Russian Journal of Bioorganic Chemistry, Vol. 30, No. 1, 2004, pp. 1–11. Translated from Bioorganicheskaya Khimiya, Vol. 30, No. 1, 2004, pp. 3–14.

REVIEW ARTICLES

Structure and Functions of Nuclear Matrix Associated Regions (S/MARs)

I. P. Chernov, S. B. Akopov, and L. G. Nikolaev*, [1]

Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, GSP Moscow, 117997 Russia

Received March 22, 2002; in final form, July 15, 2002

Abstract—Modern concepts on the chromatin loop–domain organization and the role of the DNA regions specifically binding the nuclear matrix or nuclear scaffold (S/MARs) during its formation, maintenance, and regulation are discussed. Some S/MAR structural features, properties of binding the nuclear matrix, and probable mechanisms of their involvement in the gene regulation of activity are considered.

Key words: *chromatin, loop domains, MAR, nuclear matrix, nuclear scaffold, SAR*

PACKING OF GENETIC MATERIAL WITHIN CELL NUCLEUS

General principles of DNA organization in chromatin have been formulated in the late 1970s and have not been essentially revised since the time [1–3].[2]

Three levels of DNA compaction in the interphase nucleus are usually differentiated.

1. *In nucleosome*, DNA is associated with histones H2A, H2B, H3, and H4, which form a nucleosome protein core (containing two molecules of each protein), and histone H1, which is affine to linker sites. DNA is disposed on the nucleosome surface and forms 1.75 supercoils around the core, which corresponds to a DNA fragment of 140–145-bp length. Individual nucleosomes are connected by linkers of variable lengths, and, therefore, the length of the full nucleosome repeat can vary from 155 to 210 bp [4].

2. *Fibers* of approximately 30 nm in diameter formed with the involvement of histone H1 constitute the next level [5]. There are three models of compaction of 30-nm fiber: solenoid [6], nucleomeric [7], and the model based on layered zigzag [8, 9]. The detailed structure of the 30-nm fiber is obscure up to now, although most of researches are inclined to agree with the solenoid model.

3. A *loop domain* level of chromatin organization. Electron microscopic studies suggest that 30 nm chromatin fibers are arranged in series of supercoiled loops (domains) attached by their bases to the protein component of nuclear envelope [1, 10–12]. This level of chromatin organization will be discussed below.

NUCLEAR MATRIX AND NUCLEAR SCAFFOLD

The backbone structure, which retains the general shape of cell nucleus after the removal of the major part of DNA and DNA-associated proteins, was called nuclear matrix [13]. It mainly consists of proteins, but has also a small amount of DNA (see below), RNA, and other components [14]. The nuclear matrix was presumed to be the particular structural component of the interphase nucleus that arranges the 30-nm chromatin fiber into loop domains.

Two methods of nuclear matrix preparation are widely used these days.

The first method [13, 15] is based on treatment of isolated cell nuclei with DNase I and subsequent removal of the major part of intranuclear proteins and DNA by washing with a buffer containing 2 M NaCl.

An alternative method of preparation of nuclear matrix, which was called by the authors as nuclear scaffold, is based on the action of lithium 3,5-diiodosalicylate [16]. In this case, greater part of nuclear proteins is removed under milder conditions, i.e., by washing of isolated cell nuclei with a low-salt buffer containing LIS, and restriction endonucleases are used for the DNA removal from the histone-depleted nuclei.

According to electron microscopy, the presence of three structural components is characteristic of nuclear matrix: residual nuclear lamina with pore complexes, a residual nucleolus, and a fibrillar-granular structure that fills the intranuclear space and is called internal matrix [14, 15].

[1] Corresponding author; phone: +7 (095) 330-7029; fax: +7 (095) 330-6538; e-mail: lev@humgen.siobc.ras.ru.

[2] Abbreviations: ARBP, attachment region binding protein; LAS, loop attachment site; LIS, lithium 3,5-diiodosalicylate; LTR, long terminal repeat; MAR, SAR, and S/MAR, matrix (scaffold) associated region; nmDNA, nuclear matrix DNA; MRS, MAR/SAR recognition signature; SAF-A and SAF-B, scaffold attachment factors A and B; and SATB, special AT-rich sequence-binding protein.

One more approach for the identification of structures forming nuclear matrix under the conditions close to physiological was later proposed [17, 18]. The method is based on the electrophoretic removal of the most of nuclear components from the cells embedded in 0.5% agarose. The encapsulated cells are lyzed under the conditions close to physiological, and the cell lyzate immobilized in agarose gel is treated with DNase I or restriction endonucleases and subjected to electrophoresis. The relatively low-molecular components (proteins and chromatin fragments) are easily passed through the agarose layer, whereas residual nuclei (and other large cell structures) remained fixed in the gel. The ultrastructure of residual nuclei obtained in this manner is close to that of nuclear matrix from the isolated nuclei.

Metaphase scaffold is the component of metaphase chromosomes, which is analogous in composition and probable functions to interphase nuclear matrix (scaffold) [19].

DNA WITHIN NUCLEAR MATRIX AND ITS PROPERTIES

Definition of nmDNA and Methods for Its Preparation

The DNA fragments tightly bound to the protein backbone and stable toward action of high concentrations of salts and nucleases are remained in nuclear matrix. Content of the DNA tightly bound in nuclear matrix may vary in dependence on its preparation method. A small amount (1–3%) of DNA tightly bound to nuclear matrix (nmDNA) is retained after a prolonged action of nucleases and stringent conditions of its isolation [13]. The nmDNA is supposed to be included in nuclear matrix as mono- and oligonucleosomes [20].

The appearance of loop domain hypothesis of the interphase chromatin structure [1, 10] excited the interest to studying nmDNA as a DNA fraction containing potential regions of attachment to nuclear matrix of loop domain bases. It was attempted to clone and analyze this DNA directly from nuclear matrix that was prepared by either salt extraction [21] or by LIS [22]. The major disadvantage of this approach is that a redistribution and capture of some *in vivo* extrinsic DNA fragments is possible during the process of nuclear matrix isolation. The further analysis does not allow the distinction of such fragments from actual nmDNA.

The approach using the cell encapsulation in low-melting agarose (see above) is devoid of some of the drawbacks. The presumed sites of attachment of loop domains identified by this method were called LAS (loop attachment sites) [18]. However, this method has not found a wide use probably due to its complexity and a low reproducibility.

S/MAR Elements

Another method of identification of the DNA sequences presumably responsible for the attachment of bases of the chromatin loop domains to nuclear matrix received a significantly wider distribution. It is based on binding of a specific fraction of the exogenous DNA fragments to the *in vitro* isolated nuclear matrix (scaffold) in the presence of a large excess of prokaryotic DNA. Such fragments for the matrix obtained by treatment with 2 M NaCl were called matrix-associated or matrix-attachment regions (MARs) [23]. It was later shown that the nuclear scaffold obtained by LIS extraction can also bind *in vitro* to some DNA fragments, and the corresponding genomic DNA fragments were called scaffold-associated regions (SARs) [24]. A large number of publications concerning these elements did not reveal any basic differences between SARs and MARs, and many authors currently combine these sequences under the general name of S/MAR elements or S/MARs [25]. We will also follow this terminology. The main drawback of the *in vitro* binding method consists in the fact that it is often difficult to prove whether even the DNA fragments (S/MARs) highly affine to nuclear matrix are actually its *in vivo* components. Despite these disadvantages, this method is used most widely, and it is the majority of information on the potential regions of DNA attachment to nuclear matrix that was obtained for S/MAR elements.

Special Structural Features of nmDNA

The analysis of nucleotide sequences of S/MAR elements, nmDNA, and LAS of numerous human, mammalian, and plant genes failed to reveal extended homologous regions in them. However, in some cases, special features of certain S/MARs groups were found, and the attempts to classify them were made [25–28]. One of the S/MARs characteristics is their enrichment with AT pairs [15, 27, 29, 30]. The motifs of A box ($AATAAA^{T}/_{C}AAA$), T box ($TT^{A}/_{T}T^{T}/_{A}TT^{T}/_{A}TT$), as well as the sequences ATATTT (along with its variants ATATTTTT and AATATT) [31–33], TAAT, and TAAAT [27] are abundant among the AT-rich S/MAR elements. The tendency of AT-enriched sequences to conformational transitions under the torsional stress led to the concept of the stress-induced duplex destabilization (SIDD) [34, 35]. According to this concept, S/MAR elements contain the regions prone to easy melting (BUR, base unpairing regions). Under stress, the duplex destabilization starts from the region of the so-called core unwinding element (CUE) and gradually spreads over the whole S/MAR element, which increases its affinity to nuclear matrix [36].

Many S/MAR elements are enriched with the sequences that are potentially capable of forming non-canonical secondary structures of DNA. For example, the presence of inverted repeats allows the possible formation of cruciform structures [35, 37], to which some

specific proteins can bind [38]. S/MARs also contain potential regions of triplex formation [27] and direct repeats [39]. Some authors noticed DNA bends in the regions of their interaction with nuclear matrix [40, 41].

In many cases, the cleavage sites for topoisomerase II GTN$^{A}/_{T}$A$^{G}/_{T}$ATTNATNN$^{A}/_{G}$ were found inside the AT-rich regions of S/MARs elements [23, 31], and it was also shown that topoisomerase II is actually bound to them [42, 43]. Topoisomerase I, whose cleavage site also belongs to AT-enriched regions (AAAAAGACT-T↓AGAAAAATT) [44], is indirectly bound to matrix, via the actively transcribed sequences, to which both nuclear matrix and topoisomerase I have increased affinities [27].

The efforts to search for the consensus sequences for S/MAR elements led to the discovery of the MRS motif, which contained two degenerated sequences (AATAAYAA and AWWRTAANNWWGNNNC) located in an immediate vicinity to one another. All the cases of the MRS motif detection within a DNA sequence were found to be connected with the DNA fragment that could be specifically bound to nuclear matrix. An analysis of more than 300 kbp of genomic DNA from various eukaryotes demonstrated that the MRS motif precisely predicted up to 80% of the S/MAR elements [45]. However, it is far from all of the known S/MAR elements that have the MRS sequences. It is presumed that the S/MAR properties are the most probably determined by the common structural peculiarities of their DNA, such as the presence of AT-enriched regions and/or sites inclined to bending, rather than by the similarities in the S/MAR sequences themselves [46].

On the basis of common properties of S/MARs, several computer programs were developed for the their search; these took into account some characteristic motifs of S/MAR elements [47–49]. The usefulness of these programs for the search for S/MARs within the mammalian and plant genome DNA is still obscure due to the lack of a sufficient number of identified and mapped S/MAR elements. The database of S/MAR elements (S/MARt DB) contains less than 400 S/MAR elements, and the precise position of the most of them in genome is unknown [50].

Thus, the most characteristic structural features of animal and plant S/MAR elements that affect their specificity and the strength of interaction with nuclear matrix are:

(1) the number and size of the DNA double helix destabilization sites;

(2) the presence of inverted repeats;

(3) the presence of regions capable of formation of noncanonical secondary structures (DNA bends, triple helices, Z form, etc.); and

(4) the presence of cleavage sites for topoisomerase II.

At the same time, it should be noted that a considerable part of identified S/MARs that are highly affine to nuclear matrix do not meet these rules, which implies the existence of other classes of these elements [21, 51–54].

S/MARS FUNCTIONS

Formation of Loop Domains

The main function ascribed to S/MARs and similar sequences is the formation, maintenance, and regulation of functioning of interphase chromatin loop domains as well as the formation and maintenance of compact structure of metaphase chromosomes. The average size of mammalian loop domains is 100 kbp [17, 27, 55, 56], and, one can estimate from the full human genome size that the number of S/MARs in the human genome is approximately 30000. A somewhat different situation is observed in plants. An average size of loops is 17 for tobacco, 9 for rice and sorghum, 45 for maize, and 25 kbp for Arabidopsis [57, 58]. A hypothesis was proposed that relates the size of the loop domain with the level of transcriptional activity of its genes (e.g., the cluster of histone genes has the size of its domain of about 5000 bp) or with the rate of transcription activation of this domain in response to the corresponding action (the domains of genes of interferon I have from 3000 to 14000 bp in length) [59]. Unlike these short loops, more extended domains contain genes that are transcribed only at the definite development stages or possess strict tissue specificity.

The number of attachment sites and, hence, the loop domain lengths depend on a cell type and state. In particular, some DNA sites bound to matrix in the interphase lose this ability when scaffold of metaphase chromosomes is formed, and restore it two hours after mitosis [60]. The study of dependences between the numbers of replicons and chromatin loops during the development of *Xenopus* showed that the nuclear matrix at the blastula and gastrula stages binds more S/MAR elements than the matrices at later stages of development and of adults. In other words, some of the S/MAR elements lose their functions during the processes of cell differentiation and organism development [61]. The fibroblast transformation by SV40 virus also results in the change of the domain length and the distribution of matrix-associated DNA fragments [62]. A striking example of tissue specificity is the S/MAR element binding to nuclear matrix only after the activation of T cells [63].

It was shown that a new S/MAR element was formed after heat shock in the region of the promoter of heat-shock protein gene (murine *hsp70*). This element changes the domain organization and induces expression of this gene that plays a key role in eliminating the heat-shock consequences [64]. The heat shock results in increased total number of binding sites of DNA to nuclear matrix; in particular, the 5'-S/MAR element is

stabilized and an additional S/MAR element at the 3'-end of the adenosine deaminase gene is formed [65].

Tissue specific S/MARs were also found in other genes. One of the most thoroughly studied in this respect is a 47.5-kbp domain containing the human apolipoprotein B gene (*apoB*), which is predominantly expressed in the liver and intestine cells. The domain is bordered from 3'- and 5'-ends with two S/MAR elements. The third (proximal) S/MAR element was found in this domain; it is bound to the matrix only in liver cells (HepG2), where the *apoB* gene is effectively transcribed and is unbound in the HeLa cell line [66]. Further experiments on the transient expression and stable integration into the genome confirmed that the S/MAR elements of *apoB* gene display an insulator function and have no enhancer activity [67, 68]. In addition, the S/MAR of the gene of avian *malic* enzyme is tightly bound to nuclear matrix in thymocytes (where it is active) but is not bound in reticulocytes [69].

These data led to the conclusion that S/MARs could be subdivided into two types [70, 71]. The S/MAR elements of the first type are called by various authors as *stable*, *constitutive*, or *structural* elements; they ensure the attachment of the loop bases to nuclear matrix and the formation of domains irrespective of the tissue type, cell, and the stage of cell cycle. The S/MAR elements of the second type (*dynamic*, *facultative*, *functional*, or *tissue specific*) fulfill a regulatory function by forming temporary loop domains that are necessary either at definite stages of cell life or for cells of a certain type [25, 72]. This attractive hypothesis is extensively discussed, but has not been proved until now. The information on the tissue-specific character of the S/MARs-recognizing protein expression (see below) also supports this hypothesis in addition to the above-discussed facts.

Neutralization of Position Effect

As discussed above, the DNA sites affine to nuclear matrix are often localized in noncoding genome regions and, in this manner, designate the presumed boundaries of chromatin domains. The idea of independent domains got a new impulse after S/MARs were found to be able to neutralize the effect of the position and enhance transcription of the genes randomly integrated into genome.

The level of transcription of the genes that are steadily integrated into the cell genomes in cultures and into transgenic organisms is generally unpredictable and depends on the integration site. It was shown that the presence of S/MARs at the boundaries of integrated gene enhances its transcription and made it less dependent on the integration site [73–76]. The integration of the complete domain containing the human β-interferon gene together with S/MARs surrounding it 20–30-fold increases the level of its transcription in comparison with the construct lacking S/MARs [77]. In the most cases, such effects are observed only at a stable integration of the construct into the host cell genome and are absent or are dramatically weakened in the case of transient expression [35, 75, 78–80].

The sequences called insulators display a similar biological activity; i.e., they neutralize the position effect and define the borders of functional chromatin domains [81, 82]. There are some indications that the mechanism of action of insulators involves their interaction with nuclear matrix. In fact, the same DNA fragment displayed both the S/MARs and insulator activities [83–85] at least in some cases. However, some data also demonstrate that these activities can be separated [86] or are only observed for certain genetic constructs [87]. Thus, the problem of probable S/MARs–insulator relationship has not been solved up to now and requires the identification and comparative analysis of a greater number of these elements.

Another possible mechanism of the S/MARs effect on transcription involves methylation. If S/MAR from human β-interferon is inserted into a retroviral vector based on Moloney murine leukemia virus (Mo-MuLV) directly upstream the 3'-terminal LTR, the expression of the reporter gene remains stable for a long time (more than four months), and no LTR methylation is observed. In the analogous construct lacking the S/MAR element, LTR is substantially methylated and the reporter gene is not expressed. This fact directly confirms that S/MAR elements can prevent methylation in transcriptionally active loci, thereby supporting their expression [88].

The abilities to neutralize the position effect and activate the transcription of transgene open new prospects for the use of S/MARs in transgenosis and gene therapy [89].

S/MARs as Integration Sites for Retroviral Vectors

The insertion of reporter genes into the genome of cultured cells by means of a retroviral vector using the method ensuring the single copy insertion was found to happen into the regions capable of binding to nuclear matrix *in vitro* (S/MARs) and surrounded by DNA prone to bending in all the examined transformants [36, 51]. These S/MARs had atypical sequences: they were devoid of extended AT-enriched regions and, on the contrary, were enriched with CA/TG and AG/CT pairs.

S/MARs within Retroelements

The potential nuclear matrix binding sites were found in retrotransposable repeated genome elements, particularly, in long interspersed elements (LINE) [53, 90–92] and RTVL-Ia (L.G. Nikolaev *et al.*, unpublished data). Moreover, one of the sequences found in the clone library of chromatin LAS [18] had a high homology to LINE-1. Apparently, the sequences of plant miniature inverted repeat transposable elements (MITE) contain S/MAR [57, 93]. It is possible that the

presence of S/MARs can also be a characteristic feature of other retroposons. The transposition of S/MARs together with retroelements in the process of genome evolution can cause dramatic changes in the chromatin structure and in the regulation of genes located in the domains formed by them.

PROTEINS SPECIFICALLY BINDING S/MARS

The S/MARs binding to nuclear matrix appears to occur through the nuclear matrix proteins that can recognize and tightly bind S/MARs sequences. A number of such proteins were identified and characterized in recent years.

A protein specific for thymocytes, SATB1 (special AT-rich binding protein), was found [94]. It can bind with a high efficiency to the DNA minor groove of S/MAR elements, which contain A, T, and C-enriched regions in one of their nucleotide strands and have potentially low-melting BUR regions. This protein was shown to contain a 150-aa site that displays the S/MAR-binding properties of the total protein along with a new DNA-binding motif and an atypical homeodomain [95, 96]. It was later shown that SATB1 is a potent transcriptional suppressor of the genes surrounded by S/MAR elements and does not significantly affect the transcriptional level of genes surrounded with AT-rich sequences lacking properties of S/MARs [97]. Experiments on the mouse lines containing inactivated SATB1 gene revealed numerous defects in the development of T cells and considerable changes in expression levels of various genes [98]. Protein p114 from the carcinoma cells of mammary gland (SK-BR-30) displays similar properties. Interestingly, no S/MAR-binding activity of p114 was found in the cell culture of normal mammary gland [99].

The protein ARBP is cooperatively and specifically bound to S/MARs from the locus of the chicken lysozyme gene, as well as to mouse, fruit fly, and human S/MARs. It is preferably bound to the motif 5'-GGTGT surrounded with AT-rich sequences [100–102]. The ARBP protein, whose cDNA was cloned, is highly homologous to the rat protein that binds the methyl-CpG-binding protein (MeCP2). Both proteins contain a highly conserved domain binding S/MAR elements. The ARBP/MeCP2 was shown to be a multifunctional protein involved in the chromatin organization, the formation of loop domain structure, and the recognition of methylated DNA [103]. It was later shown that this protein interacts with mSin3A protein as a corepressor and forms a histone complex containing deacetylases when being bound to S/MAR elements. The deacetylase is supposed to affect S/MAR elements, which results in the formation of a local region of inactive chromatin [104].

The SP120 protein with molecular mass of 120 kDa was initially discovered as a protein cooperatively interacting with certain DNA fragments of the murine κ-immunoglobulin and the *fushi-tarazu* gene of fruit fly. It was reported that many properties, including the capability of cooperative DNA binding, of this protein and the ARBP protein are similar. SP120 and ARBP are currently regarded as different proteins, despite their obvious similarity [105].

SAF-A factor is present in the preparations of the heterogeneous nuclear ribonucleoproteins (hnRNP complex). It is involved in the RNA packing into RNP particles, and can form globular and filament complexes with DNA and RNA [106–108]. One of the SAF-A functions probably consists in its involvement in the formation of chromatin structure and in conservation of the structure during the interphase. The SAF-A protein is prone to the formation of homopolymers *in vitro*. One molecule of the protein was shown to bind S/MARs poorly, and only a simultaneous interaction of several molecules led to the powerful and strictly specific binding. The primary structure of SAF-A reflects its dual function. Its *C*-terminal domain binds RNA and single-stranded DNA, whereas its 45-aa site at *N*-terminus called SAF-box is responsible for the S/MARs binding [109]. The SAF-box is present in many eukaryotic proteins, and at least some of these proteins can also bind S/MARs. For example, SAF-box contains the SAF-B protein binding to S/MAR elements [110]. During apoptosis, the SAF-box is cleaved with caspase 3, and, as a result, the SAF-A loses the ability to bind S/MAR elements and is separated from the nuclear matrix [111]. A similar to SAF-A proteins SAF-B (or HET, Hsp27-ERE-TATA-binding protein) displays a clear and pronounced ability to selective bind the regulatory sites of some heat-shock protein genes and the DNA region containing the estrogen receptor binding site [112].

The B-cell-specific protein that regulates the transcription of immunoglobulin H gene (Bright, B cell regulator of IgH transcription) binds the regulatory S/MARs flanking the intron enhancer of the gene of immunoglobulin heavy chains; it activates the transcription of this gene only in its natural environment and only in B type cells. The protein contains the domain required for its tetramerization and subsequent binding of S/MARs [113, 114].

The homeodomain protein Cux/CDP (CAAT displacement protein) predominantly expressed in early B cells and never in mature B cells also demonstrates the capability of binding S/MAR elements [115]. This protein is believed to be a negative regulator of intron enhancer of immunoglobulin heavy chain genes, which is specific for the cell type and differentiation stage; the effect of Cux/CDP is mediated by the S/MARs that surround the enhancer. The activity of Cux/CDP allows to Wang *et al*. to regard it as an antagonist of Bright protein [116]. The study of transcription regulation of the human *CYP7A1* gene (cholesterol-7α hydroxylase) is another example supporting the hypothesis that CDP is

a transcriptional repressor directly interacting with intron S/MARs [117].

Recently, another S/MARs binding protein, SMAR1, was identified [118]; it has common binding sites and some common structural properties with SATB1, Cux/CDP, and Bright proteins.

Some plant proteins can also bind S/MARs specifically. Protein MFP1 (MAR binding filament-like protein 1), isolated from tomatoes, contains a DNA binding domain of new type [119, 120]. A multidomain protein AHM1 (AT hook-containing MAR binding protein) was recently found in tobacco cells and characterized [121].

DOMAIN STRUCTURE OF CHROMATIN IN GENOMIC CONTEXT

S/MARs and Loop Domains

According to the loop–domain model discussed above, S/MARs should flank certain genes or gene groups and form an independently regulated unit with a definite structure and function [66, 122]. Several attempts were made to localize S/MARs on genomic sequences in order to test this hypothesis. However, the accuracy of these attempts was limited due to a rather large sizes of the S/MAR-containing fragments, which were formed after hydrolysis with the restriction endonucleases (up to 10000 bp), whose lengths were substantially larger than the estimated lengths of S/MARs [27]. Therefore, one could only state in most cases that the identified S/MAR element was within the restriction fragment, but its exact position remained unknown. Moreover, the lack of overlapping clones for rather long genome regions enabled the S/MARs mapping only in single characterized loci, which contained, as a rule, single genes or families of related genes [123]. For example, four S/MAR elements were mapped in a 200-kbp locus containing the genes of constant regions of murine immunoglobulin heavy chains [124], and eight S/MAR elements were mapped in the 90-kbp locus of human β-globin genes [125].

Nevertheless, several S/MARs were mapped on the 320-kbp fruit fly genome region containing a number of unrelated genes. In this case, the lengths of the proposed domains were 26–112 kbp. These domains contained up to eight genes, some of which were transcribed. However, no correlation between activities of the genes and their location within the revealed domains was found [126].

The construction of extended contigs of cloned DNA enabled the S/MARs (nmDNA) mapping in the more than 800 kbp long locus of the fruit fly genome [127, 128]. There were identified 85 restrictive fragments co-isolated with nuclear matrix obtained by LIS extraction; 12 of them were found in the majority of studied cells (*strong* S/MARs), 44 in a fraction of the cells, and the remaining only in single cells. *Strong* S/MARs were located between the transcriptional units and divided the locus into loops of 15–115 kbp in size.

The method of identification of loop domains developed by Razin *et al.* deserves a special attention [129, 130]. It is based on the fact that, after a salt extraction of cell nuclei, the structural backbone of nuclei retained topoisomerase II directly bound to it. This enzyme retains its activity and directly contacts the DNA responsible for the attachment of loop bases to nuclear matrix. The cells were loaded into agarose blocks, treated with a nonionic detergent to enhance the membrane permeability, extracted with 2 M NaCl, incubated in the buffer used for topoisomerase II in the presence of epipodophyllotoxin VM-26, an inhibitor of the topoisomerase ligase activity. Under these conditions, the topoisomerase-catalyzed reaction inserts double-stranded breaks into the DNA molecule, which leads to a set of DNA fragments, whose ends are determined by topoisomerase-induced cleavage sites and are close to the loop bases. This method enabled identification of 11 cleavage sites in the ~800-kbp locus of the fruit fly genome. Ten of them coincided with the previously found S/MARs; however, only one with a *strong* S/MAR element [56, 130, 131].

Later a reconstruction of the domain structure of human 9p21-22 locus containing a gene cluster of type I interferons was attempted. It was found 36 S/MAR elements; 29 of them displayed high and seven a weak binding activities toward nuclear matrix. The interferon locus was found to consist of a series of small domains (from 2 to 10 kbp), in which the encoding gene sequences were flanked with S/MAR elements [59].

An analysis of maize chromosome 1 fragment helped reveal nine potential loop domains 6 to 75 kbp long [132].

Recently, an attempt was made to map the S/MARs according to the *in vitro* binding to nuclear matrix of a 150-kbp fragment of the human chromosome 14; this fragment contains a cluster of serpin genes. Five independent S/MAR elements were found. One of them was localized 16 kbp downstream the beginning of α-1-antitrypsin gene, three of them were found between the antitrypsin-like gene (ATR) and the gene of corticosteroid-binding globulin (CBG), and one in the *CBG* gene intron. The location of the first three S/MAR elements implies that the α-1-antitrypsin and ATR genes are situated in the same 50-kbp long chromatin domain separated from *CBG* by two S/MAR elements. Two other S/MARs are localized in the promoter region and in the first intron of *CBG* gene [133].

The determination of the primary structure of extended genomic sequences of eukaryotic organisms, in particular, the first version of complete nucleotide sequence of the human genome [134], opens possibilities for the exact positioning of a large number of S/MARs relative to adjacent genes. As mentioned above, the number of S/MARs in human genome is estimated as approximately 30 000; in plant genomes,

this number seems to be considerably higher. The existing methods of fragmentation and determination of the matrix-binding activity of single DNA fragments would probably be useless in most cases for the S/MARs mapping of extended genome regions to say nothing of the complete genomes. To solve these problems, functional clone libraries containing a large number of potential S/MARs should be constructed. The exact S/MARs localization requires that an average size of fragments in a library should correspond to the average S/MARs length, i.e., to about 500 bp. Up to now, two methods of constructing such libraries have been suggested. The first of them is based on cloning DNA fragments isolated together with nuclear matrix (nmDNA) obtained by various procedures [18, 21, 22]. The clone library constructed in this way from human cell cultures contained half of the clones corresponded to only one 542-bp S/MAR element [21]. In the LAS clone library [18], only 41% of unique sequences were affine to nuclear matrix. The tobacco S/MARs library [22] was obtained using the method described in [21]. The examination of 34 random clones showed that only 30% of them preferably bound *in vitro* the nuclear matrix and, hence, were S/MARs by definition. It is reasonable that the quality of the obtained clone libraries does not allow the direct mapping of sequences they contain and requires a preliminary analysis of the clones for their ability to be bound by nuclear matrix. The mapping of sequences from the libraries has not been performed.

Another approach was used in [52, 53]. It is based on constructing the clone library of short genomic fragments (500 bp on average) presenting the sequence of human chromosome 19, followed by sorting out the subpopulation of fragments *in vitro* binding nuclear matrix, i.e., containing S/MARs. In the obtained minilibrary, 55 clones were analyzed, and 50 of them were found to be preferably bound by nuclear matrix. More than 30 sequences were mapped on the human chromosome 19 according to their primary structure. This approach was further developed in [54], in which the S/MARs clone library of 1 000-kbp long completely sequenced locus of human genome was obtained, and 16 S/MAR elements were exactly mapped. The locus is situated between *D19S208* and *COX7A1* markers and contains 22 identified genes. Eleven intergene S/MARs divide the locus into ten domains of six to 272 kbp long with an average length of 88 kbp. Another five S/MARs were found in introns of the known genes (see below). Some correlation was observed for the tissue specificities of expression of genes involved in one and the same domain.

Intron S/MARs

S/MARs must flank one or several genes to form loop domains. Nevertheless, the sequences displaying S/MARs properties are often found in introns of various genes and, hence, are transcribed. For example, such S/MARs were found in the genes of murine immunoglobulin light and heavy chains [23, 29], rat α2-macroglobulin [135], human topoisomerase I [136], and some other genes [54].

It was found that S/MARs are closely associated with intron enhancers of genes of immunoglobulin heavy [29] and κ-light chains [23] and the genes of β-T-cellular receptor [118, 137] and δ-T-cellular receptor [138]. The intron enhancer of immunoglobulin heavy chains is flanked with S/MAR elements from both sides. These S/MARs are involved in the repression of the locus transcription in other than B type cells [139]. In B cells these elements function synergistically with the enhancer and activate transcription [140–143]. The S/MAR elements flanking the enhancer of the murine immunoglobulin heavy chain activate the terminal promoter of the variable region [V(H) promoter] and tenfold increase the histone acetylation in the adjacent regions [144]. Intron S/MARs are also necessary for demethylation of regulatory sites and activation of immunoglobulin κ-chains in B cells [145].

In addition, together with the enhancer, intron S/MARs can increase the frequency of the V(D)J recombination [138, 146].The level of somatic hypermutation of V region is also affected by intron S/MARs [147, 148]. The deletion of intron S/MARs results in the hyperrecombination of Vκ–Jκ regions, which correlates with their undermethylation but not with their transcription level. The authors think that S/MARs can inhibit the V–J recombination at the stage of B cell precursors [149, 150]. Another known example of the involvement of the S/MAR element located in the intron close to the enhancer into expression regulation is a gene of keratinocyte terminal differentiation marker (SPRR2A) [151].

S/MARs and Other Genomic Regulator Elements

S/MARs can be structurally close and functionally associated not only with intron enhancers, but also with other genomic regulatory elements. S/MAR elements belong to a new class of *cis* regulatory elements, Locus Control Regions (LCTs), which are the DNA fragments that enhance and maintain the expression of the gene under control after integration into another genome [152]. It was noticed that S/MARs are necessary for the tissue-specific activity of LCRs in murine tyrosinase [153] and human apolypoprotein E/C-I genes [154].

CONCLUSIONS

The data discussed above allow the conclusion that S/MAR elements play an important, although still incompletely understood role in the formation and maintaining of the structure and regulation of functioning of the cell genetic machinery. Numerous studies devoted to the mechanisms of S/MARs functioning at the level of single genes and small gene loci report insufficient data concerning the localization of

S/MARs in extended genome regions. Actually, there are virtually no data on the tissue specificity of the S/MARs binding to nuclear matrix that could help to elucidate the mechanisms of the large-scale chromatin structure–activity regulation. The structural basis for S/MARs binding to nuclear matrix has also been studied insufficiently. These drawbacks mostly result from the lack of adequate approaches to the analysis of this type. This makes necessary the development of new methods of investigation of the structure and activity regulation of the genetic material, in particular, on the basis of the results of the Human Genome Project.

ACKNOWLEDGMENTS

The work was supported by the program Leading Scientific Schools, project no. NSh-2006.2003.4, by the Russian Foundation for Basic Research, project nos. 01-04-48933, 01-04-48980, and 02-04-48601; and by the INTAS program, project no. 2001-0279.

REFERENCES

1. Marsden, M.P. and Laemmli, U.K., *Cell* (Cambridge, Mass.), 1979, vol. 17, pp. 849–858.
2. Pienta, K.J., Partin, A.W., and Coffey, D.S., *Cancer Res.*, 1989, vol. 49, pp. 2525–2532.
3. Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., and Woodcock, C.L., *Proc. Natl. Acad. Sci. USA*, 1998, vol. 95, pp. 14 173–14 178.
4. Igo-Kemenes, T., Horz, W., and Zachau, H.G., *Annu. Rev. Biochem.*, 1982, vol. 51, pp. 89–121.
5. Thomas, J.O., *J. Cell Sci. Suppl.*, 1984, vol. 1, pp. 1–20.
6. Finch, J.T. and Klug, A., *Proc. Natl. Acad. Sci. USA*, 1976, vol. 73, pp. 1897–1901.
7. Kiryanov, G.I., Smirnova, T.A., and Polyakov, V., *Eur. J. Biochem.*, 1982, vol. 124, pp. 331–338.
8. Subirana, J.A., Munoz-Guerra, S., Aymami, J., Radermacher, M., and Frank, J., *Chromosoma*, 1985, vol. 91, pp. 377–390.
9. Woodcock, C.L., Grigoryev, S.A., Horowitz, R.A., and Whitaker, N., *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, pp. 9021–9025.
10. Cook, P.R. and Brazell, I.A., *Nucleic Acids Res.*, 1980, vol. 8, pp. 2895–2906.
11. Hancock, R. and Boulikas, T., *Int. Rev. Cytol.*, 1982, vol. 79, pp. 165–214.
12. Pienta, K.J. and Coffey, D.S., *J. Cell Sci. Suppl.*, 1984, vol. 1, pp. 123–135.
13. Berezney, R. and Coffey, D.S., *Biochem. Biophys. Res. Commun.*, 1974, vol. 60, pp. 1410–1417.
14. Berezney, R., Mortillaro, M.J., Ma, H., Wei, X., and Samarabandu, J., *Int. Rev. Cytol.*, 1995, vol. 162A, pp. 1–65.
15. Berezney, R. and Coffey, D.S., *J. Cell Biol.*, 1977, vol. 73, pp. 616–637.
16. Mirkovitch, J., Mirault, M.E., and Laemmli, U.K., *Cell* (Cambridge, Mass.), 1984, vol. 39, pp. 223–232.
17. Jackson, D.A., Dickinson, P., and Cook, P.R., *Nucleic Acids Res.*, 1990, vol. 18, pp. 4385–4393.
18. Jackson, D.A., Bartlett, J., and Cook, P.R., *Nucleic Acids Res.*, 1996, vol. 24, pp. 1212–1219.
19. Mirkovitch, J., Gasser, S.M., and Laemmli, U.K., *J. Mol. Biol.*, 1988, vol. 200, pp. 101–109.
20. Barbashov, S.F., Glotov, B.O., and Nikolaev, L.G., *Biochim. Biophys. Acta*, 1984, vol. 782, pp. 177–186.
21. Boulikas, T. and Kong, C.F., *J. Cell Biochem.*, 1993, vol. 53, pp. 1–12.
22. Michalowski, S.M., Allen, G.C., Hall, G.E., Jr., Thompson, W.F., and Spiker, S., *Biochemistry*, 1999, vol. 38, pp. 12 795–12 804.
23. Cockerill, P.N. and Garrard, W.T., *Cell* (Cambridge, Mass.), 1986, vol. 44, pp. 273–282.
24. Izaurralde, E., Mirkovitch, J., and Laemmli, U.K., *J. Mol. Biol.*, 1988, vol. 200, pp. 111–125.
25. Bode, J., Schlake, T., Rios-Ramirez, M., Mielke, C., Stengert, M., Kay, V., and Klehr-Wirth, D., *Int. Rev. Cytol.*, 1995, vol. 162A, pp. 389–454.
26. Boulikas, T., *J. Cell Biochem.*, 1993, vol. 52, pp. 23–36.
27. Boulikas, T., *Int. Rev. Cytol.*, 1995, vol. 162A, pp. 279–388.
28. Bode, J., Stengert-Iber, M., Kay, V., Schlake, T., and Dietz-Pfeilstetter, A., *Crit. Rev. Eukaryot. Gene Expr.*, 1996, vol. 6, pp. 115–138.
29. Cockerill, P.N., Yuen, M.H., and Garrard, W.T., *J. Biol. Chem.*, 1987, vol. 262, pp. 5394–5397.
30. Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klehr, D., Mielke, C., and Kohwi-Shigematsu, T., *Science*, 1992, vol. 255, pp. 195–197.
31. Gasser, S.M. and Laemmli, U.K., *Cell* (Cambridge, Mass.), 1986, vol. 46, pp. 521–530.
32. Mielke, C., Kohwi, Y., Kohwi-Shigematsu, T., and Bode, J., *Biochemistry*, 1990, vol. 29, pp. 7475–7485.
33. Pommier, Y., Cockerill, P.N., Kohn, K.W., and Garrard, W.T., *J. Virol.*, 1990, vol. 64, pp. 419–423.
34. Benham, C., Kohwi-Shigematsu, T., and Bode, J., *J. Mol. Biol.*, 1997, vol. 274, pp. 181–196.
35. Bode, J., Bartsch, J., Boulikas, T., Iber, M., Mielke, C., Schubeler, D., Seibler, J., and Benham, C., *Gene Therapy Mol. Biol.*, 1998, vol. 1, pp. 551–580.
36. Bode, J., Benham, C., Knopp, A., and Mielke, C., *Crit. Rev. Eukaryot. Gene Expr.*, 2000, vol. 10, pp. 73–90.
37. Boulikas, T., *J. Cell Biochem.*, 1993, vol. 52, pp. 14–22.
38. Broeker, P.L., Harden, A., Rowley, J.D., and Zeleznik-Le, N., *Curr. Top. Microbiol. Immunol.*, 1996, vol. 211, pp. 259–268.
39. Opstelten, R.J., Clement, J.M., and Wanka, F., *Chromosoma*, 1989, vol. 98, pp. 422–427.
40. Homberger, H.P., *Chromosoma*, 1989, vol. 98, pp. 99–104.
41. von Kries, J.P., Phi-Van, L., Diekmann, S., and Stratling, W.H., *Nucleic Acids Res.*, 1990, vol. 18, pp. 3881–3885.
42. Adachi, Y., Kas, E., and Laemmli, U.K., *EMBO J.*, 1989, vol. 8, pp. 3997–4006.
43. Blasquez, V.C., Sperry, A.O., Cockerill, P.N., and Garrard, W.T., *Genome*, 1989, vol. 31, pp. 503–509.

44. Krogh, S., Mortensen, U.H., Westergaard, O., and Bonven, B.J., *Nucleic Acids Res.*, 1991, vol. 19, pp. 1235–1241.
45. van Drunen, C.M., Sewalt, R.G., Oosterling, R.W., Weisbeek, P.J., Smeekens, S.C., and van Driel, R., *Nucleic Acids Res.*, 1999, vol. 27, pp. 2924–2930.
46. Yamamura, J. and Nomura, K., *FEBS Lett.*, 2001, vol. 489, pp. 166–170.
47. Singh, G.B., Kramer, J.A., and Krawetz, S.A., *Nucleic Acids Res.*, 1997, vol. 25, pp. 1419–1425.
48. Glazko, G.V., Rogozin, I.B., and Glazkov, M.V., *Biochim. Biophys. Acta*, 2001, vol. 1517, pp. 351–364.
49. Frisch, M., Frech, K., Klingenhoff, A., Cartharius, K., Liebich, I., and Werner, T., *Genome Res.*, 2002, vol. 12, pp. 349–354.
50. Liebich, I., Bode, J., Frisch, M., and Wingender, E., *Nucleic Acids Res.*, 2002, vol. 30, pp. 372–374.
51. Mielke, C., Maass, K., Tummler, M., and Bode, J., *Biochemistry*, 1996, vol. 35, pp. 2239–2252.
52. Nikolaev, L.G., Tsevegiyn, T., Akopov, S.B., Ashworth, L.K., and Sverdlov, E.D., *Nucleic Acids Res.*, 1996, vol. 24, pp. 1330–1336.
53. Nikolaev, L.G., Akopov, S.B., Chernov, I.P., Glotov, B.O., Eshvort, L.K., and Sverdlov, E.D., *Dokl. Ross. Akad. Nauk*, 1998, vol. 361, pp. 409–411.
54. Chernov, I.P., Akopov, S.B., Nikolaev, L.G., and Sverdlov, E.D., *J. Cell Biochem.*, 2002, vol. 84, pp. 590–600.
55. Jack, R.S. and Eggert, H., *Eur. J. Biochem.*, 1992, vol. 209, pp. 503–509.
56. Iarovaia, O., Hancock, R., Lagarkova, M., Miassod, R., and Razin, S.V., *Mol. Cell. Biol.*, 1996, vol. 16, pp. 302–308.
57. Avramova, Z., Tikhonov, A., Chen, M., and Bennetzen, J.L., *Nucleic Acids Res.*, 1998, vol. 26, pp. 761–767.
58. Paul, A.L. and Ferl, R.J., *Plant Cell*, 1998, vol. 10, pp. 1349–1359.
59. Strissel, P.L., Dann, H.A., Pomykala, H.M., Diaz, M.O., Rowley, J.D., and Olopade, O.I., *Genomics*, 1998, vol. 47, pp. 217–229.
60. Kas, E. and Chasin, L.A., *J. Mol. Biol.*, 1987, vol. 198, pp. 677–692.
61. Maric, C. and Hyrien, O., *Chromosoma*, 1998, vol. 107, pp. 155–165.
62. Tsongalis, G.J., Coleman, W.B., Smith, G.J., and Kaufman, D.G., *Cancer Res.*, 1992, vol. 52, pp. 3807–3810.
63. Cai, S. and Kohwi-Shigematsu, T., *Methods*, 1999, vol. 19, pp. 394–402.
64. Roti Roti, J.L., Wright, W.D., and VanderWaal, R., *Crit. Rev. Eukaryot. Gene Expr.*, 1997, vol. 7, pp. 343–360.
65. Sakkers, R.J., Brunsting, J.F., Filon, A.R., Kampinga, H.H., Konings, A.W., and Mullenders, L.H., *Int. J. Radiat. Biol.*, 1999, vol. 75, pp. 875–883.
66. Levy-Wilson, B. and Fortier, C., *J. Biol. Chem.*, 1989, vol. 264, pp. 21 196–21 204.
67. Wang, D.M., Taylor, S., and Levy-Wilson, B., *J. Lipid Res.*, 1996, vol. 37, pp. 2117–2124.
68. Antes, T.J., Namciu, S.J., Fournier, R.E., and Levy-Wilson, B., *Biochemistry*, 2001, vol. 40, pp. 6731–6742.
69. Brotherton, T., Zenk, D., Kahanic, S., and Reneker, J., *Biochemistry*, 1991, vol. 30, pp. 5845–5850.
70. Phi-Van, L. and Stratling, W.H., *Progr. Mol. Subcell. Biol.*, 1990, vol. 11, pp. 1–11.
71. Jackson, D.A., Dolle, A., Robertson, G., and Cook, P.R., *Cell. Biol. Int. Rep.*, 1992, vol. 16, pp. 687–696.
72. Razin, S.V. and Vassetzky, Y.S., *Cell. Biol. Int. Rep.*, 1992, vol. 16, pp. 697–708.
73. Stief, A., Winter, D.M., Stratling, W.H., and Sippel, A.E., *Nature*, 1989, vol. 341, pp. 343–345.
74. Phi-Van, L., von Kries, J.P., Ostertag, W., and Stratling, W.H., *Mol. Cell. Biol.*, 1990, vol. 10, pp. 2302–2307.
75. Allen, G.C., Hall, G.E., Jr., Childs, L.C., Weissinger, A.K., Spiker, S., and Thompson, W.F., *Plant Cell*, 1993, vol. 5, pp. 603–613.
76. Charron, G., Julien, J.P., and Bibor-Hardy, V., *J. Biol. Chem.*, 1995, vol. 270, pp. 25 739–25 745.
77. Klehr, D., Maass, K., and Bode, J., *Biochemistry*, 1991, vol. 30, pp. 1264–1270.
78. Allen, G.C., Hall, G., Jr., Michalowski, S., Newman, W., Spiker, S., Weissinger, A.K., and Thompson, W.F., *Plant Cell*, 1996, vol. 8, pp. 899–913.
79. Poljak, L., Seum, C., Mattioni, T., and Laemmli, U.K., *Nucleic Acids Res.*, 1994, vol. 22, pp. 4386–4394.
80. Allen, G.C., Spiker, S., and Thompson, W.F., *Plant Mol. Biol.*, 2000, vol. 43, pp. 361–376.
81. Geyer, P.K., *Curr. Opin. Genet. Dev.*, 1997, vol. 7, pp. 242–248.
82. Bell, A.C. and Felsenfeld, G., *Curr. Opin. Genet. Dev.*, 1999, vol. 9, pp. 191–198.
83. Kalos, M. and Fournier, R.E., *Mol. Cell. Biol.*, 1995, vol. 15, pp. 198–207.
84. Nabirochkin, S., Ossokina, M., and Heidmann, T., *J. Biol. Chem.*, 1998, vol. 273, pp. 2473–2479.
85. Namciu, S.J., Blochlinger, K.B., and Fournier, R.E., *Mol. Cell. Biol.*, 1998, vol. 18, pp. 2382–2391.
86. Scott, K.C., Taubman, A.D., and Geyer, P.K., *Genetics*, 1999, vol. 153, pp. 787–798.
87. Attal, J., Cajero-Juarez, M., Petitclerc, D., Theron, M.C., Stinnakre, M.G., Bearzotti, M., Kann, G., and Houdebine, L.M., *Mol. Biol. Rep.*, 1995, vol. 22, pp. 37–46.
88. Dang, Q., Auten, J., and Plavec, I., *J. Virol.*, 2000, vol. 74, pp. 2671–2678.
89. Bode, J., Schlake, T., Iber, M., Schubeler, D., Seibler, J., Snezhkov, E., and Nikolaev, L., *Biol. Chem.*, 2000, vol. 381, pp. 801–813.
90. Chimera, J.A. and Musich, P.R., *J. Biol. Chem.*, 1985, vol. 260, pp. 9373–9379.
91. Akopov, S.B., Nikolaev, L.G., Tyrsin, O., Ruzov, A.S., and Sverdlov, E.D., *Bioorg. Khim.*, 1997, vol. 23, pp. 727–731.
92. de Belle, I., Cai, S., and Kohwi-Shigematsu, T., *J. Cell Biol.*, 1998, vol. 141, pp. 335–348.
93. Tikhonov, A.P., Bennetzen, J.L., and Avramova, Z.V., *Plant Cell*, 2000, vol. 12, pp. 249–264.
94. Dickinson, L.A., Joh, T., Kohwi, Y., and Kohwi-Shigematsu, T., *Cell* (Cambridge, Mass.), 1992, vol. 70, pp. 631–645.

95. Nakagomi, K., Kohwi, Y., Dickinson, L.A., and Kohwi-Shigematsu, T., *Mol. Cell. Biol.*, 1994, vol. 14, pp. 1852–1860.
96. Dickinson, L.A., Dickinson, C.D., and Kohwi-Shigematsu, T., *J. Biol. Chem.*, 1997, vol. 272, pp. 11463–11470.
97. Kohwi-Shigematsu, T., Maass, K., and Bode, J., *Biochemistry*, 1997, vol. 36, pp. 12 005–12 010.
98. Alvarez, J.D., Yasui, D.H., Niida, H., Joh, T., Loh, D.Y., and Kohwi-Shigematsu, T., *Genes Dev.*, 2000, vol. 14, pp. 521–535.
99. Yanagisawa, J., Ando, J., Nakayama, J., Kohwi, Y., and Kohwi-Shigematsu, T., *Cancer Res.*, 1996, vol. 56, pp. 457–462.
100. von Kries, J.P., Buhrmester, H., and Stratling, W.H., *Cell* (Cambridge, Mass.), 1991, vol. 64, pp. 123–135.
101. von Kries, J.P., Rosorius, O., Buhrmester, H., and Stratling, W.H., *FEBS Lett.*, 1994, vol. 342, pp. 185–188.
102. Buhrmester, H., von Kries, J.P., and Stratling, W.H., *Biochemistry*, 1995, vol. 34, pp. 4108–4117.
103. Weitzel, J.M., Buhrmester, H., and Stratling, W.H., *Mol. Cell. Biol.*, 1997, vol. 17, pp. 5656–5666.
104. Stratling, W.H. and Yu, F., *Crit. Rev. Eukaryot. Gene Expr.*, 1999, vol. 9, pp. 311–318.
105. Tsutsui, K., Tsutsui, K., Okada, S., Watarai, S., Seki, S., Yasuda, T., and Shohmori, T., *J. Biol. Chem.*, 1993, vol. 268, pp. 12 886–12 894.
106. Romig, H., Fackelmayer, F.O., Renz, A., Ramsperger, U., and Richter, A., *EMBO J.*, 1992, vol. 11, pp. 3431–3440.
107. Fackelmayer, F.O., Dahm, K., Renz, A., Ramsperger, U., and Richter, A., *Eur. J. Biochem.*, 1994, vol. 221, pp. 749–757.
108. Fackelmayer, F.O. and Richter, A., *Biochim. Biophys. Acta*, 1994, vol. 1217, pp. 232–234.
109. Kipp, M., Gohring, F., Ostendorp, T., van Drunen, C.M., van Driel, R., Przybylski, M., and Fackelmayer, F.O., *Mol. Cell. Biol.*, 2000, vol. 20, pp. 7480–7489.
110. Renz, A. and Fackelmayer, F.O., *Nucleic Acids Res.*, 1996, vol. 24, pp. 843–849.
111. Kipp, M., Schwab, B.L., Przybylski, M., Nicotera, P., and Fackelmayer, F.O., *J. Biol. Chem.*, 2000, vol. 275, pp. 5031–5036.
112. Oesterreich, S., Lee, A.V., Sullivan, T.M., Samuel, S.K., Davie, J.R., and Fuqua, S.A., *J. Cell Biochem.*, 1997, vol. 67, pp. 275–286.
113. Herrscher, R.F., Kaplan, M.H., Lelsz, D.L., Das, C., Scheuermann, R., and Tucker, P.W., *Genes Dev.*, 1995, vol. 9, pp. 3067–3082.
114. Zong, R.T., Das, C., and Tucker, P.W., *EMBO J.*, vol. 19, pp. 4123–4133.
115. Banan, M., Rojas, I.C., Lee, W.H., King, H.L., Harriss, J.V., Kobayashi, R., Webb, C.F., and Gottlieb, P.D., *J. Biol. Chem.*, 1997, vol. 272, pp. 18 440–18 452.
116. Wang, Z., Goldstein, A., Zong, R.T., Lin, D., Neufeld, E.J., Scheuermann, R.H., and Tucker, P.W., *Mol. Cell. Biol.*, 1999, vol. 19, pp. 284–295.
117. Antes, T.J., Chen, J., Cooper, A.D., and Levy-Wilson, B., *J. Biol. Chem.*, 2000, vol. 275, pp. 26649–26660.
118. Chattopadhyay, S., Kaul, R., Charest, A., Housman, D., and Chen, J., *Genomics*, 2000, vol. 68, pp. 93–96.
119. Meier, I., Phelan, T., Gruissem, W., Spiker, S., and Schneider, D., *Plant Cell*, 1996, vol. 8, pp. 2105–2115.
120. Gindullis, F. and Meier, I., *Plant Cell*, 1999, vol. 11, pp. 1117–1128.
121. Morisawa, G., Han-Yama, A., Moda, I., Tamai, A., Iwabuchi, M., and Meshi, T., *Plant Cell*, 2000, vol. 12, pp. 1903–1916.
122. Bode, J. and Maass, K., *Biochemistry*, 1988, vol. 27, pp. 4706–4711.
123. Dijkwel, P.A. and Hamlin, J.L., *Mol. Cell. Biol.*, 1988, vol. 8, pp. 5398–5409.
124. Cockerill, P.N., *Nucleic Acids Res.*, 1990, vol. 18, pp. 2643–2648.
125. Jarman, A.P. and Higgs, D.R., *EMBO J.*, 1988, vol. 7, pp. 3337–3344.
126. Mirkovitch, J., Spierer, P., and Laemmli, U.K., *J. Mol. Biol.*, 1986, vol. 190, pp. 255–258.
127. Surdej, P., Got, C., Rosset, R., and Miassod, R., *Nucleic Acids Res.*, 1990, vol. 18, pp. 3713–3722.
128. Surdej, P., Brandli, D., and Miassod, R., *Biol. Cell*, 1991, vol. 73, pp. 111–120.
129. Razin, S.V., Petrov, P., and Hancock, R., *Proc. Natl. Acad. Sci. USA*, 1991, vol. 88, pp. 8515–8519.
130. Gromova, I.I., Thomsen, B., and Razin, S.V., *Proc. Natl. Acad. Sci. USA*, 1995, vol. 92, pp. 102–106.
131. Miassod, R., Razin, S.V., and Hancock, R., *Nucleic Acids Res.*, 1997, vol. 25, pp. 2041–2046.
132. Avramova, Z., SanMiguel, P., Georgieva, E., and Bennetzen, J.L., *Plant Cell*, 1995, vol. 7, pp. 1667–1680.
133. Rollini, P., Namciu, S.J., Marsden, M.D., and Fournier, R.E., *Nucleic Acids Res.*, 1999, vol. 27, pp. 3779–3791.
134. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.*, *Nature* (London), 2001, vol. 409, pp. 860–921.
135. Ito, T. and Sakaki, Y., *Biochem. Biophys. Res. Commun.*, 1987, vol. 149, pp. 449–454.
136. Romig, H., Ruff, J., Fackelmayer, F.O., Patil, M.S., and Richter, A., *Eur. J. Biochem.*, 1994, vol. 221, pp. 411–419.
137. Chattopadhyay, S., Whitehurst, C.E., and Chen, J., *J. Biol. Chem.*, 1998, vol. 273, pp. 29 838–29 846.
138. Zhong, X.P., Carabana, J., and Krangel, M.S., *Proc. Natl. Acad. Sci. USA*, 1999, vol. 96, pp. 11 970–11 975.
139. Scheuermann, R.H. and Chen, U., *Genes Dev.*, 1989, vol. 3, pp. 1255–1266.
140. Forrester, W.C., van Genderen, C., Jenuwein, T., and Grosschedl, R., *Science*, 1994, vol. 265, pp. 1221–1225.
141. Jenuwein, T., Forrester, W.C., Fernandez-Herrero, L.A., Laible, G., Dull, M., and Grosschedl, R., *Nature* (London), 1997, vol. 385, pp. 269–272.
142. Forrester, W.C., Fernandez, L.A., and Grosschedl, R., *Genes Dev.*, 1999, vol. 13, pp. 3003–3014.
143. Oancea, A.E., Berru, M., and Shulman, M.J., *Mol. Cell. Biol.*, 1997, vol. 17, pp. 2658–2668.
144. Fernandez, L.A., Winkler, M., and Grosschedl, R., *Mol. Cell. Biol.*, 2001, vol. 21, pp. 196–208.

145. Lichtenstein, M., Keini, G., Cedar, H., and Bergman, Y., *Cell* (Cambridge, Mass.), 1994, vol. 76, pp. 913–923.

146. Roch, F.A., Hobi, R., Berchtold, M.W., and Kuenzle, C.C., *Nucleic Acids Res.*, 1997, vol. 25, pp. 2303–2310.

147. Betz, A.G., Milstein, C., Gonzalez-Fernandez, A., Pannell, R., Larson, T., and Neuberger, M.S., *Cell* (Cambridge, Mass.), 1994, vol. 77, pp. 239–248.

148. Goyenechea, B., Klix, N., Yelamos, J., Williams, G.T., Riddell, A., Neuberger, M.S., and Milstein, C., *EMBO J.*, 1997, vol. 16, pp. 3987–3994.

149. Hale, M.A. and Garrard, W.T., *Mol. Immunol.*, 1998, vol. 35, pp. 609–620.

150. Yi, M., Wu, P., Trevorrow, K.W., Claflin, L., and Garrard, W.T., *J. Immunol.*, 1999, vol. 162, pp. 6029–6039.

151. Fischer, D.F., van Drunen, C.M., Winkler, G.S., van de Putte, P., and Backendorf, C., *Nucleic Acids Res.*, 1998, vol. 26, pp. 5288–5294.

152. Li, Q. and Peterson, K.R., *Trends Genet.*, 1999, vol. 15, pp. 403–408.

153. Porter, S.D., Hu, J., and Gilks, C.B., *Dev. Genet.*, 1999, vol. 25, pp. 40–48.

154. Dang, Q., Walker, D., Taylor, S., Allan, C., Chin, P., Fan, J., and Taylor, J., *J. Biol. Chem.*, 1995, vol. 270, pp. 22577–22585.