ISSN 1990-7508, Biochemistry (Moscow), Supplement Series B: Biomedical Chemistry, 2017, Vol. 11, No. 1, pp. 62–71. © Pleiades Publishing, Ltd., 2017.
Original Russian Text © A.V. Sekridova, A.M. Varizhuk, O.N. Tatarinova, V.V. Severov, N.A. Barinov, I.P. Smirnov, V.N. Lazarev, D.V. Klinov, G.E. Pozmogova, 2017, published in Biomeditsinskaya Khimiya.

Conformational Polymorphysm of G-Rich Fragments of DNA Alu-Repeats. I. Noncanonical Structures

A. V. Sekridova, A. M. Varizhuk, O. N. Tatarinova, V. V. Severov, N. A. Barinov, I. P. Smirnov, V. N. Lazarev, D. V. Klinov, and G. E. Pozmogova*

Federal Research and Clinical Center of Physical-Chemical Medicine, ul. Malaya Pirogovskaya 1a, Moscow, 119435 Russia
**e-mail: pozmge@niifhm.ru*

Received May 22, 2016

Abstract—We report results of the first systematic study of conformational polymorphism of G-rich DNA fragments of Alu-repeats. Alu retrotransposons are primate-specific short interspersed elements. Using the Alu sequence of the prooncogen *bcl2* intron and the consensus $AluS_x$ sequence as representative examples, we have determined characteristic Alu sites that are capable of adopting G-quadruplex (GQ) conformations (i.e., potential quadruplex sites—PQS_{Alu}), and demonstrated by bioinformatics methods that these sites are Alu-specific in the human genome. Genomic frequencies of PQS_{Alu} were assessed (~1/10000 bp). These sites were found to be characteristic of young (active) Alu families (Alu-Y). A recombinant DNA sequence bearing the Alu element of the human *bcl2* gene (304 bp) and its PQS-mutant (Alu-PQS) were constructed. The formation of noncanonical structures in Alu_{bcl2} dsDNA and their absence in the case of Alu-PQS have been shown using DMS-footprinting and atomic force microscopy (AFM). Expression vectors bearing wild-type and mutant Alu insertions in the promoter regions of the reporter gene have been prepared, and their regulatory effects have been compared during transfection of HEK293 and HeLa cells. We suggest that the dynamic study of the spatial organization of Alu repeats may provide insight into the mechanisms of genomic rearrangements responsible for the development of many oncological and neurodegenerative diseases.

Keywords: G-quadruplex DNA, Alu-repeats, dynamics of DNA secondary structures
DOI: 10.1134/S1990750817010097

INTRODUCTION

Alu-repeats (Alu, ~300 base pairs, bp), comprising more than 10% of the human genome, are currently considered as important regulatory elements. Short interspersed repeated sequences (SINE), which include Alu, are involved, for example, in mechanisms of the regulation of gene expression by antisense and microRNAs [1, 2]. It is believed that widespread distribution of retrotransposons (primarily, Alu) in primate genomes could result in appearance of a new regulatory level.

Analysis of nucleotide polymorphism in Alu-family sequences as well as genomic rearrangements involving deletions and insertions of Alu elements is especially important for molecular medicine, as it is known that they can lead to the development of various pathologies [3–6].

At the same time, the study of the dynamics of the secondary structures of polynucleotides (i.e., their conformational polymorphism) represents a key approach providing insights into mechanisms of some diseases. For example, the importance of the spatial organization of Alu RNA may be illustrated by A-I-editing, based on the interaction of RNA duplexes (formed in most cases by inverted Alu or their fragments) with adenosine deaminase ADAR. Decreased levels of A-I-editing were observed in amyotrophic lateral sclerosis, epilepsy, cancer, mainly glioblastomas [7–11]. The dynamics of the Alu DNA secondary structures has not been studied in detail yet.

We suggest that the G-rich fragments of DNA Alu possess conformational potential, and their features could be of interest for understanding the genomic regulation, translocation, and recombination. Attention to these sites is primarily associated with participation of PQS (Potential G-Quadruplex Site) in chromosomal rearrangements [12–15]. Recently, it has been also shown that Alu repeats are enriched with motifs, susceptible to double-strand breaks (DB) of DNA [4]. DB may lead to the development of genomic abnormalities and variations, CNV (Copy Number Variations), a known cause of neurodegenerative, oncological and other diseases [16]. It is suggested [4] that secondary non-canonical DNA structures, which are possibly forms in the G-rich tandems

typical for DB-motives, play the key role in mechanisms of CNV formation.

The aim of this study was to investigate conformational polymorphism of G-rich fragments of DNA Alu in the context of the possible role of non-canonical DNA structures in the mechanisms of biological processes.

We have studied a natural Alu-fragment of the prooncogen *bcl2* intron (18 chr.h), homologous to the $AluS_x$ consensus sequence [6, 17]. Using methods of bioinformatic analysis we have determined Alu-PQS, characterized their frequency in the human genome and belonging to different Alu-families. In order to compare regulatory effects of natural regulatory Alu and its PQS-mutant we prepared recombinant structures and corresponding expression vectors based on the pEGFP-N1 plasmid. Changes in fluorescence were registered after transfection of HeLa and HEK293 cells with these vectors. Implementation of non-canonical structures within Alu-amplicons was confirmed by atomic force microscopy (AFM) and DMS-footprinting. Detailed physical and chemical characteristics of the PQS-oligomers described in the second part of the article, revealed the structure of PQS-oligomers and demonstrated formation of stable intramolecular and intermolecular G-quadruplexes (GQ).

MATERIALS AND METHODS

Oligonucleotides were by the solid phase amidophosphite method using commercial reagents (Glen Research, USA) on an ASM-800 DNA synthesizer (Bioset, Russia) as described in [18]. Oligonucleotides purified by the chromatographic method [19] (>95%) were characterized by MALDI MS (Bruker Microflex, USA) as described in [20].

Recombinant Alu_{bcl2} and Alu-PQS were prepared by the chemical-enzymatic method using synthetic oligonucleotides (Alu1–10, Alu2m, Alu4m, Alu7m, and Alu9m, see table).

Expression plasmid vectors were prepared using a commercial plasmid pEGFP-N1 (Clontech, USA), PCR products were cloned into the pGEM-TEasy vector (Promega, USA).

Recombinant Alu_{bcl2} and Alu-PQS were amplified using PrF and PrR primers, digested with AseI restriction endonuclease and ligated with the linear pGEM-TEasy plasmid. The mixture was then heated for 10 min at 65°C and used for transformation of *Escherichia coli* DH5A cells. Plasmids were isolated from the cell suspension using a GeneJET Plasmid Midiprep Kit (Fermentas, USA) according to manufacturer's instructions. The correctness of the insertion sequence was confirmed by sequencing. Expression vectors based on pGEM-TEasy were digested with AseI restriction endonuclease. The fragment was purified and ligated to a linear pEGFP-N1 vector (AseI site). A unique cloning site was located in front of the promoter region of the reporter gene encoding EGFP. Thus, the following plasmid vectors were obtained: pAluf, pAlu-PQSf, pAlur, pAlu-PQSr (f—direct insertion, r—reverse insertion). HeLa (ATCC CCL-2) and HEK 293 (ATCC CRL-1573) cells were cultivated in plastic culture flasks (Corning-Costar, Netherlands) in DMEM medium (Sigma, USA) containing 10% fetal bovine serum, at 37°C in an atmosphere of 5% CO_2. Cells were passaged twice a week using trypsin-EDTA solution (0.05% trypsin, 0.02% EDTA).

Transfection of cells was performed using a TransPass COS/293 reagent (NEB, USA), according to manufacturer's instructions. Plasmid DNA for transfection was isolated using an EndoFree Plasmid Maxi Kit (Qiagen, Germany).

For each plasmid three replicates and two dilutions were used in experiments. The level of reporter protein expression was assessed by fluorescence intensity of the green fluorescent protein (GFP, 488 nm) using a Coulter Epix XL flow cytometer (Beckman Coulter, USA).

Atomic Force Microscopy (AFM)

The dsAlu and dsAlu-PQS amplicons were prepared using the pAlu and pAlu-PQS plasmids and primers PrF and PrR (table) under standard PCR conditions. After electrophoretic separation in an agarose gel, DNA fragments were further purified on the Performa® DTR Gel Filtration Performa cartridge. Prepared samples were diluted 100–2000 times before analysis in the buffer containing 10 mM KCl, 0.5 mM Tris-HCl (pH 7.5). Freshly cleaved highly oriented pyrolytic graphite was used as a substrate. After addition of 20 μL 0.1% graphite modifier (GM) solution (Nanotuning, Russia), followed by exposure in a humid chamber for 10 min GM was removed in a stream of compressed nitrogen gas, and substrate was dried. A DNA sample (1 μL) was applied to the substrate surface for 2–5 s, then diluted with 100 μL of water; after 10 s liquid was blown under the stream of nitrogen.

The samples surface relief was investigated in a semi-contact resonant mode by means of the research complex Integra Prima (NT-MDT, Russia) using NOVA 1.1 management software. Hyperacute high resolution silicon cantilevers with the resonance frequency of 190–325 kHz, tip curvature radius <2 nm and at its top angle <22 (Nanotuning) were used. The

Sequences of recombinant Alu and synthetic oligonucleotides

Name	5'–3'-Sequence
Alu_{bcl2}	GGCCCCGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGCGG ACCACGAGTTCAGGAGATTGAGACCATCTTGGTTAACACGGTGAAACCCTGTCTCTACTAAA AATACAAAAAATTTAGCTGGGGATGGTGGCGGGCGCCTGTAGTCCCAGCTGCTTGGGAGG CTGAGGCAGGAGAATGGCATGAACCCTGGAGGCGGAGGTTGCAGTGAGCCAAGATCACGC CACTGCACTCCAGCCTGGGCGACACAGCGAAACTCTGTCTC; 63296930–63297212 Homo sapiens chromosome 18, GRCh38 reference primary assembly. http://www.ncbi.nlm.nih.gov/
Alu-PQS	GGCCCCGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGtCtGGCGGA CCACGAGTTCAGGAGATTGAGACCATCTTGGTTAACACGGTGAAACCCTGTCTCTACTAAAA ATACAAAAAATTTAGCTGGGGATGGTGGCGGGCGCCTGTAGTCCCAGCTGCTTGGGAGGC TGAttCAGGAGAATGGCATGAACCCTttAGGCGGAGGTTGCAGTGAGCCAAGATCACGCCACT GCACTCCAGCCTGGGCGACACAGCGAAACTCTGTCTC*
Alu1	GGCCCCGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGG
Alu2	GAGGCCGAGGCGGGCGGACCACGAGTTCAGGAGATTGAGACCATCTTGGTTAACACG
Alu3	GTGAAACCCTGTCTCTACTAAAAATACAAAAAATTTAGCTGGGGATGGTGGCGGGCG-CCTGT
Alu4	AGTCCCAGCTGCTTGGGAGGCTGAGGCAGGAGAATGGCATGAACCCTGGAGGCGGA
Alu5	GGTTGCAGTGAGCCAAGATCACGCCACTGCACTCCAGCCTGGGCGACACAGCGAAACTCT GTCTC
Alu6	CTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTGCGGGGCC
Alu7	AGGGTTTCACCGTGTTAACCAAGATGGTCTCAATCTCCTGAACTCGTGGTCCGCCCGC
Alu8	AGCTGGGACTACAGGCGCCCGCCACCATCCCCAGCTAAATTTTTTGTATTTTTAGTAGAGAC
Alu9	CACTGCAACCTCCGCCTCCAGGGTTCATGCCATTCTCCTGCCTCAGCCTCCCAAGC
Alu10	GAGACAGAGTTTCGCTGTGTCGCCCAGGCTGGAGTGCAGTGGCGTGATCTTGGCT
Alu2m	GAGGCCGAGTCTGGCGGACCACGAGTTCAGGAGATTGAGACCATCTTGGTTAACACG
Alu4m	AGTCCCAGCTGCTTGGGAGGCTGATTCAGGAGAATGGCATGAACCCTTTAGGCGGA
Alu7m	AGGGTTTCACCGTGTTAACCAAGATGGTCTCAATCTCCTGAACTCGTGGTCCGCCAGA
Alu9m	CACTGCAACCTCCGCCTAAAGGGTTCATGCCATTCTCCTGAATCAGCCTCCCAAGC
PrF	CATATATTAATGGCCCCGCACGGTGG**
PrR	CTCGCATTAATGAGACAGAGTTTCGCTGTG**
DMS-PQS2	FAM-CCAGCACTTTGGGAGGCCGAGGCGGGCGGACCACGA
DMS-PQS3	FAM-CCAGCTGCTTGGGAGGCTGAGGCAGGAGAATGGCAT
DMS-PQS4	FAM-GCATGAACCCTGGAGGCGGAGGTTGCAGTGAG
PQS2_compl	TCGTGGTCCGCCCGCCTCGGCCTCCCAAAGTGCTGG
PQS3_compl	ATGCCATTCTCCTGCCTCAGCCTCCCAAGCAGCTGG
PQS4_compl	CTCACTGCAACCTCCGCCTCCAGGGTTCATGC
ss-control	FAM AAGTGTGTGCGTGAGTGCGTGAGGCCAAGT

* Lowercase letters designate mutations preventing GQ formation.

** AseI restriction sites are underlined.

amplitude of free cantilever oscillations in air was 1–10 nm, the automatically maintained amplitude (Set-Point parameter) was set at the level of 60–70% of free oscillations. Signal processing, its conversion into the digital form, and imaging were performed using NOVA 1.1 software (NT-MDT). Image processing and measurement of heights of visualized objects were carried out by the image processing module NOVA Image Analysis 2.0.

The semi-automatic measurement of the oligonucleotide length was performed using the DNA-calc program [21]. The beginning and end of each oligonucleotide were selected manually; the path length was measured automatically. Histograms of height distribution were built using MS Excel.

DMS-Footprinting

DMS-footprinting performed by the method [22] was carried out using a set of synthesized FAM-oligonucleotides bearing PQS (DMS-PQS2, DMS-PQS3, DMS-PQS4), complementary oligomers (PQS2-compl, PQS3-compl and PQS4-compl), and also a control FAM-oligomer (ss-control), unable to form a secondary structure (see table). We compared fluorescence intensity of bands formed during reaction in 18% polyacrylamide gel electrophoresis (PAGE) under denaturing conditions in the presence of potassium ions (stabilized GQ) or lithium (prevented formation of GQ-conformations).

RESULTS AND DISCUSSION

Selection of Alu Sequences and PQS Analysis

Previously, the consensus sequence $AluS_x$ (Alu_{cons}) has been proposed on the basis of detailed analysis of Alu diversity [6, 17]; this sequence has become a common and convenient object for structural studies. However, the study of the functional properties of Alu-sites clearly needs use of natural genome fragments rather than their averaged version. Therefore, we have chosen the Alu repeat sequence of the prooncogen *bcl2* intron bearing the basic Alu elements and highly homologous Alu_{cons} in the G-rich sites. The sequence flanking A-box has been excluded from consideration as it does not form non-canonical structures and is involved in [6, 17]. Sequence analysis performed by using the program for the search of potential sites GQ-ImGQ-finder (http://imgqfinder.niifhm.ru/) revealed four conservative groups of partially overlapping two-quartet PQS-sites specific to Alu_{cons} and Alu_{bcl2} (Fig. 1).

The Abundance of PQS in the Genome and Their Distribution in the Alu Families

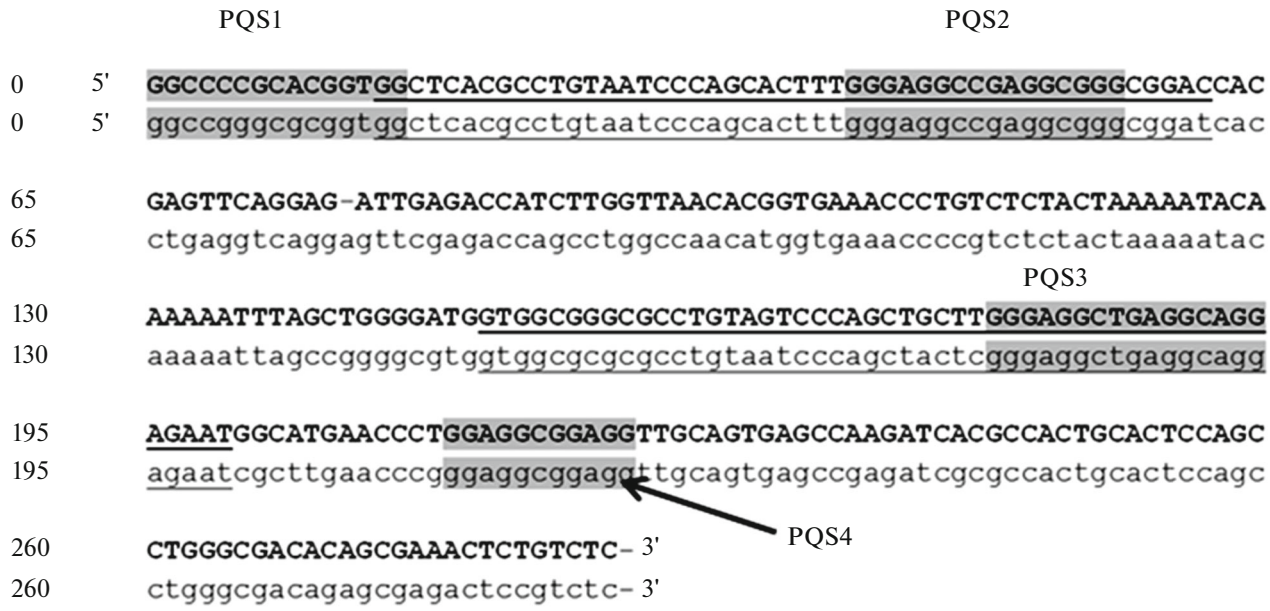
In the context of the functional role of selected sites it is known that the PQS1 fragment is an internal promoter of RNA polymerase III (so-called B-box) [6, 17]. No functional information has been published about features of PQS4.

For investigation of mechanisms of recombination PQS2 and PQS3, representing fragments of the DB-motif, are especially interesting. They are broadly spread throughout the human genome and are characterized by detection frequency of at least one per 10000 bp. PQS2, PQS3 and PQS4 share a common mode of distribution in chromosomes (Fig. 2a). The average frequency of PQS2.1 is 2.2/100000 bp, PQS3.1—9.2/100000 bp, and PQS4.1—3.9/100000 bp. The maximal values are typical for chromosomes 17 and 19, where they reach in 2.2/10000 bp (PQS3.1; chromosome 19), and the minimal value 1/100000 bp (PQS2.1) has been found in the case of the Y-chromosome.

Association of selected PQS with Alu repeats has also been shown during analysis of ~24000 Alu-sequences annotated in NCBI as Alu of human chromosome 18, in which the occurrence of Alu PQS has been close to the average value (Fig. 2a). All sites were associated almost exclusively with Alu (100% of PQS2.1 sites, 99.2% of PQS3.1 sites, and 98.2% of PQS4.1 sites).

Interestingly, the presence of PQS (e.g., PQS3.1, the most abundant in the genome) is not typical for all Alu families. In the considered set (chromosome 18) members of various families are present (Fig. 2b). Figure 2c shows a histogram of the occurrence frequency of PQS3.1 and its presence in each Alu family. This PQS3.1 is most frequently (in 85–100%) found in repeats belonging to AluYb9, AluYa5, AluYb8, and AluYg6 families. In the group of repeats (AluY, AluYd8, AluSp, AluSc, AluSq, AluSg, AluSx, AluSg/x) the occurrence frequency of this sequence varies from 15% to 49%. The third group comprises the following falimies: AluSp/q, AluJo, FAM, AluJb, AluSg1, FRAM, 7SLRNA, FLAM_A, AluJo/FLAM, FLAM_C, AluS, AluSq/x, AluJ/FLAM. In these families the occurrence frequency of PQS3.1 does not exceed 10%. Repeats referred to the Alu families AluJ, AluJ/FRAM, AluJ/monomer, AluJo/FRAM, AluYa, AluYa5/8, AluYb, AluYc, AluYd, AluYg, AluYh9, BC200, FLAM and FRAM/FAM, do not contain PQS3.1.

The considered example shows that the PQS3.1 sequence is typical for members of old (e.g. AluSx) and young (e.g. AluY) families. However, the PQS content (29%) of the main representative of the old



PQS Alu sequences

PQS1.1	ggccgggcgcggtgg	PQS2.1 PQS2.2 PQS2.3	GGGAGGCCGAGGCGG GGAGGCCGAGGCGG GGGAGGCCGAGGCGGG
PQS3.1 PQS3.2	GGGAGGCTGAGGCAGG GGAGGCTGAGGCAGG	PQS4.1 PQS4.2	GGAGGCGGAGG gggaggcggagg

Fig. 1. Sequences of the Alu-fragment of the *bcl2* gene intron, Alu_{bcl2} (uppercase), and consensus Alu_{cons} repeat (lowercase). PQS-sites are shown in gray background and their non-overlapping of their sequences are given in the table. DB-motives are underlined.

Alu families, AluSx, almost coincides with AluSx proportion in Alu-families of chromosome 18 (28%). In the young AluY family, representing 12% of all Alu-repeats in the set, the percentage of PQS reaches 27%.

Thus, PQS3.1 presence is most typical for certain young (active) families, for example, for AluYb9, AluYa5, AluYb8, AluYg6; this indirectly suggests the biological importance of this site.

Effect of Recombinant Alu_{bcl2} and Alu-PQS Insertions on Reporter Gene Expression

Biological functions of Alu repeats are actively investigated. It has been shown that these repeats are involved in stress responses, in regulation of translocation, recombination, etc. [23–26]. The Alu element located in front of the promoter has been shown to enhance expression of genes regardless of orientation and copy number [27]. In this study we have prepared recombinant Alu-DNA fragments: Alu_{bcl2} and mutant PQS (Alu-PQS). Chemical and enzymatic synthesis of the fragments (282 bp) was performed using the oligomers listed in the table.

Using commercially available pEGFP-N1 plasmid and genetic engineering methods we have prepared expression vectors bearing forward (f) and reversed (r) Alu insertions in the promoter region of the *GFP* gene (Fig. 3a). The plasmids pAlu (f/r), pAlu-PQS (f/r), and pEGFP-N1 (used as control) were delivered into HeLa and HEK293 cells by lipofection.

It was found that in all cases the Alu-insertion caused a statistically significant increase in fluorescence (30–40% in the case of HeLa, and 30–70% in HEK293 cells). This means that the investigated Alu_{bcl2} sequence bearing PQS typical for young families Alu can serve as an adequate model of the active Alu-repeat (at least in terms of regulatory properties).

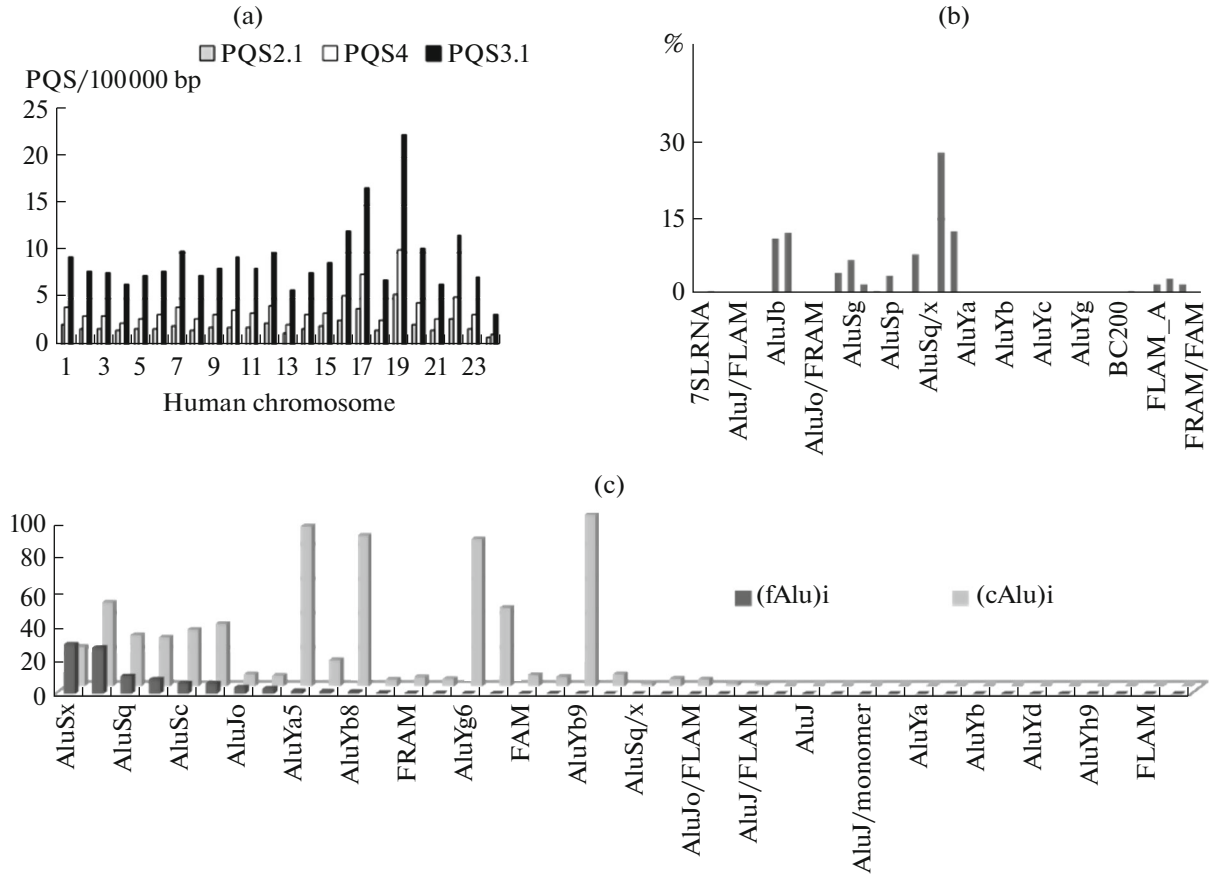


Fig. 2. Abundance of Alu-PQS in the genome and their distribution in Alu families. (a) Distribution of PQS Alu in the chromosomes of the human genome. (b) Distribution of Alu-families in chromosome 18. (c) Occurrence frequency of (fAlu)i of the PQS oligonucleotide PQS3.1 in the set of Alu-families of human chromosome 18 and abundance of (cAlu)i PQS3.1 in each family. Analysis was performed using NCBI data (GRCh38 reference primary assembly.http: //www.ncbi.nlm.nih.gov/); fAlu for Alu i-family was calculated as the ratio of PQS3.1 number in the i-family to the total number of PQS3.1 in the set; cAlu is the ratio of PQS3.1 in the i-family to the number of its members in this family.

As expected in [27] the direction of Alu insertion was not significant. Therefore, during comparative analysis of the effect of Alu and Alu-PQS we pooled together data obtained in the case of f/r plasmids. It has been shown that embryonic (HEK293) cells PQS-mutations resulted in a slight decrease of the Alu-effect (~20%). Since we failed to detect any statistically significant differences in the effects of Alu and Alu-PQS in experiments on HeLa cells (Fig. 3b), it was reasonable to suggest that the system of regulation of cancer cells was insensitive to PQS-Alu mutations.

Analysis of the Possibility of GQ Formation in Extended and Duplex Fragments of DNA_{Alu}

DMS-footprinting. A limited number of approaches is used for demonstration of GQ formation in extended sequence and/or duplex. It should be noted that the effectiveness of the DMS-footprinting and a method based on the GQ-dependent inhibition of elongation during PCR has been demonstrated only for several few highly stable GQ [22, 28]. Affinity interactions with proteins or small molecule ligands used to identify GQ-sites [29–32], apparently distort the original conformational equilibrium.

Theoretically, PQS-Alu can form only two-quartet GQ, and efficiency DMS-footprinting has not been erlier demonstrated for such cases. Nevertheless, we have attempted to detect a decrease in the rate of DMS methylation of PQS-guanines (potentially linked by Hoogsteen hydrogen bonds) in single stranded/double stranded (ss/ds) Alu-fragments compared to guanines which are not involved in proposed GQ tetrads. For this purpose we synthesized a set of oligomers and prepared duplexes, in which PQS-chains (DMS-PQS2, DMS-PQS3 and DMS-PQS4, table) carried the 5'-fluorescent label.

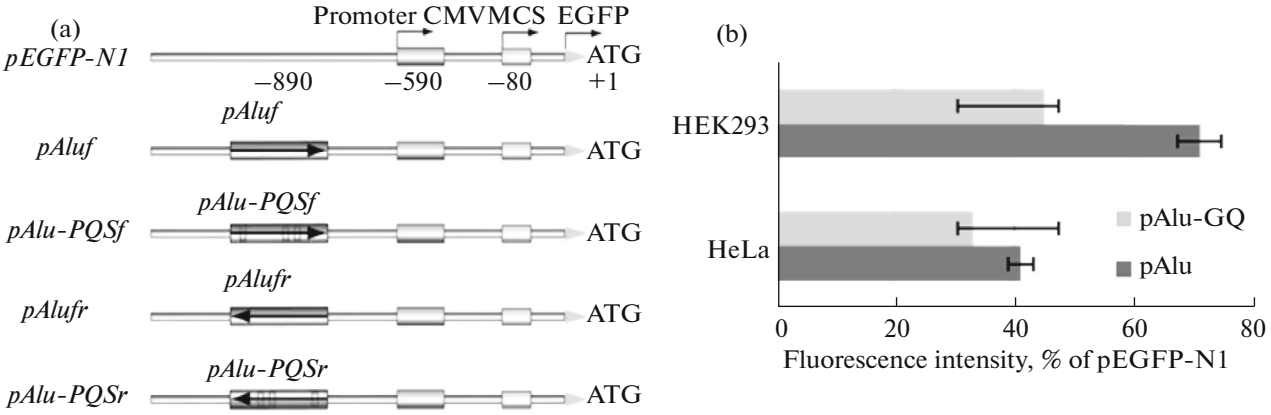


Fig. 3. The effect of Alu_{bcl2} and Alu-PQS on expression of the reporter gene. (a) Schematic structure of recombinant plasmids bearing Alu insertion in the promoter region of the reporter *GFP* gene. (b) Comparison of GFP expression levels in HeLa and HEK293 cells transfected with pAlu, pAlu-PQS, and pEGFP-N1 (control). Data for direct and inverse insertions have been pooled together. Fluorescence of transfected cells was measured by flow cytofluorimetry at 488 nm.

Figure 4 shows the results of PAGE analysis of the DMS-footprinting and histograms of relative intensity distribution (reaction in the presence of Li^+/K^+) of fluorescence of ss- and ds-oligomer fragments. Gel image data were normalized by signals of fragments marked on Fig. 4 with (Δ).

In contrast to the control (ss-control, right panel in Fig. 4), which was characterized by lack of K^+/Li^+-dependence of band intensities (deviation of fragments longer than 8 nucleotides did not exceed 3.5%), the presence of potassium ions significantly (20–60%) reduced the level of guanine methylation in some positions of PQS (2–4)-oligomers (Fig. 4). Moreover, single-stranded oligomers were characterized by constant presence of four G-blocks consisting of two or three units. In the case of DMS-PQS2 and DMS-PQS3 no statistically significant difference in signals associated with G cleavage in positions 32 and 33 was observed and therefore it was unlikely that GQ were present among equilibrium PQS conformations with inclusions of these G into tetrads.

Reduced difference between band intensities of oligomer DMS-fragments (Li^+-K^+) during G methylation in ds-PQS2 and ds-PQS3 obviously reflects a shift in the balance between non-canonical (non-B) form and duplex towards duplexes. At the same time, maintenance of similarities on ss- and ds-PQS4 histograms suggests GQ-folding of this site both in the untwisted state of DNA (for example, after the break), and within the duplex (see histograms in Fig. 4).

Although results of this study are not sufficient to make final conclusions about structure and topology of GQ, they do support the possibility of GQ formation in the considered Alu fragments. We can assume the presence of a mixture of conformers of DNA PQS loci, and most likely (for PQS2 and 3) the GQ-form is realized after melting of the DNA duplex (i.e., in the case of nicking or ds-breaks, during replication, transcription) or possibly in RNA.

A comparative study of the structure of the dsAlu and dsAlu-PQS amplicons (304 bp) by AFM. Using high-resolution AFM imaging it was possible to discriminate individual Alu DNA duplexes (Figs. 5a, 5b), obtained by PCR amplification of the plasmids dsAlu and dsAlu-PQS (dsAlu, 304 bp primers PrF and PrR, table). However, it was impossible to distinguish GQ and dsDNA bending by this method (Fig. 5b).

Formation of non-B sites within amplicons could be reliably detected by evaluation of their length in comparison with the DNA duplex [33]. Indeed, folding of GQ of 12–18 nucleotides in length could shorten the duplex by 3–5 nm.

Figure 5c shows histograms of the distribution of dsAlu and dsAlu-PQS lengths. The lengths of the mutant dsAlu-PQS-DNA follow the Gaussian distribution with maximum at 100 nm (theoretical length of dsAlu-PQS is 103.6 nm). The curve of natural dsAlu contains at least two peaks: the major peak in the region of 85–87 nm and the additional peak at ~100 nm. The length of the duplexes corresponding to the additional peak coincides with the length of the mutant duplex lacking possibility to form GQ, i.e., under our conditions of analysis the non-canonical structures are not detected in some molecules of the dsAlu amplicon. The difference between the dsAlu length of the major peak and the mutant can be estimated as 13–15 nm. Statistically significant DNA

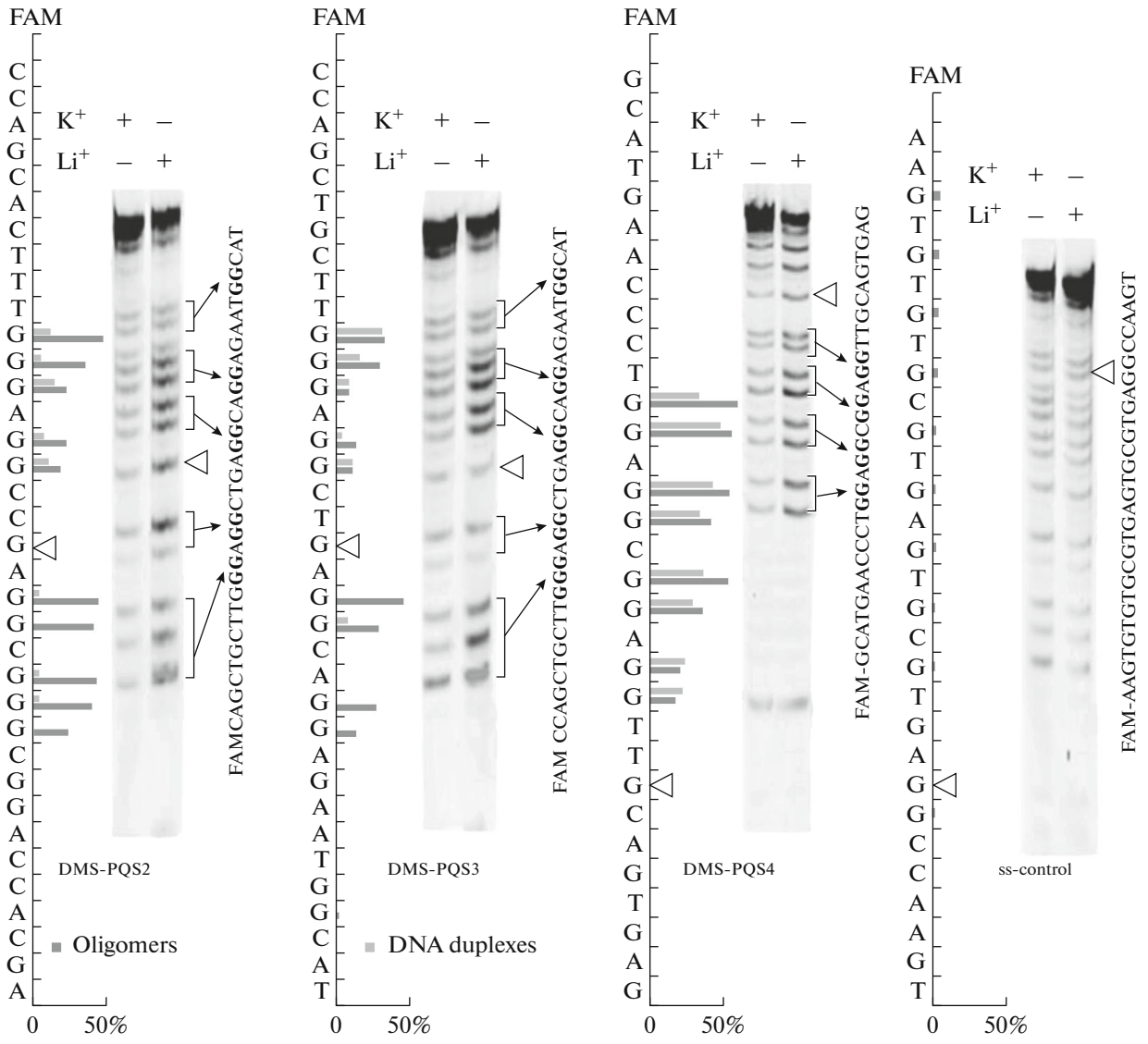


Fig. 4. DMS-footprinting of DMS-PQS2, DMS-PQS3, DMS-PQS4 and ss-control oligonucleotides (table). Histograms of relative fluorescence intensity of fragments correspond to difference of peak areas of the normalized (at positions marked by Δ) fluorescence profiles of electrophoregrams of oligomers and their duplexes after treatment with dimethyl sulfate in the presence of Li^+ and in the K^+ medium.

shortening reliably indicates the possibility of formation of the non-B structures in most $dsAlu_{bcl2}$ molecules. Moreover, based on the size, shortening, and PQS structures formation of three GQ is the most probable.

CONCLUSIONS

Analysis of G-rich fragments of the prooncogen *bcl2* intron Alu-repeat and the consensus sequence $AluS_x$ [6, 17] resulted in selection of PQS fragments that would be theoretically able to form GQ (Fig. 1). Bioinformatic analysis showed that they exclusively belonged to known Alu-repeats of the human genome. The occurrence frequencies of PQS_{Alu} were assessed as about ~1/10000 bp. It was found that they were most common for young (active) AluY-repeat families.

Experiments with recombinant plasmids bearing either Alu_{bcl2} or Alu-PQS (mutant) insertions before the *GFP* gene promoter confirmed ability of Alu-insertions to enhance expression of the reporter gene. It should be noted that the fluorescence intensity does not depend on the direction of Alu- insertion (as in the case of *CYP11B1* and *CYP11B2* genes involved in

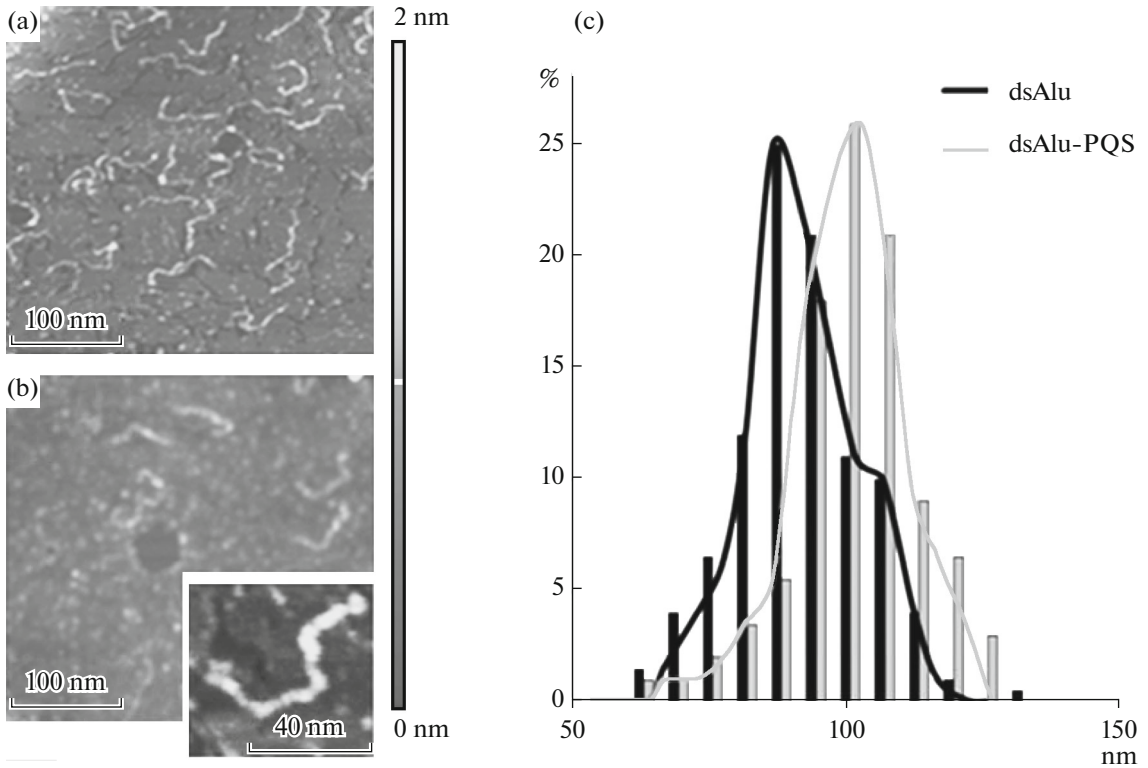


Fig. 5. AFM analysis of dsAlu-PQS (a) and dsAlu (b) preparations and the length distribution of DNA amplicons (c).

metabolism of steroids [27]. Mutations (GG→TT) preventing GQ assembly slightly decreased the effect thus suggesting possible involvement of PQS not only in the translocation processes but also in genomic regulation.

Using DMS-footprinting methods and AFM analysis we have reliably demonstrated for the first time possibility of existence of non-B structures in single-stranded and duplex Alu-fragments and DNA Alu repeat of the *bcl2* gene (304 bp).

In the second part of this study we will be consider the ability of G-rich Alu-fragments to form stable intramolecular and intermolecular GQ in vitro and describe their structure, topology, and some features.

ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation (project no. 14-25-00013).

REFERENCES

1. Hoffman, Y., Pilpel, Y., and Oren, M., *J. Mol. Cell. Biol.*, 2014, vol. 6, pp. 192–197. mju020 [pii] doi 10.1093/jmcb/mju020
2. Cui, F., Sirotin, M.V., and Zhurkin, V.B., *Biol. Direct.*, 2011, vol. 6, p. 2. 1745-6150-6-2 [pii] doi 10.1186/1745-6150-6-2
3. Spengler, R.M., Oakley, C.K., and Davidson, B.L., *Hum. Mol. Genet.*, 2014, vol. 23, pp. 1783–1793. ddt569 [pii] doi 10.1093/hmg/ddt569
4. Bose, P., Hermetz, K.E., Conneely, K.N., and Rudd, M.K., *PLoS One*, 2014, vol. 9, e101607. PONE-D-14-14652 [pii] doi 10.1371/journal.pone.0101607
5. Hoffman, Y., Dahary, D., Bublik, D.R., Oren, M., and Pilpel, Y., *Bioinformatics*, 2013, vol. 29, pp. 894–902. btt044 [pii] doi 10.1093/bioinformatics/btt044
6. Batzer, M.A. and Deininger, P.L., *Nat. Rev. Genet.*, 2002, vol. 3, pp. 370–379. [pii] doi 10.1038/nrg798
7. Chen, L.L. and Carmichael, G.G., *Cell Cycle*, 2008, vol. 7, pp. 3294–3301. 6927 [pii]
8. Kleinberger, Y. and Eisenberg, E., *BMC Genomics*, 2010, vol. 11, p. 453. 1471-2164-11-453 [pii] doi 10.1186/1471-2164-11-453
9. Wahlstedt, H. and Ohman, M., *Wiley Interdiscip. Rev. RNA*, 2011, vol. 2, pp. 761–771. doi 10.1002/wrna.89
10. Mallela, A. and Nishikura, K., *Crit. Rev. Biochem. Mol. Biol.*, 2012, vol. 47, pp. 493–501. doi 10.3109/10409238.2012.714350
11. Nishikura, K., *Nat. Rev. Mol. Cell. Biol.*, 2015. nrm.2015.4 [pii] doi 10.1038/nrm.2015.4
12. Saini, N., Zhang, Y., Usdin, K., and Lobachev, K.S., *Biochimie*, 2013, vol. 95, pp. 117–123. S0300-9084(12)00405-1 [pii] doi 10.1016/j.biochi.2012.10.005

13. Bharti, S.K., Sommers, J.A., Zhou, J., Kaplan, D.L., Spelbrink, J.N., Mergny, J.L., and Brosh, R.M., Jr., *J. Biol. Chem.*, 2014, vol. 289, pp. 29975–29993. M114.567073 [pii] doi 10.1074/jbc.M114.567073
14. Dong, D.W., Pereira, F., Barrett, S.P., Kolesar, J.E., Cao, K., Damas, J., Yatsunyk, L.A., Johnson, F.B., and Kaufman, B.A., *BMC Genomics*, 2014, vol. 15, p. 677. 1471-2164-15-677 [pii] doi 10.1186/1471-2164-15-677
15. Kejnovsky, E., Tokan, V., and Lexa, M., *Chromosome Res.*, 2015, vol. 23, pp. 615–623. [pii] doi 10.1007/s10577-015-9491-7
16. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., and Rehm, H.L., *Genet. Med.*, 2015, vol. 17, pp. 405–424. gim201530 [pii] doi 10.1038/gim.2015.30
17. Kriegs, J.O., Churakov, G., Jurka, J., Brosius, J., and Schmitz, J., *Trends Genet.*, 2007, vol. 23, pp. 158–161. S0168-9525(07)00037-6 [pii] doi 10.1016/j.tig.2007.02.002
18. Luk'yanova, T.A., Zaitseva, M.A., Karpov, V.A., and Pozmogova, G.E., *Bioorgan. Khim.*, 2008, vol. 34, pp. 83–88. doi 10.1007/s11171-008-1010-6
19. Tatarinova, O.N., Luk'yanova, T.N., Zaitseva, M.A., Veremeev, K.Yu., Karpov, V.A., Chuvilin, A.N., Petrunin, D.D., and Pozmogova, G.E., *Byull. Eksper. Biol. Med.*, 2008, vol. 145, pp. 280–284.
20. Tatarinova, O., Tsvetkov, V., Basmanov, D., Barinov, N., Smirnov, I., Timofeev, E., Kaluzhny, D., Chuvilin, A., Klinov, D., Varizhuk, A., and Pozmogova, G., *PLoS One*, 2014, vol. 9, e89383. PONE-D-13-39549 [pii] doi 10.1371/journal.pone.0089383
21. Klinov, D.V., Lagutina, I.V., Prokhorov, V.V., Neretina, T., Khil, P.P., Lebedev, Y.B., Cherny, D.I., Demin, V.V., and Sverdlov, E.D., *Nucl. Acids Res.*, 1998, vol. 26, pp. 4603–4610. gkb752 [pii]
22. Li, X.M., Zheng, K.W., Zhang, J.Y., Liu, H.H., He, Y.D., Yuan, B.F., Hao, Y.H., and Tan, Z., *Proc. Natl. Acad. Sci. USA*, 2015, vol. 112, pp. 14581–14586. 1516925112 [pii] doi 10.1073/pnas.1516925112
23. Daniel, C., Lagergren, J., and Ohman, M., *Biochimie*, 2015, vol. 117, pp. 22–27. S0300-9084(15)00170-4 [pii] doi 10.1016/j.biochi.2015.05.020
24. Luo, Y., Lu, X., and Xie, H., *Biomed. Res. Int.*, 2014, vol. 2014, p. 784706. doi 10.1155/2014/784706
25. Grandi, F.C. and An, W., *Mob. Genet. Elements*, 2013, vol. 3, e25674. 2013MGE0003R [pii] doi 10.4161/mge.25674
26. Burns, K.H. and Boeke, J.D., *Cell*, 2012, vol. 149, pp. 740–752. S0092-8674(12)00517-X [pii] doi 10.1016/j.cell.2012.04.019
27. Cheng, L.C., Pai, T.W., and Li, L.A., *Steroids*, 2012, vol. 77, pp. 100–109. S0039-128X(11)00314-X [pii] doi 10.1016/j.steroids.2011.10.010
28. Zheng, K.W., Chen, Z., Hao, Y.H., and Tan, Z., *Nucl. Acids Res.*, 2010, vol. 38, pp. 327–338. gkp898 [pii] doi 10.1093/nar/gkp898
29. Husby, J., Todd, A.K., Platts, J.A., and Neidle, S., *Biopolymers*, 2013, vol. 99, pp. 989–1005. doi 10.1002/bip.22340
30. Mela, I., Kranaster, R., Henderson, R.M., Balasubramanian, S., and Edwardson, J.M., *Biochemistry*, 2012, vol. 51, pp. 578–585. doi 10.1021/bi201600g
31. Henderson, A., Wu, Y., Huang, Y.C., Chavez, E.A., Platt, J., Johnson, F.B., Brosh, R.M., Jr., Sen, D., and Lansdorp, P.M., *Nucl. Acids Res.*, 2014, vol. 42, pp. 860–869. gkt957 [pii] doi 10.1093/nar/gkt957
32. Ma, D.L., Zhang, Z., Wang, M., Lu, L., Zhong, H.J., and Leung, C.H., *Chem. Biol.*, 2015, vol. 22, pp. 812–828. S1074-5521(15)00242-2 [pii] doi 10.1016/j.chembiol.2015.06.016
33. Rivetti, C. and Codeluppi, S., *Ultramicroscopy*, 2001, vol. 87, pp. 55–66.

Translated by A. Medvedev