612–9622.
97. Chung J.H., Bell A.C., Felsenfeld G. // Proc. Natl. Acad. Sci. USA. 1997. V. 94. P. 575–580.
98. Emery D.W., Yannaki E., Tubb J., Stamatoyannopoulos G. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. P. 9150–9155.
99. Bell A.C., West A.G., Felsenfeld G. // Cell. 1999. V. 98. P. 387–396.
100. West A.G., Huang S., Gaszner M., Litt M.D., Felsenfeld G. // Mol. Cell. 2004. V. 16. P. 453–463.
101. Huang S., Li X., Yusufzai T.M., Qiu Y., Felsenfeld G. // Mol. Cell Biol. 2007. V. 27. P. 7991–8002.
102. Dickson J., Gowher H., Strogantsev R., Gaszner M., Hair A., Felsenfeld G., West A.G. // PLoS Genet. 2010. V. 6. e1000804.
103. Giles K.E., Gowher H., Ghirlando R., Jin C., Felsenfeld G. // Cold Spring Harbor Symp. Quant. Biol. 2010. V. 75. P. 1–7.
104. Maksimenko O.G., Deikin A.V., Khodarovich Yu.M., and Georgiev P.G. // Acta Naturae. 2013. V. 5. № 1(16). P. 33–47.
105. Rival-Gervier S., Pantano T., Viglietta C., Maeder C., Prince S., Attal J., Jolivet G., Houdebine L.M. // Transgenic Res. 2003. V. 12. P. 723–730.
106. Hanawa H., Yamamoto M., Zhao H., Shimada T., Persons D.A. // Mol. Therapy. 2009. V. 17. P. 667–674.
107. Aker M., Tubb J., Groth A.C., Bukovsky A.A., Bell A.C., Felsenfeld G., Kiem H.P., Stamatoyannopoulos G., Emery D.W. // Hum. Gene Ther. 2007. V. 18. P. 333–343.
108. Kim J.M., Kim J.S., Park D.H., Kang H.S., Yoon J., Baek K., Yoon Y. // J. Biotech. 2004. V. 107. P. 95–105.
109. Girod P.A., Zahn-Zabal M., Mermod N. // Biotech. Bioeng. 2005. V. 91. P. 1–11.
110. Harraghy N., Gaussin A., Mermod N. // Curr. Gene Ther. 2008. V. 8. P. 353–366.
111. Wang T.Y., Han Z.M., Chai Y.R., Zhang J.H. // Mol. Biol. Rep. 2010. V. 37. P. 3553–2560.
112. Bode J., Schlake T., Rios-Ramirez M., Mielke C., Stengert M., Kay V., Klehr-Wirth D. // Int. Rev. Cytol. 1995. V. 162. P. 389–454.
113. Platts A.E., Quayle A.K., Krawetz S.A. // Cell Mol. Biol. Lett. 2006. V. 11. P. 191–213.
114. Evans K., Ott S., Hansen A., Koentges G., Wernisch L. // BMC Bioinformatics. 2007. V. 8. P. 71.
115. Liebich I., Bode J., Frisch M., Wingender E. // Nucl. Acids Res. 2002. V. 30. P. 372–374.
116. Girod P.-A., Nguyen D.-Q., Calabrese D., Puttini S., Grandjean M., Martinet D., Regamey A., Saugy D., Beckmann J.S., Bucher P., et al. // Nat. Methods. 2007. V. 4. P. 747–753.
117. Welcsh P.L., King M.C. // Hum. Mol. Genet. 2001. V. 10. P. 705–713.
118. Iarovaia O.V., Shkumatov P., Razin S.V. // J. Cell Sci. 2004. V. 117. P. 4583–4590.
119. Kulkarni A., Pavithra L., Rampalli S., Mogare D., Babu K., Shiekh G., Ghosh S., Chattopadhyay S. // Biochem. Biophys. Res. Commun. 2004. V. 322. P. 672–677.
120. Johnson C.N., Levy L.S. // Virol. J. 2005. V. 2. P. 68.
121. Koina E., Piper A. // J. Cell Biochem. 2005. V. 95. P. 391–402.
122. Mearini G., Chichiarelli S., Zampieri M., Masciarelli S., D'Erme M., Ferraro A., Mattia E. // FEBS Lett. 2003. V. 547. P. 119–124.
123. Piechaczek C., Fetzer C., Baiker A., Bode J., Lipps H.J. // Nucl. Acids Res. 1999. V. 27. P. 426–428.
124. Galbete J.L., Buceta M., Mermod N. // Mol. Biosyst. 2009. V. 5. P. 143–150.
125. Bode J., Benham C., Knopp A., Mielke C. // Crit. Rev. Eukaryot. Gene Expr. 2000. V. 10. P. 73–90.
126. Berezney R., Coffey D.S. // Biochem. Biophys. Res. Commun. 1974. V. 60. P. 1410–1417.
127. Albrethsen J., Knol J.C., Jimenez C.R. // J. Proteomics. 2009. V. 72. P. 71–81.

128. Cai S., Han H.J., Kohwi-Shigematsu T. // Nat. Genet. 2003. V. 34. P. 42–51.
129. Renz A., Fackelmayer F.O. // Nucl. Acids Res. 1996. V. 24. P. 843–849.
130. Agrelo R., Souabni A., Novatchkova M., Haslinger C., Leeb M., Komnenovic V., Kishimoto H., Gresh L., Kohwi-Shigematsu T., Kenner L., et al. // Dev. Cell. 2009. V. 16. P. 507–516.
131. Arope S., Harraghy N., Pjanic M., Mermod N. // PLoS One. 2013. V. 8. e79262.
132. Berrios M., Osheroff N., Fisher P.A. // Proc. Natl. Acad. Sci. USA. 1985. V. 82. P. 4142–4146.
133. Feister H.A., Onyia J.E., Miles R.R., Yang X., Galvin R., Hock J.M., Bidwell J.P. // Bone. 2000. V. 26. P. 227–234.
134. Adachi Y., Kas E., Laemmli U.K. // EMBO J. 1989. V. 8. P. 3997–4006.
135. Kohwi-Shigematsu T., Kohwi Y., Takahashi K., Richards H.W., Ayers S.D., Han H.J., Cai S. // Methods. 2012. V. 58. P. 243–254.
136. Burute M., Gottimukkala K., Galande S. // Immunol. Cell Biol. 2012. V. 90. P. 852–859.
137. Naito T., Tanaka H., Naoe Y., Taniuchi I. // Int. Immunol. 2011. V. 23. P. 661–668.
138. Yasui D., Miyano M., Cai S., Varga-Weisz P., Kohwi-Shigematsu T. // Nature. 2002. V. 419. P. 641–645.
139. Hasegawa Y., Brockdor V.N., Kawano S., Tsutsui K., Nakagawa S. // Dev. Cell. 2010. V. 19. P. 469–476.
140. Tattermusch A., Brockdorff N. // Hum. Genet. 2011. V. 130. P. 247–253.
141. Razin S.V., Iarovaia O.V., Vassetzky Y.S. // Chromosoma. 2014. V. 123. P. 217–224.
142. Harraghy N., Buceta M., Regamey A., Girod P.A., Mermod N. // Methods Mol. Biol. 2012. V. 801. P. 93–110.
143. Zahn-Zabal M., Kobr M., Girod P.A., Imhof M., Chatellard P., de Jesus M. // J. Biotech. 2001. V. 87. P. 29–42.
144. Harraghy N., Regamey A., Girod P.-A., Mermod N. // J. Biotechnol. 2011. V. 154. P. 11–20.
145. Buceta M., Galbete J.L., Kostic C., Arsenijevic Y., Mermod N. // Gene Therapy. 2011. V. 18. P. 7–13.
146. Ley D., Harraghy N., Le Fourn V., Bire S., Girod P.A., Regamey A., Rouleux-Bonnin F., Bigot Y., Mermod N. // PLoS One. 2013. V. 8. e62784.
147. Grandjean M., Girod P.A., Calabrese D., Kostyrko K., Wicht M., Yerly F., Mazza C., Beckmann J.S., Martinet D., Mermod N. // Nucl. Acids Res. 2011. V. 39. e104.
148. Puttini S., van Zwieten R.W., Saugy D., Lekka M., Hogger F., Ley D., Kulik A.J., Mermod N. // BMC Mol. Biol. 2013. V. 14. P. 26.
149. Shimizu N., Miura Y., Sakamoto Y., Tsutsui K. // Cancer Res. 2001. V. 61. P. 6987–6990.
150. Noguchi C., Araki Y., Miki D., Shimizu N. // PLoS One. 2012. V. 7. . e52990.
151. Antoniou M., Harland L., Mustoe T., Williams S., Holdstock J., Yague E., Mulcahy T., Griffiths M., Edwards S., Ioannou P.A., et al. // Genomics. 2003. V. 82. P. 269–279.
152. Williams S., Mustoe T., Mulcahy T., Griffiths M., Simpson D., Antoniou M., Irvine A., Mountain A., Crombie R. // BMC Biotech. 2005. V. 5. P. 17.
153. Benton T., Chen T., McEntee M., Fox B., King D., Crombie R., Thomas T.C., Bebbington C. // Cytotechnology. 2002. V. 38. P. 43–46.
154. Zhang F., Thornhill S.I., Howe S.J., Ulaganathan M., Schambach A., Sinclair J., Kinnon C., Gaspar H.B., Antoniou M., Thrasher A.J. // Blood. 2007. V. 110. P. 1448–1457.
155. Zhang F., Frost A.R., Blundell M.P., Bales O., Antoniou M.N., Thrasher A.J. // Mol. Ther. 2010. V. 18. P. 1640–1649.
156. Dighe N., Khoury M., Mattar C., Chong M., Choolani M., Chen J., Antoniou M.N., Chan J.K. // PLoS One. 2014. V. 9. e104805.
157. Pfaff N., Lachmann N., Ackermann M., Kohlscheen S., Brendel C., Maetzig T., Niemann H., Antoniou M.N., Grez M., Schambach A., et al. // Stem Cells. 2013. V. 31. P. 488–499.
158. Brendel C., Müller-Kuller U., Schultze-Strasser S., Stein S., Chen-Wichmann L., Krattenmacher A., Kunkel H., Dillmann A., Antoniou M.N., Grez M. // Gene Ther. 2012. V. 19. P. 1018–1029.
159. Boscolo S., Mion F., Licciulli M., Macor P., De Maso L., Brce M., Antoniou M.N., Marzari R., Santoro C., Sblattero D. // Nat. Biotechnol. 2012. V. 29. P. 477–484.
160. Lindahl A.M., Antoniou M. // Epigenetics. 2007. V. 2. P. 227–236.
161. Nair A.R., Jinger X., Hermiston T.W. // BMC Res. Notes. 2011. V. 4. P. 178.
162. Jonuschies J., Antoniou M., Waddington S., Boldrin L., Muntoni F., Thrasher A., Morgan J. // Curr. Gene Ther. 2014. V. 14. P. 276–288.
163. Otte A.P., Kwaks T.H., van Blokland R.J., Sewalt R.G., Verhees J., Klaren V.N., Siersma T.K., Korse H.W., Teunissen N.C., Botschuijver S., et al. // Biotechnol. Prog. 2007. V. 23. P. 801–807.
164. Hoeksema F., Hamer K., Siep M., Verhees J.A., Otte A.P. // Biotechnol. Res. Int. 2011. 492875.
165. Wang J., Lawry S.T., Cohen A.L., Jia S. // Cell Mol. Life Sci. 2014. V. 71. P. 4841–4852.
166. Kwaks T.H.J., Barnett P., Hemrika W., Siersma T., Sewalt R.G., Satijn D.P., Brons J.F., van Blokland R., Kwakman P., Kruckeberg A.L., et al. // Nat. Biotechnology. 2003. V. 21. P. 553–558.
167. Schwartz Y.B., Linder-Basso D., Kharchenko P.V., Tolstorukov M.Y., Kim M., Li H.B., Gorchakov A.A., Minoda A., Shanower G., Alekseyenko A.A., et al. // Genome Res. 2012. V. 22. P. 2188–2198.