OCTOBER-DECEMBER 2020 VOL. 12 № 4 (47)

ISSN 2075-8251

ActaNaturae

INSTITUTE OF GENE BIOLOGY

30 years

L-ASCORBIC ACID IN THE EPIGENETIC REGULATION OF CANCER DEVELOPMENT AND STEM CELL REPROGRAMMING

P. 5

THE FUNCTIONS AND MECHANISMS OF ACTION OF INSULATORS IN THE GENOMES OF HIGHER EUKARYOTES

P. 15

The Institute of Gene Biology Is Turning 30

The Institute of Gene Biology was established by a decree of the Presidium of the USSR Academy of Sciences thirty years ago and, since then, has successfully withstood the test of time. The Institute was founded as a "growth platform" for the numerous alumni and closest collaborators of Prof. Georgii Georgiev, one of the leading Soviet molecular biologists. Over the years, the Institute has changed in accordance with both circumstances and new trends in molecular and cellular biology retaining, however, its leading position in the study of the mechanisms of gene regulation. Although appreciably small, our institute is widely recognized for its crop of basic and translational research studies that focus on the mechanisms of gene expression regulation, spatial organization of the genome, and epigenetics.

Indeed, the Institute of Gene Biology is one of the smallest research institutions within the Russian Academy of Sciences. Its size, however, is actually an advantage, since inter-laboratory collaboration at our Institute is much more intense compared to the average. The percentage of young scientists among researchers working at the Institute is above 60%; almost 50% of all heads of research groups are young scientists. In our opinion, this state of affairs is rather unique and has its roots in the programs for the support of new research groups financially backed by the Presidium of the Russian Academy of Sciences and the concomitant long-term policy of the Institute. As a result, the mean age of the members of the Dissertation and Academic councils at the Institute is also unprecedentedly young. Such traits as keeping traditions alive and being receptive to new ideas make our Institute an attractive place to work for excellent scientists and help confront the many challenges that inevitably arise.

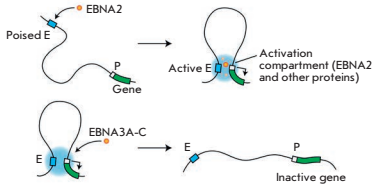
A lot has been accomplished at the Institute over the past 30 years: various discoveries were made in the field of gene functioning, new directions in research were set into motion, our research teams published articles in prestigious journals, etc. In order to outline, at least partially, today's interests and endeavors at the Institute, research teams from different laboratories have prepared a series of review articles which will be published in the current and upcoming issues of *Acta Naturae*. Our hope is that these reviews will spark interest in a broad audience and help us make new friends willing to collaborate. ●

Sincerely,
The Staff of the Institute of Gene Biology
Russian Academy of Sciences

Modification of Nuclear Compartments and the 3D Genome in the Course of a Viral Infection

S. V. Razin, A. A. Gavrilov, O. V. Iarovaia

The review addresses the question of how the structural and functional compartmentalization of the cell nucleus and the 3D organization of the cellular genome are modified during the infection of cells with various viruses. Particular attention is paid to the role of the introduced changes in the implementation of the viral strategy to evade the antiviral defense systems and provide conditions for viral replication.

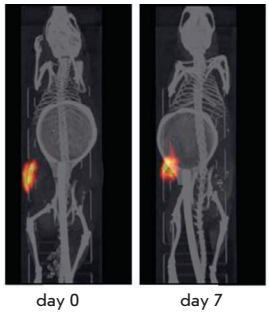


Virus-induced reorganization of the 3D genome

The Delivery of Biologically Active Agents into the Nuclei of Target Cells for the Purposes of Translational Medicine

A. S. Sobolev

A review of the author's laboratory research at the IBG RAS, which has enabled the development of modular nanotransporters – a basic technology that provides high efficiency and cell specificity to anticancer agents through their targeted delivery to the most sensitive compartment of target cells and facilitates penetration of antibodies in the desired compartment of living cells. Polycationic block copolymer of DNA complexes capable of preferential transfection of actively dividing cells have been developed.

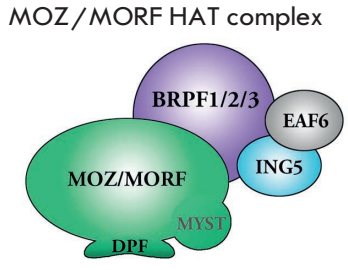


SPECT/CT visualization of radioactivity retention within the tumor

The DPF Domain As a Unique Structural Unit Participating in Transcriptional Activation, Cell Differentiation, and Malignant Transformation

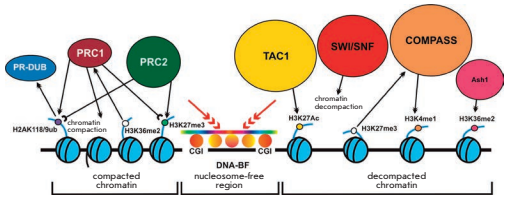
N. V. Soshnikova, A. A. Sheynov, Eu. V. Tatarskiy, S. G. Georgieva

The DPF domain consists of two PHD domains organized into a single structure. The DPF interacts with modified N-termini of H3K14ac/cr and H3K9ac histones of transcriptionally active chromatin. DPF-containing proteins are involved in the MYST histone acetyltransferase complex and SWI/SNF chromatin remodeling complex. Both of these complexes are transcriptional coactivators of genes expressed during organism development as well as differentiation and malignant transformation of mammalian cells.



Proteins and complexes containing the DPF domains

Polycomb and Trithorax Group Proteins: The Long Road from Mutations in *Drosophila* to Use in Medicine



Functional activities of the Polycomb/Trithorax group proteins

D. A. Chetverina, D. V. Lomaev, M. M. Erokhin

Polycomb group (PcG) and Trithorax group (TrxG) proteins are evolutionarily conserved factors responsible for the repression and activation of the transcription of multiple genes in Drosophila and mammals. In this review, authors focus on the major PcG and TrxG complexes, the mechanisms of PcG/TrxG action, and their recruitment to chromatin. They discuss the alterations associated with the dysfunction of a number of factors of these groups in oncology and the current strategies used to develop drugs based on small-molecule inhibitors.

ActaNaturae
OCTOBER-DECEMBER 2020 VOL. 12 № 4 (47)
since april 2009, 4 times a year

Founders
Acta Naturae, Ltd,
National Research University
Higher School of Economics

Editorial Council
Chairman: A.I. Grigoriev
Editors-in-Chief: A.G. Gabibov, S.N. Kochetkov

V.V. Vlassov, P.G. Georgiev, M.P. Kirpichnikov, A.A. Makarov, A.I. Miroshnikov, V.A. Tkachuk, M.V. Ugryumov

Editorial Board
Managing Editor: V.D. Knorre

K.V. Anokhin (Moscow, Russia)
I. Bezprozvanny (Dallas, Texas, USA)
I.P. Bilenkina (Moscow, Russia)
M. Blackburn (Sheffield, England)
S.M. Deyev (Moscow, Russia)
V.M. Govorun (Moscow, Russia)
O.A. Dontsova (Moscow, Russia)
K. Drauz (Hanau-Wolfgang, Germany)
A. Friboulet (Paris, France)
M. Issagouliants (Stockholm, Sweden)
A.L. Konov (Moscow, Russia)
M. Lukic (Abu Dhabi, United Arab Emirates)
P. Masson (La Tronche, France)
V.O. Popov (Moscow, Russia)
I.A. Tikhonovich (Moscow, Russia)
A. Tramontano (Davis, California, USA)
V.K. Švedas (Moscow, Russia)
J.-R. Wu (Shanghai, China)
N.K. Yankovsky (Moscow, Russia)
M. Zouali (Paris, France)

Project Head: N.V. Soboleva
Editor: N.Yu. Deeva
Designer: K.K. Oparin
Art and Layout: K. Shnaider
Copy Chief: Daniel M. Medjo

Phone/Fax: +7 (495) 727 38 60
E-mail: vera.knorre@gmail.com, actanaturae@gmail.com

Reprinting is by permission only.

© ACTA NATURAE, 2020

Номер подписан в печать 29 декабря 2020 г.
Тираж 100 экз. Цена свободная.
Отпечатано в типографии: НИУ ВШЭ,
г. Москва, Кочновский проезд, 3

Indexed in Pubmed, Web of Science, Scopus, and RISC

Impact Factor: 1.360

CONTENTS

RESEARCH ARTICLES

L-Ascorbic Acid in the Epigenetic Regulation of Cancer Development and Stem Cell Reprogramming

A. P. Kovina, N. V. Petrova, S. V. Razin, O. L. Kantidze*
Institute of Gene Biology Russian Academy of Sciences, Moscow, 119334 Russia
*E-mail: kantidze@gmail.com
Received June 26, 2020; in final form, September 07, 2020
DOI: 10.32607/actanaturae.11060
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

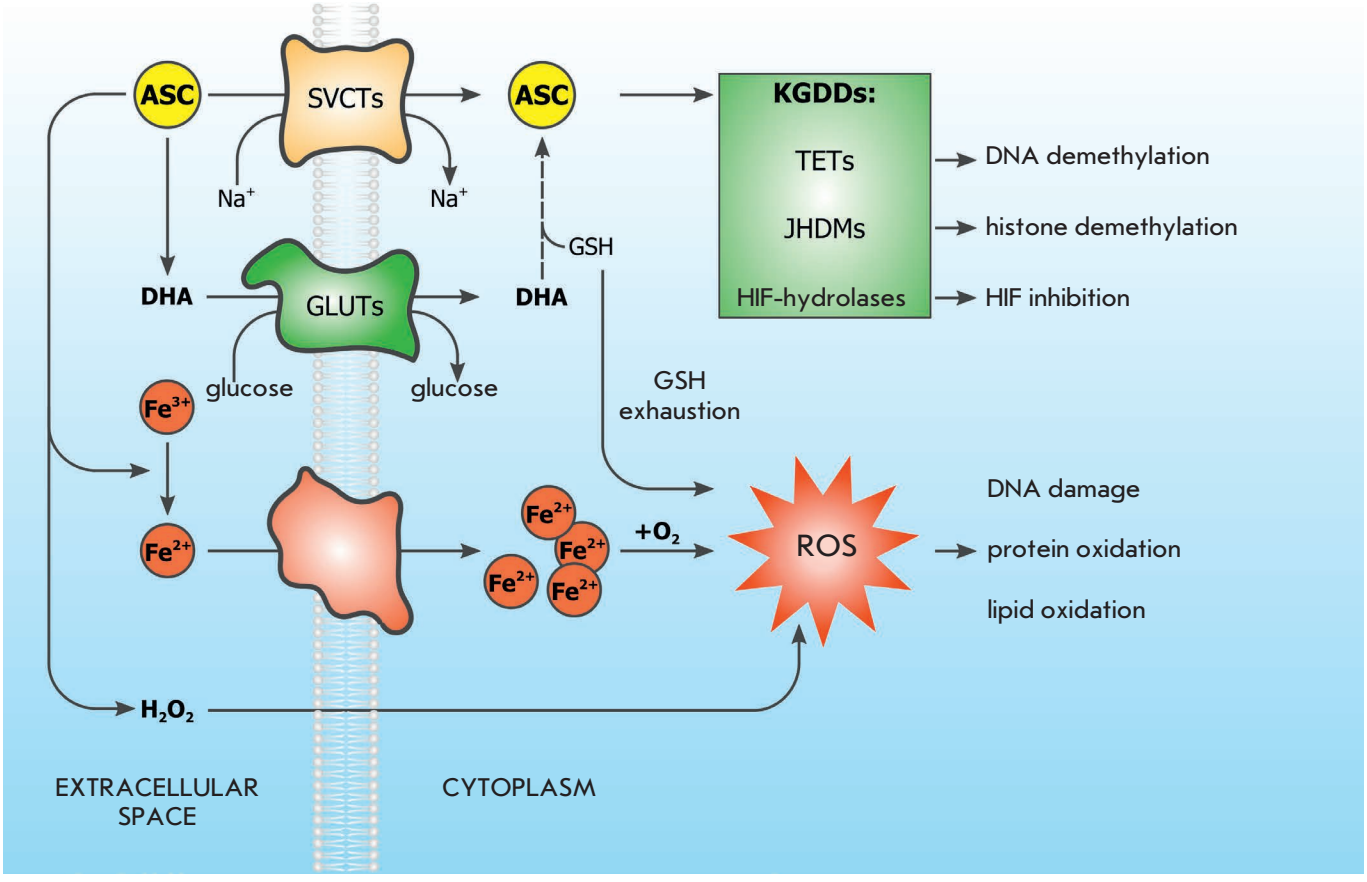
ABSTRACT **Recent studies have significantly expanded our understanding of the mechanisms of L-ascorbic acid (ASC, vitamin C) action, leading to the emergence of several hypotheses that validate the possibility of using ASC in clinical practice. ASC may be considered an epigenetic drug capable of reducing aberrant DNA and histone hypermethylation, which could be helpful in the treatment of some cancers and neurodegenerative diseases. The clinical potency of ASC is also associated with regenerative medicine; in particular with the production of iPSCs. The effect of ASC on somatic cell reprogramming is most convincingly explained by a combined enhancement of the activity of the enzymes involved in the active demethylation of DNA and histones. This review describes how ASC can affect the epigenetic status of a cell and how it can be used in anticancer therapy and stem cell reprogramming.**
KEYWORDS **vitamin C, cancer, stem cells, epigenome, chromatin.**
ABBREVIATIONS **5hmC – 5-hydroxymethylcytosine; 5mC – 5-methylcytosine; α-KG – α-ketoglutarate; AML – acute myeloid leukemia; ASC – *L*-ascorbic acid; BETi – bromodomain and extraterminal inhibitors; DHA – dehydroascorbic acid; DNMT – DNA methyltransferase; DNMTi – DNMT inhibitors; GSH – glutathione; Gulo – *L*-gulonolactone oxidase; IDH – isocitrate dehydrogenase; KGDD – α-KG-dependent dioxygenase; MEFs – mouse embryonic fibroblasts; P4H – prolyl 4-hydroxylase; PARP – poly(ADP-ribose)polymerase; TET – ten-eleven translocation dioxygenase; iPSCs – induced pluripotent stem cells.**

INTRODUCTION

L-ascorbic acid (ASC, vitamin C) belongs to a class identified as essential water-soluble vitamins. Primates, guinea pigs, and fruit bats, compared to most mammals, have lost the ability to synthesize ASC due to a mutation in the gene of the *L*-gulonolactone oxidase (Gulo) that catalyzes the last stage of ASC synthesis from glucose [1]. The concentration of ASC in the human body is regulated by several mechanisms at once, which ensure a plasma ASC level of not more than 80 μM (oral intake) [2]. In this case, most mammalian cells maintain concentrations of intracellular ASC, which can reach 1–10 mM. The sodium-dependent transporters SVCT1 and 2 (*Figure*), which are differentially expressed in different tissues, are responsible for the active transport of ASC into cells [3].

ASC is a good reducing agent: i.e., an electron donor. By donating the first electron, ASC transforms into the ascorbile radical, which is relatively stable and non-reactive. When it loses two electrons during two rounds of oxidation, ASC is converted to dehydroascorbic acid (DHA), which can be uptaken and released by the cell using the glucose transporters GLUT1, 2, 3, and 8 (*Figure*) [4]. Inside the cell, DHA can quickly get reduced to ASC by reaction with reduced glutathione (GSH) (*Figure*) [4]. In blood plasma, the reduced form of ASC predominates, while the DHA concentration is very low [5].

At micromolar concentrations, ASC can act as an antioxidant. ASC serves as a cofactor for several monooxygenases and Fe^{2+}/α-ketoglutarate (α-KG)-dependent dioxygenases (KGDDs), acting as an electron donor (*Figure*) [6]. A classic example of α-KG-dependent dioxygenases is collagen-prolyl-4-hydroxylase (P4H), which has been well studied because decreased P4H activity causes scurvy. Accumulation of Fe^{3+} ions due to the activity of this enzyme leads to the inhibition of P4H activity and, therefore, to an incomplete hydroxylation of proline residues in the collagen molecule, aberrant collagen crosslinking, and scurvy symptoms [7]. ASC can reduce oxidized Fe^{3+} ions to catalytically active Fe^{2+} and, thus, prevent the development of scurvy. ASC, as a KGDD cofactor, affects important biological functions, such as catecholamine synthesis, collagen



The role of ASC in the modulation of the epigenetic and redox statuses of the cell (see text for details). Abbreviations: ROS – reactive oxygen species; ASC – *L*-ascorbic acid; DHA – dehydroascorbic acid; GLUTs – glucose transporters; GSH – glutathione; HIFs – hypoxia-induced transcription factors; JHDM – JmjC-containing histone demethylases; KGDDs – α-ketoglutarate-dependent dioxygenases; SVCTs – Na^+ and ASC transporters; TETs – methylcytosine dioxygenases

crosslinking, alkylated DNA repair, and hypoxia-inducible factor 1α (HIF-1α) degradation. A particular KGDD group consists of enzymes that catalyze hydroxylation of methylated nucleic acids (DNA and RNA) and methylated histones. Some of these dioxygenases require ASC as a cofactor in histone and DNA demethylation. The discovery of ASC-dependent KGDDs that are involved in the hydroxylation of methylated nucleic acid bases and histone amino acid residues suggests that ASC plays a role in the epigenetic regulation of gene expression.

ASC AND DNA METHYLATION

Methylation of cytosine at the fifth position (5-methylcytosine (5mC)) is the most studied DNA modification occurring in mammals; it plays an important role in the epigenetic regulation of gene expression. Methylation of CpG nucleotides in promoters is usually associated with transcriptional repression and is involved in many processes, including X-chromosome inactivation and imprinting. 5mC is a very stable epigenetic label that can be removed in two ways: passive and active. Passive removal leads to dilution of the label during DNA replication in the absence of maintenance DNA methyltransferase (DNMT1) [8], while active demethylation is associated with the Ten-Eleven Translocation (TET) enzyme group that includes TET1–3 [9]. TETs are Fe^{2+}/α-KG-dependent dioxygenases capable of sequential oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), which are recognized and removed by DNA repair enzymes [10, 11]. Unlike 5fC and 5caC, 5hmC is relatively stable; it can perform its own epigenetic function, because there exists a group of regulatory proteins that can specifically recognize and interact with 5hmC [10].

Because ASC is known to be a cofactor of some Fe^{2+}/α-KG-dependent dioxygenases, ASC has been thought to be also a cofactor for TET-mediated DNA demethylation. Indeed, the addition of ASC to the culture medium was found to cause demethylation of several thousand genes in human embryonic stem cells (ESCs) [12]. In this regard, it is appropriate to recall that ASC promotes the formation of induced pluripotent stem cells (iPSCs) from terminally differentiated cells, which is accompanied by demethylation of the entire genome [13, 14]. *In vivo*, ASC was shown to enhance the generation of 5hmC in cultured cells. Most likely, ASC acts as a TET cofactor in the 5mC hydroxylation reaction [15, 16], because the addition of ASC dose-dependently increases the amount of 5hmC in mouse embryonic fibroblasts (MEFs) and this effect is abrogated by TET knockdown. The involvement of ASC in DNA demethylation has been observed in different cell types, as well as in model animals [17–19].

Interestingly, standard culture media lack ASC and the level of 5hmC in cultured cells is usually very low. Addition of ASC rapidly boosts the formation of 5hmC [20, 21]. This suggests that protein synthesis is not required for that task, but existing TET dioxygenases are activated [16]. According to the results of other experimental studies, ASC is required as a TET cofactor, but not just as a reducing agent. For example, addition of another reducing agent, GSH, did not change the 5hmC level; this indicates that the effect of ASC on 5hmC generation cannot be attributed to its role as a general reducing agent [16]. In mice with knockout of the *Gulo* gene ($Gulo^{-/-}$), which is necessary for ASC biosynthesis, decreased amounts of 5hmC in various tissues were observed [19]. ASC was also shown to significantly increase the levels of all 5mC oxidation products, including 5fC and 5caC [17, 19]. ASC can also directly affect the functioning of TET family proteins, interacting with the C-terminal catalytic domain of enzymes, which probably promotes their correct folding and/or reuse of Fe^{2+} [19].

Therefore, there is convincing evidence that ASC acts as a cofactor of TET dioxygenases in 5mC oxidation, which is the first stage of active DNA demethylation.

ASC AND HISTONE METHYLATION

Methylation of lysine and arginine residues in histones is an important epigenetic tool. While histone acetylation is usually considered an activating modification, methylation can be considered as a marker of both active (e.g., H3K4, H3K36, and H3K79) and inactive (e.g., H3K9, H3K27, and H4K20) chromatin [22]. Like DNA methylation, histone methylation was initially considered an irreversible post-translational modification. In the early 2000s, the lysine-specific histone demethylases KDM1A (LSD1) and KDM1B (LSD2) were discovered. They are capable of demethylating only mono- and di-, but not trimethylated lysine residues in the histone molecule [23, 24]. Later, the enzyme KDM4A (JHDM3A) was discovered, which is also capable of removing the third methyl group from the lysine residues 9 and 36 in the histone H3 molecule [25]. Further, other similar enzymes were identified, which, like KDM4A, contain the Jumonji C (JmjC) domain. This catalytic domain provides the hydroxylase activity of demethylases, which is necessary for the demethylation of the amino acid residues in histones [26]. JmjC domain-containing demethylases also belong to the family of Fe^{2+}/α-KG-dependent dioxygenases, the general functioning principles and cofactors of which were discussed earlier [25, 27].

JmjC-containing enzymes were found to require ASC. *In vitro*, ASC is required for both KDM2A and KDM3A (JHMD2A): the activity of these enzymes was in correlation with the amount of ASC in the reaction buffer [27]; in this case, KDM4A completely lost its catalytic activity in *in vitro* experiments upon removal of ASC from the medium [25].

The investigation of the differentiation of various cells has demonstrated that this process is significantly impaired in the absence of ASC due to the inability of cells to control the repressive histone modification level. For example, the absence of ASC during the endothelial-to-hematopoietic transition leads to an accumulation of H3K27me3 in the genomic loci that are important for hematopoiesis [28]. An excess of ASC underlies the loss of histone H3 lysine 9 dimethylation (H3K9me2) within extended genomic domains in mouse embryonic stem cells (LOCK domains [29]), which is apparently caused by the stimulation of the demethylases Kdm3a and Kdm3b [30]. Addition of ASC to T-lymphocytes leads to a decrease in the H3K9me3 level in the *cis*-regulatory elements of the interleukin-17 (IL-17) gene locus, due to the activation of histone demethylase KDM4A and, accordingly, to an increase in the IL-17 expression [31]. In addition, ASC was shown to stimulate histone demethylation during both the initial stages of reprogramming of somatic cells into iPSCs [32] and the transition from pre-iPSCs to completely reprogrammed iPSCs [33, 34]. All these findings suggest that ASC is a cofactor of JmjC-containing histone demethylases and that it modulates histone demethylation, most likely through the regeneration of the catalytically active Fe^{2+}.

ASC AND CANCER

Any low-molecular-weight compounds capable of modifying epigenetic profiles are considered potential

anticancer agents. The question of whether ASC may be used as an anticancer agent has been debated for decades. Interest in the possible use of ASC in cancer therapy emerged back in the 1970s, when Pauling and Cameron reported an increased survival rate in patients with late-stage cancer after an intravenous administration of ASC (10 g per day), but later attempts to repeat these results failed [35]. This was related to the method used for ASC delivery: later studies used oral administration, which prevented the achieving of therapeutically significantly high ASC concentrations in the blood [36]. Further research led to the emergence of new hypotheses about the potential mechanisms underlying the anticancer activity of ASC. As in the case of other chemotherapeutic agents, different tumor types exhibit different sensitivities to the cytotoxic effect of ASC [37]. ASC concentrations of about 2–5 mM are sufficient to reduce the survival rate of most *in vitro* cultured cancer cells by 50%. At the same time, many non-cancerous cells maintain normal activity at ASC concentrations of about 20 mM [37]. It should, however, be noted that about 10–15% of cancer cell types are insensitive to ASC even at a concentration of 20 mM.

Potential mechanisms of anticancer activity of ASC

The mechanisms of anticancer activity of ASC can be divided into two groups: mechanisms affecting redox biology, and mechanisms associated with the function of ASC as a cofactor of α-KG-dependent dioxygenases (*Figure*).

The first group includes two mechanisms that are not mutually exclusive, and their combined action may result in ASC toxicity to cancer cells. The prooxidant properties of ASC at millimolar (pharmacological) concentrations may increase the amount of non-reparable lesions to a cancer cell. ASC accelerates the Fe^{2+}-dependent production of the hydroxyl radical (· OH) from H_2O_2 through oxidation of Fe^{3+} ions to labile iron ions (Fe^{2+}), thereby continuously generating reactive oxygen species (ROS) and promoting cell death [38]. In addition, spontaneous autooxidation of ASC by oxygen can lead to the accumulation of H_2O_2, high concentrations of which cause cell death (*Figure*) [37, 39, 40].

The second mechanism from this group is extracellular oxidation of ASC to DHA that is structurally similar to glucose and is transported into cells via GLUT transporters, which promotes an increase in the intracellular DHA pool. Cancer cells can transport DHA into the cell, where it is reduced to ASC, which leads to the depletion of the pool of glutathione and NADH- and NADPH-dependent enzymes [4]. This, in turn, causes oxidative stress and inactivation of glyceraldehyde-3-phosphate dehydrogenase, inhibits glycolysis, the level of which is increased in cancer cells, and leads to an energy crisis that is fatal to cells (*Figure*) [41, 42].

As a cofactor of Fe^{2+}/α-KG-dependent dioxygenases, ASC can also significantly affect the viability of cancer cells. Hypoxia-induced transcription factors (HIFs) increase the expression of the genes responsible for a successful adaptation of cancer cells to the hypoxia caused by rapid cell division and insufficient vascularization of a growing tumor [43]. HIF activity is controlled by HIF hydroxylases that modify, at normal conditions (normoxia), subunits of these factors, which promotes their proteasomal degradation [44]. HIF-hydroxylases belong to the family of dioxygenases, and ASC may be their cofactor [45]. ASC-deficient cells exhibit reduced HIF-hydroxylase activity and, therefore, an increased level of HIF-factor transcription, in particular in mild or moderate hypoxia [46–48]. These findings suggest that the addition of ASC to cancer cells may stimulate the activity of HIF hydroxylases and decrease HIF activity, thereby slowing down the rate of tumor growth (*Figure*) [49, 50].

As a cofactor of the enzymes of the Fe^{2+}/α-KG-dependent dioxygenase family, ASC influences the epigenetic alterations that are often inextricably linked with the development of cancer (*Figure*). There are important epigenetic alterations characteristic of cancers. First, one of the cancer markers is the global DNA hypomethylation that can activate a transcription of transposons and oncogenes, which leads to changes in gene expression and, subsequently, to carcinogenesis [51]. Second, it is the hypermethylation of tumor suppressor gene promoters. As was recently shown, the hydroxymethylation level (5hmC) can also change in some cancers [10]. The possibilities of using ASC to modulate the epigenetic status of cancer cells are discussed in detail in the next section.

Biomarkers for using ASC in anticancer therapy

In recent years, growing interest has been directed at the role of ASC in the modulation of DNA and histone methylation profiles, which is due to the fact that ASC is a cofactor of the enzymes involved in the demethylation of DNA (TET) and histones (JmjC-containing demethylases) [9, 52]. Changed expression levels of these enzymes and/or mutations in their genes have been found both in various solid tumors and in hematological malignancies. Because mutations usually involve only one copy of the gene, the addition of ASC can compensate for the effect of this mutation through an increased activity of the remaining non-mutant enzyme [52].

Mutations in the *TET* genes are observed in hematological malignancies, both myeloid and lymphoid [53], and usually lead to DNA hypermethylation [54–56].

In this case, ASC acts as an epigenetic modulator: in ASC-treated cancer cells, the TET activity is increased, which leads to DNA demethylation, and expression of tumor suppressor genes, such as Smad1, is increased [55].

Mutations in cancers often involve genes that are directly associated with the TET activity. For example, the isocitrate dehydrogenases IDH1 and IDH2, which are required for the production of the TET cofactor α-KG, are often mutated in hematological malignancies, as well as in some subtypes of gliomas and solid tumors [57]. In most cases, these mutations lead to an increased level of 2-hydroxyglutarate and, as a consequence, to DNA hypermethylation and reduced 5hmC levels. Several studies have been performed on mouse and cell models of leukemia caused by mutations in the *TET2* or *IDH1* gene [52, 55, 58, 59]. Upon intravenous administration of ASC, as well as upon restoration of *TET2* expression, DNA hypermethylation was suppressed or decreased due to increased DNA demethylation [52, 55, 59]. Interestingly, after the addition of ASC, leukemia cells became more sensitive to the inhibition of poly(ADP-ribose) polymerases (PARPs), which can be used as an effective, combined strategy for the treatment of cancers with mutations in the *TET* gene [52]. The effect of ASC addition was also tested on IDH1 mutant mouse leukemia cells [59]. ASC was shown to induce a TET2-dependent increase in the amount of 5hmC, loss of 5mC, and increased expression, which was in correlation with a decreased self-renewal of leukemic stem cells and enhanced differentiation towards the mature myeloid phenotype [59]. These data indicate that ASC can, at least in part, mitigate the effect of TET and IDH loss.

Brain tissues possess the highest need in intracellular ASC, because it is involved in the enhancement of the biosynthesis of norepinephrine and acts as a cofactor of dopamine-β-hydroxylase, as well as an inhibitor of glutamate uptake in retinal neurons. An oxidized form of ASC (DHA) is able to penetrate the blood-brain barrier and then accumulate in the stem cells of the cortex and cerebellum, the neurons, and neuroblastoma cells [60, 61]. The mechanism of ASC action in glioma is believed to have something to do with its prooxidant properties. Clinical studies have shown that the combination of conventional therapies with intravenous administration of high ASC doses improves the quality of life of glioblastoma patients, increases their overall survival likelihood, and arrests the progression of the disease [62, 63].

The genes of fumarate hydratase (FH) and succinate dehydrogenase (SDH) are mutated in many cancer types [64, 65]. Mutations in these genes lead to the accumulation of succinate and fumarate, which act as oncometabolites, competitively inhibiting TET and JmjC-containing histone demethylases, even in the presence of stable α-KG levels [66]. Indeed, *FH* or *SDH* knockdown in mouse liver cells has led to a decrease in the 5hmC level [66]. The effect of ASC on cells with mutations in the *FH* or *SDH* gene has not yet been explored, but it may be suggested that enhancement of the enzymatic activity of TET or JmjC-containing demethylases may be sufficient to restore the normal epigenetic landscape, even in the presence of inhibitory oncometabolites.

ASC as adjuvant therapy

Potential interactions between ASC and chemotherapeutic agents have long been a controversial issue [67]. Animal studies have shown that the simultaneous use of high ASC doses and various chemotherapeutic agents slows the growth of a xenograft tumor [68–70]. Many *in vivo* studies have shown that orally or intravenously administered ASC decreases the level of general toxicity of chemotherapeutic agents [71]. ASC administration reduced leukocyte loss, weight loss, accumulation of ascites, hepatotoxicity, lipid oxidation, and chemotherapy-induced cardiomyopathy in [69, 72].

In clinical trials involving patients with different types of cancer, intravenous administration of high ASC doses, together with chemotherapeutic agents, showed no side effects and, in many cases, improved health and quality of life [69, 73, 74]. It has been often noted that combination therapy involving ASC increases sensitivity to certain anticancer drugs and, therefore, has the potential to reduce the required dose and side effects [52, 75]. A reduction in chemotherapy-associated toxicity was observed, e.g., in patients with stage III–IV ovarian cancer who had received carboplatin and paclitaxel, in combination with a high dose of ASC [69].

The large-scale DNA demethylation observed upon the addition of ASC to human leukemia cell lines is associated with the increased TET2 activity in them [52, 76]. DNA methyltransferase inhibitors (DNMTis) such as 5-azacytidine and decitabine reduce aberrant DNA hypermethylation by suppressing the activity of supporting and *de novo* DNA methyltransferases [77]. The synergistic action of ASC and DNMTi causes both passive and active DNA demethylation, which leads to cancer cell proliferation inhibition and apoptosis [76]. The results of clinical trials performed to date confirm, in general, the efficacy of a combined use of ASC and DNMTi [74].

ASC enhances the cytotoxic effect of a PARP1/2 inhibitor, olaparib, on human acute myeloid leukemia (AML) cells [52]. Probably, this is a case of synthetic lethality: TET-mediated DNA oxidation caused by

ASC sensitizes AML cells to PARP inhibition due to the impossibility of removing non-canonical bases from DNA.

ASC also increases the sensitivity of melanoma cells to the bromodomain and extraterminal motif-containing protein inhibitors (BETi) that cause changes in the level of histone acetylation and are considered promising agents for the treatment of cancers [75]. ASC enhances the effectiveness of BETi by decreasing the level of histone H4 acetylation via the TET-dependent suppression of the histone acetyltransferase 1 (HAT1) expression.

In the population, the average rate of ASC deficiency is low, but it is much higher in patients with advanced cancer [78]. ASC deficiency is detected in most patients with hematological malignancies [76, 79]. Even in the absence of mutations in the *TET* genes, ASC deficiency can further impair the function of TET proteins upon suppression of tumor progression. Administration of some anticancer drugs, such as cisplatin, fluorouracil, nilotinib, and interleukin-2, was shown to significantly reduce the ASC level [80, 81]. Therefore, ASC deficiency can increase the aggressiveness of the disease and increase the risk of a relapse.

ASC AND STEM CELL REPROGRAMMING

ASC and embryonic development

In the early stages of mammalian embryonic development, there are two rounds of DNA demethylation that occurs in both passive and active ways. Immediately after fertilization, 5mC in the paternal chromatin is quickly replaced by 5hmC via TET3-mediated hydroxylation, after which the formed 5hmC is diluted during the DNA replication of implanted embryos [82]. This leads to an almost complete disappearance of the 5mC pattern in the paternal chromatin as early as at the stage of 16 cells – methylation is retained only at imprinted genomic loci [82, 83]. Maternal chromatin demethylation, which occurs a little later, is also mediated by both TET3-dependent oxidation and passive demethylation [84, 85]. After embryo implantation, the internal cell mass, which gives rise to the embryo, undergoes *de novo* DNA methylation [86]. The second stage of DNA demethylation, which includes, inter alia, demethylation of imprinted loci, occurs in primary germ cells [87, 88].

A significant amount of ASC, as a cofactor, is required to satisfy the cell's TET needs, and the lack of ASC can impair embryonic development due to incomplete DNA demethylation, which may lead to congenital anomalies. ASC is required for TET-dependent demethylation of many promoters and activation of germline genes in mouse and human embryonic stem cells [12, 17]. Histone demethylation mediated by JmjC-containing histone demethylases is critical for embryonic development [89–92]. Maternal and paternal nutrition was shown to affect DNA and histone methylation patterns in offspring cells [93, 94]. As shown in a mouse model, ASC consumption is necessary for proper DNA demethylation and further development of female germ cells in the fetus [95]. ASC deficiency in the mother does not affect the overall development of the fetus, but it leads to a decreased amount of germ cells, delayed meiosis, and reduced fertility in offspring [95]. The effects of ASC deficiency in pregnancy are partially similar to those of *TET1* knockout.

In general, ASC, supporting the catalytic activity of TET and some JmjC-containing histone demethylases, especially during epigenetic reprogramming, may be required in the early stages of embryonic development.

ASC and somatic cell reprogramming

The ability to reprogram somatic cells into iPSCs that can further be used to produce various differentiated cell populations is an important tool in regenerative medicine [96, 97]. Induction of the transcription factors Oct4, Sox2, Klf4, and c-Myc (OSKM) leads to the production of iPSCs from differentiated somatic cells [96, 98, 99]. The effectiveness of the reprogramming is low due to factors such as the age of the cell donor, number of passages in the culture, and the tissue origin of the cells [100–102]. Reprogramming is based on two main processes: repression of differentiation genes and activation of the genes that regulate pluripotency. Removal of epigenetic modifications in the genome of somatic cells is critical to the success of reprogramming [103]. Numerous studies in the past decade have shown that the addition of ASC to a medium of cultured somatic cells increases the effectiveness of reprogramming and the quality of the obtained iPSCs [13, 14, 34]. By enhancing the catalytic activity of TET and JmjC-containing histone demethylases, ASC stimulates histone and DNA demethylation in somatic cells, which may simultaneously activate the expression of pluripotency genes and erase the epigenetic memory of the differentiated state in mature cells.

In the first studies, ASC was added to the culture medium for reprogramming as an antioxidant to mitigate the effects of ROS, the level of which was increased upon induced expression of OSKM [104]. However, ASC enhanced the proliferation of ESCs and generation of iPSCs from mouse and human fibroblasts more efficiently than other antioxidants [13]. ASC is supposed to promote cell reprogramming because of the increased histone demethylation that is necessary for the expression of Nanog, one of the main transcrip-

tion factors [105]. Indeed, addition of ASC-dependent KGDD inhibitors impaired iPSC formation from MEFs [34].

One of the obstacles to somatic cell reprogramming is histone H3K9 methylation [33]. Addition of ASC to the pre-iPSCs occurring in an intermediate reprogramming state leads to their transformation into fully reprogrammed iPSCs [13]. This may be explained by the fact that the presence of ASC promotes a more efficient demethylation of histone H3K9 associated with the genes of pluripotency-regulating transcription factors, which leads to an increase in their expression [33]. The effectiveness of reprogramming increases upon simultaneous addition of ASC and inhibition of H3K9-specific methyltransferases [13]. Genome-wide screening using RNA interference helped to identify histone demethylase Kdm3b (Jhdm2b) as the main target activated by ASC during cell reprogramming [33]. Also, an increase in the activity of the demethylases Kdm3a/b (Jmjd1a/b) and Kdm4b/c (Jmjd2b/c) by ASC in mouse ESCs and in pre-iPSCs was shown to lead to a specific loss of H3K9me2/me3 in the loci of the genes responsible for pluripotency [30, 33].

Another JmjC-containing enzyme from the Kdm group, Kdm6a (Utx), demethylates H3K27me3 and is the most important regulator of pluripotency induction during the reprogramming of mouse and human somatic cells [106]. Addition of ASC to the culture medium of mouse ESCs alters the distribution of H3K27me3 in their genome, and this occurs mainly locus-specifically [30], the reasons for which remain to be clarified.

An analysis of changes in the methylated H3K36 profiles during the reprogramming of MEFs into iPSCs demonstrated that ASC causes a noticeable decrease in H3K36me2/3 due to an increase in the activity of the histone demethylases Kdm2a/2b (Jhdm1a/1b) [34]. This, inter alia, decreases the expression level of cyclin-dependent kinase inhibitor genes at the INK4/ARF locus and removes restrictions on the reprogramming of somatic cells [101, 107]. Reprogramming using expression of Oct4 and histone demethylase KDM2B in the presence of ASC is known to activate the expression of the miR302/367 microRNA cluster [34]. KDM2B causes an ASC-dependent decrease in the methylation levels of H3K36 that surrounds the Oct4 binding sites located near the miR302/367 gene and promotes their expression [34]. The miR302/367 cluster regulates pluripotency by inhibiting the expression of the genes important for differentiation [108]. Because these microRNAs play a decisive role in maintaining cell pluripotency, their expression decreases during differentiation [109]. It is noteworthy that expression of the entire miR302/367 cluster is sufficient for the reprogramming of fibroblasts [110].

Expression of *TET* genes plays an important role in somatic cell reprogramming. Knockdown of *TET* genes significantly complicates, and in some cases even completely prevents, the reprogramming of MEFs into iPSCs by the expression of OSKM [20, 111, 112]. As expected, ASC increases the effectiveness of reprogramming mouse and human fibroblasts into iPSCs in a TET-dependent manner [16–19]. For a more efficient reprogramming of mouse iPSCs into the naive pluripotency state, ASC can be used, together with vitamin A (retinoic acid), which activates TET2 and TET3 transcription through specific signaling pathways [13, 113, 114].

Along with its important role in somatic cell reprogramming, ASC is also required in order to maintain proliferation and a normal differentiation potential for ESCs, iPSCs, neuronal stem cells, and mesenchymal stem cells [115]. Most likely, the involvement of ASC in the prevention of premature aging for these cell cultures and the preservation of their epigenetic plasticity is mediated by its role as a cofactor of DNA and histone demethylation enzymes.

CONCLUSION

Recent studies have significantly expanded our understanding of the mechanisms underlying ASC action, which has produced several hypotheses that validate the possibility of its use in clinical practice. ASC may be considered an epigenetic drug capable of reducing aberrant DNA and histone hypermethylation, which may be helpful in the treatment of some cancers and neurodegenerative diseases. A correct understanding of the mechanisms of ASC action and the ongoing clinical studies will help identify the types of cancer patients that may benefit from a high-dose ASC treatment. Intravenous administration of ASC can act alone, or in combination with different chemotherapeutic agents. Preclinical and clinical trials have demonstrated that the toxicity and side effects of chemotherapy in this case can be mitigated without decreasing tumor-specific cytotoxic activity. On the other hand, the clinical significance of ASC is associated with regenerative medicine, in particular with the production of iPSCs from somatic cells. The effect of ASC on somatic cell reprogramming is most convincingly explained by the combined enhancement of the activity of the enzymes involved in the active demethylation of DNA and histones. ●

This study was supported by the Russian Foundation for Basic Research (Grant No. 17-00-00098) and Russian Science Foundation (Grant No. 19-74-10009).

REFERENCES

1. Linster C.L., van Schaftingen E. // FEBS J. 2007. V. 274. № 1. P. 1–22.
2. Levine M., Conry-Cantilena C., Wang Y., Welch R.W., Washko P.W., Dhariwal K.R., Park J.B., Lazarev A., Graumlich J.F., King J., et al. // Proc. Natl. Acad. Sci. USA. 1996. V. 93. № 8. P. 3704–3709.
3. Burzle M., Hediger M.A. // Curr. Top. Membr. 2012. V. 70. P. 357–375.
4. Ferrada L., Salazar K., Nualart F. // J. Cell. Physiol. 2019. V. 234. № 11. P. 19331–19338.
5. Lykkesfeldt J. // Cancer Epidemiol. Biomarkers Prev. 2007. V. 16. № 11. P. 2513–2516.
6. Young J.I., Zuchner S., Wang G. // Annu. Rev. Nutr. 2015. V. 35. P. 545–564.
7. Gorres K.L., Raines R.T. // Crit. Rev. Biochem. Mol. Biol. 2010. V. 45. № 2. P. 106–124.
8. Bhutani N., Burns D.M., Blau H.M. // Cell. 2011. V. 146. № 6. P. 866–872.
9. Lorsbach R.B., Moore J., Mathew S., Raimondi S.C., Mukatira S.T., Downing J.R. // Leukemia. 2003. V. 17. № 3. P. 637–641.
10. Rausch C., Hastert F.D., Cardoso M.C. // J. Mol. Biol. 2019. V. 432. № 6. P. 1731–1743.
11. Kantidze O.L., Razin S.V. // Cell Cycle. 2017. V. 16. № 16. P. 1499–1501.
12. Chung T.L., Brena R.M., Kolle G., Grimmond S.M., Berman B.P., Laird P.W., Pera M.F., Wolvetang E.J. // Stem Cells. 2010. V. 28. № 10. P. 1848–1855.
13. Esteban M.A., Wang T., Qin B., Yang J., Qin D., Cai J., Li W., Weng Z., Chen J., Ni S., et al. // Cell Stem Cell. 2010. V. 6. № 1. P. 71–79.
14. Stadtfeld M., Apostolou E., Ferrari F., Choi J., Walsh R.M., Chen T., Ooi S.S., Kim S.Y., Bestor T.H., Shioda T., et al. // Nat. Genet. 2012. V. 44. № 4. P. 398–405.
15. Dickson K.M., Gustafson C.B., Young J.I., Zuchner S., Wang G. // Biochem. Biophys. Res. Commun. 2013. V. 439. № 4. P. 522–527.
16. Minor E.A., Court B.L., Young J.I., Wang G. // J. Biol. Chem. 2013. V. 288. № 19. P. 13669–13674.
17. Blaschke K., Ebata K.T., Karimi M.M., Zepeda-Martinez J.A., Goyal P., Mahapatra S., Tam A., Laird D.J., Hirst M., Rao A., et al. // Nature. 2013. V. 500. № 7461. P. 222–226.
18. Chen J., Guo L., Zhang L., Wu H., Yang J., Liu H., Wang X., Hu X., Gu T., Zhou Z., et al. // Nat. Genet. 2013. V. 45. № 12. P. 1504–1509.
19. Yin R., Mao S.Q., Zhao B., Chong Z., Yang Y., Zhao C., Zhang D., Huang H., Gao J., Li Z., et al. // J. Am. Chem. Soc. 2013. V. 135. № 28. P. 10396–10403.
20. Doege C.A., Inoue K., Yamashita T., Rhee D.B., Travis S., Fujita R., Guarnieri P., Bhagat G., Vanti W.B., Shih A., et al. // Nature. 2012. V. 488. № 7413. P. 652–655.
21. Koh K.P., Yabuuchi A., Rao S., Huang Y., Cunniff K., Nardone J., Laiho A., Tahiliani M., Sommer C.A., Mostoslavsky G., et al. // Cell Stem Cell. 2011. V. 8. № 2. P. 200–213.
22. Justin N., De Marco V., Aasland R., Gamblin S.J. // Curr. Opin. Struct. Biol. 2010. V. 20. № 6. P. 730–738.
23. Metzger E., Wissmann M., Yin N., Muller J.M., Schneider R., Peters A.H., Gunther T., Buettner R., Schule R. // Nature. 2005. V. 437. № 7057. P. 436–439.
24. Shi Y., Lan F., Matson C., Mulligan P., Whetstine J.R., Cole P.A., Casero R.A., Shi Y. // Cell. 2004. V. 119. № 7. P. 941–953.
25. Klose R.J., Yamane K., Bae Y., Zhang D., Erdjument-Bromage H., Tempst P., Wong J., Zhang Y. // Nature. 2006. V. 442. № 7100. P. 312–316.
26. McDonough M.A., Loenarz C., Chowdhury R., Clifton I.J., Schofield C.J. // Curr. Opin. Struct. Biol. 2010. V. 20. № 6. P. 659–672.
27. Tsukada Y., Fang J., Erdjument-Bromage H., Warren M.E., Borchers C.H., Tempst P., Zhang Y. // Nature. 2006. V. 439. № 7078. P. 811–816.
28. Zhang T., Huang K., Zhu Y., Wang T., Shan Y., Long B., Li Y., Chen Q., Wang P., Zhao S., et al. // J. Biol. Chem. 2019. V. 294. № 37. P. 13657–13670.
29. Chen X., Yammine S., Shi C., Tark-Dame M., Gondor A., Ohlsson R. // Epigenetics. 2014. V. 9. № 11. P. 1439–1445.
30. Ebata K.T., Mesh K., Liu S., Bilenky M., Fekete A., Acker M.G., Hirst M., Garcia B.A., Ramalho-Santos M. // Epigenetics Chromatin. 2017. V. 10. P. 36.
31. Song M.H., Nair V.S., Oh K.I. // BMB Rep. 2017. V. 50. № 1. P. 49–54.
32. Tran K.A., Jackson S.A., Olufs Z.P., Zaidan N.Z., Leng N., Kendziorski C., Roy S., Sridharan R. // Nat. Commun. 2015. V. 6. P. 6188.
33. Chen J., Liu H., Liu J., Qi J., Wei B., Yang J., Liang H., Chen Y., Chen J., Wu Y., et al. // Nat. Genet. 2013. V. 45. № 1. P. 34–42.
34. Wang T., Chen K., Zeng X., Yang J., Wu Y., Shi X., Qin B., Zeng L., Esteban M.A., Pan G., et al. // Cell Stem Cell. 2011. V. 9. № 6. P. 575–587.
35. Shenoy N., Creagan E., Witzig T., Levine M. // Cancer Cell. 2018. V. 34. № 5. P. 700–706.
36. Padayatty S.J., Levine M. // Oral. Dis. 2016. V. 22. № 6. P. 463–493.
37. Chen Q., Espey M.G., Krishna M.C., Mitchell J.B., Corpe C.P., Buettner G.R., Shacter E., Levine M. // Proc. Natl. Acad. Sci. USA. 2005. V. 102. № 38. P. 13604–13609.
38. Du J., Cullen J.J., Buettner G.R. // Biochim. Biophys. Acta. 2012. V. 1826. № 2. P. 443–457.
39. Chen Q., Espey M.G., Sun A.Y., Lee J.H., Krishna M.C., Shacter E., Choyke P.L., Pooput C., Kirk K.L., Buettner G.R., et al. // Proc. Natl. Acad. Sci. USA. 2007. V. 104. № 21. P. 8749–8754.
40. Rawal M., Schroeder S.R., Wagner B.A., Cushing C.M., Welsh J.L., Button A.M., Du J., Sibenaller Z.A., Buettner G.R., Cullen J.J. // Cancer Res. 2013. V. 73. № 16. P. 5232–5241.
41. Yun J., Mullarky E., Lu C., Bosch K.N., Kavalier A., Rivera K., Roper J., Chio II, Giannopoulou E.G., Rago C., et al. // Science. 2015. V. 350. № 6266. P. 1391–1396.
42. Ngo B., van Riper J.M., Cantley L.C., Yun J. // Nat. Rev. Cancer. 2019. V. 19. № 5. P. 271–282.
43. Semenza G.L. // Biochim. Biophys. Acta. 2016. V. 1863. № 3. P. 382–391.
44. Campbell E.J., Vissers M.C., Bozonet S., Dyer A., Robinson B.A., Dachs G.U. // Cancer Med. 2015. V. 4. № 2. P. 303–314.
45. Koivunen P., Hirsila M., Gunzler V., Kivirikko K.I., Myllyharju J. // J. Biol. Chem. 2004. V. 279. № 11. P. 9899–9904.
46. Knowles H.J., Raval R.R., Harris A.L., Ratcliffe P.J. // Cancer Res. 2003. V. 63. № 8. P. 1764–1768.
47. Kuiper C., Dachs G.U., Currie M.J., Vissers M.C. // Free Radic. Biol. Med. 2014. V. 69. P. 308–317.
48. Vissers M.C., Gunningham S.P., Morrison M.J., Dachs G.U., Currie M.J. // Free Radic. Biol. Med. 2007. V. 42. № 6. P. 765–772.
49. Kuiper C., Dachs G.U., Munn D., Currie M.J., Robinson

B.A., Pearson J.F., Vissers M.C. // Front. Oncol. 2014. V. 4. P. 10.
50. Kuiper C., Molenaar I.G., Dachs G.U., Currie M.J., Sykes P.H., Vissers M.C. // Cancer Res. 2010. V. 70. № 14. P. 5749–5758.
51. Ehrlich M., Lacey M. // Adv. Exp. Med. Biol. 2013. V. 754. P. 31–56.
52. Cimmino L., Dolgalev I., Wang Y., Yoshimi A., Martin G.H., Wang J., Ng V., Xia B., Witkowski M.T., Mitchell-Flack M., et al. // Cell. 2017. V. 170. № 6. P. 1079–1095.
53. Ko M., An J., Rao A. // Curr. Opin. Cell Biol. 2015. V. 37. P. 91–101.
54. Odejide O., Weigert O., Lane A.A., Toscano D., Lunning M.A., Kopp N., Kim S., van Bodegom D., Bolla S., Schatz J.H., et al. // Blood. 2014. V. 123. № 9. P. 1293–1296.
55. Shenoy N., Bhagat T., Nieves E., Stenson M., Lawson J., Choudhary G.S., Habermann T., Nowakowski G., Singh R., Wu X., et al. // Blood Cancer J. 2017. V. 7. № 7. P. e587.
56. Zhao Z., Chen L., Dawlaty M.M., Pan F., Weeks O., Zhou Y., Cao Z., Shi H., Wang J., Lin L., et al. // Cell Rep. 2015. V. 13. № 8. P. 1692–1704.
57. Tommasini-Ghelfi S., Murnan K., Kouri F.M., Mahajan A.S., May J.L., Stegh A.H. // Sci. Adv. 2019. V. 5. № 5. P. eaaw4543.
58. Agathocleous M., Meacham C.E., Burgess R.J., Piskounova E., Zhao Z., Crane G.M., Cowin B.L., Bruner E., Murphy M.M., Chen W., et al. // Nature. 2017. V. 549. № 7673. P. 476–481.
59. Mingay M., Chaturvedi A., Bilenky M., Cao Q., Jackson L., Hui T., Moksa M., Heravi-Moussavi A., Humphries R.K., Heuser M., et al. // Leukemia. 2018. V. 32. № 1. P. 11–20.
60. Agus D.B., Gambhir S.S., Pardridge W.M., Spielholz C., Baselga J., Vera J.C., Golde D.W. // J. Clin. Invest. 1997. V. 100. № 11. P. 2842–2848.
61. Caprile T., Salazar K., Astuya A., Cisternas P., Silva-Alvarez C., Montecinos H., Millan C., de Los Angeles Garcia M., Nualart F. // J. Neurochem. 2009. V. 108. № 3. P. 563–577.
62. Baillie N., Carr A.C., Peng S. // Antioxidants (Basel). 2018. V. 7. № 9. P. 115.
63. Schoenfeld J.D., Sibenaller Z.A., Mapuskar K.A., Wagner B.A., Cramer-Morales K.L., Furqan M., Sandhu S., Carlisle T.L., Smith M.C., Abu Hejleh T., et al. // Cancer Cell. 2017. V. 31. № 4. P. 487–500.
64. Castro-Vega L.J., Buffet A., De Cubas A.A., Cascon A., Menara M., Khalifa E., Amar L., Azriel S., Bourdeau I., Chabre O., et al. // Hum. Mol. Genet. 2014. V. 23. № 9. P. 2440–2446.
65. Oermann E.K., Wu J., Guan K.L., Xiong Y. // Semin. Cell Dev. Biol. 2012. V. 23. № 4. P. 370–380.
66. Xiao M., Yang H., Xu W., Ma S., Lin H., Zhu H., Liu L., Liu Y., Yang C., Xu Y., et al. // Genes Dev. 2012. V. 26. № 12. P. 1326–1338.
67. Lawenda B.D., Kelly K.M., Ladas E.J., Sagar S.M., Vickers A., Blumberg J.B. // J. Natl Cancer Inst. 2008. V. 100. № 11. P. 773–783.
68. Espey M.G., Chen P., Chalmers B., Drisko J., Sun A.Y., Levine M., Chen Q. // Free Radic. Biol. Med. 2011. V. 50. № 11. P. 1610–1619.
69. Ma Y., Chapman J., Levine M., Polireddy K., Drisko J., Chen Q. // Sci. Transl. Med. 2014. V. 6. № 222. P. 222ra218.
70. Xia J., Xu H., Zhang X., Allamargot C., Coleman K.L., Nessler R., Frech I., Tricot G., Zhan F. // EBioMedicine. 2017. V. 18. P. 41–49.
71. Carr A.C., Cook J. // Front. Physiol. 2018. V. 9. P. 1182.
72. Chen M.F., Yang C.M., Su C.M., Hu M.L. // Nutr. Cancer. 2014. V. 66. № 7. P. 1085–1091.
73. Polireddy K., Dong R., Reed G., Yu J., Chen P., Williamson S., Violet P.C., Pessetto Z., Godwin A.K., Fan F., et al. // Sci. Rep. 2017. V. 7. № 1. P. 17188.
74. Zhao H., Zhu H., Huang J., Zhu Y., Hong M., Zhu H., Zhang J., Li S., Yang L., Lian Y., et al. // Leuk. Res. 2018. V. 66. P. 1–7.
75. Mustafi S., Camarena V., Volmar C.H., Huff T.C., Sant D.W., Brothers S.P., Liu Z.J., Wahlestedt C., Wang G. // Cancer Res. 2018. V. 78. № 2. P. 572–583.
76. Liu M., Ohtani H., Zhou W., Orskov A.D., Charlet J., Zhang Y.W., Shen H., Baylin S.B., Liang G., Gronbaek K., et al. // Proc. Natl. Acad. Sci. USA. 2016. V. 113. № 37. P. 10238–10244.
77. Hackanson B., Robbel C., Wijermans P., Lubbert M. // Ann. Hematol. 2005. V. 84. № Suppl 1. P. 32–38.
78. Mayland C.R., Bennett M.I., Allan K. // Palliat. Med. 2005. V. 19. № 1. P. 17–20.
79. Huijskens M.J., Wodzig W.K., Walczak M., Germeraad W.T., Bos G.M. // Results Immunol. 2016. V. 6. P. 8–10.
80. Marcus S.L., Petrylak D.P., Dutcher J.P., Paietta E., Ciobanu N., Strauman J., Wiernik P.H., Hutner S.H., Frank O., Baker H. // Am. J. Clin. Nutr. 1991. V. 54. № 6. P. 1292S–1297S.
81. Weijl N.I., Hopman G.D., Wipkink-Bakker A., Lentjes E.G., Berger H.M., Cleton F.J., Osanto S. // Ann. Oncol. 1998. V. 9. № 12. P. 1331–1337.
82. Inoue A., Zhang Y. // Science. 2011. V. 334. № 6053. P. 194.
83. Mayer W., Niveleau A., Walter J., Fundele R., Haaf T. // Nature. 2000. V. 403. № 6769. P. 501–502.
84. Peat J.R., Dean W., Clark S.J., Krueger F., Smallwood S.A., Ficz G., Kim J.K., Marioni J.C., Hore T.A., Reik W. // Cell Rep. 2014. V. 9. № 6. P. 1990–2000.
85. Wang L., Zhang J., Duan J., Gao X., Zhu W., Lu X., Yang L., Zhang J., Li G., Ci W., et al. // Cell. 2014. V. 157. № 4. P. 979–991.
86. Borgel J., Guibert S., Li Y., Chiba H., Schubeler D., Sasaki H., Forne T., Weber M. // Nat. Genet. 2010. V. 42. № 12. P. 1093–1100.
87. Hackett J.A., Sengupta R., Zylicz J.J., Murakami K., Lee C., Down T.A., Surani M.A. // Science. 2013. V. 339. № 6118. P. 448–452.
88. Hajkova P., Jeffries S.J., Lee C., Miller N., Jackson S.P., Surani M.A. // Science. 2010. V. 329. № 5987. P. 78–82.
89. Casanueva E., Ripoll C., Tolentino M., Morales R.M., Pfeffer F., Vilchis P., Vadillo-Ortega F. // Am. J. Clin. Nutr. 2005. V. 81. № 4. P. 859–863.
90. Kamikawa Y.F., Donohoe M.E. // PLoS One. 2015. V. 10. № 5. P. e0125626.
91. Li Q., Wang H.Y., Chepelev I., Zhu Q., Wei G., Zhao K., Wang R.F. // PLoS Genet. 2014. V. 10. № 7. P. e1004524.
92. Welstead G.G., Creyghton M.P., Bilodeau S., Cheng A.W., Markoulaki S., Young R.A., Jaenisch R. // Proc. Natl. Acad. Sci. USA. 2012. V. 109. № 32. P. 13004–13009.
93. Dominguez-Salas P., Moore S.E., Baker M.S., Bergen A.W., Cox S.E., Dyer R.A., Fulford A.J., Guan Y., Laritsky E., Silver M.J., et al. // Nat. Commun. 2014. V. 5. P. 3746.
94.Lambrot R., Xu C., Saint-Phar S., Chountalos G., Cohen T., Paquet M., Suderman M., Hallett M., Kimmins S. // Nat. Commun. 2013. V. 4. P. 2889.
95.DiTroia S.P., Percharde M., Guerquin M.J., Wall E., Collignon E., Ebata K.T., Mesh K., Mahesula S., Agathocleous M., Laird D.J., et al. // Nature. 2019. V. 573. № 7773. P. 271–275.
96.Takahashi K., Yamanaka S. // Cell. 2006. V. 126. № 4.

P. 663–676.
97.Zhao X.Y., Li W., Lv Z., Liu L., Tong M., Hai T., Hao J., Guo C.L., Ma Q.W., Wang L., et al. // Nature. 2009. V. 461. № 7260. P. 86–90.
98. Takahashi K., Tanabe K., Ohnuki M., Narita M., Ichisaka T., Tomoda K., Yamanaka S. // Cell. 2007. V. 131. № 5. P. 861–872.
99. Yu J., Vodyanik M.A., Smuga-Otto K., Antosiewicz-Bourget J., Frane J.L., Tian S., Nie J., Jonsdottir G.A., Ruotti V., Stewart R., et al. // Science. 2007. V. 318. № 5858. P. 1917–1920.
100. Eminli S., Foudi A., Stadtfeld M., Maherali N., Ahfeldt T., Mostoslavsky G., Hock H., Hochedlinger K. // Nat. Genet. 2009. V. 41. № 9. P. 968–976.
101. Li H., Collado M., Villasante A., Strati K., Ortega S., Canamero M., Blasco M.A., Serrano M. // Nature. 2009. V. 460. № 7259. P. 1136–1139.
102. Marion R.M., Strati K., Li H., Murga M., Blanco R., Ortega S., Fernandez-Capetillo O., Serrano M., Blasco M.A. // Nature. 2009. V. 460. № 7259. P. 1149–1153.
103. Mikkelsen T.S., Hanna J., Zhang X., Ku M., Wernig M., Schorderet P., Bernstein B.E., Jaenisch R., Lander E.S., Meissner A. // Nature. 2008. V. 454. № 7200. P. 49–55.
104. Banito A., Rashid S.T., Acosta J.C., Li S., Pereira C.F., Geti I., Pinho S., Silva J.C., Azuara V., Walsh M., et al. // Genes Dev. 2009. V. 23. № 18. P. 2134–2139.
105. Cloos P.A., Christensen J., Agger K., Helin K. // Genes Dev. 2008. V. 22. № 9. P. 1115–1140.
106. Mansour A.A., Gafni O., Weinberger L., Zviran A., Ayyash M., Rais Y., Krupalnik V., Zerbib M., Amann-Zalcenstein D., Maza I., et al. // Nature. 2012. V. 488. № 7411. P. 409–413.
107. Tzatsos A., Pfau R., Kampranis S.C., Tsichlis P.N. // Proc. Natl. Acad. Sci. USA. 2009. V. 106. № 8. P. 2641–2646.
108. Houbaviy H.B., Murray M.F., Sharp P.A. // Dev. Cell. 2003. V. 5. № 2. P. 351–358.
109. Suh M.R., Lee Y., Kim J.Y., Kim S.K., Moon S.H., Lee J.Y., Cha K.Y., Chung H.M., Yoon H.S., Moon S.Y., et al. // Dev. Biol. 2004. V. 270. № 2. P. 488–498.
110. Anokye-Danso F., Trivedi C.M., Juhr D., Gupta M., Cui Z., Tian Y., Zhang Y., Yang W., Gruber P.J., Epstein J.A., et al. // Cell Stem Cell. 2011. V. 8. № 4. P. 376–388.
111. Costa Y., Ding J., Theunissen T.W., Faiola F., Hore T.A., Shliaha P.V., Fidalgo M., Saunders A., Lawrence M., Dietmann S., et al. // Nature. 2013. V. 495. № 7441. P. 370–374.
112. Hu X., Zhang L., Mao S.Q., Li Z., Chen J., Zhang R.R., Wu H.P., Gao J., Guo F., Liu W., et al. // Cell Stem Cell. 2014. V. 14. № 4. P. 512–522.
113. Hore T.A., von Meyenn F., Ravichandran M., Bachman M., Ficz G., Oxley D., Santos F., Balasubramanian S., Jurkowski T.P., Reik W. // Proc. Natl. Acad. Sci. USA. 2016. V. 113. № 43. P. 12202–12207.
114. Schwarz B.A., Bar-Nur O., Silva J.C., Hochedlinger K. // Curr. Biol. 2014. V. 24. № 3. P. 347–350.
115. Lee Chong T., Ahearn E.L., Cimmino L. // Front. Cell Dev. Biol. 2019. V. 7. P. 128.

The Functions and Mechanisms of Action of Insulators in the Genomes of Higher Eukaryotes

L. S. Melnikova*, P. G. Georgiev, A. K. Golovnin
Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia
*E-mail: lsm73@mail.ru
Received August 07, 2020; in final form, October 12, 2020
DOI: 10.32607/actanaturae.11144
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **The mechanisms underlying long-range interactions between chromatin regions and the principles of chromosomal architecture formation are currently under extensive scrutiny. A special class of regulatory elements known as insulators is believed to be involved in the regulation of specific long-range interactions between enhancers and promoters. This review focuses on the insulators of *Drosophila* and mammals, and it also briefly characterizes the proteins responsible for their functional activity. It was initially believed that the main properties of insulators are blocking of enhancers and the formation of independent transcription domains. We present experimental data proving that the chromatin loops formed by insulators play only an auxiliary role in enhancer blocking. The review also discusses the mechanisms involved in the formation of topologically associating domains and their role in the formation of the chromosomal architecture and regulation of gene transcription.**

KEYWORDS **insulator proteins, enhancer-promoter communication, chromatin loops, regulation of transcription, Su(Hw), TAD.**

ABBREVIATIONS **a.a. – amino acid; bp – base pair; kbp – kilobase pair; PRE – polycomb response element; LCR – locus control region; ANT-C – *Antennapedia* complex; TF – transcription factor; ZF – zinc finger; BX-C – *Bithorax* complex; CNS – central nervous system; BTB – Broad-complex, Tramtrack, and Bric-à-brac; POZ – poxvirus and zinc finger; ZAD – zink finger-associated domain; ICR – imprinting control region; PRC2 – polycomb repressive complex 2; TAD – topologically associating domain.**

INTRODUCTION

In higher eukaryotic cells, transcription, one of the key stages of gene expression, results from the interaction between promoters that determine transcription initiation and its basic level and the various *cis*-regulatory elements that either amplify (enhancers) or weaken (silencers) the transcription [1–3]. Enhancers and silencers may reside at a considerable distance from the genes whose transcription they regulate and be separated from them by numerous "alien" genes with their own regulation systems [4, 5]. In order to explain the mechanism of specific interactions between an enhancer/silencer and a promoter, a model has been proposed postulating that chromosomes are subdivided into transcription (chromatin) domains that strictly limit contacts between regulatory genome sequences [6].

A new class of regulatory elements called insulators was found for the first time in studies conducted using the fruit fly, *Drosophila melanogaster* [7–9]. Initially, two of the properties of insulators were described. First, insulators residing between the enhancer and the promoter prevent their interaction (an enhancer-blocking activity). Second, insulators surrounding the transgene neutralize the negative or positive effect of the adjacent chromatin on its expression (a barrier activity). Insulators have been detected in the genomes of all well-studied higher eukaryotes [10, 11]. It was initially assumed that insulators that interact with each other are responsible for the formation of isolated transcription domains. However, further research has demonstrated that insulators are multifunctional elements comprised by the regulation systems of many genes [12–18].

INSULATORS IN THE GENOMES OF HIGHER EUKARYOTES

The fruit fly *Drosophila melanogaster* was often used as a model organism in the first studies focused on insulators. By then, a system based on *P*-transposon enabling efficient transgenic modification of the fruit fly genome had already been developed [19]. It was not until

much later that the methods for *in vivo* genome modification of vertebrate animals were developed [20, 21]. *P*-dependent integration has a stochastic nature, allowing one to study the effect of different chromosomal environments on transgenic expression. The *white* gene responsible for eye pigmentation in *Drosophila melanogaster* was often used as a reporter gene [22]. In different transgenic lines carrying the *white* gene without enhancers (*mini-white*), the eye color in flies ranged from pale yellow to red, being caused by transgene integration sites. This phenomenon is known as the chromosomal position effect [22, 23]. It was assumed that expression of the *mini-white* gene depends on the chromosomal position due to the activity of genome enhancers residing near the transgene integration site. However, it was proved later that in more than 70% of cases, the *mini-white* transcription initiated in the surrounding genome regions is responsible for the activating effect of the chromosomal environment [24].

The first insulators described in the *Drosophila melanogaster* genome were the *scs* and *scs*' (specialized chromatin structure) sequences found at the cytogenetic locus 87A7 as nuclease-hypersensitive DNA regions surrounding a cluster of five genes, including two genes coding for heat shock proteins 70 (*hsp70*) [8, 9, 25]. Activation of the *hsp70* genes induces decondensation of chromomer 87A7 to form a "puff" in salivary gland polytene chromosomes. Cytological studies showed that the *scs* and *scs*' elements reside at sites where the decondensed 87A7 locus is flanked by condensed chromatin. However, it was revealed later that *scs* and *scs*' are located inside the puff rather than at its boundaries and do not restrict the 87A7 decondensation [26]. It was suggested that *scs* and *scs*' are the boundaries of the transcription domain that includes the *hsp70* genes. The *scs* and *scs*' elements within transgenes exhibited enhancer-blocking and barrier insulator properties [8, 9]. Next, it was shown that the *scs* (993 bp) and *scs*' (500 bp) insulators have a complex structure that includes the gene promoters and transcription termination signals [27–30].

The best studied insulator of *Drosophila melanogaster* was found in the regulatory region of the *gypsy* retrotransposon (Mdg4) [31]. The *gypsy* retrotransposon affects the expression of the neighboring genes by causing mutant phenotypes. The effect of *gypsy* on transcription is due to a 460-bp sequence located in its 5'-transcribed untranslated region [7, 32]. In transgenic lines, the *gypsy* insulator blocks the activity of various enhancers at all stages of Drosophila development [33–36]. The insulator was found to consist of 12 degenerated octameric sites of Su(Hw) protein binding [32, 37, 38]. The properties of the *gypsy* insulator were initially tested using the regulatory system of the *yellow* locus responsible for the pigmentation of cuticle structures in embryos, larvae, and the imago of fruit flies [39]. Enhancers controlling the transcription of *yellow* in the wing plates and body cuticle reside in the 5' gene region, while the enhancers controlling expression in bristles reside in the intron [7]. In the y^2 allele, the *gypsy* retrotransposon is integrated in the 5' region of the *yellow* gene, between the promoter and enhancers activating transcription in the wings and body. As a result, the insulator blocks the body and wing enhancers but does not affect the activity of the bristle enhancer residing in the gene intron (*Fig. 1*). A mutation inactivating the *su(Hw)* gene makes the insulator in the y^2 allele disappear, thus completely restoring *yellow* gene expression [40]. Several studies have shown that when transgene is integrated into the heterochromatin regions of the genome or in the vicinity of the Polycomb response element (PRE)-dependent silencer, the *gypsy* insulator efficiently protects the *white* reporter gene against repression [41, 42].

Another insulator was found in the long terminal repeat of the *Idefix* retrotransposon [43]. The barrier activity of the *Idefix* insulator and its ability to block

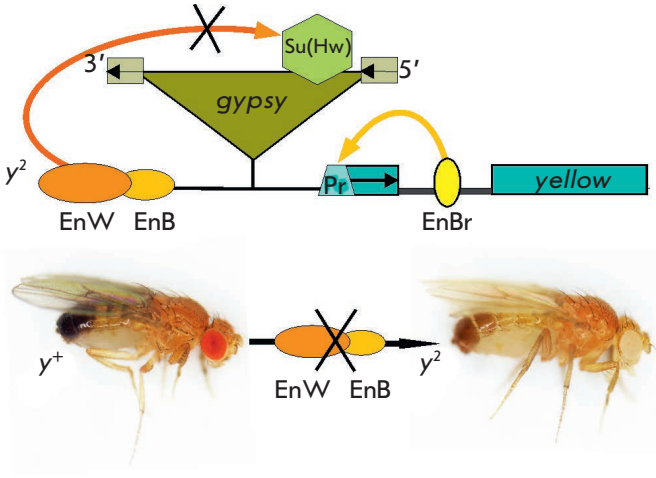


Fig. 1. Schematic representation of the y2 allele. Exons of the *yellow* gene are shown with rectangles, with an arrow indicating the direction of transcription; EnW – wing enhancer; EnB – body enhancer; EnBr – bristle enhancer; and Pr – promoter of the gene. The *gypsy* retrotransposon is depicted as a triangle; the rectangles at its ends are long terminal repeats, with their direction shown with arrows. The Su(Hw) insulator is depicted as a hexagon inside *gypsy*. The photographs show the phenotypes of flies: y^+ – wild type, the *yellow* gene is expressed in all cuticular structures; y^2 – body and wing enhancers are blocked by the Su(Hw) insulator (depicted as strikethrough); the *yellow* gene is not expressed in the body cuticle and wings but continues to be expressed in the bristles

various enhancers were detected using transgenic lines [44].

The first functional genomic insulator, 1A2, containing two Su(Hw) protein binding sites was found in the 3' region of the *yellow* gene [45, 46]. It turned out that many genome sequences, including 1–3 binding sites for Su(Hw), act as insulators in transgenes [47–49]. However, it was found using synthetic repetitive Su(Hw)-binding sites that at least four sites provide efficient insulator activity [50]. This contradiction can be attributed to the existence of proteins that have not been identified yet, which are involved in the formation of functional endogenous insulators, along with Su(Hw) [51].

The genome of *Drosophila melanogaster* was found to contain many insulator sequences not carrying binding sites for the Su(Hw) protein. These include the SF1 and SF2 insulators from the *Antennapedia* complex (ANT-C) [52, 53]; *facet-strawberry* sequences protecting the *Notch* gene against the effects of the surrounding chromatin [54]; the *Wari* insulator [55] residing at the 3' end of the *white* gene; and the ME boundary element inhibiting the activity of the enhancer from the *eyeless* gene with respect to the promoter of the neighboring *myoglianin* gene [56]. The boundaries of independent transcription domains, *Mcp*, *Fab-6*, *Fab-7*, and *Fab-8*, demonstrating properties of the insulators in transgenic lines have been revealed in the regulatory region of the *Bithorax* complex (BX-C) [57–71].

The first vertebrate insulators were found at the boundaries of clusters of transcriptionally active genes and heterochromatin regions. The HS4 insulator was detected at the 5' end of the chicken β-globin locus [72]. The core sequence of HS4 contains the CTCF-binding site [73]. Subsequently, searching for new vertebrate insulators was often based on testing DNA fragments containing CTCF-binding sites [74, 75]. Thus, an insulator containing four CTCF-binding sites and playing a crucial role in the imprinted expression of the *Igf2/H19* locus was found in mice and humans [76–78]. Many CTCF-dependent vertebrate insulators have been described, being consistent with the views on the key role played by the CTCF protein in the organization of chromatin architecture [74, 75].

THE MODELS OF THE MECHANISM OF ACTION OF INSULATORS

The data on the properties of insulators were used to propose two groups of alternative models for explaining their mechanism of action.

According to the transcription models, an insulator actively interrupts the specific long-range enhancer-promoter interactions [73, 79, 80]. Depending on the possible mechanism of enhancer-promotor interactions, different variants of insulator action were considered. According to one model, the enhancer "looks for" a promoter by moving along the chromatin fibril. In this case, the insulator acts as a physical barrier preventing enhancer motion. It was also supposed that insulators are pseudo-promoters. They do not initiate transcription but can interact with enhancers, thus inhibiting their activity (*Fig. 2A*). According to another popular model, long-range enhancer-promotor contacts are ensured by special facilitating proteins. For example, the mammalian homodimerizing protein LDB1 forms specific contacts between the enhancers and promoters of many genes [81]. The *Drosophila melanogaster* CHIP protein facilitates enhancer-promotor interactions

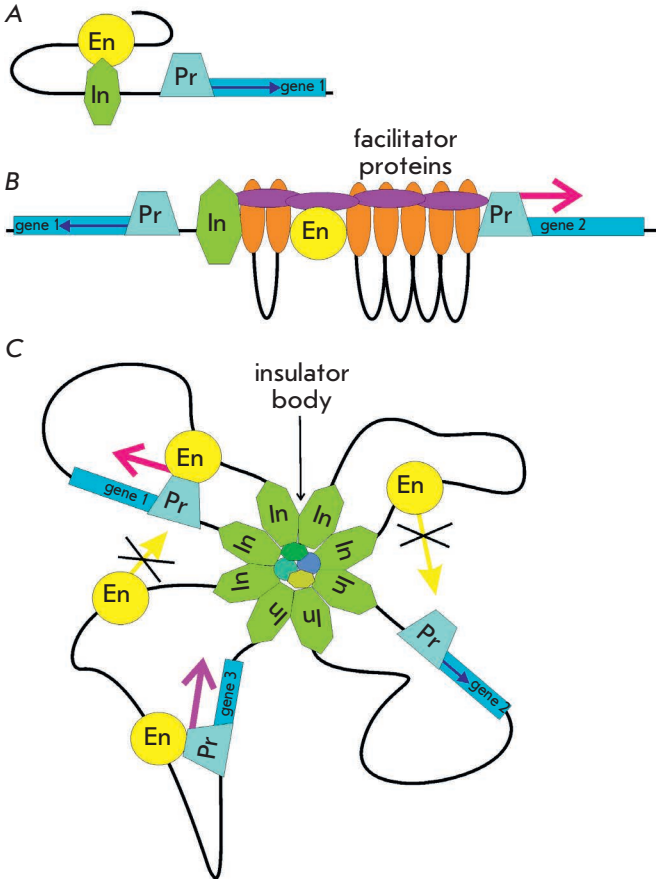


Fig. 2. The models of insulator function. (*A*) Model of an "enhancer decoy." (*B*) Blocking of facilitating proteins. (*C*) Structural model. Formation of independent transcriptional domains. Designations: En – enhancer; In – insulator; and Pr – promoter. Red arrows indicate transcription activation by a specific enhancer; blue arrows show the basic activity of the promoter. Strikethrough arrows indicate blocking interactions between enhancers and promoters from the adjacent domains

in the *cut* locus [82]. The CHIP protein was shown to interact with the components of the *gypsy* insulator [83, 84]. When the enhancer-promotor interaction is weakened by a CHIP mutation, Su(Hw)-dependent insulation becomes more efficient. Hence, the insulator can inhibit the activity of the facilitating proteins that ensure the enhancer-promotor communication (*Fig. 2B*).

The structural models of the action of insulators have gained wide popularity [85]. Initially, these models were based on the idea that chromosomes form large independent chromatin loops [6]. It was assumed that chromatin loops are independent transcription domains and block any interactions between the regulatory elements in neighboring domains.

Later, studies focused on localizing the Su(Hw) protein on chromosomes and in the nucleus substantially gained in importance. It was believed that polytene chromosome bands correspond to transcription domains, while interbands correspond to their boundaries. It was shown that the binding sites of Su(Hw) reside in some interbands (i.e., limit the transcription domains) [86]. In Drosophila cultured cells, embryos, and imaginal discs, the Su(Hw) protein was found within compact nuclear structures known as insulator bodies [86]. It was assumed that each insulator body consists of multiple individual insulators that interact with each other and divide the chromatin fibril into domain loops, thus forming rosette-like structures (*Fig. 2C*). The insulators lying in the base of the rosette can interact with the nuclear lamina (shell) or with components of the nuclear pore, thus laying the basis for the spatial organization of chromatin. The structural models postulate that the key role of insulators is to form chromatin loops, while their activity is believed to result from this organization. Chromatin looping may either topologically or physically impede interaction between enhancers and promoters located in neighboring domains [87].

Today, the structural models rely on data on the organization of higher eukaryotic chromosomes into topologically associating domains (TADs) [88–91]. A hypothesis has been put forward that insulators are the TAD boundaries. The interaction between insulators gives rise to chromatin loops limiting the enhancer activity.

Su(Hw)-DEPENDENT COMPLEX AS A MODEL FOR STUDYING INSULATORS

Insulator activity is ensured by a complex of interacting proteins that bind to the insulator DNA sequence. In many studies, the mechanisms of insulator functioning and formation in *Drosophila melanogaster* were investigated for the Su(Hw)-dependent complex.

The key protein of the complex, Su(Hw), is expressed during the entire development process and is found in most *Drosophila melanogaster* tissues. Inactivation of the *su(Hw)* gene results in female sterility [35, 92]. The Su(Hw) protein consists of the N-terminal region rich in acidic amino acids, a DNA-binding domain containing twelve C2H2-type zinc fingers (ZFs), and the C-terminal region, which is also rich in acidic amino acid residues [92]. Su(Hw) binds to a consensus sequence (~ 26 bp) consisting of three modules [93]. Cluster ZF6-9 binds to the main (central) module; cluster ZF2-4, to the CG-rich module ("down"); and cluster ZF10-12, to the AT-rich module ("up") (*Fig. 3*). The tenth ZF affects the efficiency of protein binding to some sites [93, 94]. For example, a mutation in ZF10 makes it impossible for the Su(Hw) protein to efficiently bind to the *gypsy* insulator sequence [51]. The C-terminal part of Su(Hw) carries the domain (716–892 a.a.) that is responsible for insulator activity [32, 92, 95] and the ability of the Su(Hw) protein to inhibit transcription of the central nervous system (CNS) genes in the

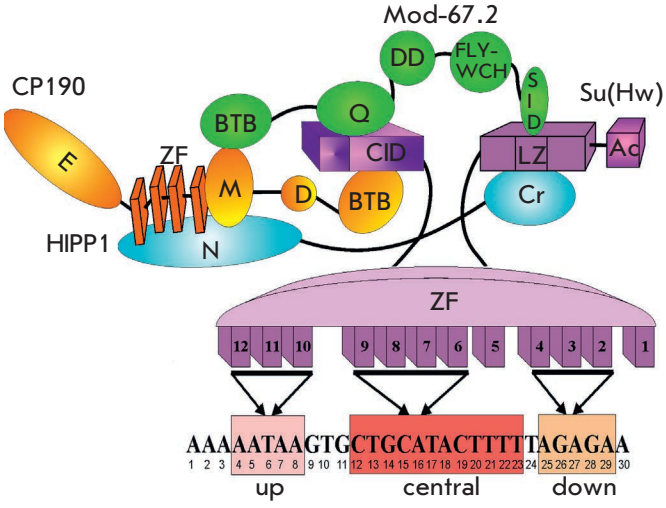


Fig. 3. A model of Su(Hw)-dependent insulator complex formation. The domains of the Su(Hw) protein are shown in lilac; Mod(mdg4)-67.2 protein domains are shown in green; CP190 protein domains, in orange; and HIPP1 protein domains, in blue. Domain abbreviations: CID – CP190 interacting domain; Ac – C-terminal acidic domain; ZF – zinc finger domain; LZ – leucine zipper; BTB – BTB/POZ domain; Q – glutamine-rich region; DD – dimerization domain; FLYWCH – FLYWCH type zinc finger; SID – Su(Hw) interacting domain; D – asparagine- rich domain; M – the microtubule and centrosome associated domain; E – glutamine-rich C-terminal domain. Below is the consensus binding sequence for the Su(Hw) protein from the *gypsy* insulator. The ZFs binding each motif are shown with arrows

ovaries [96–98]. Two more proteins, Mod(mdg4)-67.2 and CP190, are recruited to the complex via direct interaction with Su(Hw) (*Fig. 3*).

The Mod(mdg4)-67.2 protein is produced by a complex locus, *mod(mdg4)* [99, 100]. At the N-end of the Mod(mdg4)-67.2 protein, there is the BTB/POZ domain (bric-à-brac, tramtrack and broad complex/poxvirus and zinc finger), which widely occurs in higher eukaryotes and is usually homodimerized. However, the BTB domain of Mod(mdg4)-67.2 belongs to a special insect-specific group [101]. The BTB domains belonging to this group can form both homo- and heteromultimeric complexes [102]. The C-end of the Mod(mdg4)-67.2 protein carries a specific domain interacting with the C-terminal region of Su(Hw) (716–892 a.a.) [83, 103]. Furthermore, the N-terminal part of the Su(Hw) protein interacts with the glutamine-rich region of the Mod(mdg4)-67.2 protein [104] (*Fig. 3*). The Mod(mdg4)-67.2 protein is involved in the enhancer-blocking activity of the Su(Hw) insulator.

The CP190 protein simultaneously interacts with Su(Hw) and Mod(mdg4)-67.2, thus stabilizing the formation of the insulator complex. The N-end of CP190 carries the BTB domain that forms stable homodimers [102, 105–107]. The C-end of CP190 carries glutamine- and asparagine-rich domains; between them, there reside the M domain responsible for interaction with microtubules and four ZFs [108]. The CP190 BTB domain interacts with two unstructured N-terminal regions of the Su(Hw) protein located between 88 and 202 a.a. [109]. The M domain of protein CP190 simultaneously interacts with the BTB domain of the Mod(mdg4)-67.2 protein [104, 110] (*Fig. 3*).

Deletions of separate domains in the Su(Hw), Mod(mdg4)-67.2, and CP190 proteins do not affect the *in vivo* assembly of the functional complex. Therefore, the Su(Hw) insulator forms through numerous interactions between its protein components, which partially compensate for and stabilize each other. Genome-wide studies have demonstrated that the complex containing all three proteins, CP190/Mod(mdg4)-67.2/Su(Hw), is assembled only at some Su(Hw)-binding sites [48, 94, 111]. The binding of the insulator complex to these sites is largely mediated by the CP190 and Mod(mdg4)-67.2 proteins [104, 109].

A new partner of Su(Hw), the HIPP1 protein (HP1 and insulator partner protein 1), has recently been identified [112]. Highly structured regions (1–212 and 675–778 a.a., respectively) reside at the ends of the HIPP1 protein; the C-terminal region corresponds to the crotonase domain [113, 114]. The crotonase domain of HIPP1 binds to the C-terminal region of Su(Hw) (637–892 a.a.), which is simultaneously responsible for the enhancer-blocking and repressive activities of the insulator. The N-terminal domain of HIPP1 interacts with the domains M and ZF of the CP190 protein [115] (*Fig. 3*). Inactivation of the *Hipp1* gene was shown to affect neither the fertility of flies nor the Su(Hw)-dependent insulator activity [115, 116]. However, the simultaneous inactivation of the *Hipp1* and *mod(mdg4)-67.2* genes significantly changes the activity of the *gypsy* insulator and substantially weakens CP190 binding to Su(Hw)-dependent sites [115]. Therefore, the processes of HIPP1 and CP190 recruitment to the Su(Hw) insulator are mutually dependent.

It was also found that the ENY2 protein directly interacts with ZF10–12 of the Su(Hw) protein [117]. It was demonstrated for transgenic lines that the ENY2 protein is involved in the barrier activity of the Su(Hw) insulator and protects reporter gene expression against the PRE-dependent repression. Interestingly, ENY2 also binds to ZF of the dCTCF protein (CTCF ortholog in *Drosophila melanogaster*) and is involved in the barrier function of dCTCF-dependent insulators [118]. Recruitment of an unknown ENY2-dependent complex to the ZFs of various transcription factors (TFs) can potentially be regarded as the general mechanism of gene protection against PRE-dependent repression.

The RNA-binding proteins Shep and Rump, which act as negative regulators of enhancer-blocking activity, may be involved in the function of the Su(Hw)-dependent complex [119, 120]. Moreover, the activity of the Su(Hw) insulator can be regulated by the components of the RNA interference system: Ago, aub, piwi, and Rm62 [121]. However, the mechanism by which these proteins function has not been elucidated yet.

Within the nucleus, the Su(Hw), CP190, and Mod(mdg4)-67.2 proteins reside in insulator bodies [122, 123]. Post-translational modification of the CP190 and Mod(mdg4)-67.2 proteins with a small ubiquitin-like modifier (SUMO) is needed to incorporate Su(Hw)-dependent proteins into the insulator bodies [122–124]. The dCTCF protein was also revealed within insulator bodies [125]. It was shown using *in vivo* model systems that formation of insulator bodies is unrelated to insulator activity [122], while sumoylation is not a necessary condition for the manifestation of enhancer-blocking activity [123]. It can be assumed that insulator bodies act as certain "depots" for chromatin proteins. Protein complexes, which efficiently bind to DNA synthesized during replication, are pre-assembled in these depots (*Fig. 4*).

The formation of insulator bodies is regulated by the amount of matrix protein EAST [124]. Under physiological conditions, the EAST protein does not bind to chromatin [126] but interacts with the CP190 and Mod(mdg4)-67.2 proteins [124]. The EAST expression level affects binding of the Su(Hw)-dependent complex

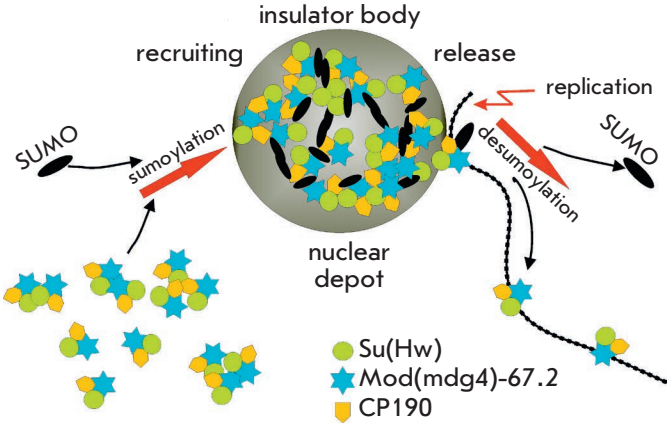


Fig. 4. The model of formation and functioning of insulator bodies. Proteins CP190/Su(Hw)/Mod(mdg4)-67.2 are recruited into insulator bodies by sumoylation. In insulator bodies, Su(Hw)-dependent complexes are pre-assembled and associated with other TFs. The "matured" insulator complex transiently interacts with chromatin fibril, leaves the insulator bodies due to desumoylation, and binds to specific chromatin sites

to chromatin and the activity of Su(Hw)-dependent insulators [124, 127]. These effects of EAST can be interpreted using the model described above, according to which the insulator complexes are pre-assembled in insulator bodies.

BRIEF CHARACTERIZATION OF INSULATOR PROTEINS

Most insulator complexes form around one or several key DNA-binding proteins. There are no clearly defined parameters according to which a protein can refer to insulator proteins. Therefore, any protein found within one or several insulators is automatically classified as belonging to the group of insulator proteins. *D. melanogaster* is known to have 11 proteins exhibiting enhancer-blocking properties that contain DNA-binding domains. Many of them (dCTCF, Su(Hw), Pita, ZIPIC, and GAF) contain C2H2-ZFs [128–130]. So far, only one conserved insulator protein, CTCF, has been described in vertebrates [131].

The CTCF protein is expressed in most mammalian tissues [132]. It is required during the early stages of mouse development and is involved in the cell cycle, apoptosis, and cell differentiation [133–135]. A CTCF ortholog having a similar domain structure (dCTCF) was found in Drosophila [136]. The dCTCF protein binds to most boundaries in the BX-C and is responsible for their insular activity. The central part of CTCF in vertebrates and Drosophila contains a cluster carrying 11 ZFs. The studies focused on the human CTCF–DNA complex have shown that ZFs 3–7 bind to the 15-bp consensus motif [137]. It was demonstrated using mutations in individual ZFs that in primary murine lymphocytes, ZFs 9–11 and ZFs 1–2 bind to the sequences flanking the consensus motif, thus stabilizing specific CTCF binding [138]. An unstructured domain forming homodimers resides at the N-end of CTCF in various organisms [139]. A motif interacting with the cohesin complex was also found at the N-end of human CTCF [140]. CTCF interacts with the cohesin complex to form chromatin loops and most of the TAD boundaries; it also mediates short-range interactions between the regulatory elements [90, 132, 141].

The ZIPIC, Pita, and Zw5 proteins carry the zinc finger-associated domain (ZAD) at their N-end and ZF clusters at their C-end [27, 68, 142, 143]. These proteins are intensively expressed at all stages of Drosophila development, especially during the embryonic stage. Mutations inactivating the *pita* and *zw5* genes cause early embryonic death, thus indicating that the *Pita* and *Zw5* proteins play an important role in gene expression regulation [27, 144]. The Zw5 protein was first detected on the *CG31211* gene promoter, a part of the *scs* insulator [27]. An analysis of whole-genome distribution of the ZIPIC, Pita, and Zw5 proteins showed that they preferentially bind to gene promoters near transcription start sites and, like the CTCF protein, are often colocalized with components of the cohesin and condensin complexes [48, 145]. Thanks to the ZAD domains, the ZIPIC, Pita and Zw5 proteins can form homodimers [145]. In transgenic lines, the multiple binding sites of these proteins form insulators inhibiting the enhancer activity and PRE-dependent repression [146].

The GAF protein is involved in the functioning of the Fab-7 insulator from the BX-C [70], SF1 insulator from the ANT-C [52], and the insulator located between the *myoglianin* and *eyeless* genes [56]. A single ZF binding to the GAGAG motif resides in the central part of the protein [147, 148]. Similar to the Mod(mdg4)-67.2 protein, the N-end of GAF carries an insect-specific BTB domain that forms homo- and heteromultimers [101, 102]. The BTB domains GAF and Mod(mdg4)-67.2 can interact with proteins from different transcription complexes [102, 149–151].

The BEAF-32 protein was initially identified as a factor interacting with the *scs*' insulator [30, 152]. To bind to DNA, BEAF-32 uses the N-terminal C2H2-like domain called BED. There is a BESS domain at the C-end of the protein, which is required for BEAF trimerization [152, 153]. Each subunit of the BEAF complex binds one CGATA motif, while BEAF trimers bind to clusters of the CGATA motif with high affinity [152]. Whole-genome analysis shows that BEAF is predominantly associated with the promoter regions of active genes and is involved in transcription stimulation [154, 155].

Identically to BEAF-32, the Ibf1 and Ibf2 proteins (insulator binding factors 1 and 2) bind to DNA through the BED domain to form hetero-oligomers [156]. A whole-genome analysis showed that Ibf1/Ibf2 is often simultaneously present with other insulator proteins, primarily with CP190 and dCTCF.

Elba1 and Elba2, the components of the recently discovered Elba (Early boundary activity) insulator complex, use conserved C-terminal BEN domains to bind to DNA [57]. The third protein, Elba3, is responsible for the formation of the Elba1/Elba2 dimer, which interacts with specific insulator sites. The Elba2 protein is expressed at most developmental stages, but two other components of the complex are present only during the early stage of embryonic development. Elba recognizes the 8-bp asymmetric CCAATAAG sequence, which is a part of the Fab-7 insulator from the BX-C. Another protein, Insv (Insensitive), binds to the Fab-7 insulator [157, 158]. Similar to the Elba protein, this protein carries the C-terminal BEN domain and is preferentially expressed in early embryos [158]. The Elba complex and Insv protein are needed to ensure *in vivo* functioning of the Fab-7 insulator [57, 157].

All the afore-listed insulator proteins found in Drosophila (except for Zw5 and the Elba complex) interact with the CP190 protein [68, 105, 108, 125, 156, 158–162]. DNA-binding insulator proteins recruit CP190 to chromatin [68, 105, 108, 161]. Meanwhile, the CP190 protein binds to most housekeeping gene promoters [108, 159, 161] and is involved in open chromatin formation [163]. The presence of the CP190 protein on insulators and promoters indicates that a functional relationship between them is possible.

DIRECT PARTICIPATION OF INSULATORS IN ENHANCER-PROMOTOR INTERACTIONS

Most binding sites of insulator proteins were detected in the promoter regions of different genes [47, 48]. It is known that generation of active promoters is one of the key functions of mammalian CTCF protein [164]. The involvement of the same proteins in the formation of promoter and insulator complexes agrees with the transcription models of insulator action.

In transgenic Drosophila lines, the *gypsy* insulator completely blocks the *yellow* gene enhancers, which are isolated by it from the promoter, while having no effect on basic promoter activity [7]. However, if the *yellow* gene promoter is weakened by a mutation, the *gypsy* and 1A2 insulators restore its activity regardless of their positions in the transgene [165]. Like active promoters, Su(Hw)-dependent insulators recruit the SAGA and Brahma complexes formed on the regulatory elements of the open chromatin domain [166]. Su(Hw) insulators potentially compensate for the partial inactivation of the *yellow* promoter by recruiting remodulating complexes to it. Therefore, the insulator-bound complexes are supposed to reside in close proximity to the promoter. Indeed, it has been shown that in transgenic lines, insulators facilitate long-range interactions between the promoters and GAL4 activators residing at the 3'-end of the reporter genes [165, 167]. ChIP and 3C assays revealed an interaction between an enhancer located upstream of the *white* gene promoter and the *gypsy* insulator at the 3'-end of the gene [168]. Short-range interactions between regulatory elements are probably ensured by the proteins binding simultaneously to insulators and promoters [47, 48, 160, 169]. It was shown that the CP190, Chromator, and BEAF-32 proteins can ensure long-range interactions between chromatin domains [107]. It is fair to assume that the main function of the endogenous insulators residing at the 3'-end of the *yellow* and *white* genes [45, 46, 55] is to enhance the activity of the promoters of these genes.

All other insulators exhibit a much weaker blocking activity against *yellow* gene enhancers compared to the *gypsy* insulator [55, 64, 68, 170]. On the other hand, the *gypsy* insulator integrated into the transgenes between the enhancer and the *white* gene promoter only slightly weakens the *white* gene expression in fruit fly eyes [168]. Interestingly, the C-terminal domain of the Su(Hw) protein is simultaneously responsible for the blocking of the *yellow* gene enhancers and repression of the promoters of the CNS genes in female gonads [171]. The Su(Hw) binding sites are located directly in the promoters of the CNS genes [98]. It is most likely that repression occurs due to the recruitment of a repressor complex specific to the germinal tissue, since no repression is observed in the eyes [28].

In the absence of the Mod(mdg4)-67.2 protein, the *gypsy* insulator becomes a repressor of the *yellow* gene promoter [83, 95, 110]. It is noteworthy that the Mod(mdg4)-67.2 protein is recruited to the insulator complex through the C-terminal domain of Su(Hw) being responsible for insulation/repression. Repression in the *yellow* locus can be attributed to the fact that the efficiency of binding between the repressor complex and the C-terminal domain of Su(Hw) increases in the absence of the Mod(mdg4)-67.2 protein. It was shown that *gypsy*-dependent repression is mediated by the promoter sequence of the *yellow* gene, same as the sequence required for ensuring long-range enhancer-promoter interactions [172]. The Su(Hw)-dependent repressor complex potentially interacts with the promoter TF, thus ensuring communication with enhancers.

The reported experimental data confirm the model according to which insulators dynamically interact

with enhancers and promoters. When an insulator is integrated between an enhancer and a promoter, the interaction between the insulator complex and TF of the promoter or the enhancer prevents efficient interaction between them. Thus, it was shown that the Mod(mdg4)-67.2 protein interacts with the Zeste protein. The Zeste protein binds to the *white* gene enhancer and promoter, thus providing communication between them [173, 174]. The interaction between the Mod(mdg4)-67 and Zeste proteins may interfere with the proper formation of enhancer-promoter contacts and reduce transcription. If the insulator recruits repressor complexes to the promoter region, enhancer activity is completely blocked.

Vertebrate CTCF protein often forms chromatin loops by interacting with active promoters [175, 176]. CTCF directly interacts with TAF3 and TFII-I, the components of the TFIID promoter complex [177, 178]. Therefore, CTCF-promoter interactions can prevent the formation of enhancer-promoter contacts. In the mammalian *Igf2/H19* locus, the genes are located so as to ensure that the *H19* gene in the maternal allele and the *Igf2* gene in the paternal allele are activated by common distal enhancers [75]. The *H19* gene is activated in the maternal allele, and the *Igf2* gene is activated in the paternal allele. The interaction between the common enhancers and gene promoters is regulated by a CTCF-dependent insulator residing in the imprinting control region (ICR). A 3C assay showed that in the maternal allele, the CTCF protein ensures direct interaction between the insulator and *Igf2* promoter, which inhibits the activation of *Igf2* by distal enhancers [179–181]. Interestingly, the CTCF protein recruits the polycomb repressive complex 2 (PRC2) repressing transcription to the *Igf2* promoter [181].

THE ROLE OF CHROMATIN LOOPS IN ENHANCER BLOCKING

The structural models of insulator action postulate that chromatin loops and TADs block interactions between the regulatory elements from adjacent domains [85, 182, 183]. However, the ability of chromatin loops to completely block the enhancer-promotor interactions has not been verified experimentally.

The functional role of the chromatin loops formed by insulators was thoroughly studied in transgenic Drosophila lines. It was found that two identical insulators integrated between the enhancer and the promoter mutually neutralize each other's activities [55, 170, 184–186]. To interpret this phenomenon, it was suggested that the same insulators interact with each other more efficiently than with an enhancer or a promoter. Therefore, they do not interfere with the enhancer-promoter interactions and even facilitate the long-range communication between the regulatory elements. This model was confirmed by the experiments where another gene surrounded by insulators was located between the enhancer and the promoter of the reporter gene [59, 186–188]. Efficient enhancer-dependent activation of the reporter gene was observed only in the presence of insulators. Therefore, the chromatin loop formed by a pair of identical insulators brought the enhancer and the promoter closer together (*Fig. 5A*). Similar results were obtained using the lines in which the enhancer was replaced with the transcription-repressing PRE [189]. The gene residing between two *gypsy* insulators was protected against PRE-dependent repression. Meanwhile, the interaction between the insulators brought PRE closer to the second gene, thus leading to its repression. The physical interaction between insulators and the approximation of PRE to the second reporter gene was confirmed by 3C assay [190].

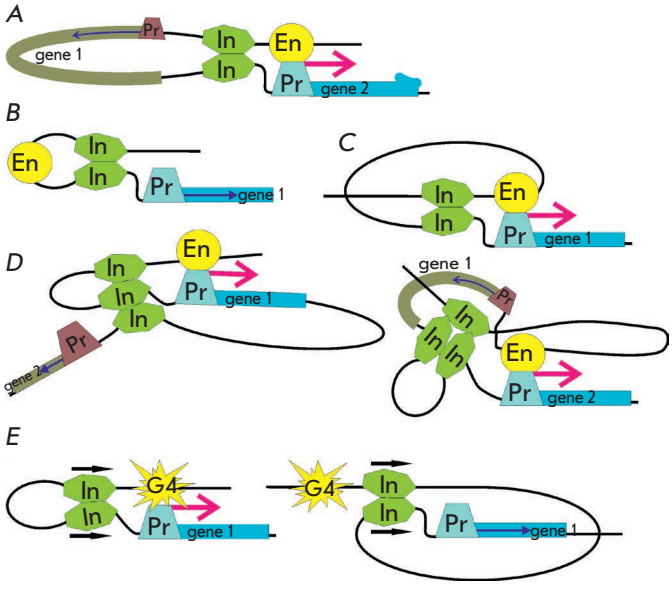


Fig. 5. Modeling chromatin loops in transgenic lines of drosophila. (*A*) A loop formed by identical insulators brings the enhancer closer to the promoter. (*B*) A tight loop between the two insulators blocks the enhancer it contains. (*C*) Increased distance between the insulators surrounding the enhancer neutralizes the insulation. (*D*) Loops formed by the three insulators do not interfere with activation of the reporter gene transcription. (*E*) Mutual orientation of insulators (indicated with arrows) determines the configuration of the chromatin loop and, as a consequence, the possibility of transcription activation. Designations: G4 – yeast activator GAL4; other designations are the same as those in *Fig. 2*.

Mutual neutralization of two identical insulators makes it possible to study the direct role played by the chromatin loop formed by them in the blocking of the enhancer-promoter contacts. As mentioned above, the integration of a single copy of the *gypsy* insulator between the enhancer and the *white* gene promoter weakens the enhancer activity only slightly [168]. However, surrounding the enhancer with a pair of *gypsy* insulators completely inactivates it. This result suggests that the formation of a small chromatin loop containing an enhancer either topologically or sterically prevents productive interaction between the enhancer and the *white* gene promoter (*Fig. 5B*). Meanwhile, in transgenic lines, a single copy of the *gypsy* insulator completely blocks the enhancers activating *yellow* gene expression in the body and wings [170, 191]. It turned out that integration of the second copy of the *gypsy* insulator upstream of the enhancers (~ 8 kbp upstream of the first enhancer) restores *yellow* gene expression. Thus, formation of the 8-bp chromatin loop neutralizes the insulator activity (*Fig. 5C*). The insulator activity was completely restored when the distance between the surrounding *yellow* insulators was decreased to 2 kbp. Therefore, only small chromatin loops containing the enhancer can completely block its activity. *In vivo*, chromatin loops are much larger than 2–3 kbp, suggesting that interactions can exist between the regulatory elements residing in neighboring loops or loops located at a distance.

Studies performed for Drosophila lines carrying three copies of Su(Hw) insulators integrated between the enhancers and two reporter genes in different combinations showed that all three copies interact with each other [170, 191]. The chromatin loop formed around the enhancer or the reporter gene did not induce insulator activity. This result confirms once again that chromatin loops do not play a crucial role in the blocking of enhancer–promoter interactions (*Fig. 5D*).

In transgenic Drosophila lines, pairs of some insulators (e.g., *gypsy*, *Mcp*, and *Fab-7*), can be involved in ultra-long-range interactions (at a distance as large as several hundred thousands of nucleotide pairs) [192, 193]. The *Homie* and *Nhomie* insulators were detected at the boundaries of the *eve* locus expressing *pair-rule* TF that is involved in embryonic development [194]. These insulators efficiently interact with each other in transgenic Drosophila lines and can maintain ultra-long-range interactions between enhancers and the promoter of the *eve* locus in the genome [194, 195].

A model has been proposed to explain the mechanism of ultra-long-range interactions between insulators [16]. According to this model, insulators consist of binding sites for several proteins; each of those can be efficiently homodimerized. Indeed, the boundary of *Mcp* from the BX-C contains binding sites for Pita, dCTCF, and two other unknown insulator proteins [143, 196]. The *Fab-7* boundary includes binding sites for GAF, Pita, Insv, Elba, the LBC complex, and several unknown proteins [57, 143, 157, 197, 198]. In transgenic Drosophila lines, paired binding sites for the Pita, ZIPIC, Zw5, dCTCF, and Su(Hw) proteins ensure long-range interactions between the reporter gene and yeast activator GAL4 [145, 146, 193]. However, any combination of the binding sites of different proteins results in a loss of interaction between insulators, thus confirming the contribution of protein homodimerization to long-range interactions.

Furthermore, the topology of chromatin loops depends on the mutual orientation of two identical insulators. This was demonstrated for the transgenic lines where GAL4 could not activate transcription of the *white* gene located at a long distance from it [146]. The identical insulators placed in close proximity to GAL4 and *white* promoter formed loops with two different configurations (*Fig. 5E*). If the insulators were oriented oppositely, GAL4 activated the *white* gene promoter. If the insulators had the same orientation, the resulting loop fully isolated GAL4 from the promoter. Similar results were obtained when the GAL4 activator was replaced with an enhancer [28, 187]. The mutual orientation of two *gypsy* insulators also affected the Flp-dependent recombination between FRT sites [199]. Oppositely oriented insulators located between the FRT sites contributed to recombination, whereas co-directional insulators inhibited it. Most likely, homodimerization of several proteins bound to identical insulators determines the direction of the interaction between them. The topology of the resulting chromatin loop regulates the interactions between the elements residing in close proximity to the insulators.

MODERN VIEWS ON CHROMOSOMAL ORGANIZATION INTO TOPOLOGICALLY ASSOCIATING DOMAINS

In all higher eukaryotes, chromosomes are organized into TADs. The size and mechanisms of formation of these domains greatly vary in different animal species [91, 200, 201]. Formation of TADs depends on the frequency of interaction between different chromatin parts: the interaction frequency within the domains is higher than that between the domains. Insulators inside TADs can form local chromatin loops, thus regulating the enhancer–promotion interactions (*Fig. 6*).

The CTCF protein and the cohesin complex interacting with it play the central role in the organization of TADs in mammals. Together with the cohesin complex, the CTCF protein resides at ~ 90% of TAD boundaries [89, 90]. The cohesin complex consisting of four subunits (SMC1, SMC3 and RAD21, SCC1) forms

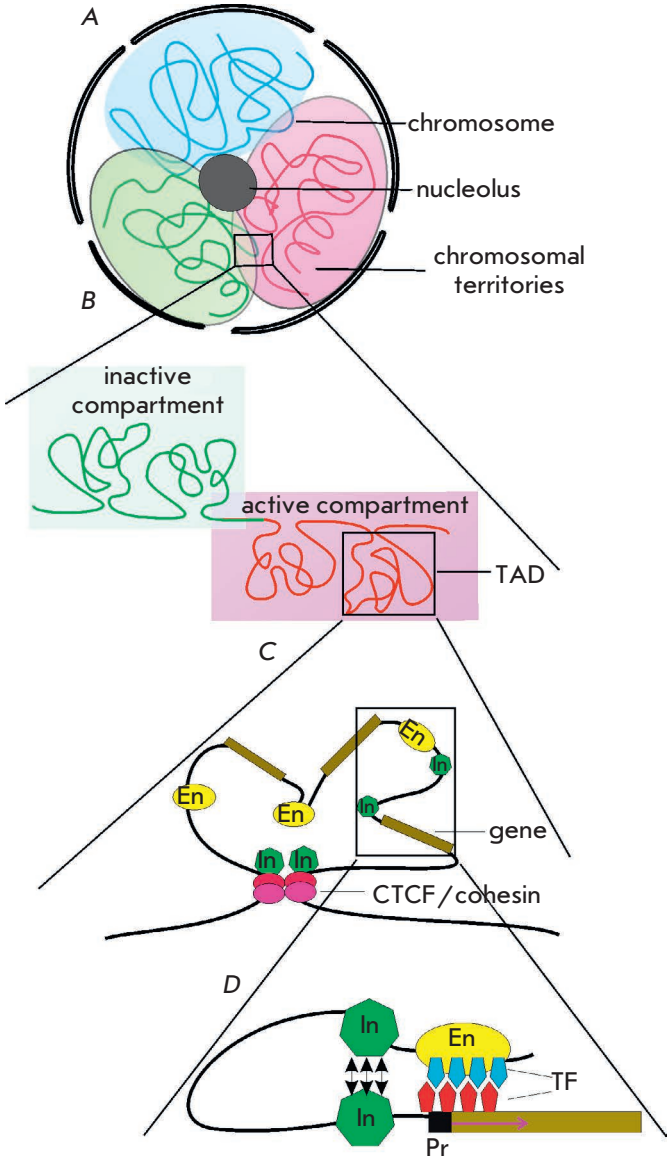


Fig. 6. The levels of chromatin organization in the nucleus: (*A*) Chromosomes within the nucleus occupy particular territories (red, green, and blue backgrounds). (*B*) Each chromosome forms TADs, which are involved in a particular nucleus compartment depending on the active/inactive chromatin state. (*C*) TADs facilitate the convergence of the regulatory elements within them and ensure synchronous gene expression. Architectural proteins can dynamically restrict the formation of TADs. (*D*) Insulators within a TAD may form local chromatin loops facilitating specific enhancer–promoter interactions. Designations: TFs – transcription factors; other designations are the same as those in *Fig. 2*

a ring-like structure around two DNA molecules [202]. It is believed that the cohesin complex can cause chromatin looping as chromatin passes through its ring-like structure (*Fig. 7A*). The cohesin complex slides along chromatin and forms loops; the binding sites of protein CTCF inverted with respect to each other act as limits for these loops [203–205]. Inactivation of CTCF or the cohesin complex components destroys most TADs, which agrees with the earlier described model [164, 206, 207]. The weak link in this model is the lack of experimental data that would confirm that the cohesin complex can cause chromatin looping *in vivo* [208].

In mammals, the role of CTCF-binding sites in the formation of TAD boundaries was studied in the murine *Hox* genes [209]. The *HoxA* and *HoxC* genes are located in the adjacent TADs and are transcribed independently. Deletion of the CTCF-binding site residing between these TADs destroyed their boundaries, thus altering the gene expression patterns and, therefore, causing homeotic transformation of the skeleton [210]. Unlike *HoxA* and *HoxC*, the *HoxD* gene is located between two TADs, each containing enhancers responsible for the function of *HoxD* in a certain tissue type. In this case, however, deletion of CTCF-binding sites in the *HoxD* gene did not destroy the TAD boundary and had a minimal impact on the gene expression pattern. The TAD boundary was destroyed, and the pattern of *HoxD* expression changed only after an extensive deletion affecting the structure of the regulatory regions of the gene. These data indicate that some additional TFs, along with CTCF and the cohesin complex, can be involved in the formation of TAD boundaries.

Unlike in vertebrates, dCTCF and the cohesin complex in Drosophila are not the key factors in TAD formation. The TADs being formed correlate well with epigenetic marks and are subdivided into classes corresponding to the specific features of chromatin: (1) the active TADs are actively transcribed and are rich in H3K4me3 and H3K36me3 histone modifications; (2) the polycomb-dependent TADs are rich in H3K27me3 histone modification and Polycomb group proteins; (3) "null" or "void" TADs have no known specific histone marks; and (4) heterochromatic TADs are rich in H3K9me2 mark and the HP1 and Su(var)3-9 proteins [91]. Chromatin regions separating the TADs are rich in genes with a high transcription level [211–213]. They actively interact with each other to form chromatin loops. There are no clearly defined sites of TAD formation such as inverted CTCF sites in mammals [211].

Hence, the TAD boundaries in Drosophila are more likely to depend on the active chromatin state and its properties rather than on the binding sites of a specific protein [213] (*Fig. 7B*). The dCTCF, CP190, Chromator, Z4, and BEAF-32 insulator proteins binding to

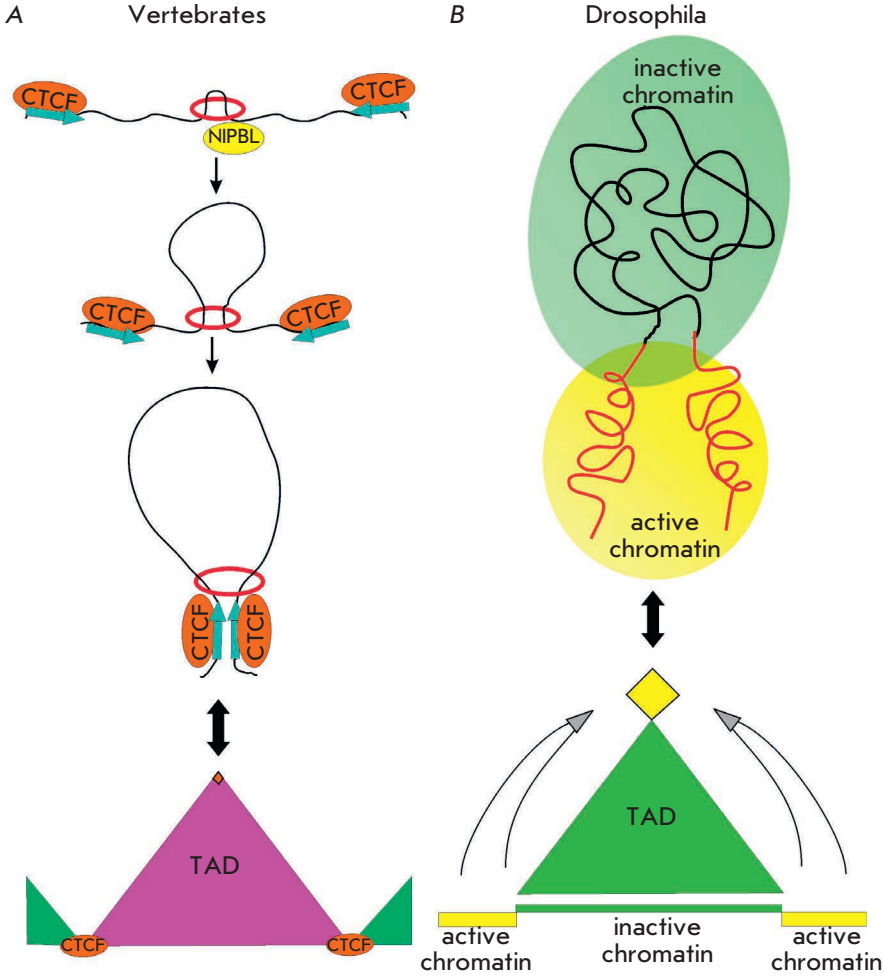


Fig. 7. The mechanism of formation of TADs in vertebrates and drosophila. (*A*) Loop formation by the cohesin complex. The cohesin complex (red ring), after being loaded onto chromatin by NIPBL, processively extrudes chromatin through its ring-shaped structure, resulting in a growing chromatin loop. Loop extrusion stops when cohesin encounters CTCF binding sites in a convergent orientation (designated by arrows). Triangles represent the neighboring TADs divided with CTCF sites. An orange rhombus at the top of the TAD designates the high frequency of interaction between CTCF-binding regions. *(B)* In drosophila, active and inactive chromatin is localized in different nuclear compartments. Inactive chromatin (a green rectangle) is confined to the areas with active transcription (yellow rectangles). The interaction of actively transcribed regions (shown with arrows) forms TAD. The yellow rhombus at the top of the TAD denotes the highest frequency of interaction between active chromatin regions

housekeeping gene promoters are often found at the TAD boundaries [201, 211, 212, 214]. However, the role played by these proteins in TAD boundary formation still needs to be elucidated.

Recent studies focused on chromatin architecture in individual mammalian cells have revealed the high heterogeneity of TAD boundary localization [215–218]. Meanwhile, DNA sites within the TADs interact on average only two to three times more frequently than sites from the adjacent TADs [89]. The transboundary interactions were confirmed by FISH [219, 220]. These results agree with the vigorous dynamics of binding/dissociation of the CTCF protein, which resides on chromatin for approximately 2 min [221]. Therefore, TAD formation is a dynamic process and TAD boundaries are not a rigid barrier limiting the enhancer-promotor interactions.

THE ROLE PLAYED BY INSULATORS AND TADS IN TRANSCRIPTION REGULATION

Drosophila insulators play a significant role in ensuring specific long-range *cis*-regulatory interactions, which has been demonstrated well for the BX-C [222]. The *Ubx*, *Abd-A*, and *Abd-B* homeotic genes within the BX-C are responsible for the formation of the third thoracic and all the abdominal segments of a fruit fly and determine its future head-to-abdomen axis. The BX-C is divided into nine regulatory domains (*iab 1–9*), each activating specific transcription of one out of three homeotic genes in a certain segment (*Fig. 8*). The BX-C contains two TADs whose shared boundary coincides with the *Fub* insulator residing between the regulatory domains of the *Ubx* and *Abd-A* genes [217] (*Fig. 8*). The *Mcp*, *Fab-6*, *Fab-7*, and *Fab-8* insulators have been the best studied. They determine the boundaries of the *iab-5*, *iab-6*, and *iab-7* domains that regulate the *Abd-B* expression level in the A5, A6, and A7 abdominal segments [222, 223]. The entire regulatory domain of the *Abd-B* gene is located within a single TAD. In the A5 segment, *iab-5* enhancers are active, while *iab-6* and *iab-7* enhancers are inactive. The *iab-6* enhancers ensuring stronger activation of *Abd-B* expression are active in the next segment (A6). Even stronger *iab-7* enhancers are active in the A7 segment. Therefore,

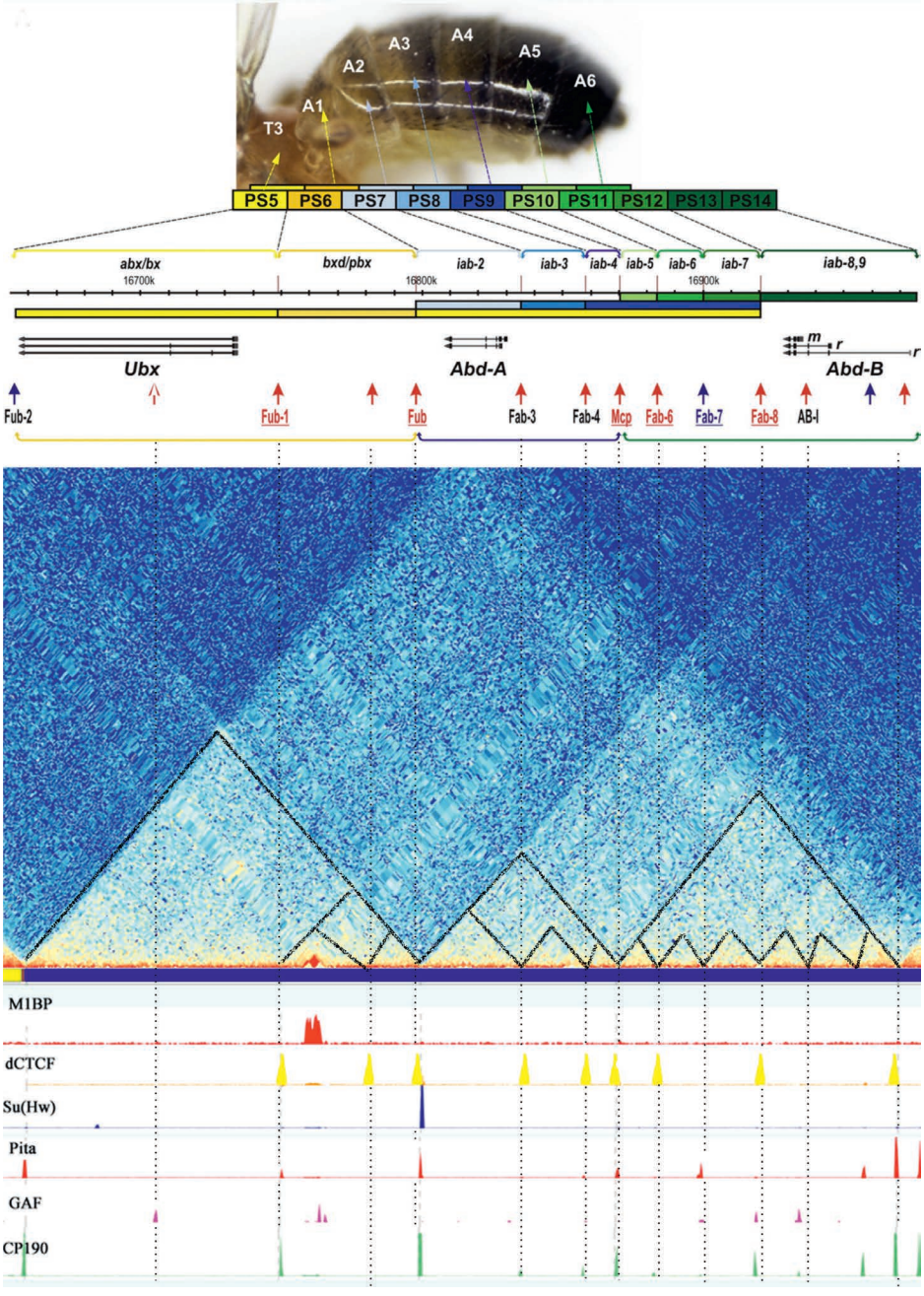


Fig. 8. Schematic representation of BX-C. The BX-C map and coordinates are taken from the FlyBase resource (R6.04). The colored rectangles represent the embryonic parasegments (PS) corresponding to the imago segments. The regulatory regions controlling the expression of the *Ubx*, *abd-A*, and *Abd-B* genes (horizontal arrows) in each PS are indicated with upper brackets. The regulatory regions are organized into three transcriptionally associated regions indicated with lower brackets. The pattern and expression level of each gene are designated by colored scale; the darker color indicates higher expression levels. BX-C insulators are indicated with arrows: red arrows denote the CTCF-dependent ones; blue arrows denote the CTCF-independent ones [223]. The distribution map of TADs and some insulator/architectural proteins in BX-C was constructed using the Chorogtnome Navigator dm3 resource [212]

Abd-B expression is enhanced in every segment thereafter, which is responsible for proper development of each abdominal segment. The interactions between the adjacent regulatory domains are blocked by insulators. For example, premature activity of *iab-6* enhancers in the A5 segment is observed when the *Fab-6* insulator is deleted.

In vivo genome editing made it possible to thoroughly study the structure and functions of insulators at the BX-C boundaries. It turned out that insulators consist of two modules: one blocking the communication between the adjacent regulatory domains (the insulator module) and the other one ensuring specific interaction between the insulator and the promoter of the *Abd-B* gene (the communicator module) [224, 225]. The Su(Hw), Pita, and dCTCF proteins, as well as the CP190 protein interacting with them, are involved in local insulation of the regulatory elements residing in the neighboring domains [143, 196, 226] (*Fig. 8*). The insulator module may consist of any combination of binding sites for these proteins, but there must be at least four sites. The communicator module of all insulators carries the binding sites of the poorly studied LBC complex comprising the GAF and CLAMP proteins

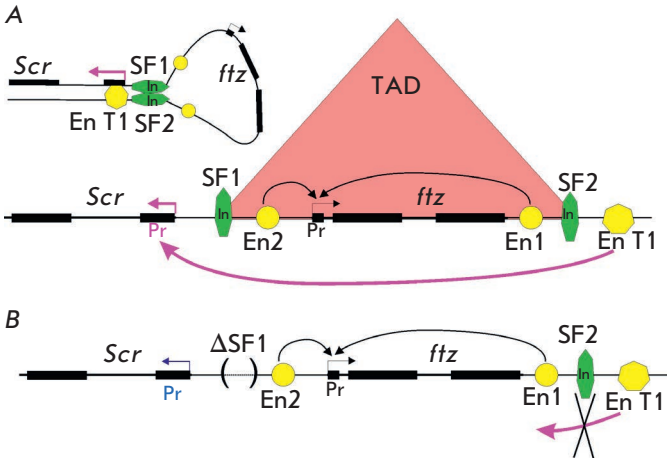


Fig. 9. The role played by the SF1 and SF2 insulators in TAD formation and transcription in ANT-C. (*A*) The boundaries of the TAD including the *ftz* gene are determined by the SF1 and SF2 insulators. Interacting insulators form a loop that brings the T1 enhancer closer to the *Scr* gene promoter. The T1 enhancer activates *Scr* transcription. (*B*) Deletion of the SF1 insulator (designated by brackets) leads to disruption of TAD but not misexpression of the *ftz* gene. However, the T1 enhancer does not activate *Scr* transcription, because a loop between insulators does not form. All designations are the same as those in *Fig. 2*

[198, 224]. The communicator modules interact with the pre-promotor domain of the *Abd-B* gene to form chromatin loops ensuring specific contacts between *iab* enhancers and the *Abd-B* promoter. SubTADs corresponding to individual *iab* domains have been revealed in embryonic cell populations [217]. The formation of subTADs correlates with activation of *iab* domains, being potentially indicative of interaction between active domains and the *Abd-B* promoter. This fact confirms that TADs are formed in *D. melanogaster* through the interaction between active chromatin sites, while insulator proteins stabilize the boundaries of the resulting domains.

The formation/destruction of TADs can only have a minimal effect on gene expression [164, 206, 227]. Thus, the TAD boundaries in the complex of homeotic ANT-C genes are determined by two insulators: SF1 and SF2 [53, 228] (*Fig. 9A*). Deletion of the SF1 insulator results in TAD destruction, while having no effect on the expression of the *fushi-tarazu* (*ftz*) gene residing inside the TAD. Interestingly, transcription of the *Scr* gene adjacent to the TAD is reduced [229] (*Fig. 9B*). In early embryos, the *Scr* gene located on one side of the TAD is activated by directly interacting with the T1 enhancer residing on the other side of the TAD [230] (*Fig. 9A*). Therefore, the interacting SF1 and SF2 insulators on the TAD boundaries bring together the T1 enhancer and the *Scr* gene. This situation fully implements the model developed for transgenic lines, according to which chromatin loop formation between insulators located at a distance contributes to enhancer-promotion interactions and transcription activation.

Furthermore, the effect of TAD boundaries on transcription was studied by performing precise deletion of different CTCF-binding sites in the *Sox9–Kcnj2* locus in mice [231]. Two TADs separated by a boundary containing inverted CTCF-binding sites resided in this locus (*Fig. 10A*). Several additional CTCF sites are also found inside each TAD. The *Sox9* and *Kcnj2* genes are activated by specific enhancers and have different expression patterns. Deletion of CTCF-binding sites on the boundary between the *Sox9* and *Kcnj2* genes did not cause merging of the TADs (*Fig. 10B*). A merged TAD was formed only after additional internal CTCF sites had been deleted (*Fig. 10C*). It is noteworthy that during TAD merging, the enhancers did not activate the nonspecific gene and expression of the *Sox9* and *Kcnj2* genes remained almost unchanged. It is possible that the high specificity of enhancer-promotor interactions did not allow the cohesion complex to form new contacts between the regulatory elements in the shared *Sox9–Kcnj2* locus. Therefore, the TAD boundary was not involved in the organization of specific enhancer-promotor interactions. Inversion, which had moved the TAD boundary to a position between the *Sox9* gene enhancers and its promoter, resulted in the formation of two new domains (*Fig. 10D*). In this case, the TAD boundary had a critical impact on transcription. The *Sox9* enhancers isolated from the promoter could not activate the specific gene but activated *Kcnj2*, which had a lethal effect.

These examples allow one to infer that chromosomal organization into topological structures and specific enhancer-promotor interactions are two different transcription regulation levels that are often independent. Only in some cases do the TAD boundaries act as insulators regulating the enhancer-promoter interactions.

The correlation between gene expression and an altered chromatin architecture was also studied in Drosophila lines carrying chromosomes with multiple inversions and deletions [232]. It was revealed that significant changes in the TAD organization have a negligible effect on gene transcription. These data once again indicate that TADs play a secondary role in gene expression regulation in higher eukaryotes.

CONCLUSIONS

Today, it is obvious that TADs form the chromosomal architecture but do not act as transcription

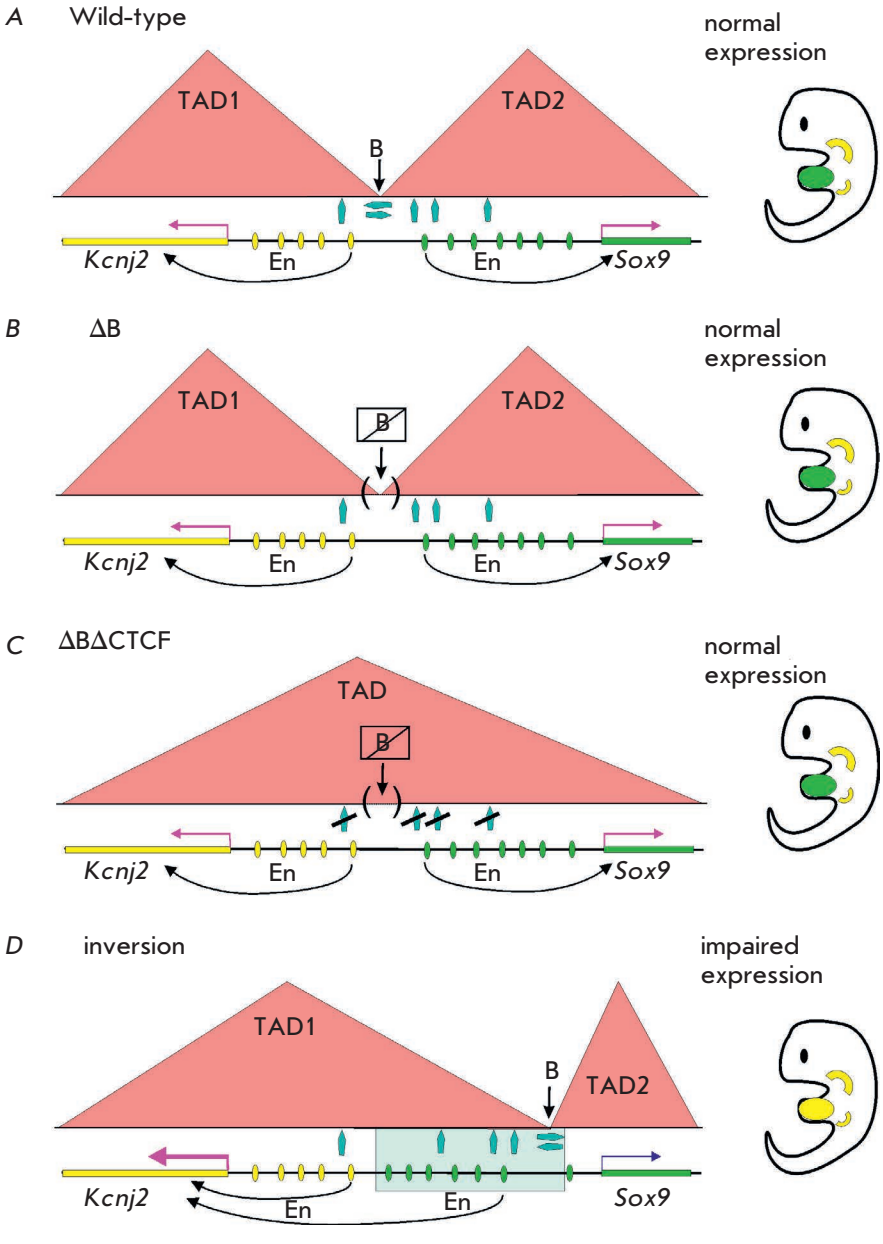


Fig. 10. The role of TADs in *Kcnj2* and *Sox9* loci expression. *(A)* Wild-type expression of the *Kcnj2* and *Sox9* genes. Two separate TADs are formed, the boundary (B) between which it colocalizes with the convergently oriented CTCF binding sites. (*B*) Deletion of CTCF-binding sites at the TAD boundary neither destroys them nor affects the gene expression patterns. (*C*) Simultaneous deletion of boundary and internal CTCF sites leads to fusion of TADs but does not affect gene expression. (*D*) Relocation of the boundary between TADs results in gene misexpression. Designations: blue arrows – CTCF-binding sites; the expression patterns of the *Kcnj2* and *Sox9* genes in the embryo are shown in yellow and green, respectively; the direction of enhancer action is shown with arrows; other designations are the same as those in *Figs. 2* and *9*

domains regulating gene expression. In *Drosophila melanogaster*, most TAD boundaries are formed by promoters of actively transcribed genes. In some cases, the TAD boundaries coincide with insulators. Interestingly, many proteins binding to insulators are also components of the complexes assembled on promoters. Insulators are the multifunctional regulatory elements. They ensure the specificity of enhancer-promotion interactions, form the boundaries between active and inactive chromatin, and form the regions containing open chromatin available for TF. The experimental data demonstrate that insulators inhibit enhancer activity by directly interacting with enhancers or promoters. Chromatin loops formed by insulators play only an auxiliary role in insulation. The question of how long-range interactions between enhancers, silencers, promoters, and insulators form and are regulated still remains open. There is little doubt that insulator proteins play a crucial role in this process. However, their mechanism of action needs further study. ●

This work was supported by the Russian Science Foundation (project No. 18-14-00295).

REFERENCES

1. Field A., Adelman K. // Annu. Rev. Biochem. 2020. V. 89. P. 213–234.
2. Klemm S.L., Shipony Z., Greenleaf W.J. // Nat. Rev. Genet. 2019. V. 20. № 4. P. 207–220.
3. Schoenfelder S., Fraser P. // Nat. Rev. Genet. 2019. V. 20. № 8. P. 437–455.
4. Frankel N., Davis G.K., Vargas D., Wang S., Payre F., Stern D.L. // Nature. 2010. V. 466. № 7305. P. 490–493.
5. Lettice L.A., Williamson I., Devenney P.S., Kilanowski F., Dorin J., Hill R.E. // Development. 2014. V. 141. № 8. P. 1715–1725.
6. Benyajati C., Worcel A. // Cell. 1976. V. 9. № 3. P. 393–407.
7. Geyer P.K., Corces V.G. // Genes Dev. 1992. V. 6. № 10. P. 1865–1873.
8. Kellum R., Schedl P. // Cell. 1991. V. 64. № 5. P. 941–950.
9. Kellum R., Schedl P. // Mol. Cell. Biol. 1992. V. 12. № 5. P. 2424–2431.
10. Phillips J.E., Corces V.G. // Cell. 2009. V. 137. № 7. P. 1194–1211.
11. Wallace J.A., Felsenfeld G. // Curr. Opin. Genet. Dev. 2007. V. 17(5). P. 400–407.
12. Ali T., Renkawitz R., Bartkuhn M. // Curr. Opin. Genet. Dev. 2016. V. 37. P. 17–26.
13. Chetverina D., Fujioka M., Erokhin M., Georgiev P., Jaynes J.B., Schedl P. // Bioessays. 2017. V. 39. № 3. doi: 10.1002/bies.201600233.
14. Ghirlando R., Felsenfeld G. // Genes Dev. 2016. V. 30. № 8. P. 881–891.
15. Hnisz D., Day D.S., Young R.A. // Cell. 2016. V. 167. № 5. P. 1188–1200.
16. Kyrchanova O., Georgiev P. // FEBS Lett. 2014. V. 588. № 1. P. 8–14.
17. Matzat L.H., Lei E.P. // Biochim. Biophys. Acta. 2014. V. 1839. № 3. P. 203–214.
18. Chen D., Lei E.P. // Curr. Opin. Cell. Biol. 2019. V. 58. P. 61–68.
19. Spradling A.C., Stern D.M., Kiss I., Roote J., Laverty T., Rubin G.M. // Proc. Natl. Acad. Sci. USA. 1995. V. 92. № 24. P. 10824–10830.
20. Carlson C.M., Largaespada D.A. // Nat. Rev. Genet. 2005. V. 6. № 7. P. 568–580.
21. Jiang F., Doudna J.A. // Annu. Rev. Biophys. 2017. V. 46. P. 505–529.
22. Pirrotta V., Steller H., Bozzetti M.P. // EMBO J. 1985. V. 4. № 13A. P. 3501–3508.
23. Levis R., Hazelrigg T., Rubin G.M. // Science. 1985. V. 229. № 4713. P. 558–561.
24. Silicheva M., Golovnin A., Pomerantseva E., Parshikov A., Georgiev P., Maksimenko O. // Nucl. Acids Res. 2010. V. 38. № 1. P. 39–47.
25. Udvardy A., Maine E., Schedl P. // J. Mol. Biol. 1985. V. 185. № 2. P. 341–358.
26. Kuhn E.J., Hart C.M., Geyer P.K. // Mol. Cell. Biol. 2004. V. 24. № 4. P. 1470–1480.
27. Gaszner M., Vazquez J., Schedl P. // Genes Dev. 1999. V. 13. № 16. P. 2098–2107.
28. Kyrchanova O., Leman D., Parshikov A., Fedotova A., Studitsky V., Maksimenko O., Georgiev P. // PLoS One. 2013. V. 8. № 4. e62690.
29. Vazquez J., Schedl P. // EMBO J. 1994. V. 13. № 24. P. 5984–5993.
30. Zhao K., Hart C.M., Laemmli U.K. // Cell. 1995. V. 81. № 6. P. 879–889.
31. Modolell J., Bender W., Meselson M. // Proc. Natl. Acad. Sci. USA. 1983. V. 80. № 6. P. 1678–1682.
32. Kim J., Shen B., Rosen C., Dorsett D. // Mol. Cell. Biol. 1996. V. 16. № 7. P. 3381–3392.
33. Bender W., Akam M., Karch F., Beachy P.A., Peifer M., Spierer P., Lewis E.B., Hogness D.S. // Science. 1983. V. 221. № 4605. P. 23–29.
34. Hoover K.K., Gerasimova T.I., Chien A.J., Corces V.G. // Genetics. 1992. V. 132. № 3. P. 691–697.
35. Parkhurst S.M., Harrison D.A., Remington M.P., Spana C., Kelley R.L., Coyne R.S., Corces V.G. // Genes Dev. 1988. V. 2. № 10. P. 1205–1215.
36. Peifer M., Bender W. // EMBO J. 1986. V. 5. № 9. P. 2293–3203.
37. Geyer P.K., Green M.M., Corces V.G. // Proc. Natl. Acad. Sci. USA. 1988. V. 85. № 22. P. 8593–8597.
38. Smith P.A., Corces V.G. // Genetics. 1995. V. 139. № 1. P. 215–228.
39. Geyer P.K., Spana C., Corces V.G. // EMBO J. 1986. V. 5. № 10. P. 2657–2662.
40. Parkhurst S.M., Corces V.G. // Mol. Cell. Biol. 1986. V. 6. № 1. P. 47–53.
41. Mallin D.R., Myung J.S., Patton J.S., Geyer P.K. // Genetics. 1998. V. 148. № 1. P. 331–339.
42. Sigrist C.J., Pirrotta V. // Genetics. 1997. V. 147. № 1. P. 209–221.
43. Conte C., Dastugue B., Vaury C. // Mol. Cell. Biol. 2002. V. 22. № 6. P. 1767–1777.
44. Brasset E., Hermant C., Jensen S., Vaury C. // Gene. 2010. V. 450. № 1. P. 25–31.
45. Golovnin A., Biryukova I., Romanova O., Silicheva M., Parshikov A., Savitskaya E., Pirrotta V., Georgiev P. // Development. 2003. V. 130. № 14. P. 3249–3258.
46. Parnell T.J., Viering M.M., Skjesol A., Helou C., Kuhn E.J., Geyer P.K. // Proc. Natl. Acad. Sci. USA. 2003. V. 100. № 23. P. 13436–13441.
47. Negre N., Brown C.D., Shah P.K., Kheradpour P., Morrison C.A., Henikoff J.G., Feng X., Ahmad K., Russell S., White R.A., et al. // PLoS Genet. 2010. V. 6. № 1. e1000814.
48. 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. № 11. P. 2188–2198.
49. Soshnev A.A., Ishimoto H., McAllister B.F., Li X., Wehling M.D., Kitamoto T., Geyer P.K. // Genetics. 2011. V. 189. № 2. P. 455–468.
50. Scott K.C., Taubman A.D., Geyer P.K. // Genetics. 1999. V. 153. № 2. P. 787–798.
51. Melnikova L., Kostyuchenko M., Parshikov A., Georgiev P., Golovnin A. // PLoS One. 2018. V. 13. № 2. e0193497.
52. Belozerov V.E., Majumder P., Shen P., Cai H.N. // EMBO J. 2003. V. 22. № 12. P. 3113–3121.
53. Li M., Ma Z., Liu J.K., Roy S., Patel S.K., Lane D.C., Cai H.N. // Mol. Cell. Biol. 2015. V. 35. № 23. P. 4018–4029.
54. Vazquez J., Schedl P. // Genetics. 2000. V. 155. № 3. P. 1297–1311.
55. Chetverina D., Savitskaya E., Maksimenko O., Melnikova L., Zaytseva O., Parshikov A., Galkin A.V., Georgiev P. // Nucl. Acids Res. 2008. V. 36. № 3. P. 929–937.
56. Sultana H., Verma S., Mishra R.K. // Nucl. Acids Res. 2011. V. 39. № 9. P. 3543–3557.
57. Aoki T., Sarkeshik A., Yates J., Schedl P. // Elife. 2012. V. 1. e00171.
58. Barges S., Mihaly J., Galloni M., Hagstrom K., Muller M.,

Shanower G., Schedl P., Gyurkovics H., Karch F. // Development. 2000. V. 127. № 4. P. 779–790.
59. Gruzdeva N., Kyrchanova O., Parshikov A., Kullyev A., Georgiev P. // Mol. Cell. Biol. 2005. V. 25. № 9. P. 3682–3689.
60. Gyurkovics H., Gausz J., Kummer J., Karch F. // EMBO J. 1990. V. 9. № 8. P. 2579–2585.
61. Hogga I., Mihaly J., Barges S., Karch F. // Mol. Cell. 2001. V. 8. № 5. P. 1145–1151.
62. Iampietro C., Cleard F., Gyurkovics H., Maeda R.K., Karch F. // Development. 2008. V. 135. № 24. P. 3983–3987.
63. Iampietro C., Gummalla M., Mutero A., Karch F., Maeda R.K. // PLoS Genet. 2010. V. 6. № 12. e1001260.
64. Rodin S., Kyrchanova O., Pomerantseva E., Parshikov A., Georgiev P. // Genetics. 2007. V. 177. № 1. P. 113–121.
65. Schweinsberg S.E., Schedl P. // Development. 2004. V. 131. № 19. P. 4743–4749.
66. Ciavatta D., Rogers S., Magnuson T. // J. Mol. Biol. 2007. V. 373. № 2. P. 233–239.
67. Hagstrom K., Muller M., Schedl P. // Genes Dev. 1996. V. 10. № 2. P. 3202–3215.
68. Maksimenko O., Bartkuhn M., Stakhov V., Herold M., Zolotarev N., Jox T., Buxa M.K., Kirsch R., Bonchuk A., Fedotova A., et al. // Genome Res. 2015. V. 25. № 1. P. 89–99.
69. Perez-Lluch S., Cuartero S., Azorin F., Espinas M.L. // Nucl. Acids Res. 2008. V. 36. № 21. P. 6926–6933.
70. Schweinsberg S., Hagstrom K., Gohl D., Schedl P., Kumar R.P., Mishra R., Karch F. // Genetics. 2004. V. 168. № 3. P. 1371–1384.
71. Zhou J., Barolo S., Szymanski P., Levine M. // Genes Dev. 1996. V. 10. № 24. P. 3195–3201.
72. Chung J.H., Whiteley M., Felsenfeld G. // Cell. 1993. V. 74. № 3. P. 505–514.
73. Bell A.C., West A.G., Felsenfeld G. // Cell. 1999. V. 98. № 3. P. 387–396.
74. Arzate-Mejia R.G., Recillas-Targa F., Corces V.G. // Development. 2018. V. 145. № 6. P. dev137729.
75. Herold M., Bartkuhn M., Renkawitz R. // Development. 2012. V. 139. № 6. P. 1045–1057.
76. Bell A.C., Felsenfeld G. // Nature. 2000. V. 405. № 6785. P. 482–485.
77. Hark A.T., Schoenherr C.J., Katz D.J., Ingram R.S., Levorse J.M., Tilghman S.M. // Nature. 2000. V. 405. № 6785. P. 486–489.
78. Kanduri C., Pant V., Loukinov D., Pugacheva E., Qi C.F., Wolffe A., Ohlsson R., Lobanenkov V.V. // Curr. Biol. 2000. V. 10. № 14. P. 853–856.
79. Dorsett D. // Curr. Opin. Genet. Dev. 1999. V. 9. № 5. P. 505–514.
80. Geyer P.K. // Curr. Opin. Genet. Dev. 1997. V. 7. № 2. P. 242–248.
81. Liu G., Dean A. // Biochim. Biophys. Acta Gene Regul. Mech. 2019. V. 1862. № 6. P. 625–633.
82. Morcillo ., Rosen C., Baylies M.K., Dorsett D. // Genes Dev. 1997. V. 11. № 20. P. 2729–2740.
83. Gause M., Morcillo P., Dorsett D. // Mol. Cell. Biol. 2001. V. 21. № 14. P. 4807–4817.
84. Torigoi E., Bennani-Baiti I.M., Rosen C., Gonzalez K., Morcillo P., Ptashne M., Dorsett D. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. № 6. P. 2686–2691.
85. Gerasimova T.I., Corces V.G. // Annu. Rev. Genet. 2001. V. 35. P. 193–208.
86. Gerasimova T.I., Byrd K., Corces V.G. // Mol. Cell. 2000. V. 6. № 5. P. 1025–1035.
87. Valenzuela L., Kamakaka R.T. // Annu. Rev. Genet. 2006. V. 40. P. 107–138.
88. Boettiger A., Murphy S. // Trends Genet. 2020. V. 36. № 4. P. 273–287.
89. Chang L.H., Ghosh S., Noordermeer D. // J. Mol. Biol. 2020. V. 432. № 3. P. 643–652.
90. Sikorska N., Sexton T. // J. Mol. Biol. 2020. V. 432. № 3. P. 653–664.
91. Szabo Q., Bantignies F., Cavalli G. // Sci. Adv. 2019. V. 5. № 4. eaaw1668.
92. Harrison D.A., Gdula D.A., Coyne R.S., Corces V.G. // Genes Dev. 1993. V. 7. № 10. P. 1966–1978.
93. Baxley R.M., Bullard J.D., Klein M.W., Fell A.G., Morales-Rosado J.A., Duan T., Geyer P.K. // Nucl. Acids Res. 2017. V. 45. № 8. P. 4463–4478.
94. Soshnev A.A., He B., Baxley R.M., Jiang N., Hart C.M., Tan K., Geyer P.K. // Nucl. Acids Res. 2012. V. 40. № 12. P. 5415–5431.
95. Georgiev P., Kozycina M. // Genetics. 1996. V. 142. № 2. P. 425–436.
96. Duan T., Geyer P.K. // Genetics. 2018. V. 209. № 3. P. 757–772.
97. Melnikova L., Elizar'ev P., Erokhin M., Molodina V., Chetverina D., Kostyuchenko M., Georgiev P., Golovnin A. // Sci. Rep. 2019. V. 9. P. 5314.
98. Soshnev A.A., Baxley R.M., Manak J.R., Tan K., Geyer P.K. // Development. 2013. V. 140. № 17. P. 3613–3623.
99. Buchner K., Roth P., Schotta G., Krauss V., Saumweber H., Reuter G., Dorn R. // Genetics. 2000. V. 155. № 1. P. 141–157.
100. Gerasimova T.I., Gdula D.A., Gerasimov D.V., Simonova O., Corces V.G. // Cell. 1995. V. 82. № 4. P. 587–597.
101. Zollman S., Godt D., Prive G.G., Couderc J.L., Laski F.A. // Proc. Natl. Acad. Sci. USA. 1994. V. 91. № 22. P. 10717–10721.
102. Bonchuk A., Denisov S., Georgiev P., Maksimenko O. // J. Mol. Biol. 2011. V. 412. № 3. P. 423–436.
103. Ghosh D., Gerasimova T.I., Corces V.G. // EMBO J. 2001. V. 20. № 10. P. 2518–2527.
104. Melnikova L., Kostyuchenko M., Molodina V., Parshikov A.., Georgiev P., Golovnin A. // Open Biol. 2017. V. 7. № 10. P. 170150.
105. Oliver D., Sheehan B., South H., Akbari O., Pai C.Y. // BMC Cell Biol. 2010. V. 11. P. 101.
106. Plevock K.M., Galletta B.J., Slep K.C., Rusan N.M. // PLoS One. 2015. V. 10. № 12. e0144174.
107. Vogelmann J., Le Gall A., Dejardin S., Allemand F., Gamot A., Labesse G., Cuvier O., Negre N., Cohen-Gonsaud M., Margeat E., et al. // PLoS Genet. 2014. V. 10. № 8. P. e1004544.
108. Pai C.Y., Lei E.P., Ghosh D., Corces V.G. // Mol. Cell. 2004. V. 16. № 5. P. 737–748.
109. Melnikova L., Kostyuchenko M., Molodina V., Parshikov A., Georgiev P., Golovnin A. // Chromosoma. 2018. V. 127. № 1. P. 59–71.
110. Golovnin A., Mazur A., Kopantseva M., Kurshakova M., Gulak P.V., Gilmore B., Whitfield W.G., Geyer P., Pirrotta V., Georgiev P. // Mol. Cell. Biol. 2007. V. 27. № 3. P. 963–974.
111. Nègre N., Brown C.D., Ma L., Bristow C.A., Miller S.W., Wagner U., Kheradpour P., Eaton M.L., Loriaux P., Sealfon R., et al. // Nature. 2011. V. 471. № 7339. P. 527–531.
112. Alekseyenko A.A., Gorchakov A.A., Zee B.M., Fuchs S.M., Kharchenko P.V., Kuroda M.I. // Genes Dev. 2014. V. 28. № 13. P. 1445–1460.
113. Caron C., Pivot-Pajot C., van Grunsven L.A., Col E., Lestrat C., Rousseaux S., Khochbin S. // EMBO Rep. 2003. V. 4. № 9. P. 877–882.

114. Lahn B.T., Tang Z.L., Zhou J., Barndt R..J, Parvinen M., Allis C.D., Page D.C. // Proc. Natl. Acad. Sci. USA. 2002. V. 99. № 13. P. 8707–8712.
115. Melnikova L., Molodina V., Erokhin M., Georgiev P., Golovnin A. // Sci. Rep. 2019. V. 9. № 1. P. 19102.
116. Glenn S.E., Geyer P.K. // G3 (Bethesda). 2019. V. 9. № 2. P. 345–357.
117. Kurshakova M., Maksimenko O., Golovnin A., Pulina M., Georgieva S., Georgiev P., Krasnov A. // Mol. Cell. 2007. V. 27. № 2. P. 332–338.
118. Maksimenko O., Kyrchanova O., Bonchuk A., Stakhov V., Parshikov A., Georgiev P. // Epigenetics. 2014. V. 9. № 9. P. 1261–1270.
119. King M.R., Matzat L.H, Dale R.K., Lim S.J., Lei E.P. // J. Cell Sci. 2014. V. 127. № 13. P. 2956–2966.
120. Matzat L.H., Dale R.K, Moshkovich N., Lei E.P. // PLoS Genet. 2012. V. 8. № 11. P. e1003069.
121. Lei E.P., Corces V.G. // Nat. Genet. 2006. V. 38. № 8. P. 936–941.
122. Golovnin A., Melnikova L., Shapovalov I., Kostyuchenko M., Georgiev P. // PLoS One. 2015. V. 10. № 10. e0140991.
123. Golovnin A., Melnikova L., Volkov I., Kostuchenko M., Galkin A.V., Georgiev P. // EMBO Rep. 2008. V. 9. № 5. P. 440–445.
124. Golovnin A., Volkov I., Georgiev P. // J. Cell Sci. 2012. V. 125. № 8. P. 2064–2074.
125. Gerasimova T.I., Lei E.P., Bushey A.M., Corces V.G. // Mol. Cell. 2007. V. 28. № 5. P. 61–72.
126. Wasser M., Chia W. // PLoS One. 2007. V. 2. № 5. e412.
127. Melnikova L., Shapovalov I., Kostyuchenko M., Georgiev P., Golovnin A. // Chromosoma. 2017. V. 126. № 2. P. 299–311.
128. Iuchi S. // Cell. Mol. Life Sci. 2001. V. 58. № 4. P. 625–635.
129. Najafabadi H.S., Mnaimneh S., Schmitges F.W., Garton M., Lam K.N., Yang A., Albu M., Weirauch M.T., Radovani E., Kim P.M., et al. // Nat. Biotechnol. 2015. V. 33. № 5. P. 555–562.
130. Razin S.V., Borunova V.V., Maksimenko O.G., Kantidze O.L. // Biochemistry (Moscow) 2012. V. 77. № 3. P. 217–226.
131. Heger P., Marin B., Bartkuhn M., Schierenberg E., Wiehe T. // Proc. Natl. Acad. Sci. USA. 2012. V. 109. № 43. P. 17507–17512.
132. 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. № 7180. P. 796–801.
133. Heath H., Ribeiro de Almeida C., Sleutels F., Dingjan G., van de Nobelen S., Jonkers I., Ling K.W., Gribnau J., Renkawitz R., Grosveld F., et al. // EMBO J. 2008. V. 27. № 21. P. 2839–2850.
134. Soshnikova N., Montavon T., Leleu M., Galjart N., Duboule D. // Dev. Cell. 2010. V. 19. № 6. P. 819–830.
135. Splinter E., Heath H., Kooren J., Palstra R.J., Klous P., Grosveld F., Galjart N., de Laat W., et al. // Genes Dev. 2006. V. 20. № 17. P. 2349–2354.
136. Moon H., Filippova G., Loukinov D., Pugacheva E., Chen Q., Smith S.T., Munhall A., Grewe B., Bartkuhn M., Arnold R., et al. // EMBO Rep. 2005. V. 6. № 2. P. 165–170.
137. Hashimoto H., Wang D., Horton J.R., Zhang X., Corces V.G., Cheng X. // Mol. Cell. 2017. V. 66. № 5. P. 711–720 e3.
138. Nakahashi H., Kieffer 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. № 5. P. 1678–1689.
139. Bonchuk A., Kamalyan S., Mariasina S., Boyko K., Popov V., Maksimenko O., Georgiev P. // Sci. Rep. 2020. V. 10. № 1. P. 2677.
140. Li Y., Haarhuis J.H.I., Sedeno Cacciatore A., Oldenkamp R., van Ruiten M.S., Willems L., Teunissen H., Muir K.W., de Wit E., Rowland B.D., et al. // Nature. 2020. V. 578. № 7795. P. 472–476.
141. 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. № 3. P. 422–433.
142. Fedotova A.A., Bonchuk A.N., Mogila V.A., Georgiev P.G. // Acta Naturae. 2017. V. 9. № 2. P. 47–58.
143. Kyrchanova O., Zolotarev N., Mogila V., Maksimenko O., Schedl P., Georgiev P. // Development. 2017. V. 144. № 14. P. 2663–2672.
144. Page A.R., Kovacs A., Deak P., Torok T., Kiss I., Dario P., Bastos C., Batista P., Gomes R., Ohkura H., et al. // EMBO J. 2005. V. 24. № 24. P. 4304–4315.
145. Zolotarev N., Fedotova A., Kyrchanova O., Bonchuk A., Penin A.A., Lando A.S., Eliseeva I.A., Kulakovskiy I.V., Maksimenko O., Georgiev P. // Nucl. Acids Res. 2016. V. 44. № 15. P. 7228–7241.
146. Kyrchanova O., Chetverina D., Maksimenko O., Kullyev A., Georgiev P. // Nucl. Acids Res. 2008. V. 36. № 22. P. 7019–7028.
147. Espinas M.L., Jimenez-Garcia E., Vaquero A., Canudas S., Bernues J., Azorin F. // J. Biol. Chem. 1999. V. 274. № 23. P. 16461–16469.
148. Lu Q., Wallrath L.L., Granok H., Elgin S.C. // Mol. Cell. Biol. 1993. V. 13. № 5. P. 2802–2814.
149. Bartoletti M., Rubin T., Chalvet F., Netter S., Dos Santos N., Poisot E., Paces-Fessy M., Cumenal D., Peronnet F., Pret A.M., et al. // PLoS One. 2012. V. 7. № 11. e49958.
150. Melnikova L., Juge F., Gruzdeva N., Mazur A., Cavalli G., Georgiev P. // Proc. Natl. Acad. Sci. USA. 2004. V. 101. № 41. P. 14806–14811.
151. Pagans S., Ortiz-Lombardia M., Espinas M.L., Bernues J., Azorin F. // Nucl. Acids Res. 2002. V. 30. № 20. P. 4406–4413.
152. Hart C.M., Zhao K., Laemmli U.K. // Mol. Cell. Biol. 1997. V. 17. № 2. P. 999–1009.
153. Gilbert M.K., Tan Y.Y., Hart C.M. // Genetics. 2006. V. 173. № 3. P. 1365–1375.
154. Emberly E., Blattes R., Schuettengruber B., Hennion M., Jiang N., Hart C.M., Kas E., Cuvier O. // PLoS Biol. 2008. V. 6. № 12. P. 2896–2910.
155. Jiang N., Emberly E., Cuvier O., Hart C.M. // Mol. Cell. Biol. 2009. V. 29. № 13. P. 3556–3568.
156. Cuartero S., Fresan U., Reina O., Planet E., Espinas M.L. // EMBO J. 2014. V. 33. № 6. P. 637–647.
157. Fedotova A., Aoki T., Rossier M., Mishra R.K., Clendinen C., Kyrchanova O., Wolle D., Bonchuk A., Maeda R. K., Mutero A., et al. // Genetics. 2018. V. 210. № 2. P. 573–585.
158. Dai Q., Ren A., Westholm J.O., Duan H., Patel D.J., Lai E.C. // Genes Dev. 2015. V. 29. № 1. P. 48–62.
159. Bartkuhn M., Straub T., Herold M., Herrmann M., Rathke C., Saumweber H., Gilfillan G.D., Becker P.B., Renkawitz R. // EMBO J. 2009. V. 28. № 7. P. 877–888.
160. Bushey A.M., Ramos E., Corces V.G. // Genes Dev. 2009. V. 23. № 11. P. 1338–1350.
161. Mohan M., Bartkuhn M., Herold M., Philippen A., Heinl N., Bardenhagen I., Leers J., White R.A., Renkawitz-Pohl R., Saumweber H., et al. // EMBO J. 2007. V. 26. № 19. P. 4203–4214.
162. Liang J., Lacroix L., Gamot A., Cuddapah S., Queille S., Lhoumaud P., Lepetit P., Martin P.G., Vogelmann J., Court F., et al. // Mol. Cell. 2014. V. 53. № 4. P. 672–681.
163. Ahanger S.H., Gunther K., Weth O., Bartkuhn M.,

Bhonde R.R., Shouche Y.S., Renkawitz R. // Sci. Rep. 2014. V. 4. P. 3917.
164. Nora E.P., Goloborodko A., Valton A.L., Gibcus J.H, Uebersohn A., Abdennur N., Dekker J., Mirny L.A., Bruneau B.G. // Cell. 2017. V. 169. № 5. P. 930–944 e22.
165. Golovnin A., Melnick E., Mazur A., Georgiev P. // Genetics. 2005. V. 170. № 3. P. 1133–1142.
166. 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. № 11. P. 5717–5730.
167. Erokhin M., Davydova A., Kyrchanova O., Parshikov A., Georgiev P., Chetverina D. // Development. 2011. V. 138. № 18. P. 4097–4106.
168. Kyrchanova O., Maksimenko O., Stakhov V., Ivlieva T., Parshikov A., Studitsky V.M., Georgiev P. // PLoS Genet. 2013. V. 9. № 7. P. e1003606.
169. Holohan E.E., Kwong C., Adryan B., Bartkuhn M., Herold M., Renkawitz R., Russell S., White R. // PLoS Genet. 2007. V. 3. № 7. P. e112.
170. Maksimenko O., Golovnin A., Georgiev P. // Mol. Cell. Biol. 2008. V. 28. № 17. P. 5469–5477.
171. Melnikova L., Elizar'ev P., Erokhin M., Molodina V., Chetverina D., Kostyuchenko M., Georgiev P., Golovnin A. // Sci. Rep. 2019. V. 9. № 1. P. 5314.
172. Melnikova L., Kostuchenko M., Silicheva M., Georgiev P. // Chromosoma. 2008. V. 117. № 2. P. 137–145.
173. Kostyuchenko M., Savitskaya E., Koryagina E., Melnikova L., Karakozova M., Georgiev P. // Chromosoma. 2009. V. 118. № 5. P. 665–674.
174. Qian S., Varjavand B., Pirrotta V. // Genetics. 1992. V. 131. № 1. P. 79–90.
175. 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. № 7. P. 630–638.
176. Sanyal A., Lajoie B.R., Jain G., Dekker J. // Nature. 2012. V. 489. № 7414. P. 109–113.
177. Liu Z., Scannell D.R., Eisen M.B., Tjian R. // Cell. 2011. V. 146. № 5. P. 720–731.
178. Pena-Hernandez R., Marques M., Hilmi K., Zhao T., Saad A., Alaoui-Jamali M.A., del Rincon S.V., Ashworth T., Roy A.L., Emerson B.M., et al. // Proc. Natl. Acad. Sci. USA. 2015. V. 112. № 7. P. E677–686.
179. Han L., Lee D.H., Szabo P.E. // Mol. Cell. Biol. 2008. V. 28. № 3. P. 1124–1135.
180. Kurukuti S., Tiwari V.K., Tavoosidana G., Pugacheva E., Murrell A., Zhao Z., Lobanenkov V., Reik W., Ohlsson R. // Proc. Natl. Acad. Sci. USA. 2006. V. 103. № 28. P. 10684–10689.
181. Li T., Hu J.F., Qiu X., Ling J., Chen H., Wang S., Hou A., Vu T.H., Hoffman A.R. // Mol. Cell. Biol. 2008. V. 28. № 20. P. 6473–6482.
182. Dekker J., Mirny L. // Cell. 2016. V. 164. № 6. P. 1110–1121.
183. Gomez-Diaz E., Corces V.G. // Trends Cell Biol. 2014. V. 24. № 11. P. 703–711.
184. Cai H.N., Shen P. // Science. 2001. V. 291. № 5503. P. 493–495.
185. Kuhn E.J., Viering M.M., Rhodes K.M., Geyer P.K. // EMBO J. 2003. V. 22. № 10. P. 2463–2471.
186. Muravyova E., Golovnin A., Gracheva E., Parshikov A., Belenkaya T., Pirrotta V., Georgiev P. // Science. 2001. V. 291. № 5503. P. 495–498.
187. Kyrchanova O., Toshchakov S., Parshikov A., Georgiev P. // Mol. Cell. Biol. 2007. V. 27. № 8. P. 3035–3043.
188. Kyrchanova O., Toshchakov S., Podstreshnaya Y., Parshikov A., Georgiev P. // Mol. Cell. Biol. 2008. V. 28. № 12. P. 4188–4195.
189. Comet I., Savitskaya E., Schuettengruber B., Negre N., Lavrov S., Parshikov A., Juge F., Gracheva E., Georgiev P., Cavalli G. // Dev. Cell. 2006. V. 11. № 1. P. 117–124.
190. Comet I., Schuettengruber B., Sexton T., Cavalli G. // Proc. Natl. Acad. Sci. USA. 2011. V. 108. № 6. P. 2294–2299.
191. 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. № 3. P. 754–761.
192. Kravchenko E., Savitskaya E., Kravchuk O., Parshikov A, Georgiev P., Savitsky M. // Mol. Cell. Biol. 2005. V. 25. № 21. P. 9283–9291.
193. Kyrchanova O., Ivlieva T., Toshchakov S., Parshikov A., Maksimenko O., Georgiev P. // Nucl. Acids Res. 2011. V. 39. № 8. P. 3042–3052.
194. Fujioka M., Mistry H., Schedl P., Jaynes J.B. // PLoS Genet. 2016. V. 12(2). P. e1005889.
195. Chen H., Levo M., Barinov L., Fujioka M., Jaynes J.B., Gregor T. // Nat. Genet. 2018. V. 50. № 9. P. 1296–1303.
196. Kyrchanova O., Maksimenko O., Ibragimov A., Sokolov V., Postika N., Lukyanova M., Schedl P., Georgiev P. // Sci. Adv. 2020. V. 6. № 13. P. eaaz3152.
197. Kaye E.G., Kurbidaeva A., Wolle D., Aoki T., Schedl P., Larschan E. // Mol. Cell. Biol. 2017. V. 37. № 21. P. e00253-17.
198. Kyrchanova O., Sabirov M., Mogila V., Kurbidaeva A., Postika N., Maksimenko O., Schedl P., Georgiev P. // Proc. Natl. Acad. Sci. USA. 2019. V. 116. № 27. P. 13462–13467.
199. Krivega M., Savitskaya E., Krivega I., Karakozova M., Parshikov A., Golovnin A., Georgiev P. // Chromosoma. 2010. V. 119. № 4. P. 425–434.
200. Sexton T., Cavalli G. // Cell. 2015. V. 160. № 6. P. 1049–1059.
201. Sexton T., Yaffe E., Kenigsberg E., Bantignies F., Leblanc B., Hoichman M., Parrinello H., Tanay A., Cavalli G. // Cell. 2012. V. 148. № 3. P. 458–472.
202. Hons M.T., Huis In 't Veld P.J., Kaesler J., Rombaut P., Schleiffer A., Herzog F., Stark H., Peters J.M. // Nat. Commun. 2016. V. 7. P. 12523.
203. Fudenberg G., Imakaev M., Lu C., Goloborodko A., Abdennur N., Mirny L.A. // Cell Rep. 2016. V. 15. № 9. P. 2038–2049.
204. Haarhuis J.H.I., van der Weide R.H., Blomen V.A., Yanez-Cuna J.O., Amendola M., van Ruiten M.S., Krijger P.H.L., Teunissen H., Medema R.H., van Steensel B., et al. // Cell. 2017. V. 169. № 4. P. 693–707. e14.
205. Sanborn A.L., Rao S.S., Huang S.C., Durand N.C., Huntley M.H, Jewett A.I., Bochkov I. D., Chinnappan D., Cutkosky A., Li J., et al. // Proc. Natl. Acad. Sci. USA. 2015. V. 112. № 47. P. E6456-6465.
206. Schwarzer W., Abdennur N., Goloborodko A., Pekowska A., Fudenberg G., Loe-Mie Y., Fonseca N.A., Huber W., Haering C.H., Mirny L., et al. // Nature. 2017. V. 551. № 7678. P. 51–56.
207. Wutz G., Varnai C., Nagasaka K., Cisneros D.A., Stocsits R.R., Tang W., Schoenfelder S., Jessberger G., Muhar M., Hossain M.J., et al. // EMBO J. 2017. V. 36. № 24. P. 3573–3599.
208. Nishiyama T. // Curr. Opin. Cell. Biol. 2019. V. 58. P. 8–14.
209. Rodriguez-Carballo E., Lopez-Delisle L., Zhan Y., Fabre P.J., Beccari L., El-Idrissi I., Huynh T.H.N., Ozadam H., Dekker J., Duboule D. // Genes Dev. 2017. V. 31. № 22. P. 2264–2281.

210. Narendra V., Bulajic M., Dekker J., Mazzoni E.O., Reinberg D. // Genes Dev. 2016. V. 30. № 24. P. 2657–2662.
211. Hug C.B., Grimaldi A.G., Kruse K., Vaquerizas J.M. // Cell. 2017. V. 169. № 2. P. 216–228 e19.
212. Ramirez F., Bhardwaj V., Arrigoni L., Lam K.C., Gruning B.A., Villaveces J., Habermann B., Akhtar A., Manke T. // Nat. Commun. 2018. V. 9. № 1. P. 189.
213. Ulianov S.V., Khrameeva E.E., Gavrilov A.A., Flyamer I.M., Kos P., Mikhaleva E.A., Penin A.A., Logacheva M.D., Imakaev M.V., Chertovich A., et al. // Genome Res. 2016. V. 26. № 1. P. 70–84.
214. Phillips-Cremins J.E., Corces V.G. // Mol. Cell. 2013. V. 50. № 4. P. 461–474.
215. Cattoni D.I., Cardozo Gizzi A.M., Georgieva M., Di Stefano M., Valeri A., Chamousset D., Houbron C., Dejardin S., Fiche J.B., Gonzalez I., et al. // Nat. Commun. 2017. V. 8. № 1. P. 1753.
216. Flyamer I.M., Gassler J., Imakaev M., Brandao H.B., Ulianov S.V., Abdennur N., Razin S.V., Mirny L.A., Tachibana-Konwalski K. // Nature. 2017. V. 544. № 7648. P. 110–114.
217. Mateo L.J., Murphy S.E., Hafner A., Cinquini I.S., Walker C.A., Boettiger A.N. // Nature. 2019. V. 568. № 7750. P. 49–54.
218. Tan L., Xing D., Chang C.H., Li H., Xie X.S. // Science. 2018. V. 361. № 6405. P. 924–928.
219. Finn E.H., Pegoraro G., Brandao H.B., Valton A.L., Oomen M.E., Dekker J., Mirny L., Misteli T. // Cell. 2019. V. 176. № 6. P. 1502–1515 e10.
220. Luppino J.M., Park D.S., Nguyen S.C., Lan Y., Xu Z., Yunker R., Joyce E.F. // Nat. Genet. 2020. V. 52. № 8. P. 840–848.
221. Hansen A.S., Pustova I., Cattoglio C., Tjian R., Darzacq X. // Elife. 2017. V. 6. P. e25776.
222. Maeda R.K., Karch F. // Chromosoma. 2015. V. 124. № 3. P. 293–307.
223. Kyrchanova O., Mogila V., Wolle D., Magbanua J.P., White R., Georgiev P., Schedl P. // Mech. Dev. 2015. V. 138. № 2. P. 122–132.
224. Kyrchanova O., Wolle D., Sabirov M., Kurbidaeva A., Aoki T., Maksimenko O., Kyrchanova M., Georgiev P., Schedl P. // Genetics. 2019. V. 213. № 3. P. 865–876.
225. Postika N., Metzler M., Affolter M., Muller M., Schedl P., Georgiev P., Kyrchanova O. // PLoS Genet. 2018. V. 14. № 12. P. e1007702.
226. Kyrchanova O., Mogila V., Wolle D., Deshpande G., Parshikov A., Cleard F., Karch F., Schedl P., Georgiev P. // PLoS Genet. 2016. V. 12. № 7. P. e1006188.
227. Rao S.S.P., Huang S.C., Glenn St Hilaire B., Engreitz J.M., Perez E.M., Kieffer-Kwon K.R., Sanborn A.L., Johnstone S.E., Bascom G.D., Bochkov I.D., et al. // Cell. 2017. V. 171. № 2. P. 305–320 e24.
228. Li M., Ma Z., Roy S., Patel S.K., Lane D.C., Duffy C.R., Cai H.N. // Sci. Rep. 2018. V. 8. № 1. P. 15158.
229. Yokoshi M., Segawa K., Fukaya T. // Mol. Cell. 2020. V. 78. № 2. P. 224–235 e5.
230. Stadler M.R., Haines J.E., Eisen M.B. // Elife. 2017. V. 6. P. e29550.
231. Despang A., Schopflin R., Franke M., Ali S., Jerkovic I., Paliou C., Chan W.L., Timmermann B., Wittler L., Vingron M., et al. // Nat. Genet. 2019. V. 51. № 8. P. 1263–1271.
232. Ghavi-Helm Y., Jankowski A., Meiers S., Viales R.R., Korbel J.O., Furlong E.E.M. // Nat. Genet. 2019. V. 51. № 8. P. 1272–1282.

Modification of Nuclear Compartments and the 3D Genome in the Course of a Viral Infection

S. V. Razin*, A. A. Gavrilov, O. V. Iarovaia
Institute of Gene Biology Russian Academy of Sciences
*E-mail: sergey.v.razin@usa.net
Received June 04, 2020; in final form, July 07, 2020
DOI: 10.32607/actanaturae.11041
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **The review addresses the question of how the structural and functional compartmentalization of the cell nucleus and the 3D organization of the cellular genome are modified during the infection of cells with various viruses. Particular attention is paid to the role of the introduced changes in the implementation of the viral strategy to evade the antiviral defense systems and provide conditions for viral replication. The discussion focuses on viruses replicating in the cell nucleus. Cytoplasmic viruses are mentioned in cases when a significant reorganization of the nuclear compartments or the 3D genome structure occurs during an infection with these viruses.**
KEYWORDS **nuclear compartmentalization, spatial organization of the eukaryotic genome, virus infection.**

INTRODUCTION

To date, there is little doubt that structural and functional compartmentalization of the cell nucleus plays an important role in the functioning of the genetic machinery. Moreover, the genome itself is a structural platform for nuclear compartmentalization [1]. Individual chromosomes occupy limited spaces within the nucleus, which are referred to as chromosome territories [2–4]. Although relatively isolated, chromosome territories form numerous interchromosomal contacts. In addition, they attach to the nuclear lamina and nucleolus, thus forming a single chromatin domain. This domain is permeated by interchromatin channels, which together constitute the interchromatin compartment [2–6]. Various functional centers, such as the nucleolus, Cajal bodies, PML bodies, speckles, and transcription factories, are located inside this compartment [1, 5, 6]. Although these functional centers, many of which are also called nuclear bodies, are located in the interchromatin compartment, it is wrong to assume that they lack DNA. DNA is found in transcription factories located in the so-called perichromatin region lining interchromatin channels [5, 6]. The nucleolus is a special form of transcription factory located around clusters of ribosomal genes [7]. Speckles and Cajal bodies are reaction centers in which post-transcriptional RNA modification takes place and the necessary enzymes accumulate [8–10]. DNA is not an integral part of these functional compartments. However, there is ample evidence that genes can be recruited to them during the processing of various RNAs [11–13].

The highest levels of spatial organization of the genome in the cell nucleus are as follows: (i) spatial segregation of active (A) and inactive (B) genomic compartments [14]; (ii) separation of chromosomes into partially insulated topologically associating domains (TADs) [15–17], which in many cases limit the areas of enhancer action [18–20]; and (iii) the establishment of spatial interactions between distant genomic elements by looping of the segments of the chromatin fiber separating them [21]. The functional significance of these spatial contacts may vary. In mammalian cells, contacts between the convergent binding sites of the insulator protein CTCF separate TADs [22]. Spatial contacts between enhancers and promoters (enhancer-promoter loops) ensure communication between these regulatory elements [23]. Changes in the spatial organization of the genome, including those resulting from chromosomal rearrangements and loss of CTCF-binding sites, alter the transcription profiles. In some cases, these changes cause cancer and other diseases [18, 24–28].

As mentioned above, the packed genome is a platform for structural and functional compartmentalization of the cell nucleus. However, the opposite is also true. The interaction between certain genomic regions and functional nuclear compartments supports the 3D organization of the genome. Thus, spatial segregation of the A and B genomic compartments is due to the re-

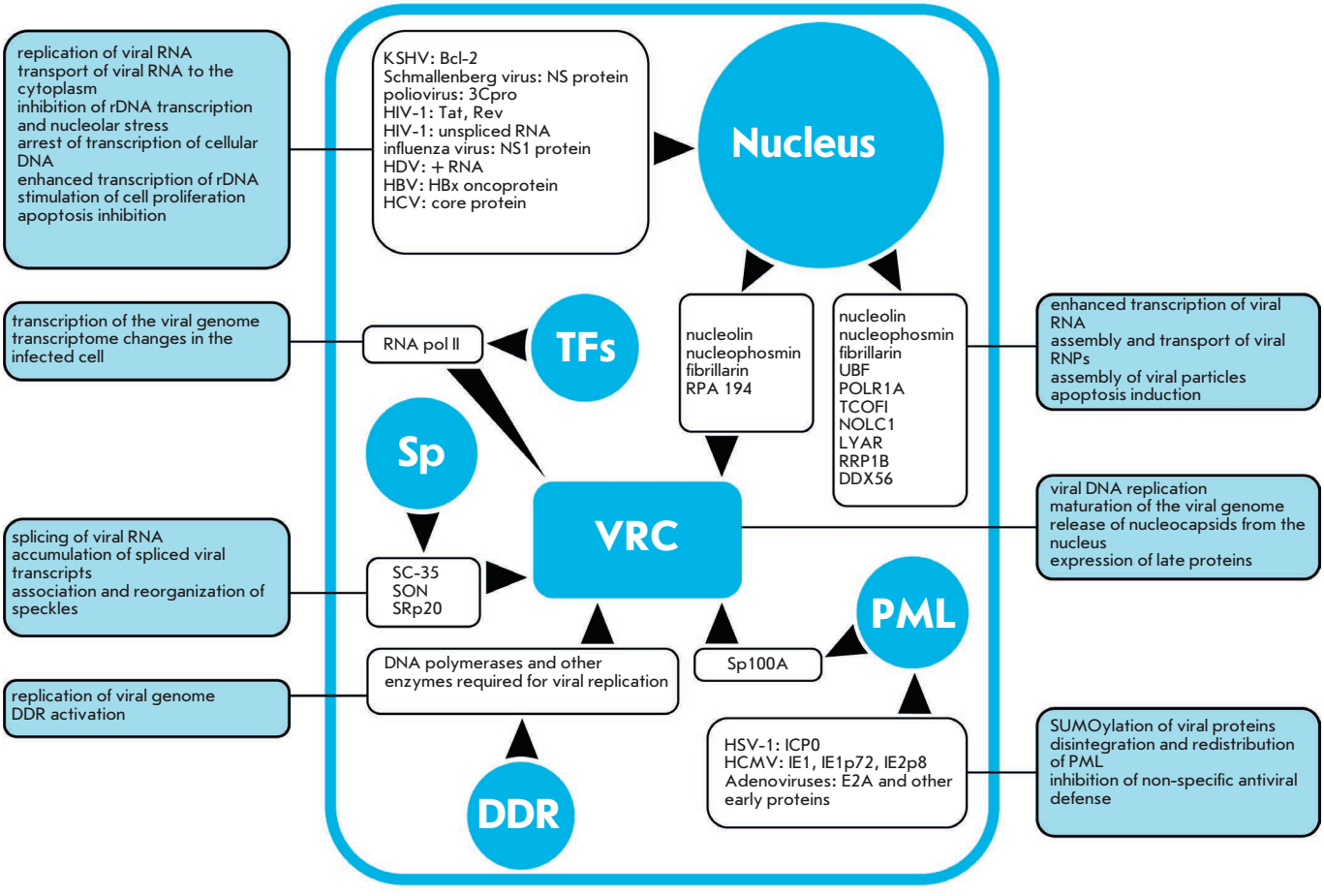


Fig. 1. Scheme of movement of cellular and viral proteins/nucleic acids between nuclear compartments during infection. Blue circles indicate nuclear compartments: the nucleolus, transcription factories (TFs), speckles (Sp), promyelocytic leukemia (PML) bodies, DNA damage repair (DDR) foci, and viral replication centers (VRCs). Within the nucleus there are viral/cellular proteins and nucleic acids that move during the infectious process. Directions of movement are marked with black arrows. The rectangles with rounded corners contain information on the effects on cellular and viral metabolism associated with the movement of proteins/nucleic acids to/from the corresponding compartment during the infections process

cruitment of active genes to speckles and the relocation of repressed genes to the nucleolus and nuclear lamina [13, 29–31]. Recruitment of various genes to Cajal bodies and common transcription factories facilitates the establishment of spatial contacts between the distant regions of the genome, as well as between different chromosomes [11, 32–36].

Viruses replicating in the cell nucleus exploit cellular systems during the infectious process. Although the features of the infectious process differ significantly for different viruses and depend on the type of infection (lytic/latent), it is apparent that viruses must adapt functional compartmentalization of the nucleus to suit their needs. Although the interaction between a virus and the host cell has been studied for decades, this aspect of the problem has not yet received enough of researchers' attention. In this review, an attempt is made to summarize current knowledge on how viruses modify the nuclear compartments and the 3D organization of the cell genome. Although our discussion mainly focuses on the viruses replicating in the cell nucleus, we will also mention cytoplasmic viruses, which somehow cause reorganization of either nuclear compartments or the 3D genome upon infection.

REORGANIZATION AND REPROFILING OF PRE-EXISTING NUCLEAR COMPARTMENTS DURING A VIRAL INFECTION

Many nuclear compartments are modified during a viral infection (*Fig. 1*). These modifications happen because viruses need to either suppress the cellular antiviral defense or use the enzymes that have accumulated in the compartments for their replication. Viruses

control the reorganization of nuclear compartments by either penetrating these compartments or directing in them proteins encoded by the viral genome (*Fig. 1*). Although viruses also interact with other nuclear compartments, the process of viral interaction with the nucleolus and PML bodies has been the most thoroughly studied. Along with this, new compartments assemble in the nuclei in which viruses replicate. All of these processes are discussed in more detail below.

Nucleolus

The nucleolus is the most recognizable functional compartment of the cell nucleus. The main function of the nucleolus is ribosome biogenesis. However, the nucleolus also has a series of other, so-called non-canonical, functions. It acts as a site for the sequestration of various proteins and participates in cell cycle regulation, response to stress, organization of the repressive genome compartment, and in a number of other functional processes [37]. Thus, it is not surprising that viruses interact closely with the nucleolus during an infection. This applies to both the viruses replicating in the cell nucleus and those replicating in the cytoplasm. The result of the interaction mediated by the transfer of various viral proteins to the nucleolus can be either complete/partial disintegration of the nucleolus, relocalization of nucleolar proteins to the nucleoplasm and cytoplasm, or relocation of nucleoplasmic proteins to the nucleolus [38–42].

Early studies have shown that the effectiveness of the infectious process directly depends on the interaction between the virus and the nucleolus [43–45]. With the development of proteomics, more complete data on the spectrum of viral and nucleolar proteins that interact with each other have been obtained [41, 46–49]. Experiments comparing the proteome of nucleoli isolated from healthy cells and cells infected with adenovirus suggest that movement from the nucleolus or into the nucleolus involves a very wide range of proteins [39–41, 50, 51]. Typical nucleolar proteins are relocated to the viral replication centers (see section 3), the nucleoplasm, and the cytoplasm. Both viral proteins and a number of cellular proteins move to the nucleolus. However, the consequences of this relocation are not always clear. The interactions between viruses and the nucleolus result from the superposition of two diametrically opposed processes: (1) cellular antiviral strategy and (2) viral strategy aimed at evading the antiviral response and maximizing the use of available cellular resources for its own purposes.

The role of nucleolin in antiviral protection has been rather fully characterized. However, it remains unclear whether the release of nucleolin from the nucleolus correlates with the implementation of its antiviral properties. Moreover, in addition to the nuclear protein nucleolin, which mainly resides in the nucleolus, the cell contains cytoplasmic nucleolin and plasma membrane-associated nucleolin [52, 53]. In some cases, it remains unclear which pool of nucleolin is used in antiviral defense. When cells are infected with a highly pathogenic strain H5N1 of the influenza virus, nucleolin expression inhibition significantly increases the activity of viral polymerase. It also enhances the synthesis of viral mRNA, as well as apoptosis and necrosis of the host cell. On the contrary, overexpression of nucleolin decreases infection intensity [54]. Antiviral activity of nucleolin has also been demonstrated in the infection of cells with the goat plague virus (peste des petits ruminants virus, PPRV). This activity is associated with the induction of the host interferon response [55]. Binding of nucleolin to G-quadruplexes in viral RNA [56] and DNA [57] inhibits the viral functions, apparently by blocking the promoters [57].

Apoptosis induction in infected cells is considered one of the mechanisms of the body's defense against an infection. In this context, it is worth mentioning that one of the elements of the host antiviral defense is sequestration of viral anti-apoptotic factors in the nucleolus and the release of cellular pro-apoptotic factors from the nucleolus. For instance, the PICT-1 protein binds to the apoptosis inhibitor KS-Bcl-2 of Kaposi's sarcoma-associated herpesvirus (KSHV) and inhibits its anti-apoptotic activity by sequestering KS-Bcl-2 in the nucleolus [58].

The specific mechanisms of induction of nucleolar stress and apoptosis upon penetration of the virus into the cell and the possibilities of reprofiling of these processes for viral reproduction are not always clear. There are many studies demonstrating the complex nature of the interaction between viral proteins and nucleolar components. For instance, the NS protein of the Schmallenberg virus induces a disruption of the nucleolus and relocalization of nucleophosmin from the nucleolus to the nucleoplasm [59]. Poliovirus protease $3C^{pro}$, which is targeted to the nucleolus, modifies UBF and SL1 involved in rDNA transcription and cleaves the transcription factor TAF110, thus inhibiting the synthesis of ribosomal RNA (rRNA) [60]. Precursors of the human rhinovirus 16 protease $3C^{pro}$ co-localize with nucleophosmin in the nucleolus. This triggers the cleavage of the OCT-1 transcription factor and complete arrest of the transcription of cellular DNA [61]. The human immunodeficiency virus (HIV) protein Tat interacts with fibrillarin and U3 small nucleolar RNA (snoRNA), resulting in impaired rRNA maturation [62]. The NS1 protein of the influenza virus H3N2 interacts with NOLC1, which regulates rDNA transcription by binding to the large subunit of RNA polymerase. This

interaction reduces NOLC1 levels, which leads to apoptosis [63]. Association of the same protein with nucleolin causes hypermethylation of the UCE (upstream control element) of rRNA genes, arrest of rRNA synthesis, and subsequent nucleolar stress [64]. The opposite process, such as the activation of rRNA gene transcription, can be observed when cells are infected with other viruses and an alternative course of the infection (latent infection) takes place. The core protein of the hepatitis C virus binds to nucleophosmin and relocates to the nucleolus, where it interacts with UBF and RNA polymerase I. This interaction enhances the association of these factors with the rRNA gene promoters and increases the level of rRNA transcription. The nucleolus grows in size and moves to the periphery of the nucleus [65]. The HBx oncoprotein of the hepatitis B virus acts in a similar way. HBx is transported to the nucleolus by nucleophosmin and acetylates nucleophosmin, which results in depletion of histones from the rDNA promoters. This, in turn, enhances the transcriptional activity of the nucleolus and the proliferative activity of the cell [66]. In combination with other mechanisms of proliferation control [67], chronic infection leads to cell transformation. The significance of all these observations in the context of viral strategy and the mechanisms of antiviral defense are yet to be elucidated.

Along with evading the antiviral response, viruses actively exploit the proteins sequestered in the nucleolus for their own purposes. In some cases, viruses also use the nucleolus as a compartment partially isolated from the nucleoplasm. During the infection, proteins of the nucleolus can be directly adopted for replication and transcription of viral nucleic acids, as well as the assembly of viral particles. Viruses with a negative-strand RNA genome (influenza virus, Thogotovirus, and Borna disease virus) replicate genomic RNA in the nucleus and closely interact with the nucleolus. Early studies showed that the Borna disease virus uses the nucleolus as a replication site [68]. The positive strand of the hepatitis delta virus RNA is transcribed in the nucleolus, while the negative strand is synthesized in the nucleoplasm [69]. Such segregation allows the virus to exploit the transcriptional machinery and compartmentalization of the host cell nucleus to its maximum efficiency. In the case of a human immunodeficiency virus (HIV-1) infection, the nucleolus is the site of assembly of the complexes providing transport of unspliced and partially spliced viral RNAs to the cytoplasm. Unspliced HIV-1 RNA acts as both genomic RNA and mRNA for the synthesis of Gag and Gag-Pol proteins. Incompletely spliced RNAs act as mRNA for the synthesis of the Vif, Vpr, Tat, Vpu, and Env proteins. Fully spliced RNAs are mRNA templates for the synthesis of the Vpr, Tat, Rev, and Nef proteins. Unspliced and incompletely spliced HIV-1 RNAs are unstable and rapidly degrade in the nucleus. The Rev protein protects these RNAs from degradation and ensures their transport to the cytoplasm. Such an intricate transport complex is formed in the nucleolus to which unspliced and partially spliced HIV-1 RNAs are relocated. Rev is synthesized in the cytoplasm from a spliced RNA and contains signals of nuclear and nucleolar localization. After being transported to the nucleus, Rev associates with nucleoporins Nup98 and Nup214, as well as with the exportin CRM1. The resulting complex is then transported to the nucleolus [70–72], where Rev multimerizes and binds to specific RRE sequences in the viral RNA [73]. Thus, in the course of an infection, the virus uses both the host cell proteins and the nucleolus as a "staging post" and a platform for the assembly of viral RNPs.

However, a more common phenomenon is the virus-induced relocalization of nucleolar proteins to the nucleoplasm with their further use for viral replication. Viral replication compartments (see section 3 and *Fig. 1*) contain various nucleolar proteins: nucleophosmin, nucleolin, fibrillarin, UBF, Nopp140, POLR1A, TCOFI, and NOLC1 [74–76]. The structure and protein composition of the nucleolus are significantly altered in cells infected with herpes viruses (HSV-1 and HCMV) [38]. The three main nucleolar proteins, namely, nucleolin, nucleophosmin, and fibrillarin, as well as RPA194, move to the virus replication compartments. There, they participate in the replication, transcription, and assembly of viral particles. A number of studies have shown that nucleolin is involved in the formation of the replication compartments of various herpes viruses [38, 42]. In combination with the viral nuclease UL12, nucleolin is responsible for the maturation of the viral genome and nucleocapsid release from the nucleus [77, 78]. In a cytomegalovirus infection, association of nucleolin with the viral DNA polymerase component UL44 is necessary for efficient DNA replication and the expression of late proteins [79].

In an infection with the influenza virus, accumulation of the multifunctional viral protein NS1 in the nucleolus is accompanied by the delocalization of nucleolin to the nuclear periphery and redistribution of fibrillarin [80]. Nucleolin is believed to ensure the transport of ribonucleoprotein complexes and participate in viral RNA replication. The nucleolar protein RRP1B, which is involved in ribosome biogenesis, relocates from the nucleolus to the nucleoplasm. There, it associates with RNA-dependent RNA polymerase, thus enhancing the transcription of viral RNA [81]. One of the multifunctional nucleolar proteins, LYAR, moves to the nucleoplasm and cytoplasm from the nucleolus

and facilitates the assembly of the ribonucleoprotein complexes of the influenza A virus [82].

Summarizing the above mentioned, one can conclude that viruses can both directly affect the ribosomal gene transcription machinery and modify the protein composition of the nucleoli, as well as use the nucleolus as a safe site for the biogenesis of new viral particles. Thus, a viral infection can affect the homeostasis of the nucleolus, as well as its morphology and compartmentalization. This, in turn, can be used to implement the most effective strategies for pathogen survival and reproduction.

Repair foci

Repair foci (DDR foci, DNA damage response) are exploited by many viruses as a source of enzymes for viral replication. These viruses include various parvoviruses, and MVM in particular. After penetrating the cell nucleus, MVM DNA preferentially localizes near the damaged regions of the cellular genome, which are associated with phosphorylated histone H2AX and repair factors [83, 84]. Viral replication centers form near the DDR foci. These centers recruit the DNA polymerases present in the DDR foci and other enzymes involved in viral replication. In the course of the infection, the number of pre-existing DDR foci proves insufficient for the assembly of new viral replication centers. For this reason, the virus stimulates the introduction of new DNA lesions, thus increasing the number of DNA repair foci to be exploited by the virus [84, 85]. Other parvoviruses apparently use a similar mechanism [86–88]. DDR activation is also typical of infection with viruses belonging to some other families [89, 90]. For instance, it has been established that, after penetration of the cell, human papillomavirus localizes at chromosomal fragile sites [91].

Transcription factories, speckles, and paraspeckles

Transcription of the genes of DNA viruses is carried out by cellular RNA polymerase II. A significant part of the RNA pol II molecules are sequestered in transcription factories [11, 32, 35, 36, 92–94]. It remains unclear what transcription factories are. According to some data, stable clusters of RNA polymerases are present in the cell regardless of active transcription. There also exists a different point of view, according to which initiated transcription complexes are assembled into clusters (see [35] for a review). In any case transcription factories are associated with the active compartment of the genome. Most viruses entering the cell nucleus preferentially interact with this very genomic compartment. Virus replication centers are assembled at subsequent stages of the infection (see section 3). It is not entirely clear whether these centers capture pre-existing transcription factories or free RNA polymerase relocates to them as the transcription factories disintegrate. A significant part of the pre-existing transcription factories are ultimately lost, while RNA polymerase II accumulates in the centers of viral replication/transcription [95–98].

Speckles are compartments where the splicing machinery is located [8, 9]. However, there is no clear information on whether these compartments simply offer storage sites for splicing factors, which are recruited to transcription sites as required, or whether splicing can occur directly in speckles [99, 100]. A viral infection leads to speckle reorganization [101–103]. The early stages of lytic infection are characterized by the redistribution of splicing factors (SC35, SON, SRp20, etc.) to the centers of viral replication/transcription [102–105] (see section 3 and *Fig. 1*). At the later stages of a lytic infection, speckles combine into larger compartments. Spliced viral transcripts can be found in these compartments [106, 107]. Fusion of speckles into larger compartments is typical of the cellular response to various stresses, including a virus infection [108, 109]. The fact that spliced transcripts concentrate in speckles at late stages of an infection suggests that accumulation of these transcripts is one of the stages in their transport to the cytoplasm [106]. A completely different picture emerges for the infection of permissive cells by the influenza virus. Splicing of one of the viral RNAs takes place in speckles [110].

In many cells, small compartments formed on the basis of non-coding RNA NEAT1 are localized next to speckles. These compartments are called paraspeckles [111]. The functions of paraspeckles are not entirely clear. They include sequestration of the RNA-editing adenosine deaminase and stress response [111–113]. The level of NEAT1 RNA and the number of paraspeckles increase significantly in case of a virus infection [114–117]. Apparently, this occurs due to the activation of the innate immunity, since NEAT1 RNA binds a repressor that inhibits transcription of genes encoding several cytokines, including interleukin-8 [114, 118]. However, one of the studies reported that the herpes simplex virus (HSV-1) adopts the proteins sequestered in paraspeckles for its replication [117]. The research has demonstrated that, during a lytic infection, the HSV-1 genome is localized in paraspeckles and that suppression of NEAT1 reduces the production of viral particles.

PML bodies

It has long been known that, at the initial stages of a viral infection, virus-specific proteins are recruited to PML bodies to stimulate their disintegration [119–123]. PML bodies contain numerous proteins. The most char-

acteristic components among them are PML, hDaxx, ATRX, and Sp100. All these proteins play an important role in non-specific antiviral immunity [124–127], which the virus must inactivate. Different viruses solve this problem in different ways. For instance, the HSV-1 ICP0 protein targeted to PML bodies is a ubiquitin ligase that selectively ubiquitinates SUMOylated proteins, including PML and Sp100. Such modification of the proteins stimulates their degradation by the proteasome system [128, 129]. The cytomegalovirus early protein IE1 suppresses PML SUMOylation, which is critical for the formation of PML bodies [130]. In both cases, the final result is the disintegration of PML bodies. Adenovirus early proteins also relocate to PML bodies and cause DAXX degradation and PML redistribution [131–133]. Disintegration of PML bodies also occurs during lytic infection of cells by other DNA viruses [134].

It should be noted that, after entering into the nucleus, the genomes of many viruses localize next to the PML bodies [135, 136]. The reasons why this happens are not entirely clear. It is also unclear whether viral genomes are transferred to the pre-existing PML bodies, or new PML bodies are formed close to the viral genomes [137, 138]. In the latter case, the assembly of PML bodies next to the viral genomes can be one of the stages of antiviral defense. The situation can be even more complicated. The virus may require a number of proteins sequestered in PML bodies, including the ubiquitination machinery. It has recently been shown that the adenovirus DNA-binding protein E2A is SUMOylated by the enzymatic machinery of the host cell and recruits the transcription factor Sp100A to viral replication centers. Sp100A is released from PML bodies after PML redistribution from bodies to tracks induced by another viral protein (E4orf3) [139]. Human cytomegalovirus proteins IE1p72 and IE2p86 are transiently localized in PML bodies, where they are SUMOylated [140].

ASSEMBLY OF NEW COMPARTMENTS: VIRAL REPLICATION CENTERS

A characteristic feature of a lytic infection with DNA viruses is the formation of a new type of functional compartments in the cell nuclei: viral replication centers (VRCs). These centers are assembled around individual viral genomes that have penetrated the cell nucleus and serve as sites of transcription and clonal replication of viral DNA [74, 141]. At the late stages of the infection, each VRC contains numerous copies of viral DNA. All these copies are replicas of the original viral DNA molecule around which the VRC is assembled [142–144]. Furthermore, areas of active replication and transcription within the VRC can be spatially segregated [145]. The protein composition of VRC is rather complex; it includes both virus-specific and cellular components [74, 141]. The latter include mainly DNA replication enzymes, RNA polymerase II, components of the transcription machinery, a wide range of repair enzymes, and chromatin remodeling factors [49, 146, 147].

The following question still remains open: what does ensure the maintenance of the VRC structure? In recent years, there has been abundant evidence that the process called liquid–liquid phase separation plays an important role in the assembly of functional nuclear compartments [148]. Separation of a compartment into a distinct phase is provided by multiple interactions between unstructured protein domains, namely, the intrinsically disordered regions (IDRs), which are present in this compartment [149]. It is worth mentioning that IDRs are present in many virus-specific proteins, including early proteins, which play a key role in the reprogramming of cellular metabolism, PML body disintegration, and VRC assembly [150–153]. The distinctive features of IDRs include their ability to interact with a large number of different partners, thus providing a platform for the assembly of functional compartments [151]. VRCs can fuse [107, 154], which is typical of liquid condensates. On the other hand, a recent study has shown that the VRCs of the herpes simplex virus are not disrupted by 1,6-hexanediol (an agent suppressing phase separation) [155]. In addition, the kinetics of the exchange of RNA polymerase II between VRC and nucleoplasm does not correspond to that expected for liquid condensates [155]. The authors suggest that nucleosome-free viral DNA serves as a platform for recruiting RNA pol II and a number of other DNA-binding proteins to VRCs. They also believe that VRCs are not typical liquid condensates, although liquid–liquid phase separation may play a certain role at the stage of their formation [155].

At least for the herpesvirus infection, it has been shown that VRCs can change their location within the cell nucleus. During the late stages of the infection they can get fused, which makes recombination between the viral genomes replicated in different VRCs possible [144]. Relocation of VRC within the nucleus is an active process, since it is suppressed by actin and myosin inhibitors. VRCs approach speckles as a result of directed relocation. This, apparently, facilitates the splicing of viral transcripts [107]. It was also shown that, during the lytic Epstein–Barr virus infection, the proteins SC35, SON, SRp20, as well as some other splicing machinery components, relocate from speckles to specific structures on the VRC surface [104]. Thus, the strategies for splicing of viral transcripts may vary for different herpes viruses.

MODIFICATION OF THE 3D GENOME IN A LYTIC AND LATENT INFECTION AND VIRAL GENOME INTEGRATION

Lytic infection: preferential association of viruses with the A compartment of the genome and an expansion of the A compartment during the later stages of the infection

In recent years, a number of studies have focused on the potential existence of regions in the host cell genome with which the virus preferentially interacts at various stages of the lytic infection. All these studies used the approaches based on the ligation of spatially proximal DNA fragments in fixed nuclei (the so-called C methods [156, 157]). By using experimental protocols that allow for the identification of the entire range of contacts between the viral and the host cell genomes, it was shown that viruses preferentially contact the active (A) genomic compartment during a lytic infection [158, 159]. Within the A compartment, adenoviruses preferentially come into contact with any promoters or enhancers [159] while the hepatitis B virus interacts with CpG islands [158]. The Epstein–Barr virus was shown to preferentially come into contact with inactive chromatin during a latent infection [160, 161] and relocate to the active chromatin compartment after induction of viral replication [161]. Association with active chromatin is also characteristic of the influenza virus, which is an RNA virus that replicates in the cell nucleus [162]. The expansion of the A compartment is stimulated by this virus and adenoviruses during a lytic infection. The mechanism of this phenomenon has been revealed for the influenza virus. The virus-specific NS1 protein prevents termination of the transcription of cellular genes at polyadenylation sites. As a result, transcription continues for significant distances beyond the gene (sometimes more than 100 kb). The authors showed that active RNA polymerase promotes cohesin removal from the CTCF-binding sites, thus leading to the loss of chromatin loops and significantly changing the genomic configuration. In addition, the enzymes associated with transcribing RNA polymerase can promote chromatin remodeling by removing repressive marks [162]. The benefits of expanding the A compartment for the virus remain to be explored. Profound inhibition of transcription termination at the gene termini also occurs in a lytic infection caused by the herpes simplex virus [163, 164]. Active chromatin is expanded to the previously inactive regions. However, it is still difficult to draw a conclusion as to how significant expansion of the active chromatin compartment in a herpesvirus infection is. This is because the effect of the infection on genome compartmentalization has not been studied yet for this virus using the Hi-C method.

Modification of the 3D genome of the host cell during a latent infection guided by viral transcription factors

As mentioned above, the Epstein–Barr virus can both cause a lytic infection and reside in cells in latent form as a circular episome associated with chromatin. There are several types of latent infections. They differ in the range of expressed viral proteins [165]. A latent infection with the Epstein–Barr virus is associated with various oncological diseases [166, 167]. For this reason, the mechanisms of epigenetic reprogramming by virus-specific proteins and microRNAs are being intensively studied. The virus-specific protein EBNA2 was shown to associate with enhancers and to modulate the expression of cellular genes by reconfiguring the spatial organization of the genome [168] (*Fig. 2A*). More specifically, EBNA2 activates the transcription of a number of genes, including *c-myc*, by stimulating the emergence of new enhancer-promoter loops [168, 169]. Activation of *c-myc* transcription leads to cell transformation. As a result of such transformation, the cells acquire the ability to unlimitedly proliferate. EBNA3A,C initiate the repression of a specific group of genes, including pro-apoptotic ones. These virus-specific proteins also bind preferentially to enhancer elements [169, 170]. In a number of cases, they prevent the establishment of enhancer-promoter contacts (anti-looping) (*Fig. 2A*). In other cases, EBNA3A,C initiate the assembly of repressive chromatin hubs. These repressive hubs form by recruiting Polycomb repressive complexes [169, 171].

The HIV-1 transcriptional regulator Tat can penetrate any cells via the cell penetration domain (CPD) [172]. Tat is secreted into the blood by T lymphocytes infected with HIV-1 and, once it has entered human B cells, it changes the mutual positions of several genes within the nucleus [173].

It remains difficult to say how widespread the mechanisms of 3D genome reorganization by viral transcriptional regulators are. This issue definitely deserves further study.

Modification of the 3D genome during integration of viral DNA into the host cell genome

The problem of insertional mutagenesis caused by the integration of retroviruses into the genome of the host cell is widely being discussed [174–178]. The discussion typically centers on the damage to the genes or the stimulation of the transcription of the cellular genes that have fallen under the control of viral promoters and enhancers [177, 179]. We suggest considering this issue in the context of the 3D genome organization.

First of all, it is worth mentioning that, after integration in the genome, viruses can use the pre-existing genomic architecture to activate the transcription of

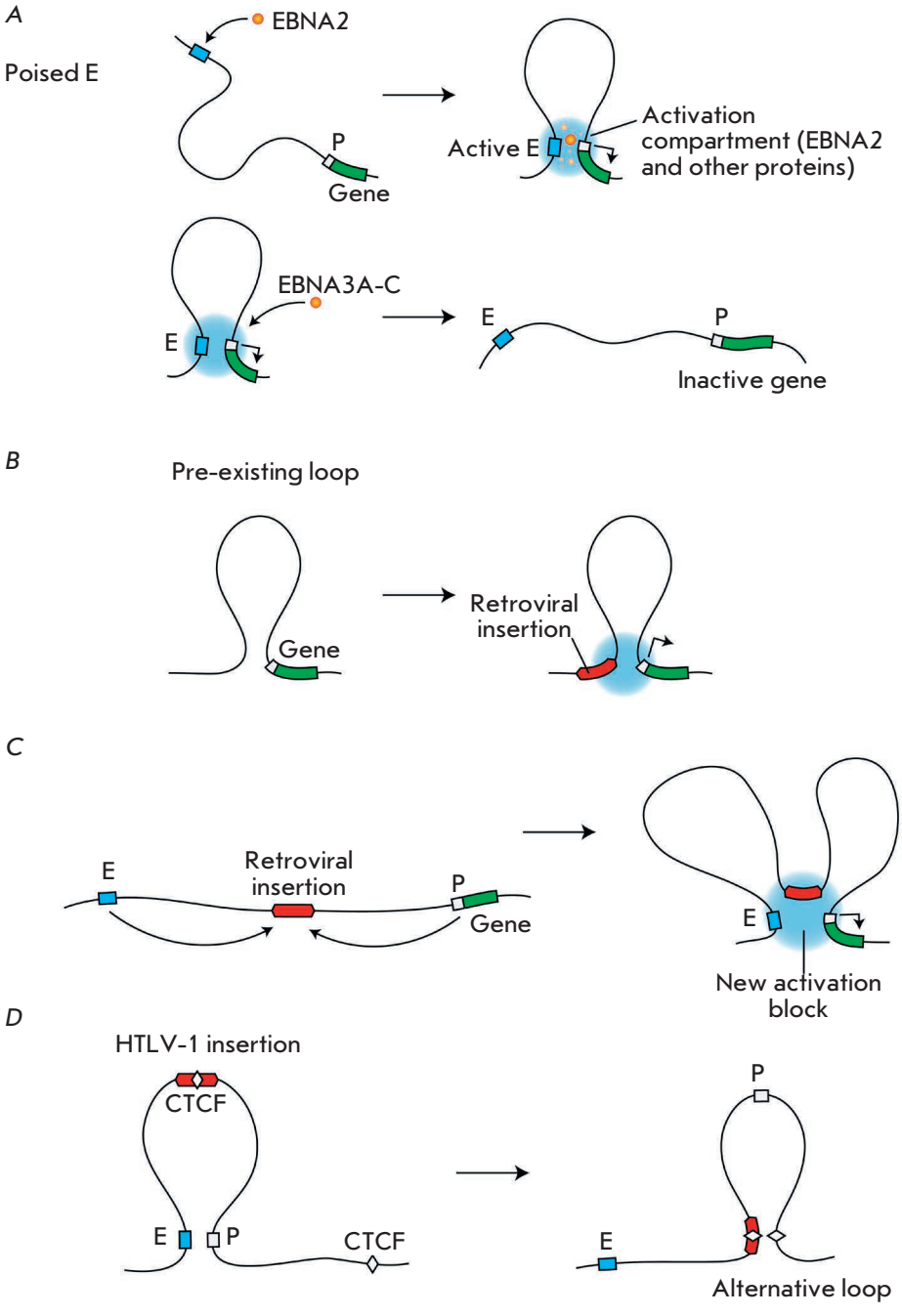


Fig. 2. Virus-induced reorganization of the 3D genome. (*A*) Induction (top) or destruction (bottom) of the promoter-enhancer contacts triggered by the viral proteins belonging to the EBNA family with the concomitant activation or repression of the host gene. (*B*) Involvement of the pre-existing genome architecture in the activation of a gene located at a considerable distance from the site of retroviral integration into the genome. (*C*) Formation of a new activator unit via the recruitment of the enchanter and promoter to the site of retroviral integration followed by activation of the host gene transcription. (*D*) Disruption of the promoter-enhancer communication as a result of the introduction of the CTCF-binding site and formation of an alternative loop. P – promoter; E – enhancer

the host's distant genes [180] (*Fig. 2B*). This mechanism has been shown, in particular, in the activation of the cyclin D1 (*Ccnd1*) gene by retroviruses integrated into the genome at a considerable distance (100 and 170 kb) upstream of this gene [180]. Activation of the *c-myb* oncogene by the mouse leukemia viruses (MLVs) integrated into the genome at considerable distances from the promoter of this oncogene is carried out in a similar manner [181]. Studies performed using genome-wide methods of analysis have shown that the preferred sites of genomic integration of various retroviruses causing tumors in mice (the so-called common insertion sites [182]) co-localize with various oncogenes within the nuclear space; i.e., in a 3D genome [183].

However, retroviruses not only exploit the pre-existing 3D organization of the genome, but they also trigger its reconfiguration (*Fig. 2C*). Thus, the occurrence of a *de novo* activator complex has been shown in HeLa cells, which carry multiple copies of the human papillomavirus (HPV) in their genome. This complex contains the *c-myc* promoter, a fragment of the HPV genome integrated at a distance of 500 kb upstream

of this promoter, and a region of chromosome 8 at a distance of 3,300 kb from the integrated HPV genome. The integrated HPV genome plays a key role in the formation of this complex, since its experimentally induced deletion leads to the loss of all interactions and abrupt reduction in the level of *c-myc* transcription [184].

Another interesting example of spatial reconfiguration of the genome directed by an integrated virus is related to the retrovirus HTLV-1. The DNA copy of its genome contains a CTCF-binding site [185]. It has been shown that in chromosomes containing an integrated HTLV-1 provirus, numerous spatial contacts arise between this provirus and distant genomic regions, which can be located at a distance of several million base pairs. [186]. The establishment of these contacts correlates with changes in the transcription profile. These changes are complex and cannot be ascribed only to the activation of the genes that spatially interact with the provirus [186]. For this reason, it is worth mentioning that the introduction of new CTCF-binding sites in the genome not only gives rise to new spatial contacts, but also disorganizes the pre-existing system of such contacts. In addition, it can also disrupt the pre-existing enhancer-promoter interactions [187, 188] (*Fig. 2D*). CTCF-binding sites are also found in the genomes of other retroviruses [189]. However, the contribution of their integration into the organization of the genome architecture has not yet been studied.

CONCLUSIONS

There is a lot of evidence on the interaction between virus-specific proteins and functional nuclear compartments in the scientific literature. In this review, we have focused on the studies that provide a mechanistic explanation for the events occurring with intranuclear compartments that are mediated by viral proteins and associated with the infectious process. Meanwhile, most of the published data do not fall under any specific theory in general. For instance, this concerns the causes for temporary deposition of various viral proteins in the nucleolus and relocation of nucleolar components to the nucleoplasm [39–41]. There has been recent evidence that the transcripts of SINE retrotransposons (aluRNA) located in the nucleolus play an important role in maintaining its structural and functional organization [190, 191]. Other studies have shown that transcription of SINE retrotransposons is activated during cell infection with a number of DNA viruses [192]. The question of whether overexpression of these RNAs has an impact on the nucleolus structure remains open. We can hope that the integrated picture will become clearer as new data are accumulated.

It was not until the past few years that virus-induced changes in the 3D genome structure started to draw researchers' attention. Considering the limited number of publications on this topic, we can only assume that these changes are part of the viral strategy to regulate the host genome. This assumption certainly needs further investigation. A promising trend is studying the possibility of reconfiguring the 3D genome by means of cellular DNA transcription induced from the promoters of the proviral genomes integrated into the host cell genome [179, 193, 194]. For now, such transcription is analyzed only in the context of the possible activation of the adjacent genes. Meanwhile, intergenic transcription was shown to promote the removal of cohesin from the CTCF-binding sites [162], which obviously results in reconfiguration of the 3D genome.

Another promising area of research is the possibility to modify the profile of chromosome splitting into TADs upon activation of proviral transcription. It is known that activation of transcription of an endogenous retrotransposon may lead to TAD separation [195]. However, it is reasonable to assume that active transcription of proviruses integrated into the cellular genome in the course of a retroviral infection has similar consequences. It is also interesting to continue the study on the modification of the spatial genome organization mediated by viral proteins binding to the regulatory regions of the host cell genome. There is no reason to assume that this phenomenon is typical only to the EBNA proteins of the Epstein–Barr virus for which this effect has been established [169, 170]. New studies on the trends mentioned above, as well as a number of other related areas, will significantly expand our understanding of the mechanisms of cell infection with various viruses. ●

This work was supported by the Russian Foundation for Basic Research (grant No. 18-29-07001).

REFERENCES

1. Ulianov S.V., Gavrilov A.A., Razin S.V. // Int. Rev. Cell Mol. Biol. 2015. V. 315. P. 183–244.
2. Cremer T., Cremer C. // Nat. Rev. Genet. 2001. V. 2. P. 292–301.
3. Cremer T., Cremer M. // Cold Spring Harb. Perspect. Biol. 2010. V. 2. P. a003889.
4. Cremer T., Kurz A., Zirbel R., Dietzel S., Rinke B., Schrock E., Speicher M.R., Mathieu U., Jauch A., Emmerich P., et al. // Cold Spring Harb. Symp. Quant. Biol. 1993. V. 58. P. 777–792.
5. Cremer T., Cremer M., Hubner B., Silahtaroglu A., Hendzel M., Lanctot C., Strickfaden H., Cremer C. // Bioessays. 2020. V. 42. P. e1900132.

6. Cremer T., Cremer M., Cremer C. // Biochemistry (Moscow). 2018. V. 83. P. 313–325.
7. Huang S. // J. Cell. Biol. 2002. V. 157. P. 739–741.
8. Lamond A.I., Spector D.L. // Nat. Rev. Mol. Cell. Biol. 2003. V. 4. P. 605–612.
9. Spector D.L., Lamond A.I. // Cold Spring Harb. Perspect. Biol. 2011. V. 3. P. a000646.
10. Machyna M., Heyn P., Neugebauer K.M. // Wiley Interdiscip. Rev. RNA. 2013. V. 4. P. 17–34.
11. Wang Q., Sawyer I.A., Sung M.H., Sturgill D., Shevtsov S.P., Pegoraro G., Hakim O., Baek S., Hager G.L., Dundr M. // Nat. Commun. 2016. V. 7. P. 10966.
12. Brown J.M., Green J., das Neves R.P., Wallace H.A., Smith A.J., Hughes J., Gray N., Taylor S., Wood W.G., Higgs D.R., et al. // J. Cell. Biol. 2008. V. 182. P. 1083–1097.
13. Chen Y., Belmont A.S. // Curr. Opin. Genet. Dev. 2019. V. 55. P. 91–99.
14. Lieberman-Aiden E., van Berkum N.L., Williams L., Imakaev M., Ragoczy T., Telling A., Amit I., Lajoie B.R., Sabo P.J., Dorschner M.O., et al. // Science. 2009. V. 326. P. 289–293.
15. 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.
16. Dixon J.R., Gorkin D.U., Ren B. // Mol. Cell. 2016. V. 62. P. 668–680.
17. Nora E.P., Lajoie B.R., Schulz E.G., Giorgetti L., Okamoto I., Servant N., Piolot T., van Berkum N.L., Meisig J., Sedat J., et al. // Nature. 2012. V. 485. P. 381–385.
18. Lupianez D.G., Kraft K., Heinrich V., Krawitz P., Brancati F., Klopocki E., Horn D., Kayserili H., Opitz J.M., Laxova R., et al. // Cell. 2015. V. 161. P. 1012–1025.
19. Symmons O., Uslu V.V., Tsujimura T., Ruf S., Nassari S., Schwarzer W., Ettwiller L., Spitz F. // Genome Res. 2014. V. 24. P. 390–400.
20. Razin S.V., Ulianov S.V. // Cell. Mol. Biol. Lett. 2017. V. 22. P. 18.
21. Rao S.S., Huntley M.H., Durand N.C., Stamenova E.K., Bochkov I.D., Robinson J.T., Sanborn A.L., Machol I., Omer A.D., Lander E.S., et al. // Cell. 2014. V. 159. P. 1665–1680.
22. Sanborn A.L., Rao S.S., Huang S.C., Durand N.C., Huntley M.H., Jewett A.I., Bochkov I.D., Chinnappan D., Cutkosky A., Li J., et al. // Proc. Natl. Acad. Sci. USA. 2015. V. 112. P. E6456–E6465.
23. Krivega I., Dean A. // Curr. Opin. Genet. Dev. 2012. V. 22. P. 79–85.
24. Krivega I., Dean A. // Nat. Cell. Biol. 2017. V. 19. P. 883–885.
25. Kaiser V.B., Semple C.A. // F1000Res. 2017. V. 6. P. F1000 Fac. Rev.-314.
26. Ibrahim D.M., Mundlos S. // Curr. Opin. Cell Biol. 2020. V. 64. P. 1–9.
27. Achinger-Kawecka J., Taberlay P.C., Clark S.J. // Cold Spring Harb. Symp. Quant. Biol. 2016. V. 81. P. 41–51.
28. Valton A.L., Dekker J. // Curr. Opin. Genet. Dev. 2016. V. 36. P. 34–40.
29. Hu S., Lv P., Yan Z., Wen B. // Epigenetics Chromatin. 2019. V. 12. P. 43.
30. Quinodoz S.A., Ollikainen N., Tabak B., Palla A., Schmidt J.M., Detmar E., Lai M.M., Shishkin A.A., Bhat P., Takei Y., et al. // Cell. 2018. V. 174. P. 744–757.
31. Chen Y., Zhang Y., Wang Y., Zhang L., Brinkman E.K., Adam S.A., Goldman R., van Steensel B., Ma J., Belmont A.S. // J. Cell. Biol. 2018. V. 217. P. 4025–4048.
32. Xu M., Cook P.R. // J. Cell. Biol. 2008. V. 181. P. 615–623.
33. Schoenfelder S., Fraser P. // Nat. Rev. Genet. 2019. V. 20. P. 437–455.
34. Schoenfelder S., Sexton T., Chakalova L., Cope N.F., Horton A., Andrews S., Kurukuti S., Mitchell J.A., Umlauf D., Dimitrova D.S., et al. // Nat. Genet. 2010. V. 42. P. 53–61.
35. Razin S.V., Gavrilov A.A., Ioudinkova E.S., Iarovaia O.V. // FEBS Lett. 2013. V. 587. P. 1840–1847.
36. Cook P.R., Marenduzzo D. // Nucl. Acids Res. 2018. V. 46. P. 9895–9906.
37. Iarovaia O.V., Minina E.P., Sheval E.V., Onichtchouk D., Dokudovskaya S., Razin S.V., Vassetzky Y.S. // Trends Cell. Biol. 2019. V. 29. P. 647–659.
38. Calle A., Ugrinova I., Epstein A.L., Bouvet P., Diaz J.J., Greco A. // J. Virol. 2008. V. 82. P. 4762–4773.
39. Salvetti A., Greco A. // Biochim. Biophys. Acta. 2014. V. 1842. P. 840–847.
40. Rawlinson S.M., Moseley G.W. // Cell. Microbiol. 2015. V. 17. P. 1108–1120.
41. Hiscox J.A., Whitehouse A., Matthews D.A. // Proteomics. 2010. V. 10. P. 4077–4086.
42. Strang B.L. // J. Gen. Virol. 2015. V. 96. P. 239–252.
43. Hiscox J.A. // Arch. Virol. 2002. V. 147. P. 1077–1089.
44. Hiscox J.A. // Nat. Rev. Microbiol. 2007. V. 5. P. 119–127.
45. Wang L., Ren X.M., Xing J.J., Zheng A.C. // Virol. Sin. 2010. V. 25. P. 151–157.
46. Calderone A., Licata L., Cesareni G. // Nucl. Acids Res. 2015. V. 43. P. D588–D592.
47. Guirimand T., Delmotte S., Navratil V. // Nucl. Acids Res. 2015. V. 43. P. D583–D587.
48. Nouri K., Moll J.M., Milroy L.G., Hain A., Dvorsky R., Amin E., Lenders M., Nagel-Steger L., Howe S., Smits S.H., et al. // PLoS One. 2015. V. 10. P. e0143634.
49. Reyes E.D., Kulej K., Pancholi N.J., Akhtar L.N., Avgousti D.C., Kim E.T., Bricker D.K., Spruce L.A., Koniski S.A., Seeholzer S.H., et al. // Mol. Cell Proteomics. 2017. V. 16. P. 2079–2097.
50. Emmott E., Wise H., Loucaides E.M., Matthews D.A., Digard P., Hiscox J.A. // J. Proteome Res. 2010. V. 9. P. 5335–5345.
51. Ni L., Wang S., Zheng C. // J. Med. Microbiol. 2012. V. 61. P. 1637–1643.
52. Jia W., Yao Z., Zhao J., Guan Q., Gao L. // Life Sci. 2017. V. 186. P. 1–10.
53. Chaudhry U., Malik D.A., Saleem N., Malik M.T. // EC Microbiology. 2018. V. 14. P. 631–640.
54. Gao Z., Hu J., Wang X., Yang Q., Liang Y., Ma C., Liu D., Liu K., Hao X., Gu M., et al. // Arch. Virol. 2018. V. 163. P. 2775–2786.
55. Dong D., Zhu S., Miao Q., Zhu J., Tang A., Qi R., Liu T., Yin D., Liu G. // J. Gen. Virol. 2020. V. 101. P. 33–43.
56. Bian W.X., Xie Y., Wang X.N., Xu G.H., Fu B.S., Li S., Long G., Zhou X., Zhang X.L. // Nucl. Acids Res. 2019. V. 47. P. 56–68.
57. Tosoni E., Frasson I., Scalabrin M., Perrone R., Butovskaya E., Nadai M., Palu G., Fabris D., Richter S.N. // Nucl. Acids Res. 2015. V. 43. P. 8884–8897.
58. Kalt I., Borodianskiy-Shteinberg T., Schachor A., Sarid R. // J. Virol. 2010. V. 84. P. 2935–2945.
59. Gouzil J., Fablet A., Lara E., Caignard G., Cochet M., Kundlacz C., Palmarini M., Varela M., Breard E., Sailleau C., et al. // J. Virol. 2017. V. 91. P. e01263–16.
60. Banerjee R., Weidman M.K., Navarro S., Comai L., Dasgupta A. // J. Gen. Virol. 2005. V. 86. P. 2315–2322.
61. Amineva S.P., Aminev A.G., Palmenberg A.C., Gern J.E. // J. Gen. Virol. 2004. V. 85. P. 2969–2979.
62. Ponti D., Troiano M., Bellenchi G.C., Battaglia P.A., Gigli-

ani F. // BMC Cell. Biol. 2008. V. 9. P. 32.
63. Zhu C., Zheng F., Zhu J., Liu M., Liu N., Li X., Zhang L., Deng Z., Zhao Q., Liu H. // Oncotarget. 2017. V. 8. P. 94519–94527.
64. Yan Y., Du Y., Wang G., Li K. // Sci. Rep. 2017. V. 7. P. 17761.
65. Kao C.F., Chen S.Y., Lee Y.H. // J. Biomed. Sci. 2004. V. 11. P. 72–94.
66. Ahuja R., Kapoor N.R., Kumar V. // Biochim. Biophys. Acta. 2015. V. 1853. P. 1783–1795.
67. Rajput P., Shukla S.K., Kumar V. // Virol. J. 2015. V. 12. P. 62.
68. Pyper J.M., Clements J.E., Zink M.C. // J. Virol. 1998. V. 72. P. 7697–7702.
69. Li Y.J., Macnaughton T., Gao L., Lai M.M. // J. Virol. 2006. V. 80. P. 6478–6486.
70. Behrens R.T., Aligeti M., Pocock G.M., Higgins C.A., Sherer N.M. // J. Virol. 2017. V. 91. P. e02107–16.
71. Michienzi A., Cagnon L., Bahner I., Rossi J.J. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. P. 8955–8960.
72. Zolotukhin A.S., Felber B.K. // J. Virol. 1999. V. 73. P. 120–127.
73. Vercruysse T., Daelemans D. // Curr. HIV Res. 2013. V. 11. P. 623–634.
74. Charman M., Weitzman M.D. // Viruses. 2020. V. 12. P. 151.
75. Hidalgo P., Gonzalez R.A. // FEBS Lett. 2019. V. 593. P. 3518–3530.
76. Shulla A., Randall G. // Curr. Opin. Microbiol. 2016. V. 32. P. 82–88.
77. Crump C. // Adv. Exp. Med. Biol. 2018. V. 1045. P. 23–44.
78. Sagou K., Uema M.,, Kawaguchi Y. // J. Virol. 2010. V. 84. P. 2110–2121.
79. Strang B.L., Boulant S., Coen D.M. // J. Virol. 2010. V. 84. P. 1771–1784.
80. Terrier O., Carron C., De Chassey B., Dubois J., Traversier A., Julien T., Cartet G., Proust A., Hacot S., Ressnikoff D., et al. // Sci. Rep. 2016. V. 6. P. 29006.
81. Su W.C., Hsu S.F., Lee Y.Y., Jeng K.S., Lai M.M. // J. Virol. 2015. V. 89. P. 11245–11255.
82. Yang C., Liu X., Gao Q., Cheng T., Xiao R., Ming F., Zhang S., Jin M., Chen H., Ma W., et al. // J. Virol. 2018. V. 92. P. e01042–18.
83. Ruiz Z., Mihaylov I.S., Cotmore S.F., Tattersall P. // Virology. 2011. V. 410. P. 375–384.
84. Majumder K., Wang J., Boftsi M., Fuller M.S., Rede J.E., Joshi T., Pintel D.J. // Elife. 2018. V. 7. P. e37750.
85. Adeyemi R.O., Landry S., Davis M.E., Weitzman M.D., Pintel D.J. // PLoS Pathog. 2010. V. 6. P. e1001141.
86. Deng X., Yan Z., Cheng F., Engelhardt J.F., Qiu J. // PLoS Pathog. 2016. V. 12. P. e1005399.
87. Schwartz R.A., Carson C.T., Schuberth C., Weitzman M.D. // J. Virol. 2009. V. 83. P. 6269–6278.
88. Luo Y., Chen A.Y., Qiu J. // J. Virol. 2011. V. 85. P. 133–145.
89. McKinney C.C., Hussmann K.L., McBride A.A. // Viruses. 2015. V. 7. P. 2450–2469.
90. Weitzman M.D., Fradet-Turcotte A. // Annu. Rev. Virol. 2018. V. 5. P. 141–164.
91. Jang M.K., Shen K., McBride A.A. // PLoS Pathog. 2014. V. 10. P. e1004117.
92. Jackson D.A., Hassan A.B., Errington R.J., Cook P.R. // EMBO J. 1993. V. 12. P. 1059–1065.
93. Eskiw C.H., Rapp A., Carter D.R., Cook P.R. // J. Cell. Sci. 2008. V. 121. P. 1999–2007.
94. Carter D.R., Eskiw C., Cook P.R. // Biochem. Soc. Trans. 2008. V. 36. P. 585–589.
95. Chen C.P., Lyu Y., Chuang F., Nakano K., Izumiya C., Jin D., Campbell M., Izumiya Y. // J. Virol. 2017. V. 91. P. e02491–16.
96. Baquero-Perez B., Whitehouse A. // PLoS Pathog. 2015. V. 11. P. e1005274.
97. Li L., Johnson L.A., Dai-Ju J.Q., Sandri-Goldin R.M. // PLoS One. 2008. V. 3. P. e1491.
98. Jenkins H.L., Spencer C.A. // J. Virol. 2001. V. 75. P. 9872–9884.
99. Girard C., Will C.L., Peng J., Makarov E.M., Kastner B., Lemm I., Urlaub H., Hartmuth K., Luhrmann R. // Nat. Commun. 2012. V. 3. P. 994.
100. Galganski L., Urbanek M.O., Krzyzosiak W.J. // Nucl. Acids Res. 2017. V. 45. P. 10350–10368.
101. Simoes M., Rino J., Pinheiro I., Martins C., Ferreira F. // Viruses. 2015. V. 7. P. 4978–4996.
102. Aspegren A., Rabino C., Bridge E. // Exp. Cell Res. 1998. V. 245. P. 203–213.
103. Bridge E., Xia D.X., Carmo-Fonseca M., Cardinali B., Lamond A.I., Pettersson U. // J. Virol. 1995. V. 69. P. 281–290.
104. Park R., Miller G. // J. Virol. 2018. V. 92. P. e01254–18.
105. Jimenez-Garcia L.F., Spector D.L. // Cell. 1993. V. 73. P. 47–59.
106. Bridge E., Riedel K.U., Johansson B.M., Pettersson U. // J. Cell. Biol. 1996. V. 135. P. 303–314.
107. Chang L., Godinez W.J., Kim I.H., Tektonidis M., de Lanerolle P., Eils R., Rohr K., Knipe D.M. // Proc. Natl. Acad. Sci. USA. 2011. V. 108. P. E136–144.
108. Melcak I., Cermanova S., Jirsova K., Koberna K., Malinsky J., Raska I. // Mol. Biol. Cell. 2000. V. 11. P. 497–510.
109. Spector D.L., Fu X.D., Maniatis T. // EMBO J. 1991. V. 10. P. 3467–3481.
110. Mor A., White A., Zhang K., Thompson M., Esparza M., Munoz-Moreno R., Koide K., Lynch K.W., Garcia-Sastre A., Fontoura B.M. // Nat. Microbiol. 2016. V. 1. P. 16069.
111. Fox A.H., Lamond A.I. // Cold Spring Harb. Perspect. Biol. 2010. V. 2. P. a000687.
112. Nakagawa S., Hirose T. // Cell. Mol. Life Sci. 2012. V. 69. P. 3027–3036.
113. Pisani G., Baron B. // Noncoding RNA Res. 2019. V. 4. P. 128–134.
114. Ma H., Han P., Ye W., Chen H., Zheng X., Cheng L., Zhang L., Yu L., Wu X., Xu Z., et al. // J. Virol. 2017. V. 91. P. e02250–16.
115. Zhang Q., Chen C.Y., Yedavalli V.S., Jeang K.T. // mBio. 2013. V. 4. P. e00596–12.
116. Saha S., Murthy S., Rangarajan P.N. // J. Gen. Virol. 2006. V. 87. P. 1991–1995.
117. Wang Z., Fan P., Zhao Y., Zhang S., Lu J., Xie W., Jiang Y., Lei F., Xu N., Zhang Y. // Cell. Mol. Life Sci. 2017. V. 74. P. 1117–1131.
118. Imamura K., Imamachi N., Akizuki G., Kumakura M., Kawaguchi A., Nagata K., Kato A., Kawaguchi Y., Sato H., Yoneda M., et al. // Mol. Cell. 2014. V. 53. P. 393–406.
119. Ahn J.H., Hayward G.S. // J. Virol. 1997. V. 71. P. 4599–4613.
120. Chelbi-Alix M.K., de The H. // Oncogene. 1999. V. 18. P. 935–941.
121. Everett R.D., Freemont P., Saitoh H., Dasso M., Orr A., Kathoria M., Parkinson J. // J. Virol. 1998. V. 72. P. 6581–6591.
122. Adamson A.L., Kenney S. // J. Virol. 2001. V. 75. P. 2388–2399.

123. Everett R.D. // Oncogene. 2001. V. 20. P. 7266–7273.
124. Tavalai N., Stamminger T. // Viruses. 2009. V. 1. P. 1240–1264.
125. Stepp W.H., Meyers J.M., McBride A.A. // mBio. 2013. V. 4. P. e00845–13.
126. Mitchell A.M., Hirsch M.L., Li C., Samulski R.J. // J. Virol. 2014. V. 88. P. 925–936.
127. Bonilla W.V., Pinschewer D.D., Klenerman P., Rousson V., Gaboli M., Pandolfi P.P., Zinkernagel R.M., Salvato M.S., Hengartner H. // J. Virol. 2002. V. 76. P. 3810–3818.
128. Boutell C., Cuchet-Lourenco D., Vanni E., Orr A., Glass M., McFarlane S., Everett R.D. // PLoS Pathog. 2011. V. 7. P. e1002245.
129. Smith M.C., Boutell C., Davido D.J. // Future Virol. 2011. V. 6. P. 421–429.
130. Lee H.R., Kim D.J., Lee J.M., Choi C.Y., Ahn B.Y., Hayward G.S., Ahn J.H. // J. Virol. 2004. V. 78. P. 6527–6542.
131. Schreiner S., Wimmer P., Sirma H., Everett R.D., Blanchette P., Groitl P., Dobner T. // J. Virol. 2010. V. 84. P. 7029–7038.
132. Leppard K.N., Emmott E., Cortese M.S., Rich T. // J. Gen. Virol. 2009. V. 90. P. 95–104.
133. Ullman A.J., Hearing P. // J. Virol. 2008. V. 82. P. 7325–7335.
134. Rivera-Molina Y.A., Martinez F.P., Tang Q. // World J. Virol. 2013. V. 2. P. 110–122.
135. Ishov A.M., Maul G.G. // J. Cell. Biol. 1996. V. 134. P. 815–826.
136. Maul G.G., Ishov A.M., Everett R.D. // Virology. 1996. V. 217. P. 67–75.
137. Everett R.D. // PLoS Pathog. 2013. V. 9. P. e1003386.
138. Everett R.D., Murray J. // J. Virol. 2005. V. 79. P. 5078–5089.
139. Stubbe M., Mai J., Paulus C., Stubbe H.C., Berscheminski J., Karimi M., Hofmann S., Weber E., Hadian K., Hay R., et al. // mBio. 2020. V. 11. P. e00049–20.
140. Reuter N., Schilling E.M., Scherer M., Muller R., Stamminger T. // J. Virol. 2017. V. 91. P. e02335–16.
141. Schmid M., Speiseder T., Dobner T., Gonzalez R.A. // J. Virol. 2014. V. 88. P. 1404–1420.
142. Nagaraju T., Sugden A.U., Sugden B. // Proc. Natl. Acad. Sci. USA. 2019. V. 116. P. 24630–24638.
143. Kobiler O., Brodersen P., Taylor M.P., Ludmir E.B., Enquist L.W. // mBio. 2011. V. 2. P. e00278–11.
144. Tomer E., Cohen E.M., Drayman N., Afriat A., Weitzman M.D., Zaritsky A., Kobiler O. // FASEB J. 2019. V. 33. P. 9388–9403.
145. Li Z., Fang C., Su Y., Liu H., Lang F., Li X., Chen G., Lu D., Zhou J. // Virol. J. 2016. V. 13. P. 65.
146. Taylor T.J., Knipe D.M. // J. Virol. 2004. V. 78. P. 5856–5866.
147. Dembowski J.A., DeLuca N.A. // PLoS Pathog. 2015. V. 11. P. e1004939.
148. Strom A.R., Brangwynne C.P. // J. Cell. Sci. 2019. V. 132. P. jcs235093.
149. Uversky V.N. // Adv. Colloid Interface Sci. 2017. V. 239. P. 97–114.
150. Tamarozzi E.R., Giuliatti S. // Int. J. Mol. Sci. 2018. V. 19. P. 198.
151. Pelka P., Ablack J.N., Fonseca G.J., Yousef A.F., Mymryk J.S. // J. Virol. 2008. V. 82. P. 7252–7263.
152. Xue B., Blocquel D., Habchi J., Uversky A.V., Kurgan L., Uversky V.N., Longhi S. // Chem. Rev. 2014. V. 114. P. 6880–6911.
153. Peng Q., Wang L., Qin Z., Wang J., Zheng X., Wei L., Zhang X., Zhang X., Liu C., Li Z., et al. // J. Virol. 2020. V. 94. P. e01771–19.
154. Taylor T.J., McNamee E.E., Day C., Knipe D.M. // Virology. 2003. V. 309. P. 232–247.
155. McSwiggen D.T., Hansen A.S., Teves S.S., Marie-Nelly H., Hao Y., Heckert A.B., Umemoto K.K., Dugast-Darzacq C., Tjian R., Darzacq X. // Elife. 2019. V. 8. P. e47098.
156. Denker A., de Laat W. // Genes Dev. 2016. V. 30. P. 1357–1382.
157. Sati S., Cavalli G. // Chromosoma. 2017. V. 126. P. 33–44.
158. Hensel K.O., Cantner F., Bangert F., Wirth S., Postberg J. // Epigenetics Chromatin. 2018. V. 11. P. 34.
159. Moreau P., Cournac A., Palumbo G.A., Marbouty M., Mortaza S., Thierry A., Cairo S., Lavigne M., Koszul R., Neuveut C. // Nat. Commun. 2018. V. 9. P. 4268.
160. Kim K.D., Tanizawa H., De Leo A., Vladimirova O., Kossenkov A., Lu F., Showe L.C., Noma K.I., Lieberman P.M. // Nat. Commun. 2020. V. 11. P. 877.
161. Moquin S.A., Thomas S., Whalen S., Warburton A., Fernandez S.G., McBride A.A., Pollard K.S., Miranda J.L. // J. Virol. 2018. V. 92. P. e01413–17.
162. Heinz S., Texari L., Hayes M.G.B., Urbanowski M., Chang M.W., Givarkes N., Rialdi A., White K.M., Albrecht R.A., Pache L., et al. // Cell. 2018. V. 174. P. 1522–1536.
163. Rutkowski A.J., Erhard F., L'Hernault A., Bonfert T., Schilhabel M., Crump C., Rosenstiel P., Efstathiou S., Zimmer R., Friedel C.C., et al. // Nat. Commun. 2015. V. 6. P. 7126.
164. Hennig T., Michalski M., Rutkowski A.J., Djakovic L., Whisnant A.W., Friedl M.S., Jha B.A., Baptista M.A.P., L'Hernault A., Erhard F., et al. // PLoS Pathog. 2018. V. 14. P. e1006954.
165. Kanda T. // Adv. Exp. Med. Biol. 2018. V. 1045. P. 377–394.
166. Farrell P.J. // Annu. Rev. Pathol. 2019. V. 14. P. 29–53.
167. Sall F.B., Germini D., Kovina A.P., Ribrag V., Wiels J., Toure A.O., Iarovaia O.V., Lipinski M., Vassetzky Y. // Biochemistry (Moscow). 2018. V. 83. P. 402–410.
168. Jiang S., Zhou H., Liang J., Gerdt C., Wang C., Ke L., Schmidt S.C.S., Narita Y., Ma Y., Wang S., et al. // Cell Host Microbe. 2017. V. 22. P. 561–573 e4.
169. Wood C.D., Veenstra H., Khasnis S., Gunnell A., Webb H.M., Shannon-Lowe C., Andrews S., Osborne C.S., West M.J. // Elife. 2016. V. 5. P. e18270.
170. McClellan M.J., Wood C.D., Ojeniyi O., Cooper T.J., Kanhere A., Arvey A., Webb H.M., Palermo R.D., Harth-Hertle M.L., Kempkes B., et al. // PLoS Pathog. 2013. V. 9. P. e1003636.
171. Paschos K., Parker G.A., Watanatanasup E., White R.E., Allday M.J. // Nucl. Acids Res. 2012. V. 40. P. 7233–7246.
172. Musinova Y.R., Sheval E.V., Dib C., Germini D., Vassetzky Y.S. // Cell. Mol. Life Sci. 2016. V. 73. P. 589–601.
173. Germini D., Tsfasman T., Klibi M., El-Amine R., Pichugin A., Iarovaia O.V., Bilhou-Nabera C., Subra F., Bou Saada Y., Sukhanova A., et al. // Leukemia. 2017. V. 31. P. 2515–2522.
174. Maeda N., Fan H., Yoshikai Y. // Rev. Med. Virol. 2008. V. 18. P. 387–405.
175. Katsura Y., Asai S. // Am. J. Med. Sci. 2019. V. 358. P. 384–388.
176. Johnson W.E. // Nat. Rev. Microbiol. 2019. V. 17. P. 355–370.
177. Bushman F.D. // Mol. Ther. 2020. V. 28. P. 352–356.
178. Uren A.G., Kool J., Berns A., van Lohuizen M. // Oncogene. 2005. V. 24. P. 7656–7672.

179. Rosewick N., Durkin K., Artesi M., Marcais A., Hahaut V., Griebel P., Arsic N., Avettand-Fenoel V., Burny A., Charlier C., et al. // Nat. Commun. 2017. V. 8. P. 15264.
180. Pattison J.M., Wright J.B., Cole M.D. // PLoS One. 2015. V. 10. P. e0120256.
181. Zhang J., Markus J., Bies J., Paul T., Wolff L. // J. Virol. 2012. V. 86. P. 10524–10532.
182. de Ridder J., Uren A., Kool J., Reinders M., Wessels L. // PLoS Comput. Biol. 2006. V. 2. P. e166.
183. Babaei S., Akhtar W., de Jong J., Reinders M., de Ridder J. // Nat. Commun. 2015. V. 6. P. 6381.
184. Shen C., Liu Y., Shi S., Zhang R., Zhang T., Xu Q., Zhu P., Chen X., Lu F. // Int. J. Cancer. 2017. V. 141. P. 540–548.
185. Satou Y., Miyazato P., Ishihara K., Yaguchi H., Melamed A., Miura M., Fukuda A., Nosaka K., Watanabe T., Rowan A.G., et al. // Proc. Natl. Acad. Sci. USA. 2016. V. 113. P. 3054–3059.
186. Melamed A., Yaguchi H., Miura M., Witkover A., Fitzgerald T.W., Birney E., Bangham C.R. // Elife. 2018. V. 7. P. e36245.
187. Hou C., Zhao H., Tanimoto K., Dean A. // Proc. Natl. Acad. Sci. USA. 2008. V. 105. P. 20398–20403.
188. Rawat P., Jalan M., Sadhu A., Kanaujia A., Srivastava M. // Mol. Cell. Biol. 2017. V. 37. P. e00557–16.
189. Goodman M.A., Arumugam P., Pillis D.M., Loberg A., Nasimuzzaman M., Lynn D., van der Loo J.C.M., Dexheimer P.J., Keddache M., Bauer T.R., Jr., et al. // J/ Virol. 2018. V. 92. P. e01639–17.
190. Caudron-Herger M., Pankert T., Rippe K. // Nucleus. 2016. V. 7. P. 308–318.
191. Caudron-Herger M., Pankert T., Seiler J., Nemeth A., Voit R., Grummt I., Rippe K. // EMBO J. 2015. V. 34. P. 2758–2774.
192. Schaller A.M., Tucker J., Willis I., Glaunsinger B.A. // J. Virol. 2020. JVI.00262–20
193. Kataoka K., Nagata Y., Kitanaka A., Shiraishi Y., Shimamura T., Yasunaga J., Totoki Y., Chiba K., Sato-Otsubo A., Nagae G., et al. // Nat. Genet. 2015. V. 47. P. 1304–1315.
194. Sokol M., Wabl M., Ruiz I.R., Pedersen F.S. // Retrovirology. 2014. V. 11. P. 36.
195. Zhang Y., Li T., Preissl S., Amaral M.L., Grinstein J.D., Farah E.N., Destici E., Qiu Y., Hu R., Lee A.Y., et al. // Nat. Genet. 2019. V. 51. P. 1380–1388.

The Delivery of Biologically Active Agents into the Nuclei of Target Cells for the Purposes of Translational Medicine

A. S. Sobolev[1,2*]
[1]Institute of Gene Biology, Russian Academy of Sciences, Moscow,119334 Russia
[2]Lomonosov Moscow State University, Moscow, 119234 Russia
*E-mail: alsobolev@yandex.ru
Received June 09, 2020; in final form, September 25, 2020
DOI: 10.32607/actanaturae.11049
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License,which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Development of vehicles for the subcellular targeted delivery of biologically active agents is very promising for the purposes of translational medicine. This review summarizes the results obtained by researchers from the Laboratory of Molecular Genetics of Intracellular Transport, Institute of Gene Biology RAS, which allowed them to design the core technology: modular nanotransporters. This approach ensures high efficacy and cell specificity for different anti-cancer agents, as they are delivered into the most vulnerable subcellular compartment within the cells of interest and makes it possible for antibody mimetics to penetrate into a compartment of interest within the target cells ("diving antibodies"). Furthermore, polyplexes, complexes of polycationic block copolymers of DNA, have been developed and characterized. These complexes are efficient both *in vitro* and *in vivo* and demonstrate predominant transfection of actively dividing cells.**
KEYWORDS **modular nanotransporters, polyplexes, drug delivery, antibody mimetics, gene therapy, photodynamic therapy, radiotherapy.**
ABBREVIATIONS **AE – Auger electron; αMSH – α-melanocyte-stimulating hormone; AP – alpha particle; ARE – antioxidant response element; DTox – a fragment of the diphtheria toxin translocation domain; EC_{50} – half maximal effective concentration; EGF – epidermal growth factor; EGFR – EGF receptor; FA – folic acid; FR – FA receptor; HMP – hemoglobin-like protein of *E. coli*; K_d – dissociation constant; MNT – modular nanotransporters; MNT_{EGF} – MNT with EGF as a ligand module; MNT_F – MNT with folic acid as a ligand module; MNT_{MSH} – MNT with αMSH as a ligand module; NLS – nuclear localization signal; Nrf2 – transcription factor regulating, in particular, expression of antioxidant response genes; PEG – polyethylene glycol; PEI – polyethyleneimine; PS – photosensitizer.**

INTRODUCTION

The cell nucleus, where the main program of cell function is stored, is a natural target for many biologically active substances. These substances can be divided into two large groups [1]. The first group includes agents (e.g., the cytotoxic ones) that can have a damaging effect anywhere in the cell, with the nucleus being the compartment most sensitive to them. In other words, if these agents reside in the nucleus, the same effect will be achieved at a minimal concentration compared to other localizations. The second group consists of agents that begin showing their impact from the instant they enter the cell nucleus (e.g., DNA). The present review focuses on both of these groups.

Photosensitizers (PSs) and radionuclides emitting particles with a short path length (such as emitters of alpha particles (APs) or Auger electrons (AEs)) exemplify agents from the first group. Both of them are cytotoxic agents that are widely used in medical practice to treat cancer, but they are not limited to this type of diseases. The cell nucleus is extremely sensitive to the cytotoxic agent of PSs, reactive oxygen species (i.e., singlet oxygen, hydroxyl radical, and a number of other free radicals) [2]. As for the emitters of APs and AEs, it has been known for over 50 years that the cell nucleus is the cellular compartment most sensitive to them [3]. Meanwhile, both PSs and emitters of APs/AEs exhibit neither tropicity with respect to the cell nucleus nor cell specificity.

There is little doubt that DNA needs to be delivered into the cell nucleus if its expression is to be achieved. The second group of biologically active substances also includes regulatory polypeptides whose effect mani-

fests upon interaction with the macromolecules of the cell nucleus.

Therefore, for the purposes of translational medicine, biologically active agents must be delivered to the nuclei of target cells so that their properties can be deployed, since most of these agents cannot reach the nuclei by themselves.

The key for delivering macromolecules or other biologically active substances (for brevity, they will hereinafter be referred to as "cargo") is to use natural intracellular transport processes, such as receptor-mediated endocytosis or nucleocytoplasmic transport (they have been described in numerous books and reviews, such as refs. [4, 5]). Accordingly, the vehicle must contain amino acid sequences or other target molecules that prompt it to move in the desired direction and to overcome the numerous barriers on its way to the nucleus (both on the target cell surface and inside it) [6].

MODULAR NANOTRANSPORTERS

The modular nanotransporters (MNTs) being developed in our laboratory meet these criteria and can be regarded as a technological platform for the delivery of therapeutic agents to a given compartment of target cells of the desired type [1, 7–16]. This platform is based on: (a) use of natural processes of specific molecular recognition; (b) the previously mentioned transport inside the cell and outside of it, and (c) the principle of modularity; i.e., the ability to change the transport or recognition units/modules to adapt MNTs to the desired type of target cells, cellular compartments, the intracellular targets, and the "cargo" being delivered. A typical MNT (*Fig. 1*) consists of a ligand module, an endosomolytic module, a nuclear localization module, and a carrier module. The ligand module ensures interaction with the internalizable surface receptor and, therefore, recognition of the target cell and transport of an agent inside this cell via receptor-mediated endocytosis. The endosomolytic module has the function of pH-dependent pore formation in the endosomes, and thus it ensures release of the MNT, with the active component being delivered from these compartments with weakly acidic contents to the target cell cytosol. The nuclear localization module

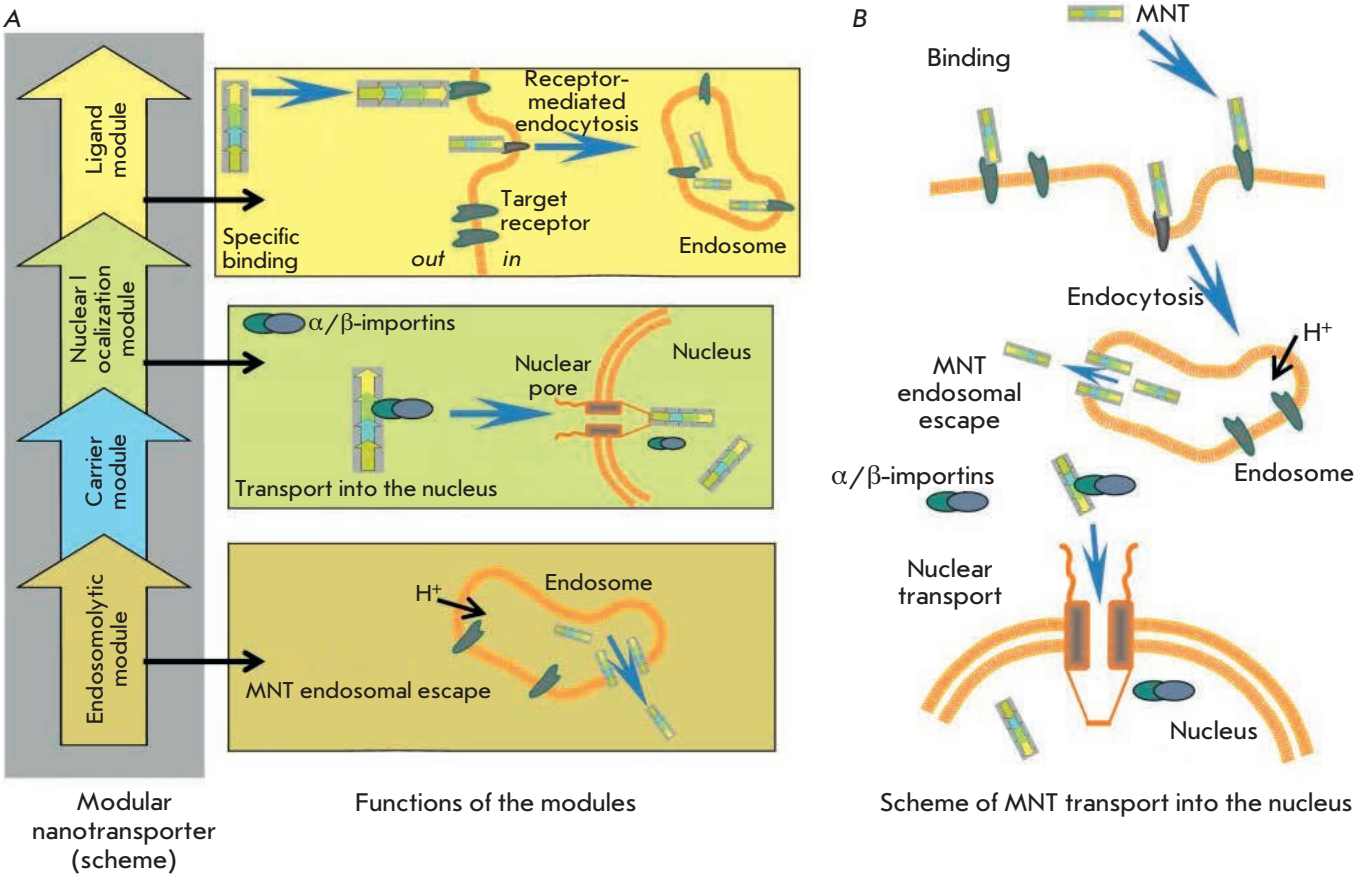


Fig. 1. Modular nanotransporters (*A*) and the schematic representation (*B*) of how they are transported into the cell nucleus (after [15])

contains an amino acid sequence that acts as a nuclear localization signal (NLS) specifically interacting with the importin complex in the cytosol and ensuring transport of the agent through the nuclear pore. The carrier module is used to join other modules into an integral whole and attach the "cargo." Along with the aforementioned four modules, the MNT can also contain other modules if interaction with some additional intra- and extracellular components is required. Thus far, the properties of the following MNTs are the ones that have been studied most thoroughly both *in vitro* and *in vivo*:

- MNTs with epidermal growth factor (EGF) as a ligand module (MHT_{EGF}), which exhibit specificity with respect to cells[1] overexpressing epidermal growth factor receptors (EGFR) [9, 19, 20],
- MNTs with the α-melanocyte-stimulating hormone as a ligand module (MHT_{MSH}), which exhibit specificity with respect to cells[2] overexpressing melanocortin-1 receptors [8, 12, 22],
- MNTs with folic acid as a ligand module (MHT_{F}) targeted at cells[3] overexpressing folate receptors [23, 24].

The compositions of the modules of these MNTs are as follows:

MNT_{EGF}: DTox-HMP-NLS-EGF;
MNT_{MSH}: DTox-HMP-NLS-αMSH;
MNT_{F}: DTox-HMP-NLS-PEG-FA,

where DTox is a fragment of the translocation domain of the diphtheria toxin (the endosomolytic module); HMP is the hemoglobin-like protein of *E. coli* (the carrier module); NLS is the optimized nuclear localization sequence of the SV40 large T antigen (the nuclear localization module); EGF, αMSH, and FA are epidermal growth factor, melanocyte-stimulating hormone, and folic acid, respectively (the ligand modules); and PEG is bifunctional polyethylene glycol.

Other MNT variants have been developed or are currently being developed. They are discussed in the sections below.

When fragments of molecules join together to form a single molecule (as is the case when creating an MNT), it is not obvious that the resulting chimeric molecule will retain the properties of these fragments. Thus, its domains may spatially mask one another and impede the interactions with cellular proteins that are required for the functioning of the chimeric molecule, so the designed artificial molecule will not have the purposive properties. Both the structure of MNTs and the ability of their modules to perform their functions were studied to test the performance of MNTs.

According to dynamic light scattering data, the dimensions of MNT_{MSH} and MNT_{EGF} are 8.3 ± 0.6 and 10.6 ± 0.5 nm [19], respectively.

Numerous attempts to crystallize MNTs for further study of their structure by high-resolution X-ray diffraction analysis have failed. However, small-angle X-ray scattering, atomic force and electron microscopy [25] have shed some light on the structures of MNT_{MSH} and MNT_{EGF}. An important conclusion drawn from this structural study is that the endosomolytic and the ligand modules are spatially separated sufficiently well. Their mutual masking and the loss of their functions are, therefore, eliminated.

This conclusion was convincingly confirmed by tests performed to evaluate the performance of the MNT modules of all developed types. To save space, let us just provide few examples. Thus, the dissociation constant (K_d) of the MNT_{EGF}-EGFR complexes was 29 nM, which is close to that of the EGF-EGFR complexes [9]. For the complexes formed between MNT_{MSH} and melanocortin receptors, K_d was approximately 20 nM [8]. The studied MNTs exhibited a membranolytic activity in two pH ranges: at pH 5.5–6.5 (the range being close to that of endosomes and mediated by DTox) and at pH range of 3–4, which is caused by the action of HMP [8, 9]. The membrane pores created by the MNTs have been characterized electrochemically and by atomic force microscopy [7, 9, 22]. After the full-length MNT_{MSH} (i.e., the ones containing all four modules) had been added to the planar lipid bilayer at pH 5.5, ion channels with a conductivity of ~ 2–5 nS appeared. Meanwhile, MNT_{MSH} without the endosomolytic module did not form ion channels at pH 5.5. The channels did not appear even under the action of full-length MNTs at a neutral pH (7.0), thus proving that the endosomolytic module exhibits its membrane activity in acidified milieus. Five to fifteen minutes after the milieu had been acidified to pH 5.5, MNT_{EGF} formed ring structures 30–50 nm in diameter in the lipid bilayer, as detected by atomic force microscopy. Fluctuating holes 50–200 nm in diameter permeating the lipid bilayer could be detected after 40–60 min. The function of the endosomolytic module was also demonstrated in living cells (Cloudman S91 mouse melanoma, the M3 clone) [8] by measuring the pH of the intracellular microenvironment of MNT_{MSH} by fluorescence ratio image microscopy. The MNT_{MSH} without the endosomolytic module (DTox) resided in vesicles with weakly acidic and acidic contents, while the full-length MNT_{MSH} (with the DTox module) was located in the neutral microenvironment. This result demonstrates that the full-length MNT_{MSH} can escape from the acidified endocytic compartments

[1] *Examples: cells of bladder cancer, head and neck cancer, glioblastoma, and colorectal cancer [17, 18].*
[2] *An example: melanoma cells [21].*
[3] *Examples: cervical and ovarian cancer cells [10].*

of living cells. The interaction between the NLS-carrying module within various MNTs and α/β-importin dimers ensuring the delivery of NLS-carrying proteins to the cell nucleus has been characterized by several methods (surface plasmon resonance and thermophoresis) [9, 26]. The measured constants of the affinity of MNTs to importin dimers were close to that of a free natural polypeptide carrying the same NLS. Hence, it was demonstrated that all the modules within a chimeric artificial MNT molecule had retained their functions.

Therefore, all the full-length MNTs penetrated the target cells via receptor-mediated endocytosis (as confirmed by the fact that the specific ligands of the respective receptors inhibited their penetration) and localized within their cell nuclei [8, 9, 14, 23, 26–28] (*Fig. 2*), as was actually planned by the authors in order to solve the problem related to "cargo" delivery into the nucleus.

PSs, which are typically used for photodynamic treatment of malignant tumors (although some other uses are also known), generate cytotoxic reactive oxygen species under irradiation in an oxygenated medium [2]. As mentioned in the Introduction section, the cell nucleus is the cellular compartment most sensitive to the damaging actions of both reactive oxygen species and the emitters of APs and AEs used in the radiation therapy of malignant tumors [29, 30]. Since MNTs can penetrate cells via the receptor-mediated pathway (this penetration is specific as MNTs penetrate cells that present these receptors), and most importantly, can accumulate in their nuclei, it was necessary to verify whether the delivery of emitters of APs, AEs, and PSs into the nuclei by MNTs can enhance their cytotoxicity.

Indeed, PSs such as the chlorin e_6 and bacteriochlorin *p* attached to MNT_{EGF} or MNT_{MSH} are hundreds and thousands of time more cytotoxic than the free ones. Thus, in the experiments on A431 cells overexpressing EGFR, the half-maximal effective concentration (EC_{50}) of MNT_{EGF}-chlorin e_6 was 0.53 nM, while EC_{50} of free chlorin e_6 was 1780 nM (i.e., 3,360-fold higher [9]). In other words, the same cytotoxic effect of the chlorin e_6 photosensitizer can be achieved by using concentrations 3,360 times as low by moving this photosensitizer to the nucleus using MNTs. The same experiments demonstrated that MNT_{EGF} made the PS specific to certain cells. Thus, whereas free chlorin e_6 was cytotoxic against both the target A431 cells and non-target NIH 3T3 fibroblasts lacking EGFR, MNT_{EGF}-chlorin e_6 affected the target cells only.

Qualitatively similar results were obtained in the experiments with MNT_{MSH} [8], where it was also demonstrated that an MNT must contain all four modules.

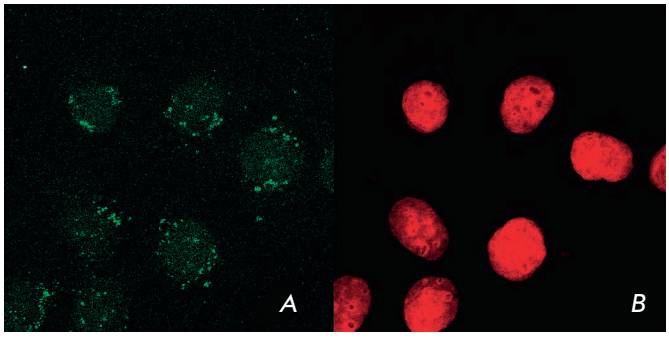


Fig. 2. Subcellular MNT_{EGF} localization within A431 human epidermoid carcinoma cells (after [9] with permission). The A431 cells were incubated for 4 h with MNT_{EGF} in a culture medium, then washed and incubated in the medium without MNT_{EGF}. (*A*) – immunocytochemical detection of MNT_{EGF}, (*B*) – nuclear DNA detection with ToPro-3 in the same group of A431 cells

Thus, an MNT lacking the endosomolytic module was 5.3 times less active than the full-length MNT, while the MNT lacking the NLS-containing module was even less cytotoxic.

It has been convincingly demonstrated in *in vivo* experiments on tumor-carrying mice that PSs are efficiently delivered to the nuclei of cancer cells using MNTs [19, 31]. An immunocytochemical analysis of the distribution of MNT_{EGF} and MNT_{MSH} injected intravenously to tumor-carrying mice revealed that MNTs preferentially accumulate in cancer cells (to be more specific, in the nuclei of cancer cells). The experiment involving intravenous injection of MNT_{MSH} for the treatment of experimentally induced melanomas showed that photosensitizer bacteriochlorin *p* delivered by MNT_{MSH} inhibits B16-F1 tumor growth 85–89% more efficiently compared to free bacteriochlorin *p*; the inhibition of the growth of Cloudman S90 melanoma was 93% more efficient. The ratio between the PS concentrations in the tumor and in the skin was as high as 9.8, some 4.5 times higher than that observed for free bacteriochlorin *p* [32]. A significant therapeutic effect was also uncovered for an intravenous injection of MNT_{EGF} with photosensitizer chlorin e_6 in a model of A431 human epidermoid carcinoma grafted into immunodeficient mice: 75% of the mice survived by day 92, while only 20% of the mice treated with free chlorin e_6 (positive control) and none of the untreated animals survived by day 23.

APs and AEs cause dense ionization and thus efficiently damage the molecules along their tracks; the path length of these particles in tissues is rather short: 50–100 μm (i.e., several cell diameters) for APs and several dozens or hundreds nanometers for AEs (i.e., they are almost equal to the dimensions of the cell nucleus).

These features of emitters of APs and AEs are rather attractive owing to the fact that in the case when they are selectively delivered to the target cells, one can expect the damage to the surrounding normal cells to be minimal. Meanwhile, both types of radiation cause multiple double-strand DNA breaks that are hardly repairable. In fact, nuclear DNA is the main target of the cytotoxic activity of these radiation types [33, 34]. Their cytotoxicity practically does not drop as the oxygen content decreases [35] (the so-called "oxygen effect" that is characteristic of sparsely ionizing radiation), so that these types of radiation have a special advantage in damaging hypoxic cancer cells. Emitters of AEs are also quite interesting, because up to several dozen AEs are produced per decay (depending on the nature of the emitter), thus ensuring a high biological efficiency for these species if their decay occurs in close proximity to DNA [36]. Taking into account the aforementioned features, APs and, especially, AEs are of interest for the treatment of malignant tumors located in such places where damage to the surrounding normal tissues must be minimized (e.g., brain tumors, especially in children) [37], or for the treatment of micrometastases [38].

The α-particle emitter ^{211}At, which has been used as a source of APs in experiments with MNTs, is considered one of the most promising radionuclides for therapeutic purposes [30]. The AP emitter ^{211}At has a relatively short half-life (7.2 hrs); the path length of APs emitted by it can reach up to 70 μm; the resulting yield of double-strand DNA breaks is rather high [39]. In experiments with A431 human epidermoid carcinoma cells, as well as two human glioblastoma lines (D247MG and U87MG.wtEGFR), the cytotoxicity exhibited by ^{211}At-MNT$_{EGF}$ was 8- to 18-fold higher than that of ^{211}At not delivered to the nuclei of these cells [40]. It also turned out that delivery of this emitter of APs to the cell nucleus enabled the effects of recoil nuclei, which are not revealed for other intracellular localizations because of their extremely short path length.

The following emitters of AEs, which are widely used in medicine as sources of gamma radiation, were employed in the experiments with MNTs: ^{125}I, ^{67}Ga, and ^{111}In. On average, they emit 24.9, 4.7, and 14.7 AEs per decay, respectively [41]. The yield of double-strand DNA breaks caused by AEs significantly depends on the distance between a DNA molecule and the emitter of AEs [42].

^{125}I or ^{67}Ga delivered by MNT$_{EGF}$ accumulated rather intensively in the nuclei of A431 human epidermoid carcinoma cells [27, 28]: by the first hour of incubation, about 60% of all the radioactivity pumped into the cells was found in their nuclei. ^{125}I-MNT$_{EGF}$ was 3,500 times more cytotoxic to A431 cells than the ^{125}I-iodinated control polypeptide, which had not penetrated the cells [27]. Similar results were obtained for ^{67}Ga [28] and ^{111}In [20]: the cytotoxicity of the emitters of AEs delivered to the cell nuclei increased abruptly. In these experiments conducted for three cell lines (A431, D247MG, and U87MG.wtEGFR), the cytotoxicities of ^{125}I and ^{67}Ga delivered into the cell nucleus by MNTs were compared to those of the radionuclides delivered mostly into the cytoplasm. As might be expected, the delivery of these emitters into the nucleus ensured a significantly higher cytotoxicity (20- to 400-fold depending on the particular radionuclide and cell line) [15].

Safety testing of MNTs during preclinical studies conducted at the National Medical Research Radiological Center of the Ministry of Health of the Russian Federation showed that the studied MNTs injected intratumorally exhibited a very low toxicity (both acute and chronic) in mice and rats, low immunogenicity/allergenicity in mice and guinea pigs, and were not pyrogenic in rabbits [19, 43–45]. In general, this therapeutic approach, involving intratumoral injection of MNTs, was considered safe [46].

^{111}In-MNT$_{EGF}$ administered as a single dose into human bladder carcinoma (EJ) grafted subcutaneously to immunodeficient Balb/c *nu/nu* mice was retained inside the tumor for a rather long time (its retention half-time in the tumor was 4.1 ± 0.5 days) [20]; no more than 0.5% of the injected dose entered the blood. When delivered intratumorally, ^{111}In-MNT$_{EGF}$ exhibited a pronounced dose-dependent therapeutic effect on EJ tumors (up to 90% compared to the untreated control (both non-labeled MNT$_{EGF}$ and free ^{111}In) at the same dose) [20] (*Fig. 3*).

Another variant of MNTs, ^{111}In-MNT$_{F}$, exhibited a similar therapeutic effect [23, 24]. ^{111}In-MNT$_{F}$ ensured a dose-dependent growth inhibition of subcutaneously grafted tumors (cervical cancer HeLa cells) in immunodeficient mice (up to 80%); the survival rate of the animals was as high as 60% (by day 90), while all the untreated animals in the control group died by day 21.

The results obtained using different cytotoxic agents (two PSs, one emitter of APs, and three emitters of AEs) have motivated researchers to view MNTs as prospective agents for the delivery of a much wider range of biologically active molecules. In this sense, bioactive polypeptides are particularly attractive because MNTs are actually chimeric polypeptides and inclusion of additional polypeptide fragments into their composition is a problem that can be solved using genetic engineering methods.

The MNT carrying a fragment of the p21 protein, p21-MNT$_{EGF}$, is one of the variants of such MNTs. The p21 protein exhibits a broad range of activities: it

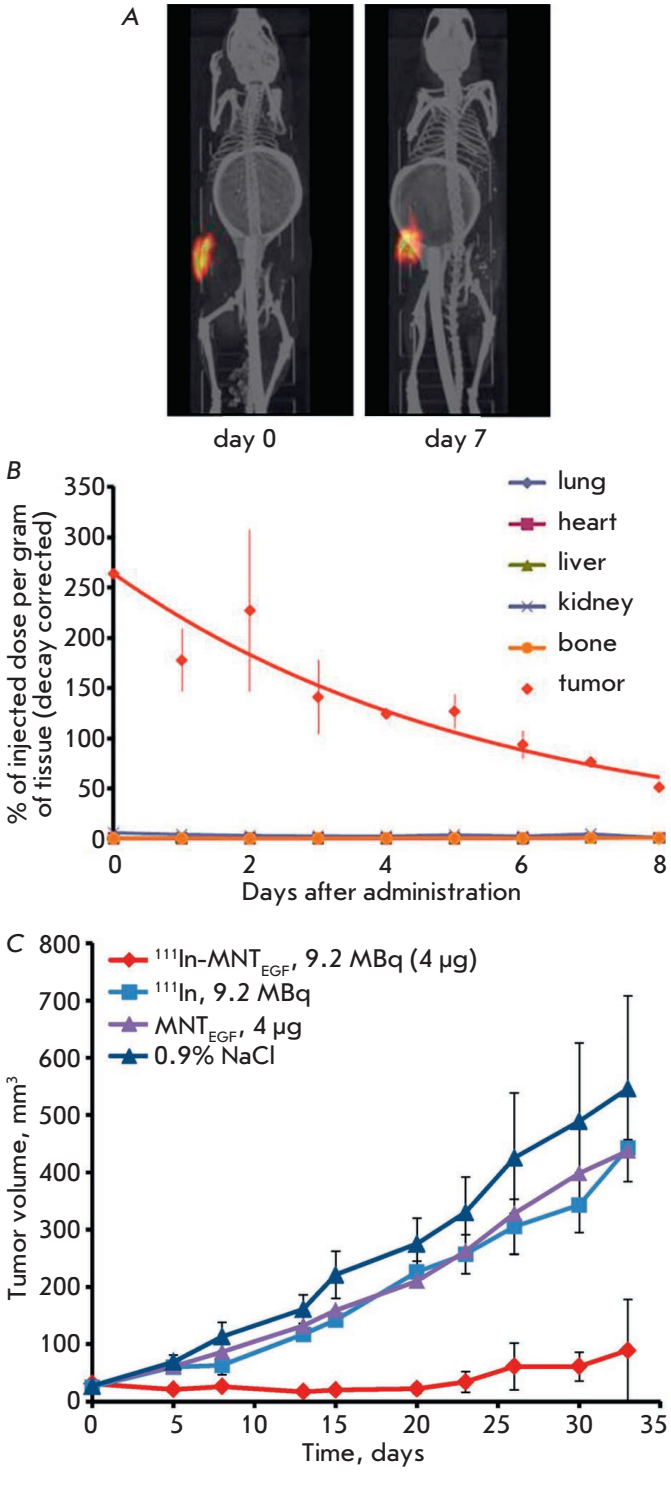


Fig. 3. Administration of ^{111}In-MNT$_{EGF}$ into subcutaneous tumors (EJ human bladder cancer) transplanted to immunodeficient Balb/c *nu/nu* mice (after [20] with changes): (*A*) – SPECT/CT visualization of radioactivity retention within the tumor; (*B*) – the kinetics of radioactivity retention by the tumor and normal tissues; (*C*) – antitumor efficiency of ^{111}In-MNT$_{EGF}$ after intratumoral administration

affects DNA repair and controls the DNA replication fork by forming a complex with the PCNA protein, and it regulates the cell cycle by interacting with cyclins and cyclin-dependent kinases [47]. This makes p21 or its fragments through which it binds to PCNA an attractive tool for modifying the action of DNA-damaging agents (e.g., those used in cancer therapy). The p21-MNT$_{EGF}$ that contained the C-terminal fragment of protein p21 (amino acid residues 87–164), with the site through which p21 binds to PCNA, was synthesized based on these starting points [48]. DNA was damaged by bleomycin, an anticancer drug that causes double-strand DNA breaks [49]. The comet assay in an alkaline medium, which allows one to detect all types of DNA breaks, was used to analyze and repair damage to DNA. Pre-incubation of A431 cells with p21-MNT$_{EGF}$ showed that p21-MNT$_{EGF}$ statistically significantly inhibits DNA repair compared to the control, MNT$_{EGF}$ (i.e., similar MNTs not carrying the p21 fragment) [14].

The encouraging results of this study have contributed to further progress towards the targeted intracellular delivery of biologically active polypeptides. MNTs carrying an antibody mimetic anti-Keap1 monobody (which activates the Nrf2/ARE signaling pathway through competition with endogenous Nrf2 for binding to the Keap1-inhibiting protein) as the effector part were designed under the Russian Science Foundation Grant No. 17-14-01304 [50]. The transcription factor Nrf2 regulates several hundred genes (some of them involved in cell defense against oxidative stress (i.e., antioxidant defense), while others participate in the defense against toxic xenobiotics and a number of other vital processes) [51]. Oxidative stress accompanies or is involved in the pathogenesis of many diseases, such as Parkinson's disease, Huntington's disease, diabetes mellitus, atherosclerosis, cell senescence, radiation-induced cell damage, etc. [51, 52]. In the absence of oxidizing agents, Nrf2, which forms a complex with its inhibitor Keap1 in the cytoplasm, undergoes ubiquitination, followed by degradation in proteasomes. The xenobiotic oxidizing agents appearing in the cell interact with the thiol groups of "cysteine sensors" within Keap1, which leads to Nrf2 release and accumulation in the nucleus, followed by its interaction with the "antioxidant response element" (ARE) within the domain of the promoters of controllable genes, thus activating their transcription [51, 53]. The experiments with MNTs containing the anti-Keap1 monobody revealed a statistically significant increase in the expression level of a number of antioxidant defense genes. Furthermore, the cells were protected against the oxidative stress induced by *tert*-butyl hydroperoxide. It was shown for the mouse model of oxidative stress induced by hepatotoxin acetaminophen that the preliminary

administration of MNTs with anti-Keap1 monobody activating the Nrf2/ARE signaling pathway inhibits the hepatotoxic action of the acetaminophen that is detected according to elevated serum aspartate aminotransferase and alanine aminotransferase activity [50]. These results indicate that MNTs can be used to deliver antibody mimetics both *in vitro* and *in vivo*.

POLYPLEXES (COMPLEXES OF CATIONIC BLOCK COPOLYMERS OF DNA) FOR DELIVERING GENETIC MATERIAL

For decades, the potential opportunity to change the function of cells by modifying their genetic program has been stimulating researchers who focus on gene therapy (e.g., for treating cancer or hereditary diseases) or bioengineering for the production of the target macromolecules, etc. As often happens when trying to solve such important problems consisting of several large tasks, finding the optimal solution to one of them is far from obvious. Targeted delivery of genetic material is one such task. A natural solution to this problem could involve viruses, as they are the supramolecular structures best suited for overcoming the barriers at the organism and cellular levels during targeted delivery of virus's own genetic material. The largest number of gene therapy preclinical and clinical studies has centered on these viruses. However, viral vectors are associated with a risk of unexpected, and often severe, adverse events, which will actually remain a problem for quite a long time [54, 55]. Therefore, simultaneously with the design of viral vectors for the delivery of genetic material, non-viral methods that arouse increasing interest among researchers are currently under development.

One of the variants of non-viral delivery is to use polycations, which form complexes with the nucleic acids known as polyplexes. In polyplexes, DNA (or RNA) is packaged and protected against hydrolytic enzymes, so that these complexes remain sufficiently stable in biological environments. Polyplexes are non-pathogenic. Many of them are also non-immunogenic and low-toxic. By modifying the original polymers, one can obtain particles with different properties, as well as attach different functional components to impart such properties as cellular specificity or other tailored properties to the complexes [56].

It is obvious that in order to achieve these favorable properties, the polymeric vehicles within the polyplexes need to be supplemented with the aforementioned functional components. Furthermore, the polymer composition also needs to be optimized to bring the properties of the polyplexes closer to those of virions capable of delivering genetic material.

Let us consider the example of the well-known polyethyleneimine (PEI)-based polyplexes. Particles of different sizes and charges form when PEI is mixed with DNA in different proportions (expressed as the N/P ratio, where N is the number of amino groups of PEI and P is the number of DNA phosphate groups). To increase the time during which the polyplexes circulate in the blood and to reduce the toxicity of PEI, PEG is attached to PEI, yielding PEG-PEI block copolymers. Since both the N/P and PEG/PEI ratios can be variegated, the resulting problem to be solved involves finding the optimal ratio between the components in the polyplex. To solve this problem, polyplex variants with different ratios between the components were tested on 11 cell lines; transfection efficiency was assessed according to the activity of the expressed reporter gene [57]. It was discovered that the resulting dependences of transfection efficiency on the N/P and PEG/PEI ratios were non-monotonous, but that their shapes were similar for all the analyzed cells. Furthermore, importantly, maximum transfection efficiencies for different cell lines were observed at the same N/P and PEG/PEI ratios. A significant, positive correlation between the transfection efficiency and the percentage of nanoparticles within polyplexes sized 50–75 nm was revealed for all the investigated cell lines. This result, obtained for more than 10 human and animal cell lines, allows one to transfect different cell lines with maximum efficiency. However, whereas the dependences of transfection efficiency on the N/P and PEG/PEI ratios were similar, there was also a significant difference for all the analyzed cell lines: the maximum achievable transfection efficiency varied from almost 100% (HeLa, HEK293, Cloudman melanoma, and B16-F1 melanoma) to 4.4% (BT-474 cells). These differences could be attributed either to the differences in reporter gene expression or to the differences in the transport and unpacking of polyplexes observed across the cell lines. Experimental testing [57] showed that the second assumption was true: the transfection efficiency showed a positive correlation with the rate of polyplex entry into the cells and a negative correlation with the rate of their unpacking in the endocytic compartments.

Modifying block copolymers with ligands specific to internalizable receptors on the target cells impart cellular specificity to the polyplexes. Thus, the polyplexes containing αMSH acquired specificity with respect to melanoma cells overexpressing melanocortin 1 receptors (αMSH is their ligand) and showed a much greater efficiency in *in vivo* transfection of these cancer cells [58].

The size of PEI-based polyplexes ensuring the most efficient transfection (50–75 nm; see the text above) casts doubt on whether nanoparticles of this size can penetrate through nuclear pores into the nucleus of a non-dividing cell, because the known size limit is ap-

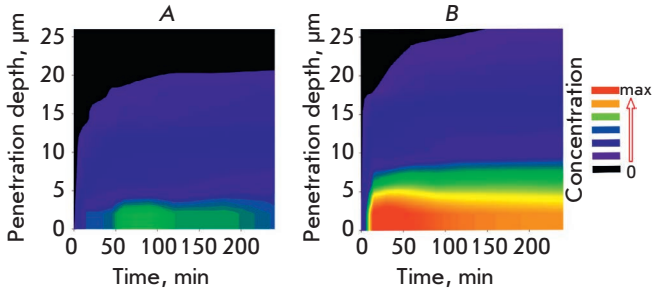


Fig. 4. Polyplex penetration into the tumor (after [62] with permission). (*A*) – control; (*B*) – after inhibition of collagen I production with losartan

proximately 40 nm even for NLS-carrying particles [59]. The experiments on transfection of cells fluorescently labeled with polyplexes showed that ~ 90% of the cells expressing the reporter gene delivered by these nanoparticles had been transfected during the cell division [60]. Therefore, it is possible to regard polyplexes as a means suitable for the transfection of actively dividing cells (first of all, the cancer ones). The average number of intact DNA molecules per nucleus of a successfully transfected cell was also estimated in this study [60]. It was found to be equal to ~ 3, which indicates that the transfection efficiency of the polyplexes was rather high. The physical properties of polyplexes also suggest that it is reasonable to use them in cancer gene therapy: thus, cancer tumors (or, to be more precise, their vessels), exhibit the so-called effect of "enhanced permeability and retention" of nanoparticles [61]. PEI-based polyplexes modified with αMSH showed different levels of efficiency in the transfection of B16-F1 and Cloudman S91 melanoma cells: the transfection efficiency was higher for B16-F1 melanoma cells compared to that for Cloudman S91 melanoma cells. As it has been shown, the reason for these differences is that B16-F1 tumors are more vascularized and their endothelium is more likely to be fenestrated, which makes the "enhanced permeability and retention effect" more pronounced [62]. Nevertheless, tumor tissues act as a barrier for polyplex nanoparticles. Although these nanoparticles penetrate tumor tissues unlike normal ones, the penetration depth is rather small (≤ 20 μm) [63] (*Fig. 4A*). Therefore, if polyplexes need to be delivered into a tumor to a greater depth, there should be some additional impact on the tumor. One of the variants allowing one to increase both the penetration depth of polyplexes and their concentration in the tumor is to modify the tumor interstitium (e.g., by inhibiting the production of collagen type I) [62, 64] (*Fig. 4B*).

PEI-based polyplexes have shown therapeutic efficacy in the case of experimentally induced tumors (S37 mouse sarcoma [65] and Cloudman S91 melanoma, clone M3 [58]). In earlier experiments involving polyplex-based mammary gland transfection in mice and sheep, the target protein was produced with their milk [66]. The same polyplexes could be used for transgenosis of early mouse and rabbit embryos [67].

CONCLUSIONS

Having summed up the results of the studies conducted on this topic, the following conclusions can be drawn.

(A) Regarding the delivery of cytotoxic agents using modular nanotransporters for cancer therapy: Modular nanotransporters (a technological platform, i.e., the core technology that serves as the basis for solving particular tasks) have been developed. This technology makes it possible to impart cellular specificity and high efficiency to a large number of antitumor agents by delivering them to the cell nucleus using the natural processes of intracellular transport.

(B) Regarding the delivery of biologically active polypeptides: Modular nanotransporters have been used to design antibody mimetics (the so-called "diving antibodies") capable of penetrating living cells and affecting the function of target molecules; furthermore, a new type of modular nanotransporters that affect the functions of transcription factors in cells both *in vitro* and *in vivo* has been designed. We believe that the approach being currently developed can lead to a breakthrough in the design of tools for the study of the function of living cells and, possibly, in the development of therapeutic agents.

(C) Regarding the delivery of genetic material using polyplexes: It has been demonstrated that polyplexes preferentially transfect dividing cells, which should be taken into account during the potential practical use of polyplexes. The efficiency of transfection using polyplexes has been demonstrated both *in vitro* and *in vivo*.

The hope is that the range of biologically active agents delivered into the cell (first of all, antibody mimetics) will be subsequently broadened: novel "diving antibodies" could be designed, and humanized MNTs for potential systemic use could be obtained. These studies have already started [68, 69]. ●

The author of this review is grateful to his colleagues A.A. Rosenkranz, Yu.V. Khramtsov, and A.V. Ulasov, for their comments made during manuscript preparation.

This work was supported by the Russian Science Foundation (grant No. 17-14-01304).

REFERENCES

1. Sobolev A.S. // Herald Russ. Acad. Sci. 2013. V. 83. № 4. P. 324–335.
2. Sobolev A.S., Jans D.A., Rosenkranz A.A. // Progress Biophys. Mol. Biol. 2000. V. 73. № 1. P. 51–90.
3. Hall E.J., Giaccia A.J. Radiobiology for the Radiologist. Philadelphia: Wolters Kluwer, 2019. 1161 p.
4. Baade I., Kehlenbach R.H. // Curr. Opin. Cell Biol. 2019. V. 58. P. 1–7.
5. Targeted Intracellular Drug Delivery by Receptor Mediated Endocytosis. / Ed. Devarajan P.V. et al. Cham: Springer, 2019. 560 p.
6. Ferrati S., Streiff A.K., Srinivasan S., Alexander J.F., Bhargava N., Peters A.M., Song N.E., Tasciotti E., Godin B., Ferrari M., et al. // Intracellular Delivery. Fundamentals and Applications. / Ed. Prokop A. Dordrecht, Heidelberg, London, New York: Springer, 2011. P. 3–56.
7. Sobolev A.S. // Bioessays. 2008. V. 30. № 3. P. 278–287.
8. Rosenkranz A.A., Lunin V.G., Gulak P.V., Sergienko O.V., Shumiantseva M.A., Voronina O.L., Gilyazova D.G., John A.P., Kofner A.A., Mironov A.F., et al. // FASEB J. 2003. V. 17. № 9. P. 1121–1123.
9. Gilyazova D.G., Rosenkranz A.A., Gulak P.V., Lunin V.G., Sergienko O.V., Khramtsov Y.V., Timofeyev K.N., Grin M.A., Mironov A.F., Rubin A.B., et al. // Cancer Res. 2006. V. 66. № 21. P. 10534–10540.
10. Slastnikova T.A., Rosenkranz A.A., Zalutsky M.R., Sobolev A.S. // Curr. Pharm. Design. 2015. V. 21. № 9. P. 1227–1238.
11. Sobolev A.S., Aliev R.A., Kalmykov S.N. // Russian Chem. Rev. 2016. V. 85. № 9. P. 1011–1032.
12. Slastnikova T., Rosenkranz A., Morozova N., Vorontsova M., Petriev V., Lupanova T., Ulasov A., Zalutsky M., Yakubovskaya R., Sobolev A. // Internat. J. Nanomed. 2017. V. 12. P. 395–410.
13. Sobolev A.S. // Encyclopedia of Cancer. / Ed. Schwab M. Berlin, Heidelberg: Springer, 2017. P. 2891–2894.
14. Kamaletdinova T.R., Rosenkranz A.A., Ulasov A.V., Khramtsov Y.V., Tsvetkova A.D., Georgiev G.P., Sobolev A.S. // Doklady Biochem. Biophys. 2018. V. 479. № 1. P. 95–97.
15. Sobolev A.S. // Front. Pharmacol. 2018. V. 9. P. 952.
16. Rosenkranz A.A., Slastnikova T.A., Georgiev G.P., Zalutsky M.R., Sobolev A.S. // Nucl. Med. Biol. 2020. V. 80. P. 45–56.
17. Haddad J., Slika S., Mahfouz R. // Meta Gene. 2017. V. 11. P. 157–163.
18. Xu M.J., Johnson D.E., Grandis J.R. // Cancer Metastasis Rev. 2017. V. 36. № 3. P. 463–473.
19. Slastnikova T.A., Rosenkranz A.A., Gulak P.V., Schiffelers R.M., Lupanova T.N., Khramtsov Y.V., Zalutsky M.R., Sobolev A.S. // Internat. J. Nanomed. 2012. V. 7. P. 467–482.
20. Rosenkranz A.A., Slastnikova T.A., Karmakova T.A., Vorontsova M.S., Morozova N.B., Petriev V.M., Abrosimov A.S., Khramtsov Y.V., Lupanova T.N., Ulasov A.V., et al. // Front. Pharmacol. 2018. V. 9. P. 1331.
21. Rosenkranz A.A., Slastnikova T.A., Durymanov M.O., Sobolev A.S. // Biochemistry (Moscow). 2013. V. 78. № 11. P. 1228–1237.
22 Khramtsov Y.V., Rokitskaya T.I., Rosenkranz A.A., Trusov G.A., Gnuchev N.V., Antonenko Y.N., Sobolev A.S. // J. Controlled Release. 2008. V. 128. № 3. P. 241–247.
23. Slastnikova T.A., Rosenkranz A.A., Khramtsov Y.V., Karyagina T.S., Ovechko S.A., Sobolev A.S. // Drug Design. Dev. Therapy. 2017. V. 11. P. 1315–1334.
24. Rosenkranz A.A., Slastnikova T.A., Khramtsov Y.V., Karyagina T.S., Georgiev G.P., Sobolev A.S. // Doklady Biochem. Biophys. 2017. V. 473. № 1. P. 85–87.
25. Khramtsov Y.V., Vlasova A.D., Vlasov A.V., Rosenkranz A.A., Ulasov A.V., Ryzhykau Y.L., Kuklin A.I., Orekhov A.S., Eydlin I.B., Georgiev G.P., Gordeliy V.I., Sobolev A.S. // Acta Cryst., 2020. V. D76. № 12. P. 1270–1279.
26. Karyagina T.S., Ulasov A.V., Slastnikova T.A., Rosenkranz A.A., Lupanova T.N., Khramtsov Y.V., Georgiev G.P., Sobolev A.S. // Front. Pharmacol. 2020. V. 11. P. 176.
27. Slastnikova T.A., Koumarianou E., Rosenkranz A.A., Vaidyanathan G., Lupanova T.N., Sobolev A.S., Zalutsky M.R. // EJNMMI Res. 2012. V. 2. № 1. P. 1–10.
28. Koumarianou E., Slastnikova T.A., Pruszynski M., Rosenkranz A.A., Vaidyanathan G., Sobolev A.S., Zalutsky M.R. // Nucl. Med. Biol. 2014. V. 41. № 6. P. 441–449.
29. Ku A., Facca V.J., Cai Z., Reilly R.M. // EJNMMI Radiopharm. Chem. 2019. V. 4. P. 1.
30. Vaidyanathan G., Zalutsky M.R. // Curr. Radiopharm. 2011. V. 4. № 4. P. 283–294.
31. Slastnikova T., Rosenkranz A., Lupanova T., Gulak P., Gnuchev N., Sobolev A. // Doklady Biochemistry and Biophysics. 2012. V. 446. № 1. P. 235–237.
32. Mironov A.F., Grin M.A., Pantushenko I.V., Ostroverkhov P.V., Ivanenkov Y.A., Filkov G.I., Plotnikova E.A., Karmakova T.A., Starovoitova A.V., Burmistrova N.V., et al. // J. Med. Chem. 2017. V. 60. № 24. P. 10220–10230.
33. Buchegger F., Perillo-Adamer F., Dupertuis Y.M., Delaloye A.B. // Eur. J. Nucl. Med. Mol. Imaging. 2006. V. 33. № 11. P. 1352–1363.
34. Baidoo K.E., Yong K., Brechbiel M.W. // Clin. Cancer Res. 2013. V. 19. № 3. P. 530–537.
35. Chadwick K.H., Leenhouts H.P. // Molecular Theory of Radiation Biology. Monographs on Theoretical and Applied Genetics. V. 5 / Ed. Frankel R. et al. Berlin, Heidelberg, New York: Springer-Verlag, 1981. 377 p.
36. Kassis A.I. // Rad. Protect. Dosimetry. 2011. V. 143. № 2–4. P. 241–247.
37. Raghavan R., Howell R.W., Zalutsky M.R. // Biomed. Phys. Engin. Express. 2017. V. 3. P. 035005.
38. Falzone N., Lee B.Q., Able S., Malcolm J., Terry S., Alayed Y., Vallis K.A. // J. Nucl. Med. 2019. V. 60. № 2. P. 250–258.
39. Claesson A.K., Stenerlöw B., Jacobsson L., Elmroth K. // Radiat. Res. 2007. V. 167. № 3. P. 312–318.
40. Rosenkranz A.A., Vaidyanathan G., Pozzi O.R., Lunin V.G., Zalutsky M.R., Sobolev A.S. // Internat. J. Radiat. Oncol. Biol. Physics. 2008. V. 72. № 1. P. 193–200.
41. Howell R.W. // Med. Phys. 1992. V. 19. № 6. P. 1371–1383.
42. Balagurumoorthy P., Xu X., Wang K., Adelstein S.J., Kassis A.I. // Internat. J. Radiat. Biol. 2012. V. 12. P. 998–1008.
43. Vorontsova M.S., Morozova N.B., Karmakova T.A., Pankratov A.A., Andreeva T.N., Plotnikova E.A., Rosenkranz A.A., Slastnikova T.A., Lupanova T.N., Sobolev.A.S. // Russian Journal of Biotherapy. 2016. V. 15. № 1. P. 20–21. (In Russian)
44. Vorontsova M.S., Morozova N.B., Karmakova T.A., Rosenkranz A.A., Slastnikova T.A., Petriev V.M., Smoryzanova O.A., Tischenko V.K., Yakubovskaya R.I., Kaprin A.D., et al. // AIP Conf. Proc. V. 1882. AIP Publ., 2017. P. 0200781-1-020078-1-4.
45. Vorontsova M.S., Morozova N.B., Karmakova T.A., Tischenko V.K., Smoryzanova O.A., Petriev V.M., Rosenkranz A.A., Slastnikova T.A., Lupanova T.N., Yakubovskaya

R.I., et al. // Research'n Practical Medicine Journal. 2017. V. 4 (special issue). P. 37. (In Russian)
46. Sobolev A.S., Rosenkranz A.A., Slastnikova T.A., Karmakova T.A., Vorontsova M.S., Morozova N.B., Petriev V.M., Lupanova T.N., Khramtsov Y.V., Ulasov A.V. et al. // Research'n Practical Medicine Journal. 2018. V. 5. S2. P. 266. (In Russian)
47. Karimian A., Ahmadi Y., Yousefi B. // DNA Repair (Amst.). 2016. V. 42. P. 63–71.
48. Bruning J.B., Shamoo Y. // Structure. 2004. V. 12. № 12. P. 2209–2219.
49. Liddle P., Lafon-Hughes L., Di Tomaso M.V., Reyes-Abalos A.L., Jara J., Cerda M., Hartel S., Folle G.A. // Chromosome Res. 2014. V. 22. № 4. P. 463–481.
50. Sobolev A.S., Ulasov A.V., Georgiev G.P., Khramtsov Y.V., Rosenkranz A.A., Slastnikova T.A., Lupanova T.N., Karyagina T.S. The development of new types of modular nanotransporters carrying biologically active macromolecules as a tool for selective adjustment of specified functions of living cells. RSF Grant #17-14-01304. 2019. URL: https://rscf.ru/contests/search-projects/17-14-01304/ (Access date: 26.05.2020).
51. Cuadrado A., Rojo A.I., Wells G., Hayes J.D., Cousin S.P., Rumsey W.L., Attucks O.C., Franklin S., Levonen A.L., Kensler T.W., et al. // Nat. Rev. Drug Discovery. 2019. V. 18. № 4. P. 295–317.
52. Sekhar K.R., Freeman M.L. // Free Rad. Biol. Med. 2015. V. 88 (Pt B). P. 268–274.
53. Hayes J.D., Dinkova-Kostova A.T. // Trends Biochem. Sci. 2014. V. 39. № 4. P. 199–218.
54. van der Eb M.M., de Leeuw B., van der Eb A.J., Hoeben R.C. // Suicide Gene Therapy: Methods and Reviews / Ed. Springer C. J. Totowa: Humana Press, 2004. P. 479–490. doi: 10.1385/1-59259-429-8:479.
55. Liu T.C., Kirn D.H. // Gene Therapy for Cancer / Ed. Hunt K.K., et al. Totowa: Humana Press, 2007. P. 351–385.
56. Rosenkranz A.A., Sobolev A.S. // Russ. Chem. Bull. 2015. V. 64. № 12. P. 2749–2755.
57. Ulasov A.V., Khramtsov Y.V., Trusov G.A., Rosenkranz A.A., Sverdlov E.D., Sobolev A.S. // Mol. Therapy. 2011. V. 19. № 1. P. 103–112.
58. Durymanov M.O., Beletkaia E.A., Ulasov A.V., Khramtsov Y.V., Trusov G.A., Rodichenko N.S., Slastnikova T.A., Vinogradova T.V., Uspenskaya N.Y., Kopantsev E.P., et al. // J. Controlled Release. 2012. V. 163. № 2. P. 211–219.
59. Panté N., Kann M. // Mol. Biol. Cell. 2002. V. 13. № 2. P. 425–434.
60. Durymanov M.O., Yarutkin A.V., Khramtsov Y.V., Rosenkranz A.A., Sobolev A.S. // J. Controlled Release. 2015. V. 215. P. 73–81.
61. Wong A.D., Ye M., Ulmschneider M.B., Searson P.C. // PLoS One. 2015. V. 10. № 5. P. e0123461.
62. Durymanov M.O., Yarutkin A.V., Bagrov D.V., Klinov D.V., Kedrov A.V., Chemeris N.K., Rosenkranz A.A., Sobolev A.S. // J. Controlled Release. 2016. V. 232. P. 20–28.
63. Durymanov M.O., Slastnikova T.A., Kuzmich A.I., Khramtsov Y.V., Ulasov A.V., Rosenkranz A.A., Egorov S.Y., Sverdlov E.D., Sobolev A.S. // Biomaterials. 2013. V. 34. № 38. P. 10209–10216.
64. Durymanov M.O., Rosenkranz A.A., Sobolev A.S. // Theranostics. 2015. V. 5. № 9. P. 1007–1020.
65. Alekseenko I.V., Snezhkov E.V., Chernov I.P., Pleshkan V.V., Potapov V.K., Sass A.V., Monastyrskaya G.S., Kopantzev E.P., Vinogradova T.V., Khramtsov Y.V., et al. // J. Translat. Med. 2015. V. 13. P. 78.
66. Sobolev A.S., Rosenkranz A.A., Smirnova O.A., Nikitin V.A., Neugodova G.L., Naroditsky B.S., Shilov I.N., Shatski I.N., Ernst L.K. // J. Biol. Chem. 1998. V. 273. № 14. P. 7928–7933.
67. Ivanova M.M., Rosenkranz A.A., Smirnova O.A., Nikitin V.A., Sobolev A.S., Landa V., Naroditsky B.S., Ernst L.K. // Mol. Reprod. Dev. 1999. V. 54. № 2. P. 112–120.
68. Ulasov A.V., Khramtsov Y.V., Lupanova T.N., Tsvetkova A.D., Rosenkranz A.A., Slastnikova T.A., Georgiev G.P., Sobolev A.S. // Doklady Biochem. Biophys. 2018. V. 479. № 1. P. 62–65.
69. Khramtsov Y.V., Ulasov A.V., Tsvetkova A.D., Rosenkranz A.A., Georgiev G.P., Sobolev A.S. // Doklady Biochem. Biophys. 2017. V. 472. № 1. P. 81–83.

The DPF Domain As a Unique Structural Unit Participating in Transcriptional Activation, Cell Differentiation, and Malignant Transformation

N. V. Soshnikova[1*], A. A. Sheynov[1], Eu. V. Tatarskiy[1], S. G. Georgieva[1,2*]
[1]Institute of Gene Biology Russian Academy of Sciences, Moscow, 119334 Russia
[2]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia
*E-mail: nsoshnikova@genebiology.ru; sofia.georgieva2021@gmail.com
Received July 17, 2020; in final form, September 28, 2020
DOI: 10.32607/actanaturae.11092
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **The DPF (double PHD finger) domain consists of two PHD fingers organized in tandem. The two PHD-finger domains within a DPF form a single structure that interacts with the modification of the N-terminal histone fragment in a way different from that for single PHD fingers. Several histone modifications interacting with the DPF domain have already been identified. They include acetylation of H3K14 and H3K9, as well as crotonylation of H3K14. These modifications are found predominantly in transcriptionally active chromatin. Proteins containing DPF belong to two classes of protein complexes, which are the transcriptional coactivators involved in the regulation of the chromatin structure. These are the histone acetyltransferase complex belonging to the MYST family and the SWI/SNF chromatin-remodeling complex. The DPF domain is responsible for the specificity of the interactions between these complexes and chromatin. Proteins containing DPF play a crucial role in the activation of the transcription of a number of genes expressed during the development of an organism. These genes are important in the differentiation and malignant transformation of mammalian cells.**
KEYWORDS **DPF domains, tandem PHD, MOZ and MORF histone acetyltransferases, DPF1, DPF2, DPF3, PHF10, BAF, PBAF.**

INTRODUCTION

The DPF (double PHD finger) domain belongs to the group of PHD (plant homeodomains) fingers, widely found in mammals. In humans, there are about two hundred PHD-containing proteins. The PHD domains have a zinc-finger (Zn-finger) structure. They consist of two antiparallel beta sheets and a C-terminal alpha helix. These structural elements are stabilized by two zinc ions coordinated by the Cys4-His-Cys3 motif [1, 2]. Although the primary structure of PHD fingers is quite diverse, their secondary structure, described for the first time in 2000, is highly conserved [3].

PHD fingers are mainly found in proteins that interact with the N-terminal fragments of histones; they regulate gene expression [4]. PHD fingers bind to the N-terminal regions of histone H3, which can exist in various modifications [5, 6].

Some proteins contain only one PHD-finger domain, while others may contain several, consecutive PHD fingers that function either independently of each other or in concert.

The DPF domain is a tandem of PHD fingers with a face-to-back orientation. Two domains form a single structure interacting with the N-terminal fragments of histones in a manner different from that for independent PHD-finger domains. Our review focuses on proteins containing the DPF domains, their organization, molecular mechanisms of recognition of histone tails, the impact on gene expression, as well as their role in mammalian development and oncogenesis.

PROTEINS AND COMPLEXES CONTAINING THE DPF DOMAIN

Proteins containing the DPF domain mostly are the subunits of large protein complexes that determine and change the epigenetic status of chromatin [6]. The specific function of these complexes is ensured by precise recognition of the epigenetic modifications of

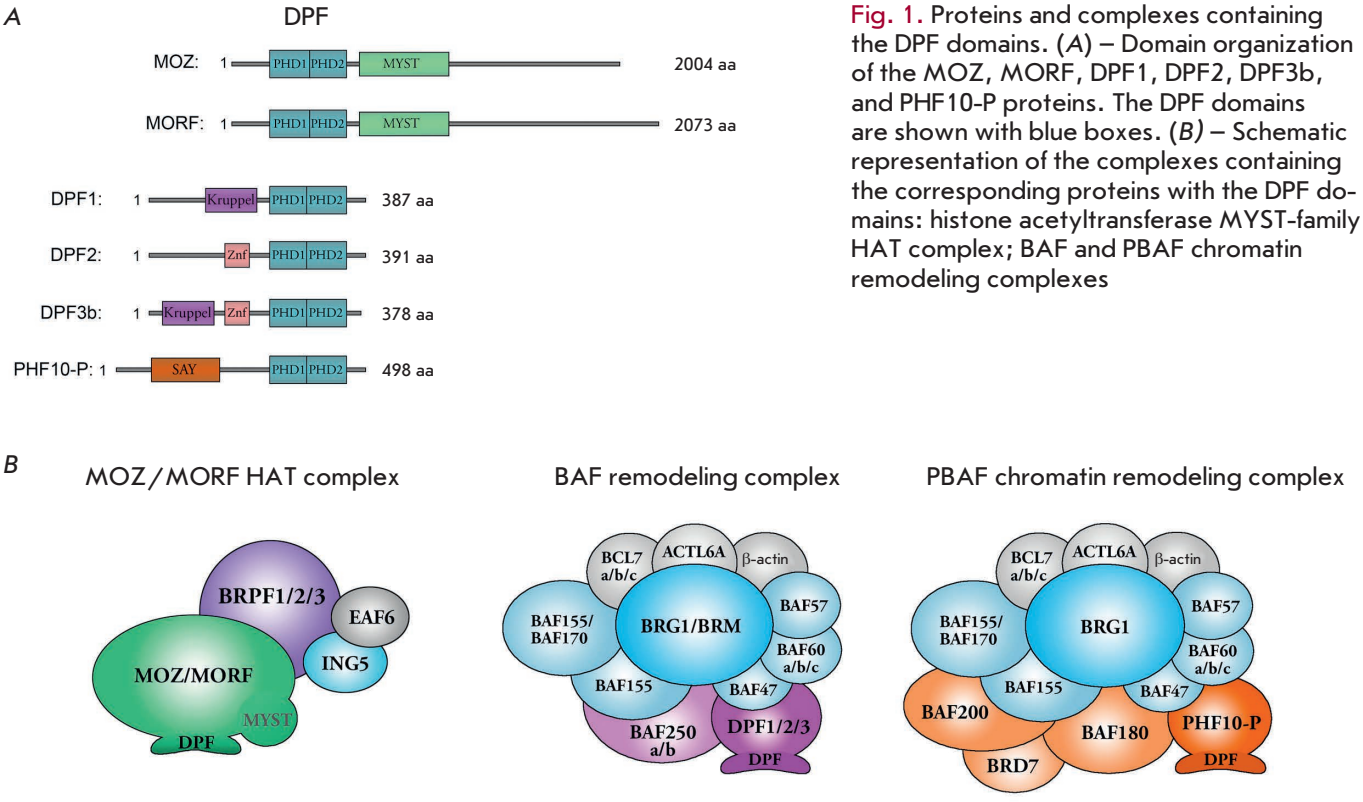


Fig. 1. Proteins and complexes containing the DPF domains. (*A*) – Domain organization of the MOZ, MORF, DPF1, DPF2, DPF3b, and PHF10-P proteins. The DPF domains are shown with blue boxes. (*B*) – Schematic representation of the complexes containing the corresponding proteins with the DPF domains: histone acetyltransferase MYST-family HAT complex; BAF and PBAF chromatin remodeling complexes

chromatin, most of which are the modified N-terminal fragments of histones. Many subunits of the complexes contain different domains that interact with histones. For instance, these domains include the bromodomain (TAF1 and BAF180 proteins), chromodomain (CHD1 protein), Tudor domain (Uhrf1 protein), and their combinations. Each of the domains can recognize a specific modification of the N-terminal histone sequence. Acting together in a combinatorial manner, they increase the number of chromatin marks that are recognized by the full complex.

The DPF domain is found in two groups of proteins. The first group includes the histone lysine acetyltransferases MOZ and MORF, while the other one is represented by proteins of the SWI/SNF chromatin remodeling complex (*Fig. 1A*). Acetyltransferases MOZ (also known as MYST3/KAT6a) and MORF (MYST4/KAT6b) are paralogs. They are alternatively contained within the MYST-family histone acetyltransferase (HAT) complex, which acetylates the N-termini of histones [7, 8] (*Fig. 1B*). The HAT complex is a transcriptional coactivator that resides in open, actively transcribed chromatin. MORF and MOZ contain the MYST domain, which acetylates the lysine residues in the N-terminal sequences of histone H3 (H3K9, H3K14ac, and H3K23). MYST-family HAT is responsible for the hyperacetylation of chromatin regions, which promotes activation of the respective genes [8–10].

Another group of proteins containing DPF is found in the SWI/SNF chromatin remodeling complex (its BAF and PBAF subfamilies) (*Fig. 1B*). This group includes the DPF1 (also known as BAF45b), DPF2 (REQ, or BAF45d), and DPF3 (BAF45c) (which are also called d4-family proteins), as well as PHF10 (BAF45a) proteins (*Fig. 1A*). The SWI/SNF complex is involved in the regulation of gene transcription, repair, and replication. Due to the ATPase activity of the major subunit of BRG1 and its homolog BRM, the complexes displace nucleosomes along the DNA strand or transfer the nucleosome to another DNA strand, remove H2A and H2B, as well as replace the canonical histone with its variant [11].

As mentioned above, the SWI/SNF family involves two types of complexes: BAF and PBAF (*Fig. 1B*). They share identical proteins in their core parts, which displace nucleosomes along the DNA strand. However, these complexes differ in the proteins within specific modules that are responsible for the interactions with chromatin. DPF proteins are components of the specific modules of the BAF and PBAF complexes and are involved in determining the specificity of complex

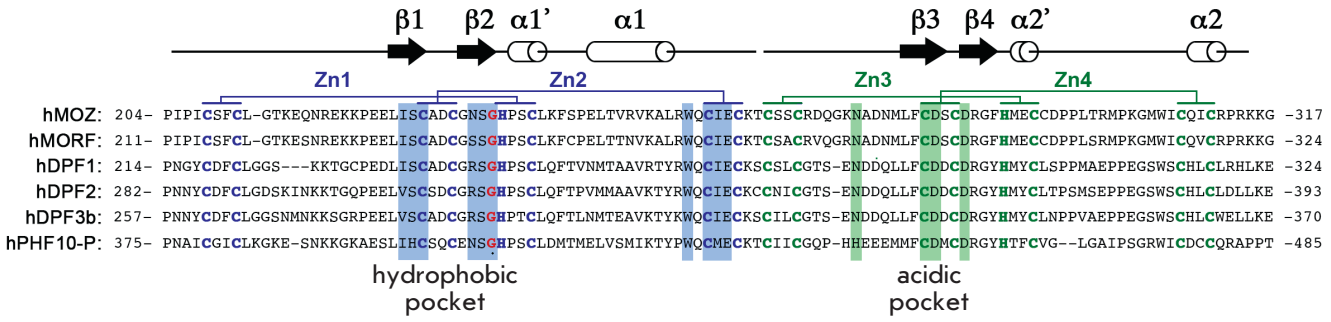


Fig. 2. Alignment of the amino acid sequences of the DPF motifs of the MOZ, MORF, DPF1, DPF2, DPF3b, and PHF10-P human proteins. Schematic representation of the secondary structures of PHD1 and PHD2 is shown above the sequences. Cysteine and histidine residues coordinating Zn ions in PHD1 and PHD2 are indicated in blue and green, respectively. Homologous amino acids in PHD1 that form a hydrophobic pocket and bind H3K14ac/cr are highlighted in blue. Homologous amino acids that form an acidic pocket and bind to the first to fourth N-terminal amino acids of histone H3 are highlighted in green

binding to chromatin. The DPF domains present in these complexes are also involved in performing this function.

THE STRUCTURAL FOUNDATIONS FOR HISTONE RECOGNITION AND THE SPATIAL ARCHITECTURE OF THE DPF DOMAINS

The DPF domains of the MOZ, MORF, DPF1, DPF2, DPF3b, and PHF10 proteins are highly homologous; their secondary structures are formed by the same key amino acids (*Fig. 2*). Therefore, the data obtained for the DPF domains of each of these proteins are likely to be true for the DPF domains of other proteins belonging to this group.

The structure of each of the two PHD domains of the MOZ, MORF, DPF1, DPF2, DPF3b, and PHF10 proteins is typical of zinc finger domains. It consists of two antiparallel beta sheets, followed by an alpha helix, which are coordinated by two zinc atoms via the Cys4-His-Cys3 motif (*Fig. 2*). However, as shown for the MOZ protein, two PHD fingers are associated with each other in a face-to-back manner through the interaction between E247 and R251 in the alpha helix of the first PHD finger, as well as interaction between S283 and R286 in the third and fourth beta sheets of the second PHD finger.

The carboxyl and carbonyl groups of E247 form two hydrogen bonds with two water molecules, which interact with the carboxyl and carbonyl hydrogen atoms in S283. R251 interacts with the nitrogen atom in the R286 side chain in a similar way. Thus, these polar interactions localize the two PHD fingers, which form a unique globular structure [12]. The DPF of the DPF2, DPF3b, and MORF proteins also form a similar integral structural unit [13, 14].

THE DPF DOMAINS INTERACT WITH ACYLATED H3K14 AND H3K9

The DPF modules of the MOZ, MORF, DPF2, and DPF3b proteins interact with unmodified N-terminal fragments of histone H3. Acetylation of H3K14 and H3K9 increases the binding constant threefold [12–15]. Methylation of H3K9me3 does not affect binding, while methylation of H3K4me3 severely inhibits DPF binding to histones (*Fig. 3*) [16]. The DPF domain of these proteins also weakly interacts with the N-terminus of histone H4. Acetylation of the H4K5, H4K8, H4K12, and H4K16 lysine residues abolishes the interaction between histone H4 and the DPF domains of MOZ and MORF (*Fig. 3*) [16].

A short time later, it was shown that the DPF domains of the MOZ and DPF2 proteins can interact with crotonylated Lys14 in histone H3 (H3K14cr) [17]. The crotonyl group has a more hydrophobic side chain residue and forms a planar spatial structure. The DPF domains of acetyltransferase MORF interact with other acyl groups, such as the butyrylated (H3K14bu), succinylated (H3K14su), and 2-hydroxyisobutyrylated H3K14 (H3K14hib) lysine residues. These groups also have longer hydrophobic side chains compared to those in acetylated modifications [18].

The molecular mechanism of interaction between DPF and various modifications of histone H3 has been studied using crystal structures of the DPF domains with either unmodified histone tails or various modifications (H3K14ac/cr). Both PHD domains form a single structural unit and bind the N-terminal fragment of histone H3 to one of the following modifications: H3K14ac, cr, or bu [12, 17, 18]. Among these, the crotonyl group is the preferential modification for binding the DPF domains in MORF. MORF DPF binds

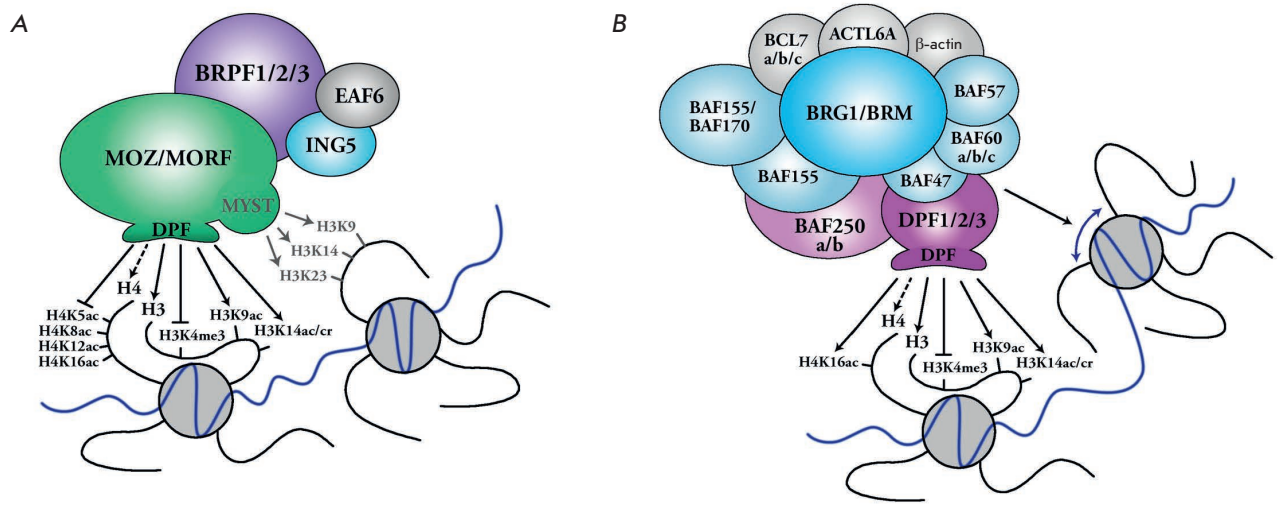


Fig. 3. Schematic representation of the activities of the MYST-family HAT (*A*) and BAF (*B*) complexes containing either the MOZ/MORF or DPF1-3 protein. The interaction between the DPF domains and histone modifications (black arrows), as well as the histone acetyltransferase activity of the MYST complex (gray arrows) (*A*) and the chromatin remodeling activity of the BAF complex (blue arrow) (*B*), is shown

to H3K14cr three times more strongly than to H3K14ac [19]. It has been shown quite recently that a small DPF region in MORF (within the R306–K309 residues) interacts with DNA. These interactions are determined by the H3K14cr modification and enhance the binding of MORF to the nucleosome [19].

THE MECHANISM THROUGH WHICH THE DPF DOMAIN RECOGNIZES POST-TRANSLATIONAL MODIFICATIONS OF HISTONES

The first PHD domain of the MOZ, MORF, DPF2, and DPF3b proteins is a unique zinc-finger type domain. It contains a hydrophobic pocket for binding acylated lysine (*Fig. 2*). Although the acylated H3K14 occupies the same pocket within the PHD1 domain of MOZ and DPF3b, different amino acids are involved in the interaction between H3K14 and DPF in the MOZ and DPF3b proteins [16]. However, the hydrophobic pocket within the beta-2 sheet in the first PHD finger is a common structural feature required for the binding of the H3K14ac, H3K14cr, and H3K14bu modifications [17, 18]. In the MOZ and MORF proteins, the hydrophobic pocket is formed by the amino-acid residues N235–G237 of the beta-2 sheet, I228–C230 of the beta-1 sheet, and amino-acid residues S210 (S217), F211 (F218), L242 (L249), W257 (W264), C259 (C266), I260 (I267), and E261, which coordinate the zinc ion (*Fig. 2*).

G237 is the most important amino-acid residue for the formation of this pocket, which recognizes acetylated and crotonylated groups (*Fig. 2*). This glycine residue is present in the DPF domains of the MOZ, MORF, DPF1, DPF2, DPF3b, and PHF10 proteins, indicating that the DPF domains in all these proteins can interact with acetylated, crotonylated, and butyrylated H3K4 (H3K4ac/cr/bu) (*Fig. 2*) [12, 17]. The F211 (F218) residue is responsible for the differences in the interaction between H3K14cr and H3K14bu as it can form π–π interactions between the aromatic ring of phenylalanine and the C=C double bond of the crotonyl group [19].

The second PHD domain of the MOZ and MORF proteins is organized in such a way that the first four H3K14ac/cr/bu residues of the peptide bind to the "acidic" pocket within the beta-1 sheet of this PHD2 domain. It is important that the amino group residues R2 and K4 are not methylated. The side chain of the H3R2 peptide is kept in its place by five hydrogen bonds of the DPF module of the MOZ protein with the C281, D282, and D285 residues. Meanwhile, the E261 and N274 residues form hydrogen bonds with the amino group in H3K4. As a result of such spatial restriction of the R2 and K4 residues, any methylation breaks the bond between DPF and H3, while binding of acetylated lysine residues preferentially occurs [12, 14, 17]. There is increasing experimental evidence that the second PHD domain of d4-family proteins is organized according to the same principle and is unable to recognize methylated H3K4 [5].

These data have been confirmed by *in vivo* experiments: it was shown that MOZ is associated with H3K14ac-rich chromatin and does not bind to

H3K4me3-marked chromatin [16]. Crotonylated H3K14cr marks were found in the same genes (*Hox9A*, *Hox7*, and *Hox5A*) to which histone acetyltransferase MOZ binds [17]. There is still no explanation for the presence of two mutually exclusive modifications, H3K14ac and H3K14cr, in the same genes. Apparently, the presence of a H3K14 modification strongly depends on the metabolic pathways active in the cell, since the percentage of crotonylated or butyrylated histones is directly related to the amount of the respective acyl-CoA available for involvement in the metabolic pathways [17, 20, 21]. Therefore, HATs can switch between substrates to change the types of acylation profile of modified histones.

THE DPF DOMAIN IN TRANSCRIPTIONAL REGULATION

As mentioned above, the DPF domains bind to the acylated (acetylated, crotonylated, and butyrylated) tail of histone H3 and act as the so-called readers: i.e., the proteins recognizing these histone modifications. The H3K14ac and H3K9ac modifications interacting with the DPF are characteristic of transcriptionally active chromatin. Histones highly enriched in these modifications are located in gene promoters and enhancers [22, 23]. Crotonylation of histone H3 (H3K14cr) is also found in transcriptionally active chromatin.

The mechanism of regulation of transcription and the epigenetic state of chromatin by HAT complexes involving MOZ and MORF has been shown for the *HOX9A* gene. The complexes are recruited to chromatin after their interaction with certain transcriptional activators (e.g., RUNX and P53) [24, 25] or due to the interaction of other subunits with various modifications [26–28]. The DPF domains of MOZ and MORF promote the localization of the complex in the H3K14ac-containing regions [12, 14], while the DNA-binding DPF motifs can stabilize these interactions with the nucleosome [19]. H3K14 acetylation is predominantly performed by histone acetyltransferase HBO1, which also contains the MYST domain [29]. However, the MOZ/MORF proteins can also induce this modification [30]. Acetylation of H3K23 and H3K9 is mainly carried out by the MYST domain of the MORF protein [31, 32] and can occur either in the same nucleosome or in an adjacent one [19]. Acetylation of the adjacent nucleosome contributes to changes in the complex localization and its transfer to the adjacent nucleosome. A similar mechanism drives the spread of histone marks between nucleosomes, thus resulting in the formation of hyperacetylated chromatin domains. The recruitment of the HAT complex to some *HOX* genes (*HoxA9*, *HoxA7*, *HoxA5*, and *HoxD13*) and formation of hyperacetylated domains in the promoters of these genes enhances their expression [12, 16, 17, 21]. Later, a genome-wide analysis of the ENCODE database revealed that the H3K23ac and H3K14ac modifications are co-localized and that the promoter regions of highly transcribed genes are rich in these modifications in IMR90, hESC, and HMEL cell lines [19, 33].

The SWI/SNF family complexes, which comprise another group of DPF-containing proteins, are more varied in terms of their protein composition than MYST acetyltransferase complexes. Combinations of different subunits are responsible for the specific composition of the complex, where the unique pattern of the domains binding DNA or histones in these complexes localizes the remodeling complex within specific chromatin regions. The BAF and PBAF complexes are recruited to certain loci by transcriptional activators and are involved in nucleosome remodeling: they shift the nucleosomes along the DNA strands and remove histones H2A and H2B [34]. Remodeling complexes abundantly occur in enhancers, which also supports the fact that the remodeling complexes are involved in transcriptional activation [35, 36].

The DPF1, DPF2, DPF3b, and PHF10 proteins also act as transcriptional coactivators. DPF3b and DPF3a bind to the NF-kB activator and are recruited together, within the BAF complex, to the *IL-6* promoter in response to TNF-α stimulation [37]. The PHF10 protein in the PBAF complex coactivates the transcription of various genes [38, 39]. Direct evidence has been obtained that the DPF domain of the PHF10 protein is required for transcriptional activation, since a protein lacking DPF cannot activate transcription [38, 40]. Interestingly, the PHF10 isoform lacking the DPF domain is also found in cells [39]. The isoform containing the DPF domain is involved in transcriptional activation, while the isoform lacking DPF is needed to maintain a steady-state transcriptional level after activation [40]. Therefore, the DPF domain in the PHF10 protein is a potent transcription coactivator.

DPF IN THE REGULATION OF CELL DIVISION AND DIFFERENTIATION. ITS ROLE IN TISSUE DEVELOPMENT IN MAMMALS

Since their discovery, MOZ and MORF have been associated with the regulation of cell proliferation. The interaction of MOZ with PML and P53 in MCF7 breast cancer cells and mouse embryonic fibroblasts (MEFs) was shown to result in acetylation of P53, followed by activation of p21 expression. The p21 protein is a cell cycle inhibitor that suppresses the cyclin E/CDK2 complex. The cyclin E/CDK2 complex phosphorylates a number of factors promoting gene activation in the G1/S checkpoint of the cell cycle. When unable to trigger gene expression for the G1/S transition, cells exit the cell cycle and stop dividing. Therefore, MOZ and

MORF inhibit proliferation and implement the subsequent scenario of cell transition to senescence [25, 41]. Meanwhile, MOZ maintains the expression level for some genes coding for senescence inhibition in the INK4/ARF locus via the H3K9ac modification [42, 43]. As described above, MOZ and MORF regulate the expression of many *HOX* genes responsible for organism development and differentiation. This partially occurs through their interaction with factor BMI1, which has been demonstrated at the genetic level [44].

MOZ plays an important role in maintaining the pool of embryonic hematopoietic stem cells in mammals. Knockout mice died at the E14.5 embryonic stage and exhibited manifestations of liver and hematopoietic pathologies [45]. The *MOZ* gene is also required for the normal development of blood B cells and progression of c-MYC-induced lymphoma. MOZ interacts with AML1 and PU.1, two important hematopoietic factors, and acts as their coactivator by ensuring accurate expression of the respective genes [46, 47].

The recent genome-wide studies of patients with congenital abnormalities (severe speech disorders, hypotension, and facial dysmorphism) revealed mutations in the *MOZ* gene [48]. MORF is actively involved in the development of neural and bone tissue. Mice with a minimal amount of MORF RNA (~10%) had dwarfism, craniofacial disorders, and cerebral defects [49]. MORF plays a crucial role in the regulation of neuronal stem cells; it is required to maintain neurogenesis in adult mice [50]. Thus, although MOZ and MORF are interchangeable *in vitro*, they play different roles *in vivo*: MOZ is important for hematopoiesis, while MORF is involved in neurogenesis and osteogenesis.

The DPF proteins within the SWI/SNF complexes are required for neurogenesis in mammals. DPF3b, a component of the BAF chromatin remodeling complex, plays a crucial role in the differentiation of muscle and cardiac tissues [51]. PHF10 is expressed in nerve cell precursors from the early embryonic stages. Its expression reduces after birth. PHF10 can maintain the proliferation of neuronal progenitor cells. As a component of the PBAF complex, it binds to the promoters of the signaling pathway genes regulating neuronal proliferation and differentiation: the Notch, SHH, and various transcription factors. Other DPFs (DPF1, 2, and 3) begin to be expressed in the mouse brain at later stages, starting from E13, and are unable to maintain proliferation of neuronal cells [38]. DPF1 is presumably important for the functioning of adult neurons, since it is expressed tissue-specifically only in the brain of adult mammals. DPF2 is also involved in the development and function of the nervous system. Single-nucleotide substitutions affecting the sequence of DPF domains and disrupting the binding of DPF2 to acetylated H3 were found in patients with the Coffin-Siris syndrome. This disease manifests itself as cognitive dysfunction and intellectual impairment of varied severity, coarse facial features, and brain abnormalities such as hypoplasia and agenesis of the corpus callosum [52].

The DPF2 (the BAF complex) and PHF10 (the PBAF complex) proteins are expressed in hematopoietic progenitor cells of E14.5 mouse embryos and regulate their differentiation [53]. DPF2 inhibits myeloid differentiation of hematopoietic progenitor cells. Its DPF domain is responsible for the recruitment of DPF2 and the entire BAF complex to specific acetylated chromatin loci: the binding sites of the transcription factor RUNX1. This factor promotes the myeloid differentiation of progenitors. DPF2 knockdown in $CD34^+$ cells reduces the expression of the genes involved in mitosis and cell cycle regulation; it also disrupts the transcription of the genes involved in differentiation [15].

The homozygous PHF10 knockout causes death of mouse embryos (E19), while conditional knockout in the hematopoietic cells of an adult mouse causes significant depletion of myeloid precursors (granulocytes). An analysis of RNA isolated from these cells showed that PHF10 significantly affects the expression of cell cycle genes [53]. A study performed using a model HL-60 cell line, which is capable of myeloid differentiation, and terminally differentiated human neutrophils showed that PHF10 isoforms containing the DPF domain play a crucial role in the maintenance of proliferating myeloid progenitors. These isoforms are also required for the activation of specific myeloid genes whose expression is activated during differentiation. In mature neutrophils, transcription of specific genes is maintained by PHF10 isoforms lacking DPF [40].

THE ROLE OF PROTEINS CONTAINING DPF DOMAINS IN THE MALIGNANT TRANSFORMATION OF CELLS

Mutant proteins containing the DPF domain are often found in tumor cells. Abnormal expression of MOZ and MORF is often associated with different types of leukemia. Chromosomal regions where the *MOZ* and *MORF* genes are located undergo various translocations, giving rise to chimeric proteins [10]. Myeloid leukemia is accompanied by translocations between the *MOZ* and *CBP* genes [54]. Acute monocytic leukemia is associated with translocations between the *MOZ* and *P300* genes [55]. Translocations between *MOZ* and *LEUTX* [56], as well as other genes [57], are observed in acute myeloid leukemia. *MORF* can also undergo translocation to form chimeric proteins. This translocation gives rise to the MORF-CBP chimeric protein, which is associated with acute myeloid leukemia [58, 59]. Chi-

meric proteins resulting from translocations carry DPF domains at their N-termini, leading to the recruitment of a new modifier, an activator or a regulator of the old chromatin environment that used to be occupied only by a MYST family acetyltransferase.

It has been found that MOZ is required to maintain the progression of lymphoma induced by the *MYC* oncogene [60]. The lack of this protein causes senescence of neural stem cells [43]. Increased MOZ expression promotes glioblastoma and breast cancer development [61–63].

D4 and PHF10 family proteins are rarely mutated in cancer cells [64, 65]. However, decreased DPF2 expression correlates with a poor survival prognosis in patients with glioma [66]. It was also shown that DPF2 maintains the proliferation of transformed MLL-AF9 myeloid progenitor cells. Upon DPF2 knockdown, the cells started to differentiate, exit the cell cycle, and undergo apoptosis [67].

No significant associations between changes in DPF1 expression and malignant cell transformation were found in cancer patients.

Decreased PHF10 expression in patients with renal cancer correlates with a higher chance of patient survival [64, 66], which may be related to the positive effect of the *c-MYC* oncogene on PHF10 expression [68].

Almost no DPF3 expression is observed in human myeloid precursors. However, due to the action of STAT5, its expression significantly increases in the granulocytes of patients with chronic lymphocytic leukemia. Apparently, this may cause transcriptional dysregulation and disease progression [69]. Decreased DPF3 expression is also associated with poor survival prognosis in breast cancer patients. Thus, reduced DPF3 expression was shown to activate the JAK2/STAT3 signaling pathway and enhance the mobility of cancer cells [70].

CONCLUSIONS

The implementation of different transcriptional pathways involves transcription factors (namely, activators and repressors), as well as various auxiliary complexes that change the chromatin structure. These complexes usually consist of a large number of subunits containing numerous diverse domains that bind DNA and specific chromatin marks, which are known as modified histone tails. Due to these domains, the complexes are positioned in a strictly defined chromatin region and further additionally modify it through their activity. The DPF domains form a unique structure that binds the histone H3 tail favoring the modified H3K14ac/cr. H3 histones with acetylated and crotonylated lysine residues mainly reside in either the promoter or enhancer regions of transcriptionally active chromatin. Therefore, they act as markers for recruiting the MYST-family HAT and BAF/PBAF complexes involving proteins containing the DPF domains. There are few proteins that contain DPF. However, these proteins have homologous DPF sequences and identical amino acids at the key positions which determine the binding to H3K14ac/cr. The MYST-family HAT and BAF/PBAF complexes acetylate other histone tails and remodel (translocate) nucleosomes, respectively; i.e., they function as coactivators and contribute to additional transcriptional activation.

Thus, DPF domains perform the important function of binding chromatin, which leads to the activation of the transcription of the genes that play a crucial role in the development of an organism. ●

This study was supported by the Russian Science Foundation (grant No. 18-14-00303 ("Studying the subunit composition of the SWI/SNF complex during cell differentiation in mammals and its role in gene expression").

REFERENCES

1. Kwan A.H.Y., Gell D.A., Verger A., Crossley M., Matthews J.M., Mackay J.P. // Structure. 2003. V. 11. № 7. P. 803–813.
2. Sanchez R., Zhou M.M. // Trends Biochem. Sci. 2011. V. 36. № 7. P. 364–372.
3. Pascual J., Martinez-Yamout M., Dyson H.J., Wright P.E. // J. Mol. Biol. 2000. V. 304. № 5. P. 723–729.
4. Musselman C.A., Kutateladze T.G. // Nucl. Acids Res. 2011. V. 39. № 21. P. 9061–9071.
5. Jain K., Fraser C.S., Marunde M.R., Parker M.M., Sagum C., Burg J.M., Hall N., Popova I.K., Rodriguez K.L., Vaidya A., et al. // Epigenetics Chromatin. 2020. V. 13. № 1. P. 1–11. https://doi.org/10.1186/s13072-020-0328-z
6. Morrison E.A., Musselman C.A. // Chromatin Signaling and Diseases. Elsevier Inc. 2016. P. 127–147. http://dx.doi.org/10.1016/B978-0-12-802389-1.00007-1
7. Ullah M., Pelletier N., Xiao L., Zhao S.P., Wang K., Degerny C., Tahmasebi S., Cayrou C., Doyon Y., Goh S.-L., et al. // Mol. Cell. Biol. 2008. V. 28. № 22. P. 6828–6843.
8. Klein B.J., Lalonde M.E., Côté J., Yang X.J., Kutateladze T.G. // Epigenetics. 2014. V. 9. № 2. P. 186–193. doi: 10.4161/epi.26792.
9. Yang X.J. // Biochim. Biophys. Acta – Mol. Cell Res. 2015. V. 1853. № 8. P. 1818–1826. http://dx.doi.org/10.1016/j.bbamcr.2015.04.014
10. Huang F., Abmayr S.M., Workman J.L. // Mol. Cell. Biol. 2016. V. 36. № 14. P. 1900–1907.
11. Wu J.I. // Acta Biochim. Biophys. Sin. 2012. V. 44. № 1. P. 54–69.
12. Qiu Y., Liu L., Zhao C., Han C., Li F., Zhang J., Wang Y., Li G., Mei Y., Wu M., et al. // Genes Dev. 2012. P. 1376–1391.
13. Zeng L., Zhang Q., Li S., Plotnikov A.N., Walsh M.J., Zhou

M.M. // Nature. 2010. V. 466. № 7303. P. 258–262.
14. Ali M., Yan K., Lalonde M.E., Degerny C., Rothbart S.B., Strahl B.D., Côté J., Yang X.J., Kutateladze T.G. // J. Mol. Biol. 2012. V. 424. № 5. P. 328–338. http://dx.doi.org/10.1016/j.jmb.2012.10.004
15. Huber F.M., Greenblatt S.M., Davenport A.M., Martinez C., Xu Y., Vu L.P., Nimer S.D., Hoelz A. // Proc. Natl. Acad. Sci. USA. 2017. V. 114. № 23. P. 6016–6021.
16. Dreveny I., Deeves S.E., Fulton J., Yue B., Messmer M., Bhattacharya A., Collins H.M., Heery D.M. // Nucl. Acids Res. 2014. V. 42. № 2. P. 822–835.
17. Xiong X., Panchenko T., Yang S., Zhao S., Yan P., Zhang W., Xie W., Li Y., Zhao Y., Allis C.D., et al. // Nat. Chem. Biol. 2016. V. 12. № 12. P. 1111–1118.
18. Klein B.J., Simithy J., Wang X., Ahn J.W., Andrews F.H., Zhang Y., Côté J., Shi X., Garcia B.A., Kutateladze T.G. // Structure. 2017. V. 25. № 4. P. 650–654.e2.
19. Klein B.J., Jang S.M., Lachance C., Mi W., Lyu J., Sakuraba S., Krajewski K., Wang W.W., Sidoli S., Liu J., et al. // Nat. Commun. 2019. V. 10. № 1. P. 4724.
20. Sabari B.R., Tang Z., Huang H., Yong-Gonzalez V., Molina H., Kong H.E., Dai L., Shimada M., Cross J.R., Zhao Y., et al. // Mol. Cell. 2018. V. 69. № 3. P. 533. https://dx.doi.org/10.1016/j.molcel.2018.01.013
21. Xie Z., Zhang D., Chung D., Tang Z., Huang H., Dai L., Qi S., Li J., Colak G., Chen Y., et al. // Mol. Cell. 2016. V. 62. № 2. P. 194–206. http://dx.doi.org/10.1016/j.molcel.2016.03.036
22. Wan J., Liu H., Chu J., Zhang H. // J. Cell. Mol. Med. 2019. V. 23. № 11. P. 7163–7169.
23. Tan M., Luo H., Lee S., Jin F., Yang J.S., Montellier E., Buchou T., Cheng Z., Rousseaux S., Rajagopal N., et al. // Cell. 2011. V. 146. № 6. P. 1016–1028. http://dx.doi.org/10.1016/j.cell.2011.08.008
24. Pelletier N., Champagne N., Stifani S., Yang X.J. // Oncogene. 2002. V. 21. № 17. P. 2729–2740.
25. Rokudai S., Laptenko O., Arnal S.M., Taya Y., Kitabayashi I., Prives C. // Proc. Natl. Acad. Sci. USA. 2013. V. 110. № 10. P. 3895–3900.
26. Qin S., Jin L., Zhang J., Liu L., Ji P., Wu M., Wu J., Shi Y. // J. Biol. Chem. 2011. V. 286. № 42. P. 36944–36955.
27. Liu L., Qin S., Zhang J., Ji P., Shi Y., Wu J. // J. Struct. Biol. 2012. V. 180. № 1. P. 165–173. http://dx.doi.org/10.1016/j.jsb.2012.06.014
28. Champagne K.S., Saksouk N., Peña P.V., Johnson K., Ullah M., Yang X.J., Côté J., Kutateladze T.G. // Proteins Struct. Funct. Genet. 2008. V. 72. № 4. P. 1371–1376.
29. Lalonde M.E., Avvakumov N., Glass K.C., Joncas F.H., Saksouk N., Holliday M., Paquet E., Yan K., Tong Q., Klein B.J., et al. // Genes Dev. 2013. V. 27. № 18. P. 2009–2024.
30. Kueh A.J., Dixon M.P., Voss A.K., Thomas T. // Mol. Cell. Biol. 2011. V. 31. № 4. P. 845–860.
31. Voss A.K., Collin C., Dixon M.P., Thomas T. // Dev. Cell. 2009. V. 17. № 5. P. 674–686. http://dx.doi.org/10.1016/j.devcel.2009.10.006
32. Simó-Riudalbas L., Pérez-Salvia M., Setien F., Villanueva A., Moutinho C., Martínez-Cardús A., Moran S., Berdasco M., Gomez A., Vidal E., et al. // Cancer Res. 2015. V. 75. № 18. P. 3936–3944.
33. Fiziev P., Akdemir K.C., Miller J.P., Keung E.Z., Samant N.S., Sharma S., Natale C.A., Terranova C.J., Maitituoheti M., Amin S.B., et al. // Cell Rep. 2017. V. 19. № 4. P. 875–889. http://dx.doi.org/10.1016/j.celrep.2017.03.078
34. Mittal P., Roberts C.W.M. // Nat. Rev. Clin. Oncol. 2020. V. 17. P. 435–448. http://dx.doi.org/10.1038/s41571-020-0357-3
35. Wang X., Lee R.S., Alver B.H., Haswell J.R., Wang S., Mieczkowski J., Drier Y., Gillespie S.M., Archer T.C., Wu J.N., et al. // Nat. Genet. 2017. V. 49. № 2. P. 289–295. http://dx.doi.org/10.1038/ng.3746
36. Alver B.H., Kim K.H., Lu P., Wang X., Manchester H.E., Wang W., Haswell J.R., Park P.J., Roberts C.W.M. // Nat. Commun. 2017. V. 8. P. 14648. doi: 10.1038/ncomms14648
37. Ishizaka A., Mizutani T., Kobayashi K., Tando T., Sakurai K., Fujiwara T., Iba H. // J. Biol. Chem. 2012. V. 287. № 15. P. 11924–11933.
38. Lessard J., Wu J.I., Ranish J.A., Wan M., Winslow M.M., Staahl B.T., Wu H., Aebersold R., Graef I.A., Crabtree G.R. // Neuron. 2007. V. 55. № 2. P. 201–215.
39. Brechalov A.V., Georgieva S.G., Soshnikova N.V. // Cell Cycle. 2014. V. 13. № 12. P. 1970–1979.
40. Viryasova G.M., Tatarskiy V.V., Sheynov A.A., Tatarskiy E.V., Sud'ina G.F., Georgieva S.G., Soshnikova N.V. // Biochim. Biophys. Acta. Mol. Cell Res. 2019. V. 1866. № 12. P. 118525. doi: 10.1016/j.bbamcr.2019.118525
41. Rokudai S., Aikawa Y., Tagata Y., Tsuchida N., Taya Y., Kitabayashi I. // J. Biol. Chem. 2009. V. 284. № 1. P. 237–244.
42. Sheikh B.N., Phipson B., El-Saafin F., Vanyai H.K., Downer N.L., Bird M.J., Kueh A.J., May R.E., Smyth G.K., Voss A.K., et al. // Oncogene. 2015. V. 34. № 47. P. 5807–5820.
43. Perez-Campo F.M., Costa G., Lie-A-Ling M., Stifani S., Kouskoff V., Lacaud G. // Stem Cells. 2014. V. 32. № 6. P. 1591–1601.
44. Sheikh B.N., Downer N.L., Phipson B., Vanyai H.K., Kueh A.J., McCarthy D.J., Smyth G.K., Thomas T., Voss A.K. // Proc. Natl. Acad. Sci. USA. 2015. V. 112. № 17. P. 5437–5442.
45. Katsumoto T., Aikawa Y., Iwama A., Ueda S., Ichikawa H., Ochiya T., Kitabayashi I. // Genes Dev. 2006. V. 20. № 10. P. 1321–1330.
46. Thomas T., Corcoran L.M., Gugasyan R., Dixon M.P., Brodnicki T., Nutt S.L., Metcalf D., Voss A.K. // Genes Dev. 2006. V. 20. № 9. P. 1175–1186.
47. Sheikh B.N., Yang Y., Schreuder J., Nilsson S.K., Bilardi R., Carotta S., McRae H.M., Metcalf D., Voss A.K., Thomas T. // Blood. 2016. V. 128. № 19. P. 2307–2318.
48. Millan F., Cho M.T., Retterer K., Monaghan K.G., Bai R., Vitazka P., Everman D.B., Smith B., Angle B., Roberts V., et al. // Am. J. Med. Genet. Part A. 2016. V. 170. № 7. P. 1791–1798.
49. Thomas T., Voss A.K., Chowdhury K., Gruss P. // Development. 2000. V. 127. № 12. P. 2537–2548.
50. Merson T.D., Dixon M.P., Collin C., Rietze R.L., Bartlett P.F., Thomas T., Voss A.K. // J. Neurosci. 2006. V. 26. № 44. P. 11359–11370.
51. Lange M., Kaynak B., Forster U.B., Tönjes M., Fischer J.J., Grimm C., Schlesinger J., Just S., Dunkel I., Krueger T., et al. // Genes Dev. 2008. V. 22. № 17. P. 2370–2384.
52. Vasileiou G., Vergarajauregui S., Endele S., Popp B., Büttner C., Ekici A.B., Gerard M., Bramswig N.C., Albrecht B., Clayton-Smith J., et al. // Am. J. Hum. Genet. 2018. V. 102. № 3. P. 468–479.
53. Krasteva V., Crabtree G.R., Lessard J.A. // Exp. Hematol. 2017. V. 48. P. 58–71.e15. http://dx.doi.org/10.1016/j.exphem.2016.11.008
54. Borrow J., Stanton V.P., Andresen J.M., Becher R., Behm F.G., Chaganti R.S.K., Civin C.I., Disteche C., Dubé I., Frischauf A.M., et al. // Nat. Genet. 1996. V. 14. № 1. P. 33–41.
55. Chaffanet M., Gressin L., Preudhomme C., Soenen-Cornu V., Birnbaum D., Pébusque M.J. // Genes Chromosom. Cancer. 2000. V. 28. № 2. P. 138–144.

56. Chinen Y., Taki T., Tsutsumi Y., Kobayashi S., Matsumoto Y., Sakamoto N., Kuroda J., Horiike Sh., Nishida K., Ohno H., et. al. // Genes. Chromosomes Cancer. 2014. V. 53. P. 299–308.
57. Liang J., Prouty L., Williams B.J., Dayton M.A., Blanchard K.L. // Blood. 1998. V. 92. № 6. P. 2118–2122.
58. Panagopoulos I. // Hum. Mol. Genet. 2001. V. 10. № 4. P. 395–404.
59. Kojima K., Kaneda K., Yoshida C., Dansako H., Fujii N., Yano T., Shinagawa K., Yasukawa M., Fujita S., Tanimoto M. // Br. J. Haematol. 2003. V. 120. № 2. P. 271–273.
60. Sheikh B.N., Lee S.C.W., El-Saafin F., Vanyai H.K., Hu Y., Pang S.H.M., Grabow S., Strasser A., Nutt S.L., Alexander W.S., et al. // Blood. 2015. V. 125. № 12. P. 1910–1921.
61. Lv D., Jia F., Hou Y., Sang Y., Alvarez A.A., Zhang W., Gao W.Q., Hu B., Cheng S.Y., Ge J., et al. // Cancer Res. 2017. V. 77. № 22. P. 6190–6201.
62. Tsai W.W., Wang Z., Yiu T.T., Akdemir K.C., Xia W., Winter S., Tsai C.Y., Shi X., Schwarzer D., Plunkett W., et al. // Nature. 2010. V. 468. № 7326. P. 927–932.
63. Yu L., Liang Y., Cao X., Wang X., Gao H., Lin S.Y., Schiff R., Wang X.S., Li K. // Oncogene. 2017. V. 36. № 20. P. 2910–2918.
64. Kadoch C., Hargreaves D.C., Hodges C., Elias L., Ho L., J.R.& G.R.C. // Nat. Genet. 2013. V. 45. № 6. P. 592–602. http://dx.doi.org/10.1038/ng.2628
65. Masliah-Planchon J., Bièche I., Guinebretière J.-M., Bourdeaut F., Delattre O. // Annu. Rev. Pathol. Mech. Dis. 2015. V. 10. № 1. P. 145–171. http://www.annualreviews.org/doi/10.1146/annurev-pathol-012414-040445
66. Savas S., Skardasi G. // Crit. Rev. Oncol. Hematol. 2018. V. 123. № 11. P. 114–131. https://doi.org/10.1016/j.critrevonc.2018.01.009
67. Cruickshank A.V., Sroczynska P., Sankar A., Miyagi S., Rundsten C.F., Johansen J.V., Helin K. // PLoS One. 2015. V. 10. № 11. P. 1–13.
68. Tatarskiy E.V., Georgiev G.P., Soshnikova N.V. // Dokl. Biochem. Biophys. 2019. V. 484. № 1. P. 66–68.
69. Theodorou M., Speletas M., Mamara A., Papachristopoulou G., Lazou V., Scorilas A., Katsantoni E. // PLoS One. 2013. V. 8. № 10. P. e76155.
70. Lin W. Hao, Dai W. Gang, Xu X. Dong, Yu Q. Hua, Zhang B., Li J., Li H. Ping. // Biochem. Biophys. Res. Commun. 2019. V. 514. № 3. P. 639–644. https://doi.org/10.1016/j.bbrc.2019.04.170

Polycomb and Trithorax Group Proteins: The Long Road from Mutations in *Drosophila* to Use in Medicine

D. A. Chetverina*, D. V. Lomaev*, M. M. Erokhin*
Institute of Gene Biology, Russian Academy of Sciences, Moscow, 119334 Russia
*E-mail: dchetverina@yandex.ru, lomaevdv@gmail.com, yermaxbio@yandex.ru
Received July 16, 2020; in final form, September 30, 2020
DOI: 10.32607/actanaturae.11090
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Polycomb group (PcG) and Trithorax group (TrxG) proteins are evolutionarily conserved factors responsible for the repression and activation of the transcription of multiple genes in *Drosophila* and mammals. Disruption of the PcG/TrxG expression is associated with many pathological conditions, including cancer, which makes them suitable targets for diagnosis and therapy in medicine. In this review, we focus on the major PcG and TrxG complexes, the mechanisms of PcG/TrxG action, and their recruitment to chromatin. We discuss the alterations associated with the dysfunction of a number of factors of these groups in oncology and the current strategies used to develop drugs based on small-molecule inhibitors.**
KEYWORDS **Polycomb, Trithorax, PRE, *Drosophila*, PRC2, cancer, oncology, PRC2 inhibitors, EZH2 inhibitors, small-molecule inhibitors.**
ABBREVIATIONS **PcG – Polycomb group; TrxG – Trithorax group; PRE – Polycomb Response Element.**

INTRODUCTION

Establishment and maintenance of precise gene expression patterns that are unique to each cell type is required for the proper functioning of multicellular organisms. Transcriptional control of gene expression is one of the key steps in this type of regulation. Polycomb group (PcG) and Trithorax group (TrxG) proteins are repressors and activators of transcription, respectively [1–8]. These proteins were first characterized in *Drosophila* as regulators of *Hox* genes expression. *Hox* genes are responsible for proper body segmentation. Their baseline expression profile is determined by the protein products of the *maternal*, *gap*, *pair-rule*, and *segment polarity* genes at the early embryonic stage of development. These proteins activate each other in a cascade-like manner [9–11]. PcG/TrxG proteins were shown to be required for the subsequent maintenance of the established expression profile [12, 13].

A *Polycomb* mutation was described in *Drosophila* in 1947 [14]. Upon this mutation, the anatomical structures called sex combs, which normally form only on the first pair of male legs, also occur on the second and third pairs of legs [14]. Dysfunction of the *Polycomb* gene was shown to cause the transformation of a number of segments [15] as a result of overexpression of *Hox* genes [12, 16, 17]. In particular, sex combs result from partial transformation of the second and third pairs of legs into the first ones due to derepression of *Scr* in the *Antennapedia* complex [7]. Later, a mutation in the *trithorax* gene was discovered; its phenotypic manifestations (reduced number of sex combs) were opposite to the phenotype of *Polycomb* mutation, which is indicative of an inactivation of *Hox* genes [18, 19]. Afterwards, all mutations in other genes manifesting themselves in a manner similar to either *Polycomb* or *trithorax* were classified into the PcG and TrxG groups, respectively [4, 7]. These groups also include genes whose mutations enhance the mutant phenotypes of other known representatives as shown by genetic tests when crossing mutant flies or whose effect was demonstrated by missexpression of the *Hox* genes determined by a direct analysis.

Evolutionarily conserved PcG and TrxG proteins are found in all multicellular organisms. In mammals, mutations in the genes encoding PcG/TrxG also have a huge impact on the development of the organism [20, 21]. In addition, it was found that the area of responsibility of the PcG/TrxG proteins is much broader than only the regulation of *Hox* genes and extends to hundreds of other targets in both *Drosophila* and mammals. In particular, PcG/TrxG factors are involved in such crucial biological processes as carcinogenesis, inactivation of the X chromosome in mammals, and maintenance of the pluripotent state of stem cells [22–24].

In this review, we discuss the structure and functions of the PcG/TrxG complexes, the mechanisms of their action, and the role of individual factors in the onset, diagnosis, and therapy of oncological diseases.

PcG AND TrxG COMPLEXES

PcG complexes

Most PcG proteins associate in several types of multisubunit complexes. The main complexes in *Drosophila* and mammals are PRC1 (Polycomb repressive complex 1), PRC2 (Polycomb repressive complex 2), and PR-DUB (Polycomb repressive deubiquitinase), as well as PhoRC in *Drosophila* (*Fig. 1*).

In *Drosophila*, PRC2 complexes contain the core components E(z), Esc, Su(z)12, and Caf1 [25, 26]. The Esc subunit has a homolog, Escl, which can replace it in the complex [27]. All *Drosophila* PRC2 subunits have direct homologs in mammals. However, there is only one copy of the Esc protein – the EED, and two copies of the E(z) and Caf1 factors – the EZH2/EZH1 and RBBP7/RBBP4, respectively. Protein Su(z)12 contains only one orthologue of the protein – SUZ12 [28, 29]. All the core PRC2 subunits were confirmed as PcG proteins in *Drosophila* by genetic tests [3, 4].

PRC2 mono-, di-, and trimethylates lysine 27 of histone H3 (H3K27me1/2/3) via the catalytic SET domain of the E(z) protein (EZH2/EZH1) [25, 26, 28, 29]. The H3K27me3 modification is a specific mark of the chromatin regions repressed by the PcG system [30, 31]. The lack of the H3K27me3 modification due to a point substitution of lysine to arginine at position 27 in histone H3 leads to the derepression of *Hox* genes in *Drosophila* [32].

Mammalian EZH2 has a higher methyltransferase activity than its homolog EZH1 in the *in vitro* system [33]. In addition, EZH1 plays a less important role in development: mouse embryos mutant for *EZH2, EED,* and *SUZ12* are non-viable and die during the post-implantation period [34–36], while *EZH1* mutants are viable and fertile [37]. In this regard, *EZH2* and *EZH1* have different expression profiles: higher *EZH2* transcription is characteristic of proliferating cells, while *EZH1* is expressed in approximately the same manner at different stages of development. However, EZH1 can replace EZH2 at later stages of development or in case of defective EZH2 [33, 38, 39].

The subunits Su(z)12/SUZ12 and Esc/EED are required for the catalytic activity of E(z)/EZH2 [26, 36, 40–42]. The interaction between Esc/EED and H3K27me3 changes the conformation of the entire PRC2 complex and stimulates its methyltransferase activity [43]. In contrast, the Caf1 subunit is not required for the methyltransferase activity of E(z) [40–42].

In *Drosophila* and mammals, the core PRC2 module can interact with additional subunits. Currently, two complexes can be distinguished: PRC2.1 and PRC2.2. The PRC2.1 complex includes the Pcl (Polycomb-like) protein in *Drosophila* and the homologous proteins PHF1, PHF19, and MTF2 in mammals. Pcl was shown to stimulate the methyltransferase activity of E(z)/EZH2 [44, 45]. The PRC2.2 complex contains the subunits JARID2 and Jing/AEPB2 [46]. JARID2 specifically binds to nucleosomes monoubiquitinated at H2AK118ub (H2AK119ub in mammals). The proteins Pcl, Jing (but not JARID2) were confirmed as PcG factors in *Drosophila* by genetic tests [3, 4].

PRC1 complexes are divided into two types: cPRC1 (canonical) and ncPRC1 (non-canonical).

Drosophila cPRC1 contains the core subunits Pc (Polycomb), Ph, Sce (also known as dRing), and Psc [47–49]. A Psc homolog, the Su(z)2 protein [50, 51], is co-purified with cPRC1 in non-stoichiometric amounts; it can replace Psc in the complex [52]. The *Drosophila* ncPRC1 complex, dRAF (dRing Associated Factors), contains the proteins Sce/dRing, Psc, and Kdm2 [53]. All cPRC1 and ncPRC1 subunits were confirmed as PcG proteins in *Drosophila* [3, 4].

Similar complexes are present in mammals, with Polycomb factors being represented by multiple paralogs [54, 55]. Mammalian cPRC1 includes homologous of all *Drosophila* cPRC1 subunits: Pc (CBX 2, 4, 6, 7, 8), Ph (PHC1–3), Sce (RING1/2), Psc (paralogs PCGF2 and PCGF4 in cPRC1). Mammalian ncPRCs contains Sce (RING1/2) and Psc homologs (paralogs PCGF1, 3, 5, and 6) and also the RYBP protein, which can be replaced by the YAF2 protein. Depending on the presence of one of the Psc paralogs, mammalian cPRC1 and ncPRC1 are further divided into the subcomplexes cPRC1.1, cPRC1.2 and ncPRC1.3, ncPRC1.4, ncPRC1.5, and ncPRC1.6 (*Fig. 1*). The ncPRC1.1 subcomplex, which contains a specific subunit, KDM2B (Kdm2 protein homolog), is the closest to the *Drosophila* dRAF complex. A RYBP homolog was found in *Drosophila* [56, 57]. Co-immunoprecipitation experiments have demonstrated that *Drosophila* RYBP can be co-purified with Sce and Kdm2; however, genetic tests showed that it has a double function and acts as a Trithorax factor as well [57].

The Sce/RING protein is a catalytic subunit of cPRC1 and ncPRC2 in *Drosophila* and mammals. Sce/RING possesses E3 ubiquitin ligase activity and is responsible for the H2AK118ub modification (H2AK119ub in mammals). As mentioned above, JARID2 of the PRC2.2 complex interacts with this modification. The enzymatic activity of *Drosophila* dRING in the dRAF complex is higher than that in the cPRC1 complex [53]. The cPRC1 complex can compact chro-

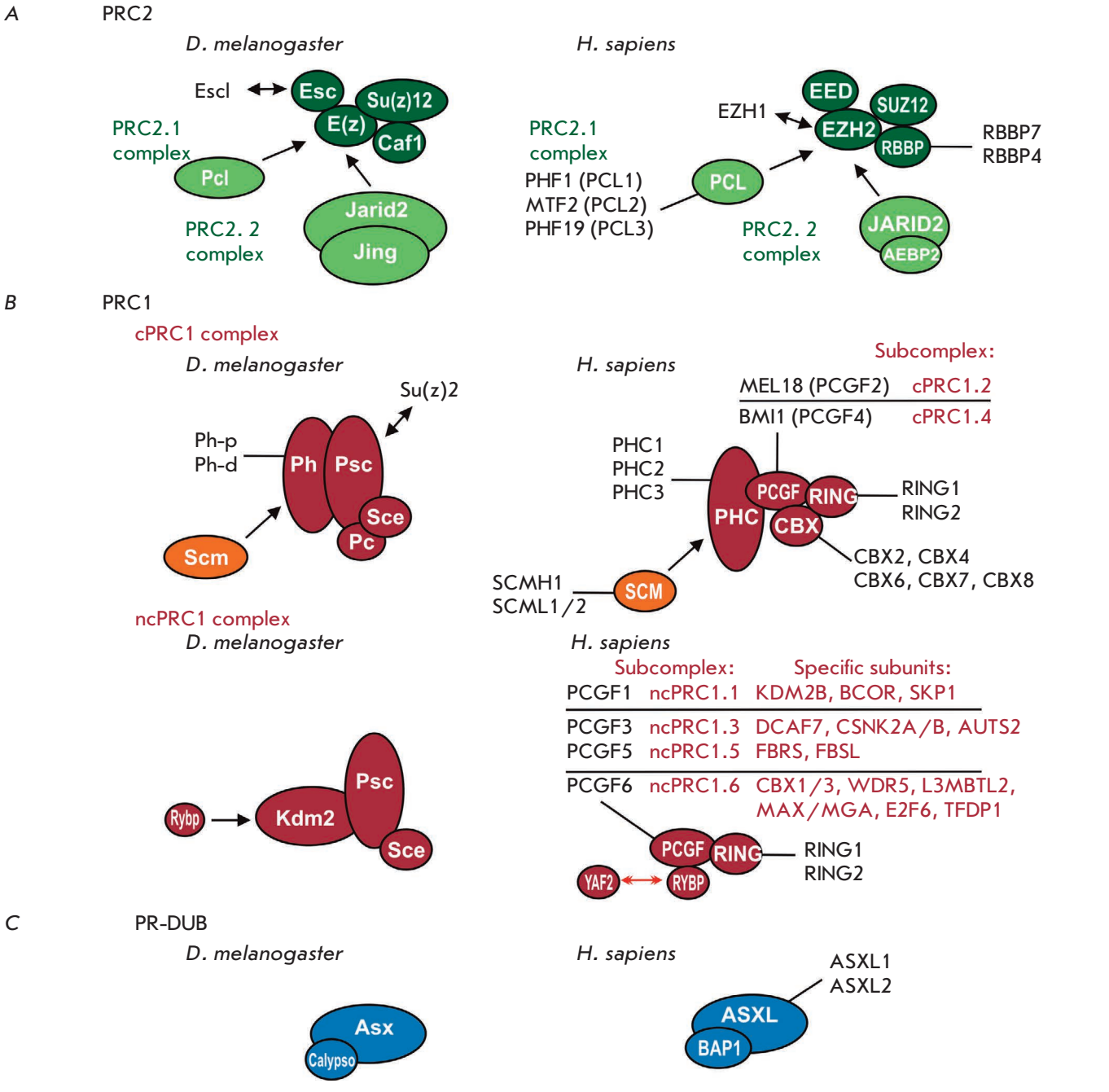


Fig. 1. Main Polycomb group complexes. (*A*) PRC2 complexes. *Drosophila* subunits are shown on the left, and their mammalian orthologs are presented on the right-hand side of the figure. In *Drosophila* and mammals, the PRC2 core is composed of E(z)–Su(z)12–Esc–Caf1 and EZH2–SUZ12–EED–RBBP7/4, respectively. *Drosophila* Esc and mammalian EZH2 can be replaced by their homologs Escl and EZH1, respectively. The PRC2.1 complex contains either Pcl or PCL1/2/3 in *Drosophila* and humans, respectively; The PRC2.2 complex contains either Jarid2/Jing or JARID2/AEBP2 in *Drosophila* and humans, respectively. (*B*) PRC1 complexes. The PRC1 core is composed of the Sce–Psc and RING–PCGF heterodimers in *Drosophila* and mammals, respectively. In cPRC1, the core subunits associate with Pc–Ph in *Drosophila* and their orthologs CBX–PHC in humans. In ncPRC1, the core subunits associate with Kdm2 in *Drosophila* and with the RYBP (or YAF2) subunit in humans. In humans, the cPRC1 and ncPRC1 complexes can be further distinguished by the presence of a specific PCGF subunit (cPRC1.2, cPRC1.4, ncPRC1.1, ncPRC1.3, ncPRC1.5, and ncPRC1.6 subcomplexes); other specific subcomplex subunits are indicated next to the complex name. (*C*) The PR–DUB complex. PR–DUB is composed of Asx–Calypso and ASXL1/2–BAP1 in *Drosophila* and humans, respectively. The size of the ovals representing the proteins corresponds to the relative size of the protein molecules

matin and repress transcription [47, 49, 58–60]. Kdm2 (KDM2B) is a histone demethylase, which removes the H3K36me2 modification characteristic of active chromatin regions [53, 61]. In addition, the Pc (CBX) protein binds nucleosomes carrying the H3K27me3 modification, which is catalyzed by PRC2 [62–64].

The Scm/SCMH1 protein, which was confirmed as a PcG factor [3], can be co-purified with *Drosophila* and mammalian cPRC1 [48, 49, 65]. In addition, Scm interacts directly with the Ph protein [66, 67]. However, at least in *Drosophila*, Scm is considered a cPRC1-independent subunit, since it can be recruited to the chromatin independently [50, 68].

The *Drosophila* PR–DUB (Polycomb repressive deubiquitinase) complex consists of the Calypso and Asx proteins [69]. Calypso is a deubiquitinase that removes the H2AK118/9ub modification, while Asx stimulates the enzymatic activity of Calypso [69]. Despite the fact that Calypso and Asx have a function opposite to the PRC1 complexes, they act as PcG factors. Mammals have two complexes with similar activity. Both complexes have a homolog of Calypso (BAP1) and include one of the Asx protein homologues – ASXL1 (which forms the PR–DUB1 complex with BAP1) or ASXL2 (which forms the PR–DUB2 complex with BAP1) [54]. The role of the simultaneous presence of ubiquitinase and deubiquitinase specific to the same histone H2A amino acid is currently unknown.

The PhoRC complex is a DNA-binding PcG complex, which includes Sfmbt and Pho [70]. Both factors are PcG proteins; their mutants are characterized by derepression of *Hox* genes [70, 71]. Pho contains a DNA-binding domain composed of C2H2-type zinc finger motifs. A Pho homolog, the Phol protein, shares the same DNA-binding site with Pho [72] and can interact with Sfmbt, instead of Pho [70]. Unlike Pho, Phol mutants do not exhibit a homeotic phenotype. Genome-wide distribution of Pho is different from that of Phol: while the main Pho peaks overlap with the PRC1 and PRC2 proteins, Phol major peaks are at the promoters of active genes [73]. Meanwhile, both factors are involved in the recruitment of PcG proteins to the chromatin (see below).

Mammals have direct homologs of the PhoRC complex subunits. However, attempts to isolate this complex have remained unsuccessful so far. The YY1 protein is a Pho/Phol homolog, while the proteins L3MBTL2, MBTD1, and SFMBT1 are Sfmbt homologs. Moreover, the YY1 protein retains the region necessary for Pho to interact with Sfmbt in *Drosophila*. *In vitro* experiments have shown that this region can interact with L3MBTL2, MBTD1, and SFMBT2. However, this interaction is 50- to 100-fold weaker than that of the Pho-Sfmbt interaction in *Drosophila* [74]. This may explain the fact that YY1 wasn't detected upon purification of the L3MBTL2 complex by co-immunoprecipitation [65, 75, 76]. Moreover, YY1 is associated with RYBP (YY1–RYBP complex) in mammals. YY1–RYBP was shown to participate in both the repression and activation of the transcription of a large number of genes [77, 78].

TrxG complexes

The TrxG is a more heterogeneous group of proteins than PcG, and genetically identified TrxG factors are subunits of different complexes involved in transcription activation [3, 4, 7, 8]. Describing the complexes, we indicate the TrxG factors that were identified by genetic tests in *Drosophila*; i.e., the factors whose mutations have phenotypes opposite to *Polycomb* mutations.

A number of Trx group factors are subunits of the ATP-dependent chromatin remodeling complexes (*Fig. 2*). Using the energy of ATP hydrolysis remodelers alter the structure, assembly, and position of nucleosomes on the DNA and, thus, facilitate the recruitment of activator complexes to the chromatin [79]. Five proteins – Osa, Brm, Mor, Snr1, and SAYP – which behave as TrxG factors in genetic tests in *Drosophila*, were shown to be subunits of the SWI/SNF subfamily of the ATP-dependent chromatin remodeler complexes: BAP (Brahma-associated proteins) and PBAP (Polybromo-associated BAP) [80–83]. The Brm (ATPase), Mor, and Snr1 proteins are subunits common to both complexes, while the Osa and SAYP proteins are specific to BAP and PBAP, respectively. All these TrxG factors have homologs in mammals which form similar complexes [79]. The Brm protein has two homologs, named SMARCA2 and SMARCA4. PBAP contains only SMARCA4, while the BAP complex can contain both homologs. Mor is homologous to the SMARCC1 and SMARCC2 proteins; Snr1 is homologous to SMARCB1. Like in *Drosophila*, homologs of the SAYP (PHF10) and OSA (ARID1A and ARID1B) factors are specific to PBAP and BAP, respectively.

The Trithorax (Trx) protein, which gave the name to the entire group, is a histone H3K4 methyltransferase. The Trx has two homologs in *Drosophila*: the Trr (Trithorax-related) and Set1 proteins. The direct mammalian orthologs of these proteins are SET1A and SET1B (Set1), MLL1 and MLL 2 (Trx), and MLL3 and MLL4 (Trr) [21, 84, 85]. All three factors and their orthologs were shown to form similar complexes: COMPASS and COMPASS-like. All complexes share common subunits: Ash2 (ASH2L), Dpy–30L1 (DPY30), Rbbp5 (RBBP5), and Wds (WDR5). In *Drosophila*, the Ash2 protein is a TrxG factor confirmed by genetic testing. These complexes catalyze H3K4me1/2/3-specific methylation of nucleosomes, which is characteristic of active

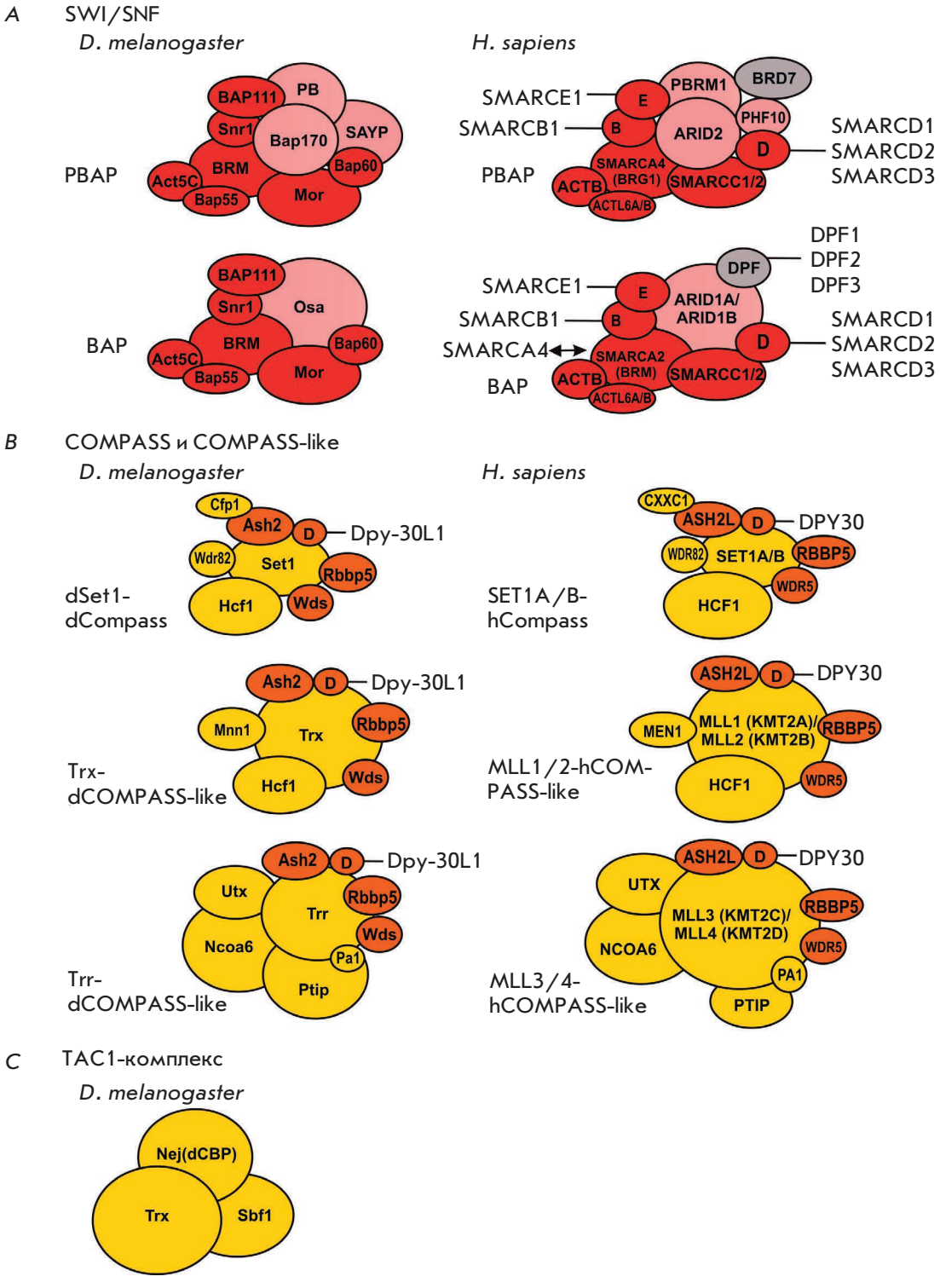


Fig. 2. Main Trithorax group complexes. (*A*) The BAP and PBAP complexes of the SWI/SNF subfamily. The subunits common to both complexes are colored in red; specific BAP and PBAP subunits are shown in pink. Human subunits whose presence in *Drosophila* BAP/PBAP complexes has not been confirmed are depicted in grey. (*B*) COMPASS and COMPASS-like complexes. The subunits common to the three complexes are colored in orange; the specific subunits are presented in yellow. (*C*) *Drosophila* TAC1 complex (not confirmed as present in humans). The size of the ovals representing proteins corresponds to the relative size of the protein molecules

chromatin regions [8, 21, 86]. According to a number of publications, the H3K4me2 modification is associated with both enhancers and gene promoters; H3K4me3 is associated with promoters of actively transcribed genes, while H3K4me1 has higher specificity to enhancers [87]. In addition, the modifications H3K4me1 and H3K4me2 in *Drosophila* overlap with the known sites of PcG complexes recruitment to chromatin (PRE, see below) [88, 89]. According to current data, Set1 is responsible for the majority of the di- and trimethylation of H3K4 in *Drosophila* cells [90, 91], while the main function of the Trx/Trr proteins is monomethyl-

ation of H3K4 [89, 92]. Similarly, the mammalian proteins MLL1/MLL2/MLL3/MLL4 mediate H3K4me1 [89, 93]. It is noteworthy that UTX, which is a subunit of Trr (MLL3/MML4)-Compass-like, is a H3K27me2/3 demethylase [94–99].

The TAC1 complex contains the TrxG factor – histone acetylase dCBP (also known as Nejire) – and the Sbf1 protein [100]. Mammals have two homologs of the Nej protein: P300 (EP300) and CBP (CREBBP). However, a TAC1-like complex has not yet been characterized. The Proteins dCBP/P300/CREBBP catalyze the H3K27Ac modification in active chromatin [101, 102]. Acetylation of histones weakens the interaction between nucleosomes and DNA and leads to chromatin decompaction [103]. The dCBP, Trx, and Trr proteins, as well as the modifications H3K4me1 and H3K27Ac catalyzed by them, respectively, were shown to colocalize at active enhancers and at the regions of PcG proteins recruitment in *Drosophila*. Moreover, acetylation of H3K27 *in vitro* is enhanced in the presence of H3K4me1, Trr, and Trx [89, 92, 101].

Another protein genetically characterized as TrxG is the Ash1 protein, which, like its mammalian homolog ASH1L, methylates histone H3 lysine 36 (H3K36me2) [104–106] and, respectively, has an activity opposite to that of Kdm2 (KDM2B). Furthermore, effective methylation of H3K36 requires the TrxG protein Kis [107], which is a homolog of the ATP-dependent chromatin remodeling proteins belonging to the CHD subfamily in mammals [108].

Rad21 (RAD21 in mammals) [109], a subunit of the cohesin complex, also belongs to the Trithorax group of proteins identified by genetic tests. The cohesin complex stabilizes long-range interactions in the nucleus, including enhancer-promoter contacts, which are necessary for the activation of transcription [110, 111].

MECHANISMS OF POLYCOMB AND TRITHORAX COMPLEXES RECRUITMENT TO CHROMATIN

PRE elements in Drosophila

Specific DNA elements that serve as PcG binding fragments were found in the *Drosophila* genome: Polycomb Response Elements (PRE) [3, 112–115]. PRE elements can be located both at a distance from the target gene (tens and hundreds of thousands of base pairs) and in the immediate vicinity of the transcription start site (TSS).

PREs were shown to act as the memory elements of the repressed state; they ensure a proper enhancer activity profile established at the early stages of embryogenesis. In transgene constructs outside of the genomic environment (in the absence of PRE elements), embryonic enhancers exhibit proper segment-specific activity only at the early stages of development (0–6 hrs), after which they become active in other parasegments, where they are normally inactive. However, a nearby inserted PRE element can maintain the correct pattern of enhancer activity at later stages of embryogenesis and suppress gene activation in unnecessary segments [116] (see details in reviews [3, 115]).

It was shown that PREs lack predetermined tissue specificity and that the enhancer determines the region of PRE activity. A number of studies have shown that PRE can either be switched from the repressing to activating state or become inactivated in the presence of an activator/enhancer. The dual nature of PREs can be also witnessed in a series of transgenic lines that carry the same transgenic construct inserted in different regions of the genome. It turns out that PRE activity is very sensitive to the insertion site, and that repression is observed only in half of the lines. In non-repressing cases, PREs presumably either exist in a neutral state or activate transcription. In addition, a number of embryonic enhancers that regulate developmental genes possess the PRE property in adults [117], indicating that at least some PREs in the activating state can potentially act as classical enhancers. In accordance with their dual activity, PREs can recruit not only PcG but TrxG proteins as well. It is important to note that PcG/TrxG proteins can be recruited to PRE DNA regardless of the PRE state [30, 118–120], which suggests direct competition between the PcG/TrxG proteins in PREs functioning. In accordance, the core subunits of the PcG complexes can be associated with active chromatin regions and potential enhancers [121–125]. It is believed that PcG proteins can inhibit the excessive activity of enhancers and promoters in these regions.

Recruitment of PcG/TrxG proteins to PRE elements

The minimum length of the PRE element required for the repression of reporter genes in transgenes is several hundred base pairs. For instance, the minimum DNA fragments sufficient for repression are 217 bp, 181 bp and 152 bp in case of *Fab7*PRE, *en*PRE and *eve*PRE, respectively [126–128].

The core PRE sequences contain sites for various DNA-binding factors. The characterized PRE DNA-binding factors in *Drosophila* are the Pho, Phol, GAF, Combgap, Spps, Zeste, Psq, Adf1, Grh, and Dsp1 proteins [112, 113]. The exact combination, number, and relative position of these protein binding sites vary in different PREs, which indicates their unequal role in the functioning of an individual PRE. It was shown that 90% of PRC1/PRC2 binding peaks overlap with Pho and Combgap [129–132]. Half (50%) of the GAF

and Dsp1 peaks overlap with PRC1 [132]. The Zeste and Phol colocalize with approximately 25% and 21% of PRC1 peaks, respectively [132].

It is important to note that most, if not all, DNA-binding factors associated with PRE participate in both repression and activation of transcription and have other targets in the genome, including promoters of active genes and potential enhancers [112, 113]. Except for Pho, mutations in the genes encoding the proteins of this group do not have a clear PcG phenotype. It is important to note that the binding sites responsible for the recruitment of a single DNA-binding protein, including Pho, as well as the combinations of the sites for different proteins, cannot ensure the recruitment of PcG proteins. This suggests the existence of a combinatorial component in the functioning of PREs, in which DNA-binding proteins form a platform for the recruitment of PcG proteins [112].

Despite the fact that the analyzed combinations of DNA-binding sites do not recruit PcG factors, these proteins/or their binding sites are significant in functional tests [112]. The role of Pho and its homolog Phol in the recruitment of PcG proteins is the best studied. Inhibition of Pho by RNA interference (RNAi) in a *Drosophila* cell line lacking Phol expression was shown to diminish the binding of Pc (PRC1), E(z), and Su(z)12 (PRC2) to one of the well-characterized PRE (*bxd*PRE) [68, 133]. At the larval stage, when both homologs are expressed, inactivation of both factors is required for the loss of PcG proteins binding [133]. The factors Pho/Phol were found to establish direct contacts with PRC1 and PRC2. Pho interacts directly with the E(z), Esc (PRC2) [133] and Ph, Pc (PRC1) proteins [134], while Phol associates with Esc [133]. The dependence of PcG protein recruitment on Pho/Phol varies between different PREs. For example, a genome-wide study has demonstrated that, in addition to Pho, the DNA-binding factors Spps and Combgap play an important role in the recruitment of PcG proteins to a number of PREs [125, 135].

Moreover, in accordance with the combinatorial basis of PRE functioning, the Spps, Dsp1, GAF, and Grh proteins can foster interactions between Pho and PcG [125, 135–138].

Grh was shown to interact directly with Sce (PRC1) [139] and Pho [136]. According to two-hybrid screening results, Spps directly interacts with Scm [140], which, in turn, can associate with the proteins Ph [66, 67] and Sfmbt [141, 142]. These interactions can stabilize the recruitment of PcG proteins to the chromatin.

In addition to the recruitment of PcG proteins, DNA-binding factors can participate in the binding of TrxG proteins to PRE. Pho was shown to interact directly with Brm ATPase [134], Zeste associates with MOR [143], while GAF is required for the recruitment of Brm and Polybromo to *bxd*PRE [144].

Thus, DNA-binding proteins in *Drosophila* can recruit both Pc and Trx group proteins. Apparently, the commonality of DNA-binding factors between the PcG/TrxG complexes increases the plasticity of transcription regulation processes by facilitating, if necessary, a rapid switch from repression to activation and vice versa.

PRE elements in mammals

A number of PRE-like elements, as well as a number of DNA-binding proteins associated with the PcG/TrxG complexes, have been described in mammals [145–149]. Among DNA-binding proteins are the AEBP2, REST, SNAIL, RUNX1, E2F6, and MGA/MAX factors [1].

However, it should be noted that in mammals, in addition to sequence-specific DNA-binding proteins, a large role in PRC2 recruitment belongs to the CpG islands (CGI) [150–152].

Apparently, as in the case of the *Drosophila* genome, there is no universal DNA-binding recruiter responsible for the binding of all of the Polycomb or Trithorax proteins to the chromatin. The existence of numerous PcG paralogs indicates the possibility of a wide variety of DNA-binding factors as well. This, taking into account the tendency of DNA-binding factors to partially functionally substitute for each other, creates obstacles for their identification. We suppose that combinations of binding sites for different DNA-binding proteins can play the primary role in the recruitment of the PcG/TrxG factors to the chromatin that form fairly extended regions for stable PcG/TrxG recruitment.

Epigenetic modifications

A number of studies indicate the impact of nucleosome modifications on the recruitment of Polycomb/Trithorax complexes. Methyltransferase E(z)/EZH2/EZH1 of the PRC2 complex creates the H3K27me3 modification, which is bound by the Pc/CBX protein of the PRC1 complex. On the other hand, the Sce/RING subunit of the PRC1 complex mediates H2AK118/9ub that is recognized by the JARID2 subunit of the PRC2 complex. The identified activities and interactions suggest the existence of positive feedback facilitating the recruitment of the PRC1 and PRC2 complexes to the chromatin. However, disruption of the PRC2 activity does not completely eliminate the binding of PRC1 subunits [153, 154]. This indicates that histone modifications can rather increase the affinity of PcG complexes to the chromatin than serve as the main recruitment factor. The role of the H2AK118/9ub modification is of great interest. Impaired ubiquitination activity of Sce/RING1 in *Drosophila* and mice does not lead to a

significant loss of Polycomb-dependent repression [155, 156]. However, it should be indicated that interrelation between PRC1 and PRC2 recruitment can depend on the object of study, since it has recently been shown that elimination of the catalytic activity of RING1 leads to a significant loss of PRC2 binding in mouse embryonic stem cells [157, 158]. Moreover, binding of the JARID2-containing PRC2.2 complex, which specifically associates with the H2AK119ub modification, was affected more strongly compared to the PRC2.1 variant, which contains PCL.

Long Non-coding RNAs

Long, non-coding RNAs (lncRNAs) are found at many mammalian genomic loci regulated by Polycomb repressors. Mutations in the Polycomb group genes were shown to suppress the activity of some lncRNAs. For instance, damage to the PRC2 core component, the EED protein, disrupts the activity of Xist lncRNA, which is required for X-chromosome inactivation in mammals [159], and of the lncRNAs involved in genomic imprinting [160]. This has led to the hypothesis that the fundamental step in the recruitment of Polycomb group repressors is the binding of PRC2 to non-coding RNAs that can attract this complex to chromatin [161]. However, it was subsequently established that PRC2 can associate randomly with various RNAs, including short RNAs, actively transcribed mRNAs, and even bacterial RNAs [162–164].

Recent studies have shown that the non-canonical PRC1 complex containing PCGF3/5 components can interact with Xist lncRNA [165–167]. This interaction is mediated by the RNA-binding factor hnRNP K, which efficiently recognizes C-rich motifs in RNA [168]. These data suggest a more specific binding of PRC1-Xist in comparison with PRC2-Xist (discussed in [169]). However, it has not been established whether this mechanism can be extended to the recruitment of Polycomb factors in the case of other regions of the mammalian genome and lncRNAs.

Attempts have also been made to elucidate the potential role of lncRNAs in *Drosophila*. However, no stable association of lncRNAs with PRC1 and PRC2 has yet been found.

Summing up, it can be assumed that in *Drosophila*, as well as in mammals, the DNA-binding factors and specific combinations of their binding sites play an important role in the targeted recruitment of Polycomb/Trithorax proteins to chromatin. Currently, especially in the mammalian genome, there is only limited information about PcG-associated DNA-binding factors and identification of these factors is one of the important tasks in the near future [170]. Epigenetic modifications of histones (and DNA modifications in mammals), as well as RNA-protein interactions, can play an important role in stabilizing interactions between Polycomb/Trithorax factors and chromatin. However, the specificity of a set of genomic targets is, apparently, determined by particular DNA sequences and the proteins that bind them.

MECHANISMS OF POLYCOMB/TRITHORAX PROTEINS ACTION

Competition between the PcG and TrxG proteins

Many known functions of the TrxG proteins counteract the activities of the PcG proteins (*Fig. 3*). Competition between PcG and TrxG proteins can occur at the PREs, enhancers, and gene promoters.

TrxG activators mediate the H3K36me2 modification [104–106], H3K27me3 demethylation [94–99], and the acetylation of H3K27 [89, 101, 102].

PcG repressors catalyze the H3K27me3 modification [25, 26, 28, 29] and demethylation of H3K36me2 [53, 61]. In addition, a number of studies indicate that PcG proteins can function in tandem with the histone deacetylase Rpd3/HDAC1 which is responsible for the deacetylation of H3K27Ac [171–173].

Histone modifications that are markers of active chromatin were shown to inhibit PcG modifications. For instance, the modifications H3K4me3 [174], H3K36me2/3 [106, 174], and H3K27Ac [101] inhibit the methylation of H3K27me3.

The competition between PcG and TrxG proteins also influences the chromatin structure. While PRC1 can compact chromatin [58], the TrxG BAP and PBAP complexes, as well as the acetylation of nucleosomes by CBP, promote chromatin decompaction [79, 103].

Spatial interactions and PcG/TrxG function

In multicellular organisms, nuclear DNA is organized into highly ordered structures that possess several levels. The first level of DNA packaging is the nucleosomes, which are assembled into chromatin fibers. At a higher level, the fibers form loop structures folded into topologically associated domains (TADs). Interactions between TADs lead to the formation of active and inactive chromosome compartments, which are partitioned into chromosome territories [175–178]. Individual genomic loci located at a great distance from each other on the same chromosome or even different chromosomes can physically interact with each other.

Some of the first evidence of the importance of spatial interactions in the activity of PRE/TRE elements was obtained on transgenic *Drosophila* lines. In the lines, repression of the marker gene by PRE increased in flies homozygous for the transgenic construct. This effect, known as Pairing Sensitive Silencing (PSS), is

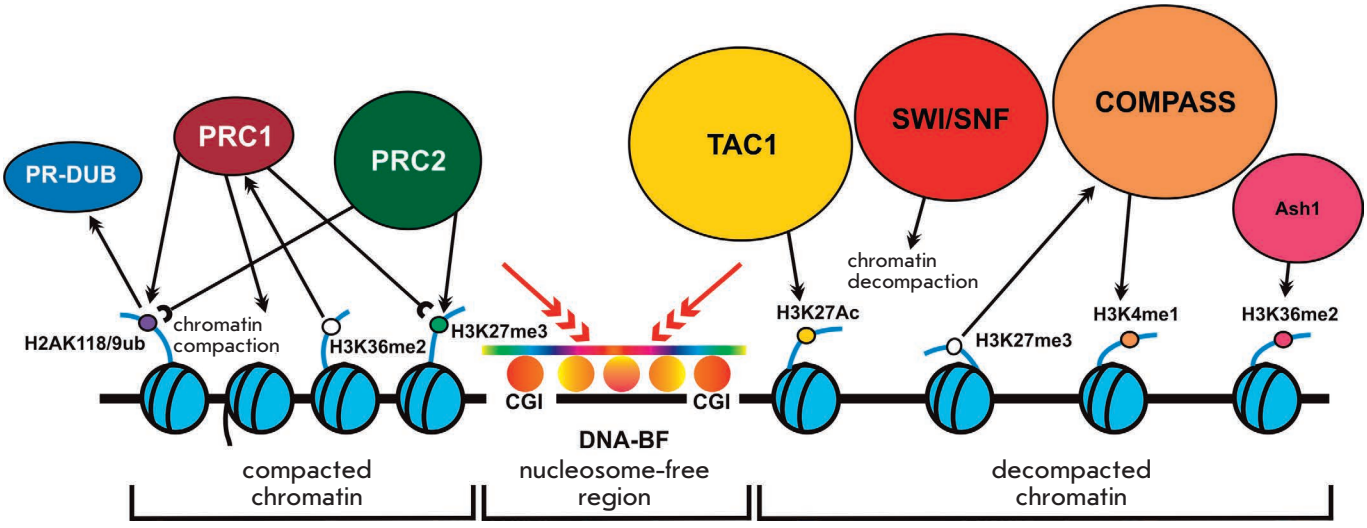


Fig. 3. Functional activities of the Polycomb/Trithorax group proteins. The PRC1 complex compacts chromatin, mediates nucleosome ubiquitination (H2AK118Ub in *Drosophila*, H2AK119Ub in mammals), and also specifically binds to the nucleosome tri-methylated at H3K27. The PRC2 complex is responsible for the H3K27me3 histone modification; it interacts with the H2AK118/9Ub nucleosomes. The PR-DUB complex deubiquitinates H2AK118/9Ub nucleosomes. Trithorax group activator complexes decompact chromatin (SWI/SNF), acetylate histones (TAC1), and catalyze H3K4 methylation (COMPASS). DNA-binding factors (DNA-BF) and hypomethylated CpG islands (CGIs) are involved in the recruitment of the PcG/TrxG complexes to chromatin

assumed to depend on the ability of the two PREs copies located on homologous chromosomes to interact and enhance each other's activity [179, 180]). In addition, PRE elements are able to repress target genes at a long distance, and this activity can be blocked by insulators [181–183]. In this aspect, the PRE/TRE is akin to the activity of enhancers, which are also able to initiate long-distance and ultralong-distance interactions with the target gene promoters, all this regulated by insulator elements [3, 184].

DYSFUNCTION OF POLYCOMB/TRITHORAX PROTEINS IN ONCOLOGICAL DISEASES

Misregulation of the activity of PcG/TrxG factors has been described in many pathological conditions, including cancer. The proteins of these groups play an important role in various cellular processes and can act as either tumor suppressors or oncogenes, depending on the tumor and tissue context. It has been shown that violation of at least one of the BAP/PBAP subunits occurs in about 25% of cancers [185]. An essential role in carcinogenesis was also revealed for H3K4-specific methyltransferases of the COMPASS and COMPASS-like complexes [186, 187]. The role of PRC1 complex PcG factors in carcinogenesis and the possibility of creating small-molecule inhibitors to block their activity are under active studies [22, 188].

In this review, we focus on the role of the PRC2 complex in oncological diseases. Studies in the last decade have demonstrated a wide variety of changes in EZH2 and its partners in cancer. This has led to the development of a number of small-molecule inhibitors to block PRC2 activity. One of them, tazemetostat, was approved in January 2020 for clinical use in medical practice in the United States [189–191].

Dysfunction of the activity of the PRC2 complex in carcinogenesis

It is now a demonstrated fact that both enhancement and suppression of PRC2 activity can lead to cancer. Basically, changes are detected in the core PRC2 subunits: the EZH2, SUZ12, and EED proteins [21, 23, 192–194]. The most studied to date case is the enhancement of PRC2 activity, i.e., situations in which PRC2-encoding genes act as oncogenes. In many types of malignant tumors, overexpression of PRC2 components is observed. Activation of the PRC2 function can also be a result of gain-of-function (GOF) mutations, which increase the catalytic activity of EZH2/EZH1. On the other hand, tumors associated with *EZH2*-, *SUZ12*-, or *EED*-impaired activity have also been described, which suggests a tumor-suppressive role for PRC2 in these cases.

Oncogenic role of PRC2

Overexpression of EZH2, SUZ12, and EED. *EZH2* is the gene whose transcription level changes most commonly during carcinogenesis, compared to other PRC2 subunits. In normal cells the transcription level of *EZH2*

is regulated by the RB–E2F signaling pathway, and the high level of *EZH2* expression in proliferative cells is significantly lower in differentiated cells [33, 195, 196]. However, *EZH2* overexpression is observed in many malignant neoplasms [131, 195, 197–236] (*Fig. 4*). *EZH2* overexpression was found to be associated with an increased level of H3K27me3 and, often, associated with an amplification of the EZH2-encoding gene [195, 221, 226]. In some cases, a correlation between a high expression of *EZH2* and poor survival prognosis was noted [131, 197, 199, 202, 204, 206–209, 212–214, 218, 219, 222–225, 227–231, 236].

Overexpression of *SUZ12* and *EED* was detected in some types of cancer as well [202, 209, 211, 212, 214, 215, 217, 233, 237, 238]. However, there is currently significantly less clinical data regarding these genes. In a number of studies, an increased level of *SUZ12* and *EED* transcription is associated with poor survival prognosis [202, 209, 212, 215, 233, 237].

EZH2 GOF mutations. In addition to *EZH2*, *SUZ12*, and *EED* overexpression, the activity of the PRC2 complex can be enhanced by GOF mutations in the methyltransferase domain of *EZH2*. Such mutations have been described in specific types of non-Hodgkin lymphomas (diffuse large B-cell lymphoma (DLBCL) and follicular lymphomas) [129, 221, 239–245] (*Fig. 4*). Point Y641→F,N,H,S substitutions relative to isoform C (denoted as Y646 relative to isoform A) are the most common mutations [129, 221, 239, 242–245]. There are also functionally similar mutations at the A677 and A687 positions [129, 221, 241, 243, 246, 247].

Mutant forms of EZH2 were shown to more efficiently methylate histone H3 (H3K27me2), which leads to an increased degree of H3K27me3 modification. In lymphoid tumors with monoallelic GOF mutations, wild-type EZH2 prefers H3K27me0/me1 nucleosomes as a substrate for methylation, while the mutant form shows enhanced catalytic activity against H3K27me2 [248, 249]. However, GOF mutations in *EZH2*, despite their widespread occurrence in lymphomas, are not associated with a poor survival prognosis in follicular lymphomas [239] and DLBCL [244].

GOF mutations were also detected in the EZH2 homolog EZH1 (Q571R) in thyroid adenoma [250]. This mutation also results in an increased level of H3K27me3.

Tumor-suppressive role of PRC2

LOF mutations in the PRC2 complex subunits. LOF (loss-of-function) mutations that disrupt PRC2 activity have been described in all three core components: EZH2, SUZ12, and EED [251–268] (*Fig. 4*). LOF mutations in *EZH2* and *EED* have been shown to be associated with a negative prognosis in myelodysplastic syndrome/myeloproliferative neoplasm [251, 253–255, 261, 265, 268].

Thus, the consequences of an inactivation of the PRC2 function observed in a number of tumors remain insufficiently studied, while the data on PRC2 hyperactivity are more substantive. Further studies will help elucidate the significance and frequency of LOF PRC2 mutations in different types of tumors.

H3K27M mutation of histones. Another type of mutations affecting the activity of PRC2 are point substitutions in the *H3F3A* and *HIST1H3B* genes (which encode for the histone variants H3.3 and H3.1, respectively). These mutations lead to the substitution of lysine for methionine at position 27 of H3 (H3K27) and are designated as H3K27M. It was found that such mutant histone variants interact with EZH2 and inhibit the methyltransferase activity of the PRC2 complex, decreasing the H3K27me3 level both *in vivo* and *in vitro* [269–272]. H3K27M mutations are found in 80% of pediatric gliomas [273–275] and in 6% of secondary acute myeloid leukemias [276]. It was recently demonstrated that EZH2 can be automethylated at positions EZH2-K510 and EZH2-K514. This automethylation stimulates the histone activity of EZH2 and is impaired in lines carrying the H3K27M mutation [277].

Suppression of PRC2 activity by the factor EZHIP (EZH2 Inhibitory Protein) has recently been discovered in ependymoma cells (CNS tumor) [278–281]. The EZHIP region is considered to mimic the H3K27M structure and inhibit PRC2 activity in a similar way.

Mechanisms of the oncogenic and tumor-suppressive PRC2 roles

The mechanisms that underlie the opposite PRC2 roles in different types of tumors are currently being studied vigorously. In general, these differences are driven by the PRC2-mediated suppression of either oncogenes or tumor suppressors in different type of cells.

Overexpression of *EZH2* has been shown to enhance cell proliferation both *in vitro* [195, 208] and *in vivo* [282–284]. An increased EZH2 level stimulates metastasis [285], cell invasion [208], and affects DNA repair [283]. GOF mutations in *EZH2* accelerate MYC- and BCL-2-mediated lymphomagenesis in mice [282, 286]. The available data indicate that the oncogenic effect of PRC2 consists in inhibiting the transcription of a number of tumor suppressors, while the specific set of tumor suppressors to be inhibited is strongly dependent on the type of tumor. For instance, PRC2 suppresses *CDKN2A* transcription in prostate and endometrial cancer cells, as well as in lymphoid tumors (see details

Type of cancer	Oncogenic function of PRC2				Tumor-suppressor function of PRC2			REFERENCES
	Overexpression			GOF	LOF			
	EZH2	SUZ12	EED	EZH2	EZH2	SUZ12	EED	
Bladder cancer	+PP	+PP						[195, 209, 228, 237]
Follicular lymphoma	+			+				[129, 221, 239, 242, 243, 245]
Diffuse large B-cell lymphoma (DLBCL)				+				[242–245]
Mantle cell lymphoma	+PP	+						[206, 213, 217]
Natural killer/T-cell lympoma (NKTL)	+							[234]
Myeloma	+PP							[223]
T-cell acute lymphoblastic leukemia					+	+	+	[262, 266]
Acute myeloid leukemia					+	+		[263]
MDS/MPN					+PP	+	+PP	[251, 253–255, 261, 264, 265, 268]
Glioblastoma	+					+	+	[195, 232, 252]
Malignant peripheral nerve sheath tumors						+	+	[252, 257, 267]
Breast cancer	+PP	+	+					[131, 195, 196, 201, 203, 208, 224, 225, 235]
Colorectal cancer	+PP	+PP	+PP					[195, 215, 220, 228, 231]
Gastric cancer	+PP	+PP						[195, 204, 222, 223]
Retinoblastoma	+							[210]
Renal cell carcinoma	+PP							[195, 230]
Laryngeal cancer	+PP							[195, 236]
Hepatocellular carcinoma	+							[227]
Cholangiocarcinoma	+PP							[219]
Lung cancer	+PP	+	+					[195, 199, 207, 214, 228, 238]
Rhabdomyosarcoma	+							[200]
Ovarian cancer	+PP	+PP	+					[211, 212, 216]
Prostate cancer	+PP	+PP						[131, 202, 226, 229, 232]
Melanoma	+PP					+	+	[131, 195, 252]
Testis cancer	+							[195]
Thyroid cancer	+PP							[195, 198, 218]
Cervical cancer, endometrial cancer	+PP					+		[131, 195, 197, 205, 256, 258, 259, 260]

Fig. 4. Disruption of the PRC2 core subunits' activity in carcinogenesis. The "+" sign against the pink background stands for cases of hyperactivation of the PRC2 enzyme (overexpression or GOF mutations); "+"against the blue background stands for cancer associated with a loss of the PRC2 function (LOF mutations); "PP" (Poor Prognosis) indicates that the PRC2 subunit dysfunction was shown to correlate with a poor survival prognosis. MDS/MPN – myelodysplastic/myeloproliferative neoplasm

in [192]). It should be noted that PRC2 inactivation suppresses the growth of some tumor cells *in vitro* and *in vivo*, which has allowed for the development of small-molecule inhibitors (see below for details).

The mechanisms of tumor-suppressive PRC2 action have been less studied. However, such PRC2 targets as, for instance, the oncogenes *HOXA9* and *MYC* are overexpressed in many types of tumors [252, 255, 262, 287]. In transgenic mice, somatic deletions of *EZH2* and *EED* interact with the Q61K mutation of the *NRAS* oncogene and hyperactivate the STAT3 signaling pathway, leading to acute myeloid leukemias [288]. The combination of mutations in the *EZH2/RUNX1* or *EZH2/p53* gene leads to the formation of therapy-resistant

myeloid lymphocytic leukemias in mouse models [289, 290]. A *SUZ12* deletion interacts with the JAK3 factor mutation, leading to acute lymphoblastic T-cell leukemia [291]. Inactivating mutations in *EZH2* contribute to the development of a myelodysplastic syndrome that is induced by mutations in *RUNX1* [292]. The loss of SUZ12 activity is associated with a *NF1* mutation in tumors of the peripheral nervous system: glioma and melanoma [252]. *NF1* encodes a GTPase, which activates the *ras* gene; and mutations in this factor result in Ras-dependent activation of carcinogenesis [293]. A loss of PRC2 activity is also observed in the case of other gene disfunctions: for example, mutations in *ASXL1* [294], or upon HMGN1 overexpression [295] in leukemia.

Thus, depending on the mutations or changes in the expression of other genes, PRC2 inactivation can lead to malignant cell transformation [296, 297]. Carcinogenesis can be associated with a loss of function by all three core components of the PRC2: EZH2, SUZ12, and EED.

Small-molecule PRC2 inhibitors

The discovery of numerous abnormalities associated with PRC2 hyperactivity stimulated scientists to develop small-molecule inhibitors that suppress the activity of this complex (*Fig. 5*).

The first of such substance was DZNep. This inhibitor reduced the level of H3K27me3 modification in tumor cell cultures [298]. However, it was later found that treatment of cells with DZNep decreases the overall level of nucleosome methylation at different positions [299]. At the next step, three inhibitors, named EPZ005687 [300], GSK126 [301] and EI1 [302], were developed. They specifically inhibited the methyltransferase activity of both the native and GOF mutant (at position Y641) forms of the EZH2 protein. The mechanism of these inhibitors' action is based on the competition with the cofactor S-adenosylmethionine (SAM) for selective binding to the SET domain of EZH2. The treatment of a cell culture with the EI1 inhibitor has a comparable effect on the level of H3K27me3 as the complete deletion of the *EZH2* gene [302], while the GSK126 is able to suppress the *in vivo* growth of a tumor obtained by xenotransplantation of human lymphoma KARPAS422 cells in mice [301].

The EZH2 inhibitor EPZ-6438 (which was later registered as Tazemetostat) also targets the methyltransferase domain of EZH2. The activity of EPZ-6438 has been demonstrated on malignant rhabdoid tumors (MRTs). These cells carry the mutant *SMARCB1* gene, which encodes the subunit of the SWI/SNF chromatin remodeling complex [303]. Mutations in this gene are often found in rhabdoid tumors [304] and confer a high sensitivity of the tumor cells to the

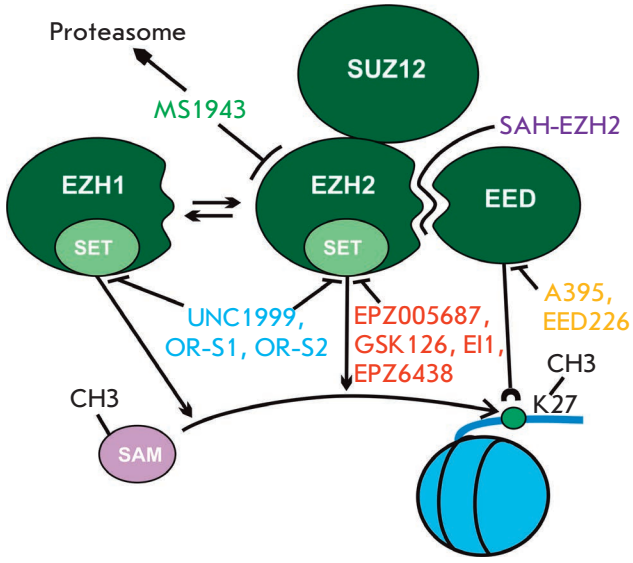


Fig. 5. Schematic representation of the mechanisms of suppression of PRC2 hyperactivity by small-molecule inhibitors. **EPZ005687**, **GSK126**, **EI1**, and **EPZ6438 (tazemetostat)** target the EZH2 SET domain and inhibit the transfer of a methyl group from S-adenosylmethionine (SAM) to histone H3 lysine 27 (H3K27). **UNC1999**, **OR-S1**, and **OR-S2** suppress the activity of EZH2 and its close homolog, EZH1. **SAH-EZH2** inhibits the interaction between EZH2 and EED, which leads to destabilization of the PRC2 complex. **A-395** and **EED226** suppress the recruitment of the EED protein to the H3K27me3 modification, eliminating the stimulation of the PRC2 methyltransferase activity. **MS1943** recognizes the unique three-dimensional structure of the EZH2 protein and directs it to the proteasome degradation pathway

suppression of the PRC2 activity [305]. Treatment of xenograft-bearing mice with EPZ-6438 was shown to decrease the total level of H3K27me3 and reactivate a number of repressed genes. Further experiments also confirmed the ability of EPZ-6438 to suppress the proliferation of tumor cells derived from lymphoid tumors [306].

All of the described inhibitors were highly specific to EZH2 and much less active against EZH1. However, it had previously been shown that EZH1 can replace EZH2 if the latter is damaged. Thus, in some cases, when using highly specific inhibitors that block the activity of EZH2, the PRC2 complex can retain partial activity thanks to EZH1. Therefore, a series of inhibitors were further developed to solve this problem.

The inhibitors UNC1999 [307], OR-S1, and OR-S2 [308, 309] target the methyltransferase activity of both EZH2 and EZH1. Suppression of cell proliferation by OR-S2 was analyzed on a big set of tumor cell lines [309]. OR-S2 was shown to inhibit the growth of

33 out of 68 tumor cell lines of hematopoietic origin (lymphoma, myeloma, and leukemia) and 26 out of 124 solid tumors. The cytotoxic effects of OR-S1 and OR-S2 have also been confirmed on models of gastric cancer, rhabdoid tumors, and acute myeloid leukemia [308, 309].

High-throughput screening methods can be used to search for the novel small-molecule inhibitors of PRC2. For example, screening of approximately 250,000 substances allowed scientists to identify 162 that are able to inhibit EZH2 [310].

Determination of the spatial structure of the PRC2 complex allowed scientists to develop a new approach to the creation of small-molecule inhibitors [43, 311, 312].

First, the SAH-EZH2 peptide, which mimics the EED region required for interaction with EZH2, was synthesized [313]. Treatment of cells with SAH-EZH2 impairs PRC2 complex formation, reduces the level of H3K27me3, and inhibits the proliferation of malignant blood and retinoblastoma cells [313, 314].

Second, a region in the EED protein that specifically interacts with the H3K27me3 modification and is important for the recruitment of PRC2 to chromatin was used as a target. Two small-molecule substances have been developed: A-395 and EED226. Their activity against lymphoid tumor cells is comparable to that of EZH2 methyltransferase domain inhibitors [315, 316]. It should be noted that A-395 displays a cytotoxic activity against tumor cells that have acquired resistance to GSK126, which is an inhibitor of the EZH2 methyltransferase domain [315]. Thus, a combination of inhibitors targeting different regions of the PRC2 complex can be used to avoid the emergence of resistance to chemotherapeutic drugs.

Third, the hydrophobic tagging (HyT) method is used to suppress PRC2 activity. In this case, the chimeric molecule is created, one part of which binds to the target protein, and the other one directs the bound complex to proteasome degradation [317]. This method allowed researchers to develop the MS1943 inhibitor, which is specific to EZH2 [318]. MS1943 was shown to suppress the growth of the triple-negative breast cancer MDA-MB-468 cell line, which is resistant to EZH2 methyltransferase domain inhibitors. Thus, inhibition of the methyltransferase activity and degradation of the PRC2 complex can have different therapeutic effects, something that can be implemented in medical practice by using combinations of different drugs.

CONCLUSION

It has been more than 70 years since the discovery of the *Polycomb* mutation. Tremendous progress has been made in the study of how the PcG/TrxG system functions. The global role of these factors in the transcription regulation and maintenance of cellular homeostasis is becoming clearer. There is also a growing body of data concerning PcG/TrxG dysfunctions in various pathologies. However, a number of questions needing to be addressed for a more complete understanding of the system's functioning remain open. The details of PcG/TrxG complexe's recruitment to specific genomic regions remain unclear. The exact contribution of various factors to these processes, such as the activity of specific DNA-binding factors, epigenetic marks, non-coding RNAs, as well as other unknown biological processes, has not yet been established. The increasing complexity of the PcG/TrxG system in the process of evolution from invertebrates to mammals and the emergence of numerous paralogs of these proteins represent a challenge for researchers: to what extent the currently understood composition of protein complexes is characteristic of all types of cells? Are they unique only to a number of tissues and/or developmental stages and do they differ in others? Numerous recent studies assign a crucial role to the spatial organization of genes in the nucleus. These processes were also shown to be closely related to the functioning of PcG/TrxG complexes. It is important to determine the extent to which the spatial organization determines the functions of DNA regulatory elements or whether it is a consequence of transcriptional complexes' recruitment to chromatin. Much needs to be done to further elucidate the significance of PcG/TrxG factors in medicine, in the development of improved small-molecule inhibitors, and in the creation of optimal therapeutic protocols. At the same time, despite the tremendous progress achieved in the study of PcG/TrxG proteins in mammals, the *Drosophila* remains an indispensable model organism for studying the details of transcription control by PcG/TrxG proteins.

The emergence of biological methods (genome editing, high-throughput sequencing, mass spectrometry approaches etc.) provides hope that many of these questions will be answered in the future. However, what remains quite clear is that the long road in the study of PcG/TrxG factors is far from being completed. ●

This work was supported by the Russian Foundation for Basic Research (grant No. 18-34-20046).

REFERENCES

1. Bauer M., Trupke J., Ringrose L. // Chromosoma. 2015. V. 125. P. 471–496.
2. Brand M., Nakka K., Zhu J., Dilworth F.J. // Cell Stem Cell. 2019. V. 24. № 4. P. 518–533.
3. Chetverina D.A., Elizar'ev P.V., Lomaev D.V., Georgiev P.G., Erokhin M.M. // Genetika. 2017. V. 53. № 2. P. 133–154.
4. Kassis J.A., Kennison J.A., Tamkun J.W. // Genetics. 2017. V. 206. № 4. P. 1699–1725.
5. Kuroda M.I., Kang H., De S., Kassis J.A. // Annu. Rev. Biochem. 2020. V. 89 P. 235–253.
6. Schuettengruber B., Bourbon H.M., Di Croce L., Cavalli G. // Cell. 2017. V. 171. № 1. P. 34–57.
7. Kingston R.E., Tamkun J.W. // Cold Spring Harb. Perspect. Biol. 2014. V. 6. № 10. P. a019349.
8. Schuettengruber B., Martinez A.M., Iovino N., Cavalli G. // Nat. Rev. Mol. Cell. Biol. 2011. V. 12. № 12. P. 799–814.
9. Pankratz M.J., Jackle H. // Trends Genet. 1990. V. 6. № 9. P. 287–292.
10. Small S., Levine M. // Curr. Opin. Genet. Dev. 1991. V. 1. № 2. P. 255–260.
11. Akam M. // Development. 1987. V. 101. № 1. P. 1–22.
12. McKeon J., Brock H.W. // Roux Arch. Dev. Biol. 1991. V. 199. № 7. P. 387–396.
13. Struhl G., Akam M. // EMBO J. 1985. V. 4. № 12. P. 3259–3264.
14. Lewis P.H. // Dros. Inf. Serv. 1947. V. 21. P. 69.
15. Lewis E.B. // Nature. 1978. V. 276. № 5688. P. 565–570.
16. Wedeen C., Harding K., Levine M. // Cell. 1986. V. 44. № 5. P. 739–748.
17. Beuchle D., Struhl G., Muller J. // Development. 2001. V. 128. № 6. P. 993–1004.
18. Ingham P.W. // Wilehm. Roux. Arch. Dev. Biol. 1981. V. 190. № 6. P. 365–369.
19. Ingham P.W., Whittle J.R.S. // Mol. Gen. Genet. 1980. V. 179. P. 607–614.
20. Laugesen A., Helin K. // Cell Stem Cell. 2014. V. 14. № 6. P. 735–751.
21. Piunti A., Shilatifard A. // Science. 2016. V. 352. № 6290. P. aad9780.
22. Chan H.L., Morey L. // Trends Biochem. Sci. 2019. V. 44. № 8. P. 688–700.
23. Comet I., Riising E.M., Leblanc B., Helin K. // Nat. Rev. Cancer. 2016. V. 16. № 12. P. 803–810.
24. Brockdorff N. // Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2017. V. 372. № 1733. P. 1–5.
25. Czermin B., Melfi R., McCabe D., Seitz V., Imhof A., Pirrotta V. // Cell. 2002. V. 111. № 2. P. 185–196.
26. Muller J., Hart C.M., Francis N.J., Vargas M.L., Sengupta A., Wild B., Miller E.L., O'Connor M.B., Kingston R.E., Simon J.A. // Cell. 2002. V. 111. № 2. P. 197–208.
27. Kurzhals R.L., Tie F., Stratton C.A., Harte P.J. // Dev. Biol. 2008. V. 313. № 1. P. 293–306.
28. Cao R., Wang L., Wang H., Xia L., Erdjument-Bromage H., Tempst P., Jones R.S., Zhang Y. // Science. 2002. V. 298. № 5595. P. 1039–1043.
29. Kuzmichev A., Nishioka K., Erdjument-Bromage H., Tempst P., Reinberg D. // Genes Dev. 2002. V. 16. № 22. P. 2893–2905.
30. Papp B., Muller J. // Genes Dev. 2006. V. 20. № 15. P. 2041–2054.
31. Schwartz Y.B., Kahn T.G., Nix D.A., Li X.Y., Bourgon R., Biggin M., Pirrotta V. // Nat. Genet. 2006. V. 38. № 6. P. 700–705.
32. Pengelly A.R., Copur O., Jackle H., Herzig A., Muller J. // Science. 2013. V. 339. № 6120. P. 698–699.
33. Margueron R., Li G., Sarma K., Blais A., Zavadil J., Woodcock C.L., Dynlacht B.D., Reinberg D. // Mol. Cell. 2008. V. 32. № 4. P. 503–518.
34. Faust C., Schumacher A., Holdener B., Magnuson T. // Development. 1995. V. 121. № 2. P. 273–285.
35. O'Carroll D., Erhardt S., Pagani M., Barton S.C., Surani M.A., Jenuwein T. // Mol. Cell. Biol. 2001. V. 21. № 13. P. 4330–4336.
36. Pasini D., Bracken A.P., Jensen M.R., Lazzerini Denchi E., Helin K. // EMBO J. 2004. V. 23. № 20. P. 4061–4071.
37. Ezhkova E., Lien W.H., Stokes N., Pasolli H.A., Silva J.M., Fuchs E. // Genes Dev. 2011. V. 25. № 5. P. 485–498.
38. Shen X., Liu Y., Hsu Y.J., Fujiwara Y., Kim J., Mao X., Yuan G.C., Orkin S.H. // Mol. Cell. 2008. V. 32. № 4. P. 491–502.
39. Son J., Shen S.S., Margueron R., Reinberg D. // Genes Dev. 2013. V. 27. № 24. P. 2663–2677.
40. Ketel C.S., Andersen E.F., Vargas M.L., Suh J., Strome S., Simon J.A. // Mol. Cell Biol. 2005. V. 25. № 16. P. 6857–6868.
41. Nekrasov M., Wild B., Muller J. // EMBO Rep. 2005. V. 6. № 4. P. 348–353.
42. Cao R., Zhang Y. // Mol. Cell. 2004. V. 15. № 1. P. 57–67.
43. Jiao L., Liu X. // Science. 2015. V. 350. № 6258. P. aac4383.
44. Nekrasov M., Klymenko T., Fraterman S., Papp B., Oktaba K., Kocher T., Cohen A., Stunnenberg H.G., Wilm M., Muller J. // EMBO J. 2007. V. 26. № 18. P. 4078–4088.
45. Sarma K., Margueron R., Ivanov A., Pirrotta V., Reinberg D. // Mol. Cell. Biol. 2008. V. 28. № 8. P. 2718–2731.
46. Herz H.M., Mohan M., Garrett A.S., Miller C., Casto D., Zhang Y., Seidel C., Haug J.S., Florens L., Washburn M.P., et al. // Mol. Cell. Biol. 2012. V. 32. № 9. P. 1683–1693.
47. Francis N.J., Saurin A.J., Shao Z., Kingston R.E. // Mol. Cell. 2001. V. 8. № 3. P. 545–556.
48. Saurin A.J., Shao Z., Erdjument-Bromage H., Tempst P., Kingston R.E. // Nature. 2001. V. 412. № 6847. P. 655–660.
49. Shao Z., Raible F., Mollaaghababa R., Guyon J.R., Wu C.T., Bender W., Kingston R.E. // Cell. 1999. V. 98. № 1. P. 37–46.
50. Kang H., McElroy K.A., Jung Y.L., Alekseyenko A.A., Zee B.M., Park P.J., Kuroda M.I. // Genes Dev. 2015. V. 29. № 11. P. 1136–1150.
51. Strubbe G., Popp C., Schmidt A., Pauli A., Ringrose L., Beisel C., Paro R. // Proc. Natl. Acad. Sci. USA. 2011. V. 108. № 14. P. 5572–5577.
52. Lo S.M., Ahuja N.K., Francis N.J. // Mol. Cell. Biol. 2009. V. 29. № 2. P. 515–525.
53. Lagarou A., Mohd-Sarip A., Moshkin Y.M., Chalkley G.E., Bezstarosti K., Demmers J.A., Verrijzer C.P. // Genes Dev. 2008. V. 22. № 20. P. 2799–2810.
54. Hauri S., Comoglio F., Seimiya M., Gerstung M., Glatter T., Hansen K., Aebersold R., Paro R., Gstaiger M., Beisel C. // Cell Rep. 2016. V. 17. № 2. P. 583–595.
55. Levine S.S., Weiss A., Erdjument-Bromage H., Shao Z., Tempst P., Kingston R.E. // Mol. Cell. Biol. 2002. V. 22. № 17. P. 6070–6078.
56. Bejarano F., Gonzalez I., Vidal M., Busturia A. // Mech. Dev. 2005. V. 122. № 10. P. 1118–1129.
57. Fereres S., Simon R., Mohd-Sarip A., Verrijzer C.P., Busturia A. // PLoS One. 2014. V. 9. № 11. P. e113255.
58. Francis N.J., Kingston R.E., Woodcock C.L. // Science. 2004. V. 306. № 5701. P. 1574–1577.
59. King I.F., Emmons R.B., Francis N.J., Wild B., Muller J., Kingston R.E., Wu C.T. // Mol. Cell. Biol. 2005. V. 25. № 15.

P. 6578–6591.
60. King I.F., Francis N.J., Kingston R.E. // Mol. Cell. Biol. 2002. V. 22. № 22. P. 7919–7928.
61. Tsukada Y., Fang J., Erdjument-Bromage H., Warren M.E., Borchers C.H., Tempst P., Zhang Y. // Nature. 2006. V. 439. № 7078. P. 811–816.
62. Fischle W., Wang Y., Jacobs S.A., Kim Y., Allis C.D., Khorasanizadeh S. // Genes Dev. 2003. V. 17. № 15. P. 1870–1881.
63. Kaustov L., Ouyang H., Amaya M., Lemak A., Nady N., Duan S., Wasney G.A., Li Z., Vedadi M., Schapira M., et al. // J. Biol. Chem. 2011. V. 286. № 1. P. 521–529.
64. Min J., Zhang Y., Xu R.M. // Genes Dev. 2003. V. 17. № 15. P. 1823–1828.
65. Gao Z., Zhang J., Bonasio R., Strino F., Sawai A., Parisi F., Kluger Y., Reinberg D. // Mol. Cell. 2012. V. 45. № 3. P. 344–356.
66. Kim C.A., Sawaya M.R., Cascio D., Kim W., Bowie J.U. // J. Biol. Chem. 2005. V. 280. № 30. P. 27769–27775.
67. Peterson A.J., Kyba M., Bornemann D., Morgan K., Brock H.W., Simon J. // Mol. Cell. Biol. 1997. V. 17. № 11. P. 6683–6692.
68. Wang L., Jahren N., Miller E.L., Ketel C.S., Mallin D.R., Simon J.A. // Mol. Cell. Biol. 2010. V. 30. № 11. P. 2584–2593.
69. Scheuermann J.C., de Ayala Alonso A.G., Oktaba K., Ly-Hartig N., McGinty R.K., Fraterman S., Wilm M., Muir T.W., Muller J. // Nature. 2010. V. 465. № 7295. P. 243–247.
70. Klymenko T., Papp B., Fischle W., Kocher T., Schelder M., Fritsch C., Wild B., Wilm M., Muller J. // Genes Dev. 2006. V. 20. № 9. P. 1110–1122.
71. Fritsch C., Brown J.L., Kassis J.A., Muller J. // Development. 1999. V. 126. № 17. P. 3905–3913.
72. Brown J.L., Fritsch C., Mueller J., Kassis J.A. // Development. 2003. V. 130. № 2. P. 285–294.
73. Kahn T.G., Stenberg P., Pirrotta V., Schwartz Y.B. // PLoS Genet. 2014. V. 10. № 7. P. e1004495.
74. Alfieri C., Gambetta M.C., Matos R., Glatt S., Sehr P., Fraterman S., Wilm M., Muller J., Muller C.W. // Genes Dev. 2013. V. 27. № 21. P. 2367–2379.
75. Trojer P., Cao A.R., Gao Z., Li Y., Zhang J., Xu X., Li G., Losson R., Erdjument-Bromage H., Tempst P., et al. // Mol. Cell. 2011. V. 42. № 4. P. 438–450.
76. Zhang J., Bonasio R., Strino F., Kluger Y., Holloway J.K., Modzelewski A.J., Cohen P.E., Reinberg D. // Genes Dev. 2013. V. 27. № 7. P. 749–766.
77. Bajusz I., Henry S., Sutus E., Kovacs G., Pirity M.K. // Genes (Basel). 2019. V. 10. № 11. P. 1–33.
78. Zhan S., Wang T., Ge W., Li J. // J. Cell. Mol. Med. 2018. V. 22. № 4. P. 2046–2054.
79. Bracken A.P., Brien G.L., Verrijzer C.P. // Genes Dev. 2019. V. 33. № 15–16. P. 936–959.
80. Chalkley G.E., Moshkin Y.M., Langenberg K., Bezstarosti K., Blastyak A., Gyurkovics H., Demmers J.A., Verrijzer C.P. // Mol. Cell. Biol. 2008. V. 28. № 9. P. 2920–2929.
81. Mashtalir N., D'Avino A.R., Michel B.C., Luo J., Pan J., Otto J.E., Zullow H.J., McKenzie Z.M., Kubiak R.L., St Pierre R., et al. // Cell. 2018. V. 175. № 5. P. 1272–1288.
82. Mohrmann L., Langenberg K., Krijgsveld J., Kal A.J., Heck A.J., Verrijzer C.P. // Mol. Cell. Biol. 2004. V. 24. № 8. P. 3077–3088.
83. Vorobyeva N.E., Soshnikova N.V., Nikolenko J.V., Kuzmina J.L., Nabirochkina E.N., Georgieva S.G., Shidlovskii Y.V. // Proc. Natl. Acad. Sci. USA. 2009. V. 106. № 27. P. 11049–11054.
84. Mohan M., Herz H.M., Smith E.R., Zhang Y., Jackson J., Washburn M.P., Florens L., Eissenberg J.C., Shilatifard A. // Mol. Cell. Biol. 2011. V. 31. № 21. P. 4310–4318.
85. van Nuland R., Smits A.H., Pallaki P., Jansen P.W., Vermeulen M., Timmers H.T. // Mol. Cell. Biol. 2013. V. 33. № 10. P. 2067–2077.
86. Shilatifard A. // Annu. Rev. Biochem. 2012. V. 81. P. 65–95.
87. Erokhin M., Vassetzky Y., Georgiev P., Chetverina D. // Cell. Mol. Life Sci. 2015. V. 72. № 12. P. 2361–2375.
88. Rickels R., Hu D., Collings C.K., Woodfin A.R., Piunti A., Mohan M., Herz H.M., Kvon E., Shilatifard A. // Mol. Cell. 2016. V. 63. № 2. P. 318–328.
89. Tie F., Banerjee R., Saiakhova A.R., Howard B., Monteith K.E., Scacheri P.C., Cosgrove M.S., Harte P.J. // Development. 2014. V. 141. № 5. P. 1129–1139.
90. Ardehali M.B., Mei A., Zobeck K.L., Caron M., Lis J.T., Kusch T. // EMBO J. 2011. V. 30. № 14. P. 2817–2828.
91. Hallson G., Hollebakken R.E., Li T., Syrzycka M., Kim I., Cotsworth S., Fitzpatrick K.A., Sinclair D.A., Honda B.M. // Genetics. 2012. V. 190. № 1. P. 91–100.
92. Herz H.M., Mohan M., Garruss A.S., Liang K., Takahashi Y.H., Mickey K., Voets O., Verrijzer C.P., Shilatifard A. // Genes Dev. 2012. V. 26. № 23. P. 2604–2620.
93. Hu D., Gao X., Morgan M.A., Herz H.M., Smith E.R., Shilatifard A. // Mol. Cell. Biol. 2013. V. 33. № 23. P. 4745–4754.
94. Agger K., Cloos P.A., Christensen J., Pasini D., Rose S., Rappsilber J., Issaeva I., Canaani E., Salcini A.E., Helin K. // Nature. 2007. V. 449. № 7163. P. 731–734.
95. De Santa F., Totaro M.G., Prosperini E., Notarbartolo S., Testa G., Natoli G. // Cell. 2007. V. 130. № 6. P. 1083–1094.
96. Hong S., Cho Y.W., Yu L.R., Yu H., Veenstra T.D., Ge K. // Proc. Natl. Acad. Sci. USA. 2007. V. 104. № 47. P. 18439–18444.
97. Lan F., Bayliss P.E., Rinn J.L., Whetstine J.R., Wang J.K., Chen S., Iwase S., Alpatov R., Issaeva I., Canaani E., et al. // Nature. 2007. V. 449. № 7163. P. 689–694.
98. Lee M.G., Villa R., Trojer P., Norman J., Yan K.P., Reinberg D., Di Croce L., Shiekhattar R. // Science. 2007. V. 318. № 5849. P. 447–450.
99. Smith E.R., Lee M.G., Winter B., Droz N.M., Eissenberg J.C., Shiekhattar R., Shilatifard A. // Mol. Cell. Biol. 2008. V. 28. № 3. P. 1041–1046.
100. Petruk S., Sedkov Y., Smith S., Tillib S., Kraevski V., Nakamura T., Canaani E., Croce C.M., Mazo A. // Science. 2001. V. 294. № 5545. P. 1331–1334.
101. Tie F., Banerjee R., Stratton C.A., Prasad-Sinha J., Stepanik V., Zlobin A., Diaz M.O., Scacheri P.C., Harte P.J. // Development. 2009. V. 136. № 18. P. 3131–3141.
102. Jin Q., Yu L.R., Wang L., Zhang Z., Kasper L.H., Lee J.E., Wang C., Brindle P.K., Dent S.Y., Ge K. // EMBO J. 2011. V. 30. № 2. P. 249–262.
103. Barnes C.E., English D.M., Cowley S.M. // Essays Biochem. 2019. V. 63. № 1. P. 97–107.
104. An S., Yeo K.J., Jeon Y.H., Song J.J. // J. Biol. Chem. 2011. V. 286. № 10. P. 8369–8374.
105. Tanaka Y., Katagiri Z., Kawahashi K., Kioussis D., Kitajima S. // Gene. 2007. V. 397. № 1–2. P. 161–168.
106. Yuan W., Xu M., Huang C., Liu N., Chen S., Zhu B. // J. Biol. Chem. 2011. V. 286. № 10. P. 7983–7989.
107. Dorighi K.M., Tamkun J.W. // Development. 2013. V. 140. № 20. P. 4182–4192.
108. Daubresse G., Deuring R., Moore L., Papoulas O., Zakrajsek I., Waldrip W.R., Scott M.P., Kennison J.A., Tamkun J.W. // Development. 1999. V. 126. № 6. P. 1175–1187.
109. Hallson G., Syrzycka M., Beck S.A., Kennison J.A., Dorsett D., Page S.L., Hunter S.M., Keall R., Warren W.D., Brock H.W., et al. // Proc. Natl. Acad. Sci. USA. 2008. V. 105.

№ 34. P. 12405–12410.
110. Dorsett D. // Trends Genet. 2019. V. 35. № 7. P. 542–551.
111. Putlyaev E.V., Ibragimov A.N., Lebedeva L.A., Georgiev P.G., Shidlovskii Y.V. // Biochemistry (Moscow). 2018. V. 83. № 4. P. 423–436.
112. Erokhin M., Georgiev P., Chetverina D. // Epigenomes. 2018. V. 2. № 1. P. 1–24.
113. Kassis J.A., Brown J.L. // Adv. Genet. 2013. V. 81. P. 83–118.
114. McElroy K.A., Kang H., Kuroda M.I. // Open Biol. 2014. V. 4. P. 140006.
115. Steffen P.A., Ringrose L. // Nat. Rev. Mol. Cell. Biol. 2014. V. 15. № 5. P. 340–356.
116. Muller J., Bienz M. // EMBO J. 1991. V. 10. № 11. P. 3147–3155.
117. Erceg J., Pakozdi T., Marco-Ferreres R., Ghavi-Helm Y., Girardot C., Bracken A.P., Furlong E.E. // Genes Dev. 2017. V. 31. № 6. P. 590–602.
118. Erokhin M., Elizar'ev P., Parshikov A., Schedl P., Georgiev P., Chetverina D. // Proc. Natl. Acad. Sci. USA. 2015. V. 112. № 48. P. 14930–14935.
119. Langlais K.K., Brown J.L., Kassis J.A. // PLoS One. 2012. V. 7. № 11. P. e48765.
120. Elizar'ev P.V., Lomaev D.V., Chetverina D.A., Georgiev P.G., Erokhin M.M. // Acta Naturae. 2016. V. 8. № 2. P. 79–86.
121. Arora M., Packard C.Z., Banerjee T., Parvin J.D. // Nucleic Acids Res. 2016. V. 44. № 5. P. 2136–2144.
122. Chan H.L., Beckedorff F., Zhang Y., Garcia-Huidobro J., Jiang H., Colaprico A., Bilbao D., Figueroa M.E., LaCava J., Shiekhattar R., et al. // Nat. Commun. 2018. V. 9. № 1. P. 3377.
123. Frangini A., Sjoberg M., Roman-Trufero M., Dharmalingam G., Haberle V., Bartke T., Lenhard B., Malumbres M., Vidal M., Dillon N. // Mol. Cell. 2013. V. 51. № 5. P. 647–661.
124. Schaaf C.A., Misulovin Z., Gause M., Koenig A., Gohara D.W., Watson A., Dorsett D. // PLoS Genet. 2013. V. 9. № 6. P. e1003560.
125. Brown J.L., Sun M.A., Kassis J.A. // Proc. Natl. Acad. Sci. USA. 2018. V. 115. № 8. P. E1839–E1848.
126. Dejardin J., Cavalli G. // EMBO J. 2004. V. 23. № 4. P. 857–868.
127. DeVido S.K., Kwon D., Brown J.L., Kassis J.A. // Development. 2008. V. 135. № 4. P. 669–676.
128. Fujioka M., Yusibova G.L., Zhou J., Jaynes J.B. // Development. 2008. V. 135. № 24. P. 4131–4139.
129. Bodor C., Grossmann V., Popov N., Okosun J., O'Riain C., Tan K., Marzec J., Araf S., Wang J., Lee A.M., et al. // Blood. 2013. V. 122. № 18. P. 3165–3168.
130. Kwong C., Adryan B., Bell I., Meadows L., Russell S., Manak J.R., White R. // PLoS Genet. 2008. V. 4. № 9. P. e1000178.
131. Bachmann I.M., Halvorsen O.J., Collett K., Stefansson I.M., Straume O., Haukaas S.A., Salvesen H.B., Otte A.P., Akslen L.A. // J. Clin. Oncol. 2006. V. 24. № 2. P. 268–273.
132. Schuettengruber B., Ganapathi M., Leblanc B., Portoso M., Jaschek R., Tolhuis B., van Lohuizen M., Tanay A., Cavalli G. // PLoS Biol. 2009. V. 7. № 1. P. e13.
133. Wang L., Brown J.L., Cao R., Zhang Y., Kassis J.A., Jones R.S. // Mol. Cell. 2004. V. 14. № 5. P. 637–646.
134. Mohd-Sarip A., Venturini F., Chalkley G.E., Verrijzer C.P. // Mol. Cell. Biol. 2002. V. 22. № 21. P. 7473–7483.
135. Ray P., De S., Mitra A., Bezstarosti K., Demmers J.A., Pfeifer K., Kassis J.A. // Proc. Natl. Acad. Sci. USA. 2016. V. 113. № 14. P. 3826–3831.
136. Blastyak A., Mishra R.K., Karch F., Gyurkovics H. // Mol. Cell. Biol. 2006. V. 26. № 4. P. 1434–1444.
137. Dejardin J., Rappailles A., Cuvier O., Grimaud C., Decoville M., Locker D., Cavalli G. // Nature. 2005. V. 434. № 7032. P. 533–538.
138. Mahmoudi T., Zuijderduijn L.M., Mohd-Sarip A., Verrijzer C.P. // Nucleic Acids Res. 2003. V. 31. № 14. P. 4147–4156.
139. Tuckfield A., Clouston D.R., Wilanowski T.M., Zhao L.L., Cunningham J.M., Jane S.M. // Mol. Cell. Biol. 2002. V. 22. № 6. P. 1936–1946.
140. Shokri L., Inukai S., Hafner A., Weinand K., Hens K., Vedenko A., Gisselbrecht S.S., Dainese R., Bischof J., Furger E., et al. // Cell Rep. 2019. V. 27. № 3. P. 955–970.
141. Frey F., Sheahan T., Finkl K., Stoehr G., Mann M., Benda C., Muller J. // Genes Dev. 2016. V. 30. № 9. P. 1116–1127.
142. Grimm C., Matos R., Ly-Hartig N., Steuerwald U., Lindner D., Rybin V., Muller J., Muller C.W. // EMBO J. 2009. V. 28. № 13. P. 1965–1977.
143. Kal A.J., Mahmoudi T., Zak N.B., Verrijzer C.P. // Genes Dev. 2000. V. 14. № 9. P. 1058–1071.
144. Nakayama T., Shimojima T., Hirose S. // Development. 2012. V. 139. № 24. P. 4582–4590.
145. Arnold P., Scholer A., Pachkov M., Balwierz P.J., Jorgensen H., Stadler M.B., van Nimwegen E., Schubeler D. // Genome Res. 2013. V. 23. № 1. P. 60–73.
146. Cameron S.R., Nandi S., Kahn T.G., Barrasa J.I., Stenberg P., Schwartz Y.B. // J. Biol. Chem. 2018. V. 293. № 37. P. 14342–14358.
147. Corley M., Kroll K.L. // Cell Tissue Res. 2015. V. 359. № 1. P. 65–85.
148. Dietrich N., Lerdrup M., Landt E., Agrawal-Singh S., Bak M., Tommerup N., Rappsilber J., Sodersten E., Hansen K. // PLoS Genet. 2012. V. 8. № 3. P. e1002494.
149. Woo C.J., Kharchenko P.V., Daheron L., Park P.J., Kingston R.E. // Cell. 2010. V. 140. № 1. P. 99–110.
150. Jermann P., Hoerner L., Burger L., Schubeler D. // Proc. Natl. Acad. Sci. USA. 2014. V. 111. № 33. P. 3415–3421.
151. Lynch M.D., Smith A.J., De Gobbi M., Flenley M., Hughes J.R., Vernimmen D., Ayyub H., Sharpe J.A., Sloane-Stanley J.A., Sutherland L., et al. // EMBO J. 2012. V. 31. № 2. P. 317–329.
152. Mendenhall E.M., Koche R.P., Truong T., Zhou V.W., Issac B., Chi A.S., Ku M., Bernstein B.E. // PLoS Genet. 2010. V. 6. № 12. P. e1001244.
153. Kahn T.G., Dorafshan E., Schultheis D., Zare A., Stenberg P., Reim I., Pirrotta V., Schwartz Y.B. // Nucleic Acids Res. 2016. V. 44. № 21. P. 10132–10149.
154. Tavares L., Dimitrova E., Oxley D., Webster J., Poot R., Demmers J., Bezstarosti K., Taylor S., Ura H., Koide H., et al. // Cell. 2012. V. 148. № 4. P. 664–678.
155. Illingworth R.S., Moffat M., Mann A.R., Read D., Hunter C.J., Pradeepa M.M., Adams I.R., Bickmore W.A. // Genes Dev. 2015. V. 29. № 18. P. 1897–1902.
156. Pengelly A.R., Kalb R., Finkl K., Muller J. // Genes Dev. 2015. V. 29. № 14. P. 1487–1492.
157. Blackledge N.P., Fursova N.A., Kelley J.R., Huseyin M.K., Feldmann A., Klose R.J. // Mol. Cell. 2020. V. 77. № 4. P. 857–874 e859.
158. Tamburri S., Lavarone E., Fernandez-Perez D., Conway E., Zanotti M., Manganaro D., Pasini D. // Mol. Cell. 2020. V. 77. № 4. P. 840–856.
159. Silva J., Mak W., Zvetkova I., Appanah R., Nesterova T.B., Webster Z., Peters A.H., Jenuwein T., Otte A.P., Brock-

dorff N. // Dev. Cell. 2003. V. 4. № 4. P. 481–495.
160. Mager J., Montgomery N.D., de Villena F.P., Magnuson T. // Nat. Genet. 2003. V. 33. № 4. P. 502–507.
161. Rinn J.L., Chang H.Y. // Annu. Rev. Biochem. 2012. V. 81. P. 145–166.
162. Hendrickson D.G., Kelley D.R., Tenen D., Bernstein B., Rinn J.L. // Genome Biol. 2016. V. 17. P. 28.
163. Davidovich C., Zheng L., Goodrich K.J., Cech T.R. // Nat. Struct. Mol. Biol. 2013. V. 20. № 11. P. 1250–1257.
164. Wang X., Goodrich K.J., Gooding A.R., Naeem H., Archer S., Paucek R.D., Youmans D.T., Cech T.R., Davidovich C. // Mol. Cell. 2017. V. 65. № 6. P. 1056–1067.
165. Almeida M., Pintacuda G., Masui O., Koseki Y., Gdula M., Cerase A., Brown D., Mould A., Innocent C., Nakayama M., et al. // Science. 2017. V. 356. № 6342. P. 1081–1084.
166. Colognori D., Sunwoo H., Kriz A.J., Wang C.Y., Lee J.T. // Mol. Cell. 2019. V. 74. № 1. P. 101–117.
167. Pintacuda G., Wei G., Roustan C., Kirmizitas B.A., Solcan N., Cerase A., Castello A., Mohammed S., Moindrot B., Nesterova T.B., et al. // Mol. Cell. 2017. V. 68. № 5. P. 955–969.
168. Paziewska A., Wyrwicz L.S., Bujnicki J.M., Bomsztyk K., Ostrowski J. // FEBS Lett. 2004. V. 577. № 1–2. P. 134–140.
169. Almeida M., Bowness J.S., Brockdorff N. // Curr. Opin. Genet. Dev. 2020. V. 61. P. 53–61.
170. Fedotova A.A., Bonchuk A.N., Mogila V.A., Georgiev P.G. // Acta Naturae. 2017. V. 9. № 2. P. 47–58.
171. Chang Y.L., Peng Y.H., Pan I.C., Sun D.S., King B., Huang D.H. // Proc. Natl. Acad. Sci. USA. 2001. V. 98. № 17. P. 9730–9735.
172. Tie F., Prasad-Sinha J., Birve A., Rasmuson-Lestander A., Harte P.J. // Mol. Cell. Biol. 2003. V. 23. № 9. P. 3352–3362.
173. van der Vlag J., Otte A.P. // Nat. Genet. 1999. V. 23. № 4. P. 474–478.
174. Schmitges F.W., Prusty A.B., Faty M., Stutzer A., Lingaraju G.M., Aiwazian J., Sack R., Hess D., Li L., Zhou S., et al. // Mol. Cell. 2011. V. 42. № 3. P. 330–341.
175. Bonev B., Cavalli G. // Nat. Rev. Genet. 2016. V. 17. № 11. P. 661–678.
176. Cheutin T., Cavalli G. // Crit. Rev. Biochem. Mol. Biol. 2019. V. 54. № 5. P. 399–417.
177. Razin S.V., Ulianov S.V. // Cell. Mol. Biol. Lett. 2017. V. 22. P. 18.
178. Szabo Q., Bantignies F., Cavalli G. // Sci. Adv. 2019. V. 5. № 4. P. eaaw1668.
179. Kassis J.A. // Genetics. 1994. V. 136. № 3. P. 1025–1038.
180. Kassis J.A. // Adv. Genet. 2002. V. 46. P. 421–438.
181. Mallin D.R., Myung J.S., Patton J.S., Geyer P.K. // Genetics. 1998. V. 148. № 1. P. 331–339.
182. Sigrist C.J., Pirrotta V. // Genetics. 1997. V. 147. № 1. P. 209–221.
183. Ogiyama Y., Schuettengruber B., Papadopoulos G.L., Chang J.M., Cavalli G. // Mol. Cell. 2018. V. 71. № 1. P. 73–88.
184. Chetverina D., Aoki T., Erokhin M., Georgiev P., Schedl P. // Bioessays. 2014. V. 36. № 2. P. 163–172.
185. Mittal P., Roberts C.W.M. // Nat. Rev. Clin. Oncol. 2020. V. 17. № 4. P. 435–448.
186. Fagan R.J., Dingwall A.K. // Cancer Lett. 2019. V. 458. P. 56–65.
187. Zhao Z., Shilatifard A. // Genome Biol. 2019. V. 20. № 1. P. 245.
188. Jangal M., Lebeau B., Witcher M. // Expert. Opin. Ther. Targets. 2019. V. 23. № 7. P. 565–578.
189. Hoy S.M. // Drugs. 2020. V. 80. № 5. P. 513–521.
190. Italiano A. // J. Hematol. Oncol. 2020. V. 13. № 1. P. 33.
191. Rothbart S.B., Baylin S.B. // Cell. 2020. V. 181. № 2. P. 211.
192. Kim K.H., Roberts C.W. // Nat. Med. 2016. V. 22. № 2. P. 128–134.
193. Lue J.K., Amengual J.E. // Curr. Hematol. Malig. Rep. 2018. V. 13. № 5. P. 369–382.
194. Yamagishi M., Uchimaru K. // Curr. Opin. Oncol. 2017. V. 29. № 5. P. 375–381.
195. Bracken A.P., Pasini D., Capra M., Prosperini E., Colli E., Helin K. // EMBO J. 2003. V. 22. № 20. P. 5323–5335.
196. Iliopoulos D., Lindahl-Allen M., Polytarchou C., Hirsch H.A., Tsichlis P.N., Struhl K. // Mol. Cell. 2010. V. 39. № 5. P. 761–772.
197. Azizmohammadi S., Azizmohammadi S., Safari A., Kaghazian M., Sadrkhanlo M., Behnod V., Seifoleslami M. // Oncol. Res. 2017. V. 25. № 4. P. 495–501.
198. Borbone E., Troncone G., Ferraro A., Jasencakova Z., Stojic L., Esposito F., Hornig N., Fusco A., Orlando V. // J. Clin. Endocrinol. Metab. 2011. V. 96. № 4. P. 1029–1038.
199. Cao W., Ribeiro Rde O., Liu D., Saintigny P., Xia R., Xue Y., Lin R., Mao L., Ren H. // PLoS One. 2012. V. 7. № 12. P. e52984.
200. Ciarapica R., Russo G., Verginelli F., Raimondi L., Donfrancesco A., Rota R., Giordano A. // Cell Cycle. 2009. V. 8. № 1. P. 172–175.
201. Collett K., Eide G.E., Arnes J., Stefansson I.M., Eide J., Braaten A., Aas T., Otte A.P., Akslen L.A. // Clin. Cancer Res. 2006. V. 12. № 4. P. 1168–1174.
202. Crea F., Hurt E.M., Mathews L.A., Cabarcas S.M., Sun L., Marquez V.E., Danesi R., Farrar W.L. // Mol. Cancer. 2011. V. 10. P. 40.
203. Gonzalez M.E., Moore H.M., Li X., Toy K.A., Huang W., Sabel M.S., Kidwell K.M., Kleer C.G. // Proc. Natl. Acad. Sci. USA. 2014. V. 111. № 8. P. 3098–3103.
204. He L.J., Cai M.Y., Xu G.L., Li J.J., Weng Z.J., Xu D.Z., Luo G.Y., Zhu S.L., Xie D. // Asian. Pac. J. Cancer Prev. 2012. V. 13. № 7. P. 3173–3178.
205. Jia N., Li Q., Tao X., Wang J., Hua K., Feng W. // Oncol. Lett. 2014. V. 8. № 5. P. 2049–2054.
206. Kienle D., Katzenberger T., Ott G., Saupe D., Benner A., Kohlhammer H., Barth T.F., Holler S., Kalla J., Rosenwald A., et al. // J. Clin. Oncol. 2007. V. 25. № 19. P. 2770–2777.
207. Kikuchi J., Kinoshita I., Shimizu Y., Kikuchi E., Konishi J., Oizumi S., Kaga K., Matsuno Y., Nishimura M., Dosaka-Akita H. // Cancer. 2010. V. 116. № 12. P. 3015–3024.
208. Kleer C.G., Cao Q., Varambally S., Shen R., Ota I., Tomlins S.A., Ghosh D., Sewalt R.G., Otte A.P., Hayes D.F., et al. // Proc. Natl. Acad. Sci. USA. 2003. V. 100. № 20. P. 11606–11611.
209. Lee S.R., Roh Y.G., Kim S.K., Lee J.S., Seol S.Y., Lee H.H., Kim W.T., Kim W.J., Heo J., Cha H.J., et al. // Clin. Cancer Res. 2015. V. 21. № 23. P. 5391–5403.
210. Lei Q., Shen F., Wu J., Zhang W., Wang J., Zhang L. // Oncol. Rep. 2014. V. 32. № 1. P. 261–269.
211. Li H., Cai Q., Godwin A.K., Zhang R. // Mol. Cancer Res. 2010. V. 8. № 12. P. 1610–1618.
212. Li H., Cai Q., Wu H., Vathipadiekal V., Dobbin Z.C., Li T., Hua X., Landen C.N., Birrer M.J., Sanchez-Beato M., et al. // Mol. Cancer Res. 2012. V. 10. № 11. P. 1462–1472.
213. Lin Y.L., Zou Z.K., Su H.Y., Huang Y.Q. // Zhongguo. Shi. Yan. Xue. Ye. Xue. Za. Zhi. 2019. V. 27. № 3. P. 820–826.
214. Liu H., Li W., Yu X., Gao F., Duan Z., Ma X., Tan S., Yuan Y., Liu L., Wang J., et al. // Oncotarget. 2016. V. 7. № 35. P. 56338–56354.
215. Liu Y.L., Gao X., Jiang Y., Zhang G., Sun Z.C., Cui B.B.,

Yang Y.M. // J. Cancer Res. Clin. Oncol. 2015. V. 141. № 4. P. 661–669.
216. Lu C., Han H.D., Mangala L.S., Ali-Fehmi R., Newton C.S., Ozbun L., Armaiz-Pena G.N., Hu W., Stone R.L., Munkarah A., et al. // Cancer Cell. 2010. V. 18. № 2. P. 185–197.
217. Martin-Perez D., Sanchez E., Maestre L., Suela J., Vargiu P., Di Lisio L., Martinez N., Alves J., Piris M.A., Sanchez-Beato M. // Am. J. Pathol. 2010. V. 177. № 2. P. 930–942.
218. Masudo K., Suganuma N., Nakayama H., Oshima T., Rino Y., Iwasaki H., Matsuzu K., Sugino K., Ito K., Kondo T., et al. // In Vivo. 2018. V. 32. № 1. P. 25–31.
219. Nakagawa S., Okabe H., Sakamoto Y., Hayashi H., Hashimoto D., Yokoyama N., Sakamoto K., Kuroki H., Mima K., Nitta H., et al. // Ann. Surg. Oncol. 2013. V. 20 Suppl 3. P. S667–675.
220. Ohuchi M., Sakamoto Y., Tokunaga R., Kiyozumi Y., Nakamura K., Izumi D., Kosumi K., Harada K., Kurashige J., Iwatsuki M., et al. // Oncol. Lett. 2018. V. 16. № 4. P. 5275–5281.
221. Okosun J., Bodor C., Wang J., Araf S., Yang C.Y., Pan C., Boller S., Cittaro D., Bozek M., Iqbal S., et al. // Nat. Genet. 2014. V. 46. № 2. P. 176–181.
222. Pan Y.M., Wang C.G., Zhu M., Xing R., Cui J.T., Li W.M., Yu D.D., Wang S.B., Zhu W., Ye Y.J., et al. // Mol. Cancer. 2016. V. 15. № 1. P. 79.
223. Pawlyn C., Bright M.D., Buros A.F., Stein C.K., Walters Z., Aronson L.I., Mirabella F., Jones J.R., Kaiser M.F., Walker B.A., et al. // Blood Cancer J. 2017. V. 7. № 3. P. e549.
224. Pietersen A.M., Horlings H.M., Hauptmann M., Langerod A., Ajouaou A., Cornelissen-Steijger P., Wessels L.F., Jonkers J., van de Vijver M.J., van Lohuizen M. // Breast Cancer Res. 2008. V. 10. № 6. P. R109.
225. Puppe J., Drost R., Liu X., Joosse S.A., Evers B., Cornelissen-Steijger P., Nederlof P., Yu Q., Jonkers J., van Lohuizen M., et al. // Breast Cancer Res. 2009. V. 11. № 4. P. R63.
226. Saramaki O.R., Tammela T.L., Martikainen P.M., Vessella R.L., Visakorpi T. // Genes Chromosomes Cancer. 2006. V. 45. № 7. P. 639–645.
227. Sudo T., Utsunomiya T., Mimori K., Nagahara H., Ogawa K., Inoue H., Wakiyama S., Fujita H., Shirouzu K., Mori M. // Br. J. Cancer. 2005. V. 92. № 9. P. 1754–1758.
228. Takawa M., Masuda K., Kunizaki M., Daigo Y., Takagi K., Iwai Y., Cho H.S., Toyokawa G., Yamane Y., Maejima K., et al. // Cancer Sci. 2011. V. 102. № 7. P. 1298–1305.
229. Varambally S., Dhanasekaran S.M., Zhou M., Barrette T.R., Kumar-Sinha C., Sanda M.G., Ghosh D., Pienta K.J., Sewalt R.G., Otte A.P., et al. // Nature. 2002. V. 419. № 6907. P. 624–629.
230. Wagener N., Macher-Goeppinger S., Pritsch M., Husing J., Hoppe–Seyler K., Schirmacher P., Pfitzenmaier J., Haferkamp A., Hoppe-Seyler F., Hohenfellner M. // BMC Cancer. 2010. V. 10. P. 524.
231. Wang C.G., Ye Y.J., Yuan J., Liu F.F., Zhang H., Wang S. // World J. Gastroenterol. 2010. V. 16. № 19. P. 2421–2427.
232. Wilson B.G., Wang X., Shen X., McKenna E.S., Lemieux M.E., Cho Y.J., Koellhoffer E.C., Pomeroy S.L., Orkin S.H., Roberts C.W. // Cancer Cell. 2010. V. 18. № 4. P. 316–328.
233. Xia R., Jin F.Y., Lu K., Wan L., Xie M., Xu T.P., De W., Wang Z.X. // Tumour Biol. 2015. V. 36. № 7. P. 5341–5351.
234. Yan J., Ng S.B., Tay J.L., Lin B., Koh T.L., Tan J., Selvarajan V., Liu S.C., Bi C., Wang S., et al. // Blood. 2013. V. 121. № 22. P. 4512–4520.
235. Yu H., Simons D.L., Segall I., Carcamo-Cavazos V., Schwartz E.J., Yan N., Zuckerman N.S., Dirbas F.M., Johnson D.L., Holmes S.P., et al. // PLoS One. 2012. V. 7. № 12. P. e51239.
236. Zhang M.J., Chen D.S., Li H., Liu W.W., Han G.Y., Han Y.F. // Int. J. Clin. Exp. Pathol. 2019. V. 12. № 6. P. 2184–2194.
237. Abudurexiti M., Xie H., Jia Z., Zhu Y., Zhu Y., Shi G., Zhang H., Dai B., Wan F., Shen Y., et al. // Front. Oncol. 2019. V. 9. P. 856.
238. Liu C., Shi X., Wang L., Wu Y., Jin F., Bai C., Song Y. // Tumour Biol. 2014. V. 35. № 6. P. 6073–6082.
239. Bodor C., O'Riain C., Wrench D., Matthews J., Iyengar S., Tayyib H., Calaminici M., Clear A., Iqbal S., Quentmeier H., et al. // Leukemia. 2011. V. 25. № 4. P. 726–729.
240. Lohr J.G., Stojanov P., Lawrence M.S., Auclair D., Chapuy B., Sougnez C., Cruz-Gordillo P., Knoechel B., Asmann Y.W., Slager S.L., et al. // Proc. Natl. Acad. Sci. USA. 2012. V. 109. № 10. P. 3879–3884.
241. Majer C.R., Jin L., Scott M.P., Knutson S.K., Kuntz K.W., Keilhack H., Smith J.J., Moyer M.P., Richon V.M., Copeland R.A., et al. // FEBS Lett. 2012. V. 586. № 19. P. 3448–3451.
242. Morin R.D., Johnson N.A., Severson T.M., Mungall A.J., An J., Goya R., Paul J.E., Boyle M., Woolcock B.W., Kuchenbauer F., et al. // Nat. Genet. 2010. V. 42. № 2. P. 181–185.
243. Morin R.D., Mendez-Lago M., Mungall A.J., Goya R., Mungall K.L., Corbett R.D., Johnson N.A., Severson T.M., Chiu R., Field M., et al. // Nature. 2011. V. 476. № 7360. P. 298–303.
244. Reddy A., Zhang J., Davis N.S., Moffitt A.B., Love C.L., Waldrop A., Leppa S., Pasanen A., Meriranta L., Karjalainen-Lindsberg M.L., et al. // Cell. 2017. V. 171. № 2. P. 481–494.
245. Ryan R.J., Nitta M., Borger D., Zukerberg L.R., Ferry J.A., Harris N.L., Iafrate A.J., Bernstein B.E., Sohani A.R., Le L.P. // PLoS One. 2011. V. 6. № 12. P. e28585.
246. McCabe M.T., Graves A.P., Ganji G., Diaz E., Halsey W.S., Jiang Y., Smitheman K.N., Ott H.M., Pappalardi M.B., Allen K.E., et al. // Proc. Natl. Acad. Sci. USA. 2012. V. 109. № 8. P. 2989–2994.
247. Ott H.M., Graves A.P., Pappalardi M.B., Huddleston M., Halsey W.S., Hughes A.M., Groy A., Dul E., Jiang Y., Bai Y., et al. // Mol. Cancer Ther. 2014. V. 13. № 12. P. 3062–3073.
248. Sneeringer C.J., Scott M.P., Kuntz K.W., Knutson S.K., Pollock R.M., Richon V.M., Copeland R.A. // Proc. Natl. Acad. Sci. USA. 2010. V. 107. № 49. P. 20980–20985.
249. Yap D.B., Chu J., Berg T., Schapira M., Cheng S.W., Moradian A., Morin R.D., Mungall A.J., Meissner B., Boyle M., et al. // Blood. 2011. V. 117. № 8. P. 2451–2459.
250. Calebiro D., Grassi E.S., Eszlinger M., Ronchi C.L., Godbole A., Bathon K., Guizzardi F., de Filippis T., Krohn K., Jaeschke H., et al. // J. Clin. Invest. 2016. V. 126. № 9. P. 3383–3388.
251. Bejar R., Stevenson K., Abdel-Wahab O., Galili N., Nilsson B., Garcia-Manero G., Kantarjian H., Raza A., Levine R.L., Neuberg D., et al. // N. Engl. J. Med. 2011. V. 364. № 26. P. 2496–2506.
252. De Raedt T., Beert E., Pasmant E., Luscan A., Brems H., Ortonne N., Helin K., Hornick J.L., Mautner V., Kehrer-Sawatzki H., et al. // Nature. 2014. V. 514. № 7521. P. 247–251.
253. Ernst T., Chase A.J., Score J., Hidalgo-Curtis C.E., Bryant C., Jones A.V., Waghorn K., Zoi K., Ross F.M., Reiter A., et al. // Nat. Genet. 2010. V. 42. № 8. P. 722–726.
254. Guglielmelli P., Biamonte F., Score J., Hidalgo-Curtis C., Cervantes F., Maffioli M., Fanelli T., Ernst T., Winkelman N., Jones A.V., et al. // Blood. 2011. V. 118. № 19. P. 5227–5234.

255. Khan S.N., Jankowska A.M., Mahfouz R., Dunbar A.J., Sugimoto Y., Hosono N., Hu Z., Cheriyath V., Vatolin S., Przychodzen B., et al. // Leukemia. 2013. V. 27. № 6. P. 1301–1309.
256. Koontz J.I., Soreng A.L., Nucci M., Kuo F.C., Pauwels P., van Den Berghe H., Dal Cin P., Fletcher J.A., Sklar J. // Proc. Natl. Acad. Sci. USA. 2001. V. 98. № 11. P. 6348–6353.
257. Lee W., Teckie S., Wiesner T., Ran L., Prieto Granada C.N., Lin M., Zhu S., Cao Z., Liang Y., Sboner A., et al. // Nat. Genet. 2014. V. 46. № 11. P. 1227–1232.
258. Li H., Ma X., Wang J., Koontz J., Nucci M., Sklar J. // Proc. Natl. Acad. Sci. USA. 2007. V. 104. № 50. P. 20001–20006.
259. Ma X., Wang J., Wang J., Ma C.X., Gao X., Patriub V., Sklar J.L. // Oncotarget. 2017. V. 8. № 3. P. 4062–4078.
260. Makise N., Sekimizu M., Kobayashi E., Yoshida H., Fukayama M., Kato T., Kawai A., Ichikawa H., Yoshida A. // Virchows Arch. 2019. V. 475. № 4. P. 527–531.
261. Nikoloski G., Langemeijer S.M., Kuiper R.P., Knops R., Massop M., Tonnissen E.R., van der Heijden A., Scheele T.N., Vandenberghe P., de Witte T., et al. // Nat. Genet. 2010. V. 42. № 8. P. 665–667.
262. Ntziachristos P., Tsirigos A., van Vlierberghe P., Nedjic J., Trimarchi T., Flaherty M.S., Ferres-Marco D., da Ros V., Tang Z., Siegle J., et al. // Nat. Med. 2012. V. 18. № 2. P. 298–301.
263. Puda A., Milosevic J.D., Berg T., Klampfl T., Harutyunyan A.S., Gisslinger B., Rumi E., Pietra D., Malcovati L., Elena C., et al. // Am. J. Hematol. 2012. V. 87. № 3. P. 245–250.
264. Score J., Hidalgo-Curtis C., Jones A.V., Winkelmann N., Skinner A., Ward D., Zoi K., Ernst T., Stegelmann F., Dohner K., et al. // Blood. 2012. V. 119. № 5. P. 1208–1213.
265. Ueda T., Sanada M., Matsui H., Yamasaki N., Honda Z.I., Shih L.Y., Mori H., Inaba T., Ogawa S., Honda H. // Leukemia. 2012. V. 26. № 12. P. 2557–2560.
266. Zhang J., Ding L., Holmfeldt L., Wu G., Heatley S.L., Payne-Turner D., Easton J., Chen X., Wang J., Rusch M., et al. // Nature. 2012. V. 481. № 7380. P. 157–163.
267. Zhang M., Wang Y., Jones S., Sausen M., McMahon K., Sharma R., Wang Q., Belzberg A.J., Chaichana K., Gallia G.L., et al. // Nat. Genet. 2014. V. 46. № 11. P. 1170–1172.
268. Zhang Q., Han Q., Zi J., Ma J., Song H., Tian Y., McGrath M., Song C., Ge Z. // Genes Dis. 2019. V. 6. № 3. P. 276–281.
269. Bender S., Tang Y., Lindroth A.M., Hovestadt V., Jones D.T., Kool M., Zapatka M., Northcott P.A., Sturm D., Wang W., et al. // Cancer Cell. 2013. V. 24. № 5. P. 660–672.
270. Chan K.M., Fang D., Gan H., Hashizume R., Yu C., Schroeder M., Gupta N., Mueller S., James C.D., Jenkins R., et al. // Genes Dev. 2013. V. 27. № 9. P. 985–990.
271. Justin N., Zhang Y., Tarricone C., Martin S.R., Chen S., Underwood E., De Marco V., Haire L.F., Walker P.A., Reinberg D., et al. // Nat. Commun. 2016. V. 7. P. 11316.
272. Lewis P.W., Muller M.M., Koletsky M.S., Cordero F., Lin S., Banaszynski L.A., Garcia B.A., Muir T.W., Becher O.J., Allis C.D. // Science. 2013. V. 340. № 6134. P. 857–861.
273. Schwartzentruber J., Korshunov A., Liu X.Y., Jones D.T., Pfaff E., Jacob K., Sturm D., Fontebasso A.M., Quang D.A., Tonjes M., et al. // Nature. 2012. V. 482. № 7384. P. 226–231.
274. Sturm D., Witt H., Hovestadt V., Khuong-Quang D.A., Jones D.T., Konermann C., Pfaff E., Tonjes M., Sill M., Bender S., et al. // Cancer Cell. 2012. V. 22. № 4. P. 425–437.
275. Wu G., Broniscer A., McEachron T.A., Lu C., Paugh B.S., Becksfort J., Qu C., Ding L., Huether R., Parker M., et al. // Nat. Genet. 2012. V. 44. № 3. P. 251–253.
276. Boileau M., Shirinian M., Gayden T., Harutyunyan A.S., Chen C.C.L., Mikael L.G., Duncan H.M., Neumann A.L., Arreba-Tutusaus P., De Jay N., et al. // Nat. Commun. 2019. V. 10. № 1. P. 2891.
277. Lee C.H., Yu J.R., Granat J., Saldana-Meyer R., Andrade J., LeRoy G., Jin Y., Lund P., Stafford J.M., Garcia B.A., et al. // Genes Dev. 2019. V. 33. № 19–20. P. 1428–1440.
278. Hubner J.M., Muller T., Papageorgiou D.N., Mauermann M., Krijgsveld J., Russell R.B., Ellison D.W., Pfister S.M., Pajtler K.W., Kool M. // Neuro Oncol. 2019. V. 21. № 7. P. 878–889.
279. Jain S.U., Do T.J., Lund P.J., Rashoff A.Q., Diehl K.L., Cieslik M., Bajic A., Juretic N., Deshmukh S., Venneti S., et al. // Nat. Commun. 2019. V. 10. № 1. P. 2146.
280. Piunti A., Smith E.R., Morgan M.A.J., Ugarenko M., Khaltyan N., Helmin K.A., Ryan C.A., Murray D.C., Rickels R.A., Yilmaz B.D., et al. // Sci. Adv. 2019. V. 5. № 7. P. eaax2887.
281. Ragazzini R., Perez-Palacios R., Baymaz I.H., Diop S., Ancelin K., Zielinski D., Michaud A., Givelet M., Borsos M., Aflaki S., et al. // Nat. Commun. 2019. V. 10. № 1. P. 3858.
282. Beguelin W., Popovic R., Teater M., Jiang Y., Bunting K.L., Rosen M., Shen H., Yang S.N., Wang L., Ezponda T., et al. // Cancer Cell. 2013. V. 23. № 5. P. 677–692.
283. Chang C.J., Yang J.Y., Xia W., Chen C.T., Xie X., Chao C.H., Woodward W.A., Hsu J.M., Hortobagyi G.N., Hung M.C. // Cancer Cell. 2011. V. 19. № 1. P. 86–100.
284. Herrera-Merchan A., Arranz L., Ligos J.M., de Molina A., Dominguez O., Gonzalez S. // Nat. Commun. 2012. V. 3. P. 623.
285. Min J., Zaslavsky A., Fedele G., McLaughlin S.K., Reczek E.E., De Raedt T., Guney I., Strochlic D.E., Macconaill L.E., Beroukhim R., et al. // Nat. Med. 2010. V. 16. № 3. P. 286–294.
286. Berg T., Thoene S., Yap D., Wee T., Schoeler N., Rosten P., Lim E., Bilenky M., Mungall A.J., Oellerich T., et al. // Blood. 2014. V. 123. № 25. P. 3914–3924.
287. Abdel-Wahab O., Dey A. // Leukemia. 2013. V. 27. № 1. P. 10–15.
288. Danis E., Yamauchi T., Echanique K., Zhang X., Haladyna J.N., Riedel S.S., Zhu N., Xie H., Orkin S.H., Armstrong S.A., et al. // Cell Rep. 2016. V. 14. № 8. P. 1953–1965.
289. Booth C.A.G., Barkas N., Neo W.H., Boukarabila H., Soilleux E.J., Giotopoulos G., Farnoud N., Giustacchini A., Ashley N., Carrelha J., et al. // Cancer Cell. 2018. V. 33. № 2. P. 274–291.
290. Wang C., Oshima M., Sato D., Matsui H., Kubota S., Aoyama K., Nakajima-Takagi Y., Koide S., Matsubayashi J., Mochizuki-Kashio M., et al. // J. Clin. Invest. 2018. V. 128. № 9. P. 3872–3886.
291. Broux M., Prieto C., Demeyer S., Vanden Bempt M., Alberti-Servera L., Lodewijckx I., Vandepoel R., Mentens N., Gielen O., Jacobs K., et al. // Blood. 2019. V. 134. № 16. P. 1323–1336.
292. Sashida G., Harada H., Matsui H., Oshima M., Yui M., Harada Y., Tanaka S., Mochizuki-Kashio M., Wang C., Saraya A., et al. // Nat. Commun. 2014. V. 5. P. 4177.
293. Maertens O., Cichowski K. // Adv. Biol. Regul. 2014. V. 55. P. 1–14.
294. Abdel-Wahab O., Adli M., LaFave L.M., Gao J., Hricik T., Shih A.H., Pandey S., Patel J.P., Chung Y.R., Koche R., et al. // Cancer Cell. 2012. V. 22. № 2. P. 180–193.
295. Lane A.A., Chapuy B., Lin C.Y., Tivey T., Li H., Townsend E.C., van Bodegom D., Day T.A., Wu S.C., Liu H., et al. // Nat. Genet. 2014. V. 46. № 6. P. 618–623.

296. Simon C., Chagraoui J., Krosl J., Gendron P., Wilhelm B., Lemieux S., Boucher G., Chagnon P., Drouin S., Lambert R., et al. // Genes Dev. 2012. V. 26. № 7. P. 651–656.
297. Souroullas G.P., Jeck W.R., Parker J.S., Simon J.M., Liu J.Y., Paulk J., Xiong J., Clark K.S., Fedoriw Y., Qi J., et al. // Nat. Med. 2016. V. 22. № 6. P. 632–640.
298. Tan J., Yang X., Zhuang L., Jiang X., Chen W., Lee P.L., Karuturi R.K., Tan P.B., Liu E.T., Yu Q. // Genes Dev. 2007. V. 21. № 9. P. 1050–1063.
299. Miranda T.B., Cortez C.C., Yoo C.B., Liang G., Abe M., Kelly T.K., Marquez V.E., Jones P.A. // Mol. Cancer Ther. 2009. V. 8. № 6. P. 1579–1588.
300. Knutson S.K., Wigle T.J., Warholic N.M., Sneeringer C.J., Allain C.J., Klaus C.R., Sacks J.D., Raimondi A., Majer C.R., Song J., et al. // Nat. Chem. Biol. 2012. V. 8. № 11. P. 890–896.
301. McCabe M.T., Ott H.M., Ganji G., Korenchuk S., Thompson C., Van Aller G.S., Liu Y., Graves A.P., Della Pietra A., 3rd, Diaz E., et al. // Nature. 2012. V. 492. № 7427. P. 108–112.
302. Qi W., Chan H., Teng L., Li L., Chuai S., Zhang R., Zeng J., Li M., Fan H., Lin Y., et al. // Proc. Natl. Acad. Sci. USA. 2012. V. 109. № 52. P. 21360–21365.
303. Knutson S.K., Warholic N.M., Wigle T.J., Klaus C.R., Allain C.J., Raimondi A., Porter Scott M., Chesworth R., Moyer M.P., Copeland R.A., et al. // Proc. Natl. Acad. Sci. USA. 2013. V. 110. № 19. P. 7922–7927.
304. Versteege I., Sevenet N., Lange J., Rousseau-Merck M.F., Ambros P., Handgretinger R., Aurias A., Delattre O. // Nature. 1998. V. 394. № 6689. P. 203–206.
305. Kohashi K., Oda Y. // Cancer Sci. 2017. V. 108. № 4. P. 547–552.
306. Knutson S.K., Kawano S., Minoshima Y., Warholic N.M., Huang K.C., Xiao Y., Kadowaki T., Uesugi M., Kuznetsov G., Kumar N., et al. // Mol. Cancer Ther. 2014. V. 13. № 4. P. 842–854.
307. Konze K.D., Ma A., Li F., Barsyte–Lovejoy D., Parton T., Macnevin C.J., Liu F., Gao C., Huang X.P., Kuznetsova E., et al. // ACS Chem. Biol. 2013. V. 8. № 6. P. 1324–1334.
308. Fujita S., Honma D., Adachi N., Araki K., Takamatsu E., Katsumoto T., Yamagata K., Akashi K., Aoyama K., Iwama A., et al. // Leukemia. 2018. V. 32. № 4. P. 855–864.
309. Honma D., Kanno O., Watanabe J., Kinoshita J., Hirasawa M., Nosaka E., Shiroishi M., Takizawa T., Yasumatsu I., Horiuchi T., et al. // Cancer Sci. 2017. V. 108. № 10. P. 2069–2078.
310. Zhou Y., Du D.H., Wang J., Cai X.Q., Deng A.X., Nosjean O., Boutin J.A., Renard P., Yang D.H., Luo C., et al. // Chem. Biol. Drug. Des. 2020 May 11. doi: 10.1111/cbdd.13702. Online ahead of print.
311. Kasinath V., Faini M., Poepsel S., Reif D., Feng X.A., Stjepanovic G., Aebersold R., Nogales E. // Science. 2018. V. 359. № 6378. P. 940–944.
312. Margueron R., Justin N., Ohno K., Sharpe M.L., Son J., Drury W.J., 3rd, Voigt P., Martin S.R., Taylor W.R., De Marco V., et al. // Nature. 2009. V. 461. № 7265. P. 762–767.
313. Kim W., Bird G.H., Neff T., Guo G., Kerenyi M.A., Walensky L.D., Orkin S.H. // Nat. Chem. Biol. 2013. V. 9. № 10. P. 643–650.
314. Khan M., Walters L.L., Li Q., Thomas D.G., Miller J.M., Zhang Q., Sciallis A.P., Liu Y., Dlouhy B.J., Fort P.E., et al. // Lab. Invest. 2015. V. 95. № 11. P. 1278–1290.
315. He Y., Selvaraju S., Curtin M.L., Jakob C.G., Zhu H., Comess K.M., Shaw B., The J., Lima-Fernandes E., Szewczyk M.M., et al. // Nat. Chem. Biol. 2017. V. 13. № 4. P. 389–395.
316. Qi W., Zhao K., Gu J., Huang Y., Wang Y., Zhang H., Zhang M., Zhang J., Yu Z., Li L., et al. // Nat. Chem. Biol. 2017. V. 13. № 4. P. 381–388.
317. Lai A.C., Crews C.M. // Nat. Rev. Drug. Discov. 2017. V. 16. № 2. P. 101–114.
318. Ma A., Stratikopoulos E., Park K.S., Wei J., Martin T.C., Yang X., Schwarz M., Leshchenko V., Rialdi A., Dale B., et al. // Nat. Chem. Biol. 2020. V. 16. № 2. P. 214–222.

Involvement of the N Domain Residues E34, K35, and R38 in the Functionally Active Structure of *Escherichia coli* Lon Protease

A. G. Andrianova[1], A. M. Kudzhaev[1], V. A. Abrikosova[1], A. E. Gustchina[2], I. V. Smirnov[1], T. V. Rotanova[1*]

[1]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, 117997 Russia

[2]Macromolecular Crystallography Laboratory, NCI-Frederick, P.O. Box B, Frederick, MD 21702, USA

*E-mail: tatyana.rotanova@ibch.ru

Received September 09, 2020; in final form, October 21, 2020

DOI: 10.32607/actanaturae.11197

Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **ATP-dependent Lon protease of *Escherichia coli* (*Ec*Lon), which belongs to the superfamily of AAA$^+$ proteins, is a key component of the cellular proteome quality control system. It is responsible for the cleavage of mutant, damaged, and short-lived regulatory proteins that are potentially dangerous for the cell. *Ec*Lon functions as a homooligomer whose subunits contain a central characteristic AAA$^+$ module, a C-terminal protease domain, and an N-terminal non-catalytic region composed of the actual N-terminal domain and the inserted α-helical domain. An analysis of the N domain crystal structure suggested a potential involvement of residues E34, K35, and R38 in the formation of stable and active *Ec*Lon. We prepared and studied a triple mutant LonEKR in which these residues were replaced with alanine. The introduced substitutions were shown to affect the conformational stability and nucleotide-induced intercenter allosteric interactions, as well as the formation of the proper protein binding site.**

KEYWORDS **cellular proteome quality control, AAA$^+$ proteins, ATP-dependent proteolysis, LonA proteases, N domain.**

ABBREVIATIONS **AMPPNP – adenosine 5'-(β,γ-imido)triphosphate; DTDP – 4,4'-dithiodipyridine; Nu – nucleotide; PepTBE – Suc-Phe-Leu-Phe-SBzl; Suc – succinyl; OD – optical density.**

INTRODUCTION

ATP-dependent Lon proteases (MEROPS: clan SJ, family S16) are key components of the cellular protein quality control system that ensures proteome homeostasis in all kingdoms of nature. Along with Lon and other ATP-dependent proteases, the protein quality control (PQC) system includes molecular chaperones that are responsible for correct protein folding, formation of protein assemblies, and prevention of aggregate accumulation in the cell. In turn, ATP-dependent proteases and multisubunit bifunctional complexes, proteasomes, degrade damaged, mutant, and short-lived regulatory proteins that are potentially dangerous for the cell [1–6].

Lon proteases are homooligomeric enzymes. Their subunits include the ATPase (AAA$^+$) module formed by the nucleotide binding (NB) and α-helical (H) domains, the protease (P) domain that is a serine-lysine peptide hydrolase, and either the N-terminal or the inserted non-catalytic extra domain (ED) (*Fig. 1*) [7, 8].

Because Lon proteases, as well as other PQC proteases, contain the AAA$^+$ module in their structure, they belong to the superfamily of AAA$^+$ proteins (ATPases Associated with a variety of cellular Activities) that are abundant in nature and involved in important processes, such as DNA replication, transcription, cell division, intracellular transport, folding, proteolysis, etc. [9–12]. AAA$^+$ proteases are highly selective enzymes. Their main features are coupling of proteolytic activity with ATP hydrolysis and processive hydrolysis of protein targets to form extremely low-molecular-weight products (5–15 amino acid (aa) residues) [13–15].

ATP-dependent proteases select their substrates from a variety of cellular proteins based on the pres-

ence of special structural elements: exposed hydrophobic protein regions or labels called degrons. Degrons are specific amino acid sequences located at the end or inside of a substrate polypeptide chain [16–18]. Protein called ubiquitin serves as a label of substrates for eukaryotic proteasomes [19, 20]. The processive mechanism of substrate hydrolysis by AAA$^+$ proteases is implemented through a barrel-like quaternary structure of these enzymes. Their cylindrical oligomers use ATP energy for binding, denaturation, and translocation of protein substrates through the central pore, which is formed by stacked rings of ATPase modules and protease domains, to peptidase centers hidden within the enzyme oligomer [21–23].

To date, three subfamilies (A, B, and C; *Fig. 1A*) have been identified in the total pool of ATP-dependent Lon proteases in the MEROPS database. Differences in the environment of the catalytically active serine and lysine residues of proteolytic centers and the localization of extra domains controlling the ATPase component architecture serve as the basis for allocation of Lon enzymes into subfamilies [7, 8, 24]. Two types of proteolytic centers have been identified in the Lon protease family: the P_A type located in the P domains of the enzymes of the largest LonA subfamily comprising bacterial and eukaryotic enzymes [7, 8, 24, 25], and the P_B type detected in the enzymes of the archaeal LonB subfamily [8, 26] and a small bacterial subfamily, LonC (*Fig. 1A*) [27, 28].

The extra domain of LonA proteases is an extended N-terminal region that provides a distinctive feature of members of this subfamily. LonB and LonC proteases contain inserted extra domains located in their nucleotide-binding domains, between the Walker A and B motifs. A specific feature of the extra domain of LonB enzymes is its transmembrane segment. The extra domain of LonC proteases is characterized by being longer compared to that of the LonB extra domain and by degeneration of the ATPase function due to a replacement of some essential residues of the ATPase site

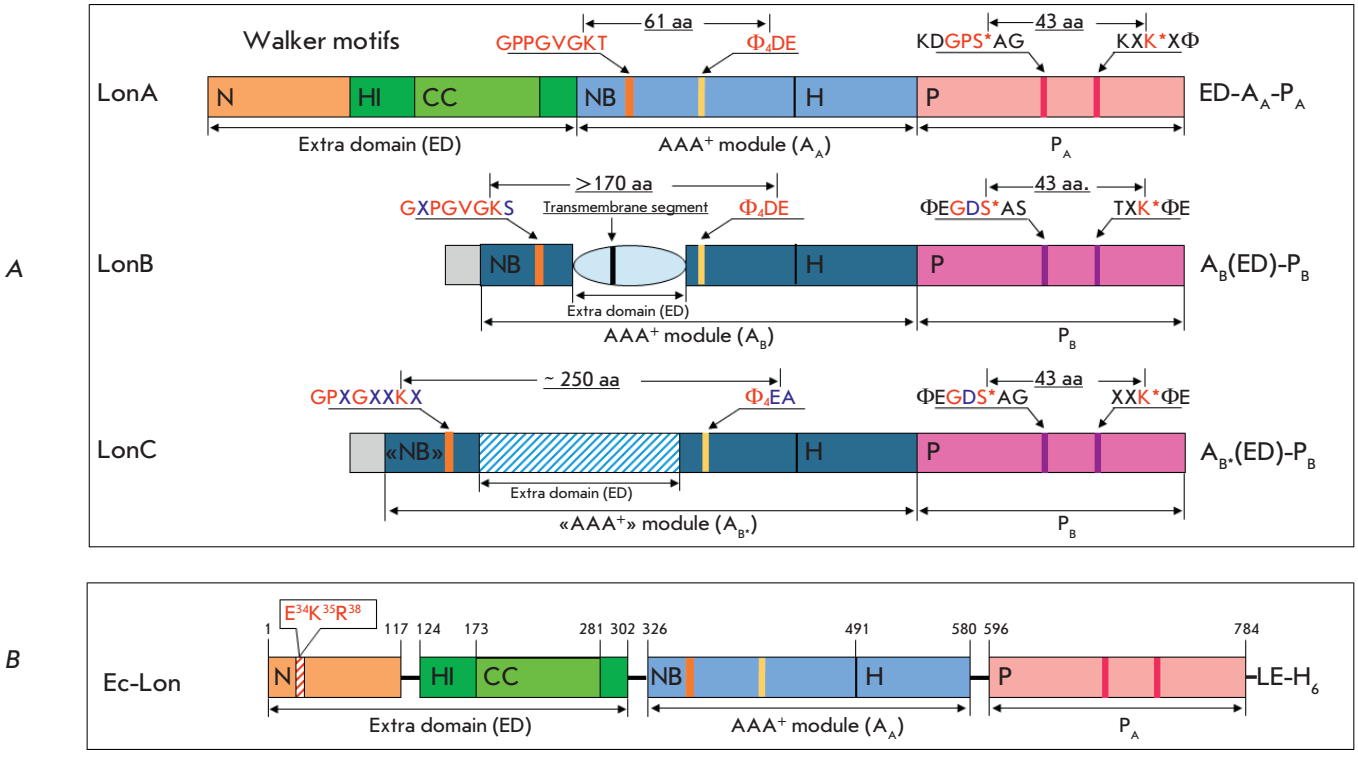


Fig. 1. Domain organization of Lon proteases from different subfamilies (*A*) and domain boundaries in the subunit of *E. coli* Lon protease (*B*). (*A*) **S*** and **K*** – catalytic residues of the proteolytic active site; Φ – hydrophobic amino acid residue; X – any amino acid residue; **P$_A$** and **P$_B$** – A-type (pink) and B-type (purple) protease domains; **A$_A$**, **A$_B$**, and **A$_B$·** – AAA$^+$ modules of A-type (light blue), B-type (blue), and "degenerate" B*-type (blue), respectively; **NB** – nucleotide-binding domain; **H** – α-helical domain; **ED** – extra domains represented by the **N**-domain (brown) and inserted α-helical **HI(CC)** domain (green) with a coiled-coil region (light green) in LonA proteases, a transmembrane domain (light blue) in LonB, and an inserted domain (shaded) in LonC; aa – amino acid residue; amino acid substitutions in conserved fragments are highlighted in blue. (*B*) *E. coli* Lon protease subunit with a C-terminal 6His-tag; the N domain region comprising E34, K35, and R38 residues is shaded

(*Fig. 1A*). However, LonC proteases are also involved in the protein quality system because regulation of their proteolytic activity is mediated by their retained ability to bind nucleotides [27].

Members of the LonA subfamily have been explored most extensively. Their N-terminal region has a two-domain structure [21, 29]. In the LonA protease of *E. coli* (*Ec*Lon), this region includes 325 aa and is formed by the "true" N-terminal (M1–Y117) and α-helical-inserted HI(CC) (E124–P302) domains (*Fig. 1B*) [29, 30]. The former has a twisted β-sheet structure and is topologically similar to RNA-binding PUA domains [31, 32]. The latter domain is formed by eight α-helices. It includes a region with a specific coiled-coil (CC) conformation, and moreover it is highly similar to the H domain of its own AAA^+ module, as well as to the α-helical domain of the first AAA^+ module of chaperone disaggregases ClpB/Hsp104, which contains an inserted M domain with a CC conformation [30, 31, 33].

To date, a lot of evidence has been accumulated showing the role of the AAA^+ module and protease domain in the functioning of LonA proteases. However, the functions of the N-terminal region of LonA proteases have not yet been fully characterized. According to published data, this region of the molecule is involved in the recognition and binding of substrate proteins [34–37]. Recently, the N-terminal region has been shown to participate in the formation of dodecameric structures from *E. coli* LonA protease hexamers [38, 39]. In addition, difference in the functions of the N and HI(CC) domains in the full-length *Ec*Lon protease has been revealed [40–45], confirming the two-domain organization of the enzyme's N-terminal region. Results of various studies indicate a crucial role played by the N-terminal region of LonA proteases in maintaining their functionally active conformation. In this case, it remains unclear which fragments of the N domain are important for the structural organization and are involved in the stabilization of enzymes.

The aim of this study was to identify the N-terminal domain residues involved in the formation of a stable, functionally active structure of the *Ec*Lon protease (hereinafter referred to as Lon protease), perform site-directed mutagenesis of these residues in order to produce a mutant enzyme, and investigate the structural and enzymatic characteristics of the mutant compared to those of intact *Ec*Lon.

EXPERIMENTAL

Materials

We used commercial reagents from Sigma, Bio-Rad, Thermo Scientific (USA), Fluka, Bachem (Switzerland), Boehringer Mannheim (Germany), Pharmacia (Sweden), Difco (England), Panreac (Spain), and Reakhim (Russia).

Preparation of recombinant *Ec*Lon protease (Ec-Lon) and its mutant form, LonEKR

Recombinant *Ec*Lon protease containing a hexahistidine fragment within the LEHHHHHH octapeptide at the C terminus of the protein (Ec-Lon) was produced according to a previously described procedure [40].

A triple mutant LonEKR was produced based on a megaprimer approach using the nucleotide sequence of Ec-Lon protease with the following primers: Lon_E34K35R38/AAA, T7 promoter, and f9 (5'-CCATCGCCGCTTCCAGACA AGCGATAGATGCTGCCCGCCCGACAAATAAGGGG-3', 5'-TTAATACGACTCACTATAGGGGA-3', and 5'-CGTTTACACCCGGCTCATCC-3', respectively). The gene fragment was amplified in two stages using plasmid DNA pET28-Ec-*lon* as a template. At the first stage, Lon_E34K35R38/AAA and T7 promoter primers were used to prepare a PCR fragment that, together with the f9 primer, was used as a primer at the second stage. The produced DNA fragment of about 250 bp was cloned into the pET28-Ec-*lon* vector at the unique XbaI and HindIII restriction sites.

Cloned DNA sequencing and primer synthesis were performed by EVROGEN (www.evrogen.ru). Restriction and ligation procedures were performed according to the protocols of the enzyme's manufacturers.

E. coli BL21(DE3) cells carrying the pET28-*lonEKR* plasmid were cultured in a LB medium with kanamycin at 37°C with vigorous stirring until OD_{600} reached 0.5, then the cell culture temperature was lowered to 25°C, and induction at 0.1/1 mM IPTG was performed for 3 h.

Ec-Lon and LonEKR were isolated and purified using Ni^{2+}-chelate affinity chromatography (HisTrap FF column, 5 mL, GE Healthcare, USA) and anion exchange chromatography (HiTrap™ Q FF column, 5 mL, GE Healthcare) according to the previously described procedure [40], followed by two-stage gel filtration on HiPrep™ 16/60 Sephacryl S-300 HR (120 mL, GE Healthcare) with the following buffers: 50 mM imidazole, pH 7.5, 0.5 M NaCl (stage 1) and 50 mM Tris-HCl, pH 7.5, 0.5 M NaCl (stage 2).

Protein concentrations were determined using the Bradford method [46].

The homogeneity of protein samples was tested electrophoretically [47] using a commercial set of markers (kDa): β-galactosidase (116.0), bovine serum albumin (66.2), ovalbumin (45.0), lactate dehydrogenase (35.0), restriction enzyme Bsp98I (25.0), β-lactalbumin (18.4), and lysozyme (14.4).

Determination of the enzymatic properties of Ec-Lon and its triple mutant LonEKR

ATPase activity was tested based on the kinetics of inorganic phosphate accumulation in the ATP hydrolysis reaction in 50 mM Tris-HCl buffer, pH 8.1, containing 200 mM NaCl, 2.5 mM ATP, 2.5 or 20 mM $MgCl_2$, and 0.1–1.0 μM enzyme, with and without β-casein (1 mg/mL), at 37°C [48]. In the control experiment, the enzyme was replaced with the buffer. The initial reaction rates were determined using the OD value of a mixture of 200 μL of the reaction medium and 600 μL of the reagent (100 mM $Zn(AcO)_2$, 15 mM $(NH_4)_6Mo_7O24$, 1% SDS, pH 4.5–5.0) at a wavelength of 350 nm (ε_{350} = 7,360 M^{-1} cm^{-1}).

Thioesterase activity. Hydrolysis of a thiobenzyl ester of the N-protected tripeptide Suc-Phe-Leu-Phe-SBzl (PepTBE) was monitored spectrophotometrically at a wavelength of 324 nm using the OD value of 4-thiopyridone (ε_{324} = 16,500 M^{-1} cm^{-1}) formed in the reaction between a hydrolysis product (benzyl thiolate, $BzlS^-$) and 4,4'-dithiodipyridine (DTDP) [49]. PepTBE was hydrolyzed at 37°C in 50 mM Tris-HCl buffer, pH 8.1, containing 200 mM NaCl, 10% DMSO, 0.2 mM DTDP, 0.1 mM PepTBE, and 0.1–1.0 μM enzyme. When studying the influence of effectors, a nucleotide, up to 2.5 mM, and $MgCl_2$, up to 20 mM, were added to the mixture.

Proteolytic activity of enzymes was tested electrophoretically [47]. The reaction was conducted at 37°C in 50 mM Tris-HCl buffer, pH 8.1, containing 200 mM NaCl, 20 μM β-casein, and 1 μM enzyme, with and without 5 mM Nu and 20 mM $MgCl_2$. In the control experiment, the enzyme was replaced with the buffer. An aliquot of the reaction or control mixture was mixed with the lysis buffer (0.2 M Tris-HCl, pH 8.9, 4% SDS, 20% glycerol, 0.5 mM EDTA, 0.8% bromophenol blue, 3% β-mercaptoethanol) at a 3 : 1 ratio, boiled for 5 min, and applied to a 12% polyacrylamide gel for electrophoresis.

The autolytic activity of enzymes was tested electrophoretically [47] under conditions similar to those for determining the proteolytic activity, but in the absence of β-casein.

Limited chymotrypsin proteolysis of Ec-Lon protease and its triple mutant LonEKR was carried out at 30°C in 50 mM Tris-HCl buffer, pH 8.1, containing 300 mM NaCl, 11 μM enzyme, and 0.2 μM chymotrypsin, with and without *Ec*Lon protease effectors.

RESULTS AND DISCUSSION

Identification of *Ec*Lon protease N domain residues presumably involved in formation of the functionally active enzyme

Previously, we have shown that the HI(CC)-inserted domain plays the key role in the correct binding of a protein substrate by the *Ec*Lon protease, efficient functioning of its ATPase and peptidase centers, implementation of intercenter allosteric interactions, and the processive mechanism of proteolysis [40–45]. In this case, the (E124–H172) and (M281–N302) fragments flanking the CC region were critically important for the interaction with a protein substrate and its hydrolysis [41–43].

We found [44] that the N-terminal domain ensures the conformational stability of the *Ec*Lon protease upon coupling of proteolysis with ATP hydrolysis, because a truncated enzyme (G107–K784) produced by the removal of the (M1–N106) fragment undergoes intensive autolysis, despite the preserved ability for processive proteolysis. In addition, the N-terminal domain residues R33, E34, and K35 were shown to be involved in the specific binding of *Ec*Lon substrates containing the so-called sul20-degron (a fragment of the cell division inhibitor SulA), which, in turn, affects the activities of ATPase and proteolytic centers [34].

At the same time, the results of the X-ray diffraction analysis of the N-terminal region of *E. coli* LonA

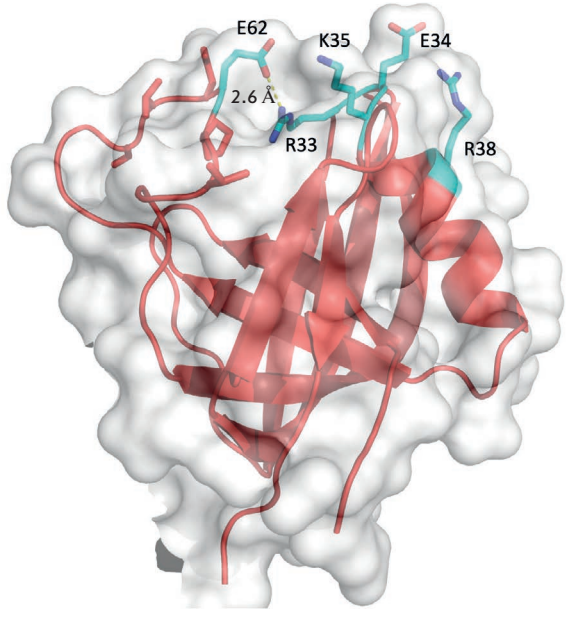


Fig. 2. Cartoon representation of the *Ec*Lon N domain comprising residues 7–118, with side chains of the R33, E34, K35, R38, and E62 residues shown in sticks. The solvent accessible surface of the protein is shown in light gray

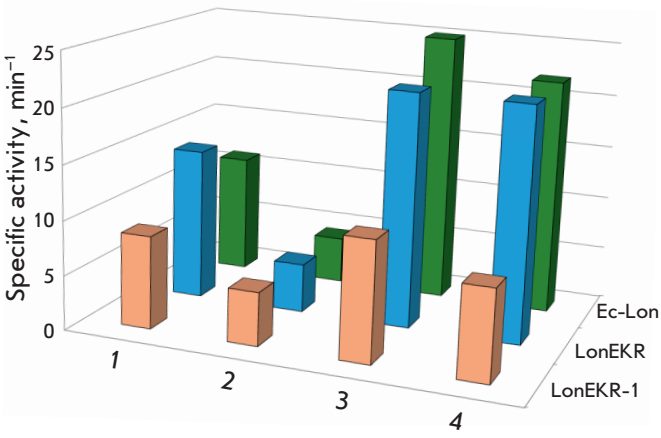


Fig. 3. ATPase activity of intact Ec-Lon protease and its LonEKR and LonEKR-1 mutants. Experimental conditions: 50 mM Tris-HCl buffer, pH 8.1; 0.2 M NaCl; 37°C; concentrations: 2.5 mM ATP; 2.5 (*1*, *3*) or 20 mM (*3*, *4*) $MgCl_2$; 0 (*1*, *2*) or 1.0 mg/mL (*3*, *4*) β-casein; 0.1–1.0 μM enzyme. The root-mean-square deviation R^2 in the experiments was 0.98–1.00

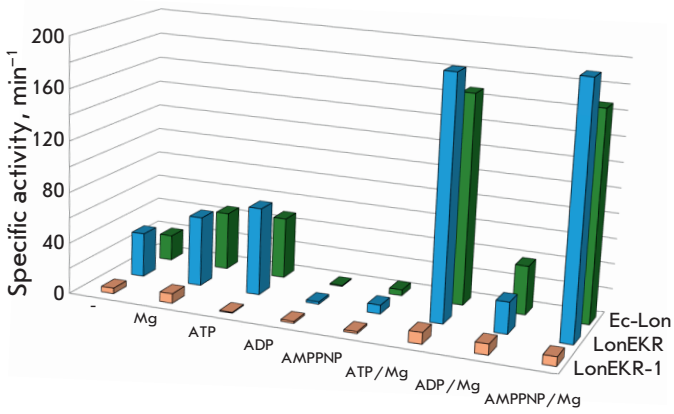


Fig. 4. Peptidase activity of intact Ec-Lon protease and its LonEKR and LonEKR-1 mutants. Experimental conditions: 50 mM Tris-HCl buffer, pH 8.1; 0.2 M NaCl; 10% DMSO; 0.2 mM DTDP; 37°C; concentrations: 0.1 mM PepTBE; 2.5 mM nucleotides; 20 mM $MgCl_2$; 0.1–1.0 μM enzyme. The root-mean-square deviation R^2 in the experiments was 0.98–1.00

protease [31] suggest that the region containing residues R33, E34, K35, and R38 may be important for interdomain and/or intersubunit interactions in the enzyme. This region is located on the surface of the *Ec*Lon protease N domain (*Fig. 2*), and, therefore, these residues may be directly involved in both the interactions with the substrate and the interactions between the protomers within the *Ec*Lon oligomers. The suggestion about the involvement of this region in the active structure and functioning of *Ec*Lon can be verified by studying the properties of a mutant enzyme with substitutions of potentially significant residues.

However, *Fig. 2* shows that the R33 residue forms an ion pair with the E62 residue located at the end of an 18 aa surface loop. This interaction restricts the mobility of this loop and, thereby, maintains its conformation. Mutation of the R33 residue may impair the topology of the studied region. For this reason, in this study, we investigated an *Ec*Lon protease mutant (LonEKR) in which only three residues, namely E34, K35, and R38, were substituted with alanine.

Preparation of the LonEKR triple mutant of *E. coli* Lon protease

The LonEKR mutant containing the E34A, K35A, and R38A substitutions was produced using recombinant *Ec*Lon containing a hexahistidine fragment at the C-terminus of the protein (Ec-Lon) [40]. The intact enzyme and its triple mutant were isolated according to a scheme including affinity chromatography on Ni-Sepharose, ion-exchange chromatography on Q-Sepharose, and gel filtration on Sephacryl S-300. The ATPase, peptidase, proteolytic, and autolytic activities were determined for the intact and mutant enzymes. When studying ATP hydrolysis, the effects of excess magnesium ions and of the protein substrate were evaluated. The peptidase (substrate, Suc-Phe-Leu-Phe-SBzl (PepTBE)), proteolytic (model protein substrate, β-casein), and autolytic activities were tested with and without Lon protease effectors – nucleotides and magnesium ions.

ATPase activity of the LonEKR mutant

Previously, intact Ec-Lon protease was shown to exhibit maximum ATPase activity in the reaction medium at pH 8.0–8.2 and at 2.5 mM equimolar ATP and magnesium ion concentrations. An increased concentration of Mg^{2+} ions, which is typical of physiological conditions (20 mM), results in a decrease in the ATPase activity. A protein substrate can restore the rate of ATP hydrolysis to its optimal values [40, 43].

The efficiency of ATP hydrolysis by the triple Ec-Lon protease mutant is close to that of the intact enzyme; in this case, the mutant retains its functional features, including inhibition by an excess of magnesium ions and subsequent activation of ATPase centers by β-casein (hereinafter referred to as casein). However, activation of the centers in the mutant in response to any interaction with casein is less effective than that in the intact Ec-Lon protease (*Fig. 3*), which may be due to weaker binding of the protein target caused by mutations of the E34, K35, and R38 residues.

In a separate experiment, producer strain cultivation conditions, in particular the induction condition, were shown to affect the efficiency of LonEKR ATPase centers. For the Ec-Lon protease and its modified forms,

the optimal conditions were chosen as those reducing the crowding effect during expression of the target gene: fermentation was performed in the presence of 0.1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at a temperature of 25°C. As the inducer concentration increased to 1 mM, the baseline ATPase activity of an isolated mutant (LonEKR-1) decreased by 40% compared to that of the intact enzyme (*Fig. 3*). The efficiency of LonEKR-1 ATPase activity recovery upon interaction with a protein substrate was also noticeably lower than that of the intact Lon protease and LonEKR mutant (*Fig. 3*). This suggests that IPTG at a concentration of 1 mM adversely affects the folding of the Ec-Lon protease mutant, which is also confirmed by LonEKR-1 gel filtration experiments demonstrating broadening and tailing of the protein peak compared to Lon and LonEKR.

Peptidase center activity of the LonEKR mutant

The efficiency of the peptidase centers of the intact Ec-Lon protease and its LonEKR mutant was assessed by the hydrolysis of a thiobenzyl ester of the N-protected tripeptide Suc-Phe-Leu-Phe-SBzl (PepTBE) [40]. During hydrolysis of the peptide substrate in the absence of nucleotide effectors, the LonEKR mutant was found to be more efficient (1.7-fold) than the intact Lon (*Fig. 4*). In this case, magnesium ions do not significantly activate the peptidase centers of both forms. Among free nucleotides, only ATP exhibits a weak but similarly efficient activating effect, whereas ADP and AMPPNP equally inhibit the peptidase activity, which indicates a similar affinity of nucleotides for the intact and mutant enzymes. The ATP/Mg and AMPPNP/Mg complexes exert the strongest activating effect on the peptidase sites of both Lon forms (*Fig. 4*). This indicates that the peptide hydrolase centers of the triple mutant act, in general, like centers of the intact enzyme.

These findings suggest that mutations in the E34, K35, and R38 residues of the *Ec*Lon N-terminal domain do not lead to significant changes in the functioning of enzyme peptidase centers. However, because ATP/Mg- and AMPPNP/Mg-based activation of the intact Lon noticeably exceeds that of the LonEKR form, it may be assumed that transmission of allosteric signals from the ATPase center to the peptidase center changes in the mutant, probably due to the differences in the efficiency of binding of Nu/Mg complexes.

It should be noted that the LonEKR-1 enzyme form produced upon expression of the mutant Lon protease gene in the presence of 1 mM IPTG exhibits a drastically decreased peptidase activity compared to that of the LonEKR mutant (*Fig. 4*). In the absence of effectors, hydrolysis of a low-molecular-weight substrate by LonEKR-1 is 8-fold slower than that by LonEKR,

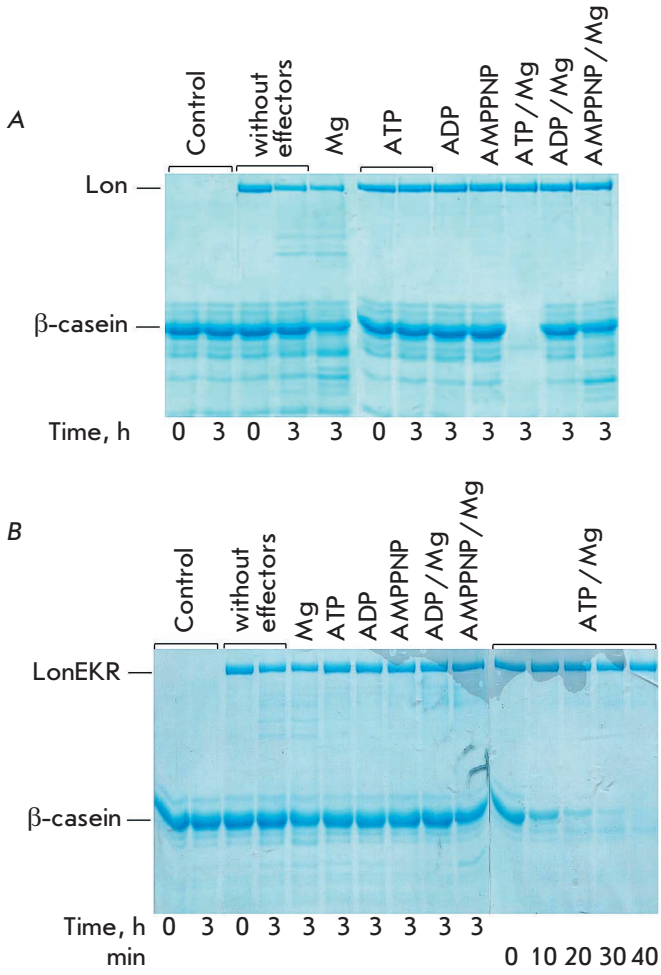


Fig. 5. Hydrolysis of β-casein by Ec-Lon protease (*A*) and its LonEKR mutant (*B*) with and without effectors (electrophoresis in 12% PAAG). Experimental conditions: 50 mM Tris-HCl buffer, pH 8.1; 0.2 M NaCl; 37°C; concentrations: 20 μM β-casein; 5 mM nucleotides; 20 mM $MgCl_2$; 1.0 μM enzyme

but the activating effect of magnesium ions remains. In contrast to the effect on LonEKR, any free nucleotides inhibit the peptidase activity of LonEKR-1 and their complexes with Mg^{2+} accelerate peptide hydrolysis only 2-fold, on average, which differs little from the effect of magnesium ions. Thus, as in the case of ATPase activity, these findings indicate that induction in the presence of 1 mM IPTG leads to significant conformational disruption in the enzyme structure, which affects its functional activity.

Proteolytic activity and autolytic properties of the LonEKR mutant

The proteolytic activity of Ec-Lon protease and its mutant was assessed by hydrolysis of β-casein (*Fig. 5*), similarly to refs. [40–45]. The LonEKR mutant retains the ability, characteristic of PQC enzymes, to hydrolyze

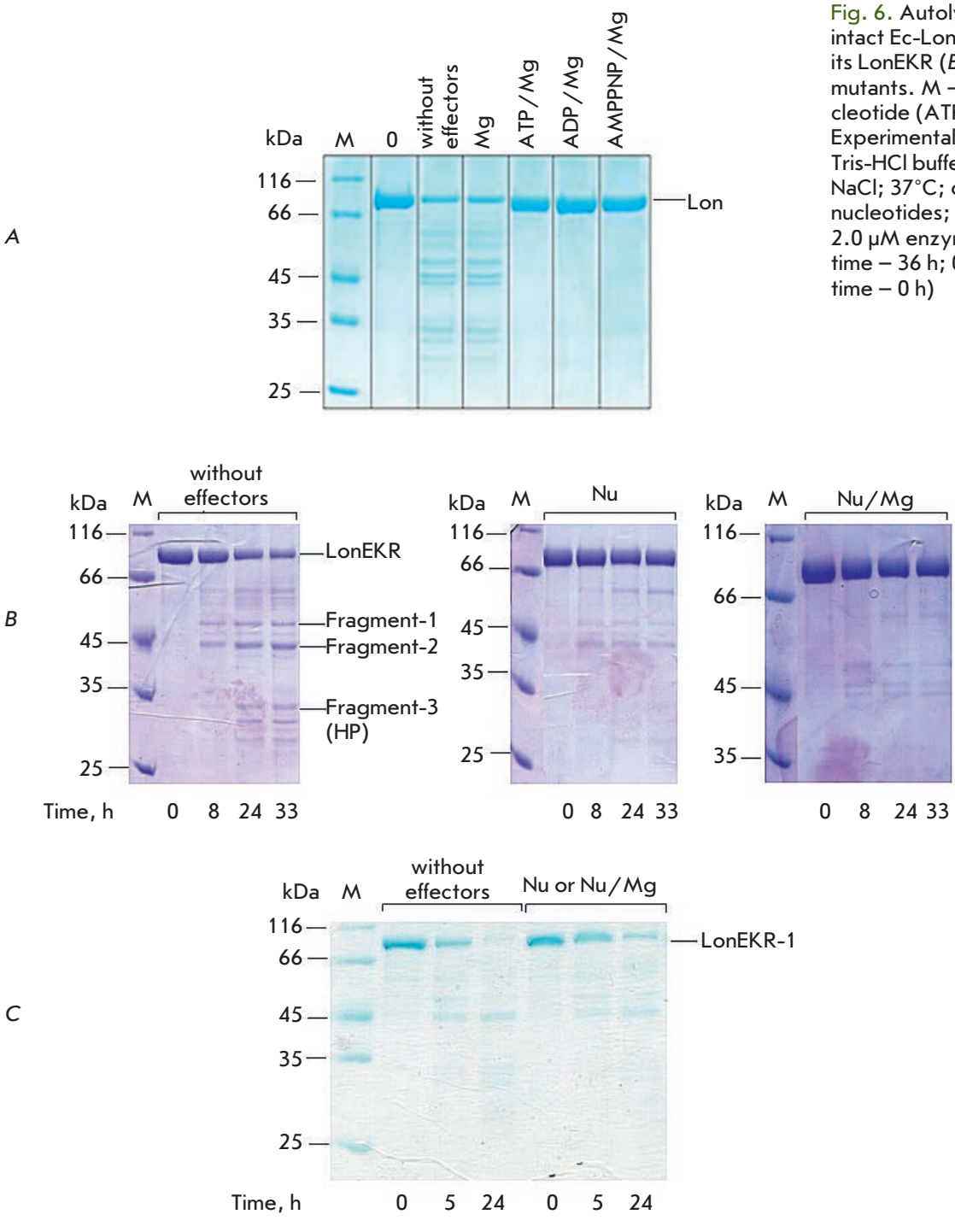


Fig. 6. Autolytic properties of intact Ec-Lon protease (*A*) and its LonEKR (*B*) and LonEKR-1 (*C*) mutants. M – markers; Nu – nucleotide (ATP, ADP, or AMPPNP). Experimental conditions: 50 mM Tris-HCl buffer, pH 8.1; 0.2 M NaCl; 37°C; concentrations: 5 mM nucleotides; 20 mM $MgCl_2$; 2.0 μM enzyme. (*A*) reaction time – 36 h; 0 – control (reaction time – 0 h)

a protein target via the processive mechanism (without releasing large intermediate products) upon coupling of proteolysis with ATP hydrolysis (*Fig. 5B*). This mechanism is implemented via the hexameric LonEKR structure, the formation of which was confirmed by gel filtration (data not shown). In the presence of the ATP/Mg complex, more than 50% of casein is degraded by the mutant in the first 10 min of reaction, which is comparable to the known efficiency of the ATP-dependent hydrolysis of this substrate by the native *Ec*Lon protease [43]. The intact enzyme is also characterized by an ability to degrade a protein substrate in the presence of the complex of a non-hydrolysable ATP analog, AMPPNP, with magnesium ions. In this case, the reaction products are high-molecular-weight fragments; i.e., proteolysis occurs by a non-processive

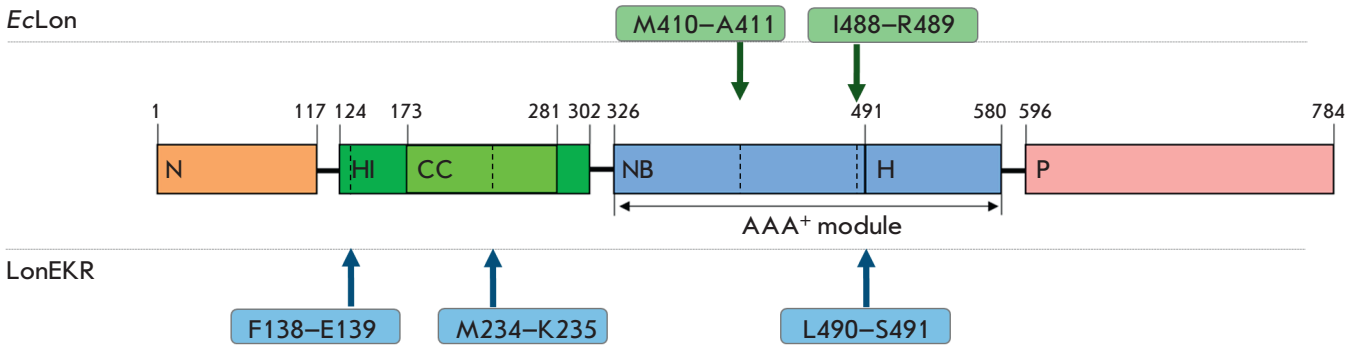


Fig. 7. Location of autolysis sites in native *Ec*Lon protease and LonEKR mutant

mechanism and with low efficiency (*Fig. 5A*). Magnesium ions may also be considered separately as activators of non-processive hydrolysis of casein by Ec-Lon protease (*Fig. 5A*). In contrast to the intact enzyme, the proteolytic activity of the LonEKR triple mutant in the presence of both magnesium ions and the AMPPNP/Mg complex proves to be almost absent over the same period of time (*Fig. 5B*). These results may reflect both a decreased efficiency in the binding of a protein substrate to LonEKR and disruption of allosteric interactions between the ATPase and proteolytic centers in the mutant enzyme.

As seen in *Fig. 5*, interaction between the enzyme and a protein substrate in the absence of effectors and in the presence of magnesium ions is accompanied by pronounced autolysis of the intact Lon protease and weak autolysis of the mutant. Investigation of the autolytic function of the native and mutant Lon forms in the absence of a target protein showed that the amounts of both enzymes significantly decreased over the experimental time interval (36 h for Lon and 33 h for LonEKR) (*Fig. 6A, B*). In this case, autolysis of the intact Lon occurred only in the absence of nucleotide effectors while autolysis of the LonEKR mutant occurred under any conditions, but nucleotides and their complexes with magnesium ions significantly stabilized the mutant enzyme.

N-terminal sequencing revealed that stable LonEKR fragments were formed by autolysis of the enzyme at bonds located in the inserted HI(CC) domain (F138–E139 and M234–K235) and at the boundary between the NB and H domains (L490–S491) (*Figs. 1B* and 7). The products of autolysis at the F138–E139 and M234–K235 bonds are a 50 kDa Fragment-1 and a 44 kDa Fragment-2, respectively, (*Fig. 6B*). In these products, the C-terminal regions of the LonEKR sequence (presumably P domains) are probably also cleaved. Autolytic cleavage of the triple mutant at the L490–S491 bond leads to formation of a Fragment-3 (33 kDa) that includes H and P domains (*Figs. 6B* and 7).

Stable fragments of native Lon protease were formed during autolysis in the NB domain at the M410–A411 and I488–R489 bonds and only in the absence of nucleotide effectors [43] (*Fig. 7*). In the latter case, as in LonEKR, a 33 kDa fragment comprising α-helical and protease domains (HP) was formed. Thus, the autolysis results indicate a difference in the conformations of the intact Lon protease and its triple mutant LonEKR, as well as the potential effect of the introduced mutations on the efficiency of binding of Nu/Mg complexes.

Cleavage of the native enzyme at the M234–K235 bond located in the characteristic "long helix" of the CC region is also possible, but this degradation pattern occurs only upon limited chymotryptic proteolysis of Lon in the presence of nucleotides or Nu/Mg complexes [50]. Thus, it may be suggested that the M234–K235 and L490–S491 (or I488–R489) bonds are located in Lon subunit regions accessible to various proteases. However, cleavage of the F138–E139 bond in the N-terminal α-helix of the HI(CC) domain has not yet been found either in native Lon protease or in any of its modified forms.

Autolysis sites in the HI(CC) domain (aa 124–302), which are not typical of intact Lon protease, were previously found in three N-terminal domain-truncated enzymes in the presence of the ATP/Mg complex. For example, under these conditions, a Lon-d106 form lacking the first 106 aa undergoes intense cleavage of the A267–K268 bond located at the N-terminus of the last helix of the CC region [44]. Because Lon-d106 is the only truncated enzyme retaining an ability for ATP-dependent processive hydrolysis of a protein substrate, it was concluded that the Lon protease N domain is not involved in the processive proteolysis mechanism, but its presence ensures the conformational stability of the enzyme under classical conditions of its functioning [44]. A Lon-d172 form lacking the first 172 residues is also unstable in the presence of the ATP/Mg complex and undergoes autolysis of the D245–D246 bond (central

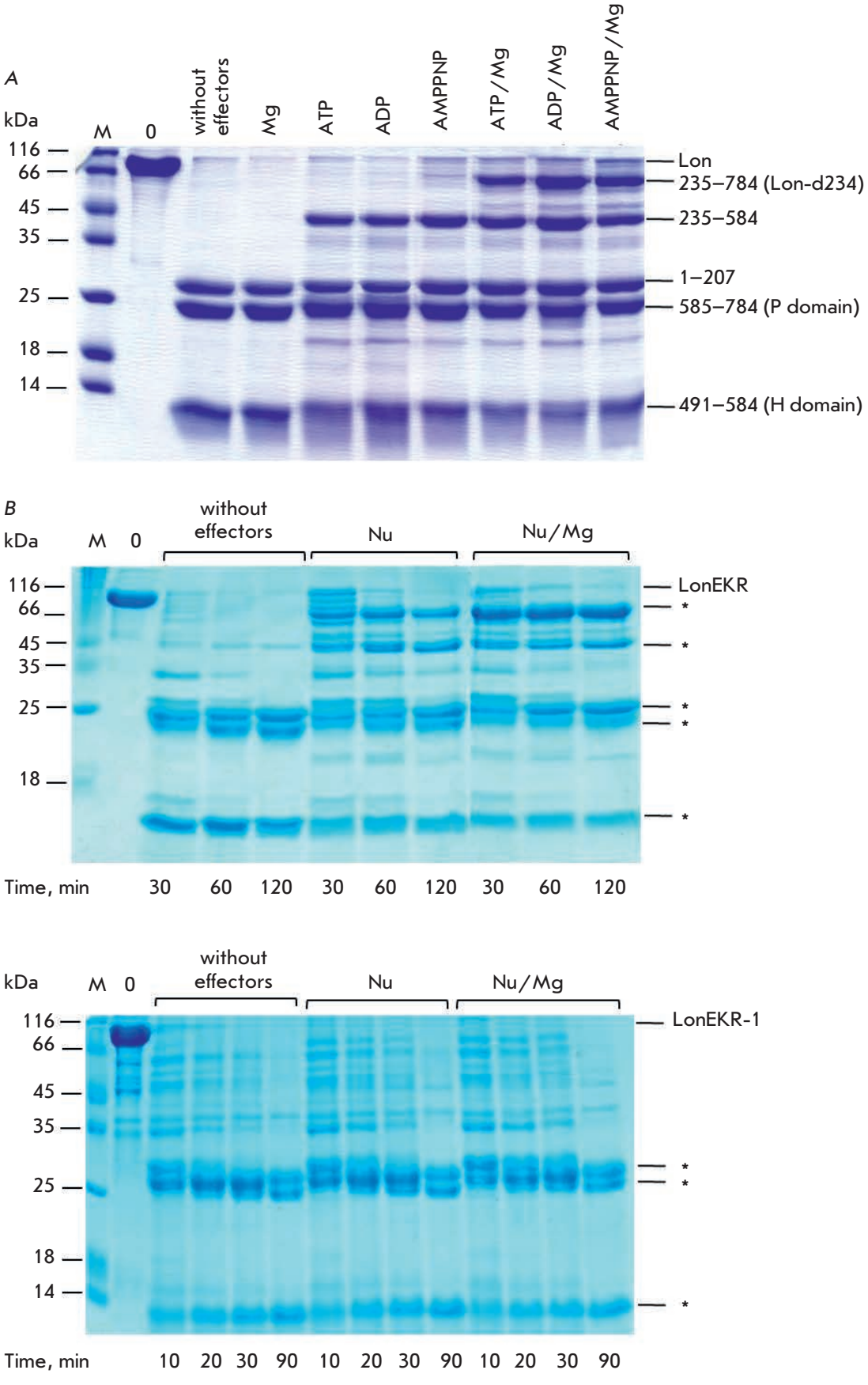


Fig. 8. Chymotrypsinolysis of native *Ec*Lon protease (*A*) and LonEKR (*B*) and LonEKR-1 (*C*) mutants. M – markers; 0 – reaction mixture sample at initial time; Nu – nucleotide (ATP, ADP, or AMPPNP). * – Products of LonEKR and LonEKR-1 chymotrypsinolysis whose N-termini are not confirmed by sequence analysis. Experimental conditions: 50 mM Tris-HCl buffer, pH 8.1; 0.3 M NaCl; 30°C; concentrations: 11 μM Lon (LonEKR or LonEKR-1); 5 mM nucleotides; 20 mM $MgCl_2$; 0.2 μM chymotrypsin. (*A*) reaction time – 2 h

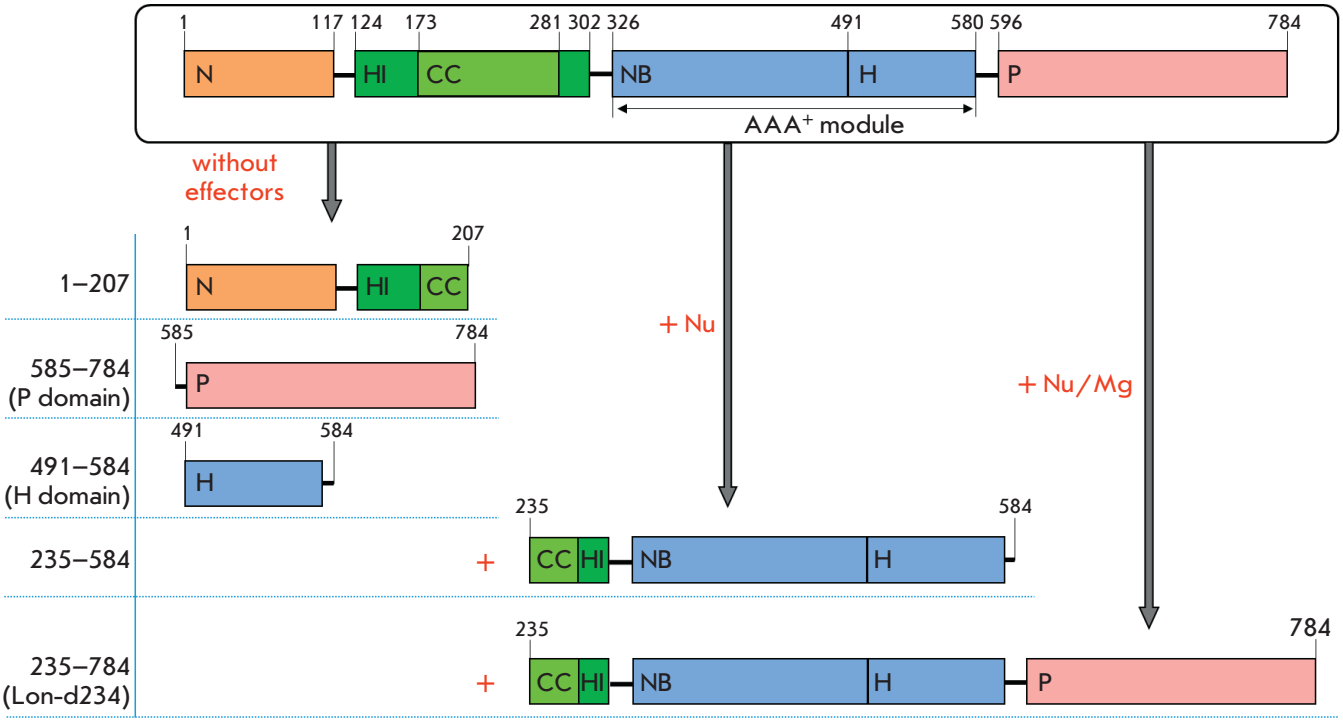


Fig. 9. Structures of the products of *Ec*Lon limited proteolysis by chymotrypsin

part of the CC region) [43]. A Lon protease fragment, Lon-d234 (aa 235–784), produced by limited proteolysis also exhibits increased autolytic activity upon coupling with ATP hydrolysis: autolysis amounts to 50% just after 20 min, with the cleavage occurring immediately after the CC region at the A286–E287 bond [50].

Thus, the introduction of three mutations into the Lon protease N-terminal domain was shown to noticeably destabilize the enzyme and cause conformational changes permitting exposure to the environment of a natively hidden region comprising the N terminus of the α-helical HI(CC) domain.

It should be noted that these LonEKR features become even more evident when the mutant gene is induced under conditions not optimal for this enzyme (1 mM IPTG). The LonEKR-1 mutant produced in this way undergoes almost complete autolysis within a day, regardless of the presence of nucleotides or nucleotide-magnesium complexes in the reaction mixture (*Fig. 6C*).

To further characterize the conformational stability of Lon protease and its LonEKR mutant, we also used limited chymotryptic proteolysis. The result of chymotrypsinolysis of native Lon protease is effector-dependent [50]. In the absence of effectors, only the N-terminal fragment (1–207) and P and H domains are formed, whereas the presence of a nucleotide leads to stabilization of the central NB domain and, as a result, to formation of an additional fragment (235–584) involving the AAA$^+$ module (326–584) and also a HI(CC) domain portion (235–302) with a linker (303–325) (*Figs. 1B* and *8A*). The products of Lon protease chymotrypsinolysis are shown schematically in *Fig. 9*. The presence of nucleotide-magnesium complexes stabilizes the region between the ATPase module and the protease domain, which leads to formation of the fragment (235–784), referred to above as Lon-d234 (*Figs. 8A* and *9*).

Limited chymotryptic proteolysis of the LonEKR form occurs in a similar way (*Figs. 8A* and *B*), and it may be assumed that the resulting fragments do not differ from the products of chymotrypsinolysis of the intact enzyme. However, in the case of the LonEKR-1 form produced with 1 mM IPTG, no stable NB domain-containing fragments of the sequence were detected either in the presence of nucleotides or in the presence of their complexes with magnesium ions (*Fig. 8B*). The chymotrypsinolysis results indicating that nucleotides and nucleotide-magnesium complexes do not stabilize the LonEKR-1 mutant structure are in full agreement with the autolysis data for this mutant. Therefore, induction of the *lonEKR* gene (1 mM IPTG) causes formation of an unstable conformation of the LonEKR-1 enzyme, which leads to its rapid autolytic cleavage under experimental conditions.

CONCLUSION

We previously established that the N-terminal domain provides conformational stability to *Ec*Lon protease. In this study, on the basis of X-ray structural data, we

proposed testing the role of residues E34, K35, and R38 of the N domain as amino acids involved in maintaining a stable structure of the functional enzyme through intersubunit and/or interdomain interactions. The replacement of these residues with alanine resulting in the triple LonEKR mutant was shown not to cause significant changes in the functioning of the ATPase and peptide hydrolase centers of the enzyme, but reduced binding of a protein substrate.

Like the native enzyme, the LonEKR mutant forms hexameric structures, but its ability to form dodecamers still remains unclear. Thus the LonEKR form retains the main property of ATP-dependent proteases – the ability to processively degrade a target protein when proteolysis is coupled with ATP hydrolysis, despite the detected disruption in intercenter allosteric interactions. However, in contrast to the intact enzyme, the LonEKR form is somewhat destabilized by the introduced substitutions because nucleotides and their complexes with magnesium ions, which are stabilizers of the Lon protease structure, are unable to completely prevent autolytic cleavage of the mutant.

It should be emphasized that gene induction and subsequent folding of the protein molecule play the key role in the formation of a stable structure of the functionally active Lon protease under crowding conditions. The LonEKR-1 form produced at a relatively high inducer concentration (1 mM IPTG) is not stabilized at all by nucleotides and exhibits an increased autolysis rate compared to the intact Lon and LonEKR form.

Therefore, this study has revealed that the N-terminal domain residues E34, K35, and R38 in the *Ec*Lon protease affect the formation of the correct binding site for a protein substrate, participate in the enzyme transformations caused by interaction with nucleotides, and maintain the conformational stability of the enzyme. Putative involvement of the studied residues in the formation of *Ec*Lon protease dodecameric forms may be a subject for future structural research into the properties of the LonEKR mutant. ●

The authors are grateful to Yu.F. Leonova for N-terminal sequencing of the EcLon protease fragments and LonEKR mutant.

This study was supported by the Russian Foundation for Basic Research (Project No. 19-04-00646).

REFERENCES

1. Gottesman S., Wickner S., Maurizi M.R. // Genes Dev. 1997. V. 11. P. 815–823.
2. Mogk A., Haslberger T., Tessarz P., Bukau B. // Biochem. Soc. Trans. 2008. V. 3. P. 120–125.
3. Tyedmers J., Mogk A., Bukau B. // Nat. Rev. Mol. Cell Biol. 2010. V. 11. P. 777–788.
4. Balchin D., Hayer-Hartl M., Hartl F.U. // Science. 2016. V. 353 (6294). aac4354.
5. Jeng W., Lee S., Sung N., Lee J., Tsai F.T.F. // F1000Research. 2015. V. 4. (F1000 Faculty Rev). 1448.
6. Finka A., Mattoo R.U.H., Goloubinoff P. // Annu. Rev. Biochem. 2016. V. 85. P. 715–742.
7. Rotanova T.V., Tsirulnikov K.B., Melnikov E.E. // Russ. J. Bioorg. Chem. 2003. V. 29. P. 85–87.
8. Rotanova T.V., Melnikov E.E., Khalatova A.G., Makhovskaya O.V., Botos I., Wlodawer A., Gustchina A. // Eur. J. Biochem. 2004. V. 271. P. 4865–4871.
9. Steinman J.B., Kapoor T.M. // Curr. Opin. Chem. Biol. 2019. V. 50. P. 45–54.
10. Miller J.M., Enemark E.J. // Archaea. 2016. V. 2016. P. 1–12.
11. Puchades C, Sandate C.R., Lander G.C. // Nat. Rev. Mol. Cell Biol. 2020. V. 21. P. 43–58.
12. White S.R., Lauring B. // Traffic. 2007. V. 8. P. 1657–1667.
13. Bittner L.M., Arends J., Narberhaus F. // Biopolymers. 2016. V. 105. P. 505–517.
14. Striebel F., Kress W., Weber-Ban E. // Curr. Opin. Struc. Biol. 2009. V. 19. P. 209–217.
15. Gottesman S. // Annu. Rev. Cell Dev. Biol. 2003. V. 19. P. 565–587.
16. Anthony J.R., Steven E.G. // J. Mol. Biol. 2017. V. 429. P. 873–885.
17. Baker T.A., Sauer R.T. // Trends Biochem. Sci. 2006. V. 31. P. 647–653.
18. Gur E., Sauer R.T. // Proc. Natl. Acad. Sci. USA. 2009. V. 106. P. 18503–18508.
19. Ciechanover A., Stanhill A. // Biochim. Biophys. Acta. 2014. V. 1843. P. 86–96.
20. Lopez-Castejon G. // FEBS J. 2020. V. 287(1). P. 11–26.
21. Sauer R.T., Baker T.A. // Annu. Rev. Biochem. 2011. V. 80. P. 587–612.
22. Gur E., Vishkautzan M., Sauer R.T. // Protein Sci. 2012. V. 21. P. 268–278.
23. Francis T.T., Christopher P.H. // eLife. 2020. V. 9. P. 1–3.
24. Rotanova T.V., Botos I., Melnikov E.E., Rasulova F., Gustchina A., Maurizi M.R., Wlodawer A. // Protein. Sci. 2006. V. 15. P. 1815–1828.
25. Botos I., Melnikov E.E., Cherry S., Tropea J.E., Khalatova A.G., Rasulova F., Dauter Z., Maurizi M.R., Rotanova T.V., Wlodawer A., Gustchina A. // J. Biol. Chem. 2004. V. 279. P. 8140–8148.
26. Botos I., Melnikov E.E., Cherry S., Kozlov S., Makhovskaya O.V., Tropea J.E., Gustchina A., Rotanova T.V., Wlodawer A. // J. Mol. Biol. 2005. V. 351. P. 144–157.
27. Liao J.H., Kuo C.I., Huang Y.Y., Lin Y.C., Lin Y.C., Yang C.Y., Wu W.L., Chang W.H., Liaw Y.C., Lin L.H., et al. // PLoS One. 2012. V. 7(7). P. 1–13.
28. Liao J.H., Ihara K., Kuo C.I., Huang K.F., Wakatsuki S., Wu S.H., Chang C.I. // Acta Crystallogr. D Biol. Crystallogr. 2013. V. 69. P. 1395–1402.
29. Rotanova T.V., Melnikov E.E. // Biochemistry (Moscow) Suppl. Series B: Biomed. Chem. 2010. V. 4. P. 404–408.
30. Rotanova T.V., Dergousova N.I., Morozkin A.D. // Russ. J. Bioorg. Chem. 2013. V. 39. P. 268–282.
31. Li M., Gustchina A., Rasulova F., Melnikov E.E., Maurizi

M.R., Rotanova T.V., Dauter Z., Wlodawer A. // Acta Crystallogr. Sec. D Biol. Crystallogr. 2010. V. 66. P. 865–873.
32. Bertonati C., Punta M., Fischer M., Yachdav G., Forouhar F., Zhou W., Kuzin A.P., Seetharaman J., Abashidze M., Ramelot T.A., et al // Proteins. 2009. V. 75. P. 760–773.
33. Rotanova T.V., Andrianova A.G., Kudzhaev A.M., Li M., Botos I., Wlodawer A., Gustchina A., et al. // FEBS OpenBio. 2019. V. 9. P. 1536–1551.
34. Wohlever M.L., Baker T.A., Sauer R.T. // Mol. Microbiol. 2014. V. 91. P. 66–78.
35. Rudyak S.G., Shrader T.E. // Protein Sci. 2000. V. 9. P. 1810–1817.
36. Adam C., Picard M., Déquard-Chablat M., Sellem C.H., Hermann-Le Denmat S., Contamine V. // PLoS One. 2012. V. 7. P. 1–10.
37. Ebel W., Skinner M.M., Dierksen K.P., Scott J.M., Trempy J.E. // J. Bacteriol. 1999. V. 181. P. 2236–2243.
38. Botos I., Lountos G.T., Wu W., Cherry S., Ghirlando R., Kudzhaev A.M., Rotanova T.V., de Val N., Tropea J., Gustchina A., Wlodawer A. // Curr. Res. Struct. Biol. 2019. V. 1. P. 13–20.
39. Vieux E.F., Wohlever M.L., Chen J.Z., Sauer R.T., Baker T.A. // Proc. Natl. Acad. Sci. USA. 2013. V. 110. P. 2002–2008.
40. Andrianova A.G., Kudzhaev A.M., Serova O.V., Dergousova N.I., Rotanova T.V. // Russ. J. Bioorg. Chem. 2014. V. 40. P. 620–627.
41. Kudzhaev A.M., Dubovtseva E.S., Serova O.V., Andrianova A.G., Rotanova T.V. // Russ. J. Bioorg. Chem. 2018. V. 44. P. 518–527.
42. Kudzhaev A.M., Andrianova A.G., Dubovtseva E.S., Serova O.V., Rotanova T.V. // Acta Naturae. 2017. V. 9. № 2. P. 75–81.
43. Andrianova A.G., Kudzhaev A.M., Dubovtseva E.S., Rotanova T.V. // Russ. J. Bioorg. Chem. 2017. V. 43. P. 368–376.
44. Kudzhaev A.M., Dubovtseva E.S., Serova O.V., Andrianova A.G., Rotanova T.V. // Russ. J. Bioorg. Chem. 2016. V. 42. P. 381–388.
45. Kudzhaev A.M., Andrianova A.G., Serova O.V., Arkhipova V.A., Dubovtseva E.S., Rotanova T.V. // Russ. J. Bioorg. Chem. 2015. V. 41. P. 518–524.
46. Bradford M.M. // Anal. Biochem. 1976. V. 72. P. 248–254.
47. Laemmli U.K. // Nature. 1970. V. 227. P. 680–685.
48. Bencini D.A., Wild J.R., O'Donovan G.A. // Anal. Biochem. 1983. V. 132. P. 254–258.
49. Castillo M.J., Nakajima K., Zimmerman M., Powers J.C. // Anal. Biochem. 1979. V. 99. P. 53–64.
50. Melnikov E.E., Andrianova A.G., Morozkin A.D., Stepnov A.A., Makhovskaya O.V., Botos I., Gustchina A., Wlodawer A., Rotanova T.V. // Acta Biochim. Pol. 2008. V. 55. P. 281–296.

A Simplified Streptozotocin-Induced Diabetes Model in Nude Mice

I. G. Gvazava[1,3], A. V. Kosykh[1,3], O. S. Rogovaya[1,3], O. P. Popova[2], K. A. Sobyanin[3], A. C. Khrushchev[3], A. V. Timofeev[3*], E. A. Vorotelyak[1,3]

[1]Koltsov Institute of Developmental Biology, Russian Academy of Sciences, Moscow, 119334 Russia

[2]National Medical Research Treatment and Rehabilitation Centre, Ministry of Health of the Russian Federation, Moscow, 125367 Russia

[3]Pirogov Russian National Research Medical University, Ministry of Health of the Russian Federation, Moscow, 117997 Russia

*E-mail: alvaltim@gmail.com

Received September 14, 2020; in final form, October 02, 2020

DOI: 10.32607/actanaturae.11202

Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Preclinical studies of human cellular and tissue-based products (HCT/Ps) for transplantation therapy of type 1 diabetes mellitus (T1DM) necessarily involve animal models, particularly mouse models of diabetes induced by streptozotocin (STZ). These models should mimic the clinical and metabolic manifestations of T1DM in humans (face validity) and be similar to T1DM in terms of the pathogenetic mechanism (construct validity). Furthermore, since HCT/Ps contain human cells, modeling of diabetes in immune-deficient animals is obligatory. Here we describe the most simplified diabetes model in Nude mice. Diabetes was induced in 31 males by a single intraperitoneal injection of STZ in normal saline at a medium-to-high dose of 150 mg/kg body weight. Fourteen control animals received only saline. Non-fasting plasma glucose (PG) levels were measured periodically for 50 days. All STZ-treated mice survived beyond 50 days. By day 15 after STZ administration, 22 of 31 (71%) mice developed stable diabetes based on the following criteria: (1) non-fasting PG ≥ 15 mmol/L on consecutive measurements up until day 50; (2) no diabetes remission. The mean non-fasting PG in mice with stable diabetes over the period of 35 days was equal to 25.7 mmol/L. On day 50, mean plasma insulin concentration, mean pancreatic insulin content, and the average number of β-cells in pancreatic islets were 2.6, 8.4, and 50 times lower, respectively, than in the control animals. We consider that our Nude mouse model of diabetes meets face validity and construct validity criteria and can be used in preclinical studies of HCT/Ps.**

KEYWORDS **animal model, Nude mice, diabetes mellitus, streptozotocin.**

ABBREVIATIONS **HCT/P – human cellular and tissue-based product; IPGTT – intraperitoneal glucose tolerance test; PG – plasma glucose level; STZ – streptozotocin; T1DM – type 1 diabetes mellitus.**

INTRODUCTION

Over the past two decades, considerable progress has been made in the development of human cellular and tissue-based products (HCT/Ps) for the transplantation therapy of type 1 diabetes mellitus (T1DM) [1]. Preclinical studies of these HCT/Ps require the assessment of their antidiabetic (glucose-lowering) effect in animal models of diabetes. Streptozotocin (STZ)-induced diabetic mouse models are the ones used most commonly. This is due to their simplicity, low cost, and, most importantly, their pathogenetic and phenotypic adequacy [2, 3]. Pathogenetic adequacy implies similarity between the developmental mechanisms of STZ-induced diabetes in mice and T1DM in humans. In both cases, the disease is caused by the destruction of β-cells, resulting in insulin deficiency. Phenotypic adequacy refers to the similarity between the manifestations of STZ-induced diabetes and type 1 diabetes: mice develop hyperglycemia; the number of β-cells in the islets of Langerhans decreases sharply; polyuria, polydipsia, weight loss, and decreased viability are observed.

There are two main methods for diabetes induction by streptozotocin in mice: repeated low-dose administration of streptozotocin (40–60 mg/kg of animal weight) for 4–5 days and a single administration of a medium to high dose (100–250 mg/kg). The first method is slightly more efficient, though more laborious [2]. STZ is injected intraperitoneally or intravenously via either one of the tail veins or the penile vein (for males). For the intraperitoneal injection, there is a risk

of accidentally injuring the intestine, which leads to animal death. At the same time, possible penetration of STZ into the subcutaneous tissue rather than the peritoneal cavity weakens the diabetogenic effect of STZ [4]. Nevertheless, intraperitoneal administration of STZ is used much more often than intravenous injection, as the former method is simpler.

Being structurally and conformationally similar to glucose, STZ enters murine β-cells via the glucose transporter GLUT2. Since STZ competes with glucose for the uptake by this transporter, it is recommended not to feed the animals for at least 4 h prior to STZ administration in order to increase the efficiency of diabetes induction [5]. However, Chaudhry et al. showed that the effectiveness of diabetes induction by STZ is the same in both fed and fasting C57BL/6 and NOD/SCID mice [6]. Administration of STZ to fed mice is preferable, since it allows one to eliminate the stress caused by starvation.

STZ is believed to rapidly lose its activity in neutral pH solutions. For this reason, many protocols recommend dissolving STZ in citrate buffer with a pH of 4–4.5 to induce diabetes [5, 7]. Even a small volume of citrate buffer at such low pH can cause peritoneal irritation and significantly shift the acid-base equilibrium. Therefore, many researchers use pH-neutral media (phosphate-buffered saline, Hanks' balanced salt solution, and 0.9% NaCl) to dissolve STZ [4, 6, 8].

It is important to note that mice challenged with high-dose STZ (> 200 mg/kg) rapidly develop dehydration (due to hyperglycemia and the general toxic effect of STZ) and severe hypoglycemia (caused by a massive release of insulin from destroyed β-cells). Subcutaneous injections of saline solutions are used to correct the water-electrolyte imbalance; a sucrose solution is administered orally to eliminate hypoglycemia [5, 9]. These measures are not necessary when using lower-dose STZ.

Studying the antidiabetic effect of HCT/Ps in diabetic mice involves a number of challenges:

· manifestation of the HCT/P effect usually requires quite a long time, from several weeks to several months. During this period, mice should maintain stable diabetes; i.e., the rate of spontaneous remission of the disease should be as low as possible;

· blood glucose levels in diabetic mice should be much higher than those in intact animals: it is the only way to confidently determine the effect of HCT/Ps;

· in order to study the effects of different doses of HCT/Ps and/or different methods of their transplantation, it is obligatory to have many groups of animals with stable diabetes while the size of each group should ensure the statistical reliability of the results. Therefore, the effectiveness of diabetes induction (morbidity) should be maximized, while the diabetes mortality rate should be minimized;

· increasing the STZ dose to enhance the effectiveness of diabetes induction raises the mortality rate among mice. Mortality can be reduced by constant therapy with low-dose insulin administration [9, 10]; however, this complicates the handling of the animals and makes it difficult to assess the effects of HCT/Ps;

· any HCT/P contains human cells, which are xenogeneic to recipient mice. For this reason, animals resistant to xenoantigens (and Nude mice in particular) are used to study antidiabetic HCT/Ps. The data on the suitability of Nude mice for modeling diabetes with STZ are rather controversial. Some researchers consider that these mice are especially vulnerable to the toxic effect of STZ because of their genetic aberrations [7]. Others believe that Nude mice are quite convenient for diabetes modeling with streptozotocin but still use insulin therapy to improve animal survival [9].

The aim of our study was to find the simplest and most reliable Nude mouse model of diabetes. The main problem needing a solution before any work could start was choosing the proper STZ dose. An analysis of the published data showed that stable diabetes can be induced in Nude mice from different breeders by a single administration of STZ at a dose range of 160–240 mg/kg. However, high animal mortality was observed when using such doses; it ranged from 7% to 100% for a period of 30 days after STZ injection [4, 9, 11, 12]. For this reason, we decided to use a lower dose of STZ. We conducted preliminary experiments in C57BL/6 mice and found that STZ at a dose of 150 mg/kg provides an acceptable incidence of diabetes and almost a 100% survival rate (unpublished data). This was the dose used to induce diabetes in Nude mice in the present study.

EXPERIMENTAL

Animals

Male Nude Crl:NU(NCr)-*Foxn1*^{nu} mice (age, 15–18 weeks; average weight, 31.5 ± 3.3 g) purchased from Charles River Breeding Laboratories (Germany) were used. All work with mice was performed under SPF conditions. The animals received sterilized chaw and water *ad libitum*. Mice were maintained at a temperature of 20–25°C on a 12:12 h light/dark cycle. All the experiments were carried out in accordance with the Guide for the Care and Use of Laboratory Animals of Pirogov Russian National Research Medical University dated March 27, 2019, in compliance with European Directive 2010/63/EU on the protection of experimental animals.

Method of diabetes induction

The animals were divided into two groups: the experimental (D, $n = 31$) and control ones (C, $n = 14$). In group D mice, diabetes was induced by a single intraperitoneal injection of STZ (Sigma S0130, USA) at a dose of 150 mg/kg; mice were deprived of food 4 h prior to administration. STZ was dissolved in cold 0.9% NaCl immediately before the injection; the injection volume was 450–550 μL. Group C mice were injected with 0.9% NaCl.

Methods for assessing the diabetogenic effect of STZ

In all animals, non-fasting PG was determined prior to STZ administration (on day 0), as well as on days 8, 10, and then every 5 days until day 50 after STZ administration in the time period between 13:00 and 15:00. A Contour TS glucose meter and corresponding test strips (Bayer, Switzerland) were used to measure PG. The diagnostic performance of the glucose meter and test strips was assessed periodically using control Contour solutions with low, normal, and high glucose concentrations. Blood samples for PG measurements were taken from tail tips. The High symbol was displayed on the screen at PG > 33.3 mmol/L. In such cases, the PG was considered equal to 33.3 mmol/L.

Diabetes was diagnosed when PG was equal to or exceeded 15 mmol/L for two consecutive readings (e.g., on days 8 and 10). Diabetes was considered stable if PG ≥ 15 mmol/L was obtained in all measurements between days 15 and 50. Diabetes remission was established if PG was below 15 mmol/L in at least one measurement on days 40 through 50.

On day 50, the intraperitoneal glucose tolerance test (IPGTT) was performed in group D mice with stable diabetes and in group C mice. Glucose dissolved in 500 μL of 0.9% NaCl was injected at a dose of 2 g/kg. At minutes 0 (prior to glucose injection), 15, and 60 of the test, mice were anesthetized with isoflurane (Baxter Healthcare Corporation, USA). Next, thoracotomy was performed, and 200–400 μL of blood was collected from the heart chambers into a lithium heparin tube (Microvette 500-LH, Sarstedt, Germany) using a 25G needle. PG was measured in the whole blood. The sample was then centrifuged, and the plasma insulin level was measured by ELISA (Mercodia, Sweden). After blood sampling, the mice were sacrificed by cervical dislocation.

Simultaneously with blood sampling at minute 0 of IPGTT, the pancreas was removed from the sacrificed mice and divided into three fragments. The first fragment was fixed in 10% neutral formalin (BioVitrum, Russia), embedded in paraffin, and then cut into 4- to 5-μm-thick sections. The sections were incubated with mouse anti-insulin antibodies (1 : 1000; catalog # 035K4884, Merck/Sigma, USA). Insulin-positive cells were detected using an EnVision FLEX kit (Agilent/Dako K8000, Denmark). The second fragment was frozen in liquid nitrogen, and 4-μm cryostat sections were prepared. These sections were sequentially incubated with rabbit anti-insulin antibodies (1 : 200; catalog # ab181547, Abcam, UK) and anti-rabbit Ig antibodies (1 : 500; Invitrogen Alexa Fluor Plus 488, A32790; ThermoFisher Scientific, USA). Next, the sections were mounted in Vectashield Antifade Mounting Medium with the DAPI fluorescent dye (H-1200, Vector Laboratories, USA). Immunomorphological studies were performed using a Nikon Eclipse 80i microscope (Nikon, Japan). The third fragment of the pancreas was used to assess the insulin content in the pancreatic tissue. The fragment was dried, weighed, minced with scissors in a minimal volume of water, and then sonicated. Insulin was extracted from the resulting suspension with a mixture of ethanol and hydrochloric acid [13]. Insulin concentration in the extract was measured by ELISA and normalized to the weight of the fragment.

The weight of the mice was measured in all groups at the beginning and end of the observation period.

Methods of statistical data processing and analysis

We used the MedCalc Statistical Software (version 19.4.0, MedCalc Software Ltd, Ostend, Belgium; https://www.medcalc.org; 2020). The normal distribution of data was assessed using the Shapiro–Wilk test. Intergroup differences were analyzed using the two-tailed Student's *t*-test in the case of a normal distribution of data and homogeneity of variance. Welch's *t*-test was used in case of a normal data distribution and heterogeneity of variance. In all cases, the level of significance of the differences was considered equal to 5% (α error = 0.05). The Kaplan–Meier plot analysis was used to estimate the diabetes incidence. The results of our measurements of PG, animal weight, plasma insulin levels, and insulin content in the pancreas are presented as a mean ± standard deviation with 95% confidence intervals for the means in the text and as mean ± standard deviations in figures.

RESULTS AND DISCUSSION

Effectiveness of diabetes induction

During the entire observation period, diabetes developed in 25 mice in group D (*Fig. 1*). However, stable diabetes was noted in only 22 animals. Thus, the effectiveness of induction of stable diabetes amounted to 71%. One mouse with late onset of diabetes developed remission; no remission was observed in mice with stable diabetes. The median incidence was 10 (10–15) days. None of the group D animals died within 50 days after STZ administration.

It is difficult to compare our data on the effectiveness of diabetes induction and survival rate to the results of other studies, since Nude mice from other breeders and administered different doses of STZ were used in these studies. For instance, Deeds et al. [4] conducted experiments in mice obtained from Taconic Farms (USA). After having received an STZ dose of 220 mg/kg, 92.5% of the animals developed severe diabetes on day 5; however, the mortality rate by day 20 was 20%. In a study by Graham et al. [9], Charles River mice (USA) developed stable diabetes on day 5 after administration of 240 mg/kg of STZ, while the mortality rate by day 30 was as low as 8%. However, such a low mortality rate might be explained by the fact that the animals received insulin therapy during the study period. In the study by Zhao et al. [12], the effectiveness of diabetes induction in mice purchased from the Shanghai Slac-ass breeding nursery (China) was 100% on day 8 after injection of 200 mg/kg of STZ; however, all mice died on day 30. Thus, our medium-dose model of diabetes is inferior to high-dose models in terms of the effectiveness of disease induction but superior to them in such an important parameter as animal survival.

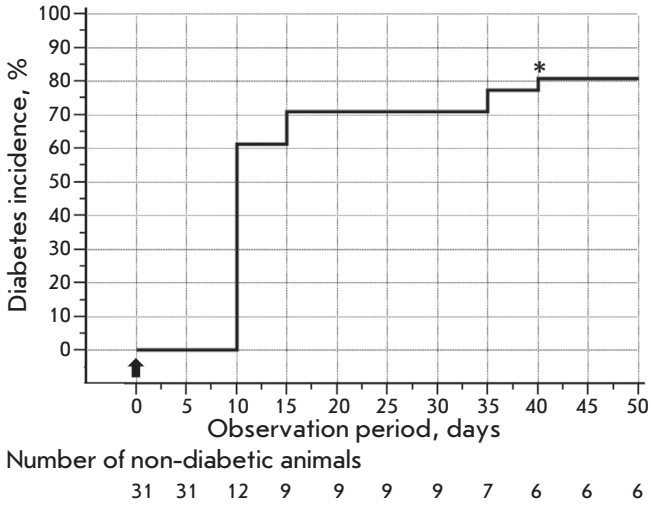


Fig. 1. Diabetes incidence in group D (Kaplan–Meyer analysis). The arrow indicates STZ injection; the asterisk marks the onset of diabetes remission in one of the animals

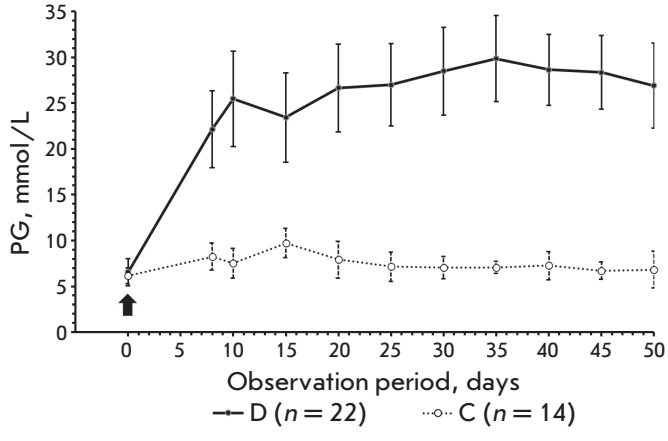


Fig. 2. Non-fasting PG values in group D mice with stable diabetes and group C mice during the observation period. The arrow indicates STZ injection

Changes in PG

Hyperglycemia in the diabetic range (PG ≥ 15 mmol/L) was observed in group D mice with stable diabetes starting from day 8 after STZ administration (*Fig. 2*). The mean group levels of PG for the entire observation period in group D mice with stable diabetes and in group C mice were 25.7 ± 3.5 (24.1–27.2) mmol/L and 7.5 ± 0.3 (7.1–7.8) mmol/L, respectively. The areas under the PG curves for the entire observation period were 1,258 ± 172 (1,184–1,332) and 365 ± 13 (349–382) mmol/L × 50 days, respectively ($P < 0.0001$ in both cases; Student's *t*-test). Our results of PG evaluation in the groups D and C are similar to those obtained by Deeds et al. [4]. In this study, the mean baseline PG in fed Nude mice was 7.7 ± 1.1 mmol/L. It increased to 28.6 ± 5.3 mmol/L seven days after STZ administration and remained at this level for 20 days.

Weight changes in mice

By the end of the observation period, the weight of mice with stable diabetes had decreased by an average of 4.8 ± 0.9%, while the weight of group C mice increased by 13 ± 5.8% (*Fig. 3*). Weight loss in STZ-induced diabetic rodents has been well documented and needs no further discussion.

The results of the intraperitoneal glucose tolerance test (IPGTT)

The basal insulin levels (at minute 0 of IPGTT) in group D mice with stable diabetes were 2.6 times lower than those in group C mice, equal to 67 ± 17 (49–85) pmol/L and 174 ± 31 (141–207) pmol/L, respectively ($P < 0.0001$; Student's *t*-test) (*Fig. 4*).

The areas under the PG curves in mice with stable diabetes and intact mice were 1,870 ± 108 (1,757–1,982) and 996 ± 160 (827–1,163) mmol/L × 60 min, respectively; the areas under the insulin level curves were 3,770 ± 849 (2,879–4,661) and 20,008 ± 4,052 (15,755–24,260) pmol/L × min, respectively ($P < 0.0001$ in both cases; Student's *t*-test).

The changes in PG and insulin levels observed by us during IPGTT in intact Nude mice were close to those

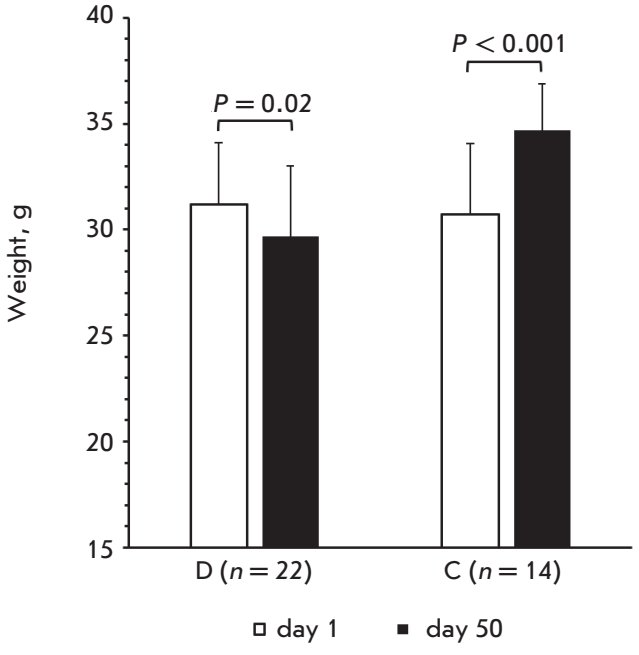


Fig. 3. Body weight of group D mice with stable diabetes and group C mice at the beginning and at the end of the observation period

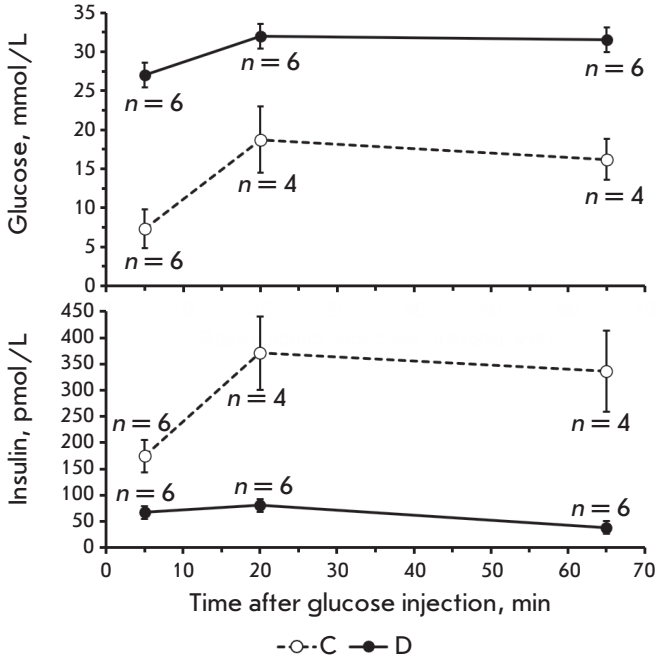


Fig. 4. Changes in PG and plasma insulin content during IPGTT in group D mice with stable diabetes and group C mice

in similar tests carried out both in Nude mice and mice of other strains. Thus, in a study by Christoffersson et al. [14], the highest PG in intact Nude mice recorded 15 min after the intraperitoneal glucose injection at a dose of 2.5 mg/kg was approximately 17 mmol/L, while the area under the PG curve was approximately 800 mmol/L × 60 min. Harper et al. [15] showed that insulin levels in intact outbred mice obtained from different breeders at minute 0 varied between 120 and 200 pmol/L, and maximum insulin levels were attained at minute 15 after glucose administration and ranged from 165 pmol/L to 280 pmol/L.

In our study, the plasma insulin levels in mice with stable diabetes were quite significant at all stages of IPGTT. Therefore, even in the presence of severe diabetes, Nude mice retain a certain number of functionally active β-cells. Residual insulin secretion is also observed in patients with type 1 diabetes for several years after clinical manifestation of the disease [16]. Thus, the presence of insulin in the plasma of group D mice with stable diabetes confirms the phenotypic similarity of our diabetic model to T1DM.

Insulin content in the pancreas

On day 50, the mean insulin levels in the pancreas of group D mice with stable diabetes and group C mice were 0.7 ± 0.3 (0.2–1.1) and 5.9 ± 0.6 (4.2–7.7) μg/g of the gland weight, respectively; $P < 0.0001$; Welch's *t*-test (*Fig. 5*).

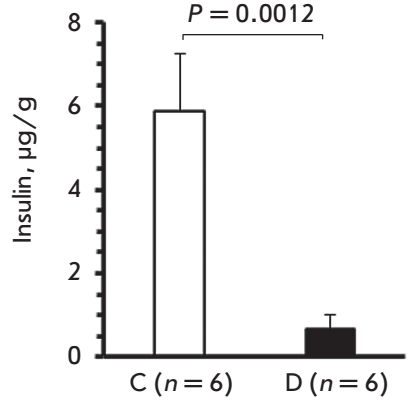


Fig. 5. Insulin content in the pancreatic tissue of group D mice with stable diabetes and group C mice on observation day 50

According to the published data, the insulin content in the mouse pancreas varies widely: it ranges from 2.5 to 80 μg/g of pancreatic weight in healthy animals and from 0.2 to 20 μg/g of pancreatic weight in animals with STZ-induced diabetes [3, 8, 12]. This wide fluctuation is due to interlinear, age, and sex differences in animals, different duration and severity of the diabetes, as well as the variety of the samples (entire pancreas, individual pancreatic lobes) and methods used for insulin extraction. Ultimately, when assessing the degree

of damage to β-cells, it is not the absolute amount of insulin in the pancreas of diseased and healthy animals but the ratio between these amounts that is of importance. For instance, in our study, the pancreatic insulin content in mice with stable diabetes on day 50 after a STZ injection was 8.9-fold lower than that in intact animals. In the study by Zhao et al. [12], the insulin content in the pancreas of diabetic mice on day 25 after the injection of STZ at a dose of 200 mg/kg was 18-fold lower than that in healthy animals.

Microscopic studies of the pancreas

By observation day 50, the number of islet β-cells had greatly decreased, and foci of intra- and peri-insular sclerosis occurred in animals with stable diabetes (*Fig. 6*). By having directly counted β-cells (*Fig. 7*), we found that their number in the islets of mice with stable diabetes had decreased about 50-fold compared to the control. A similar pathomorphological pattern is typical of diabetes induced by the administration of a single medium or high dose of STZ to mice [4, 12].

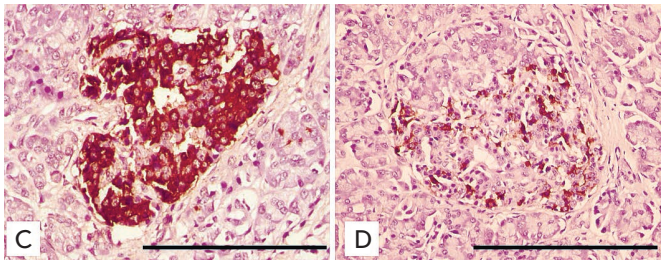


Fig. 6. Pancreatic islets in group D mice with stable diabetes and group C mice on day 50. Light microscopy, immunohistochemical staining for insulin, 400× magnification. Scale bar, 100 μm

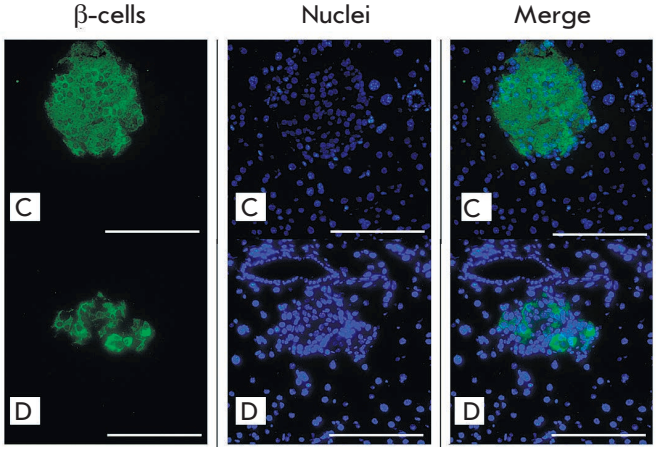


Fig. 7. Pancreatic islets in group D mice with stable diabetes and group C mice on day 50. Fluorescent microscopy; β-cells were stained for insulin (green); nuclei were stained with DAPI (magenta); 200× magnification. Scale bar, 100 μm

CONCLUSIONS

The advantages of our diabetes model are as follows:

· the use of Nude mice allows for transplantation of xenogeneic HCT/Ps containing human cells in animals;

· the method of diabetes induction is simplified as much as possible: STZ is administered once intraperitoneally;

· since 0.9% NaCl is used instead of a low-pH buffer solution to dissolve STZ, the risk of peritoneal irritation is eliminated while the general toxic effect of STZ is reduced;

· the effectiveness of diabetes induction is approximately 71%, while the survival rate is 100%. This makes it possible to form several experimental groups of mice with a group size sufficient to obtain statistically reliable experimental data;

· the use of medium-to-high doses of STZ requires neither correction of the water-electrolyte balance nor maintenance of insulin therapy;

· stable diabetes persists for a long time: from day 15 to day 50 after STZ administration. This period is sufficient to assess the antidiabetic effect of HCT/Ps;

· PG values are measured in fed animals. This eliminates the stress caused by prolonged starvation to animals;

· in animals with stable diabetes, PG is much higher than that in the control animals and there is also no spontaneous remission of the disease, thus simplifying the assessment of the antidiabetic effect of HCT/Ps;

· the model is phenotypically and pathogenetically similar to T1DM in humans; and

· the model allows one to conduct the biochemical, hormonal, and pathomorphological studies required in order to assess the antidiabetic effect of HCT/Ps.

We believe that our Nude mouse model of diabetes is well suited for preclinical studies of antidiabetic HCT/Ps and convenient for researchers. Its only drawback consists in the relatively low effectiveness of diabetes induction. This fact should be taken into account when settling on an initial number of animals. ●

This study was supported by the grant from the Ministry of Science and Higher Education of the Russian Federation.
Conflict of interests: the authors declare that there is no conflict of interest.

REFERENCES

1. Loretelli C., Assi E., Seelam A.J., Ben Nas M., Fiorina P. // Expert Opin. Biol. Ther. 2020. № 3. P. 1744–1768.
2. Gvazava I.G., Rogovaya O.S., Borisov M.A., Vorotelyak E.A., Vasiliev A.V. // Acta Naturae. 2018. V. 10. № 1. P. 24–33.
3. King A.J.F., Estil-les E., Montanya E. // Methods Mol. Biol. 2020. V. 2128. P. 135–147.
4. Deeds M.C., Anderson J.M., Armstrong A.S., Gastineau D.A., Hiddinga H.J., Jahangir A., Eberhardt N.L., Kudva Y.C. // Lab. Anim. 2011. V. 45. № 3. P. 131–140.
5. Diabetic Complications Consortium. High-dose streptozotocin induction protocol (mice). 2015. https://www.diacomp.org/shared/protocols.aspx.
6. Chaudhry Z.Z., Morris D.L., Moss D.R., Sims E.K., Chiong Y., Kono T., Evans-Molina C. // Lab. Anim. 2015. V. 47. № 4. P. 257–265.
7. Estil-les E., Tellez N., Nacher M., Montanya E. // Cell Transplant. 2018. V. 27. № 11. P. 1684–1691.
8. Kintoko K., Xu X., Lin X., JiaoY., Wen Q., Chen Z., Wei J., Liang T., Huang R. // Arch. Med. Sci. 2018. V. 14. № 5. P. 1163–1172.
9. Graham M.L., Janecek J.L., Kittredge J.A., Hering B.J., Schuurman H.J. // Comp. Med. 2011. V. 61. № 4. P. 356–360.
10. Cavelti-Weder C., Li W., Zumsteg A., Stemann-Andersen M., Zhang Y., Yamada T., Wang M., Lu J., Jermendy A., Bee Y.M., et al. // Diabetologia. 2016. V. 59. № 3. P. 522–532.
11. Ricordi C., Kneteman N.M., Scharp D.W., Lacy P.E. // World J. Surg. 1988. V. 12. № 6. P. 861–865.
12. Zhao T., Luo D., Sun Y., Niu X., Wang Y., Wang C., Jia W. // J. Mol. Histol. 2018. V. 49. № 4. P. 419–428.
13. Mercodia Technical Note No 334-0137 v.4.0. Analysis of insulin, C-peptide or proinsulin from acid ethanol extractions from islets, cells or tissue. Mercodia AB, 2019.
14. Christoffersson G., Henriksnas J., Johansson L., Rolny C., Ahlstrom H., Caballero-Corbalan J., Segersvard R., Permert J., Korsgren O., Carlsson P.O., et al. // Diabetes. 2010. V. 59. № 10. P. 2569–2578.
15. Harper J.M., Durkee S.J., Smith-Wheelock M., Miller R.A. // Exp. Gerontol. 2005. V. 40. № 4. P. 303–314.
16. Miller R.G., Yu L., Becker D.J., Orchard T.J., Costacou T. // Diabet. Med. 2020. V. 37. № 8. P. 1386–1394.

Intracellular Acidification Suppresses Synaptic Vesicle Mobilization in the Motor Nerve Terminals

A. L. Zefirov[1,2], R. D. Mukhametzyanov[1], A. V. Zakharov[1,3], K. A. Mukhutdinova[2], U. G. Odnoshivkina[1], A. M. Petrov[1,2,4*]

[1]Kazan State Medical University, Department of Normal Physiology, Kazan, 420012 Russia
[2]Institute of Neuroscience, Kazan State Medical University, Kazan, 420012 Russia
[3]Kazan Federal University, Kazan, 420008 Russia
[4]Laboratory of Biophysics of Synaptic Processes, Kazan Institute of Biochemistry and Biophysics, Federal Research Center "Kazan Scientific Center of RAS", Kazan, 420111 Russia
*E-mail: fysio@rambler.ru

Received June 18, 2020; in final form, September 07, 2020
DOI: 10.32607/actanaturae.11054

Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **Intracellular protons play a special role in the regulation of presynaptic processes, since the functioning of synaptic vesicles and endosomes depends on their acidification by the H^+-pump. Furthermore, transient acidification of the intraterminal space occurs during synaptic activity. Using microelectrode recording of postsynaptic responses (an indicator of neurotransmitter release) and exo-endocytic marker FM1-43, we studied the effects of intracellular acidification with propionate on the presynaptic events underlying neurotransmitter release. Cytoplasmic acidification led to a marked decrease in neurotransmitter release during the first minute of a 20-Hz stimulation in the neuromuscular junctions of mouse diaphragm and frog cutaneous pectoris muscle. This was accompanied by a reduction in the FM1-43 loss during synaptic vesicle exocytosis in response to the stimulation. Estimation of the endocytic uptake of FM1-43 showed no disruption in synaptic vesicle endocytosis. Acidification completely prevented the action of the cell-membrane permeable compound 24-hydroxycholesterol, which can enhance synaptic vesicle mobilization. Thus, the obtained results suggest that an increase in $[H^+]_{in}$ negatively regulates neurotransmission due to the suppression of synaptic vesicle delivery to the sites of exocytosis at high activity. This mechanism can be a part of the negative feedback loop in regulating neurotransmitter release.**

KEYWORDS **exocytosis, synaptic vesicle translocation, neurotransmission, acidification, neuromuscular junction.**
ABBREVIATIONS **AZ – active zone; EPP – end-plate potential; NE – nerve ending; SV – synaptic vesicle.**

INTRODUCTION

Synaptic transmission is based on neurotransmitter release from synaptic vesicles (SVs) via exocytosis in response to the arrival of an action potential from an axon to the nerve ending (NE). This mechanism is universal and depends on the transport (mobilization) of SVs to the sites of exocytosis, the so-called active zones (AZs), where the proteins involved in exocytosis and voltage-gated Ca^{2+} channels are concentrated [1]. In turn, the mobilization depends on the number of available SVs located near the AZs and the supply of SVs newly formed via endocytosis immediately after exocytosis. Under conditions of continuous rhythmic or moderate-frequency activity, the mobilization rate is responsible for the level of neurotransmitter release and, consequently, the reliability of neurotransmission [2, 3].

The fundamental mechanisms regulating SV mobilization remain insufficiently understood. The cytoskeleton, motor proteins, and small GTPases were shown to play an important role in the regulation of SV transport [4]. However, the significance of such an important factor as the cytoplasmic pH has not been clarified. Meanwhile, it is known that changes in pH_{in} related to proton pumping and the function of vesicular neurotransmitter transporters occur in NEs during synaptic activity. Neurotransmitter transporters exchanging a neurotransmitter molecule for proton(s) are incorporated into the presynaptic membrane after exocytosis of SVs, while maintaining their functional

activity [5, 6]. The Ca^{2+}/H^+ exchange by Ca^{2+} ATPase of the presynaptic plasma membrane can also take part in cytoplasmic acidification in response to increased $[Ca^{2+}]_{in}$ during depolarization, while the Na^+/H^+ exchanger is involved in pH_{in} restoration in NEs [7]. Intense stimulation was shown to decrease pH of the NE cytosol in the neuromuscular junctions of a fruit fly, mouse, and rat [5–7]. However, it remains unclear what effect intracellular acidosis caused by synaptic activity has on presynaptic processes.

The early studies showed that an abrupt drop in pH of cells can inhibit clathrin-mediated endocytosis [8]. This can be a result of impaired clathrin coat assembly, dysfunction of adapter proteins, or decreased synthesis of phosphatidylinositol-4,5- bisphosphates [9, 10]. However, the same cannot be extrapolated to the synaptic machinery, since endocytosis in the synapse is highly specific and requires a specific set of proteins to become involved. In addition, several types of endocytosis, including clathrin-independent ones, coexist in the synapse [11]. For example, a carbonic anhydrase inhibitor switches the type of endocytosis in the neuromuscular junctions of mice to a clathrin-independent one by lowering the cytosolic pH [12].

In general, it remains unclear how a reduced cytoplasmic pH can affect neurotransmitter release and the mobilization of SVs during continuous activity. In the present study, using electrophysiological detection of neurotransmitter release and a fluorescence-based method for tracking the exocytosis and endocytosis, we were able to demonstrate for the first time that intracellular acidification can significantly inhibit SV mobilization in the neuromuscular junctions of cold- and warm-blooded animals. We suggest that this phenomenon may be a new physiological mechanism regulating the SV transport during synaptic activity.

EXPERIMENTAL

Our experiments were carried out using isolated neuromuscular preparations from the diaphragm muscle of white laboratory mice and the cutaneous pectoris muscle of frogs (*Rana ridibunda*) in autumn and winter, in compliance with the Guide for the Care and Use of Laboratory Animals. The experiment protocol complied with European Directive 2010/63/EU on the protection of animals used for scientific purposes and was approved by the Ethics Committee of the Kazan Medical University.

Solutions and reagents

The muscle was attached to the bottom of a 5-mL chamber under continuous perfusion. An oxygenated Krebs solution of the following composition was used in the experiments performed on the mouse muscle: 144.0 mM NaCl, 5.0 mM KCl, 0.1 mM $MgCl_2$, 2.0 mM $CaCl_2$, 1.0 mM NaH_2PO_4, 2.4 mM $NaHCO_3$, and 11.0 mM glucose. Ringer's solution (115.0 mM NaCl, 2.5 mM KCl, 1.8 mM $CaCl_2$, and 2.4 mM $NaHCO_3$) was used in the experiments performed on the frog muscle. The pH of the solutions was maintained at 7.3–7.4 at a temperature of 20°C. D-tubocurarine (2–5 μM) was used to avoid muscle contraction. Modified Krebs and Ringer solutions, with sodium chloride partially replaced with sodium propionate (namely, 72 mM), were used to induce intracellular acidification. The resulting concentration of sodium propionate in the modified solutions was 72 mM; pH and osmolality were maintained identical to those of normal saline. Reagents procured from Sigma-Aldrich (USA) were used. The experiments were started after perfusion of the preparations with propionate solutions for 45–50 min; 24-hydroxycholesterol (0.4 μM) was applied for 15 min.

Electrophysiology

End-plate potentials (EPPs) were recorded intracellularly using glass microelectrodes (tip diameter < 1 μm; resistance, 5–20 mΩ) filled with 3 M KCl. The amplifier Model 1600 (A-M Systems, USA) and an LA-2USB analog-to-digital converter were used to amplify and record EPPs under the control of the Elph software [13]. The motor nerve was stimulated by rectangular 0.1–0.2 ms pulses at a frequency of 20 Hz for 3 min (Model 2100 Stimulator, A-M Systems, USA). The stimulation frequency was then reduced to 0.3 Hz, and the recovery of the EPP amplitude was recorded [14, 15].

The quantum content of EPPs was calculated using the modified method of variations described earlier in details [3]. For this purpose, the area of each EPP in the series was determined. Further, the region on the diagram showing the reduction in the EPP area during high-frequency stimulation in which the average EPP area remained practically unchanged was identified (the plateau phase, which usually lasts the first 10–30 s). The quantum value (i.e., the average area of EPPs produced by a single neurotransmitter quantum) can be calculated from the fluctuations in the EPP area within this region (q): $q = \sigma^2/\langle V \rangle$, where σ is the standard deviation of the EPP area, and $\langle V \rangle$ is the average EPP area in this region. Next, one can determine the quantum content of each EPP in the series: $m_i = V_i/q$, where m_i is the quantum content of the i^{th} EPP, and V_i is the area of the i^{th} EPP.

Fluorescence microscopy

Fluorescence was observed using an Olympus BX-51WI microscope. An Olympus UPLANSapo lens (60× magnification) and a LumPlanPF lens (100× magnification) were used. Images were recorded using an

Olympus DP71 camera and processed using the CellSens software (Olympus). The ImagePro (Media Cybernetics) software was employed for the fluorescence analysis.

A FM1-43 dye (5 μM) was used to assess the endo-/exocytosis of SVs. FM1-43 binds reversibly to the presynaptic membrane and is loaded into the newly formed SVs during endocytosis. Fluorescent spots appear as NEs are loaded with the dye, indicating that FM1-43 containing SVs are clustered in the AZ regions [16, 17]. To assess the endocytosis of SVs, FM1-43 was kept present during the stimulation and for 10 min afterwards to ensure that endocytosis caused by exocytosis stimulation had ended by that time. Next, the preparation was washed for 40 min in saline containing the ADVASEP-7 reagent (5 μM), which promotes the dissociation of FM1-43 from surface membranes [18]. The SVs that are formed during endocytosis and retain the FM1-43 dye start losing the dye in a new round of exocytosis. To assess the dynamics of the exocytosis, the preparations preloaded with FM1-43 (20-Hz stimulation for 3 min) were re-stimulated (at 20 Hz frequency, 10–20 min) and the decrease in the fluorescence intensity due to dye unloading was analyzed [19]. The properties of the FM1-43 marker are independent of pH in a pH range of 5–9 [20]; 0.4 μM 24-HC also does not affect the fluorescence of FM1-43 [21, 22].

The fluorescence of FM1-43 was detected using an excitation filter (480/10 nm), a dichroic mirror (505 nm), and an emission filter (535/40 nm). Fluorescence was evaluated as the average pixel intensity in the region of interest after subtracting the background fluorescence. When determining the rate of dye loss during unloading, the initial fluorescence of NE before the stimulation had started was assumed to be equal to 1.0.

The BCECF-AM ratiometric fluorescent probe (Molecular Probes, USA) was used as an indicator of cytoplasmic pH. The muscle specimens were incubated with the dye at a concentration of 5 μM for 15 min and then perfused for 30 min to reduce the background fluorescence. The dye-loaded synaptic contacts were exposed to intermittent light (1 s, 505/10 and 450/10 nm). Fluorescence was detected in the synaptic region using a 530-nm broadband emission filter. The I^{505}/I^{450} fluorescence ratio upon excitation by two wavelengths was used to estimate the intracellular pH. The decreased I^{505}/I^{450} ratio indicates that the intracellular pH had dropped. At the end of each experiment, the muscle specimens were perfused with phosphate buffer (138 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4) containing 10 μM nigericin to equalize the extra- and intracellular pH. The I^{505}/I^{450} ratio was evaluated when the preparation was exposed to buffer with different pH values (7.4–7.1) for calibration [5].

Statistical analysis

The results are presented as a mean ± standard error; n is the number of independent experiments performed on individual animals (indicated in the figure legends). The Mann–Whitney U test was used to compare two independent samples. Differences at $P < 0.05$ were considered statistically significant.

RESULTS

Monitoring of intracellular acidification

Anions of weak acids are known to acidify the intracellular environment. Propionate is widely used to mimic intracellular acidosis [23]. The undissociated form of propionic acid enters the cytoplasm and dissociates, thus decreasing the pH_{in}. Indeed, measurements of cytoplasmic pH using BCECF showed that application of propionate reduced the I^{505}/I^{450} ratio. This was an indication that pH in the synaptic zone of the mouse and frog preparations decreased (*Fig. 1A, B*). After 40 min of exposure to propionate, the I^{505}/I^{450} ratio fell to its steady-state level, being ~ 60–65% of the baseline (by ~0.25 pH units). This is comparable to the previously estimated change in pH_{in} in rat synapses in the presence of propionate [5]. In the control, the I^{505}/I^{450} ratio remained unchanged for 40 min (*Fig. 1A, B*), indicating that pH_{in} is stable at rest. Stimulation (at a frequency of 20 Hz) transiently reduced the I^{505}/I^{450} ratio in the synaptic region (*Fig. 1A, B*). This is consistent with the concept of intracellular acidification of NEs taking place during synaptic activity; once there remains no activity, pH_{in} is slowly restored [5–7].

Dynamics of neurotransmitter release and the effect of intracellular acidification

Long-term synaptic activity is maintained due to SV mobilization from the recycling and reserve pools to the AZ, followed by subsequent neurotransmitter release [1, 3]. In the control, the quantum content of EPPs during stimulation of the mouse phrenic nerve by electric pulses at a frequency of 20 Hz dropped rapidly during the first 5–10 pulses, to 20–25% of the baseline (155 ± 20 quanta). The quantum content was then stabilized and slowly decreased down to 10–15% of the baseline by the 3rd min of stimulation. Once the stimulation had been completed, the quantum content of EPPs was quickly restored to 50% of the baseline (6 ± 2 s, *Fig. 2A*). These changes in EPPs in response to 20-Hz stimulation and the rapid recovery of secretion are consistent with the view that neurotransmitter release in mouse motor NEs during 20-Hz stimulation

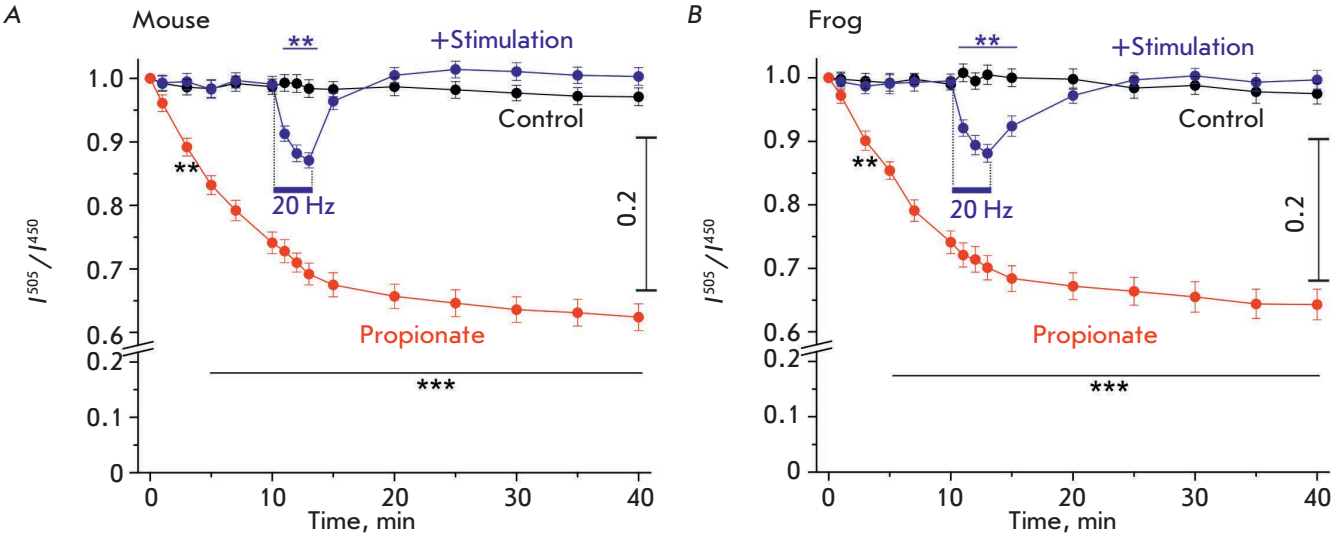


Fig. 1. Monitoring of cytosolic pH in the synaptic regions. The ratio between BCECF dye fluorescence intensities upon 505 and 450 nm excitation (I^{505}/I^{450}) is an indicator of pH and decreases in response to pH_{in} reduction. (*A*), (*B*) – the measurements of the I^{505}/I^{450} ratio in the synaptic regions of mouse (*A*) and frog (*B*) preparations at rest (control), exposed to 20-Hz stimulation for 3 min (shown in blue), and in the presence of propionate (72 μM; shown in red). Y axis: the I^{505}/I^{450} ratio at the initial instant is set at 1.0; $n = 7$ for each curve. The right scale shows the decrease in the I^{505}/I^{450} ratio in response to pH drop by 0.2 pH units. Data are presented as a mean ± SEM. $^{**}P < 0.01$, $^{***}P < 0.001$ is the statistical significance of the differences between the curves

is first maintained by the vesicles constituting the readily releasable pool and then by the vesicles of the recycling pool [2, 3]. SVs of the recycling pool are rapidly recovered by endocytosis and then re-used in the neurotransmitter release.

Application of propionate caused no statistically significant changes in the quantum content of the first EPP (128 ± 17 quanta, $P < 0.05$). However, it significantly accelerated the rundown of EPPs in the neuromuscular junctions of mice (*Fig. 2A*). The quantum content eventually decreased to 3–5% of the baseline by the 3rd min of 20-Hz stimulation. Recovery of the quantum content after the stimulation had ended was slower (up to 50% from the baseline within 13 ± 3 s) than that in the control. In order to quantify the neurotransmitter release, the cumulative curves were plotted by summing up the quantum contents of each EPP during 20-Hz stimulation for 3 min and the total number of quanta released from NEs was determined. A total of (90 ± 3.9) × 10^3 quanta were released in the control during the 3-min stimulation. This value was significantly lower in the presence of propionate ($P < 0.01$): (51 ± 2.8) × 10^3 quanta (*Fig. 2B*). Therefore, intracellular acidification of NEs in mice significantly inhibits neurotransmission during high synaptic activity by weakening the release of the neurotransmitter from the recycling pool of SVs.

Stimulation of the motor nerve of the frog cutaneous pectoris muscle at a frequency of 20 Hz was first accompanied by a reduction (during 30–40 stimuli) in the quantum content of EPPs to approximately 80% of the baseline (272 ± 30 quanta). The quantum content was stabilized by 3–5 s (the plateau) and then gradually decreased to 10–15% of the baseline value by the 3rd min of stimulation (*Fig. 2C*). It took 18 ± 3 s for the quantum content after the 20-Hz stimulation had completed to recover, reaching 50% of the baseline. These dynamics indicate that not only the readily releasable and the recycling pools are involved in neurotransmitter release, but the reserve pool as well [2, 3, 24].

The quantum content of the first EPP in the frog cutaneous pectoris muscle exposed to propionate had no difference compared to the control and was equal to 227 ± 35 quanta. The initial EPP inhibition was more pronounced, while the plateau phase was longer. The recovery of the quantum content after the 20-Hz stimulation was slower than that in the control (up to 50% within 21 ± 3 s) (*Fig. 2C*). By comparing the cumulative curves of the quantum contents of EPPs, one can see that neurotransmitter release decreases in the presence of propionate, being pronounced during the first minute of the 20-Hz stimulation (*Fig. 2D*). By the 3rd min of stimulation, the neurotransmitter release reached control values and amounted to (347 ± 13) × 10^3 quanta (vs. (355 ± 17) × 10^3 quanta in the control). Thus, intracellular acidification inhibits neurotransmitter release in frog NEs during the period when secretion is mediated by the recycling pool SVs.

Endo- and exocytosis during intracellular acidification

Endocytosis. Considering that acidification in NEs inhibited the neurotransmitter release, which was dependent on the recycling pool of SVs, endocytosis dysfunction is quite possible. In order to test this hypothesis, we evaluated the endocytic uptake of FM1-43 in NEs. SV endocytosis follows exocytosis and is carried out at a 1 : 1 ratio. Therefore, we selected the duration of 20-Hz stimulation, at which the same neurotransmitter release level was observed in both the control and experimental series (and, therefore, the same number of SVs underwent exocytosis). An analysis of the cumulative curves of neurotransmitter release showed that approximately the same number of neurotransmitter quanta was released during a 70-s stimulation in the control and a 100-s stimulation in the presence of propionate (*Fig. 2B, D*). If endocytosis is not disrupted, a similar level of FM1-43 loading can be expected under the chosen conditions. Indeed, the fluorescence intensities of mouse and frog NEs in the presence of sodium propionate did not significantly differ from those in the control (*Fig. 3A*).

The exocytosis kinetics. The dynamics of SV exocytosis during long-term 20-Hz stimulation was assessed by measuring the unloading of the FM1-43 dye from NEs. In the synapses of the control mice, fluorescence gradually decreased to 25–30% within 10 min of stimulation and then changed slowly (*Fig. 3B*). FM1-43 unloading was significantly impeded upon cytoplasmic acidification, and the fluorescence intensity decreased only to ~70% of the baseline by the 10th min of stimulation (*Fig. 3B*). Therefore, propionate inhibits the involvement in exocytosis of SVs, which maintain neurotransmission in mice at 20-Hz stimulation.

In control, the decrease in the fluorescence intensity occurred in two phases in the frog NEs: a rapid drop during the first 2 min (up to 70% of the baseline), followed by a slower decrease (*Fig. 3C*). By the 20th min of stimulation, the fluorescence intensity had dropped to 25–30%. FM1-43 unloading was inhibited upon intracellular acidification. The rate of FM1-43 unloading decreased most significantly within the first 2 min of stimulation (the fluorescence intensity declined only to ~95% of the baseline). By the 20th min of the 20-Hz stimulation, the fluorescence intensity had decreased to ~70% of the baseline (*Fig. 3C*). Therefore, propionate markedly inhibits the involvement in exocytosis of the recycling pool SVs, which maintain neurotransmission during the first several minutes of the 20-Hz stimulation of the motor nerve in frog NEs.

Intracellular acidification and the effect of 24-hydroxycholesterol on the changes in exocytosis

We have previously shown that 24-hydroxycholesterol can enhance SV mobilization in neuromuscular junctions upon 20-Hz stimulation [21]. Exposure to 24-hydroxycholesterol (0.4 μM) accelerated FM1-43 unloading upon the 20-Hz stimulation (*Fig. 4A, B*). A similar effect was observed in the NEs of mice and frogs. In the presence of propionate, 24-hydroxycholesterol completely lost its ability to accelerate the rate of FM1-43 release during exocytosis (*Fig. 4A, B*). Hence, intracellular acidification rendered acceleration of SV mobilization upon the 20-Hz stimulation impossible.

DISCUSSION

Numerous regulatory circuits acting on exocytosis, mobilization, and endocytosis of SVs establish the proper level of neurotransmitter release during synaptic activity. In the present study, we obtained data on the suppression of SV mobilization upon intracellular acidification for the first time. Furthermore, this phenomenon was observed in the NEs of both mice and frogs, thus indicating that the general mechanisms of intracellular acidosis action are identical.

Propionate efficiently reduced the intracellular pH by ~ 0.25 pH units, which is twofold higher than the degree of acidification caused by motor nerve stimulation with 20-Hz pulses. An analysis of postsynaptic responses showed that propionate did not significantly change the quantum content in response to the first stimulus, while accelerating the depression of neurotransmitter release in response to the 20-Hz stimulation. Under these conditions, the neurotransmitter release depends on the delivery of SVs to the AZ. In mouse synapses, the effect of propionate was clearly observable throughout the entire stimulation. On the contrary, the effect was observed only during the first minute of stimulation in frog synapses. These features of the effect of propionate are probably related to the specific involvement of SV pools in neurotransmission upon a 20-Hz stimulation. In particular, the recycling pool sustains long-term neurotransmitter release in mouse motor NEs upon a 20-Hz stimulation. Meanwhile, in frog motor NEs, this pool maintains the release mainly during the first minute of the stimulation, after which the reserve pool SVs become involved in neurotransmission. Hence, propionate seems to inhibit the involvement of the recycling pool SVs in the release. This selectivity of intracellular acidosis is consistent with the concept that there are independent pathways that regulate the recycling and reserve pools [15, 19, 25–27]. Moreover, the rate of propionate-mediated inhibition of neurotransmitter release in the frog NEs decreased after 60 s. As a result, the number of released transmit-

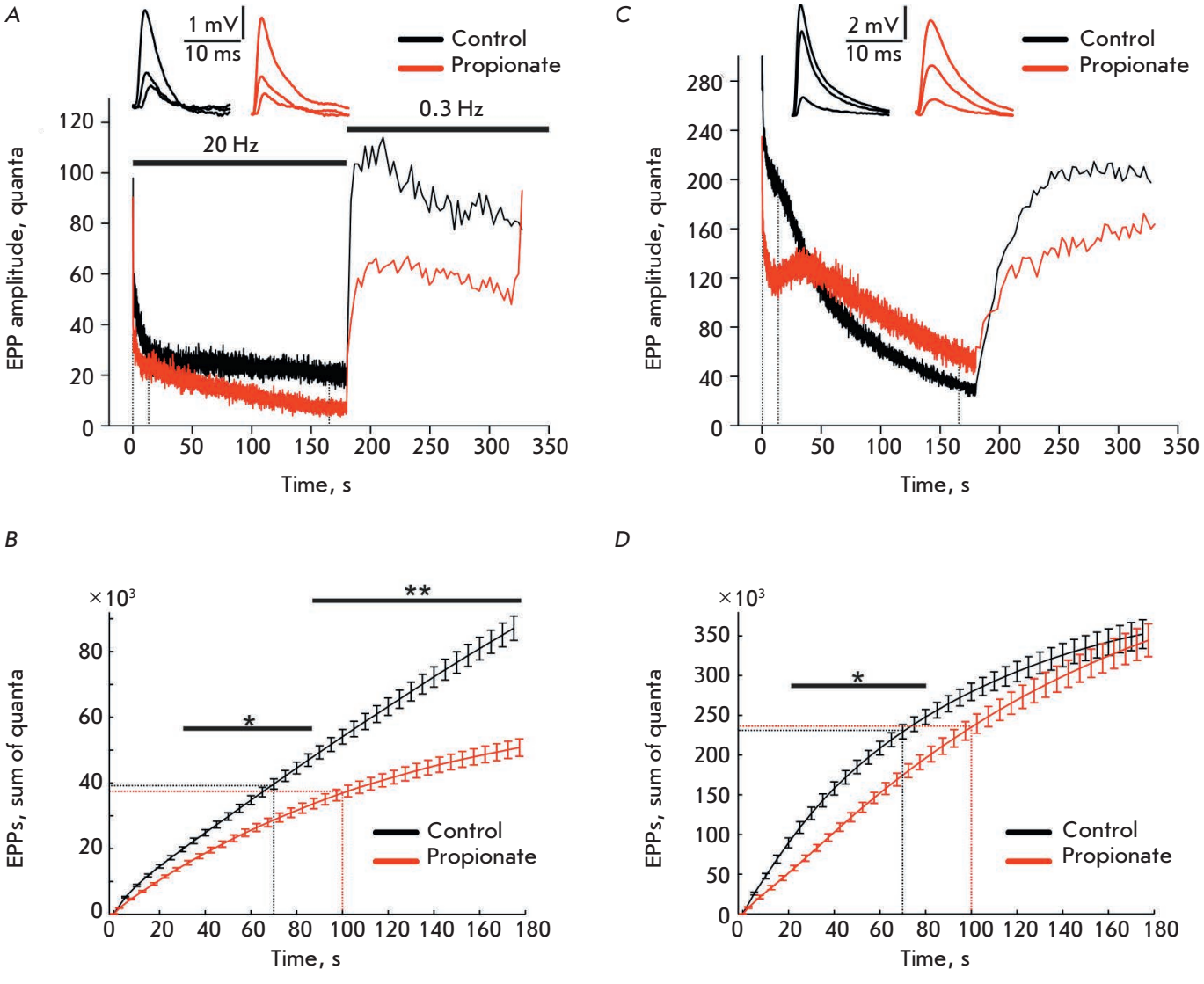


Fig. 2. The kinetics of neurotransmitter release at a 20-Hz stimulation. (*A*), (*C*) – Stimulation-induced changes in the quantum content of EPPs in the neuromuscular junctions in a mouse (*A*) and a frog (*C*) under control conditions and upon intracellular acidification with propionate. The recovery of the quantum content of EPPs after the 20-Hz stimulation is also shown. The native EPPs recorded at the moment of stimulation marked with dashed lines on the graph are shown at the top. Averaged curves are presented; $n = 5$. (*B*), (*D*) – The cumulative curves of the quantum content of EPPs in a 20-Hz stimulation in the neuromuscular junctions of a mouse (*B*) and a frog (*D*). Data are presented as a mean ± SEM. $^{*}P < 0.05$, $^{**}P < 0.01$ is the statistical significance of the differences between the curves; $n = 5$. The dashed lines denote the time points (70 and 100 s) at which the same number of quanta was released in the control and in the presence of propionate

ter quanta differed little from that in the control after stimulation for 3 min. Inhibition of the recruitment of SVs from the recycling pool seems to contribute to the release of the neurotransmitter from SVs belonging to the reserve pool.

The involvement of recycling pool SVs depends on both their mobilization to the sites of exocytosis and vesicle formation by endocytosis. Evaluation of the FM1-43 uptake showed that propionate does not disturb the endocytosis responsible for SV reformation after exocytosis. However, propionate markedly reduces the rate of FM1-43 dye release from SVs during a 20-Hz stimulation. This directly indicates that the delivery of SVs to the AZ is inhibited. The rate of FM1-43 release was markedly low in frog NEs during the first minute of the 20-Hz stimulation. This fact is consistent

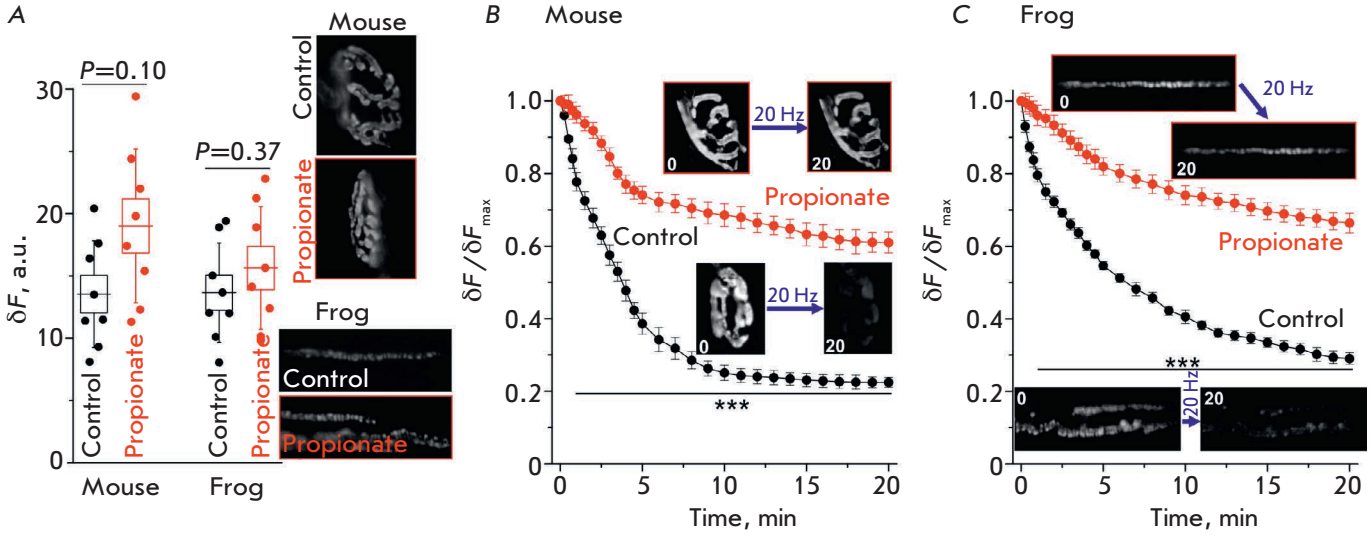


Fig. 3. Endocytosis and exocytosis in response to the 20-Hz stimulation. (*A*) – FM1-43 uptake by endocytosis in motor NEs in the control and in the presence of propionate. Boxes and whiskers show SEM and SD, respectively. Right, representative fluorescence images after the background was subtracted, $n = 8$ for each group. Y axis: fluorescence in arbitrary units (a.u.) after background subtraction. (*B*), (*C*) – FM1-43 unloading due to exocytosis during the 20-Hz stimulation in the control and in the presence of propionate in the neuromuscular junctions of a mouse (*B*) and a frog (*C*). $n = 8$ for each curve. The images illustrate a decrease in the NE fluorescence after 20 min of stimulation. Data are presented as a mean ± SEM. $^{***}P < 0.001$ – statistical significance of the differences between the curves. Y axis: normalized fluorescence, where 1.0 is the fluorescence before the onset of stimulation

with our assumption about an impaired mobilization of the SVs from the recycling pool upon intracellular acidosis.

The mechanisms regulating SV mobilization are organized hierarchically and in a coordinated manner. Cholesterol, its content in membranes, and its metabolites act as potent regulators of SV transport in both the CNS and neuromuscular junctions [22, 28–31]. Previously, we found that the main cholesterol metabolite in the brain (namely, 24-hydroxycholesterol), which is predominantly produced by neurons (including in synaptic regions), can enhance the involvement of the recycling pool SVs in neurotransmitter release in the mouse neuromuscular junctions [21]. The effect of the hydroxycholesterol depends on protein kinase G, which controls the function of the SV recycling pool in frog NEs [19]. 24-Hydroxycholesterol was found to accelerate the FM1-43 release during exocytosis in mouse and frog NEs, while propionate completely prevents its effect. Therefore, intracellular acidification may be the predominant factor precluding the increased neurotransmitter release in response to humoral stimuli.

The relationship between intracellular acidosis and SV mobilization possibly has a physiological and (or) pathological significance. A transient decrease in the pH in NEs [5–7] can suppress the mobilization of SVs in order to inhibit neurotransmitter release upon high activity. Thus, a negative feedback loop can form, limiting the release upon intense activity. Inhibition of SV delivery to the AZ can also provide sufficient time for endocytosis of SV to complete and, therefore, replenishment of the recycling pool. Decreased intracellular pH in neurons is observed in patients with a wide range of diseases (metabolic disorders, ischemia, epileptic activity, and neurodegenerative diseases) and in aging [32–34]. Excess glutamate and acetylcholine release cause damage to the central and neuromuscular synapses [35, 36]. Olesoxime, which can increase the influx of chloride anions (and, therefore, protons as well) into NEs and inhibit neurotransmitter release, exhibits a pronounced neuroprotective effect [31]. Similarly, the antiepileptic drug levetiracetam reduces pH in neocortical neurons, thereby contributing to the anticonvulsant properties of the drug [37]. A slight intracellular acidification may also underlie the anticonvulsant effect of short-chain monocarboxylates and ketone bodies [23]. Systemic ketoacidosis has the potential to affect motor performance by inhibiting neuromuscular transmission at the level of SV mobilization.

The molecular mechanism underlying the effect of intracellular pH on SV translocation to the AZ is unknown. This mechanism can be associated with the changes in mitochondrial function [38], calcium signaling [39], and protein-protein interactions [40]. The fact that the mobilization stage of SVs from the recycling pool is highly sensitive to pH changes suggests that

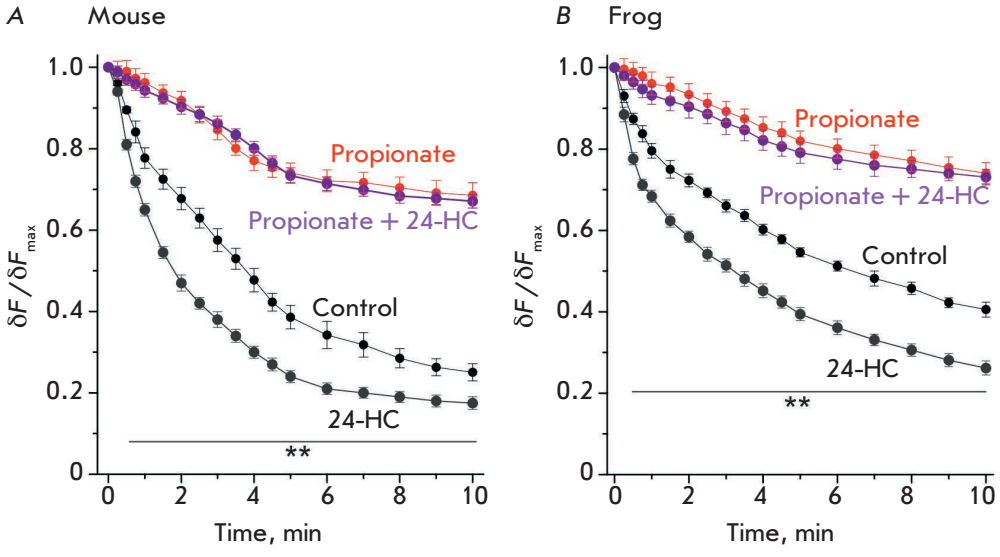


Fig. 4. The influence of cytoplasmic acidification on the effect of 24-hydroxycholesterol (24-HC) on exocytosis in a 20-Hz stimulation. (*A*), (*B*) – changes in the kinetics of FM1-43 unloading due to the administration of 24-HC in the control and in the presence of propionate in the NEs of a mouse (*A*) and a frog (*B*). The control and propionate curves (from *Fig. 3B, C*) are also shown. Data are presented as a mean ± SEM. $^{**}P < 0.01$ – statistical significance of the differences between the control and the effect of 24-HC. Y axis: normalized fluorescence, where 1.0 is the fluorescence prior to the onset of stimulation

there is a pH sensor. Probably, mobilization of SVs into the recycling pool along actin filaments using motor proteins is suppressed as pH in NEs decreases. So, local interactions between F-actin and myosin strongly depend on pH [41].

CONCLUSIONS

The pH of the cytoplasm of presynaptic nerve terminals decreases in response to increased synaptic activity under physiological conditions. A more pronounced drop in the pH can occur in disease. In the present study, we obtained data for the first time indicating that intracellular acidification can suppress neurotransmission by inhibiting the synaptic vesicle mobilization in the active zone upon high activity. This phenomenon can be part of a complex mechanism that regulates neurotransmission based on the acid-base processes in neurons. ●

This study was supported in part by the Russian Foundation for Basic Research (grants # 20-015-00507 and 20-04-00077) and by the state assignment of the Federal Research Center "Kazan Scientific Center of Russian Academy of Sciences".

REFERENCES

1. Dittman J.S., Ryan T.A. // Nat. Rev. Neurosci. 2019. V. 20. № 3. P. 177–186.
2. Zefirov A.L., Zakharov A.V., Mukhamedzianov R.D., Petrov A.M. // Zh. Evol. Biokhim. Fiziol. 2008. V. 44. № 6. P. 603–612.
3. Zakharov A.V., Petrov A.M., Kotov N.V., Zefirov A.L. // Biophysics. 2012. V. 57. № 4. P. 508–518.
4. Rizzoli S.O. // EMBO J. 2014. V. 33. № 8. P. 788–822.
5. Petrov A.M., Naumenko N.V., Uzinskaya K.V., Giniatullin A.R., Urazaev A.K., Zefirov A.L. // Neuroscience. 2011. V. 186. P. 1–12.
6. Zhang Z., Nguyen K.T., Barrett E.F., David G. // Neuron. 2010. V. 68. № 6. P. 1097–1108.
7. Rossano A.J., Kato A., Minard K.I., Romero M.F., Macleod G.T. // J. Physiol. 2017. V. 595. № 3. P. 805–824.
8. Sandvig K., Olsnes S., Petersen O.W., van Deurs B. // J. Cell Biochem. 1988. V. 36. № 1. P. 73–81.
9. Brown C.M., Petersen N.O. // Biochem. Cell Biol. 1999. V. 77. № 5. P. 439–448.
10. Dejonghe W., Kuenen S., Mylle E., Vasileva M., Keech O., Viotti C., Swerts J., Fendrych M., Ortiz-Morea F.A., Mishev K., et al. // Nat. Commun. 2016. V. 7. P. 11710.
11. Watanabe S., Boucrot E. // Curr. Opin. Cell Biol. 2017. V. 47. P. 64–71.
12. Bertone N.I., Groisman A.I., Mazzone G.L., Cano R., Tabares L., Uchitel O.D. // Synapse. 2017. V. 71. № 12. P. e22009.
13. Zakharov A.V. // Uchenye Zapiski Kazanskogo Universiteta. Seriya Estestvennye Nauki 2019. V. 161. № 2. P. 245–254.
14. Kasimov M.R., Giniatullin A.R., Zefirov A.L., Petrov A.M. // Biochim. Biophys. Acta. 2015. V. 1851. № 5. P. 674–685.

15. Kasimov M.R., Zakyrjanova G.F., Giniatullin A.R., Zefirov A.L., Petrov A.M. // Biochim. Biophys. Acta. 2016. V. 1861. № 7. P. 606–616.
16. Zefirov A.L., Grigor'ev P.N., Petrov A.M., Minlebaev M.G., Sitdikova G.F. // Tsitologiia. 2003. V. 45. № 12. P. 1163–1171.
17. Betz W.J., Bewick G.S. // J. Physiol. 1993. V. 460. № 1. P. 287–309.
18. Kay A.R., Alfonso A., Alford S., Cline H.T., Holgado A.M., Sakmann B., Snitsarev V.A., Stricker T.P., Takahashi M., Wu L.G. // Neuron. 1999. V. 24. № 4. P. 809–817.
19. Petrov A.M., Giniatullin A.R., Sitdikova G.F., Zefirov A.L. // J. Neurosci. 2008. V. 28. № 49. P. 13216–13222.
20. Betz W.J., Mao F., Smith C.B. // Curr. Opin. Neurobiol. 1996. V. 6. № 3. P. 365–371.
21. Kasimov M.R., Fatkhrakhmanova M.R., Mukhutdinova K.A., Petrov A.M. // Neuropharmacology. 2017. V. 117. P. 61–73.
22. Mukhutdinova K.A., Kasimov M.R., Zakyrjanova G.F., Gumerova M.R., Petrov A.M. // Neuropharmacology. 2019. V. 150. P. 70–79.
23. Bonnet U., Bingmann D., Speckmann E.J., Wiemann M. // Life Sci. 2018. V. 204. P. 65–70.
24. Mukhametshina A.R., Fedorenko S.V., Petrov A.M., Zakyrjanova G.F., Petrov K.A., Nurullin L.F., Nizameev I.R., Mustafina A.R., Sinyashin O.G. // ACS Appl. Mater. Interfaces. 2018. V. 10. № 17. P. 14948–14955.
25. Petrov A.M., Giniatullin A.R., Zefirov A.L. // Neurochem. J. 2008. V. 2. № 3. P. 175–182.
26. Petrov A.M., Kasimov M.R., Giniatullin A.R., Zefirov A.L. // Neuroscience and Behavioral Physiology. 2014. V. 44. № 9. P. 1020–1030.
27. Marra V., Burden J.J., Thorpe J.R., Smith I.T., Smith S.L., Hausser M., Branco T., Staras K. // Neuron. 2012. V. 76. № 3. P. 579–589.
28. Teixeira G., Vieira L.B., Gomez M.V., Guatimosim C. // Neurochem. Int. 2012. V. 61. № 7. P. 1151–1159.
29. Wasser C.R., Ertunc M., Liu X., Kavalali E.T. // J. Physiol. 2007. V. 579. Pt 2. P. 413–429.
30. Krivoi I.I., Petrov A.M. // Int. J. Mol. Sci. 2019. V. 20. № 5. P. 1046.
31. Zakyrjanova G.F., Gilmutdinov A.I., Tsentsevitsky A.N., Petrov A.M. // Biochim. Biophys. Acta. Mol. Cell. Biol. Lipids. 2020. V. 1865. № 9. P. 158739.
32. Xiong Z.Q., Saggau P., Stringer J.L. // J. Neurosci. 2000. V. 20. № 4. P. 1290–1296.
33. Bonnet U., Bingmann D., Speckmann E.J., Wiemann M. // J. Neural. Transm. (Vienna). 2018. V. 125. № 10. P. 1495–1501.
34. Majdi A., Mahmoudi J., Sadigh-Eteghad S., Golzari S.E., Sabermarouf B., Reyhani-Rad S. // J. Neurosci. Res. 2016. V. 94. № 10. P. 879–887.
35. Meldrum B.S. // Neurology. 1994. V. 44. № 11 Suppl 8. P. S14–23.
36. Sugita S., Fleming L.L., Wood C., Vaughan S.K., Gomes M.P., Camargo W., Naves L.A., Prado V.F., Prado M.A., Guatimosim C., Valdez G. // Skelet. Muscle. 2016. V. 6. P. 31.
37. Bonnet U., Bingmann D., Speckmann E.J., Wiemann M. // Brain Res. 2019. V. 1710. P. 146–156.
38. Pekun T.G., Lemeshchenko V.V., Lyskova T.I., Waseem T.V., Fedorovich S.V. // J. Mol. Neurosci. 2013. V. 49. № 1. P. 211–222.
39. Trudeau L.E., Parpura V., Haydon P.G. // J. Neurophysiol. 1999. V. 81. № 6. P. 2627–2635.
40. Sinning A., Hubner C.A. // FEBS Lett. 2013. V. 587. № 13. P. 1923–1928.
41. Kohler S., Schmoller K.M., Crevenna A.H., Bausch A.R. // Cell Rep. 2012. V. 2. № 3. P. 433–439.

Drosophila Zinc Finger Protein CG9890 Is Colocalized with Chromatin Modifying and Remodeling Complexes on Gene Promoters and Involved in Transcription Regulation

N. A. Fursova, M. Y. Mazina, J. V. Nikolenko, N. E. Vorobyova, A. N. Krasnov*
*E-mail: krasnov@genebiology.ru
Institute of Gene Biology Russian Academy of Sciences, Moscow, 119334 Russia
Received May 22, 2020; in final form, June 19, 2020
DOI: 10.32607/actanaturae.11056
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **In this work, we conducted a genome-wide study of the zinc finger protein CG9890 and showed that it is localized mostly on the promoters of active genes. The CG9890 binding sites are low-nucleosome-density regions and are colocalized with the chromatin modifying and remodeling complexes SAGA and dSWI/SNF, as well as with the ORC replication complex. The CG9890 protein was shown to be involved in the regulation of the expression of some genes on the promoters of which it is located, with the ecdysone cascade genes accounting for a significant percentage of these genes. Thus, the CG9890 protein is a new member of the transcriptional network which is localized on active promoters, interacts with the main transcription and replication complexes, and is involved in the regulation of both basal and inducible transcription.**
KEYWORDS **ENY2, CG9890, Drosophila, zinc fingers, ChIP-Seq.**
ABBREVIATIONS **ENY2 – enhancer of yellow 2; C2H2 – zinc fingers of C2H2 type; SAGA – histone acetyltransferase complex; SWI/SNF – chromatin remodeler; AMEX – mRNA export complex; ORC – origin recognition complex.**

INTRODUCTION

Previously, our laboratory isolated and characterized the ENY2 protein that was found to be a component of many of the protein complexes involved in the regulation of transcription and replication. ENY2 is involved in the SAGA, AMEX, and THO transcriptional complexes and connects various stages of gene expression: transcriptional domain organization and chromatin modification, transcription activation and elongation, export of mRNA, and regulation of spatial gene arrangement in the nucleus [1–7]. Also, ENY2 was found to be involved in the ORC replication complex responsible for positioning the replication origin [8–11].

An analysis of the ENY2-Su(Hw) two-hybrid interaction revealed that Su(Hw) recruits the ENY2 protein to the Su(Hw)-dependent insulators of Drosophila, which is necessary for barrier function [5]. Then, Su(Hw) was shown to recruit the histone-acetyltransferase complex SAGA (containing ENY2) [12] and the chromatin remodeling complex dSWI/SNF [13–15] on Su(Hw)-dependent insulators, causing the formation of a low-nucleosome-density region and creating the conditions for the binding of the ORC replication complex. Knockdown of Su(Hw) almost completely disrupts the recruitment of the SAGA, dSWI/SNF, and ORC complexes to Su(Hw)-dependent insulators and significantly increases the nucleosome density on these regulatory elements [1, 2]. Su(Hw) was shown to be the first example of a protein responsible for positioning the replication origin. Su(Hw) is required for the formation of 6% of the replication origins in the Drosophila genome; therefore, some other, not yet identified, proteins are responsible for the formation of the remaining 94% origins.

Previously, we discovered that there is an interaction between ENY2 and another protein, CG9890, that contains a C2H2-type zinc finger domain, just like Su(Hw) [16]. We reckon that, like Su(Hw), CG9890 is a DNA-binding protein that recruits ENY2-containing complexes to their binding sites, thereby organizing the regulatory genome elements necessary for cell functioning. The CG9890 protein was shown to be localized in the cell nucleus. Biochemical studies revealed an interaction between the CG9890 protein and

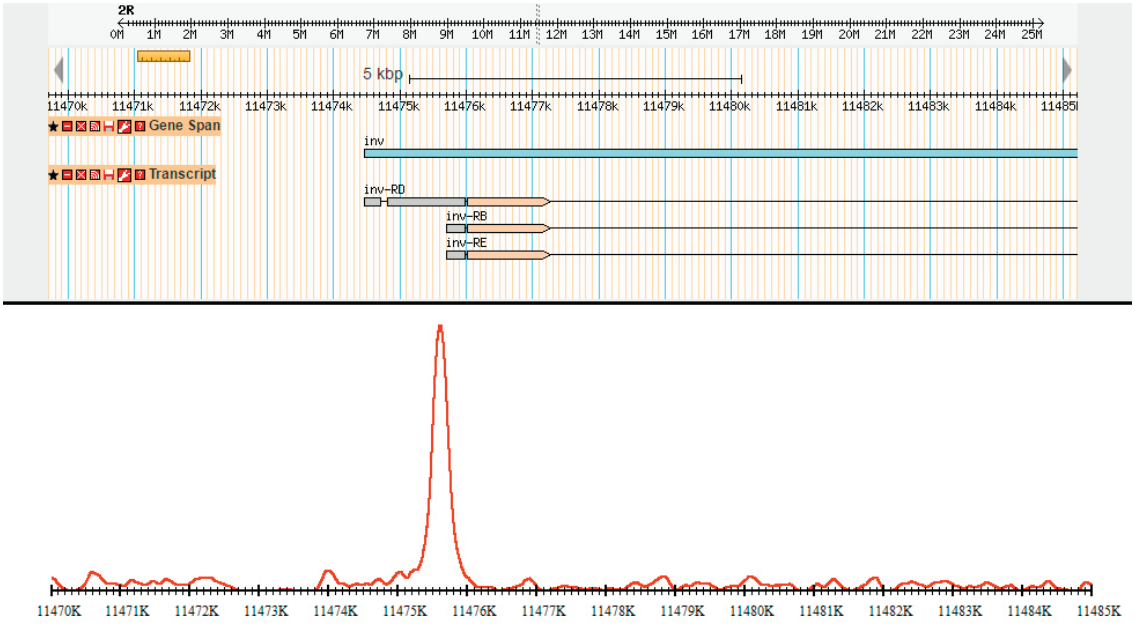


Fig. 1. Typical ChIP-Seq profile of the CG9890 protein. The Figure shows a genomic region corresponding to the *Inv* gene promoter. Information about this region from the genome browser is shown on the top panel; the ChIP-Seq profile is shown on the bottom panel

the ENY2-containing complexes SAGA, ORC, dSWI/SNF, TFIID, and THOC [16]. CG9890 interacts with the transcriptional complexes involved in transcription initiation and elongation but does not interact with the AMEX complex involved in the export of mRNA from the nucleus to the cytoplasm, which indicates activity of CG9890 at the first stages of the transcription cycle.

In this study, we performed a genome-wide analysis of the CG9890 protein to identify and characterize the regulatory elements for which CG9890 may be responsible.

EXPERIMENTAL

Antibodies and cell lines

In this study, we used the *Drosophila melanogaster* S2 cell line. α-CG9890 polyclonal antibodies were derived from the blood serum of a rabbit immunized with the full-length protein CG9890 expressed in *Escherichia coli* cells [16].

Chromatin immunoprecipitation and whole genome sequencing

Chromatin immunoprecipitation was performed according to [1]. For this, we used the CG9890 polyclonal antibodies produced previously [16]. ChIP-Seq libraries were prepared using a NEBNext DNA library preparation kit (New England Biolabs). The quality of the libraries was checked using a Bioanalyzer. For high-throughput sequencing, 200–500 bp fragments were used. The libraries were sequenced on an Illumina HiSeq 2000 genomic sequencer. The produced sequences were mapped to the Drosophila reference genome using the Bowtie2 software. Only uniquely mapped reads were used for further analysis. Identification of the peak coordinates and generation of a full genome profile (WIG file) for the CG9890 protein were performed using the SPP software (FDR < 5%) [17]. A genomic interval of +/−100 bp from the peak position was considered the peak region.

Bioinformatics analysis

D. melanogaster gene annotations were taken from the official FlyBase website. The genome was divided into the following regions: transcription start sites (TSSs), transcription end sites (TESs), transcribed regions (gene regions except for TSS and TES), and intergenic regions (the others). The ChIP-Seq peak was identified as belonging to one of these categories provided that genomic intervals overlapped at least 10 bp. During peak annotation, the following priority of genomic categories was used: TSS, TES, transcribed and intergenic regions.

RESULTS AND DISCUSSION

Protein CG9890 is localized mainly on gene promoters

To determine the localization of the studied protein in the genome, we performed chromatin immunoprecipitation from S2 cells using polyclonal antibodies to the CG9890 protein, followed by high-throughput sequencing (ChIP-Seq). A typical ChIP-Seq profile of the CG9890 protein at one of its binding sites is shown in *Fig. 1*. A total of 4,709 binding sites of the CG9890 protein were identified in the Drosophila genome (FDR < 5%).

We annotated the identified sites based on their localization in one of the following Drosophila genome elements: promoters, gene ends, gene bodies, and intergenic regions. According to the obtained data (*Fig. 2*), the largest number of ChIP-Seq peaks of the CG9890 protein (73.2%) is localized in the promoter regions of Drosophila genes. We reckon that, being localized predominantly on gene promoters, the CG9890 protein may participate in the functioning of regulatory genetic elements of this type.

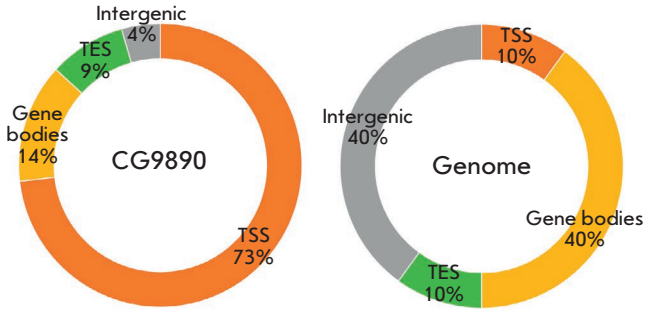


Fig. 2. Distribution of the CG9890 protein binding sites relative to the annotated elements of the Drosophila genome (left). For comparison, the relative representation of all annotated elements in the genome is shown (right). TSS – promoter region, TES – end of the gene, Gene bodies – gene region between TSS and TES, Intergenic – intergenic regions

The CG9890 protein is colocalized with chromatin modifying and remodeling complexes in low nucleosome density regions

Previously, we confirmed the interaction between CG9890 and the ENY2 protein and revealed the interaction between CG9890 and the ENY2-containing protein complexes SAGA, ORC, dSWI/SNF, TFIID, and THO. Therefore, we studied genomic colocalization of the CG9890 protein with the SAGA and ORC complexes, as well as with the dSWI/SNF remodeling complex that, together with the SAGA complex, participates in the formation of the chromatin structure required for the correct functioning of regulatory elements, including promoters. For this purpose, we used software of our own design for generating an averaged profile of the investigated factor at specified genomic sites [1]. The genomic profiles of the ORC2, GCN5, and OSA proteins and histone H3 were previously obtained in our laboratory. Averaged profiles of these proteins were calculated at all 4,709 binding sites of the CG9890 protein, as well as at 4,709 random promoters and 4,709 random genomic sites (*Fig. 3*).

Because the CG9890 protein is located predominantly on gene promoters, there may be enrichment

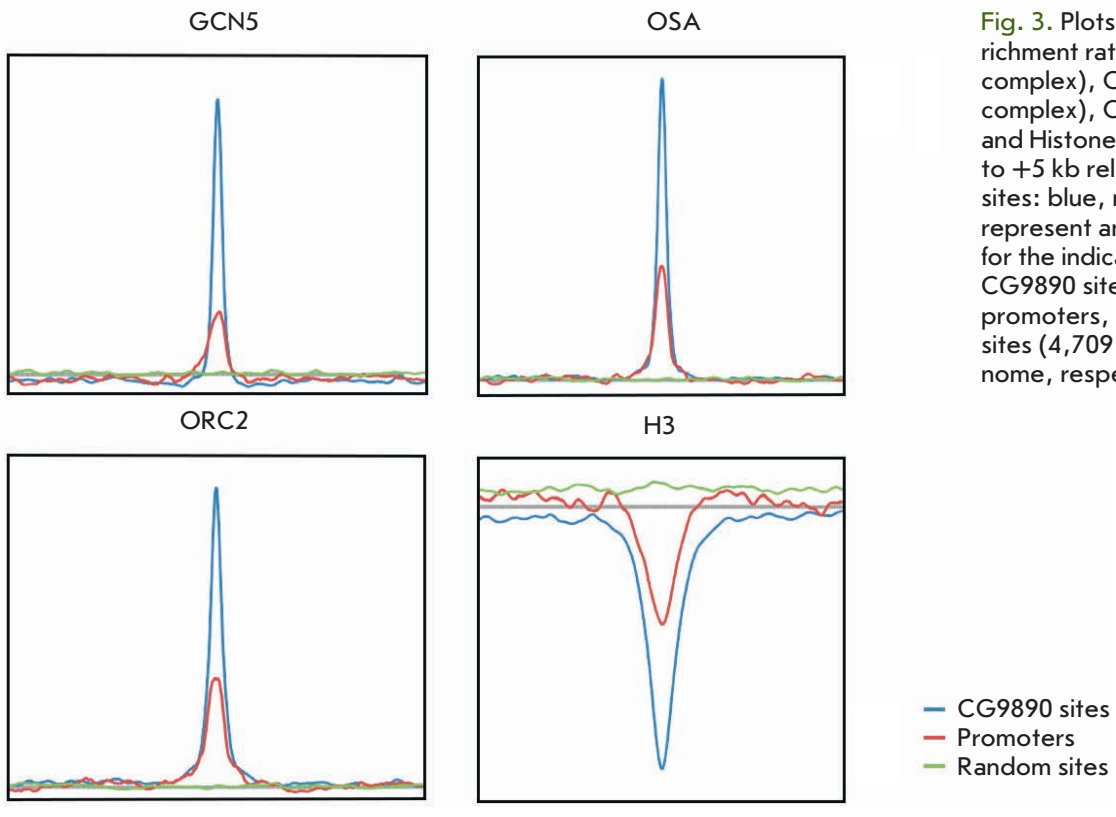


Fig. 3. Plots of averaged log2 enrichment ratios for GCN5 (SAGA complex), OSA (dSWI/SNF complex), ORC2 (ORC complex), and Histone H3 at positions –5 to +5 kb relative to the following sites: blue, red, and green plots represent an averaged profile for the indicated factors on the CG9890 sites, randomly selected promoters, and random genome sites (4,709 sites each) in the genome, respectively

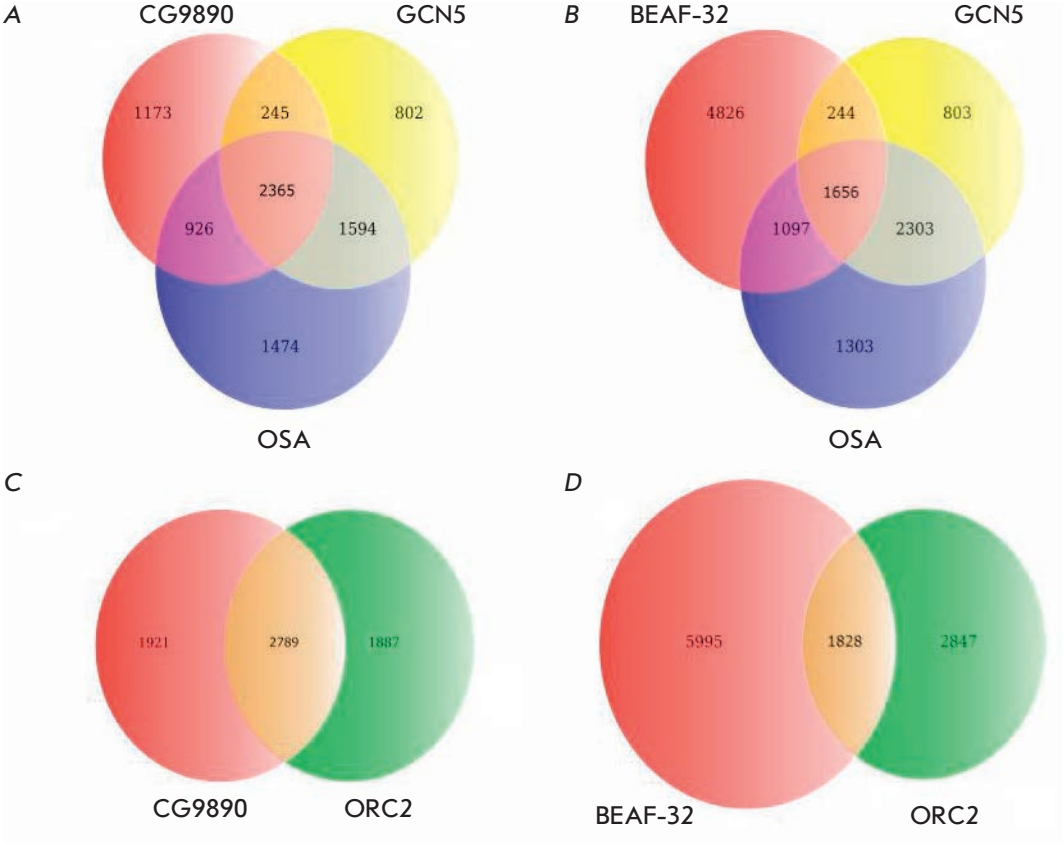


Fig. 4. Euler-Venn diagrams showing overlap of the binding sites of the CG9890 and BEAF-32 proteins with the binding sites of GCN5 (SAGA complex), OSA (dSWI/SNF complex), and ORC2 (ORC complex). (*A*) CG9890, GCN5, OSA; (*B*) BEAF-32, GCN5, OSA; (*C*) CG9890 and ORC2; (*D*) BEAF-32 and ORC2

of any promoter factors, including the SAGA, dSWI/SNF, and ORC complexes, at the CG9890 binding sites relative to the average genome level. However, as seen from *Fig. 3*, the GCN5 (SAGA complex), OSA (dSWI/SNF complex), and ORC2 (ORC complex) proteins are enriched at the binding sites of the CG9890 protein not only in comparison with the average genome level, but also in comparison with random promoters. This result indicates that this colocalization is associated not just with random coincidence on gene promoters, but with the fact that it is the CG9890 protein binding site that promotes the localization of the SAGA, dSWI/SNF, and ORC complexes. In addition, as follows from *Fig. 3*, the binding sites of the CG9890 protein are characterized by a lower nucleosome density (enrichment of histone H3) than the genome average and on the promoters, which indicates an active state of these regulatory elements.

Using the second approach, we calculated the number of CG9890 protein sites overlapping with the GCN5, OSA, and ORC2 protein sites. The well-known protein BEAF-32 was chosen as a control factor [18]. The coordinates of the ChIP-Seq peaks for the BEAF-32 protein were obtained from NCBI GEO (GSE35648). The peaks of two proteins were considered to overlap if their genomic intervals overlapped by at least 10 bp. The obtained data are shown in *Fig. 4*.

As seen from *Fig. 4*, about 60% of the CG9890 protein sites overlap with the sites of the ORC2 protein, a subunit of the ORC complex, which, in turn, accounts for about 60% of the ORC2 protein sites. The level of overlapping of the ORC2 sites with the sites of BEAF-32, another factor localized on the promoters, is significantly lower despite the fact that the number of BEAF-32 protein binding sites in the genome is much higher. *Figure 4* shows that the CG9890 protein is colocalized with the GCN5 and OSA proteins at half of the CG9890 binding sites in the genome, which is significantly higher than an analogous value for the control BEAF-32 protein.

The CG9890 protein is involved in the regulation of gene expression

Previously, we demonstrated that the CG9890 protein interacts with the ENY2 protein that coordinates many steps in the regulation of gene expression. The interaction between the CG9890 protein and the ENY2-containing complexes SAGA, ORC, dSWI/SNF, TFIID, and THOC, i.e. the complexes involved in the initiation and elongation of transcription, was

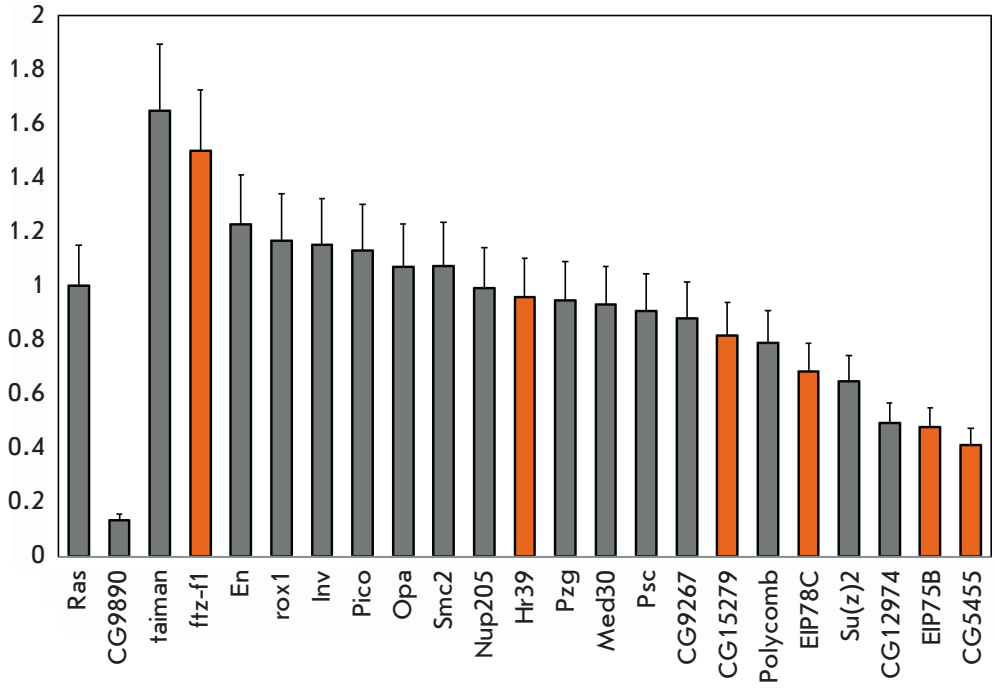


Fig. 5. Changes in the expression levels of CG9890-associated genes after RNA interference of the CG9890 protein. The vertical axis shows a change in the mRNA level for the indicated genes after RNA interference relative to the initial level. The *Ras* gene was used for normalization. Error bars correspond to a standard error of the mean. Orange bars correspond to ecdysone cascade genes

revealed [16]. Given that CG9890 was found predominantly on gene promoters, we decided to investigate, by RNA interference, what changes in the expression of CG9890-associated genes would result from a decrease in the intracellular level of the CG9890 protein. By optimizing the conditions for RNA interference, we achieved an effective decrease in the expression of the studied protein in cells by more than 5 times in terms of the mRNA amount and almost complete depletion of the protein (below the Western blotting detection limit).

By using RT-qPCR, we analyzed the changes in the level of mRNA 21 of the CG9890-associated gene in the cells after RNA interference compared to the control samples. The results of this experiment are shown in *Fig. 5*. After knockdown of the CG9890 protein, the amount of mRNA of seven of these genes decreased by at least 20% and the amount of mRNA in three genes increased by at least 20%. Thus, the CG9890 protein is indeed involved in the regulation of the expression of at least some of the genes on whose promoters it is localized.

Among the 10 genes whose expression changed statistically significantly upon RNA interference of the CG9890 protein, five are ecdysone cascade genes. Their transcription is significantly activated during the response to ecdysone. This enables the use of a convenient model system for cell induction by ecdysone to study in detail the functioning of the CG9890 protein in the regulation of the expression of these genes. The advantage of this system is that it may be used to study the dynamic processes of inducible regulation of gene expression, but not simply to maintain basal transcription [19].

CONCLUSION

In previous studies, we found that the zinc finger insulator protein Su(Hw) interacts with ENY2 and recruits ENY2-containing complexes to the Su(Hw)-dependent insulators of Drosophila, participating simultaneously in the regulation of transcription and the positioning of replication origins. We also established an interaction between ENY2 and another protein, CG9890, containing a zinc finger domain, like Su(Hw). Biochemical studies revealed an interaction between the CG9890 protein and the ENY2-containing complexes SAGA, ORC, dSWI/SNF, TFIID, and THOC. We suggest that, like Su(Hw), the CG9890 protein is a DNA-binding protein that recruits ENY2-containing complexes to their binding sites, thereby organizing the genome regulatory elements necessary for cell functioning. In this study, we identified the CG9890 binding sites in the genome and showed that they are located mainly on gene promoters. We found a genome-wide correlation between CG9890 binding sites and the ENY2-containing complexes SAGA, ORC, and dSWI/SNF. The CG9890 protein binding sites are characterized by a lower nucleosome density (enrichment of histone H3) than the genome and promoter averages, which indicates an active state of these regulatory elements. The CG9890 protein is involved in the regulation of the expression of some genes on the promoters of which it occurs, with

the ecdysone cascade genes accounting for a significant percentage of these genes. Thus, the CG9890 protein is a new member of the cell transcriptional network which is localized on active promoters, interacts with the main transcription and replication complexes, and is involved in the regulation of both basal and inducible transcription. ●

This study was supported by the Russian Science Foundation (Grant No. 20-14-00269).

REFERENCES

1. 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. № 11. P. 5717–5730.
2. Mazina M., Vorob'eva N.E., Krasnov A.N. // Tsitologiia. 2013. V. 55. № 4. P. 218–224.
3. Vorobyeva N.E., Nikolenko J.V., Krasnov A.N., Kuzmina J.L., Panov V.V., Nabirochkina E.N., Georgieva S.G., Shidlovskii Y.V. // Cell Cycle. 2011. V. 10. № 11. P. 1821–1827.
4. Kopytova D.V., Krasnov A.N., Orlova A.V., Gurskiy D.Y., Nabirochkina E.N., Georgieva S.G., Shidlovskii Y.V. // Cell Cycle. 2010. V. 9. № 3. P. 479–481.
5. Kurshakova M., Maksimenko O., Golovnin A., Pulina M., Georgieva S., Georgiev P., Krasnov A. // Mol. Cell. 2007. V. 27. № 2. P. 332–338.
6. Krasnov A.N., Kurshakova M.M., Ramensky V.E., Mardanov P.V., Nabirochkina E.N., Georgieva S.G. // Nucl. Acids Res. 2005. V. 33. № 20. P. 6654–6661.
7. Gurskiy D., Orlova A., Vorobyeva N., Nabirochkina E., Krasnov A., Shidlovskii Y., Georgieva S., Kopytova D. // Nucl. Acids Res. 2012. V. 40. № 21. P. 10689–10700.
8. Kopytova D., Popova V., Kurshakova M., Shidlovskii Y., Nabirochkina E., Brechalov A., Georgiev G., Georgieva S. // Nucl. Acids Res. 2016. V. 44. № 10. P. 4920–4933.
9. Eaton M.L., Prinz J.A., MacAlpine H.K., Tretyakov G., Kharchenko P.V., MacAlpine D.M. // Genome Res. 2011. V. 21. № 2. P. 164–174.
10. Masai H., Matsumoto S., You Z., Yoshizawa-Sugata N., Oda M. // Annu. Rev. Biochem. 2010. V. 79. P. 89–130.
11. MacAlpine H.K., Gordan R., Powell S.K., Hartemink A.J., MacAlpine D.M. // Genome Res. 2010. V. 20. № 2. P. 201–211.
12. Baptista T., Grunberg S., Minoungou N., Koster M.J.E., Timmers H.T.M., Hahn S., Devys D., Tora L. // Mol. Cell. 2017. V. 68. № 1. P. 130–143.
13. Shi J., Zheng M., Ye Y., Li M., Chen X., Hu X., Sun J., Zhang X., Jiang C. // Nucl. Acids Res. 2014. V. 42. № 15. P. 9730–9739.
14. Harikrishnan K.N., Chow M.Z., Baker E.K., Pal S., Bassal S., Brasacchio D., Wang L., Craig J.M., Jones P.L., Sif S., et al. // Nat. Genet. 2005. V. 37. № 3. P. 254–264.
15. Elfring L.K., Daniel C., Papoulas O., Deuring R., Sarte M., Moseley S., Beek S.J., Waldrip W.R., Daubresse G., DePace A., et al. // Genetics. 1998. V. 148. № 1. P. 251–265.
16. Fursova N.A., Nikolenko J.V., Soshnikova N.V., Mazina M.Y., Vorobyova N.E., Krasnov A.N. // Acta Naturae. 2018. V. 10. № 4. P. 110–114.
17. Kharchenko P.V., Tolstorukov M.Y., Park P.J. // Nat. Biotechnol. 2008. V. 26. № 12. P. 1351–1359.
18. Zhao K., Hart C.M., Laemmli U.K. // Cell. 1995. V. 81. № 6. P. 879–889.
19. Mazina M.Y., Nikolenko J.V., Fursova N.A., Nedil'ko P.N., Krasnov A.N., Vorobyeva N.E. // Cell Cycle. 2015. V. 14. № 22. P. 3593–3601.

The Influence of an Elevated Production of Extracellular Enveloped Virions of the Vaccinia Virus on Its Properties in Infected Mice

S. N. Shchelkunov*, S. N. Yakubitskiy, T. V. Bauer, A. A. Sergeev, A. S. Kabanov, L. E. Bulichev, I. A. Yurganova, D. A. Odnoshevskiy, I. V. Kolosova, S. A. Pyankov, O. S. Taranov
State Research Center of Virology and Biotechnology VECTOR, Rospoterbnadzor, Novosibirsk region, Koltsovo, 630559 Russia
*E-mail: snshchel@rambler.ru; snshchel@vector.nsc.ru
DOI: 10.32607/actanaturae.10972
Copyright © 2020 National Research University Higher School of Economics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT **The modern approach to developing attenuated smallpox vaccines usually consists in targeted inactivation of vaccinia virus (VACV) virulence genes. In this work, we studied how an elevated production of extracellular enveloped virions (EEVs) and the route of mouse infection can influence the virulence and immunogenicity of VACV. The research subject was the LIVP strain, which is used in Russia for smallpox vaccination. Two point mutations causing an elevated production of EEVs compared with the parental LIVP strain were inserted into the sequence of the VACV *A34R* gene. The created mutant LIVP-A34R strain showed lower neurovirulence in an intracerebral injection test and elevated antibody production in the intradermal injection method. This VACV variant can be a promising platform for developing an attenuated, highly immunogenic vaccine against smallpox and other orthopoxvirus infections. It can also be used as a vector for designing live-attenuated recombinant polyvalent vaccines against various infectious diseases.**
KEYWORDS **smallpox, vaccine, immunogenicity, virulence.**
ABBREVIATIONS **pfu – plaque-forming unit; NS – normal saline; CEV – cell-associated enveloped virion; EEV – extracellular enveloped virion; IMV – intracellular mature virion; VACV – vaccinia virus.**

INTRODUCTION

The vaccinia virus (VACV) belongs to the genus *Orthopoxvirus* of the family Poxviridae. This genus includes animal viruses such as the variola virus (VARV), the monkeypox virus (MPXV), the cowpox virus (CPXV), and others [1, 2]. Orthopoxviruses are the largest complexly organized DNA-containing mammalian viruses; their entire life cycle takes place in the cytoplasm of infected cells. The members of this genus are morphologically indistinguishable and antigenically closely related to each other. Therefore, infection with one species of orthopoxvirus provides protective immunity against other members of its genus [3]. For this very reason, the use of a live attenuated vaccine based on different VACV strains has made it possible to eradicate smallpox [1, 4].

Like other species of orthopoxviruses, VACV exists in two infectious forms. The virus progeny mostly consists of intracellular mature virions (IMVs) and a much smaller number of extracellular enveloped virions (EEVs) [5, 6]. IMVs accumulate in large amounts in an infected cell and are released into the environment only after the cell is destroyed. A small percentage of synthesized viral particles get enveloped with an additional lipoprotein coating and are released on the cell surface at the early stage of the viral replication cycle, where they are associated with the cell (cell-associated enveloped virions, CEVs). Some of these particles detach from the cells and exist in their free form (EEVs) [7]. EEVs make up less than 1% of all progeny of most VACV strains [5]. Meanwhile, the efficiency of EEV penetration into the cell is higher than that for IMVs [7, 8]; so, the virus quickly disseminates throughout the organism [5, 9]. No detailed studies of the effect of an elevated EEV production on the immunogenicity of VACV have been performed yet.

The VACV strains can differ substantially in terms of their level of EEV production [6, 10]. The IHD-J (International Health Department-J) strain is the most

thoroughly studied variant of VACV that ensures a high yield of EEVs in the infected cell culture [6]. The *A34R* gene is one of the genes that regulate the release of CEVs to free EEVs [10]. Protein A34, contained in the lipoprotein envelope of EEVs, is not found in IMVs. The amino acid sequence of protein A34 of the neurovirulent mouse-adapted VACV Western Reserve (WR) strain (< 1% EEVs among infectious virus progeny in the cell culture) differs from the amino acid sequence of this protein for the IHD-J strain (up to 30% EEVs) by only two point substitutions: Asp110 → Asn and Lys151 → Glu [10]. It was shown that replacement of the *A34R* gene in the VACV WR strain with the gene from the IHD-J strain significantly increases the yield of EEVs [9, 10].

It has been proved experimentally that the elevated EEV production caused by the insertion of mutations into the *A34R* gene leads to a more efficient dissemination of oncolytic variants of VACV and improves the *in vivo* antitumor activity of these viruses [9, 11]. However, the effect of these mutations on the virulence and immunogenicity of VACV has not been studied.

The cessation of smallpox vaccination after 1980 [2–4] has led to a situation where the contemporary human population is unprotected against the re-emerging orthopoxvirus infections [12]. Therefore, research focused on the development of novel, attenuated and highly immunogenic VACV-based vaccines becomes especially important [4].

VACV is extensively used not only to produce safe, next-generation live-attenuated vaccines against human orthopoxvirus infections, but also as a molecular vector in designing live recombinant polyvalent vaccines against various infectious diseases [3, 6, 12–14]. An important direction in research is the study of the effect of different viral genes and their mutant variants on the immunogenicity and safety of the vaccines being developed.

The objective of this study was to produce a LIVP VACV strain carrying mutations in the *A34R* gene that result in an elevated production of EEVs and to investigate the virulent and immunogenic properties of the LIVP-A34R variant compared to the parent LIVP strain when mice are infected via different routes.

EXPERIMENTAL

Virus, cell cultures

In this study, we used the clonal variant 14 of the VACV LIVP strain (earlier described in [15]) and the African green monkey kidney cell cultures CV-1 and Vero from the cell culture collection of the State Research Center of Virology and Biotechnology (SRC VB) VECTOR, Rospotrebnadzor (Russian Federal Service for Surveillance, Consumer Rights Protection and Human Welfare). The viruses were grown and titrated on the CV-1 cell culture according to the procedure described in [16]. The Vero cell culture was used for the virus neutralization test conducted in the serum of mice.

Production of VACV with point mutations in the *A34R* gene

Two point mutations were inserted into the nucleotide sequence of the *A34R* gene by PCR with synthetic oligonucleotide primers; these mutations caused the synthesis of the protein corresponding to protein A34 of the IHD-J VACV strain (Asp110 → Asn and Lys151 → Glu substitutions) [17]. The recombinant LIVP-A34R strain carrying the mutant *A34R* gene was produced on the basis of the clonal variant 14 of the LIVP strain, using plasmid pMGCgpt-A34R* according to the procedure described previously [15].

Animals

Inbred BALB/c mice, both males and females, procured from the husbandry farm of SRC VB VECTOR, Rospotrebnadzor, were used in this study. The animals were fed a standard diet with a sufficient amount of water according to veterinary laws and regulations and in compliance with the National Research Council Guidelines on Laboratory Animal Care and Use [18]. All the manipulations on the animals were approved by the Bioethics Committee of the SRC VB VECTOR, Rospotrebnadzor.

Assessment of the neurovirulence of VACV strains

Suckling (2- to 3-day old) mice were challenged intracerebrally with the recombinant LIVP-A34R strain or the parent LIVP clonal variant diluted in normal saline (NS) at a dose of 10 pfu/10 μL/mouse. The animals in the control group received an identical volume of NS. The mice were followed up for 12 days; the number of animals that died was counted.

Infecting mice

The 3- to 5-week old BALB/c mice weighing 13–16 g were used. The animals were challenged with preparations of LIVP and LIVP-A34R viruses or normal saline intranasally (i.n.), subcutaneously (s.c.) or intradermally (i.d.) according to [16]. Infectious doses of 10^8, 10^7 or 10^6 pfu/30 μL/animal were used. Each group consisted of 5–6 experimental animals. The mice were weighed daily, and external clinical signs of the disease (adynamia, tremor, and ruffled hair coat) were documented during 14 days.

Collecting blood samples from the experimental animals

Blood samples were collected from the retro-orbital venous sinus using sterile disposable capillaries on 28 dpi; then, the mice were euthanized by cervical dislocation. Serum was isolated from mouse blood by precipitating blood cells via centrifugation. Individual mouse serum samples were stored at 20°C.

Enzyme-linked immunosorbent assay (ELISA)

ELISA of individual mouse blood serum samples was performed according to [16]. A purified VACV LIVP preparation was used as an antigen. The geometric means of log reciprocal titer of VACV-specific IgG in experimental groups were calculated; the confidence intervals for a 95% matching between each sample and the total population were determined.

Measuring the serum titers of virus neutralizing antibodies

The titers of antibodies against VACV LIVP in mouse serum samples were quantified using the plaque reduction neutralization test (PRNT), according to the decrease in virus plaque count in a monolayer Vero cell culture, as described in [19]. Prior to performing PRNT, serum samples were inactivated at 56°C for 30 min. Four- to fivefold dilution series of serum samples, starting from a 1 : 10 dilution, in the cell maintenance medium were prepared. The dilution where 0.1 mL of the cell culture contained 30–60 pfu was used as the working dilution of VACV. The diluted serum samples and VACV solutions were mixed in equal volumes and incubated at 37.0 ± 0.5°C for 1 h. This mixture (0.2 mL) was placed onto the Vero cell monolayer in 24-well plates; 0.8 mL of the cell maintenance medium was added to each well, and the cells were cultured for 3 days in a CO_2 incubator. After culturing, the monolayer was stained with a gentian violet solution and the plaque number in the wells was counted.

Pathomorphological and virological analyses of the organs

The mouse organs (lungs, brain, liver, kidneys, and spleen) and tissue samples (nasal septum or skin samples from the injection site) were collected from mice euthanized by cervical dislocation 3, 7, and 10 days post inoculation (dpi), with viral preparations or a normal saline solution. At each time point, organ and tissue samples from three animals were collected and analyzed individually.

To perform a postmortem analysis, mouse organs were fixed in a 4% paraformaldehyde solution (Sigma, USA) for 48 h. The samples were treated using the conventional procedure: sequential dehydration in alcohol solutions in increasing concentrations, impregnation in the xylene–paraffin mixture, and embedding into paraffin. Paraffin-embedded sections 4–5 μm thick were prepared on a HM-360 automated rotary microtome (Germany). The sections were stained with hematoxylin and eosin. Optical microscopy studies and photomicrography were carried out on an AxioImager Z1 microscope (Carl Zeiss, Germany) using the AxioVision 4.8.2 software package (Carl Zeiss, Germany).

To perform the virological analysis, 10% homogenates of mouse organs and tissues were prepared by mechanical disintegration on a stainless-steel ball homogenizer, with a DMEM medium added subsequently. After several freeze–thaw cycles, the viral titers in the homogenates were determined on the CV-1 cell culture monolayer by viral plaque assay [15].

RESULTS

Production of the EEVs by LIVP and LIVP-A34R VACV strains

Since VACV EEVs are released from the cell before the primarily infected cell is lysed and all the infectious viral forms get into the extracellular space, we conducted experiments where the viral titers in the infected cells and the extracellular fluid were quantified depending on the time post-infection. The CV-1 cell monolayer in a six-well plate was inoculated with LIVP or mutant LIVP-A34R produced from it with a multiplicity of 1 pfu/cell. Aliquots of the extracellular fluid were collected every 3 h for 1 day, and the cells contained in the growth medium were subjected to two freeze–thaw cycles. Viral titer in the samples was determined by viral plaque assay. Three replicates were recorded for each sampling point.

The results of these experiments (*Fig. 1*) demonstrate that the mutant LIVP-A34R does not differ from the parent LIVP in terms of the level of synthesis of the IMVs (*Fig. 1A*), while the EEVs are produced in a statistically significantly greater amount compared to LIVP (*Fig. 1B*). The exact percentage of the IMV and EEV forms in the total viral yield was not determined.

The pathogenicity of VACV strains for different routes of inoculation of mice

To perform a comparative analysis of the effect of the route of inoculation and the dose of the administered viral preparation on the pathogenic properties of the LIVP and LIVP-A34R strains, mice were infected via three of the most popular routes (the closest to the natural ones): intranasally (i.n.), intradermally (i.d.), or subcutaneously (s.c.). The infective doses of each virus were 10^6, 10^7 or 10^8 pfu/animal. Since inoculation of

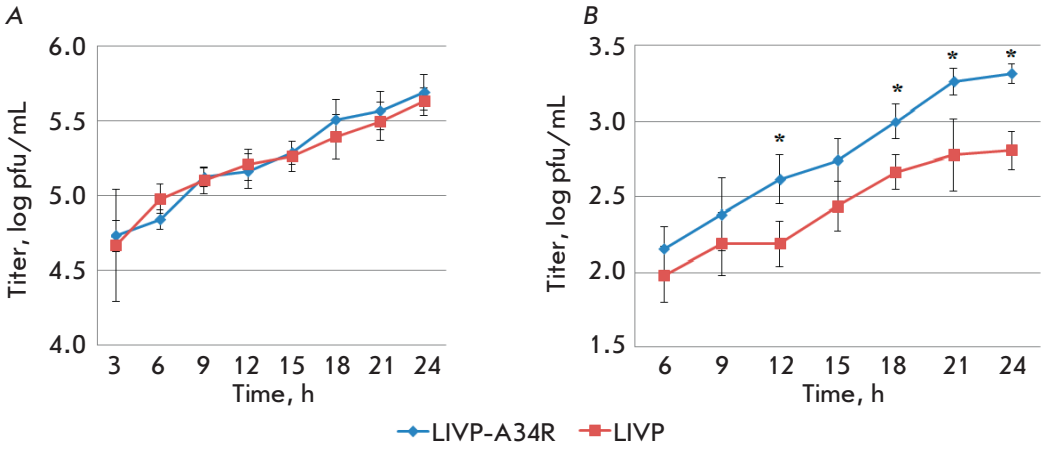


Fig. 1. The dynamics of replication of different variants of VACV in the CV-1 cell culture. (*A*) – biosynthesis of the intracellular IMV form; (*B*) – accumulation of the EEV form in the extracellular medium. * – differences are statistically significant at $p > 0.1$

adult mice with most of the VACV strains usually does not cause animal death, the pathogenicity of the variants of this virus are studied according to the changes in body weight after inoculation and the clinical manifestations of the disease (ruffled hair coat, adynamia, and tremor) [20, 21].

Marked clinical manifestations of infection and transient body weight loss were observed only after i.n. inoculation of mice with both VACV strains (*Fig. 2*). The peak of the disease occurred on 6–8 dpi. With increasing virus dose, clinical manifestations of infection became more pronounced and the decline in mouse body weight was more significant. *Figure 2* demonstrates that the mutant LIVP-A34R strain at doses of 10^6 and 10^7 pfu was characterized by the lowest pathogenicity.

No clinical manifestations of infection were observed in mice i.d. inoculated with both VACV strains. The dynamics of the body weight of infected mice were almost the same as in those of the controls (*Fig. 3*). Identical results were also observed for s.c. inoculated mice (data not shown).

Virus dissemination in the mouse organism

The organ and tissue samples from animals i.n., i.d., or s.c. inoculated with the LIVP or LIVP-A34R strain at doses of 10^6, 10^7 or 10^8 pfu/animal collected on 3, 7, and 10 dpi were used to prepare 10% homogenates; viral titers were determined on the CV-1 cell culture monolayer by viral plaque assay. The lung, brain, liver, spleen, and kidney samples were analyzed. In i.n. inoculated mice, the nasal septum mucosa was additionally examined; skin flaps from the site of the virus inoculation were also analyzed in the animal groups that had received an i.d. or s.c. injection.

In i.n. inoculated mice, the viruses were detected in all examined organs: the highest titers were revealed in the nasal septum mucosa (the primary virus replication focus), and in decreasing order, in the lungs, brain, liver (*Fig. 4*), kidneys, and spleen.

In i.d. inoculated mice, the viruses were detected in skin samples from the injection site of the viral preparation; in animals that had been inoculated with a maximal infective dose (10^8 pfu), the viruses were also detected in the lungs and liver of some animals on 3 and 7 dpi (data not shown). More of the LIVP-A34R mutant variant was found in the lungs compared to the parent LIVP. No detectable viral titers were revealed in the brain, spleen, or kidney samples.

In animals s.c. inoculated with the analyzed VACV variants, the viruses were detected only in the skin flap samples collected from the site of injection of the viral suspensions at the maximal infective dose. No viruses were detected in the internal organ samples.

Pathomorphological analysis of mouse organs

In general, the pathological changes in the organs of experimental animal groups correspond to the histological pattern of the changes observed in the laboratory animals infected with orthopoxviruses [22], thus confirming the adequacy of the selected model. The severity and extension of the pathological changes varied depending on the virus strain, the infective dose, and the route of administration of the viral preparation.

The most typical pathomorphological manifestations of the infection were observed in the organs of the respiratory system, mostly in the lungs. The following manifestations were revealed in the respiratory tissue: profound swelling of the interalveolar septa, capillary hyperemia, and active release of blood cells and blood plasma into the alveolar space. In the most severe cases, exudation was accompanied by dystrophic and necrobiotic changes in the alveolar epithelium, fibrin accumulation, and mixed inflammatory cell infiltration (neutrophils, lymphocytes and a small amount of

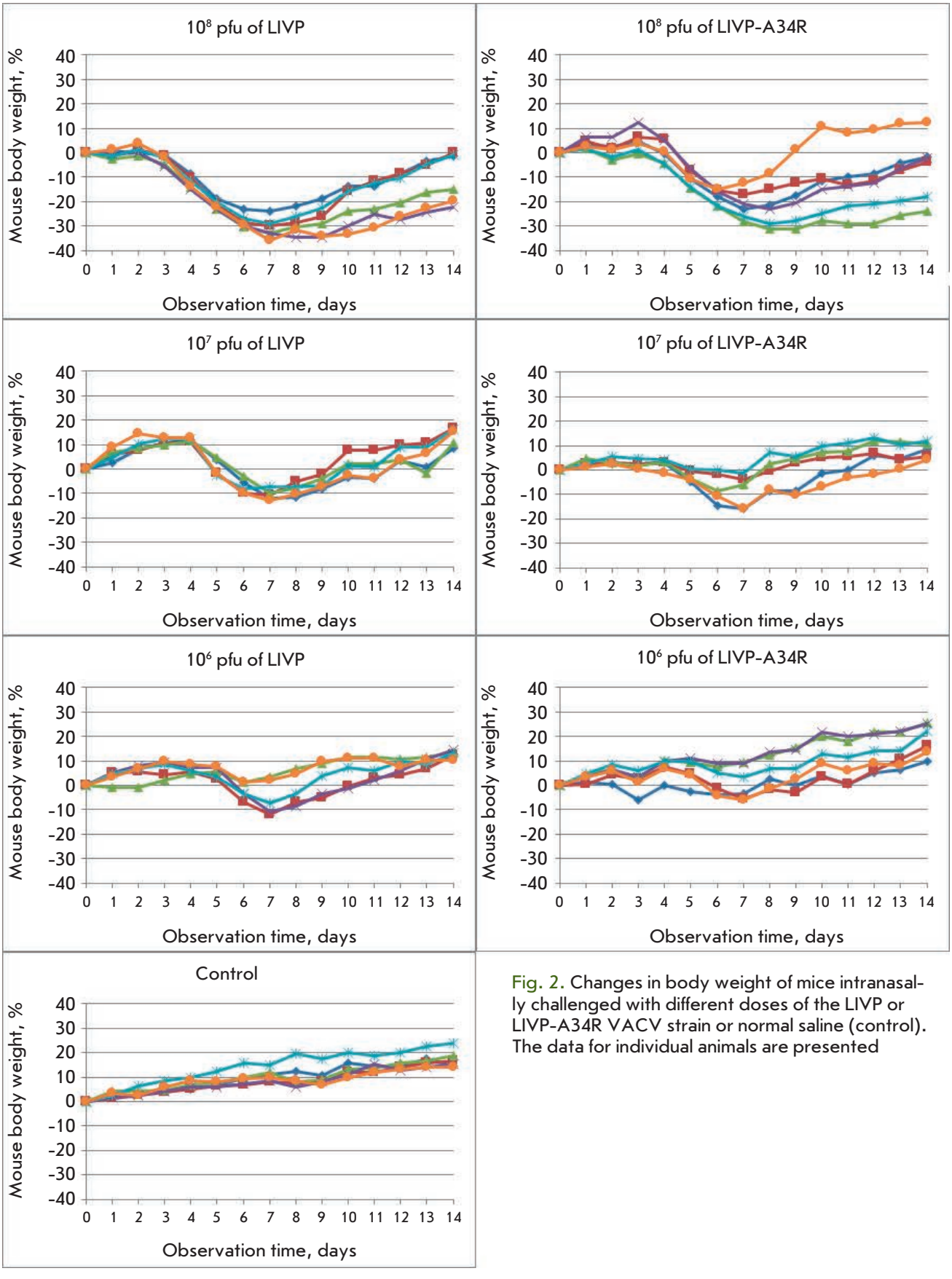


Fig. 2. Changes in body weight of mice intranasally challenged with different doses of the LIVP or LIVP-A34R VACV strain or normal saline (control). The data for individual animals are presented

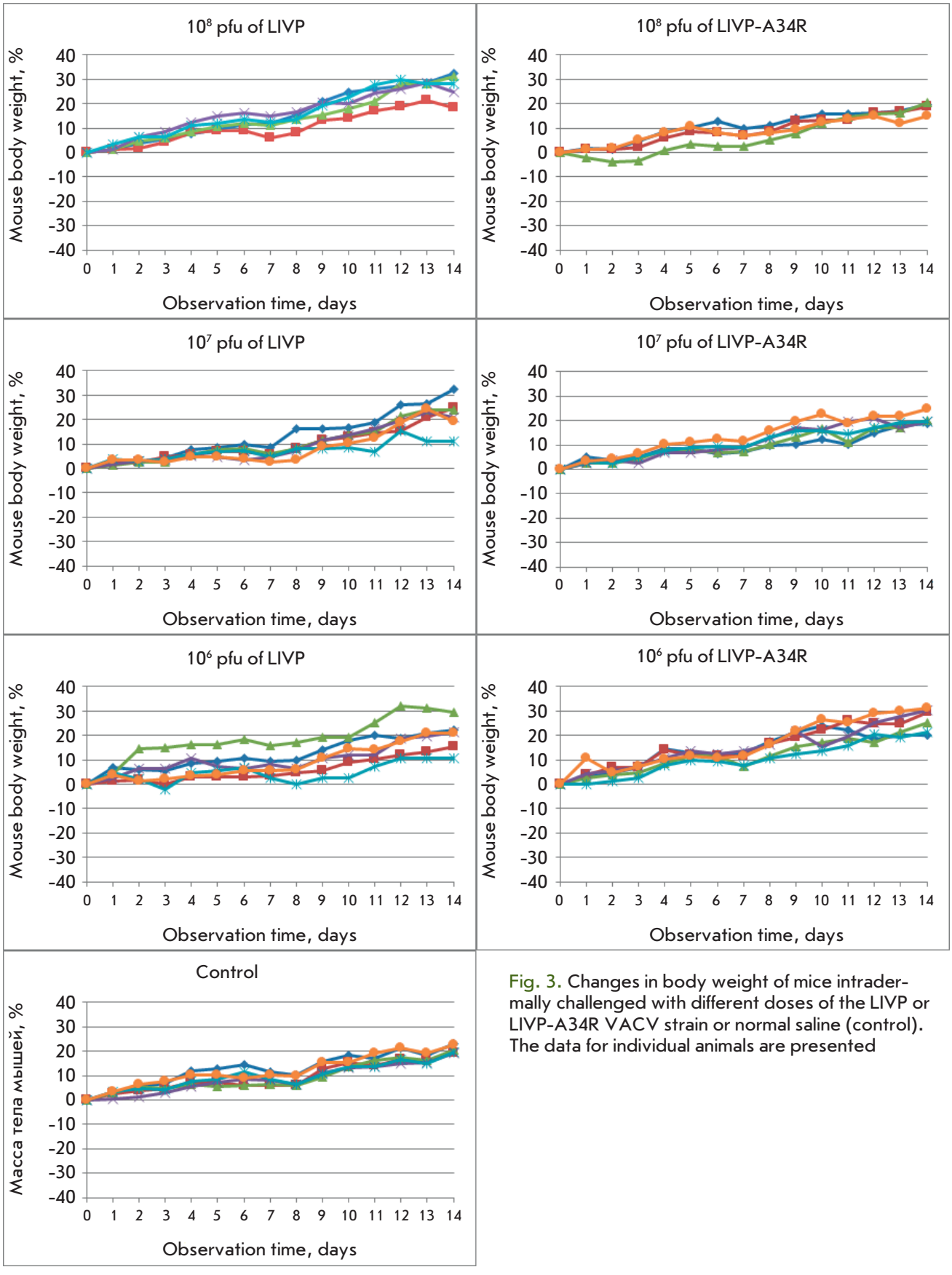


Fig. 3. Changes in body weight of mice intradermally challenged with different doses of the LIVP or LIVP-A34R VACV strain or normal saline (control). The data for individual animals are presented

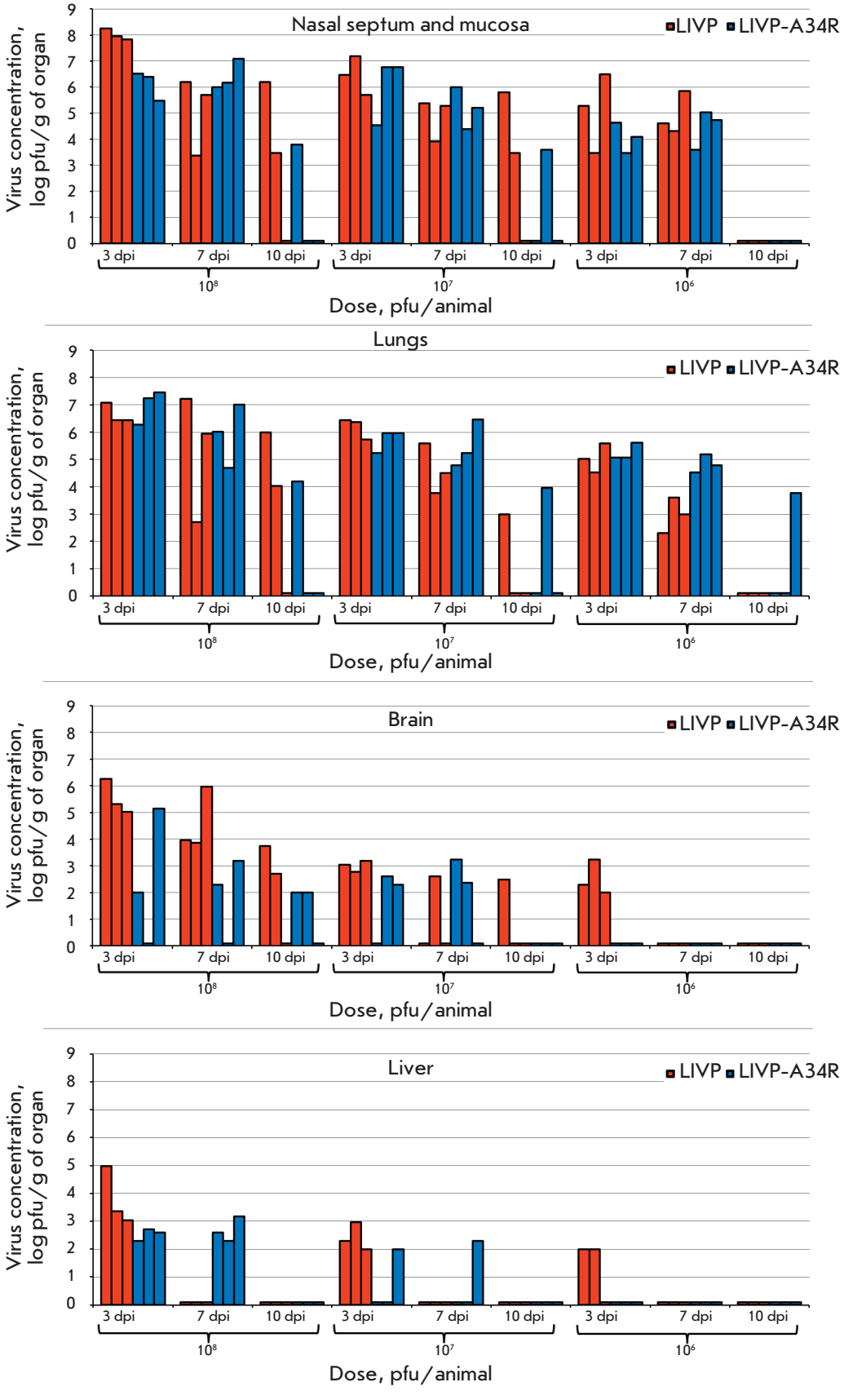


Fig. 4. Accumulation of the VACV in organs and tissues in mice intranasally infected with different doses of the LIVP or LIVP-A34R strain. The data for individual animals are presented. dpi – days post-infection

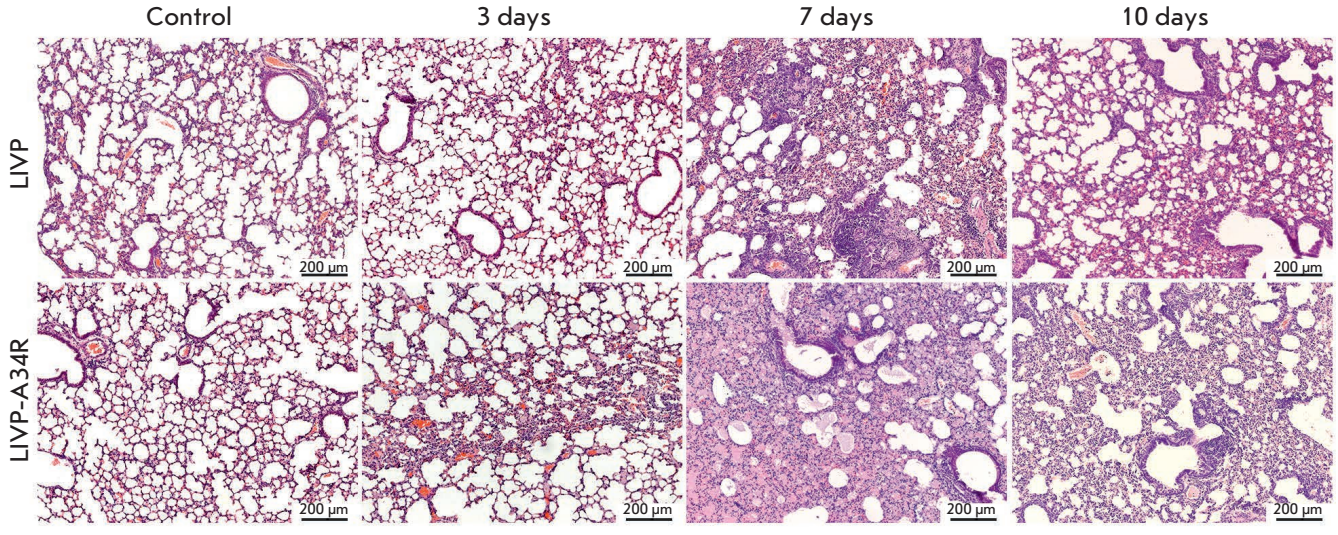


Fig. 5. The dynamics of development of the pathological process in respiratory tissue after intranasal infection at a dose of 10^7 pfu (see explanation in the text). Histological lung specimen. Hematoxylin & eosin staining

eosinophils). Macrophages were detected rarely, predominantly in mice i.n. inoculated with the virus at a dose of 10^8 pfu.

In mice i.n. inoculated with the LIVP-A34R strain, reduced lung tissue airness, mild swelling, hyperemia, moderate infiltration of the stroma with lymphocytes and neutrophils were observed for the 1/5–1/4 of the section area on 3 dpi. In mice infected with the LIVP strain, changes in the lungs were minimal on 3 dpi (*Fig. 5*). The most pronounced pathomorphological signs of the disease were observed on 7 dpi: advanced severe swelling of interalveolar septa, polymorphic cellular infiltration, acute hyperemia and thrombosis of the microcirculatory vessels, and necrotic foci in the connective tissue surrounding major bronchi and blood vessels. Plasmorrhagia and fibrin exudation in the alveoli were more pronounced when mice were infected with the LIVP-A34R strain. The pathological manifestations were moderate on 10 dpi.

The pathological changes in the trachea and bronchi were mild, mostly manifesting themselves as sparse loci of epithelial dystrophy, thickening of the walls of small bronchi, moderate swelling of the intercellular spaces, and rarely, as epithelial desquamation, accompanied by the development of an erosive surface. Moderate peribronchial and perivascular polymorphic cellular infiltration was observed rather rarely and only in i.n. inoculated mice. The bronchi remained almost uninvolved in the pathological process in i.d. or s.c. inoculated mice.

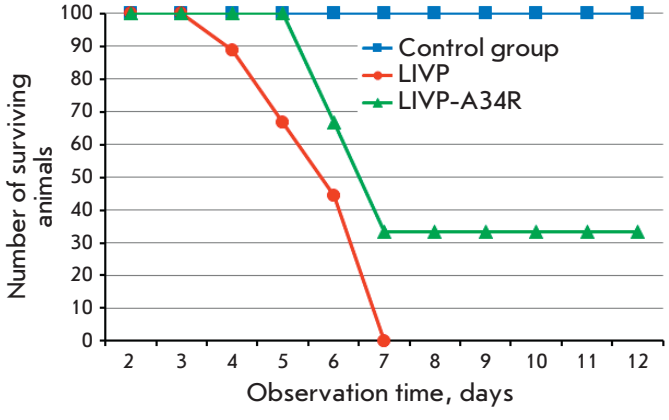


Fig. 6. The death rate dynamics in newborn mice after intracerebral infection with the LIVP or LIVP-A34R VACV strain

Neurovirulence of the VACV variants

The ability of the viruses to cause death upon intracerebral inoculation was studied in three groups of newborn mice (10 animals per group), which were followed during 12 dpi. In the group of animals inoculated with LIVP VACV at a dose of 10 pfu/mouse, the animals started to die on 4 dpi; all of them had died by 7 dpi. In the group of mice inoculated with the same dose of the mutant LIVP-A34R variant, the animals started to die on 6 dpi; after 7 pdi, no animal death was observed and a third of them survived (*Fig. 6*). No animal deaths

were observed in the control group (mice that received an injection of NS).

Immunogenicity of VACV strains

The immunogenicity of the LIVP and LIVP-A34R VACV variants was assessed according to the titers of the virus-specific (ELISA) and virus-neutralizing antibodies (based on a reduction of the infectivity of the VACV preparation) induced by them in the mouse serum samples collected on 28 dpi via three different routes with different virus doses (10^6, 10^7 or 10^8 pfu/mouse).

A purified LIVP VACV preparation (the IMV form of the virus) was used as an antigen in ELISA tests. The results of ELISA (*Fig. 7*) demonstrate that, in i.n. inoculated mice, high titers of antibodies against the virion proteins of the IMV of VACV were detected for both the high (10^8 pfu) and lower infective doses. No statistically significant differences between the LIVP and LIVP-A34R strains were revealed for this parameter.

The lowest titers of VACV-specific antibodies were detected in mice s.c. inoculated with the viruses. Antigen production was strongly dependent on the virus dose. The mutant LIVP-A34R strain ensured a higher production of virus-specific antibodies compared to the parent LIVP VACV strain (*Fig. 7*).

The titer of VACV-specific antibodies in mice i.d. inoculated with the virus at an infective dose of 10^8 pfu was comparable to the titer of antibodies elicited by i.n. inoculation with the viruses at the same dose. When the dose of i.d. inoculated virus decreased, biosynthesis of virus-specific antibodies fell more noticeably compared to the i.n. route of inoculation (*Fig. 7*). Meanwhile, the mean antibody titers were higher in the serum samples collected from mice inoculated with LIVP-A34R.

The findings obtained by analyzing the titer of virus neutralizing antibodies were similar to the ELISA data, but there were some differences (*Fig. 8*). The highest level of production of virus-neutralizing antibodies was detected in mice i.n. inoculated with the viruses. The LIVP strain produced higher levels of antibodies compared to the mutant LIVP-A34R variant. The weakest immune response was revealed in mice that received s.c. injections of the viruses. In i.d. inoculated mice, neutralizing antibodies were synthesized at a relatively high level, while the LIVP-A34R strain ensured a more efficient production of these antibodies compared to the LIVP strain.

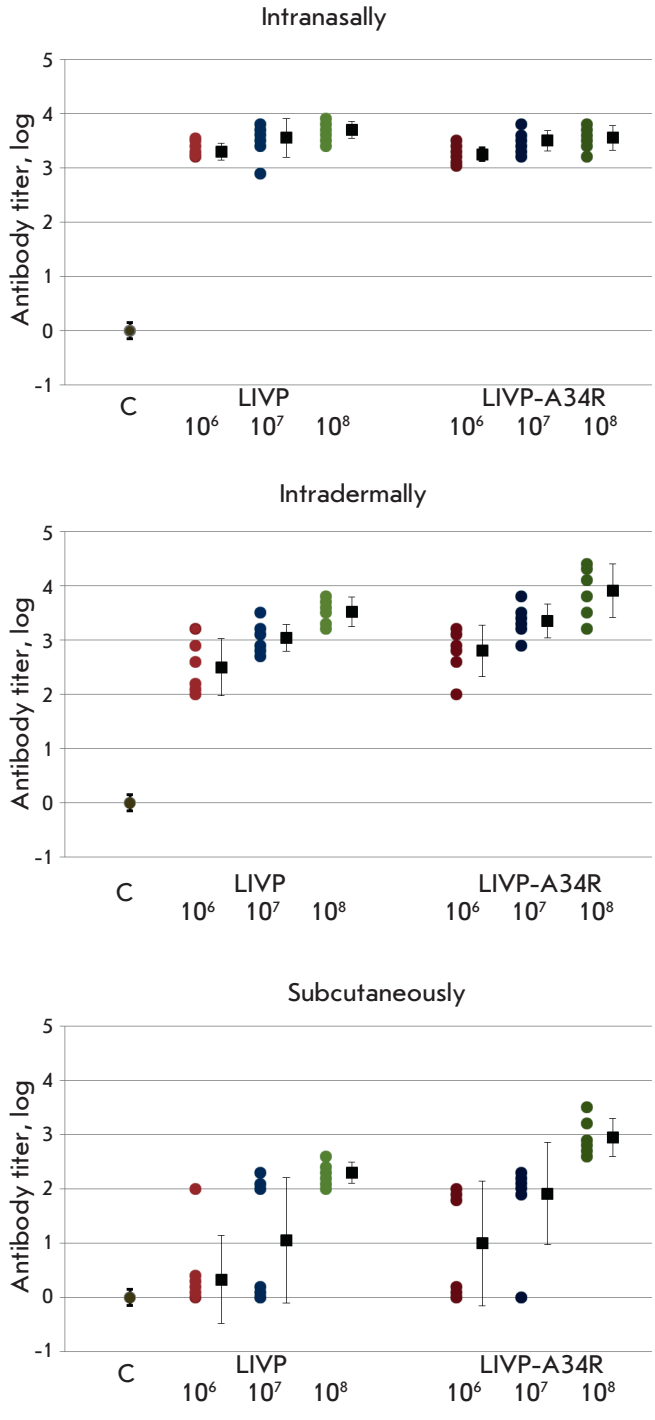


Fig. 7. The titers of ELISA-determined VACV-specific antibodies in the serum samples of mice inoculated with the LIVP or LIVP-A34R strain at different doses through different routes. The data for individual animals and the geometric means of log reciprocal titer of VACV-specific IgG and confidence levels for the 95% matching between each sample and the total population are presented for each group

DISCUSSION

Vaccination was introduced over 200 years ago when ways to protect against smallpox were being developed. It remains the most reliable means for preventing viral

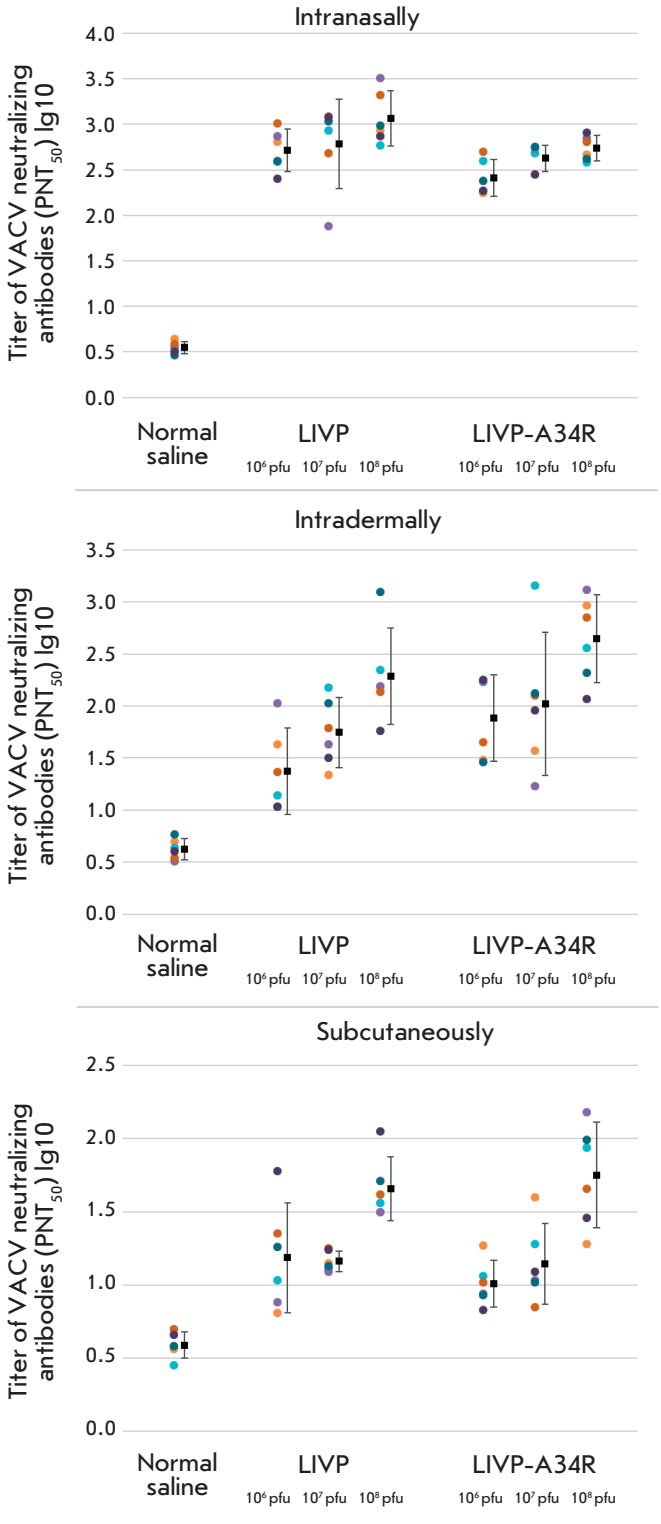


Fig. 8. Levels of VACV-neutralizing activities of the serum samples collected on day 28 after the mice had been infected with LIVP or LIVP-A34R strains through different routes. The data for individual animals and the geometric means of log reciprocal antivirus neutralization titer and confidence levels for the 95% matching are presented for each group

infections [1, 6]. The conventional live smallpox vaccine was produced by replicating the VACV on the skin of calves or other domestic animals. This vaccine provided reliable protection against smallpox but in some cases caused severe post-vaccination adverse reactions (including encephalitis and encephalomyelitis, sometimes lethal) [1, 2]. Therefore, after the World Health Organization announced in May 1980 that smallpox had been eradicated, vaccination against this extremely dangerous human infectious disease was discontinued. As a result, today a large share of the world's population has immunity neither against smallpox nor against other zoonotic orthopoxvirus infections such as the monkeypox, cowpox, camelpox, and vaccinia. The lack of herd immunity significantly facilitates the circulation of zoonotic orthopoxviruses in the human population [23–27]. Special concern is related to the human monkeypox, as its clinical presentation in humans is similar to that of smallpox, and the lethality of this infection can reach 10%. Furthermore, the efficiency of human-to-human transmission of the monkeypox virus has recently witnessed a manifold increase [23, 25].

In order to preclude a development of small outbreaks of orthopoxvirus infections into massive epidemics and reduce the risk associated with them due to the natural evolution of the orthopoxvirus, which is highly pathogenic for humans, researchers have focused their efforts on developing safe, next-generation live VACV-based vaccines [3, 6, 28].

The modern approach to designing attenuated, highly immunogenic vaccines usually involves a targeted inactivation of the VACV virulence genes [6, 15, 29–31]. Furthermore, the pathogenicity and immunogenicity of VACV depend on the virus strain and its route of inoculation into the animal organism [2, 6, 32–35].

Earlier, it was shown using laboratory VACV strains that the VACV *A34R* gene is one of the key genes that regulate the detachment of the extracellular CEVs bound to the infected cell surface from the cell and their release into the environment (the so-called EEV form) [7, 10]. The *A34R* gene encodes a glycoprotein carrying a lectin-like domain within the outer membrane of VACV EEVs [7, 10, 36]. It turns out that protein A34 of the WR VACV strain producing less than 1% of virus progeny in the form of EEV in the cell culture differs from this same protein in the IHD-J strain (with the EEVs constituting up to 30–40% of its progeny) by only two amino acid residues at positions 110 and 151. Substitution of the *A34R* gene in the WR strain with a variant of this gene in the IHD-J strain significantly increases production of the EEV form, but the yield of EEVs typical of the VACV IHD-J strain is not attained [10]. Protein A34 performs its function (ensuring the release of EEVs from cells) by interacting

with a number of other viral and cellular proteins [37] and provides efficient binding of EEVs to the cell and their penetration into the cell [8, 38].

It is believed that the C-terminal lectin-like domain of viral glycoprotein A34, which resides on the surface of extracellular virions (CEVs and EEVs) provides a highly specific interaction between this protein (and virions) and the carbohydrates on the cell surface [10]. The Lys151 → Glu substitution within this domain of protein A34 presumably reduces the efficiency of binding of the VACV virions released from the cell to the cell surface and increases the release of EEVs into the environment [10, 39].

The complex formed between the viral proteins B5 and A34 plays a crucial role in the binding of EEVs to the cell surface. The (80–130 a.a.) domain in protein A34 is the region where these proteins interact [38]. The Asp110 → Asn mutation in glycoprotein A34 of the IHD-J strain probably reduces the efficiency of complex formation between the proteins A34 and B5, thus decreasing the efficiency of CEV binding to the cell surface and additionally increasing the yield of EEVs.

We studied for the first time how an elevated production of the EEV form of VACV can affect the virulence and immunogenicity of the virus depending on its route of inoculation into laboratory mice.

The studies were conducted using the VACV LIVP strain, which is conventionally utilized for smallpox vaccination in Russia [2]. The clonal LIVP variant described earlier was used as the parent strain [15]. Two point mutations typical of the *A34R* gene in the IHD-J VACV strain were inserted. It was demonstrated for the CV-1 cell culture that the designed mutant LIVP-A34R variant produces a statistically significantly greater amount of EEVs compared to the parent LIVP strain (*Fig. 1*).

Intranasal inoculation of both VACV strains to BALB/c mice was shown to have a pathogenic effect, which was revealed through clinical signs and a reduction of mouse body weight (*Fig. 2*). Peak of the infection occurred on 6–8 dpi. The pathogenicity of the mutant LIVP-A34R VACV strain (assessed using these signs) was somewhat lower. Intradermal (*Fig. 3*) or subcutaneous inoculation of the viruses even at the maximal dose of 10^8 pfu neither reduced the mouse body weight nor led to the emergence of clinical signs of the disease.

An analysis of viral dissemination within the organisms of the experimental animals demonstrated that the highest *in vivo* viral dissemination was observed after i.n. inoculation; the viral load in the internal mouse organs depended on the infective dose (*Fig. 4*). It should be mentioned that more of the mutant VACV variant accumulated in the lungs of mice inoculated with the virus at doses of 10^7 and 10^6 pfu on 7 dpi (the peak of infection) compared to the parent LIVP strain.

Histological examination of mouse organs revealed the most typical manifestations of the infection in the organs of the respiratory system, mostly in the lungs. Pathological changes in the lungs in mice i.n. inoculated with LIVP-A34R appeared earlier than in mice infected with LIVP, being more severe because of the more significant involvement of microcirculatory vessels. Therefore, edema, plasmorrhagia and hemorrhagia developed to a greater extent on 3 dpi (*Fig. 5*). In both cases, the pathological manifestations decreased on 10 dpi.

Only the LIVP-A34R strain was detected in the liver in i.n. inoculated mice on 7 dpi (*Fig. 4*). All these findings demonstrate that LIVP-A34R is disseminated in the mouse organism more efficiently compared to LIVP. However, a lower level of LIVP-A34R accumulated in the brain compared to LIVP.

Intracerebral inoculation of the virus to newborn mice also demonstrated that the LIVP-A34R strain was characterized by reduced neurovirulence compared to the parent LIVP (*Fig. 6*).

For i.d. inoculated mice, the viruses were detected only in skin samples collected from the injection site of the viral preparation and at the maximum infective dose in the lungs and liver of some animals on 3 and 7 dpi. The amount of the mutant LIVP-A34R variant detected in the lungs was larger compared to that for the parent LIVP strain. No viruses were detected in the brain, spleen, or kidney samples.

In animals s.c. inoculated with the analyzed VACV variants, the viruses were detected only in the skin flap samples harvested from the injection site for the maximum infective dose. No viruses were detected in the internal organ samples.

It is known that the level of antibody response to vaccination against orthopoxvirus infections plays a decisive role in ensuring protection against a subsequent viral infection [6]. Therefore, we confined ourselves to studying the induction of biosynthesis of antiviral antibodies depending on the dose and route of inoculation of VACV into laboratory mice. The serum titers of the antiviral antibodies in mice inoculated with the viral strains under study were measured using two methods. The enzyme-linked immunosorbent assay was used to test the antibodies specifically interacting with the virion proteins of VACV IMVs. The titers of the antibodies that were bound to IMVs *in vitro*, thus suppressing their infectivity (plaque formation) upon the subsequent inoculation to the cell culture, were quantified in the same serum samples using the second method. Correlated results were obtained after i.d. inoculation of the viruses (*Figs. 7* and *8*). Both viruses

induced high production levels of both VACV-specific and VACV neutralizing antibodies; but LIVP-A34R elicited a stronger immune response.

The lowest level of production of antiviral antibodies was observed in mice s.c. inoculated with both LIVP and LIVP-A34R.

The maximum levels of production of both VACV-specific and virus – neutralizing antibodies were revealed in mice i.n. inoculated with the LIVP strain (*Figs. 7, 8*). For this route of inoculation, the mutant LIVP-A34R virus induced the formation of VACV-specific antibodies at concentrations comparable to those observed for the parent LIVP strain and a lower amount of VACV-neutralizing antibodies.

Hence, when inoculated through the i.n. route, both viruses exhibit pronounced pathogenicity, disseminate through internal organs, and therefore ensure a high level of induction of antiviral antibodies. The pathogenicity and neurovirulence of the LIVP-A34R strain are lower than those of LIVP. However, taking into account the high total virulence of the infection, this route of virus inoculation is hardly acceptable for the LIVP or LIVP-A34R strain being used as a live smallpox vaccine or a platform for designing a recombinant polyvalent vaccine.

When inoculated s.c. to mice, both LIVP and LIVP-A34R exhibit low virulence and produce a low level of virus-specific antibodies.

Intradermal injection can be considered as the optimal route for inoculating both the parent LIVP strain and the mutant LIVP-A34R variant based on the pathogenicity/immunogenicity ratio. In addition to being detected in the skin at the injection site, after i.d. injection viruses are revealed by titration (the detection level, $\geq 10^2$ pfu/g of the organ) on 3 and 7 dpi only in the lungs and liver of some animals solely at the maximum infection dose used (10^8 pfu). Meanwhile, a larger amount of the LIVP-A34R variant accumulated in the lungs compared to LIVP. The increased ability to disseminate in the organism could be the reason why the LIVP-A34R strain inoculated i.d. exhibited a higher immunogenicity compared to the parent LIVP variant.

CONCLUSIONS

The results of this study demonstrate that, through the insertion of two point substitutions in the sequence of the A34 viral protein, the LIVP-A34R VACV strain produces more extracellular enveloped virions (EEVs) compared to the parent LIVP strain, is less neurovirulent, and induces an enhanced production of antiviral antibodies when administered intradermally. This variant of VACV can be used as a platform to develop a highly immunogenic, attenuated vaccine against smallpox and other re-emerging human orthopoxvirus infections. This variant of VACV can also be used as a molecular vector to design live recombinant polyvalent vaccines against various infectious diseases and oncolytic VACV variants. ●

This work was supported by the Russian Science Foundation (grant No. 19-14-00006).

REFERENCES

1. Fenner F., Henderson D.A., Arita I., Jezek Z., Ladnyi I.D. Smallpox and its eradication. Geneva: World Health Organization, 1988. 1460 p.
2. Shchelkunov S.N., Marennikova S.S., Moyer R.W. Orthopoxviruses pathogenic for humans. Berlin, Heidelberg, New York: Springer, 2005. 425 p.
3. Moss B. // Immunol. Rev. 2011. V. 239. P. 8–26.
4. Shchelkunova G.A., Shchelkunov S.N. // Acta Naturae. 2017. V. 9. P. 4–12.
5. Payne L.G. // J. Gen. Virol. 1980. V. 50. P. 89–100.
6. Shchelkunov S.N., Shchelkunova G.A. // Acta Naturae. 2020. V. 44. P. 33–41.
7. Smith G.L. // J. Gen. Virol. 2002. V. 83. P. 2915–2931.
8. Locker J.K., Kuehn A., Schleich S., Rutter G., Hohenberg H., Wepf R., Griffiths G. // Mol. Biol. Cell. 2000. V. 11. P. 2497–2511.
9. Kirn D.H., Wang Y., Liang W., Contag C.H., Thorne S.H. // Cancer Res. 2008. V. 68. P. 2071–2075.
10. Blasco R., Sisler J.R., Moss B. // J. Virol. 1993. V. 67. P. 3319–3325.
11. Thirunavukarasu P., Sathaiah M., Gorry M.C., O'Malley M.E., Ravindranathan R., Austin F., Thorne S.H., Guo Z.S., Bartlett D.I. // Mol. Ther. 2013. V. 21. P. 1024–1033.
12. Olson V.A., Shchelkunov S.N. // Viruses. 2017. V. 9. e242.
13. Sanchez-Sampedro L., Perdiguero B., Mejias-Perez E., Garcia-Arriaza J., Di Pilato M., Esteban M. // Viruses. 2015. V. 7. P. 1726–1803.
14. Prow N.A., Jimenez Martinez R., Hayball J.D., Howley P.M., Suhrbier A. // Exp. Rev. Vaccines. 2018. V. 17. P. 925–934.
15. Yakubitskiy S.N., Kolosova I.V., Maksyutov R.A., Shchelkunov S.N. // Acta Naturae. 2015. V. 7. P. 113–121.
16. Shchelkunov S.N., Sergeev A.A., Kabanov A.S., Yakubitskiy S.N., Bauer T.V., Pyankov S.A. // Russian Journal of Infection and Immunity. 2020. V. 10. doi:10.15789/2220-7619-PAI-1375.17.
17. Bauer T.V., Tregubchak T.V., Shchelkunov S.N., Maksyutov R.A., Gavrilova E.V. // Medical Immunology (Russia). 2020. V. 22. P. 371–378.
18. National Research Council Guidelines on Laboratory Animal Care and Use, 8th ed. National research Council of the National Academies. Washington: The National Acad. Press, 2011. 135 p.
19. Leparc-Goffart I., Poirier B., Garin D. Tissier M.-H., Fuchs F., Crance J.-M. // J. Clin. Virol. 2005. V. 32. P. 47–52.
20. McIntosh A.A.G., Smith G.L. // J. Virol. 1996. V. 70. P. 272–281.

21. Paran N., Lustig S., Zvi A., Erez N., Israely T., Melamed S., Politi B., Ben-Nathan D., Schneider P., Lachmi B., et al. // Virol. J. 2013. V. 10. e229.
22. Stabenow J., Buller R. M., Schriewer J., West C., Sagartz J.E., Parker S. // J. Virol. 2010. T. 84. C. 3909–3920.
23. Shchelkunov S.N. // PLoS Path. 2013. V. 9. e1003756.
24. Peres M.G., Bacchiega T.S., Appolinario C.M., Vicente A.F., Mioni M.S.R., Ribeiro B.L.D., Fonseca C.R.S., Pelicia V.C., Ferreira F., Oliveira G.P., et al. // Viruses. 2018. V. 10. e42.
25. Shchelkunov S.N., Shchelkunova G.A. // Voprosy Virusologii (Russia). 2019. V. 64. P. 206-214.
26. Reynolds M.G., Doty J.B., McCollum A.M., Olson V.A., Nakazawa Y. // Expert Rev. Anti. Infect. Ther. 2019. V. 17. P. 129–139.
27. Styczynski A., Burgado J., Walteros D., Usme-Ciro J., Laiton K., Farias A.P., Nakazawa Y., Chapman C., Davidson W., Mauldin M., et al. // Emerg. Infect. Dis. 2019. V. 25. P. 2169–2176.
28. Shchelkunov S.N. // Vaccine. 2011. V. 29S. P. D49–D53.
29. Yakubitskiy S.N., Kolosova I.V., Maksyutov R.A., Shchelkunov S.N. // Dokl. Biochem. Biophys. 2016. V. 466. P. 35–38.
30. Di Pilato M., Mejias-Perez E., Sorzano C.O.S., Esteban M. // J. Virol. 2017. V. 91. e00575–17.
31. Albarnaz J.D., Torres A.A., Smith G.L. // Viruses. 2018. V. 10. e101.
32. McClain D.J., Harrison S., Yeager C.L., Cruz J., Ennis F.A., Gibbs P., Wright M.S., Summers P.L., Arthur J.D., Graham J.A. // J. Infect. Dis. 1997. V. 175. P. 756–763.
33. Phelps A., Gates A.J., Eastaugh L., Hillier M., Ulaeto D.O. // Vaccine. 2017. V. 35. P. 3889–3896.
34. Roy S., Jaeson M.I., Li Z., Mahboob S., Jackson R.J., Grubor-Bauk B., Wijesundara D.K., Gowans E.J., Ranasinghe C. // Vaccine. 2019. V. 37. P. 1266–1276.
35. Xie L., Zai J., Yi K., Li Y. // Vaccine. 2019. V. 37. P. 3335–3342.
36. McIntosh A.A.G., Smith G.L. // J. Virol. 1996. V. 70. P. 272–281.
37. McNulty S., Powell K., Erneux C., Kalman D. // J. Virol. 2011. V. 85. P. 7402–7410.
38. Monticelli S.R., Earley A.K., Tate J., Ward B.M. // J. Virol. 2019. V. 93. e01343–18.
39. Earley A.E., Chan W.M., Ward B.M. // J. Virol. 2008. V. 82. P. 2161–2169.

GENERAL RULES

Acta Naturae publishes experimental articles and reviews, as well as articles on topical issues, short reviews, and reports on the subjects of basic and applied life sciences and biotechnology.

The journal *Acta Naturae* is on the list of the leading periodicals of the Higher Attestation Commission of the Russian Ministry of Education and Science. The journal Acta Naturae is indexed in PubMed, Web of Science, Scopus and RCSI databases.

The editors of *Acta Naturae* ask of the authors that they follow certain guidelines listed below. Articles which fail to conform to these guidelines will be rejected without review. The editors will not consider articles whose results have already been published or are being considered by other publications.

The maximum length of a review, together with tables and references, cannot exceed 60,000 characters with spaces (approximately 30 pages, A4 format, 1.5 spacing, Times New Roman font, size 12) and cannot contain more than 16 figures.

Experimental articles should not exceed 30,000 symbols (approximately 15 pages in A4 format, including tables and references). They should contain no more than ten figures.

A short report must include the study's rationale, experimental material, and conclusions. A short report should not exceed 12,000 symbols (8 pages in A4 format including no more than 12 references). It should contain no more than four figures.

The manuscript and all necessary files should be uploaded to www.actanaturae.ru:

1) text in Word 2003 for Windows format;
2) the figures in TIFF format;
3) the text of the article and figures in one pdf file;
4) the article's title, the names and initials of the authors, the full name of the organizations, the abstract, keywords, abbreviations, figure captions, and Russian references should be translated to English;
5) the cover letter stating that the submitted manuscript has not been published elsewhere and is not under consideration for publication;
6) the license agreement (the agreement form can be downloaded from the website www.actanaturae.ru).

MANUSCRIPT FORMATTING

The manuscript should be formatted in the following manner:

- Article title. Bold font. The title should not be too long or too short and must be informative. The title should not exceed 100 characters. It should reflect the major result, the essence, and uniqueness of the work, names and initials of the authors.
- The corresponding author, who will also be working with the proofs, should be marked with a footnote *.
- Full name of the scientific organization and its departmental affiliation. If there are two or more scientific organizations involved, they should be linked by digital superscripts with the authors' names. Abstract. The structure of the abstract should be very clear and must reflect the following: it should introduce the reader to the main issue and describe the experimental approach, the possibility of practical use, and the possibility of further research in the field. The average length of an abstract is 20 lines (1,500 characters).
- Keywords (3 – 6). These should include the field of research, methods, experimental subject, and the specifics of the work. List of abbreviations.
- INTRODUCTION
- EXPERIMENTAL PROCEDURES
- RESULTS AND DISCUSSION
- CONCLUSION

  The organizations that funded the work should be listed at the end of this section with grant numbers in parenthesis.
- REFERENCES

  The in-text references should be in brackets, such as [1].

RECOMMENDATIONS ON THE TYPING AND FORMATTING OF THE TEXT

- We recommend the use of Microsoft Word 2003 for Windows text editing software.
- The Times New Roman font should be used. Standard font size is 12.
- The space between the lines is 1.5.
- Using more than one whole space between words is not recommended.
- We do not accept articles with automatic referencing; automatic word hyphenation; or automatic prohibition of hyphenation, listing, automatic indentation, etc.
- We recommend that tables be created using Word software options (Table → Insert Table) or MS Excel. Tables that were created manually (using lots of spaces without boxes) cannot be accepted.
- Initials and last names should always be separated by a whole space; for example, A. A. Ivanov.
- Throughout the text, all dates should appear in the "day.month.year" format, for example 02.05.1991, 26.12.1874, etc.
- There should be no periods after the title of the article, the authors' names, headings and subheadings, figure captions, units (s – second, g – gram, min – minute, h – hour, d – day, deg – degree).
- Periods should be used after footnotes (including those in tables), table comments, abstracts, and abbreviations (mon. – months, y. – years, m. temp. – melting temperature); however, they should not be used in subscripted indexes (T_m – melting temperature; $T_{p.t}$ – temperature of phase transition). One exception is mln – million, which should be used without a period.
- Decimal numbers should always contain a period and not a comma (0.25 and not 0,25).
- The hyphen ("-") is surrounded by two whole spaces, while the "minus," "interval," or "chemical bond" symbols do not require a space.
- The only symbol used for multiplication is "×"; the "×" symbol can only be used if it has a number to its

right. The "·" symbol is used for denoting complex compounds in chemical formulas and also noncovalent complexes (such as DNA·RNA, etc.).
- Formulas must use the letter of the Latin and Greek alphabets.
- Latin genera and species' names should be in italics, while the taxa of higher orders should be in regular font.
- Gene names (except for yeast genes) should be italicized, while names of proteins should be in regular font.
- Names of nucleotides (A, T, G, C, U), amino acids (Arg, Ile, Val, etc.), and phosphonucleotides (ATP, AMP, etc.) should be written with Latin letters in regular font.
- Numeration of bases in nucleic acids and amino acid residues should not be hyphenated (T34, Ala89).
- When choosing units of measurement, SI units are to be used.
- Molecular mass should be in Daltons (Da, KDa, MDa).
- The number of nucleotide pairs should be abbreviated (bp, kbp).
- The number of amino acids should be abbreviated to aa.
- Biochemical terms, such as the names of enzymes, should conform to IUPAC standards.
- The number of term and name abbreviations in the text should be kept to a minimum.
- Repeating the same data in the text, tables, and graphs is not allowed.

GUIDENESS FOR ILLUSTRATIONS

- Figures should be supplied in separate files. Only TIFF is accepted.
- Figures should have a resolution of no less than 300 dpi for color and half-tone images and no less than 500 dpi.
- Files should not have any additional layers.

REVIEW AND PREPARATION OF THE MANUSCRIPT FOR PRINT AND PUBLICATION

Articles are published on a first-come, first-served basis. The members of the editorial board have the right to recommend the expedited publishing of articles which are deemed to be a priority and have received good reviews.

Articles which have been received by the editorial board are assessed by the board members and then sent for external review, if needed. The choice of reviewers is up to the editorial board. The manuscript is sent on to reviewers who are experts in this field of research, and the editorial board makes its decisions based on the reviews of these experts. The article may be accepted as is, sent back for improvements, or rejected.

The editorial board can decide to reject an article if it does not conform to the guidelines set above.

The return of an article to the authors for improvement does not mean that the article has been accepted for publication. After the revised text has been received, a decision is made by the editorial board. The author must return the improved text, together with the responses to all comments. The date of acceptance is the day on which the final version of the article was received by the publisher.

A revised manuscript must be sent back to the publisher a week after the authors have received the comments; if not, the article is considered a resubmission.

E-mail is used at all the stages of communication between the author, editors, publishers, and reviewers, so it is of vital importance that the authors monitor the address that they list in the article and inform the publisher of any changes in due time.

After the layout for the relevant issue of the journal is ready, the publisher sends out PDF files to the authors for a final review.

Changes other than simple corrections in the text, figures, or tables are not allowed at the final review stage. If this is necessary, the issue is resolved by the editorial board.

FORMAT OF REFERENCES

The journal uses a numeric reference system, which means that references are denoted as numbers in the text (in brackets) which refer to the number in the reference list.

For books: the last name and initials of the author, full title of the book, location of publisher, publisher, year in which the work was published, and the volume or issue and the number of pages in the book.

For periodicals: the last name and initials of the author, title of the journal, year in which the work was published, volume, issue, first and last page of the article. Must specify the name of the first 10 authors. Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McLay K., Muzny D., Platzer M., Howell G.R., Burrows C., Bird C.P., et al. // Nature. 2005. V. 434. № 7031. P. 325–337.

References to books which have Russian translations should be accompanied with references to the original material listing the required data.

References to doctoral thesis abstracts must include the last name and initials of the author, the title of the thesis, the location in which the work was performed, and the year of completion.

References to patents must include the last names and initials of the authors, the type of the patent document (the author's rights or patent), the patent number, the name of the country that issued the document, the international invention classification index, and the year of patent issue.

The list of references should be on a separate page. The tables should be on a separate page, and figure captions should also be on a separate page.

The following e-mail addresses can be used to contact the editorial staff: vera.knorre@gmail.com, actanaturae@gmail.com,
tel.: (495) 727-38-60, (495) 930-87-07