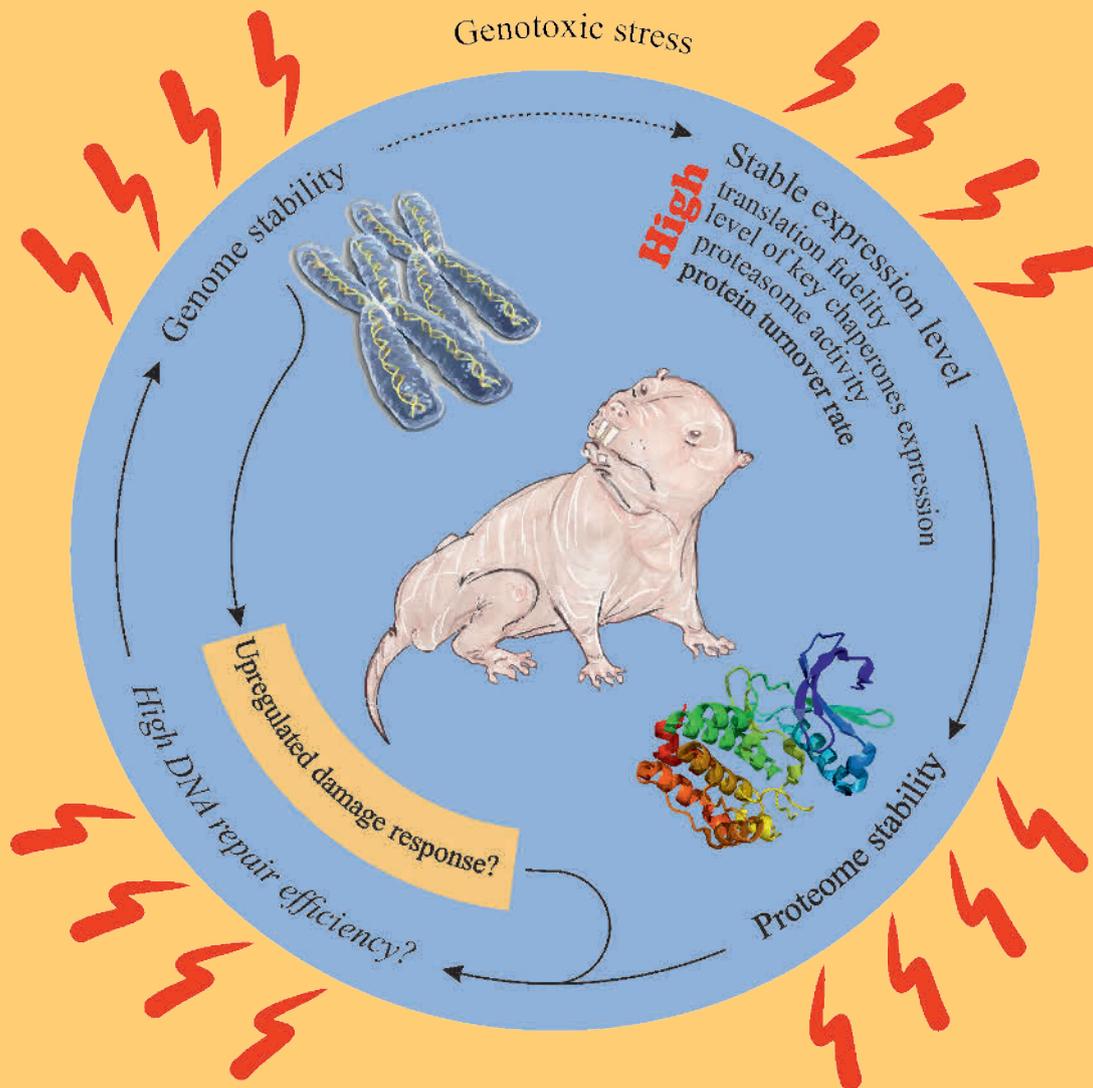


Acta Naturae

Genome Stability Maintenance in Naked Mole-Rat



NON-CODING RNAs AS TRANSCRIPTIONAL
REGULATORS IN EUKARYOTES
P. 13

40 YEARS WITHOUT
SMALLPOX
P. 4



MERCK

Comprehensive solutions for cell analysis



- Cell lines and primary cells
- Traditional and specialized culture media
- Sterilizing filtration



- Biochemical reagents
- Water purification systems
- Cell counting and analysis
- Cryopreservation



An extensive range and top quality of cell lines from our partner, the European Collection of Authenticated Cell Cultures (ECACC):

- 4000 animal and human cell lines;
- Cells of 45 animal species and 50 tissue types;
- 370 B-lymphoblastoid cell lines for which human leukocyte antigen (HLA) typing data are available;
- 480 hybridoma cell lines secreting monoclonal antibodies;
- DNA, RNA, and cDNA extracted from the cell lines from our collection;

[SIGMAaldrich.com/ECACC](https://www.sigmaaldrich.com/ECACC)

LLC Merck

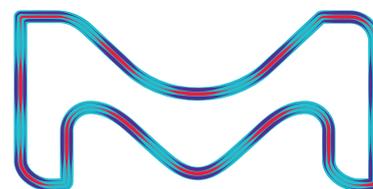
Valovaya Str., 35, Moscow, 115054, Russia;

Tel. +7 (495) 937-33-04

E-mail: mm.russia@merckgroup.com, ruorder@sial.com

[SIGMAaldrich.com/cellculture](https://www.sigmaaldrich.com/cellculture)

[MERCKmillipore.com/cellculture](https://www.merckmillipore.com/cellculture)

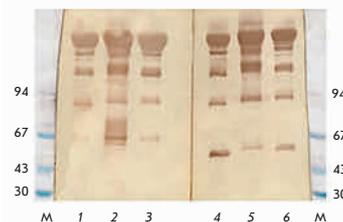


SIGMA-ALDRICH is now **MERCK**

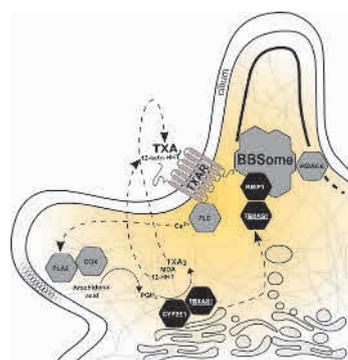
Recombinant Antibodies to the Ebola Virus Glycoprotein

A. A. Panina, I. G. Dementieva, T. K. Aliev, V. A. Toporova, D. S. Balabashin, M. N. Bokov, L. P. Pozdnyakova, O. B. Shemchukova, D. A. Dolgikh, P. G. Sveshnikov, M. P. Kirpichnikov

Currently, there are no approved therapies for targeted prevention and treatment of Ebola hemorrhagic fever. In the present work, authors described the development of a eukaryotic expression system for production of three full-length chimeric antibodies (IgG1-kappa isotypes) GPE118, GPE325 and GPE534 to the recombinant glycoprotein of the Ebola virus (EBOV GP), which is a key factor in the pathogenicity of the disease.



Immunoblot of recombinant chimeric antibodies



Schematic representation of thromboxane A₂ biosynthesis

Direct Molecular Fishing of New Protein Partners for Human Thromboxane Synthase

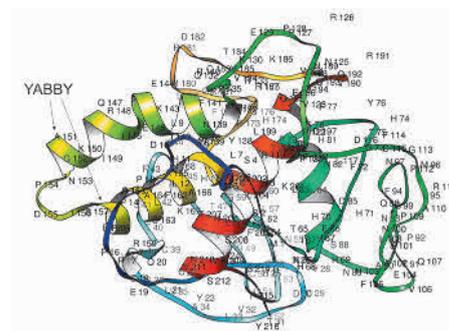
A. V. Svirid, P. V. Ershov, E. O. Yablokov, L. A. Kaluzhskiy, Yu. V. Mezentsev, A. V. Florinskaya, T. A. Sushko, N. V. Strushkevich, A. A. Gilep, S. A. Usanov, A. E. Medvedev, A. S. Ivanov

The objective of this study was to identify protein partners of human thromboxane synthase (TBXAS1) using a complex approach based on the direct molecular fishing technique, LC-MS/MS protein identification and validation of protein-protein interactions by surface plasmon resonance. Twelve potential TBXAS1 protein partners were identified, including the components regulating cytoskeleton organization, components of the coagulation cascade, and the enzyme involved in the metabolism of xenobiotics and endogenous bioregulators.

YABBY3-Orthologous Genes in Wild Tomato Species: Structure, Variability, and Expression

M. A. Filyushin, M. A. Slugina, A. V. Shchennikova, E. Z. Kochieva

Complete sequences of *YABBY3*-orthologous genes were identified in 13 samples of cultivated and wild tomato species. These genes encode transcription factors that play a key role in determining the abaxial-adaxial lateral asymmetry of all aboveground lateral plant organs. The structure of *YABBY3* genes and proteins they encode is similar to that of the previously characterized members of the *YABBY* family. Phylogenetic and expression analysis confirmed that the identified genes belong to the *YABBY1/3* subfamily and may have conserved function in different tomato species.



Founders

Ministry of Education and
Science of the Russian Federation,
Lomonosov Moscow State University,
Park Media Ltd

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, 2017

Номер подписан в печать 27 декабря 2017 г.

Тираж 200 экз. Цена свободная.

Отпечатано в типографии «МИГ ПРИНТ»

CONTENTS

FORUM

G. A. Shchelkunova, S. N. Shchelkunov
40 Years without Smallpox 4

REVIEWS

O. Y. Burenina, T. S. Oretskaya, E. A. Kubareva
**Non-coding RNAs As Transcriptional
Regulators In Eukaryotes** 13

Z. S. Kutlubaeva, H. V. Chetverina,
A. B. Chetverin
**The Contribution of Ribosomal
Protein S1 to the Structure and Function
of Q β Replicase** 26

I. O. Petrusseva, A. N. Evdokimov, O. I. Lavrik
**Genome Stability Maintenance
in Naked Mole-Rat** 31

S. S. Ryabichko, A. N. Ibragimov,
L. A. Lebedeva, E. N. Kozlov, Y. V. Shidlovskii
**Super-Resolution Microscopy
in Studying the Structure and Function
of the Cell Nucleus** 42

RESEARCH ARTICLES

M. V. Arkhipenko, N. A. Nikitin,
E. K. Donchenko, O. V. Karpova,
J. G. Atabekov
**Translational Cross-Activation
of the Encapsidated RNA of Potexviruses** ... 52

I. V. Zelepukin, V. O. Shipunova, A. B. Mirkasymov, P. I. Nikitin, M. P. Nikitin, S. M. Deyev Synthesis and Characterization of Hybrid Core-Shell Fe₃O₄/SiO₂ Nanoparticles for Biomedical Applications	58
S. V. Kalinichenko, M. V. Shepelev, P. N. Vikhrev, I. V. Korobko A Novel Hybrid Promoter ARE-hTERT for Cancer Gene Therapy	66
N. G. Kukava, B. V. Titov, G. J. Osmak, N. A. Matveeva, O. G. Kulakova, A. V. Favorov, R. M. Shakhnovich, M. Ya. Ruda, O. O. Favorova Multilocus Analysis of Genetic Susceptibility to Myocardial Infarction in Russians: Replication Study	74
A. A. Panina, I. G. Dementieva, T. K. Aliev, V. A. Toporova, D. S. Balabashin, M. N. Bokov, L. P. Pozdnyakova, O. B. Shemchukova, D. A. Dolgikh, P. G. Sveshnikov, M. P. Kirpichnikov Recombinant Antibodies to the Ebola Virus Glycoprotein	84
A. V. Svirid, P. V. Ershov, E. O. Yablokov, L. A. Kaluzhskiy, Yu. V. Mezentsev, A. V. Florinskaya, T. A. Sushko, N. V. Strushkevich, A. A. Gilep, S. A. Usanov, A. E. Medvedev, A. S. Ivanov Direct Molecular Fishing of New Protein Partners for Human Thromboxane Synthase	92
M. A. Filyushin, M. A. Slugina, A. V. Shchennikova, E. Z. Kochieva YABBY3-Orthologous Genes in Wild Tomato Species: Structure, Variability, and Expression	101

SHORT REPORTS

A. E. Gaydukov, O. P. Balezina CaMKII Is Involved in the Choline-Induced Downregulation of Acetylcholine Release in Mouse Motor Synapses	110
Guidelines for Authors	114

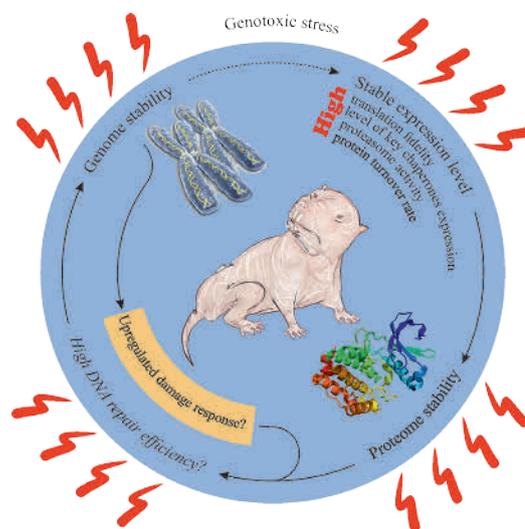


IMAGE ON THE COVER PAGE
(see the article by Petrusseva *et al.*)

40 Years without Smallpox

G. A. Shchelkunova¹, S. N. Shchelkunov^{1,2*}

¹State Research Center of Virology and Biotechnology VECTOR, Koltsovo, Novosibirsk region, 630559, Russia

²Novosibirsk State University, Pirogov Str. 2, Novosibirsk, 630090, Russia

*E-mail: snshchel@rambler.ru; snshchel@vector.nsc.ru

ABSTRACT The last case of natural smallpox was recorded in October, 1977. It took humanity almost 20 years to achieve that feat after the World Health Organization had approved the global smallpox eradication program. Vaccination against smallpox was abolished, and, during the past 40 years, the human population has managed to lose immunity not only to smallpox, but to other zoonotic orthopoxvirus infections as well. As a result, multiple outbreaks of orthopoxvirus infections in humans in several continents have been reported over the past decades. The threat of smallpox reemergence as a result of evolutionary transformations of these zoonotic orthopoxviruses exists. Modern techniques for the diagnostics, prevention, and therapy of smallpox and other orthopoxvirus infections are being developed today.

KEYWORDS smallpox, variola (smallpox) virus, evolution, DNA diagnostics, vaccine, chemotherapeutic drugs.

ABBREVIATIONS BPXV – buffalopox virus; CMLV – camelpox virus; CPXV – cowpox virus; MPXV – monkeypox virus; PCR – polymerase chain reaction; RPXV – rabbitpox virus; VACV – vaccinia virus; VARV – variola virus; WHA – World Health Assembly; WHO – World Health Organization.

INTRODUCTION

Smallpox (also known by its Latin name *variola vera*, derived from *varius* (spotted) or *varus* (pimple)) got its current name as early as the 16th century, although the disease had been known since ancient times and had inflicted a heavier toll on humans than many other infections and numerous wars. In the 20th century alone, over the almost 80 years during which mass vaccination against smallpox and an intensive anti-epidemic campaign were being conducted, 300 million people still died as a result of the disease [1].

In 1796, the English physician Edward Jenner proposed a method for protection against smallpox via the inoculation of the infectious material obtained from cows with a smallpox-like disease. This method became known as *vaccination* (from the Latin *vacca* for cow). This breakthrough event took place almost a century before the kingdom of viruses was discovered [1–3].

After its introduction in 1919, mandatory smallpox vaccination in Russia (and then, in the Soviet Union), a vast country with mani-

fold geographic conditions ranging from high-mountain expanses and deserts to northern tundra and outlandish taiga areas and home to dozens of nationalities differing in traditions, rites, and religious practices, made it possible to eliminate smallpox morbidity by 1936 [3].

This dangerous, highly contagious disease was eradicated in many developed countries in the first half of the 20th century. However, smallpox outbreaks were still recorded each year in 50–80 countries even in the 1950s. In addition, the foci of endemic smallpox in Asia, Africa, and South America posed a constant threat of importation to countries already free of the disease.

Based on an analysis of the tremendous scientific and organizational expertise on smallpox eradication accumulated in the Soviet Union, V.M. Zhdanov suggested initiating a worldwide program of smallpox eradication at the 9th World Health Assembly (WHA). The corresponding resolution implying complete smallpox eradication was adopted at the 7th WHA

plenary meeting on June 12, 1958 [1, 2].

This marked the start of an unprecedented international program of global smallpox eradication under the aegis of the World Health Organization (WHO). The Soviet Union not only played a key role in initiating the eradication program, but it was also an important backer at all stages of its implementation in subsequent years. In 1958, the year of its inception, the Soviet Government offered 25 million doses of dry smallpox vaccines to the WHO, which was then delivered to different countries. In 1960, a laboratory for large-scale production of the vaccine, in compliance with the WHO requirements, was organized at the Institute of Viral Preparations (IVP, Moscow). This laboratory subsequently became a center where professionals from different countries came to master smallpox vaccine manufacturing. A total of over 1.5 billion doses of the smallpox vaccine produced in the Soviet Union were used for mass vaccination in 45 countries over 20 years of the international smallpox eradica-

tion program. This represents one of the key roles played by the Soviet Union in the global smallpox eradication [2].

The Laboratory for Smallpox Prevention at the Institute of Viral Preparations played an important role in global smallpox eradication and led to the establishment of the International Reference Center for Smallpox. Numerous Soviet experts were trained at the Center before visiting smallpox-endemic countries and received the necessary preparation for practical work.

Thanks to the joint efforts of the world community in anti-epidemic control and mass anti-smallpox vaccination under the Intensified Smallpox Eradication Programme approved by WHO in 1966, the last natural case of smallpox was recorded in Somalia in October, 1977. Based on the statement made by the Global Commission for the Certification of Smallpox Eradication on May 8, 1980, the 33rd WHA declared that people throughout the globe had overcome smallpox. This was the first, and yet only, victory of the world community over a highly dangerous infectious human disease [1–3].

THE GENOME PROJECT

Once smallpox had been eradicated, the number of laboratories that stored a smallpox virus named variola virus (VARV) was reduced to stave off the risk of its accidental spread. As early as 1981, only four such laboratories remained (in the United States, the Soviet Union, the Republic of South Africa, and the United Kingdom) and their number was reduced to two in 1984. The latter two laboratories, namely, the Institute of Viral Preparations (Moscow, Soviet Union) and Centers for Disease Control and Prevention (CDC, Atlanta, United States) got the status of WHO Collaborating Centers for Smallpox and Other Poxvirus Infections [1].

Despite strict WHO control, these two repositories of live VARV strains were regarded as a source of potential biological threat. Correspondingly, a decision was made at the 4th meeting of the WHO Committee on Orthopoxvirus Infections (Geneva, 1986) to destroy the collections of VARV strains and their genomic DNAs. Taking into account the planned destruction of VARV collections, it was necessary to reliably preserve the genetic material of various VARV isolates in biologically safe form, as an issue of paramount importance for future research. In order to preserve information about this unique virus, the WHO Advisory Board deemed it necessary to sequence the VARV genome [4].

Correspondingly, A.I. Kondrusev, a deputy minister of public health in the Soviet Union, and Yu.T. Kalinin, a deputy minister of medical industry, approved the *National Program for Conservation of Genetic Material of the Russian Collection of Variola Virus Strains* with L.S. Sandakhchiev, director general of the Scientific and Production Association Vector (VECTOR), and O.G. Andzhaparidze, the director of IVP, as scientific supervisors and S.N. Shchelkunov and S.S. Marennikova as principal researchers.

In December 1990, the 5th meeting of the WHO Committee on Orthopoxvirus Infections approved the national programs for research into the VARV genome proposed by Russia (VECTOR, Koltsovo, Novosibirsk region, and IVP, Moscow) and the United States (CDC, Atlanta, Georgia, and Institute for Genomic Research, Gaithersburg, Maryland). In May 1991, the WHO Commission inspected the VECTOR's laboratory headed by S.N. Shchelkunov and officially approved the cloning of VARV DNA fragments and their sequencing.

The species specific name of VARV is *Variola virus*. Two sub-

species are commonly distinguished: *V. major*, causing the disease with a mortality rate of 5–40% and *V. minor*, with a lethal outcome of less than 2% [1]. VARV is a member of the genus *Orthopoxvirus* belonging to the family Poxviridae. This genus also includes the zoonotic species *Monkeypox virus* (MPXV), *Cowpox virus* (CPXV), *Vaccinia virus* (VACV), *Buffalopox virus* (BPXV, a subspecies of VACV), and *Camelpox virus* (CMLV), all able to infect humans [2–4]. Orthopoxviruses are closely related in their antigenic and immunological characteristics and provide cross-protection when they infect humans or animals [1].

By mid-1992, Russian scientists were first to successfully complete the genome sequencing of a highly virulent VARV *major* strain isolated in India in 1967 during a smallpox outbreak with a mortality rate of 31%, perform a computer analysis of the sequencing data [13–16], and compare them to the then recently published genome sequence of VACV [17, 18]. The results of that work were for the first time reported as an oral presentation at the opening the 9th International Conference on Poxviruses and Iridoviruses [19]. One year later, an American team completed the sequencing and analysis of the whole genome of another highly virulent VARV *major* strain, Bangladesh-1975, isolated during a smallpox outbreak with a mortality rate of 18.5% [20]. Subsequent comparison of the genomes of these strains revealed that they were highly conserved [21, 22].

It was decided at the 6th Meeting of the WHO Committee on Orthopoxvirus Infections (September 1994, Geneva, Switzerland) that the VARV DNA stocks should be stored in two international repositories, namely, VECTOR (by that time with the status of State Research Center of Virology and Biotechnology) and CDC (United States).

The complete genome of a low virulent VARV *minor* strain, Garcia-1966, was sequenced and analyzed (*table*) by collaborating teams from VECTOR and the CDC [23].

Taking into account the potential threat related to manipulations of live VARV in Moscow, the collection of VARV strains was transferred from the Institute of Viral Preparations to VECTOR (Koltsovo, Novosibirsk region) in September, 1994, under a joint order from the Russian Ministry of Health and Medical Industry, the Ministry of Science, the State Committee on Sanitary and Epidemiology Surveillance, and the Academy of Medical Sciences.

The WHO officially registered the organization of the WHO Collaborating Center for Orthopoxvirus Diagnosis and Repository for Variola Virus Strains and DNA at SRC VB VECTOR in June 1997, after a WHO Commission, in 1995, had inspected the laboratory facilities providing the highest degree of physical safety intended for this purpose. The right of VECTOR to keep the repository of VARV strains and their genomic DNAs was officially approved by WHA Resolution no. 49.01 and confirmed by later Resolutions nos. 52.10, 55.15, and 60.1.

The WHO Advisory Committee on Variola Virus Research was organized in 1999 to oversee manipulations with VARV and holds annual meetings for all experts involved in relevant studies and in the development of diagnostic, prevention, and therapeutic tools for smallpox and other human orthopoxvirus infections.

To gain insight into the evolutionary interactions of different orthopoxvirus species, it became necessary to compare their genomes. The VECTOR's team was the first to sequence the genome DNAs of CPXV [24] and MPXV [25, 26] isolated from sick individuals (*table*). The analysis of the complete VARV, MPXV, CPXV, VACV, and CMLV genomes made it possible to establish that the CPXV DNA is not only the longest among the studied orthopoxviruses, but also contains all the genetic elements characteristic of the remaining orthopoxvirus species [24, 27–31]. VARV, MPXV, and VACV can be regarded as CPXV variants with deletions, rearrangements, and point mutations specific to each individual species. This suggested to us that a CPXV-like virus is the ancestor of all extant orthopoxvirus species pathogenic for humans [24, 26, 32].

The accumulated data laid the basis for the pioneering comparative analysis of the genomic strategies utilized by all orthopoxvirus species pathogenic for humans, the first phylogenetic studies of this virus group, and the discovery of their evolutionary relationships. However, that data has not yet allowed us to date the molecular evolution of orthopoxviruses and, in particular, VARV [26, 33–36].

The issue of dating the VARV molecular evolution essentially shifted when the VECTOR and CDC teams designed a method allowing one to detect genetic differences between VARV strains. The method utilizes complete VARV genomes and includes long-distance polymerase chain reaction (LPCR) of overlapping genomic segments of the virus' DNA (with a length of 10 kbp and longer) with subsequent hydrolysis of the synthesized amplicons by frequently cutting restriction endonucleases, electrophoresis, and a computer-aided analysis of restriction fragment length polymorphism (RFLP). This relatively simple approach (the LPCR–RFLP assay), which is close to sequencing in its information content (analyzing the positions of over 300 recognition sites for several restriction endonucle-

Table 1. The first sequenced orthopoxvirus genomes

Species	Strain	Genome size, bp	Number of potential genes	Organization which made sequencing	Year of sequencing
Vaccinia virus	Copenhagen	191636	198	Virogenetics, USA	1990
Variola major virus	India-1967	185578	199	SRC VB VECTOR, Russia	1992
Variola major virus	Bangladesh-1975	186103	196	CDC, USA	1993
Variola minor virus	Garcia-1966	186986	206	SRC VB VECTOR, Russia; CDC, USA	1995
Cowpox virus	GRI-90	223666	212	SRC VB VECTOR, Russia	1997
Vaccinia virus	Ankara	177923	157	Biomedical Research Center, Austria	1998
Monkeypox virus	Zaire-96-I-16	196858	191	SRC VB VECTOR, Russia	2001
Cowpox virus	Brighton Red	224499	218	Duke University Medical Center, USA	2002
Vaccinia virus	WR	194711	206	CDC, USA	2003

ases in a virus DNA sequence), has for the first time made it possible to discover detailed differences between the genomes of 63 VARV strains from the Russian and U.S. collections isolated in several geographic regions and in different years. Phylogenetic analysis of the RFLP data for viral DNAs allowed us to pioneer the discovery that the West African and South American VARV strains form a separate subtype (clade) that significantly differ in their genome organization from the remaining, studied geographic variants of VARV [37]. It is essential here that the West African and South American VARV strains within the discovered subtype form two distinct phylogenetic groups (subclades), which suggests their independent evolution over a certain time period. The results of this analysis and the historical facts that VARV had been several times imported from West Africa to South America in the 16th–18th centuries through slaves allowed us to quantitatively estimate the rate of poxvirus evolution for the first time [38].

Sequencing of the complete genomes of a large set of VARV strains isolated in different years and geographic regions [39], as well as extended genome segments of several additional VARV strains [40], made it possible to more precisely date the key events in VARV evolution [41, 42].

POSSIBLE SMALLPOX REEMERGENCE

Taking into account the fact that smallpox vaccination in several cases had adverse side effects, the WHO recommended ceasing vaccination after 1980 in all countries. The result of this decision was that mankind lost its collective immunity not only to smallpox, but also to other zoonotic orthopoxvirus infections. The ever more frequently recorded human DNA cases of zoonotic

orthopoxvirus infections force us to revisit the problem of possible smallpox reemergence resulting from a natural evolution of these viruses [32, 43].

An important feature of VARV is its ability to infect only humans and the absence of a natural reservoir (a sensitive animal species). One should keep in mind that VARV infection of a human can in many cases (up to 40% and more) result in a lethal outcome [1–4].

MPXV causes a human disease that resembles smallpox in its clinical manifestations and also may result in a lethal outcome in up to 10% of cases. The major difference between the human monkeypox and smallpox consists in a low human-to-human transmission efficiency of the former, which so far has prevented the development of local monkeypox outbreaks into epidemics [44]. However, some recent data demonstrate that the efficiency of MPXV spread in human populations is growing [45, 46], which should cause concern in both the medical communities in Central and West Africa and at the WHO.

MPXV in the long-term absence of population-scale vaccination and an increased rate of human monkeypox cases can well acquire the ability to spread from human to human more efficiently, as is characteristic of VARV. If this happens, mankind will face a much more complex problem as compared to smallpox eradication. First and foremost, this will have to do with the fact that MPXV, unlike VARV, has a natural reservoir; namely, an abundant African rodent species [32].

Other zoonotic orthopoxvirus species typically cause sporadic human infections (small-scale outbreaks) with a benign outcome in most cases [6, 9, 12]. However, it is known that human infection with CPXV can lead to a generalized disease resembling smallpox with a

lethal outcome in immunodeficient individuals [47, 48].

As mentioned above, the comparative analysis of the genomes of VARV and the zoonotic orthopoxviruses pathogenic for humans has shown that CPXV has the largest genome containing all the genes characteristic of the remaining orthopoxvirus species. Part of the genes in other orthopoxviruses is broken or deleted, and individual orthopoxviruses have species-specific differences in their set of retained genes. These data support the concept of reductive evolution of orthopoxviruses, according to which the loss of genes plays an important role in the evolutionary adaptation of an ancestral virus to a certain host species, as well as in the emergence of new virus species [49, 50]. VARV, the virus most pathogenic to humans, possess the smallest genome among all orthopoxviruses. This indicates a possibility that a VARV-like virus can evolve from extant zoonotic orthopoxviruses with a longer genome as a result of natural evolution [32, 42].

An analysis of available archive data on smallpox epidemics, the history of ancient civilizations, and the most recent data on the evolutionary relationship between orthopoxviruses has allowed us to hypothesize that VARV could have repeatedly reemerged via evolutionary changes in a zoonotic ancestor virus and then disappeared because of an insufficient population size of isolated ancient civilizations [43]. Only the historically latest smallpox pandemic raged for a long time and was contained and stopped in the 20th century thanks to the joint efforts of medical professionals and scientists from many countries under the aegis of the WHO.

Therefore, the reemergence of smallpox or a similar human disease in the future in the course of a natural evolution of currently ex-

isting zoonotic orthopoxviruses is not impossible. Correspondingly, it is of utmost importance to develop and widely adopt state-of-the-art methods for an efficient and rapid species-specific diagnosis of all orthopoxvirus species pathogenic for humans, including VARV. It is also important to develop new safe methods for the prevention and therapy of human orthopoxvirus infections.

SPECIES-SPECIFIC DNA DIAGNOSTICS OF ORTHOPOXVIRUSES

The characteristics of orthopoxvirus infections are similar external manifestations, including skin lesions; however, experience has shown that clinical diagnosis of these diseases is frequently erroneous [3, 4].

The advent of the polymerase chain reaction technique has resulted in cutting-edge methods that allow for the detection and identification of trace amounts of microorganisms in assayed samples with a high specificity and over a short time [51]. Moreover, and most importantly, these methods require no manipulations with live specific pathogens, including VARV and MPXV.

In the case of orthopoxviruses pathogenic for humans, test kits that provide genus-specific DNA identification for an assayed virus with concurrent species-specific differentiation are a priority. The VECTOR team was the first to elaborate such methods based on classical multiplex PCR [52, 53] and multiplex real-time PCR [54–58].

The method that utilizes oligonucleotide microarrays is also based on PCR; in the assay, the synthesized DNA amplicons are identified by hybridization with specific oligonucleotides immobilized on a support in a particular order. The DNA preparations to be assayed in hybridization are fluorescently

labeled. After hybridization and washing, the microarray is analyzed with the help of a laser scanner and the recorded fluorescence data for each cell of the microarray are processed using specialized software. Similar to classical PCR, this method can allow one to detect trace amounts of the analyte in a specimen. One of the important advantages of oligonucleotide microarrays is the possibility to simultaneously analyze a multitude of genetic loci, thereby considerably increasing the reliability of the method [4].

Different variants of diagnostic oligonucleotide microarrays have been designed for species-specific diagnosis of orthopoxviruses [59–62].

The development of next-generation sequencing technologies would make it possible to obtain the complete genome nucleotide sequence of a research subject in short order. Genome-wide sequencing of isolated viruses in the case of unusual orthopoxvirus infections is an ever more frequent situation [63, 64]. These studies demonstrate that laboratory diagnostic techniques for orthopoxvirus infections, as well as epidemiological surveillance, need further upgrades. Naturally circulating zoonotic orthopoxviruses pathogenic for humans require a comprehensive study and monitoring for the emergence of new species that can potentially lead to the emergence of new orthopoxvirus variants highly pathogenic for humans while routine smallpox immunization is absent.

MODERN ANTI-SMALLPOX VACCINES

The first-generation smallpox vaccine was a VACV preparation produced by propagating the virus on calf (or other animal) skin. Today, VARV vaccine strains are produced in mammalian cell cultures and are referred to as second-generation

smallpox vaccines [65]. Although vaccine production in cell cultures meets current standards, second-generation smallpox vaccines, similar to first-generation ones, can cause adverse side effects and, thus, are of limited use [66].

Third-generation attenuated smallpox vaccines are produced via multiple passaging of a VACV strain in the cell culture of a heterologous host. For example, the best studied third-generation vaccine, MVA, is produced by multiple passages of the VACV strain Ankara in a chick fibroblast culture. The MVA strain genome has accumulated numerous mutations and long deletions that distinguish it from the initial VACV strain. MVA is unable to replicate in most mammalian cells, including human cells [67].

The vaccine based on the VACV strain MVA (Imvanex/Imvamune) has undergone numerous clinical trials, including studies in subjects with atopic dermatitis and HIV [68–70]. This vaccine is shown to induce an antibody profile that is similar to that induced by the conventional first-generation vaccine and to protect various laboratory animals against zoonotic orthopoxviruses [71–73]. Imvanex/Imvamune has been licensed in European countries, Canada, and the United States. First and foremost, this vaccine is intended for primary vaccination of subjects with contraindications for using first- and second-generation smallpox vaccines.

Another third-generation smallpox vaccine, LC16m8, licensed in Japan, was produced from VACV strain Lister via multiple passages in a primary rabbit kidney cell culture at a decreased temperature (30°C). Clinical studies have demonstrated a considerable reduction in the number of adverse side effects compared to the conventional Lister-based vaccine. The resulting attenuation of this vaccine strain is mainly due to a mutation (single

nucleotide deletion) in the *B5R* gene that encodes a protein essential for extracellular enveloped virion formation [74, 75]. The protective efficacy of LC16m8 in animal model experiments is comparable to that of the parental strain Lister [76, 77].

A new approach to the production of fourth-generation attenuated smallpox vaccines consists in genetic engineering of variants with impaired genes that control the host's protective response to virus infection, the range of sensitive hosts, etc. by introducing targeted deletions/insertions. The best studied variant of such VACV is strain NYVAC, with a deleted block of 12 genes and six additional individual damaged genes. The NYVAC strain induces considerably weaker immunity in humans as compared to the classical Lister or Dryvax vaccine, including the inability to induce A27-specific antibodies, which are necessary for efficient neutralization of a VACV infectious form, the intracellular mature virus [78, 79].

In Russia, a highly attenuated VACV variant was produced by successive introduction of targeted deletions/insertions into five individual genes of strain LIVP [80]. Additional targeted deletion introduced into the *A35R* gene yielded another highly immunogenic attenuated strain, VACdelta6 [81], which is currently under preclinical trials as a fourth-generation smallpox vaccine candidate. This vaccine can be used in combination with the smallpox DNA vaccine [82].

ANTI-SMALLPOX CHEMOTHERAPEUTICS

Chemotherapeutics are no less important in the treatment of human orthopoxvirus infections, and the search for such drugs over the past 20 years has been a success. Since there are no adequate animal models for smallpox, potential anti-smallpox drugs are tested in surrogate smallpox animal models

[83]. The inhibitors of orthopoxvirus reproduction were initially screened in cell cultures to further study the compounds with high *in vitro* antiviral activity using animal models, first and foremost, intranasal or aerosol infection of mice with CPXV and monkeys with MPXV [84, 85]. Rabbits infected with the rabbitpox virus (RPXV) and ground squirrels infected with MPXV have recently been actively used [86–88]. However, none of the surrogate animal models of orthopoxvirus infection precisely reproduces human smallpox. Correspondingly, the candidate compounds are examined in parallel, using several animal models.

Cidofovir, an antiviral nucleotide analog (brand name Vistide) officially approved for clinical use against cytomegalovirus retinitis and acting as an inhibitor of virus DNA polymerase, was the first compound intensively studied as an anti-orthopoxvirus drug [83]. Cidofovir proved efficient against orthopoxvirus infections in different animal models; however, its essential shortcomings are poor water solubility and mandatory intravenous administration. Correspondingly, a lipid cidofovir conjugate, CMX001 (Brincidofovir), has been synthesized [86, 89]. It is a broad-spectrum drug with pronounced anti-orthopoxvirus activity and is also administrable in tablet form.

ST-246, a compound that blocks the final stage in the assembly of intracellular enveloped virions and prevents the release of the virus from an infected cell [83, 90], is of the greatest interest. ST-246 was identified by screening a library comprising over 350 thousand unique compounds for antiviral activity. ST-246 (Tecovirimat) has shown low toxicity and high antiviral efficacy in mice infected with ectromelia virus, VACV, and CPXV; rabbits infected with RPXV and ground squirrels infected with

MPXV; and monkeys infected with MPXV or VARV [90–92]. This compound is currently undergoing clinical trials. NIOCH-14, an analog of ST-246, also showed high activity in different animal models of orthopoxvirus infections [93].

The search for new anti-orthopoxvirus chemotherapeutic agents with other molecular targets is in progress [90, 94].

CONCLUSIONS

The analysis of the genome organization of orthopoxviruses pathogenic for humans and their patterns of evolution suggest the fundamental possibility that smallpox or a similar human disease can emerge in the future via a natural evolution of extant zoonotic orthopoxviruses. Cessation of anti-smallpox vaccination and the resulting loss of collective population immunity not only to smallpox, but also to other orthopoxvirus infections creates conditions that promote the spread of zoonotic orthopoxviruses among people, thereby potentially enhancing the selection of virus variants highly pathogenic for humans and epidemically dangerous. However, the situation today does not look irredeemable and radically differs from the events far past, when man had no control over infections. Today, most outbreaks of orthopoxvirus infections in domestic animals and humans are registered and investigated; in addition, the efficient international system for clinical sampling and identification of infectious agents has been validated and anti-epidemic activities and protocols for mass vaccination were developed during the implementation of the global smallpox eradication program [1].

The recent efforts at the WHO are directed towards developing state-of-the-art methods for rapid VARV identification and designing next-generation safe anti-smallpox vaccines and chemotherapeutic

agents against VARV and other orthopoxviruses [94].

The studied vaccines and chemotherapeutics are not strictly species-specific with respect to orthopoxviruses pathogenic for humans and, thus, are applicable to outbreaks caused by any orthopoxvirus species. Taking into account the above succinct information, it results that diagnostic methods should be focused on rapid identification not only of VARV, but also of

MPXV, CPXV, VACV, and CMLV [32]. The recent increase in the number of outbreaks of orthopoxvirus infections in animals and humans and the potential danger they pose demonstrate the importance of constant monitoring of these infections all over the world aimed at insuring against the development of small outbreaks into epidemics and, thus, decreasing the risk of an emergence of a new orthopoxvirus highly pathogenic for humans.

Phenomenal advance in synthetic biology has made it possible to *de novo* synthesize the complete horsepox virus genome and obtain a live virus [95]. This suggests that any orthopoxvirus, including VARV, can be recreated in a laboratory. That is why the development and wide clinical application of the most advanced methods for the diagnosis, prevention, and therapy of orthopoxvirus infections pose a vital challenge. ●

REFERENCES

- Fenner F., Henderson D.A., Arita I., Jezek Z., Ladnyi I.D. Smallpox and Its Eradication. Geneva: World Health Organization, 1988. 1460 p.
- Ladnyi I.D. Eradication of Smallpox and Prevention of Its Recurrence. Moscow: Meditsina, 1985. 224 p.
- Marennikova S.S., Shchelkunov S.N. Orthopoxviruses Pathogenic for Humans. Moscow: KMK Scientific Press Ltd., 1998. 386 p.
- Shchelkunov S.N., Marennikova S.S., Moyer R.W. Orthopoxviruses Pathogenic for Humans. New York: Springer, 2005. 425 p.
- Di Giulio D.B., Eckburg P.B. // *Lancet Infect. Dis.* 2004. V. 4. P. 15–25.
- Essbauer S., Pfeffer M., Meyer H. // *Vet. Microbiol.* 2010. V. 140. P. 229–236.
- Bhanuprakash V., Venkatesan G., Balamurugan V., Hosamani M., Yogisharadhya R., Gandhale P., Reddy K.V., Damle A.S., Kher H.N., Chandel B.S. et al. // *Zoonoses Public Health.* 2010. V. 57. P. e149–155.
- Popova A.Y., Maksyutov R.A., Taranov O.S., Tregubchak T.V., Zaikovskaya A.V., Sergeev A.A., Vlashchenko I.V., Bodnev S.A., Ternovoi V.A., Alexandrova N.S. et al. // *Epidemiol. Infect.* 2017. V. 145. P. 755–759.
- Singh R.K., Balamurugan V., Bhanuprakash V., Venkatesan G., Hosamani M. // *Indian J. Virol.* 2012. V. 23. P. 1–11.
- Abrahao J.S., Campos R.K., Trindade G.S., Guimaraes da Fonseca F., Ferreira P.C., Kroon E.G. // *Emerg. Infect. Dis.* 2015. V. 21. P. 695–698.
- Bera B.C., Shanmugasundaram K., Barua S., Venkatesan G., Virmani N., Riyesh T., Gulati B.R., Bhanuprakash V., Vaid R.K., Kakker N.K. et al. // *Vet. Microbiol.* 2011. V. 152. P. 29–38.
- Balamurugan V., Venkatesan G., Bhanuprakash V., Singh R.K. // *Indian J. Virol.* 2013. V. 24. P. 295–305.
- Shchelkunov S.N., Marennikova S.S., Totmenin A.V., Blinov V.M., Chizhikov V.E., Gutorov V.V., Safronov P.F., Pozdnyakov S.G., Shelukhina E.M., Gashnikov P.V. et al. // *Doklady Akademii Nauk.* 1991. V. 321. P. 402–406.
- Shchelkunov S.N., Blinov V.M., Totmenin A.V., Marennikova S.S., Kolikhailov A.A., Frolov I.V., Chizhikov V.E., Gutorov V.V., Gashnikov P.V., Belanov E.F. et al. // *Molecular Biology.* 1992. V. 26. P. 731–744.
- Shchelkunov S.N., Marennikova S.S., Blinov V.M., Resenchuk S.M., Totmenin A.V., Chizhikov V.E., Gutorov V.V., Safronov P.F., Kurmanov R.K., Sandakhchiev L.S. // *Doklady Akademii Nauk.* 1993. V. 328. P. 629–632.
- Shchelkunov S.N., Resenchuk S.M., Totmenin A.V., Blinov V.M., Marennikova S.S., Sandakhchiev L.S. // *FEBS Lett.* 1993. V. 327. P. 321–324.
- Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P., Paoletti E. // *Virology.* 1990. V. 179. P. 247–266, 517–563.
- Shchelkunov S.N. // *Virus Genes.* 1995. V. 10. P. 53–71.
- Shchelkunov S.N., Marennikova S.S., Blinov V.M., Totmenin A.V., Chizhikov V.E., Netesov S.V., Andzhaparidze O.G., Sandakhchiev L.S. // In: Poxviruses and Iridoviruses. Abstr. of the 9th International Conference. Les Diablerets. Switzerland, 1992. P. 31.
- Massung R.F., Liu L.I., Qi J., Knight J.C., Yuran T.E., Kerlavage A.R., Parsons J.M., Venter J.C., Esposito J.J. // *Virology.* 1994. V. 201. P. 215–240.
- Shchelkunov S.N., Massung R.F., Esposito J.J. // *Virus Res.* 1995. V. 36. P. 107–118.
- Massung R.F., Loparev V.N., Knight J.C., Totmenin A.V., Chizhikov V.E., Parsons J.M., Safronov P.F., Gutorov V.V., Shchelkunov S.N., Esposito J.J. // *Virology.* 1996. V. 221. P. 291–300.
- Shchelkunov S.N., Totmenin A.V., Loparev V.N., Safronov P.F., Gutorov V.V., Chizhikov V.E., Knight J.C., Parsons J.M., Massung R.F., Esposito J.J. // *Virology.* 2000. V. 266. P. 361–386.
- Shchelkunov S.N., Safronov P.F., Totmenin A.V., Petrov N.A., Ryazankina O.I., Gutorov V.V., Kotwal G.J. // *Virology.* 1998. V. 243. P. 432–460.
- Shchelkunov S.N., Totmenin A.V., Babkin I.V., Safronov P.F., Ryazankina O.I., Petrov N.A., Gutorov V.V., Uvarova E.A., Mikheev M.V., Sisler J.R. et al. // *FEBS Lett.* 2001. V. 509. P. 66–70.
- Shchelkunov S.N., Totmenin A.V., Safronov P.F., Mikheev M.V., Gutorov V.V., Ryazankina O.I., Petrov N.A., Babkin I.V., Uvarova E.A., Sandakhchiev L.S. et al. // *Virology.* 2002. V. 297. P. 172–194.
- Shchelkunov S.N., Totmenin A.V. // *Virus Genes.* 1995. V. 9. P. 231–245.
- Uvarova E.A., Shchelkunov S.N. // *Virus Res.* 2001. V. 81. P. 39–45.
- Shchelkunov S., Totmenin A., Kolosova I. // *Virus Genes.* 2002. V. 24. P. 157–162.
- Shchelkunov S.N. // *Virus Genes.* 2010. V. 41. P. 309–318.

31. Shchelkunov S.N. // *Adv. Virol.* 2012. V. 2012. Article ID 524743.
32. Shchelkunov S.N. // *PLoS Pathog.* 2013. V. 9. P. e1003756.
33. Upton C., Slack S., Hunter A.L., Ehlers A., Roper R.L. // *J. Virol.* 2003. V. 77. P. 7590–7600.
34. McLysaght A., Baldi P.F., Gaut B.S. // *Proc. Natl. Acad. Sci. USA.* 2003. V. 100. P. 15655–15660.
35. Gubser C., Hue S., Kellam P., Smith G.L. // *J. Gen. Virol.* 2004. V. 85. P. 105–117.
36. Xing K., Deng R., Wang J., Feng J., Huang M., Wang X. // *Intervirology.* 2006. V. 49. P. 207–214.
37. Babkina I.N., Babkin I.V., Li Y., Ropp S., Kline R., Damon I., Esposito J., Sandakhchiev L.S., Shchelkunov S.N. // *Dokl. Biochem. Biophys.* 2004. V. 398. P. 316–320.
38. Babkin I.V., Shchelkunov S.N. // *Molecular Biology.* 2006. V. 40. P. 16–19.
39. Esposito J.J., Sammons S.A., Frace A.M., Osborne J.D., Olsen-Rasmussen M., Zhang M., Govil D., Damon I.K., Kline R., Laker M. et al. // *Science.* 2006. V. 313. P. 807–812.
40. Babkin I.V., Nepomnyashchikh T.S., Maksyutov R.A., Gutorov V.V., Babkina I.N., Shchelkunov S.N. // *Molecular Biology.* 2008. V. 42. P. 543–555.
41. Babkin I.V., Shchelkunov S.N. // *Russ. J. Genet.* 2008. V. 44. P. 895–908.
42. Shchelkunov S.N. // *Arch. Virol.* 2009. V. 154. P. 1865–1871.
43. Shchelkunov S.N. // *Vaccine.* 2011. V. 29S. P. D49–53.
44. Breman J.G. // *Emerg. Infect.* 2000. V. 4. P. 45–67.
45. Rimoin A.W., Mulembakani P.M., Johnston S.C., Lloyd Smith J.O., Kisalu N.K., Kinkela T.L., Blumberg S., Thomassen H.A., Pike B.L., Fair J.N. et al. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. P. 16262–16267.
46. Reynolds M.G., Carroll D.S., Karem K.L. // *Curr. Opin. Virol.* 2012. V. 2. P. 335–343.
47. Czerny C.P., Eis-Hubinger A.M., Mayr A., Schneweis K.E., Pfeiff B. // *J. Vet. Med.* 1991. V. 1338. P. 421–431.
48. Fassbender P., Zange S., Ibrahim S., Zoeller G., Herbstreit F., Meyer H. // *Emerg. Infect. Dis.* 2016. V. 22. P. 553–555.
49. Hendrickson R.C., Wang C., Hatcher E.L., Lefkowitz E.J. // *Viruses.* 2010. V. 2. P. 1933–1967.
50. Coulson D., Upton C. // *Virus Genes.* 2011. V. 42. P. 171–177.
51. Mullis K.B., Faloona F. // *Meth. Enzymol.* 1987. V. 155. P. 335–350.
52. Gavrilova E.V., Babkin I.V., Shchelkunov S.N. // *Molekular. Genet. Microbiol. Virusol.* 2003. № 1. P. 45–52.
53. Shchelkunov S.N., Gavrilova E.V., Babkin I.V. // *Mol. Cell. Probes.* 2005. V. 19. P. 1–8.
54. Olson V.A., Laue T., Laker M.T., Babkin I.V., Drosten C., Shchelkunov S.N., Niedrig M., Damon I.K., Meyer H. // *J. Clin. Microbiol.* 2004. V. 42. P. 1940–1946.
55. Kostina E.V., Gavrilova E.V., Ryabini V.A., Shchelkunov S.N., Sinyakov A.N. // *Vopr. Virusol.* 2009. V. 54. P. 28–33.
56. Shchelkunov S.N., Shcherbakov D.N., Maksyutov R.A., Gavrilova E.V. // *J. Virol. Meth.* 2011. V. 175. P. 163–169.
57. Shcherbakov D.N., Gavrilova E.V., Maksyutov R.A., Shchelkunov S.N. // *Klinich. Laborat. Diagnost.* 2011. № 12. P. 39–42.
58. Maksyutov R.A., Gavrilova E.V., Shchelkunov S.N. // *J. Virol. Methods.* 2016. V. 236. P. 215–220.
59. Lapa S., Mikheev M., Shchelkunov S., Mikhailovich V., Sobolev A., Blinov V., Babkin I., Guskov A., Sokunova E., Zasedatelev A. et al. // *J. Clin. Microbiol.* 2002. V. 40. P. 753–757.
60. Mikheev M.V., Lapa S.A., Shchelkunov S.N., Chikova A.K., Mikhailovich M.V., Sobolev A.Y., Babkin I.V., Gryadunov D.A., Bulavkina M.A., Guskov A.A. et al. // *Vopr. Virusol.* 2003. V. 48. P. 4–9.
61. Laassri M., Chizhikov V., Mikheev M., Shchelkunov S., Chumakov K. // *J. Virol. Methods.* 2003. V. 112. P. 67–78.
62. Ryabini V.A., Shundrin L.A., Kostina E.B., Laassri M., Chizhikov V., Shchelkunov S.N., Chumakov K., Sinyakov A.N. // *J. Med. Virol.* 2006. V. 78. P. 1325–1340.
63. Mavian C., Lopez-Bueno A., Alcamí A. // *Genome Announc.* 2014. V. 2. P. e01086–13.
64. Vora N.M., Li Y., Geleishvili M., Emerson G.L., Khmaladze E., Maghlakelidze G., Navdarashvili A., Zakhshvili K., Kokhreizidze M., Endeladze M. et al. // *N. Engl. J. Med.* 2015. V. 372. P. 1223–1230.
65. Sanchez-Sampedro L., Perdiguero B., Mejias-Perez E., Garcia-Arriaza J., Di Pilato M., Esteban M. // *Viruses.* 2015. V. 7. P. 1726–1803.
66. Frey S.E., Newman F.K., Kennedy J.S., Ennis F., Abate G., Hoft D.F., Monath T.P. // *Vaccine.* 2009. V. 10. P. 1637–1644.
67. Meyer H., Sutter G., Mayr A. // *J. Gen. Virol.* 1991. V. 72. P. 1031–1038.
68. Sonnenburg F., Perona P., Darsow U., Ring J., von Krempelhuber A., Vollmar J., Roesch S., Baedeker N., Kollaritsch H., Chaplin P. // *Vaccine.* 2014. V. 32. P. 5696–5702.
69. Zitzmann-Roth E.M., von Sonnenburg F., de la Motte S., Arndt-Wiedemann N., von Krempelhuber A., Uebler N., Vollmar J., Virgin G., Chaplin P. // *PLoS One.* 2015. V. 10. P. e0122653.
70. Greenberg R.N., Hay C.M., Stapleton J.T., Marbury T.C., Wagner E., Kreitmeier E., Roesch S., von Krempelhuber A., Young P., Nichols R. et al. // *PLoS One.* 2016. V. 11. P. e0157335.
71. Earl P.L., Americo J.L., Wyatt L.S., Eller L.A., Whitbeck J.C., Cohen G.H., Eisenberg R.J., Hartmann C.J., Jackson D.L., Kulesh D.A. et al. // *Nature.* 2004. V. 428. P. 182–185.
72. Jones D.I., McGee C.E., Sample C.J., Sempowski G.D., Pickup D.J., Staats H.F. // *Clin. Vaccine Immunol.* 2016. V. 23. P. 648–651.
73. Volz A., Sutter G. // *Adv. Virus Res.* 2017. V. 97. P. 187–243.
74. Kidokoro M., Tashiro M., Shida H. // *Proc. Natl. Acad. Sci. USA.* 2005. V. 102. P. 4152–4157.
75. Eto A., Saito T., Yokote H., Kurane I., Kanatani Y. // *Vaccine.* 2015. V. 33. P. 6106–6111.
76. Empig C., Kenner J.R., Perret-Gentil M., Youree B.E., Bell E., Chen A., Gurwith M., Higgins K., Lock M., Rice A.D. et al. // *Vaccine.* 2006. V. 24. P. 3686–3694.
77. Yokote H., Shinmura Y., Kanehara T., Maruno S., Kuranaga M., Matsui H., Hashizume S. // *Vaccine.* 2015. V. 33. P. 6112–6119.
78. Tartaglia J., Perkus M.E., Taylor J., Norton E.K., Audonnet J.C., Cox W.L., Davis S.W., van der Hoeven J., Meignier B., Riviere M. et al. // *Virology.* 1992. V. 188. P. 217–232.
79. Midgley C.M., Putz M.M., Weber J.N., Smith G.L. // *J. Gen. Virol.* 2008. V. 89. P. 2992–2997.
80. Yakubitskiy S.N., Kolosova I.V., Maksyutov R.A., Shchelkunov S.N. // *Acta Naturae.* 2015. V. 7. P. 113–121.
81. Yakubitskiy S.N., Kolosova I.V., Maksyutov R.A., Shchelkunov S.N. // *Dokl. Biochem. Biophys.* 2016. V. 466. P. 35–38.
82. Maksyutov R.A., Yakubitskiy S.N., Kolosova I.V., Shchelkunov S.N. // *Acta Naturae.* 2017. V. 9. P. 88–93.
83. Smee D.F. // *Antivir. Chem. Chemother.* 2008. V. 19. P. 115–124.
84. Bray M., Martinez M., Smee D.F., Kefauer D., Thompson E., Huggins J.W. // *J. Infect. Dis.* 2000. V. 181. P. 10–19.

85. Zaucha G.M., Jahrling P.B., Geisbert T.W., Swearingen J.R., Hensley L. // *Lab. Invest.* 2001. V. 81. P. 1581–1600.
86. Rice A.D., Adams M.M., Wallace G., Burrage A.M., Lindsey S.F., Smith A.J., Swetnam D., Manning B.R., Gray S.A., Lambert B. et al. // *Viruses*. 2011. V. 3. P. 47–62.
87. Sbrana E., Jordan R., Hruba D.E., Mateo R.I., Xiao S.Y., Siirin M., Newman P.C., Da Rosa A.P., Tesh R.B. // *Am. J. Trop. Med. Hyg.* 2007. V. 76. P. 768–773.
88. Sergeev A.A., Kabanov A.S., Bulychev L.E., Sergeev A.A., Pyankov O.V., Bodnev S.A., Galahova D.O., Zamedyanskaya A.S., Titova K.A., Glotova T.I. et al. // *Transbound. Emerg. Dis.* 2017. V. 64. P. 226–236.
89. Parker S., Crump R., Foster S., Hartzler H., Hembrador E., Lanier E.R., Painter G., Schriewer J., Trost L.C., Buller R.M. // *Antiviral Res.* 2014. V. 111. P. 42–52.
90. Smee D.F. // *Future Virol.* 2013. V. 8. P. 891–901.
91. Mucker E.M., Goff A.J., Shamblin J.D., Grosenbach D.W., Damon I.K., Mehal J.M., Holman R.C., Carroll D., Gallardo N., Olson V.A. et al. // *Antimicrob. Agents Chemother.* 2013. V. 57. P. 6246–6253.
92. Berhanu A., Prigge J.T., Silvera P.M., Honeychurch K.M., Hruba D.E., Grosenbach D.W. // *Antimicrob. Agents Chemother.* 2015. V. 59. P. 4296–4300.
93. Mazurkov O.Y., Kabanov A.S., Shishkina L.N., Sergeev A.A., Skarnovich M.O., Bormotov N.I., Skarnovich M.A., Ovchinnikova A.S., Titova K.A., Galahova D.O. et al. // *J. Gen. Virol.* 2016. V. 97. P. 1229–1239.
94. Scientific Review of Variola Virus Research, 1999–2010. Geneva: World Health Organization, 2010. 128 p.
95. WHO Advisory Committee on Variola Virus Research: Report of the Eighteenth Meeting. Geneva: World Health Organization, 2017. 58 p.

Non-coding RNAs As Transcriptional Regulators In Eukaryotes

O. Y. Burenina^{1,2*}, T. S. Oretskaya^{2,3}, E. A. Kubareva³

¹Skolkovo Institute of Science and Technology, Nobel Str. 3, Moscow, 143026, Russia

²Lomonosov Moscow State University, Chemistry Department, Leninskie Gory 1, bld. 3, Moscow, 119991, Russia

³Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Leninskie Gory 1, bld. 40, Moscow, 119991, Russia

*E-mail: alunit@inbox.ru

Received: November 02, 2016; in final form August 30, 2017

Copyright © 2017 Park-media, Ltd. 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 Non-coding RNAs up to 1,000 nucleotides in length are widespread in eukaryotes and fulfil various regulatory functions, in particular during chromatin remodeling and cell proliferation. These RNAs are not translated into proteins: thus, they are non-coding RNAs (ncRNAs). The present review describes the eukaryotic ncRNAs involved in transcription regulation, first and foremost, targeting RNA polymerase II (RNAP II) and/or its major proteinaceous transcription factors. The current state of knowledge concerning the regulatory functions of SRA and TAR RNA, 7SK and U1 snRNA, GAS5 and DHFR RNA is summarized herein. Special attention is given to murine B1 and B2 RNAs and human Alu RNA, due to their ability to bind the active site of RNAP II. Discovery of bacterial analogs of the eukaryotic small ncRNAs involved in transcription regulation, such as 6S RNAs, suggests that they possess a common evolutionary origin.

KEYWORDS noncoding RNAs; RNA polymerase; transcription regulation.

ABBREVIATIONS ncRNA – noncoding RNA; nt – nucleotide residues; PIC – preinitiating complex; RNAP – RNA polymerase; RNP – ribonucleoprotein; snRNA – small nuclear RNA; SINE – short interspersed elements.

INTRODUCTION

According to transcriptome analysis, only 1.5% of the total amount of RNAs in eukaryotic cells encodes proteins, while other transcripts are non-coding (ncRNAs). Apparently, the "repertoire" of genes that encode proteins has remained relatively static in the course of evolution, and the number of ncRNA genes has increased when proceeding to more complex organisms. Ribosomal, transfer, small nuclear, and small nucleolar ncRNAs, which are constantly expressed in the cells, are conventionally classified as housekeeping ncRNAs, by analogy with the name of the most important cellular genes [1]. However, the majority of ncRNAs fulfil regulatory functions and participate in the equally important and often conversely directed molecular processes such as DNA demethylation and imprinting, activation and repression of gene transcription, as well as chromatin remodeling, RNA interference, and alternative splicing [2–4]. The level of synthesis of many ncRNAs varies under different stress conditions, during cancer and neurologic diseases [5, 6]. ncRNAs play a major role in cell differentiation [7]. Taking into account the fact that this is only a small part of the currently known properties and functions of ncRNAs, one can

assume that their contribution to the maintenance of a normal functioning of the cell is no less significant than the contribution of protein factors.

Usually ncRNAs are classified into short (~20–30 nt), which include microRNAs (miRs), small interfering (siRNAs), and PIWI-interacting RNAs (P-element-induced wimpy testis, piRNA) [8]; small ncRNAs up to 200 nt; and long ncRNAs (> 200 nt). Among small ncRNAs, promoter-associated RNAs (paRNAs) are the most well-known, although this class includes representatives of various lengths [9]. The term "long non-coding RNAs" (lncRNAs) is widely used for the transcripts that are several thousand nucleotides in length and belong to long intergenic ncRNAs (lincRNAs) and enhancer RNAs (eRNAs) [10]. However, there are also extremely lengthy ncRNAs, consisting of several hundred thousand nucleotides, such as very long intergenic ncRNAs (vlincRNAs) and macroRNAs [11].

Considering the diversity of the classes and functions of ncRNAs, it is no surprise that many of them are involved in the regulation of transcription in eukaryotes. This occurs primarily through various epigenetic mechanisms; in particular, chromatin remodeling (this area of ncRNA functioning has been studied

much better than others) [12, 13]. The well-known examples of such ncRNAs include XIST RNA (X-inactive specific transcript), roX RNA, HOTAIR (Hox transcript antisense intergenic RNA); enhancer RNAs NRIP1, GREB1, KLK; and NEAT1 RNA (nuclear enriched abundant transcript 1), which is responsible for paraspeckle formation in tumor cell nuclei. Regulation of co-transcriptional splicing involves MALAT1 RNA (metastasis-associated lung adenocarcinoma transcript 1) and H19 RNA, which serve as therapeutic targets for various diseases, including cancer [14]. In addition, there are ncRNAs that interact with RNA polymerase II (RNAP II) or the transcription factors incorporated in the preinitiating complex (PIC) or elongation complex. The latter include 7SK small nuclear RNA (snRNA) and TAR RNA, which regulate the activity of the transcription elongation factor P-TEFb; U1 snRNA, which interacts with the initiation factor TFIIF; SRA RNA that activates steroid receptors, and some others (Fig. 1). These ncRNAs are involved in complex multistep regulatory mechanisms and usually interact with cascades of proteins, indirectly affecting a transcription process. On the contrary, the murine B1 and B2 RNAs and human Alu RNA encoded by mobile genetic elements (SINE) are able to directly bind RNAP II [15]. To date the X-ray data of their complexes with the enzyme has not been obtained. FC RNA, the synthetic aptamer consisting of two short hairpins, is the only ncRNA whose complex with RNAP II has been solved by X-ray analysis [16]. Since the secondary structure of almost all of the aforementioned regulatory ncRNAs includes short hairpin elements, which interact with the active site of RNAP II, they are often regarded as aptamers for the enzyme [17].

According to the conventional criteria, TAR RNA, B1 RNA, B2 RNA, and U1 snRNA should be classified as small ncRNAs, while Alu RNA, 7SK snRNA, DHFR RNA, SRA RNA, and GAS5 RNA are lncRNAs. The specific properties and functions of each of these ncRNAs are discussed in detail in this review. The structural features of the interaction between FC RNA and RNAP II are comprehensively described in [16].

REGULATORY RNAs ENCODED BY GENETIC ELEMENTS OF THE SINE FAMILY

SINE (short interspersed elements) are retrotransposons 80 to 500 bp in length located randomly in the genome of higher eukaryotes. The nucleotide sequences of SINE characterized by 65–90% similarity are clustered into families, and the number of homologous SINE can vary from 10^3 to 10^6 copies per cell [18]. Historically, SINE had been considered as “genetic garbage” used to establish phylogenetic relationships and study speciation in mammals until it was found that

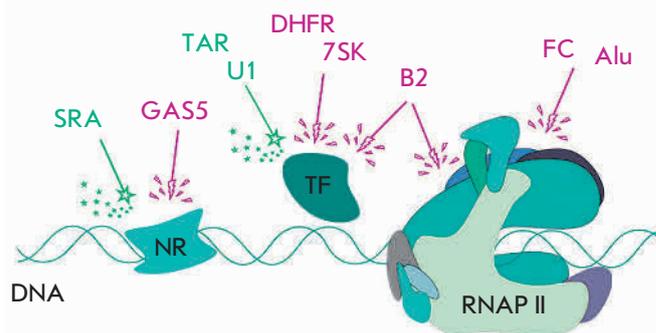


Fig. 1. The best known ncRNAs acting as transcriptional activators (green) or inhibitors (purple) via interactions with RNAP II and/or its general transcription factors (TF) or with other regulatory proteins, in particular nuclear receptors (NR).

transcription of SINE-“genes” is activated in cells in response to heat shock [19]. It is believed that this is due to the increased accessibility of SINE for transcription in chromatin remodeling and activation of the transcription factor TFIIC binding to the promoter regions of SINE. As it turned out, SINE are involved in the regulation of gene expression, localization of mRNA, and they can act as enhancers or mobile promoters for RNAP II [20]. To date, it is known that SINE do not encode proteins and are transcribed by RNAP III into the corresponding SINE RNAs. Unexpectedly, it was discovered that some SINE RNAs are capable of binding RNAP II and inhibiting transcription. The main results were obtained for murine B1 and B2 RNAs and human Alu RNA [14, 21]. Exposure to UV and γ -radiation, viral infections, ethanol, antibiotics and anticancer drugs induces an increase in the expression level of these ncRNAs in cells [14]. These data certainly suggest an important functional role for B1, B2, and Alu RNAs in cell life.

Human Alu RNA and murine B1 RNA

SINE element Alu was named due to the presence of recognition sites of restriction endonuclease from *Arthrobacter luteus* (R.AluI). The human genome contains more than 1 million copies of Alu encoding Alu RNAs, which accounts for about 10.6% of nuclear DNA. B1 RNA-encoding SINE are more rare in the murine genome, less than 550,000 per cell. Both of these RNAs belong to the family of retroelements of small cytoplasmic 7SL RNA [22] and possess a similar secondary structure (Fig. 2). The full-length Alu RNA, sized ~280 nt in length, is a tandem repeat of two B1-like elements connected by a 20-nt A-rich linker. Alu RNA processing produces scAlu RNA of 118 nt in length, which is localized in the cytoplasm and is a complete

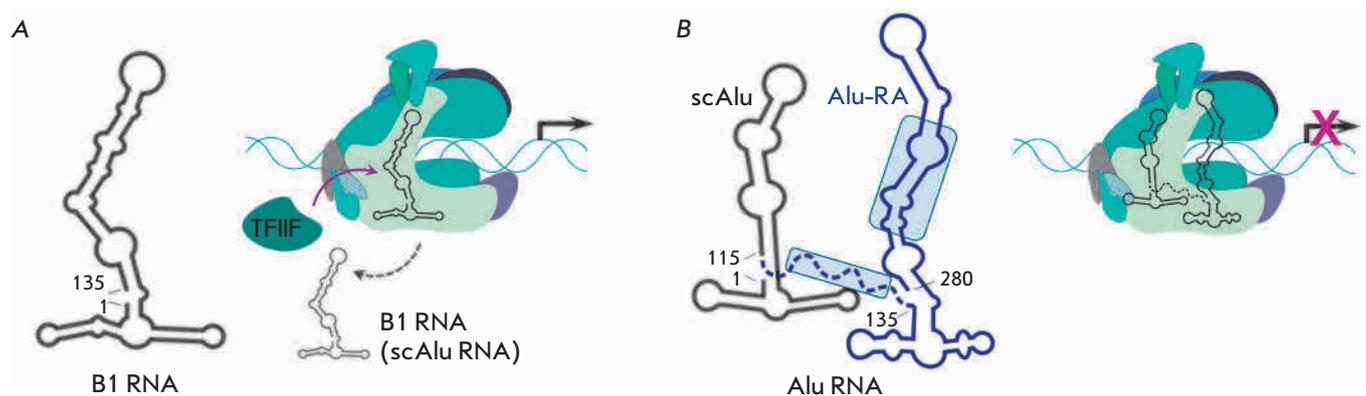


Fig. 2. Scheme of the functioning of murine B1 RNA (**A**) and human Alu RNA (**B**). Secondary structures of ncRNAs are schematically shown on the left. Alu RNA structural elements responsible for transcription inhibition are in light-blue frames, the functional domain (Alu-RA) is colored in blue, the A-rich linker is shown by a dash line. A schematic view of interactions between RNAP II and B1, scAlu, or Alu RNA is shown on the right. Transcription is indicated by a black arrow. B1 and scAlu RNA are displaced by TFIIF from their complexes with RNAP II, thus, they are unable to inhibit transcription, in contrast to Alu RNA.

analog of murine B1 RNA (*Fig. 2*) [23]. Alu RNA has an unusual shape, hence its structured parts were named “left arm” (identical to scAlu RNA) and “right arm” (Alu-RA, 135–280 nt of Alu RNA). Each Alu RNA domain can bind one RNAP II molecule, but only interaction of Alu-RA (or full-sized Alu RNA) with the enzyme results in the inhibition of transcription. Murine B1 RNA, despite its high affinity to RNAP II, cannot affect transcription (*Fig. 2A*), although chimeric RNA consisting of B1 RNA and Alu-RA has all the properties of the full-size Alu RNA [24, 25].

Besides the two RNAP binding domains located in the “left arm” and “right arm,” Alu RNA has two domains disposed in the central region of the “right arm” and in the A-rich linker which are responsible for transcription inhibition (*Fig. 2B*). Correspondingly, B1 RNA and scAlu RNA possess only a RNAP binding domain. According to cryoelectron microscopy, both Alu and B1 RNA interact with the “clamp” domain of RNAP II near the active site of the enzyme [26]. So how does the transcription occur in case of non-functional B1 and scAlu RNAs? It has been shown that RNAP II is released from its complexes with B1 and scAlu RNAs under the action of the transcription factor TFIIF, causing a dissociation of these ncRNAs from PIC, while Alu RNA remains bound to the polymerase (*Fig. 2*). At the same time, no direct contact between TFIIF and B1 or scAlu RNA was detected [27]. The disruption of RNA-protein contacts is likely to occur during conformational changes in RNAP caused by TFIIF binding. Since *in vivo* TFIIF is usually associated with RNAP II prior to PIC assembly on the promoter, “useless” binding of non-regulatory ncRNAs (having no effect

on the transcriptional activity of RNAP) probably does not occur.

The precise mechanism of interaction between Alu RNA and PIC is not fully understood. Transcription inhibition was observed *in vitro* only when Alu RNA was added prior to the initiation of transcription from the promoter, although the efficiency of abortive transcript synthesis in the presence of Alu RNA was ~ ten-fold lower. At the same time, electrophoretic mobility shift assay has shown that Alu RNA comigrates together with the DNA as a part of PIC RNAP II [23]. Thus, transcription is inhibited not owing to competition with DNA, but rather as a result of altered enzyme activity due to the formation of specific ncRNA-protein contacts. However, Alu RNA cannot stop active transcription and fulfils its functions before the initiation step.

Murine B2 RNA

B2 RNA is transcribed by RNAP III in the presence of TFIIB and TFIIC factors from the respective B2 SINE (belonging to the tRNA^{Ala} retroelement family), whose number is estimated to be about 10^5 copies per cell [28]. This RNA can be isolated in complex with RNAP II during immunoprecipitation of the nuclear extracts of cells exposed to heat shock [29], and it is capable of inhibiting transcription *in vitro* [25]. Knockdown of B2 RNA in murine cells leads to increased expression of actin and hexokinase II, whereas their genes are repressed under heat shock conditions [24]. Increase in the amount of B2 RNA has been observed during cell response to various stress factors, as well as in embryonic and tumor cells [30]. Thus, the important role of this ncRNA as a transcription inhibitor is undoubted.

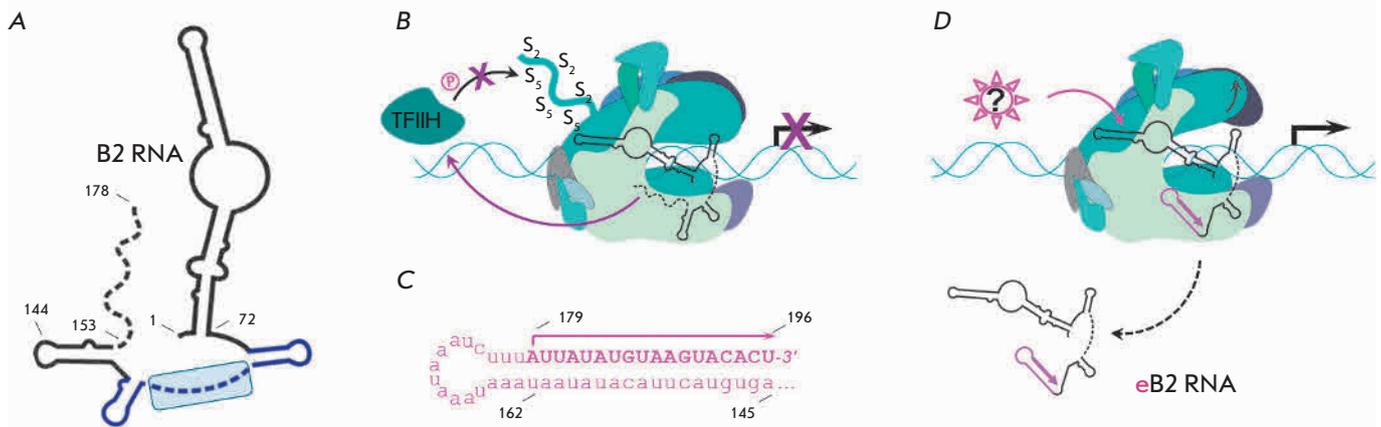


Fig. 3. Scheme of the functioning of murine B2 RNA. **(A)** Simplified view of the B2 RNA secondary structure. Disordered parts are shown by a dash line; functional part – by a blue line; inhibitory domain is in a light-blue frame. **(B)** B2 RNA prevents initiation of transcription by “switching off” TFIIF kinase activity. The Ser2 and Ser5 amino acid residues are marked as “S₂” and “S₅,” phosphorylation – as “P” in a circle. **(C)** The sequence of the 3'-end fragment (145–178 nt) of B2 RNA extended by additional 18 nt [34]. **(D)** Elongation of B2 RNA 3'-end (eB2 RNA) leads to the formation of a new hairpin (pink), which causes conformational changes in RNAP II (shown by a grey arrow) and dissociation of B2 RNA. Transcription on a DNA template is indicated by a black arrow, transcription on a B2 RNA template – by a pink arrow.

Unfortunately, there is only scarce data on the nature of B2 RNA functioning *in vivo*, while, the mechanism of action of this ncRNA has been studied in detail.

Murine cells contain at least four variants of B2 transcripts of different lengths: ~150, ~180, ~240 and ~500 nt. The two longest variants are very stable ($\tau_{1/2} = 60$ min) due to polyadenylation, whereas the degradation time of a 180 nt-long transcript is just 3–4 minutes. The shortest 150-nt variant of B2 RNA is more stable and characterized by a $\tau_{1/2}$ value of 20 min [31]. In 2004, the secondary structure of the ~180 nt-long transcript was determined [25], consisting of three nominal fragments (Fig 3A): (1) a long double-stranded region (1–72 nt) with unwound moiety in the center; (2) a poorly structured region (73–153 nt) containing three small hairpins, and (3) a short 3'-terminal unstructured AU-rich region (154–178 nt) which is conserved in all SINE.

Footprinting studies have shown that RNAP II binds to the least structured part of the molecule (73–155 nt), and a 5'-terminal hairpin is required neither for B2 RNA binding to RNAP II nor for transcription repression. Analysis of various deletion mutants of B2 RNA has determined the 51 nt-long region (81–131 nt) which directly interacts with RNAP II and inhibits transcription *in vitro* with the same efficiency as full-length B2 RNA [31]. Notably, the unstructured region of B2 RNA (99–115 nt) flanked by two hairpins plays the most important role in transcription inhibition. Removal of any

of them leads to a loss of inhibitory activity of a fully functional deletion derivative of B2 RNA (81–131 nt). At the same time, absence of these hairpins in full-length B2 RNA has no effect on its properties. Besides, all these deletion mutants of B2 RNA demonstrated specific binding to RNAP during PIC assembly on the promoter [31, 32]. Therefore, transcriptional repression requires correct positioning of a single-stranded region (99–115 nt) of B2 RNA complexed with RNAP II, which, apparently, can be achieved by any of the present hairpins.

The similarity of the structural organization of B2 RNA and Alu RNA indicates that, besides the active site (whose blocking leads to global inhibition of mRNA synthesis), RNAP II contains an additional docking site highly specific to ncRNAs. As in the case of Alu RNA, RNAP II forms a ternary complex *in vitro* with B2 RNA and the promoter simultaneously [25]. Therefore, B2 RNA can also bind to RNAP after the formation of a stable complex with the promoter and inhibit transcription at the stage of initiation. This disables not only the synthesis of full-length mRNA, but also abortive transcripts. Experiments on crosslinking and footprinting of PIC associated with B2 RNA have shown that this ncRNA hinders proper coordination of the promoter in the active site of the polymerase and, thereby, switches PIC into its inert form. In fact, B2 RNA alters the conformation of the “closed” RNAP complex and prevents its conversion into an “open” complex and,

all the more, into an initiation complex. At the same time, all factors associated with PIC, including TBP and TFIIB, remain bound to the promoter and hold the complex on the DNA [25].

Let us recall that murine cells also express B1 RNA binding to RNAP II, but are incapable of inhibiting transcription. B1 RNA possesses affinity to polymerase comparable to that of B2 RNA and can displace B2 RNA from PIC. Therefore, B1 RNA must prevent the functioning of B2 RNA. However, it has been shown *in vitro* that B2 RNA can inhibit transcription even in the case when PIC had been previously bound to B1 RNA [27]. The mechanisms of competition between these two ncRNAs *in vivo* have not been established. Since human Alu RNA also has a non-functional analogue, it can be assumed that these inactive ncRNAs, B1 and scAlu RNA, in certain circumstances may replace B2 and Alu RNA, respectively, and re-stimulate transcription.

Interestingly, in addition to direct “physical” RNAP II active site blocking, B2 RNA specifically inhibits the kinase activity of the TFIIH transcription factor (*Fig. 3B*). TFIIH contains cyclin-dependent kinase 7 (CDK7), which, under normal conditions, phosphorylates serine residues within the heptapeptide repeats YSPTSPS (mainly Ser5) at the C-terminal domain (CTD) of the large subunit (Rpb1) of RNAP II. Modification of Ser2 and Ser5 at the CTD of Rpb1 is extremely important for transcription. It occurs at various stages of transcription: the domain is not phosphorylated in the initiation complex and, on the contrary, hyperphosphorylated during transcription elongation [33]. Thus, B2 RNA not only creates conformational constraint in RNAP II itself, but also disables the proceeding to the elongation stage, affecting the functioning of the transcription factor. Although TFIIH is not the primary target of B2 RNA and its repression is likely due to the interaction between B2 RNA and PIC, this is a unique phenomenon for ncRNAs.

The more surprising fact is that B2 RNA can promote self-elongation in a complex with RNAP II [34]. The enzyme uses the 3'-end of the B2 RNA molecule as a template for *de novo* transcription and synthesizes 18 additional nucleotide residues, forming a stable extended hairpin (*Fig. 3C, D*). Elongation of the B2 RNA leads to the dissociation of the molecule from PIC and, apparently, enables reversibility of inhibition. The released extended B2 RNA undergoes degradation. An analysis of computer modeling data suggests that the elongation of the B2 RNA strand (or any other RNA located in the active site of the polymerase) should lead to the partial opening of the “clamp” domain of RNAP and, as a consequence, the weakening of ligand binding to the enzyme. In fact, a newly formed structural ele-

ment of the elongated B2 RNA “extrudes” the molecule from PIC. Notably, B2 RNA elongation was observed *in vitro* only after treatment of the B2 RNA-RNAP II complex with the cell extract [34]. It is believed that the RNA-dependent transcription of RNAP II is initiated by a protein factor whose nature is yet unknown.

Since most RNA polymerases are DNA-dependent (except for retroviral RNAPs), elongation of B2 RNA is a kind of exception to the rule, since the enzyme modifies its substrate specificity. To date, there are only a few such examples, which are also related to the functioning of ncRNAs. For instance, the same mechanism is used by the hepatitis delta virus and plant viroids to replicate their own genome. These pathogenic circular ncRNAs lack inherent RNA polymerases and use host cell RNAPs, reprogramming them for RNA synthesis on RNA templates [35]. This phenomenon might be more clearly exemplified by prokaryotic 6S RNA, which inhibits transcription due to interaction with RNAP, similarly to B2 RNA. Under certain conditions, bacterial RNAP can synthesize short transcripts (pRNAs) up to 30 nt in length on 6S RNA as a template. In this case, the enzyme dissociates from the complex with 6S RNA and resumes transcription from gene promoters [36]. Therefore, despite the huge differences in the transcription processes in prokaryotes and eukaryotes, there is undoubted similarity between the functioning of bacterial 6S RNA and murine B2 RNA.

NON-CODING RNAs REGULATING THE ACTIVITY OF GENERAL TRANSCRIPTION FACTORS

U1 snRNA

U1 snRNA is one of the five major snRNAs forming the core of the spliceosome. Human U1 snRNA, 164 nt in length, is associated with U1-A, U1-C, and U1-70k proteins, as well as with eight Sm proteins, together forming the U1 snRNP complex (~245 kDa). The main function of U1 snRNP is to recognize pre-mRNA at the first (initial) stage of spliceosome assembly, which occurs due to complementary interactions between the 5'-end region of U1 snRNP and the intron splice site [37]. However, in addition to its primary role, U1 snRNA can interact with cyclin H (CycH) as a part of TFIIH, which in turn leads to an increased kinase activity of another subunit of this factor, CDK7 (*Fig. 4A*). Studies of *in vitro* transcription showed that the presence of U1 snRNA in the reaction mixture increases the rate of formation of the first phosphodiester bond and that the efficiency of transcription initiation increases more than tenfold. Furthermore, U1 snRNA stimulates abortive initiation and re-initiation of transcription from the promoter preceding the 5'-terminal splicing site [38]. Besides TFIIH, U1 snRNA may interact with other

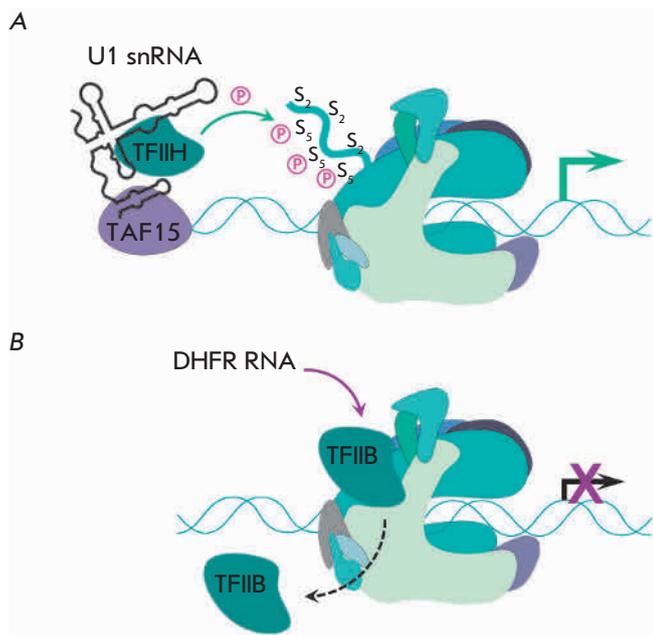


Fig. 4. Scheme of the functioning of U1 snRNA (**A**) and DHFR ncRNA (**B**). The simplified secondary structure of U1 snRNA is adapted from [37]. There is no structural data for DHFR ncRNA. U1 snRNA activates transcription (shown by green arrow) by stimulating the TFIIB-dependent phosphorylation of RNAP II Rpb1 CTD. DHFR ncRNA inhibits transcription by displacing the TFIIB transcription factor from PIC.

transcription factor – TAF15 – associated with TFIID in PIC and presumably involved in the elongation step [39]. For all intent and purposes, U1 snRNA activates the transcription process unlike the other aforementioned regulatory ncRNAs.

DHFR ncRNA

The *DHFR* gene encodes dihydrofolate reductase, one of the key enzymes of folate metabolism. About 99% of DHFR mRNAs are transcribed from the major promoter and contain six exons. During serum starvation and cell growth retardation, an alternate promoter located ~450 nt upstream from the main transcription initiation site is activated. Early transcription termination at the second intron results in the formation of a short product from the minor promoter, DHFR ncRNA, whose length varies from 800 nt to 2–3 thousand nt [40]. The functional part of the molecule is likely to be a ~400 nt-long fragment complementary to the promoter region of its own gene and containing long poly-(dG)-sequences. The latter are involved in the formation of an H-shaped purine-purine-pyrimidine triplex between DHFR ncRNA and the promoter that impedes PIC as-

sembly [41]. Thus, DHFR ncRNA belongs to the class of promoter-associated ncRNAs [9]. Above that, it can interact with the transcription factor TFIIB incorporated in PIC, resulting in its dissociation [42]. Since TFIIB binding to the promoter is a key stage of PIC assembly and DHFR ncRNAs totally prevents this process, transcription is inhibited. It is yet unknown which region of DHFR ncRNA is responsible for the interaction with TFIIB, as well as the details of the processing of this ncRNA.

7SK and TAR RNA

Human 7SK snRNA and TAR RNA of HIV are probably the most well-known eukaryotic ncRNAs involved in the regulation of transcription elongation. Both of these ncRNAs act as platforms for the assembly of protein associates, modulating the activity of the RNAP II elongation complex, and they interact with the factor P-TEFb [43–45].

P-TEFb is a key transcription factor that stimulates the proceeding of RNAP II arrested at the promoter (the so-called transcription pause required for 5'-capping of the nascent mRNA strand) to the activation of elongation. P-TEFb consists of cyclin-dependent kinase 9 (CDK9) and cyclin T1 or its analogs, CycT1b and CycT2b (hereinafter CycT). Its main function includes phosphorylation of Ser2 in the CTD of Rpb1 RNAP II, as well as the transcription repressors NELF and DSIF [46] (*Fig. 5A*). P-TEFb is attracted to the polymerase by various DNA-binding proteins, first of all Brd4, but also by some general transcriptional factors, such as NF-κB, HSF, p53, c-Myc, etc. After overcoming the pause stage, P-TEFb binds several other proteins, forming a super elongation complex (SEC) of RNAP II [47].

In the absence of P-TEFb, RNAP II can transcribe only short 5'-terminal sequences of pre-mRNA; i.e., this factor is required for the synthesis of most cellular mRNAs. When interacting with P-TEFb, 7SK snRNA inhibits its activity, which is an important regulatory mechanism of gene expression in eukaryotic cells. On the other hand, the release of P-TEFb from its complex with 7SK snRNA may serve as a signal for cell growth and proliferation [48]. On the contrary, HIV TAR RNA activates P-TEFb, which facilitates the initiation of transcription of the 5'-end of the viral promoter (5'-LTR) [45]. This review describes only the main features of these ncRNAs and their functioning principles.

7SK snRNA is 332 nt in length and consists of four main long hairpin structures connected by disordered areas and additional small hairpins. Although the human genome contains hundreds of 7SK snRNA pseudogenes, this RNA (~2 × 10⁵ copies per cell) is transcribed by RNAP III from a single genuine gene located on the sixth chromosome. The nucleotide sequence of

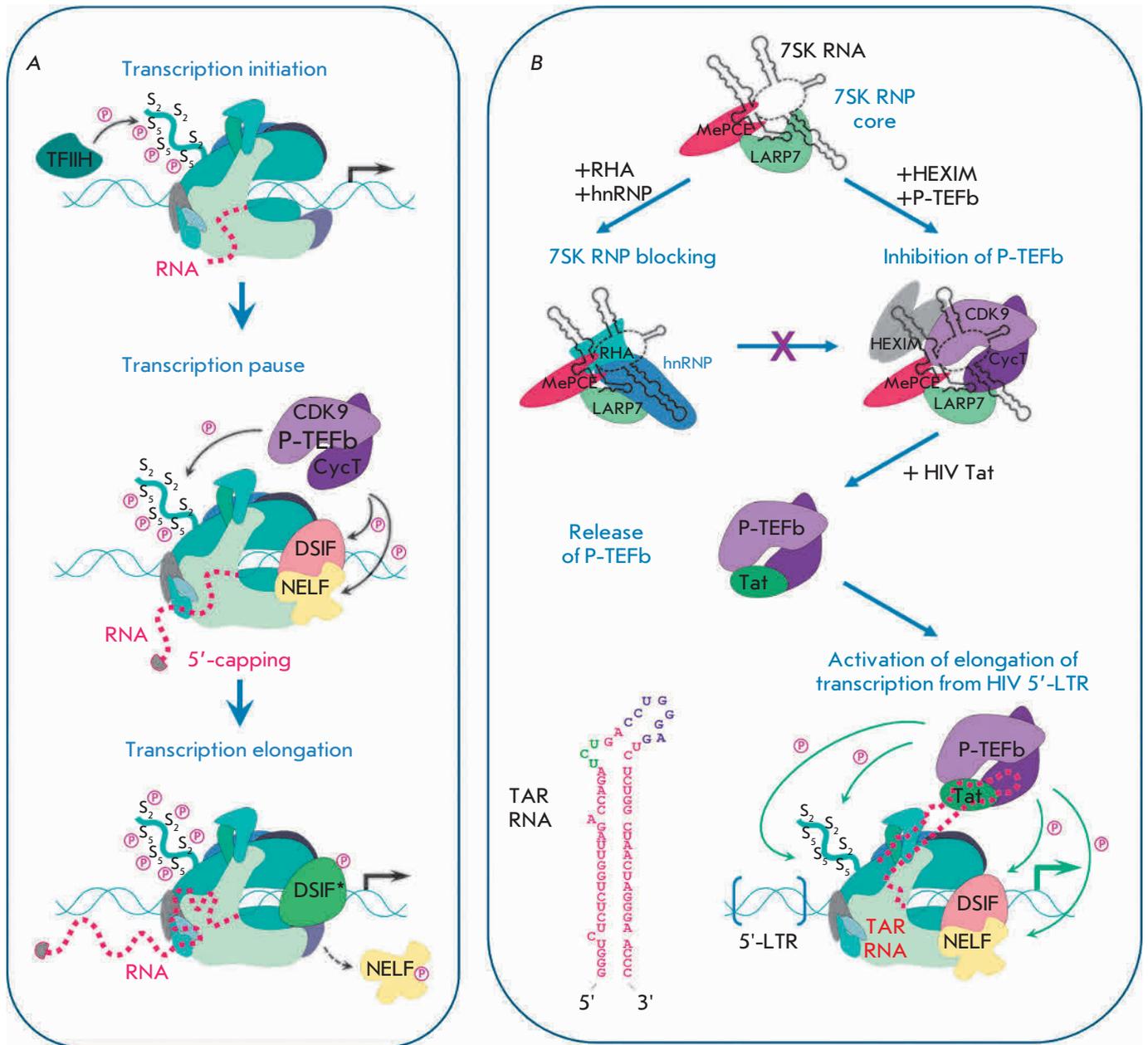


Fig. 5. Transcription regulation involving human 7SK snRNA and HIV TAR RNA. **(A)** General steps of transcription conducted by RNAP II [46]. For transcription initiation, TFIIH phosphorylates Ser5 residues in RNAP Rpb1 CTD. The enzyme stops after the synthesis of a small transcript and the negative elongation factors NELF and DSIF bind to RNAP, resulting in transcription pause. After 5'-capping of a nascent RNA strand transcription restarts: DNA-binding proteins attract the P-TEFb factor, which further phosphorylates RNAP Rpb1 CTD and the factors NELF and DSIF. The latter turns into a transcription activator (DSIF*, shown in green), and modified NELF dissociates from the complex, that enables RNAP to proceed to the transcription elongation step. **(B)** Assembly of alternative protein complexes on 7SK snRNA. Binding of RHA and hnRNP to the 7SK snRNP core prevents the inhibition of P-TEFb. The HIV protein Tat can displace P-TEFb from 7SK snRNP and attract it to RNAP, arrested near the transcription start site. TAR RNA interacts with Tat and CycT and activates the kinase activity of P-TEFb, resulting in hyperphosphorylation of RNAP Rpb1 CTD and NELF/DSIF, followed by elongation of a viral transcript.

this gene is highly conserved in vertebrates [49]. During post-transcriptional modifications, nucleases cleave 1–3 nt from the 3'-end of 7SK RNA, before adenylation occurs, resulting in three different isoforms of 7SK snRNA present in the cell: 330, 331 and 332 nt in length, of which 331-mer is the most stable. In addition, 7SK snRNA is capped at the 5'-end: methyltransferase MePCE methylates the γ -phosphate group of the 5'-terminal guanosine residue. This process is not characteristic of transcripts synthesized by RNAP III, and it has been previously described only for U6 and 7SK snRNAs [50].

Approximately 90% of 7SK snRNA in the cell is bound to MePCE and, together with the LARP7 protein, forms a so-called core of the ribonucleoprotein complex, 7SK snRNP (*Fig. 5B*). MePCE and LARP7 also interact with each other, further stabilizing snRNP; in this state, 7SK RNA is protected from degradation. The complex further binds the HEXIM protein in the form of a dimer consisting of alternative HEXIM1 and/or HEXIM2 paralogs. The arginine-rich RNA-binding domain (ARM) of HEXIM binds the 5'-terminal hairpin of 7SK snRNA, resulting in conformational change in the protein, so that it can interact with CycT of P-TEFb. Additionally, the C-terminal domain of LARP7 binds to CDK9, providing a stable structure of the whole complex. Apparently, 7SK RNA is also involved in the formation of contacts with P-TEFb. As a result, the factor loses its kinase activity, which prevents it from promoting transcription elongation [48, 51, 52].

However, not all P-TEFb molecules are bound to 7SK snRNP. A wealth of experimental data shows that there is continuous equilibrium between the free and bound forms of P-TEFb in the cell nucleus, which is controlled through various signaling pathways. For example, a number of heterogeneous nuclear ribonucleoproteins (hnRNPs) and RNA helicase A (RHA) block access to P-TEFb through binding to the 7SK snRNP core (*Fig. 5B*). Another mechanism concerns temporary inactivation of P-TEFb, since only the activated form of the protein (bearing the phosphorylated T186 residue in the so-called T-loop of the CDK9) can interact with 7SK snRNP. Serine-threonine phosphatases are responsible for this process, including PPM1G, attracted by the NF- κ B. Some proteins may also acetylate CycT, phosphorylate HEXIM, demethylate 7SK snRNA at the 5'-end, or carry out MePCE proteolysis, which leads to a destabilization of the complex and dissociation of P-TEFb. After the release from 7SK snRNP, the factor is again modified [52]. Let us note that the bulk of the P-TEFb that forms a complex with 7SK snRNP is associated with chromatin (e.g., through the Brd4 protein, interacting with the acylated histones H3 and

H4), and the described mechanisms are often realized co-transcriptionally. Brd4 can also bind to P-TEFb in its complex with 7SK snRNP and initiate conformational changes in CycT and dissociation of CDK9 [53].

The best known mechanism of P-TEFb dissociation from its complex with 7SK snRNP is represented by TAR RNA in HIV-infected cells (*Fig. 5B*). TAR RNA is a 5'-terminal structural element (hairpin) of the nascent strand of viral RNA synthesized from 5'-LTR. In the absence of additional activation, RNAP II is incapable of synthesizing transcripts longer than 60–80 nt from 5'-LTR, and TAR RNA consisting of 59 nt is the smallest fragment, followed by a transcription pause. In order to stimulate elongation, TAR RNA binds to a viral protein, Tat, which attracts various transcription factors to 5'-LTR, including P-TEFb [54]. This interaction is a result of specific contacts between the arginine-rich RNA-binding domain (ARM) of Tat and the trinucleotide side loop 5'-UCU-3' in the TAR RNA. At the same time, the apical loop of TAR RNA and the flanking region are associated with CycT (*Fig. 5B*). This region of the molecule mimics the 5'-terminal hairpin of the 7SK snRNA, which enables TAR RNA binding to the ARM HEXIM, thus preventing the activation of P-TEFb in the absence of Tat. Moreover, Tat directly interacts with CycT and CDK9, forming a stable complex: that crystal structure was resolved in 2010 [55]. When binding to the so-called T-loop of CDK9, Tat changes the substrate specificity of the kinase, which then phosphorylates not only Ser2 in CTD of the Rpb1, but also Ser5 residues [56]. This allows HIV to activate transcription elongation even without the involvement of TFIIH (*Fig. 5B*). Formation of a Tat-TAR-P-TEFb ternary complex is regulated by a number of enzymes that perform the acetylation, phosphorylation, methylation, and ubiquitination of Tat [54].

Obviously, there should be competition between the Tat-TAR RNA and 7SK snRNP complexes for binding to P-TEFb. As it turns out, Tat can displace the elongation factor from its complex with 7SK snRNP due to a direct interaction between Tat and CycT and conformational changes in P-TEFb [57, 58]. A similar mechanism was described for the RNA-binding proteins SRSF1 and SRSF2 that are involved in RNA splicing and metabolism in mammalian cells and are usually associated with the promoter regions of actively transcribed genes. Both proteins are capable of binding the 5'-terminal hairpin of 7SK snRNA, forming an alternative 7SK snRNP. If the nascent RNA stand contains the ESE sequence (exonic-splicing enhancer), SRSF1 and SRSF2 bind to it, resulting in a release of the active P-TEFb in close vicinity of RNAP II and the stimulation of transcription elongation of the required gene [59].

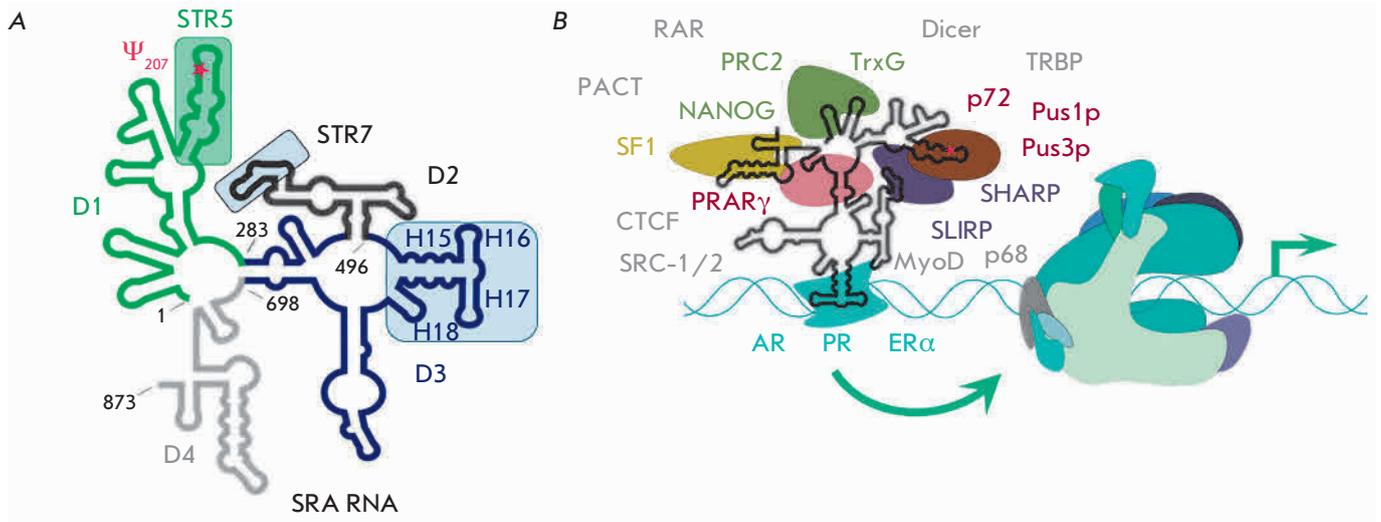


Fig. 6. Schematic representation of the secondary structure of human SRA RNA (**A**) and its currently known protein partners (**B**) according to Liu *et al.* [61]. SRA RNA domains are colored: D1 – green, D2 – black, D3 – blue, D4 – grey. The U207 residue subjected to pseudouridilation is marked by an asterisk. Panel **A** shows the main structural elements of SRA RNA that bind several proteins, shown in color frames. Panel **B** shows a schematic representation of the proteins which directly bind to SRA RNA (they are labeled in corresponding colors). Nuclear receptors (AR – androgen, PR – progesterone, ER α – estrogen α) are colored in blue. All other proteins known to interact with SRA RNA are denoted in grey (without animation).

NONCODING RNAs INTERACTING WITH OTHER TRANSCRIPTION FACTORS

SRA RNA

Human SRA RNA (steroid receptor activator) is a long ncRNA involved in the activation of estrogen (ER), progesterone (PR), glucocorticoid (GR), and other nuclear receptors. Similarly to 7SK snRNA, SRA RNA serves as a platform for the binding (including competitive) of various transcription factors. CTCF, SLIRP, and SHARP, as well as the RNA helicases p68 and p72, are the most important among them [60, 61]. Besides, SRA RNA modulates the activity of the transcription factor MyoD, which plays a key role in the differentiation of muscle cells [62]. SRA RNA is present in all human tissues, although its highest level is observed in the liver, heart, and skeletal muscles [63]. SRA RNA expression increases in females with polycystic ovarian and breast cancer, which justifies the rising interest in this RNA as a therapeutic target [61].

The *sra1* gene encoding SRA RNA is highly conserved in the genome of mice, rats, and humans. It is about 6,500 bp in length and consists of five exons. At least 20 different isoforms of SRA RNA, from 700 to 1,500 nt, have been detected in human cells. Most transcripts contain a core element of 687 nt, which corresponds to exons 2–5, and differ in their 5'- and 3'-end regions [64]. In 2012, chemical and enzymatic probing

studies revealed the secondary structure of the 873-nt variant of RNA SRA, consisting of 25 hairpins (H1–H25) characterized by various lengths and shapes, which were conventionally divided into four domains, D1–D4, and 12 main structural elements, STR1–STR12 (Fig. 6A) [65]. Analysis of the deletion mutants of SRA RNA determined the six most important STRs that are responsible for the binding to certain proteins. Furthermore, removal of any STR results in complete or partial loss of the molecule's functional properties; i.e., all of the main interactions occur owing to the multiplet structure of SRA RNA [66].

Nevertheless, the D3 domain (494–699 nt) is the most important one for the interaction with nuclear receptors. Its constituent element H15–H18 (505–575 nt) is highly conserved in vertebrates. Interestingly, individual expression of this element leads, on the contrary, to the inhibition of the transcription of ER α -dependent genes [67]. Switching of the SRA RNA function from activation to repression of nuclear receptors was also observed after replacement of the extremely important U207 residue in STR5 by adenosine. U207 is a site of pseudouridilation, which is carried out by the pseudouridine-synthetases Pus1p and Pus3p, coactivators of nuclear receptors. For example, direct interaction between SRA RNA and Pus1p in murine cells activates the transcription of genes dependent on retinoic acid receptors (mRARc) [68]. Synthetic oligonucleotide

identical to STR5 can compete with full-length SRA RNA and block Pus1p, preventing a modification of this ncRNA, which results in the inhibition of the transcription of AR- and ER α -dependent genes [69].

STR7 of SRA RNA interacts with the RNA-recognition motifs (RRM) of the factors SHARP and SLIRP and, thereby, initiates both activation and inhibition of the transcription of various genes. Another SRA RNA partner, the receptor PPAR γ , regulates the expression of the genes involved in the control of adipogenesis and insulin sensitivity [60]. Clearly, SRA RNA is one of the most important components among those engaged in the control of the activity of nuclear receptors, and it participates in various regulatory mechanisms, frequently accompanied by protein cascades (*Fig. 6B*).

Similarly to other ncRNAs, some SRA RNA suppression mechanisms exist that in particular involve the SRAP protein. The latter is encoded in 39% of SRA RNA transcripts and represents an example of specific self-regulation. In fact, SRA RNA is a coding RNA, although the key function is fulfilled by a non-coding variant of the *sra1* gene transcript. At the moment, it remains unclear whether SRAP is capable of direct binding to SRA RNA and inhibition of its interaction with other proteins, or whether this process is carried out through the transcription factors or nuclear receptors associated with this RNA [70]. Nevertheless, the ratio between the amount of translated and untranslated products of *sra1* gene transcription is one of the key factors of transcriptional regulation in the cell. For example, in the case of myocyte differentiation, equilibrium is strongly shifted towards the non-coding SRA RNA and it can smoothly interact with transcriptional activators, attracting them to the MyoD-dependent promoter and activating the transcription of the respective genes [71].

GAS5 RNA

GAS5 RNA (growth arrest-specific 5) is another long ncRNA that controls transcription through regulation of nuclear receptors. Under normal conditions, GAS5 RNA is rapidly degraded. However, in case of serum starvation in cells, arrested at a certain stage of growth, or after treatment with translation inhibitors, expression of GAS5 RNA is induced and its stability enhances, that gave the name for this ncRNA [72]. The main function of GAS5 RNA is to inhibit glucocorticoid receptor GR, the transcription factor responsible for activation of glucocorticoid genes. GR is a DNA-binding protein that recognizes the nucleotide sequence of GRE (glucocorticoid responsive element) in the promoter regions of controlled genes. The functional region of GAS5 RNA mimics GRE and, when binding to GR, blocks its access to promoters, thereby preventing activation of

their transcription (*Fig. 7*). GAS5 RNA can interact not only with GR, but also with other nuclear receptors binding GRE, in particular androgen, progesterone receptors, etc. [73]. Recent *in vitro* and *in vivo* studies clearly indicate the important role of GAS5 RNA in the initiation of apoptosis in various types of tumor cells and inhibition of proliferation and metastasis, as well as in immune response regulation in various inflammatory, bacterial, and viral diseases [74, 75].

Human GAS5 RNA is encoded by the *gas5* gene, which comprises 12 exons interspersed with 10 introns encoding small nucleolar RNAs. After transcription performed by RNAP II, pre-mRNA undergoes polyadenylation and alternative splicing, which results in different isoforms of GAS5 RNA. The most important of them are GAS5a (612 nt) and GAS5b (651 nt) RNAs, containing 7a or 7b exons, respectively. Longer variants of GAS5 RNA (~1200–1800 nt) are less common and contain one or more sequences encoding small nucleolar RNAs [72, 76]. Under normal conditions, GAS5 RNA is localized in the cytoplasm and remains associated with the ribosome. However, in the case of arrested cell growth, it is translocated to the nucleus, where it interacts with the GR receptor [76].

The GAS5 RNA secondary structure is represented by several hairpins. The functional region (identified by deletion analysis) is located at the 3'-terminal region of the molecule (400–598 nt, *Fig. 7A*) and is found in all GAS5 RNA isoforms. The main contacts are formed between the GR and the GRE-1/GRE-2 hairpin stem (539–544 and 553–559 nt), which mimic the conformation of the palindromic GRE-sequence of DNA: d(5'-AGAACANNNTGTTCT-3'/3'-TCTTGTTNNACAAGA-5'), where N = A, T, C, G). The G540 and C554 residues in GAS5 RNA are conserved among the human consensus GRE-sequences and interact with K442 and R447 in the DNA-binding domain of the GR protein, respectively. C554U substitution in GAS5 RNA, maintaining the stability of the double helix, results in a loss of the ability of this RNA to inhibit GR-dependent transcription from the MMTV (mouse mammary tumor virus) promoter *in vivo* [76]. Thus, GAS5 RNA competes with GRE-containing promoters for binding to GR similarly to bacterial 6S RNA, which also mimics the promoter and inhibits RNAP [36]. It has been shown that transfection of tumor cell lines with oligodeoxyribonucleotides identical to the 538–560 nt region of GAS5 RNA leads to effective apoptosis induction and a decrease of the cell survival rate [77], which could possibly enable using them for therapeutic purposes in the future.

The amount of GAS5 RNA in the cell is regulated by the NMD (nonsense-mediated RNA decay) system implementing the degradation of “nonsense” mRNA

sequences and the mTOR (mammalian target of rapamycin) kinase-dependent signaling pathway. It is assumed that during active cell growth, mTOR-dependent translation of the short open reading frame (ORF) located in GAS5 RNA can occur (let us recall that this ncRNA is associated with the ribosome under these conditions). However, the large number of stop codons in the ORF and the short length of the potentially synthesized peptide lead to the activation of NMD and degradation of GAS5 RNA. In the case of cellular arrest and low level of the mTOR complex, GAS5 RNA is not translated and its concentration increases [73].

In addition to its primary function, GAS5 RNA binds oncogenic miR-21, miR-222, and miR-103, thereby serving as a siRNA-sponge and preventing them from impacting gene expression [74, 78]. Moreover, recent studies have shown that the 5'-terminal part of GAS5 RNA (1–250 nt) can bind the NS3 protein of the hepatitis C virus (HCV) and inhibit its function, thereby repressing HCV replication [79]. Obviously, multifunctionality of GAS5 RNA is achieved via different domains of the molecule: each one is responsible for interaction with a certain target (*Fig. 7*).

OTHER RNAs INVOLVED IN THE REGULATION OF TRANSCRIPTION FACTOR ACTIVITY

Dozens of ncRNAs modulating the activity of transcription factors are currently known. Most of them

“work” only under certain (stress) conditions and are often tissue-specific [13, 17]. For example, NRSE RNA (neuron-restrictive silencer element) is expressed in stem cells and associated with the transcriptional repressor NRSF/REST responsible for neuron-specific gene silencing. The mechanism of this RNA-protein interaction is similar to the GAS5 RNA functional mechanism: NRSE RNA is a short (~20 bp) double stranded RNA that mimics the structure of the promoter. The NRSE-bound NRSF/REST factor turns into a transcription activator and “switches on” neuron-specific gene expression [80]. TSU RNA (trophoblast STAT utron), the 5'-untranslated end of the mRNA of the gene encoding the transcription factor STAT1, binds to its own protein and, thus, mimics the STAT-binding promoter and thereby inhibits gene expression of the major histocompatibility complex [81]. Long noncoding HSR1 RNA (~600 nt) activates HSF1, the main heat shock transcription factor, which initiates the functioning of the RNAP II elongation complex retained at stress gene promoters [82]. Other ncRNAs affect the transcription factor activity, altering their cellular localization; e.g., NRON RNA and lncRNA-p21 [83].

Special attention should be paid to circular RNAs (circRNAs or ciRNAs), products of alternative splicing, which results in closure of the 5'- and 3'-ends of the molecule. According to recent data, more than

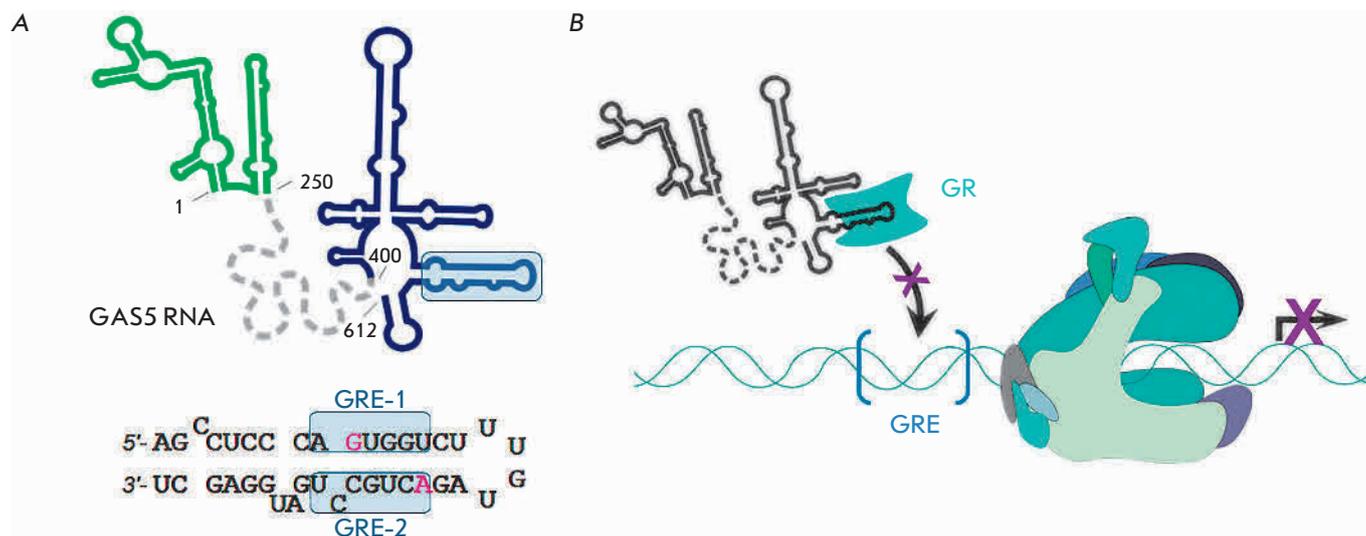


Fig. 7. Inhibition of the transcription of glucocorticoid-dependent genes by GAS5 RNA. **(A)** Predicted GAS5 RNA secondary structure (top) and nucleotide sequence of the functional element (hairpin), containing the GRE-1 and GRE-2 regions that mimic DNA promoter (bottom). The key residues G540 and C554 are shown in pink. The GAS5 RNA domain responsible for the interaction with the HCV NS3 protein is colored in green. Central part of the molecule which is a target for miR binding is shown as a grey dash line (secondary structure is unknown). Both domains apparently do not participate in transcription regulation. **(B)** GAS5 RNA inhibits transcription by binding nuclear receptors and preventing their interaction with GRE-containing promoters.

100 circRNAs are associated with RNAP II and at least some of them activate transcription of their own genes [84]. It is assumed that circRNAs indirectly interact with RNAP II through the U1 snRNP of spliceosome, but the exact mechanism of their action is unknown [13].

CONCLUSION

In recent years, increasingly detailed information has appeared regarding the various ncRNAs involved in transcription regulation in eukaryotic cells both at the level of specific genes and on a more global scale. Most often, they are long ncRNAs that regulate transcription during chromatin remodeling. In this review, we described those ncRNAs whose action mechanisms are closely related to the supervision of the functioning of the RNAP II transcription complex. The aforementioned ncRNAs have a number of other features which are equally important. For example, Alu RNA binds SRP9/14 (signal recognition particle) proteins and, as part of this RNP, inhibits translation initiation. In

contrast, free Alu RNA is able to activate this process [85]. U1 snRNA is one of the main components of the spliceosome, and its participation in the activation of transcription factors is not that significant. At the same time, the ability of GAS5 RNA to interact with oncogenic miRs may be no less important than the ability to bind GR-receptors. Finally, it was shown that B2 RNA regulates transcription not only by interacting with RNAP II and inhibiting its activity, but also by direct binding to heat shock protein genes and inhibiting their expression in the absence of stress. Increased temperature results in the degradation of B2 RNA initiated by the EZH2 protein incorporated in the PRC2 complex and releases these genes for active transcription [86]. These and other facts attest to the diversity of the properties and functions of ncRNAs and undoubtedly demonstrate their importance in cell activity. ●

This work was supported by a Russian Scientific Foundation grant (No 14-24-00061).

REFERENCES

1. Yang Z., Li X., Yang Y., He Z., Qu X., Zhang Y. // Cell Death Dis. 2016. V. 7. № 9. P. e2389.
2. Francia S. // Front Genet. 2015. V. 13. № 6. P. 320.
3. Kaikkonen M.U., Lam M.T.Y., Glass C.K. // Cardiovasc. Res. 2011. V. 90. № 3. P. 430–440.
4. Castelnuovo M., Stutz F. // Curr. Opin. Cell Biol. 2015. V. 34. P. 16–22.
5. Esteller M. // Nat. Rev. Genet. 2011. V. 12. № 12. P. 861–874.
6. Khurana E., Fu Y., Chakravarty D., Demichelis F., Rubin M.A., Gerstein M. // Nat. Rev. Genet. 2016. V. 17. № 2. P. 93–108.
7. Lopez-Pajares V. // Pflug. Arch. 2016. V. 468. № 6. P. 971–981.
8. Iwasaki Y.W., Siomi M.C., Siomi H. // Annu. Rev. Biochem. 2015. V. 84. P. 405–433.
9. Yan B.X., Ma J.X. // Cell Mol. Life Sci. 2012. V. 69. № 17. P. 2833–2842.
10. Lam M.T., Li W., Rosenfeld M.G., Glass C.K. // Trends Biochem. Sci. 2014. V. 39. № 4. P. 170–182.
11. St. Laurent G., Wahlestedt C., Kapranov P. // Trends Genet. 2015. V. 31. № 5. P. 239–251.
12. Rutenberg-Schoenberg M., Sexton A.N., Simon M.D. // Annu. Rev. Genomics Hum. Genet. 2016. V. 17. P. 69–94.
13. Eidem T.M., Kugel J.F., Goodrich J.A. // J. Mol. Biol. 2016. V. 428. № 12. P. 2652–2659.
14. Boon R.A., Jaé N., Holdt L., Dimmeler S. // J. Am. Coll. Cardiol. 2016. V. 67. № 10. P. 1214–1226.
15. Walters R.D., Kugel J.F., Goodrich J.A. // IUBMB Life. 2009. V. 61. № 8. P. 831–837.
16. Kettenberger H., Eisenführ A., Brueckner F., Theis M., Famulok M., Cramer P. // Nat. Struct. Mol. Biol. 2006. V. 13. № 1. P. 44–48.
17. Mondragón E., Maher L.J. // Nucl. Acids Ther. 2016. V. 26. № 1. P. 29–43.
18. Kramerov D.A., Vassetzky N.S. // Heredity. 2011. V. 107. № 6. P. 487–495.
19. Fornace A.J.Jr., Alamo I.Jr., Hollander M.C., Lamoreaux E. // Exp. Cell Res. 1989. V. 182. № 1. P. 61–74.
20. Elbarbary R.A., Lucas B.A., Maquat L.E. // Science. 2016. V. 351. № 6274. P. aac7247.
21. Ponicsan S.L., Kugel J.F., Goodrich J.A. // Curr. Opin. Genet. Dev. 2010. V. 20. № 2. P. 149–155.
22. Vassetzky N.S., Ten O.A., Kramerov D.A. // Gene. 2003. V. 319. P. 149–160.
23. Mariner P.D., Walters R.D., Espinoza C.A., Drullinger L.F., Wagner S.D., Kugel J.F., Goodrich J.A. // Mol. Cell. 2008. V. 29. № 4. P. 499–509.
24. Allen T.A., Von Kaenel S., Goodrich J.A., Kugel J.F. // Nat. Struct. Mol. Biol. 2004. V. 11. № 9. P. 816–821.
25. Espinoza C.A., Allen T.A., Hieb A.R., Kugel J.F., Goodrich J.A. // Nat. Struct. Mol. Biol. 2004. V. 11. № 9. P. 822–829.
26. Kassube S.A., Fang J., Grob P., Yakovchuk P., Goodrich J.A., Nogales E. // J. Mol. Biol. 2013. V. 425. № 19. P. 3639–3648.
27. Wagner S.D., Kugel J.F., Goodrich J.A. // Mol. Cell Biol. 2010. V. 30. № 1. P. 91–97.
28. Kramerov D.A., Vassetzky N.S. // Wiley Interdiscip. 2011. Rev. RNA. V. 2. P. 772–786.
29. Allen T.A., Von Kaenel S., Goodrich J.A., Kugel J.F. // Nat. Struct. Mol. Biol. 2004. V. 11. № 9. P. 816–821.
30. Bladon T.S., Frégeau C.J., McBurney M.W. // Mol. Cell Biol. 1990. V. 10. № 8. P. 4058–4067.
31. Espinoza C.A., Goodrich J.A., Kugel J.F. // RNA. 2007. V. 13. № 4. P. 583–596.
32. Ponicsan S.L., Kugel J.F., Goodrich J.A. // Noncoding RNA. 2015. V. 1. P. 4–16.
33. Sansó M., Fisher R.P. // Transcription. 2013. V. 4. № 4. P. 146–152.
34. Wagner S.D., Yakovchuk P., Gilman B., Ponicsan S.L., Drullinger L.F., Kugel J.F., Goodrich J.A. // EMBO J. 2013. V. 32. № 6. P. 781–790.
35. Flores R., Ruiz-Ruiz S., Serra P. // Semin Liver Dis. 2012. V. 32. № 3. P. 201.

36. Burenina O.Y., Elkina D.A., Hartmann R.K., Oretskaya T.S., Kubareva E.A. // *Biochemistry (Mosc)*. 2015. V. 80. P. 1429–1446.
37. Guiro J., O'Reilly D. // *WIREs RNA*. 2015. V. 6. № 1. P. 79–92.
38. Kwek K.Y., Murphy S., Furger A., Thomas B., O'Gorman W., Kimura H., Proudfoot N.J., Akoulitchev A. // *Nat. Struct. Biol.* 2002. V. 9. № 11. P. 800–805.
39. Leichter M., Marko M., Ganou V., Patrinou-Georgoula M., Tora L., Guialis A. // *Biochim. Biophys. Acta*. 2011. V. 1814. № 12. P. 1812–1824.
40. Masters J.N., Attardi G. // *Mol. Cell. Biol.* 1985. V. 5. № 3. P. 493–500.
41. Blume S.W., Meng Z., Shrestha K., Snyder R.C., Emanuel P.D. // *J. Cell Biochem.* 2003. V. 88. № 1. P. 165–180.
42. Martianov I., Ramadass A., Serra Barros A., Chow N., Akoulitchev A. // *Nature*. 2007. V. 445. № 7128. P. 666–670.
43. Yang Z., Zhu Q., Luo K., Zhou Q. // *Nature*. 2001. V. 414. № 6861. P. 317–322.
44. Nguyen V.T., Kiss T., Michels A.A., Bensaude O. // *Nature*. 2001. V. 414. № 6861. P. 322–325.
45. Wei P., Garber M.E., Fang S.M., Fischer W.H., Jones K.A. // *Cell*. 1998. V. 20. № 4. P. 451–462.
46. Jonkers I., Lis J.T. // *Nat. Rev. Mol. Cell. Biol.* 2015. V. 16. № 3. P. 167–177.
47. McNamara R.P., Bacon C.W., D'Orso I. // *Cell Cycle*. 2016. V. 15. № 16. P. 2115–2123.
48. Peterlin B.M., Brogie J.E., Price D.H. // *Wiley Interdiscip. Rev. RNA*. 2012. V. 3. № 1. P. 92–103.
49. Wassarman D.A., Steitz J.A. // *Mol. Cell. Biol.* 1991. V. 11. № 7. P. 3432–3445.
50. Gupta S., Busch R.K., Singh R., Reddy R. // *J. Biol. Chem.* 1990. V. 265. № 31. P. 19137–19142.
51. D'Orso I. // *RNA Biol.* 2016. V. 13. № 6. P. 545–553.
52. Quaresma A.J.C., Bugai A., Barboric M. // *Nucl. Acids Res.* 2016. V. 44. № 16. P. 7527–7539.
53. Liu W., Ma Q., Wong K., Li W., Ohgi K., Zhang J., Aggarwal A., Rosenfeld M.G. // *Cell*. 2013. V. 155. № 7. P. 1581–1595.
54. Karn J., Stoltzfus M.C. // *Cold Spring Harb. Perspect. Med.* 2012. V. 2. № 2. P. a006916.
55. Tahirov T.H., Babayeva N.D., Varzavand K., Cooper J.J., Sedore S.C., Price D.H. // *Nature*. 2010. V. 465. № 7299. P. 747–751.
56. Zhou M., Halanski M.A., Radonovich M.F., Kashanchi F., Peng J., Price D.H., Brady J.N. // *Mol. Cell. Biol.* 2000. V. 20. № 14. P. 5077–5086.
57. Krueger B.J., Varzavand K., Cooper J.J., Price D.H. // *PLoS One*. 2010. V. 5. № 8. P. e12335.
58. Ott M., Geyer M., Zhou Q. // *Cell Host Microbe*. 2011. V. 10. № 5. P. 426–435.
59. Ji X., Zhou Y., Pandit S., Huang J., Li H., Lin C.Y., Xiao R., Burge C.B., Fu X.D. // *Cell*. 2013. V. 153. № 4. P. 855–868.
60. Leygue E. // *Nucl. Recept. Signal.* 2007. V. 5. P. e006.
61. Liu C., Wu H.-T., Zhu N., Shi Y.-N., Liu Z., Ao B.-X., Liao D.-F., Zheng X.-L., Qin L. // *Clin. Chim. Acta*. 2016. V. 459. P. 137–146.
62. Caretti G., Schiltz R.L., Dilworth F.J., Di Padova M., Zhao P., Ogryzko V., Fuller-Pace F.V., Hoffman E.P., Tapscott S.J., Sartorelli V. // *Dev. Cell*. 2006. V. 11. № 4. P. 547–560.
63. Colley S.M., Leedman P.J. // *Biochimie*. 2011. V. 93. № 11. P. 1966–1972.
64. Lanz R.B., Razani B., Goldberg A.D., O'Malley B.W. // *Proc. Natl. Acad. Sci. USA*. 2002. V. 99. № 25. P. 16081–16086.
65. Novikova IV., Hennelly S.P., Sanbonmatsu K.Y. // *Nucl. Acids Res.* 2012. V. 40. № 11. P. 5034–5051.
66. Sanbonmatsu K.Y. // *Biochim. Biophys. Acta*. 2016. V. 1859. № 1. P. 41–45.
67. Jung E., Jang S., Lee J., Kim Y., Shin H., Park H.-S., Lee Y. // *Mol. Biol. Rep.* 2016. V. 43. № 10. P. 1019–1025.
68. Zhao X., Patton J.R., Ghosh S.K., Fischel-Ghodsian N., Shen L., Spanjaard R.A. // *Mol. Endocrinol.* 2007. V. 21. № 3. P. 686–699.
69. Ghosh S.K., Patton J.R., Spanjaard R.A. // *Biochemistry*. 2012. V. 51. № 41. P. 8163–8172.
70. McKay D.B., Xi L., Barthel K.K., Cech T.R. // *J. Mol. Biol.* 2014. V. 426. № 8. P. 1766–1785.
71. Hube F., Velasco G., Rollin J., Furling D., Francastel C. // *Nucl. Acids Res.* 2011. V. 39. № 2. P. 513–525.
72. Tani H., Torimura M., Akimitsu N. // *PLoS One*. 2013. V. 8. № 1. P. e55684.
73. Pickard M.R., Williams G.T. // *Genes*. 2015. V. 6. № 3. P. 484–499.
74. Yu X., Li Z. // *Oncol Lett.* 2015. V. 10. № 3. P. 1953–1958.
75. Mayama T., Marr A.K., Kino T. // *Horm. Metab. Res.* 2016. V. 48. № 8. P. 550–557.
76. Kino T., Hurt D.E., Ichijo T., Nader N., Chrousos G.P. // *Sci. Signal.* 2010. V. 3. № 107. P. ra8.
77. Pickard M.R., Williams G.T. // *Oncotarget*. 2016. V. 7. № 9. P. 10104–10116.
78. Guo C., Song W., Sun P., Jin L., Dai H. // *J. Biomed. Sci.* 2015. V. 22. P. 100.
79. Qian X., Xu C., Zhao P., Qi Z. // *Virology*. 2016. V. 492. P. 155–165.
80. Kuwabara T., Hsieh J., Nakashima K., Taira K., Gage F.H. // *Cell*. 2004. V. 116. № 6. P. 779–793.
81. Peyman J.A. // *Biol. Reprod.* 1999. V. 60. № 1. P. 23–31.
82. Sharnovsky I., Ivannikov M., Kandel E.S., Gershon D., Nudler E. // *Nature*. 2006. V. 440. № 7083. P. 556–560.
83. Meng X., Li X., Zhang P., Wang J., Zhou Y., Chen M. // *Brief Bioinform.* 2017. V. 18. № 4. P. 547–557.
84. Li Z., Huang C., Bao C., Chen L., Lin M., Wang X., Zhong G., Yu B., Hu W., Dai L. // *Nat. Struct. Mol. Biol.* 2015. V. 22. № 3. P. 256–264.
85. Chen L.L., Yang L. // *Trends Cell Biol.* 2017. V. 27. № 7. P. 480–490.
86. Zovoilis A., Cifuentes-Rojas C., Chu H.P., Hernandez A.J., Lee J.T. // *Cell*. 2016. V. 167. № 7. P. 1788–1802.

The Contribution of Ribosomal Protein S1 to the Structure and Function of Q β Replicase

Z. S. Kutlubaeva, H. V. Chetverina, A. B. Chetverin*

Institute of Protein Research, Institutskaya Str. 4, Pushchino, Moscow, 142290, Russia

*E-mail: alexch@vega.protres.ru

Received: July 7, 2016; in final form November 9, 2016

Copyright © 2017 Park-media, Ltd. 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 high resolution crystal structure of bacterial ribosome was determined more than 10 years ago; however, it contains no information on the structure of the largest ribosomal protein, S1. This unusual protein comprises six flexibly linked domains; therefore, it lacks a fixed structure and this prevents the formation of crystals. Besides being a component of the ribosome, protein S1 also serves as one of the four subunits of Q β replicase, the RNA-directed RNA polymerase of bacteriophage Q β . In each case, the role of this RNA-binding protein has been thought to consist in holding the template close to the active site of the enzyme. In recent years, a breakthrough was made in studies of protein S1 within Q β replicase. This includes the discovery of its paradoxical ability to displace RNA from the replicase complex and determining the crystal structure of its fragment capable of performing this function. The new findings call for a re-examination of the contribution of protein S1 to the structure and function of the ribosome.

KEYWORDS Bacteriophage Q β , initiation, crystal structure, OB domain, ribosomal protein S1, RNA replication, RNA-directed RNA polymerase, termination.

INTRODUCTION

Located on a small (30S) subparticle, protein S1 is not just the largest protein of the *Escherichia coli* ribosome, but it also has an unusual structure [1]. While other ribosomal proteins form compact globules [2, 3], protein S1 comprises a flexible strand [4, 5] almost as long as the ribosome [1] and consisting of six structurally similar units [6] called OB domains (for Oligonucleotide / oligosaccharide Binding [7]).

Protein S1 is vital to the cell, since deletions of or amber mutations in its *rpsA* gene are lethal [8, 9]. However, the exact functions of protein S1 and the position of its structural domains within the ribosome remain unknown. We only know that the N-terminal segment of the protein interacts with protein S2, which is located between the head and the platform of the 30S subparticle [10]. Establishing the crystal structure of the ribosome has not helped us clarify the matter, since only ribosomes completely devoid of protein S1 have proved to be crystallizable [3]. Apparently, crystallization of the ribosomes is hindered by this protein, which lacks a fixed conformation.

Undoubtedly, protein S1, with its strong RNA-binding capacity, is important for the initiation of translation [1, 11]. However, the contribution of protein S1 to translation is not limited to the initiation step, since,

unlike the initiation factors, S1 is present in the ribosome in stoichiometric amounts and remains bound to the ribosome during the elongation of a nascent polypeptide [1].

In addition to protein synthesis, protein S1 contributes to other processes that occur in the cell, both on the ribosome and outside of it [11]. One of the most known non-ribosomal functions of protein S1 is its recruitment in the synthesis of RNA as an α subunit of Q β replicase, the RNA-dependent RNA polymerase of the bacteriophage Q β . In addition to S1, Q β replicase contains the phage genome-encoded catalytic β subunit and the translation elongation factors EF-Tu and EF-Ts (γ and δ subunits, respectively) [12]. The β , γ , and δ subunits constitute the Q β replicase core, to which protein S1 is bound relatively weakly (as to the ribosome [1]) and partially dissociates during enzyme isolation [13].

Recent results of studies of protein S1 within Q β replicase have significantly advanced our understanding of the structure and function of this protein. These results are the subject of this review.

DOMAIN STRUCTURE OF PROTEIN S1

The “classic” OB domain consists of ≈ 70 amino acid residues and comprises a Greek key barrel of five β strands usually covered by an α -helix [7, 14]. OB do-

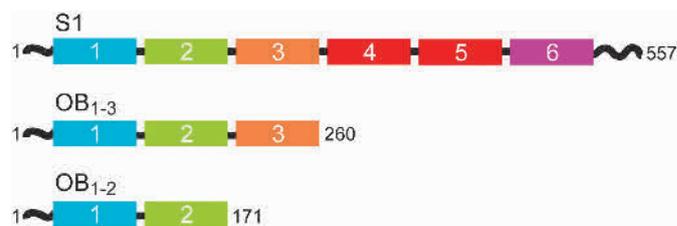
mains are present in the structure of many proteins capable of binding polynucleotides and polysaccharides [15]. An analysis of the amino acid sequence [6] has led to the conclusion that there are six OB domains in protein S1 (Figure). Subsequent structural studies have confirmed this conclusion with further refinement: the N-terminal OB domain (OB₁) contains 4 rather than 5 β -strands [10, 16–18].

The name OB domain suggests that it would display affinity for polynucleotides. Indeed, RNA-binding properties were demonstrated for domains OB₃ to OB₆ [1, 16, 19]. At the same time, it was thought that the domains OB₁ and OB₂ do not bind RNA and are involved in the protein-protein interactions responsible for the binding of protein S1 to the ribosome and to the Q β replicase core [1]. Recent studies have demonstrated that the OB₁ and OB₂ domains indeed form contacts with the Q β replicase core [17, 18], but, in addition, domain OB₂ can interact with RNA due to the high density of positively charged residues on the surface area not involved in the protein-protein interaction [18]. Thus, all five classical OB domains of S1 possess RNA-binding properties.

A special role belongs to the N-terminal segment preceding domain OB₁ and consisting of 20 amino acid residues. In the unbound state, this segment is unstructured [20], but upon interaction with ribosomal protein S2 [10] or the Q β replicase core [17, 18] it forms an α helix, which is slightly longer in the latter case. Apparently, the N-terminal α -helix makes the main contribution to the interaction of protein S1 with the ribosome [19]. It is also important for the interaction of S1 with the Q β replicase core, as suggested by both crystallographic data [17, 18] and gel filtration of protein complexes: the yield of the complex of protein S1 or its fragments with the Q β replicase core drops sharply if this helix is removed (Z. S. Kutlubayeva, P. Seweryn and A. B. Chetverin, unpublished data).

PROTEIN S1 AS A TERMINATION FACTOR OF RNA REPLICATION

Q β replicase is famous for its unique ability to rapidly amplify RNA. Similar to PCR, the reaction follows an exponential kinetics, since both the original template and its complementary copy serve as templates in the next amplification round. Therefore, the number of templates increases two-fold in each round, as long as the replicase remains in molar excess over RNA. However, unlike PCR, the reaction is isothermal: there is no need to increase the temperature to melt the duplex, since the immediate reaction product comprises a single-stranded RNA. How does Q β replicase manage to copy RNA according to the principle of complementarity, yet preserve the single strandedness of the template



Schematic presentation of the domain structure of protein S1 and its functional fragments. OB domains are shown as numbered rectangles colored according to [16]; also indicated are the numbers of terminal amino acid residues.

and the nascent strand, remains one of the unsolved mysteries of Q β phage replication [12].

In 1972, Weissmann and colleagues published a paper [21] arguing that Q β replicase only needs protein S1 to initiate the copying of the phage Q β genomic (plus) RNA strand, and does not need it to copy other templates, including the Q β RNA minus strand and the small replicating RNAs (“6S” or RQ RNAs, termed so for being Replicable by Q β replicase). Soon after, the lead author of that paper published a review in which, by referring to unpublished results, he claimed that protein S1 is neither needed at the steps of elongation and termination of the minus strand produced by copying the Q β RNA plus strand [22]. This view of the role of protein S1 in RNA replication persisted for the next 40 years.

Weissmann and his colleagues obtained their results when the Q β replicase and template concentrations were similar and no exponential synthesis of RNA was possible. We found that an entirely different result was obtained when the replicase was in large excess. In that case, protein S1 dramatically stimulated the replication of both the Q β RNA and RQ RNAs. In the presence of protein S1, most of the product was found to consist of single-stranded RNA, whereas in its absence the product was double stranded [23]. It seemed likely that protein S1 helped the replicase maintain the single strandedness of the template and the nascent RNA strand due to its known ability to bind single-stranded RNA and melt duplexes [24].

To verify this assumption, we examined the proportion between the single-stranded and double stranded forms of RNA during the elongation in the presence and absence of protein S1. As a template, we used the 4217 nt-long Q β RNA plus strand, whose copying takes about 4 min at 30°C. To avoid an overestimation of the amount of double-stranded RNA due to its formation upon denaturation of the replicative complex [25], we tested the RNA sensitivity to ribonuclease T1

before the phenol extraction step. We found that, irrespective of the presence of protein S1, the nascent strand remained single stranded throughout the elongation step and even some time after its completion. However, while a rapid release of the single-stranded full-length product from the replicative complex was observed in the presence of protein S1, the synthesized strand remained associated with the replicative complex in its absence. Over time, a minor amount of the product spontaneously left the complex in single stranded form, whereas the major portion formed a duplex with the template and acquired resistance to ribonuclease [23].

This result showed us that protein S1 catalyzes the release of a single-stranded product from the active site of Q β replicase: in other words, it acts as a termination factor. This function seems to be performed by protein S1 during the replication of any legitimate template [26] of Q β replicase. As a result, both the original template and its complementary copy become available for copying in the next replication round, which provides for the exponential accumulation of RNA.

TWO FUNCTIONS OF PROTEIN S1 HAVE DIFFERENT STRUCTURAL BASES

Thus, protein S1 performs two distinct functions during the replication of Q β RNA: the function of a termination factor common to all legitimate templates, and a special function performed during initiation on the plus strand, previously thought to be its only function.

Our direct measurements of the rate of initiation on the plus strand of Q β RNA showed that the requirement for protein S1 is not absolute. In a low salt buffer (50 mM NaCl), initiation in the absence of protein S1 occurred almost as rapidly as in its presence. However, the addition of 50 mM ammonium sulfate resulted in an almost complete inhibition of initiation in the absence of protein S1 and only in a two-fold inhibition in its presence (initiation on other legitimate templates was inhibited approximately two-fold regardless of the S1 presence) [23]. The addition of 100 mM of any other monovalent cation had the same effect, regardless of the nature of the anion (Z. S. Kutlubaeva, H. V. Chetverina and A. B. Chetverin, unpublished data). In view of the above, we conclude that at the initiation step protein S1 performs an anti-salt function. Apparently, when the Q β phage RNA is replicated in *E. coli* cells, the anti-salt function is as important as the termination function because the cytoplasmic concentration of monovalent cations is even higher than 150 mM [27].

In order to determine whether all domains of protein S1 are necessary for it to perform its functions, we cloned and purified a series of N-terminal S1 fragments

containing an increasing number of OB domains (*Figure*). It turned out that fragment OB₁₋₂ can replace protein S1 at the termination step, while fragment OB₁₋₃ can replace it in the protection of the initiation step against salt [23].

THE STRUCTURE OF COMPLEXES OF THE Q β -REPLICASE CORE WITH FUNCTIONAL FRAGMENTS OF PROTEIN S1

As with the ribosome, the failure of protein S1 to acquire a fixed conformation prevented the crystallization of the Q β replicase holoenzyme. This problem was overcome when it was discovered that relatively short (and therefore less flexible) fragments can replace protein S1 in all its functions [23]. Previously, two teams (Danish-Russian and Japanese) had solved independently of each other the crystal structure of the Q β replicase core [28, 29]. Recently, the same teams independently solved the structure of a complex containing the core and the first two OB domains of protein S1 [17, 18, 30]. Although the Japanese group investigated the crystals of the replicase core complexed with fragment OB₁₋₃ [17] while the Danish-Russian team studied the complex with fragment OB₁₋₂ [18], the same structural information was obtained in each case, since the third OB domain was not visible, due to the fact that it was unfixed within the structure of the complex [17]. In addition to contributing to our understanding of the mechanism of RNA replication, these studies are interesting in that they have established the crystal structure of a 1/3 of the protein S1 molecule, precisely the part whose structure was the least studied.

The binding of domains OB₁ and OB₂ produces almost no effect on the structure of the Q β replicase core. These domains interact with the β -subunit in the region of the “fingers” domain, which participates in the formation of the active site of replicase, the binding of RNA, and the unwinding of the complementary strands of the replicative complex [28, 31]. The N-terminal α -helix of protein S1 is located between domains OB₁ and OB₂ and forms a number of contacts with the β -subunit and EF-Tu [18]. These contacts are similar to those the helix forms with the ribosomal protein S2 [10].

Although the two research groups reported nearly identical structural data [17, 18], they drew somewhat different conclusions. Thus, the Japanese group claimed that domains OB₁ and OB₂ do not have basic and aromatic amino acid residues capable of forming bonds with the phosphates and nitrogenous bases of RNA, and therefore cannot interact with RNA [17]. On the contrary, the other group discovered an extended, positively charged region on the surface of domain OB₂ and presented NMR data demonstrating the ability of this domain to bind RNA [18].

THE CONTRIBUTION OF PROTEIN S1 TO THE INITIATION OF RNA REPLICATION

This function is similar to the one commonly assigned to protein S1 when considering its role in translation. The difference is that, while protein S1 stimulates initiation of the translation of the vast majority of mRNAs [1, 11], it promotes the initiation of replication of only one of a variety of Q β replicase templates. Unlike other templates, the initiation on the Q β RNA plus strand requires that the replicase binds the RNA not only at the 3' end wherein it begins copying, but also at the "M site," an internal template site spaced by $\approx 1,500$ nt from its 3' end. In the absence of protein S1, the replicase cannot bind the M site [32]. Apparently, protein S1 stimulates the initiation on the Q β RNA plus strand by increasing the replicase affinity to the M site. This is supported by the fact that initiation becomes sensitive to the elevated salt concentration both in the absence of S1 protein [23] and in its presence, if certain mutations are introduced into the M site [33]. The ability of fragment OB₁₋₃, rather than OB₁₋₂, to replace protein S1 in the anti-salt function means that the third OB domain plays a major role in the interaction with the M site, whereas domains OB₁ and OB₂ are needed as far as they form a link between domain OB₃ and the replicase core.

THE CONTRIBUTION OF S1 PROTEIN TO THE TERMINATION OF RNA REPLICATION

The Japanese group proposed a mechanism for protein S1 action in which the mobile domain OB₃ plays a key role in both the initiation and termination steps, by interacting with the M site of the Q β RNA plus strand during initiation, and with the newly synthesized RNA strand during termination [17, 34]. Although the authors referred to our paper [23], they apparently read it inattentively, since the paper directly showed that during termination protein S1 was replaced by its fragment OB₁₋₂, in which domain OB₃ was absent and, therefore, could not participate. Incidentally, neither did they notice the fact that the same paper demonstrated the key role of domain OB₃ at the initiation step one and a half year prior to their publication.

The Danish-Russian paper reports that the positively charged region of domain OB₂ adjoins a simi-

larly charged region on the surface of the β subunit and forms a continuous, positively charged tract leading from the opening through which the synthesized strand is thought to be released from the active site [18]. Probably, this tract is essential for the release of the synthesized strand from the replicative complex.

However, it would be premature to propose a detailed mechanism of termination, since the product is terminated from the closed conformation of Q β replicase [26], whereas the reported structure of the core : fragment OB₁₋₂ complex represents the open conformation. In this regard, we would note that protein S1 catalyzes the termination step even if it is added during the elongation step, but before its completion [23]. In other words, there is a "no return point" somewhere at the end of the elongation after which protein S1 cannot promote the release of the synthesized strand. What is that point?

As a result of the initiation on a legitimate template in the presence of GTP, Q β replicase acquires a closed conformation from which neither the template nor its complementary copy can dissociate until the elongation is complete [26]. This ensures high processivity of the replicase, but it hinders the evacuation of its active site after the copy is completed. To ensure "recycling" of the enzyme, the closed conformation must be converted back to the open one. Probably, it is this transition that is induced by the mysterious untemplated 3'-terminal adenylation of the synthesized strand, which precedes its termination [12] and represents the very moment by which protein S1 has to be embedded into the replicase molecule in order to fulfill the function of a termination factor.

In conclusion, we would like to note that the discovery of the ability of protein S1 to displace RNA from a complex changes the basic paradigm according to which the only purpose of this protein is to hold RNA near the active site of an enzyme, whether it is a replicase or a ribosome [1], and calls for a re-evaluation of the possible role of S1 in translation and other cellular processes. ●

This work was supported by the Russian Science Foundation (project No. 14-14-00350).

REFERENCES

1. Subramanian A.R. // Prog. Nucl. Acid Res. Mol. Biol. 1983. V. 28. P. 101–142.
2. Spirin A.S., Serdyuk I.N., Shpungin J.L., Vasiliev V.D. // Proc. Natl. Acad. Sci. USA. 1979. V. 76. № 10. P. 4867–4871.
3. Schuwirth B.S., Borovinskaya M.A., Hau C.W., Zhang W., Vila-Sanjurjo A., Holton J.M., Cate J.H. // Science. 2005. V. 310. № 5749. P. 827–834.
4. Chu Y.G., Cantor C.R. // Nucl. Acids Res. 1979. V. 6. № 6. P. 2363–2379.
5. Moore P.B., Laughrea M. // Nucl. Acids Res. 1979. V. 6. № 6. P. 2355–2361.
6. Gribskov M. // Gene. 1992. V. 119. № 1. P. 107–111.
7. Murzin A.G. // EMBO J. 1993. V. 12. № 3. P. 861–867.
8. Kitakawa M., Isono K. // Mol. Gen. Genet. 1982. V. 185. № 3. P. 445–447.

9. Sorensen M.A., Frick J., Pedersen S. // *J. Mol. Biol.* 1998. V. 280. № 4. P. 561–569.
10. Byrgazov K., Grishkovskaya I., Arenz S., Coudeville N., Temmel H., Wilson D.N., Djinovic-Carugo K., Moll I. // *Nucl. Acids Res.* 2015. V. 43. № 1. P. 661–673.
11. Hajnsdorf E., Boni I.V. // *Biochimie.* 2012. V. 94. № 7. P. 1544–1553.
12. Chetverin A.B. // *Mol. Biol. (Mosk.)* 2011. V. 45. № 1. P. 127–137.
13. Blumenthal T. // *Methods Enzymol.* 1979. V. 60. P. 628–638.
14. Bycroft M., Hubbard T.J., Proctor M., Freund S.M., Murzin, A.G. // *Cell.* 1997. V. 88. № 2. P. 235–242.
15. Draper D.E., Reynaldo L.P. // *Nucl. Acids Res.* 1999. V. 27. № 2. P. 381–388.
16. Salah P., Bisaglia M., Aliprandi P., Uzan M., Sizun C., Bontems F. // *Nucl. Acids Res.* 2009. V. 37. № 16. P. 5578–5588.
17. Takeshita D., Yamashita S., Tomita K. // *Nucl. Acids Res.* 2014. V. 42. № 16. P. 10809–10822.
18. Gytz H., Mohr D., Seweryn P., Yoshimura Y., Kutlubaeva Z., Dolman F., Chelchessa B., Chetverin A.B., Mulder F.A., Brodersen D.E., et al. // *Nucl. Acids Res.* 2015. V. 43. № 22. P. 10893–10906.
19. Aliprandi P., Sizun C., Perez J., Mareuil F., Caputo S., Leroy J.L., Odaert B., Laalami S., Uzan M., Bontems F. // *J. Biol. Chem.* 2008. V. 283. № 19. P. 13289–13301.
20. Giraud P., Crechet J., Uzan M., Bontems F., Sizun C. // *Biomol. NMR Assign.* 2015. V. 9. № 1. P. 107–111.
21. Kamen R., Kondo M., Römer W., Weissmann C. // *Eur. J. Biochem.* 1972. V. 31. № 1. P. 44–51.
22. Kamen R. // *RNA Phages.* / Ed. Zinder N.D. Cold Spring Harbor, N.Y.; Cold Spring Harbor Lab. Press, 1975. P. 203–234.
23. Vasilyev N.N., Kutlubaeva Z.S., Ugarov V.I., Chetverina H.V., Chetverin A.B. // *Nat. Commun.* 2013. V. 4. P. 1781.
24. Kolb A., Hermoso J.M., Thomas J.O., Szer W. // *Proc. Natl. Acad. Sci. USA.* 1977. V. 74. № 6. P. 2379–2383.
25. Feix G., Slor H., Weissmann C. // *Proc. Natl. Acad. Sci. USA.* 1967. V. 57. № 5. P. 1401–1408.
26. Ugarov V.I., Demidenko A.A., Chetverin A.B. // *J. Biol. Chem.* 2003. V. 278. № 45. P. 44139–44146.
27. Richey B., Cayley D.S., Mossing M.C., Kolka C., Anderson C.F., Farrar T.C., Record M.T., Jr. // *J. Biol. Chem.* 1987. V. 262. № 15. P. 7157–7164.
28. Kidmose R.T., Vasiliev N.N., Chetverin A.B., Andersen G.R., Knudsen C.R. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. № 24. P. 10884–10889.
29. Takeshita D., Tomita K. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. № 36. P. 15733–15738.
30. Olesen H., Knudsen C., Seweryn P., Brodersen D., Kutlubaeva Z., Chetverin A., Mulder F., Jensen L., Yoshimura Y. // *Acta Cryst.* 2014. V. A70. C1602.
31. Takeshita D., Tomita K. // *Nat. Struct. Mol. Biol.* 2012. V. 19. № 2. P. 229–237.
32. Meyer F., Weber H., Weissmann C. // *J. Mol. Biol.* 1981. V. 153. № 3. P. 631–660.
33. Schuppli D., Miranda G., Qiu S., Weber H. // *J. Mol. Biol.* 1998. V. 283. № 3. P. 585–593.
34. Tomita K. // *Int. J. Mol. Sci.* 2014. V. 15. № 9. P. 15552–15570.

Genome Stability Maintenance in Naked Mole-Rat

I. O. Petrusheva¹, A. N. Evdokimov¹, O. I. Lavrik^{1,2,3*}

¹Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Lavrentjeva Ave. 8, Novosibirsk, 630090, Russia

²Novosibirsk State University, Ministry of education and science, Pirogova Str. 1, Novosibirsk, 630090, Russia

³Altai State University, Ministry of education and science, Lenina Ave. 61, Barnaul, 656049, Russia

*E-mail: lavrik@niboch.nsc.ru

Received: May 12, 2017; in final form August 30, 2017

Copyright © 2017 Park-media, Ltd. 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 naked mole-rat (*Heterocephalus glaber*) is one of the most promising models used to study genome maintenance systems, including the effective repair of damage to DNA. The naked mole-rat is the longest living rodent species, which is extraordinarily resistant to cancer and has a number of other unique phenotypic traits. For at least 80% of its lifespan, this animal shows no signs of aging or any increased likelihood of death and retains the ability to reproduce. The naked mole-rat draws the heightened attention of researchers who study the molecular basis of lengthy lifespan and cancer resistance. Despite the fact that the naked mole-rat lives under genotoxic stress conditions (oxidative, etc.), the main characteristics of its genome and proteome are high stability and effective functioning. Replicative senescence in the somatic cells of naked mole-rats is missing, while an additional p53/pRb-dependent mechanism of early contact inhibition has been revealed in its fibroblasts, which controls cell proliferation and its mechanism of *arf*-dependent aging. The unique traits of phenotypic and molecular adaptations found in the naked mole-rat speak to a high stability and effective functioning of the molecular machinery that counteract damage accumulation in its genome. This review analyzes existing results in the study of the molecular basis of longevity and high cancer resistance in naked mole-rats.

KEYWORDS *Heterocephalus glaber*, cancer resistance, genome and proteome stability, DNA repair.

ABBREVIATIONS *H. glaber* – *Heterocephalus glaber*; *Spalax* – *Spalax galili*; *M. musculus* – *Mus Musculus*; **MLS** – maximum life span; **OD** – oxidative damages; **GPX1** – glutathione peroxidase; **GSH** – reduced glutathione; **ECI** – early contact inhibition; **HA** – hyaluronic acid; **HAS2** – hyaluronan synthase 2; **PARP** – poly(ADP-ribose) polymerase; **PARG** – poly(ADP-ribose) glycohydrolase; **NMR** – naked mole-rat.

INTRODUCTION

DNA damage caused by environmental stress and normal metabolic processes occur daily at a frequency ranging from 1,000 to 1·10⁶ per living cell [1]. As a result, only 0.00017% of the human genome consisting of 3·10⁹ base pairs is damaged, but lesions in essential genes, such as the genes that code for tumor-suppressor proteins, can significantly disturb cellular function. The efficient DNA repair mechanisms that counteract DNA damage accumulation substantially contribute to genome stability maintenance, which is one of the crucial cellular functions. Accumulation of DNA lesions and mutations increases the risk of cancer and is related to aging [2–4]. The defects in DNA repair mechanisms in humans are associated with a number of hereditary diseases [1–4]. Furthermore, the high conservatism of repair pathways allows one to regard the efficiency of DNA repair mechanisms as one of the underlying rea-

sons behind longevity [2–7]. Only a few experimental studies have focused on the search for a correlation between the activity of DNA repair systems and maximum lifespan [8, 9]. The complexity of these studies and their controversial findings may stem from both the imperfect methods used for activity assessment and improper selection of model systems [10].

The naked mole-rat (NMR, *Heterocephalus glaber*) is one of the most promising models used to study genome maintenance systems, including effective repair of DNA damage. The NMR is the longest living of small burrowing mammals. It's native to Southeastern Africa (Ethiopia, Kenya, Somalia) and slightly larger than a mouse. NMR colonies are housed in about 60 zoos worldwide and a number of laboratories. It is one of the ~50 known burrowing herbivorous rodents, a representative of the exceptionally rare true eusocial mammals [11]. Due to the keen interest in the NMR, the journal

Science named this species “Vertebrate of the Year” for 2013. The lifespan of the NMR can reach 32 years, ten times longer than that of the mouse. For most of its lifespan (at least 80%), this animal shows no signs of aging and retains the ability to reproduce [12–14]. It possesses a very efficient mechanism of resistance to cancer, including cancer induced by different stressors [15]. Initial case reports of cancer in naked mole-rats kept in captivity were published in 2016 [16]. The NMR draws the heightened attention of researchers engaged in the study of the molecular basis of lengthy lifespan and cancer resistance.

Noticeable progress in this area was achieved through research performed using laboratory-generated naked mole-rat lineages and bioinformatics and omics approaches [17–21]. The unique features of the metabolism and its regulation attendant to the NMR have been revealed.

In this review, we have made an attempt to analyze the results of these studies, as well as those of research that employed biochemical and molecular genetic approaches, to paint an overview of the possible features of the DNA repair systems in the NMR.

STUDYING THE NMR GENOME AND TRANSCRIPTOME USING BIOINFORMATICS APPROACHES

The advances achieved in high-performance whole-genome sequencing have offered us an unprecedented opportunity to reveal the genetic differences of the NMR that underlie the unique traits of this species. An analysis of the data obtained by primary sequencing of the NMR genome revealed a number of typical and important traits; in particular, ones pointing to its enhanced stability [17]. Another version of the genome was subsequently obtained and analyzed [18], and the web portal Naked Mole Rat Genome Resource (<http://www.nakedmole-rat.org>) was developed. A comparative analysis of the complete NMR and mouse transcriptomes revealed a substantially higher transcription activity for some genes in the NMR. These genes are mainly associated with oxidation/reduction and the mitochondrial function. A record-setting 300- and 140-fold higher expression of the *Epcam* and *α2m* genes coding for the extracellular protein was revealed. The difference between the expression levels of the genes encoding repair proteins in mouse and NMR was not that significant [19].

The first results of a deep sequencing (98.6%) of the genome of a male naked mole-rat were published in 2011 [17]. Back then, the difference in the expression levels of mitochondrial genes and the genes related to the redox system in the NMR and mouse was reported [19]. The sequences of 22,000 NMR genes were predicted using the sequencing data. An analysis of syntenic

regions in NMR and human chromosomes identified 750 gained and 320 lost genes; 739 gained and 448 lost genes were revealed in NMR as compared to a mouse. Among the gained genes, 75.5% showed evidence of transcription, while the list of lost genes included many genes related to ribosome and nucleoside biosynthesis functions. Pseudogenes associated with the visual system, olfaction, spermatogenesis, and protein ubiquitination are predominant among all pseudogenes in NMR. Conversion of these genes to pseudogenes (non-functional genes) correlates with weakened and suppressed physiological functions in NMR [13] and the accumulation of ubiquitinated proteins with age that is less intense than in mouse [22].

A total of 1.87 million heterozygous single-nucleotide polymorphisms (SNPs) were also identified using the Genome Analysis Toolkit (GATK, <https://software.broadinstitute.org/gatk/>). The estimated nucleotide diversity (mean per nucleotide heterozygosity) was 7×10^{-4} , which is much lower than that in mouse and rat populations and is comparable to the nucleotide diversity in humans. The low level of nucleotide diversity may reflect a low effective size of the NMR population, but it may also be due to a high level of inbreeding, a reduced mutation rate, or high efficiency of the repair systems [17]. Genome stability is believed to correlate with a reduced transposon level. Kim et al. [17] demonstrated that only 25% of the NMR genome is represented by transposon-derived repeats (vs. 40% in the human, 37% in mouse, and 35% in rat genomes).

The *Tep1* and *Terf1* genes involved in telomere length regulation belong to the set of positively selected genes in NMR, unlike those in the rat and mouse [23]. Telomere length is short in NMR: the telomeres are shorter than those in laboratory mice or rats and are approximately as long as human telomeres. The *Tert* gene coding for the telomerase catalytic subunit is stably expressed in the somatic cells of NMR at any age. Meanwhile, the telomerase activity is low. A comparative study showed that there is a negative correlation between the levels of telomerase expression and rodent size, since no correlation between telomere length and lifespan has been found [21, 24, 25]. The recent detailed comparison of the genetic structure of telomerase RNA (*hgTerc*) in NMR and other species has revealed two main differences: the A→G replacement in the first loop of pseudoknot P2b-p3 (an equivalent of nucleotide 111 in human telomerase RNA) and the G→A replacement in the CR7-p8b domain (an equivalent of nucleotide 421 in *hTERC*). Two transcription factor binding sites were identified in the promoter regions of the *hgTerc* gene: the ETS family site, which was found to be a conserved element for all the analyzed TER promoters, and the binding site for the SOX17 transcription factor, which

was unique to the NMR gene. The absence of one Sp1 binding site was an additional specific feature of the *Terc* gene in NMR [26]. Hence, the NMR *Terc* gene has a unique polymorphism and promoter structure.

The results of a sequencing of RNA isolated from the brain, liver, and kidneys of a newborn, young (4 year-old), and old (20 year-old) NMRs showed that the expression level changes with age in a very small number of genes. In the human brain, the expression level decreased in 33 genes, while increasing in 21 genes [27]. In NMR, the expression level of 32 of these genes did not significantly vary with age: it was stable for 30 genes and increased to some extent in only two genes (*Cyp46a1* and *Smad3*) [17]. The transcription activity of these human genes decreased with age [27].

Furthermore, Kim et al. [17] performed a bioinformatics analysis of 39 NMR genes encoding a number of proteins associated with G1/S transition, thermogenesis, and the visual function, including cyclin E1 (*Ccne1*), uncoupling protein 1 (*Ucp1*), and γ -crystallin, as well as the proteins that code for the proteins directly involved in DNA metabolism: multifunctional DNA repair enzyme AP endonuclease (APE1), the large subunit of the replication/repair factor RFC1, and topoisomerase TOP2A. TOP2A controls the topologic states of DNA during transcription and, along with TEP1 and TERF1, is part of a 5-protein complex of alternate lengthening of the telomere pathway. A comparison to the orthologs present in the genomes of 36 mammals revealed a divergence in the NMR genome, attesting to the existence of 45 unique amino acid substitutions in the respective proteins [17].

Hence, the first attempted sequencing [17] revealed the important features of the NMR genome, although some of the results were later refined and reconsidered [18, 28, 29]. Thus, the hairless phenotype of NMR was attributed to a replacement of the conserved amino acid residue in the protein associated with hair growth (HR) [17]. This interpretation was based on the fact that such mutations in this codon cause hair loss in mice, rats, and humans. However, two other rodents, the Damaraland mole-rat and guinea pig, also carry this mutation in the *HR* gene but have pelage [29]. The differences between the *HR* genes in NMR and mouse/humans more likely show the phylogenetic divergence from mouse to humans [29, 30]. The differences in the structure of HAS2 (hyaluronan synthase 2) in NMR are attributed to the exceptional resistance of NMR to cancer [31]. However, some of the presumably important mutations found in the gene encoding HAS2 are identical in several species, including guinea pig. These mutations are not always associated with cancer resistance, and their functional sequelae are unknown as of yet [32]. Interestingly, high-molecular-weight hyaluronans

are also synthesized in cancer-resistant long-lived blind mole-rat *Spalax galili*, but its genome carries none of the mutations considered to be key ones in NMR [33, 34]. Furthermore, the conclusion [17] regarding the reduced level of instability source (transportosons) in the NMR genome as compared to those in the mouse and human genomes remains to be adjudicated [28].

A comparative analysis of a group of the genes involved in genome stability maintenance in humans, mouse, and mole-rat has demonstrated that an elevated gene copy number is not typical of the NMR genome [20]. Meanwhile, the *Cebpg* gene coding for the transcription factor involved in DNA repair regulation is represented by three copies; and the *Tinf2* gene of the shelterin complex component, by two copies. Furthermore, the NMR and human genomes, as opposed to the mouse genome, were found to carry the *Rpa4* gene coding for an analogue of the second subunit of the RPA protein that consists of three subunits (RPA1, RPA2, and RPA3) and is involved in many processes related to DNA conversions. Full-length coding sequences of this gene were previously revealed only in the genomes of apes and horse [35]. The RPA4 and RPA2 proteins can be expressed simultaneously, while the ratio between their levels depends on the tissue type. The α RPA heterotrimer (an alternative RPA containing the RPA4 subunit instead of RPA2) cannot maintain SV40 replication (the common model to study replication *in vitro*) but exhibits an increased affinity for damaged DNA and participates in the repair and activation of cell-cycle control (the G2/M stages) [36–38].

The higher quality of genome annotation has made it possible to identify ~1,800 non-coding and ~42,000 coding DNA regions and approximately the same amount of proteins using sequencing data. As a result, NMR was found to exhibit a number of features of the gene sequences associated with cancer resistance and aging [18]. Unique replacements in the fragment of the *p53* gene that encodes the region involved in apoptosis regulation, as well as in the hyaluronan receptor genes CD44 and HMMR, were revealed. Furthermore, NMR *p53* carries the PXXP motifs (P – proline and X – any other amino acid), similar to the PXXP motifs in human *p53*.

Investigation of the genomes and transcriptomes of nine African naked mole rat species has demonstrated that the genes related to tumor suppression, telomere regulation, cell division, RNA repair, and response to stress have been under positive selection in these species [30].

Modern bioinformatics approaches allow one to perform a full-scale targeted comparison of the transcriptomes of gene groups in different animal species. The liver is an organ characterized by a high level of oxida-

tive metabolism and a significant number of spontaneous lesions. MacRae et al. [39] performed a targeted comparison of the expression levels of the genes encoding repair proteins in the liver tissues of long-lived species (humans and NMR) and short-lived mouse. A comparison of a sample consisting of 130 genes revealed that the transcription activity of these genes was higher in the long-lived species. The gene of tumor suppressor p53, the key regulator of excision repair pathways was among the 12 genes whose expression level was at least twice as high both in human and in NMR. Higher expression levels were also shown for the genes encoding the mismatch repair proteins (MSH3) and base excision repair proteins – DNA glycosylase (MUTYH, MBD4, NEIL1, NEIL2 and TDG), the proteins partaking in nonhomologous recombination (NHEJ1, Ku70, DNA polymerase λ – POLL and κ – POLK), and ubiquitin ligase UBE2N.

Most genes encoding DNA repair proteins are constitutively expressed and regulated by post-transcriptional modifications. Nevertheless, transcription of some genes in this group is induced upon genotoxic stress, including the genes coding for the key components of the nucleotide excision repair (NER) pathway: DDB1, DDB2, ERCC1, XPC, ERCC4 (XPF), and ERCC5 (XPG) [40]. A specialized algorithm for signaling pathways [41] was used to demonstrate that the strongest response to genotoxic stress is provided by the pathways controlled by ATM, BRCA1, p53, and PTEN [39].

EARLY CONTACT INHIBITION

A vast body of results of studies focused on the biochemical features of NMR and aimed at searching for the mechanisms underlying the unique phenotypic traits of NMR, including its cancer resistance, has been published. The unique system of early contact inhibition of cell growth discovered in 2009 is one of these mechanisms [31]. Contact inhibition is a key mechanism that arrests cell division when cells reach a density at which they begin to enter into contact with each other or the extracellular matrix [42]. In humans and mice, regular contact inhibition is mediated by membrane proteins and takes place at an upregulated expression of the cyclin-dependent kinase (CDK) inhibitor p27^{Kip1}. P27^{Kip1} binds to cyclin-CDK complexes and arrests cell division at the G1 phase of the cell cycle. The key tumor-suppressor pathways, the Rb and p53 pathways, are activated by products of the *Ink4a* and *Arf* genes [43–46]. Protein p16^{INK4a}, the *Ink4a* gene product, binds to and inhibits CDK 4/6, thus activating Rb [43]. The *Arf* gene product activates p53 by binding to and activating the MDM2 protein. Hence, the *Ink4a* and *Arf* genes play a crucial role in senescence and resistance to cancer [44–48].

Replicative senescence is not typical of NMR fibroblasts, but in cell culture, the latter grow slowly and arrest at a much lower density, thus showing hypersensitivity to the emergence of intercellular contacts. It was shown that there is an additional mechanism that controls cell proliferation termed “early contact inhibition” (ECI). ECI in NMR was initially believed to be associated with increased p16^{INK4a} protein levels [31]. This hypothesis was based on the fact that p16^{INK4a} is not expressed in the NMR mutant cells SFMut that spontaneously form after long-term culture and lose their capability of early contact inhibition. Recombinant DNA (plasmid) carrying the genes encoding mutant forms of the large T-antigen antibody SV40 that inactivates either p53 (LTK1; pSG5 LTK1), or pRb (LTK Δ 434-444, pSG5 LTK Δ 434-444), or the wild-type protein gene (wtLT; pSG5 LT) suppressing the activity of both p53 and pRb were used to demonstrate that, as opposed to mouse fibroblasts, the ability of NMR fibroblasts to ECI after transfection with these DNA decreases when the activity of both suppressor proteins is inhibited. The possibility of standard contact inhibition mediated by p27^{Kip1} only backs up ECI mediated by the kinase inhibitor p16^{INK4a} [31]. Later, it was shown using RNA sequencing data that the protein termed pALTINK4a/b appears in cultured NMR cells and tissues upon expression of the product of alternative splicing of the *p15a*, *p15b* genes and the *Ink/Arf* locus. The pALTINK4a/b protein was revealed in neither mice nor humans. Expression of pALTINK4a/b is induced upon ECI and under stress, such as UV or ionizing radiation, loss of adherence to the substrate, and oncogene expression. Furthermore, pALTINK4a/b is more efficient at inducing cell cycle arrest, thus leaving more time for the cells to overcome the consequences of genotoxic stress, including DNA-damage repair before replication starts. The two-tiered contact inhibition typical of NMR cells (as opposed to mouse and human cells) may contribute to the maintenance of the stability of its genome [49] (Fig. 1).

HIGH-MOLECULAR-WEIGHT HYALURONIC ACID AND ONCOTRANSFORMATION OF NMR CELLS

In accordance with the data reported in [32], early contact inhibition is related to ultra-high-molecular-weight (6–12 MDa) hyaluronans (HA, hyaluronic acid), which are synthesized in NMR tissues and cells and released into the extracellular space. This polysaccharide was previously better known as a component of the extracellular matrix associated with inflammation and cancer. HA fragments of different molecular weights vary in their functions: medium-sized molecules (30–500 kDa) can stimulate cell division, while smaller fragments (< 50 kDa)

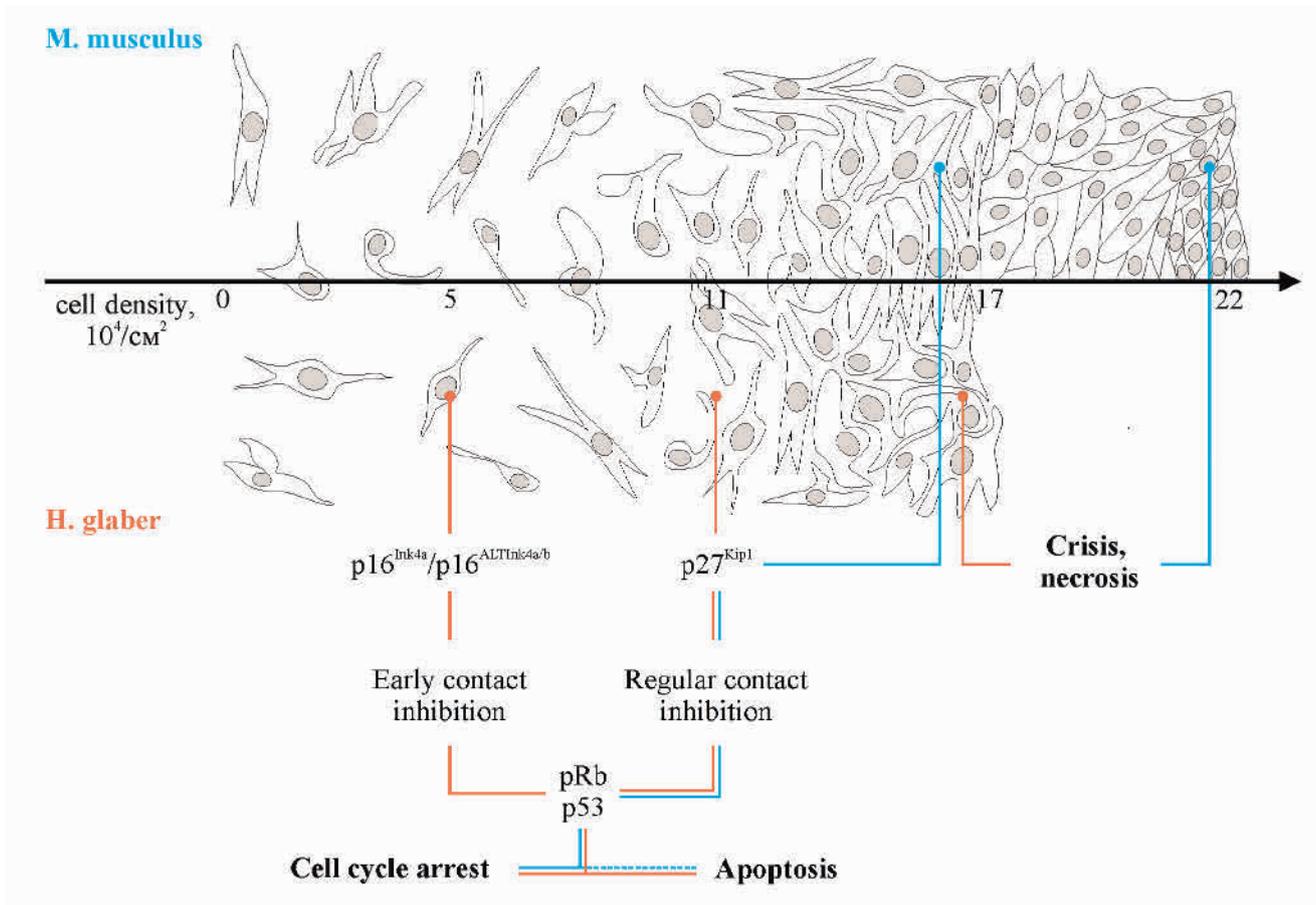


Fig. 1. Two tiers of contact inhibition in naked mole-rat *H. glaber* (based on the data presented in papers [31, 49]). In contrast, mouse only has regular contact inhibition.

can stimulate their migration. Short HA fragments bind to HA receptors, such as CD44 and HMMR, induce inflammation, and activate the signaling pathways that promote survival, migration, and invasion of both tumor and normal cells. Normal human body fluids contain HA 1–8 MDa [50, 51]. In NMR, ultra-high-molecular-weight molecules accumulate due to the low activity of its hyaluronidases and high processivity of hyaluronan synthase 2 (HAS2), whose active site has a specific structure. Substitution of asparagine residues at positions 188 and 301 with serine in HAS2 facilitates the synthesis of ultra-high-molecular-weight HA polymers. Disruptions within the signaling pathways relieving the limitations for initiation of mouse fibroblast oncotransformation do not cause a transformation of NMR cells. If synthesis of high-molecular-weight HA is arrested as a result of HAS2 knockdown or HA degrades rapidly due to an elevated expression of hyaluronidase, NMR cells become susceptible to transformation [32].

EARLY CONTACT INHIBITION AND THE NEW TYPE OF SENESCENCE IN NMR CELLS: *Arf* SUPPRESSION-INDUCED SENESCENCE

Discovered in 2009, the phenomenon of early contact inhibition of fibroblast growth in NMR remains interesting to researchers. In addition to the recently revealed pALTINK4a/b protein, the product of expression of the alternatively spliced form of *Ink4* partaking in ECI [49], another new effect specific to NMR cells has been discovered: *Arf* suppression-induced senescence. The coding sequences of the NMR *Ink4a* and *Arf* *H* genes were identified using conventional cloning procedures and subsequent Sanger sequencing; lentiviral vectors carrying these genes and high-specificity polyclonal antibodies against the respective proteins were constructed. Endogenous *Ink4a* and *Arf* expression in NMR fibroblasts was shown to be upregulated following exposure to DNA-damaging factors or serial passaging [52]. The upregulated *Ink4a* or *Arf* expression caused cell cycle arrest in NMR fibroblasts. Hence,

it was experimentally proved that the genes involved in producing the effect of early contact inhibition play a conserved function of cell cycle inhibitors in NMR [52]. These results were used when studying the mechanisms that suppress tumor development from induced pluripotent stem cells (iPSCs) in NMR [53]. Tumorigenicity of iPSCs was tested for its teratoma-forming potential. NMR iPSCs transplanted into a tested mouse, unlike a number of other stem cells, did not form teratomas; i.e., they were not tumorigenic. This unique feature is based on species-specific activation of the *Arf* oncosuppressor gene and a unique frameshift mutation in the *RAS* (*ERAS*) oncogene expressed by stem cells. The upregulated expression of the *Arf* gene in mouse iPSCs noticeably reduced their tumorigenic potential. The mechanism related to NMR cells that can protect iPSCs and somatic cells against *arf* suppression and tumor formation was found. A special type of senescence, *Arf* suppression-induced senescence, was also revealed in NMR iPSCs. The *Arf*-dependent senescence specific to NMR can act as a backup protection method inducing cell senescence and following death by suppressing *Arf* expression in the cells where this gene used to be suppressed under stress [53].

APOPTOSIS

Apoptosis is one of the mechanisms used for resisting the oncotransformation of cells. The ability of NMR cells to undergo apoptosis in response to genotoxic stress has been insufficiently studied. When investigating the mechanism of ECI, Seluanov et al. [31] demonstrated that the spontaneous apoptosis level in NMR fibroblasts is low (no higher than 7% in skin fibroblasts and 15%, in cultured lung fibroblasts) and is characterized by specific regulation. The count of apoptotic cells in these cultures abruptly increased approximately twofold after transfection with plasmids carrying the genes coding for the mutant forms of the SV40 large T-antigen, pSG5 LTK1 and pSG5 LTA434-444. Transfection of NMR fibroblast cultures with plasmid pSG5 LT carrying the wild-type gene reduced the count of apoptotic cells in them below the control level, while LT had no effect on mouse fibroblasts [31]. In a mouse and humans, apoptosis is also induced to a certain extent when the cell cycle regulator pRb loses its activity [54, 55]. In order to elucidate the mechanism ensuring inhibition of fibroblast growth in NMR upon inactive p53, NMR fibroblasts transfected with these recombinant plasmids were cultured in the presence of the caspase inhibitor Z-Vad-FMK. Growth of fibroblasts transfected with pSG5 LTA434-444 increased in the presence of the apoptosis inhibitor. The mutant protein LTA434-444 inactivates pRb, thus disturbing the mechanism of cell cycle arrest. A combination of pRb

inactivation and apoptosis inhibition in the presence of Z-Vad-FMK results in cell growth, to achieve high confluent density. The growth pattern of cells transfected with pSG5 LTK1 in the presence of the apoptosis inhibitor remained unchanged. Z-Vad-FMK and LTK1 inactivate p53, while pRb remains active: it induces cell cycle arrest and controls cell proliferation [31].

The necrotic cell death pathway is also typical of the cancer-resistant *Spalax* genus of blind mole rats (*Spalax ehrenbergi* and *S. galili*) [56]. In *Spalax*, p53 differs from that in most related mammals by having an arginine-to-lysine substitution at position 174. This specific mutation is frequently detected in human tumors [57]. The arginine-to-lysine substitution affects the properties of the DNA-binding domain of p53. The protein carrying this substitution can induce cell cycle arrest but cannot induce apoptosis. R174K mutation in p53 reduces its ability to activate the apoptotic cascade and activates immuno-inflammatory processes stimulating the development of necrosis induced by interferon- β 1 [55, 56]. Nevertheless, the pathway associated with the activity of p53 is also needed for necrotic cell death in *Spalax* [57–60]. As opposed to *Spalax*, the arginine residue occupies position 174 in *H. glaber* p53, as well as in normal human and mouse cells [18].

The study by Salmon *et al.* focused on the effect of toxic stressors on NMR fibroblasts demonstrated that these cells are more resistant to methyl methanesulfonate, paraquat, and low-glucose media but more sensitive to H₂O₂, UV light, and rotenone compared to mouse fibroblasts [61]. Labinskyy et al. compared the apoptotic response of the cultured arterial endothelial cells of NMR and laboratory mouse to oxidation with H₂O₂ at a concentration ranging from 10⁻⁶ to 10⁻³ M and heat (42°C). The apoptotic response of NMR cells to exposure to H₂O₂ was 3- to 10-fold weaker, while their resistance to heat was higher than that in mouse endothelial cells [62].

TRANSLATIONAL FIDELITY AND SPLIT 28S rRNA

Translational fidelity is one of the key features of the functioning of key NMR systems. With the translation rates close, the number of misincorporated amino acids in NMR fibroblasts is fourfold lower than that in mouse fibroblasts [63]. The translational fidelity in NMR is attributed to the fact that 28S rRNA split into two fragments (that is what NMR 28S rRNA specimens look like after electrophoresis under denaturing conditions) optimizes the folding and/or dynamics of the large ribosomal subunit [63]. Comparison of transcriptomes in a number of rodents showed that degradation of NMR 28S rRNA results from the deletion of a fragment of specific sequence located in the D6 domain of 28S pre-rRNA [64]. In NMR and the Talas tuco-tu-

co (*Ctenomys talarum*), these sequences are characterized by a high degree of sequence conservation. Its 28S rRNA also looks like it is split, but highly accurate protein biosynthesis is not typical of the tuco-tuco [64]. Quite a few species with splitting in RNA are known, but no correlation with their lifespan was revealed. It is also unclear if 28S rRNA is split as a result of specific splicing and fragments are linked into one structure by hydrogen bonds only or if splitting is an artifact emerging under high temperature during RNA isolation or analysis [65–69]. Therefore, attributing the exceptional translational fidelity in NMR to the structural features 28S rRNA is controversial. The high accuracy of protein biosynthesis undoubtedly contributes to the stability of the NMR proteome; however, the features of the underlying molecular mechanisms are yet to be studied. In particular, the first translation stage that is significantly responsible for the accuracy of protein synthesis (tRNA aminoacylation) remains completely unstudied in NMR [70].

OXIDATIVE DAMAGE AND PROTEIN STRUCTURE STABILITY

The oxidative stress theory considers the accumulation of oxidative damage in cells to be one of the factors behind aging. For this reason, the questions regarding the level of oxidative damage and the features of the mechanisms of antioxidant protection in the “long-live” NMR draw researchers’ attention.

Proteins are the main target where oxidative damage emerges. Oxidative events may disrupt the protein structure and functions, in particular by inactivating enzymes and facilitating the formation of protein aggregates containing covalent cross-links. The cysteine thiol groups are characterized by high sensitivity to oxidation, since they can form both reversible (disulfide S–S, sulfenic acid) and irreversible lesions (sulfinic and sulfonic acids) [22]. Other common types of oxidative damage to proteins include carbonylation, irreversible modification of the side chains of proline, arginine, lysine, threonine, cysteine, and histidine residues [71]. Lysates of tissues of different organs of NMR and laboratory mice of respective physiological age are mostly used as model systems to study oxidative damage in proteins [22, 72–76]. The level of oxidative damage in cysteine, the carbonylated protein level, the effect of oxidative damage on the protein structure and function and the activity of a number of the enzymes involved in resisting accumulation of oxidative damage were studied.

Comparison of the activities of glutathione synthetase, catalase, superoxide dismutases, and glutathione peroxidase (GPX1) demonstrated that the activity of all enzymes but GPX1 in liver extract from a young

naked mole-rat was 1.3- to 2-fold higher than that in the extract from the liver of a C57BL/6 mouse of respective physiological age. GPX1 activity in the NMR extract was lower by almost an order of magnitude [72]. In accordance with more recent data, the levels of mRNA *Gpx1* and the respective protein also abruptly decrease in NMR [19, 73].

According to [22], the level of free thiol groups and reversible oxidative damage such as S–S and sulfenic cysteine derivatives in the proteins of young NMR is 1.6-fold higher than that in mouse proteins (C57BL/6). Furthermore, the level of oxidative damage to cysteine in mice increases 3.4-fold with age and the levels of irreversible oxidative damage in cysteine and carbonyl lesions increase, while such changes were not observed in NMR [22, 72–76]. This demonstrates that the performance of the systems counteracting oxidative stress in NMR is more efficient.

An analysis of the levels of protein carbonylation in NMR and mouse tissues demonstrated that triose phosphate isomerase (TPI) and peroxy redoxin 1 (Prdx1) are the main targets for carbonylation in all samples. NMR proteins are characterized by a 1.5-fold higher level of carbonyl damage but better retain enzymatic activity.

The level of carbonyl damage in NMR proteins is 1.5-fold higher, but the proteins better retain their enzyme activity. The specific activity of TPI in the cytosol fraction of a NMR kidney tissue lysate was three times higher than that of mouse. Furthermore, NMR TPI and Prdx1 form fewer covalently cross-linked protein oligomers under oxidative stress (ascorbate/Fe²⁺) [73, 74]. 4,4'-Dianilino-1,1'-binaphthyl-5,5'-disulfonic acid (BisANS) was used as a nonpolar fluorescent probe that interacts with hydrophobic amino acid residues on the surface of protein globules to demonstrate that NMR proteins are much more resistant to the denaturing effect of 1 M urea than mouse proteins. In particular, glyceraldehyde 3-phosphate dehydrogenase (GAPDH), whose active site contains thiol groups, retains 60% of its activity in NMR, as opposed to 10% in mouse GAPDH [22].

Comparison of the distribution of carbonylated proteins over subcellular fractions in long-lived (including NMR) and short-lived mammals demonstrated that the relative level of proteins with oxidative damage in the nucleus in long-lived animals is lower than that in the cytoplasm [9, 76]. This gave grounds for suggesting an inverse correlation between the level of oxidative damage to nuclear proteins and lifespan [76]. However, no detailed studies have been performed to verify this speculation. No data on the level of damage to the proteins involved in DNA repair are available.

The actual situation may be disguised by the results of evaluations that do not distinguish between

the types of damage or damaged molecules (proteins or DNA) originating from different cellular compartments. Lack of consistency in the nature of the methods and agents used in different publications also impedes any analysis. One of the reasons for this can be the fact that a high level of oxidative damage is typical of only certain molecules (molecular classes) and/or cellular compartments [76].

Furthermore, the data regarding the antioxidant status are controversial. Thus, a lack of consistency was observed for the GSH levels in NMR tissue evaluated in different studies: it was reported to be 1.4-fold lower than that in mouse [77], while the level reported in [22] was 1.4-fold higher. This discrepancy does not allow one to compare the antioxidant status of these organisms. Furthermore, the phenomenon of eusociality may also affect the results of the experiments with organ tissue extracts and body fluids from NMR [78].

THE UBIQUITIN-PROTEASOME SYSTEM AND THE UNIVERSAL PROTEASE INHIBITOR, ALPHA-2-MACROGLOBULIN

The ubiquitin-proteasome system plays a crucial role in maintaining the required level of active proteins with a proper structure (proteostasis) in the cell [79].

Evaluation of the proteolytic activity in combination with the results of Western blotting revealed a higher chymotrypsin-like (ChT-L) and trypsin-like (TL) protease activity in 26S and 20S proteasomes in liver tissue extracts of NMR. Specific ChT-L activity of NMR proteasomes was shown to be 3–5 times higher than that of mouse proteasomes [80]. Most of this activity is provided by the activity of 26S proteasomes. Furthermore, 20S proteasomes can perform ubiquitin-independent hydrolysis of proteins containing oxidative damage, such as carbonylated proteins [79]. This may contribute to the maintenance of stable functioning of the NMR proteome in which the level of protein ubiquitination is low and does not increase with age. The levels of the 19S regulatory subunits and immunoproteasome catalytic subunits ($\beta 5i$ and $\beta 2i$) in NMR are also higher than those in mice [80]. Furthermore, the expression level of the key chaperons HSP72, HSP40, and HSP25 is higher in NMR. Two of these chaperons are components of the so-called cytosolic protein factor that protects proteasomes against inhibitors and increases their efficiency [81]. The increased peptidase activity and involvement of chaperones in the protection of proteasomes against inhibitors observed in NMR are chaperone functions that had been previously unknown. All these facts could be indicative of the fact that NMR is characterized by high proteome quality control.

The multifunctional blood plasma protein alpha-2-macroglobulin ($\alpha 2m$) is also associated with proteo-

stasis maintenance. Human $\alpha 2m$ can bind to various cytokines, growth factors (TGF- $\beta 1$, TNF- α , and IL- 1β), and it is a universal inhibitor of proteinases (trypsin, chymotrypsin, elastase, and metalloproteinases). Binding of $\alpha 2m$ -proteinase complexes to the LRP1 (CD91) receptor triggers their quick elimination from the blood and tissues via receptor-mediated endocytosis. This protein is believed to act as a chaperone preventing protein aggregation and to facilitate the retention of zinc in cells (in humans, reduction of the zinc level with age is accompanied by the development of a number of diseases) [82–85]. The level of transcription of the gene coding for $\alpha 2m$ in NMR liver is elevated 140-fold compared to that in mouse liver [19]. Blood plasma concentration of the $\alpha 2m$ protein in NMR is 2–3 times higher than that in humans. This fact is potentially responsible for the proteolytic activity of blood plasma in NMR, which is lower compared to that in humans [86].

Another important feature of NMR is the constant activity of the signaling pathway regulated by the Nrf2 factor, which activates the transcription of over 200 genes involved in the antioxidant and anti-inflammatory response of the organism to endogenous and exogenous stressors [87].

CONCLUSIONS: STABLE GENOME, A STABLE GENE EXPRESSION LEVEL, STABLE PROTEOME, AND EFFICIENT DNA REPAIR

Efficient “functioning” of DNA repair systems is believed to be one of the basics of genome stability maintenance. The typical features of the NMR genome include an increased stability of its structure and function, which are maintained during the entire lifespan. Its protein system (the proteome) is also stable. Translational fidelity, upregulated expression of key chaperons, and permanently active proteasomes in combination with a high $\alpha 2m$ expression level facilitate the maintenance of a pool of efficiently functioning proteins in NMR cells. Resistance to denaturing conditions and the ability to retain their functional activity under permanent oxidative stress were experimentally proved for a number of NMR proteins. All the aforementioned factors and the upregulated expression of a number of genes coding for repair proteins, as well as the intensity of the response of the signaling pathways to damage, provide grounds to expect a high efficiency of the DNA repair system in NMR. In particular, this speculation is consistent with the results of studies using cells of mammals with different lifespans. The rate of UV-induced DNA synthesis in the fibroblasts of the white-footed mouse *Peromyscus leucopus* is 2.5-fold higher than that in mouse (*Mus musculus*) fibroblasts [8]. UV-induced damage in the fibroblasts of long-lived Snell dwarf mice is

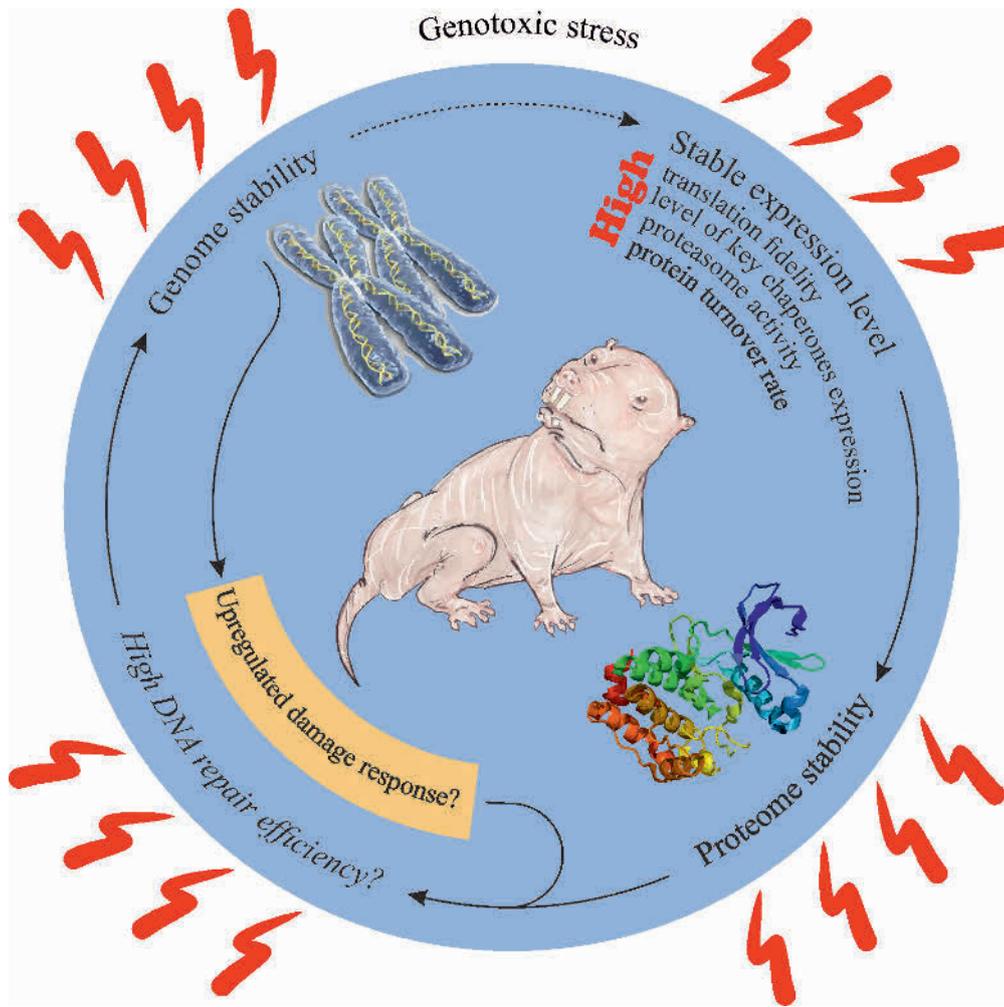


Fig. 2. Stable expression level, proteome stability and upregulated damage response maintain *H. glaber* genome stability upon genotoxic stress.

repaired more efficiently than in the fibroblasts of a mouse with a normal lifespan [9]. Comparison of the activities of poly(ADP-ribose) polymerases (PARPs) in the mononuclear blood leukocytes of 13 mammalian species revealed a positive correlation between the PARP activity level and the maximum lifespan typical of these mammals. In particular, PARP activity in human cells was shown to be five times as high as that in rat cells. Meanwhile, no difference in the levels of the respective proteins was observed and no significant poly(ADP-ribose) polymer degradation was detectable under the experimental conditions, ruling out any interference by poly(ADP-ribose) glycohydrolase (PARG) activity. A hypothesis that a higher poly(ADP-ribosyl)ation capacity might contribute to the efficient maintenance of genome integrity and stability in long-lived species was put forward [88].

It is quite possible that the activity of the poly(ADP-ribosyl)ation processes that regulate different repair mechanisms [89] is also elevated in the extremely long-

lived NMR; however, no experimental evidence to this fact has been obtained thus far.

The unique phenotypic traits of NMR [90] are obviously based on the structural and regulation features of its genome and proteome.

Model systems of different complexities are employed to study these features using an increasingly broad range of methods [91]. Ma et al. [92] have recently conducted a study using fibroblasts from 16 mammalian species to demonstrate that upregulated expression of the genes encoding the proteins associated with DNA repair is typical of long-lived mammals. Modeling and analysis of the stability of the genetic networks linking age, stress resistance, and decelerated physiological senescence have demonstrated that the stability of the simplest model genetic network increases sharply when such a parameter as “efficient repair” is added to the calculations. Furthermore, according to modeling results, the contributions of DNA repair and the processes ensuring the presence of efficiently func-

tioning proteins in the cell (proteostasis maintenance and proteome repair) to the stability of the genetic network are equally significant, while these processes are interrelated [7].

Hence, there is good reason to believe that the molecular machinery counteracting the accumulation of damage in the NMR genome, including the mechanisms of DNA repair, is very efficient. We have made an attempt to illustrate this conclusion with a scheme shown in Fig. 2. However, the lack of studies focused on apoptosis induction under various genotoxic stressors and experimental data regarding the function of DNA

repair systems leave a gap in knowledge regarding the real contribution of these processes to the longevity and cancer resistance of NMR. In this connection, a comparative evaluation of the functional activities of DNA repair systems is a rather important and topical task. ●

The authors are grateful to academician V.P. Skulachev for careful reading of the manuscript and providing valuable comments.

This work was supported by the Russian Science Foundation (grant no. 14-24-00038).

REFERENCES

- Scharer O.D. // *Angew. Chem. Int.* 2003. V. 42. № 26. P. 2946–2974.
- Vijg J., Suh Y. // *Annu. Rev. Physiol.* 2013. V. 75. P. 645–668.
- Hoeijmakers J.H.J. // *N. Engl. J. Med.* 2009. V. 361. № 15. P. 1475–1485.
- Friedberg E.C., Aguilera A., Gellert M., Hanawalt P.C., Hays J.B., Lehmann A.R., Lindahl T., Lowndes N., Sarsin A., Wood R.D. // *DNA Repair (Amst.)*. 2006. V. 5. № 8. P. 986–996.
- Hanawalt P.C. // *Mech. Ageing Dev.* 2008. V. 129. P. 503–505.
- Promislow D.E. // *J. Theor. Biol.* 1994. V. 170. P. 291–300.
- Kogan V., Molodtsov I., Menshikov L.I., Reis R.J.S., Fedichev P. // *Sci. Repts.* 2015. V. 5. № 13589. P. 1–12.
- Hart R.W., Sacher G.A., Hoskins T.L. // *J. Gerontol.* 1979. V. 34. P. 808–817.
- Salmon A.B., Ljungman M., Miller R.A. // *J. Gerontol. A Biol. Sci. Med. Sci.* 2008. V. 63. № 3. P. 219–231.
- Hanawalt P.C. // *Environ. Mol. Mutagen.* 2001. V. 38. № 23. P. 89–96.
- Begall S., Burda H., Schleich C. *Subterranean Rodents: News from Underground*. Berlin Heidelberg: Springer, 2007.
- Gorbunova V., Seluanov A., Mao Z., Hine C. // *Nucl. Acids Res.* 2007. V. 35. № 22. P. 7466–7474.
- Buffenstein R. // *J. Comp. Physiol. B.* 2008. V. 178. № 4. P. 439–445.
- Gorbunova V., Bozzella M.J., Seluanov A. // *AGE*. 2008. V. 30. № 2. P. 111.
- Liang S., Mele J., Wu Y., Buffenstein R., Hornsby P.J. // *Aging Cell.* 2010. V. 9. № 4. P. 626–635.
- Delaney M.A., Ward J.M., Walsh T.F., Chinnadurai S.K., Kerns K., Kinsel M.J., Treuting P.M. // *Vet. Pathol.* 2016. V. 53. № 3. P. 691–696.
- Kim E.B., Fang X., Fushan A.A., Huang Z., Lobanov A.V., Han L., Marino S.M., Sun X., Turanov A.A., Yang P., et al. // *Nature*. 2011. V. 479. № 7372. P. 223–227.
- Keane M., Craig T., Alfoldi J., Berlin A.M., Johnson J., Seluanov A., Gorbunova V., Di Palma F., Lindblad-Toh K., Church G.M., et al. // *Bioinformatics*. 2014. V. 30. № 24. P. 3558–3560.
- Yu C., Li, Y., Holmes A., Szafranski K., Faulkes C.G., Coen C.W., Buffenstein R., Platzer M., de Magalhães J., Church G.M. // *PLoS One*. 2011. V. 6. № 11. P. e26729.
- MacRae S.L., Zhang Q., Lemetre C., Seim I., Calder R.B., Hoeijmakers J., Suh Y., Gladyshev V.N., Seluanov A., Gorbunova V., et al. // *Aging Cell*. 2015. V. 14. № 2. P. 288–291.
- Seluanov A., Hine C., Bozzella M., Hall A., Sasahara T.H., Ribeiro A.A., Catania K.C., Presgraves D.C., Gorbunova V. // *Aging Cell*. 2008. V. 7. № 6. P. 813–823.
- Perez V.I., Buffenstein R., Masamsetti V., Leonard S., Salmon A.B., Mele J., Andziak B., Yang T., Edrey Y., Friguet B., et al. // *Proc. Natl. Acad. Sci. USA*. 2009. V. 106. № 9. P. 3059–3064.
- Rubtsova M.P., Vasilkova D.P., Malyavko A.N., Naraikina Y., Zvereva M.I., Dontsova O.A. // *Acta Naturae*. 2012. V. 4. № 2. P. 44–61.
- Gomes N.M.V., Ryder O.A., Houck M.L., Charter S.J., Walker W., Forsyth N.R., Austad S.N., Venditt C., Pagel M., Shay J.W., et al. // *Aging Cell*. 2011. V. 10. № 5. P. 761–768.
- Gorbunova V., Seluanov A. // *Mech. Ageing Dev.* 2009. V. 130. № 12. P. 3–9.
- Evratov S.A., Smekalova E.M., Golovin A.V., Logvina N.A., Zvereva M.I., Dontsova O.A. // *Acta Naturae*. 2014. V. 6. № 2. P. 41–47.
- Hong M.G., Myers A.J., Magnusson P.K., Prince J.A. // *PLoS One*. 2008. V. 3. № 8. P. e3024.
- Lewis K.N., Soifer I., Melamud E., Roy M., McIsaac R.S., Hibbs M., Buffenstein R. // *Mammalian Genome*. 2016. V. 27. P. 259–278.
- Delsuc F., Tilak M.K. // *Genome Biol. Evol.* 2015. V. 7. № 3. P. 768–774.
- Davies K.T., Bennett N.C., Tsagkogeorga G., Rossiter S.J., Faulkes C.G. // *Mol. Biol. Evol.* 2015. V. 32. № 12. P. 3089–3097.
- Seluanov A., Hine C., Azpurua J., Feigenson M., Bozzella M., Mao Z., Catania K.C., Gorbunova V. // *Proc. Natl. Acad. Sci. USA*. 2009. V. 106. № 46. P. 19352–19357.
- Faulkes C.G., Davies K.T.J., Rossiter S.J., Bennett N.C. // *Biol. Lett.* 2015. V. 11. № 20150185. P. 1–8.
- Tian X., Azpurua J., Hine C., Vaidya A., Myakishev-Rempel M., Ablaeva J., Mao Z., Nevo E., Gorbunova V., Seluanov A. // *Nature*. 2013. V. 499. № 7458. P. 346–349.
- Manov I., Hirsh M., Iancu T.C., Malik A., Sotnichenko N., Band M., Avivi A., Shams I. // *BMC Biol.* 2013. V. 11. № 91. P. 1–17.
- Mullins D.N., Crawford E.L., Khuder S.A., Hernandez D.A., Yoon Y., Willey J.C. // *BMC Cancer*. 2005. V. 5. P. 141.
- Haring S.J., Humphreys T.D., Wold M.S. // *Nucl. Acids Res.* 2010. V. 38. № 3. P. 846–858.
- Kemp M.G., Mason A.C., Carreira A., Reardon J.T., Haring S.J., Borgstahl G.E.O., Kowalczykowski S.C., Sancar A., Wold M.S. // *J. Biol. Chem.* 2010. V. 285. № 7. P. 4788–4797.
- Mason A.C., Roy R., Simmons D.T., Wold M.S. // *Biochemistry*. 2010. V. 49. № 28. P. 5919–5928.

39. MacRae S.L., Croken M.M., Calder R.B., Aliper A., Millholland B., White R.R., Zhavoronkov A., Gladyshev V.N., Seluanov A., Gorbunova V., et al. // *Aging* (Albany NY). 2015. V. 7. № 12. P. 1171–1184.
40. Christmann M., Kaina B. // *Nucl. Acids Res.* 2013. V. 41. № 18. P. 8403–8420.
41. Buzdin A.A., Zhavoronkov A.A., Korzinkin M.B., Venkova, L.S., Zenin A.A., Smirnov P.Y., Borisov N.M. // *Front. Genet.* 2014. V. 5. P. 1–20.
42. Abercrombie M. // *Nature.* 1979. V. 281. № 5729. P. 259–262.
43. Serrano M., Hannon G.J., Beach D. // *Nature.* 1993. V. 366. № 6456. P. 704–707.
44. Hannon G.J., Beach D. // *Nature.* 1994. V. 371. № 6494. P. 257–261.
45. Quelle D.E., Zindy F., Ashmun R.A., Sherr C.J. // *Cell.* 1995. V. 83. № 6. P. 993–1000.
46. Serrano M., Lin A.W., McCurrach M.E., Beach D., Lowe S.W. // *Cell.* 1997. V. 88. № 5. P. 593–602.
47. Gil J., Peters G. // *Nat. Rev. Mol. Cell Biol.* 2006. V. 7. № 9. P. 667–677.
48. Campisi J. // *Aging Cell.* 2008. V. 7. № 3. P. 281–284.
49. Tian X., Azpurua J., Ke Z., Augereau A., Zhang Z.D., Vijj J., Gladyshev V.N., Gorbunova V., Seluanov A. // *Proc. Natl. Acad. Sci. USA.* 2015. V. 112. № 4. P. 1053–1058.
50. Cowman M.K., Lee H.-G., Schwertfeger K.L., McCarthy J.B., Turley E.A. // *Front. Immunol.* 2015. V. 6. № 261. P. 1–8.
51. Schwertfeger K.L., Cowman M.K., Telmer P.G., Turley E.A., McCarthy J.B. // *Front. Immunol.* 2015. V. 6. P. 236.
52. Miyawaki S., Kawamura Y., Hachiya T., Shimizu A., Miura K. // *Inflammation Regeneration.* 2015. V. 35. № 1. P. 42–50.
53. Miyawaki S., Kawamura Y., Oiwa Y., Shimizu A., Hachiya T., Bono H., Koya I., Okada Y., Kimura T., Tsuchiya Y., et al. // *Nat. Commun.* 2016. V. 7. P. 11471.
54. Martel C., Batsche E., Harper F., Cremisi C. // *Cell Death Differ.* 1996. V. 3. P. 285–298.
55. Morgenbesser S.D., Williams B.O., Jacks T., DePinho R.A. // *Nature.* 1994. V. 371. P. 72–74.
56. Shams I., Malik A., Manov I., Joel A., Band M., Avivi A. // *J. Mol. Biol.* 2013. V. 425. № 7. P. 1111–1118.
57. Ashur-Fabian O., Avivi A., Trakhtenbrot L., Adamsky K., Cohen M., Kajakaro G., Joel A., Amariglio N., Nevo E., Rechavi G. // *Proc. Natl. Acad. Sci. USA.* 2004. V. 101. № 33. P. 12236–12241.
58. Avivi A., Ashur-Fabian O., Joel A., Trakhtenbrot L., Adamsky K., Goldstein I., Amariglio N., Rechavi G., Nevo E. // *Oncogene.* 2007. V. 26. № 17. P. 2507–2512.
59. Malik A., Korol A., Weber M., Hankeln T., Avivi A., Band M. // *BMC Genomics.* 2012. V. 13. P. 1–20.
60. Band M., Ashur-Fabian O., Avivi A. // *Cell Cycle.* 2010. V. 9. № 16. P. 3347–3452.
61. Salmon A.B., Sadighi A.A., Buffenstein R., Miller R.A. // *J. Gerontol. A Biol. Sci. Med. Sci.* 2008. V. 63. P. 232–241.
62. Labinskyy N., Csiszar A., Orosz Z., Smith K., Rivera A., Buffenstein R. // *Am. J. Physiol. Heart Circ. Physiol.* 2006. V. 291. P. 2698–2704.
63. Azpurua J., Ke Z., Chen I.X., Zhang Q., Ermolenko D.N., Zhang Z.D., Gorbunova V., Seluanov A. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 43. P. 17350–17355.
64. Fang X., Seim I., Huang Z., Gerashchenko M.V., Xiong Z., Turanov A.A., Zhu Y., Lobanov A.V., Fan D., Yim S.H., et al. // *Cell Reports.* 2014. V. 8. № 5. P. 1354–1364.
65. Melen G.J., Pesce C.G., Rossi M.S., Kornbliht A.R. // *EMBO J.* 1999. V. 18. № 11. P. 3107–3118.
66. Winnebeck E.C., Millar C.D., Warman G.R. // *J. Insect Sci.* 2010. V. 10. P. 1–7.
67. McCarthy S.D., Dugon M.M., Power A.M. // *Peer J.* 2015. V. 3. P. e1436.
68. Ishikawa H., Newburgh R.W. // *J. Mol. Biol.* 1972. V. 64. № 1. P. 135–144.
69. Fujiwara H., Ishikawa H. // *Nucl. Acids Res.* 1986. V. 14. № 16. P. 6393–6401.
70. Yadavalli S.S., Ibba M. // *Adv. Protein Chem. Struct. Biol.* 2012. V. 86. P. 1–43.
71. Nystrom T. // *EMBO J.* 2005. V. 24. P. 1311–1317.
72. Andziak B., O'Connor T.P., Buffenstein R. // *Mech. Ageing Dev.* 2005. V. 126. № 11. P. 1206–1212.
73. Kasaikina M.V., Lobanov A.V., Malinowski M.Y., Lee B.C., Seravalli J., Fomenko D.E., Turanov A.A., Finney L., Vogt S., Park, T.J., et al. // *J. Biol. Chem.* 2011. V. 286. № 19. P. 17005–17014.
74. Andziak B., O'Connor T.P., Qi W., DeWaal E.M., Pierce A., Chaudhuri A.R., van Remmen H., Buffenstein R. // *Aging Cell.* 2006. V. 5. № 6. P. 463–471.
75. De Waal E.M., Liang H., Pierce A., Hamilton R.T., Buffenstein R., Chaudhuri A.R. // *Biochem. Biophys. Res. Commun.* 2013. V. 434. № 4. P. 815–819.
76. Bhattacharya A., Leonard S., Tardif S., Buffenstein R., Fischer K.E., Richardson A., Austad S.N., Chaudhuri A.R. // *Aging Cell.* 2011. V. 10. № 4. P. 720–723.
77. Andziak B., Buffenstein R. // *Aging Cell.* 2006. V. 5. № 6. P. 525–232.
78. Novikov E.A., Kondratyuk E. Yu., Burda H. // *Zoologicheskii zhurnal.* 2015. V. 94. № 1. P. 119–124.
79. Sorokin A.V., Kim E.R., Ovchinnikov L.P. // *Biochemistry (Moscow).* 2009. V. 74. № 13. P. 1411–1442.
80. Rodriguez K.A., Edrey Y.H., Osmulski P., Gaczynska M., Buffenstein R. // *PLoS One.* 2012. V. 7. № 5. P. e35890.
81. Rodriguez K.A., Osmulski P.A., Pierce A., Weintraub S.T., Gaczynska M., Buffenstein R. // *Biochim. Biophys. Acta.* 2014. V. 1842. № 11. P. 2060–2072.
82. Sottrup-Jensen L. // *J. Biol. Chem.* 1989. V. 264. № 20. P. 11539–11542.
83. Birkenmeier G., Muller R., Huse K., Forberg J., Glaser C., Hedrich H., Nicklisch S., Reichenbach A. // *Exp. Neurol.* 2003. V. 184. № 1. P. 153–161.
84. Borth W. // *FASEB J.* 1992. V. 6. № 15. P. 3345–3353.
85. Isaac L., Florido M.P., Fecchio D., Singer L.M. // *Inflamm Res.* 1999. V. 48. № 8. P. 446–452.
86. Thieme R., Kurz S., Kolb M., Debebe T., Holtze S., Morhart M., Huse K., Szafranski K., Platzer M., Hildebrandt T.B., et al. // *PLoS One.* 2015. V. 10. № 6. P. e0130470.
87. Lewis K.N., Wason E., Edrey Y.H., Kristan D.M., Nevo E., Buffenstein R. // *Proc. Natl. Acad. Sci. USA.* 2015. V. 112. № 12. P. 3722–3727.
88. Grube K., Bürkle A. // *Proc. Natl. Acad. Sci. USA.* 1992. V. 89. № 24. P. 11759–11763.
89. Hodyreva S.N., Lavrik O.I. // *Molecular Biology.* 2016. V. 50. № 4. P. 655–673.
90. Skulachev V.P., Holtze S., Vyssokikh M.Y., Bakeeva L.E., Skulachev M.V., Markov A.V., Hildebrandt T.B., Sadovnichii V.A. // *Physiol. Rev.* 2017. V. 97. № 2. P. 699–720.
91. Dziegielewska M., Holtze S., Vole C., Wachter U., Menzel U., Morhart, M., Groth M., Szafranski K., Sahm A., Sponholz C., Dammann, P., Huse, K., Hildebrandt, T., Platzer M. // *Redox Biol.* 2016. V. 8. P. 192–198.
92. Ma S., Upneja A., Galecki A., Tsai Y.M., Burant C.F., Ras-kind S., Zhang Q., Zhang Z.D., Seluanov A., Gorbunova V., et al. // *Elife.* 2016. V. 5. pii: e19130.

Super-Resolution Microscopy in Studying the Structure and Function of the Cell Nucleus

S. S. Ryabichko¹, A. N. Ibragimov¹, L. A. Lebedeva¹, E. N. Kozlov¹, Y. V. Shidlovskii^{1,2*}

¹Institute of Gene Biology RAS, Vavilova Str. 34/5, Moscow, 119334, Russia

²I.M. Sechenov First Moscow State Medical University, Trubetskaya Str. 8, bldg. 2, Moscow, 119048, Russia

*E-mail: yul.biogen@gmail.com

Received: April 03, 2017; in final form November 11, 2017

Copyright © 2017 Park-media, Ltd. 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 recent decades, novel microscopic methods commonly referred to as super-resolution microscopy have been developed. These methods enable the visualization of a cell with a resolution of up to 10 nm. The application of these methods is of great interest in studying the structure and function of the cell nucleus. The review describes the main achievements in this field.

KEYWORDS histone, DNA, super-resolution microscopy, chromatin, chromosome, cell nucleus.

ABBREVIATIONS SRM – super-resolution microscopy; BALM – binding-activated localization microscopy; FISH – fluorescence *in situ* hybridization; PALM – photoactivated localization microscopy; SIM – structured illumination microscopy; SMLM – single-molecule localization microscopy; STED – stimulated emission depletion; STORM – stochastic optical reconstruction microscopy.

INTRODUCTION

The cell nucleus performs the cardinal function of storing and processing hereditary information. For a long time, progress in nuclear structure and function research has been hampered by the lack of adequate techniques to explore intranuclear structures. Today, the most popular method used for studying the nuclear structure is scanning confocal microscopy, which has a resolution limit of about 200 and 500 nm (lateral and axial resolution, respectively). Researchers also use other microscopic techniques that enable them to explore nuclear processes at the molecular level [1]. However, studying the nucleus in a range of 20–200 nm remains technically arduous. Only the development of 3C methods and introduction of the Hi-C method improved understanding of the principles of chromatin organization at this level [2, 3]. In recent decades, another approach, known as super-resolution microscopy (SRM), has been actively advancing and has enabled researchers to achieve a breakthrough in the exploration of the structure and function of the cell nucleus. In SRM, the issue of the diffraction limit is overcome by the use of various technologies, including both technical solutions in the microscope design and computational methods for image reconstruction [4].

THE PRINCIPLES AND CAPABILITIES OF SRM

The most significant advances in super-resolution imaging were achieved using far-field microscopy [5]. 4Pi microscopy uses two opposing objective lenses focused to the same point, which improves axial resolution to 100 nm. In other SRM methods, the diffraction limit is overcome in two ways: The first one is spatial and/or temporal modulation of the transition between two molecular states of a fluorophore, and the second one is a narrowing of the point spread function of an ensemble image of many fluorophores located near each other. The main methods of the first group include stimulated emission depletion (STED), ground state depletion (GSD), structured illumination microscopy (SIM), and some of their combinations with I³M microscopy (4Pi variant) [6, 7]. The main methods of the second group are photoactivated localization microscopy (PALM), fluorescence photoactivation localization microscopy (FPALM), stochastic optical reconstruction microscopy (STORM), and binding-activated localization microscopy (BALM) [7, 8].

The SIM technique uses specifically structured illumination of the object, which doubles the resolution on each axis and provides 100 nm for lateral resolution and 300 nm for axial resolution. In 3D, the resolution increases eightfold. Frequent use of this microscopy

is related to the possibility of preparing samples in a standard way and to apply standard fluorophores. In STED, the sample is simultaneously illuminated by two lasers at different wavelengths, exciting and depleting; this combination of lasers excites fluorescence in a narrow region of a focal spot 50–80 nm in size. This method achieves a resolution of 20 nm in the focal plane and 45 nm in all three directions.

PALM, FPALM, BALM, and STORM belong to single-molecule localization microscopy (SMLM). The sample is excited multiple times with weak laser pulses. Each pulse activates and precisely localizes (based on the point spread function) a small fraction of fluorophores. The image resolution in these methods depends on the ratio of the photon number from an individual fluorophore to the total fluorescence background and, in theory, can reach 1 nm [9].

More detailed information on the principles of these methods, as well as their advantages and disadvantages, can be found in recent reviews [10, 11]. One of the latest SRM techniques is W-4PiSMSN that enables visualization of the whole cell with a resolution of 10–20 nm, with the cell thickness attaining 10 μm [12].

Global nuclear structure is usually visualized via the labeling of histones with fluorescent proteins [13] or binding of fluorescent labels to DNA and RNA [14–16]. Histones can be fused to fluorescent proteins and, also, chemically modified [17, 18]. In SRM experiments, DNA can be labeled with classical dyes, such as Hoechst and DAPI [19]; also, more photostable and photoswitchable fluorophores have been developed [15, 20, 21]. SRM is compatible with click chemistry, which enables the introduction of fluorescent labels into DNA [22, 23]. A technique based on intrinsic nucleotide fluorescence was proposed for the visualization of DNA in [24].

In SRM, specific DNA sequences can be detected using classical fluorescence *in situ* hybridization (FISH) that preserves, in general, the fine nuclear structure [25]. The DNA-PAINT method has been proposed. It enables multicolor imaging with a resolution of more than 10 nm [26]. What's more, methods of multiplexed FISH with probe exchange have also been put forth [27, 28]. Specific DNA sequences can also be localized with programmable gRNA/dCas9 complexes [29].

In SRM-based DNA visualization, the resolution usually hovers around 25 nm, which corresponds to about 70 bp linear double-stranded DNA [30, 31]. A resolution of 70 nm has been achieved in living cells [32]. At present, a resolution of 5 nm in densely packed samples is achievable using oligonucleotide hybridization [33]. Therefore, SRM has opened perspectives for high-resolution optical genome mapping [34]. For example, DNA repeats in the Yq12 heterochromatin

region of the human genome have been studied [35]. The number of trinucleotide repeats in the 5'-untranslated region of the *FMR1* gene has been counted [36]. Single-molecule microscopy combined with Oligopaint probes can help researchers detect single-nucleotide polymorphisms and, thus, distinguish between homologous chromosomes [37]. A complementary method of molecular beacons has been proposed, which enabled the visualization of a 2.5 kb unique DNA sequence [38]. SRM allows one to very accurately measure the volume occupied by individual gene loci in the nuclear space, evaluate their level of compactification [39], and detect the exact spatial relationships between individual loci [40, 41]. Also, epigenetic markers can be very accurately localized in the nucleus [42, 43].

A protein that ends up being detected in the nucleus by SRM is usually synthesized in the cell in fusion form with another fluorescent protein. It is important to note that SRM allows one not only to localize a fluorophore, but also to determine its amount at the localization site with an accuracy of up to one molecule [44–46]. There are methods that allow for simultaneous detection of several proteins (currently, up to nine) in a single sample [47, 48].

SRM has been used in many studies. Below are the main results obtained using the method when exploring the structure and functions of the cell nucleus.

GENERAL ARCHITECTURE OF THE CELL NUCLEUS AND CHROMATIN PACKAGING

The general nuclear architecture has been studied with high resolution in different cell types [49]. For example, there are reports on the reorganization of the cell nucleus structure during myelopoiesis in humans [50], neurogenesis in mice [51], in cells of early embryos and cloned bovine cells [52, 53], and in the cell cycle of yeast [54]. Such studies are often descriptive, because very little is known about the mechanisms of fine nuclear structure formation.

In one of the early studies, the fine organization of the mammalian cell nucleus was explored using 3D-SIM [55]. In the nucleoplasm, channels and lacunas (interchromatin compartment) starting at the nuclear pore and expanding throughout the nuclear space were revealed. Within each chromosome territory, internal regions of higher order chromatin domains are separated from interchromatin by a 100- to 200-nm-thick layer of decondensed transcriptionally active chromatin (peripheral chromatin). The latter is enriched in markers of active transcription and replication, while splicing markers are predominantly localized in the interchromatin compartment. Clusters of RNA polymerase II were revealed, with a significant fraction of transcription being observed outside the clusters. The

SRM and 3C data enabled researchers to generate a nuclear organization model where active chromatin forms clusters and aligns at the boundaries of the network of the intracellular channels pervading most of inactive chromatin [56]. The localization of transcription regulation sites at the periphery of chromatin domains was confirmed by FISH [57].

Exploration of the nuclear periphery in human cells using SIM demonstrated that the number of heterochromatin markers is generally reduced near nuclear pores, but both euchromatin and heterochromatin may form contacts with the pore. Chromatin-modifying enzymes are associated with nuclear pores, which may be indicative of the pore's role in the organization of the nuclear chromatin structure [58].

One of the first studies that used labeled histones demonstrated that the total histone density is different depending on the type of human cells. Global fluctuations in the nuclear histone density on a scale of 1–2 μm were detected using dSTORM [59]. The total histone H2B density in the nuclei of a human U2OS cell line measured by PALM was used to determine the parameters of a nuclear chromatin distribution model and, also, to confirm the fractal globule model [60]. A study of the histone H2B dynamics in HeLa cells using dSTORM showed that histones form clusters that are separated by a distance of about 100 nm and move at a rate of 3 nm/s in the interphase nucleus [18]. In human cells, histone clusters with a mean diameter of 160 nm have been observed using PALM. The formation of clusters was shown to depend on cohesin and internucleosomal interactions, but not on transcription. These domains are present both in mitotic chromosomes and in the interphase nucleus. Perhaps, they are the building blocks of chromosomes [61].

Using PALM, filaments with a diameter of 70 nm were shown to be present in metaphase chromatin in *Drosophila* [62]. The use of STED microscopy to study the chromatin structure in cardiomyocytes revealed 40–70 nm chromatin domains [63]. In mouse and human cells, the nucleosome distribution in DNA has been analyzed in detail using STORM [64]. Nucleosomes were shown to form heterogeneous, different-size groups (clutches) in the nucleus. The mean number of nucleosomes in a group and their density depend on the cell type: in pluripotent cells, these groups are less dense and contain less nucleosomes. RNA polymerase II is predominantly associated with the smallest clutches, while histone H1 and heterochromatin are rich in the largest clutches.

A large-scale study of DNA packing in various chromatin types in *Drosophila* cells was performed using 3D-STORM [65]. Transcriptionally active, inactive, and Polycomb domains were studied. The for-

mer were the least compact, while the last were the densest. The packing degree of these domains varied depending on the domain size: the larger the active domain size, the less dense the domain, whereas the larger the Polycomb domain size, the higher its density. Polycomb domains, unlike the other two types of domains, were demonstrated to be characterized by a high degree of DNA intermixing within the domain and almost complete absence of intermixing with neighboring domains of a different type. Models of chromatin organization domains of different types have been proposed.

Thus, intranuclear chromatin domains differing in size have been described to date. Further research is required to clarify the correlation between intranuclear domains themselves and with the topologically associated domains detected by Hi-C analysis.

HETEROCHROMATIN

The structural details of satellite DNA heterochromatin in aging human cells have been studied using STED microscopy [66]. In both aging and dividing cells, satellite DNA is packed into a set of compact globules separated by linkers. But during aging, the distance between the globules increases. Light-sheet Bayesian super-resolution microscopy (LSBM) helped reveal that the major heterochromatin protein HP1 forms a network in human nuclei [67].

A SIM study demonstrated that, in the interphase, transcriptionally inactive chromatin in yeast cells is less compact than euchromatin. At the same time, highly condensed bodies spanning 50 kb in sequence and flanking inactive telomeric regions were found. The formation of these bodies was independent of the HP1 protein, but it depended on methylation of the H3K36 residue [68].

3D-STORM combined with Oligopaint probes was used to visualize chromatin associated with the Polycomb factor in embryonic stem cells and mouse neuronal progenitor cells [69]. The formation of compact regions in the *Hox* gene locus was established; in this case, removal of the Phc1 protein led to chromatin decompaction. These regions consist of small, discrete domains containing 20–140 kb DNA, which differ from topologically associated domains (TADs). Another study revealed several hundred Polycomb clusters in the nuclei of *Drosophila* cells; the clusters were different from previously studied Polycomb bodies. The number of clusters was dependent on the integrity of the polymerizable SAM-motif of the Ph protein and its content in the cell. It is probable that these clusters form a network of long-range interactions throughout the genome, thereby maintaining the global nuclear architecture [70].

The structure of Barr bodies in mice cells was determined using 3D-SIM [71]. Despite a significant decrease in size, an inactivated X chromosome retained the general structure characteristic of normal chromosome territory. But in this case, the volume of the internal channels was significantly smaller. The same technique was used to determine the localization of the RNA-binding proteins Rbm15, Spen, and Wtap in Xist RNA [72]. A STORM analysis of an inactivated X chromosome in mouse fibroblasts revealed that only 50–100 Xist molecules and about 50 clusters of the PRC2 complex are present on the X chromosome [73].

STRUCTURE OF CONDENSED CHROMOSOMES

Chromatin is located not only along a chromosome (bands on polytene chromosomes), but also radially. A peripheral localization of actively transcribed loci, beginning with the prophase, on the mitotic X chromosome of *Drosophila* males was demonstrated using SIM. In general, in a condensed mitotic chromosome, silent regions are localized closer to the chromosome's axis, while active regions are closer to the surface [74]. 3D-SIM was used to study the differential accessibility of the different loci of a metaphase chromosome in human lymphoblasts [75].

Similarly, SMLM-based visualization of mouse chromosomes at the pachytene stage allowed researcher to identify three chromatin types [76]. The first one is located radially (it carries the active transcription marker H3K4me3) and forms loop-like structures. The second is located along the chromosome axis and carries the H3K27me3 marker. Finally, there is centromeric chromatin carrying the H3K9me3 marker.

3D-SIM was used to investigate the involvement of various condensin forms in chromosome scaffold formation in chicken cells. In condensin I subunit knockout cells, mitotic chromosomes had become shorter and wider and had a diffuse scaffold; in condensin II subunit knockout cells, the scaffold was more defined and the chromosomes were more stretched and lacked axial rigidity [77]. The spatial organization of a meiotic chromosome was studied using various SRM methods [78]. In yeast cells, 3D-SIM demonstrated the formation of a chromosome axis from meiotic cohesin, as well as impairment in the chromosome structure in the absence of cohesin [79, 80].

The so-called T-loops were visualized on mouse telomeres using STORM. The investigation of the loop structure in the setting of various mutations showed the importance of the TRF2 factor in the formation of this structure [81].

Single-molecule microscopy was used to study centromere organization in chicken cells: centromeric chromatin in the chicken was represented by lay-

ers with alternating domains enriched in CENP-A or H3 histones. During mitosis, the CENP-C-dependent mechanism links CENP-A-blocks [82]. In yeast cells, PALM was used to count CENP-A molecules in the centromere; the histone was shown to have accumulated at the centromere in the G2 phase in yeast, unlike metazoan cells [83].

As was shown with SIM in the cells of various plant species, histone CENH3 and modified histone H2AThr120P are incorporated into various nucleosomes that form different domains [84]. In barley, two centromeric histone variants – alpha- and beta-CENH3 – are incorporated into various domains of centromeric chromatin in the interphase, with the incorporation pattern being tissue-specific [85].

TRANSCRIPTION

SRM was used to visualize the processes occurring during the activation of gene transcription. STORM imaging revealed that, upon activation of the mammalian *Hoxd* gene, its locus becomes de-compacted and acquires an elongated configuration [86]. In mouse cells, the beta-globin locus was visualized using 3D-SIM: the inactive locus had several different conformations; during the differentiation of the cells, the locus decreased in size and its structure became more arranged [87]. Local de-condensation of chromatin in active transcription loci was visualized using BALM in chromatin spreads from HeLa cells. Stimulation of the cells led to the appearance of open chromatin regions of about 388 nm in length and about 60 nm in width enriched in the active form of RNA polymerase II [88]. Using the same technique, long-range contacts forming in the genome of an individual cell with the participation of the transcription factors YAP, SRF, and NF-kappaB [89] were visualized.

Single-molecule microscopy, combined with light sheet microscopy, enabled us to perform a quantitative analysis of transcription factories in mammalian cells [90]. More than 70% of the transcription sites were found to contain only one RNA polymerase II molecule, which contradicts the existing models. The dynamics of transcription factories has been studied [31, 91]: these factories are not static structures and can form in the nucleus upon stimulation of the cell. Formation of factories is not blocked by the inhibitors of RNA polymerase II elongation; i.e. the factories form at the transcription initiation stage.

The concept of transcription on an immobilized transcription factory was confirmed using SRM. Induction of human cells with a cytokine leads to closer approximation of two genes in the nucleus, which are remotely located in the genome, with their transcripts also being located near each other [92]. In the

same model, the process of induced transcription on a 221-kb-long gene template was visualized [93]. The transcription factory initially associates with the gene promoter, and then the DNA template is pulled through this transcription factory. In this case, the promoter remains near RNA polymerase for a while, but then it can lose contact, and re-initiation can occur in another factory.

Visualization of transcription on lampbrush chromosomes using dSTORM enabled the imaging of newly synthesized RNA packaging. The splicing and tight packing of mRNA result in the fact that the thickness of a transcribed chromatin loop remains almost unchanged along the active gene [94].

In plants, a study of the distribution of different RNA polymerases II using SIM and PALM revealed networks of these molecules in euchromatin, with different forms constituting different networks [95, 96]. An increased number of polymerase molecules in the nucleus of polyploid cells was found, which was in general proportional to the total number of genes. Probably, plants have another form of transcription organization, different from that in mammalian transcription factories.

The localization of transcription factors in the nucleus has been primarily studied using SMLM. The binding sites of the Sox2 transcription factor in mice are found to form clusters in the nucleus, which plays an important role in the regulation of the transcription of its target genes [97]. Similarly, the human STAT1 factor forms clusters, with their size and number increasing significantly during transition from the G1 to the G2 phase of the cell cycle, as well as upon stimulation of the cell with cytokines [98]. The transcription factor FoxP3 in T cells forms two types of complexes with other factors: the activation complex located closer to the nucleus center and the repression complex located at the nuclear periphery [99]. SRM was used to study the distribution of histone H2A and the chromatin remodeling factor Snf2H subunit [100], as well as the euchromatin protein MAD2L2 [101].

4Pi-microscopy was used to study the internal structure of PML bodies in human cells. The transcription factors Sp100 and PML form a 50- to 100-nm-thick shell of bodies, which is permeable to proteins. The inner region of the bodies contains polymeric chains of the SUMO protein, which serves to concentrate SUMO-binding factors in the bodies [102].

Exploration of the nucleolus in human cells using SMLM showed that, like the nucleus, it has a very heterogeneous structure: the nucleolus contains regions with high and low concentrations of newly synthesized RNAs [103].

REPLICATION

The structure of replication factories in human cells was studied using STEM microscopy with labeled PCNA and RPA proteins [104]. The diameter of the replication factories was found to be below 160 nm, on average. In the early S phase, up to 1,400 factories with two or three replication forks each were detected. The factory size in mouse cells, estimated using SIM and SMI, was 125 nm [105]. In mammalian cells, about 5,000 replication factories, with each of them being a separate replication site, were detected using SIM in the S phase [106].

A study of replication in yeast cells using SIM revealed that clustering of replicons in single replication factories was a stochastic process and varied greatly from cell to cell, but once associated in one factory, replicons remained stably linked [107]. As demonstrated using SIM, replication of peripheral heterochromatin in mammalian cells, which was tightly associated with the lamina, occurred without disassembly of the latter [108].

SRM helped to reveal a new function for the replication factor Cdt1 in the formation of an extended conformation of the kinetochore complex Ndc80 and its stable association with microtubules in mitotic human cells [109]. Data on the location of kinetochore components of a *Drosophila* chromosome and on the role of the Spc105 protein in the assembly of this structure were obtained [110]. Also, the kinetochore structure in yeast [111] and human [112] cells was studied. Visualization of mitosis enabled us to track the switching dynamics of the directional movement of sister kinetochores in the metaphase [113].

Coordination of replication and transcription was studied in mammalian cells [114]. In nucleoli where the rRNA genes are actively transcribed, there is a strong negative correlation between these processes. At the same time, this correlation is absent in the nucleoplasm.

REPAIR AND RECOMBINATION

The repair process was also studied using SRM [115]. The localization pattern of histone gamma-H2AX (a double-strand DNA break marker) and the dynamics of its localization sites have been explored in a number of studies using different methods [116–118]. Also, the mutual arrangement of gamma-H2AX and the Ku repair complex was studied, and the number of molecules in this complex at the repair site was counted [119]. dSTORM was used to elucidate the molecular basis of nonhomologous end joining in human cells: it was found that DNA ends first interact with each other through protein filaments, then the two ends position themselves relative to one another, and ligation occurs [120].

As shown using 3D-SIM in HeLa cells, the pattern of the proteins BRCA1 and 53BP1 inside repair foci was mutually exclusive. Probably, their mutual arrangement determines the choice of a repair pathway [121]. STED-microscopy was used to visualize the repair factors gamma-H2AX, 53BP1, and Rad51 in HeLa cells after exposure to ionizing radiation. The first two proteins form regions whose size depends on the radiation energy (540 nm for strong and 412 nm for weaker radiation); furthermore, these regions display an internal structure and a negative correlation in the distribution of the two proteins. Rad51 forms regions that lack an internal structure, and their size (135 nm) is independent of the radiation energy [122]. As revealed by dSTORM in human cells, the partners BRCA2 and Rad51 are localized in different sites of the repaired DNA, which indicates different dynamics of their interaction with DNA [123]. As shown using STED and 3D-SIM in human cells, the gamma-H2AX break marker is distributed over nearby chromatin loops, and this process is controlled by the CTCF protein. Therefore, the repair focus is a close group of nano-foci, with each of them being a chromatin loop [124].

The synaptonemal complex formed between homologous chromosomes during meiosis was visualized and studied using various SRM methods in barley and wheat [125, 126], mouse [127], *Caenorhabditis elegans* [128], and yeast [129] cells.

NUCLEAR MEMBRANE

The nuclear membrane contains a variety of transmembrane proteins. For their localization, SRM methods with an axial resolution of more than 10 nm have been developed [130, 131].

The structure of the most important nuclear periphery component, the nuclear pore, has been studied with high accuracy using various SRM methods. In different species, the pore size [132] and the position of its main subunits [133, 134] have been determined, and the pore assembly process has been visualized [135]. Accumulation of several thousand images of individual nuclear pores has enabled a structural analysis of the pore with an accuracy of more than 1 nm [136]. The localization and distribution of individual nuclear pore subunits over the membrane have been determined [137–141]. Using SRM, the contacts between nucleoporins and transport receptors have been studied [142], and contact between the nuclear pore and the active locus has been demonstrated [143].

SPEED microscopy with a spatial resolution of 8 nm and a temporal resolution of 2 ms was used to trace transport through the nuclear pore in human cells [144]. Only 36% of mRNA molecules entering the pore

were shown to be successfully exported to the cytoplasm, with a transport time of about 12 ms. The kinetics of mRNA transport through nuclear pores in mouse cells was also studied with a spatial resolution of 26 nm and a temporal resolution of 20 ms [145]. Export was shown to be a three-stage process, including docking to the pore (80 ms), transport (5–20 ms), and release (80 ms) of the transcript.

The investigation of the nuclear periphery demonstrated that there are invaginations in the lamina, which had been indistinguishable using confocal microscopy [146]. Invaginations have been described in the interphase nucleus in different cell types, but their functions remain unclear. They may be involved in mRNA transport [147].

SIM microscopy has enabled an investigation of the relationship between the nuclear lamina and the actin cytoskeleton in human cells. Mechanical pressure of the actin filament has been found to cause the formation of invaginations in the lamina and the emergence of condensed chromatin domains [148].

CONCLUSION

The topicality of light microscopy for the investigation of intranuclear processes has always been high. An intensively developing group of SRM methods opens new perspectives in this field [149]. Theoretically, SRM may achieve a resolution of up to 1 nm: therefore, it is a unique tool to be used to study processes on a scale ranging from an individual molecule to a whole cell. Apart from helping solve basic problems, SRM is applied in the study of processes that occur in cell nuclei in various pathologies. SRM has been used to study Hutchinson-Gilford progeria [150], Alzheimer's disease [151], hypoxia and fasting in cardiomyocytes, oncogenesis [21, 152], and viral infections [150, 151]. SRM-based techniques have been developed for the diagnosis of diseases, in particular cancers [153].

The data presented in this review demonstrate that SRM has significantly expanded our knowledge about how the nucleus functions at various levels of its organization. At the same time, it is obvious how little the potential of this powerful method has been used to date. Undoubtedly, the combination of SRM with other modern methods is poised to become the basis for new discoveries in the biology of the cell nucleus in the near future. ●

This work was supported by a grant of the Ministry of Education and Science of the Russian Federation 14.613.21.0036 (unique identifier RFMEFI61315X0036).

REFERENCES

1. Vivante A., Brozgol E., Bronshtein I., Garini Y. // *Methods*. 2017. V. 123. P. 128–137.
2. 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.
3. Sati S., Cavalli G. // *Chromosoma*. 2017. V. 126. № 1. P. 33–44.
4. Candès E.J., Fernandez-Granda C. // *Commun. Pure Appl. Mathematics*. 2014. V. 67. № 6. P. 906–956.
5. Hell S.W. // *Science*. 2007. V. 316. № 5828. P. 1153–1158.
6. Nienhaus K., Nienhaus G.U. // *J. Mol. Biol.* 2016. V. 428. № 2 Pt A. P. 308–322.
7. Turkowyd B., Virant D., Endesfelder U. // *Analyt. Bioanalyt. Chem.* 2016. V. 408. № 25. P. 6885–6911.
8. Han R., Li Z., Fan Y., Jiang Y. // *J. Genet. Genom. = Yi chuan xue bao*. 2013. V. 40. № 12. P. 583–595.
9. Yildiz A., Forkey J.N., McKinney S.A., Ha T., Goldman Y.E., Selvin P.R. // *Science*. 2003. V. 300. № 5628. P. 2061–2065.
10. Cremer C., Szczurek A., Schock F., Gourram A., Birk U. // *Methods*. 2017. V. 123. P. 11–32.
11. Klementieva N.V., Zagaynova E.V., Lukyanov K.A., Mishin A.S. // *Sovrem. Tehnol. Med.* 2016. V. 8. № 2. P. 130–138.
12. Huang F., Sirinakis G., Allgeyer E.S., Schroeder L.K., Duim W.C., Kromann E.B., Phan T., Rivera-Molina F.E., Myers J.R., Irnov I., et al. // *Cell*. 2016. V. 166. № 4. P. 1028–1040.
13. Watanabe S., Punge A., Hollopeter G., Willig K.I., Hobson R.J., Davis M.W., Hell S.W., Jorgensen E.M. // *Nat. Methods*. 2011. V. 8. № 1. P. 80–84.
14. Benke A., Olivier N., Gunzenhauser J., Manley S. // *Nano Lett.* 2012. V. 12. № 5. P. 2619–2624.
15. Flors C. // *Biopolymers*. 2011. V. 95. № 5. P. 290–297.
16. Flors C. // *J. Microsc.* 2013. V. 251. № 1. P. 1–4.
17. Klein T., Loschberger A., Proppert S., Wolter S., van de Linde S., Sauer M. // *Nat. Methods*. 2011. V. 8. № 1. P. 7–9.
18. Wombacher R., Heidbreder M., van de Linde S., Sheetz M.P., Heilemann M., Cornish V.W., Sauer M. // *Nat. Methods*. 2010. V. 7. № 9. P. 717–719.
19. Szczurek A.T., Prakash K., Lee H.K., Zurek-Biesiada D.J., Best G., Hagmann M., Dobrucki J.W., Cremer C., Birk U. // *Nucleus*. 2014. V. 5. № 4. P. 331–340.
20. Sreedharan S., Gill M.R., Garcia E., Saeed H.K., Robinson D., Byrne A., Cadby A., Keyes T.E., Smythe C., Pellett P., et al. // *J. Am. Chem. Soc.* 2017. V. 139. № 44. P. 15907–15913.
21. Szczurek A., Klewes L., Xing J., Gourram A., Birk U., Knecht H., Dobrucki J.W., Mai S., Cremer C. // *Nucl. Acids Res.* 2017. V. 45. № 8. P. e56.
22. Vranken C., Deen J., Dirix L., Stakenborg T., Dehaen W., Leen V., Hofkens J., Neely R.K. // *Nucl. Acids Res.* 2014. V. 42. № 7. P. e50.
23. Zessin P.J., Finan K., Heilemann M. // *J. Struct. Biol.* 2012. V. 177. № 2. P. 344–348.
24. Dong B., Almossalha L.M., Stypula-Cyrus Y., Urban B.E., Chandler J.E., Nguyen T.Q., Sun C., Zhang H.F., Backman V. // *Proc. Natl. Acad. Sci. USA*. 2016. V. 113. № 35. P. 9716–9721.
25. Markaki Y., Smeets D., Fiedler S., Schmid V.J., Schermelleh L., Cremer T., Cremer M. // *Bioessays*. 2012. V. 34. № 5. P. 412–426.
26. Jungmann R., Avendano M.S., Woehrstein J.B., Dai M., Shih W.M., Yin P. // *Nat. Methods*. 2014. V. 11. № 3. P. 313–318.
27. Schueder F., Strauss M.T., Hoerl D., Schnitzbauer J., Schlichthaerle T., Strauss S., Yin P., Harz H., Leonhardt H., Jungmann R. // *Angew. Chem. Int. Ed. Engl.* 2017. V. 56. № 14. P. 4052–4055.
28. Wang Y., Woehrstein J.B., Donoghue N., Dai M., Avendano M.S., Schackmann R.C.J., Zoeller J.J., Wang S.S.H., Tillberg P.W., Park D., et al. // *Nano Lett.* 2017. V. 17. № 10. P. 6131–6139.
29. Anton T., Bultmann S., Leonhardt H., Markaki Y. // *Nucleus*. 2014. V. 5. № 2. P. 163–172.
30. Baday M., Cravens A., Hastie A., Kim H., Kudeki D.E., Kwok P.Y., Xiao M., Selvin P.R. // *Nano Lett.* 2012. V. 12. № 7. P. 3861–3866.
31. Kamiyama D., Huang B. // *Developmental Cell*. 2012. V. 23. № 6. P. 1103–1110.
32. Benke A., Manley S. // *Chembiochem*. 2012. V. 13. № 2. P. 298–301.
33. Dai M., Jungmann R., Yin P. // *Nat. Nanotechnol.* 2016. V. 11. № 9. P. 798–807.
34. Jeffet J., Kobo A., Su T., Grunwald A., Green O., Nilsson A.N., Eisenberg E., Ambjornsson T., Westerlund F., Weinholt E., et al. // *ACS Nano*. 2016. V. 10. № 11. P. 9823–9830.
35. Weiland Y., Lemmer P., Cremer C. // *Chromosome Res.* 2011. V. 19. № 1. P. 5–23.
36. Stuhlmüller M., Schwarz-Finsterle J., Fey E., Lux J., Bach M., Cremer C., Hinderhofer K., Hausmann M., Hildenbrand G. // *Nanoscale*. 2015. V. 7. № 42. P. 17938–17946.
37. Beliveau B.J., Boettiger A.N., Avendano M.S., Jungmann R., McCole R.B., Joyce E.F., Kim-Kiselak C., Bantignies F., Fonseca C.Y., Erceg J., et al. // *Nat. Commun.* 2015. V. 6. P. 7147.
38. Ni Y., Cao B., Ma T., Niu G., Huo Y., Huang J., Chen D., Liu Y., Yu B., Zhang M.Q., et al. // *Elife*. 2017. V. 6. P. e21660.
39. Hildenbrand G., Rapp A., Spori U., Wagner C., Cremer C., Hausmann M. // *Biophys. J.* 2005. V. 88. № 6. P. 4312–4318.
40. Esa A., Edelmann P., Kreth G., Trakhtenbrot L., Amarioglio N., Rechavi G., Hausmann M., Cremer C. // *J. Microsc.* 2000. V. 199. № 2. P. 96–105.
41. Rauch J., Knoch T.A., Solovei I., Teller K., Stein S., Buiting K., Horsthemke B., Langowski J., Cremer T., Hausmann M., et al. // *Differentiation; Res. Biolog. Diversity*. 2008. V. 76. № 1. P. 66–82.
42. Hewitson T.D., Holt S.G., Tan S.J., Wigg B., Samuel C.S., Smith E.R. // *Front. Pharmacol.* 2017. V. 8. P. 307.
43. Tajbakhsh J., Stefanovski D., Tang G., Wawrowsky K., Liu N., Fair J.H. // *Exp. Cell. Res.* 2015. V. 332. № 2. P. 190–201.
44. Karathanasis C., Fricke F., Hummer G., Heilemann M. // *Chemphyschem: Eur. J. Chem. Phys. Phys. Chem.* 2017. V. 18. № 8. P. 942–948.
45. Wollman A.J., Leake M.C. // *Faraday Discuss.* 2015. V. 184. P. 401–424.
46. Zanicchi F.C., Manzo C., Alvarez A.S., Derr N.D., Garcia-Parajo M.F., Lakadamyali M. // *Nat. Methods*. 2017. V. 14. № 8. P. 789–792.
47. Agasti S.S., Wang Y., Schueder F., Sukumar A., Jungmann R., Yin P. // *Chem. Sci.* 2017. V. 8. № 4. P. 3080–3091.
48. Georgieva M., Cattoni D.I., Fiche J.B., Mutin T., Chamouset D., Nollmann M. // *Methods*. 2016. V. 105. P. 44–55.
49. Ricci M.A., Cosma M.P., Lakadamyali M. // *Curr. Opin. Genet. Dev.* 2017. V. 46. P. 186–193.
50. Hubner B., Lomiento M., Mammoli F., Illner D., Markaki Y., Ferrari S., Cremer M., Cremer T. // *Epigenet. Chromatin*. 2015. V. 8. P. 47.
51. Patel N.S., Rhinn M., Semprich C.I., Halley P.A., Dolle P.,

- Bickmore W.A., Storey K.G. // *PLoS Genet.* 2013. V. 9. № 7. P. e1003614.
52. Popken J., Graf A., Krebs S., Blum H., Schmid V.J., Strauss A., Guengoer T., Zakhartchenko V., Wolf E., Cremer T. // *PLoS One.* 2015. V. 10. № 5. P. e0124619.
53. Popken J., Koehler D., Brero A., Wuensch A., Guengoer T., Thormeyer T., Wolf E., Cremer T., Zakhartchenko V. // *Nucleus.* 2014. V. 5. № 6. P. 542–554.
54. Wang R., Kamgoue A., Normand C., Leger-Silvestre I., Mangeat T., Gadal O. // *J. Cell. Sci.* 2016. V. 129. № 24. P. 4480–4495.
55. Markaki Y., Gunkel M., Schermelleh L., Beichmanis S., Neumann J., Heidemann M., Leonhardt H., Eick D., Cremer C., Cremer T. // *Cold Spring Harb. Symp. Quant Biol.* 2010. V. 75. P. 475–492.
56. Cremer T., Cremer M., Hubner B., Strickfaden H., Smeets D., Popken J., Sterr M., Markaki Y., Rippe K., Cremer C. // *FEBS Lett.* 2015. V. 589. № 20 Pt A. P. 2931–2943.
57. Cremer M., Schmid V.J., Kraus F., Markaki Y., Hellmann I., Maiser A., Leonhardt H., John S., Stamatoyannopoulos J., Cremer T. // *Epigenetics Chromatin.* 2017. V. 10. № 1. P. 39.
58. Fiserova J., Efenberkova M., Sieger T., Maninova M., Uhlirova J., Hozak P. // *J. Cell Sci.* 2017. V. 130. № 12. P. 2066–2077.
59. Bohn M., Diesinger P., Kaufmann R., Weiland Y., Muller P., Gunkel M., von Ketteler A., Lemmer P., Hausmann M., Heermann D.W., et al. // *Biophys. J.* 2010. V. 99. № 5. P. 1358–1367.
60. Recamier V., Izeddin I., Bosanac L., Dahan M., Proux F., Darzacq X. // *Nucleus.* 2014. V. 5. № 1. P. 75–84.
61. Nozaki T., Imai R., Tanbo M., Nagashima R., Tamura S., Tani T., Joti Y., Tomita M., Hibino K., Kanemaki M.T., et al. // *Mol. Cell.* 2017. V. 67. № 2. P. 282–293.
62. Matsuda A., Shao L., Boulanger J., Kervrann C., Carlton P.M., Kner P., Agard D., Sedat J.W. // *PLoS One.* 2010. V. 5. № 9. P. e12768.
63. Mitchell-Jordan S., Chen H., Franklin S., Stefani E., Bentolila L.A., Vondriska T.M. // *J. Mol. Cell. Cardiol.* 2012. V. 53. № 4. P. 552–558.
64. Ricci M.A., Manzo C., Garcia-Parajo M.F., Lakadamyali M., Cosma M.P. // *Cell.* 2015. V. 160. № 6. P. 1145–1158.
65. Boettiger A.N., Bintu B., Moffitt J.R., Wang S., Beliveau B.J., Fudenberg G., Imakaev M., Mirny L.A., Wu C.T., Zhuang X. // *Nature.* 2016. V. 529. № 7586. P. 418–422.
66. Swanson E.C., Rapkin L.M., Bazett-Jones D.P., Lawrence J.B. // *Nucleus.* 2015. V. 6. № 4. P. 254–260.
67. Hu Y.S., Zhu Q., Elkins K., Tse K., Li Y., Fitzpatrick J.A., Verma I.M., Cang H. // *Opt. Nanoscopy.* 2013. V. 2. № 1. P. e7.
68. Matsuda A., Chikashige Y., Ding D.Q., Ohtsuki C., Mori C., Asakawa H., Kimura H., Haraguchi T., Hiraoka Y. // *Nat. Commun.* 2015. V. 6. P. 7753.
69. Kundu S., Ji F., Sunwoo H., Jain G., Lee J.T., Sadreyev R.I., Dekker J., Kingston R.E. // *Mol. Cell.* 2017. V. 65. № 3. P. 432–446.
70. Wani A.H., Boettiger A.N., Schorderet P., Ergun A., Munger C., Sadreyev R.I., Zhuang X., Kingston R.E., Francis N.J. // *Nat. Commun.* 2016. V. 7. P. 10291.
71. Smeets D., Markaki Y., Schmid V.J., Kraus F., Tattermusch A., Cerase A., Sterr M., Fiedler S., Demmerle J., Popken J., et al. // *Epigenetics Chromatin.* 2014. V. 7. P. 8.
72. Moindrot B., Cerase A., Coker H., Masui O., Grijzenhout A., Pintacuda G., Schermelleh L., Nesterova T.B., Brockdorff N. // *Cell. Rep.* 2015. V. 12. № 4. P. 562–572.
73. Sunwoo H., Wu J.Y., Lee J.T. // *Proc. Natl. Acad. Sci. USA.* 2015. V. 112. № 31. P. E4216–4225.
74. Strukov Y.G., Sural T.H., Kuroda M.I., Sedat J.W. // *PLoS Biol.* 2011. V. 9. № 1. P. e1000574.
75. Khan W.A., Rogan P.K., Knoll J.H. // *Mol. Cytogenet.* 2015. V. 8. P. 65.
76. Prakash K., Fournier D., Redl S., Best G., Borsos M., Tiwari V.K., Tachibana-Konwalski K., Ketting R.F., Parekh S.H., Cremer C., et al. // *Proc. Natl. Acad. Sci. USA.* 2015. V. 112. № 47. P. 14635–14640.
77. Green L.C., Kalitsis P., Chang T.M., Cipetic M., Kim J.H., Marshall O., Turnbull L., Whitchurch C.B., Vagnarelli P., Samejima K., et al. // *J. Cell. Sci.* 2012. V. 125. № 6. P. 1591–1604.
78. Carlton P.M. // *Biophys. Rev.* 2013. V. 5. № 4. P. 313–322.
79. Ding D.Q., Haraguchi T., Hiraoka Y. // *Curr. Genet.* 2016. V. 62. № 3. P. 499–502.
80. Ding D.Q., Matsuda A., Okamasa K., Nagahama Y., Haraguchi T., Hiraoka Y. // *Chromosoma.* 2016. V. 125. № 2. P. 205–214.
81. Doksani Y., Wu J.Y., de Lange T., Zhuang X. // *Cell.* 2013. V. 155. № 2. P. 345–356.
82. Ribeiro S.A., Vagnarelli P., Dong Y., Hori T., McEwen B.F., Fukagawa T., Flors C., Earnshaw W.C. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. № 23. P. 10484–10489.
83. Lando D., Endesfelder U., Berger H., Subramanian L., Dunne P.D., McColl J., Klenerman D., Carr A.M., Sauer M., Allshire R.C., et al. // *Open Biol.* 2012. V. 2. № 7. P. 120078.
84. Demidov D., Schubert V., Kumke K., Weiss O., Karimi-Ashtiyani R., Buttler J., Heckmann S., Wanner G., Dong Q., Han F., et al. // *Cytogenet. Genome Res.* 2014. V. 143. № 1–3. P. 150–156.
85. Ishii T., Karimi-Ashtiyani R., Banaei-Moghaddam A.M., Schubert V., Fuchs J., Houben A. // *Chromosome Res.* 2015. V. 23. № 2. P. 277–284.
86. Fabre P.J., Benke A., Joye E., Nguyen Huynh T.H., Manley S., Duboule D. // *Proc. Natl. Acad. Sci. USA.* 2015. V. 112. № 45. P. 13964–13969.
87. van de Corput M.P., de Boer E., Knoch T.A., van Cappellen W.A., Quintanilla A., Ferrand L., Grosveld F.G. // *J. Cell Sci.* 2012. V. 125. № 19. P. 4630–4639.
88. Wang Y., Maharana S., Wang M.D., Shivashankar G.V. // *Sci. Rep.* 2014. V. 4. P. 4477.
89. Wang Y., Ratna P., Shivashankar G.V. // *Sci. Rep.* 2017. V. 7. P. 42422.
90. Zhao Z.W., Roy R., Gebhardt J.C., Suter D.M., Chapman A.R., Xie X.S. // *Proc. Natl. Acad. Sci. USA.* 2014. V. 111. № 2. P. 681–686.
91. Chen X., Wei M., Zheng M.M., Zhao J., Hao H., Chang L., Xi P., Sun Y. // *ACS Nano.* 2016. V. 10. № 2. P. 2447–2454.
92. Papantonis A., Larkin J.D., Wada Y., Ohta Y., Ihara S., Kodama T., Cook P.R. // *PLoS Biol.* 2010. V. 8. № 7. P. e1000419.
93. Larkin J.D., Papantonis A., Cook P.R., Marenduzzo D. // *Nucl. Acids Res.* 2013. V. 41. № 4. P. 2216–2227.
94. Kaufmann R., Cremer C., Gall J.G. // *Chromosome Res.* 2012. V. 20. № 8. P. 1009–1015.
95. Schubert V. // *Cytogenet. Genome Res.* 2014. V. 143. № 1–3. P. 69–77.
96. Schubert V., Weisshart K. // *J. Exp. Botany.* 2015. V. 66. № 6. P. 1687–1698.
97. Liu Z., Legant W.R., Chen B.C., Li L., Grimm J.B., Lavis L.D., Betzig E., Tjian R. // *Elife.* 2014. V. 3. P. e04236.
98. Gao J., Wang F., Liu Y., Cai M., Xu H., Jiang J., Wang H. // *Sci. Rep.* 2015. V. 5. P. 9045.

99. Kwon H.K., Chen H.M., Mathis D., Benoist C. // *Nat. Immunol.* 2017. V. 18. № 11. P. 1238–1248.
100. Gunkel M., Erdel F., Rippe K., Lemmer P., Kaufmann R., Hormann C., Amberger R., Cremer C. // *Biotechnol. J.* 2009. V. 4. № 6. P. 927–938.
101. Rahjouei A., Pirouz M., Di Virgilio M., Kamin D., Kessel M. // *Stem Cell Repts.* 2017. V. 8. № 4. P. 813–821.
102. Lang M., Jegou T., Chung I., Richter K., Munch S., Udvarhelyi A., Cremer C., Hemmerich P., Engelhardt J., Hell S.W., et al. // *J. Cell. Sci.* 2010. V. 123. № 3. P. 392–400.
103. Szczurek A., Xing J., Birk U.J., Cremer C. // *Front. Genet.* 2016. V. 7. P. 114.
104. Cseresnyes Z., Schwarz U., Green C.M. // *BMC Cell. Biol.* 2009. V. 10. P. 88.
105. Baddeley D., Chagin V.O., Schermelleh L., Martin S., Pombo A., Carlton P.M., Gahl A., Domaing P., Birk U., Leonhardt H., et al. // *Nucl. Acids Res.* 2010. V. 38. № 2. P. e8.
106. Chagin V.O., Casas-Delucchi C.S., Reinhart M., Schermelleh L., Markaki Y., Maiser A., Bolius J.J., Bensimon A., Fillies M., Domaing P., et al. // *Nat. Commun.* 2016. V. 7. P. 11231.
107. Saner N., Karschau J., Natsume T., Gierlinski M., Retkute R., Hawkins M., Nieduszynski C.A., Blow J.J., de Moura A.P., Tanaka T.U. // *J. Cell. Biol.* 2013. V. 202. № 7. P. 1001–1012.
108. Zhironkina O.A., Kurchashova S.Y., Bratseva A.L., Cherepanynets V.D., Strelkova O.S., Belmont A.S., Kireev, II. // *Tsitologiya.* 2014. V. 56. № 12. P. 899–906.
109. Varma D., Chandrasekaran S., Sundin L.J., Reidy K.T., Wan X., Chasse D.A., Nevis K.R., DeLuca J.G., Salmon E.D., Cook J.G. // *Nat. Cell. Biol.* 2012. V. 14. № 6. P. 593–603.
110. Venkei Z., Przewloka M.R., Ladak Y., Albadri S., Sossick A., Juhasz G., Novak B., Glover D.M. // *Open Biol.* 2012. V. 2. № 2. P. 110032.
111. Joglekar A.P., Bloom K., Salmon E.D. // *Curr. Biol.* 2009. V. 19. № 8. P. 694–699.
112. Wan X., O'Quinn R.P., Pierce H.L., Joglekar A.P., Gall W.E., DeLuca J.G., Carroll C.W., Liu S.T., Yen T.J., McEwen B.F., et al. // *Cell.* 2009. V. 137. № 4. P. 672–684.
113. Burroughs N.J., Harry E.F., McAinsh A.D. // *Elife.* 2015. V. 4. P. e09500.
114. Smirnov E., Borkovec J., Kovacic L., Svidenska S., Schrofel A., Skalnikova M., Svindrych Z., Krizek P., Ovesny M., Hagen G.M., et al. // *J. Struct. Biol.* 2014. V. 188. № 3. P. 259–266.
115. Uphoff S., Kapanidis A.N. // *DNA Repair (Amst.)*. 2014. V. 20. P. 32–40.
116. Bach M., Savini C., Krufczik M., Cremer C., Rosl F., Hausmann M. // *Int. J. Mol. Sci.* 2017. V. 18. № 8. P. 1726.
117. Bewersdorf J., Bennett B.T., Knight K.L. // *Proc. Natl. Acad. Sci. USA.* 2006. V. 103. № 48. P. 18137–18142.
118. D'Abrantes S., Gratton S., Reynolds P., Kriechbaumer V., McKenna J., Barnard S., Clarke D., Botchway S.W. // *Radiat. Res.* 2017. In press.
119. Britton S., Coates J., Jackson S.P. // *J. Cell. Biol.* 2013. V. 202. № 3. P. 579–595.
120. Reid D.A., Keegan S., Leo-Macias A., Watanabe G., Strande N.T., Chang H.H., Oksuz B.A., Fenyo D., Lieber M.R., Ramsden D.A., et al. // *Proc. Natl. Acad. Sci. USA.* 2015. V. 112. № 20. P. E2575–2584.
121. Chapman J.R., Sossick A.J., Boulton S.J., Jackson S.P. // *J. Cell. Sci.* 2012. V. 125. № 15. P. 3529–3534.
122. Reindl J., Girst S., Walsh D.W., Greubel C., Schwarz B., Siebenwirth C., Drexler G.A., Friedl A.A., Dollinger G. // *Sci. Rep.* 2017. V. 7. P. 40616.
123. Sanchez H., Paul M.W., Grosbart M., van Rossum-Fikkert S.E., Lebbink J.H.G., Kanaar R., Houtsmuller A.B., Wyman C. // *Nucl. Acids Res.* 2017. V. 45. № 8. P. 4507–4518.
124. Natale F., Rapp A., Yu W., Maiser A., Harz H., Scholl A., Grulich S., Anton T., Horl D., Chen W., et al. // *Nat. Commun.* 2017. V. 8. P. 15760.
125. Colas I., Darrier B., Arrieta M., Mittmann S.U., Ramsay L., Sourdille P., Waugh R. // *Front. Plant Sci.* 2017. V. 8. P. 1235.
126. Phillips D., Nibau C., Wnetrzak J., Jenkins G. // *PLoS One.* 2012. V. 7. № 6. P. e39539.
127. Qiao H., Chen J.K., Reynolds A., Hoog C., Paddy M., Hunter N. // *PLoS Genet.* 2012. V. 8. № 6. P. e1002790.
128. Sato-Carlton A., Li X., Crawley O., Testori S., Martinez-Perez E., Sugimoto A., Carlton P.M. // *PLoS Genet.* 2014. V. 10. № 10. P. e1004638.
129. Lao J.P., Cloud V., Huang C.C., Grubb J., Thacker D., Lee C.Y., Dresser M.E., Hunter N., Bishop D.K. // *PLoS Genet.* 2013. V. 9. № 12. P. e1003978.
130. Chizhik A.M., Ruhlandt D., Pfaff J., Karedla N., Chizhik A.I., Gregor I., Kehlenbach R.H., Enderlein J. // *ACS Nano.* 2017. In press.
131. Mudumbi K.C., Schirmer E.C., Yang W. // *Nat. Commun.* 2016. V. 7. P. 12562.
132. Huve J., Wesselmann R., Kahms M., Peters R. // *Biophys. J.* 2008. V. 95. № 2. P. 877–885.
133. Loschberger A., Franke C., Krohne G., van de Linde S., Sauer M. // *J. Cell. Sci.* 2014. V. 127. № 20. P. 4351–4355.
134. Loschberger A., van de Linde S., Dabauvalle M.C., Rieger B., Heilemann M., Krohne G., Sauer M. // *J. Cell. Sci.* 2012. V. 125. № 3. P. 570–575.
135. Otsuka S., Szymborska A., Ellenberg J. // *Meth. Cell. Biol.* 2014. V. 122. P. 219–238.
136. Szymborska A., de Marco A., Daigle N., Cordes V.C., Briggs J.A., Ellenberg J. // *Science.* 2013. V. 341. № 6146. P. 655–658.
137. Chatel G., Desai S.H., Mattheyses A.L., Powers M.A., Fahrenkrog B. // *J. Struct. Biol.* 2012. V. 177. № 1. P. 81–89.
138. Dahan-Pasternak N., Nasereddin A., Kolevzon N., Pe'er M., Wong W., Shinder V., Turnbull L., Whitchurch C.B., Elbaum M., Gilberger T.W., et al. // *J. Cell. Sci.* 2013. V. 126. № 14. P. 3055–3069.
139. Gottfert F., Wurm C.A., Mueller V., Berning S., Cordes V.C., Honigmann A., Hell S.W. // *Biophys. J.* 2013. V. 105. № 1. P. L01–03.
140. Kinoshita Y., Kalir T., Dottino P., Kohtz D.S. // *PLoS One.* 2012. V. 7. № 4. P. e36137.
141. Ma J., Kelich J.M., Junod S.L., Yang W. // *J. Cell. Sci.* 2017. V. 130. № 7. P. 1299–1306.
142. Ma J., Goryaynov A., Yang W. // *Nat. Struct. Mol. Biol.* 2016. V. 23. № 3. P. 239–247.
143. Rohner S., Kalck V., Wang X., Ikegami K., Lieb J.D., Gasser S.M., Meister P. // *J. Cell. Biol.* 2013. V. 200. № 5. P. 589–604.
144. Ma J., Liu Z., Michelotti N., Pitchiaya S., Veerapaneni R., Androsavich J.R., Walter N.G., Yang W. // *Nat. Commun.* 2013. V. 4. P. 2414.
145. Grunwald D., Singer R.H. // *Nature.* 2010. V. 467. № 7315. P. 604–607.
146. Schermelleh L., Carlton P.M., Haase S., Shao L., Winoto L., Kner P., Burke B., Cardoso M.C., Agard D.A., Gustafsson M.G., et al. // *Science.* 2008. V. 320. № 5881. P. 1332–1336.
147. Schoen I., Aires L., Ries J., Vogel V. // *Nucleus.* 2017. P. 506–514.
148. Versaavel M., Braquenier J.B., Riaz M., Grevesse T., Lan-

REVIEWS

- toine J., Gabriele S. // *Sci. Rep.* 2014. V. 4. P. 7362.
149. Cremer C., Birk U. // *Front. Physics.* 2016. V. 4. P. 11.
150. Chojnowski A., Ong P.F., Wong E.S., Lim J.S., Mutalif R.A., Navasankari R., Dutta B., Yang H., Liow Y.Y., Sze S.K., et al. // *Elife.* 2015. V. 4. P. e07759.
151. Garcia A., Huang D., Righolt A., Righolt C., Kalaw M.C., Mathur S., McAvoy E., Anderson J., Luedke A., Itorralba J., et al. // *J. Cell. Physiol.* 2016. V. 232. № 9. P. 2387–2395.
152. Kirmes I., Szczurek A., Prakash K., Charapitsa I., Heiser C., Musheev M., Schock F., Fornalczyk K., Ma D., Birk U., et al. // *Genome Biol.* 2015. V. 16. P. 246.
153. Hausmann M., Ilic N., Pilarczyk G., Lee J.H., Logeswaran A., Borroni A.P., Krufczik M., Theda F., Waltrich N., Bestvater F., et al. // *Int. J. Mol. Sci.* 2017. V. 18. № 10. P. 2066.

Translational Cross-Activation of the Encapsidated RNA of Potexviruses

M. V. Arkhipenko*, N. A. Nikitin, E. K. Donchenko, O. V. Karpova, J. G. Atabekov

Biology Department, Lomonosov Moscow State University, 1 bld. 12 Leninskie gory, Moscow, 119234, Russia

*E-mail: armar@genebee.msu.ru

Received: March 02, 2017; in final form October 03, 2017

Copyright © 2017 Park-media, Ltd. 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 We had shown the genomic RNA of potexviruses potato virus X and the alternanthera mosaic virus to be inaccessible *in vitro* to ribosomes while in intact virion form, but the RNAs can be translationally activated following the binding of movement protein 1 (MP1) to virus particles. Here, we present the results of the follow-up study targeting two more potexvirus species – the Narcissus mosaic virus and the Potato aucuba mosaic virus. We found encapsidated potexviral RNA to share common translational features *in vitro* and the MP1 to be potent over homological virions of its “own” species and over heterological virions of other species, as well exhibiting selective specificity. Reciprocal cross-activation is observed among viral species phylogenetically either close or distant. There is direct evidence that MP1 binding to the end of the virion is necessary, but not sufficient, for translational activation of encapsidated RNA.

KEYWORDS plant viruses, genomic RNA, cross-activation, potexviruses, translational activation, movement protein 1.

ABBREVIATIONS CP – coat protein, PAMV – potato aucuba mosaic virus, AltMV – alternanthera mosaic virus, NMV – narcissus mosaic virus, MP1 – movement protein 1, PVX – potato virus X.

INTRODUCTION

We have shown in our previous study that the genomic RNA of the Potexvirus-belonging potato virus X (PVX) is inaccessible to ribosomes in intact virions *in vitro* while being rendered into translatable form upon phosphorylation of the PVX coat protein (CP), forming the virion or binding to PVX movement protein 1 (MP1). We suggest that two mechanisms of translation activation operate at different stages of the infection. The encapsidated RNA becomes accessible to ribosomal translation in primary infected cells following the phosphorylation of the PVX coat protein. The MP1 produced during the course of the infection binds to the PVX virions' end, producing the virus transport form and, thus, activating the encapsidated RNA [1–3].

We clarified subsequently that MP1 interacts with the terminal CP molecules of the PVX helical virions corresponding to the 5' end of PVX RNA, while not contacting the genomic RNA itself. Phosphorylation and the MP1 interaction engage different regions of the PVX CP: the former targets the 19 N-terminal amino acids, whereas the latter is restricted to the 10/18 amino acids long C-terminal fragment accessible to the MP1 at the end of the virion [4, 5].

The MP1 binding to intact PVX virions results in a destabilization of the entire helical PVX structure, leading in turn to a conformational shift from the sta-

ble non-translatable form to a metastable form, where the 5' end of the PVX RNA is accessible to ribosomes. So, the MP1, being a component of the PVX transport form, may be perceived as a mediator of the virion-packed genomic RNA translation [1, 6, 7].

A set of deletion mutants was used to reveal the virion-binding MP1 motif located between the amino acid positions 112 and 122. Noteworthy, the MP1-CP interaction is necessary but not sufficient to translationally activate encapsidated RNA. The MP1-dependent translational activation is abolished upon the removal of a MP1 fragment not interfering with virion binding or the protein phosphorylation, probably due to conformational changes in the protein molecule [5].

The genomic RNA of another potexvirus, the alternanthera mosaic virus (AltMV, strain AltMV-MU), is non-translatable *in vitro* in the intact virion but, similar to PVX RNA, can be rendered translatable by the phosphorylation of the virion coat protein or by interaction with AltMV MP1 [8, 9]. The observation of such a similarity raised a question: whether PVX MP1 can activate genomic virion-packed AltMV RNA and vice versa. It was shown that encapsidated AltMV RNA is efficiently translationally activated by PVX MP1 [8]. Moreover, the opposite proved true also – the virion-packed PVX RNA is translationally activated by AltMV MP1 [10].

One may consider, based on the presented data, the translational features of encapsidated RNA, as well as the translation activation pathway, to be shared across the genus *Potexvirus*.

The present study deals with the translational features of two more potexviruses: the narcissus mosaic virus (NMV) and the potato aucuba mosaic virus (PAMV). We looked into whether the corresponding MP1s are capable of translationally activating the encapsidated RNA of the four *Potexvirus* members: namely, PVX, AltMV, NMV, and PAMV.

MATERIALS AND METHODS

Virus isolation and viral RNA extraction

The PVX, NMV, and PAMV preparations were isolated from infected *Datura stramonium* L. plants as described in [1]. The AltMV preparations were isolated from infected *Portulaca grandiflora* plants as previously described in [8]. The viral RNA was prepared using the phenol method with modifications [11].

Production of mutant MP1

Recombinant PVX and AltMV MP1 molecules were constructed as described in [1, 10].

The NMV MP1 and PAMV MP1 expressing plasmids were constructed using the pQE30 vector (Qiagen). The corresponding coding regions, supplemented with His₆ tags, were amplified on the the NMV and PAMV viral RNA templates using the following primers: NMV-forward-BamHI(+) 5'-acacggatccatggactgtaagta-3', NMV-reverse-PstI(-) 5'-acacctgcagcgtagtaacaggtg-3', Auc-forward-BamHI(+) 5'-acatggatccggaatggaatat-3', and Auc-reverse-PstI(-) 5'-acacctgcagatcagtctaaat-3'. The *Escherichia coli* M15 [pREP4] strain was transformed with the constructs. The recombinant proteins were purified following expression induction by chromatography on a Ni-NTA agarose. The SDS PAGE analysis of the protein samples using a 8-20% gel revealed a single band corresponding to either NMV MP1 (26.7 kDa) or PAMV MP1 (27.2 kDa).

Translation in vitro

The translation was performed in a wheat germ extract cell-free system (Promega), following a modified manufacturer protocol as described [12]. The RNA input was 40 µg/µl. The study of translational activation was performed with a PVX to recombinant MP1 molar ratio of 1 : 100, i.e., 1 µg RNA (20 µg of virus) per 1.4 µg MP1.

Immune electron microscopy

The immune electron microscopic observation was done as described in [13]. Polyclonal antisera to the PVX MP1, AltMV MP1, and PAMV MP1 were used as

primary antibodies according to [10]. Gold-conjugated (12 nm) antibodies were used as secondary antibodies. The samples were contrasted with 2% aqueous uranyl acetate. The stained samples were examined under a JEOL JEM-1011 transmission electron microscope (JEOL, Japan) at 80 kV. Images were acquired using a Gatan Erlangshen ES500W digital camera and the Gatan Digital Micrograph software.

RESULTS AND DISCUSSION

To study the translational features of *Potexvirus* encapsidated RNA, we propagated, harvested, and purified NMV and PAMV viral preparations, accompanied by these viruses' recombinant MP1 samples. We found the translational properties of encapsidated NMV and PAMV RNA to not differ from those of PVX and AltMV. Our findings show virion-packed NMV and PAMV RNA to be nontranslatable *in vitro* (Fig. 1, lane 2), but it could be rendered translatable following exposure to its own MP1 (Fig. 1, lane 3). The same had been demonstrated for AltMV and PVX RNA (Fig. 1, positive controls).

We reported previously on the translational cross-activation of encapsidated AltMV and PVX RNA after the interaction of PVX with AltMV MP1 and vice versa [8, 10]. We aimed to further investigate in this study whether cross-activation exists among other potexviruses. Our experiments revealed that PAMV virion-packed RNA can be rendered translationally active upon exposure to NMV MP1 (Fig. 2, 5). Worthwhile, the treatment of encapsidated PAMV RNA with PVX MP1 (Fig. 2, 4) does not render it translationally active contrary to AltMV RNA [8]. The same is true in the inverse situation: PVX virion-packed RNA treated with PAMV MP1 remains inaccessible to ribosomes (Fig. 2, 6).

A different pattern was observed when analyzing the translation activation of NMV and PVX virion-packed RNA, while NMV MP1 appeared incapable of activating PVX RNA (Fig. 3, 4), whereas PVX MP1 activated encapsidated NMV RNA (Fig. 3, 3). No reciprocal cross-activation exists between NMV and PVX, as is the case for NMV and AltMV; i.e., NMV virion-packed RNA remains nontranslatable when exposed to AltMV MP1 (Fig. 3, 7); AltMV virion-packed RNA, in contrast, does undergo activation under treatment with NMV MP1 (Fig. 4, 4) or PVX MP1 (Fig. 4, 3) [8]. These observations are the first pieces of evidence of nonreciprocal cross-activation of encapsidated potexviral RNA.

We also tested whether PAMV MP1 is capable of activating encapsidated NMV or AltMV RNA. This movement protein proved to activate the translation of NMV viral RNA but not AltMV viral RNA. Similarly,

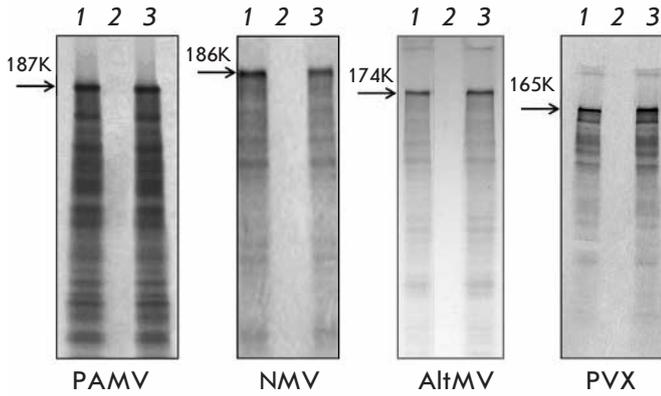


Fig. 1. Translational activation of encapsulated PAMV and NMV RNA *in vitro*. PAMV (A), NMV (B), AltMV, (C) and PVX (D) genomic RNA (lanes 1); encapsidated RNA (lanes 2); encapsidated RNA incubated with MP1 (lanes 3). The arrowheads indicate the position of the PAMV replicase (187K), NMV replicase (186K), AltMV replicase (174K), and PVX replicase (165K). Electrophoretic analysis of ³⁵S-labeled translation products.

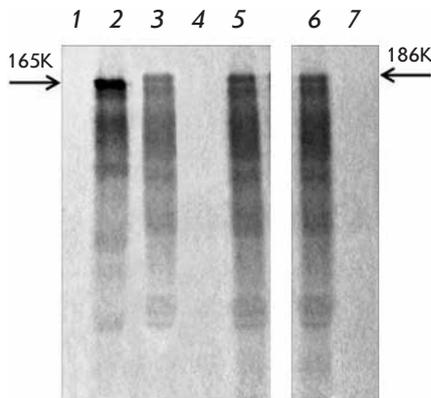


Fig. 3. Translational activation of genomic NMV RNA within viral particles *in vitro*. Encapsidated NMV RNA (lane 1); encapsidated NMV RNA with PVX MP1 (lane 3), NMV MP1 (lane 5), and AltMV MP1 (lane 7). Purified NMV genomic RNA genome as a positive control (lane 6). Translation of encapsidated PVX RNA incubated with NMV MP1 (lane 4) and PVX MP1 as a positive control (lane 2). The arrowheads indicate the position of NMV (186K) and PVX (165K) RNA polymerases. Electrophoretic analysis of ³⁵S labeled translation products.

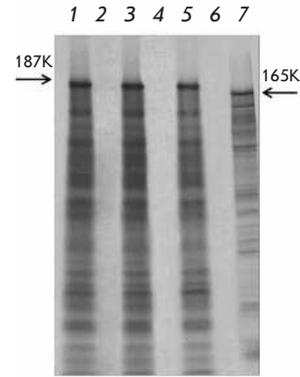


Fig. 2. Translational activation of genomic PAMV RNA within viral particles *in vitro*. PAMV genomic RNA (lane 1); encapsidated PAMV RNA (lane 2); encapsidated PAMV RNA incubated with MP1 of PAMV (lane 3), PVX MP1 (lane 4), NMV MP1 (lane

5); and PVX-viral-particles-incubated PAMV MP1 (lane 6) or PVX MP1 (lane 7). The arrowheads indicate the position of the PAMV replicase (187K) and PVX replicase (165K). Electrophoretic analysis of ³⁵S-labeled translation products.

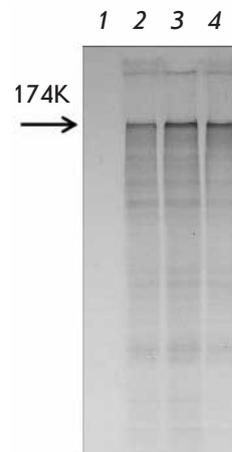


Fig. 4. Translational activation of encapsidated AltMV RNA *in vitro*. AltMV virion-packed RNA (lane 1); AltMV genomic RNA as a positive control (lane 2); encapsidated AltMV RNA with PVX MP1 (lane 3) or NMV MP1 (lane 4). The arrowhead indicates the position of the AltMV replicase (174K). Electrophoretic analysis of ³⁵S-labeled translation products.

Table 1. Potexvirus-encapsidated RNA translationally activated by MP1.

Virus	Protein			
	PVX MP1	NMV MP1	PAMV MP1	AltMV MP1
PVX	+	-	-	+
NMV	+	+	+	-
PAMV	-	+	+	-
AltMV	+	+	-	+

Note. The color indicates the phylogenetic subgroups of MP1 according to Wong et al. [16]: Ia (pink); Ib (green); Ic (yellow). "+" – translational activation, "-" – no translational activation.

		motif I	motif IA			
AltMV MP1	--MNHFINLLIEEGYVRTNEIL--SDT	LVVHAVAGAGKSTLIRKF	IHQLPQARAYTHGVP	56		
PVX MP1	--MDILISSLKSLGYSRTSKSLD--SGPL	LVVHAVAGAGKSTALRKL	ILRHPTFTVHTLGVP	57		
PAMV MP1	MEYSFLVRLLDHYGFERTTEKIVPGQPI	IVVQGIAGCGKTTLLRN	FHQEYPSIPIYSC-FP	59		
NMV MP1	MDCKYLLELLDSYSFIRSSRSFS--SPI	IIHGVAGCGKSTIIQK	IALAFPELLIGSFTPA	58		
	motif IA	motif II	motif III	motif IV		
AltMV MP1	DPPNLEGRFIQAFKSPDPNH	FNILDEYCAEPLSG-GWN	VLIADPLQHR	SQ-ALRPHYIKR	114	
PVX MP1	DKVSIRTRGIQKPGPIPEGNF	AILDEYTLDN	TTRNSYQALFAD	PYQAPF-SLEPHFYLE	116	
PAMV MP1	QKISENSEELQLLAKARFTAS	AILDEYLAHKFDYQK	CLAVFADPLQY	SHLGALRPHYQTS	119	
NMV MP1	LLDSNSGRKQLAVTSDP---	LDILDEYLG-GPNP	VVRLAKFCDPLQY	SCEQPEVPHFTSL	114	
	motif IV					
AltMV MP1	ESHRLGVATCELLTRVGLPVLSHK	TED--QVDYQGI	FEGPLFGTVIALD	STVRALLLKHG	172	
PVX MP1	TSFRVPRKVADLIAGCGDFETNSQ	EEG-HLEITGIF	KGPLLGKVI	AIDEESETLSRHG	175	
PAMV MP1	KHIGLVLVLLILSLRNWIPIESLL	SEEKTILKECD	PYATDPIGQII	ASNHEVLNYIKPQA	179	
NMV MP1	LTWRFCVTRTALLNGIFGCQIKSR	REDLCHL	THENPYTTDP	KGVVVAHEQEVINLLLQHG	174	
	motif V	motif VI				
AltMV MP1	ISPLCPAEVLGSEFEETT	VVSEVS--LSQVKFK	HALYIALTRHKK	SLHVRAPPLPDTPSR	230	
PVX MP1	VEFVKPCQVTGLELKVVTIV	SAAP--IEEIGQ	STAFYNAITRSK	GLTYVRAGT-----	226	
PAMV MP1	VEAICSEVLGKEFQTVSCYYQ	SHKLEDSAEERR	GLYIAISRAK	SAVLLFDLD-----	232	
NMV MP1	CPVTP	QHLWGLTIPVVS	VYITSIASLSTVDRAN-	LFLSLTRDSKALHIFE	DAWSHATC	233
AltMV MP1	SL	232				
PVX MP1	--					
PAMV MP1	--					
NMV MP1	--					

Fig. 5. Alignment of the amino acid sequences of AltMV MP1, PVX MP1, PAMV MP1, and NMV MP1. The grey shading denote the MP1 conservative motifs (motifs I, IA, II, III, IV, V, VI) [15].

the PAMV virion-packed RNA does not translationally become activated upon treatment with AltMV MP1 (data not shown).

The results described above are summarized in Table.

The findings of cross-activation studies point out that NMV MP1 is capable of translationally activating encapsidated PAMV and AltMV RNA but not encapsidated PVX RNA. Furthermore, PAMV MP1 can translationally activate NMV virion-packed RNA but not AltMV and PVX RNA. We have already observed cross-activation of encapsidated AltMV and PVX RNA. Now, we have found evidence of PVX MP1 activating NMV virion-packed RNA, as well as encapsidated AltMV RNA, but not PAMV RNA. Moreover, AltMV MP1 was found to lack the capability of activating PAMV and NMV virion-packed RNA (Table 1). Reciprocal cross-activation was discovered in pairs: PVX – AltMV and NMV – PAMV.

Broadly, potexviral movement proteins were shown to be able to translationally activate the encapsidated

RNA of kin species showing, however, selective specificity.

The amino acid sequence of the corresponding MP1 was compared to shed some light on the data obtained (Fig. 5). *Potexvirus* movement proteins are known to belong to superfamily I helicases, which contain seven highly conservative NTPase/RNA helicase motifs constituting a NTPase/RNA helicase domain [14, 15]. We have previously produced a set of PVX MP1 variants that carry deletions in different functional regions. The deletions proved to be negligible in regard to the MP1-CP interaction unless those involve the motif IV of the NTPase/RNA helicase domain (amino acid residues 112–122). Hence, our sequence analysis considered mainly motif IV [5].

The MP1 motif IV sequences of PVX, AltMV, PAMV, and NMV show a high degree of variability: hence, the prediction that these MP1s bind to the end of heterologous virions in the absence of cross-activation is questionable (Fig. 5). Additional experiments were performed to clarify this question.

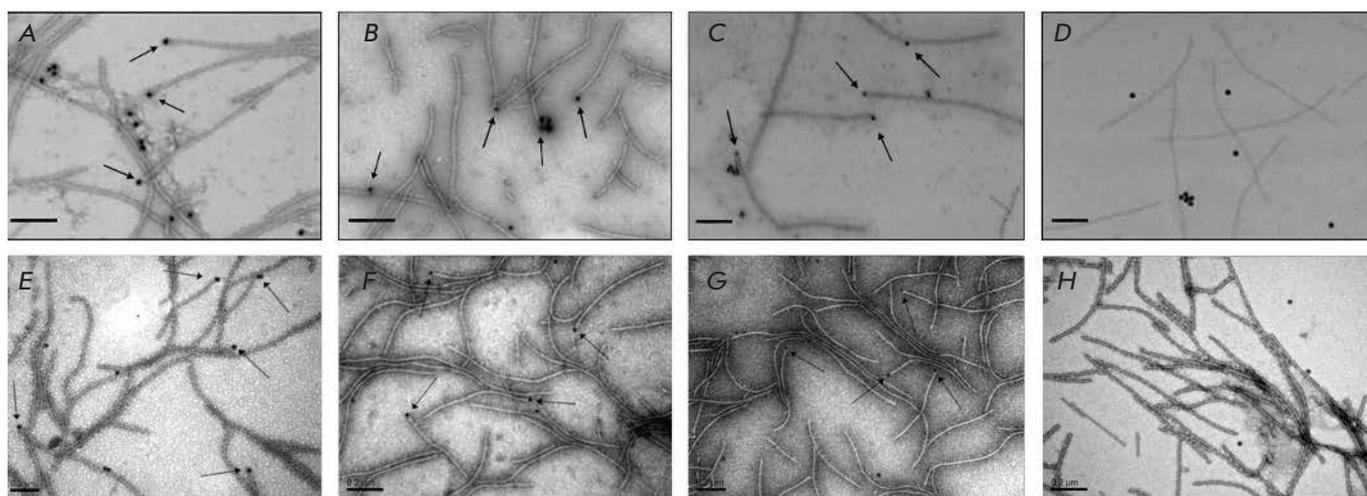


Fig. 6. Immunoelectron microscopy of PVX virions interaction with PVX MP1 (A), AltMV MP1 (B), or PAMV MP1 (C) and AltMV virions interactions with AltMV MP1 (E), NMV MP1 (F), or PAMV MP1 (G). Primary antibodies against PVX MP1 (A – positive control), AltMV MP1 (B,E), PAMV MP1 (C,G), or NMV MP1 (F) and secondary gold-conjugated (12 nm) antibodies. Negative controls: PVX (D) or AltMV (H) virions treated with primary and secondary gold-labeled antibodies in the absence of MP1. The arrowheads indicate the position of MP1 bound to the end of virions. The scale bars represent 200 nm.

We used immune electron microscopy employing primary antibodies against the MP1 of PVX, AltMV, and PAMV and gold-conjugated secondary antibodies. This approach helped us clearly visualize the binding to the PVX virion end of PVX MP1 (Fig. 6A, positive control) and AltMV MP1 (Fig. 6B) that render encapsidated PVX RNA translatable, as well as the PAMV MP1 that lacks such capability (Fig. 6C). In a control experiment (Fig. 6D, negative control), no binding of gold beads was observed in the absence of MP1. Similar results were obtained when studying AltMV viral particles (Fig. 6E–M): virion end binding was observed for PAMV MP1 (Fig. 6G), AltMV MP1 (Fig. 6E), and NMV MP1 (Fig. 6F), while only AltMV MP1 was able to activate AltMV virion-packed RNA.

Our results provide the first piece of direct evidence of MP1 physical binding to a heterologous virion end. Such binding is necessary, but not sufficient, to the translational activation of encapsidated RNA.

A phylogenetic analysis of potexviral MP1 [16] split the *Potexvirus* genus members into three subclusters, Ia, Ib, and Ic. AltMV MP1 was assigned to Ia; NMV and PAMV MP1, to Ib; and PVX MP1, to Ic.

We have demonstrated in this study a reciprocal cross-activation of translation in the pairs PAMV–NMV and AltMV–PVX, while other pairs show either non-reciprocal cross-activation (NMV–PVX and NMV–AltMV) or a total absence of cross-activation (PAMV–PVX) (Table). The PAMV and NMV cluster-

ing in the same subgroup, Ib, tracks well with the fully reciprocal cross-activation observed for this pair. The observed cross-activation of PVX and AltMV virion-packed RNA corresponds probably to a lesser branching between the subgroups Ia and Ic compared to the one between Ia and Ib [16]. We suggest that MP1 conformation and interaction with CP C-terminal region are key features that determine the specificity of translational activation. We suggest that the MP1–CP interaction destabilizes the proteinaceous helical virion. The activation that is abolished by MP1 phosphorylation further supports this assumption [5]. Our previous data on MP1 deletion mutants successfully binding to a virion does not support the idea of the MP1 conformation being crucial to virion binding [5]. In addition, Rodionova *et al.* showed that the full-length movement proteins of potexviruses can bind to heterologous virions but fail to activate the translation of its genomic RNA. We consider phylogenetically related movement proteins to share conformational features and, thus, able to loosen the helical proteinaceous virion of either virus, hence exposing the 5'-end of genomic RNA to ribosomes [7].

CONCLUSIONS

Our results clearly indicate that encapsidated potexviral RNAs share translational features *in vitro*. Direct evidence of MP1 binding to the virion end being essential but not sufficient to induce translation of genomic

RNA was obtained. Potexviral movement proteins 1 are capable of translationally activating heterologous potexviruses virion-packed RNA with unequal specificity. Reciprocal cross-activation is observed for potexviruses of the same subgroup (NMV–PAMV, subgroup *Ib*) or closely related subgroups (PVX – AltMV, subgroups *Ic* and *Ia*). The movement proteins 1 capable

of translationally activating heterologous potexviruses encapsidated RNA are likely to share protein molecule conformation; this suggestion, however, is a matter for further research. ●

This work was supported by the Russian Science Foundation (grant no. 14-24-00007).

REFERENCES

- Atabekov J.G., Rodionova N.P., Karpova O.V., Kozlovsky S.V., Poljakov V.Yu. // *Virology*. 2000. V. 271. № 2. P. 259–263.
- Atabekov J.G., Rodionova N.P., Karpova O.V., Kozlovsky S.V., Novikov V.K., Arkhipenko M.V. // *Virology*. 2001. V. 286. P. 466–474.
- Karpova O.V., Zayakina O.V., Arkhipenko M.V., Sheval E.V., Kiselyova O.I., Poljakov V.Yu., Yaminsky I.V., Rodionova N.P., Atabekov J.G. // *J. Gen. Virol.* 2006. V. 87. № 9. P. 2731–2740.
- Karpova O.V., Arkhipenko M.V., Zayakina O.V., Nikitin N.A., Kiselyova O.I., Kozlovsky S.V., Rodionova N.P., Atabekov J.G. // *Molecular biology*. 2006. V. 40. № 4. P. 628–634.
- Zayakina O., Arkhipenko M., Kozlovsky S., Nikitin N., Smirnov A., Susi P., Rodionova N., Karpova O., Atabekov J. // *Mol. Plant. Pathol.* 2008. V. 9. № 1. P. 37–44.
- Kiseleva O.I., Yaminsky I.V., Karpova O.V., Rodionova N.P., Kozlovsky S.V., Arkhipenko M.V., Atabekov J.G. // *J. Mol. Biol.* 2003. V. 332. № 2. P. 321–325.
- Rodionova N.P., Karpova O.V., Kozlovsky S.V., Zayakina O.V., Arkhipenko M.V., Atabekov J.G. // *J. Mol. Biol.* 2003. V. 333. № 3. P. 565–572.
- Mukhamedzhanova A.A., Karpova O.V., Rodionova N.P., Atabekov J.G. // *Doklady Biochemistry and Biophysics*. 2009. V. 428. № 2. P. 239–241.
- Ivanov P.A., Mukhamedzhanova A.A., Smirnov A.A., Rodionova N.P., Karpova O.V., Atabekov J.G. // *Virus Genes*. 2011. V. 42. № 2. P. 239–241.
- Mukhamedzhanova A.A., Smirnov A.A., Arkhipenko M.V., Ivanov P.A., Chirkov S.N., Rodionova N.P., Karpova O.V., Atabekov J.G. // *Open Virol. J.* 2011. V. 5. P. 136–140.
- Karpova O.V., Tyulkina L.G., Atabekov K.J., Rodionova N.P. // *J. Gen. Virol.* 1989. V. 70. P. 2287–2297.
- Arkhipenko M.V., Petrova E.K., Nikitin N.A., Protopopova A.D., Dubrovin A.V., Yaminskii I.V., Rodionova N.P., Karpova O.V., Atabekov J.G. // *Acta Naturae*. 2011. V. 3. № 3. P. 40–46.
- Nikitin N., Trifonova E., Karpova O., Atabekov J. // *Microsc. Microanal.* 2013. V. 19. № 4. P. 808–813.
- Koonin E.V., Dolja V.V. // *Crit. Rev. Biochem. Mol. Biol.* 1993. V. 28. № 5. P. 375–430.
- Morozov S.Y., Solovyev A.G. // *J. Gen. Virol.* 2003. V. 84. P. 1351–1366.
- Wong S.M., Lee K.C., Yu H.H., Leong W.F. // *Virus Genes*. 1998. V. 16. № 3. P. 295–302.

Synthesis and Characterization of Hybrid Core-Shell Fe₃O₄/SiO₂ Nanoparticles for Biomedical Applications

I. V. Zelepukin^{1,2,3*}, V. O. Shipunova^{1,2,3}, A. B. Mirkasymov^{1,2}, P. I. Nikitin^{3,4}, M. P. Nikitin^{1,2}, S. M. Deyev^{1,3,5}

¹Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya Str. 16/10, Moscow, 117997, Russia

²Moscow Institute of Physics and Technology, Institutskiy Per. 9, Moscow Region, Dolgoprudny, 141700, Russia

³National Research Nuclear University MEPhI, Kashirskoe shosse 31, Moscow, 115409, Russia

⁴Prokhorov General Physics Institute, Russian Academy of Sciences, Vavilov Str. 38, Moscow, 119991, Russia

⁵Lobachevsky State University of Nizhny Novgorod, Gagarina prospekt 23, Nizhny Novgorod, 603950, Russia

*E-mail: zelepukin@phystech.edu

Received: July 15, 2017; in final form October 30, 2017

Copyright © 2017 Park-media, Ltd. 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 creation of markers that provide both visual and quantitative information is of considerable importance for the mapping of tissue macrophages and other cells. We synthesized magnetic and magneto-fluorescent nanomarkers for the labeling of cells which can be detected with high sensitivity by the magnetic particle quantification (MPQ) technique. For stabilization under physiological conditions, the markers were coated with a dense silica shell. In this case, the size and zeta-potential of nanoparticles were controlled by a modified Stober reaction. Also, we developed a novel facile two-step synthesis of carboxylic acid-functionalized magnetic SiO₂ nanoparticles, with a carboxyl polymer shell forming on the nanoparticles before the initiation of the Stober reaction. We extensively characterized the nanomarkers by transmission electron microscopy, electron microdiffraction, and dynamic and electrophoretic light scattering. We also studied the nanoparticle cellular uptake by various eukaryotic cell lines.

KEYWORDS magnetic nanoparticles, surface modification, magnetic detection, silicium dioxide, cell labeling.

ABBREVIATIONS MPI – magnetic particle imaging; MPQ – magnetic particle quantification; TEOS – tetraethyl orthosilicate; m-cit – citrate-coated magnetic particles; m-CMD – carboxymethyl dextran-coated magnetic particles; m-cit-SiO₂ – magnetic particles coated by SiO₂ via an intermediate citrate coating; m-CMD-SiO₂ – magnetic particles coated by SiO₂ via an intermediate carboxymethyl dextran coating.

INTRODUCTION

There is growing interest at the moment in the use of nanoparticles as theranostic objects (agents that combine diagnostic and therapeutic functions on one single platform) [1–3] and in the development of nanocomplexes capable of performing a therapeutic function or binding to cells only in response to certain signals from the body or to the absence of such signals [4], or external stimuli [5]. For an early diagnosis of diseases and the monitoring of ongoing therapy, it is important to be able to visualize the distribution of nanoagents in the body by means of various markers.

Many magnetic nanoparticles are superparamagnetic, which makes them detectable by magnetic resonance imaging [6], MPI visualization [7], ferromag-

netic resonance [8], giant magnetic resistance [9], etc. [10–12]. Of particular interest is the detection of non-linear magnetic materials, which is based on the exposure of a sample to a magnetic field at two frequencies and the monitoring of the response at combinatorial frequencies of the applied field (MPQ detection) [13]. This method enables highly sensitive and quantitative detection of superparamagnetic nanoparticles in a wide range of concentrations, in particular non-invasively in a living organism, which opens up broad prospects for their use in biomedicine.

Nonstabilized magnetic nanoparticles do not have colloidal stability under physiological conditions, and they are susceptible to oxidation, which can decrease their detection limit [14]. An effective dense coating

that can protect magnetic particles from oxidation and aggregation is a silica shell. Such a coating is highly stable and inert, and its surface can be modified by the desired functional groups. In addition, the mesoporous silica structure is used to deliver therapeutic agents and genetic vectors [15].

Silica-containing nanoparticles are often synthesized using the Stober method. This is a simple and convenient one-step method that avoids surfactants or toxic organic solvents, and a relatively low rate of inorganic layer formation enables one to control the resulting nanoparticle's size [16].

We synthesized magnetic and magneto-fluorescent markers coated with a silica shell. The particle surface was functionalized with amino and carboxyl groups to ensure use of these markers for conjugation with other nanoagents, proteins, and targeting moieties. We also proposed a method for coating magnetic particles with silica without the need for further modification of the surface by functional groups. We studied the synthesized nanoparticles by transmission electron microscopy, electron microdiffraction, and dynamic and electrophoretic light scattering, and we measured the detection limit of the nanoparticles as magnetic markers for biomedical research. We demonstrated effective quantitative and optical labeling of various eukaryotic cells by the nanoparticles and found a relatively low cytotoxicity of the markers at the tested concentrations.

The produced markers are promising for use *in vivo*: e.g., to identify tissue macrophages and determine their activity for the diagnosis of atherosclerosis, cancers, myocardial infarction, and other human diseases [17, 18].

EXPERIMENTAL

In the study, we used iron (II) chloride tetrahydrate, iron (III) chloride hexahydrate, tetraethyl orthosilicate (TEOS), (3-aminopropyl)triethoxysilane, tris(2,2'-bipyridyl)ruthenium (II) chloride hexahydrate, succinic anhydride, carboxymethyl dextran sodium salt, *L*-glutamine, dye Hoechst 33342 (Sigma-Aldrich), aqueous ammonia, nitric acid, trisodium citrate dihydrate, isopropyl alcohol, ethyl alcohol, dimethyl sulfoxide (Chimmed), ninhydrin, MTT solution (Dia-m), dry methyl alcohol (Merck), Concanavalin A (lectin from *Canavalia ensiformis*) (Vector Laboratories), phosphate buffered saline (PBS) pH 7.4, carbonate buffer pH 9, McCoy's 5A medium (Life Technologies), fetal bovine serum (FBS) (HyClone), and BT-474, SK-BR-3 (human mammary gland), HEK 293T (human embryonic kidney), and CHO (Chinese hamster ovary) cell lines. For magnetic separation, a permanent cylindrical Neodymium Iron Boron magnet D 25 × 10 mm (Ningbo Ketian Magnet Co.) was used.

Synthesis of magnetite nanoparticles

A mixture containing 2.9 mmol of $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$, 10.1 mmol of $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$, and 40 mL of distilled water was added with 5 mL of 30% NH_4OH under constant stirring. The solution was kept at 80 °C for 2 h. The resulting particles were treated with a 2M HNO_3 solution and then repeatedly washed with distilled water by magnetic separation on a 25 mm Neodymium Iron Boron magnet. The particles not attracted to the magnet were sequentially collected for 15 min, thereby forming different fractions of magnetic particles. The first two fractions had low pH values, which led to a rapid degradation of particles. Magnetic nanoparticles of the third fraction were used in the experiments.

Coating of nanoparticles with a silica shell

To stabilize magnetic nanoparticles under the reaction conditions, the particles were pre-coated with a citrate anion by adding trisodium citrate (a concentration of 25 g/L) to a colloidal solution of magnetic particles. Alternatively, the particles were coated with a polymeric carboxymethyl dextran layer. For this, carboxymethyl dextran was added (to get a final concentration of 50 g/L) to a colloidal nanoparticle solution under heating to 80 °C. After preliminary stabilization, the magnetic particles were washed three times with distilled water.

Fifty microliters of magnetic particles was added to 1 mL of alcohol. The reaction mixture pH was adjusted to 9, and then 10 – 200 μL of TEOS was added. After 1 day, the nanoparticles were washed by centrifugation with distilled water.

Functionalization of the silica nanoparticle surface

A 1% solution of (3-aminopropyl)triethoxysilane in ethanol was added to the synthesized particles, which led to exposure of the primary amino groups on their surface. The particles were then washed twice with ethanol. Further, the amino groups were modified into carboxyl groups by the addition of succinic anhydride in a carbonate buffer (pH 9) to a concentration of 4 g/L. After 3 h, the particles were washed from the reaction products with distilled water.

Preparation of magneto-luminescent silica nanoparticles

Magneto-luminescent nanoparticles were prepared analogously to the magnetic particles coated with silica in ethanol by adding 0.03 mg of tris(2,2'-bipyridyl)ruthenium (II) chloride hexahydrate to 1 mL of the reaction mixture 5 min after the start of TEOS hydrolysis. After the synthesis, the nanoparticles were stored in the dark at +4 °C.

Characterization of particles

The hydrodynamic nanoparticle size and zeta potential were determined by dynamic light scattering and electrophoretic light scattering using a Zetasizer Nano ZS analyzer (Malvern Instruments Ltd.). We used the mean particle size and the mean zeta potential value. To measure the zeta potential, the particles were transferred to PBS, pH 7.4, before measurement.

The morphology of the nanoparticles was examined using a JEM-2100 transmission electron microscope (JEOL Ltd.) with an accelerating voltage of 200 kV. The nanoparticle samples were applied to a carbon-coated copper grid and then dried in air.

The phase composition of the particles was determined by the electron microdiffraction method.

The magnetic signal of iron oxide markers was determined by MPQ detection of nonlinear magnetic materials [13]. For measurement, 20 μL of the nanoparticle sample in the cylindrical tube was placed into the coil of the MPQ reader.

Fluorescence and absorption spectra were acquired using an Infinite M1000PRO Microplate reader (Tecan Group Ltd.).

Cell labeling with nanoparticles

Cells of the BT-474, SK-BR-3, HEK 293T, and CHO lines were cultured in a McCoy's 5A medium supplemented with 10% heat-inactivated fetal bovine serum and 2 mM *L*-glutamine at +37°C in a humidified atmosphere with 5% CO₂. The cells were passaged 2 to 3 times a week at 80–90% confluence. The cells removed from the culture plastic surface (0.7×10^6) were washed twice with PBS, incubated with nanoparticles at a concentration of 0.01 g/L at room temperature for 2 h, and washed from unbound particles under constant stirring. The number of cell-bound particles was determined by MPQ-cytometry [18].

Cell viability assay

Nanoparticle cytotoxicity was assessed using an MTT test. Cells were seeded on a 96-well plate, 10⁴ cells/well into 200 μL of McCoy's 5A medium with 10% FBS. The cells were cultured at 37°C in a CO₂ incubator overnight. Then, the medium was removed and the cells were sterilely added with a serum-free medium (negative control) and a serum-free medium containing the tested particles at various concentrations at a volume of 100 μL per well. The cells were incubated at room temperature for 2 h, then washed with the serum-free medium, added with McCoy's 5A medium containing 10% FBS, and incubated in a CO₂ incubator (24 h, 37°C). The medium was then shaken off, and the cells were washed once with the medium. After this, 100 μL of a MTT solution (0.5 g/L in McCoy's 5A) was added per

well and incubated at 37 °C in a 5% CO₂ atmosphere for 1 h. Then, the MTT solution was removed, 100 μL of dimethyl sulfoxide was added per well, and the plate was shaken until complete dissolution of formazan crystals. The solution absorbance in each well was measured using an Infinite M1000PRO Microplate reader - (Tecan Group Ltd.) at a wavelength of $\lambda=540$ nm.

Fluorescence microscopy

Cells were plated into a 96-well plate, 10⁴ cells/well in 200 μL of McCoy's 5A medium with 10% FBS. After culturing at 37°C in a CO₂ incubator overnight, the tested particles were sterilely added to the cells and the cells were incubated at room temperature for 2 h, washed with serum-free medium, added with McCoy's 5A medium with 10% FBS, and incubated in a CO₂ incubator at 37°C for 24 h. Cell nuclei were stained with the Hoechst 33342 dye at room temperature for 10 min and then washed three times with PBS. Cell samples were analyzed with a Leica DMI 6000B inverted fluorescent microscope (Leica Microsystems) in transmitted light and fluorescence channels corresponding to nanoparticle fluorescence (excitation at 545/30; emission at 610/75) and Hoechst 33342 dye fluorescence (excitation at 360/40; emission at 470/40).

RESULTS AND DISCUSSION

Magnetite nanoparticles were synthesized by co-precipitation of iron (II) and (III) chlorides under alkaline conditions. The synthesis was optimized to produce magnetic markers with a minimum detection limit. Because many iron oxyhydroxides produced in the reaction were not superparamagnetic and reduced the detectable magnetic signal of the entire nanoparticle sample [19], it was very important to determine the optimum ratio of iron salts in the reaction mixture. The maximum, normalized signal of particles was found to occur at a salt ratio of $[\text{FeCl}_2]/[\text{FeCl}_3]=1/3.5$. In this case, the maximum magnetic signal was observed in the third and fourth fractions of nanoparticles (*Fig. 1A*). The detection limit of these nanoparticles determined with MPQ was found to be 2.7 ng in 20 μL of solution.

Then, the nanoparticles were coated with a silica shell. The zeta potential of magnetic nanoparticles at pH 9 was near zero, which led to their aggregation under the reaction conditions. The agglomerates that formed at high pH lost colloidal stability. Therefore, it was necessary to modify the particles before the synthesis of the silica coating. For this purpose, as in [20], we used a citrate coating (hereinafter, these particles are designated as m-cit). In this case, the zeta potential of the nanoparticles became strongly negative, and the particles remained stable over a wide range of pH values. For the first time, a polymeric carboxymethyl

dextran was used as an alternative intermediate coating. The magnetic nanoparticles coated with carboxymethyl dextran (hereinafter m-CMD) were stable under the reaction conditions. In addition, the polymer bounded several magnetite particles together, which resulted in polymer-coated particles with a high content of magnetic nuclei, and, hence, a lower detection limit. Then, hydrolysis of tetraethyl orthosilicate with polycondensation of the reaction products on the magnetite surface was performed.

Synthesis of silica nanoparticles lacking a magnetic core was used as a model system for exploring the main dependencies of the synthesis process. We studied the effect of parameters such as the solvent type, $[\text{H}_2\text{O}]/[\text{TEOS}]$ ratio, and the reaction pH on the size of the resulting SiO_2 nanoparticles.

Increase in the carbon chain length in the used alcohol was found to result in a substantial increase in the size of the synthesized particles. The mean silica particle size was ~ 10 nm in methanol, 100 nm in ethanol, and 500 nm in isopropanol. Solvents with a longer carbon chain are hydrophobic, which is not compatible with the standard Stober reaction. A change in the $[\text{H}_2\text{O}]/[\text{TEOS}]$ ratio in the reaction enabled a more accurate control of the silica particle size (*Fig. 1B*). The dependence of the hydrodynamic particle size on the reagent ratio had a characteristic profile with a pronounced maximum, which was probably related to a decrease of tetraethyl orthosilicate or water quantities due to the reaction within the tested concentration range. The Stober process proceeded at alkaline pH, and increase in the pH significantly accelerated the reaction, which adversely affected the particle size dispersion. Most of the experiments were carried out at pH 9, with the particle synthesis time being approximately 1 h (*Fig. 1C*).

We synthesized both magnetic and magnetic-luminescent silica-coated nanoparticles. The growth of magnetic silica particles depended on the reaction conditions in the same pattern as that of SiO_2 particles. In particular, the $[\text{TEOS}]/[\text{H}_2\text{O}]$ ratio in the synthesis of m-cit- SiO_2 influenced the nanoparticle size in the same pattern as was previously determined for SiO_2 nanoparticles.

The use of methanol or ethanol as a solvent resulted in magnetic particles with mean sizes ranging from 50 to 80 nm and 100 to 200 nm, respectively. In isopropanol, citrate-coated magnetic particles aggregated: therefore, before starting the reaction, the particles were first coated with a thin SiO_2 layer in methanol and then used as nucleation centers in the next step of the Stober process in isopropanol. This procedure resulted in particles of 300–500 nm in size.

Using carboxymethyl dextran as an intermediate coating, we obtained nanoparticles with mean sizes of

200 ± 60 nm, with the mean size of initial m-CMD particles being 44 ± 12 nm. Nanoparticles of 200 ± 60 nm in size were used twice as “seeds” in the Stober reaction in isopropanol to produce particles of 764 ± 187 nm in size. The use of the described particles as nucleation centers in a multistage variant of the Stober reaction resulted in large-size particles that were not appropriate for *in vivo* experiments [21] but interesting for *ex vivo* and *in vitro* diagnostics.

To synthesize fluorescent nanoparticles, we, as in [22], added $[\text{Ru}(\text{bipy})_3]\text{Cl}_2$ to the reaction mixture 5 min after starting the reaction to avoid the aggregation of magnetic nuclei through an increase in the ionic strength of the solution. Tris(2,2'-bipyridyl)ruthenium (II) was incorporated into the forming amorphous silica lattice, which induced fluorescent properties in the particles. The excitation and emission spectra of the particles are shown in *Fig. 2*. The particles retained their colloidal stability and ability to fluoresce for at least 1 year.

For various biological applications, conjugation of nanoparticles with proteins or other objects is often necessary. Bioconjugation chemistry allows one to couple objects of different nature via certain functional groups. One of the most convenient techniques is carbodiimide conjugation of a carboxyl group to a primary amino group, resulting in the formation of a stable peptide bond [23]. The synthesized m-cit- SiO_2 nanoparticles initially exposed surface hydroxyl groups, so we performed a two-step modification of their surface to obtain carboxyl groups. First, the surface of SiO_2 particles was treated with (3-aminopropyl)triethoxysilane. The presence of surface amino groups was proved by a change in the color of a nanoparticle solution upon its interaction with a 5% ninhydrin solution and also by a change in the mean particle zeta potential from -36 to $+12$ mV. After that, the particles were treated with succinic anhydride and surface amino groups were converted to carboxyl groups by the ring opening reaction. The logarithmic dependence of the resulting zeta potential of particles on the succinic anhydride concentration in the mixture enabled the production of particles with different zeta potentials ranging from $+12$ to -58 mV (*Fig. 1D*).

Surface-exposed amino groups cause particle aggregation, but after the second stage of modification, the hydrodynamic particle size becomes equal to that of the original particles without aggregate formation, which made it possible to use the method in producing colloidally stable solutions of differently charged magnetic silica particles.

It should be noted that the use of carboxymethyl dextran for intermediate stabilization of magnetite eliminates the need for an additional modification of

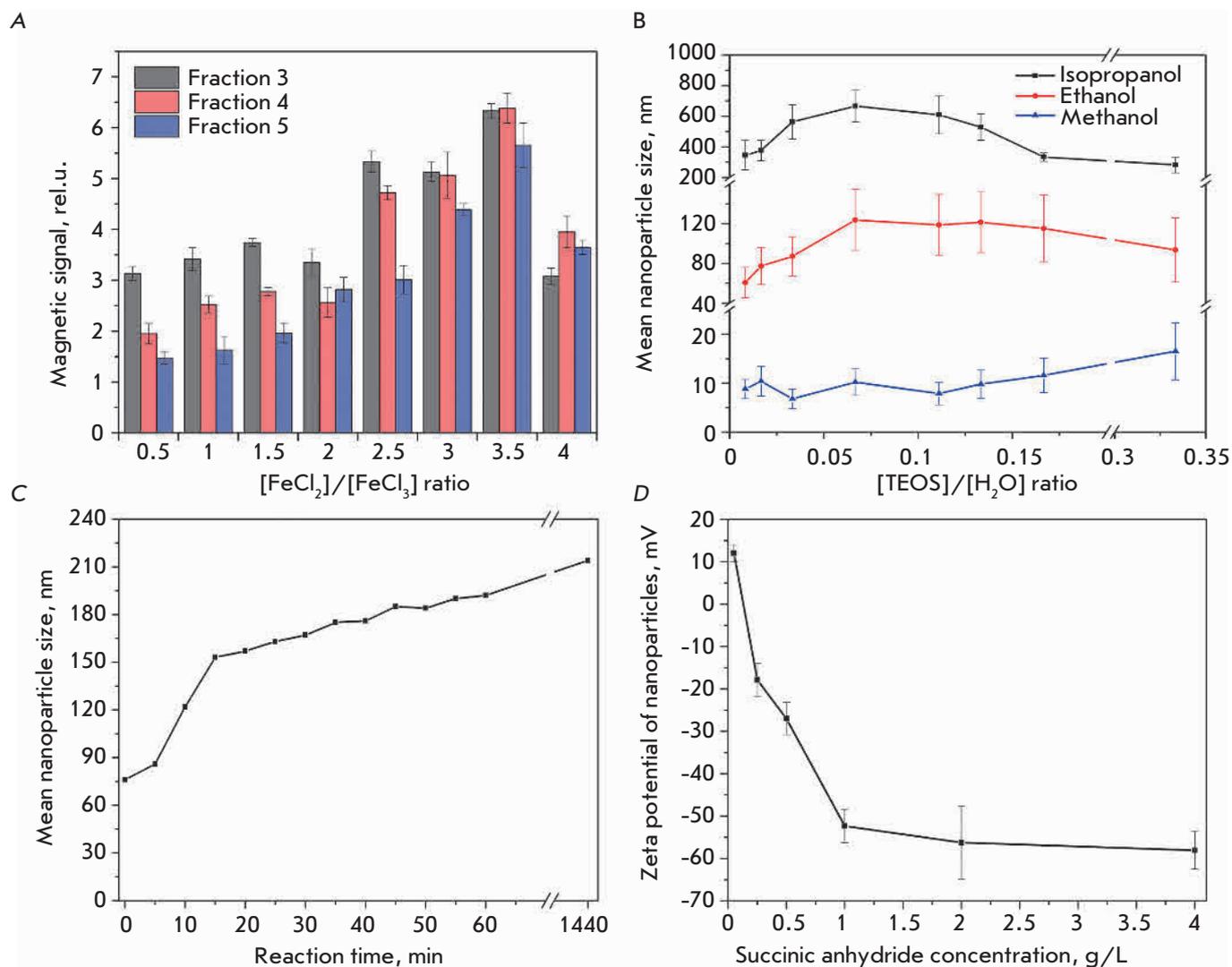


Fig. 1. Control of the physico-chemical properties of magnetic silica nanoparticles: (A) the dependence of the normalized nanoparticle magnetic signal on the ratio of iron salts; (B) the dependence of the hydrodynamic particle size on the [TEOS]/[H₂O] ratio in the Stober process for various solvents; (C) the dependence of the hydrodynamic particle size on time in the Stober reaction; (D) the effect of the succinic anhydride concentration in the reaction mixture on the particle zeta potential. Error bars indicate the standard deviation from the mean particle size.

the surface of magnetic silica particles, because the carbohydrate polymer with carboxyl groups occurs on the surface immediately after the synthesis. The presence of dextran on the surface was confirmed by sedimentation of particles in the presence of Concanavalin A that bound carbohydrates and, consequently, the polysaccharide on the particles surface [24]. Thus, the use of carboxymethyl dextran accelerates the synthesis and provides, immediately after the Stober reaction, markers ready for conjugation with proteins.

The morphology of m-cit-SiO₂ magnetic silica nanoparticles with surface carboxyl groups, synthe-

sized in ethanol and methanol, and m-CMD-SiO₂ particles synthesized in methanol was studied by transmission electron microscopy and electron microdiffraction (Fig. 3).

The obtained electron micrographs revealed that all synthesized nanoparticles were multinuclear structures containing 2 to 30 magnetite nuclei and having a solid silica shell with a thickness of 2 (Fig. 3A) to 30 nm (Fig. 3C). The m-CMD-SiO₂ particles had more iron oxide nuclei than m-cit-SiO₂ ones, on average, and had a detection limit determined with MPQ of 2.7 ng in 20 μL of solution, which is comparable or superior to many

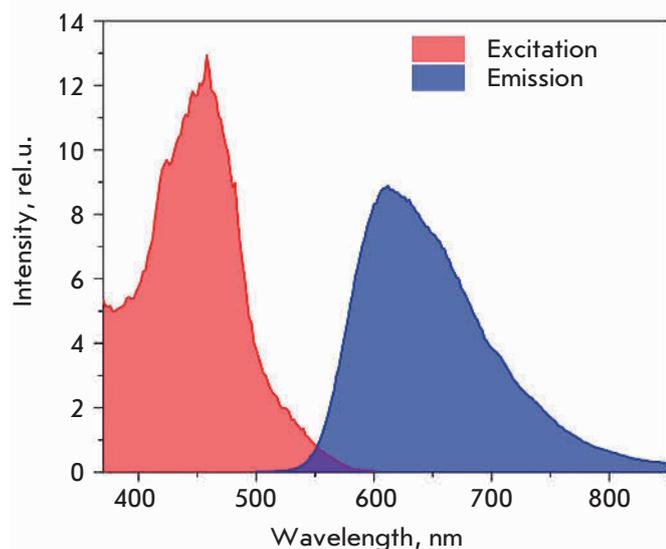


Fig. 2. Excitation and emission spectra of luminescent magnetic silica nanoparticles.

methods widely used in magnetometry. It should be noted that the nanoparticle's size determined from image analysis correlates with the dynamic light scattering data. An analysis of the diffraction spectrum of the magnetic nuclei demonstrated that they consisted of magnetite Fe_3O_4 , the crystallographic Fd3m space group (cubic system), with the main interplanar distances in the crystal being 0.49, 0.29, 0.25, 0.21, 0.17, and 0.15 nm. The strongest line was 0.25 nm.

The synthesis of magnetic nanoparticles capable of effectively interacting with the surface of living cells is important in such areas as MRI monitoring of stem cells, magnetic tissue engineering, magnetofection of eukaryotic cells, and some others. Previous studies have demonstrated that one of the important characteristics that determine the interaction of particles with proteins and cells is the zeta potential [25]. While a positive charge on a particle's surface leads to a more active adsorption of proteins, a strong negative charge significantly increases the efficiency of cellular uptake by particles [25].

In this paper, we have demonstrated effective labeling of eukaryotic cells from different tissues and species by negatively charged m-CMD-SiO₂-particles. The cell lines BT-474, SK-BR-3, HEK 293T, and CHO were incubated with nanoparticles and washed from unbound particles for further analysis. Using fluorescence microscopy, we found that these particles were able to effectively visualize eukaryotic cells (Fig. 4A), with the cell membrane integrity being preserved. Upon interaction with the cell membrane, the particles, despite

their high colloidal stability, tended to form bright and visually detectable large conglomerates.

Nanoparticles used as markers of the cell surface should possess high biocompatibility. Therefore, we compared the cytotoxicity of the particles in the MTT test. At the nanoparticle concentration used for cell visualization, namely 0.01 g/L, more than 85% of the cells (except the HEK 293T line) retained their viability (Fig. 4C). In this case, IC₅₀ of m-CMD-SiO₂ particles for all four cell lines was in the range of 63–125 mg/L, which indicates their low cytotoxicity, comparable to that of other magnetic nanoparticles used *in vivo* [26].

It is interesting to note that ruthenium (II)-based fluorescent compounds can be considered for use as chemotherapeutic agents [27]. But in our case, the presence of ruthenium (II) did not significantly affect particle toxicity, probably due to the strong fixation of ruthenium in a silica shell.

The physico-chemical properties of these particles, such as fluorescence and magnetism, as well as the opportunity of their effective modification by biomolecules, make the particles very promising for diagnostic purposes. These nanoparticles can be simultaneously visualized and quantified in explored sites of their uptake. For example, we used MPQ-cytometry to quantify interactions between m-CMD-SiO₂ nanoparticles and the mentioned cell lines and revealed statistically different uptakes of the nanoparticles in different cells, expressed in the mass content of particles per cell

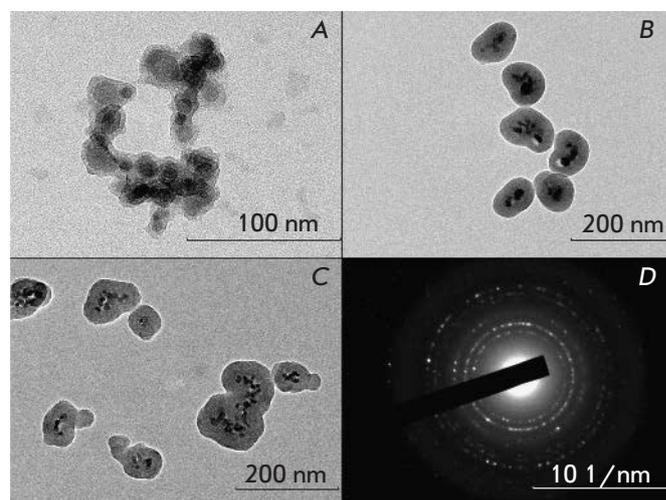


Fig. 3. Electron micrographs showing the typical architectures of magnetic silica nanoparticles. Nanoparticles were synthesized in methanol (A) or ethanol (B) through intermediate stabilization with citrate and in ethanol through intermediate coating with carboxymethyl dextran (C). (D) shows a microdiffraction pattern of magnetic nanomarkers.

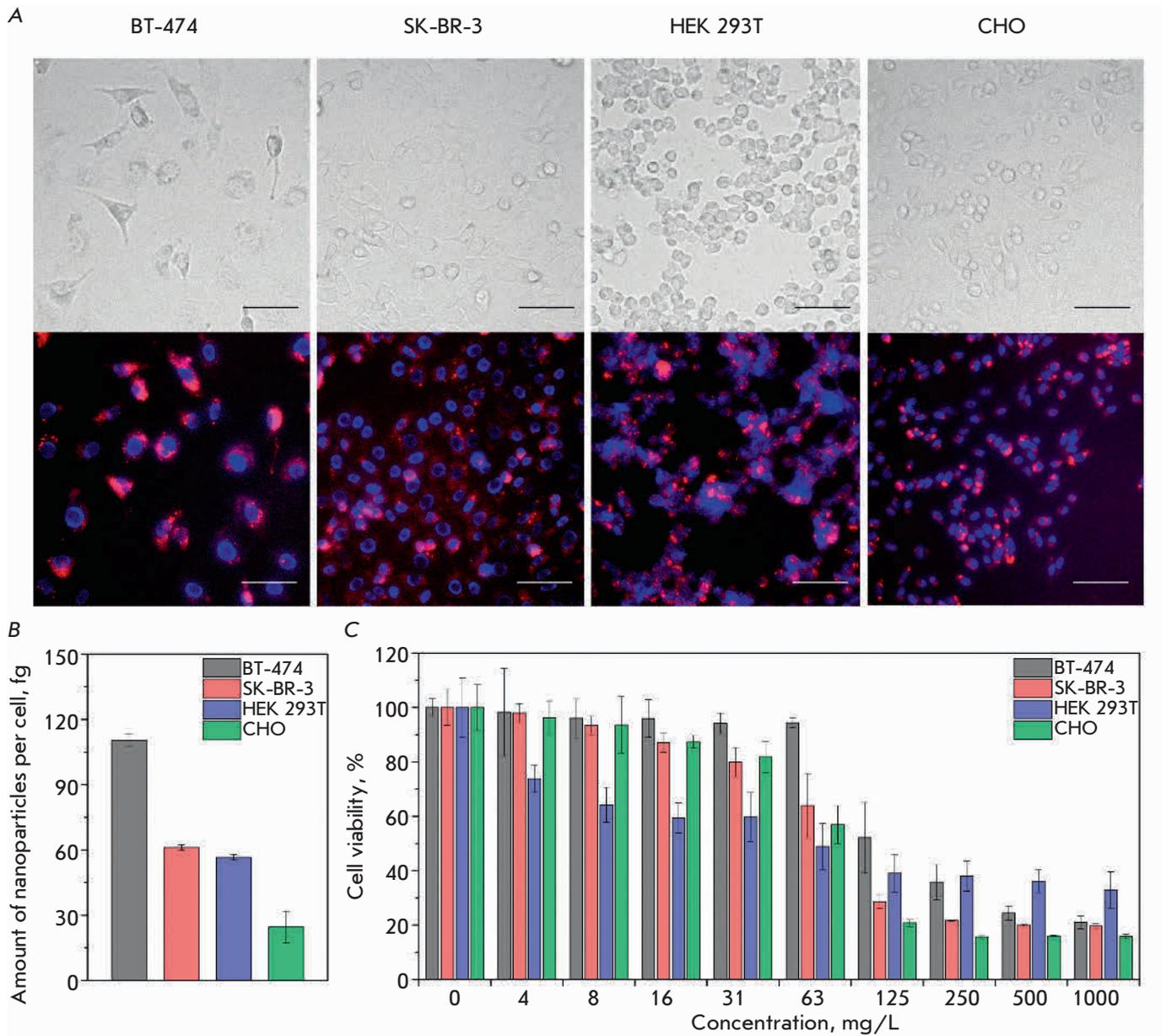


Fig. 4. Labeling of eukaryotic cells by magneto-fluorescent m-CMD-SiO₂ nanoparticles. (A) Fluorescence microscopy: visualization of different cells with m-CMD-SiO₂ nanoparticles. Top panel: transmitted light images; bottom panel: overlaying of fluorescence light images of nanoparticles (excitation at 545/30, emission at 610/75) and the nuclear dye Hoechst 33342 (excitation at 360/40, emission at 470/40). Scale bars – 75 μm. (B) Cellular uptake of m-CMD-SiO₂ nanoparticles measured by MPQ-cytometry. (C) Cytotoxicity of m-CMD-SiO₂ nanoparticles.

(BT-474: 110.4 ± 1.3; SK-BR-3: 61.1 ± 1.2; HEK 293T: 56.6 ± 1.3; CHO: 24.6 ± 7.2 fg/cell). It should be noted that even a smaller amount of magnetic particles associated with cells is sufficient not only for *in vitro* imaging of cells, but also for tracking cells in a living organism [28].

Therefore, we synthesized magnetic and magneto-fluorescent particles with the desired features: mag-

netism, fluorescence, and controlled surface properties. These particles were effectively used for the labeling of eukaryotic cells, with the integrity and viability of the cells being preserved. The particles can be detected with high sensitivity using the original method for the detection of nonlinear magnetic materials. The synthesized SiO₂-coated nanoparticles may be further linked to various biopolymer molecules [29] and used for tar-

geted drug delivery. In addition, they are promising cell surface markers for such biological and biomedical applications as tissue engineering, various immunoassays, as well as different nanobiotechnology aspects where highly efficient labeling of cells with magnetic particles is necessary in order to further affect the resulting cell-nanoparticle complexes [30]. ●

This study was partially supported by the Russian Science Foundation (project # 14-24-00106, Synthesis, modification, and investigation of magnetic silica nanoparticle stability), Russian Foundation for Basic

Research (project # 17-02-01415, Characterization of the nanoparticle structure and quantitative detection of magnetic markers), and Russian Foundation for Basic Research and by the National Intellectual Development Foundation (project # 17-34-80105 mol_ev_a, Work with the cell cultures, analysis of cytotoxicity, and fluorescence microscopy).

The study was partially performed using equipment provided by the IBCH core facility (CKP IBCH, supported by Russian Ministry of Education and Science, grant RFMEFI62117X0018).

REFERENCES

1. Yoo D., Lee J. H., Shin T. H., Cheon J. // *Accounts Chem. Res.* 2011. V. 44. № 10. P. 863–874.
2. Aghayeva U.F., Nikitin M.P., Lukash S.V., Deyev S.M. // *ACS Nano*. 2013. V. 7. № 2. P. 950–961.
3. Grebenik E.A., Kostyuk A.B., Deyev S.M. // *Russ. Chem. Rev.* 2016. V. 85. № 12. P. 277–296.
4. Nikitin M.P., Shipunova V.O., Deyev S.M., Nikitin P.I. // *Nature Nanotechnol.* 2014. V. 9. № 9. P. 716–722.
5. Wang Y., Shim M.S., Levinson N.S., Sung H.W., Xia Y. // *Adv. Funct. Materials*. 2014. V. 24. № 27. P. 4206–4220.
6. Huang J., Zhong X., Wang L., Yang L., Mao H. // *Theranostics*. 2012. V. 2. № 1. P. 86–102.
7. Gleich B., Weizenecker J. // *Nature*. 2005. V. 435. № 7046. P. 1214–1217.
8. Wilhelm C., Gazeau F., Bacri J.C. // *Eur. Biophys. J.* 2002. V. 31. № 2. P. 118–125.
9. Devkota J., Kokkinis G., Berris T., Jamalieh M., Cardoso S., Cardoso F., Srikanth H., Phan M.H., Giouroudi I. // *RSC Advances*. 2015. V. 5. № 63. P. 51169–51175.
10. Gutierrez L., Mejias R., Barber D.F., Veintemillas-Verdaguer S., Serna C.J., Lazaro F.J., Morales M.P. // *J. Physics D: Appl. Physics*. 2011. V. 44. № 25. P. 255002.
11. Quini C.C., Prospero A.G., Calabresi M.F., Moretto G.M., Zufelato N., Krishnan S., Pina D.R., Oliveira G.M., Baffa O., Bakuzis A.F., et al. // *Nanomedicine: NBM*. 2017. V. 13. № 4. P. 1519–1529.
12. Levy M., Luciani N., Alloyeau D., Elgrabli D., Deveaux V., Pechoux C., Chat S., Wang G., Vats N., Gendron F., et al. // *Biomaterials*. 2011. V. 32. № 16. P. 3988–3999.
13. Nikitin M.P., Vetoshko P.M., Brusnetsov N.A., Nikitin P.I. // *J. Magnetism Magnetic Materials*. 2009. V. 321. № 10. P. 1658–1661.
14. Xi Z., Zheng B., Wang C. // *Nanosci. Nanotechnol. Lett.* 2016. V. 8. № 12. P. 1061–1066.
15. Slowing I.I., Vivero-Escoto J.L., Wu C.W., Lin V.S.Y. // *Adv. Drug Delivery Rev.* 2008. V. 60. № 11. P. 1278–1288.
16. Stober W., Fink A., Bohn E. // *J. Colloidal Interface Sci.* 1968. V. 26. № 1. P. 62–69.
17. Weissleder R., Nahrendorf M., Pittet M.J. // *Nat. Materials*. 2014. V. 13. № 2. P. 125–138.
18. Shipunova V.O., Nikitin M.P., Nikitin P.I., Deyev S.M. // *Nanoscale*. 2016. V. 8. № 25. P. 12764–12772.
19. Sakurai S., Namai A., Hashimoto K., Ohkoshi S.I. // *J. Am. Chem. Soc.* 2009. V. 131. № 51. P. 18299–18303.
20. Yang D., Hu J., Fu S. // *J. Phys. Chem. C*. 2009. V. 113. P. 7646–7651.
21. Mansour H.M., Rhee Y.S., Wu X. // *Int. J. Nanomed.* 2009. V. 4. P. 299–319.
22. Qian L., Yang X.R. // *Adv. Funct. Materials*. 2007. V. 17. № 8. P. 1353–1358.
23. Wu X., Liu H., Liu J., Haley K.N., Treadway J.A., Larson J.P., Ge N., Peale F., Bruchez M.P. // *Nat. Biotechnol.* 2003. V. 21. № 1. P. 41–46.
24. Zhang J., Roll D., Geddes C.D., Lakowicz J.R. // *J. Phys. Chem. B*. 2004. V. 108. № 32. P. 12210–12214.
25. Patil S., Sandberg A., Heckert E., Self W., Seal S. // *Biomaterials*. 2007. V. 28. № 31. P. 4600–4607.
26. Bahadar H., Maqbool F., Niaz K., Abdollahi M. // *Iranian Biomed. J.* 2016. V. 20. № 1. P. 1–11.
27. Han Ang W., Dyson P.J. // *Eur. J. Inorganic Chem.* 2006. V. 2006. № 20. P. 4003–4018.
28. Lu C.W., Hung Y., Hsiao J.K., Yao M., Chung T.H., Lin Y.S., Wu S.H., Hsu S.H., Liu H.M., Mou C.Y., et al. // *Nano Lett.* 2007. V. 7. № 1. P. 149–154.
29. Deyev S.M., Lebedenko E.N., Petrovskaya L.E., Dolgikh D.A., Gabibov A.G., Kirpichnikov M.P. // *Russ. Chem. Rev.* 2015. V. 84. P. 1–26.
30. Merkoci A. // *Biosensors Bioelectronics*. 2010. V. 26. № 4. P. 1164–1177.

A Novel Hybrid Promoter ARE-hTERT for Cancer Gene Therapy

S. V. Kalinichenko^{1#}, M. V. Shepelev^{1#*}, P. N. Vikhreva^{1,2}, I. V. Korobko¹

¹Institute of Gene Biology, Russian Academy of Sciences, Vavilova Str. 34/5, Moscow, 119334, Russia

²MRC Toxicology Unit, University of Leicester, Leicester, UK

#These authors equally contributed to this work.

*E-mail: mshepelev@mail.ru

Received: July 09, 2017; in final form October 30, 2017

Copyright © 2017 Park-media, Ltd. 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 describe a novel hybrid tumor-specific promoter, ARE-hTERT, composed of the human *TERT* gene promoter (hTERT) and the antioxidant response element (ARE) from the human *GCLM* gene promoter. The hybrid promoter retains the tumor specificity of the basal *hTERT* promoter but is characterized by an enhanced transcriptional activity in cancer cells with abnormal activation of the Nrf2 transcription factor and upon induction of oxidative stress. In the *in vitro* enzyme-prodrug cancer gene therapy scheme, ARE-hTERT promoter-driven expression of CD:UPRT (yeast cytosine deaminase:uracil phosphoribosyltransferase) chimeric protein induced a more pronounced death of cancer cells either upon treatment with 5-fluorouracil (5FC) alone or when 5FC was combined with chemotherapeutic drugs as compared to the hTERT promoter. The developed hybrid promoter can be considered a better alternative to the hTERT promoter in cancer gene therapy schemes. **KEYWORDS** hTERT promoter, ARE elements, oxidative stress, hybrid promoter, cancer gene therapy, tumor-specific promoter.

ABBREVIATIONS ARE – antioxidant response element; CMV – cytomegalovirus; *GCLM* – glutamate-cysteine ligase modifier subunit; *TERT* – telomerase reverse transcriptase; tBHQ – tert-butylhydroquinone; ROS – reactive oxygen species; RLU – relative light units; SD – standard deviation; CD:UPRT – yeast cytosine deaminase:uracil phosphoribosyltransferase chimeric protein; 5FC – 5-fluorouracil.

INTRODUCTION

Gene therapy is a strategy that is witnessing dynamic development; several drugs have already been approved for clinical use, many are undergoing various phases of clinical trials, and a vast number of drugs are under development at laboratories. Various approaches to provide specific activity of gene therapy agents in cancer cells have been proposed and validated, including post-transcriptional regulation of the therapeutic transgene expression level in cancer cells via the selective stabilization of the transcript in tumor cells [1] or destabilization of the transcript in normal cells [2]. Along with these relatively new approaches, activation of transgene expression predominantly in tumor cells by tumor-specific promoters remains one of the most frequently used, well-explored, and justified strategies to provide tumor specificity in cancer gene therapy [3].

Tumor-specific promoters have been successfully validated in many cases, but their usage is associated with several drawbacks, including a far from absolute tumor specificity and low promoter activity, which af-

fects the transgene expression level and, consequently, the therapeutic effect. In particular, the human telomerase reverse transcriptase (hTERT) gene promoter is one of the best characterized tumor-specific promoters and is active in a wide variety of tumors, providing the advantage of targeting cancer cells of different origins [3–7]. Nevertheless, the hTERT promoter is relatively weak, which might affect the overall efficiency of the therapy in clinical settings. So, several attempts to improve its activity have been made. Since the hTERT promoter is “TATA-less,” two modifications were proved to increase promoter activity: linking the promoter to a synthetic TATA-box or to a minimal early/immediate cytomegalovirus promoter to provide conventional basal promoter elements [8, 9]. However, the activity of the hTERT promoter varies greatly among different tumor cell lines, which might compromise the advantage of its universality [10]. Therefore, taking into account the above considerations, there is a definite need to increase hTERT promoter activity in tumor cells while retaining its tumor cell specificity.

Many promoters like hTERT show tumor-specific activity due to reactivation in tumor cells, and the tumor specificity of transgene expression can be further increased by exploiting genetic regulatory elements that respond to a perturbed tumor microenvironment or are abnormally active due to somatic mutations in tumor cells. This strategy can be exemplified by exploiting *cis*-acting regulatory elements that provide a transcriptional response to oxidative stress or hypoxia, which are the hallmarks of many tumors. In particular, antioxidant response elements (ARE), the binding sites for the Nrf2 transcription factor, which is a master activator of the oxidative stress response, suffice to support tumor-specific transgene expression when linked to a basal promoter [11]. In such a setting, transcription is maintained due to the abnormal Nrf2 activation occurring in response to intrinsic oxidative stress in tumor cells or because of somatic mutations resulting in constitutive Nrf2 activation.

In this paper, we show that combining the tumor-specific hTERT promoter with ARE results in increased activity of the hybrid promoter in tumor cells compared to the hTERT promoter. At the same time, this modification did not affect promoter activity in non-cancerous cells, in which Nrf2 is not activated under normal conditions. This approach can be used to increase the transgene expression level and activity of therapeutic proteins in tumor cells without an appreciable loss of tumor specificity.

MATERIALS AND METHODS

Cell culture

Human lung epidermoid carcinoma Calu-1 (ECACC #93120818), nonsmall cell lung carcinomas NCI-H1299 (ATCC #CRL-5803) and A549 (ATCC #CRL-185), and nonsmall cell lung bronchioalveolar carcinoma NCI-H358 (ATCC #CRL-5807) cell lines were cultured in a DMEM/F12 (1 : 1) medium (HyClone, USA) supplemented with 10% fetal bovine serum (HyClone), penicillin (100 U/ml), and streptomycin (100 µg/ml) (Gibco, UK). Human bronchial epithelial cells HBEpC (ECACC #502-05) were cultured in the Bronchial Epithelial Growth Medium (Lonza, Switzerland). For cell viability or luciferase reporter gene assays, cells were seeded into 24-well plates at an indicated density (NCI-H1299, 20000 cells/well; A549, 30000 cells/well; Calu-1, 40000 cells/well; NCI-H358, 150000 cells/well; HBEpC, 80000 cells/well) and transfected with a Unifectin-56 transfection reagent (Rusbiolink, Russia) the next day.

Plasmids

The plasmid pH_{TERT}-Luc encoding firefly luciferase under the control of a -206...+37 nt *hTERT* promoter

was described earlier [10]. Plasmid pARE-hTERT-Luc containing firefly luciferase cDNA under the control of the hybrid ARE-hTERT promoter was generated by cloning 56 bp ARE (tgagtaacggttacgaagcactttctcggtacgatttctgcttagtcattgtctt) from the human glutamate-cysteine ligase modifier (*GCLM*) gene promoter to the 5'-end of the hTERT promoter in pH_{TERT}-Luc plasmid [12]. pH_{TERT}-CD : UPRT and pARE-hTERT-CD:UPRT plasmids for the expression of the yeast cytosine deaminase:uracil phosphoribosyltransferase (CD : UPRT) chimeric protein under the control of hTERT or ARE-hTERT promoter, respectively, were constructed on the backbone of pBluescriptII SK(-) vector (Stratagene, USA) [13]. The SV40 signal for transcription termination and polyadenylation derived from pBK-CMV expression vector (Stratagene) was cloned at the 3'-end of CD : UPRT cDNA.

Chemicals

Tert-butylhydroquinone (tBHQ), doxorubicin, cisplatin, etoposide, and 5-fluorocytosine (5FC) were purchased from Sigma-Aldrich (USA).

Luciferase reporter gene assay

Cells were transfected in triplicate for each plasmid combination by a mixture of the firefly reporter plasmid (pH_{TERT}-Luc, pARE-hTERT-Luc or promoterless pGL3-Basic plasmid (Promega)) with pRL-CMV (Promega, USA) plasmid (encoding the *Renilla* luciferase reporter gene under the control of the CMV immediate early enhancer/promoter). If indicated, cells were treated with 100 µM tBHQ for 24 hrs prior to harvesting for luciferase activity analysis. Luciferase activities were quantified 2 days after transfection using the Dual-Luciferase[®] Reporter Assay System (Promega). Firefly luciferase activity was normalized to the *Renilla* luciferase activity, and the average values of relative light units (RLU) and standard deviation (SD) were calculated.

Cell viability assay

Cells were transfected with CD : UPRT-encoding plasmids or mock-transfected with the pBK-CMV vector and plated into the wells of a 96-well plate 24 hrs after transfection (2,000 cells/well for NCI-H1299, A549 and Calu-1 cell lines, and 5,000 cells/well for NCI-H358 cells). 5FC and/or etoposide, cisplatin, or doxorubicin were added 24 hrs after plating. Cell culture medium containing 5FC and/or chemotherapeutic drugs was changed for a fresh one after 24 and 96 hrs of incubation. If indicated, tBHQ was present in the medium at a concentration of 100 µM between 24 and 96 hrs of incubation. Cell viability was determined after 120 hrs of incubation using the CellTiter96[®] AQueous One

Solution Cell Proliferation Assay (Promega) according to the manufacturer's instruction. Each experimental point was analyzed in triplicate. Cell viabilities were normalized to the viability of cells incubated in the absence of 5FC and chemotherapeutic drugs, which was taken as 100%.

RESULTS

Design of the hybrid ARE-hTERT promoter

Hurttila et al. previously analyzed the potency of AREs derived from several Nrf2-responsive genes to support transgene expression under oxidative stress conditions and showed that the highest expression level among the tested AREs was provided by ARE from the *GCLM* gene promoter [12]. Based on these findings, we used ARE from the human *GCLM* gene promoter and placed it at the 5'-end relative to the transcription start site of the -206...+37 nt fragment of the hTERT promoter, which is sufficient to support tumor-specific transcription [5].

Activity of the hybrid ARE-hTERT promoter in cancer and normal cells

We used a luciferase reporter assay to compare the activities of the conventional and the hybrid hTERT promoters in cancer and normal cells. In three out of four tested lung cancer cell lines (NCI-H1299, Calu-1 and A549), the ARE-hTERT promoter showed 2- to 3-fold higher activity compared to the unmodified hTERT promoter, while in NCI-H358 cell line the introduction of ARE did not significantly improve the promoter activity (*Fig. 1* "-tBHQ" samples). Importantly, a similar lack of effect on the transcriptional activity of the hybrid promoter was observed in non-cancerous HBEpC cells: relative luciferase activities in pARE-hTERT-Luc- and phTERT-Luc-transfected cells were 1.41 ± 0.45 and 1.00 ± 0.214 ($P = 0.2272$, two-tailed Student's *t*-test). Therefore, the ARE-hTERT promoter outperformed the unmodified hTERT promoter in three out of four tested lung cancer cell lines, while modification did not affect the promoter activity in normal cells.

Induction of oxidative stress can stimulate hybrid promoter activity

Next, we studied if induction of oxidative stress could increase ARE-hTERT promoter activity in cancer cells. Treatment of NCI-H358 cells with tBHQ resulted in ~2.5-fold induction of luciferase reporter gene activity under the control of the ARE-hTERT promoter, while having no effect on the activity of the hTERT promoter (*Fig. 1*, "+ tBHQ" samples), which is well in line with the responsiveness of *GCLM* ARE to external oxida-

tive stress [12]. At the same time, tBHQ treatment did not affect ARE-hTERT promoter activity in Calu-1, A549, and NCI-H1299 cell lines in which ARE-hTERT promoter activity significantly outperformed that of the unmodified hTERT promoter without an external oxidative stressor (*Fig. 1*, "- tBHQ" samples). Taken together, the hybrid ARE-hTERT promoter possesses higher activity in cancer cells under basal conditions compared to the hTERT promoter, likely owing to the activation of the Nrf2 pathway that frequently occurs in cancer cells due to somatic mutations or the increased level of reactive oxygen species (ROS) [14–19]. In cells with low basal activity of the hybrid promoter, its transcriptional activity can be boosted by oxidative stress inducers.

A hybrid promoter improves the efficiency of enzyme-prodrug suicide cancer gene therapy *in vitro*

The increased activity of the hybrid promoter compared to the conventional hTERT promoter observed in NCI-H1299, Calu-1, and A549 cells in luciferase reporter gene assay leads one to assume that the modification of the hTERT promoter with ARE will also improve the performance of cancer gene therapy vectors. In order to directly address this issue, we compared the capacities to induce cancer cell death in the enzyme-prodrug CD : UPRT-5FC suicide cancer gene therapy scheme when CD : UPRT expression was driven by either an unmodified or ARE-modified hTERT promoter [20]. As expected, ARE-hTERT promoter-driven CD : UPRT expression resulted in a more pronounced level of cell death in the presence of the same 5FC concentrations compared to hTERT-driven expression (*Fig. 2*). Alike, in agreement with the results of a promoter activity analysis with the reporter gene, ARE modification of the hTERT promoter directing CD : UPRT expression did not affect cytotoxicity for NCI-H358 cells. At the same time, simultaneous treatment of NCI-H358 cells with tBHQ significantly augmented the cytotoxic effect only when the ARE-modified promoter was used (*Fig. 3*), while tBHQ did not increase cytotoxicity when the ARE-hTERT promoter was used instead of the unmodified promoter in NCI-H1299, Calu-1, and A549 cells, where the hybrid promoter is intrinsically more active than the hTERT promoter according to the reporter gene assay (data not shown).

ARE-hTERT promoter-driven enzyme-prodrug suicide cancer gene therapy more efficiently sensitizes cancer cells to conventional chemotherapeutic drugs

ARE-driven enzyme-prodrug suicide cancer gene therapy was reported to increase the sensitivity of

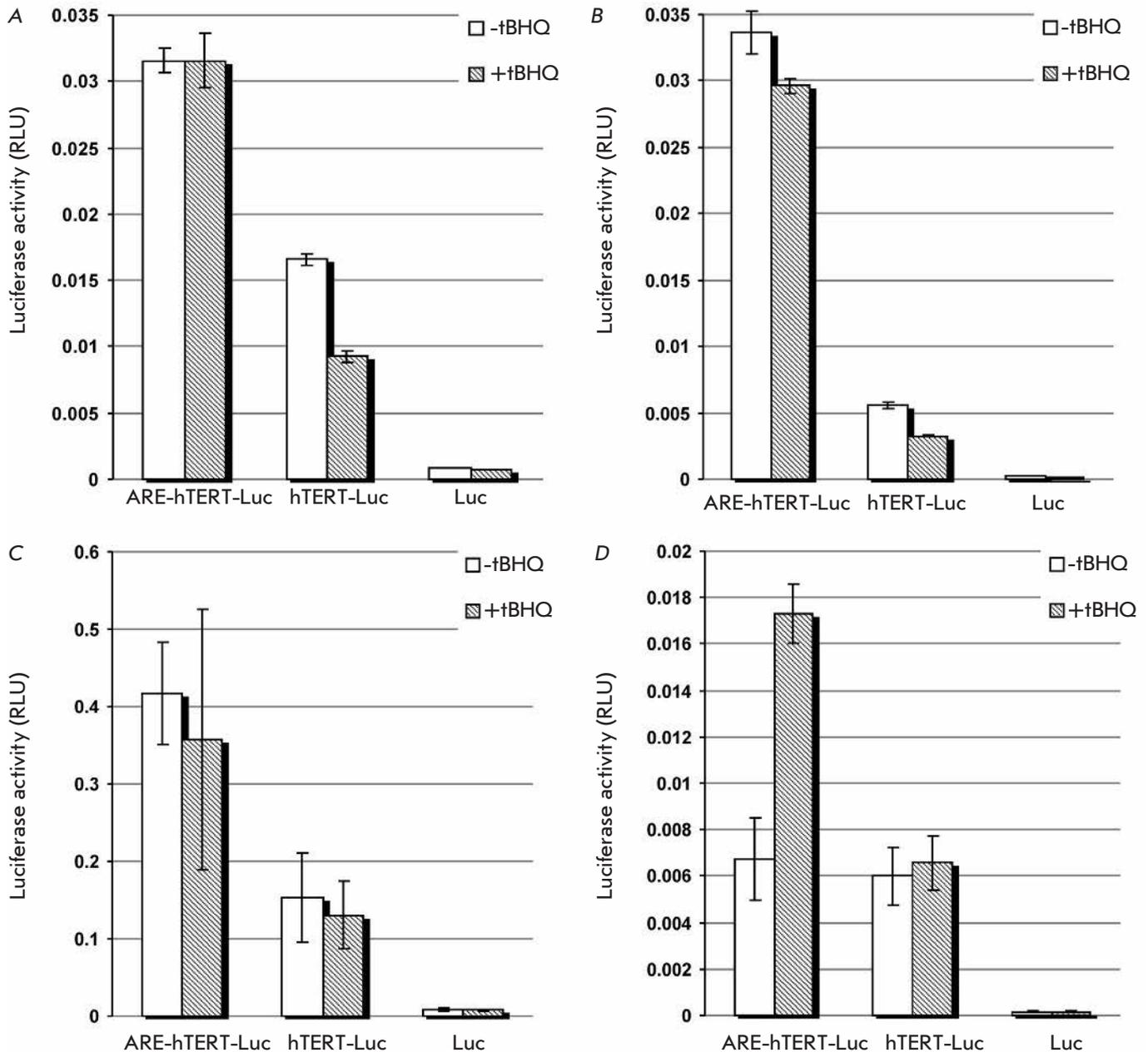


Fig. 1. Effect of ARE-modification of the hTERT promoter on luciferase reporter gene activity in lung cancer cell lines. The luciferase reporter gene activity was measured in NCI-H1299 (A), Calu-1 (B), A549 (C), and NCI-H358 (D) cells transfected with pTERT-Luc (*hTERT-Luc*), pARE-hTERT-Luc (*ARE-hTERT-Luc*), or promoterless pGL3-Basic (*Luc*) plasmid, together with pRL-CMV plasmid for normalization. If indicated (*hatched bars*), cells were treated with 100 μ M tBHQ for 24 hrs. The data are shown as average RLU values \pm SD.

cancer cells to chemotherapeutic drugs: in particular, to doxorubicin [11]. Therefore, we questioned if ARE-hTERT promoter-driven CD : UPRT-5FC enzyme-prodrug suicide cancer gene therapy would also result in enhanced cytotoxicity when combined with chemotherapeutic drugs. *Figure 4A* demonstrates that

hTERT-driven CD : UPRT expression did not result in NCI-H1299 cell death in the presence of 10 μ M 5FC. Also, treatment with 0.1 μ M doxorubicin resulted only in marginal NCI-H1299 cell death. Notably, under the same settings, application of the ARE-modified hTERT promoter resulted in substantial cell death (~40%),

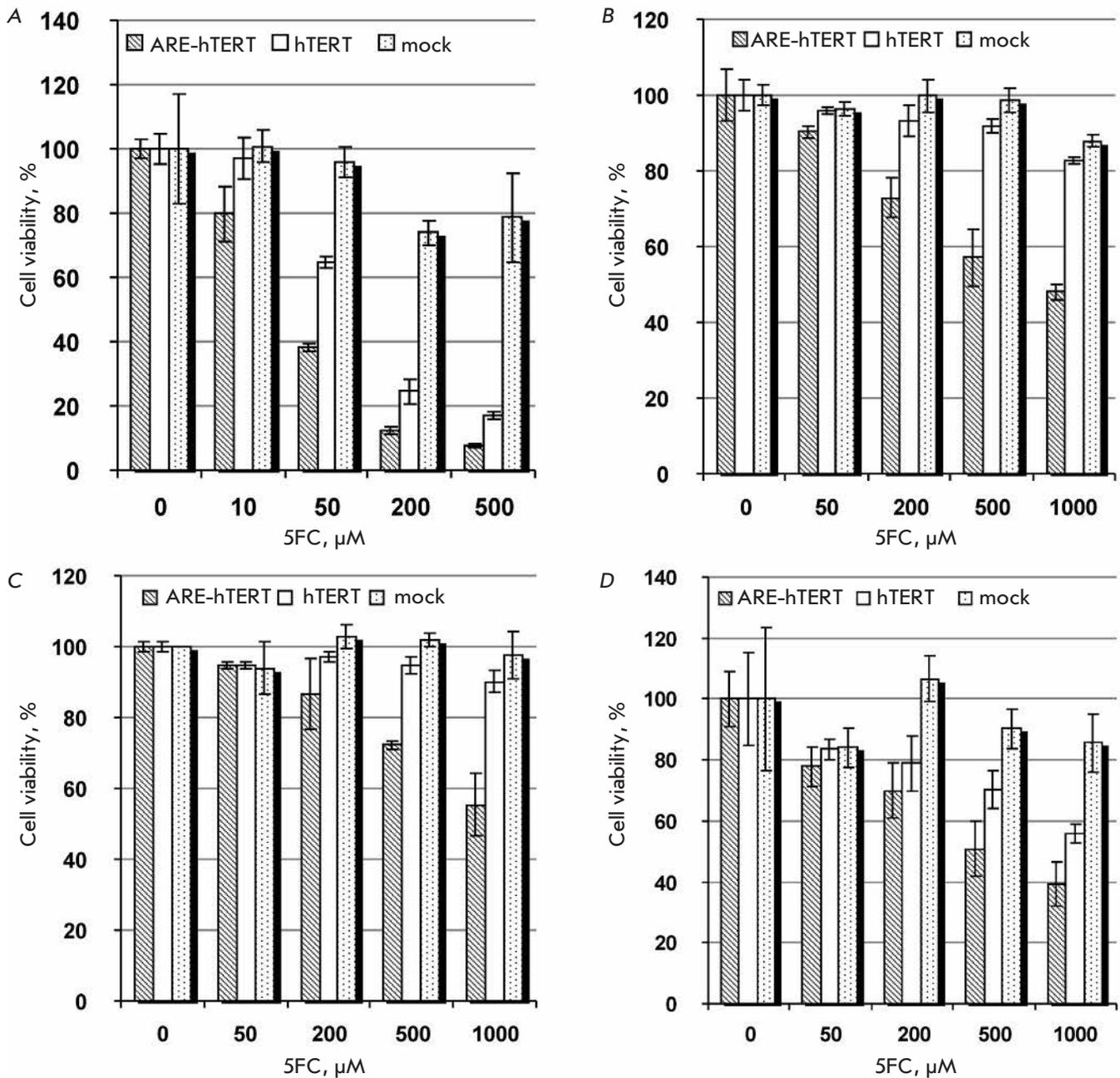


Fig. 2. Comparison of the cytotoxic effects of *hTERT* and *ARE-hTERT* promoter-driven CD:UPRT expression in lung cancer cell lines in the presence of 5FC. Relative viabilities of NCI-H1299 (**A**), Calu-1 (**B**), A549 (**C**), and NCI-H358 (**D**) cells transfected with pARE-hTERT-CD : UPRT (*hatched bars*), pHTERT-CD : UPRT (*blank bars*), and pBK-CMV (mock, *dotted bars*) plasmids after incubation with indicated concentrations of 5FC are shown as average values \pm SD of the percentage of viable cells relative to the viability of similarly transfected cells incubated in the absence of 5FC.

which was further significantly potentiated by combined treatment with doxorubicin (*Fig. 4A*). Similar observations were made for A549 and Calu-1 cells treated with doxorubicin, etoposide or cisplatin (*Fig. 4B* and data not shown). Importantly, under our experimental settings, *hTERT* promoter-driven cancer gene thera-

py showed no effect alone and failed to potentiate the cell death elicited by chemotherapeutic drugs, while modification of the promoter with ARE resulted in pronounced cytotoxicity both in the case of cancer gene monotherapy and when it was combined with chemotherapeutic drugs.

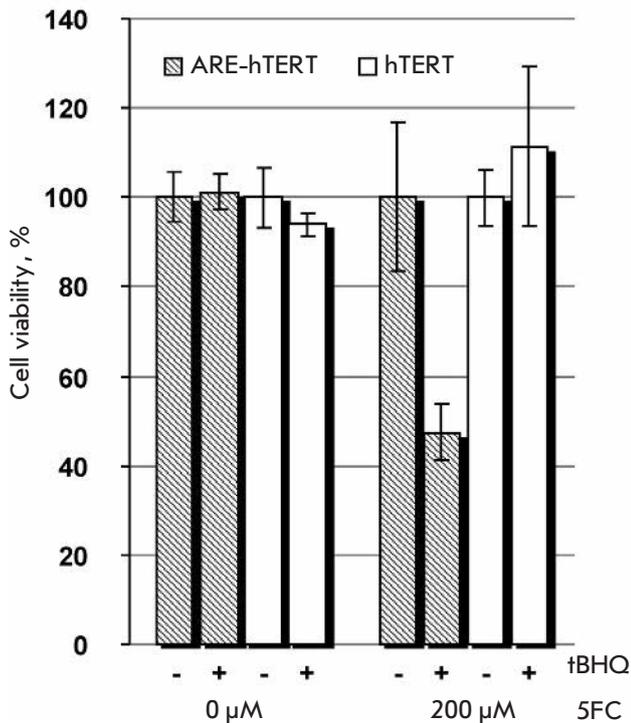


Fig. 3. The cytotoxic effect of ARE-hTERT promoter-driven CD : UPRT expression in the presence of 5FC is enhanced by an oxidative stress inducer in NCI-H358 cells. NCI-H358 cells were transfected with pARE-hTERT-CD : UPRT (hatched bars) or pHTERT-CD : UPRT (blank bars) plasmid, and cell viabilities were determined after incubation in the absence or presence of 200 μ M of 5FC and/or 100 μ M tBHQ as indicated. Data are shown as average values \pm SD of the percentage of viable cells relative to the viability of similarly treated cells incubated in the absence of 5FC.

DISCUSSION

We tested the hypothesis that introduction of ARE into a promoter with intrinsic tumor specificity, which is routinely used to target cancer cells in cancer gene therapy, will enhance promoter activity without an appreciable loss of specificity toward tumor cells. Indeed, ARE linked to a nonselective minimal promoter was previously shown to provide tumor-specific expression owing to an aberrantly activated Nrf2 transcription factor [14–19] or intrinsically higher ROS levels in tumor cells [11, 12]. As we demonstrated, a hybrid promoter containing the human *TERT* gene promoter and ARE derived from the human *GCLM* gene promoter showed better performance in 3 of 4 tested cancer cell lines both in reporter gene assay and in the CD : UPRT-5FC suicidal cancer gene therapy scheme. In NCI-H358 cells, where hTERT promoter modification did not affect the promoter activity (which sug-

gests a lack of abnormal Nrf2 regulation), the activity of the hybrid promoter could be boosted by oxidative stress inducers such as tBHQ. It is important to mention that an analysis in primary epithelial HBEpC cells showed a lack of any appreciable increase in promoter activity after the inclusion of ARE in the promoter, thus demonstrating that the introduced modification did not affect the cancer-cell specificity of transcription.

Our results indicate that the novel hybrid promoter, while retaining a high cancer cell specificity, will outperform the conventional hTERT promoter in a substantial proportion of tumors where Nrf2 is activated due to a somatic mutation. In addition, cancer cells are generally characterized by an increased ROS level both *in vitro* and *in vivo*, which is caused by several factors, such as altered metabolism and inadequate vascularization [21]. In addition, many conventional chemotherapeutic drugs are known to induce oxidative stress; therefore, combination of ARE-hTERT driven cancer gene therapy with conventional chemotherapeutic drugs *in vivo* might further potentiate the overall efficiency of the treatment through the promotion of therapeutic transgene expression.

The efficiency of a cancer gene therapy is primarily determined by the therapeutic transgene expression level, which should be high enough to elicit a therapeutic effect. In this work, we used the CD : UPRT-5FC enzyme-prodrug cancer gene therapy approach, in which the overall efficiency of the therapy is determined by the efficiency of plasmid delivery into cancer cells, promoter activity, and cell sensitivity to the cytotoxic agent obtained from the prodrug conversion. These parameters will obviously vary for specific cell types, potentially resulting in a loss of treatment efficiency. Indeed, under the experimental settings used (Fig. 4), the hTERT-driven gene therapy failed to result in tumor cell elimination or to enhance the cytotoxic effect of chemotherapeutic drugs. However, under identical conditions, modification of the hTERT promoter with ARE restored the cytotoxic effect of the gene therapy and significantly potentiated the chemotherapeutic drug-induced cytotoxicity. These results explicitly demonstrate that the application of a hybrid promoter, instead of the conventional hTERT promoter, would broaden the therapeutic efficiency of gene therapy, thus demonstrating the advantages of the reported hybrid promoter.

CONCLUSIONS

We have created a novel tumor-specific promoter that retains the tumor specificity of the basal hTERT promoter but is characterized by an enhanced transcriptional activity in cancer cells due to either abnormal

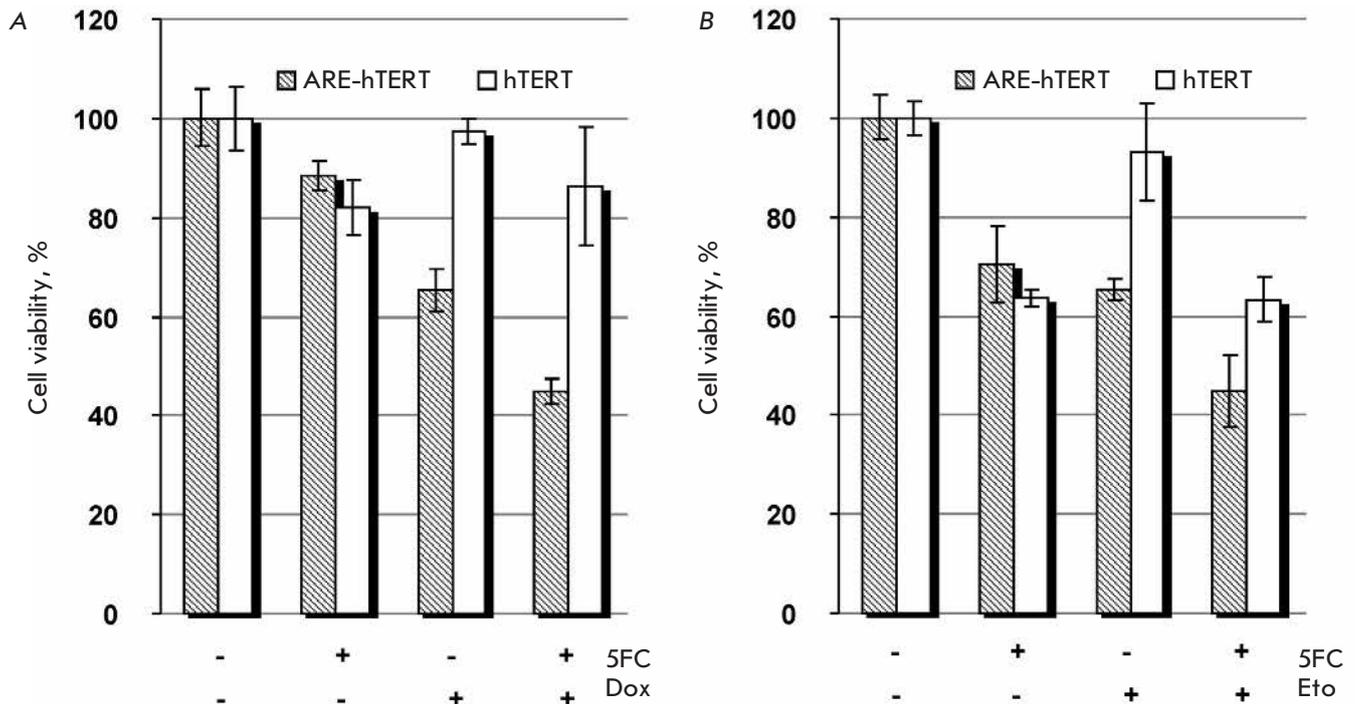


Fig. 4. Comparison of the cytotoxic effects of *hTERT* and *ARE-hTERT* promoter-driven CD : UPRT expression in lung cancer cell lines in the presence of 5FC when combined with chemotherapeutic agents. NCI-H1299 (**A**) or A549 (**B**) cells were transfected with pARE-*hTERT*-CD : UPRT (*hatched bars*) or pH*TERT*-CD : UPRT (*blank bars*) plasmid and incubated in the presence or absence (as indicated) of 10 μ M of 5FC and 0.1 μ M of doxorubicin (Dox) (**A**) or 500 μ M of 5FC and 2 μ M of etoposide (Eto) (**B**). Data are shown as average values \pm SD of the percentage of viable cells relative to the viability of similarly treated cells incubated in the absence of 5FC and chemotherapeutic agents.

Nrf2 transcription factor activation or stimulation with ROS inducers. Owing to the above characteristics, the ARE-*hTERT* hybrid promoter can be considered a better alternative to the *hTERT* promoter in cancer gene therapy schemes. In addition, the combination of ARE with other tumor- or tissue-specific promoters used to

develop vectors for cancer gene therapy can be regarded as a way to improve their performance without an appreciable loss of specificity. ●

This work was supported by the Russian Foundation for Basic Research (grant No. 13-04-40173).

REFERENCES

- Shepelev M.V., Korobko E.V., Georgiev G.P., Sverdlov E.D., Korobko I.V. // *Cancer Gene Ther.* 2011. V. 18. № 9. P. 682–684.
- Korobko I.V. // *Curr. Cancer Ther. Rev.* 2014. V. 10. № 3. P. 271–276.
- Gu J., Fang B. // *Cancer Biol. Ther.* 2003. V. 2. № 4. Suppl. 1. P. S64–70.
- Horikawa I., Cable P.L., Afshari C., Barret J.C. // *Cancer Res.* 1999. V. 59. № 4. P. 826–830.
- Takakura M., Kyo S., Kanaya N., Hirano H., Takeda J., Yutsudo M., Inoue M. // *Cancer Res.* 1999. V. 59. № 3. P. 551–557.
- Horikawa I., Barrett J.C. // *Carcinogenesis.* 2003. V. 24. № 7. P. 1167–1176.
- Poole J.C., Andrews L.G., Tollefsbol T.O. // *Gene.* 2001. V. 269. № 1–2. P. 1–12.
- Wirth T., Zender L., Schulte B., Mundt B., Plentz R., Rudolph K.L., Manns M., Kubicka S., Kühnel F. // *Cancer Res.* 2003. V. 63. № 12. P. 3181–3188.
- Davis J.J., Wang L., Dong F., Zhang L., Guo W., Teraishi F., Xu K., Ji L., Fang B. // *Cancer Gene Ther.* 2006. V. 13. № 7. P. 720–723.
- Shepelev M.V., Kopantzev E.P., Vinogradova T.V., Sverdlov E.D., Korobko I.V. // *Oncol. Lett.* 2016. V. 12. № 2. P. 1204–1210.
- Leinonen H.M., Ruotsalainen A.K., Määttä A.M., Laitinen H.M., Kuosmanen S.M., Kansanen E., Pikkarainen J.T., Lappalainen J.P., Samaranayake H., Lesch H.P., et al. // *Cancer Res.* 2012. V. 72. № 3. P. 6227–6235.
- Hurttila H., Koponen J.K., Kansanen E., Jyrkkänen H.K., Kivelä A., Kylätie R., Ylä-Herttuala S., Levonen A.L. // *Gene Ther.* 2008. V. 15. № 18. P. 1271–1279.
- Kuzmin D.V., Vinogradova T.V., Kopantzev E.P., Sverdlov

- ED. // *Open Gene Ther. J.* 2010. V. 3. P. 31–39.
14. Hayes J.D., McMahon M. // *Trends Biochem. Sci.* 2009. V. 34. № 4. P. 176–188.
15. Kim Y.R., Oh J.E., Kim M.S., Kang M.R., Park S.W., Han J.Y., Eom H.S., Yoo N.J., Lee S.H. // *J. Pathol.* 2010. V. 220. № 4. P. 446–451.
16. Konstantinopoulos P.A., Spentzos D., Fountzilas E., Francoeur N., Sanisetty S., Grammatikos A.P., Hecht J.L., Canistra S.A. // *Cancer Res.* 2011. V. 71. № 15. P. 5081–5089.
17. Seng S., Avraham H.K., Birrane G., Jiang S., Li H., Katz G., Bass C.E., Zagodzoon R., Avraham S. // *Oncogene.* 2009. V. 28. № 3. P. 378–389.
18. Zhang P., Singh A., Yegnasubramanian S., Esopi D., Kombaraju P., Bodas M., Wu H., Bova S.G., Biswal S. // *Mol. Cancer Ther.* 2010. V. 9. № 2. P. 336–346.
19. Cairns R.A., Harris I.S., Mak T.W. // *Nat. Rev. Cancer.* 2011. V. 11. № 2. P. 85–95.
20. Adachi Y., Tamiya T., Ichikawa T., Terada K., Ono Y., Matsumoto K., Furuta T., Hamada H., Ohmoto T. // *Hum. Gene Ther.* 2000. V. 11. № 1. P. 77–89.
21. Brown N.S., Bicknell R. // *Breast Cancer Res.* 2001. V. 3. № 5. P. 323–327.

Multilocus Analysis of Genetic Susceptibility to Myocardial Infarction in Russians: Replication Study

N. G. Kukava^{1#}, B. V. Titov^{1,2#}, G. J. Osmak^{1,2}, N. A. Matveeva^{1,2}, O. G. Kulakova^{1,2},
A. V. Favorov³, R. M. Shakhnovich¹, M. Ya. Ruda¹, O. O. Favorova^{1,2*}

¹National Medical Scientific Center for Cardiology, 3rd Cherepkovskaya Str. 15a, Moscow, 121552, Russia

²Pirogov Russian National Research Medical University, Ostrovitjanova Str. 1, Moscow, 117997, Russia

³Oncology Biostatistics and Bioinformatics, Johns Hopkins School of Medicine, Baltimore, MD 21205, US

#These authors contributed equally to this work

*E-mail: olga.favorova@gmail.com

Received: April 14, 2017; in final form September 28, 2017

Copyright © 2017 Park-media, Ltd. 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 search of genetic markers of myocardial infarction (MI) risk, which have prognostic significance for Russians, we performed a replication study of MI association with genetic variants of *PCSK9* (rs562556), *APOE* (epsilon polymorphism, rs7412 and rs429358), *LPL* (rs320), *MTHFR* (rs1801133), *eNOS* (rs2070744), and the 9p21 region (rs1333049) in 405 patients with MI and 198 controls. Significant MI association was observed with variants of the lipid metabolism genes (*PCSK9*, *APOE* and *LPL*), and of *eNOS*. The SNPs in the *MTHFR* gene and the 9p21 region were not significantly associated with MI one by one but were included in several different MI-associated allelic combinations identified by multilocus analysis. Since we have not revealed nonlinear epistatic interactions between the components of the identified combinations, we postulate that the cumulative effect of genes that form a combination arises from the summation of their small independent contributions. The prognostic significance of the additive composite model built from the *PCSK9*, *APOE*, *LPL*, and *eNOS* genes as genetic markers was assessed using ROC analysis. After we included these markers in the previously published composite model of individual genetic risk of MI, the prognostic efficacy in our sample reached AUC = 0.676. However, the results obtained in this study certainly need to be replicated in an independent sample of Russians.

KEYWORDS myocardial infarction, Russians, genes, allelic polymorphism, multilocus analysis, genetic markers

ABBREVIATIONS AUC – area under curve in ROC analysis; GWAS – genome-wide association study; NO – nitrogen oxide; ROC – receiver operating characteristic; SNP – single nucleotide polymorphism; CI – confidence interval; CAD – coronary artery disease; MI – myocardial infarction; OR – odds ratio; PCR – polymerase chain reaction; PCR-SSP – polymerase chain reaction using allele-specific primers; CVD – cardiovascular disease, m. a. – mean age.

INTRODUCTION

Myocardial infarction (MI) is the most severe form of coronary artery disease (CAD). Although significant progress in prevention and treatment of cardiovascular diseases (CVDs) has been achieved over the past decades in the developed countries, MI still remains the leading cause of death worldwide.

Both MI and CAD are polygenic and multifactorial diseases; the non-Mendelian inheritance pattern characterizing them results from the interplay between genetic variants. The genetic predisposition to CAD has been well investigated in genome-wide association

studies (GWAS), while the number of GWASs for MI as a particular phenotype is relatively small [1, 2]. The rather poor replicability of the few MI-associated loci identified in separate studies may be due to ethnic differences between the samples. Although the objective of the GWASs is to identify genetic variants that would enable assessing the risk of MI, no progress in predicting the risk of this disease has been made yet [3].

It is no wonder that the conventional candidate gene approach still remains relevant. Vast amounts of data have been accumulated on the association of individual candidate genes with MI in Russian population; much of

Table 1. Genes included in the study and their polymorphic regions

Gene	Chromosomal locus	Gene product(s)	Polymorphic region*	
			SNP	rs ID
<i>PCSK9</i>	1p32.3	Proprotein convertase subtilisin/kexin type 9	1420G > A	rs562556
<i>APOE</i>	19q13.2	Apolipoprotein E	epsilon polymorphism	(rs7412, rs429358)
<i>LPL</i>	8p22	Lipoprotein lipase	495T > G (HindIII H+>H-)	rs320
<i>MTHFR</i>	1p36.22	5,10-Methylene tetrahydrofolate reductase	677C > T	rs1801133
<i>eNOS</i> (also known as <i>NOS3</i>)	7q36	Endothelial nitric oxide synthase	-786T > C	rs2070744
the <i>ANRIL-CDKN2A/2B</i> gene cluster	9p21.3	Long non-coding RNA (the <i>ANRIL</i> gene); cyclin-dependent kinase inhibitors 2A and 2B (the <i>CDKN2A</i> and <i>CDKN2B</i> genes)	C > G	rs1333049

*The examined single nucleotide polymorphism (SNP) and its designation according to the reference nucleotide sequence of the human genome (rs ID).

these data have been obtained by Russian participants of the MONICA [4] and HAPIEE [5] international research projects. Replication of the results obtained both for the independent samples consisting of subjects belonging to the same ethnic group and other ethnic populations is believed to play a special role in identifying factors of genetic predisposition. We have earlier found that the genetic variants of the *FGB*, *TGFB1*, *CRP*, *IFNG*, and *PTGS1* genes, whose products are involved in the inflammation and coagulation systems, are associated with the risk of MI development in ethnic Russians and replicated these results in an independent sample of Russians [6]. The prognostic significance of the identified markers has been demonstrated; summing up the contributions of individual genes significantly increased the prognostic efficacy. However, the identified loci explain only a small contribution to the risk of MI. In pursuit of other genetic markers for MI risk that would have prognostic significance for Russians, we have broadened the candidate gene list under study by including the lipid metabolism genes (*PCSK9*, *LPL*, and *APOE*), the *MTHFR* and the *eNOS* genes, and the 9p21 locus.

The products of the selected lipid metabolism genes are known to be involved in the development of CVDs. The *PCSK9* protein (proprotein convertase subtilisin/kexin type 9) encoded by the *PCSK9* gene partakes in degradation of low-density lipoprotein receptors and is used as a target in treatment of dyslipidemia and related CVDs [7]. The *APOE* gene product, apolipoprotein E, is involved in lipid transport and plays a crucial role in the development of CVD [8]. Lipoprotein lipase encoded by the *LPL* gene is a key enzyme in lipid metabolism and transport; it also participates in pathogenesis of atherosclerosis [9].

The role of products of the *MTHFR* and *eNOS* genes in CVD pathogenesis is also well-known. The *MTHFR* gene codes for methylenetetrahydrofolate reductase, the enzyme involved in conversion of homocysteine to methionine. Homocysteinaemia may cause endothelial dysfunction, which is a risk factor for atherosclerosis and CVDs related to it [10]. Endothelial nitric oxide (NO) synthase encoded by the *eNOS* gene catalyzes production of NO involved in regulation of vascular tone and permeability; disturbances in the NO system may lead to atherosclerosis, hypertension, and thrombosis [11].

The MI association with the rs1333049 polymorphism on chromosome 9p21 was revealed in several GWASs and has been validated in a number of ethnic groups. This locus carries the gene of non-coding regulatory RNA *ANRIL*. This RNA may regulate the expression of cyclin-dependent kinase inhibitors p15^{INK4a} and p16^{INK4b}, which are encoded by the *CDKN2A* and *CDKN2B* genes residing within the same region. It is believed that the 9p21 region can participate in pathogenesis of atherosclerosis by regulating proliferation and apoptosis of smooth muscle cells [12].

The aim of our study was to conduct a replication study of the association of the polymorphic variants in the *PCSK9*, *APOE*, *LPL*, *MTHFR*, and *eNOS* genes and the 9p21 region with the risk of MI development in Russians. *Table 1* lists the characteristics of the selected genes and single nucleotide polymorphisms (SNPs). We also carried out a multilocus analysis of the association between the combinations of variants of these genes/loci and MI, since the cumulative genetic effect can be identified using this approach [13]. The nature of this effect was also studied. Furthermore, we evaluated the

prognostic efficacy of the identified markers both one by one and along with the markers identified previously [6].

EXPERIMENTAL

Genomic DNA samples collected from patients receiving treatment at the Emergency Cardiology Department (National Medical Research Center of Cardiology, the Ministry of Health of the Russian Federation) were used in the case-control study. The study group consisted of 405 ethnic Russians (mean age (m. a.) \pm standard deviation, 57.5 ± 12.8 years): 271 males (m.a., 53.4 ± 11.9 years) and 134 females (m.a., 65.6 ± 10.3 years). The diagnosis of MI was made using the criteria described in [14]. The control group consisted of 198 Russian subjects with no past history of CVD (mean age, 59.8 ± 13.3 years): 112 males (m.a., 57.1 ± 11.9 years) and 86 females (m.a., 63.2 ± 14.2 years). All patients provided informed consent for participating in the study.

Genomic typing was performed using the polymerase chain reaction (PCR)-based methods. Restriction fragment length polymorphism analysis of the PCR products was carried out to detect the *APOE* gene epsilon polymorphism (rs7412, rs429358) [15], 495T > G in an *LPL* gene (rs320) [16], 677C > T in the *MTHFR* gene (rs1801133) [17], and -786T>C in the *eNOS* gene (rs2070744) [18]. Genome typing of the polymorphisms rs562556 in the *PCSK9* gene and rs1333049 in the 9p21.3 region was performed by real-time PCR using a TaqMan® SNP Genotyping Assay kit (Applied Biosystems).

Statistical analysis

The deviations of genotype frequencies from the Hardy-Weinberg equilibrium were analyzed using Haploview 4.2 software [19]. APSampler software was used to search for the associations between carriage of alleles and genotypes of individual polymorphisms or their combinations and development of MI [20]. The significance of the revealed associations was assessed using the Fisher's exact test and the odds ratio (OR). The Bonferroni correction for the number of tests (multiple comparisons) was used for the p values calculated using the Fisher's exact test (p_{corr}). The p and p_{corr} values < 0.05 were considered significant when the 95% confidence interval (CI) values for OR did not cross 1. An SNP was considered to be myocardial infarction-associated when the association was significant either in the recessive or the dominant model.

The earlier proposed approach [6] was used to reveal possible non-linear interactions (epistasis) between alleles in the identified biallelic combinations: the syn-

ergy factor (SF) was determined [22] and the p values were calculated using the exact three-way interaction test [21]. The interaction between the alleles was considered to be epistatic if the p value was less than 0.05 and the 95% CI value for SF did not cross 1.

The prognostic models were built using the stepwise logistic regression method (Stats v.3.3.1 for R). The prognostic efficacy was assessed by ROC (receiver operating characteristic) analysis by measuring the area under the curve (AUC) using pROC v.1.8 for R software package; pairwise comparisons were made using the method described in [23]. The probability threshold was calculated using the procedure described in [24] to assess sensitivity and specificity of the prognostic models.

RESULTS

All studied polymorphic regions in the control group were in Hardy-Weinberg equilibrium ($p > 0.05$). *Figure 1* shows the allele frequencies for all the examined loci in the control group compared to minor allele frequencies (global MAF) in the SNP Database under the 1000 Genomes Project (Phase 3) [25]. The absolute differences between the observed allele frequencies and those deposited in the SNP database are $\leq 10\%$.

Table 2 summarizes the data on carriage of the alleles and genotypes of the *PCSK9* (rs562556), *APOE* (epsilon polymorphism, rs7412 and rs429358), *LPL* (rs320), *MTHFR* (rs1801133), and *eNOS* (rs2070744) genes and the 9p21 region (rs1333049) in 405 MI patients and 198 controls. Significant differences were revealed in carriage frequencies of alleles and genotypes of polymorphic regions for all three lipid metabolism genes: *PCSK9*, *APOE*, and *LPL*. Significant differences were also found for the *eNOS* gene but not for the *MTHFR* gene or the 9p21 region. The *PCSK9**A/A ($p = 0.013$, OR = 1.45), *APOE** $\epsilon 3/\epsilon 3$ ($p = 0.034$, OR = 1.52), and *LPL** G/G ($p = 0.032$, OR = 1.96) genotypes and carriage of the *eNOS**C allele ($p = 0.0034$, OR = 1.63) were found to be the risk factors of MI. However, the p_{corr} value calculated using the Bonferroni correction for the number of tests (multiple comparisons) was significant only for the genetic variants of the *APOE* and *eNOS* genes.

We used the APSampler software employing the dynamic Monte Carlo method to carry out a multilocus analysis aimed at identifying the cumulative contribution of combinations of the alleles and genotypes of the genes under study to predisposition to MI. The revealed bi- and triallelic combinations associated with the risk of MI are characterized by a stronger effect and a greater significance level of association with MI than their individual components. Along with the *PCSK9*, *APOE*, *LPL*, and *eNOS* genetic variants, the combina-

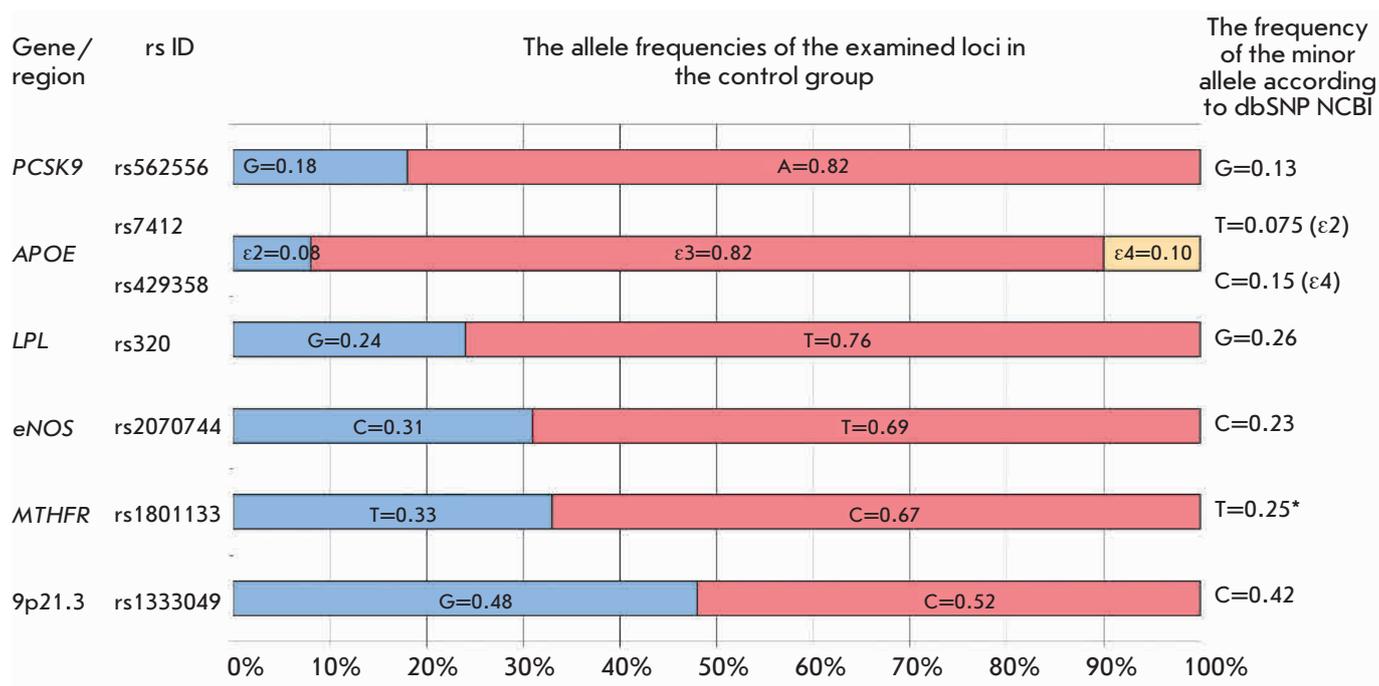


Fig. 1. The allele frequencies of the examined loci in the control group (ethnic Russians) as compared to the allele frequencies from the dbSNP NCBI database [25].

*In the dbSNP NCBI database, minor allele frequency (MAF) is shown for the complementary chain.

tions also include the alleles/genotypes of the *MTHFR* gene and SNP rs1333049 in the 9p21 region. *Figure 2 (A–C)* illustrates the OR and 95% CI values for the combinations containing the variants of the latter two loci: *MTHFR**C, rs1333049*C, and rs1333049*C/G. One can see in all these cases that the variants shown at the bottom of each figure are not significant. However, the combination of the *MTHFR**C and *eNOS**C alleles is significant ($p = 0.0006$; OR = 1.80): more significant than carriage of a single *eNOS**C allele (*Fig. 2A*). The triallelic combination (*LPL**G/G + *MTHFR**C + rs1333049*C) ($p = 0.018$; OR = 2.83) and the biallelic combination being a part of it (*LPL**G/G + *MTHFR**C) ($p = 0.021$; OR = 2.30) are also associated with the risk of MI; the association of the biallelic combination with the risk of MI is less significant than MI association with the triallelic combination but more significant than MI association with the single *LPL**G/G genotype (*Fig. 2B*). Another triallelic combination (*APOE**ε4 + *eNOS**T + rs1333049*C/G) is negatively associated with the risk of MI ($p = 0.00041$; OR = 0.30) (*Fig. 2C*). The association of the biallelic combinations being a part of it (*APOE**ε4 + rs1333049*C/G) and (*APOE**ε4 + *eNOS**T) with the risk of MI is less significant. However, it is stronger than MI association with carriage of the *APOE**ε4 allele, the only component of the combination that is significant alone. Hence, we have used multilocus analysis to

identify that the genetic variants of the *MTHFR* gene (rs1801133) and 9p21 locus (rs1333049) within several allelic combinations are involved in predisposition to MI, while the genetic variants one by one showed no significant association with MI.

In order to answer the question what is the reason for the cumulative effect of the alleles of different genes (whether it is the summation of small mutually independent contributions of individual alleles or epistatic interactions between these alleles), we analyzed the three-way interactions using the statistical approach described earlier [6]. The synergy factor (SF) with 95% CI and the p values calculated using the exact three-way test, similar to the OR with 95% CI and the p value determined by standard evaluation of the associations between the phenotype and the genotype (i.e., using the two-way Fisher's test), were not significant. Therefore, no significant epistatic interactions were revealed between the components of all the identified combinations.

In order to assess the prognostic significance of the identified genetic risk factors, we calculated the individual risk of MI in each subject depending on carriage of the *PCSK9*, *APOE*, *LPL*, and *eNOS* genetic variants using logistic regression. The contribution of carriage of a combination of these risk alleles/genotypes was evaluated using ROC analysis (*Fig. 3A*)

Table 2. Distribution of alleles and genotypes of polymorphic regions of the examined genes in MI patients ($n = 405$) and controls ($n = 198$)

Carriage of alleles and genotypes	Patients, n (%)	Controls, n (%)	p	p_{corr}^*	OR (95% CI) for significant differences**
rs562556 in PCSK9					
A	389(96)	193(97)	NS	NS	
G	102(25)	65(33)	0.013	NS	0.69 (0.47–1.00)
A/A	303(75)	133(67)	0.013	NS	1.45 (1.00–2.10)
A/G	86(21)	60(30)	0.010	NS	0.62 (0.42–0.91)
G/G	16(4)	5(3)	NS	NS	
rs7412, rs429358 (epsilon polymorphism) in APOE					
$\epsilon 2$	63(16)	30(15)	NS	NS	
$\epsilon 3$	393(98)	194(98)	NS	NS	
$\epsilon 4$	40(10)	38(19)	0.0013	0.0091	0.46 (0.28–0.75)
$\epsilon 2/\epsilon 2$	5(1)	2(1)	NS	NS	
$\epsilon 2/\epsilon 3$	55(14)	26(13)	NS	NS	
$\epsilon 2/\epsilon 4$	3(1)	2(1)	NS	NS	
$\epsilon 3/\epsilon 3$	305(75)	132(67)	0.017	NS	1.52 (1.05–2.2)
$\epsilon 3/\epsilon 4$	33(8)	36(18)	0.00033	0.0023	0.40 (0.24–0.66)
$\epsilon 4/\epsilon 4$	4(1)	0(0)			
rs320 in LPL					
G	192(47)	85(43)	NS	NS	
T	363(90)	187(94)	0.032	NS	0.51 (0.25–0.99)
G/G	42(10)	11(6)	0.032	NS	1.96 (1.00–3.91)
G/T	150(37)	74(37)	NS	NS	
T/T	213(53)	113(57)	NS	NS	
rs1801133 in MTHFR					
C	369(91)	174(87)	NS	NS	
T	206(51)	106(54)	NS	NS	
C/C	199(49)	92(46)	NS	NS	
C/T	170(42)	82(41)	NS	NS	
T/T	36(9)	24(13)	NS	NS	
rs2070744 in eNOS					
C	253(62)	100(50)	0.0034	0.024	1.63 (1.16–2.30)
T	343(85)	174(88)	NS	NS	
C/C	62(15)	24(12)	NS	NS	
C/T	191(47)	76(38)	NS	NS	
T/T	152(38)	98(50)	0.0034	0.024	0.61 (0.43–0.86)
rs1333049 in the 9p21 region					
C	313(78)	155(78)	NS	NS	
G	305(75)	146(74)	NS	NS	
C/C	100(25)	52(26)	NS	NS	
C/G	213(53)	103(52)	NS	NS	
G/G	92(22)	43(22)	NS	NS	

NS – not significant.

*The Bonferroni correction for the number of tests (multiple comparisons) was applied to the p values.** $p < 0.05$.

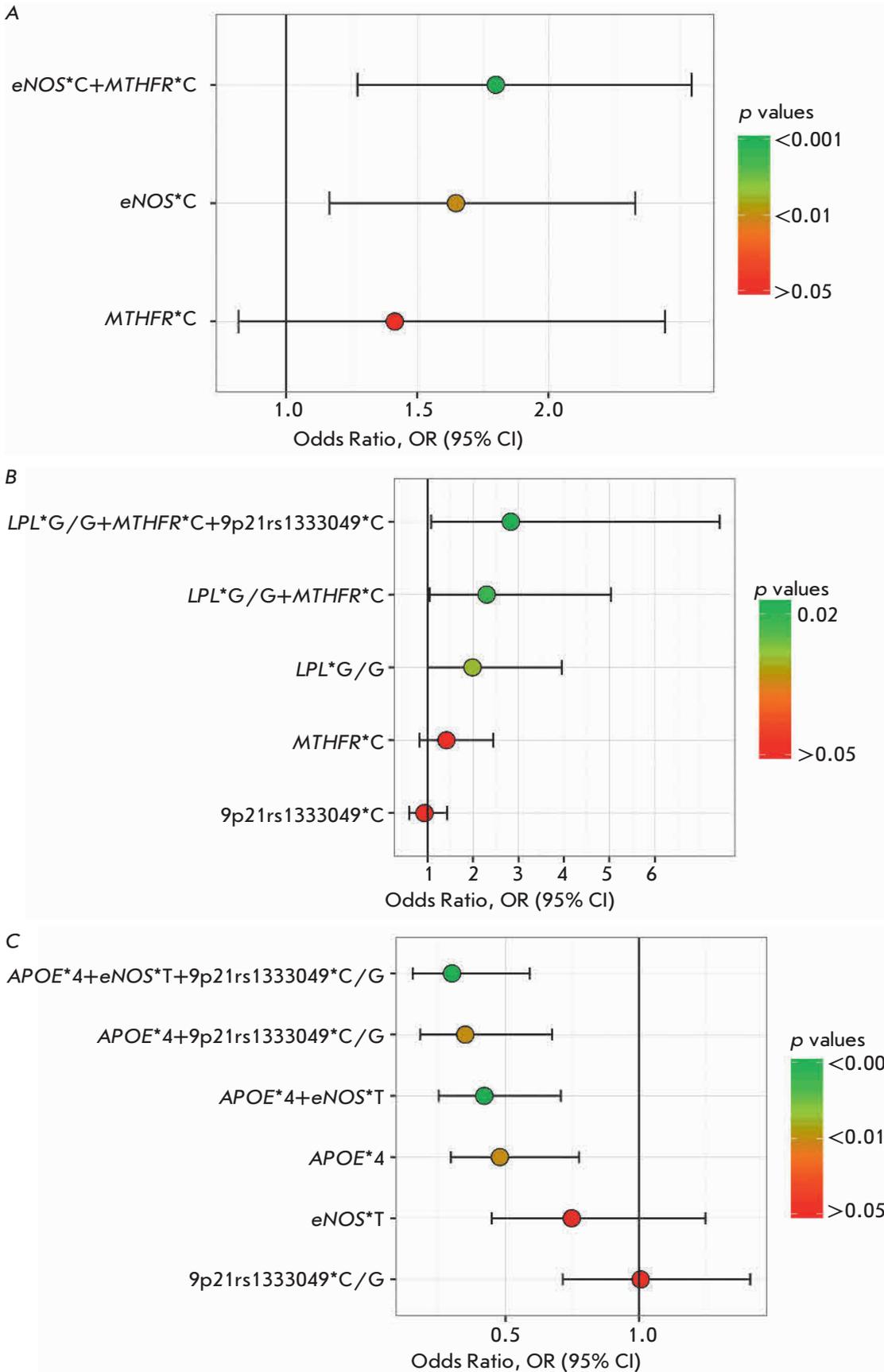


Fig.2. Multilocus analysis made it possible to identify the association between MI and rs1801133 in the *MTHFR* gene and rs1333049 in the 9p21 region, which one by one are not significantly associated with MI. The odds ratios (ORs), the confidence intervals (CIs) and the significance levels (qualitatively, by color of each circle, which corresponds to the OR value) are graphically presented for MI-associated combinations, which include the variants of the *MTHFR* gene and/or rs1333049, and for the components of these combinations.

A. Biallelic combination (*MTHFR**C + eNOS*C) that is positively associated with MI, and its components.

B. Triallelic combination (*LPL**G/G + *MTHFR**C + rs1333049*C) that is positively associated with MI, and its components.

C. Triallelic combination (*APOE**ε4+eNOS*T +rs1333049*C/G) that is negatively associated with MI, and its components.

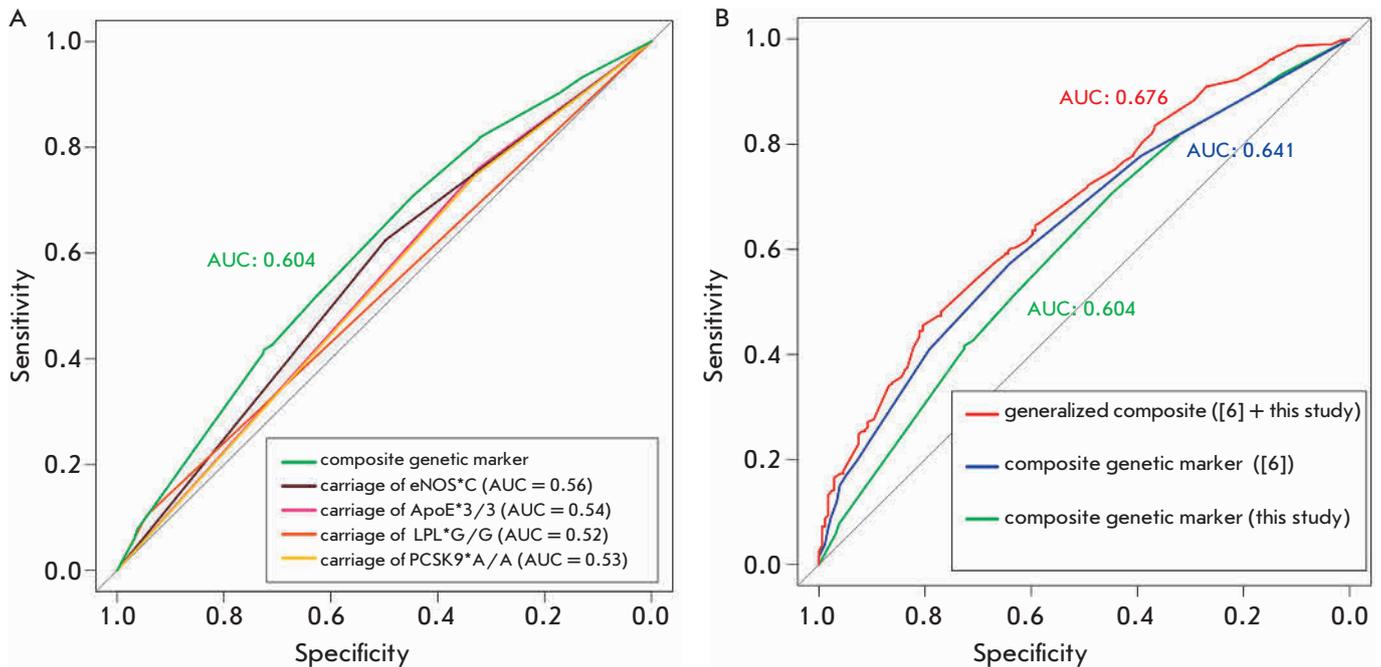


Fig. 3. ROC analysis for efficiency of the models that are based on different genetic markers of individual risk of MI. A. The efficiency of classification of individuals using the models based on the carriage of individual genetic markers (variants of the *PCSK9*, *APOE*, *LPL*, and *eNOS* genes) and the model that takes into account the carriage of variants of all four genes (composite genetic marker, green line). B. Prognostic efficacy of the generalized composite model of individual genetic risk of MI (red line) obtained by supplementing the previously described model [6] (blue line) with data on the carriage of one or more variants of the *PCSK9*, *LPL*, *eNOS* and *APOE* genes (green line). The AUC (area under the curves) values for different models are shown in the same color as the corresponding curve.

according to the efficiency of classifying the subjects into MI patients and healthy individuals. One can see that the genetic factors considered one by one are poor classifiers of the risk of MI (AUC < 0.60). However, satisfactory prognostic efficacy is achieved (AUC = 0.604) when taking into account the data on carriage of a combination of the *PCSK9*, *APOE*, *LPL*, and *eNOS* alleles/genotypes. We would like to mention that the model does not become more efficient if the *MTHFR* and the 9p21 region alleles, which are the components of the combinations identified by APsampler, are added one by one.

These findings were used to improve the earlier built composite genetic model of the risk of MI, which includes the *TGFB1*, *FGB*, and *CRP* genetic variants and the epistatic combination of the *IFNG* and *PTGS1* genes as predictors [6]. Figure 3B shows three ROC curves: the ROC curve obtained in the analyzed sample for the composite model described in [6]; the ROC curve for the carriage of a combination of the newly identified markers shown in Fig. 3A; and the ROC curve for the generalized composite model that includes the

markers from both the previous and the present studies. One can see that the prognostic efficacy has significantly increased ($p = 0.014$): from AUC = 0.641 in the model without new markers to 0.676 in the generalized composite model.

DISCUSSION

The case-control multilocus analysis of the association of the polymorphic variants of *PCSK9* (rs562556), *APOE* (epsilon polymorphism, rs7412 and rs429358), *LPL* (rs320), *MTHFR* (rs1801133), *eNOS* (rs2070744), and the 9p21 region (rs1333049) with the risk of MI revealed the *PCSK9*, *APOE*, *LPL*, and *eNOS* alleles/genotypes significantly associated with MI and the bi- and triallelic combinations that carried the *MTHFR* and the 9p21 region alleles/genotypes along with the variants of the aforelisted genes.

Each of the examined lipid metabolism genes (*PCSK9*, *LPL*, and *APOE*) turns out to be associated with the risk of MI; the OR of the risk genotypes ranges from 1.45 to 1.96 but the significance level is rather low ($p = 0.013$ – 0.032). The involvement of the lipid metabo-

lism genes in the development of MI shows good agreement with the well-known fact that disorders of lipid metabolism, high cholesterol level and elevated atherogenic index lead to formation of atheromatous plaques in the arterial tunica intima. However, the published data regarding the involvement of the examined variants of the lipid metabolism genes in development of CVDs are rather controversial.

SNP rs562556 in the *PCSK9* gene is responsible for the Ile-474-to-Val substitution in the encoded protein, which apparently does not affect its expression level [26]. We observed the MI association with the A/A genotype in this SNP in Russians. However, the distributions of rs562556 variants in Japanese subjects with MI were not different from those in the control population, although this SNP was associated with cholesterol level [27]. An association between polymorphism rs562556, the presence of anti-phospholipid antibodies, and development of thrombosis (the risk factor of MI) was revealed in subjects carrying these antibodies [28]. The association of other *PCSK9* gene variants (namely, rs11206510 [29] and rs11591147 [30]) with MI was demonstrated in different populations. Hence, our findings regarding the association of the *PCSK9* gene with MI are consistent with the data published earlier.

Individual allelic variants of the *APOE* gene epsilon polymorphism are associated with CVDs (and with MI in particular) in almost all populations. The meta-analyses demonstrate that the $\epsilon 4$ allele is associated with the risk of MI, while the $\epsilon 2$ allele has a protective effect [31, 32]. However, the conclusions drawn in the meta-analysis involving different ethnicities cannot be automatically extrapolated to separate populations, where the roles of individual alleles in predisposition to MI vary significantly. The considerable difference in allele frequencies in different populations and even within the same population residing in different regions is potentially the key reason for the poor replicability of the results obtained in individual studies [33]. In particular, $\epsilon 3/\epsilon 3$ turned out to be the risk genotype in our study involving Russians living in Central Russia, while $\epsilon 2/\epsilon 3$ was the risk genotype among Siberian males [34].

The HindIII polymorphism (rs320) is responsible for the T-to-G substitution in intron 8 of the *LPL* gene. This polymorphism is believed to reside in the regulatory sequence and to regulate *LPL* expression [35]. We revealed an association between rs320 and MI in our sample. The association between this polymorphism and the development of MI was demonstrated in a number of previous publications [36, 37], including the studies involving Russian populations [38]. Other polymorphisms of this gene associated with MI have been reported for the Japanese population [39]. However,

the data on association of individual alleles with MI are sometimes inconsistent.

Polymorphism rs2070744 in the *eNOS* gene is another locus whose variant showed significant association with MI in our study. This polymorphism resides within the promoter region. The C allele associated with the risk of MI in our study is related to downregulation of mRNA expression and, correspondingly, the eNOS protein level [40]. Our data are consistent with the findings reported in other publications [41].

The C-to-T substitution in the rs1801133 variant of the *MTHFR* gene caused Ala-222-to-Val substitution in the protein [42] and reduction of methylenetetrahydrofolate reductase activity by almost 50% [43]. No association between rs1801133 and the risk of MI was found in different ethnic groups, including Caucasians [44] and Russians [45], in most studies. We also revealed no association of SNP rs1801133 with the risk of MI in Russians; however, carriage of the C allele in combination with the C allele of the *eNOS* gene (Fig. 2A) or carriage of the G/G genotype of the *LPL* gene (Fig. 2B) showed significant association with the risk of MI. We believe that these data can be interpreted as an argument in favor of the involvement of the *MTHFR* gene in predisposition to MI.

The 9p21 region is the only genomic region whose association with the risk of MI has been replicated in several GWASs at a genome-wide significance level ($p < 5 \times 10^{-8}$) [2]. These findings have been confirmed in a number of validation studies, including those for rs10757278 and rs1333049 in a sample consisting of MI patients and controls from the Siberian population (of unspecified ethnicity) [46]. In our study, no significant MI association with rs1333049 was found in Russians, but the multilocus analysis revealed a number of combinations containing this SNP. Carriage of the rs1333049*C allele within the triallelic combination is associated with the risk of MI (Fig. 2B), while the rs1333049*C/G genotype within the bi- and triallelic combinations was found to have a protective effect (Fig. 2C) and showed good agreement with the results reported in [46].

The statistical analysis of three-way interactions revealed no epistatic interactions between the components of all the identified combinations. Meanwhile, all the biallelic combinations associated with the risk of MI (OR > 1) are characterized by higher significance levels and higher OR values compared to those of alleles/genotypes within these combinations considered one by one (correspondingly, lower OR values for the protective combinations with OR < 1). A similar regularity is observed for the triallelic combinations as compared to the biallelic ones. Therefore, the cumulative effects observed in this study result from the additivity of the

contributions from individual genes. The reason for this additivity is that statistical significance in a relatively small sample for the association between a combination of unidirectional weak genetic factors and the disease is higher compared to the association observed for each factor one by one. Therefore, there is every reason to believe that the *MTHFR* (rs1801133) and 9p21 (rs1333049) loci are independent risk factors of MI having weak effects. Statistical powder was insufficient to reveal significant associations of these factors with MI in the examined sample, while multilocus analysis compensated for this drawback.

Identically, since the effects of the *MTHFR* (rs1801133) and 9p21 (rs1333049) loci are weak, adding them to the composite genetic model of the risk of MI did not increase the prognostic efficacy of the model. It is worth mentioning that the epistatic combinations are better risk classifiers than the additive combinations. Indeed, the epistatic combination of the *IFNG* and *PTGS1* genes was found to be one of the MI risk factors, while both components were not associated with the disease one by one [6]. Meanwhile, identifying the genetic variants within additive combinations may indicate the possibility of identifying the association of these genetic variants one by one with the disease in larger samples.

Although being statistically significant, the prognostic efficacy of the composite genetic model of the risk of MI built using the findings obtained in our study is rather low. This is also true for the earlier obtained model [6]. The AUC value of 0.676 was achieved by combining the two models; at the cut-off equal to 0.74 it corresponds to sensitivity of 0.80 and specificity of 0.45. Overall, neither the results of GWASs nor the findings

obtained using the candidate gene approach currently make it possible to effectively predict development of MI using genetic analysis.

CONCLUSIONS

The analysis of the association between the polymorphic regions of six candidate genes and MI showed that they are significantly associated with MI, either one by one or within combinations. We have replicated the association of the polymorphic variants of the *PCSK9*, *APOE*, *LPL*, *MTHFR*, and *eNOS* genes and the 9p21 region with MI in independent samples of Russians living in Central Russia. Since the variants of the same genes (rs1801133 in the *MTHFR* gene or rs1333049 in the 9p21 region), which are not significant one by one, were components of several different combinations (with no epistatic interactions between their components revealed), it is fair to conclude that the cumulative effect of the genes within a combination identified using multilocus analysis results from summation of their small independent contributions.

Inclusion of the identified markers to the previously reported model of individual genetic risk [6] significantly increases its prognostic efficacy, although our findings need to be replicated for an independent sample of Russians. In order to further increase the predictive power of the composite model, it should be improved by including other genetic predictors of risk and refining the regression coefficients for larger samples. ●

This work was supported by the Russian Science Foundation (grant No. 16-14-10251).

REFERENCES

- Dai X., Wiernek S., Evans J.P., Runge M.S. // *World J. Cardiol.* 2016. V. 8. № 1. P. 1–23.
- <http://www.ebi.ac.uk/gwas>
- Dehghan A., Bis J.C., White C.C., Smith A.V., Morrison A.C., Cupples L.A., Trompet S., Chasman D.I., Lumley T., Völker U., et al. // *PLoS One.* 2016. V. 11. № 3. e0144997.
- <http://www.thl.fi/monica/>
- <http://www.ucl.ac.uk/easteurope/hapiee.html>
- Barsova R.M., Lvovs D., Titov B.V., Matveeva N.A., Shakhnovich R.M., Sukhinina T.S., Kukava N.G., Ruda M.Y., Karamova I.M., Nasibullin T.R., et al. // *PLoS One.* 2015. V. 10. № 12. e0144190.
- Catapano A.L., Papadopoulos N. // *Atherosclerosis.* 2013. V. 228. № 1. P. 18–28.
- Mahley R.W. // *J. Mol. Med. (Berlin).* 2016. V. 94. № 7. P. 9–46.
- Li Y., He P.P., Zhang D.W., Zheng X.L., Cayabyab F.S., Yin W.D., Tang C.K. // *Atherosclerosis.* 2014. V. 237. № 2. P. 597–608.
- Ueland P.M., Refsum H. // *J. Lab. Clin. Med.* 1989. V. 114. № 5. P. 473–501.
- Liu D., Jiang Z., Dai L., Zhang X., Yan C., Han Y. // *Gene.* 2014. V. 545. № 1. P. 175–183.
- Holdt L.M., Teupser D. // *Arterioscler. Thromb. Vasc. Biol.* 2012. V. 32. № 2. P. 196–206.
- Lvovs D., Favorova O.O., Favov A.V. // *Acta Naturae.* 2012. V. 4. № 3. P. 62–75.
- Third Universal definition of myocardial infarction // *Eur. Heart J.* 2012. V. 33. P. 2551–2567.
- Hixson J.E., Vernier D.T. // *J. Lipid. Res.* 1990. V. 31. P. 545–548.
- Shimo-Nakanishi Y., Urabe T., Hattori N., Watanabe Y., Nagao T., Yokochi M., Hamamoto M., Mizuno Y. // *Stroke.* 2001. V. 32. P. 1481–1486.
- Genest J., Rozen R. // *Arterioscler. Thromb. Vasc. Biol.* 1997. V. 17. № 3. P. 569–573.
- Augeri A.L., Tsongalis G.J., van Heest J.L., Maresh C.M., Thompson P.D., Pescatello L.S. // *Atherosclerosis.* 2009. V. 204. № 2. e28–34.
- <http://www.broad.mit.edu/mpg/haploview>
- <https://sourceforge.net/projects/apsampler/>
- White D.R., Pesner R., Reitz K.P. // *Behavior Sci. Res.* 1983. V. 18. P. 103–122.

- 22 Cortina-Borja M., Smith A.D., Combarros O., Lehmann D.J. // *BMC Res. Notes*. 2009. V. 2. P. 105–111.
- 23 DeLong E.R., DeLong D.M., Clarke-Pearson D.L. // *Biometrics*. 1988. V. 44. P. 837–845.
- 24 Youden W.J. // *Cancer*. 1950. V. 3. № 1. P. 32–35.
- 25 www.ncbi.nlm.nih.gov/snp/
- 26 Astrakova (Benimetskaya) K.S., Shakhtshneider E.V., Ivanoshchuk D.E., Orlov P.S., Ragino Y.I., Voevoda M.I. // *Atherosclerosis*. 2016. V. 12. № 2. P. 18–24.
- 27 Shioji K., Mannami T., Kokubo Y., Inamoto N., Takagi S., Goto Y., Nonogi H., Iwai N. // *J. Hum. Genet.* 2004. V. 49. № 2. P. 109–114.
- 28 Ochoa E., Iriondo M., Manzano C., Fullaondo A., Villar I., Ruiz-Irastorza G., Zubiaga A.M., Estonba A. // *PLoS One*. 2016. V. 11. № 1. e0146990.
- 29 Myocardial Infarction Genetics Consortium, Kathiresan S., Voight B.F., Purcell S., Musunuru K., Ardissino D., et al. // *Nat. Genet.* 2009. V. 41. № 3. P. 334–341.
- 30 Kathiresan S., Myocardial Infarction Genetics Consortium // *N. Engl. J. Med.* 2008. V. 358. № 21. P. 2299–2300.
- 31 Wang Y.L., Sun L.M., Zhang L., Xu H.T., Dong Z., Wang L.Q., Wang M.L. // *FEBS Open Bio*. 2015. V. 5. P. 852–858.
- 32 Xu H., Li H., Liu J., Zhu D., Wang Z., Chen A., Zhao Q. // *PLoS One*. 2014. V. 9. № 8. e104608.
- 33 Borinskaia S.A., Kal'ina N.R., Sanina E.D., Kozhebaeva Zh.M., Gupalo E.I., Garmash I.V., Ogurtsov P.P., Parshukova O.N., Boiko S.G., Veselovskii E.M., et al. // *Russ J Genet.* 2007. V. 43. № 10. P. 1201–1207.
- 34 Voevoda M.I., Shakhtshneider E.V., Maksimov V. N., Kulikov I.V., Romaschenko A.G. // *Atherosclerosis*. 2008. V. 4. № 1. P. 11–26.
- 35 Chen Q., Razzaghi H., Demirci F.Y., Kamboh M.I. // *Atherosclerosis*. 2008. V. 200. № 1. P. 102–108.
- 36 Tanguturi P.R., Pullareddy B., Rama Krishna B.S., Murthy D.K. // *Indian Heart J.* 2013. V. 65. № 6. P. 653–657.
- 37 Gigeck Cde O., Chen E.S., Cendoroglo M.S., Ramos L.R., Araujo L.M., Payão S.L., Smith Mde A. // *Clin. Chem. Lab. Med.* 2007. V. 45. № 5. P. 599–604.
- 38 Malygina N.A., Melent'ev A.S., Kostomarova I.V., Melent'ev I.A., Saégitov R.T., Smirnova Iu.B., Serova L.D. // *Mol Biol (Mosk)*. 2001. V. 35. № 5. P. 787–791.
- 39 Matsuoka R., Abe S., Tokoro F., Arai M., Noda T., Watanabe S., Horibe H., Fujimaki T., Oguri M., Kato K., et al. // *Int. J. Mol. Med.* 2015. V. 35. № 5. P. 1451–1459.
- 40 Doshi A.A., Ziolo M.T., Wang H., Burke E., Lesinski A., Binkley P. // *J. Card. Fail.* 2010. V. 16. № 4. P. 314–319.
- 41 Kong X.Z., Zhang Z.Y., Wei L.H., Li R., Yu J. // *Med. Sci. Monit.* 2017. V. 11. № 23. P. 759–766.
- 42 Dayakar S., Goud K.I., Reddy T.P., Rao S.P., Sesikeran S.B., Sadhnani M. // *Genet. Test. Mol. Biomarkers*. 2011. V. 15. № 11. P. 765–769.
- 43 Xuan C., Bai X.Y., Gao G., Yang Q., He G.W. // *Arch. Med. Res.* 2011. V. 42. № 8. P. 677–685.
- 44 Alizadeh S., Djafarian K., Moradi S., Shab-Bidar S. // *Int. J. Cardiol.* 2016. V. 217. P. 99–108.
- 45 Nazarenko G.I., Skvortsova V.I., Kleimenova E.B., Konstantinova M.V. // *Zh Nevrol Psikhiatr Im S S Korsakova*. 2009. V. 109. № 10. Suppl 2. P. 19–25.
- 46 Maximov V.N., Kulikov I.V., Orlov P.S., Gafarov V.V., Malyutina S.K., Romashchenko A.G., Voevoda M.I. // *Vestn. Ross Akad Med Nauk*. 2012. V. 67. № 5. P. 24–29.

Recombinant Antibodies to the Ebola Virus Glycoprotein

A. A. Panina^{1*}, I. G. Dementieva², T. K. Aliev³, V. A. Toporova¹, D. S. Balabashin^{1,4},
M. N. Bokov², L. P. Pozdnyakova², O. B. Shemchukova², D. A. Dolgikh^{1,4}, P. G. Sveshnikov²,
M. P. Kirpichnikov^{1,4}

¹Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Mikluho-Maclay Str. 16/10, Moscow, 117997, Russia

²Russian Scientific Center for Molecular Diagnosis and Treatment, Simferopol Blvd. 8, Moscow, 117149, Russia

³Lomonosov Moscow State University, Department of Chemistry, Leninskie gory 1, bldg. 3, Moscow, 119991, Russia

⁴Lomonosov Moscow State University, Faculty of Biology, Leninskie gory 1, bldg. 12, Moscow, 119991, Russia

*E-mail: paniann07@yandex.ru, panina@mx.ibch.ru

Received: April 27, 2017; in final form October 24, 2017

Copyright © 2017 Park-media, Ltd. 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 Currently, there are no approved therapies for the targeted prevention and treatment of Ebola hemorrhagic fever. In the present work, we describe the development of a eukaryotic expression system for the production of three full-length chimeric antibodies (IgG1-kappa isotypes) GPE118, GPE325, and GPE534 to the recombinant glycoprotein of the Ebola virus (EBOV GP), which is a key factor in the pathogenicity of the disease. The immunochemical properties of the obtained antibodies were studied by immunoblotting and indirect, direct, and competitive ELISA using the recombinant EBOV proteins rGPdTM, NP, and VP40. The authenticity of the antibodies and the absence of cross-specificity with respect to the structural proteins NP and VP40 of the Ebola virus were proved. The epitope specificity of the resulting recombinant antibodies was studied using commercial neutralizing antibodies against the viral glycoprotein. The recombinant antibodies GPE118, GPE325, and GPE534 were shown to recognize glycoprotein epitopes that coincide or overlap with the epitopes of three well-studied neutralizing anti-Ebola virus antibodies.

KEYWORDS Ebola hemorrhagic fever virus, therapeutic recombinant chimeric antibodies.

ABBREVIATIONS mAbs – monoclonal antibodies; EBOV – Ebola hemorrhagic fever virus; EBOV GP – glycoprotein of the Ebola hemorrhagic fever virus; rGPdTM – recombinant glycoprotein of the Ebola hemorrhagic fever virus lacking transmembrane domain; NP – Ebola virus nucleoprotein; VP40 – structural protein of the Ebola virus; SOE-PCR – splicing by overlap extension PCR; IEDB – the Immune Epitope Database; Kd – the dissociation constant; PBS – phosphate buffered saline.

INTRODUCTION

The number of people that got infected during the most recent Ebola hemorrhagic fever outbreak exceeded 28,000, and more than 11,000 deaths were reported. Although the epidemic was stopped, the World Health Organization deems a new outbreak possible. In this regard, there is the urgent task of developing effective agents for the prevention and therapy of the disease.

Several approaches to the therapy of Ebola hemorrhagic fever have been proposed over the past decade. Thus, intravenous injection of inhibitors of blood coagulation, such as recombinant human-activated protein C, increases the survival rate in patients by 18% [1]. Intravenously administered small interfering RNAs increase the survival rate from 66 to 100% depending on the number of injections made [2]. However, the thera-

py needs to be started promptly, within the first 30–60 min post-infection.

Therapeutic antibodies against the Ebola virus have been under development since the 1990s. Among the eight viral proteins, the major pathogenicity factor, glycoprotein (GP), is considered to be the key immunotherapy target. Unstable results have been obtained when using certain monoclonal antibodies (mAT) or antibodies collected from the blood of Ebola survivors [3]. Meanwhile, G.G. Olinger Jr. et al. demonstrated in primate experiments that, unlike nonspecific antiviral therapeutics, passive immunization with antibodies injected 24 h post-infection has a therapeutic effect [4]. There are ongoing attempts to find sources of neutralizing antibodies in the blood of Ebola survivors to develop agents that can be administered as monotherapy

[5]. However, it was revealed during the production and study of anti-GP mAbs that the ZMapp antibody cocktail, a combination of antibodies specific to different GP epitopes, exhibits the strongest protective effect and significantly reduces lethality in model animals [6]. Different combinations of three GP-specific chimeric antibodies were tested in guinea pigs and rhesus macaques. The animals (18 rhesus macaques) were infected with lethal doses of the virus and were treated with ZMapp (50 mg/kg body weight) on days 3, 4, and 5 post-infection. ZMapp was found to ensure survival of all 18 rhesus macaques, including those who had strongly marked signs of the disease. The antibodies within the ZMapp cocktail were expressed in tobacco leaves and used during the 2014 Ebola epidemic in West Africa [7]. Compared to the group of patients receiving palliative care only, the mortality rate in the group receiving additional therapy with the study drug decreased from 37% (13 out of 35) to 22% (8 out of 36). Administration of the study drug in a larger patient sample was limited, because the trial was conducted at the final stage of the epidemic, when the number of newly infected patients was rather small, making it difficult to recruit Ebola virus carriers.

We previously obtained murine mAbs against the recombinant glycoprotein of the Ebola hemorrhagic fever virus (Zaire strain) lacking the transmembrane domain (rGPdTM) and selected three mAbs exhibiting different epitope specificities: GPE118, GPE325, and GPE534. The nucleotide and amino acid sequences of the variable domains were identified. The framework and hypervariable regions of immunoglobulin heavy and light chains were identified [8, 9].

The aim of this study was to design full-length recombinant chimeric antibodies against EBOV GP based on murine mAbs and to investigate their immunochemical properties: to determine the authenticity, specificity, and immunoreactivity of full-length chimeric antibodies and to measure the dissociation constants and the epitope specificity.

EXPERIMENTAL

We used: the recombinant EBOV protein rGPdTM (IBT Bioservices, USA); recombinant proteins NP and VP40 (Fitzgerald Industries International, USA); mAb against human Ig kappa chain 4G7 and mAb against human IgG gamma-1 chain region (Hytest, Turku, Finland); anti-EBOV GP mAbs h13F6, c13C6FR1, c6D8, KZ52 and 4F3 (IBT Bioservices, USA); ExtrAvidin-Peroxidase conjugate (E2886, Sigma-Aldrich, USA); TMB substrate (3,3',5,5'-tetramethylbenzidine, BioTest Systems, Russia); biotinyl-N-hydroxy-succinimide (H1759, Sigma-Aldrich); nitrocellulose membrane (Membrane filters, cellulose nitrate, pore size 0.45 μm ,

S045A330R, Advantec MFS, Inc., USA); Vivaflow 200 membrane (Sartorius Stedim Biotech, Germany); 96-well plates with high binding capacity (Corning-Costar, the Netherlands); and Tween 20.

Construction of the expression system for chimeric mAbs in mammalian cells

The chimeric sequences of light (L) or heavy (H) chains of mAbs were constructed by successively attaching the constant domain of the human kappa light chain (for the L chain) or the constant domains C_{H1} - C_{H3} of the human IgG1 heavy chain (for the H chain) to DNA encoding the variable domain of the antibody. DNA fragments encoding the 5'-untranslated region and carrying a Kozak sequence, sequences of native leader peptides ensuring immunoglobulin secretion in the culture medium, and the 3'-untranslated DNA fragment carrying a polyadenylation site were also attached to the sequences listed above.

DNA fragments carrying the sequence coding for the variable domain of the heavy chain were PCR-synthesized on the template of plasmids produced earlier [8] and encoding the variable domain of the heavy chain of the antibodies GPE118, GPE534, and GPE325. SacI and ApaI restrictase recognition sites were simultaneously inserted at the 3'-end. For light chains, the mAb fragments forming a single transcription unit were joined using splicing by overlap extension PCR (SOE-PCR); the 5'-untranslated region and the leader peptide of heavy chains were attached using the same procedure. The constant domains C_{H1} - C_{H3} of the human IgG1 heavy chain and the 3'-untranslated region carrying the polyadenylation site were ligated at the Apa I site that is present at the 5'-end of the human C_{H1} domain and was inserted to the 3'-end of the variable domain by PCR. The dual-promoter (the hEF1-HTLV promoter for the heavy chain and the CMV promoter for the antibody light chain) expression vector was constructed according to the procedure described in [10].

Generation of cell lines producing recombinant antibodies

Full-length recombinant antibodies were produced in CHO DG44 cells. The cells were cultured in a CD DG44 medium (Invitrogen) supplemented with 8 mM *L*-glutamine and 0.18% Pluronic F-68 (Invitrogen). A cell suspension (30 ml, 3×10^5 cells/ml) was placed in 125 ml Erlenmeyer flasks and cultured under constant stirring on an orbital shaker at a rate of 130 rpm in a Sanyo MCO-18AIC CO₂ incubator (Sanyo, Japan) at 37°C, under 8% CO₂ and maximum humidity. Transfection was performed 24 h after the initiation of cultivation using Freestyle MAX reagent (Invitrogen). 15 μl of Freestyle

MAX and 18 µg of the plasmid carrying the light and heavy chains of the antibodies under the control of the CMV and EF-1 alpha promoters, respectively, and the DHFR selection marker were used for transfection. Cell culture selection was carried out 48 h post-transfection. The cell culture was seeded into a nucleotide-free CD OptiCHO medium (Invitrogen). The cells were seeded into 125 Erlenmeyer flasks with 30 ml of the OptiCHO medium supplemented with 8 mM *L*-glutamine and cultured under the conditions described above. The first stage of selection in this medium was considered completed when culture viability of at least 95% was attained, while the cell population doubling time was 24 h. At the next stage of the selection of cell lines, 10 nM methotrexate (MTX) was added to the same culture medium. After this selection stage, experimental samples of recombinant antibodies started to be generated.

After selection using MTX, the producer cell lines were cloned on a ClonePIX FL cell sorter (Genetix, UK) to increase specific production of recombinant antibodies and stabilize the producers. The cell suspension was seeded onto a CloneMedia-CHO complete semi-solid medium supplemented with 8 mM *L*-glutamine and 10 mg/ml FITC-labeled human IgG secondary antibodies (Genetix, UK). Clones were selected according to the recommendations provided by the cell sorter manufacturer and cultured in a XP Media medium (Genetix, UK) supplemented with 8 mM *L*-glutamine. After the cell count had increased in each clone, the cells were transferred to the CD OptiCHO medium (Invitrogen, USA) for further growth, cryoconservation, and analysis of expression and growth properties.

Generation of recombinant antibodies

The cultures were seeded into 125 Erlenmeyer flasks containing 30 ml of the OptiCHO medium supplemented with *L*-glutamine to a concentration of 8 mM. Cell culture was started at a concentration of 3×10^5 cells/ml under the conditions described above. Cultivation was terminated when cell culture viability decreased to 50%. The supernatant (0.5 l) with the verified presence of the antibody, which was obtained after cultivation of the CHO cell line, was centrifuged at 4,000 rpm for 30 min and passed through a filter with a 0.45 µm pore size. The supernatant was concentrated by tangential ultrafiltration on a Vivaflow setup (Sartorius, Germany) with a molecular weight cut-off < 50 kDa. The antibodies were isolated from the concentrated supernatant by affine chromatography on columns packed with protein A-agarose (GE Healthcare, USA) according to the manufacturer's recommendations.

The content of recombinant antibodies in eluates after affine chromatography was evaluated by indirect ELISA.

Immunoblotting of recombinant antibodies with mAbs 4G7 and 2C11

Electrophoretic separation of recombinant antibodies (6 µg per lane) was carried out in 12% PAAG under reducing and nonreducing conditions. The proteins were then subjected to electrophoretic transfer (electroblotting) from the gel onto a nitrocellulose membrane with a 0.45 µm pore size (Advantec MFS, Inc., USA). The membrane was blocked with a solution of 5% casein in PBS overnight at 4°C and washed three times with PBS-T (10 mM K₂HPO₄, pH 7.5, 0.145 M NaCl, 0.05% Tween 20). The membrane was then cut into strips and incubated in antibody solutions (mAb 4G7 or mAb 2C11, 10 µg/ml, PBS) at 37°C for 1 h. After repeated five-fold washing with a PBS-T buffer, incubation in the presence of HRP-conjugated goat anti-mouse IgG secondary antibodies at a 1 : 15,000 dilution at 37°C was carried out for 1 h. After repeated three-fold washing in PBS-T, the substrate (3,3-diaminobenzidine, 4-chloro-1-naphthol, H₂O₂) was added and the system was incubated for 4–10 min. The reaction was stopped by washing the strips with water.

Immunoblotting of the recombinant EBOV rGPdTM with recombinant antibodies

After electrophoretic separation of EBOV rGPdTM (12 µg) in 12% PAAG under reducing conditions, electroblotting of the proteins from the gel to the nitrocellulose membrane and blocking as described above, the transferred proteins were detected on the nitrocellulose membrane by indirect ELISA (immunoblotting). For this purpose, the membrane was cut into strips, which were placed in a 10 µg/ml solution of the recombinant antibodies GPE118, GPE325, and GPE534 and incubated at 37°C for 1 h. After the incubation, the strips were incubated with HRP-conjugated mAb 4G7 at a dilution of 1 : 25,000 at 37°C for 1 h. The final wash and the reaction development were carried out as described above.

Indirect ELISA with the immobilized recombinant Ebola virus proteins rGPdTM, NP, and VP40

The antigen (1 µg/ml, PBS) was sorbed in the wells of a 96-well plate with high binding capacity at 4°C overnight. The plates were washed five times with PBS and 0.05% Tween 20. A solution of the recombinant antibodies under study (PBS, 2% BSA) was then added by threefold serial dilutions starting from a concentration of 3 µg/ml, incubated at 37°C for 1 h, washed five times with PBS and 0.05% Tween 20, and then incubated with HRP-conjugated mAb 4G7 against the human Ig kappa chain at a 1 : 50,000 working dilution (PBS, 2% BSA) at 37°C for 1 h. To develop the reaction, the wells were washed five times with PBS and 0.05% Tween 20;

100 μ l of the TMB substrate was added to each well; and the plate was incubated at room temperature on a shaker for 15 min. The reaction was stopped with 0.5 M H_2SO_4 and absorbance was measured at a wavelength of 450 nm.

Determination of K_d for the complex of recombinant antibodies and EBOV GP

At the first stage, recombinant antibodies at a constant concentration of 7 pM (1 ng/mL) were incubated with the EBOV GP antigen in a concentration range of 0.1–10 nM (10–1000 ng/ml) at room temperature for 2 h under constant stirring to achieve thermodynamic equilibrium in the three-component system: free antigen, free antibody, and the antigen–antibody complex. At the second stage, the concentration of free recombinant antibodies was measured by solid-phase ELISA with EBOV GP immobilized on the plate. At the final stage, K_d was calculated from the Klotz equation using the values of the total antigen concentration and the concentration of free recombinant antibodies [11].

Competitive ELISA for epitope mapping of recombinant chimeric antibodies

Parental murine mAbs, control commercial mAbs, and the full-length chimeric antibodies under study were sorbed at a concentration of 5 μ g/ml at 4°C overnight in PBS, pH 7.2 on 96-well plates with high binding capacity (100 μ l per well). The plate was washed 5 times with PBS-T (200 μ l per well). The control mAbs KZ52, h13F6, c13C6FR1, c6D8, and 4F3 were titrated starting from a concentration of 6 μ g/ml by three-fold serial dilutions in PBS-T (0.01 M KH_2PO_4 , 0.1 M NaCl, 0.2% BSA, and 0.1% Tween 20), 50 μ l per well. The recombinant viral protein was biotinylated using a fivefold excess of (+)-biotin N-hydroxysuccinimide ester (H1759 Sigma). Biotinylated EBOV rGPdTM was added at a concentration of 800 ng/ml in PBS-T, 50 μ l per well. The plate was incubated at 37°C for 1 h with stirring. The plate was washed five times with PBS-T, 200 μ l per well. ExtrAvidine–peroxidase conjugate at a concentration of 500 ng/mL was added to each well at a volume of 100 μ l and incubated at 37°C for 1 h. The plate was washed five times with PBS-T, 200 μ l per well. TMB substrate (100 μ l) was added to each well. The plate was incubated at room temperature for 15 min on a shaker. The reaction was stopped by adding 100 μ l of 0.5 M H_2SO_4 per well. The absorbance was measured at a wavelength of 450 nm.

RESULTS AND DISCUSSION

Recombinant antibodies were produced using the suspension-adapted dihydrofolate reductase-deficient Chinese hamster ovary cells CHO DG44. The expres-

sion system was preliminarily constructed, and cell lines producing recombinant antibodies were generated for this purpose. A dual-promoter system based on the commercial vector pOptiVEC-TOPO containing the cytomegalovirus promoter (CMV) and hEF1-HTLV hybrid promoter for the translation of cloned genes, as well as the gene conferring resistance to geneticin, was used as an expression system in eukaryotic cell lines [10]. The vector also carries the gene coding for dihydrofolate reductase (DHFR), whose expression is regulated by the CMV promoter through the independent internal ribosome entry site (IRES). In case of coexpression of the target antibody gene, DHFR can be used as a selective marker to generate a stable cell line. DHFR expression makes it possible to neutralize the effect of the potentially cytotoxic methotrexate (MTX), thus maintaining cell proliferation. The *DHFR* gene is amplified at elevated MTX concentrations, which, in turn, increases the antibody gene copy number. Under selective pressure of methotrexate, genomic amplification of the genes of heavy and light chains occurs simultaneously. Proper dual-promoter constructs carrying nucleotide sequences coding for the kappa light chain and the IgG1 heavy chain were produced for each antibody.

The resulting CHO cell lines were used to generate recombinant antibodies by suspension cultivation in a serum-free medium. The quality of the experimental samples of anti-Ebola recombinant antibodies was controlled electrophoretically and chromatographically.

The authenticity of the antibodies was proved by immunoblotting using monoclonal antibodies specific to the characteristic regions of the heavy and light chains: murine mAbs 4G7 against human Ig kappa chain and mAbs 2C11 against human IgG gamma-1 chain. Immunoblotting was performed under nonreducing and reducing conditions. The results are shown in *Fig. 1*.

The immunoblotting data demonstrate that all the samples of recombinant chimeric antibodies under non-reducing conditions yield a major band with mobility corresponding to a molecular weight of ~150 kDa and become virtually identically stained with mAbs 4G7 and 2C11, which corresponds to the composition and the anticipated molecular weight of an IgG molecule. All antibody samples also give three or four additional minor bands that become stained with both mAbs and contain fragments of the heavy and light chains, being indicative of potential proteolytic degradation of antibodies, incomplete assembly (the absence of closing of disulfide bonds in the antibody hinge region), and presence of free light chains. Under reducing conditions, the antibody 4G7 stains only the kappa free light chain with electrophoretic mobility corresponding to ~25 kDa. Antibody 2C11 against the human IgG1 heavy chain produces no bands in antibody samples under re-

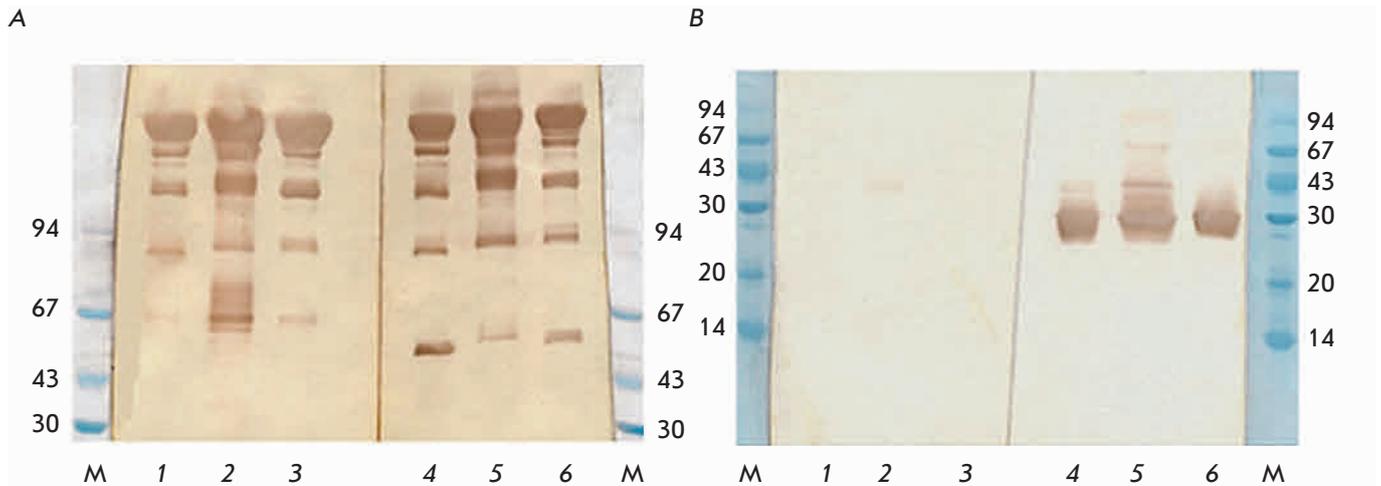


Fig. 1. Immunoblot of recombinant chimeric antibodies under nonreducing (A) and reducing (B) conditions after 7% SDS-PAGE with conjugates based on mAbs 2C11 (lanes 1–3) and mAb 4G7 (lanes 4–6). Lanes 1, 4 – GPE 118; lanes 2, 5 – GPE 325; lanes 3, 6 – GPE 534. Lane M – molecular weight standards, kDa.

ducing conditions, which demonstrates that binding of this antibody depends on closure of intrachain disulfide bonds in the antibodies under study. Mobility and the band pattern revealed by immunoblotting using 2C11 and 4G7 antibodies unambiguously prove the authenticity of the samples of the full-length recombinant anti-EBOV GP antibodies GPE118, GPE325, and GPE534.

The immunoreactivity of recombinant chimeric anti-EBOV GP IgGs was determined by indirect ELISA with the immobilized EBOV proteins rGPdTM, NP, and VP40 (Fig. 2).

The results of indirect ELISA of full-length chimeric antibodies with the immobilized structural EBOV proteins rGPdTM, NP, and VP40 demonstrate that all these antibodies are targeted against the viral glycoprotein only.

The immunoreactivity (determining the end point titer, EPT) of full-length chimeric antibodies was evaluated by indirect ELISA with sorption of EBOV rGPdTM onto the solid phase. The results of titration (Fig. 3) of full-length chimeric antibodies in indirect ELISA with immobilized EBOV rGPdTM demonstrate that the experimental samples have a high affinity for the target protein: the EPT values for xGPE118, xGPE325, and xGPE534 correspond to a concentration of 0.3 ng/ml (2 pM).

The specificity of three experimental samples of full-length recombinant chimeric anti-EBOV GP antibodies was analyzed by immunoblotting (Fig. 4) using recombinant EBOV rGPdTM, which is the extracellular part of the viral glycoprotein with a deleted transmembrane

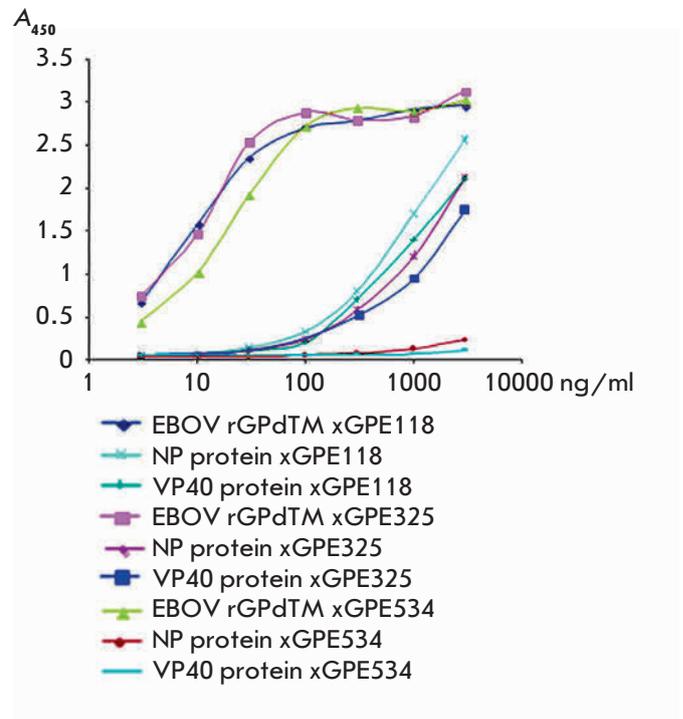


Fig. 2. Titration curves of full-length chimeric antibodies in indirect ELISA (OD 450 nm) with the immobilized EBOV proteins rGPdTM, NP, and VP40. xGPE118, xGPE325, and xGPE 534 – chimeric recombinant antibodies against rGPdTM; NP – nucleoprotein; VP40 – structural protein; and rGPdTM – the Ebola virus glycoprotein.

domain (this domain was expressed in insect cells and has a natural glycosylation pattern).

The immunoblot results demonstrate that the full-length recombinant antibodies GPE118, GPE325, and GPE534 give a single major band with a molecular weight of ~95 kDa under nonreducing conditions, which corresponds to full-length glycoprotein GP. Under re-

ducing conditions, these antibodies yield two bands corresponding to the GP1 and GP2 subunits, thus unambiguously proving the specificity of the resulting antibodies against EBOV rGPdTM.

The dissociation constants of the antigen-antibody complex were determined in Klotz coordinates using the method suggested by Friguet et al. [11]. Table 1 compares the dissociation constants of parental murine mAbs [9] and the full-length recombinant chimeric proteins produced.

Comparison of the affinities of the parental murine mAbs and recombinant chimeric antibodies shows that this parameter remains unchanged as one proceeds from natural full-length mAbs to full-length recombinant chimeric antibodies, although the constant antibody domains have been modified, thus being indicative of proper identification of the amino acid sequences of murine mAbs and the proper folding of recombinant proteins in the selected expression system. We would like to mention that affinity of the IgG1 antibody xGPE325 was even somewhat higher than that of parental murine IgM mAbs.

Full-length recombinant antibodies were used for epitope mapping by competitive ELISA using commercial antibodies with the known epitope specificity (Table 2). This analysis is needed to theoretically assess the potential protective activity. The first step in verifying proper selection of three and more anti-EBOV GP monoclonal antibodies is to demonstrate that mAbs are bound to or interact with three nonoverlapping epitopes of GP and that these epitopes are close to those of the known neutralizing antibodies.

When conducting solid-phase competitive ELISA using a monomeric antigen of biotinylated EBOV rGPdTM, the capture antibody under study was immobilized on a solid phase carrier; the biotinylated antigen

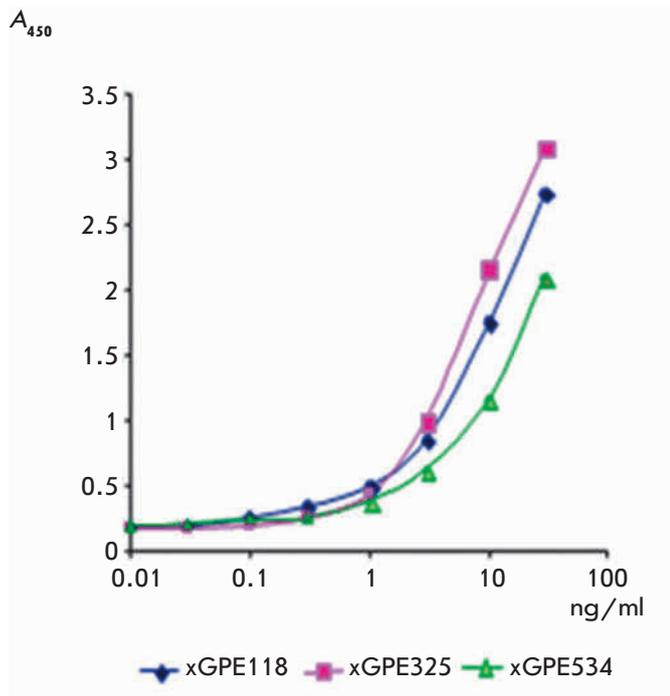


Fig. 3. Absorbance at 450 nm as a function of the concentration of full-length chimeric antibodies in indirect ELISA with immobilized EBOV rGPdTM. xGPE118, xGPE325, xGPE 534 – chimeric recombinant antibodies against rGPdTM

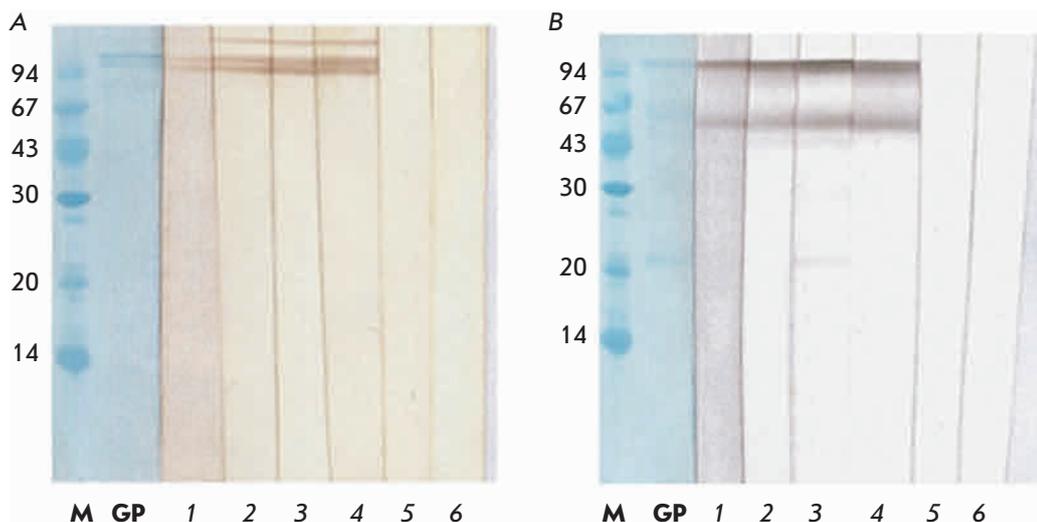


Fig. 4. Immunoblot of recombinant chimeric antibodies against EBOV GP after 12% SDS-PAGE under nonreducing (A) and reducing (B) conditions. After transfer to the membrane, each strip was incubated with a separate antibody. Lane 1 – GPE 118, lane 2 – GPE 325, lane 3 – GPE 534, lane 4 – mAT c6D8 (positive control), lane 5 – h13F6 (negative control), lane 6 – Fl6v3, lane M – molecular weight standards, kDa

and control mAb with the known epitope specificity were added simultaneously. If each mAb within a pair is targeted against different (nonoverlapping) recognition sites (epitopes), the three-component complex capture antibody–antigen–control mAb is formed. No three-component complex is formed on the solid phase if both antibodies are targeted against the same epitope.

Each commercial antibody was immobilized on a solid-phase carrier. Competitive ELISA was performed using the recombinant chimeric antibodies GPE118, GPE325, and GPE534, as well as commercial antibodies with a known epitope specificity [12]. The results of competitive ELISA for the produced full-length recombinant chimeric and commercial antibodies upon binding to biotinylated EBOV rGPdTM allow one to qualitatively characterize the epitopes of the antibodies under study. The antibody GPE534 competes with the neutralizing antibody KZ52; the antibodies GPE118, GPE325, and GPE534 strongly, although to a different extent, compete with the neutralizing antibody h13F6. The antibodies GPE118 and GPE534 compete rather weakly with the neutralizing antibody c13C6 and weakly compete with the neutralizing antibody c6D8. None of the antibodies under study competes with non-neutralizing murine mAb 4F3. The competitive ELISA data (not shown) allow one to calculate the coefficient

of inhibition (CI) for the binding of full-length recombinant antibodies with biotinylated EBOV rGPdTM in the presence of control commercial antibodies at different concentrations (Table 3). CI is the ratio between absorbance in competitive ELISA in the presence (3 µg/ml) and in the absence of the control mAbs. At $CI \geq 1$, there is no competition between the control mAb and the full-length antibodies under study; i.e., the antibodies are targeted against different epitopes. If CI values are below 1, the control mAb and the full-length antibodies under study interact with the same or the closely located epitopes. The smaller the CI value, the closer the epitopes are located.

Table 1. Comparison of K_d for parental murine mAbs and full-length chimeric antibodies

Sample	Subisotopes	K_d , nM
GPE118	IgG1 kappa, mouse	1.7–2.0
xGPE118	IgG1 kappa, human	2.5–4.0
GPE325	IgM kappa, mouse	1.8–3.4
xGPE325	IgG1 kappa, human	1.2–2.5
GPE534	IgG2b kappa, mouse	0.8–1.0
xGPE534	IgG1 kappa, human	1.3–1.9

Note. xGPE118, xGPE325, xGPE534 – chimeric recombinant antibodies against rGPdTM.

Table 2. Properties of the commercial anti-EBOV GP antibodies used in this study

Antibody	Species	Epitope	Polypeptide	Component of the antibody cocktail	Neutralizing activity	Reference
KZ52	Human	conformational	GP1–GP2	none	+	[11]
h13F6	Mouse/human	404–412	GP1	MB-003	+	[4]
c13C6 FR1	Mouse/human	33–295	GP1	MB-003, ZMapp	+	[4, 5]
c6D8	Mouse/human	393–401	GP1	MB-003	+	[4]
4F3	Mouse	NA		none	-	

Table 3. The coefficient of inhibition of experimental samples of the full-length antibodies GPE 118, GPE 325, and GPE 534 by the control mAbs according to the data of competitive ELISA with biotinylated EBOV rGPdTM

Control mAbs	CI of the control mAb	GPE118 IgG/Fab*/mAb	GPE325 IgG/Fab*/mAb	GPE534 IgG/Fab*/mAb	Epitope, a.a.r.	Polypeptide
h13F6	0.39	0.35/0.69/0.31	0.41/0.57/0.30	0.96/1.00/0.97	404–412	GP1
c13C6FR1	0.20	0.86/0.93/0.94	1.13/0.94/1.18	0.96/1.01/0.88	33–295	GP1
c6D8	0.39	0.76/0.99/0.77	0.60/0.64/0.70	0.67/0.99/0.97	393–401	GP1
KZ52	0.17	1.09/1.02/1.03	1.04/1.10/1.11	0.63/0.92/0.86	Conf.	GP1–GP2
4F3	0.43	1.05/0.90/1.23	1.31/1.10/1.44	1.0/0.87/0.96	NA	NA

*According to the findings reported in [13].

Note. contr. mAbs – control mAbs; IgG – full-length chimeric antibodies; conf. – conformational; NA – no data available.

An analysis of the CI values allows one to infer that the epitopes of the full-length recombinant epitopes GPE118, GPE325, and GPE534, chimeric Fab fragments [13], and the parental murine mAbs coincide. All three candidate anti-EBOV GP antibodies are also targeted against different epitopes. None of them competes with nonneutralizing mAb 4F3. The epitope of the antibodies GPE118 and GPE325 overlaps with the epitope of mAb h13F6 localized in the EBOV glycoprotein mucin-like domain between the amino acid residues 404–412 of the GP1 glycoprotein subunit. GPE118 shows stronger competition for binding to the antigen than the antibody h13F6 competes with itself. The epitope of the antibody GPE325 overlaps with the epitope of mAb c6D8 localized between the amino acid residues 393–401 of the GP1 subunit and is slightly shifted towards the N-end of the GP1 subunit with respect to the epitope of GPE118. The epitope of GPE534, having a linear nature according to the immunoblotting data, resides near the conformational epitope of mAb KZ52 formed by numerous amino acid residues of the GP1 and GP2 subunits of the viral glycoprotein. Hence, all three full-length recombinant chimeric anti-EBOV GP antibodies have epitopes that either coincide or overlap with the epitopes of the three well-studied neutralizing anti-EBOV mAbs.

CONCLUSIONS

Full-length recombinant chimeric antibodies against the Ebola virus glycoprotein GPE118, GPE325, and GPE534 were constructed and produced in CHO cells. The immunochemical properties of the full-length re-

combinant chimeric antibodies were studied by immunoblotting and indirect, direct, and competitive ELISA using the recombinant EBOV proteins rGPdTM, NP, and VP40. The authenticity of the full-length recombinant chimeric IgG1 antibodies and their specificity with respect to EBOV GP were proven by immunoblotting using mAbs 4G7 against the human Ig light chain kappa and mAbs 2C11 against the human IgG1 heavy chain. The results of indirect ELISA demonstrate that there is no cross-specificity with respect to the proteins NP and VP40 of the Ebola virus. The affinity of the full-length antibodies shows that the experimental samples exhibit high affinity to EBOV GP.

An analysis of the coefficients of inhibition determined by competitive ELISA using a panel of commercial neutralizing antibodies allows us to draw the conclusion that all the antibodies under study are targeted against different glycoprotein regions. The epitopes of the recombinant antibodies either coincide or partially overlap with the epitopes of three commercial neutralizing anti-Ebola virus antibodies. This finding demonstrates that the resulting recombinant antibodies exhibit a potentially high neutralizing activity. ●

This work was financially supported by the subsidy (Agreement no. 14.607.21.0096) of the Ministry of Education and Science of the Russian Federation under the Federal Targeted Program for Research and Development in Priority Areas of Advancement of the Russian Scientific and Technological Complex for 2014–2020 (unique project identifier RFMEFI60714X0096).

REFERENCES

- Geisbert T.W., Young H.A., Jahrling P.B., Davis K.J., Kagan E., Hensley L.E. // *J. Infect. Dis.* 2003. V. 188. № 11. P. 1618–1629.
- Warren T.K., Warfield K.L., Wells J., Swenson D.L., Donner K.S., van Tongeren S.A., Garza N.L., Dong L., Mourich D.V., Crumley S., et al. // *Nat. Med.* 2010. V. 16. № 9. P. 991–994.
- Oswald W.B., Geisbert T.W., Davis K.J., Geisbert J.B., Sulivan N.J., Jahrling P.B., Parren P.W., Burton D.R. // *PLoS Pathog.* 2007. V. 3. № 1. e9.
- Olinger G.G. Jr., Pettitt J., Do Kim, Working C., Boharov O., Bratcher B., Hiatt E., Hume S., Johnson A., Morton J., et al. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 44. P. 18030–18035.
- Corti D., Misasi J., Mulangu S., Stanley D.A., Kanekiyo M., Wollen S., Ploquin A., Doria-Rose N.A., Staube R.P., Bailey M., et al. // *Science.* 2016. V. 351. № 6279. P. 1339–1342.
- Qiu X., Wong G., Audet J., Bello A., Fernando L., Alimonti J.B., Fausther-Bovendo H., Wei H., Aviles J., Hiatt E., et al. // *Nature.* 2014. V. 514. № 7520. P. 47–53.
- PREVAIL II Writing Group.; Multi-National PREVAIL II Study Team, Davey R.T. Jr., Dodd L., Proschan M.A., Neaton J., Neuhaus Nordwall J., Koopmeiners J.S., Beigel J., Tierney J., Lane H.C., Fauci A.S., et al. // *N. Engl. J. Med.* 2016. V. 375. № 15. P. 1448–1456.
- Panina A.A., Aliev T.K., Shchemchukova O.B., Dementieva I.G., Varlamov N.E., Pozdnyakova L.P., Bokov M.N., Dolgikh D.A., Sveshnikov P.G., Kirpichnikov M.P. // *Dokl Biochem Biophys.* 2016. V. 467. № 1. P. 117–120.
- Shchemchukova O.B., Dementieva I.G., Varlamov N.E., Pozdnyakova L.P., Bokov M.N., Aliev T.K., Panina A.A., Dolgikh D.A., Kirpichnikov M.P., Sveshnikov P.G. // *Moscow University Biological Sciences Bulletin.* 2016. V. 71. № 1. P. 24–28.
- Balabashin D.S., Aliev T.K., Toporova V.A., Panina A.A., Dolgikh D.A., Kirpichnikov M.P. Patent RF № 2555533.
- Friguet B., Chaffotte A.F., Djavadi-Ohanian L., Goldberg M.E. // *J. Immunol. Meth.* 1985. V. 77. № 2. P. 305–319.
- Lee J.E., Fusco M.L., Hessell A.J., Oswald W.B., Burton D.R., Saphire E.O. // *Nature.* 2008. V. 454. P. 177–182.
- Panina A.A., Dementieva I.G., Aliev T.K., Toporova V.A., Bokov M.N., Pozdnyakova L.P., Shchemchukova O.B., Dolgikh D.A., Sveshnikov P.G., Kirpichnikov M.P. // *Bioorg Khim.* 2017. V. 43. № 4. P. 398–401

Direct Molecular Fishing of New Protein Partners for Human Thromboxane Synthase

A. V. Svirid¹, P. V. Ershov², E. O. Yablokov², L. A. Kaluzhskiy², Yu. V. Mezentsev^{2*},
A. V. Florinskaya², T. A. Sushko^{1,3}, N. V. Strushkevich¹, A. A. Gilep¹, S. A. Usanov¹,
A. E. Medvedev², A. S. Ivanov²

¹Institute of Bioorganic Chemistry, National Academy of Sciences of Belarus, Akad. Kuprevicha Str. 5/2, Minsk, 220141, Belarus

²Institute of Biomedical Chemistry, Pogodinskaya Street 10, bldg. 8, Moscow, 119121, Russia

³Department of Bioengineering, School of Engineering, The University of Tokyo, Tokyo, 108-8639, Japan

*E-mail: yu.mezentsev@gmail.com

Received: April 20, 2017; in final form October 11, 2017

Copyright © 2017 Park-media, Ltd. 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 Thromboxane synthase (TBXAS1) catalyzes the isomerization reaction of prostaglandin H₂ producing thromboxane A₂, the autocrine and paracrine factor in many cell types. A high activity and metastability by these arachidonic acid derivatives suggests the existence of supramolecular structures that are involved in the regulation of the biosynthesis and directed translocation of thromboxane to the receptor. The objective of this study was to identify TBXAS1 protein partners from human liver tissue lysate using a complex approach based on the direct molecular fishing technique, LC-MS/MS protein identification, and protein-protein interaction validation by surface plasmon resonance (SPR). As a result, 12 potential TBXAS1 protein partners were identified, including the components regulating cytoskeleton organization (BBIP1 and ANKMY1), components of the coagulation cascade of human blood (SERPINA1, SERPINA3, APOH, FGA, and FN1), and the enzyme involved in the metabolism of xenobiotics and endogenous bioregulators (CYP2E1). SPR validation on the Biacore 3000 biosensor confirmed the effectiveness of the interaction between CYP2E1 (the enzyme that converts prostaglandin H₂ to 12-HHT/thromboxane A₂ proantagonist) and TBXAS1 ($K_d = (4.3 \pm 0.4) \times 10^{-7}$ M). Importantly, the TBXAS1 · CYP2E1 complex formation increases fivefold in the presence of isatin (indole-2,3-dione, a low-molecular nonpeptide endogenous bioregulator, a product of CYP2E1). These results suggest that the interaction between these hemoproteins is important in the regulation of the biosynthesis of eicosanoids.

KEYWORDS Thromboxane synthase (CYP5A1, TBXAS1), cytochrome P450, surface plasmon resonance, direct molecular fishing, protein partners, isatin.

ABBREVIATIONS Protein-protein interaction (PPI), surface plasmon resonance (SPR), association rate constant (k_{on}), dissociation rate constant (k_{off}), equilibrium dissociation constant (K_d).

INTRODUCTION

Human thromboxane synthase (TBXAS1) belongs to the cytochrome P450 superfamily (CYP5A1). However, it functions differently from “classical” cytochromes P450, which catalyze various monooxygenase reactions, involving redox partners as the electron donors [1]. TBXAS1 catalyzes the reaction of prostaglandin H₂ (PGH₂) isomerization, which requires no redox partners and produces thromboxane A₂ (TXA₂)[2]. The latter acts as a paracrine and autocrine regulator and is an important mediator of platelet aggregation and contraction of blood vessels, which contributes to increase in blood pressure.

Apart from PGH₂ isomerization, TBXAS1 catalyzes the alternative PGH₂ transformation reaction, resulting in its cleavage to 12-hydroxy-5,8,10-heptatrienic acid (12-HHT) and malondialdehyde (MDA) [3]. There is currently no accurate information on the functional role of MDA and 12-HHT. MDA can form adducts with the protein amino groups or polar groups of phospholipids and thus plays a role in the molecular mechanisms of atherosclerosis, cancer, and some genetic diseases [4, 5]. 12-HHT and its metabolites can block the action of leukotriene receptors and act as a partial TXA₂ antagonist by enhancing the synthesis of prostacyclin and antagonizing the thromboxane receptor (TXAR) [6, 7]. It

is possible that TBXAS1 also performs other functions: catalyze monooxygenase reactions characteristic of cytochrome P450 and involving redox partners.

TBXAS1 was first isolated from human platelets [3] and pig lungs [8]. TBXAS1 is mostly synthesized in prothrombocytes and monocyte precursor hematopoietic stem cells, leukocytes, and macrophages, where TXA₂ is involved in the regulation of cell differentiation [9]. Synthesis of TBXAS1 was also detected in the cells of lungs, kidneys, the stomach, intestine, spleen, thymus, pancreas, and the liver [10]. TXAR, which belongs to the class of G-protein-coupled receptors (GPCRs), is expressed in many tissues (lung, spleen, liver, uterus, placenta, aorta, heart muscle, intestine, thymus, kidney, brain, and spinal cord) [11]. This may be indicative of other possible functions of TBXAS1 or the versatility of the mechanisms underlying its basic function.

One approach to elucidating the unknown functions of a protein is based on studying its interactions with other proteins whose functions are known [12]. This approach is based on the concept that the functions of the interacting protein partners must be either interrelated or form a single protein complex that performs interrelated functions. The substrate and product of the reaction catalyzed by TBXAS1 are extremely short-living and active lipophilic molecules, whose diffuse transport is complicated, while the TXA₂ receptor is located on the outside of the plasma membrane. This suggests the existence of a specific transport mechanism or, most likely, an interaction with the associated protein complexes responsible for the transportation of these short-living compounds.

To date, information on experimental validation of protein-protein interactions (PPI) involving TBXAS1 remains scarce. The BioGRID database includes only two records of identified PPIs involving TBXAS1 (<https://thebiogrid.org/112778/summary/homo-sapiens/tbxas1.html?sort=bait>): (1) interaction with an eukaryotic elongation factor 1 α -2 (EEF1A2) citing unpublished data [13], and (2) interaction with ubiquitin C (UBC) [14]. It is most likely that both these interactions are nonspecific, since the same BioGRID database includes records of 132 potential interactions of EEF1A2 with 124 partners and 2,332 interaction of UBC with 1,440 partners. In 2016, Meling D.D. cited in the abstract of his dissertation (Protein-protein interactions and mechanistic insights for CYP2J2 and TBXAS1) unpublished data on interaction between TBXAS1 and cytochrome P450 reductase (CPR) (<http://hdl.handle.net/2142/90774>), which undoubtedly may be functionally significant, since CPR is a known protein partner of the microsomal cytochromes P450.

Previously, we developed an integrated approach to the discovery of novel protein partners interacting

with a target protein which is based on the use of direct molecular fishing on the affinity sorbent with the immobilized target protein (or peptide) as a ligand, mass spectrometric identification of the isolated proteins, and validation of the potential PPIs by surface plasmon resonance (SPR) [15–17].

The objective of the present study was to search for novel potential TBXAS1 protein partners in the human liver tissue lysate using this approach. As a result, 12 potential TBXAS1 protein partners were isolated on the affinity column with immobilized TBXAS1 using a LC-MS/MS-analysis, one of which was cytochrome P450 (CYP2E1). SPR validation confirmed its interaction with TBXAS1 immobilized on the optical chip and identified another potential protein partner (CYP11B2). SPR experiments with five control cytochromes P450 (CYP2C19, CYP11A1, CYP11B1, CYP3A4, CYP3A5) were negative, indicating the high specificity of the detected PPIs. Since CYP2E1 is involved in the metabolism of various indole derivatives [18], we further investigated the possible influence of the well-known endogenous bioregulator isatin (indole-2,3-dione) [19–22] on the interaction of CYP2E1 and CYP11B2 with TBXAS1. We found that isatin results in a fivefold increase in affinity of the TBXAS1 · CYP2E1 interaction and does not affect the TBXAS1 · CYP11B2 interaction.

EXPERIMENTAL

Protein preparations

Highly purified (> 95% according to denaturing polyacrylamide gel electrophoresis (SDS-PAGE)) preparations of the recombinant proteins, TBXAS1, cytochromes P450 (limonene 6-monooxygenase (CYP2C19), steroid-20,22-lyase (CYP11A1), steroid-11 β -hydroxylase (CYP11V1), aldosterone synthase (CYP11B2), taurochenodeoxycholate-6 α -monooxygenase (CYP3A4), cyclic hydrocarbon hydroxylase (CYP3A5), 4-nitrophenol-2-hydroxylase (CYP2E1), microsomal cytochrome b₅ (CYB5A), NADPH-cytochrome-P450-reductase (CPR), NADPH-adrenodoxin reductase (ADR), adrenodoxin (ADX), and ferrochelatase (FECH), SMAD4, RAB27B) were prepared at the Institute of Bioorganic Chemistry (Republic of Belarus) by molecular cloning and heterologous expression in a bacterial system (*E. coli*), followed by purification using metal-affinity and ion exchange chromatography [23, 24]. The preparation of retinol-binding protein 4 (RBP4) was obtained from Cayman chemical (USA).

Human liver tissue lysate

Human liver tissue samples were obtained from the ILSbio LLC (www.ilsbio.com). The lysate was prepared

by homogenization of a 100-mg tissue liver sample in a Potter mortar with 1 mL of the CellLytic Mammalian Tissue Lysis/Extraction Reagent (Sigma, USA) and 10 μ L of a protease inhibitor cocktail (Sigma, USA). After centrifugation at 13,400 g and 4°C for 25 min, the supernatant was collected, glycerol was added to a final concentration of 25%, and the resulting solution was stored at -80°C . The total protein concentration in the lysate samples was 10–20 mg/mL, as determined spectrophotometrically using a Bradford assay.

Direct molecular fishing

An affinity sorbent with covalently immobilized TBXAS1 as the bait protein was prepared by covalent protein binding to CNBr-Sepharose 4B (GE Healthcare, USA) according to the manufacturer's protocol. It was found that the 0.5 mg/1 g protein to sorbent ratio was optimal for binding of the used TBXAS1 preparation to the sorbent. The remaining active groups of the sorbent were inactivated by incubation in a buffer containing 100 mM Tris-HCl (pH 7.4) and 150 mM NaCl. Direct molecular fishing was carried out in the original microcolumn (volume 200 μ L) filled with the affinity sorbent. In the control experiments, a similar microcolumn filled with "empty" (no bait protein) inactivated CNBr-Sepharose 4B was used. HBS-EP+ buffer (10 mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.05% Tween 20) passed through the microcolumn at a flow rate of 50 μ L/min at 15°C was used as a running buffer. Affinity isolation of the TBXAS1 protein partners was carried out by passing 2 mL of lysate (0.5 mg/mL of protein) twofold diluted with the running buffer through the column for 80 minutes using the 10 AKTA Purifier (GE Healthcare, USA) system. Proteins from the lysate bound to the sorbent were eluted with 4% HCOOH (pH 2.5) at a flow rate of 50 μ L/min for 100 min. The total protein content in the eluates determined by the Bradford assay was 25–35 μ g/mL (average value 30 μ g/mL). The experiments on the affinity isolation of potential TBXAS1 protein partners were repeated in triplicates.

LC-MS/MS-analysis

Special sample preparation was used for the mass-spectrometric identification of the proteins. An aliquot containing 30 μ g of total protein was sampled from each chromatographic fraction and subjected to the standard trypsinolysis procedure with preliminary reduction and alkylation of the sulfhydryl groups of the proteins. All procedures were carried out in Vivaspin 500 Centrifugal Concentrators, 10 kDa MWCO (GE Healthcare, USA) using the FASP method [25]. A lyophilized trypsin preparation obtained from porcine pancreas (activity 15600 IU/mg, V5111, Promega, USA) was used for trypsin digestion of the proteins.

Mass-spectrometric analysis of the samples was carried out in three technical replicates using a Agilent 1200 chromatograph and Agilent 6300 mass-detector with Ion Trap LC/MS (Agilent Technologies, USA). Peptides were separated using a reversed-phase HPLC-column ZORBAX Extend-C18 (2.1 \times 150 mm, 1.8 μ m) (Agilent Technologies, USA) in a gradient of solvent A (0.2% formic acid in water) and Solvent B (0.2% formic acid in acetonitrile) for 55 min at a flow rate of 350 μ L/min. The sample volume applied to the column was 15 μ L (~ 7–8 mg of material). The gradient was as follows: from 0 to 20% of solvent B in 5 minutes, from 20 to 80% in 40 minutes, from 80 to 95% in 5 min, and 95% for 5 min. Column temperature was 50°C. Mass spectra were acquired in the positive ionization mode (APESI-ionization) with the following parameters: gas temperature 400°C, gas flow rate 9 L/min, capillary voltage 2 kV, and fragmentor voltage 360 V. The mass analyzer was operated in the auto-MS/MS-mode with the following parameters: m/z range from 50 to 2200 m/z , fragmentation energy was calculated according to the following formulas: $(3.1(m/z)/100 + 1.0)$ V for $z = 2$ and $(3.6(m/z)/100 - 4.8)$ V for $z \geq 3$. The proteins were identified using the Mascot software (www.matrixscience.com) and SwissProt database (www.uniprot.org). The following search parameters were used: proteolytic enzyme trypsin, acceptable mass deviation of monoisotopic peptides ± 2.6 Da, acceptable MS/MS deviation ± 0.6 Da, acceptable number of omitted trypsin cleavage sites is 2, variable modification – "oxidized methionine," and fixed modifications – "carbamidomethyl." The resulting list of reliably detected proteins included only those proteins which were identified in three technical replicates with a significance of 0.01 and Mascot Score > 50 .

Surface plasmon resonance (SPR)

PPIs were analyzed on the four-channel optical biosensor Biacore 3000 (GE Healthcare, USA), whose operation is based on the surface plasmon resonance effect controlled by the Biacore Software v. 1.0. Biosensor signals were recorded in resonance units, RU (1 RU corresponds to binding of about 1 pg of the protein on the optical chip surface). The values of the equilibrium dissociation constants (K_d), association rate constants (k_{on}), and dissociation rate constants (k_{off}) of the complexes were calculated using the BiaEvaluation v. 4.1 software package.

TBXAS1 was immobilized by the formation of covalent bonds between the carboxyl groups on the surface of the optical chip CM5 and the free amino groups of the protein. For this purpose, we used the Amine Coupling Kit (GE Healthcare, USA). The TBXAS1 sample (50 μ g/mL) in 10 mM acetate buffer (pH 5.0) was inject-

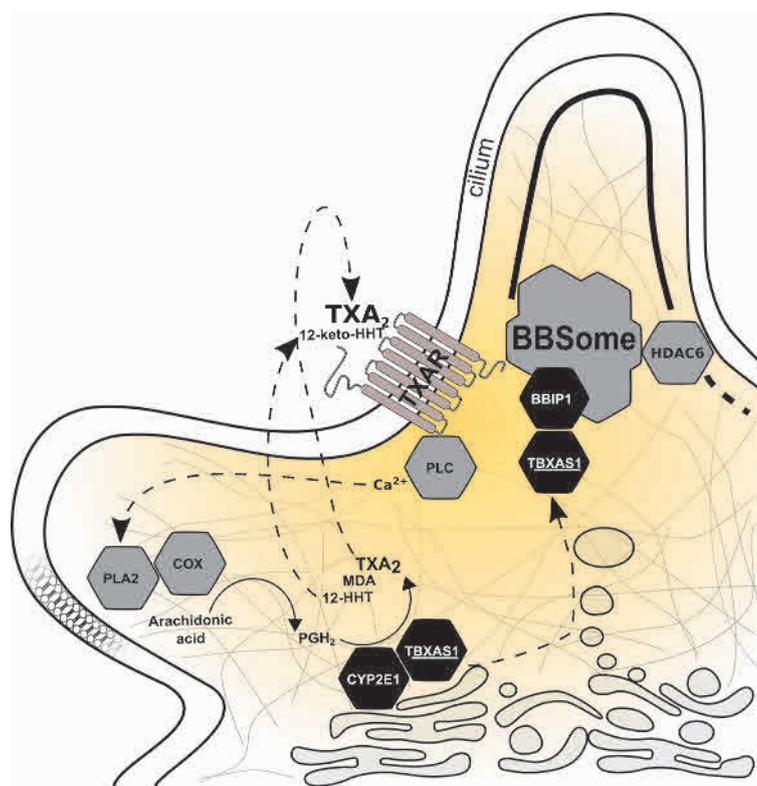


Fig. 1. Schematic representation of TXA₂ biosynthesis supplemented by our experimental results. PLC – phospholipase C, PLA2 – phospholipase A₂, COX – cyclooxygenase. Biosynthesis of TXA₂ begins with the release of arachidonic acid from membrane phospholipids assisted by PLA2. Then, COX catalyzes the transformation of arachidonic acid into prostaglandin H₂, which is then metabolized by TBXAS1 to form TXA₂, 12-HHT, and MDA. At the same time, prostaglandin H₂ is transformed by CYP2E1 to 12-HHT and MDA. TXA₂ binding to TXAR causes signal transmission via the inositol phosphate pathway with PLC activation and mobilization of intracellular Ca²⁺, which has a stimulating effect on PLA2. Further, 12-HHT is metabolized by 15-hydroxyprostaglandin dehydrogenase to form 12-keto-HHT, which has a partial antagonistic effect on TXAR. TBXAS1 also presumably interacts with BBIP1, which is a component of the protein transport complex of cilia (BBSome). BBIP1 can influence the stability of the microtubulin cytoskeleton, indirectly inhibiting HDAC6 (microtubule deacetylase).

ed at a flow rate of 5 $\mu\text{L}/\text{min}$ for 20 min. The TBXAS1 immobilization level in the working channel of the optical biosensor averaged 7,500 RU (7.5 ng/mm^2).

Interactions of test proteins with the immobilized TBXAS1 were recorded in real time mode while injecting protein samples at a concentration ranging from 50 nM to 5 μM through the control channel (without the protein) and then through the channel with immobilized TBXAS1 for 10 min at a flow rate of 5 $\mu\text{L}/\text{min}$. Each measurement was followed by a regeneration of the optical chip surface by injecting buffer containing 2 M NaCl and 0.4% CHAPS for 30 s at a flow rate of 20 $\mu\text{L}/\text{min}$. All measurements were performed at least 4 times, which provided adequate accuracy and reproducibility (CV value was less than 10%). In the experiments assessing the possible impact of nonpeptide low-molecular-weight endogenous bioregulator isatin (2,3-dioxindole) on PPIs involving TBXAS1, isatin was added to the samples of analyzed proteins at a final concentration of 100 μM and the mixture was incubated for 15 min.

RESULTS AND DISCUSSION

An integrated approach based on direct molecular fishing on an affinity sorbent with a 4B target protein immobilized on CNBr-Sepharose, mass-spectrometric identification of isolated proteins, and validation

of PPIs with SPR enables the isolation and identification of 12 potential TBXAS1 protein partners from the lysates of human liver tissue (*Table*). To date, the scientific literature still provides no information on the interaction between these proteins and TBXAS1. However, some assumptions about their possible functional relationship with TBXAS1 can be made. For example, the BBIP1 protein (BBSome component of the transport protein complex of cilia) is involved in the regulation of cellular cytoskeleton stability [26, 27]. Information on ANKMY1 is available only at the transcript level. However, its structure includes ankyrin repeats, which form one of the most common interfaces for PPIs. These repeats were found in proteins characterized by various functions [28]. Based on this fact, we assumed that ankyrin repeats of ANKMY1 can specifically recognize certain structural motifs of TBXAS1, facilitating interaction between these proteins.

However, it should be noted that proteins identified by direct molecular fishing can only be considered as potential protein partners of TBXAS1, since not only real partner proteins, but also simultaneously “fished” extraneous proteins composing micelles or supramolecular complexes may be isolated from the lysate due to the features of this techniques [17]. Among our “fished” proteins (*table*), these are SERPINA1, SERPINA3, APOH, FGA, and FN1, which are involved in the blood

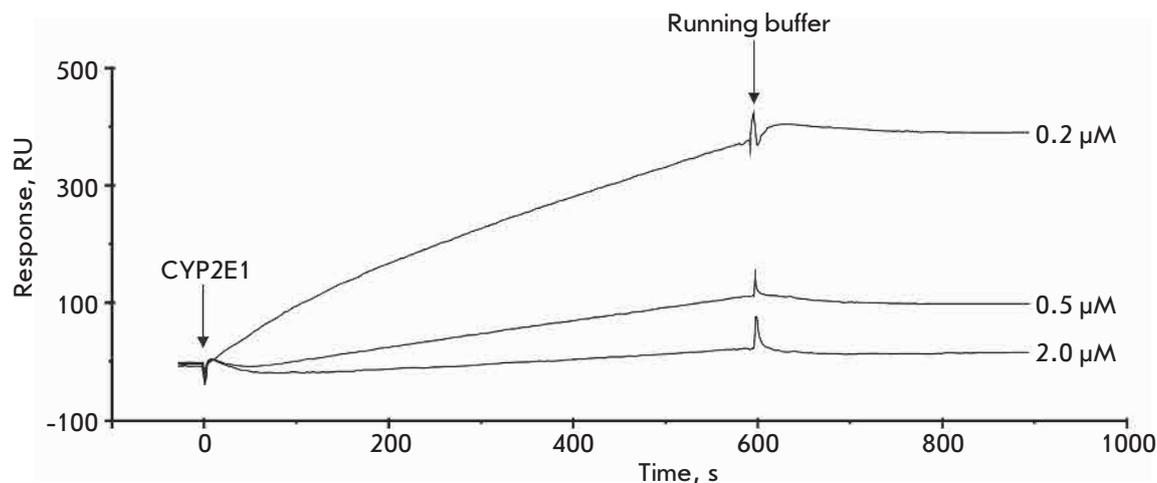


Fig. 2. Typical sensorgrams of the interaction between various concentrations of CYP2E1 and TBXAS1 immobilized on a CM5 optical chip

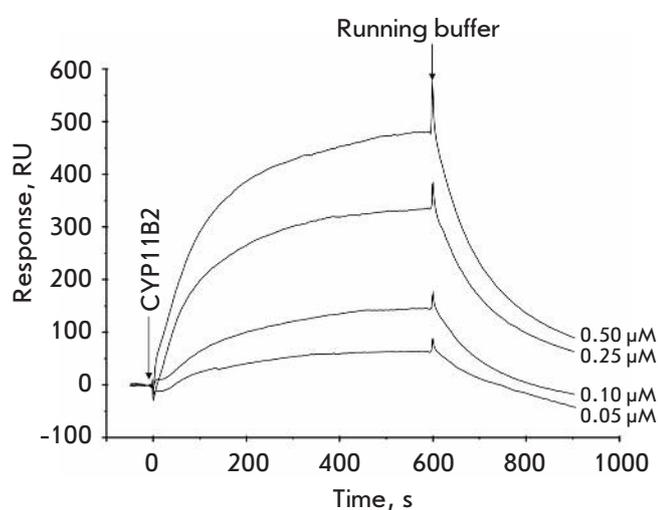


Fig. 3. Typical sensorgrams of the interaction between various concentrations of CYP11B2 and TBXAS1 immobilized on a CM5 optical chip

clotting cascade [29, 30], as well as serum proteins (HP, SAA1, CP), which can have a high nonspecific adsorption level.

The presence of CYP2E1, which belongs to the cytochromes P450 superfamily, in the list of “fished” proteins is of particular interest. A functional relationship between CYP2E1 and TBXAS1 may be important in the context of the complementary enzymatic conversion reactions of common substrates. As it is known, CYP2E1 is characterized by broad substrate specificity and a broad tissue localization profile, including the liver [31]. For example, CYP2E1 can oxidize arachidonic acid (via ω -1-hydroxylation) and prostaglandin H_2 [32] to form side metabolites, which, in turn, are formed in the prostaglandin H_2 to thromboxane A_2 isomerization reactions. 12-keto-HHT is a further metabolite of

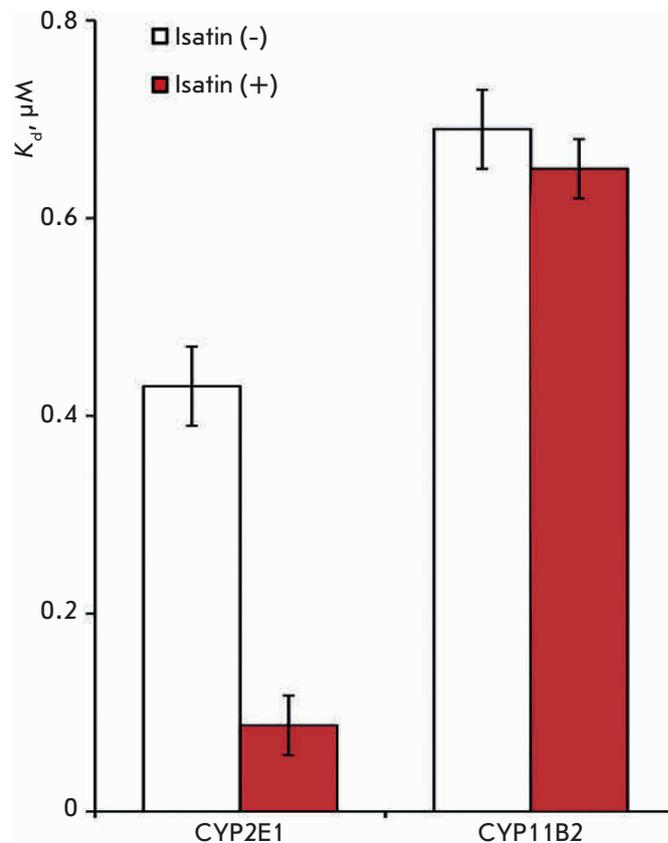


Fig. 4. Diagram representation of equilibrium dissociation constant (K_d) values of the TBXAS1 · CYP2E1 and TBXAS · CYP11B2 complexes in the absence and presence of 100 μ M of isatin; $M \pm m$, $n = 3$

one of the reaction products and can influence TXA_2 by increasing prostacyclin production and antagonistic action on $TXAR$ [4, 5]. Therefore, colocalization of TBXAS1, synthesizing thromboxane A_2 , and CYP2E1 could serve as an additional mechanism regulating the effectiveness of enzymatic conversion of common sub-

Mass-spectrometric identification of the proteins in eluates from chromatographic microcolumns filled with an affinity sorbent

S/N	Gene	Protein	MW, Da	Uniprot number ^a	Score ^b	Peptides ^c	emPAI ^d
Sorbent with immobilized thromboxane (only Test). (TBXAS1 potential protein partners)							
1	FGA	Fibrinogen alpha chain	95656	P02671	97	19 (2)	0.04
2	FN1	Fibronectin	266052	P02751	95	14 (2)	0.01
3	CP	Ceruloplasmin	122983	Q24478	85	10 (5)	0.03
4	SERPINA3	Alpha1-antichymotrypsin	47792	P01011	67	9 (3)	0.08
5	SAA1	Serum amyloid A1 protein	13581	P0DJI8	67	7 (3)	0.29
6	CYP2E1	Cytochrome P450 2E1	56849	P05181	61	11 (4)	0.12
7	ANKMY1	Ankyrin repeat and MYND domain-containing protein 1	107101	Q9P2S6	59	25 (7)	0.03
8	ACTB	Actin, cytoplasmic 1	42052	P60709	56	17 (2)	0.13
9	BBIP1	BBSome-interacting protein 1	10557	A8MTZ0	54	10 (4)	0.38
10	SERPINA1	Alpha 1-antitrypsin	46878	P01009	50	3 (3)	0.13
11	APOH	Beta2-glycoprotein 1	39584	P02749	52	3 (3)	0.09
12	HP	Haptoglobin	45861	Q61687	50	42 (6)	0.21
Sorbent without protein immobilization (only Control)							
1	ACY1	Aminoacylase-1	46084	Q03154	59	17 (3)	0.08
2	ADH1A	Alcohol dehydrogenase 1A	40745	P07327	189	85 (19)	0.42
3	SLC25A4	ADP/ATP translocase 1	33271	P12235	59	16 (5)	0.24
4	SLC25A5	ADP/ATP translocase 2	33059	P05141	71	27 (5)	0.11
5	MAOB	Amine oxidase [flavin-containing] B	59238	P27338	147	31 (8)	0.20
6	ASL	Argininosuccinate lyase	51910	P04424	89	9 (3)	0.07
7	ASS1	Argininosuccinate synthase	46786	P00966	59	48 (6)	0.17
8	ATP5B	ATP synthase subunit beta, mitochondrial	56525	P25705	108	27 (4)	0.07
9	ATP5C1	ATP synthase subunit gamma, mitochondrial	33032	P36542	95	16 (4)	0.24
10	CALR	Calreticulin	48283	P27797	68	10 (4)	0.25
11	HSPD1	60 kDa heat shock protein, mitochondrial	61187	P10809	126	28 (7)	0.13
12	CPS1	Carbamoyl-phosphate synthase [ammonia], mitochondrial	165975	P31327	249	119 (27)	0.17
13	DEFA1	Neutrophil defensin 1	10536	P59665	58	8 (5)	0.38
14	FBP1	Fructose-1,6-bisphosphatase 1	37218	P09467	78	4 (3)	0.10
15	FABP1	Fatty acid-binding protein, liver	14256	P07148	184	29 (17)	1.08
16	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	36201	P04406	169	23 (5)	0.22
17	AGL	Glycogen debranching enzyme	176819	P35573	77	18 (7)	0.06
18	SHMT1	Serine hydroxymethyltransferase, cytosolic	53619	P34896	88	16 (5)	0.14
19	HSPA5	78 kDa glucose-regulated protein	72402	P11021	75	19 (4)	0.11
20	GSTA1	Glutathione S-transferase A1	25672	P08263	94	74 (13)	1.29
21	IDH1	Isocitrate dehydrogenase [NADP] cytoplasmic	46915	O75874	274	15 (9)	0.17
22	IDH2	Isocitrate dehydrogenase [NADP], mitochondrial	51333	Q8IQA7	105	10 (4)	0.15
23	LDHA	L-lactate dehydrogenase A chain	36950	P00338	240	9 (8)	0.21
24	NONO	Non-POU domain-containing octamer-binding protein	54311	Q15233	102	27 (6)	0.14
25	PGM1	Phosphoglucomutase-1	61696	P36871	120	10 (6)	0.12
26	SFPQ	Splicing factor, proline- and glutamine-rich	76216	P23246	106	27 (7)	0.10
27	ACAT1	Acetyl-CoA acetyltransferase, mitochondrial	45456	P24752	60	11 (4)	0.17
28	TPI1	Triosephosphate isomerase	31057	P60174	150	12 (4)	0.26
29	UGT2B10	UDP-glucuronosyltransferase 2B10	61190	P36537	55	8 (3)	0.06
30	UGP2	UTP-glucose-1-phosphate uridylyltransferase	57076	Q16851	72	15 (4)	0.13

RESEARCH ARTICLES

Control - upper line, Test - bottom line							
1	UGP2	Alcohol dehydrogenase 1B	40684	P00325	408	117 (34)	1.02
					90	53 (6)	0.30
2	ADH4	Alcohol dehydrogenase 4	41108	P08319	222	36 (16)	0.42
					85	13 (8)	0.30
3	ALB	Serum albumin	71317	P02768	2790	360 (150)	1.90
					959	186 (50)	0.66
4	ALDH2	Aldehyde dehydrogenase, mitochondrial	56859	P05091	400	33 (15)	0.29
					94	7 (3)	0.07
5	ALDOB	Fructose-bisphosphate aldolase B	39961	P05062	167	25 (8)	0.31
					220	28 (8)	0.09
6	APOA1	Apolipoprotein A-I	30759	P02647	98	12 (6)	0.42
					53	25 (6)	0.59
7	ATP5F1	ATP synthase subunit b, mitochondrial	28947	P24539	192	17 (10)	0.45
					97	13 (5)	0.28
8	ATP5L	ATP synthase subunit g, mitochondrial	11421	O75964	265	7 (6)	0.35
					157	5 (5)	0.35
9	DCXR	L-xylulose reductase	26182	Q7Z4W1	194	27 (6)	0.15
					161	8 (5)	0.15
10	DECR1	2,4-dienoyl-CoA reductase, mitochondrial	36330	Q16698	250	24 (11)	0.22
					222	13 (8)	0.10
11	HSD17B4	Peroxisomal multifunctional enzyme type 2	80092	P51659	1754	179 (87)	1.26
					112	16 (5)	0.09
12	SORD	Sorbitol dehydrogenase	38927	Q00796	171	17 (13)	0.44
					73	7 (4)	0.10
13	CES1	Liver carboxylesterase 1	62766	P23141	88	28 (4)	0.06
					83	20 (4)	0.06
14	HBA1	Hemoglobin subunit alpha	15305	P69905	70	48 (10)	2.90
					82	16 (6)	0.25
15	HBB	Hemoglobin subunit beta	16102	P68871	153	30 (11)	0.54
					222	29 (14)	0.54
16	HMGCS2	Hydroxymethylglutaryl-CoA synthase, mitochondrial	57113	P22791	249	26 (13)	0.37
					98	10 (3)	0.07
17	HRG	Histidine-rich glycoprotein	60510	P04196	60	14 (6)	0.13
					52	9 (5)	0.13
18	PHB2	Prohibitin-2	33276	Q99623	107	11 (4)	0.11
					99	7 (3)	0.11
19	ACAA2	3-ketoacyl-CoA thiolase, mitochondrial	42354	P42765	82	20 (6)	0.18
					71	15 (3)	0.18
20	TF	Serotransferrin	79294	P02787	138	20 (6)	0.15
					142	22 (7)	0.20
21	SLC25A1	Tricarboxylate transport protein, mitochondrial	34333	P53007	70	9 (4)	0.23
					54	10 (3)	0.23

^a – Numbers in the Uniprot database (<http://www.uniprot.org>).

^b – The reliability of peptide identification by mass spectrometry (MASCOT score).

^c – The number of MASCOT peptides; the number of unique peptides (in parentheses).

^d – emPAI, Exponentially Modified Protein Abundance Index.

The names of the identified proteins are listed in the same form as they appear in the Uniprot database used for their identification.

strates. On the other hand, oligomerization of various cytochromes P450 can also lead to change in the catalytic parameters of enzymatic reactions: e.g., the affinity of the enzymes to the substrate [33]. Schematic representation of a TXA₂ biosynthesis system complemented by our experimental data is shown in *Fig. 1*.

We confirmed the formation of the heteromeric TBXAS1 · CYP2E1 complex in direct SPR experiments (*Fig. 2*). The specificity of the TBXAS1 and CYP2E1 interaction was tested by running control SPR experiments using both microsomal (CYP2C19, CYP3A4, CYP3A5) and mitochondrial (CYP11A1, CYP11B1, CYP11B2) cytochromes P450 as analytes. Other well known cytochrome P450 protein partners (CYB5A, CPR, ADR, ADX) and several proteins unrelated to the cytochrome P450 monooxygenase system (FECH, SMAD4, RAB27B, RBP4) were also used for the specificity test. It was shown that all protein analytes, except for CYP11B2 (*Fig. 3*), did not bind to TBXAS1 immobilized on the optical chip even at micromolar concentrations. A similar control experiment using TBXAS1 as a protein analyte showed no dimerization or oligomerization process. Thus, we can confidently state that the interaction of CYP2E1 and CYP11B2 with TBXAS1 is highly specific.

The calculated K_d values of TBXAS1 · CYP11B2 and TBXAS1 · CYP2E1 complex formation were $(6.9 \pm 0.3) \times 10^{-7}$ M and $(4.3 \pm 0.4) \times 10^{-7}$ M, respectively. These values are comparable to K_d of the complexes of various cytochromes P450 with their functional partners (CPR, CYB5A, ADX) [23, 34–37]. It is important to note that, while the difference in the K_d values of complex formation is about twofold, TBXAS1 · CYP11B2 and TBXAS1 · CYP2E1 interactions are very different in their kinetic parameters. Association and dissociation of TBXAS1 · CYP2E1 occur about an order of magnitude slower compared to TBXAS1 · CYP11B2. Association rate constants (k_{on}) are at a 10-fold difference, and dissociation rate constants of the complexes (k_{off}) are at a 15-fold difference.

The revealing of specific TBXAS1 · CYP11B2 complex formation was a new and unexpected result, since CYP11B2 was not identified as a protein partner of TBXAS1 used as a bait in the experiments on molecular fishing from liver tissue lysate (*Table*). These data are quite comparable, they are not due to the false-negative results of molecular fishing, and can be explained in terms of the tissue-specific CYP11B2 expression profile (preferential expression in the adrenal tissue), which appears from the information in the open Internet resources Proteinatlas (<http://www.proteinatlas.org>) and Genecards (<http://www.genecards.org>) and publications [38]. It is currently difficult to deduce the functional consequences and causes of this PPI, so this

paper reports only on the fact of experimental confirmation of a direct interaction between TBXAS1 and CYP11B2.

It is known that indole is oxidized to isatin by some cytochromes P450 (CYP2A6, CYP2C19, and CYP2E1) which are responsible for the metabolism of various xenobiotics [18]. Isatin is an endogenous bioregulator with a wide range of biological and pharmacological activities which are implemented when it interacts with many intracellular isatin-binding proteins [19–22, 39–41]. Since CYP2E1 turned out to be one of the proteins that interact with TBXAS1, we assumed that isatin can affect TBXAS1 · CYP2E1 complex formation. This hypothesis was tested via SPR analysis of the interaction between TBXAS1 and CYP2E1 in the absence and presence of isatin. We found that isatin really affects TBXAS1 · CYP2E1 complex formation, but it has no effect on the TBXAS1 · CYP11B2 interaction (*Fig. 4*) used as a control. The effect of the fivefold increase in the affinity of TBXAS1 · CYP2E1 in the presence of isatin is due to both a twofold increase in k_{on} values and a 2.5-fold decrease in k_{off} .

CONCLUSION

TBXAS1 potential protein partners were isolated from a human liver tissue lysate by direct molecular fishing and mass-spectrometric identification. Using the SPR biosensor technique, it was for the first time shown that TBXAS1 interacts with cytochrome P450 CYP2E1 and CYP11B2, while the affinity of TBXAS1 · CYP2E1 complex formation is fivefold higher in the presence of low-molecular-weight nonpeptide endogenous bioregulator isatin (2,3-dioxindole). Overall, our results suggest that TBXAS1 has other functions, such as participation in the functioning of the cytoskeleton and regulation of the biosynthesis of biologically active molecules. ●

Preparation of recombinant proteins was carried out at the Institute of Bioorganic Chemistry of the Belarusian Academy of Sciences and supported by the Belarusian Republican Foundation for Fundamental Research (Grant No X16P-062). The studies on molecular fishing and identification of potential TBXAS1 protein partners were carried out at the Institute of Biomedical Chemistry and supported by the Russian Foundation for Basic Research (grant No 16-54-00097 Bel_a), and the study of the effect of isatin on PPI involving TBXAS1 was supported by the Russian Science Foundation (grant No 16-14-10327). LC-MS/MS protein identification was performed using equipment from the “Human Proteome” Core Facilities of the Institute of Biomedical Chemistry (Russia)

REFERENCES

1. Waterman M.R., Pikuleva I.A. *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, Weinheim, Germany: Wiley-VCH., 2006. 716 p.
2. Hamberg M., Svensson J., Samuelsson B. // *Proc. Natl. Acad. Sci. USA*. 1975. V. 72. № 8. P. 2994–2998.
3. Haurand M., Ullrich V. // *J. Biol. Chem.* 1985. V. 260. P. 15059–15067.
4. Uchida K. // *Trends Cardiovasc. Med.* 1999. V. 9. № 5. P. 109–113.
5. Chaudhary A.K., Nokubo M., Reddy G.R., Yeola S.N., Morrow J.D., Blair I.A., Marnett L.J. // *Science*. 1994. V. 265. № 5178. P. 1580–1582.
6. Sadowitz P.D., Setty B.N., Stuart M. // *Prostaglandins*. 1987. V. 34. № 5. P. 749–763.
7. Ruf A., Mundkowski R., Siegle I., Hofmann U., Patscheke H., Meese C.O. // *Br. J. Haematol.* 1998. V. 101. № 1. P. 59–65.
8. Shen R.F., Tai H.H. // *J. Biol. Chem.* 1986. V. 261. № 25. P. 11592–11599.
9. Ullrich V., Nüsing R. // *Stroke*. 1990. V. 21. № 12 Suppl. P. IV 134–138.
10. Nüsing R., Sauter G., Fehr P., Dürmüller U., Kasper M., Gudat F., Ullrich V. // *Virchows Arch. A. Pathol. Anat. Histopathol.* 1992. V. 421. № 3. P. 249–254.
11. Huang J.-S., Ramamurthy S.K., Lin X., Le Breton G.C. // *Cell. Signal.* 2004. V. 16. № 5. P. 521–533.
12. Ivanov A.S., Zgoda V.G., Archakov A.I. // *Russ. J. Bioorganic Chem.* 2011. V. 37. № 1. P. 4–16.
13. Huttlin E.L., Ting L., Bruckner R.J., Gebreab F., Gygi M.P., Szpyt J., Tam S., Zarraga G., Colby G., Baltier K., et al. // *Cell*. 2015. V. 162. № 2. P. 425–440.
14. Nathan J.A., Kim H.T., Ting L., Gygi S.P., Goldberg A.L. // *EMBO J.* 2013. V. 32. № 4. P. 552–565.
15. Ershov P., Mezentsev Y., Gnedenko O., Mukha D., Yantsevich A., Britikov V., Kaluzhskiy L., Yablokov E., Molnar A., Ivanov A., et al. // *Proteomics*. 2012. V. 12. № 22. P. 3295–3298.
16. Ivanov A.S., Medvedev A., Ershov P., Molnar A., Mezentsev Y., Yablokov E., Kaluzhsky L., Gnedenko O., Buneeva O., Haidukevich I., et al. // *Proteomics*. 2014. V. 14. № 20. P. 2261–2274.
17. Ivanov A.S., Ershov P.V., Molnar A.A., Mezentsev Y.V., Kaluzhskiy L.A., Yablokov E.O., Florinskaya A.V., Gnedenko O.V., Medvedev A.E., Kozin S.A., et al. // *Russ. J. Bioorganic Chem.* 2016. V. 42. № 1. P. 14–21.
18. Gillam E.M., Notley L.M., Cai H., De Voss J.J., Guengerich F.P. // *Biochemistry*. 2000. V. 39. № 45. P. 13817–13824.
19. Medvedev A.E., Clow A., Sandler M., Glover V. // *Biochem. Pharmacol.* 1996. V. 52. № 3. P. 385–391.
20. Medvedev A., Igosheva N., Crumeyrolle-Arias M., Glover V. // *Stress*. 2005. V. 8. № 3. P. 175–183.
21. Medvedev A., Buneeva O., Glover V. // *Biologics*. 2007. V. 1. № 2. P. 151–162.
22. Pandeya S.N., Smitha S., Jyoti M., Sridhar S.K. // *Acta Pharm.* 2005. V. 55. № 1. P. 27–46.
23. Sergeev G.V., Gilep A.A., Usanov S.A. // *Biochemistry (Moscow)*. 2014. V. 79. № 5. P. 406–416.
24. Dormeshkin D.O., Svirid A.V., Gilep A.A., Usanov S.A. // *Proceedings of the National Academy of sciences of Belarus*. 2015. V. 59. № 2. P. 53–60.
25. Wiśniewski J.R., Zougman A., Nagaraj N., Mann M. // *Nat. Methods*. 2009. V. 6. № 5. P. 359–362.
26. Loktev A.V., Zhang Q., Beck J.S., Searby C.C., Scheetz T.E., Bazan J.F., Slusarski D.C., Sheffield V.C., Jackson P.K., Nachury M.V. // *Dev. Cell*. 2008. V. 15. № 6. P. 854–865.
27. Cerecedo D. // *Blood Coagul. Fibrinolysis*. 2013. V. 24. P. 798–808.
28. Mosavi L.K., Cammett T.J., Desrosiers D.C., Peng Z. // *Protein Sci.* 2004. V. 13. № 6. P. 1435–1448.
29. Moussavi-Harami S.F., Annis D.S., Ma W., Berry S.M., Coughlin E.E., Strotman L.N., Maurer L.M., Westphall M.S., Coon J.J., Mosher D.F., et al. // *J. Proteome Res.* 2013. V. 12. № 7. P. 3393–3404.
30. Pankov R., Yamada K.M. // *J. Cell Sci.* 2002. V. 115. Pt 20. P. 3861–3863.
31. Lu Y., Cederbaum A.I. // *Free Radic. Biol. Med.* 2008. V. 44. № 5. P. 723–738.
32. Arnold C., Konkel A., Fischer R., Schunck W.-H. // *Pharmacol. Rep.* 2010. V. 62. № 3. P. 536–547.
33. Davydov D.R., Davydova N.Y., Sineva E. V., Halpert J.R. // *J. Biol. Chem.* 2015. V. 290. № 6. P. 3850–3864.
34. Gnedenko O.V., Yablokov E.O., Usanov S.A., Mukha D.V., Sergeev G.V., Bulko T.V., Kuzikov A.V., Moskaleva N.E., Shumyantseva V.V., Ivanov A.S., et al. // *Chem. Phys. Lett.* 2014. V. 593. P. 40–44.
35. Bridges A., Gruenke L., Chang Y.T., Vakser I.A., Loew G., Waskell L. // *J. Biol. Chem.* 1998. V. 273. № 27. P. 17036–17049.
36. Strushkevich N., MacKenzie F., Cherkesova T., Grabovec I., Usanov S., Park H.-W. // *Proc. Natl. Acad. Sci. USA*. 2011. V. 108. № 25. P. 10139–10143.
37. Lewis D.F., Hlavica P. // *Biochim. Biophys. Acta*. 2000. V. 1460. № 2–3. P. 353–374.
38. Fallo F., Pezzil V., Barzon L., Mulatero P., Veglio F., Sonino N., Mathis J. M. // *Eur. J. Endocrinol.* 2002. V. 147. P. 795–802.
39. Buneeva O.A., Kopylov A.T., Tikhonova O. V., Zgoda V.G., Medvedev A.E., Archakov A.I. // *Biochemistry (Moscow)*. 2012. V. 77. № 11. P. 1326–1338.
40. Buneeva O., Gnedenko O., Zgoda V., Kopylov A., Glover V., Ivanov A., Medvedev A., Archakov A. // *Proteomics*. 2010. V. 10. № 1. P. 23–37.
41. Medvedev A.E., Buneeva O.A., Kopylov A.T., Gnedenko O.V., Medvedeva M.V., Kozin S.A., Ivanov A.S., Zgoda V.G., Makarov A.A. // *Int. J. Mol. Sci.* 2014. V. 16. № 1. P. 476–495.

YABBY3-Orthologous Genes in Wild Tomato Species: Structure, Variability, and Expression

M. A. Filyushin¹, M. A. Slugina^{1,2}, A. V. Shchennikova¹, E. Z. Kochieva^{1,2}

¹Federal State Institution «Federal Research Centre «Fundamentals of Biotechnology» of the Russian Academy of Sciences», Leninsky pr. 33, bldg. 2, Moscow, 119071, Russia

²Department of Biotechnology, Faculty of Biology, Moscow State University, Leninskie Gory 1, bldg. 12, Moscow, 119991, Russia

*E-mail: michel7753@mail.ru

Received: April 24, 2017; in final form October 17, 2017

Copyright © 2017 Park-media, Ltd. 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 Evolution of the genes encoding YABBY transcription factors is believed to be one of the key reasons for flat leaf emergence from the radially symmetrical stem and gynoeceum diversity. YABBY genes determine the identity of the abaxial surface of all aboveground lateral organs in seed plants. In the present study, complete sequences of YABBY3-orthologous genes were identified and characterized in 13 accessions of cultivated and wild tomato species with diverse morphophysiology of leaves, flowers, and fruits. The obtained gene sequences showed high homology (95–99%) and an identical exon-intron structure with the known *S. lycopersicum* YABBY3 gene, and they contained sequences that encode the conserved HMG-like YABBY and Cys2Cys2-zinc-finger domains. In total, in the analyzed YABBY3 genes, 317 variable sites were found, wherein 8 of 24 exon-specific SNPs were nonsynonymous. In the vegetative and reproductive organs of red-fruited and green-fruited tomato species, YABBY3 gene expression was similar to that in *S. pimpinellifolium* described earlier, but it demonstrated interspecies differences at the leaf-, bud- and flower-specific expression levels.

KEYWORDS YABBY3, polymorphism, qRT-PCR, *Solanum* section *Lycopersicon*, adaxial-abaxial asymmetry.

ABBREVIATIONS CRC – CRABS CLAW; INO – INNER NO OUTER; FIL – FILAMENTOUS FLOWER; qRT-PCR – quantitative real-time PCR.

INTRODUCTION

Plant growth and development processes are controlled by transcription factors, whose evolution is one of the major causes of morphological diversity in the plant kingdom [1–4]. The origin of the flower and reproductive organs is believed to be related to the duplication and changes in MADS-box transcription factor genes [5, 6]. At the same time, flat leaf emergence from the radially symmetrical stem, as well as gynoeceum diversity, is considered to be a consequence of YABBY transcription factor genes evolution [7]. The presence of these genes in angiosperm and gymnosperm plants and their absence in moss and lycopodium [8–10] suggest that YABBY genes originate from one or two predecessors in the last common ancestor of seed plants [10–12]. Diversification of YABBY genes led to the occurrence of individual family members with unique roles in leaf, carpel, and ovule development [8, 11, 13, 14], including the YABBY2 and YABBY5 genes, which were presumably involved in the evolutionary divergency of the pistil stalk and stamen filament morphology [15, 16]. Other

YABBY gene families, INNER NO OUTER (INO) and CRABS CLAW (CRC), apparently developed in parallel with the evolution of the carpel and ovule during leaf-like reproductive sporophyll modification [11, 17].

In dicots and monocots, YABBY genes play similar roles in leaf and leaf-like organs development, specifying their abaxial-adaxial asymmetry and lamina growth, as well as leaf boundaries [4, 10, 18]. Additionally, YABBY genes are involved in the formation of such flower organs as nectaries, carpels, etc. [19–21]. To date, the functions of certain YABBY proteins have only been described in the model plant *Arabidopsis thaliana*. Thus, it has been shown that YABBY1 (syn. FILAMENTOUS FLOWER, FIL), YABBY3, and YABBY5, along with other components of the transcription complex, support the identity of abaxial leaf surface cells and are also involved in the initiation of embryonic shoot apical meristem and its postembryonal maintenance [22]. Activation of a certain YABBY gene expression in the nectaries and carpels involves MADS-domain proteins [23]. In turn, YABBY1, togeth-

Table 1. The cultivated and wild tomato species used in the present study

Species/subspecies/cultivar	VIR Ref. No	Crossing system	Color of the ripe fruit
<i>S. cheesmaniae</i> (Riley) Fosberg	3969	self-compatible	Red
<i>S. galapagense</i> Darwin & Peralta	3970	self-compatible	Red
<i>S. lycopersicum</i> var. <i>humboldtii</i> (Willd.) Dunal	2912	self-compatible	Red
<i>S. lycopersicum</i> L., cv. Silvestre recordo	1580	self-compatible	Red
<i>S. pimpinellifolium</i> var. <i>racemigerum</i> (Lange) Brezhnev	1018	self-compatible	Red
<i>S. chmielewskii</i> (Rick, Kesicki, Fobes & Holle) Spooner, Anderson & Jansen	13725	self-compatible	Green
<i>S. neorickii</i> Spooner, Anderson & Jansen	5033	self-compatible	Green
<i>S. arcanum</i> Peralta	13958	self-incompatible	Green
<i>S. chilense</i> (Dunal) Reiche	4300	self-incompatible	Green
<i>S. corneliomulleri</i> Macbr.	4367	self-incompatible	Green
<i>S. habrochaites</i> Knapp & Spooner	13964	self-incompatible	Green
<i>S. peruvianum</i> L.	4361	self-incompatible	Green
<i>S. peruvianum</i> var. <i>dentatum</i> (Dunal) Dunal	3966	self-incompatible	Green

er with other transcription factors, controls the spatial activity of MADS-box genes and, thus, is involved in floral organ primordia initiation in the correct position and number, determining the corresponding cell's fate [24–26].

YABBY genes encode small proteins (180–250 amino acid residues) containing two conserved domains [27, 28]. The N-terminal part of the protein includes the Cys2Cys2-zinc-finger motif, and the C-terminus includes the YABBY domain.

In plant genomes, the YABBY genes number differs. In *A. thaliana*, six YABBY genes were found; four of them (YABBY1, YABBY2, YABBY3, and YABBY5) are mainly expressed in leaves and leaf-like organs (cotyledons, sepals, petals, stamens, and carpels), while the other two (*CRC* and *INO*) are expressed in some parts of floral reproductive organs [10, 23, 27]. Eight genes were identified in rice *Oryza sativa*; moreover, each *OsYABBY2* and *OsYABBY7* has two alternatively spliced transcripts [29].

Nine YABBY genes (*YABBY1*, *YABBY2*, *YABBY3*, *YABBY5a*, *YABBY5b*, *CRCa*, *CRCb*, *FAS*, and *INO*) were identified in cultivated tomato (*Solanum lycopersicum*), which is one of the major vegetable crops [30, 31]. *S. lycopersicum*, along with 12 wild related species, comprises the Lycopersicon section of the *Solanum* genus [32]. Tomato species widely vary in their morphophysiological characteristics, including leaf and flower morphology. Depending on the mating system structure, tomatoes are divided into self-compatible and self-incompatible species. The latter are characterized by high polymorphism, large flowers, and exerted stigma [32]. It is known that the plant reproductive system, which depends on the flower's morphophysiology, as

well as the differences in the leaf structure, can result from the different activities of YABBY transcription factors [7]. First and foremost, this relates to YABBY1/YABBY3 proteins, which are expressed in almost all asymmetric aboveground plant organs.

The present study was focused on the identification of YABBY3-orthologous genes in wild tomato species and an evaluation of their polymorphism. To date, complete YABBY3 sequences are determined only for two tomato species: *S. lycopersicum* and *S. pennellii*, and YABBY3 expression patterns were characterized only in *S. lycopersicum* [31] and *S. pimpinellifolium* [30]. Therefore, the present results, based on an analysis of a large number of tomato species, will contribute to our knowledge of YABBY genes and their possible functions.

EXPERIMENTAL

A set of 13 accessions of 11 tomato species from the collection of the All-Russian Institute of Plant Genetic Resources n.a. N.I. Vavilov (VIR) was selected for this study. The analyzed species differed both in the mating system and fruit morphology (Table 1).

The plants were grown from seeds in a greenhouse (8/16 h night/day; 23/28°C night/day, light intensity 300–400 μmol/m²). Genomic DNA was isolated from leaves using ZR-96 Plant/Seed DNA Kit (Zymo research, Irvine, USA). Five weeks after planting in the greenhouse, as fruit formation started, tissue samples were collected simultaneously from each plant, including leaves, young buds, open flowers, and immature green fruits, at 9.00–12.00 a.m. The sampled material was immediately frozen and ground in liquid nitrogen. Total RNA was isolated using a RNeasy

Plant Mini Kit (QIAGEN, Hilden, Germany) and used for cDNA synthesis with a GoScript kit (Promega, Madison, USA).

Specific primers, sYB3F (5'-AATCAAATCAATCA-CAAAARCAG-3') and sYB3R (5'-CACATTAATTG-GTTAGACTTA-3'), were designed based on the complete *YABBY3* gene sequence of *S. lycopersicum* (GeneID: 101247051) and *S. pennellii* (GeneID: 107026918) for an amplification of the full-length copies of this gene in the examined species. Additional internal primers, sYB3ex2R (5'-ATTAGTGCAGTGTTCCACATC-3') and sYB3ex4R (5'-TTGATGAATCGGTTG-TAAGC-3'), were designed for sequencing. The genes were amplified using LongAmp® polymerase Hot Start Taq DNA Polymerase (USA) under the following conditions: initial denaturation (10 min, 94°C); 35 cycles of denaturation (30 sec, 94°C); annealing (30 sec, 58°C) and elongation (4 min, 65°C); and final elongation (10 min, 65°C). PCR fragments were purified using a QIAEX® II Gel Extraction kit (QIAGEN, Hilden, Germany), cloned into the plasmid vector pGEMT-easy (Promega, Madison, USA), and sequenced using the BigDye system and an Applied Biosystems 3730 DNA analyzer (Applied Biosystems, Waltham, United States; Core Facility "Bio-engineering").

The obtained sequences were aligned and analyzed using the MEGA 7.0 [33]. The comparative analysis was carried out using known *YABBY3* complete sequences of two tomato species, *S. lycopersicum* cv. Heinz (GeneID:101247051) and *S. pennellii* (GeneID: 107026918), potato *S. tuberosum* (GeneID: 102577797), and *A. thaliana* (GeneID : 827 914). The positions of nucleotide and amino acid substitutions were determined in comparison with the *S. lycopersicum* cv. Heinz *YABBY3* (GeneID: 101247051). The structural domains of *YABBY3* orthologs were determined using NCBI-CDD (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and published data [27, 28]. Known sequences of *YABBY* genes, cDNAs, and the proteins of *S. lycopersicum* (*SIYABBY1* (XM_004229745), *SIYABBY2* (XM_004241308), *SIYABBY3* (XM_004245689), *SIYABBY5a* (XM_004242730), *SIYABBY5b* (XM_004251674), *SIFAS* (NM_001247461), *SLINO* (XM_004239291), *SLCRCa* (XM_004238984), *SLCRCb* (XM_004228801)), and *A. thaliana* (*AtYABBY1* (AF136538), *AtYABBY2* (AF136539), *AtYABBY3* (AF136540), *AtYABBY5* (NM_179750), *AtINO* (AF195047), *AtCRC* (AF132606)) were subjected to a phylogenetic analysis performed using the MEGA 7.0 Maximum Likelihood method (ML), preassigned by the Modeltest program. The possible effects of the amino acid substitutions on the protein structure and function were assessed using the Grantham matrix [34] and PROVEAN [35]. The three-dimensional protein structure was analyzed us-

ing the Phyre2 program [36] and visualized by Chimera 1.11.2 (<http://www.cgl.ucsf.edu/chimera/>).

YABBY3-orthologous genes expression was determined in young leaves, young buds, open flowers, and green immature fruits by quantitative real-time PCR (qRT-PCR) using the Reaction mixture for qRT-PCR in the presence of SYBR GreenI and a ROX kit (Syntol, Moscow, Russia) on a CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, USA). qRT-PCR was carried out using a gene-specific primer pair: tY3rt1F (5'-GTCACACTTACTTCTCTCCTTCAC-3') and tY3rtR (5'-CAGGAGGTCTGTTAACAACGG-3'). The reactions were carried out in two biological and three technical replicates under the following conditions: 95°C – 5 min; 40 cycles (95°C – 15 sec, 62°C – 50 sec). The relative expression level was assessed using the *CAC* gene as a reference [37]. The statistical analysis was performed using GraphPad Prism v. 7.02, including the assessment of the statistical significance of the expression differences in various organs of each analyzed tomato species using the unpaired t-test with Welch's correction (Table 3).

RESULTS AND DISCUSSION

Complete sequences of the *YABBY3*-orthologous genes were determined in 13 accessions of 11 tomato species (*Solanum* section *Lycopersicon*). The comparative analysis of these sequences showed that they are highly homologous (95–99% similarity) to the known tomato *YABBY3* gene (ID: 101247051). The total length of the gene varied from 2622 bp in *S. neorickii* to 2713 bp in *S. cheesmaniae*. The genes were composed of seven exons and six introns (Table 2) and included sequences that encoded the conserved HMG-like *YABBY* (125–176 aa) and the Cys2Cys2-zinc-finger (18–62 aa) domains (Fig. 1).

In the 9 analyzed accessions, including all red-fruited and three green-fruited (*S. chmielewskii*, *S. chilense*, and *S. habrochaites*) species, *YABBY3* cDNA was 651 bp (Table 2). In *S. neorickii*, cDNA was 654 bp due to TCA duplication in the second exon (N66_H67insH in amino acid sequence). In *S. arcanum*, *S. corneliomulleri*, *S. peruvianum*, and *S. peruvianum* var. *dentatum*, it was 660 bp due to 9 bp insertion in the first exon (P17_S18insPPP). In *S. pennellii*, which is known to be the most ancient species [32], cDNA of 645 bp was due to 6 bp deletion in the second exon (H67del, H68del). Accordingly, the length of the *YABBY3* orthologs was 217 aa (*S. neorickii*), 219 aa (*S. arcanum*, *S. corneliomulleri*, *S. peruvianum*, and *S. peruvianum* var. *dentatum*), and 216 aa (other accessions). Interestingly, among the previously described conserved *YABBY1/3*-characteristic motifs, *Solanum YABBY3* orthologs included the clade-specific motifs FIL-A, -D,

Table 2. Characteristics of the exon-intron structure of the YABBY3 gene in the examined tomato accessions

Species/subspecies/cultivar	NCBI number	Exon-intron structure of YABBY3												Total length, bp	cDNA, nt	Protein, aa residues	
		Exon I	Intron I	Exon II	Intron II	Exon III	Intron III	Exon IV	Intron IV	Exon V	Intron V	Exon VI	Intron VI				Exon VII
<i>S. cheesmaniae</i>	KY952537	102	546	150	223	127	314	49	374	76	425	75	180	72	2,713	651	216
<i>S. galapagense</i>	KY952538	102	531	150	222	127	316	49	373	76	426	75	180	72	2,699	651	216
<i>S. lycopersicum</i> cv. Heinz *	ID:101247051	102	536	150	222	127	316	49	373	76	426	75	179	72	2,703	651	216
<i>S. lycopersicum</i> var. <i>humboldtii</i>	KY952544	102	538	150	221	127	313	49	373	76	427	75	180	72	2,703	651	216
<i>S. lycopersicum</i> cv. <i>Silvestre recordo</i>	KY952543	102	537	150	222	127	316	49	372	76	426	75	180	72	2,704	651	216
<i>S. pimpinellifolium</i> var. <i>racemigerum</i>	KY952549	102	537	150	222	127	316	49	373	76	426	75	180	72	2,705	651	216
<i>S. chmielewskii</i>	KY952540	102	537	150	222	127	314	49	371	76	426	75	180	72	2,701	651	216
<i>S. neorickii</i>	KY952545	102	469	153	218	127	313	49	376	76	426	75	166	72	2,622	654	217
<i>S. arcanum</i>	KY952547	111	530	150	228	127	316	49	374	76	426	75	166	72	2,700	660	219
<i>S. chilense</i>	KY952539	102	533	150	224	127	311	49	374	76	426	75	165	72	2,684	651	216
<i>S. corneliomulleri</i>	KY952541	111	535	150	227	127	314	49	361	76	425	75	166	72	2,688	660	219
<i>S. habrochaites</i>	KY952542	102	545	150	222	127	274	49	361	76	421	75	178	72	2,652	651	216
<i>S. pennellii</i> *	ID:107026918	102	522	144	222	127	312	49	372	76	414	75	179	72	2,666	645	214
<i>S. peruvianum</i>	KY952546	111	539	150	228	127	316	49	365	76	427	75	169	72	2,704	660	219
<i>S. peruvianum</i> var. <i>dentatum</i>	KY952548	111	485	150	221	127	312	49	374	76	426	75	180	72	2,658	660	219
<i>S. tuberosum</i> *	ID:102577797	114	552	147	271	127	314	49	398	76	417	75	175	72	2,787	660	219
<i>A. thaliana</i> *	ID:827914	102	97	138	101	151	93	49	119	76	136	99	440	81	1,682	696	231

* Sequences from the NCBI database.

-E, and -G, but no FIL-B and -C, which are usually localized in the inter-domain region [12] (Fig. 1).

When compared with the previously characterized *S. lycopersicum* cv. Heinz YABBY3 (ID: 101247051), in the YABBY3 genes of the analyzed accessions, 317 variable sites, mostly localized in introns, were revealed. In the exons, 24 substitutions were detected, and 8 of them were nonsynonymous. Substitutions detected in cDNA were localized mainly in the sequence encoding the inter-domain region and at the 3'-terminus. In the region encoding the zinc-finger domain, only one substitution was detected: A59G transition in *S. galapagense*, which leads to a glutamine on arginine substitution, Q20R (Fig. 1). The sequence encoding the YABBY-domain revealed five nucleotide substitutions, and only one of them, A434G transition in *S. peruvianum* var. *dentatum* (3966), leads to a glutamic acid on glycine substitution, E145G (Fig. 1).

In YABBY3 proteins, 4 out of 11 aa substitutions (S64C, Y76C, D116G, and E145G) (Fig. 1) are considered to be radical (physicochemical distance according to Grantham's matrix <57.9). At the same time, an assessment using PROVEAN, generalizing known algorithms for a charge of aa substitutions and indels, revealed only one radical substitution (E145G in the *S. peruvianum* var. *dentatum* YABBY-domain), whereas the other substitutions, deletions, and insertions were rated as neutral. The possible effect of substitutions on the protein function needs further experimental analysis.

Modeling (Phyre2) of the YABBY3 three-dimensional structures showed a disordered organization of more than 60% of the sequence, while 29% were predicted with a confidence of more than 90% based on the known HMG-like protein structures (PDB: d1qrva, d1k99a etc.). The reliably predicted sequence was represented by a HMG-like YABBY domain [10] consisting of two α -helices connected by a loop (helix-loop-helix) (Fig. 2). The HMG-domain presumably binds to the DNA minor groove and bends the double helix at that point [38].

The phylogenetic analysis showed that all known *S. lycopersicum* YABBY genes are clustered with the corresponding *A. thaliana* orthologs (Fig. 3). On the cDNA-based dendrogram, YABBY genes formed four sub-clusters: YAB1/3 (YABBY1- and YABBY3-like genes); YAB2/5 (YABBY2-, YABBY5-, and FAS-like genes); CRC (CRC-like genes); and INO (INO-like genes) (Fig. 3A). The clusters resulting from the analysis of the amino acid sequences (Fig. 3B) were similar to those described above, except for YABBY2 and YABBY5, which formed separate sub-clusters corresponding to the previously proposed classification of the YABBY family into five subfamilies [10,

Table 3. The ANOVA analysis of YABBY3 gene expression in tomato species using Welch's t-test.

<i>S. lycopersicum</i> cv. <i>Silvestre recordo</i>			
	Leaf	Bud	Flower
Bud	0.0012		
Flower	0.6189	0.0007	
Fruit	<0.0001	<0.0001	<0.0001
<i>S. chmielewskii</i>			
	Leaf	Bud	Flower
Bud	0.0242		
Flower	0.1117	0.5025	
Fruit	<0.0001	<0.0001	<0.0001
<i>S. peruvianum</i> var. <i>dentatum</i>			
	Leaf	Bud	Flower
Bud	<0.0001		
Flower	0.1014	<0.0001	
Fruit	<0.0001	0.3049	<0.0001
<i>S. habrochaites</i>			
	Leaf	Bud	Flower
Bud	<0.0001		
Flower	<0.0001	<0.0001	
Fruit	<0.0001	<0.0001	<0.0001

* *p*-values <0.05 are considered as significant.

23]. The phylogenetic analysis based on the YABBY3 genomic sequences clustered the analyzed tomato accessions into two groups with a branch of the most ancient *S. pennellii* and potato *S. tuberosum* (Fig. 4). The results generally agreed with the tomato division into green-fruited and red-fruited, as well as self-compatible and self-incompatible, groups. At the same time, two self-compatible green-fruited species, *S. chmielewskii* and *S. neorickii*, fell into opposing clusters, which apparently corresponds to an evolutionary boundary point where red-fruited self-compatible species originated from green-fruited self-incompatible ones.

The YABBY genes expression in angiosperms suggests that YABBY1/3 genes preserved their ancient expression pattern [12], transcribing in the abaxial portion of the primordia of all aboveground lateral organs (except for ovules) [25, 41]. This is confirmed by our data on YABBY3 expression in the vegetative and reproductive organs of *S. chmielewskii*, *S. lycopersicum* cv. *Silvestre recordo*, *S. habrochaites*, and *S. peruvianum* var. *dentatum*. In *S. habrochaites*, gene expression in

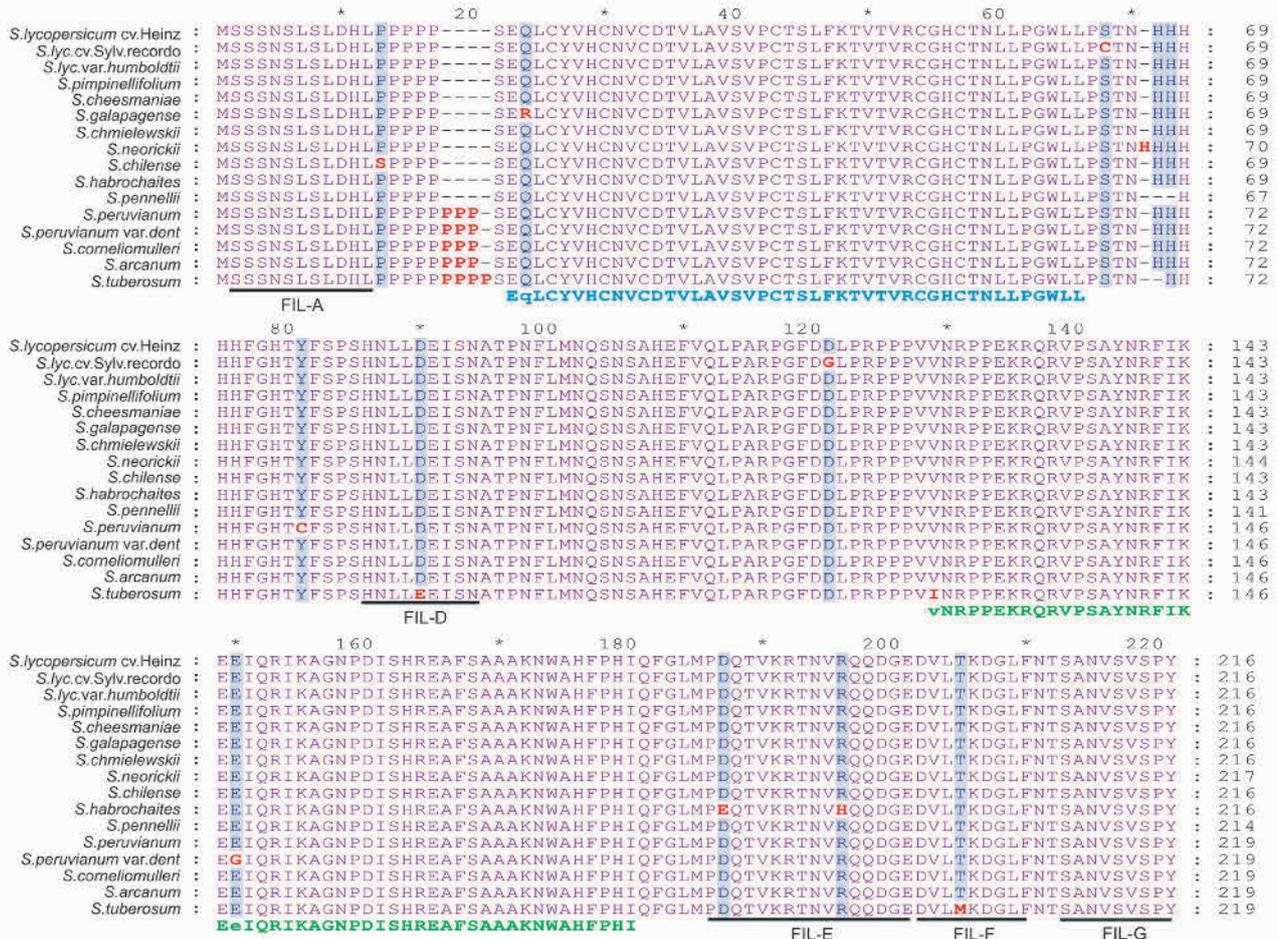


Fig. 1. Alignment of YABBY3 amino acid sequences from accessions of tomato and potato (*S. tuberosum*) species. The Zinc-finger and YABBY domains are indicated by blue and green letters, respectively, under alignment. Indels and substitutions are highlighted in red. Conserved motifs specific to YABBY1/YABBY3 clade are underlined and named

leaves is somewhat higher than that in flowers, while the other three species have no statistically significant differences in YABBY3 expression levels in leaves and flowers (Fig. 5, Tab. 3). At the same time, almost no YABBY3 expression was detected in the fruits of the studied species, except for *S. peruvianum* var. *dentatum* (Fig. 5). These four species were selected for expression analysis, since they belong to four groups that are evolutionarily distant from each other. *S. lycopersicum* is a red-fruited, self-compatible species of relatively recent origin; *S. chmielewskii* is a green-fruited, but self-compatible, species, and its position on the evolutionary tree is between red-fruited self-compatible and green-fruited self-incompatible species; *S. peruvianum* is a representative of the green-fruited self-incompatible species; and, finally, *S. habrochaites* (green-fruited, self-incompatible) is considered as one of the most

ancient tomato species [32]. The YABBY3 expression pattern in *S. peruvianum* var. *dentatum* is somewhat different from that in other analyzed accessions, although the reason for the low-level expression in buds is not fully understood (Fig. 5). In the analyzed organs of *S. habrochaites*, the YABBY3 expression dynamics is similar, but the transcription level is almost twice lower than that in *S. lycopersicum* and *S. chmielewskii*. In general, the identified YABBY3 expression patterns in *S. lycopersicum*, *S. chmielewskii*, and *S. habrochaites* were similar to those in *S. pimpinellifolium*, where in the YABBY3 expression level is maximal in young buds and decreases along with flower-to-fruit development [30].

It has been shown that, in *A. thaliana*, both the YABBY3 constitutive expression and its knockout lead to an abnormal development of leaves and flowers due

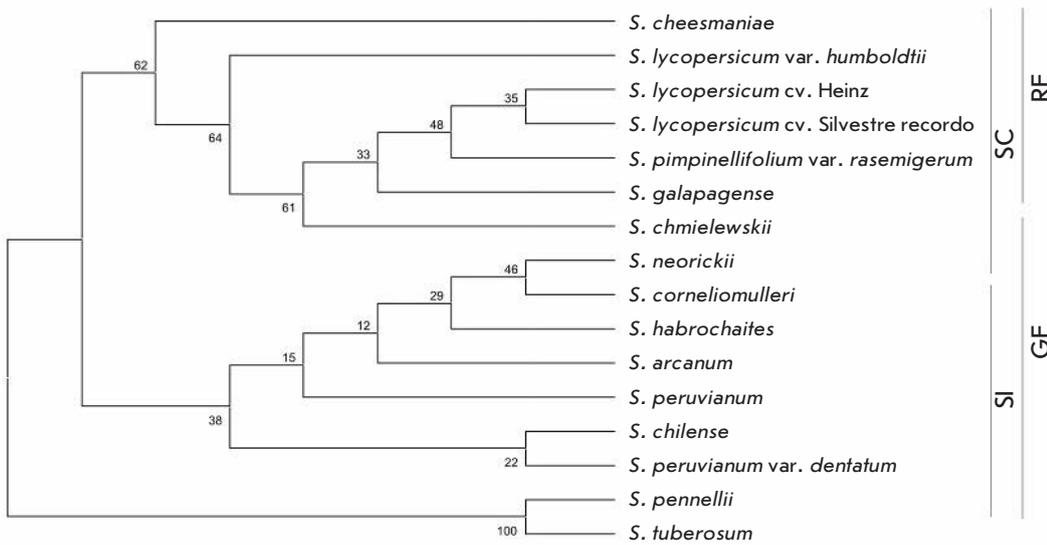


Fig. 4. Phylogenetic tree based on YABBY3 genomic sequences of the accessions of cultivated and wild tomato species and rooted with *S. tuberosum* YABBY3 (MEGA7.0, ML method, model HKY + G+I). RF – red-fruited accessions; GF – green-fruited accessions; SC – self-compatible accessions; SI – self-incompatible accessions

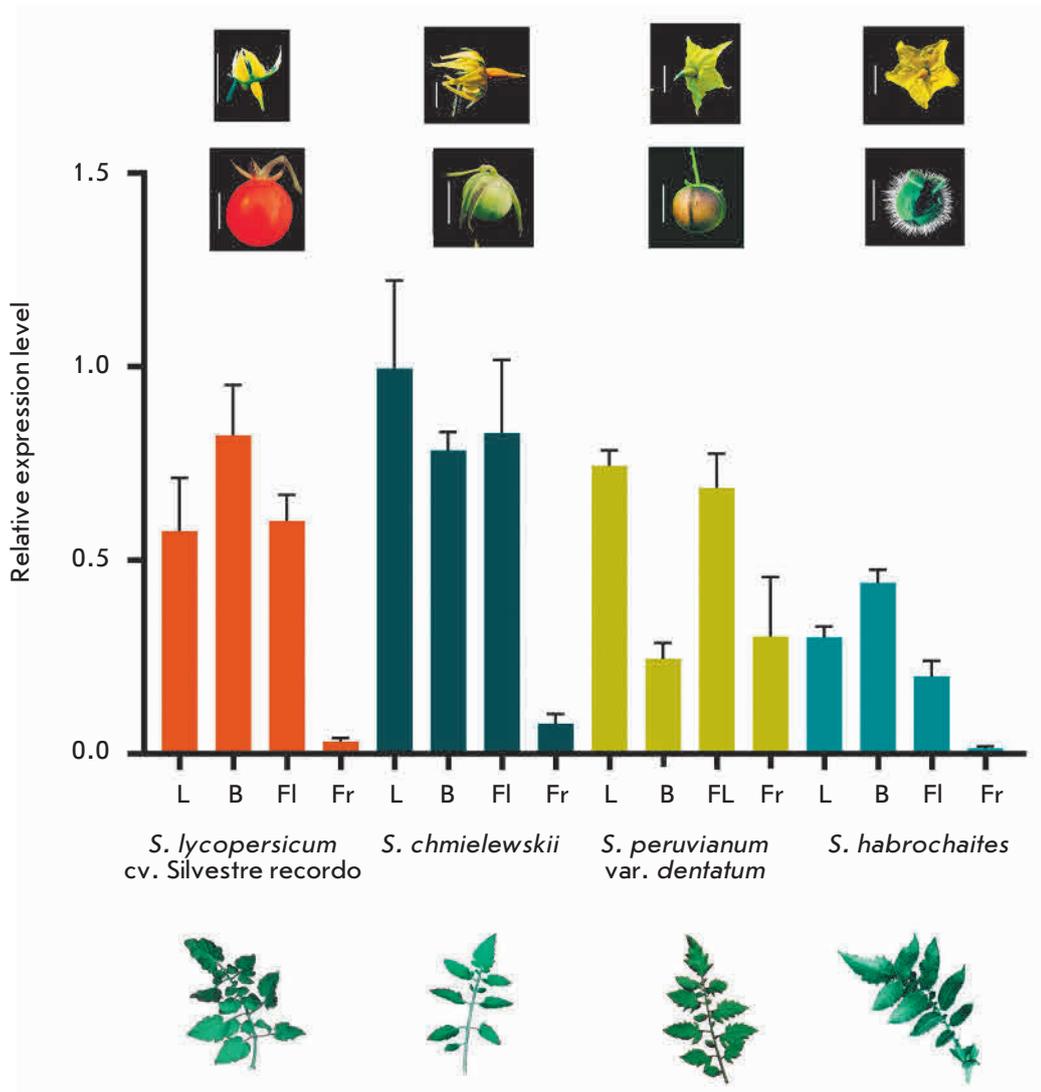


Fig. 5. Relative YABBY3 expression in the leaves (L), young buds (B), opened flowers (FI), and green immature fruits (Fr) of four tomato accessions

to the lack of polar differentiation in the organs [18]. The variability of this gene expression level can also affect the organ structure and morphophysiology; in particular, the leaves, flowers, and fruits of the analyzed tomato accessions. Significant levels of gene expression in *S. peruvianum* var. *dentatum* fruits may be indicative of a possible preservation of abaxial tissue identity in the fruit skin.

CONCLUSION

In this study, *YABBY3*-orthologous genes were detected in 13 accessions of cultivated and wild tomato species. These genes encode transcription factors that play a key role in determining the abaxial-adaxial asymme-

try of all aboveground plant lateral organs. The structure of *YABBY3* genes and the encoded proteins is similar to that of the previously characterized members of the *YABBY* family. A phylogenetic and expression analysis confirmed that the identified genes belong to the *YABBY1/3* subfamily and may have conserved functions in different tomato species. ●

This work was supported by the Russian Science Foundation (grant No. 16-16-10022), and was done with the use of the experimental climate control facility and Core Facility “Bioengineering” (Federal Research Center of Biotechnology, Russian Academy of Sciences).

REFERENCES

- Albert E.V., Ezhova T.A. // Russian Journal of Genetics. 2013. V. 49. № 2. P. 127–140.
- Lutova L.A., Dodueva I.E., Lebedeva M.A., Tvorogova V.E. // Russian Journal of Genetics. 2015. V. 51. № 5. P. 449–466.
- Castrillo G., Turck F., Leveugle M., Lecharny A., Carbone-ro P., Coupland G., Paz-Ares J., Oñate-Sánchez L. // PLoS One. 2011. V. 6. № 6. e21524.
- Yang C., Ma Y., Li J. // J. Exp. Bot. 2016. V. 67. № 18. P. 5545–5556.
- Gramzow L., Ritz M.S., Theissen G. // Trends Genet. 2010. V. 26. № 4. P. 149–153.
- Melzer R., Theissen G. // Methods Mol. Biol. 2011. V. 754. P. 3–18.
- Sarojam R., Sapp P.J., Goldshmidt A., Efroni I., Floyd S.K., Eshed Y., Bowman J.L. // Plant Cell. 2010. V. 22. № 7. P. 2113–2130.
- Floyd S.K., Bowman J.L. // Int. J. Plant Sci. 2007. V. 168. № 1. P. 1–35.
- Rensing S.A., Lang D., Zimmer A.D., Terry A., Salamov A., Shapiro H., Nishiyama T., Perroud P.F., Lindquist E.A., Kamisugi Y., et al. // Science. 2008. V. 319. № 5859. P. 64–69.
- Finet C., Floyd S.K., Conway S.J., Zhong B., Scutt C.P., Bowman J.L. // Evol. Dev. 2016. V. 18. № 2. P. 116–126.
- Yamada T., Yokota S., Hirayama Y., Imaichi R., Kato M., Gasser C.S. // Plant J. 2011. V. 67. № 1. P. 26–36.
- Bartholmes C., Hidalgo O., Gleissberg S. // Plant Biol (Stuttg). 2012. V. 14. № 1. P. 11–23.
- Meyerowitz E.M. // Cell. 1997. V. 88. № 3. P. 299–308.
- Bowman J.L., Eshed Y., Baum S.F. // Trends Genet. 2002. V. 18. № 3. P. 134–141.
- de Almeida A.M.R., Yockteng R., Schnable J., Alvarez-Buylla E.R., Freeling M., Specht C.D. // Sci. Rep. 2014. V. 4. Article number 6194.
- Morioka K., Yockteng R., Almeida A.M., Specht C.D. // Front Plant Sci. 2015. V. 6. Article 1106.
- Kelley D.R., Skinner D.J., Gasser C.S. // Plant J. 2009. V. 57. № 6. P. 1054–1064.
- Siegfried K.R., Eshed Y., Baum S.F., Otsuga D., Drews G.N., Bowman J.L. // Development. 1999. № 126. P. 4117–4128.
- Villanueva J.M., Broadhvest J., Hauser B.A., Meister R.J., Schneitz K., Gasser C.S. // Genes Dev. 1999. V. 13. № 23. P. 3160–3169.
- Fourquin C., Vinauger-Douard M., Fogliani B., Dumas C., Scutt C.P. // Proc. Natl. Acad. Sci. USA. 2005. V. 102. № 12. P. 4649–4654.
- Lee J.Y., Baum S.F., Alvarez J., Patel A., Chitwood D.H., Bowman J.L. // Plant Cell. 2005. V. 17. № 1. P. 25–36.
- Stahle M.L., Kuehlich J., Staron L., von Arnim A.G., Golz J.F. // Plant Cell. 2009. V. 21. № 10. P. 3105–3118.
- Lee J.Y., Baum S.F., Oh S.H., Jiang C.Z., Chen J.C., Bowman J.L. // Development. 2005. № 132. P. 5021–5032.
- Sawa S., Ito T., Shimura Y., Okada K. // Plant Cell. 1999. V. 11. № 1. P. 69–86.
- Sawa S., Watanabe K., Goto K., Kanaya E., Morita E.H., Okada K. // Genes Dev. 1999. V. 13. № 9. P. 1079–1088.
- Chen Q., Atkinson A., Otsuga D., Christensen T., Reynolds L., Drews G.N. // Development. 1999. № 126. P. 2715–2726.
- Bowman J.L. // Curr. Opin. Plant Biol. 2000. V. 3. P. 17–22.
- Kanaya E., Nakajima N., Okada K. // J. Biol. Chem. 2002. V. 277. № 14. P. 11957–11964.
- Toriba T., Harada K., Takamura A., Nakamura H., Ichikawa H., Suzaki T., Hirano H.Y. // Mol. Genet. Genomics. 2007. V. 277. № 5. P. 457–468.
- Huang Z., van Houten J., Gonzalez G., Xiao H., van der Knaap E. // Mol. Genet. Genomics. 2013. V. 288. № 3–4. P. 111–129.
- Han H.Q., Liu Y., Jiang M.M., Ge H.Y., Chen H.Y. // Genet. Mol. Res. 2015. V. 14. № 2. P. 7079–7091.
- Peralta I.E., Spooner D.M., Knapp S. // Systematic Botany Monographs Am. Soc. Plant Taxonomists, USA. 2008. V. 84. 186 p.
- Kumar S., Stecher G., Tamura K. // Mol. Biol. Evol. 2016. V. 33. № 7. P. 1870–1874.
- Grantham R. // Science. 1974. V. 185. P. 862–864.
- Choi Y., Sims G.E., Murphy S., Miller J.R., Chan A.P. // PLoS One. 2012. V. 7. № 10. e46688.
- Kelley L.A., Mezulis S., Yates C.M., Wass M.N., Sternberg M.J. // Nat. Protoc. 2015. V. 10. № 6. P. 845–858.
- Expósito-Rodríguez M., Borges A.A., Borges-Pérez A., Pérez J.A. // BMC Plant Biol. 2008. V. 8. № 131. P. 1–12.
- Bowman J.L., Smyth D.R. // Development. 1999. № 126. P. 2387–2396.
- Hasegawa M., Kishino H., Yano T. // J. Mol. Evol. 1985. V. 22. № 2. P. 160–174.
- Dayhoff M.O., Schwartz R.M., Orcutt B.C. // Atlas Protein Sequence and Structure. 1978. V. 5. P. 345–352.
- Golz J.F., Roccaro M., Kuzoff R., Hudson A. // Development. 2004. № 131. P. 3661–3670.

CaMKII Is Involved in the Choline-Induced Downregulation of Acetylcholine Release in Mouse Motor Synapses

A. E. Gaydukov*, O. P. Balezina

Department of Human and Animal Physiology, Biological Faculty, Lomonosov Moscow State University, Leninskie Gory 1, bldg. 12, Moscow, 119234, Russia

*E-mail: gaydukov@gmail.com

Received: July 07, 2017; in final form October 03, 2017

Copyright © 2017 Park-media, Ltd. 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 We investigated the involvement of calcium-dependent enzymes, protein kinase C (PKC) and calcium-calmodulin-dependent protein kinase II (CaMKII) in the signaling pathway triggered by the activation of presynaptic alpha7-type nicotinic acetylcholine receptors by exogenous choline, leading to downregulation of the evoked acetylcholine (ACh) release in mouse motor synapses. Blockade of PKC with chelerythrine neither changed the evoked release of ACh by itself nor prevented the inhibitory effect of choline. The CaMKII blocker KN-62 did not affect synaptic activity but fully prevented the choline-induced downregulation of ACh release.

KEYWORDS calcium-calmodulin-dependent protein kinase II, neuromuscular synapse, alpha7-nicotinic acetylcholine receptors, choline.

ABBREVIATIONS ACh – acetylcholine; alpha7-nAChR – nicotinic acetylcholine receptor of alpha7-type; MEPP – miniature endplate potential; EPP – endplate potential; PKC – protein kinase C; CaMKII – calcium-calmodulin-dependent protein kinase II.

INTRODUCTION

Choline comes from the products of the hydrolysis of acetylcholine (ACh) neurotransmitter by acetylcholinesterase in cholinergic synapses. Along with choline reuptake into the nerve terminals where it is recycled to synthesize ACh, choline plays an important role in the auto-regulation of ACh release by the feedback mechanism. This mechanism is associated with the ability of choline to selectively activate presynaptic alpha7-type nicotinic acetylcholine receptors (alpha7-nAChR) [1]. These receptors are abundant in central and peripheral synapses. Alpha7-nAChR permit the influx of sodium and calcium ions into the cell upon activation by choline and other agonists, leading to membrane depolarization, and also trigger diverse intracellular signaling cascades with the involvement of enzymes and channels [2]. In addition, it has been recently established that an alpha7-nAChR molecule contains an amino acid cluster that enables a functional interaction between alpha7-nAChR and G-proteins. This broadens the potential functions of alpha7-nAChR both as rapidly desensitizing ionotropic receptors and as special metabotropic receptors that trigger long-term signaling with long-term effects [3]. Therefore, these ambiguous consequences of presynaptic alpha7-nAChR activation in various types of synapses pose an im-

portant challenge that remains poorly studied. We have established recently that choline (0.1 mM) downregulates the evoked ACh release in mouse neuromuscular synapses via Ca²⁺-dependent Ca²⁺ efflux from the store through ryanodine receptors and the activation of the SK-type K_{Ca} channels of terminals, resulting in downregulation of ACh release [4]. Meanwhile, it was unclear whether Ca²⁺-dependent enzymes, such as protein kinase C (PKC) and/or calcium-calmodulin-dependent protein kinase type II (CaMKII), are involved in this cascade. Therefore, the aim of this study was to assess the changes in choline-evoked ACh release in mouse motor synapses, together with the blockade of calmodulin and Ca²⁺-dependent enzymes, protein kinase C (PKC), and calcium-calmodulin-dependent protein kinase II (CaMKII).

MATERIALS AND METHODS

The experiments were conducted using isolated neuromuscular preparations from the diaphragm (*m. diaphragma – n. phrenicus*) of mature (P30) male mice of the 129/Sv strain provided by the Anokhin Institute of Normal Physiology, Russian Academy of Sciences (Moscow, Russia). A total of 16 mice were used. The mice were euthanized by quick decapitation. The mice were handled in accordance with Directive 86/609/

EEC that regulates the use of laboratory animals. The procedure was approved by the Bioethics Commission of the Department of Biology, Moscow State University. All the experiments were conducted at room temperature of 20–22°C. Dissection of the neuromuscular preparation of the left half of the diaphragm with the phrenic nerve was performed according to the earlier described standard protocol [4]. Miniature endplate potentials (MEPPs) and multiquantal endplate potentials (EPPs) upon stimulation of the phrenic nerve were recorded using intracellular glass microelectrodes filled with 2.5 M KCl (the resistance at the microelectrode tip was 15–20 MΩ). First, MEPPs were recorded for 100 s, followed by recording of the EPPs in each synapse. The phrenic nerve was then stimulated with short trains of stimuli (50 stimuli 0.1 ms long each, frequency of 50 Hz). Signals were recorded using the Neuroprobe Amplifier Model 1600 (A-M Systems) and recorded using an L-Card E-154 analog-to-digital converter (with Power-Graph interface) into the PC hard drive. The data were then processed using the MiniAnalysis software (Synaptosoft). Controls included MEPP and EPP recordings from 5 or more different synapses; next, the substances under study were added to the perfusion solution in a particular order. The synaptic activity was registered during 1–1.5 h. At least 3 neuromuscular preparations were used in each experimental series. Choline, chelerythrine (Sigma, USA), W-7, KN-62 (Enzo Life Sciences, USA) were used. The amplitude, time parameters of MEPPs and EPPs, the MEPP frequency, and the quantal content of EPPs were estimated (the latter was calculated as the ratio between the mean EPP amplitude corrected for non-linear summation [5] and the mean MEPP amplitude). The statistical significance of the difference between the sample groups was assessed using the Student's t-test and Mann-Whitney test. The significance level of the differences between two sample groups was 0.05 (n is the number of synapses studied).

RESULTS AND DISCUSSION

Similar to our previous study [4], we used a 100-μM choline concentration to assess the presynaptic action of choline. This concentration is close to the choline concentration in the synaptic cleft during the hydrolysis of ACh and slightly exceeds the EC_{50} for activating alpha7-nAChR [6].

Choline significantly changed neither the membrane potential of muscle fibers nor the spontaneous MEPP frequency. The mean amplitude of MEPPs in the presence of choline (1.08 ± 0.09 mV ($n = 17$)) also did not change significantly compared to that of the control (1.05 ± 0.08 mV ($n = 15$), $p > 0.05$). Short rhythmic stimulus trains (50 Hz, 1 s) led to characteristic changes in

the amplitude and quantal content of EPPs in the train. The short-term facilitation of the synaptic transmission was followed by a depression in the form of a decreased amplitude of EPPs compared to the first EPP in the train, continuing into a lower stable level of EPPs (and the quantal content) compared to the first EPP (Fig. 1). When pauses (at least 2 min long) were made between stimulus trains, the patterns of repeated trains were steadily reproduced in an individual synapse or other tested synapses. Application of choline reduced the EPP amplitude in the train because of the decay in the quantal content of EPPs. The quantal content of EPPs in the train in the presence of choline decreased significantly to 64–71% compared to the control ($p < 0.05$). In addition, the general pattern of the train remained unchanged (Fig. 1). The amplitude and the quantal content of EPPs decreased within 10–15 min after the administration of choline and remained at a lower level over the whole period during which choline was applied (for 45–60 min).

The effects of PKC chelerythrine blocker

The application of PKC blocker chelerythrine on the muscle (4 μM) for 30–40 min did not significantly change the behavior of the bursting synaptic activity: neither the quantal content of EPPs in the train nor the train pattern (the initial facilitation, subsequent depression, and a plateau) changed significantly in the presence of chelerythrine ($p > 0.05$). In addition, chelerythrine had no impact on the inhibitory effects of choline in terms of the EPPs quantal content during bursting synaptic activity (Fig. 2A). Therefore, (1) the Ca^{2+} -signals that enter the terminal upon choline-induced activation of alpha7-nAChR with subsequent release of the stored calcium and (2) the possible metabotropic signaling from alpha7-nAChR coupled with G_q -protein shown on other study objects [3] both do not activate PKC, and, thus PKC is not involved in the downregulation of ACh release in motor synapses. This agrees with our data and published citations showing that in motor terminals, PKC activation can be triggered by a calcium influx into the nerve endings via other Ca^{2+} -channels, more specifically, via L-type Ca^{2+} -channels, and also lead to the facilitation of ACh release [7].

The effects of calmodulin blocker W-7

Next, we studied the effects of choline subsequent to a preliminary inhibition of the regulatory activity of calmodulin using the W-7 calmodulin blocker (10 μM). The W-7 calmodulin blocker neither had a direct influence on synaptic transmission nor influenced significantly the downregulation effect of choline on the evoked ACh release. At the same time, choline-induced downregulation of ACh release in the presence of the

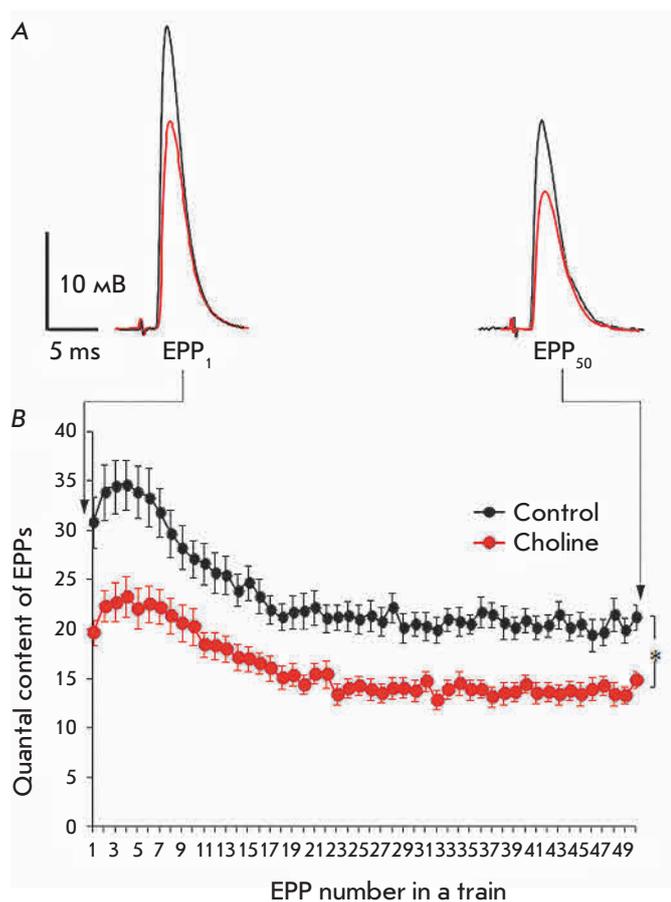


Fig. 1. Downregulation of evoked ACh release by exogenous choline ($100 \mu\text{M}$) during rhythmic synaptic activity with a frequency of 50 Hz (1 s): **A** – the averaged recordings of the first (EPP_1) and last (EPP_{50}) EPPs in the trains in the control (black) and in the presence of choline (red). **B** – changes in the quantal content of EPPs during short-term rhythmic trains with a frequency of 50 Hz in the control and after the addition of choline ($100 \mu\text{M}$). The Y axis shows the quantal content of EPPs; the X axis shows the number of EPPs in the short train. * $p < 0.05$ with respect to the control values

calmodulin blocker was weaker than when only choline was added (*Fig. 2B*).

Effects of CaMKII blocker KN-62

In the final series of the experiments, we studied the possible activation and involvement of CaMKII in the inhibitory effects of choline. KN-62 ($3 \mu\text{M}$), a selective CaMKII blocker, was used. Neither statistically significant increments in the MEPP amplitude nor changes in the quantal content of EPPs in short trains were revealed during KN-62 solution perfusion of neuromuscular preparations for 30–40 min. Thus, the amplitude

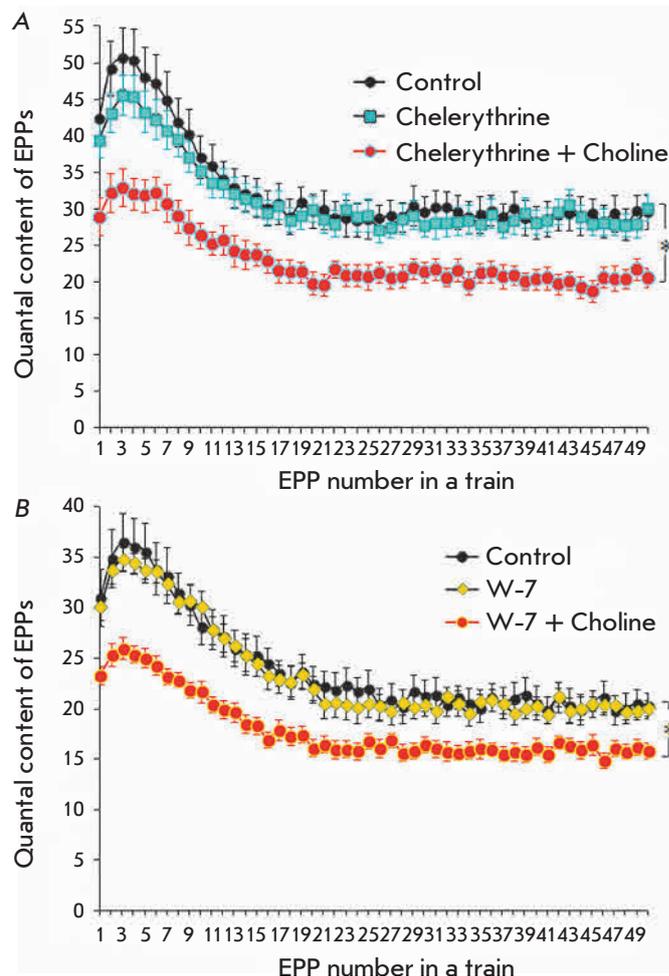


Fig. 2. Changes in the quantal content during short-term rhythmic trains of EPPs with a frequency of 50 Hz: **A** – in the control, after chelerythrine ($4 \mu\text{M}$) was applied and after choline ($100 \mu\text{M}$) was added in the presence of preapplied chelerythrine. **B** – in the control ($n = 17$), after application of W-7 ($10 \mu\text{M}$) ($n = 15$) and when choline ($100 \mu\text{M}$) was added in the presence of pre-applied W-7 ($n = 18$). The Y axis shows the quantal content of EPPs; the X axis shows the EPP number in a short train. * $p < 0.05$ with respect to the control values

of MEPPs was $0.91 \pm 0.05 \text{ mV}$ ($n = 20$) in the control; it was $0.85 \pm 0.04 \text{ mV}$ ($n = 23$, $p > 0.05$) in the presence of KN-62 and was $0.83 \pm 0.6 \text{ mV}$ ($n = 25$) in the presence of choline and KN-62. However, in motor synapses when choline was added in the presence of preapplied KN-62, there was no significant decline in the amplitude and the quantal content of EPPs in a train compared to the control (*Fig. 3*).

We had previously revealed a choline-induced downregulation of ACh release triggered by the activation of presynaptic $\alpha 7\text{-nAChR}$, which suggests that activation of CaMKII can be involved in the downreg-

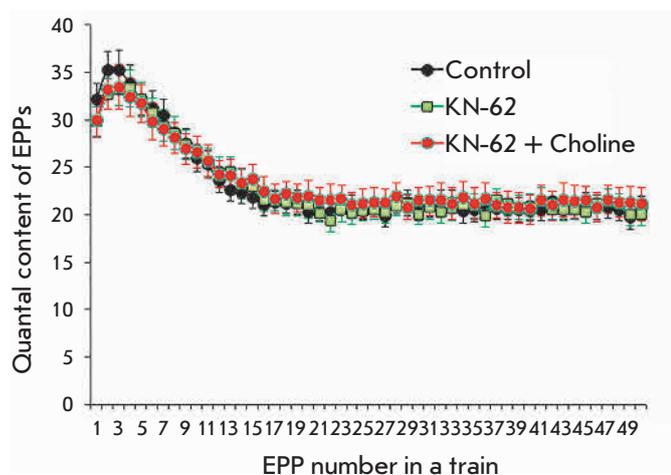


Fig. 3. Changes in the quantal content during short-term rhythmic trains of EPPs with a frequency of 50 Hz in the control, after KN-62 (3 μ M) was applied and when choline (100 μ M) was added in the presence of pre-applied KN-62. The Y axis shows the quantal content of EPPs, and the X axis shows the EPP number in a short train

ulation of neurotransmitter release, along with other processes [4].

External calcium influx [8] and intracellular calcium efflux [9] have been shown to activate presynaptic CaMKII in the terminals of central and peripheral synapses, and possible downregulating and upregulating influences of CaMKII activation on transmitter and cotransmitter release have been revealed [9, 10]. A modulation of the evoked transmitter release occurs upon the activation of $\alpha 7$ -nAChR by either endogenous or exogenous choline in CNS synapses. The generation of a calcium signal was recently described in response to the influence of choline on presynaptic $\alpha 7$ -nAChR in hippocampal synapses; the calcium signals increased the amplitude of excitatory postsynaptic potentials, but the effects did not cause CaMKII activation and persisted in the presence of the KN-62 blocker [11]. We had first shown previously that in mouse peripheral synapses, the

choline-induced activation of $\alpha 7$ -nAChR downregulates the evoked ACh transmitter release and that this downregulation can be fully prevented by the blocking of ryanodine receptors or SK-channels [4]. This study is an important supplementation to these concepts. CaMKII was also found to participate in the auto-regulation of ACh release that occurs with the involvement of choline and $\alpha 7$ -nAChR. With the revealed role of CaMKII in the auto-regulation of ACh release, we can add this kinase to the already described list of enzymes that play different roles in the signal transmission following $\alpha 7$ -nAChR activation in different types of cells [3, 11]. Therefore, it is necessary to take into account the possibility of CaMKII activation when studying the role of $\alpha 7$ -nAChR in the regulation of cellular processes.

We recently revealed CaMKII activation and its contribution to the enhancement of ACh release during calcium influx via L-type calcium channels, and this has been so far the only example of CaMKII involvement in the functions of neuromuscular synapses in rodents [12]. This study describes for the first time a fundamentally different way of CaMKII activation and participation in nerve end functions: i.e., the activation of $\alpha 7$ -nAChR is associated with the involvement of CaMKII in downregulating ACh release. The role of CaMKII molecules residing close to $\alpha 7$ -nAChR and intraterminal calcium stores can be to enhance and prolong the calcium signal, coupled with the function of ryanodine receptors, which is necessary for the activation of SK-type potassium channels.

Therefore, we have revealed for the first time a cascade of reactions in mouse motor nerve terminals that are triggered by a choline-induced activation of presynaptic $\alpha 7$ -nAChR that downregulates ACh release. This cascade has been shown to rely on calcium release from stores, calcium-activated SK-type K^+ -channels, and the activity of the Ca^{2+} -dependent enzyme CaMKII. ●

This study was supported by the Russian Foundation for Basic Research (grant No. 13-04-00413a).

REFERENCES

- Albuquerque E.X., Pereira E.F., Alkondon M., Rogers S.W. // *Physiol. Rev.* 2009. V. 89. № 1. P. 73–120.
- Cheng Q., Yakel J.L. // *Biochem. Pharmacol.* 2015. V. 97. № 4. P. 439–444.
- King J.R., Nordman J.C., Bridges S.P., Lin M.K., Kabbani N. // *J. Biol. Chem.* 2015. V. 290. № 33. P. 20060–20070.
- Gaydukov A.E., Bogacheva P.O., Tarasova E.O., Balezina O.P. // *Acta Naturae.* 2014. V. 6. № 4. P. 110–115.
- McLachlan E.M., Martin A.R. // *J. Physiol.* 1981. V. 311. P. 307–324.
- Papke R.L., Porter Papke J.K. // *Br. J. Pharmacol.* 2002. V. 137. № 1. P. 49–61.
- Gaydukov A.E., Marchenkova A.A., Balezina O.P. // *Bull. Exp. Biol. Med.* 2012. V. 153. № 4. P. 415–418.
- Zhong C., Talmage D.A., Role L.W. // *PLoS One.* 2013. V. 8. № 12. e82719.
- de Jong A.P., Verhage M. // *Curr. Opin. Neurobiol.* 2009. V. 19. № 3. P. 245–253.
- Shakiryanova D., Kloze M.K., Zhou Y., Gu T., Deitcher D.L., Atwood H.L., Hewes R.S., Levitan E.S. // *J. Neurosci.* 2007. V. 27. № 29. P. 7799–7806.
- Cheng Q., Yakel J.L. // *J. Neurosci.* 2014. V. 34. № 1. P. 123–144.
- Tarasova E.O., Gaydukov A.E., Balezina O.P. // *Neurochem. J.* 2015. V. 9. № 2. P. 101–107.

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 is published by the Park Media publishing house in both Russian and English.

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 the accompanying documents should be sent to the Editorial Board in electronic form:

- 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. Ab-

stract. 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_{pt} – 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 “x” 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 accept-

ed 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