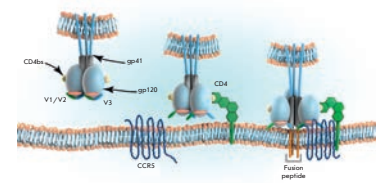


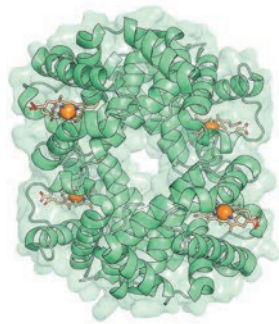
Antibodies Neutralizing a Wide Range of HIV-1 Isolates: A New Side of the Immune System

D. N. Shcherbakov, A. Yu. Bakulina, L. I. Karpenko, A. A. Il'ichev

Broadly neutralizing antibodies (bNAbs) differ from common antibodies by a high level of somatic mutations and unusually elongated variable loops, which enables them to bind to the conservative, but hard-to-access regions of Env HIV-1. This review describes anti-HIV-1 broadly neutralizing antibodies classified according to their interaction with vulnerability areas on the surface of viral glycoproteins.



Stages of interaction between the Env trimer and the cell membrane



Three-dimensional structure of human hemoglobin

A Study of Functional and Allosteric Sites in Protein Superfamilies

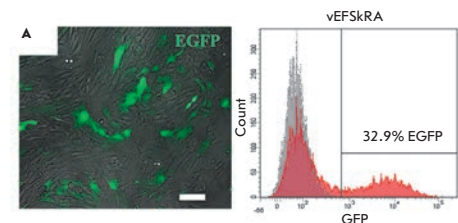
D. A. Suplatov, V. K. Shvyadas

Methods of computational biology, bioinformatics, and molecular modeling enable studying the structural organization features of functional and allosteric sites in proteins, searching for new sites, and investigating the molecular mechanisms of allostery. Discovery of new functional, allosteric, and regulatory sites and establishment of evolutionary relationships among them should help better understand the role of structure in functioning and regulating and provide new opportunities in drug development and effective biocatalyst design.

Induced Pluripotent Stem Cells of Vole Hybrids *Microtus levis* × *Microtus arvalis*: Conditions for Production and Maintenance

E. V. Grigor'eva, A. I. Shevchenko, S. P. Medvedev, N. A. Mazurok, A. I. Zhelezova, S. M. Zakiyan

We generated induced pluripotent stem cells (iPSCs) of hybrids produced by breeding of two species *Microtus levis* × *Microtus arvalis* from a group of common voles that are the object for studying the molecular genetic organization of the genome and the mechanisms of X chromosome inactivation. Maintenance of the undifferentiated state in iPSC lines of common voles is provided by activation of their pluripotency genes: *Nanog*, *Oct4*, *Sox2*, *Sall4*, and *Esrrb*.



Generation and characterization of iPSCs of hybrids of common voles *M. levis* × *M. arvalis*

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
Publisher: K.V. Kiselev

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)
K. Nierhaus (Berlin, Germany)
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: S.B. Nevskaya

Editor: N.Yu. Deeva

Designer: K.K. Oparin

Art and Layout: K. Shnaider

Copy Chief: Daniel M. Medjo

Address: 119234 Moscow, Russia, Leninskiye Gory, Nauchny
Park MGU, vlad.1, stroeniye 75G.
Phone/Fax: +7 (495) 727 38 60
E-mail: vera.knorre@gmail.com, mmorozova@strf.ru,
actanaturae@gmail.com

Reprinting is by permission only.

© ACTA NATURAE, 2015

Номер подписан в печать 20 октября 2015 г.

Тираж 200 экз. Цена свободная.

Отпечатано в типографии «МИГ ПРИНТ»

CONTENTS

FORUM

A. N. Petrov, N. G. Kurakova

**Assessment of the Volume of State Funding for
the Development of Biomedicine in Russia and
in the USA**4

REVIEWS

D. N. Shcherbakov, A. Y. Bakulina, L. I. Karpenko,
A. A. Ilyichev

**Broadly Neutralizing Antibodies against HIV-1
As a Novel Aspect of the Immune Response** .11

P. Pletnev, I. Osterman, P. Sergiev, A. Bogdanov,
O. Dontsova

**Survival guide: *Escherichia coli*
in the stationary phase**22

D. Suplatov, V. Švedas

**Study of Functional and Allosteric Sites in
Protein Superfamilies**34

T. V. Vinogradova, I. P. Chernov,
G. S. Monastyrskaya, L. G. Kondratyeva,
E. D. Sverdlov

**Cancer Stem Cells: Plasticity Works against
Therapy**46

RESEARCH ARTICLES

E. V. Grigor'eva, A. I. Shevchenko,
S. P. Medvedev, N. A. Mazurok,
A. I. Zhelezova, S. M. Zakian

**Induced Pluripotent Stem Cells of *Microtus levis*
x Microtus arvalis Vole Hybrids:
Conditions Necessary for Their Generation
and Self-Renewal**56

S. S. Efimova, O. S. Ostroumova
**Modifiers of the Dipole Potential
 of Lipid Bilayers**70

O. S. Petrakova, V. V. Ashapkin, V. Y.
 Shtratnikova, L. I. Kutueva, E. A. Vorotelyak,
 M. A. Borisov, V. V. Terskikh, I. G. Gvazava,
 A. V. Vasiliev
**Valproic Acid Increases the Hepatic
 Differentiation Potential of Salivary Gland Cells** 80

E. A. Sokolova, O. A. Stremovskiy, T.
 A. Zdobnova, I. V. Balalaeva, S. M. Deyev
**Recombinant Immunotoxin 4D5scFv-PE40 for
 Targeted Therapy of HER2-Positive Tumors.** .93

M. Tikhonov, N. B. Gasanov, P. Georgiev,
 O. Maksimenko
**A Model System in S2 Cells to Test the
 Functional Activities of Drosophila Insulators.**97

G. Y. Fisunov, D. V. Evsyutina,
 A. A. Arzamasov, I. O. Butenko, V. M. Govorun
**Profiling of *Mycoplasma gallisepticum*
 Ribosomes**107

S. N. Yakubitskiy, I. V. Kolosova,
 R. A. Maksyutov, S. N. Shchelkunov
Attenuation of Vaccinia Virus113

A. S. Kuznetsov, P. E. Volynsky, R. G. Efremov
**Role of the Lipid Environment in the
 Dimerization of Transmembrane Domains
 of Glycophorin A**122

N. V. Anufrieva, E. A. Morozova,
 V. V. Kulikova, N. P. Bazhulina, I. V. Manukhov,
 D. I. Degtev, E. Yu. Gnuchikh, A. N. Rodionov,
 G. B. Zavilgelsky, T. V. Demidkina
**Sulfoxides, Analogues of L-Methionine and
 L-Cysteine As Pro-Drugs against Gram-Positive
 and Gram-Negative Bacteria**128

S. S. Terekhov, I. V. Smirnov,
 O. G. Shamborant, T. V. Bobik,
 D. G. Ilyushin, A. N. Murashev, I. A. Dyachenko,
 V. A. Palikov, V. D. Knorre, A. A. Belogurov Jr.,
 N. A. Ponomarenko, E. S. Kuzina, D. D. Genkin,
 P. Masson, A. G. Gabibov
**Chemical Polysialylation and *In Vivo*
 Tetramerization Improve Pharmacokinetic
 Characteristics of Recombinant Human
 Butyrylcholinesterase-Based Bioscavengers** 136

SHORT REPORTS

D. A. Babkov, M. P. Paramonova,
 A. A. Ozerov, A. L. Khandazhinskaya,
 R. Snoeck, G. Andrei, M. S. Novikov
**2-(2,4-Dioxy-1,2,3,4-Tetrahydropyrimidin-
 1-yl)-N-(4-Phenoxyphenyl)-Acetamides As a
 Novel Class of Cytomegalovirus Replication
 Inhibitors**142

E. V. Konovalova, O. M. Lopacheva,
 I. A. Grivennikov, O. S. Lebedeva,
 E. B. Dashinimaev, L. G. Khaspekov,
 E. Yu. Fedotova, S. N. Illarionov
**Mutations in Parkinson's Disease-Associated
 PARK2 Gene Are Accompanied by Imbalance in
 Programmed Cell Death Systems**146

Guidelines for Authors..... 150

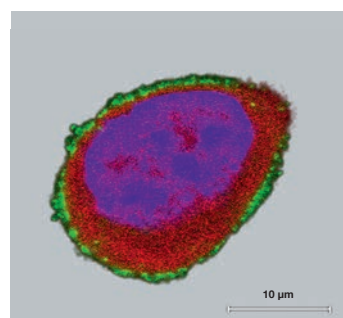


IMAGE ON THE COVER PAGE

Human ovarian adenocarcinoma cell expressing the red fluorescent protein (see an article by Sokolova *et al.*)

Assessment of the Volume of State Funding for the Development of Biomedicine in Russia and in the USA

A. N. Petrov*, N. G. Kurakova

«Directory of Scientific-Technological Programmes» Russian Ministry of Science and Education, Presnenskiy Val., 19/1, 123557, Moscow, Russia

*E-mail: petrov@fcntp.ru

ABSTRACT This article demonstrates that Russia's funding for research and development is less than 2.5 % of global funding, whereas the amount of financing of just three countries, the USA, China, and Japan amounts to 50%. It is argued that the inadequacy of Russia's domestic financing for the development of the science sector vis a vis that of developed countries allows the country to prioritize only a limited number of research fields in its scientific and technological development. We have compared and contrasted expenditures on research and development in biomedicine in the USA and Russia. It has been demonstrated that in 2014, basic funding for 27 research centers included in the US National Health Institutes' network exceeded the amount of financing for 104 Russian medical scientific and research institutes subordinated to the Russian Ministry of Health and Federal Agency of Scientific Organizations by 173 times. We have concluded that a substantial increase in state funding for fundamental, exploratory, and applied research in the field of biomedicine is required if life sciences are to be preserved as one of the priorities in the scientific-technological and social development of Russia. It is also necessary to eliminate all administrative and tax barriers that prevent active participation of domestic industrial entities in the co-financing of the development of Russian drugs and medical equipment.

KEYWORDS Biomedical scientific research and development, government funding, priority areas, scientific and technological development, Russia, choice criteria.

In 2014, basic funding for 27 research centers of the National Institute of Health of the USA amounted to 173 times the amount of funding for 104 medical research institutions subordinated to the Ministry of Health of the Russian Federation and the Federal Agency for Scientific Organizations (FANO). The contribution of the U.S. private sector to the development of technologies and products for the health care market was estimated at 92.6 billion dollars in 2014. In Russia, almost half (45%) of the budget of the Federal Target Program (FTP) "Pharma 2020" was allocated to the development of corporate sector R&D. The President of the Russian Federation has called for "full support" of the development of cutting-edge medical technologies.

HOW MUCH FUNDING DOES ONE NEED FOR IT? THIS ARTICLE OFFERS AN ANSWER TO THIS QUESTION

Current science and technology policy focuses on a radical increase in the efficiency of federal budget funding for civilian science and, above all, for the development of its priorities. For example, on June 25, 2015, at a meeting of the Council for Science and Education the President of the Russian Federation said that the current system of budget planning in the field of science and research was "highly opaque in the absence of any clear criteria for efficiency in the use of resources."

Simultaneously, the President set the goal of selecting several priority areas "to ensure their funding to the fullest extent," and named "cutting-edge medical technologies" as one of the key priorities [1].

The purpose of this study was to evaluate the volume of state sup-

port for the development of biomedical technologies within the "Life sciences" priorities in the Russian Federation and to compare it with what is being done in other countries.

ASSESSMENT OF THE SHARE OF RF DOMESTIC EXPENDITURE IN THE GLOBAL BUDGET FOR RESEARCH AND DEVELOPMENT

In 2014, the total global budget for research and development (R&D) was estimated at 1.6 trillion dollars [2]. The countries with the biggest R&D expenditures were the USA (31.1% of the global R&D budget), China (17.5%), Japan (10.2%), and the EU countries (21.7%), of which Germany alone contributed an estimated 5.7%. The share of these countries accounted for 78% of the global budget in 2014.

The same year, the share of Russian domestic R&D expenditure

was 2.5% of the global budget (40 out of 1,618 billion USD) [3]. According to the federal law “on the federal budget for 2015, and the planning period of 2016, and 2017” [4], 315.08 billion rubles will be allocated for R&D in the year 2016, which at the current exchange rate (60 rubles per US dollar) amounts to an estimate of a maximum of 5.3 billion USD of state budget funding for the domestic R&D sector in 2016.

In February 2015, during the discussion of the R&D budget for 2016, in response to an increase in Asian countries R&D budgets over the past 5 years, which has exceeded the corresponding rate in the U.S., the President of the United States proposed “putting an end to budget austerity” and planning a 6% increase in federal R&D funding for 2016 [5]. Therefore, the expected volume of U.S. state budget funding, allocated for the development of the civilian R&D sector, will exceed 62 billion USD (i.e. almost 12 times the corresponding expenditure in the budget of the Russian Federation.) It should be noted that the share of the federal budget allocated for civil R&D in 2015 is 1.7% in the U.S. and 2% in Russia.

Let us examine how much funding is considered to be sufficient for the development of medical technologies within the framework of such a priority as life sciences, according to federal budget managers in the United States and Russia.

THE AMOUNT OF FUNDING FOR BIOMEDICAL RESEARCH AND ITS INDIVIDUAL KEY AREAS IN RUSSIA AND THE USA

Our analysis of the planned amount of funding shows that biomedical research programs are given priority in the United States. Overall, the already substantial funding of the National Institute of Health (“NIH”) in the U.S. (which consists

of 27 research centers) to 30.2 billion USD in 2015 will be increased by 1 billion USD in 2016. In addition, the National Science Foundation, whose budget will also be increased to 7.72 billion USD (a 5.2% increase compared to 2015), will increase its funding for the “Understanding the Brain” program (one of the sub-projects of the BRAIN Initiative) by 35.2% compared to 2015.

The analogue of the NIH in the Russian Federation is a cluster of medical institutions, of which 50 belong to the FANO (the former Russian Academy of Medical Sciences) and 54 are within the jurisdiction of the Russian Ministry of Health. The consolidated budget for basic and project financing at these organizations is defined in the State Program for the Development of Health Care in the Russian Federation (sub-program 3) [6], and in 2015, it is set to 11.8 billion rubles, which at the current exchange rate of 60 rubles per U.S. dollar is equivalent to 200 million USD.

In 2016, the gap in funding for the two reference groups of institutions with a similar scope of research will only grow as the USA plans to increase funding for the National Institutes of Health, whereas the Russian Federation will cut funding for the civilian R&D sector in general and biomedical research in particular.

The gap in funding for certain key areas of biomedical research in the United States and Russia is equally huge.

For example, in 2016, the United States plans to increase funding for studies of the mechanisms of antibiotic resistance to 1.2 billion USD, which is an almost twofold increase for this area of priority in comparison with 2015. Financial support for the “Brain Research through Advancing Innovative Neurotechnologies” initiative will increase more than twofold in 2016, from 64 million USD in 2015 to 136 million USD

in 2016. The National Childhood Disease Research Program, which has a 2014-2016 annual budget of 1.2 billion USD, will get additional funding in 2016 (165 million USD) to assess the impact of the environment on children’s health [5]. A total of 215 million USD is pledged for the development of personalized medicine within the framework of the new “Precision Medicine Initiative,” which includes the development of a joint database of the health and genetic data of millions of volunteers. A total of 699 million USD are allocated to the new research program of natural foci of infections.

In the Russian Federation, competitive and project financing for fundamental, exploratory, and applied biomedical research in 2015 originated from several public funds: the FTP “Research and development in priority areas of science and technology complex of the Russian Federation for 2014-2020” (hereinafter, FTP “R&D”) and the FTP “Development of the Pharmaceutical and Medical Industry of the Russian Federation for the period up to 2020 and beyond” (hereinafter, FTP “Pharma-2020”).

The following table shows allocated and planned amounts of federal budget support for the development of fundamental, translational, and personalized medicine, as well as support for some priority areas in the field of life sciences in the United States and Russia in 2014-2016. It should be emphasized that the list of key managers of biomedical R&D budgets in the Russian Federation and the United States is not exhaustive due to the lack of data.

These data reveal the absolute incomparability of the amount of funding allocated by the United States and Russia for fundamental, exploratory, and applied research in the biomedical field, as well as for priority support of the most

Funding for biomedical research and selected life sciences priorities in the United States and Russia in 2014-2016

USA	Russian Federation
Federal budget funds allocated for the development of fundamental, translational, and personalized medicine	
2015: Health care services, including the National Institutes of Health (27 research centers) – 30.2 billion USD^a	2015: Russian Ministry of Health (54 research centers) – 1.566 billion rubles – <i>fundamental research^b</i> 2015: FANO (50 research centers) – 5.976 billion rubles – <i>fundamental research^b</i> 2015: Russian Ministry of Health (54 research centers) – 3.195 billion rubles – <i>applied research^b</i> 2015: Russian Ministry of Health (63 State Medical Academies) – 1.110 billion rubles – <i>applied research^b</i> 2015: Russian Science Foundation, 3.8 billion rubles – <i>fundamental and exploratory research in biomedicine^d</i> 2015: FTP “R&D” – 1.6 billion rubles – <i>applied research and pilot projects in biomedicine^d</i> 2015: Russian Foundations for Basic Research – 1.6 billion rubles – <i>fundamental research in biomedicine^d</i> 2015: Small Enterprise Assistance Fund – 0.65 billion rubles – <i>applied research and pilot projects in biomedicine^d</i> 2015: Decree of the Government of the Russian Federation № 218 of April 9, 2010 – 0.6 billion rubles – <i>applied research and pilot projects in biomedicine^d</i> 2015: Decree of the Government of the Russian Federation № 220 of April 9, 2010 – 0.2 billion rubles – <i>fundamental research in biomedicine^d</i> FTP “PHARMA-2020” – 12.8 billion rubles, including 1.936 billion rubles for funding of pre-clinical trials ^d
Private sector R&D budget for life sciences	
2014: – 92.6 billion USD^c	The contribution of industrial enterprises to the consolidated national R&D budget of the Russian Federation did not exceed 12% in 2012-2014 ^f .
TOTAL 123.7 billion USD	TOTAL 33.1 billion rubles = 600 million USD
Funding for some priority areas of biomedical research	
The study of mechanisms of antibiotic resistance: 2015: 600 million USD 2016: 1.2 billion USD	The RSF competition “New approaches to the fight against infectious diseases” 2015: 100 thousand USD/year (6 million rubles/year) ^e
“Brain Research through Advancing Innovative Neurotechnologies” initiative: 2015: 64 million USD^c 2015: 136 million USD^c <u>National Program for Childhood Research:</u> 1.2 billion USD in 2015 ^c +165 million USD for the study of the impact of the environment on children’s health in 2016 ^c	2014: A special RAS program “Fundamental research for the development of biomedical technologies” 3.3 million USD/year (200 million rubles/year, or 1–4 million rubles per project) ^f

Sources:

^aThe 2015 Budget: Science, Technology, and Innovation for Opportunity and Growth [7].

^bState program of Russian Federation “Development of Healthcare” [6].

^c2014 Global R&D Funding Forecast [2].

^dKazeev K V. Report of the Head of Department of Science and Technology of the Ministry of Education of Russian Federation [8].

^ePublic analytical report “Biomedicine” [9].

^fAnalytical report “Annual monitoring of R&D expenditures (including priority areas of innovation development in Russia)” [3].

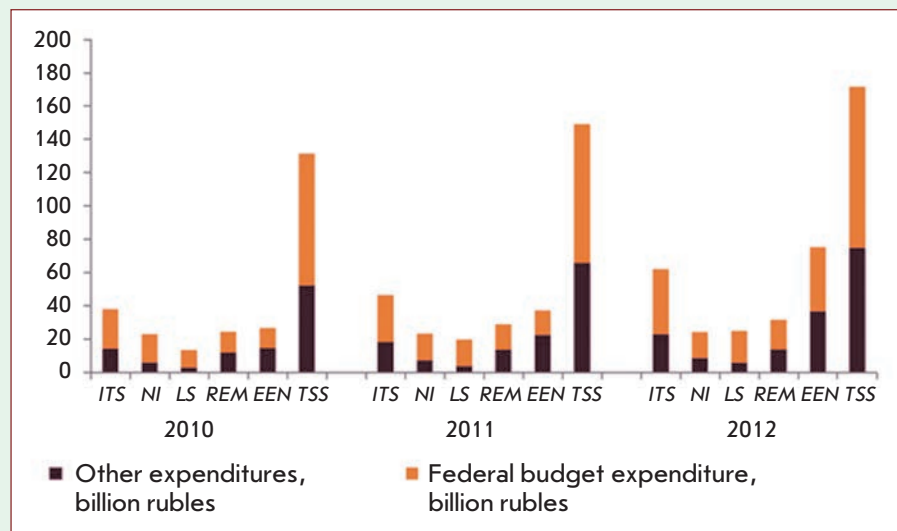


Fig. 1. Structure of Russian domestic expenditures on Research and Development in priority areas of development in science, technology, and engineering (Source: *Science Indicators 2012–2014*)

Abbreviations: ITS, information and telecommunications systems; NI, nano systems industry; LS, life sciences; REM, Rational Environmental Management; EEN, energy efficiency, energy conservation, nuclear power; TSS, transport and space systems

promising and socially important topics.

Another noteworthy topic is the extremely low amount of financing of research activities in medical schools: 63 medical colleges of the Russian Ministry of Health allocate only 1.110 billion rubles or (18.5 million USD) to applied research.

An even more drastic difference is observed in the funding of projects addressing thematically identical issues in life sciences, such as deciphering the mechanisms of antibiotic resistance of pathogenic microorganisms. In the United States, 600 million USD was allocated to finance the program for antibiotic resistance research in 2015, and for 2016, planned federal support for the program has been doubled and will amount to 1.2 billion USD. In Russia, the Russian Science Foundation announced in 2014 an annual competition: “New approaches to the fight against infectious diseases.” It has awarded 63 research grants of up to 6 million rubles per year, which is equivalent to about

6.3 million USD in total. Therefore, the difference in the level of funding for similar priorities in the field of biomedicine in the U.S. and Russia is a hundredfold!

THE VOLUME OF PRIVATE SECTOR INVESTMENT IN BIOMEDICAL R&D IN THE U.S. AND RUSSIA

The estimated contribution of national industrial enterprises involved in the development of technologies and products for the health care market to biomedical R&D budgets deserves a separate discussion. The annual R&D budgets of U.S. companies in the field of “life sciences” account for almost half of global corporate budget in this field. In 2014, this figure amounted to 92.6 out of 201.3 billion USD of the global corporate R&D expenditures on the “life sciences industry.” Remarkably, this “almost half” contribution by the U.S. biomedical industrial sector has been maintained over the last five years; in 2011, it was estimated to be 84.5 out of 184.2 billion USD of global expenditures (i.e. 46% [2].)

As a result, the ratio of public and private sector contribution in the 2014 U.S. consolidated budget for fundamental and applied research for fundamental and applied research to the biomedical field was *ca.* 3 to 1 (32 billion USD of the health and human services budget, including the National Institutes of Health vs. 92.6 billion USD of corporate contribution to R&D in life sciences) [2].

We have been unable to find information on the volume of Russian industrial companies’ investment into the development of drugs and medical equipment in 2014. According to the “Annual monitoring of federal budget R&D expenditures (including priority areas of innovation development of Russia),” the contribution of Russian industrial companies to the consolidated national budget for R&D did not exceed 12% [3].

However, domestic industrial companies actively use public funding, first of all FTP “Pharma-2020”, to carry out corporate R&D. For example, in 2015, the FTP funds for scientific research were allocated as follows: 8.64 billion rubles for 312 projects in publicly funded institutions and 7.15 billion rubles for R&D in 215 commercial companies. Thus, almost half (45%) of the FTP “Pharma-2020” budget in 2015 was allocated to the development of corporate sector R&D [10].

These data highlight the fact that the domestic industrial sector not only fails to provide an additional and significant source of research funding for biomedicine, but also creates competition for public funds for R&D in this area.

ASSESSMENT OF THE DYNAMICS OF FEDERAL BUDGET EXPENDITURES IN LIFE SCIENCES IN THE CONTEXT OF R&D FUNDING IN AREAS OF PRIORITY

In 2010, 2011, and 2012, support for the priority areas accounted for 49%, 50%, and 56%, respectively, of

the total volume of federal budget expenditures on R&D. [3]. An analysis of the dynamics of federal budget expenditures on R&D in six priority areas of science, technology, and engineering in the Russian Federation, approved by Presidential Decree № 899 of 07.02.2011 [11], reveals that between 2010-2012 “Life sciences” received the smallest share of the federal budget (Fig. 1) [12].

In contrast, the quotas of public financing in the United States show that life sciences are one of the main scientific and technological priorities of the country. *Figure 2* outlines the distribution of the U.S. federal budget on R&D between the main recipients in 2012-2014 [2].

The U.S. National Institute of Health received the most significant amount of money compared to all other entities. For example, in 2014 the National Institute of Health received more than half of all the money allocated for the civil sector of U.S. science: 32.0 out of 58.8 billion USD. It is not surprising that more than half of the Nobel Prize laureates in medicine since 2000 are affiliated with U.S. universities [2].

According to the R&D Magazine of the Battelle analytical agency, the U.S. is the global technological leader in the field of biomedicine. The UK ranks second, Germany, third, and Japan fourth with China rounding up the top five (Fig. 3)

DISCUSSION

According to the list of orders issued on July 14, 2015, following a meeting of the Council for Science and Education [13], the Presidential Administration was charged with defining the principles of establishing priorities in scientific and technological development. It appears that in the current economic situation, it would be prudent to base this choice not only on the assessment of the socio-economic impor-

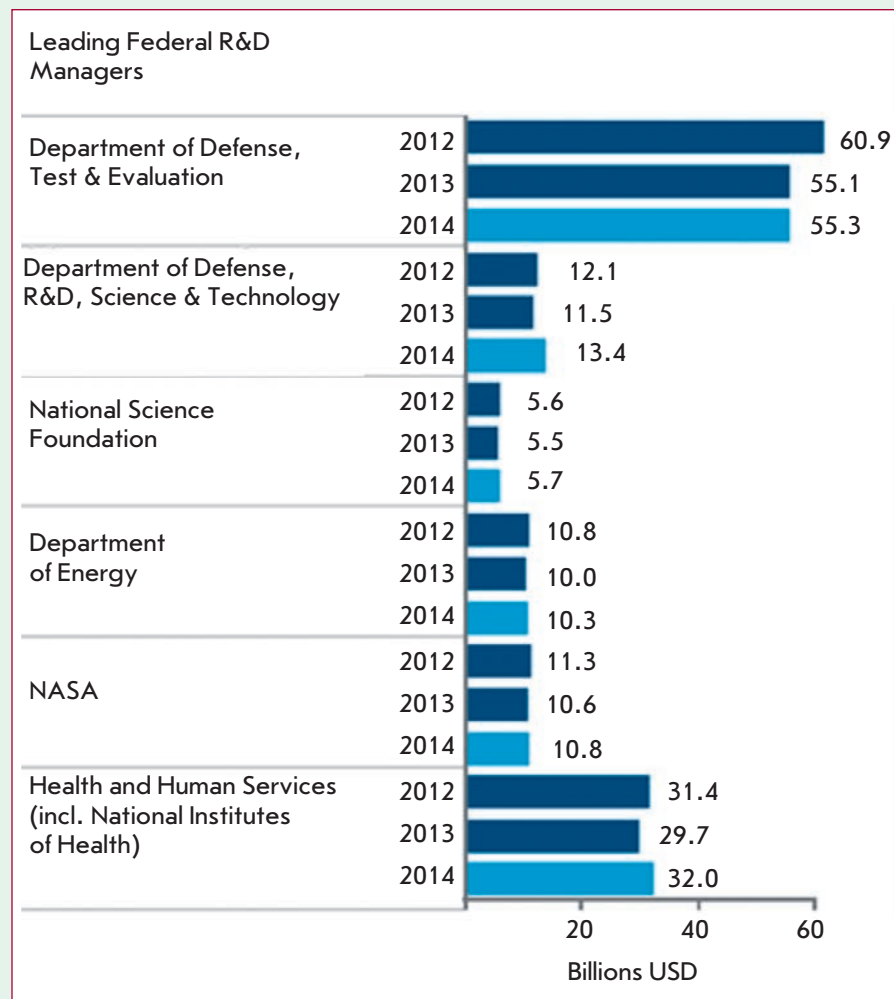


Fig. 2. Distribution of U.S. Federal budget funds between main budget controllers in 2012-2014 (Source: 2014 Global R&D Funding Forecast)

tance of a scientific and technological area for the country, but also on the volume of state budget funding necessary to achieve competitive scientific and technological breakthroughs on the world stage.

The data suggest that the inadequacy of the Russian Federation expenditures on the development of the civil sector of science in comparison with industrialized countries necessitates the selection of a very limited number of research areas as priorities for the scientific and technological development of the country. Only a

marked decrease in the number of priorities would “ensure that they are fully funded,” as noted by the President of the Russian Federation at the Meeting on Science and Education [1].

When it is declared that priority should be given to technologies that ensure a quality of life and, above all, advanced medical technologies [1], the objectives and expected outcomes of the implementation of these priorities are absolutely clear. However, one should take into account the fact that the development of new medical technologies

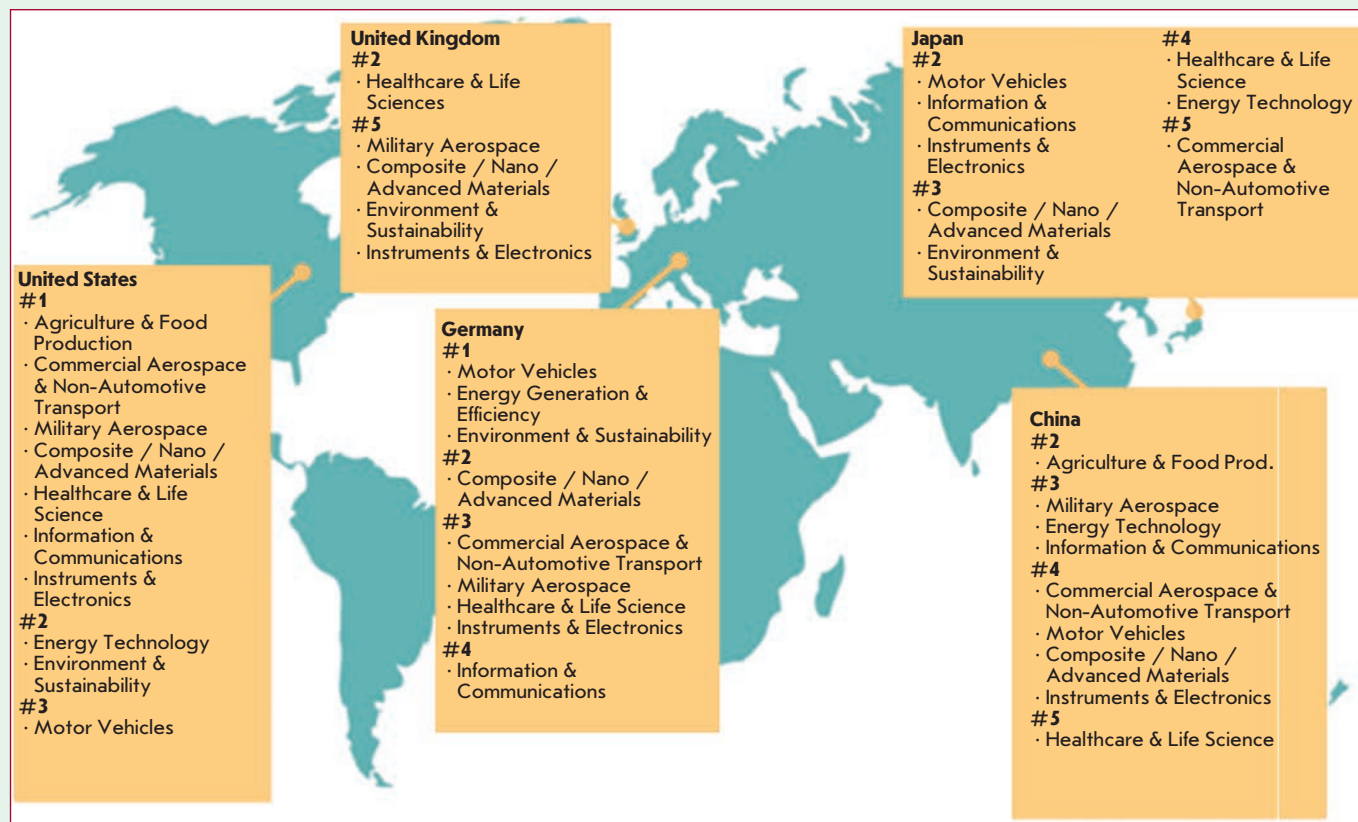


Fig. 3. Five world leading countries in life sciences (Source: 2014 Global R&D Funding Forecast)

is one of the most expensive scientific fields in the world, and that countries striving for a position of technological leadership in biomedical fundamental and applied research are pledging budgets which are hundreds of times bigger than funding for biomedical research in the Russian Federation.

An additional factor impeding growth in competitive national biomedical research is the lack of interest on the part of the industrial sector in financing exploratory and applied research aimed at creating new domestic drugs and medical equipment. In the United States, the share of “life sciences” industrial companies’ contribution reached 75% of nationwide R&D expenditures in the field of biomedicine in 2014, and in absolute terms their investment in biomedical R&D amounted to 92.6 billion USD, hav-

ing increased from 84.5 billion dollars back in 2011.

CONCLUSION

Based on the facts outlined above, it seems appropriate to implement the following set of measures aimed at improving the mechanisms of biomedical research funding in Russia.

First of all, given that Russian domestic R&D expenditures in 2016 will amount to no more than 2% of the global budget, we believe that there is no need to include areas of research that are already a topic of major international projects into the list of priorities, since the results of those projects are publicly available and there are no obvious barriers and restrictions to their use in practice in Russian health care. Examples of such projects include the international “Human

Proteome” project, the program of study of the mechanisms of antibiotic resistance (USA), the program for the study of the natural foci of infections (USA), etc. That does not mean, of course, that funding for domestic research in these areas should be shut down, but it does not make sense to give them a “priority” status - which implies a concentration of significant financial resources - if those areas can be developed in the Russian Federation through the active use of the results of major international projects.

Secondly, we should eliminate unnecessary duplication of funding of biomedical projects with similar subjects by different managers of the state R&D budgets (both in public funding and in contest-based and targeted financing), and simultaneously achieve a balance be-

tween the amount of funding, numbers of staff at research institutes, and the number of compulsory and competitive research topics.

Thirdly, since the major beneficiary of fundamental and applied research in the biomedical field is the population of the Russian Federation and since the maintenance and restoration of its health is the responsibility of the Russian Ministry of Health, this agency should adjust R&D themes based on the health status of the population, including contests organized by state

foundations and institutes for development.

Fourthly, development of biomedical technologies that are relevant to practical aspects of health care in the Russian Federation and making research and industrial production in this area competitive at the global level is impossible without active participation of major industrial companies, both Russian and International. Today, Russian pharmaceutical companies and companies producing medical equipment basically do not invest

in the scientific and technological groundwork in the field of biomedicine and do not implement the outcomes of domestic research in production, citing regulatory barriers and lack of guaranteed demand for products from the Russian Ministry of Health. Therefore, it is necessary to strengthen coordination between state entities, scientific organizations, and all other participants involved in the production and consumption of drugs and equipment for medical care in the Russian Federation.

REFERENCES

1. RF Presidential Board Meeting dedicated to science and technology on "New challenges and priorities in development of science and technology in Russian Federation". Transcript (2015) // Official website of Russian President, 24.06.2015, Available from: <http://kremlin.ru/events/councils/by-council/6/49755>.
2. 2014: Global R&D Funding Forecast (2013). Battelle, R&D Magazine. Available from: http://www.battelle.org/docs/tpp/2014_global_rd_funding_forecast.pdf.
3. Analytical report "Annual monitoring of federal R&D expenditures (including priority areas of innovation development of Russia)" (2014) // Analyst Centre affiliated with the government of Russian Federation, December 2014, Moscow. Available from: <http://ac.gov.ru/files/publication/a/4889.pdf>.
4. Federal Law dated 01.12.2014 N 384-FL (Ed.13.07.2015) "On Federal budget for 2015 and new planned period of 2016 -2017 years" // Internet portal of "Rossiyskaiya Gazeta" December 5, 2014. Available from: <http://www.rg.ru/2014/12/05/budget-dok.html>
5. Investing in America's future (2014). White House Office of Science and technology policy. Available from: <https://www.whitehouse.gov/sites/default/files/omb/budget/fy2016/assets/investing.pdf>.
6. Russian Government Resolution dated 15 April 2014. № 294 (2014) "On adopting a state programme of Russian Federation "Development of Healthcare" // Internet portal of "Rossiyskaiya Gazeta", April 24, 2014. Available from: <http://www.rg.ru/2014/04/24/zdravooxr-site-dok.html>
7. The 2015 Budget: Science, Technology, and Innovation for Opportunity and Growth. Science, Technology, Innovation, and STEM Education in the 2015 Budget (2014). White House Office of Science and technology policy. March 2014. URL: <https://www.whitehouse.gov/sites/default/files/microsites/ostp/Fy%202015%20R&D.pdf>.
8. Kazeev K V. Report of the Head of the Department of Science and Technology of Ministry of Education of the Russian Federation (2015). Materials from the Committee Meeting of Expert Council on Healthcare 05.06.2015 // Internet-portal of Federal Assembly Committee of Russian Federation, 09.06.2015. Available from: http://social.council.gov.ru/activity/expert_activities/56304.
9. Analytical report "Biomedicine" (2015). Moscow. Available from: <https://reestr.extech.ru/docs/analytic/reports/medicine.pdf>.
10. List of contracts for Research and Development Programmes "Development of pharmaceutical and medical industry of Russian Federation in the period until 2020 and future perspective". // Internet-portal of Department of State Grant Programmes and Capital Investments of Russian Ministry of Economics. Available from: <http://fcp.economy.gov.ru/cgi-bin/cis/fcp.cgi/Fcp/ViewFcp/View/2015/350/>.
11. Russian President Decree № 899 of 07.02.2011 (2011) "On recognition of priority areas in development of science, technology and engineering in Russian Federation and listing critical technologies for Russian Federation" // Official Website of Russian President. Available from: <http://kremlin.ru/acts/bank/33514>.
12. Science indicators: 2014: statistics summary (2014) M.: National Research University «Higher School of Economics», 400 c.
13. List of Decrees dated 14 July 2015 following the results of Board Meeting on Science and Education dated 24 June 2015 (2015) // Official Website of Russian President. Available from: <http://kremlin.ru/acts/assignments/orders/50006>.

Broadly Neutralizing Antibodies against HIV-1 As a Novel Aspect of the Immune Response

D. N. Shcherbakov^{1,2*}, A. Y. Bakulina^{1,3}, L. I. Karpenko¹, A. A. Ilyichev¹

¹State research center of virology and biotechnology "Vector", Koltsovo, 630559, Novosibirsk region, Russia

²Altai State University, 61 Lenin St., 656049, Barnaul, Russia

³Novosibirsk State University, 2 Pirogova St., 630090, Novosibirsk, Russia

*E-mail: scherbakov_dn@vector.nsc.ru

Copyright © 2015 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 human immunodeficiency virus-1 (HIV-1) has the ability to evade the adaptive immune response due to high mutation rates. Soon after the discovery of HIV-1, it was originally proposed that neutralizing of antibodies to the virus occurs rarely or cannot be elicited at all. In the 1990s, there appeared reports that sera of select HIV-1-infected individuals contained antibodies capable of neutralizing different virus subtypes. Such antibodies were named broadly neutralizing antibodies (bNAbs). Since 2009, the development of new cell technologies has intensified research efforts directed at identifying new bNAbs with a neutralization potency of over 90% of primary HIV-1 isolates. These antibodies have unique characteristics which include high levels of somatic mutations and unusually long variable loops that penetrate through the glycan shield of HIV-1 Env to contact the protein surface. In this review, we will attempt to summarize the latest data on bNAbs against HIV-1 in terms of their interactions with the sites of vulnerability on HIV-1 glycoproteins.

KEYWORDS HIV-1, gp120, gp41, bNAbs, Broadly neutralizing antibodies.

ABBREVIATIONS HIV-1 - human immunodeficiency virus type 1; AIDS - acquired immune deficiency syndrome; bNAbs - broadly neutralizing antibodies; Env - HIV-1 viral envelope protein; gp - glycoprotein; CD4 - transmembrane glycoprotein receptor for an HIV-1; CD4bs - CD4 binding site; CCR5 - C-C chemokine receptor type 5; CXCR4 - C-X-C chemokine receptor type 4; MPER - membrane-proximal external region; RSC - Resurfaced Stabilized Core; RT-PCR - reverse transcription polymerase chain reaction; CDR - complementarity determining region; CDR H3 - third complementarity-determining regions of the heavy chain.

INTRODUCTION

A distinctive hallmark of modern-day medicine in the last decade has been the increasing use of monoclonal antibodies offering targeted therapeutic effects for a range of disorders. A successful outcome with monoclonal treatment has been reported for dozens of commercial products over the past 15 years. Experimental data on the design and application of monoclonal antibodies have been reviewed in detail elsewhere [1, 2]. Although the mechanisms by which the humoral response is triggered and maintained remain elusive, new insight into broadly neutralizing HIV-1 antibodies (bNAbs) has expanded our understanding of the antibody response.

The human immunodeficiency virus type 1 (HIV-1), which causes the acquired immunodeficiency syndrome (AIDS), was discovered over 30 years ago. According to the WHO, > 78 million people were diagnosed as HIV-1 positive by the end of 2013, over half of whom have been reported dead. A safe and potent

vaccine against HIV-1 could limit the spread of HIV-1 and subsequently eradicate the disease. The tendency of HIV-1 to rapidly accumulate mutations to escape host immune responses represents a major hurdle to the development of effective vaccines. HIV-1 has now been classified into 9 distinct subtypes and their recombinant forms [3].

Prior to 1990, it was considered that antibody-mediated neutralization of HIV-1 in the host was reduced or even abolished. In the 90s, it was found that sera of HIV-1-infected individuals contained antibodies that could recognize and neutralize different subtypes of HIV-1. These antibodies were called broadly neutralizing antibodies (bNAbs) [4]. Since 2009, with the advent of new cell-based assays, there has been a surge in the number of publications pertaining to the application of novel bNAbs. This review summarizes current literature on bNAbs, which suggests new possibilities for anti-HIV-1 vaccine design.

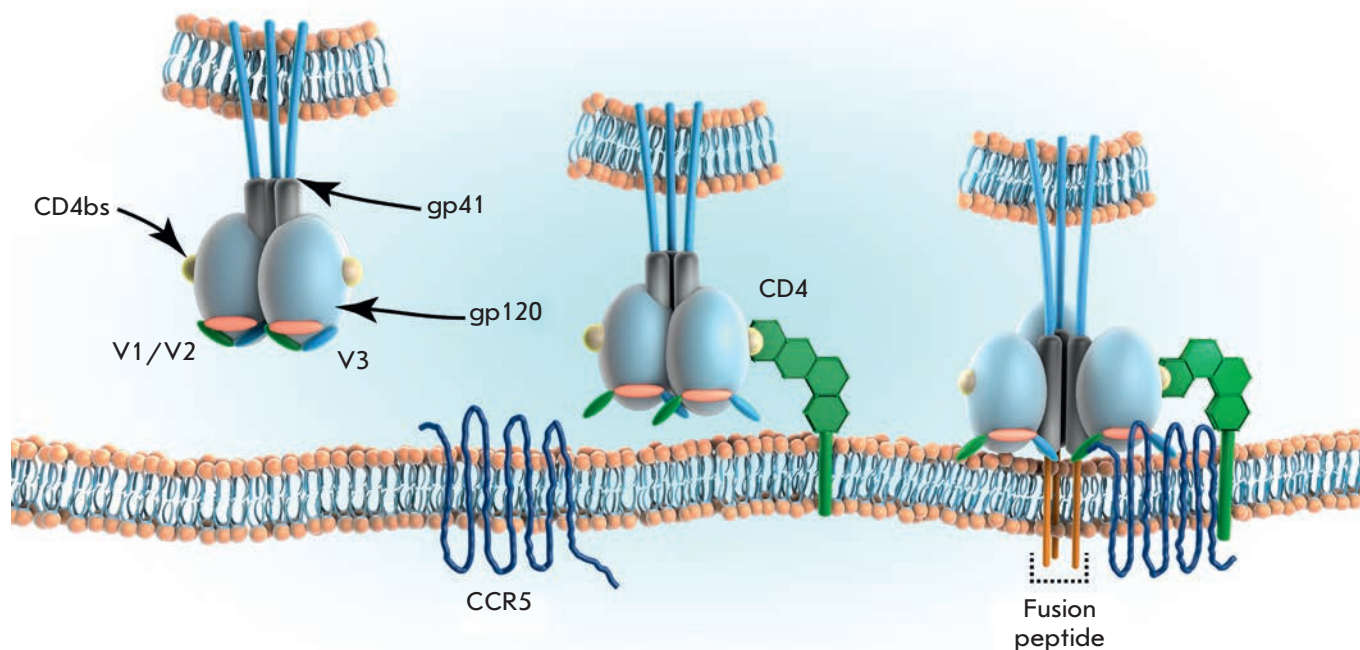


Fig. 1. Trimeric Env interaction with the host cell membrane is illustrated. The gp120 subunit binds to the CD4 receptors, triggering conformational rearrangements to unmask the coreceptor binding site originally hidden by the V3 and V1/V2 loops. Engagement with CCR5 or the other coreceptor drives viral fusion and entry

STRUCTURAL AND FUNCTIONAL ORGANIZATION OF HIV-1 SURFACE GLYCOPROTEINS

HIV-1 is a spherical enveloped virus with a diameter of 140 nm. The viral envelope consists of a lipid bilayer derived from the plasma membrane of infected cells, with glycoprotein spikes anchored in it. Each viral spike is a trimeric heterodimer containing the external glycoprotein gp120 and the transmembrane glycoprotein gp41, with about 70–79 trimers on the virion surface [5]. Of all viral proteins, only gp120 and gp41 have epitopes for antibody recognition. These proteins play an essential role in virus entry into host cells.

The glycoproteins gp120 and gp41, which are encoded by the *env* gene, are called Env proteins and translate to a full-length gp160 polyprotein, followed by trimerization and cleavage by a furin-like protease in a Golgi compartment. The cleaved gp120–gp41 molecule is trapped in a metastable state until a transition to an energetically more favorable state. Like other Type 1 fusion proteins, these trimetric structures undergo receptor-induced conformational changes to increase the exposure of the gp 41 ectodomain for the fusion of viral and cellular membranes (*Fig. 1*). The crystallography on individual gp120 and gp41 components, as well as in the context of trimeric gp120/gp41, has been obtained in recent years, alongside mapping of gp120 CD4 and co-receptor binding sites [6].

HIV-1 infects cells through interaction with CD4 and chemokine receptors via transmembrane domains, such as CCR5 or CXCR4. Susceptible cells include T helper cells (Th), macrophages, follicular dendritic cells, Langerhans cells, and microglial cells. Certain CD4-negative cell types carrying chemokine receptors can also be infected. They include astrocytes, cervical cells, rectal and bowel mucosal cells, brain capillary and cervical endothelial cells, and corneal cells. CD4 serves as an adhesion molecule that stabilizes the viral contact with the host cell membrane [7]. The lack of attachment to the coreceptors prevents fusion from taking place; the virus enters by endocytosis and is typically inactivated upon uptake [7].

The ability of HIV-1 to rapidly accumulate mutations enhances the sequence variability of viral proteins. However, the domains within the proteins binding to CD4 and CCR5 are conserved. gp120 contains five conserved regions (C1–C5) that are interspersed between 5 variable regions (V1–V5). The variable loop regions occlude the constant regions to escape from antibody attack [8]. Following infection, antibodies are primarily raised against variable regions and, due to hypervariability, HIV-1 evades immune surveillance [9, 10]. Another mechanism by which the virus overcomes the immune defences is the glycosylation of surface proteins. It has been demonstrated that gp120

contains approximately 25 N-glycosylation sites, which form a glycan shield [8]. Mutations induce changes in the positioning of glycosylation sequences on gp120, thus altering the antigenic makeup of the viral envelope [11]. A virus mutant lacking certain variable loops and glycosylation sites becomes more susceptible to neutralization by polyclonal sera. This leads one to suggest that hypervariable loops mask the conserved epitopes of the Env protein [12]. With this in mind, it was a long-standing view that neutralization antibodies against Env antigens cannot be elicited in the course of the disease.

Indeed, knowledge on antibodies capable of neutralizing HIV-1 was lacking during the first years of HIV-1 research. Past evidence had posited that the human organism by itself was unable to limit viral replication, owing to its failure to raise neutralizing antibodies or, if mounted, their poor neutralizing capacity [18–20]. Recently, there have appeared reports on sera of HIV-1-infected individuals containing antibodies that neutralize both laboratory-adapted strains and primary isolates [21–26]. It was first suggested that broadly neutralizing HIV-1 antibodies (bNAbs) occur in a small proportion of HIV-1 infected patients [20, 27]. From then onwards, bNAbs were detected in some 30% of infected individuals diagnosed within one year of infection [23, 28, 29]. More recently, bNAbs have been found in over 50% of HIV-1 carriers [30]. Importantly, 1% of infected individuals elicit neutralizing antibodies with strong affinity for a wide array of primary HIV-1 isolates, as well as up to 99% of the HIV-1 isolates known to date [31].

Insights into bNAbs and their interaction with HIV-1 could provide fundamental clues in our understanding of this phenomenon and may also be useful in the rational design of effective vaccine.

The HIV-1 trimeric complex gp41-gp120 has 5 sites of vulnerability to neutralizing bNAbs. Each site carries overlapping epitopes recognized by different bNAbs. These sites of vulnerability include the CD4 binding site (CD4bs) of gp120, the site within gp120 targeted by the PG9 and PG16 antibodies, and the membrane proximal external region (MPER) of gp41, an epitope adjacent to the V3 loop and spanning the gp120/gp41 interface. The five major sites of vulnerability are represented in *Fig. 2*.

Brief characteristics of bNAbs with the history of discovery and sites of binding are given in Table. The antibodies highlighted in grey are first-generation bNAbs.

FIRST GENERATION bNAbs

The history of broadly neutralizing antibodies can be divided into two periods. The first studies of bNAbs ap-

peared in the early 1990's, reporting on b12, 2G12, 2F5, Z13, and 4E10 antibodies.

The first bNAb produced using phage display was b12, which binds to conserved gp120 CD4bs [4]. It was obtained as an antibody Fab-fragment generated from a phage display antibody library from the bone marrow of an HIV-1-infected non-progressor. The cloning of the variable region of the Ig heavy chain and light chain was random; therefore, such combinations may not occur naturally [32].

The 2G12 bNAb recognizes the $\alpha 1 \rightarrow 2$ mannose residues on gp120, located close to the V3 and V4 loops [33], and has a unique structure. The heavy chains are intersected, with each light chain bound to the constant region of one heavy chain and the variable region of the other heavy chain. Due to this arrangement, Fab-fragments are unusually closely aligned. Such an antibody was obtained from only one donor library. The epitope recognized by 2G12 is conformationally sensitive, strongly depending on asparagine glycosylation in the C2-, C3-, C4-domains, and the V4 loop.

The 2F5 and 4E10 bNAbs interact with linear overlapping epitopes based around the MPER-region of gp41, exhibiting polyreactivity with bivalent heterologation [34]. They have the ability to strongly bind to MPER with one Fab-fragment, while the other Fab-fragment demonstrates low affinity for another molecule target on the HIV-1 surface. The heterogeneous ligand binding seems to increase the neutralization activity against primary HIV-1 isolates [34].

Like b12, Z13 was generated from a combinatorial phage display library. To enhance affinity, amino acid substitutions were introduced to the paratope to generate a clonal variant, Z13e1, with a 35-fold increase in the binding capacity [19, 35, 36].

Studies carried out with HIV-1 pseudoviruses of different subtypes have demonstrated that first-generation bNAbs exhibit moderate breadth and neutralization potency. Achieving the desired efficacy against a wide range of HIV-1 isolates requires high concentrations of these bNAbs, which impedes progress in this field. At the same time, passive immunization of macaques with a combination of neutralizing Abs b12, 4E10, 2F5, and 2G12 confers complete protection against challenge by SHIV89.6P [37]. These findings spurred further studies aimed at identifying new bNAbs.

SECOND GENERATION bNAbs

Numerous attempts to produce bNAbs with high, excellent characteristics had not met with success. The first bNAbs with enhanced efficacy and potency against a broad spectrum of primary HIV-1 isolates were only identified in 2009. The successful outcome

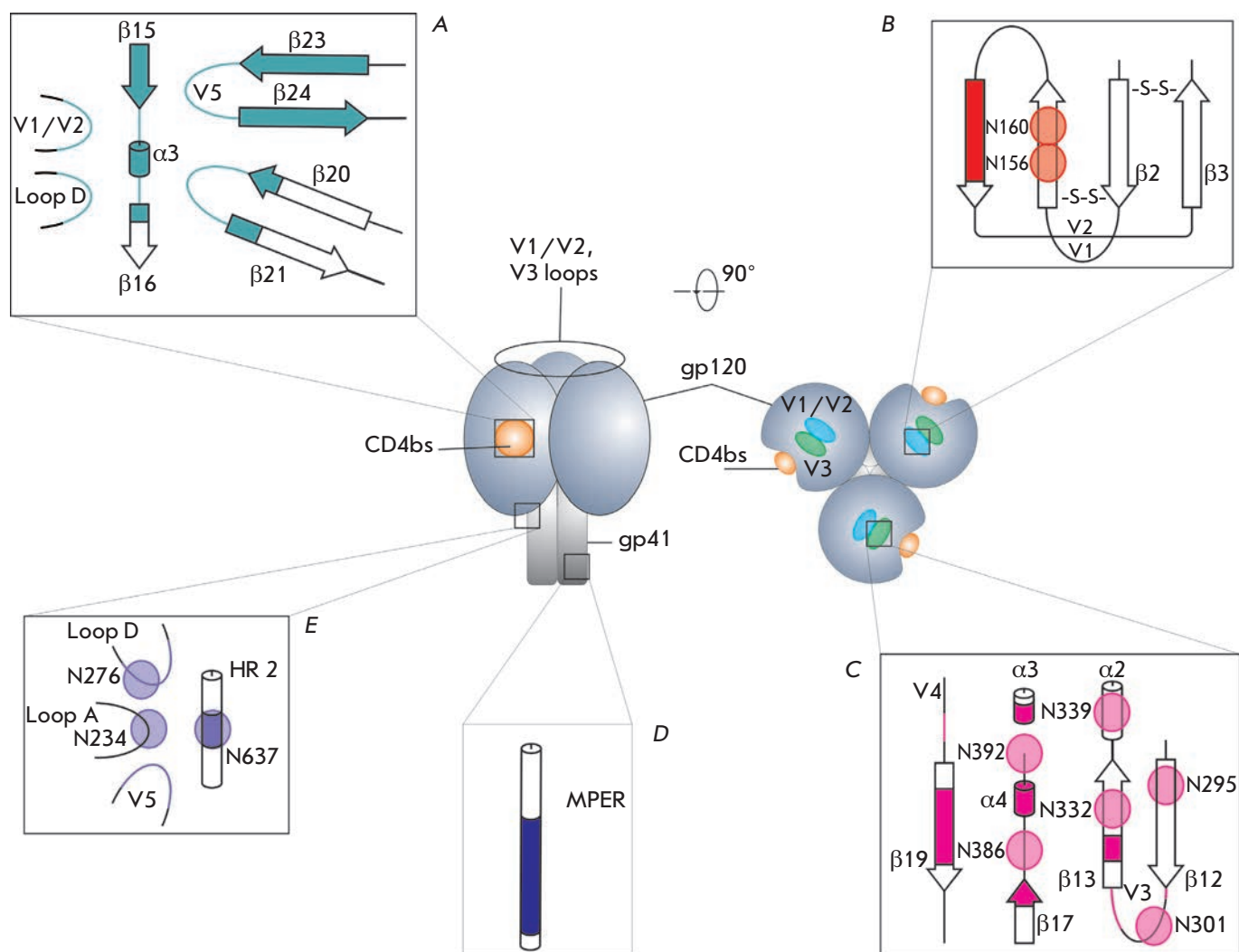


Fig. 2. Schematic of the trimeric HIV-1 envelope glycoprotein structure and sites of vulnerability recognized by bNAbs are shown. The α -helices are shown with cylinders, β -sheets, with arrows; loops, with thin lines; glycosylated amino acid residues, with circles. Detailed characteristics of bNAbs are given in Table.

A – CD4bs on gp120 is involved in CD4 attachment. The major epitopes are the D loop, the V1/V2 loops, the V5 loops and the flanking β -sheets 23 and 24, an epitope within β -sheet 15, the α -helix3, and an epitope within β -sheet 16. The epitope structure is reconstituted based on data from [13].

B – the epitope made up of V1/V2. The antibody recognition site is a region in a β -conformation, including glycans at N156 and N160. The epitope structure is drawn based on Ref [14].

C – the epitope is on gp120. The sites involved in binding are: regions of β -sheets 19, 17, 13, V3, and V4 regions, 4 and 3 α -helices, glycans at N392, N386, N339, N332, N301, and N295. The epitope structure is drawn based on Ref [15].

D – MPER-site, a linear epitope on gp41. A region within the MPER-site is amenable to recognition. The epitope structure is drawn based on Ref [16].

E – the epitope at the gp120/ gp41 interface. N-linked glycans within gp41, a glycan moiety at N637, N276, and N234. V5 and D regions are targeted. The epitope structure is drawn based on Ref [17].

was achieved through the use of three strategies: (i) Screening of sera from chronically infected HIV-1 individuals which contained high affinity and cross-reactive antibodies, (ii) application of new approaches to B-cell selection and sorting, (iii) development of high-throughput procedures for generating human monoclonal antibodies.

The identification of VRC01 [38] and PG9/PG16 [39] was a breakthrough in the field of bNAbs in 2009–2010. The distinctive features of these antibodies are the strong neutralization profiles of a wide array of primary HIV-1 isolates and enhanced efficacy; 10-fold lower levels of antibodies are needed for protection with regard to first-generation bNAbs. VRC01, for example, shows neutralization activity of up to 93%; PG9/PG16, up to 80% of primary HIV-1 isolates, whereas b12 (first-generation bNAbs) neutralizes only 35% [40].

VRC01 was produced using a novel strategy employed by Mascola *et al.* [38] that generated an antigenically resurfaced glycoprotein representing a substituted gp120 core, called the resurfaced stabilized core (RSC). To facilitate epitope selectivity using RSC as a probe, the CD4bs was preserved, the variable regions 1 to 3 removed, and other antigenic regions altered to reduce recognition. In addition, a Δ RSC probe with impaired b12 binding was used as a negative control. Sera containing NAbs to CD4bs were identified, followed by isolation of individual B-cells using the RSC and Δ RSC probes conjugated to fluorochromes. Single B-cells bound to the RSC probe were sorted. Single-cell RT-PCR [41] was applied to amplify cDNA encoding light and heavy chains of individual cells, followed by cloning into expression vectors that reconstituted the heavy- and light-chain constant regions [38].

PG9 and PG16 were identified earlier than VRC01, using a high-throughput strategy [39]. Activated B-cells were screened for antibodies with neutralizing activity against the primary HIV-1 isolates JR-CSF and SF162 and binding to the recombinant gp120 and gp41 proteins. The desired antibody genes were obtained from five B-cell clones. All five antibodies were tested for neutralization activity against a panel of pseudoviruses, and PG9 and PG16 demonstrating exceptional neutralization breadth and potency were selected.

ENV SITES OF VULNERABILITY TARGETED BY bNAbs

CD4 binding site

Following the discovery of second-generation bNAbs, VRC01 enjoyed much attention due to its remarkable affinity for CD4bs of the HIV gp120 trimeric molecule. CD4bs, one of the prominent sites of vulnerability, harbors epitope for bNAbs (*Fig. 2A*). The existence of broadly reacting antibodies was hypothesized earlier

[24, 42]. However, besides the monoclonal antibody b12, other broadly neutralizing antibodies escaped identification. Use of high-throughput strategies yielded three novel bNAbs (VRC01, VRC02, and VRC03) that recognize CD4bs. All of them were shown to be somatic variants with shared characteristics, with VRC03 displaying a limited neutralization breadth [38]. VRC01 has a number of distinctive features. First, it has a high level of somatic mutations in its variable regions. Somatic mutations usually account for 5–20% of V_H genes, whereas VRC01 can carry up to 40%. Second, The variable domain of VRC01 is closely related in structure to the CD4 receptor on T helper cells. VRC01 displays structural mimicry of CD4 interaction with CD4bs on gp120 [13]. The precise targeting is a key determinant of high neutralization potency. Despite the close resemblance to the CD4 receptor, VRC01 interaction with gp120 considerably differs. Upon binding by CD4 to trimeric Env, the gp120 subunit undergoes structural conformations; by contrast, in the same context VRC01 traps gp120 in a state that prevents viral entry [43].

Following VRC01 identification, a myriad of CD4bs-binding antibodies, for example, PGV04, CH30–34 [44], 3BNC117, 3BNC60, 3BNC55, 12A21, 12A12, 8ANC195, 8ANC131, 8ANC134, NIH45-46, 1NC9, and 1B2530 [45], were obtained using the same strategy.

Notwithstanding that all these antibodies target a CD4bs epitope on gp120, considerable differences in the mode of action are observed [45]. For instance, certain antibodies bound to monomeric gp120 trigger conformational changes, reminiscent of those that take place upon binding by CD4, which is not shown for other antibodies [46]. Despite shared structural characteristics, CD4bs-binding antibodies could be encoded by different genes, allowing for a subclassification of VRC01-like antibodies [47].

Of particular interest is NIH45-46^{G54W}, whose identification was made possible owing to the structure-based design based on NIH45-46, an antibody with exceptional potency and breadth against CD4bs. X-ray crystallographic data for the structure of NIH45-46 bound to gp120 revealed that a glycine to tryptophan substitution at position 54 increases the interactive surface between the antibody and the viral glycoprotein. Efforts to pursue the substitution yielded NIH45-46^{G54W}, which displays enhanced potency and breadth [48].

PG9 and PG16 recognition site

PG9 and PG16 antibodies, which are somatic variants, show excellent neutralization breadth. PG9 neutralizes 78% of pseudoviruses; PG16, 73%. Importantly, the neutralization potency exhibited by both antibodies could vary by two orders of magnitude. These differ-

REVIEWS

Characteristics of bNAbs against HIV-1

Envelope site	Epitope (specificity)	Antibody designation	Year of generation	Neutralization breadth, %	Neutralization potency*, $\mu\text{g/ml}$	The length of CDR H3, a.a.	Somatic mutations, %, aa substitutions
gp41 MPER	ELDKWA [18]	2F5**	1992	55–67 [39, 40, 57, 58, 59]	1.44 [40]	24	15.2
	WFD(I/L)(T/S) NW(L/I)WYIK [60]	4E10**	1994	85–100 [36, 39, 57, 58, 61, 62]	1.62 [40]	20	15.6
	SLWNWFDITN [63]	Z13**	2001	35 [62]	40 [62]	19	21
	WNWFDITN [63]	Z13e1**	2007	50 [36]			
	WFDITNWIWYIL/R [57]	10E8	2012	98–99 [40, 57, 58]	0.25 [40]	22	22.1
gp120 CD4bs	The loops D, V1/V2, V5 and CD4-binding loop	b12**	1991	35–75 [32, 38, 39, 61]	2.82 [39]	18	17.3
	The core epitope between the outer and inner domains, D474, M475 and R476 residues are important for recognition [64]	HJ16	2010	36 [61]	8.01	21	36.7
	The loops D, V1/V2, V5 and CD4-binding loop	VRC01	2010	88–93 [38, 40, 44, 45, 51, 57, 58, 65]	0.09 [45] 0.92 [48]	14	38.8
	The loops D, V1/V2, V5 and CD4-binding loop	VRC02	2010	90–91 [38, 40]	0.13 [40]	14	34.9
	The loops D, V1/V2, V5 and CD4-binding loop	VRC03	2010	51–59 [38, 40, 58]	0.08 [44]	16	34.9
	The loops D, V1/V2, V5 and CD4-binding loop	PGV04 (VRC-PG04)	2011	77–88 [40, 44, 46, 51]	0.14 [40]	16	38.2
	The loops D, V1/V2, V5 and CD4-binding loop	CH31 (VRC-CH31)	2011	84–91 [40, 44, 66]	0.02 [44]	15	31.9
	The loops D, V1/V2, V5 and CD4-binding loop	CH33 (VRC-CH33)	2011	90 [44]	0.24 [44]	15	31.9
	The loops D, V1/V2, V5 and CD4	NIH45-46	2011	84–86 [40, 45, 48]	0.08 [45] 0.41 [48]	18	44
	The loops D, V1/V2, V5 and CD4	45-46 ^{G54W}	2011	92 [48]	0.04 [48]	18	44
	The loops D, V1/V2, V5 and CD4	3BNC117	2011	86–92 [40, 45, 58]	0.06 [40]	12	36.9
	The loops D, V1/V2, V5 and CD4	12A12	2011	92–96 [40, 45]	0.07 [40]	15	34
	The loops D, V1/V2, V5 and CD4	VRC23	2013	65–80 [40, 58]	0.58 [40]	No data	No data

REVIEWS

V1/V2 gp120 loop	Glycans at N160 and N156 and a β -sheet region within the V1/V2 loop	PG9	2009	77–83 [39, 40, 51, 57, 58]	0.08 [58]	30	15.4
		PG16	2009	73–79 [39, 40, 57, 58]	0.02 [57]	30	16.8
		PGT145	2011	78 [51]	0.29	33	22.8
		CH01	2011	46 [50]	3.75 [50]	24	23.3
gp120 V3 loop	Three glycans at N332, N339, N392	2G12	1994	28–39 [39, 40, 61]	1.45 [40]	16	33.6
	Complex-type N-glycans at N332 and V3 loop region	PGT121	2011	70 [51]	0.03	26	21.2
	High-mannose glycans and β -sheet region at the C-end of the V3 loop	PGT128	2011	72 [51]	0.02	21	27.9
CD4i/V3 gp120	The V3 loop	3BC176	2012	64 [67]	12.8 [67]	19	29.4
Epitope at the gp120 / gp41 interface	Glycan-dependent epitope (a cluster of N-glycans terminated with a galactose residue at N611 and N637)	PGT151-155	2014	64–66 [56]	0.008–0.012 [56]	28	No data
	gp120 D and V5 loops	8ANC195	2011	67 [45]	0.87 [45]	9	No data

Note. The rows shaded in grey are for first-generation bNAbs.

* - the percentage of neutralization is expressed as the amount of virus neutralized at IC_{50} values lower than 50 $\mu\text{g/ml}$.

** – polyreactive.

ences can be explained by the slight variation in the epitope binding sites recognized by PG9 and PG16.

Glycosylation of N156 and N160 can affect PG9 and PG16 interaction with trimeric gp120. Unlike 2G12, whose binding requires glycans at N332, N339, and N392 [33], PG9 and PG16 involve both the N-glycosylation sites and amino acids of gp120 encompassed by the V2 and V3 loops [39]. Artificial proteins mimicking PG9 recognition in the gp120 context shed light on the structure of PG9 bound to the target. The CDR H3 loop of HIV broadly neutralizing antibody PG9 plays an essential role in stabilizing the PG9-gp120 complex. Its exceptionally long loop of 30 amino acid residues penetrates the glycan shield on gp120 to allow access to the protein surface around the V2 and V3 loops. The tip of the loop has a hammerhead structure, formed by two β -sheets. The outer β -sheet of the CDR H3 antibody makes four hydrogen bonds to the β -sheet at the base of the V2 loop (*Fig. 3*). Beside the hydrogen bonds, the negatively charged CDR H3 interaction with asparag-

ine-linked sugar moieties at N160 and N156 also contributes to the interaction of PG9 with the epitope. The V3 loop and sugar moieties make up more than 50% of the contact-surface area [14]. This structure partially mimics the natural conformation of trimeric gp120 bound to the cell membrane [49].

PG9 and PG16 were among the first bNAbs used to target the second Env site of vulnerability spanning amino acid residues in the V1/V2 loops and oligomannose moieties at positions 160 and 156 (or 176) (*Fig. 2B*). Later on, CH01-04 [50] and PGT141-145 [51], which recognize the same epitope, were obtained [14]. The hallmark of these antibodies is their exceptionally long CDR H3, allowing penetration between the glycans and interaction with them, which contributes to enhanced binding to gp120 [14].

V3 loop region

The key residues of another site of vulnerability on HIV-1 Env are high-mannose glycans on N332 and a

region within the V3 loop (*Fig. 2C*) [15]. The spectrum of antibodies recognizing this epitope includes the first-generation antibody 2G12, because the sugar moieties of gp120 that comprise the epitope 2G12 form a conformation similar to that of the 2G12 antibody. All other related bNAbs recognize not only carbohydrates, but also amino acids (*Table*). The structural shape of these antibodies and PG9-related antibodies allows them to recognize and pierce the gp120 glycan shield to interact with the protein surface beneath. PGT127-128 antibodies exhibit an elongated CDR H2 loop, and the PGT135 antibody has an extended CDR H1 loop [52].

PGT135 use long loops (CDR H1 and CDR H3) to penetrate the gp120 glycan shield, with CDR H3 playing the critical role. Beside the contact with the protein backbone of gp120, the flanking glycans also contribute to the interaction. CDR H3 interacts with glycans at N332, N386, and N392; CDR H1, only at N386. Like PG9, PGT135 and the carbohydrates of the epitope make up less than half of the overall interaction contacts which considerably contribute to the binding energy [15]. Sugar moieties play a minor role in recognition by another bNAb, PGT128, that interacts with glycan at N332, while N301 employs CDR H3 and CDR H2. The antibody's elongated CDRH2 loop forms extensive interactions with gp120. Importantly, the moderate β -sheet at the tip of CDR H3 and the β -sheet structure of V3 on gp120 make hydrogen bonds critical for PGT128 recognition of Env [53].

Of note are PGT121 and 10-1074, which, in contrast to 2G12, PGT135, and PGT128, bind to complex-type N-glycans present in a low percentage on gp120, rather than high-mannose N-glycans [54].

If arranged in an order according to involvement of carbohydrates in the interaction, 2G12 ranks first for being fully carbohydrate-dependent, followed by PGT135, which is focused on binding to sugar moieties and to a lesser extent to amino acid residues, and PGT128, whose binding is strongly dependent on the CDR H3 contact with V3 amino acids. Interestingly, the neutralization breadth increases with the contribution of protein-protein interactions to epitope binding (*Table*). Following from this, the neutralization breadth seems to be driven by protein contribution to the contact surface. Although the glycan-recognizing antibody 2G12 can neutralize HIV-1 isolates, it demonstrates moderate breadth. By contrast, antibodies targeting the protein surface have increased breadth. The elicitation of glycan-recognizing bNAbs will likely exhibit poor potency.

MPER region

The fourth major site of vulnerability is gp41 MPER. Like CD4bs, it is very conserved and has been pursued

as a target for bNAbs. MPER is critical for fusion and cell entry, hence MPER conservation is required to maintain its functions. Indeed, bNAbs, which recognize this target, were the first to be identified. However, sera of HIV-infected individuals that contained bNAbs against multiple HIV isolates showed that MPER-binding bNAbs are not common. In addition, the monoclonal MPER-directed antibodies 2F5, 4E10, and Z13 are polyreactive. At the same time, Huang *et al* (2012) reported a MPER-specific antibody, named 10E8 (*Table*), which neutralized ~98% of the HIV-1 isolates examined and did not show polyreactivity.

gp120/gp41 interface region

Broadly neutralizing antibodies which recognize both gp120 and gp41 have only recently been discovered [17, 55]. Like the majority of bNAbs (except for gp41 MPER specific antibodies), these antibodies bind Env through glycan-mediated interactions, but by contrast, critical glycan contacts are located on gp41. All the antibodies are structurally similar to other known bNAbs. 8ANC195 has a long CDR H3 loop and a protruding FWR3 (third framework region) of its heavy chain. This structure overcomes the glycan masking contributed by sugars at N234 and N276 to contact the gp120 D and V5 loops [17]. PGT151 has elongated CDR H3 and CDR L1 loops that are capable of recognizing Env in a conformation adopted prior to fusion of the viral membrane with the target cell membrane. Interestingly, PGT151-158-like antibodies can mediate antibody-dependent cellular cytotoxicity [56].

CONCLUSIONS

The identification and investigation of bNAbs against HIV-1 has been a breakthrough in the understanding of the humoral immune response. Although these antibodies fail to prevent AIDS developing from HIV and virus clearance, bNAbs guide us through the remarkable adaptations of B-cell immunity in response to a sophisticated agent such as HIV-1.

Numerous studies of bNAbs have shown that the host's immune system can accommodate HIV-1 escape mutations by generating unusual antibodies directed at hidden conserved epitopes.

The exceptional neutralization potency of bNAbs is mainly due to their structure. First, their hypervariable loops carry amino acid insertions in CDR loops, particularly CDR H3, which allow access to the gp120 protein surface through the glycan canopy. Second, bNAbs have the ability to accommodate epitope diversity by altering the conformation of their loops.

A distinctive feature of multiple bNAbs against HIV-1 is the engagement of the sugar moieties decorating the gp120/gp41 HIV-1 envelope glycoprotein. In

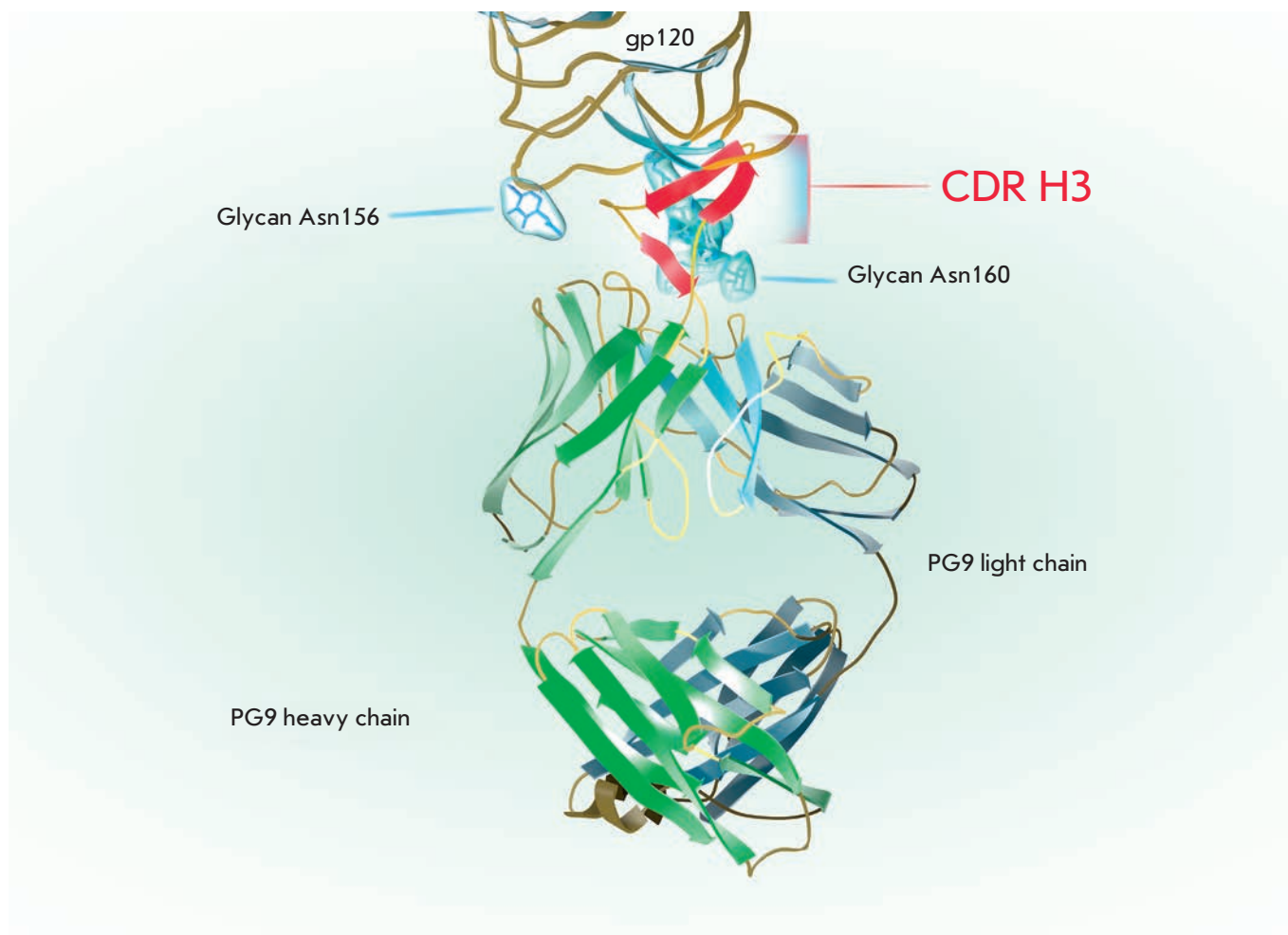


Fig. 3. The PG9 Fab-fragment in complex with HIV-1 gp120 is shown. The variable and conserved domains of the heavy and light chains are in green and grey, respectively. The extended CDR H3 loop reaching through to the gp120 surface is highlighted. The β -sheets of CDR H3, critical for binding, are in red. The N-glycans at N160 and N156 through which the PG9 CDR H3 penetrates are shown as blue clouds. The schematic is reproduced based on the structures of 3U4E and 3DNN from Protein Data Bank

addition to protein-protein interactions, carbohydrates effectively complement the antibody epitope recognition. Interestingly, glycan moieties could contribute equally or more than half to the overall binding energy.

Another important strategy is structural mimicry. There have been generated antibodies towards the site of CD4 attachment on HIV-1 gp120 that mimic important molecular details of the CD4-gp120 interaction.

These insights lend support, on one hand, to the po-

tential of the immune system to address the challenge of pathogen diversity and, on the other hand, to vaccine design leading to the elicitation of potent bNAbs against HIV-1

The authors would like to thank A.I. Shapoval for critical reading of the manuscript and E.A. Kolosova for her excellent illustrations. This work was supported by RSF (contract № 14-14-00660).

REFERENCES

1. Deyev S.M., Lebedenko E.N. // Acta Naturae. 2009. V. 1. P. 32–50.
2. Deyev S.M., Lebedenko E.N., Petrovskaya L.E.E., Dolgikh D.A., Gabibov A.G., Kirpichnikov M. P. // Russian Chemical Reviews. 2015. V. 84. № 1. P. 1–26.
3. Peeters M., D'Arc M., Delaporte E. // AIDS Rev. 2014. V. 16. № 1. P. 23–34.
4. Burton D.R., Barbas C.F., Persson M.A., Koenig S., Chanoock R.M., Lerner R.A. // Proc. Natl. Acad. Sci. U. S. A. 1991. V. 88. № 22. P. 10134–10137.
5. Zhu P., Chertova E., Bess J., Lifson J.D., Arthur L.O., Liu

- J., Taylor K.A., Roux K.H. // *Proc. Natl. Acad. Sci. U. S. A.* 2003. V. 100. № 26. P. 15812–15817.
6. Pancera M., Zhou T., Druz A., Georgiev I.S., Soto C., Gorman J., Huang J., Acharya P., Chuang G.-Y., Ofek G., et al. // *Nature*. 2014. V. 514. № 7523. P. 455–461.
7. Nowak S.A., Chou T. // *Biophys. J.* 2009. V. 96. № 7. P. 2624–36.
8. Johnson W.E., Desrosiers R.C. // *Annu. Rev. Med.* 2002. V. 53. № 1. P. 499–518.
9. Gray E.S., Moore P.L., Choge I.A., Decker J.M., Bibollet-Ruche F., Li H., Leseka N., Treurnicht F., Mlisana K., Shaw G.M., et al. // *J. Virol.* 2007. V. 81. № 12. P. 6187–6196.
10. Frost S.D.W., Wrin T., Smith D.M., Kosakovsky Pond S.L., Liu Y., Paxinos E., Chappey C., Galovich J., Beauchaine J., Petropoulos C.J., et al. // *Proc. Natl. Acad. Sci. U. S. A.* 2005. V. 102. № 51. P. 18514–18519.
11. Moore P.L., Gray E.S., Wibmer C.K., Bhiman J.N., Nonyane M., Sheward D.J., Hermanus T., Bajimaya S., Tumba N.L., Abrahams M.-R., et al. // *Nat. Med.* 2012. V. 18. № 11. P. 1688–1692.
12. Malenbaum S.E., Yang D., Cavacini L., Posner M., Robinson J., Cheng-Mayer C. // *J. Virol.* 2000. V. 74. № 23. P. 11008–11016.
13. Zhou T., Georgiev I., Wu X., Yang Z.-Y., Dai K., Finzi A., Kwon Y. Do, Scheid J.F., Shi W., Xu L., et al. // *Science*. 2010. V. 329. № 5993. P. 811–817.
14. McLellan J.S., Pancera M., Carrico C., Gorman J., Julien J.-P., Khayat R., Louder R., Pejchal R., Sastry M., Dai K., et al. // *Nature*. 2011. V. 480. № 7377. P. 336–343.
15. Kong L., Lee J.H., Doores K.J., Murin C.D., Julien J.-P., McBride R., Liu Y., Marozsan A., Cupo A., Klasse P.-J., et al. // *Nat. Struct. Mol. Biol.* 2013. V. 20. № 7. P. 796–803.
16. Guenaga J., Wyatt R.T. // *PLoS Pathog.* 2012. V. 8. № 7. P. e1002806.
17. Scharf L., Scheid J.F., Lee J.H., West A.P., Chen C., Gao H., Gnanapragasam P.N.P., Mares R., Seaman M.S., Ward A.B., et al. // *Cell Rep.* 2014. V. 7. № 3. P. 785–795.
18. Muster T., Steindl F., Purtscher M., Trkola A., Klima A., Himmler G., Rucker F., Kattinger H. // *J. Virol.* 1993. V. 67. № 11. P. 6642–6647.
19. Buchacher A., Predl R., Strutzenberger K., Steinfellner W., Trkola A., Purtscher M., Gruber G., Tauer C., Steindl F., Jungbauer A. // *AIDS Res. Hum. Retroviruses.* 1994. V. 10. № 4. P. 359–369.
20. Montefiori D.C., Pantaleo G., Fink L.M., Zhou J.T., Zhou J.Y., Bilska M., Miralles G.D., Fauci A.S. // *J. Infect. Dis.* 1996. V. 173. № 1. P. 60–67.
21. Dhillon A.K., Donners H., Pantophlet R., Johnson W.E., Decker J.M., Shaw G.M., Lee F.-H., Richman D.D., Doms R.W., Vanham G., et al. // *J. Virol.* 2007. V. 81. № 12. P. 6548–6562.
22. Gray E.S., Taylor N., Wycuff D., Moore P.L., Tomaras G.D., Wibmer C.K., Puren A., DeCamp A., Gilbert P.B., Wood B., et al. // *J. Virol.* 2009. V. 83. № 17. P. 8925–8937.
23. Simek M.D., Rida W., Priddy F.H., Pung P., Carrow E., Laufer D.S., Lehrman J.K., Boaz M., Tarragona-Fiol T., Miiro G., et al. // *J. Virol.* 2009. V. 83. № 14. P. 7337–7348.
24. Li Y., Svehla K., Louder M.K., Wycuff D., Phogat S., Tang M., Migueles S.A., Wu X., Phogat A., Shaw G.M., et al. // *J. Virol.* 2009. V. 83. № 2. P. 1045–1059.
25. Walker L.M., Simek M.D., Priddy F., Gach J.S., Wagner D., Zwick M.B., Phogat S.K., Poignard P., Burton D.R. // *PLoS Pathog.* 2010. V. 6. № 8. P. e1001028.
26. Thèze J., Chakrabarti L.A., Vingert B., Porichis F., Kaufmann D.E. // *Clin. Immunol.* 2011. V. 141. № 1. P. 15–30.
27. Cao Y., Qin L., Zhang L., Safrin J., Ho D.D. // *N. Engl. J. Med.* 1995. V. 332. № 4. P. 201–208.
28. Sather D.N., Armann J., Ching L.K., Mavrantoni A., Sellhorn G., Caldwell Z., Yu X., Wood B., Self S., Kalams S., et al. // *J. Virol.* 2009. V. 83. № 2. P. 757–769.
29. Doria-Rose N.A., Klein R.M., Manion M.M., O'Dell S., Phogat A., Chakrabarti B., Hallahan C.W., Migueles S.A., Wrarmert J., Ahmed R., et al. // *J. Virol.* 2009. V. 83. № 1. P. 188–199.
30. Hraber P., Seaman M.S., Bailer R.T., Mascola J.R., Montefiori D.C., Korber B.T. // *AIDS*. 2014. V. 28. № 2. P. 163–169.
31. Euler Z., van den Kerkhof T.L.G.M., van Gils M.J., Burger J.A., Edo-Matas D., Phung P., Wrin T., Schuitemaker H. // *J. Virol.* 2012. V. 86. № 4. P. 2045–2055.
32. Burton D.R., Pyati J., Koduri R., Sharp S.J., Thornton G.B., Parren P.W., Sawyer L.S., Hendry R.M., Dunlop N., Nara P.L. // *Science*. 1994. V. 266. № 5187. P. 1024–1027.
33. Scanlan C.N., Pantophlet R., Wormald M.R., Ollmann Saphire E., Stanfield R., Wilson I.A., Kattinger H., Dwek R.A., Rudd P.M., Burton D.R. // *J. Virol.* 2002. V. 76. № 14. P. 7306–7321.
34. Mouquet H., Scheid J.F., Zoller M.J., Krogsgaard M., Ott R.G., Shukair S., Artyomov M.N., Pietzsch J., Connors M., Pereyra F., et al. // *Nature*. 2010. V. 467. № 7315. P. 591–595.
35. Zwick M.B., Jensen R., Church S., Wang M., Stiegler G., Kunert R., Kattinger H., Burton D.R. // *J. Virol.* 2005. V. 79. № 2. P. 1252–1261.
36. Nelson J.D., Brunel F.M., Jensen R., Crooks E.T., Cardoso R.M.F., Wang M., Hessel A., Wilson I.A., Binley J.M., Dawson P.E., et al. // *J. Virol.* 2007. V. 81. № 8. P. 4033–4043.
37. Ferrantelli F., Rasmussen R.A., Hofmann-Lehmann R., Xu W., McClure H.M., Ruprecht R.M. // *Vaccine*. 2002. V. 20 Suppl 4. P. A61–A65.
38. Wu X., Yang Z.-Y., Li Y., Hogerkorp C.-M., Schief W.R., Seaman M.S., Zhou T., Schmidt S.D., Wu L., Xu L., et al. // *Science*. 2010. V. 329. № 5993. P. 856–861.
39. Walker L.M., Phogat S.K., Chan-Hui P.-Y., Wagner D., Phung P., Goss J.L., Wrin T., Simek M.D., Fling S., Mitcham J.L., et al. // *Science*. 2009. V. 326. № 5950. P. 285–289.
40. Georgiev I.S., Doria-Rose N.A., Zhou T., Kwon Y. Do, Staube R.P., Moquin S., Chuang G.-Y., Louder M.K., Schmidt S.D., Altae-Tran H.R., et al. // *Science*. 2013. V. 340. № 6133. P. 751–756.
41. Tong J., Bendahhou S., Chen H., Agnew W.S. // *Nucleic Acids Res.* 1994. V. 22. № 15. P. 3253–3254.
42. Binley J.M., Lybarger E.A., Crooks E.T., Seaman M.S., Gray E., Davis K.L., Decker J.M., Wycuff D., Harris L., Hawkins N., et al. // *J. Virol.* 2008. V. 82. № 23. P. 11651–11668.
43. Li Y., O'Dell S., Walker L.M., Wu X., Guenaga J., Feng Y., Schmidt S.D., McKee K., Louder M.K., Ledgerwood J.E., et al. // *J. Virol.* 2011. V. 85. № 17. P. 8954–8967.
44. Wu X., Zhou T., Zhu J., Zhang B., Georgiev I., Wang C., Chen X., Longo N.S., Louder M., McKee K., et al. // *Science*. 2011. V. 333. № 6049. P. 1593–1602.
45. Scheid J.F., Mouquet H., Ueberheide B., Diskin R., Klein F., Oliveira T.Y.K., Pietzsch J., Fenyo D., Abadir A., Velinzon K., et al. // *Science*. 2011. V. 333. № 6049. P. 1633–1637.
46. Falkowska E., Ramos A., Feng Y., Zhou T., Moquin S., Walker L.M., Wu X., Seaman M.S., Wrin T., Kwong P.D., et al. // *J. Virol.* 2012. V. 86. № 8. P. 4394–4403.
47. Kwong P.D., Mascola J.R. // *Immunity*. 2012. V. 37. № 3. P. 412–425.
48. Diskin R., Scheid J.F., Marcovecchio P.M., West A.P., Klein F., Gao H., Gnanapragasam P.N.P., Abadir A., Sea-

REVIEWS

- man M.S., Nussenzweig M.C., et al. // *Science*. 2011. V. 334. № 6060. P. 1289–1293.
49. Moulard M., Lortat-Jacob H., Mondor I., Roca G., Wyatt R., Sodroski J., Zhao L., Olson W., Kwong P.D., Sattentau Q.J. // *J. Virol.* 2000. V. 74. № 4. P. 1948–1960.
50. Bonsignori M., Hwang K.-K., Chen X., Tsao C.-Y., Morris L., Gray E., Marshall D.J., Crump J.A., Kapiga S.H., Sam N.E., et al. // *J. Virol.* 2011. V. 85. № 19. P. 9998–10009.
51. Walker L.M., Huber M., Doores K.J., Falkowska E., Pejchal R., Julien J.-P., Wang S.-K., Ramos A., Chan-Hui P.-Y., Moyle M., et al. // *Nature*. 2011. V. 477. № 7365. P. 466–470.
52. Corti D., Lanzavecchia A. // *Annu. Rev. Immunol.* 2013. V. 31. P. 705–742.
53. Pejchal R., Doores K.J., Walker L.M., Khayat R., Huang P.-S., Wang S.-K., Stanfield R.L., Julien J.-P., Ramos A., Crispin M., et al. // *Science*. 2011. V. 334. № 6059. P. 1097–1103.
54. Mouquet H., Scharf L., Euler Z., Liu Y., Eden C., Scheid J.F., Halper-Stromberg A., Gnanapragasam P.N.P., Spencer D.I.R., Seaman M.S., et al. // *Proc. Natl. Acad. Sci. U. S. A.* 2012. V. 109. № 47. P. E3268–E3277.
55. Blattner C., Lee J.H., Slieden K., Derking R., Falkowska E., de la Peña A.T., Cupo A., Julien J.-P., van Gils M., Lee P.S., et al. // *Immunity*. 2014. V. 40. № 5. P. 669–680.
56. Falkowska E., Le K.M., Ramos A., Doores K.J., Lee J.H., Blattner C., Ramirez A., Derking R., van Gils M.J., Liang C.-H., et al. // *Immunity*. 2014. V. 40. № 5. P. 657–668.
57. Huang J., Ofek G., Laub L., Louder M.K., Doria-Rose N.A., Longo N.S., Imamichi H., Bailer R.T., Chakrabarti B., Sharma S.K., et al. // *Nature*. 2012. V. 491. № 7424. P. 406–412.
58. Chuang G.-Y., Acharya P., Schmidt S.D., Yang Y., Louder M.K., Zhou T., Kwon Y. Do, Pancera M., Bailer R.T., Doria-Rose N.A., et al. // *J. Virol.* 2013. V. 87. № 18. P. 10047–10058.
59. Binley J.M., Wrinn T., Korber B., Zwick M.B., Wang M., Chappey C., Stiegler G., Kunert R., Zolla-Pazner S., Katinger H., et al. // *J. Virol.* 2004. V. 78. № 23. P. 13232–13252.
60. Cardoso R.M.F., Brunel F.M., Ferguson S., Zwick M., Burton D.R., Dawson P.E., Wilson I.A. // *J. Mol. Biol.* 2007. V. 365. № 5. P. 1533–1544.
61. Corti D., Langedijk J.P.M., Hinz A., Seaman M.S., Vanzetta F., Fernandez-Rodriguez B.M., Silacci C., Pinna D., Jarrossay D., Balla-Jhaghoorsingh S., et al. // *PLoS One*. 2010. V. 5. № 1. P. e8805.
62. Zwick M.B., Labrijn A.F., Wang M., Spenlehauer C., Saphire E.O., Binley J.M., Moore J.P., Stiegler G., Katinger H., Burton D.R., et al. // *J. Virol.* 2001. V. 75. № 22. P. 10892–10905.
63. Pejchal R., Gach J.S., Brunel F.M., Cardoso R.M., Stanfield R.L., Dawson P.E., Burton D.R., Zwick M.B., Wilson I.A. // *J. Virol.* 2009. V. 83. № 17. P. 8451–8462.
64. Pietzsch J., Scheid J.F., Mouquet H., Klein F., Seaman M.S., Jankovic M., Corti D., Lanzavecchia A., Nussenzweig M.C. // *J. Exp. Med.* 2010. V. 207. № 9. P. 1995–2002.
65. Liao H.-X., Lynch R., Zhou T., Gao F., Alam S.M., Boyd S.D., Fire A.Z., Roskin K.M., Schramm C.A., Zhang Z., et al. // *Nature*. 2013. V. 496. № 7446. P. 469–476.
66. Bonsignori M., Montefiori D.C., Wu X., Chen X., Hwang K.-K., Tsao C.-Y., Kozink D.M., Parks R.J., Tomaras G.D., Crump J.A., et al. // *J. Virol.* 2012. V. 86. № 8. P. 4688–4692.
67. Klein F., Gaebler C., Mouquet H., Sather D.N., Lehmann C., Scheid J.F., Kraft Z., Liu Y., Pietzsch J., Hurley A., et al. // *J. Exp. Med.* 2012. V. 209. № 8. P. 1469–1479.

Survival Guide: *Escherichia coli* in the Stationary Phase

P. Pletnev*, I. Osterman, P. Sergiev, A. Bogdanov, O. Dontsova

Moscow State University, Chemistry Department, Moscow, 119991, Russia

*E-mail: philippetnev@gmail.com

Received 13.07.2015

Copyright © 2015 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 This review centers on the stationary phase of bacterial culture. The basic processes specific to the stationary phase, as well as the regulatory mechanisms that allow the bacteria to survive in conditions of stress, are described.

KEYWORDS stationary phase, stress, starvation, survival, *Escherichia coli*.

ABBREVIATIONS GASP – the growth advantage in stationary phase (GASP) phenotype; UTR – untranslated region; TLD – tRNA-like domain; MLD – mRNA-like domain; tmRNA – transport and messenger RNA; PCD – programmed cell death; TA – toxin-antitoxin; QS – bacterial intercellular communication system; VBNC – viable but nonculturable bacteria.

INTRODUCTION

The conditions that sustain constant bacterial growth are seldom found in nature, in contrast to when bacteria are cultured under optimal laboratory conditions in rich media and at an optimum temperature. The influence of harsh environmental factors, accumulation of toxic metabolic waste products during starvation, and antibiotics – all this threatens the survival of *Escherichia coli* and other bacteria. For protection against harsh environmental influences, bacterial culture can enter a stationary phase where its internal systems of protection against stress become activated. In order to survive under adverse conditions, the bacterial culture can dramatically change its organization both at the molecular and cellular levels.

Knowledge of the processes occurring in the stationary phase is necessary for both fundamental and practical viewpoints. Cells in the stationary phase are orders of magnitude more resistant to antimicrobials and acquire the ability to survive even under extremely adverse environmental settings. This review focuses on the basic processes characteristic of the stationary phase.

PHASES OF BACTERIAL CULTURE GROWTH

The growth of a bacterial culture represents a process of sequential division of the cells of the culture to form two identical daughter cells.

Study of *E. coli* cells survival during cultivation for several days revealed a characteristic growth curve pattern comprising five phases. Despite differences in growth conditions, measurements, and even species-

specific features, the general shape of the curve always remains the same, except for some parameters (*Fig. 1*) [1].

The moment when the cells enter into the nutrient medium after being in the stationary phase, one can observe what is usually called the lag phase. This phase is characterized by an almost absence of bacterial growth in culture for some time, which can be attributed to the need for adaptation of cellular metabolism to the new habitat conditions. The duration of the lag phase is determined not only by the bacterial species, but also by the length of time that the cells have remained in starved conditions [2].

Once the cells are adapted to the new cultivation conditions, they start to divide exponentially and enter a logarithmic growth phase. Since bacterial cells divide asexually by binary fission, an increase in the number of cells in the medium per unit of time is well approximated by an exponential function. The growth rate of culture in the logarithmic phase is characterized by the number of doublings. It is worth noting that this rate depends directly on the culture medium, being slower in nutrient-poor conditions and faster in nutrient-rich conditions. The standard laboratory strain *E. coli* MG1655 K-12 has a doubling time of about 30 min at 37°C.

Once nutrients in the medium are exhausted, bacterial culture enters a stationary phase, which is characterized by equilibrium between the numbers of dividing and dying cells and represents a plateau in the growth curve. It should be noted that the term “stationary phase” refers specifically to the region on the growth curve which is characterized by equilibrium

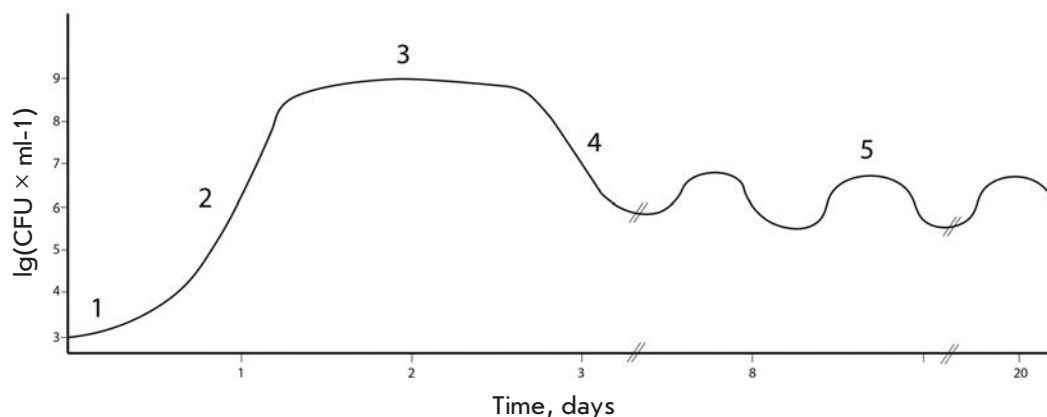


Fig. 1. Bacterial growth curve. 1 – lag-phase, 2 – logarithmic phase, 3 – stationary phase, 4 – death phase, 5 – long-term stationary phase

between dividing and dying cells, but not to the mechanism of defense during starvation. The stationary phase begins in the cell population not only due to the exhaustion of the external environment, but also because of the various other stress factors. Over time, cultures in the stationary phase accumulate toxic products of catabolism in the environment, leading to a decline in the number of viable cells, known as the death phase. Stochastic death and programmed cells death have been postulated to be responsible for this phase.

The end of the death phase comes after the majority of the population dies and the dead cells release nutrients into the environment. Survivors may use these substances for their survival that brings the bacterial culture to a state of long-term stationary phase, with viability remaining for several weeks and even months. One of the characteristics of the long-term stationary phase is the successive increase and decrease in the titer of viable cells in the population. This phenomenon is referred to as the growth advantage in the stationary phase (GASP) phenotype and is explained by the fact that mutant cells, which are more adapted to grow under these conditions than the parent strain, appear among the bacterial population [3].

CHANGES IN DNA STRUCTURE AND TOPOLOGY

Genomic DNA of *E. coli* is represented by a single circular chromosome, which forms a structure in the cytoplasm of bacteria called the nucleoid. This structure also includes nucleoid proteins (regulatory and structural) and RNA [4].

A multitude of proteins are responsible for maintaining the nucleoid structure, whose expression depends on the growth phase of the bacterial culture. The IHF, HU, Dps, Fis, and H-NS proteins are considered the most essential structural proteins of the nucleoid.

The active form of H-NS is a dimer characterized by the presence of two oppositely directed DNA-binding domains making it possible for the protein to act as a

“bridge” between two DNA duplexes. H-NS is not sequence-specific, but it exhibits greater selectivity for bent DNA rather than linear DNA [4]. In the logarithmic growth phase, one molecule of H-NS in the cell accounts for 1,400 bp DNA [5].

The integration host factor (IHF) is a heterodimeric protein exhibiting specificity for consensus regions of DNA of approximately 30 bp. Binding of IHF causes DNA bending, which is stabilized by the interaction of the negatively charged DNA backbone and the mostly positively charged surface of the protein. It is shown that IHF binding can reduce the length of the DNA by 30% [4]. Expression of IHF is at a maximum in the stationary phase, and one molecule of IHF accounts for 335 bp genomic DNA [5]. Probably, IHF is responsible for the organization of the nucleoid structure in the early stationary phase.

The histone-like HU homodimer protein is composed of either two HU α or two HU β subunits. HU exhibits a high (40%) structural similarity to the protein IHF. HU binds to DNA nonspecifically, but it has selectivity for overwound and unordered DNA. HU seems to be able to induce and stabilize the bend of the double DNA helix with a variety of angles of rotation. Random HU binding leads to a large number of “mobile” bends of DNA (with bending angles of 180°), which eventually reduces the length of linear DNA by 50% [4]. HU content is highest in cells during the logarithmic phase, where one molecule of the protein accounts for approximately 550 bp DNA [5].

The factor for inversion stimulation (Fis) is a homodimer DNA-binding protein capable of recognizing certain consensus sequences of 15 bp in length, but it can also efficiently bind DNA at random sites. Binding of Fis produces DNA bending by 50–90°. Many Fis binding sites are located in the promoter regions of operons, where binding of the protein plays a regulatory role. Fis is believed to be the “sensor” of DNA supercoiling. Depending on the topology of DNA, Fis exhibits the ability to

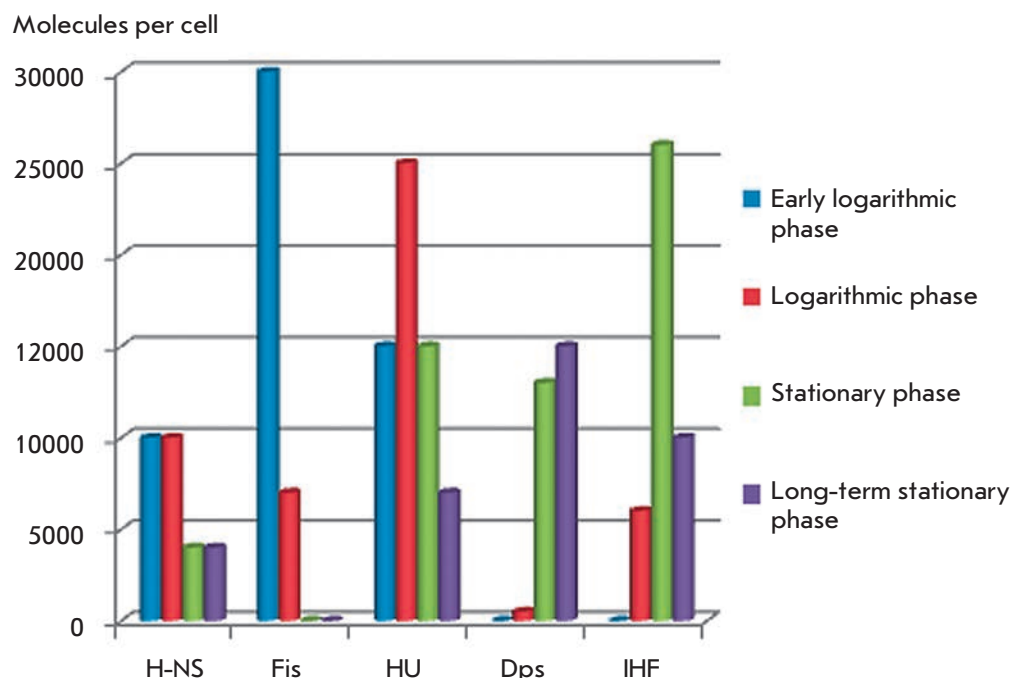


Fig. 2. Normalized (H-NS, Fis, HU, IHF – dimers, Dps – dodecamers) amounts of structural nucleoid proteins during different growth phases. Based on data from [5]

down regulate the expression of the DNA gyrase gene [4]. Fis is one of the structural proteins of the nucleoid most represented in the logarithmic phase, the content of which reaches 1 molecule per 450 bp DNA [5].

In the stationary phase, the nucleoid becomes more condensed to protect DNA from damage. This mechanism is implemented by means of the Dps protein (DNA-binding protein from starved cells), which is capable of non-specific binding with DNA and is active exactly in periods of starvation [6]. Under oxidative stress in the logarithmic phase, the expression of the Dps-encoding gene is under the control of σ^{70} -subunit of RNA polymerase and OxyR protein, and during periods of starvation it is regulated by σ^{38} -subunit [7]. After the induction of synthesis in the stationary phase, Dps becomes the most represented protein in *E. coli* cells [7]. Monomers of Dps form ring-like dodecamer structures, which bind to DNA in the presence of Mg^{2+} and promote the formation of a highly ordered and stable nucleoprotein complex called “biocrystal” [8]. It is the formation of this complex that leads to the condensation of the nucleoid. It seems that the global protective role of Dps against different forms of stress (starvation, oxidative stress, UV and γ -irradiation, thermal stress and pH) is implemented via a combination of several of its properties: the ability to condense DNA, chelate iron ions, and exert ferroxidase activity, as well as the ability to regulate gene expression [6, 9].

CbpA (Curved DNA binding protein) is another protein which helps protect DNA from damage in the sta-

tionary phase. In the log phase, CbpA is absent in the cells, but upon onset of the stationary phase its amount increases up to 10,000 copies per cell. Transcription of the *cbpA* gene also depends on σ^{38} -subunit of RNA polymerase [5].

CbpA binds DNA being in the form of a dimer, resulting in DNA compactization. This complex protects DNA from *in vitro* degradation by endonucleases [10].

The structural proteins of the nucleoid influence directly not only the structure of the bacterial chromosome, but they are also actively involved in the regulation of gene expression. It should be noted that these proteins have similar functions in compactization and protection of DNA from damage, and the prevalence of each of them depends on the growth phase (Fig. 2). However, due to differences in the mechanisms and ways of compactization the use of either structural protein of the nucleoid allows the cell to adapt, making access to DNA easy under the most favorable conditions and providing maximum protection of the genetic material against stress.

It has been shown recently that the methylation of cytosine residues in bacterial DNA can influence the regulation of protein synthesis in the stationary phase. In the study of a DNA methyltransferase Dcm gene knockout strain, it was found that the strain exhibits considerably increased synthesis of proteins of the stationary phase, and particularly, the RNA polymerase σ^{38} -subunit [11].

TRANSCRIPTIONAL REGULATION IN THE STATIONARY PHASE

σ^{38} – Stationary phase sigma-factor

Transcription is a central and vital process when RNA is synthesized on a DNA template. This reaction is catalyzed by RNA polymerase.

In bacteria, transcription is initiated by the RNA polymerase holoenzyme that is formed by a multisubunit ($\alpha_2\beta\beta'\omega$) core, which contains the active center and performs RNA synthesis, and the initiation factor – σ -subunit, or σ -factor. In order to form an active holoenzyme capable of synthesizing RNA, σ -subunit, which is responsible for promoter recognition, must bind to RNA polymerase [12]. It is known that the *E. coli* genome encodes seven different σ -factors (Table), each capable of recognizing only a certain group of promoters. It is the σ -subunit which is the primary regulator of cellular transcription. The use of different σ -factors allows the cell to dramatically change their transcriptome in response to various signals.

Under the conditions of the stationary phase, a bacterial cell has to regulate transcription in such a way as to activate the expression of the genes required for survival under stress and starvation and to suppress the transcription of “unnecessary” genes. *E. coli* uses σ^{38} (σ^S)-factor encoded by the *rpoS* gene for this purpose, which acts as the main regulator of transcription in response to various forms of stress. Genome-wide analysis of gene expression dependent on σ^{38} -factor showed that σ^{38} directly or indirectly regulates the transcription of about 10% of *E. coli* genes [20].

σ^S -factor is responsible for the transcription of the genes involved in stress response and secondary metabolism. Most of the genes regulated by σ^{38} -subunit undergo additional regulation. This factor is required for the transition of the bacterial culture to the stationary

phase. There are some σ^{38} -dependent genes whose expression is maximal at the time of transition of the cultures from the exponential into the stationary phase [14].

σ^S -factor is a homologue of the main cellular sigma factor – σ^{70} , which is responsible for the transcription of housekeeping genes and has the highest affinity for RNA polymerase. It is shown that σ^{38} -subunit recognizes the same consensus sequences as σ^{70} , the most significant of which are the -10 and -35 elements. It is assumed that the difference between the promoters that are recognized by these sigma factors consists in single-base substitutions in the region of consensus hexamers: i.e., σ^{38} specificity can be determined by a small deviation of the hexamer sequence from the consensus sequence [21]. σ^{38} -RNA polymerase is reported to be selective for promoters with non-optimal sequences for σ^{70} at the region of the -10 and -35 elements [22]. Based on the known promoter sequences, it was shown that an A/T-rich region in the -10/+1 region can improve promoter recognition by σ^{38} -RNA polymerase [14].

Expression of σ^{38} -subunit undergoes complex regulation at all levels (transcription, translation, factor activity, stability of σ^{38} and its mRNA), which apparently allows for enhancement of the sensitivity of the cellular response to a variety of stress signals (Fig. 3).

rpoS transcription is repressed by the phosphorylated form of the ArcA regulator [23], and a complex of cAMP-CRP also represses transcription [24]. The BarA protein is necessary for inducing expression in the logarithmic phase [25]. The signaling molecule ppGpp positively influences the basal level of synthesis of σ^S [26]. The expression of *rpoS* increases by an order of magnitude in the transition from the logarithmic to the stationary phase [24].

It was shown that a multitude of factors influence the secondary structure, stability, and ability of the *rpoS* mRNA of translation. The *rpoS* gene mRNA contains a long 5'-UTR [27], which plays an important role in the regulation of translation and stability of this mRNA. The regulator H-NS binds to mRNAs and promotes its decay [28]. In turn, the small DsrA RNA stabilizes mRNA and promotes translation initiation via unfolding of the secondary structure of the mRNA in the ribosome entry site region [29]. The Hfq protein is necessary for translation of *rpoS* mRNA [30], and small OxyS RNA represses translation of σ^S -subunit most likely due to changes in the activity of Hfq [31]. Small RprA RNA and HU protein stimulate translation of σ^{38} -subunit. Under phosphorus starvation, σ^S accumulates due to increased amounts of *rpoS* mRNA. Protein LrhA, together with Hfq, is able to repress the translation of *rpoS* mRNA [32].

It is known that 6S RNA activates transcription from certain σ^S -dependent promoters without influencing

List of σ -factors of *E. coli*

Sigma-factor	Function	Reference
RpoD (σ^{70})	Housekeeping gene expression	[13]
RpoS (σ^{38})	Initiation of the stationary phase and stress response	[14]
RpoF (σ^{28})	Synthesis of flagella and chemotaxis	[15]
RpoN (σ^{54})	Activation of nitrogen metabolism	[16]
RpoH (σ^{32})	Response to heat shock	[17]
RpoE (σ^{24})	Response to stress associated with membrane damage	[18]
FecI (σ^{19})	Expression of nitrate transport genes	[19]

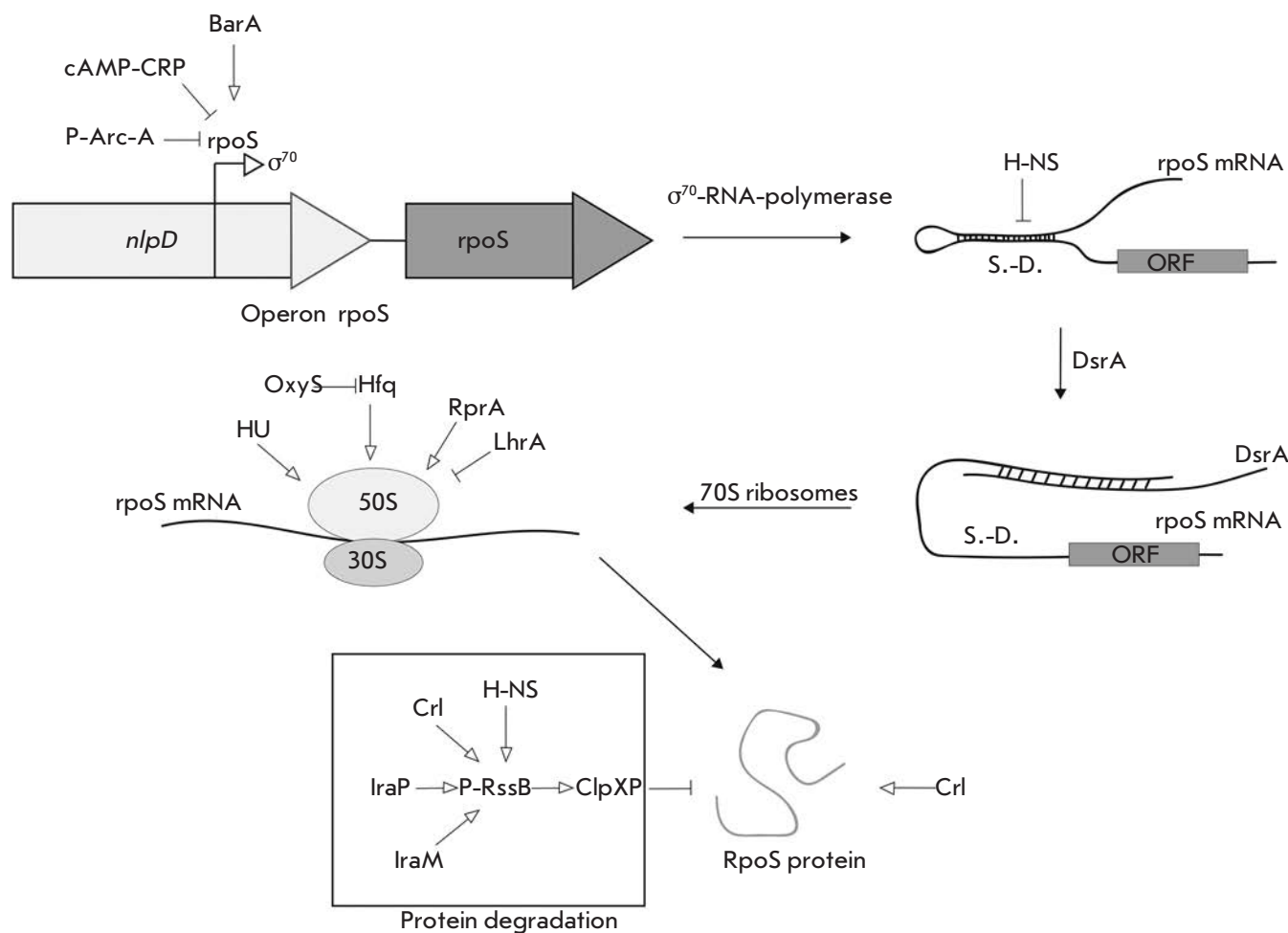


Fig. 3. Regulation of σ^{38} expression and activity. The upper left side of the figure shows the regulation of σ^{38} (*rpoS*) gene transcription. Effect of transcription regulatory proteins shown by arrows. The right-hand part of the figure illustrates the effect of the H-NS protein and DsrA RNA on the secondary structure of *rpoS* mRNA. Shine-Dalgarno sequence designated as "S.-D.", open read frame designated as ORF. Proteins affecting the translation of *rpoS* mRNA are shown in the center of the picture. Proteins affecting the stability of σ^{38} -protein are shown in the lower part of the picture

the level of *rpoS* expression [33]. The signaling molecule ppGpp enhances the ability of σ^S to displace σ^{70} from the minimum enzyme RNA polymerase [34]. It has been shown that the protein Rsd performs a similar function [35]. Lack of nitrogen leads to the expression of genes regulated by σ^S ; however, the level of expression of the subunit itself increases only 2-fold: i.e., nitrogen scarcity apparently influences the activity of the sigma factor more than its expression [36]. RNA polymerase assembly factor Crl increases the activity of σ^S , influencing its ability to bind RNA polymerase [37].

In the logarithmic growth phase, σ^S -subunit is degraded by energy-dependent ClpXP protease, which instantly cleaves factor σ^S under excess energy in the cell [38]. Degradation of σ^S by ClpXP protease requires

an additional protein, RssB, to promote rapid proteolysis of σ^S . RssB recognizes σ^S -factor. Poly (A)-polymerase and the IraP, IraM, H-NS, and Crl proteins enhance the RssB-mediated effect [37, 39–41]. Transcription of *rssB* is under the control of σ^S . Lack of carbon sources leads to the accumulation of the sigma factor due to an increase in its stability. The molecular mechanism of this process has not been investigated [42].

In summary, it should be noted that *rpoS* is expressed in response to sudden adverse changes in environmental conditions. In this, the complex system regulating the expression of this gene and momentary decay of σ^S under optimal growth conditions permit the cell to efficiently change its transcriptional profile in response to stress and quickly return to the use of σ^{70} when adverse conditions no longer exist.

Some transcription regulators in the stationary phase

Transcriptional regulation in the stationary phase is not limited to the change of sigma factors. The bacterial cell contains many regulators that are specific to the stationary phase to alter the expression of certain genes.

One of these regulators is a highly conserved bacterial protein, Lrp (leucine-responsive regulatory protein), which can act both as a transcriptional repressor and activator. This protein is one of the main regulators in the stationary phase influencing more than 400 genes of *E. coli*, with about 75% of these genes active exactly in the stationary growth phase. Among these are genes whose products are responsible for the biosynthesis of amino acids, catabolism, the transport system of nutrients, pili synthesis, and the use of various carbon sources [43]. The main role of Lrp is adaptation of cellular metabolism to environmental settings. Interestingly, Lrp increases the amino acid anabolism level but reduces the level of their catabolism [44].

lrp gene expression is upregulated by means of a signaling molecule, ppGpp. Binding of leucine may also affect its activity. Lrp is able to activate the expression of the genes required during starvation, and to repress genes that are active in the logarithmic growth phase. It is assumed that the mechanism of sensitivity to starvation is based on the binding of leucine molecules, reduction in the intracellular concentration of which may be a sign of this condition [45].

Mutations that disrupt the function of the DNA binding domain of Lrp enhance the effect of GASP-phenotype, in particular, because the cells with this mutation are able to more efficiently metabolize certain amino acids [3].

Not only protein regulators, but also small RNAs affect gene expression in the stationary phase. These RNAs can stimulate translation and affect the stability of specific mRNAs. The genome of *E. coli* contains more than 60 genes of small RNAs, a part of which is responsible for the regulation of the stress response. Bacterial small RNAs are short RNAs of 80–100 nucleotides. The activities of many of them require binding to the chaperone Hfq [46] capable of forming a complex with AU-rich regions of RNAs, whereby it can stabilize the mRNA or, alternatively, enhance the hydrolysis and inhibit its translation. Small DsrA and RprA RNAs stimulate translation of σ^S -factor. Under optimal growth conditions, the 5'-UTR of *rpoS* mRNA has a secondary structure that blocks the ribosome entry site. Small RNAs of DsrA and RprA are capable of interacting with the 5'-UTR of *rpoS* mRNA via complementary regions, which changes the secondary structure of mRNA and opens the ribosome entry site [47].

Another small RNA, OxyS, appears under oxidative stress and represses translation of RpoS by competitive binding of RNA chaperone Hfq [31]. The other small RNAs active in the stationary phase are MicA and RybB, which are involved in the regulation of the outer membrane permeability. It is the outer membrane that serves as the first line of defense in the contact with the environment. To protect the cells from damage, the composition of the membrane changes to allow the cell to endure periods of stress. MicA and RybB with Hfq are understood??? to cause antisense inhibition of translation. RybB RNA controls the expression of two proteins, the outer membrane components – OmpC and OmpW. In turn, the small MicA RNA causes the decay of mRNA of the OmpA outer membrane protein [48].

The stringent response

Inhibition of rRNA synthesis under amino acid starvation was one of the first mechanisms of gene expression regulation in bacteria ever described at the molecular level. Genetic analysis made it possible to identify a mutation that leads to the absence of reduced rRNA synthesis in response to amino acid starvation. This mutation was described as “attenuating the strictness” of the influence of the number of amino acids on the biosynthesis of RNA [49]. Later, it was shown that this mutation inactivates the *relA* gene, encoding (p)ppGpp-synthase [50]. It was shown further that the synthesis of this nucleotide regulator is a response of the cell to stress. This regulatory system controls replication, translation, transcription, and the activity of the enzymes of the stress response [51].

Synthesis of ppGpp is performed by two proteins with similar functions – RelA and SpoT. RelA, or ppGpp-synthase I, only synthesizes guanosyltetraphosphate, while SpoT shows double catalytic activity: ppGpp synthesis (ppGpp-synthase II) and its degradation (ppGpp-hydrolase). The activity of RelA and SpoT is regulated by different mechanisms. RelA is responsible for the transmission of the starvation signal of one or more amino acids, and SpoT senses carbon, phosphorus, iron, or fatty acids scarcity [3].

Under the abundance of nutrients, RelA is associated with 70S ribosomes. In the case of amino acid starvation, deacylated tRNA accumulates in the cell. This tRNA, being in excess, can enter into the A-site of the ribosome, the ribosome then stops, which leads to dissociation of the RelA ribosome complex. In the free form, RelA is capable of catalyzing the transfer of pyrophosphate from ATP or GTP to GDP [52]. One feature of this mechanism is that RelA responds to the absence of a single amino acid, even if other amino acids are present in sufficient amounts [50].

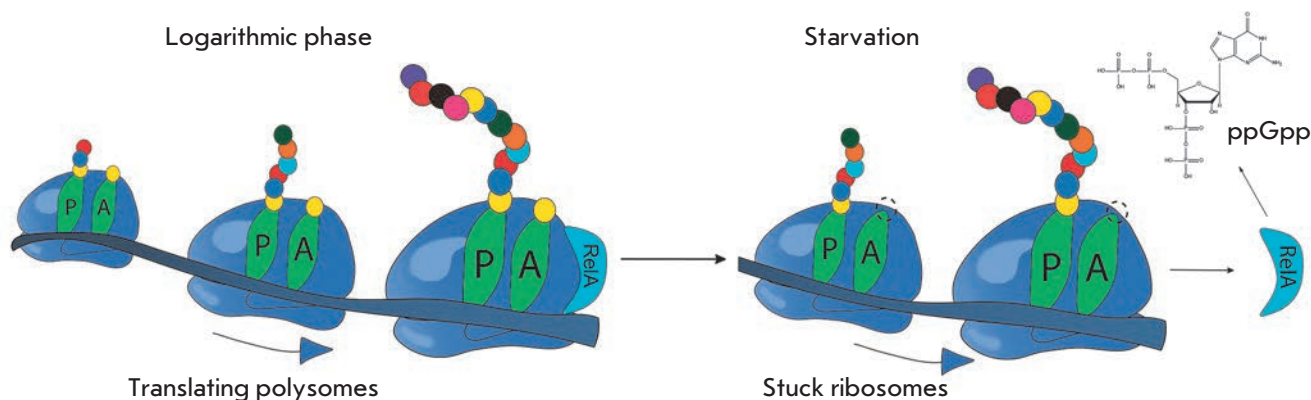


Fig. 4. Scheme of RelA-dependent synthesis of ppGpp during amino acid starvation

It is known that a bacterial cell contains very few molecules of RelA that, for a long time, could not be correlated to the experimentally observed rate of ppGpp accumulation under amino acid starvation. It was proposed that when stuck ribosomes appear and RelA dissociates from ribosomes, a single act of dissociation is accompanied by the synthesis of the ppGpp molecule. After this, free RelA can “hop” to the neighboring ribosome, translating mRNA. If it is also unable to conduct the synthesis because of deacylated tRNA in the A site, the cycle is repeated, and if it is active, RelA remains bound to this ribosome in the inactive form (Fig. 4) [52].

Under favorable growth conditions, SpoT exerts only hydrolytic activity towards ppGpp, which leads to absence of guanosyltetraphosphate in the cell. The hydrolytic activity of SpoT is repressed upon binding to deacylated tRNAs. Because of this mechanism of activation, SpoT possesses sufficient activity only under deficiency of a large number of different amino acids. SpoT is believed to be associated with an acyl-carrying protein that allows it to control the amount of fatty acids in the cell. Probably, this mechanism also allows it to “feel” carbon starvation [53]. At a low concentration of fatty acids, synthesis of ppGpp is induced. With the help of the DksA partner protein, ppGpp binds to β -subunit of RNA polymerase, directly affecting the affinity to different promoters and thus altering the expression level of more than 80 genes. Particularly important is the suppression of the expression of all components of the protein biosynthesis system: rRNA, ribosomal proteins, and translation factors [54].

ppGpp, together with antisigma factor Rsd, helps σ^{38} -subunit to compete for the enzyme base of RNA polymerase by reducing the affinity of σ^{70} to RNA polymerase. When environmental conditions become favorable, the hydrolytic activity of SpoT is restored

and the ppGpp level decreases, which means the end of the stringent response [34].

Translation process in *E. coli* in the stationary growth phase

Protein synthesis is one of the most important processes in the cell. The key actor here is the ribosome; a compound nucleoprotein complex capable of synthesizing proteins according to information encoded in mRNA. In the stationary phase of a bacterial culture, there is a sharp decrease in the level of protein synthesis, which is not surprising, since translation is considered the most energy-consuming processes in the cell, and under deficiency of amino acids and other resources in a bacterial cell it is necessary to suppress translation. It should be noted that the processes affecting translation in the stationary phase are highly dependent on the duration and grade of the state of starvation. In response to starvation, a variety of mechanisms to rescue a single cell and, later, the entire bacterial population, are activated.

Defense mechanisms at minor starvation

When nutrients in the cell are exhausted, deacylated tRNA and truncated mRNA accumulate. Ribosomes can become stuck on these mRNA, since during the synthesis the ribosome reaches the end of mRNA and does not find a stop codon. This ribosome remains bound to mRNA and cannot be released due to the fact that the mechanism for termination of translation depends on the presence of a stop codon. This ribosome can be “rescued” through the mechanism of trans-translation.

Trans-translation is performed by a complex of an unusual tmRNA (transport and messenger RNA) with a small protein, SmpB. tmRNA consists of two domains: the tRNA-like domain (TLD) and mRNA-like domain

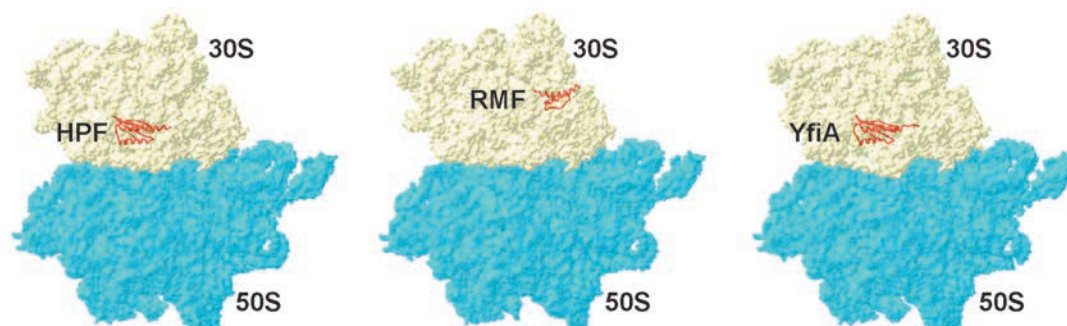


Fig. 5. Binding sites of the hibernation factors RMF (PDB: 4V8G), HPF (PDB: 4V8H), and YfiA (PDB: 4V8I)

(MLD). The similarity of the TLD-domain and tRNA is not only structural, but also functional. This domain is recognized by alanyl-tRNA synthetase, and the 3'-end of tmRNA is charged with an alanine residue [55].

In the case of stuck ribosomes, cooperation between tmRNA, SmpB, and the EF-Tu elongation factor allows the cell to recognize these ribosomes [56], after which the complex SmpB and tmRNA enters the A-site of the ribosome, where SmpB takes a shape mimicking the anticodon structure of alanine tRNA. Then, by means of hydrolysis of the GTP molecule the peptide from the tRNA in the P-site is carried to the alanine residue of tmRNA in the A-site, and translation resumes on the MLD-domain template of tmRNA [56]. C-terminal peptide is encoded in the MLD, signaling the need for degradation of this polypeptide chain. As a result of trans-translation, the stuck ribosome is freed and the potentially harmful polypeptide is degraded by proteases [57].

Another way to rescue a ribosome which has arrived at the 3'-end of mRNA and has not met a stop codon is the use of the ArfA and ArfB proteins. The *ArfA* gene encodes a short polypeptide consisting of 72 amino acid residues. The functional form of the protein consists of 55 amino acids and is translated from a fragment of the mRNA truncated by the ribonuclease III. Due to the lack of a stop codon, emergence of this polypeptide is possible only in case of a disturbed trans-translation mechanism. ArfA binds to the stuck ribosome and recruits the RF2 termination factor thereto, which leads to cleavage of the polypeptide chain from the peptidyl-tRNA and ribosome release [58]. Factor ArfB also acts similarly, binding to the empty A site of ribosomes and catalyzing the hydrolysis of peptidyl-tRNA independently of translation termination factors [59].

Hibernating ribosomes as a response to growing starvation

Synthesis of ribosomes is a highly energy- and resource-consuming process. That is why there must be mechanisms to suppress translation impermissible for starved cells, but all the while retain existing ribosome until better times.

It turns out that a mechanism for storing inactive ribosomes does exist and represents a temporary “switch off” of ribosomes. This process, referred to as ribosome hibernation, is activated within the stringent response, when deacylated tRNAs accumulate in the cell under a deficiency of amino acids, which serves as a signal to the synthesis of the molecule ppGpp using the enzyme RelA associated with ribosomes [52]. It is ppGpp that regulates the expression of the genes that encode the proteins of ribosome hibernation.

The main route of hibernation is the formation of “sleeping” 100S dimers and 70S inactive monomers from active 70S ribosomes. HPF, RMF, and YfiA are the major hibernation proteins of *E. coli* (Fig. 5).

The first two proteins are responsible for the formation of inactive 100S dimers. In the first step, the RMF protein binds to the 16S rRNA that interacts with the Shine-Dalgarno sequence of the mRNA. It is this area that is critical for the initiation of translation in prokaryotic cells, and the binding itself of this protein inhibits translation. It is particularly important that the RMF cannot bind to the translating ribosomes that prevent formation of under-synthesized proteins that may be toxic to a cell. Binding of the hibernation factor in this region leads to a rotation of the head of the small subunit. This change in conformation contributes to the formation of 90S ribosomal dimers. After dimerization of two ribosomes and 90S dimer formation, the HPF hibernation factor binds, which further stabilizes this structure and closes access for tRNA to the A- and P-sites of ribosomes. The resulting complex of two 70S ribosomes, together with the HPF and RMF proteins, has a sedimentation coefficient of 100S. Dimerization of ribosomes is reversible. When a cell enters a nutrient-rich environment, the complex of hibernation factors and ribosomes dissociates to form two active 70S ribosomes [55, 60].

Another path of ribosome hibernation is binding of the Y (YfiA) protein by 70S ribosomes, which leads to the formation of inactive monomers 70S. The N-terminal part of the Y protein is similar to protein HPF; it also binds within the entry site of tRNA. Moreover, the

C-terminus of the protein prevents RMF protein binding to the anti-Shine-Dalgarno sequence on 16S rRNA. After binding to the ribosome, YfiA inhibits its activity and prevents dissociation into individual subunits. When a cell enters favorable settings, the Y protein leaves the ribosome and translation restarts [60].

Literature data provide evidence of the complex influence of the stringent response on protein synthesis in a bacterial cell. Under the effect of ppGpp, not only is the expression of the biosynthetic apparatus suppressed, but translation itself is inhibited via hibernation factors.

Translation under conditions of prolonged starvation. Death phase and programmed cell death

When nutrients almost completely disappear from the environment and a critical amount of toxic metabolic waste products accumulate, programmed cell death (PCD) is activated in order to destroy most of the population and retain a small number of living cells. The population by this “sacrifice” reduces the load on the remaining bacteria, which will resume their generation when new resources appear.

This is the translation system which is particularly sensitive to intensifying starvation. As already mentioned, using the components of the translation machinery, in particular, ribosomes, a starvation signal is transmitted to all components of a bacterial cell [55]. The PCD mechanism is based on a toxin-antitoxin (TA) system: a set of two genes that together encode both a protein “poison” and a corresponding “antidote” antitoxin. Degradation of the antitoxin and downregulation of its expression leads to accumulation of active toxin and cell death.

The most studied and important unit of TA is the *mazEF* system revealed in a multitude of prokaryotes. This unit consists of two genes located in the same operon – *mazF* and *mazE*. The first gene encodes a stable cytotoxic protein and the second a labile antitoxin protein, which is easily degraded by ATP-dependent protease ClpAP. Under normal conditions, genes of both proteins are expressed, which does not allow the toxin to influence the cell because of the formation of a toxin-antitoxin complex. The starvation activates the synthesis of ppGpp, which suppresses transcription of the *mazEF* operon. After that, *mazE* is rapidly degraded and *mazF* is released. Toxicity of *mazF* is caused by its endoribonuclease activity, specific for the ACA sequence in mRNA, as well as a 3'-end fragment of the 16S rRNA [61]. After cleavage of the 3'-end of 16S rRNA, ribosomes “lose touch” with the anti-Shine-Dalgarno sequence that drives the translation of canonical mRNA. It has been previously shown that these ribosomes exhibit high selectivity towards the synthesis of

small proteins, including both “death” proteins, killing the cell and the proteins necessary to retain small cell populations. It is assumed that in this way the system *mazEF* mediates programmed cell death, leading to the death of most of the population and the continued survival of a small sub-population [62].

Later, it was showed that this system of PCD is regulated by the signal peptide referred to as EDF (Extracellular Death Factor) and having the sequence NNWNN. This peptide is a component of a system of intercellular communication among bacteria called quorum sensing (QS). At a critical population density in a bacterial culture, EDF peptide, which is able to easily penetrate the cells, appears. This peptide significantly increases the activity of *mazE* and reduces the ability of *mazF* to inhibit the toxin. Thus, programmed cell death is activated in all cells in which EDF enters, and only a small population of cells remains intact and survives [63].

Structural features of cells in the stationary growth phase

The transition into the stationary growth phase, as mentioned above, is accompanied by the accumulation of factor σ^{38} . Transition to σ^{38} influences not only metabolic and regulatory pathways, but also dramatically alters the physiology of the bacterial cell. It has been shown that the genes whose expression is controlled by σ^{38} are involved in the change of cell morphology, stress response, metabolic adaptation to starvation, and long-term survival in the stationary phase.

The bacteria in the stationary phase undergo adaptation of morphology critical for survival. The cells become smaller as a result of two processes: reductive division and formation of dwarf cells [64].

Reductive division is caused by the fact that the processes of DNA replication and cell division are initiated at the time of entrance of cells into the stationary phase, when, due to deficiency of resources, further growth of the cells is blocked at all levels of regulation. The cells cannot divide in the stationary phase, and because of the possibility of accidental initiation of DNA replication, cells with a duplicated number of chromosomes appear. As a result, the bacterial culture in the stationary phase becomes extremely heterogeneous in its chromosomal composition. Bacterial cells may comprise even more than two chromosomes [65]. The reason for the heterogeneity of the cell by the chromosome content may be that, in some cells during growth, a minimal distance between the nucleoids had not been achieved, which does not allow the cell division machinery to act. Another reason may be that, in the stationary growth phase, the termination of replication and DNA decatenation processes are disrupted in the cells

[65]. When considering cell survival, mutation rate, and genomic stability, one should consider the fact that the cells in the stationary phase contain different numbers of chromosomes. The cells formed by reductive division are saphenous in morphology. This morphological feature could be explained by the influence of the *bolA* gene, which is actively expressed in the stationary phase under the control of RpoS.

It is most likely that the influence of BolA on cell morphology is determined by the transcription regulation of the *dacA* (gene encoding the penicillin binding protein 5, PBP5), *dacC* (a gene coding for a penicillin binding protein 6, PBP6), and *ampC* (the gene encoding β -lactamase) genes. These proteins have a D,D-carboxypeptidase activity and are involved in the formation of precursors of the peptidoglycan membrane layer affecting the degree of cross-linking of cell wall peptidoglycans by regulating the number of components available for crosslinking [66]. Reductive division is a form of cell adaptation to adverse environmental conditions. Reductive division allows cells to gain advantages under conditions of starvation due to increased the surface to volume ratio, producing more spherical cells. It is worth noting that reductive division is not induced by starvation and is determined by starvation onset at the time of active cell division.

The formation of dwarf cells, unlike reductive division, is activated by starvation. This process is characterized by a steady decrease in the size of the cells due to degradation not only of endogenous resources, but also of the cell wall, particularly the cytoplasmic membrane and the cell wall. In some Gram-negative bacteria such as *E. coli*, in the stationary phase the outer membrane is not degraded and is not compressed, like the inner one, which leads to an increase in the periplasmic space [67].

A distinctive feature of the adaptation to the stationary phase is the formation of a cell wall that can efficiently withstand harsh environmental settings. The formation of this intensified barrier includes extensive changes at all levels of bacterial membranes: the inner and outer membranes, periplasm, and peptidoglycans. In the outer membrane, polysaccharide concentration increases, the amount of proteins decreases, and the number of molecular crosslinks between lipoproteins of the outer membrane and the peptidoglycan layer increases. Oligosaccharides, such as trehalose, accumulate in the periplasm to act as osmoprotectants. The peptidoglycan layer (peptidoglycan is a strong and elastic polymer that serves as the stress-bearing component of the bacterial cell wall) increases its thickness. Recently, it has been shown that in the stationary phase, D-amino acids modify the peptidoglycan layer by means of their incorporation into the peptidoglycan

polymer. Some significant changes occur in the structure of the inner membrane. The amount of monounsaturated fatty acids falls with simultaneous increase in the proportion of polyunsaturated fatty acids. Unsaturated fatty acids are also converted into cyclopropyl derivatives, and the ratio of the number phosphoglycerine to phosphoethanolamine increases when the cell enters the stationary growth phase. The consequence of all these changes is the formation of a rigid structure of the inner membrane and its reduced fluidity [64].

Microevolution of bacteria in the stationary phase

The bacterial population can adapt to adverse conditions in unexpected ways. During the observation of cell survival in the stationary phase, it was discovered that during quite a long incubation of the culture without change of the nutrient medium the growth curve had a characteristic successive rise and fall in the number of viable cells. This phenomenon was defined as the growth advantage in the stationary phase (GASP) phenotype. This behavior of the culture is accounted for by the emergence of mutant cells among the population, which are better adapted to these conditions than the parent strain [1].

GASP-phenotype is mediated by several key mutations which benefit in the stationary phase. One of these mutations leads to reduced activity of σ^{38} , and *rpoS* deletion mutant strains showed no characteristic phenotype. It is supposed that the advantages of this mutation are a result of its pleiotropic effect. This effect may be determined by imbalance in the competition among sigma factors for the RNA polymerase [3].

Another mutation, more exactly a group of mutations that leads to GASP-phenotype, is mutation in the *lrp* and *sgaC* genes, as well as the genomic rearrangement inactivating the *cstA* gene and activating operon *ybeJ-gltJKL*. These mutants actively feed on the amino acids debris entering the environment from the dead cells. A genomic rearrangement leads to the deactivation of a gene encoding oligopeptide permease and activation of operon encoding annotated transporter-proteins of glutamic and aspartic acids. Thus, due to lost ability to degrade oligopeptides the cell acquires enhanced ability to feed on monomeric amino acids entering the medium from dead cells. After additional incubation in cells with a mutation in the *rpoS* gene, the *bgl* operon is activated, which leads to the appearance of a population that can use aryl- β -glycosides arbutin and salicin as a resource [3].

Interestingly, in the stationary phase a population of cells appears with a viable but nonculturable (VBNC) phenotype. This phenotype manifests as a response to a variety of stresses and occurs in many bacteria. The molecular nature of the mechanism of VBNC-phenotype

notype has not been identified, but it is clear that it goes beyond a single regulatory pathway and includes a global change in cell metabolism. A feature of the VBNC-phenotype is the huge reduction in metabolism and changes in cell morphology. Probably, in this way the cell is trying to cope with stress via “hibernation” and fence itself from the harsh environment by an impermeable barrier. The mechanism of exit from this state is little known [68].

Antibiotic resistance in the stationary phase

Rapid growth in the frequency of use of antibiotics against bacterial infections results in rapid emergence of bacterial strains resistant to antibiotics. Therefore, an “arms race” with bacteria has become one of the most important issues in modern medicine. A search for newer and newer antibiotics is necessary in order to overcome the problem of resistance for a while. One of the directions in this “arms race” involves the study of the cellular mechanisms that confer resistance, instead of a search for new antibiotics, since through suppression of these mechanisms it is possible to overcome the problem of bacterial resistance to antibiotics.

It has long been observed that bacterial resistance to the action of different classes of antibiotics increases significantly under starvation. As already mentioned, the cell cycle and all stages of the genetic information implementation are suppressed in the stationary phase. Accordingly, antibiotic resistance has been mostly accounted for by the absence of growth of bacteria during starvation [69].

The molecular mechanism of bacterial resistance to antibiotics of different classes in the stationary growth phase during growth arrest has relatively recently been identified. It has been shown that upon deactivation of the *relA* and *spoT* genes (which makes it impossible to develop a stringent response), bacterial resistance to antibiotics decreases significantly and the intracellular concentration of hydroxyl radicals increases. Since it is known that the lethal action of almost all classes of antimicrobials is eventually determined by the accumulation of reactive oxygen species

in the cell [70], the authors of the research decided to assess the level of catalase activity in the cells, which proved to be significantly reduced. Thus, it was shown that active cell response to stress, rather than the absence of growth, is important in tolerance to antibiotics in the stationary phase [71].

CONCLUSIONS

The cell deals with survival in harsh settings in various ways. For protection against mechanical damage and stress factors, the cell wall is strengthened and rebuilt and the shape of cells changes. In turn, the nucleoid becomes condensed and is included in the nucleoprotein complex to protect it from damage.

To save resources, the translation process is inhibited in particular by downregulation of the expression of genes encoding components of the protein biosynthetic machinery. Particularly interesting is the variety of regulatory pathways through which translation is suppressed. It is the translational apparatus, as the most energy-consuming process, that is the key member of stress signal transmission to other components of the cell. Depending on the extent of starvation, the cell passes the pathway from reduced expression of ribosomal operons to complete suppression of translation and degradation of ribosomes.

The cell’s ability to use an alternative sigma factor for the regulation of gene expression of stress response is also important. A complex system of regulation of synthesis and stability of the sigma factor allows the cell to respond immediately to the occurrence of stress and quickly return to normal growth.

It becomes clear that the transition to the stationary growth phase is a natural defense mechanism of bacterial culture to cope with stress and starvation. Under these conditions, the cell structure changes at all levels of the organization directed at the survival of both individual cells and the whole population.

This study was supported by a grant from the Russian Science Foundation (№ 14-14-00072).

REFERENCES

1. Finkel S.E. // *Nat. Rev. Microbiol.* 2006. V. 4. № 2. P. 113–120.
2. Pin C., Baranyi J. // *Appl. Environ. Microbiol.* 2008. V. 74. № 8. P. 2534–2536.
3. Llorens J.M.N., Tormo A., Martínez-García E. // *FEMS Microbiology Rev.* 2010. V. 34. № 4. P. 476–495.
4. Luijsterburg M.S., Noom M.C., Wuite G.J.L., Dame R.T. // *J. Structural Biol.* 2006. V. 156. № 2. P. 262–272.
5. Azam T.A., Iwata A., Nishimura A., Ueda S., Ishihama A. // *J. Bacteriol.* 1999. V. 181. № 20. P. 6361–6370.
6. Nair S., Finkel S.E. // *J. Bacteriol.* 2004. V. 186. № 13. P. 4192–4198.
7. Almirón M., Link A.J., Furlong D., Kolter R. // *Genes Dev.* 1992. V. 6. № 12b. P. 2646–2654.
8. Wolf S.G., Frenkiel D., Arad T., Finkel S.E., Kolter R., Minisky A. // *Nature.* 1999. V. 400. № 6739. P. 83–85.
9. Ilari A., Ceci P., Ferrari D., Rossi G.L., Chiancone E. // *J. Biol. Chem.* 2002. V. 277. № 40. P. 37619–37623.
10. Cosgriff S., Chintakayala K., Chim Y.T.A., Chen X., Allen S., Lovering A.L., Grainger D.C. // *Mol. Microbiol.* 2010. V. 77. № 5. P. 1289–1300.
11. Kahramanoglou C., Prieto A.I., Khedkar S., Haase B., Gupta A., Benes V., Fraser G.M., Luscombe N.M., Seshasayee A.S.N. // *Nat. Commun.* 2012. V. 3. P. 886.
12. Ishihama A. // *Annu. Rev. Microbiol.* 2000. V. 54. № 1. P. 499–518.

REVIEWS

13. Reznikoff W.S., Siegele D.A., Cowing D.W., Gross C.A. // *Annu. Rev. Genet.* 1985. V. 19. № 1. P. 355–387.
14. Maciag A., Peano C., Pietrelli A., Egli T., Bellis G.D., Landini P. // *Nucl. Acids Res.* 2011. V. 39. № 13. P. 5338–5355.
15. Arnosti D.N., Chamberlin M.J. // *Proc. Natl. Acad. Sci. USA.* 1989. V. 86. № 3. P. 830–834.
16. Hunt T.P., Magasanik B. // *Proc. Natl. Acad. Sci. USA.* 1985. V. 82. № 24. P. 8453–8457.
17. Straus D.B., Walter W.A., Gross C.A. // *Nature.* 1987. V. 329. № 6137. P. 348–351.
18. Rhodius V.A., Suh W.C., Nonaka G., West J., Gross C.A. // *PLoS Biol.* 2005. V. 4. № 1. P. e2.
19. Enz S., Braun V., Crosa J.H. // *Gene.* 1995. V. 163. № 1. P. 13–18.
20. Weber H., Polen T., Heuveling J., Wendisch V.F., Hengge R. // *J. Bacteriol.* 2005. V. 187. № 5. P. 1591–1603.
21. Gaal T., Ross W., Estrem S.T., Nguyen L.H., Burgess R.R., Gourse R.L. // *Mol. Microbiol.* 2001. V. 42. № 4. P. 939–954.
22. Typas A., Hengge R. // *Mol. Microbiol.* 2006. V. 59. № 3. P. 1037–1051.
23. Mika F., Hengge R. // *Genes Dev.* 2005. V. 19. № 22. P. 2770–2781.
24. Lange R., Hengge-Aronis R. // *Genes Dev.* 1994. V. 8. № 13. P. 1600–1612.
25. Mukhopadhyay S., Audia J.P., Roy R.N., Schellhorn H.E. // *Mol. Microbiol.* 2000. V. 37. № 2. P. 371–381.
26. Gentry D.R., Hernandez V.J., Nguyen L.H., Jensen D.B., Cashel M. // *J. Bacteriol.* 1993. V. 175. № 24. P. 7982–7989.
27. Cuning C., Brown L., Elliott T. // *J. Bacteriol.* 1998. V. 180. № 17. P. 4564–4570.
28. Brescia C.C., Kaw M.K., Sledjeski D.D. // *J. Mol. Biol.* 2004. V. 339. № 3. P. 505–514.
29. Lease R.A., Belfort M. // *Proc. Natl. Acad. Sci. USA.* 2000. V. 97. № 18. P. 9919–9924.
30. Muffler A., Fischer D., Hengge-Aronis R. // *Genes Dev.* 1996. V. 10. № 9. P. 1143–1151.
31. Zhang A., Altuvia S., Tiwari A., Argaman L., Hengge-Aronis R., Storz G. // *EMBO J.* 1998. V. 17. № 20. P. 6061–6068.
32. Majdalani N., Hernandez D., Gottesman S. // *Mol. Microbiol.* 2002. V. 46. № 3. P. 813–826.
33. Trotochaud A.E., Wassarman K.M. // *J. Bacteriol.* 2004. V. 186. № 15. P. 4978–4985.
34. Jishage M., Kvint K., Shingler V., Nyström T. // *Genes Dev.* 2002. V. 16. № 10. P. 1260–1270.
35. Jishage M., Ishihama A. // *J. Bacteriol.* 1999. V. 181. № 12. P. 3768–3776.
36. Gyaneshwar P., Paliy O., McAuliffe J., Jones A., Jordan M.I., Kustu S. // *Proc. Natl. Acad. Sci. USA.* 2005. V. 102. № 9. P. 3453–3458.
37. Typas A., Barenbruch C., Possling A., Hengge R. // *EMBO J.* 2007. V. 26. № 6. P. 1569–1578.
38. Schweder T., Lee K.H., Lomovskaya O., Matin A. // *J. Bacteriol.* 1996. V. 178. № 2. P. 470–476.
39. Santos J.M., Freire P., Mesquita F.S., Mika F., Hengge R., Arraiano C.M. // *Mol. Microbiol.* 2006. V. 60. № 1. P. 177–188.
40. Bougdour A., Wickner S., Gottesman S. // *Genes Dev.* 2006. V. 20. № 7. P. 884–897.
41. Zhou Y., Gottesman S. // *J. Bacteriol.* 2006. V. 188. № 19. P. 7022–7025.
42. Zgurskaya H.I., Keyhan M., Matin A. // *Mol. Microbiol.* 1997. V. 24. № 3. P. 643–651.
43. Tani T.H., Khodursky A., Blumenthal R.M., Brown P.O., Matthews R.G. // *Proc. Natl. Acad. Sci. USA.* 2002. V. 99. № 21. P. 13471–13476.
44. Zinser E.R., Kolter R. // *J. Bacteriol.* 2000. V. 182. № 15. P. 4361–4365.
45. Calvo J.M., Matthews R.G. // *Microbiol. Rev.* 1994. V. 58. № 3. P. 466–490.
46. Gottesman S. // *Trends Genet.* 2005. V. 21. № 7. P. 399–404.
47. Majdalani N., Cuning C., Sledjeski D., Elliott T., Gottesman S. // *Proc. Natl. Acad. Sci. USA.* 1998. V. 95. № 21. P. 12462–12467.
48. Johansen J., Rasmussen A.A., Overgaard M., Valentin-Hansen P. // *J. Mol. Biol.* 2006. V. 364. № 1. P. 1–8.
49. Borek E., Rockenbach J., Ryan A. // *J. Bacteriol.* 1956. V. 71. № 3. P. 318–323.
50. Cashel M., Kalbacher B. // *J. Biol. Chem.* 1970. V. 245. № 9. P. 2309–2318.
51. Boutte C.C., Crosson S. // *Trends Microbiol.* 2013. V. 21. № 4. P. 174–180.
52. English B.P., Haurlyuk V., Sanamrad A., Tankov S., Dekker N.H., Elf J. // *Proc. Natl. Acad. Sci. USA.* 2011. V. 108. № 31. P. E365–E373.
53. Murray D.K., Bremer H. // *J. Mol. Biol.* 1996. V. 259. № 1. P. 41–57.
54. Barker M.M., Gaal T., Josaitis C.A., Gourse R.L. // *J. Mol. Biol.* 2001. V. 305. № 4. P. 673–688.
55. Starosta A.L., Lassak J., Jung K., Wilson D.N. // *FEMS Microbiol. Rev.* 2014. V. 38. № 6. P. 1172–1201.
56. Zvereva M.I., Ivanov P.V., Teraoka Y., Topilina N.I., Dontsova O.A., Bogdanov A.A., Kalkum M., Nierhaus K.H., Shpanchenko O.V. // *J. Biol. Chem.* 2001. V. 276. P. 47702–47708.
57. Neubauer C., Gillet R., Kelley A.C., Ramakrishnan V. // *Science.* 2012. V. 335. № 6074. P. 1366–1369.
58. Chadani Y., Ono K., Ozawa S., Takahashi Y., Takai K., Nanamiya H., Tozawa Y., Kutsukake K., Abo T. // *Mol. Microbiol.* 2010. V. 78. № 4. P. 796–808.
59. Chadani Y., Ono K., Kutsukake K., Abo T. // *Mol. Microbiol.* 2011. V. 80. № 3. P. 772–785.
60. Polikanov Y.S., Blaha G.M., Steitz T.A. // *Science.* 2012. V. 336. № 6083. P. 915–918.
61. Zhang J., Zhang Y., Inouye M. // *J. Biol. Chem.* 2003. V. 278. № 34. P. 32300–32306.
62. Amitai S., Kolodkin-Gal I., Hananya-Meltabashi M., Sacher A., Engelberg-Kulka H. // *PLoS Genet.* 2009. V. 5. № 3. P. e1000390.
63. Moll I., Engelberg-Kulka H. // *Trends Biochem. Sci.* 2012. V. 37. № 11. P. 493–498.
64. Nyström T. // *Annu. Rev. Microbiol.* 2004. V. 58. № 1. P. 161–181.
65. Akerlund T., Nordström K., Bernander R. // *J. Bacteriol.* 1995. V. 177. № 23. P. 6791–6797.
66. Santos J.M., Lobo M., Matos A.P.A., De Pedro M.A., Arraiano C.M. // *Mol. Microbiol.* 2002. V. 45. № 6. P. 1729–1740.
67. Reeve C.A., Bockman A.T., Matin A. // *J. Bacteriol.* 1984. V. 157. № 3. P. 758–763.
68. Hayes C.S., Low D.A. // *Curr. Opin. Microbiol.* 2009. V. 12. № 6. P. 667–673.
69. Levin B.R., Rozen D.E. // *Nat. Rev. Microbiol.* 2006. V. 4. № 7. P. 556–562.
70. Kohanski M.A., Dwyer D.J., Hayete B., Lawrence C.A., Collins J.J. // *Cell.* 2007. V. 130. № 5. P. 797–810.
71. Nguyen D., Joshi-Datar A., Lepine F., Bauerle E., Olakanmi O., Beer K., McKay G., Siehnel R., Schafhauser J., Wang Y., et al. // *Science.* 2011. V. 334. P. 982–986.

Study of Functional and Allosteric Sites in Protein Superfamilies

D. Suplatov, V. Švedas*

Lomonosov Moscow State University, Belozersky Institute of Physicochemical Biology, Vorobjev hills 1-40, Moscow 119991, Russia

Lomonosov Moscow State University, Faculty of Bioengineering and Bioinformatics, Vorobjev hills 1-73, 119991, Moscow, Russia

*E-mail: vytaš@belozersky.msu.ru

Received:

Copyright © 2015 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 interaction of proteins (enzymes) with a variety of low-molecular-weight compounds, as well as protein-protein interactions, is the most important factor in the regulation of their functional properties. To date, research effort has routinely focused on studying ligand binding to the functional sites of proteins (active sites of enzymes), whereas the molecular mechanisms of allosteric regulation, as well as binding to other pockets and cavities in protein structures, remained poorly understood. Recent studies have shown that allostery may be an intrinsic property of virtually all proteins. Novel approaches are needed to systematically analyze the architecture and role of various binding sites and establish the relationship between structure, function, and regulation. Computational biology, bioinformatics, and molecular modeling can be used to search for new regulatory centers, characterize their structural peculiarities, as well as compare different pockets in homologous proteins, study the molecular mechanisms of allostery, and understand the communication between topologically independent binding sites in protein structures. The establishment of an evolutionary relationship between different binding centers within protein superfamilies and the discovery of new functional and allosteric (regulatory) sites using computational approaches can improve our understanding of the structure-function relationship in proteins and provide new opportunities for drug design and enzyme engineering.

KEYWORDS binding sites; catalytic site; allosteric site; function; regulation; structure-function relationship; bioinformatics.

ABBREVIATIONS PDB - Protein Data Bank; RNAP – DNA-dependent RNA polymerase; MD – Molecular Dynamics; NAD – nicotinamide adenine dinucleotide; SSP – subfamily-specific position.

INTRODUCTION

Understanding the relationship between protein structure and function is one of the most challenging problems of modern biochemistry. It is complicated due to the fact that similarity of structures does not imply a common function – proteins with different properties can share a common structural framework [1, 2], while the same function can be performed by proteins with different folds [3]. Specific protein-protein interactions and recognition of low-molecular-weight compounds are crucial to all living systems. To understand the molecular mechanisms of these processes and the structure-function relationship in proteins, it is necessary to study the structural organization of the specific sites responsible for the binding of various ligands (substrates, inhibitors, effectors) [4]. Analysis and functional classification of pockets and cavities on the protein surface which form binding sites with unique properties can lead to a better understanding of the molecu-

lar mechanisms of protein functions, facilitate function prediction of recently discovered enzymes, and provide new opportunities for protein/enzyme engineering and drug design.

When protein function is investigated the functional sites – active sites of enzymes, channels of membrane transport proteins, DNA- and protein-binding motifs of different regulatory proteins – attract the most attention. However, during recent years we have witnessed the increasing role of the allostery phenomenon – regulation of protein functions at the binding of low-molecular-weight effectors in regulatory sites which are topologically independent of functional sites [5]. These facts have stimulated research activities to understand the regulation of the biological macromolecules' function caused by interaction with different ligands in allosteric centers. Several experimental and computational approaches have been developed to search for new regulatory sites in protein structures.

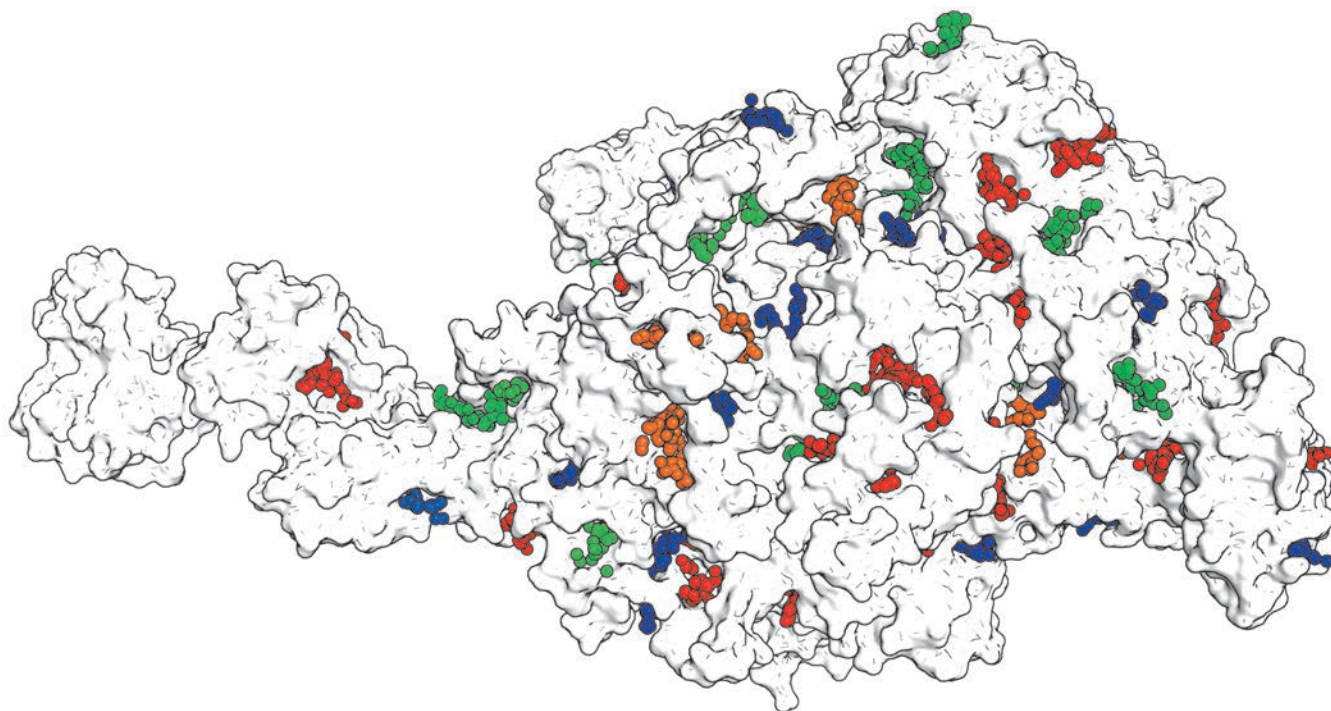


Fig. 1. Potential binding sites for low-molecular-weight effectors in the structure of bacterial RNAP. Clusters of same-colored spheres mark the potential sites on the protein surface and represent centers of α -spheres that fill in the volume of the corresponding binding pocket (see Appendix). The figure was prepared using PyMol based on the crystal structure 1YNN from PDB

Attempts have been made to understand the relationship between the functional and regulatory centers located at a considerable distance from each other and explore the molecular mechanisms of their interaction [6]. Low-molecular-weight inhibitors have been discovered that are capable of selective interaction with allosteric sites in various proteins associated with human diseases [7]. However, the particular attention to this issue is not so much because of the unique features of specific proteins, but is rather due to the general importance of these processes for functional regulation in living organisms. There are reasons to believe that allostery is a universal phenomenon common to most proteins [8], which in addition to our interest into fundamental mechanisms draws attention due to its potential applications in biotechnology and biomedicine. Recent studies have shown that proteins and enzymes, along with quite well-studied functional sites (active sites) and allosteric centers, contain a significant amount of virtually unexplored potential binding pockets. *Figure 1* shows the structure of DNA-dependent RNA polymerase – a key enzyme of RNA synthesis in all living organisms [9, 10]. The surface of this large multi-subunit protein is covered by a large number of cavities – potential

binding sites. These include the active center containing catalytic residues and the DNA binding motifs, as well as several known allosteric sites capable of binding a variety of low-molecular-weight ligands [11, 12]. The role of other binding sites, i.e. the majority of existing pockets in this case, remains unknown. How important are these sites for enzyme function? Which binding centers play a physiological role and which can be used to create a protein with new properties for practical applications? How to evaluate the potential role of each specific site for the regulation of protein function?

In this review, we discuss the study of the structure-function relationship in proteins based on the analysis of various binding sites in their structures. In this context, experimental and computational approaches are considered which allow us to search for new binding centers capable of interacting with regulatory ligands and study the molecular mechanisms of allostery and the relationship between function and regulation in proteins. The rapid expansion of public databases makes genomic, structural, and functional information widely available for a large number of proteins. In this respect, bioinformatic methods provide an opportunity to study protein functions within the cor-

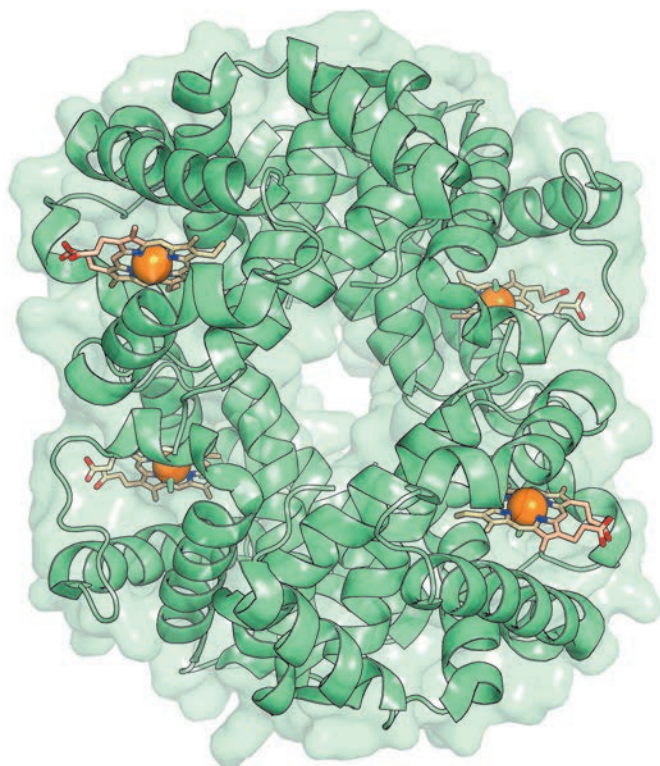


Fig. 2. Three-dimensional structure of human hemoglobin. Heme molecules (orange) are shown as sticks in each subunit of the tetramer. The figure was prepared using PyMol based on the crystal structure 1GZX from PDB

responding superfamilies systematically, rather than individually. Analysis of the structural information and experimental data concerning individual proteins, as well as in their relationship with close and distant evolutionary relatives, should contribute to a better understanding of the structure–function relationship in proteins/enzymes and unveil new mechanisms of regulation of their functional properties.

THE PHENOMENON OF ALLOSTERY

Allostery is generally defined as the process of regulation of protein function due to the binding of an effector – a ligand or another protein – in a site on the protein surface referred to as an allosteric center [6]. The term “allosteric” comes from the Greek roots *allos* (other) and *stereos* (solid), and can be translated as “different shape” in order to emphasize the relationship of conformational states between structurally remote sites in proteins [8]. It is known that allosteric regulation of metabolism is important for all living cells, and allosteric effectors can be either inhibitors or activators with respect to protein function [13].

Historically, allostery has typically referred to a cooperative effect in multi-subunit proteins that function at the quaternary structural level. The first “concerted” molecular model of allostery, known as the MWC (Monod-Wyman-Changeux) model, was proposed in 1965 based on the 24 then known examples [14]. The abrupt increase in hemoglobin oxygen affinity described by an S-shaped curve suggested a cooperative effect. However, a crystal structure of hemoglobin obtained in 1960 with 5.5Å resolution showed that heme molecules that bind oxygen are located in different subunits of the protein at a considerable distance from each other (*Fig. 2*) [15]. This led to the conclusion that allosteric proteins have a symmetrical arrangement of subunits which can adopt at least two conformational states – R (relaxed) or T (tense), characterized, in the case of hemoglobin, by high and low affinity for oxygen, respectively. The transition from one conformational state to another as a result of ligand binding proceeds in a coordinated manner among all subunits so that the oligomeric protein does not exist in a hybrid state RT. This simplified model was used to kinetically characterize hemoglobin saturation by oxygen [14]. However, the molecular mechanism of this phenomenon remained unclear until a series of structural studies [16, 17]. It was shown that the binding (release) of oxygen is accompanied by significant changes in the spatial organization of the functional center and disruption (formation) of a few salt bridges. This leads to displacement of subunits relative to one another so that binding of the first oxygen molecule affects the whole tetramer. In other words, binding (release) of one molecule of the substrate to one subunit changes oxygen affinity in other subunits, making hemoglobin an effective oxygen carrier along the pressure gradient. Such cooperative effects in homo-oligomeric proteins/enzymes are one of the most known examples of allostery. In this case, the active center of one subunit can function as an allosteric center with respect to the active center of another subunit. Therefore, the binding of the second substrate molecule (or the corresponding analogue) may not be accompanied by its catalytic conversion but leads to the allosteric effect on the binding site of the first substrate molecule. According to the “sequential” or KNF model (Koshland-Nemethy-Filmer), subunits within the multimer change their conformation one at a time; i.e. binding of a ligand changes the conformation and properties of the corresponding subunit and affects its neighbors [18]. In other words, ligand binding causes consecutive conformational changes in protein subunits; e.g., this model describes negative cooperativity in the enzyme glyceraldehyde-3-phosphate dehydrogenase. Binding of a coenzyme NAD⁺

to the active site of one subunit weakens its binding to the adjacent subunit due to the rearrangement of intra- and inter-subunit contacts [19, 20]. This property maintains enzyme activity at a constant level regardless of the concentration of a ligand in the environment. Although an attempt was made to combine the MWC and KNF models into a more general one [21], further studies showed that molecular mechanisms of allosteric regulation are so complex and diverse that none of the proposed simplified models can exhaustively describe the phenomenon of allostery.

Nowadays, it is nearly generally accepted that not just multi-subunit proteins, but also monomeric ones are subjected to allosteric regulation, and allosteric ligands are mainly considered as low-molecular-weight compounds that bind to regulatory sites topologically independent of functional centers. In addition, the regulatory effects caused by protein-protein interactions, phosphorylation, and even point mutations are sometimes also considered as allosteric ones. The diversity of allosteric mechanisms in various proteins and enzymes is well illustrated in recent publications [5, 7, 22]. It has been suggested that allostery may be an intrinsic property of virtually all proteins [8]. The exceptions, probably, draw up structural proteins with rigid conformations that limit their flexibility and opportunities for regulation. Indeed, there is growing experimental evidence of allostery in enzymes that were previously considered as non-allosteric.

Phosphofructokinase catalyzes one of the key steps in glycolysis and offers an example of a protein whose function can be regulated by various effectors. Allostery has been described in this superfamily for enzymes from both prokaryotes [23] and eukaryotes, the latter being characterized by much larger globules due to duplications, insertions, and mutations of the ancestral prokaryotic gene, which contributed to the emergence of new allosteric centers [24]. At the same time, phosphofructokinase from fungus *Dictyostelium discoideum* is different from its homologs and considered as non-allosteric. However, it was shown that deletion of one C-terminal leucine residue leads to the emergence of allosteric properties in this enzyme similar to other superfamily members [25]. A different example is allosteric regulation in pyruvate kinases [26]. Four isoforms of this enzyme have been characterized in mammalian tissues – L, R, M₁, and M₂. All isoforms, except for M₁, are allosteric enzymes and show positive homotropic cooperativity with respect to the substrate, as well as positive heterotropic cooperativity with respect to fructose-1,6-bisphosphate. Isoforms M₁ and M₂ were shown to be produced from a common gene by alternative splicing. The corresponding primary sequences are different in 23 amino acid residues which are lo-

cated at the intersubunit interface and are involved in the formation of the binding site of fructose-1,6-bisphosphate. It has been shown that two point mutations introduced into the structure of M₁ isoform – at the intersubunit interface [27] and in the binding site of fructose-1,6-bisphosphate [26] – lead to the emergence of allosteric properties similar to those of other homologs. In a different study, myoglobin, a paralog¹ of hemoglobin, has been shown to exist in three major conformational states with different catalytic properties – the so-called taxonomic substates – and each of these assumes a very large number of slightly different conformations or statistical substates [28]. Based on this observation, it has been further assumed that bimolecular reactions with diatomic molecules (e.g., NO and O₂) can be allosterically controlled in myoglobin due to changes in the geometry of conservative cavities adjacent to the active center. It is interesting to note that in all these cases allosteric regulation has been discovered in proteins which are evolutionary related to other allosteric enzymes. These examples speak not only to the wide occurrence of allostery, but also underline the general mechanisms of this phenomenon within protein superfamilies. They indicate the possibility of fine-tuning allostery by only several point mutations, but also emphasize the complex relationship between function and regulation.

The current concept of protein structure assumes that proteins exist as complex statistical ensembles of conformers that fold and unfold continuously by making local rearrangements [6, 29]. In this context, the allosteric effect is a result of the redistribution of conformational states [8]. In other words, binding of an allosteric effector leads to a population shift toward conformational states that are significantly different in functional terms from the native state [30]. On the other hand, if a protein is considered to be non-allosteric, this can simply be an indication that alternative conformations of binding sites and functionally important conformational transitions have not been discovered yet. It does not mean, however, that one cannot choose a ligand or specific environmental conditions that would be able to cause a conformational redistribution and trigger allosteric behavior in otherwise not-allosteric proteins. In principle, almost any substance bound to the protein surface can cause population shift of conformational states, the question being only in the effectiveness of the shift and its impact on protein function [8]. Further studies with various proteins should be performed to understand this problem. However, the conformational changes associated with

¹ *Paralogs – evolutionarily related proteins that occurred as a result of duplication of the ancestral gene.*

Online services to predict binding sites in protein structures and rank them by functional significance

Name	On-line address	Algorithm used to identify the binding sites	Algorithm used to rank the binding sites
Fpocket [35]	Programs>Structure>Pockets>fpocket">http://mobylye.rpbs.univ-paris-diderot.fr/>Programs>Structure>Pockets>fpocket	Geometric, based on Voronoi tessellation and detection of α -spheres	Statistical, by estimating similarity to known ligand binding sites
POCASA [33]	http://altair.sci.hokudai.ac.jp/g6/service/pocasa/	Geometric, by rolling spherical probes along the protein surface	Geometric, taking into account the position and size of the pocket
pocketZebra [39]	http://biokinet.belozersky.msu.ru/pocketzebra	Geometric, based on Voronoi tessellation and detection of α -spheres	Bioinformatic, analysis of subfamily-specific positions in protein superfamilies
SiteHound [38]	http://scbx.mssm.edu/sitehound/sitehound-web/Input.html	Energy-based, by estimating the interaction energy of amino acids at the protein surface with carbon or phosphate chemical probes	Energy-based, by estimating the interaction energy of amino acids at the protein surface with carbon or phosphate chemical probes
LIGSITE ^{csc} [41]	http://projects.biotec.tu-dresden.de/pocket/	Geometric, based on the calculation of the Connolly surface	Bioinformatic, analysis of conserved positions

allosteric regulation are difficult to detect using current experimental techniques. The recently developed bioinformatic and computational biology approaches provide new opportunities for solving this problem.

IDENTIFICATION OF BINDING SITES IN PROTEIN STRUCTURES

The prediction of binding sites in proteins based on information about their structures is a new challenging field in computational biology [31]. Various geometry-based structural approaches to the search for pockets and cavities on the protein surface have been developed (*table*). Often, several pockets are found in a protein structure and an attempt is further made to select the most relevant sites that are likely to bind a ligand – by implementing various geometric criteria (size, depth, and orientation of a potential binding cavity [32-34]) or a statistical analysis that takes into account the physicochemical properties of the known ligand binding sites [35, 36]. Alternatively, energy-based approaches have been proposed that predict and rank binding pockets by calculating the binding energy of small organic molecules (probes) on the protein surface [37, 38]. All these approaches basing on the analysis of the available protein structure can quickly and efficiently detect cavities and pockets that form potential binding sites, but they give no idea about their functional significance and structure of the complementary ligands.

In the course of evolution of proteins from a common ancestor, some functional properties were preserved, while others underwent changes as a result of natural selection, which led to functional diversity. For example, homologous enzymes within a superfamily can share a common fold and reaction chemistry but differ

in other functional properties (e.g., substrate specificity, enantio- and regio-selectivity, and type of catalyzed chemical transformation), as well as principles of their regulation. The continuous growth in public databases providing access to genomic and structural information on various proteins and enzymes opens new perspectives for a large-scale comparative analysis of both evolutionarily close and distant relatives within protein superfamilies. Not all positions in protein structures are equally susceptible to variation in the course of evolution, reflecting differing selection pressure on amino acids residues with different functional roles. That makes it possible to apply a bioinformatic analysis of protein superfamilies to the study of the evolutionary relationship of amino acid residues in functional and regulatory binding sites [39] (*table*). Totally conserved positions play a key role in a function common to all proteins within a superfamily; e.g., they are involved in the enzyme’s catalytic mechanism. It should be noted, however, that catalytically important amino acids are not always conserved throughout enzyme superfamilies and can even migrate within a common structural framework of homologous proteins [2, 40]. Catalytic nucleophile in α/β -hydrolases can be represented by serine, cysteine or aspartate, and the catalytic acid can be located in at least two alternative positions of the main polypeptide chain. Nevertheless, it was shown that conservation of residues in pockets and cavities in the protein structure is an efficient criterion for annotation of functional centers [41-43] and can be used to characterize a wide range of enzymes [44-46]. In fact, when characterizing a new protein with an unknown function it is reasonable to begin with a comparative analysis of its closest homologs in order to identify the conserved positions in columns of the corresponding

multiple alignment. The role of the most conserved residues can be further studied experimentally by introducing point mutations and evaluating their impact on the protein function or enzyme catalytic properties. Annotation of functional sites in a new protein by residue conservation can be performed even in the absence of structural data given that the appropriate information is available for evolutionary relatives. This type of data for various proteins from different superfamilies is in constant growth and is being accumulated in public databases (see next chapter). To sum up, integration of geometry-based structural methods with bioinformatics approaches can provide more efficient annotation of functional centers in proteins.

A comparative study of various proteins allowed researchers to conclude that allosteric sites in enzyme superfamilies are characterized by a lower content of conserved positions and a higher content of variable positions [47]. It has been further shown that mutagenesis of variable positions in allosteric centers leads to a change in the allosteric effect, while substitution of conserved positions in these centers leads, as a rule, to a loss of catalytic function. These results demonstrate that residue conservation may not be a suitable criterion to annotate regulatory centers but indicate the important role of variable positions in the binding of ligands and allosteric regulation of functional properties in proteins superfamilies. In this regard, the subfamily-specific positions (SSPs) – conserved within functional subfamilies, but different between them – attract special attention [48, 49]. SSPs are observed in both catalytic and allosteric sites, and their presence can be a very powerful factor for the identification of functional and regulatory centers in protein structures [39]. Identification of statistically significant subfamily-specific positions can help understand the difference in the organization of binding sites within evolutionarily related proteins. DNA-dependent RNA polymerase (RNAP) is a key enzyme in DNA transcription crucial to all living systems. The catalytic core of the bacterial enzyme consists of subunits $\alpha_2\beta\beta'\omega$, which are characterized by a high degree of structural and functional similarity among homologs from different organisms. Bacterial RNAP is also a confirmed target for antimicrobial drugs [12]. The first-line anti-tuberculosis drug rifampicin selectively inhibits transcription in *Mycobacterium tuberculosis* due to interaction with the allosteric center located in the β -subunit of this enzyme. Interaction of the inhibitor with the bacterial enzyme directly blocks the elongation path in the pathogen, without affecting the homologous human enzyme [50]. Bioinformatic analysis of the RNAP superfamily shows that the selective rifampicin binding by the bacterial enzyme is caused by the presence of different amino

acid residues in prokaryotic and human proteins at the subfamily-specific positions of the corresponding binding site (Fig. 3). This example shows that the role of SSPs in binding sites which interact with regulatory ligands should be further evaluated to better understand the molecular mechanisms of specific recognition of allosteric effectors and reveal patterns of functional regulation in protein superfamilies.

PUBLIC DATABASES

The speed and availability of versatile information from open sources through the Internet is an important driving force in modern science. In this context, a special role belongs to the numerous public databases.

The Catalytic Site Atlas (CSA) database is one of the main resources related to enzymes [51]. The core of CSA is the experimental data on 1,000 catalytically active proteins with different properties, and bioinformatic methods are further implemented to search for similar amino acid sequences and accurately annotate catalytic residues basing on their conservation patterns. As a result, the database provides information about tens of thousands of proteins.

Information relating to allosteric proteins is not as abundant due to the lack of corresponding structural data and difficulties in determining the allosteric sites; however, during the last decade progress has been achieved in this area. Compared to the 24 allosteric proteins discovered more than 50 years ago (when the first model of cooperativity was proposed), today hundreds of documented cases are reported. The recently founded Allosteric Database (ASD) has been the first attempt to generalize relevant data from the literature [52]. Today ASD contains nearly 2,000 sites. However, not all entries are supported by structural data about a protein complex with an effector, and some annotations seem controversial. Nevertheless, it shall be expected that the systematization of experimental data related to the structure and function of allosteric centers will be continued, also in other public databases.

THE RELATIONSHIP BETWEEN FUNCTION AND REGULATION

There is no doubt that a conformational change in protein structure caused by binding of a ligand to an allosteric center eventually leads to a change in its functional properties. However, little is known about the particular molecular mechanisms of this phenomenon. How to explain the observed cooperativity at binding of various ligands and how to predict the relationship between independent sites in proteins which are not ascribed yet as allosteric ones? Several attempts have been undertaken during recent years to understand the relationship between function and regulation [6].

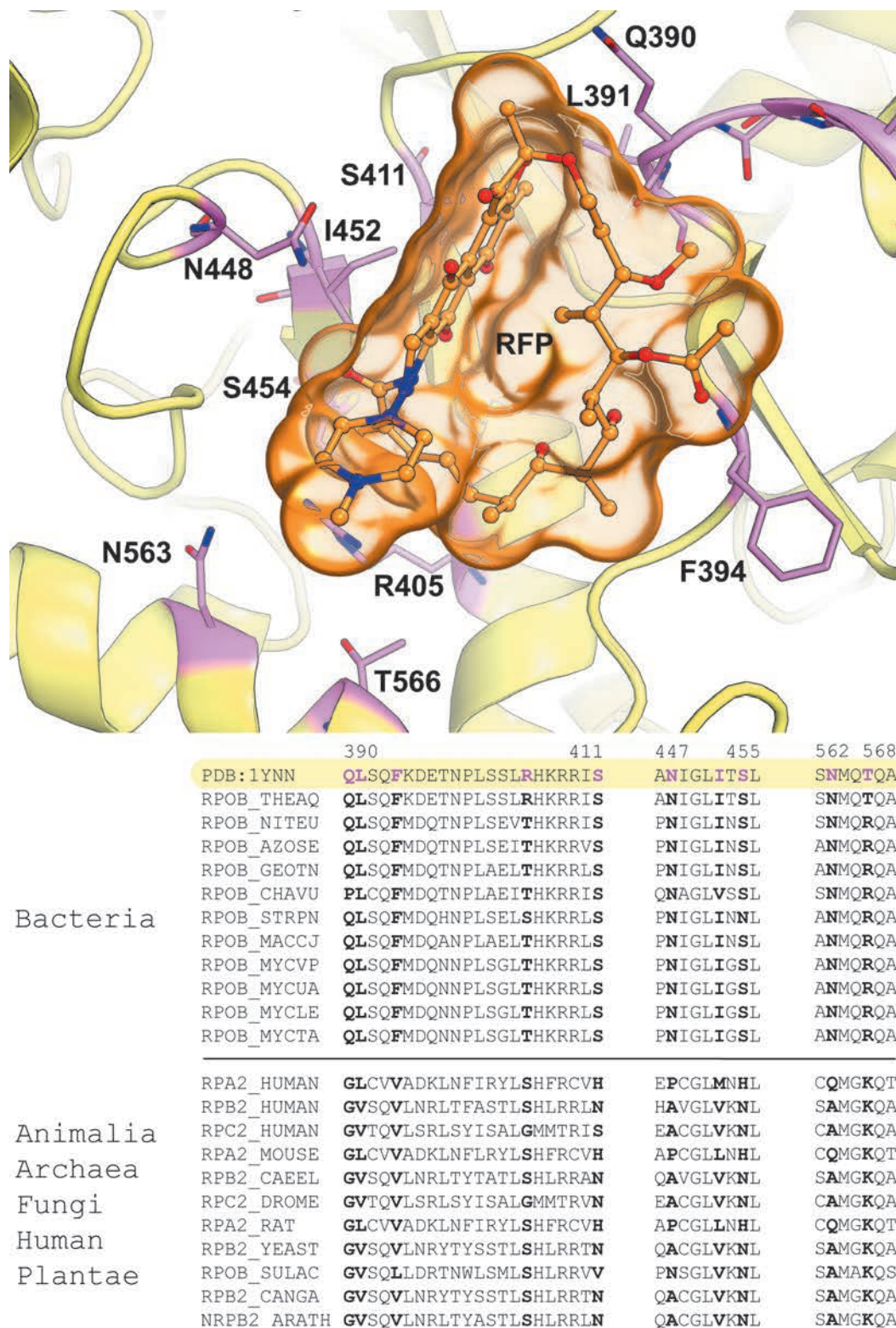


Fig. 3. Binding of rifampicin (RFP) to the allosteric site in β -subunit of bacterial RNAP. Sticks represent subfamily-specific positions (pink) identified by the bioinformatic analysis of 271 RNAPs from different sources, and corresponding fragments of the multiple alignment are shown. An interactive version of this illustration is available online (see Appendix). The figure was prepared using PyMol based on the crystal structure 1YNN from PDB

These studies were aimed at a computational search for correlations – structural or evolutionary – between events occurring in topologically independent centers on the protein surface upon binding of ligands. Let us consider some such examples.

Structural changes that occur as a result of ligand binding are directly related to the conformational mobility of the protein globule. Molecular dynamics has proven efficient in studying structural changes in proteins [53, 54], including correlated fluctuations of atoms occurring as a result of collective movement [55]. Covariance maps of atomic fluctuations along MD trajectories have been calculated and used to study the molecular mechanisms of allosteric regulation in the von Hippel-Lindau tumor suppressor protein (pVHL) [56]. Free pVHL is only marginally stable and exists in the so-called “molten globule” state. pVHL stability is greatly improved after binding to elongin C and elongin B, and these proteins together form a substrate-recognition component for the hypoxia-inducible factor (HIF) within the E3 ubiquitin protein ligase complex (Fig. 4). MD study has shown that the interface between the pVHL α and β domains is the most unstable region of the protein. Amino acid residues in pVHL have been selected whose motions were strongly correlated with this unstable motif. Molecular modeling has shown that introduction of amino acids from a more stable *Caenorhabditis elegans* pVHL into human pVHL at the selected positions results in significant stabilization of the protein in both the free state and within the complex. In other words, mutation of pVHL residues which are located away from elongin C and HIF binding sites has led to stabilization of the pVHL-elongin C complex and lowered the binding free energy of pVHL with HIF. The authors of [56] conclude that the stability and efficiency of binding to pVHL could be regulated allosterically by drugs mimicking the effect of the introduced mutations.

A different example demonstrates the value of evolutionary correlations obtained from a statistical analysis of genomic sequences [57, 58]. The approach is based on the assumption that if two sites in a protein structure are functionally related, then the corresponding amino acid residues in homologous proteins should have been coevolving during evolution from a common ancestor, and therefore this correlation can be detected by a statistical comparison of amino acid sequences. In such a case, the correlation of amino acid occurrence at two sites in structures of related proteins can indicate the existence of a functional dependence between them. This approach has been used to analyze membrane protein FecA – a member of the TonB-dependent transporters family whose function is to pump iron through the outer membrane into the

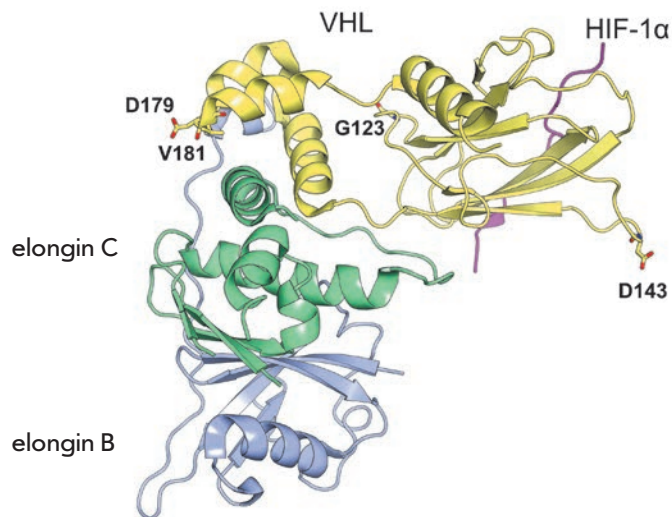


Fig. 4. The complex of the transcription factor HIF (magenta) – von Hippel-Lindau tumor suppressor protein pVHL (yellow) – elongin C (green) – elongin B (blue). The α domain of pVHL interacts with elongin C, while the β domain binds HIF. Amino acid residues in pVHL whose motions were strongly correlated with the unstable inter-domain region are shown as sticks. The figure was prepared using PyMol based on the crystal structure 1LM8 from PDB to illustrate the results of [56]

cells of gram-negative bacteria [59]. The interaction of a periplasmic domain of the TonB protein, which is involved in maintaining the proton gradient across the cytoplasmic membrane, with a conserved N-terminal TonB-binding motif of the transporter is an essential step of iron transport. It was suggested that TonB binding at the periplasmic surface is somehow dependent on a siderophore¹ binding at the extracellular surface and causes conformational changes in the transporter protein that drive iron import. However, the specific mechanisms of the allosteric communication between the two binding sites located in different cell compartments at a considerable distance from each other are unknown. Statistical sequence analysis of TonB-dependent transport proteins revealed a sparse, but structurally connected network of evolutionarily correlated residues which provide functional communication between the periplasmic and extracellular binding sites in FecA (Fig. 5). Mutation of the selected residues, which were not directly involved in binding of either TonB or siderophore, has led to the disruption of FecA transport function.

¹Siderophores – low-molecular-weight compounds with high affinity to ferric ions (e.g., ferric citrate).

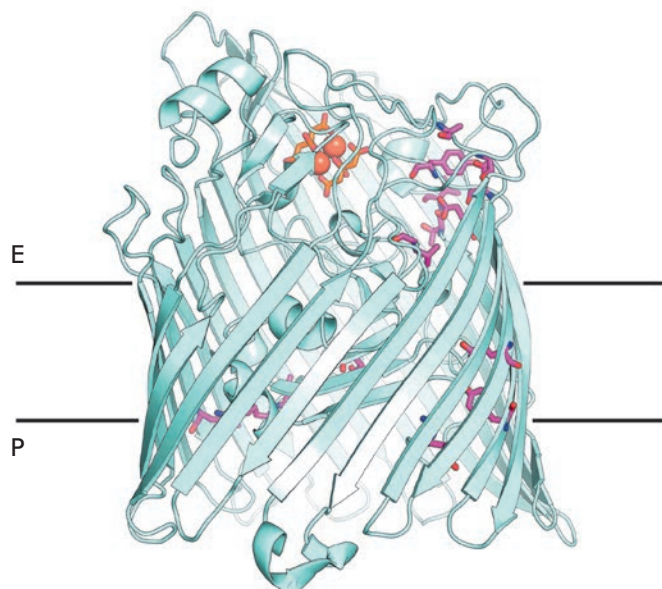


Fig. 5. The structure of the outer membrane transport protein FecA. The relative position of the periplasmic surface (P) and extracellular surface (E) is shown. Ferric ions solubilized by ferric citrate (orange) are shown in the siderophore binding site. Magenta sticks correspond to the correlated residues involved in the formation of a network of interactions that provide functional communication between the periplasmic and extracellular binding sites in FecA. The figure was prepared using PyMol based on the crystal structure 1KMP from PDB to illustrate the results of [59]

To sum up, the available results suggest that identification of evolutionary and structural correlations presents an important tool to study the molecular mechanisms of allosteric regulation in proteins.

BINDING SITES IN BIOTECHNOLOGY AND BIOMEDICINE

Binding sites of substrates/ligands in enzymes/proteins have been extensively studied to create new biocatalysts for industrial use (e.g., see review [60]), as well as inhibitors to treat human diseases [61, 62]. Although the choice of particular methods is specific to each case, the core principle of the most successful research projects can be described as “stochastic analysis.” The stochastic techniques, which are usually referred to as Directed evolution approaches, have been developed to produce enzymes with improved functional properties [63, 64]. These methods mimic the Darwinian process by combining random mutagenesis with screening and selection of the desired phenotype. Mutations are randomly introduced into the whole protein structure or in particular regions, and then their effect is evaluated experimentally to select the most

promising substitutions which lead to an improvement in the desired properties. The stochastic approaches got much faster during the last decades due to implementation of the statistical analysis and computational tools [65, 66]. However, they remain resource-demanding, still require large mutant libraries, efficient screening techniques, and yet are able to scan only a small part of the sequence space. To sum up, random evolutionary methods are hampered by a high frequency of deleterious mutations and a low frequency of beneficial mutations. Similarly, the design of new drugs is usually based on a blind experimental screening of huge libraries of low-molecular-weight compounds in an attempt to find potential inhibitors of a target protein [67, 68]. Although the structure of the lead discovered by chance can be further optimized using experimental and computational methods, this approach in general is very resource-consuming and inefficient. Embracing this approach in 1995–2001, GlaxoSmithKline performed 70 high-performance experimental screening campaigns (US\$1 million each) of selected target proteins from different pathogenic bacteria using original collections of potential inhibitors (which consisted of 260,000–530,000 compounds). Only 5 leads were found after seven years of research [69]. A review of the literature shows that similar studies were conducted between 1996 and 2004 by at least 34 different companies on 60 targets and are generally considered to have been unsuccessful [70]. The high costs and poor performance have eventually dampened interest in this empirical methodology.

Despite the apparent multipurposeness of stochastic approaches, they are usually aimed at studying the functional centers in proteins. To alter the catalytic properties of an enzyme random mutations are introduced into the structure of the active site [71]. Similarly, the majority of the developed drugs bind to functional centers of proteins (see [72] as an example). Practical application of allosteric sites is arguably rare, although some examples are known. It has been shown that introduction of a single mutation in the structure of glyceraldehyde-3-phosphate dehydrogenase leads to the destruction of a salt bridge near the active site and consequent loss of cooperativity in the binding of NAD^+ [73]. The corresponding enzyme variant has been characterized by a two-fold increase in specific activity. Certain drugs are known that interact with the regulatory sites of proteins. Rifampicin and Myxopyronin bind to pockets within the β and β' subunits of RNAP, which are topologically independent from the active site, and block enzyme operation [12]. Binding of doramapimod at the allosteric center of human p38 MAP kinase and consequent conformational rearrangements impose steric hindrance on the

ATP binding [74]. Inhibitors of HIV-1 reverse transcriptase – efavirenz, nevirapine and delavirdine – bind to an allosteric site at a significant distance from the active site [75]. Summing up the existing experience, it should be noted that due to the higher sequence variability of allosteric sites within superfamilies these regulatory centers should be considered as no less attractive targets for selective inhibition than the catalytic sites [76].

The low efficiency of stochastic methods has stimulated the development of computational approaches to rationally design effective biocatalysts and find selective inhibitors of key metabolic enzymes. It has been shown that a bioinformatic analysis of the evolutionary relationships in functionally diverse protein superfamilies can be used not just to detect the key “hotspots” in enzyme structures, but also determine the specific amino acid substitutions to produce mutants with improved properties [54, 77-79]. Use of computational approaches in protein design has been recently reviewed [80-82]. The whole-genome sequencing of bacterial pathogens, including *Mycobacterium tuberculosis* [83], marked the beginning of computer genomics in medicine. Genomic approaches can be used to make a list of all target proteins in a particular organism and to identify the most promising ones for further experimental evaluation [84]. An important advantage of the post-genomic analysis is the ability to select taxonomically widely distributed molecular targets, as well as the ones specific to a particular organism. It was assumed that in this way one could create drugs with a broad therapeutic activity, as well as with high specificity to a particular pathogen. In addition, a comparative genomic analysis of bacteria and animals can be used to exclude proteins which have human homologs from the list of potential molecular targets. In such a way, it could be possible to avoid the toxicity of the drug [85]. It should be noted, however, that currently used postgenomic methods in drug discovery do not get into too much details when choosing molecular targets for new antibiotics; e.g., the first choice at selection of targets are proteins conserved in bacteria and absent from the human organism. The structural peculiarities of these proteins and the architecture of their binding sites are frequently left out of consideration. In general, exclusion from the list of potential molecular targets of those proteins of bacterial pathogens that have human homologs (in order to avoid the toxicity of the designed inhibitors) should be considered as quite unreasonable. The major metabolic pathways are mostly conserved, and the corresponding key enzymes are present in both pathogenic bacteria and man; e.g., we have already mentioned the first-line anti-tuberculosis drug rifampicin, which inhibits replication due to selective

binding to the β -subunit of RNAP in bacteria, although the enzyme has a human homolog [50].

In summary, we can note the general trend away from inefficient stochastic approaches towards more rational and focused strategies. In this context, the role of bioinformatics and molecular modeling in biotechnology and biomedicine continues to grow steadily. Development of new approaches for a systematic analysis of various binding sites in large protein superfamilies will help, on the one hand, to establish a relationship between structure, function, and regulation of proteins/enzymes, and, on the other hand, to detect binding sites for new substrates and inhibitors/effectors with a previously unknown mechanism of action.

CONCLUSIONS

Protein-protein interactions, as well as interaction of proteins (enzymes) with a variety of low-molecular-weight compounds, are a crucial factor in the regulation of their functional properties. To date, research efforts have typically focused on studying ligand binding to the functional sites of proteins (active sites of enzymes), whereas the molecular mechanisms of allosteric regulation, as well as binding to other pockets and cavities in protein structures, remained poorly understood. In this context, it is of great interest not only to study the interaction between functional and allosteric centers, but also to identify and characterize new binding sites and their role in the regulation of protein function. Despite the first steps being made towards a better understanding of the relationship between structure, function, and regulation, the issue remains far from resolved and requires continued attention. Analysis of the available literature allows one to conclude that the role of bioinformatic methods and molecular modeling in investigating the role of different binding centers in protein function, as well as allosteric effects, will continue to grow. Establishment of an evolutionary relationship between different binding sites within protein superfamilies and discovery of new functional and allosteric (regulatory) sites using computational approaches will improve our understanding of the structure-function relationship in proteins and provide opportunities for creating new drugs and designing more effective biocatalysts.

This work was supported by the Russian Science Foundation (grant #15-14-00069)

APPENDIX

Multiple alignment of the RNAP superfamily of proteins

The structure of bacterial RNAP (PDB code: 1YNN) was used as a seed. To reconstruct evolutionarily dis-

tant RNAPs from various organisms (bacteria, animal, human, etc.) structural similarity search versus PDB database was carried out using the superpose algorithm from the CCP4 package [86]. The lower bound for matching of secondary structure elements was set to 30% in each protein. For each detected remote homolog, a list of evolutionarily close proteins was reconstructed using the BLAST algorithm versus the Swiss-Prot database [87]. The resulting samples were filtered to remove redundant sequences at a 95% identity upper bound, as well as outliers with similarity of less than 0.5 bit score per column [88] compared to the respective homolog with a known structure. Structural alignments were carried out by the Matt algorithm [89], and sequence alignments were performed using T-coffee [90]. The resulting structural alignment of distant homologs was used as a core to align samples of closely related sequences. The resulting structure-guided multiple alignment of the RNAP superfamily contains 271 protein sequences.

Binding site prediction

Identification of pockets and cavities in the RNAP structure (code in PDB: 1YNN), which are potentially capable of binding small molecules, was performed by the Fpocket algorithm [35].

Bioinformatic analysis

Identification of the subfamily-specific positions and subfamily-specific binding sites in the RNAP superfamily was performed using the Zebra [91] and pocket-Zebra [39] algorithms.

Structural analysis

Visualization and analysis of protein structural information was performed using PyMol (Schrödinger LLC).

Dissemination

Online access to the results of the bioinformatics analysis of the superfamily of DNA-dependent RNA polymerases is provided at <http://biokinet.belozersky.msu.ru/pocketzebra> (see “Examples”).

REFERENCES

- Thornton J.M., Todd A.E., Milburn D., Borkakoti N., Orengo C.A. // *Nat. Struct. Mol. Biol.* 2000. V. 7. P. 991–994.
- Todd A.E., Orengo C.A., Thornton J.M. // *Curr Opin Chem Biol.* 1999. V. 3. № 5. P. 548–556.
- Martin A.C., Orengo C.A., Hutchinson E.G., Jones S., Karmirantzou M., Laskowski R.A., Mitchell J., Taroni C., Thornton J.M. // *Structure.* 1998. V. 6. № 7. P. 875–884.
- Jones S., Thornton J.M. // *Curr Opin Chem Biol.* 2004. V. 8. № 1. P. 3–7.
- Laskowski R.A., Gerick F., Thornton J.M. // *FEBS Lett.* 2009. V. 583. № 11. P. 1692–1698.
- Goodey N.M., Benkovic S.J. // *Nat. Chem. Biol.* 2008. V. 4. № 8. P. 474–482.
- Hardy J.A., Wells J.A. // *Curr. Opin. Struct. Biol.* 2004. V. 14. № 6. P. 706–715.
- Gunasekaran K., Ma B., Nussinov R. // *Proteins.* 2004. V. 57. № 3. P. 433–443.
- Campbell E.A., Pavlova O., Zenkin N., Leon F., Irschik H., Jansen R., Severinov K., Darst S.A. // *EMBO J.* 2005. V. 24. № 4. P. 674–682.
- Esyunina D., Klimuk E., Severinov K., Kulbachinskiy A. // *Proc. Natl. Acad. Sci. USA.* 2015. V. 112. № 7. P. 2017–2022.
- Sousa R. // *Cell.* 2008. V. 135. № 2. P. 205–207.
- Darst S.A. // *Trends Biochem. Sci.* 2004. V. 29. № 4. P. 159–162.
- Monod J., Changeux J.P., Jacob F. // *J. Mol. Biol.* 1963. V. 6. № 4. P. 306–329.
- Monod J., Wyman J., Changeux J.P. // *J. Mol. Biol.* 1965. V. 12. № 1. P. 88–118.
- Perutz M.F., Rossmann M.G., Cullis A.F., Muirhead H., Will G. // *Nature.* 1960. V. 185. P. 416–422.
- Perutz M.F., Wilkinson A.J., Paoli M., Dodson G.G. // *Annu Rev Biophys Biomol Struct.* 1998. V. 27. № 1. P. 1–34.
- Eaton W.A., Henry E.R., Hofrichter J., Mozzarelli A. // *Nat. Struct. Mol. Biol.* 1999. V. 6. № 4. P. 351–358.
- Koshland D.E., Nemethy G., Filmer D. // *Biochemistry.* 1966. V. 5. № 1. P. 365–385.
- Conway A., Koshland D.E. // *Biochemistry.* 1968. V. 7. № 11. P. 4011–4023.
- Makshakova O.N., Semenyuk P.I., Kuravsky M.L., Ermakova E.A., Zuev Y.F., Muronetz V.I. // *J. Struct. Biol.* 2015. V. 190. № 2. P. 224–235.
- Eigen M. // *Nobel Symp.* 1967. V. 5. P. 333–369.
- Arkin M.R., Wells J.A. // *Nat Rev Drug Discov.* 2004. V. 3. № 4. P. 301–317.
- Schirmer T., Evans P.R. // *Nature.* 1990. V. 343. № 6254. P. 140–145.
- Poorman R.A., Randolph A., Kemp R.G., Henrikson R.L. // *Nature.* 1984. V. 309. № 5967. P. 467–469.
- Santamaría B., Estévez A.M., Martínez-Costa O.H., Aragón J.J. // *J. Biol. Chem.* 2002. V. 277. № 2. P. 1210–1216.
- Ikeda Y., Taniguchi N., Noguchi T. // *J. Biol. Chem.* 2000. V. 275. № 13. P. 9150–9156.
- Ikeda Y., Tanaka T., Noguchi T. // *J. Biol. Chem.* 1997. V. 272. № 33. P. 20495–20501.
- Frauenfelder H., McMahon B.H., Austin R.H., Chu K., Groves J.T. // *Proc. Natl. Acad. Sci. U.S.A.* 2001. V. 98. № 5. P. 2370–2374.
- Hilser V.J., Thompson E.B. // *Proc. Natl. Acad. Sci. U.S.A.* 2007. V. 104. № 20. P. 8311–8315.
- Chennubhotla C., Yang Z., Bahar I. // *Mol Biosyst.* 2008. V. 4. № 4. P. 287–292.
- Henrich S., Salo-Ahen O.M., Huang B., Rippmann F.F., Cruciani G., Wade R.C. // *J. Mol. Recognit.* 2010. V. 23. № 2. P. 209–219.
- Weisel M., Proschak E., Schneider G. // *Chem Cent J.* 2007. V. 1. № 7. P. 1–17.
- Yu J., Zhou Y., Tanaka I., Yao M. // *Bioinformatics.* 2010. V. 26. № 1. P. 46–52.
- Yaffe E., Fishelovitch D., Wolfson H.J., Halperin D., Nussinov R. // *Nucleic Acids Res.* 2008. V. 36. № suppl 2. P. W210–W215.
- Le Guilloux V., Schmidtke P., Tuffery P. // *BMC bioinformatics.* 2009. V. 10. № 1. P. 168.
- Volkamer A., Kuhn D., Grombacher T., Rippmann F.,

- Rarey M. // *Journal of chemical information and modeling*. 2012. V. 52. № 2. P. 360–372.
37. Laurie A.T.R., Jackson R.M. // *Bioinformatics*. 2005. V. 21. № 9. P. 1908–1916.
38. Hernandez M., Ghersi D., Sanchez R. // *Nucleic Acids Res.* 2009. V. 37. № suppl 2. P. W413–W416.
39. Suplatov D., Kirilin E., Arbatsky M., Takhaviev V., Švedas V. // *Nucleic Acids Res.* 2014. V. 42. № W1. P. W344–W349.
40. Todd A.E., Orengo C.A., Thornton J.M. // *J. Mol. Biol.* 2001. V. 307. № 4. P. 1113–1143.
41. Huang B., Schroeder M. // *BMC Struct. Biol.* 2006. V. 6. № 1. P. 19.
42. Glaser F., Morris R.J., Najmanovich R.J., Laskowski R.A., Thornton J.M. // *Proteins*. 2006. V. 62. № 2. P. 479–488.
43. Kalinina O.V., Gelfand M.S., Russell R.B. // *BMC bioinformatics*. 2009. V. 10. № 1. P. 174.
44. Varfolomeev S.D., Gurevich K.G., Poroykov V.V., Sobolev B.N., Fomenko A.E. // *Dokl. Biochem. Biophys.*, 2001. V. 379. № 1. P. 252–254.
45. Khaliullin I.G., Suplatov D.A., Shalaeva D.N., Otsuka M., Asano Y., Švedas V.K. // *Acta Naturae*. 2010. V. 2. № 2. P. 66–71.
46. Suplatov D.A., Arzhanik V.K., Švedas V.K. // *Acta Naturae*. 2011. V. 3. № 1. P. 93–98.
47. Yang J.S., Seo S.W., Jang S., Jung G.Y., Kim S. // *PLoS Comput Biol.* 2012. V. 8. № 7. P. e1002612.
48. Kalinina O.V., Mironov A.A., Gelfand M.S., Rakhmanova A.B. // *Protein Sci.* 2004. V. 13. № 2. P. 443–456.
49. Suplatov D., Shalaeva D., Kirilin E., Arzhanik V., Švedas V. // *J. Biomol. Struct. Dyn.* 2014. V. 32. № 1. P. 75–87.
50. Campbell E.A., Korzheva N., Mustaev A., Murakami K., Nair S., Goldfarb A., Darst S.A. // *Cell*. 2001. V. 104. № 6. P. 901–912.
51. Porter C.T., Bartlett G.J., Thornton J.M. // *Nucleic Acids Res.* 2004. V. 32. № suppl 1. P. D129–D133.
52. Huang Z., Zhu L., Cao Y., Wu G., Liu X., Chen Y., Wang Q., Shi T., Zhao Y., Wang Y., et al. // *Nucleic Acids Res.* 2011. V. 39. № suppl 1. P. D663–D669.
53. Seifert A., Tatzel S., Schmid R.D., Pleiss J. // *Proteins*. 2006. V. 64. № 1. P. 147–155.
54. Suplatov D., Panin N., Kirilin E., Shcherbakova T., Kudryavtsev P., Švedas V. // *PLoS One*. 2014. V. 9. № 6. P. e100643.
55. Ichiye T., Karplus M. // *Proteins*. 1991. V. 11. № 3. P. 205–217.
56. Liu J., Nussinov R. // *Proc. Natl. Acad. Sci. U.S.A.* 2008. V. 105. № 3. P. 901–906.
57. Dima R.I., Thirumalai D. // *Protein Sci.* 2006. V. 15. № 2. P. 258–268.
58. Reynolds K.A., McLaughlin R.N., Ranganathan R. // *Cell*. 2011. V. 147. № 7. P. 1564–1575.
59. Ferguson A.D., Amezcua C.A., Halabi N.M., Chelliah Y., Rosen M.K., Ranganathan R., Deisenhofer J. // *Proc. Natl. Acad. Sci. U.S.A.* 2007. V. 104. № 2. P. 513–518.
60. Bornscheuer U.T., Huisman G.W., Kazlauskas R.J., Lutz S., Moore J.C., Robins K. // *Nature*. 2012. V. 485. № 7397. P. 185–194.
61. Imming P., Sinning C., Meyer A. // *Nat Rev Drug Discov.* 2006. V. 5. № 10. P. 821–834.
62. Overington J.P., Al-Lazikani B., Hopkins A.L. // *Nat Rev Drug Discov.* 2006. V. 5. № 12. P. 993–996.
63. Arnold F.H. // *Acc. Chem. Res.* 1998. V. 31. № 3. P. 125–131.
64. Reetz M.T. // *Enzyme Catalysis in Organic Synthesis*, Third Edition. 2012. P. 119–190.
65. Reetz M.T., Carballeira J. D. // *Nature protocols*. 2007. V. 2. № 4. P. 891–903.
66. Kazlauskas R.J., Bornscheuer U.T. // *Nature chemical biology*. 2009. V. 5. № 8. P. 526–529.
67. Knowles J., Gromo G. // *Nature Reviews Drug Discovery*. 2003. V. 2. № 1. P. 63–69.
68. Roses A.D., Burns D.K., Chisoso S., Middleton L., Jean P.S. // *Drug Discov. Today*. 2005. V. 10. № 3. P. 177–189.
69. Payne D.J., Gwynn M.N., Holmes D.J., Pompliano D.L. // *Nat Rev Drug Discov.* 2007. V. 6. № 1. P. 29–40.
70. Chan P.F., Holmes D.J., Payne D.J. // *Drug Discovery Today: Therapeutic Strategies*. 2004. V. 1. № 4. P. 519–527.
71. Reetz M.T., Bocola M., Carballeira J.D., Zha D., Vogel A. // *Angew. Chem. Int. Ed. Engl.* 2005. V. 44. № 27. P. 4192–4196.
72. Nilov D.K., Prokhorova E.A., Švedas V.K. // *Acta Naturae*. 2015. V. 7. № 2. P. 57–63.
73. Kuravsky M.L., Barinova K.V., Asryants R.A., Schmalhausen E.V., Muronetz V.I. // *Biochimie*. 2015. V. 115. P. 28–34.
74. Pargellis C., Tong L., Churchill L., Cirillo P.F., Gilmore T., Graham A.G., Grob P.M., Hickey E.R., Moss N., Pav S., et al. // *Nat. Struct. Mol. Biol.* 2002. V. 9. № 4. P. 268–272.
75. Esnouf R., Ren J., Ross C., Jones Y., Stammers D., Stuart D. // *Nat. Struct. Mol. Biol.* 1995. V. 2. № 4. P. 303–308.
76. Conn P.J., Christopoulos A., Lindsley C.W. // *Nat Rev Drug Discov.* 2009. V. 8. № 1. P. 41–54.
77. Vázquez-Figueroa E., Chaparro-Riggers J., Bommaris A.S. // *ChemBioChem*. 2007. V. 8. № 18. P. 2295–2301.
78. Jochens H., Aerts D., Bornscheuer U.T. // *Protein Eng. Des. Sel.* 2010. V. 23. № 12. P. 903–909.
79. Suplatov D.A., Besenmatter W., Švedas V.K., Svendsen A. // *Protein Eng. Des. Sel.* 2012. V. 25. № 11. P. 689–697.
80. Pleiss J. // *Curr. Opin. Biotechnol.* 2011. V. 22. № 5. P. 611–617.
81. Damborsky J., Brezovsky J. // *Curr Opin Chem Biol.* 2014. V. 19. P. 8–16.
82. Suplatov D., Voevodin V., Švedas V. // *Biotechnology J.* 2015. V. 10. № 3. P. 344–355.
83. Cole S., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eglmeier K., Gas S., Barry C.E., et al. // *Nature*. 1998. V. 393. № 6685. P. 537–544.
84. Galperin M.Y., Koonin E.V. // *Curr. Opin. Biotechnol.* 1999. V. 10. № 6. P. 571–578.
85. Moir D.T., Shaw K.J., Hare R.S., Vovis G.F. // *Antimicrob. Agents Chemother.* 1999. V. 43. № 3. P. 439–446.
86. Krissinel E., Henrick K. // *Acta Crystallogr. D Biol. Crystallogr.* 2004. V. 60. № 12. P. 2256–2268.
87. Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. // *Nucleic Acids Res.* 1997. V. 25. № 17. P. 3389–3402.
88. Fischer J.D., Mayer C.E., Söding J. // *Bioinformatics*. 2008. V. 24. № 5. P. 613–620.
89. Menke M., Berger B., Cowen L. // *PLoS Comput Biol.* 2008. V. 4. № 1. P. e10.
90. Notredame C., Higgins D.G., Heringa J. // *J. Mol. Biol.* 2000. V. 302. № 1. P. 205–217.
91. Suplatov D., Kirilin E., Takhaviev V., Švedas V. // *J. Biomol. Struct. Dyn.* 2014. V. 32. № 11. P. 1752–1758.

Cancer Stem Cells: Plasticity Works against Therapy

T. V. Vinogradova*, I. P. Chernov, G. S. Monastyrskaya, L. G. Kondratyeva, E. D. Sverdlov
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences
E-mail: vintv56@gmail.com
Submitted: 02.04.2015

Copyright © 2015 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 Great successes in identification and deciphering of mechanisms of the adult stem cells regulation have given rise to the idea that stem cells can also function in tumors as central elements of their development, starting from the initial stage and continuing until metastasis. Such cells were called cancer stem cells (CSCs). Over the course of intense discussion, the CSCs hypothesis gradually began to be perceived as an obvious fact. Recently, the existence of CSCs has been indeed confirmed in a number of works. However, when are CSCs universal prerequisites of tumors and to what extent their role is essential for tumor evolution remains an issue far from resolved. Likewise, the problem of potential use of CSCs as therapeutic targets remains unsolved. The present review attempts to analyze the issue of cancer stem cells and the potential of targeting them in tumor therapy.

KEYWORDS hierarchical structure of the tumor, clonal evolution, cancer, cancer stem cells, stem cells.

ABBREVIATIONS CSCs, cancer stem cells; SPM, stemness phenotype model.

INTRODUCTION

Major advances in the identification and decoding of the mechanisms of adult stem cells regulation have given rise to the idea that stem cells can also function in tumors, acting as a “driving force” behind their development, all the way from the initial stage to metastasis. Such cells were called cancer stem cells (CSCs). The importance of the issue has led to numerous publications devoted to these hypothetical central players in the development of cancer. With time, the hypothesis has begun to be perceived as a fact, as a kind of dogma that is accepted without question [1]. Indeed, the existence of CSCs has been recently confirmed in a number of studies. However, when their role is crucial for tumor evolution and when their presence is a prerequisite for malignant tumor evolution is far from resolved. Likewise, the issue of the potential use of CSCs as targets for tumor therapy remains unsolved.

The present review attempts to analyze the issue of cancer stem cells and the feasibility of targeting them in tumor therapy. We will not address a number of pertinent issues related to the regulation of the molecular mechanisms of oncogenesis. Neither will we discuss the very important issue of resistance to therapy, which has been covered in recent reviews [2–11].

Most tumors are of monoclonal origin. However, by the time they are detected, they consist of genetically, phenotypically, and epigenetically heterogeneous clones. Two key hypotheses have been put forward to

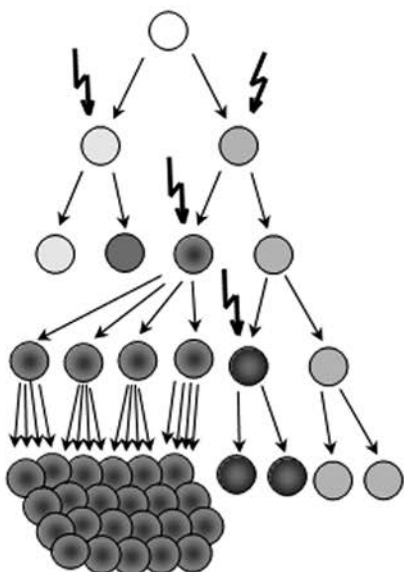
account for this heterogeneity: cancer stem cells and the clonal (stochastic) evolution model [12–19]. Even though these two concepts share common provisions, they are fundamentally different and imply different approaches to the treatment of tumors [12]. A number of theories combining these two concepts have been developed in recent years [20, 21]. A major driving force behind this unification approach has been the data obtained by large-scale sequencing of cancer cell genomes [22]

CLASSICAL STOCHASTIC CLONAL EVOLUTION MODEL

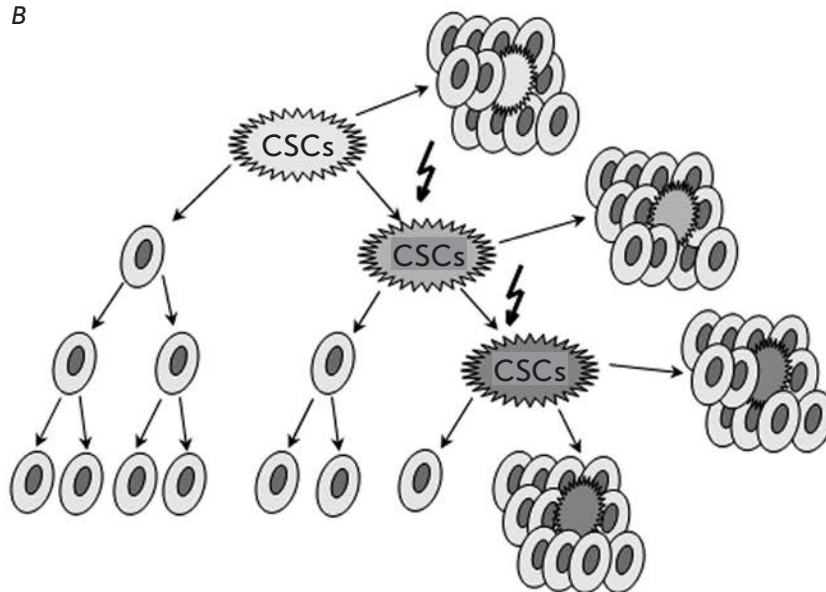
We will start with the classical model that considers evolution of cancer in terms of Darwinian evolution, where cells more or less adapted to survival in a tumor are competing with each other [1, 12, 22–25]. Chronologically, the first model to consider is a stochastic evolution model, which is represented schematically in *Figure A*.

Intratumoral heterogeneity has been traditionally assessed using a stochastic model proposed by Peter Nowell in 1976 [26] (e.g., recent reviews [23, 27, 28]). Nowell relied on data on chromosomal heterogeneity in evolving tumors available at the time. His central idea is a concept of a clone defined as a group of cells derived from the same progenitor cell. Clonal expansion occurs when cells with advantage in fitness (e.g., growth rate) are selected in the course of evolution. Therefore, it implies that genetic and epigenetic changes arising from mutations in

A



B



Models of heterogeneity in solid cancer cells. (A) The clonal evolution model assumes that every cell in a tumor is potentially tumor-initiating. Progression is governed by rare stochastic events operating in all cells. Cells with mutations that acquire a growth advantage will dominate over all other cells in the tumor and will originate a new clone containing cells characterized by a different phenotype and having different proliferative potentials; in a clonogenicity or tumorigenicity assay, some of these cells would have a low probability of exhibiting this potential (modified from: <https://egtheory.wordpress.com/2014/10/25/stochastic-cancer>). (B) The cancer stem cell model states that a particular subset of tumor cells with stem cell-like properties, called “cancer stem cell” (CSC), drives tumor initiation, progression, and recurrence. CSCs are able to self-renew indefinitely and to differentiate, leading to the production of all cell types that make up the rest of the tumor. In clonogenic assays, CSCs have the potential to proliferate extensively and can form new tumors on transplantation (modified from: <http://www.hindawi.com/journals/jo/2008/492643/fig1>). Stochastic mutations are shown by broken arrows

cells can be selected and produce clones with different numbers of cells. Since many different clones are selected in the course of tumor evolution, the tumor becomes polyclonal, even though all cells and clones are derived from a single progenitor (Figure A). This type of polyclonal structure has been confirmed by the sequencing of the genomes of various types of tumors [28].

These cells are genetically and epigenetically different, creating a huge heterogeneity within the tumor [29]. In terms of further evolution, they can also have different functionality, with some of them being more aggressive [30] and, ultimately, leading to metastasis. It is important to stress that mutations occur stochastically; there are no cells or regions in the genome which are favorable to mutations. It has been assumed that cancer cells have a mutator phenotype [9, 31, 32], which accounts for the general non-selective increase in the

mutation rate in “cancerous” genomes compared to normal ones.

It is important to remember that all these processes occur in a particular environment, a niche that can be referred to as a ‘within-a-body’ ecosystem [14]. This ecosystem has a major impact on close selection [33]. It is different in each individual, and this factor, apparently, largely defines the unpredictability of individual tumor development in each patient.

It should also be noted that each malignant tumor is characterized by a large variety of mutations that are different in each individual and can have a different impact on the development of the tumor and emergence of resistance to therapy [26]. The clonal evolution model and CSCs model are not mutually exclusive, since the evolution of CSCs most likely also follows the laws of clonal evolution [27].

Already back in 1976, Nowell offered a very interesting insight into the individual differences in tumors: “One may ultimately have to consider each advanced malignancy as an individual therapeutic problem after as many cells as possible have been eliminated through the nonspecific modalities of surgery, radiation, and chemotherapy. Then, perhaps, immunotherapy becomes a leading candidate for the easiest means of destroying the remainder of the neoplastic clone [26].”

Recently, an interesting stochastic clonal model has been proposed in which all cells are regarded as phenotypically similar to stem cells (stemness phenotype model, SPM) [34]. The term “stemness” [35, 36] has been gaining popularity in recent years and includes all the properties attributed to stem cells; in particular, the ability to self-renew and differentiate. There are attempts to apply this term to CSCs in a more general sense, referring to the ability to maintain and regulate the state of a stem cell [21]. SPM cells possess the properties of stem cells to a greater or lesser degree; upon implantation, they can initiate the development of a tumor; however, in principle, any cancer cell can be tumorigenic. Therefore, according to this model all tumor cells must be considered as targets for elimination in order to defeat the cancer.

There are other models, closely resembling the SPM (for review, see [37]). We will return to this model in the discussion of CSCs-related issues.

It should be noted that the use of any model must take into account that a tumor is a stochastic complex dynamic structure with unpredictable behavior, or, in other words, trajectory of development [38].

DO CSCS EXIST! THE DEFINITION IS ALREADY A CHALLENGE

The general concept of CSCs is represented in *Figure B*. First of all, we should define CSCs and their distinctive characteristics. CSCs-related issues were discussed in recent reviews [6, 21, 35, 39–43], and in other publications cited in this paper.

The obvious role of stem cells during embryogenesis [39, 44–47] and the hypothesis that a normal stem cell can become one in which the process of malignant transformation begins [39, 48, 49] were logically extended to tumors and has become the foundation of the CSCs model [50, 51]. A large number of articles have been devoted to CSCs. Sometimes, CSCs are referred to as tumor-initiating cells (TIC) or tumor-propagating cells (TPC) [1, 43]. We will use the term CSCs.

Eventually, both the hypothesis of CSCs and their existence began to be perceived as self-evident, as a kind of dogma which was accepted without question [1]. Nevertheless, the existence of CSCs and the exact

criteria that distinguish CSCs from other cancer cells remain unresolved and intensely debated issues.

The concept of cancer stem cells was first proposed in the middle of the year 1990 (for review, see [21, 52–56]) and has since become the subject of intense discussions and clarifications, both in terms of its substance and, which is very important, in terms of nomenclature. This concept is based on the trajectory which a normal stem cell follows during differentiation. Initially, it was believed that mutations leading to cancer occur in a normal stem cell and result in its transformation into a cancer stem cell. The ability of a normal stem cell to self-renew and differentiate was automatically attributed to putative cancer stem cells. In this case, CSCs must always produce new cells: cancer stem cells and more differentiated progenies that are capable of only a limited number of divisions. Mutations accumulate during CSCs division. More differentiated cells have a limited ability to mutate, and their genome reflects the state of the cancer stem cell from which they originate. Clones are produced. Although it is generally believed that a cancerous cell is derived from a normal stem cell, during the development of a tumor CSCs can arise from various tumor cells, including differentiated ones, by the process of dedifferentiation (see below).

In 2011, a Workshop on CSCs developed key recommendations on definitions set forth in [43]. According to these recommendations, the term CSCs refers to neoplastic cells that can propagate or maintain an invasive solid tumor or leukemia over an indefinite or prolonged period of time.

This definition is apparently the result of major efforts to reach a consensus, since there are many other definitions in use that include, for example, such an important property as self-renewal.

Here are some definitions from the latest reviews published in the most prestigious journals:

Nature Review Cancer, 2012 [57]: “We have chosen to define CSCs as the cells within a malignant clonal population that can propagate the cancer ... This definition assumes that not all of the cells within a population of malignant cells have this property ... This definition also implies that CSCs are responsible for generating all of the cells within the malignant population that lack cancer-propagating ability (as well as those cells that perpetuate it). It also implies that the choice of these alternative fates by CSCs is embedded in an intrinsically established intracellular molecular response network that is likely to be related to the tissue from which the CSCs originate and that the loss of cancer-propagating ability is not readily reversible *in vivo*.”

Cell Cycle, 2013 [58]: “[The concept of CSCs states] that as in the normal somatic stem cells (SSC) ... a small population of cells, the cancer stem cells (CSC), would

reproduce *ad infinitum* and generate the very diverse, limited lifespan, multilineage differentiated majority of cells in a cancer, called the derived population cells.”

Cell Stem Cell, 2012. [42]: “The cancer stem cell (CSC) model postulates a hierarchical organization of cells such that only a small subset is responsible for sustaining tumorigenesis and establishing the cellular heterogeneity inherent in the primary tumor. Although CSCs exhibit the stem cell properties of self-renewal and differentiation, they do not necessarily originate from the transformation of normal tissue stem cells.”

Nature Review Drug Discovery, 2014 [6]: “The cancer stem cell (CSC) hypothesis posits the existence of subpopulations of neoplastic cells within a tumor that exhibit an elevated ability to seed new tumors upon experimental implantation in appropriate animal hosts. Implicit in this power is the ability of such cells to divide asymmetrically, yielding daughters that remain as CSCs (the trait of self-renewal) as well as daughters that differentiate into the neoplastic cells forming the bulk of the tumor... The existence of multiple subpopulations within a tumor with distinct tumor-initiating powers is no longer a matter of speculation and hypothesis. Accordingly, the use of the term “cancer stem cell paradigm” now seems to be more appropriate... To date most CSCs are not known to differentiate into more than a single cell type – the cells composing the bulk of the tumor... The phenotypes of CSCs are complex, variable from one tumor to another ... hence CSCs are often difficult to rigorously define by associating them with traits beyond their shared functional trait of tumor-initiating ability. Moreover, the existence of CSCs within tumors implies that cancer cells sharing a common genetic make-up can nevertheless exist in at least two alternative phenotypic states – CSCs and non-CSCs.”

Cell Stem Cell, 2014 [21]: “For many cancers, CSCs represent a distinct population that can be prospectively isolated from the remainder of the tumor cells and can be shown to have clonal long-term repopulation and self-renewal capacity—the defining features of a CSC. However, in some cancer types it has not been possible to distinguish CSCs from non-CSCs because most cells have CSC function. Such tumors seem to be homogeneous or possess a very shallow hierarchy.”

Self-renewal and long-term ability to generate more differentiated cancerous cells were added to the requirements only in 2015 [59].

It should be emphasized that according to the CSCs concept, CSCs form a *separate* population in a tumor, which differs from the bulk of the tumor in their ability to initiate new tumors when implanted, self-renew and exhibit the presence of phenotypic markers, which distinguish them from the bulk. In general, the concept of

CSCs is a hierarchical model with CSCs as the source; therefore, the CSCs model is sometimes referred to as hierarchical. It is very similar to the hierarchy observed in adult stem cells, which produce a sequence of more differentiated cells.

Thus, according to the concept, the CSCs population has the following features:

1. It constitutes a small proportion of the total population of tumor cells.
2. It expresses a specific set of characteristic surface markers that distinguish it from the bulk of other cells.
3. In contrast to other tumor cell populations, it selectively retains the ability to initiate tumors.
4. It supports the growth of a heterogeneous mass containing the full repertoire of partially (or fully) differentiated cancer cells capable of several differentiations or at the stage of final differentiation.
5. It forms a separate pool of cells that can be identified by biological and physico-chemical methods. (There should be at least two pools of cells in tumors: CSCs and their derivative cells that are differentiated to different extents [1].)
6. Just like regular stem cells, it displays an ability for unlimited self-renewal and differentiation in many directions [37].
7. It exhibits high resistance to standard therapy. The possible reasons for the increased resistance of putative CSCs have been detailed in recent reviews [37, 60]. They include selective expression of some members of the family of multidrug resistance transporters, increased expression of anti-apoptotic molecules, increased DNA repair capacity, activation of specific stem cell survival signals (prosurvival signaling), in particular, Notch, Hedgehog (Hh), Wnt, JAK/STAT, and others. However, this issue remains a subject of debate and more studies are required, primarily those that would confirm that the observed effects are associated with CSCs.

It is important to understand that CSCs (and cancer cells in general) are characterized by higher intratumoral heterogeneity. They form subclones in the tumor, but within a subclone each cell differs from the other in the structure of its genome, the nature of transcriptome, proteome, etc. In the course of tumor development, this heterogeneity leads to different and constantly changing combinations of molecular defects in the subclones. Since the CSCs that drive the reproduction of various subclones are different, the subclones have different malignant potentials, can exist in different intratumoral microenvironments, and have different ways of interacting with these microenvironments.

Consequently, every existing and dividing subclone must contain its own CSCs compartment with unique genomic and epigenetic characteristics and each sub-

Table 1. Typical surface antigens that occur with increased frequency in putative CSCs and were targeted in clinical trials [59] (see also Ref. [105])

Surface antigen	Type of cancer, containing the antigen in CSCs	Drug used
CD20	ALL, CLL	Rituximab
CD33	CML, AML	GO
CD44	AML	mAb
CD52	5q-AML, CLL	Alemtuzumab
CD123	AML	mAb
EGFR	Colon-Ca	Cetuximab
HER2/neu	Breast/Gastric/Ovarian-Ca	Trastuzumab

Note. ALL, acute lymphoblastic leukemia; CLL, chronic lymphocytic leukemia; CML, chronic myeloid leukemia; AML, acute myeloid leukemia; mAb, monoclonal antibody; Ca, carcinoma.

clone can give rise to new subclones, whose properties will be different from the original subclone. If a new cell with increased malignant potential appears in a new subclone, it will become a new CSC and will be able to produce new CSCs.

The existence of human CSCs has been confirmed experimentally by demonstrating the ability of cells obtained directly from a patient to produce a malignant derivative population when transplanted to immunodeficient mice. An inherent phenotypic difference allowing to physically separate CSCs from the bulk of cells, for example, a surface antigen, is used to demonstrate that CSCs are truly different from the bulk of tumor cells.

Identification of surface antigens characteristic of CSCs is the subject of many experimental papers and reviews, and the list of such markers is constantly growing. *Table 1* lists some of the markers and intracellular proteins that are characteristic of putative CSCs. They will be discussed below.

The putative CSCs can be separated from the bulk of tumor cells by flow cytometry using specific surface markers [61, 62]. The isolated cells have higher tumorigenicity in case of xenotransplantation to immunocompetent animals. The essential biochemical characteristic of these cells is overexpression of cytoprotective enzymes, such as aldehyde dehydrogenase (ALDH), and pumps that remove toxic compounds from the cells; for example, ABC-transporters [61, 62].

This distinction, however, is not universal, because there are CSCs that do not possess these markers and those with non-CSCs markers [35, 63]. Thus, the markers are not stable phenotypic characteristics; they can vary from one individual to another and change at different stages of tumor development. Therefore, phenotypic data alone cannot be used as sufficient proof of the presence of a separate CSCs population. It is assumed that the mechanisms that define the specific

properties of CSCs are unstable because they can be associated with changes in the epigenome, which are frequent in tumor populations [43]. There are other reasons to believe that the CSCs phenotype can be very unstable [43].

Some aforementioned CSCs characteristics, in particular asymmetric division and irreversible process of transition to a differentiated state, are either not proven or seem unlikely [1], and the irreversibility of the transition has been disproved in recent experiments (see below). The claim that CSCs represent a relatively small fraction of the general population of cells should also be treated with caution, since a fraction of CSCs can vary greatly (from 0.1 to 30%) depending on the tumor type and design of the experiment [27, 64–66].

It seems that the authors of the review [27] are right to believe that some types of cancers develop through CSCs, while others do not. It has even been suggested that this may apply to the same type of cancer in different patients.

If the concept of CSCs is true, then in practical terms this implies that all the measurements that we perform, e.g., full genome sequencing, refer to the bulk of the tumor which has already ceased active division and accumulation of mutations. And the CSCs that represent an absolute minority and continue to actively divide after the separation of the bulk of the tumor and acquire mutations which give them new functionalities are factored out. Any positive outcome in a study of the total mass should be attributed to the fact that all the cells of this mass are derived from CSCs and share common genetic elements with them. However, many CSCs elements, which may be important for diagnosis and prognosis, may be overlooked by researchers. This is especially true of epigenetic changes. One can predict that heterogeneity within the CSCs fraction must be higher than in the bulk, which consists mostly of differentiated and non-dividing tumor cells. A recent

Table 2. Signaling pathways targeted in clinical trials for the inhibition of various cancers (modified from [105])

Cancer type	Targeted signaling pathway	Therapy: mono- or combination
Colon cancer	STAT/ β - β -catenin /Nanog	Combination
Stomach cancer	STAT/ β - β -catenin /Nanog	Monotherapy Combination
GI cancer	STAT/ β - β -catenin /Nanog	Combination
Hepatocellular carcinoma	STAT/ β - β -catenin /Nanog Wnt/FZD8Fc/FAK	Combination Combination
Mesothelioma	FAK	Monotherapy
Squamous cell lung cancer	Notch FAK	Combination Monotherapy
Testicular cancer	Notch FAK	Combination Combination
Pancreas cancer	STAT/ β - β -catenin /Nanog Notch	Combination Combination
Small cell lung cancer	Notch	Combination

analysis of exomes [67] showed that the majority of somatic mutations in the CSCs fraction and in the bulk of the tumor are identical. These data can be interpreted as the result of constant dynamic transitions from CSCs to the differentiated state, and vice versa. This will be discussed below.

Many aspects of the CSC model remain controversial, but several experiments employing modern technologies to trace cells during development, including lineage-tracing studies, provide strong evidence in favor of a more or less stable existence of CSCs and a hierarchical organization of the tumor at least in some cases. These experiments were conducted in murine models of brain tumors [68], small intestine [69] and skin cancers [70], human colorectal adenomas [71], and gliomas [72] (see also comment in [73]). These studies showed that the majority of the investigated tumor cells had limited proliferative capacity and, apparently, originated from subpopulations with properties similar to those of CSCs.

CSCS NICHES: DEFINITIONS AND CHALLENGES

It has been established that normal stem cells have so-called niches; physiological microenvironments composed of specialized cells that are involved in the regulation of stem cells, functioning through various types of signaling. In this case, the definition of a niche in which they exist is fairly straightforward [44, 45, 74-77]. It is a quite well-defined area around discreetly localized functioning stem cells in the tissue, although the existence of stem cells and their respective niches have not been demonstrated for all normal tissues [39]. The classic definition states that “stem-cell populations are established in ‘niches’ — specific anatomic locations that regulate how they participate in tissue generation, maintenance and repair. The niche saves stem cells from depletion, while protecting the host

from over-exuberant stem-cell proliferation. It constitutes a basic unit of tissue physiology.” [78].

Using adult stem cells as an example, it has been proposed that CSCs also have niches and that interaction of CSCs with niches may regulate self-renewal, proliferation, and differentiation of CSCs [77, 79, 80]. However, in the case of CSC there are more questions than answers (see recent reviews [1, 81, 82]). Many authors (e.g., [79]) consider a solid tumor to be an “organ” which consists of cancer cells and the stroma, which occupies most of the tumor volume and creates a microenvironment that can be considered an analogue of normal cell niches.

The CSCs niche is a microenvironment which has no morphological structure [83]. However, according to some authors (see, e.g., review [84]), the CSCs niches differ from the overall microenvironment. The cells inside the niche produce factors that stimulate CSCs self-renewal, induce angiogenesis, and recruit immune and other stromal cells, which secrete additional factors that contribute to the invasion and metastasis of cancer cells.

The interaction of tumor stromal elements with putative CSCs is covered in a large number of publications. For example, there is evidence that anti-tumor agents affect putative CSCs *in vitro* and *in vivo* differently [43], and this may mean that some important components of CSCs regulation are provided by the microenvironment (niche). The microenvironment includes the extracellular matrix, mesenchymal and endothelial cells, immune system cells, adhesion molecules, various growth factors, and cytokines and their receptors [57, 85]. It is assumed that blood vessels can also play a role in the creation of niches [81], as they do in the case of normal stem cells. The tumor stroma secretes factors that regulate the behavior of cancer cells [80] and actively support the growth of the tumor via

neo- and by angiogenesis [86]. The microenvironment determines the fate of a tumor, serves as a barrier for the therapeutic intervention, and can affect the plasticity of tumor cells, for example, transitions from CSCs to non-CSCs [8].

Two-way interactions between cancer cells and the stroma are widely discussed, especially, the role of stroma in tumor development and, in particular, the acquisition of such highly important qualities as invasiveness and metastatic potential [87]. It is believed that not only does the stromal niche affect cancer cells, but the reverse is also true: cancer cells (which mainly means CSCs) can affect the stroma as well and use it for their development [81], in particular, to create pre-metastatic niches [80, 85, 86, 88].

Despite the widely discussed importance of stromal niches in tumor development, there are very little data reliably confirming their function at the molecular level, as well as at the level of information transport; e.g., whether the transport is carried out in paracrine, autocrine, or in any other way [43].

More detailed information on this matter will help to better evaluate the real role of the niche in the development of tumors and to develop rational therapeutic strategies.

NORMAL AND CANCEROUS NON-STEM CELLS CAN SPONTANEOUSLY TRANSFORM INTO A STATE SIMILAR TO STEM CELLS. APPARENTLY, THERE IS NO STRICT BARRIER BETWEEN CSCS AND NON-CSCS CANCER CELLS

Different types of cancers and, maybe, even the same type of cancer in different patients may follow either the CSCs model or the stochastic evolution model. In dealing with such labile systems as stem cells, cancer cells, or CSCs one should always consider the possibility of their phenotypic restructuring as a result of epigenetic processes. In 2011, three groups [89, 90] (see reviews [35, 91]) of researchers used cancer cell lines and primary tumors to describe the acquisition of a self-renewing capacity in non-CSCs populations. For example, they described mammary epithelial cells capable of spontaneous de-differentiation into a “stem-like” state [92]. The malignant transformation enhanced this ability, making it possible for ordinary cancer cells to transform into a state similar to CSCs’ *in vitro* and *in vivo*. These data demonstrated the high plasticity of stem cells in general and CSCs in particular, as well as the ease of interconversion of non-stem cells into stem cells and vice versa, especially in malignant tumors.

These findings provide ground to believe that such reversibility and lack of rigid hierarchy in stem and non-stem cells may be commonplace, making cancer therapy, in particular therapy aimed at CSCs, more challenging.

The observed plasticity of cancer cells and the possibility of transition between stem and non-stem cells introduce additional complexity in the study of the role of cancer stem cells in carcinogenesis. This plasticity may depend on a number of factors, with signals from the microenvironment and intercellular interactions in the niches playing an important role [8]. These transitions most likely are stochastic epigenetic or genetic events influenced by the microenvironment of cancer cells and intercellular interactions in the niches [20]. For example, CSCs can be converted into a non-CSC, and *vice versa* [8], existing in a dynamic equilibrium [93].

Therefore, one of the important aspects of the CSCs concept, i.e. the presence of an irreversibly separated CSCs fraction in the tumor mass (see above), is untenable overall [88, 91]. Differentiated cancer cells and CSCs are in a constant state of mutual transformation [94]. Environmental factors, including growth factors, can cause transitions between the states. Moreover, the plasticity of CSCs may cause a transformation of epithelial cancerous cells into mesenchymal (and *vice versa*) [6, 41, 67, 95-98]. This phenotypic plasticity is caused by both mesenchymal-epithelial and epithelial-mesenchymal transitions and, apparently, by genetic, epigenetic, and intracellular and intercellular signaling programs [58]. It can be assumed that the existence of CSCs and their plasticity depend on both internal factors (genetic and epigenetic architecture of cancer cells) and on the microenvironment. Since both types of factors are highly specific for the type of tumor and each individual patient, it is easy to imagine a situation where in some cases CSCs are present, and the barrier between CSCs and non-CSCs is high enough to produce an isolated permanent CSCs fraction, while in other cases this barrier is low and there is a constant interconversion of CSCs and non-CSCs cells. Finally, in some types of cancers and in some patients CSCs groups are not produced and the tumor evolution proceeds by the stochastic clonal mechanism.

CONCLUSION: CSCS AS THERAPEUTIC TARGETS

Numerous publications are dedicated to the potential use of CSCs for therapeutic purposes. This is not surprising.

The bulk of the tumor mass has a limited proliferative capacity, and CSCs, if they really are the driving force behind the development of the tumor, represent only a small part of it. As a result, therapeutic agents act mainly on the highly differentiated part of the tumor and their efficiency is low. Targeting the active CSCs minority would greatly increase the efficiency of the therapy. Subsequently, in recent years many authors (see e.g. [59]) have considered (and still do consider) CSCs as a rather promising therapeutic target. New

strategies and approaches to therapy focused on CSCs are being developed very intensively. The literature devoted to the principles and methods developed in this area is very abundant and contradictory. Readers can get acquainted with the proposed ideas and methods in a number of reviews published in 2015 [8, 59, 99–105]. CSCs features, which were targeted by therapeutic interventions, are listed in *Table 1*.

Despite the multitude of proposed approaches, there are only two basic strategies, which differ in the choice of target (see [35] and reviews, cited above). The first strategy targets surface antigens, which are presumably characteristic of CSCs. The second strategy relies on the fact that self-renewing putative CSCs are in many respects similar to embryonic stem cells (see, e.g. [35, 59]) and should express embryonic signaling pathways, which are not typical of adult cells [59].

Speaking of surface antigens, there are known cases of CSCs, normal cells, and normal stem cells having the same antigens. As a result, a therapy targeting surface antigens often causes serious side effects. For example, the therapeutic monoclonal antibodies gemtuzumab ozogamicin (GO, anti-CD33) and alemtuzumab (anti-CD52) have been recently withdrawn from the oncological market because of their toxicity. Therefore, the search for surface CSCs markers with higher selectivity continues [59].

Various components of the signaling pathways known to be involved in embryogenesis, such as Notch, Hedgehog (HH) and Wnt have been also investigated as targets for influencing CSCs. Some of them are listed in *Table 1* (see [59, 99] and others cited in reviews). It is still too early to judge the success of any particular therapy, because the trials are in their early stages (*Table 2*). However, it should be noted that treatment targeting embryonic signaling systems that are active in adult stem cells can cause serious side effects.

An article published in *Science* in early 2015 [102] covers various points of view on the existence of cancer stem cells and their significance for the evolution and treatment of cancer. In particular, it lists both practical and theoretical considerations on the feasibility of the use of cancer stem cells as the only therapeutic target.

In our review, we have attempted to address these issues from different perspectives. We believe that these data allow us to express doubt that a therapy aimed at CSCs will be successful. For example, it seems that easy transition between CSC and undifferentiated cells, and *vice versa*, will make it impossible to completely eradicate CSCs. Moreover, there are concerns about the potential emergence of cells that will initiate a new tumor from the remaining non-differentiated cancer cells. Clinical trials of drugs that target CSCs are in too early a stage to pass judgment on their success. Kaiser J. [102] in the conclusion of his article writes: “For now, cancer patients, researchers and physicians, and investors [of companies using CSCs as targets] ... will anxiously wait for data to roll in from the clinical trials. For those with a stake in treatments, the results could bring hope. For researchers debating the reality of cancer stem cells, though, they may not bring resolution. Says Jeremy Rich of the Cleveland Clinic in Ohio, who is studying stem cells in brain cancer, “Even if we’re wildly successful, which I don’t think we will be, I don’t think there will be a black-and-white answer.”

In 1209 during the siege of Beziers, which was defended by true Catholics and heretics, Arnold Amalric, the papal legate and a prominent participant of the Albigensian Crusade, allegedly answered a question on how to distinguish true believers from heretics with “Kill them all and let God sort them out.” This cannot be applied to people, of course, but in our opinion it is the only right strategy for cancer therapy: there is no need to look for differences between cancer cells, just kill them all. In our work, we adhere to this principle, combining two strategies of total eradication of cancer cells: suicide gene therapy and immunotherapy. The results are very promising, at least in preclinical trials [106].

The analysis of clinical trials [105], examples of which are listed in *Table 2*, shows that the majority of researchers are also using a combination therapy, which is absolutely justified given the extreme plasticity of CSCs.

This work was supported by the Russian Science Foundation (project No. 14-50-00131).

REFERENCES

1. Maenhaut C., Dumont J.E., Roger P.P., van Staveren W.C. // *Carcinogenesis*. 2010. V. 31. № 2. P. 149–158.
2. Moitra K., Lou H., Dean M. // *Clin. Pharmacol. Ther.* 2011. V. 89. № 4. P. 491–502.
3. Cunningham J.J., Gatenby R.A., Brown J.S. // *Mol. Pharmacol.* 2012. V. 8. № 6. P. 2094–2100.
4. Rycak K., Tang D.G. // *Int. J. Radiat. Biol.* 2014. V. 90. № 8. P. 615–621.
5. Dean M. // *J. Mammary Gland Biol. Neoplasia*. 2009. V. 14. № 1. P. 3–9.
6. Pattabiraman D.R., Weinberg R.A. // *Nat. Rev. Drug Discov.* 2014. V. 13. № 7. P. 497–512.
7. Colak S., Medema J.P. // *FEBS J.* 2014. V. 281. № 21. P. 4779–4791.
8. Cabrera M.C., Hollingsworth R.E., Hurt E.M. // *W. J. Stem Cells*. 2015. V. 7. № 1. P. 27–36.
9. Roberts S.A., Gordenin D.A. // *Nat. Rev. Cancer*. 2014. V. 14. № 12. P. 786–800.
10. Du W., Elemento O. // *Oncogene*. 2015. V. 34. P. 3215–3225.
11. Tam W.L., Weinberg R.A. // *Nat. Med.* 2013. V. 19. № 11. P. 1438–1449.

REVIEWS

12. Campbell L.L., Polyak K. // *Cell Cycle*. 2007. V. 6. № 19. P. 2332–2338.
13. Ding L., Ellis M.J., Li S., Larson D.E., Chen K., Wallis J.W., Harris C.C., McLellan M.D., Fulton R.S., Fulton L.L., et al. // *Nature*. 2010. V. 464. № 7291. P. 999–1005.
14. Greaves M., Maley C.C. // *Nature*. 2012. V. 481. № 7381. P. 306–313.
15. Marusyk A., Polyak K. // *Science*. 2013. V. 339. № 6119. P. 528–529.
16. Navin N., Kendall J., Troge J., Andrews P., Rodgers L., McIndoo J., Cook K., Stepansky A., Levy D., Esposito D., et al. // *Nature*. 2011. V. 472. № 7341. P. 90–94.
17. Yates L.R., Campbell P.J. // *Nat. Rev. Genet.* 2012. V. 13. № 11. P. 795–806.
18. Collisson E.A., Cho R.J., Gray J.W. // *Nat. Rev. Clin. Oncol.* 2012. V. 9. № 11. P. 621–630.
19. Swanton C., Burrell R.A., Futreal P.A. // *Breast Cancer Res.* 2011. V. 13. № 1. P. 104.
20. Shah M., Allegrucci C. // *Breast Cancer (Dove Med. Press)*. 2012. V. 4. P. 155–166.
21. Kreso A., Dick J.E. // *Cell Stem Cell*. 2014. V. 14. № 3. P. 275–291.
22. Ma Q.C., Ennis C.A., Aparicio S. // *Curr. Opin. Genet. Dev.* 2012. V. 22. № 1. P. 3–9.
23. Merlo L.M., Pepper J.W., Reid B.J., Maley C.C. // *Nat. Rev. Cancer*. 2006. V. 6. № 12. P. 924–935.
24. Brosnan J.A., Iacobuzio-Donahue C.A. // *Semin. Cell Dev. Biol.* 2012. V. 23. № 2. P. 237–242.
25. Sabaawy H.E. // *J. Leukemai (Los Angel)*. 2013. V. 1. № 4. P. 1000124.
26. Nowell P.C. // *Science*. 1976. V. 194. № 4260. P. 23–28.
27. Shackleton M., Quintana E., Fearon E.R., Morrison S.J. // *Cell*. 2009. V. 138. № 5. P. 822–829.
28. Aparicio S., Caldas C. // *N. Engl. J. Med.* 2013. V. 368. № 9. P. 842–851.
29. Sverdlov E.D. // *Curr. Gene Ther.* 2011. V. 11. № 6. P. 501–531.
30. Podlaha O., Riester M., De S., Michor F. // *Trends Genet.* 2012. V. 28. № 4. P. 155–163.
31. Loeb L.A., Bielas J.H., Beckman R.A. // *Cancer Res.* 2008. V. 68. № 10. P. 3551–3557; discussion 3557.
32. Prindle M.J., Fox E.J., Loeb L.A. // *Curr. Drug Targets*. 2010. V. 11. № 10. P. 1296–1303.
33. Polyak K., Haviv I., Campbell I.G. // *Trends Genet.* 2009. V. 25. № 1. P. 30–38.
34. Cruz M.H., Siden A., Calaf G.M., Delwar Z.M., Yakisich J.S. // *ISRN Oncol*. 2012. V. 2012. P. 392647.
35. O'Connor M.L., Xiang D., Shigdar S., Macdonald J., Li Y., Wang T., Pu C., Wang Z., Qiao L., Duan W. // *Cancer Lett*. 2014. V. 344. № 2. P. 180–187.
36. Snippert H.J., Clevers H. // *EMBO Rep*. 2011. V. 12. № 2. P. 113–122.
37. Luo J., Zhou X., Yakisich J.S. // *Onco Targets Ther*. 2014. V. 7. P. 1129–1134.
38. Sverdlov E.D. // *Ross Fiziol Zh Im I M Sechenova*. 2014. V. 100. № 5. P. 505–541. (in Russian).
39. White A.C., Lowry W.E. // *Trends Cell Biol*. 2015. V. 25. № 1. P. 11–20.
40. Allegra A., Alonci A., Penna G., Innao V., Gerace D., Rotondo F., Musolino C. // *Cancer Invest*. 2014. V. 32. № 9. P. 470–495.
41. van de Stolpe A. // *Am. J. Cancer Res*. 2013. V. 3. № 1. P. 107–116.
42. Visvader J.E., Lindeman G.J. // *Cell Stem Cell*. 2012. V. 10. № 6. P. 717–728.
43. Valent P., Bonnet D., De Maria R., Lapidot T., Copland M., Melo J.V., Chomienne C., Ishikawa F., Schuringa J.J., Stassi G., et al. // *Nat. Rev. Cancer*. 2012. V. 12. № 11. P. 767–775.
44. Rempel P., Mesa K.R., Greco V. // *Nature*. 2013. V. 502. № 7472. P. 513–518.
45. O'Brien L.E., Bilder D. // *Annu. Rev. Cell Dev. Biol*. 2013. V. 29. P. 107–136.
46. Fuchs E. // *Cell*. 2009. V. 137. № 5. P. 811–819.
47. Alison M.R., Islam S. // *J. Pathol*. 2009. V. 217. № 2. P. 144–160.
48. Shibata M., Shen M.M. // *Bioessays*. 2013. V. 35. № 3. P. 253–260.
49. Blanpain C. // *Nat. Cell Biol*. 2013. V. 15. № 2. P. 126–134.
50. Shackleton M. // *Semin. Cancer Biol*. 2010. V. 20. № 2. P. 85–92.
51. Beck B., Blanpain C. // *Nat. Rev. Cancer*. 2013. V. 13. № 10. P. 727–738.
52. Dick J.E. // *Blood*. 2008. V. 112. № 13. P. 4793–4807.
53. Dick J. // *Cancer Discov*. 2013. V. 3. № 2. P. 131.
54. Ailles L.E., Weissman I.L. // *Curr. Opin. Biotechnol*. 2007. V. 18. № 5. P. 460–466.
55. Dalerba P., Cho R.W., Clarke M.F. // *Annu. Rev. Med*. 2007. V. 58. P. 267–284.
56. Schatton T., Frank N.Y., Frank M.H. // *Bioessays*. 2009. V. 31. № 10. P. 1038–1049.
57. Nguyen L.V., Vanner R., Dirks P., Eaves C.J. // *Nat. Rev. Cancer*. 2012. V. 12. № 2. P. 133–143.
58. Antoniou A., Hebrant A., Dom G., Dumont J.E., Maenhaut C. // *Cell Cycle*. 2013. V. 12. № 24. P. 3743–3748.
59. Schulenburg A., Blatt K., Cerny-Reiterer S., Sadovnik I., Herrmann H., Marian B., Grunt T., Zielinski C., Valent P. // *J. Hematol. & Oncol*. 2015. V. 8. № 1. P. 16.
60. Sotiropoulou P.A., Christodoulou M.S., Silvani A., Herold-Mende C., Passarella D. // *Drug Discov Today*. 2014. V. 19. № 10. P. 1547–1562.
61. Alison M.R., Lim S.M., Nicholson L.J. // *J. Pathol*. 2011. V. 223. № 2. P. 147–161.
62. Alison M.R., Lin W.R., Lim S.M., Nicholson L.J. // *Cancer Treat. Rev*. 2012. V. 38. № 6. P. 589–598.
63. Medema J.P. // *Nat. Cell Biol*. 2013. V. 15. № 4. P. 338–344.
64. Quintana E., Shackleton M., Sabel M.S., Fullen D.R., Johnson T.M., Morrison S.J. // *Nature*. 2008. V. 456. № 7222. P. 593–598.
65. Hill R.P. // *Cancer Res*. 2006. V. 66. № 4. P. 1891–1895; discussion P. 1890.
66. Adams J.M., Strasser A. // *Cancer Res*. 2008. V. 68. № 11. P. 4018–4021.
67. Klevebring D., Rosin G., Ma R., Lindberg J., Czene K., Kere J., Fredriksson I., Bergh J., Hartman J. // *Breast Cancer Res*. 2014. V. 16. № 4. P. R72.
68. Chen J., Li Y., Yu T.S., McKay R.M., Burns D.K., Kernie S.G., Parada L.F. // *Nature*. 2012. V. 488. № 7412. P. 522–526.
69. Schepers A.G., Snippert H.J., Stange D.E., van den Born M., van Es J.H., van de Wetering M., Clevers H. // *Science*. 2012. V. 337. № 6095. P. 730–735.
70. Driessens G., Beck B., Caauwe A., Simons B.D., Blanpain C. // *Nature*. 2012. V. 488. № 7412. P. 527–530.
71. Humphries A., Cereser B., Gay L.J., Miller D.S., Das B., Gutteridge A., Elia G., Nye E., Jeffery R., Poulson R., et al. // *Proc. Natl. Acad. Sci. USA*. 2013. V. 110. № 27. P. E2490–2499.
72. Wang J., Ma Y., Cooper M.K. // *Transl. Cancer Res*. 2013. V. 2. № 5. P. 429–441.
73. Gilbertson R.J., Graham T.A. // *Nature*. 2012. V. 488. № 7412. P. 462–463.

74. Xie T., Li L. // *Development*. 2007. V. 134. № 11. P. 2001–2006.
75. Scadden D.T. // *Cell*. 2014. V. 157. № 1. P. 41–50.
76. Clevers H., Loh K.M., Nusse R. // *Science*. 2014. V. 346. № 6205. P. 1248012.
77. Sneddon J.B., Werb Z. // *Cell Stem Cell*. 2007. V. 1. № 6. P. 607–611.
78. Scadden D.T. // *Nature*. 2006. V. 441. № 7097. P. 1075–1079.
79. Ye J., Wu D., Wu P., Chen Z., Huang J. // *Tumour Biol*. 2014. V. 35. № 5. P. 3945–3951.
80. Saltarella I., Lamanuzzi A., Reale A., Vacca A., Ria R. // *W. J. Stem Cells*. 2015. V. 7. № 1. P. 84–95.
81. Takakura N. // *Cancer Sci*. 2012. V. 103. № 7. P. 1177–1181.
82. Borovski T., De Sousa E.M.F., Vermeulen L., Medema J.P. // *Cancer Res*. 2011. V. 71. № 3. P. 634–639.
83. Junttila M.R., de Sauvage F.J. // *Nature*. 2013. V. 501. № 7467. P. 346–354.
84. Plaks V., Kong N., Werb Z. // *Cell Stem Cell*. 2015. V. 16. № 3. P. 225–238.
85. Ghiabi P., Jiang J., Pasquier J., Maleki M., Abu-Kaoud N., Rafii S., Rafii A. // *PLoS One*. 2014. V. 9. № 11. P. e112424.
86. Horimoto Y., Polanska U.M., Takahashi Y., Orimo A. // *Cell Adh. Migr*. 2012. V. 6. № 3. P. 193–202.
87. Park T.S., Donnenberg V.S., Donnenberg A.D., Zambidis E.T., Zimmerlin L. // *Curr. Pathobiol. Rep*. 2014. V. 2. № 1. P. 41–52.
88. Fessler E., Dijkgraaf F.E., De Sousa E.M.F., Medema J.P. // *Cancer Lett*. 2013. V. 341. № 1. P. 97–104.
89. Chaffer C.L., Brueckmann I., Scheel C., Kaestli A.J., Wiggins P.A., Rodrigues L.O., Brooks M., Reinhardt F., Su Y., Polyak K., et al. // *Proc. Natl. Acad. Sci. USA*. 2011. V. 108. № 19. P. 7950–7955.
90. Iliopoulos D., Hirsch H.A., Wang G., Struhl K. // *Proc. Natl. Acad. Sci. USA*. 2011. V. 108. № 4. P. 1397–1402.
91. Xie X., Teknos T.N., Pan Q. // *Stem Cells Transl. Med*. 2014. V. 3. № 10. P. 1111–1115.
92. Chaffer C.L., Marjanovic N.D., Lee T., Bell G., Kleer C.G., Reinhardt F., D'Alessio A.C., Young R.A., Weinberg R.A. // *Cell*. 2013. V. 154. № 1. P. 61–74.
93. Li Y., Laterra J. // *Cancer Res*. 2012. V. 72. № 3. P. 576–580.
94. Vermeulen L., Snippert H.J. // *Nat. Rev. Cancer*. 2014. V. 14. № 7. P. 468–480.
95. Biddle A., Liang X., Gammon L., Fazil B., Harper L.J., Emich H., Costea D.E., Mackenzie I.C. // *Cancer Res*. 2011. V. 71. № 15. P. 5317–5326.
96. Marjanovic N.D., Weinberg R.A., Chaffer C.L. // *Cell Cycle*. 2013. V. 12. № 17. P. 2713–2714.
97. Xie G., Ji A., Yuan Q., Jin Z., Yuan Y., Ren C., Guo Z., Yao Q., Yang K., Lin X., et al. // *Br. J. Cancer*. 2014. V. 110. № 10. P. 2514–2523.
98. Varga J., De Oliveira T., Greten F.R. // *FEBS Lett*. 2014. V. 588. № 15. P. 2422–2427.
99. Takebe N., Miele L., Harris P.J., Jeong W., Bando H., Kahn M., Yang S.X., Ivy S.P. // *Nat. Rev. Clin. Oncol*. 2015. V. 12. № 8. P. 445–464.
100. Maccalli C., De Maria R. // *Cancer Immunol. Immunother*. 2015. V. 64. № 1. P. 91–97.
101. Liu H., Lv L., Yang K. // *Am. J. Cancer Res*. 2015. V. 5. № 3. P. 880–893.
102. Kaiser J. // *Science*. 2015. V. 347. № 6219. P. 226–229.
103. Jung Y., Kim W.Y. // *Arch. Pharm. Res*. 2015. V. 38. № 3. P. 414–422.
104. Jeter C.R., Yang T., Wang J., Chao H.P., Tang D.G. // *Stem Cells*. 2015. V. 33. № 8. P. 2381–2390.
105. Ajani J.A., Song S., Hochster H.S., Steinberg I.B. // *Semin. Oncol*. 2015. V. 42 Suppl 1. P. S3–17.
106. Alekseenko I.V., Snezhkov E.V., Chernov I.P., Pleshkan V.V., Potapov V.K., Sass A.V., Monastyrskaya G.S., Koptantzev E.P., Vinogradova T.V., Khramtsov Y.V., et al. // *J. Transl. Med*. 2015. V. 13. № 1. P. 78.

EDITORIAL NOTE

Oncological diseases kill millions of people each year. Despite the significant progress made in recent years, cancer remains far from defeated. What's more, that is unlikely to happen in the next ten years. The more we learn about the molecular and cellular basis of the malignant transformation, the more intractable the problems, whose solution is required for successful treatment, seem.

From the prevailing point of view, cancer is a set of diseases un-

derlain by disturbances in the basic processes of cell metabolism. A baffling variety of these processes determines both the origin and course of the malignant transformation and the set of approaches to treatment. Investigation of these processes in the last decades has been littered with victories and defeats, and it is arduous to say which will prevail.

Today, there is a wide variety of sometimes opposing viewpoints on cancer. In our opinion, a discussion of this topic should be of

interest to everyone. In this issue, we are opening the discussion with an article by T.V. Vinogradova, I.P. Chernov, G.S. Monastyrskaya, L.G. Kondratyeva, and E.D. Sverdlov "Cancer stem cells: plasticity versus therapy." We emphasize once again the disputable nature of the publication, but it reflects the authors' opinion. At the same time, we believe that this publication will give rise to further discussion of the issue.

We are awaiting your reactions!

Induced Pluripotent Stem Cells of *Microtus levis* x *Microtus arvalis* Vole Hybrids: Conditions Necessary for Their Generation and Self-Renewal

E. V. Grigor'eva^{1,2,3}, A. I. Shevchenko^{1,2,3}, S. P. Medvedev^{1,2,3,4}, N. A. Mazurok^{1,2,3},
A. I. Zhelezova¹, S. M. Zakian^{1,2,3,4*}

¹Federal Research Center, Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Lavrentiev Ave., 10, 630090, Novosibirsk, Russia

²Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Lavrentiev Ave., 8, 630090, Novosibirsk, Russia

³State Research Institute of Circulation Pathology, Ministry of Healthcare of the, Rechkunovskaya Str., 15, 630055, Novosibirsk, Russia

⁴Novosibirsk State University, Pirogova Str., 2, 630090, Novosibirsk, Russia

*E-mail: zakian@bionet.nsc.ru

Copyright © 2015 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 Every year, the list of mammalian species for which cultures of pluripotent stem cells (PSCs) are generated increases. PSCs are a unique tool for extending the limits of experimental studies and modeling different biological processes. In this work, induced pluripotent stem cells (iPSCs) from the hybrids of common voles *Microtus levis* and *Microtus arvalis*, which are used as model objects to study genome organization on the molecular-genetic level and the mechanisms of X-chromosome inactivation, have been generated. Vole iPSCs were isolated and cultured in a medium containing cytokine LIF, basic fibroblast growth factor (bFGF), ascorbic acid, and fetal bovine serum. Undifferentiated state of vole iPSCs is maintained by activation of their endogenous pluripotency genes – *Nanog*, *Oct4*, *Sox2*, *Sall4*, and *Esrrb*. The cells were able to maintain undifferentiated state for at least 28 passages without change in their morphology and give rise to three germ layers (ectoderm, mesoderm and endoderm) upon differentiation.

KEYWORDS reprogramming, induced pluripotent stem cells, common voles.

ABBREVIATIONS iPSCs – induced pluripotent stem cells; mESCs – mouse embryonic stem cells; PSCs – pluripotent stem cells; ESCs – embryonic stem cells; AP – alkaline phosphatase; OSKM – transcription factors *Oct4*, *Sox2*, *Klf4* and *c-Myc*; RT-PCR – reverse transcription followed by polymerase chain reaction.

INTRODUCTION

Currently, in addition to the traditional method of isolating pluripotent stem cells (PSCs) from early mammalian embryos, it has also become possible to induce pluripotency by reprogramming different types of terminally differentiated somatic cells [1-4]. Reprogramming of somatic cells to a pluripotent state allows one to obtain an unlimited amount of autologous iPSCs of any mammal, including humans. Reprogramming technology holds tremendous prospects not only for a personalized approach to the treatment of various diseases, but also serves as a tool for genetic modeling of many biological processes, including the study of early embryonic development, the signaling pathways and factors involved in pluripotency maintenance, and the triggering of differentiation.

Reprogramming of differentiated cells of various mammalian species to a pluripotent state is possible due to the overexpression of four transcription factors. These factors are *Oct4*, *Sox2*, *Klf4*, and *c-Myc* (OSKM). The genes are evolutionarily conserved in mammals [5, 6], and their products have a substantially overlapping set of key target genes, and thus genetic constructs expressing human or mouse OSKM can be frequently applied for cell reprogramming in different mammalian species [7-10]. To date, iPSCs of mouse, human, macaque, rat, dog, many farm animals and some other mammalian species, including prairie vole *Microtus ochrogaster*, have been obtained. The conditions for induction and maintenance of pluripotency vary among the species [11-19]. This is partly due to the species-specificity of the signaling cascades involved

in the activation and maintenance of an undifferentiated state *in vitro*, and due to various requirements for the composition of the culture medium, e.g. the presence or absence of bovine serum in the medium. Induction and maintenance of pluripotency are facilitated by the presence in the medium of inhibitors of various signaling pathways, inhibitors of histone deacetylases, histone and DNA methyltransferases, as well as anti-oxidants.

In the current study, the conditions for obtaining and maintaining cells with induced pluripotency of the common vole species of genus *Microtus* have been selected. Four closely related cryptic species, *M. arvalis*, *M. levis*, *M. transcaspicus*, and *M. kirgisorum*, comprising a group of common voles are the subject of a molecular genetic study of genome organization and the mechanisms of X chromosome inactivation [20–28]. The genes involved in the establishment and maintenance of pluripotency in these species have been studied, and their conservation has been demonstrated, including the expression pattern [5, 6, 29]. The existence of pluripotent cells of common voles might become an appropriate tool for molecular-genetic studies of these species.

Previously, we have undertaken numerous attempts to obtain PSCs of common vole species of the genus *Microtus* from early pre-implantation blastocysts and germinal cells [30]. Multipotent cell lines of the pre-implantation embryo, such as trophoblast stem cells and extraembryonic endoderm cells, have been derived [25, 31–33]. However, PSCs of common voles have not been obtained yet. In the experiments on somatic cell reprogramming presented in this work, we managed for the first time to determine the conditions that allow one to obtain and maintain PSCs of common voles.

MATERIAL AND METHODS

Cell cultures and media

Skin fibroblasts and brain cells were isolated from *M. levis* × *M. arvalis* hybrid embryos on day 19 (E19) of embryonic development and then cultured. The fibroblasts were grown in DMEM/F12 (F12, Nutrient Mixture/Dulbecco's Modified Eagle Medium; Gibco) 1:1 supplemented with 10% fetal bovine serum (FBS; Gibco), 1× Non-Essential Amino Acids (NEAA; Gibco), 1× Pen Strep (100 u/ml penicillin, 100 µg/ml streptomycin; Gibco), 1× GlutaMAX (Gibco). Brain cells were grown in DMEM/F12 (1:1) with 10% FBS, 1× NEAA, 1× Pen Strep, 1× GlutaMAX for the first 3 days, then transferred to Schneider's medium: DMEM/F12 (1:1), 10 ng/ml bFGF (StemCell), 10 ng/ml EGF (Sigma), 2 µg/ml heparin (Sigma), 1× N2 Supplement (Gibco), 1× Pen Strep, 1× GlutaMAX.

In the first experiment, two media were used for the induction of pluripotency in *M. levis* × *M. arvalis* hybrid cells, which were previously used to obtain *M. ochrogaster* iPSCs [16]. The media were prepared from Advanced DMEM/F12 (1:1) (Gibco) containing 15% KSR (Knockout Serum Replacement; Gibco), 1× NEAA, 0.1 mM 2-mercaptoethanol, 1× Pen Strep, and 1× GlutaMAX. The first medium was supplemented with three inhibitors of signaling pathways (3iR conditions): CHIR99021 (3 µM, StemRD), PD0325901 (1 µM, StemRD), A83-01 (0.5 µM, Stemgent), and ROCK inhibitor (Y-27632, 10 µM, StemRD); the second medium did not contain any inhibitor. Both media contained 1,000 u/ml mouse LIF (mLIF, Invitrogen) and 2 µg/ml doxycycline (Sigma), which was added only at the initial stages of cultivation. Mouse embryonic fibroblasts mitotically inactivated by a mitomycin C solution

Table 1. List of the antibodies used in the study

Antibodies	Source	Catalogue number	Type	Working dilutions
SSEA1	Santa Cruz Biotechnology	sc-21702	mouse monoclonal IgM	1 : 50
OCT4	Santa Cruz Biotechnology	sc-5279	mouse monoclonal IgG2b	1 : 100
SOX2	Santa Cruz Biotechnology	sc-20088	rabbit polyclonal IgG	1 : 100
KLF4	Abcam	ab104846	mouse monoclonal IgG1	1 : 200
β-III-tubulin	Covance	MMS-435P-100	mouse monoclonal IgG2a	1 : 1000
Nestin	Abcam	ab6142	mouse monoclonal IgG1	1 : 400
α-SMA	Dako	M0851	mouse monoclonal IgG2a	1 : 100
CD90	Millipore	MAB1406	mouse monoclonal IgG2b	1 : 100
KRT18	Millipore	MAB3234	mouse monoclonal IgG	1 : 200
SOX17	Millipore	09-038	rabbit polyclonal IgG	1 : 100

(10 mg/ml, Sigma) for 2 h were used as a feeder substrate for the generation and cultivation of vole iPSCs. Transfer of individual primary colonies at the first passage was performed with a glass capillary, and further using TrypLE (Gibco).

In further experiments on reprogramming in the medium based on Advanced DMEM/F12 (1:1) supplemented with 1× NEAA, 0.1 mM 2-mercaptoethanol, 1× Pen Strep, and 1× GlutaMAX, we varied the composition of 3iR, mLIF, PKCi (Gö6983, Tocris, 5 µM), bFGF, ascorbic acid, KSR, and FBS. iPSCs of hybrid *M. levis* × *M. arvalis* were derived in the medium, which included: Advanced DMEM/F12 (1 : 1), 7% KSR (Gibco), 7% FBS (Gibco), 0.1 mM 2-mercaptoethanol, 1× Pen Strep, 1× GlutaMAX, 1000 u/ml mLIF (Invitrogen), 10 ng/ml bFGF (StemCell), 50 mkg/ml ascorbic acid (Wako). The medium initially contained 2 µg/ml doxycycline (Sigma), which was abrogated on day 14 of reprogramming. Vole iPSCs were frozen in 90% FBS and 10% DMSO.

Plasmid constructs, generation of lentiviruses, transduction scheme

Three plasmids were used in the experiments: 1) TetO-FUW-OSKM (Addgene Plasmid 20321) encoding mouse reprogramming factors OSKM; 2) FUDeltaGW-rtTA (Addgene Plasmid 19780) carrying tetracycline transactivator cDNA necessary for regulation of the transcriptional activity of the construct with OSKM by supplementing the growth media with doxycycline; 3) pGpur expressing a green fluorescent protein (EGFP) gene for monitoring transduction efficiency. Lentiviruses were obtained by cotransfection of the plasmids into 293FT cells with pxPAX2 (Addgene Plasmid 12260) and pMD2.G (Addgene Plasmid 12259) plasmids encoding the proteins required for packaging of viral

particles [34, 35]. Human embryo kidney cells, 293FT, were seeded at a density of 6×10^6 into T75 culture flasks and incubated overnight. Transfection was performed by the calcium phosphate method [36]. The medium with viral particles was harvested and filtered (0.45 µm; Millipore) 48, 72, and 96 h after transfection. Viruses were concentrated using an ultracentrifuge (Beckman Coulter, Optima XE-90 Ultracentrifuge) for 90 min at 70,000 g. A viral pellet was dissolved in 200 µl of F12/DMEM and kept in aliquots at -70°C.

The fibroblasts and cells isolated from the brain were seeded in 12-well plates at a density of 50×10^3 to 75×10^3 cells per well, respectively, 24 h before transduction. The cells were transduced for 4-6 passages. On the day of transduction, the medium with lentiviruses obtained using TetO-FUW-OSKM, FUDeltaGW-rtTA, or pGpur plasmids was added for 14-16 h, with a titer of about 3×10^7 infectious units per 1 ml (MOI 5-7.5) for each of the lentiviruses and 4 µg/ml polybrene (Hexadimethrine bromide, Sigma). After 4 days, cells transduced by lentiviruses with reprogramming factors and tetracycline transactivator were passaged using TrypLE at a dilution of 1:10 to 1:20 (depending on cell density in a well) on mitotically inactivated mouse embryonic fibroblasts in culture media varying in composition. For determination of transduction efficiency, cells transduced with lentiviruses containing pGpur were used, and the assessment of the percentage of green cells was performed using a fluorescent microscope and/or flow cytometry 4 days after transduction.

Histochemical detection of endogenous alkaline phosphatase activity

The activity of endogenous alkaline phosphatase (AP) was measured histochemically according to [37]. Cells

Table 2. Primers and PCR conditions

Gene	Nucleotide sequence	Primers	Mg ²⁺ concentration, mM	Annealing temperature, °C
<i>β-actin</i>	gacggggtcaccacactgt gag-tacttgcgctcaggaggag	β-actin-1 β-actin-2	3	60
<i>Nanog</i>	agtgtcttaaggacgcagaa atctcctaattgccaatacc	Nanog_QF1 Nanog_QR1	3	60
<i>Oct4</i>	ccaagctgctgaagcagaaga tt-gaatgcatgggagagccag	OCT4-2F OCT-5R	4	53
<i>Sox2</i>	tccatgaccagctcgcagacctac cctcccaattccctgtttctct	Sox2F Sox2R2	3	60
<i>Sall4</i>	tcaccaacgcccgtcatgttacagc ggtgggctgtgctcggataaatgt	Sall4F Sall4R	2	60
<i>Errβ</i>	agctgcggctccttcatcaag cttg-tacttctggcgctcctc	ERRB1F ERRB4R	1.5	63

were fixed by air-drying and incubated in a dye solution: 100 μ M Tris-HCl pH 9.0; 100 μ M NaCl; 5 μ M $MgCl_2$; 0.4 μ g/ml naphthol phosphate (Sigma); 1 μ g/ml Fast Violet B Salt (Sigma) for 15–20 min in the dark at room temperature.

Immunofluorescence assay

Cells were fixed with 4% formaldehyde for 10–15 min at room temperature, washed with PBS, and incubated with 0.1% Triton X-100 and 2.5% FBS (or BSA) dissolved in PBS for 30 min at room temperature. Immunoprecipitation was performed by using primary antibodies overnight at 4°C. A list of the primary antibodies used in this study is provided in *Table 1*. Localization of primary antibodies was visualized using secondary anti-rabbit or -mouse antibodies conjugated to the fluorescent dyes Alexa 488 and Alexa 568 (Life Technologies). Nuclei were stained with DAPI (Vector Laboratories).

In vitro differentiation of cell lines in embryoid bodies

Colonies of undifferentiated cells were mechanically transferred onto Petri dishes coated with a thin layer of 1% agarose in DMEM/F12 (1:1) with 10% FBS, 1 \times Pen Strep and 1 \times GlutaMAX. The formed embryoid bodies were cultured for 7 days in suspension and transferred onto Petri dishes coated with 0.1% gelatin for attachment. Attached embryoid bodies were grown for 14–20 days, followed by analysis of differentiated cells by immunofluorescence assay or disaggregation using TrypLE with further RNA isolation and RT-PCR.

RNA isolation, RT-PCR

RNA was isolated with Trizol Reagent (Ambion). Samples were treated with DNase I (Turbo DNA-free, Ambion) in order to prevent DNA contamination; cDNA was synthesized using Super-ScriptIII (Invitrogen). Primer sequences and reaction conditions for RT-PCR are shown in *Table 2*. For the transcription analysis of the exogenous lentiviral cDNA of OSKM in mouse cells, we used the primers and RT-PCR conditions listed in [34]. A negative control reaction (RT-), with the reaction solution containing all the components necessary for cDNA synthesis, except for reverse transcriptase, was performed for every primer pair.

Bisulfite sequencing of vole Oct4 gene promoter DNA

Bisulfite modification and purification of genomic DNA (500 ng) were conducted using EZ DNA Methylation – Direct Kit (Zymo Research). Modified DNA was further used for PCR with primers:

Oct4_Reg2_f2 (5'-TAAGAGGATGGGGGGG-TAGTGAAAG-3')

Oct4_Reg2_r2 (5'-GAAATCTAAAACCAAATATC-CAACCATAAA-3').

The obtained PCR products were cloned using pGEM-T Easy Vector System I (Promega). A total of 10 randomly selected plasmid clones of each DNA sample were sequenced using a universal M13 Reverse primer. Nucleotide sequences were analyzed with the QUMA software (Quantification tool for Methylation Analysis, <http://quma.cdb.riken.jp/>, [38]).

Cytogenetic analysis

Cells were fixed according to standard protocols [39] with several modifications: time of incubation with ethidium bromide (0.05%) – 3 h; with colcemid (10 μ g/ml, Gibco) – 2 h; and hypotonic incubation – 15–20 min.

Prior to staining, samples were incubated overnight at 50°C. A Hoechst 33258 (Sigma) solution at a concentration of 0.05 μ g/ml was prepared using Hank's balanced solution (HBSS, Gibco). Metaphase spreads were stained in the Hoechst 33258 solution for 10 min, and then the samples were rinsed with water and placed in a acetate buffer (pH 5.5). The metaphase spreads were visualized using the microscope Ni-E (Nikon), Lucia software.

Cell analysis by flow cytometry

The number of EGFP- and SSEA1-positive cells was estimated using the BD FACS Aria and BD FACSCanto II systems with the BD FACS Diva software. Surface antigen SSEA1 was determined using antibodies (sc-21702, Santa Cruz Biotechnology) according to the manufacturer's protocol.

RESULTS

Obtaining doxycycline-dependant iPSC-like vole cells and their characterization

Common vole iPSCs were obtained using lentiviruses expressing cDNA of four key mouse reprogramming factors OSKM under a doxycycline-regulated promoter. Transcription from this promoter can be activated by adding an antibiotic in the culture medium; however, the cells should be previously transduced with lentiviruses expressing cDNA of a tetracycline-dependant transactivator. This system, where mouse OSKM expression is modulated by the addition of doxycycline, was successfully utilized previously for obtaining human and mouse iPSCs [34]. Mice and voles belong to the same family of rodents (Muridae), order Rodentia, and exhibit high similarity of OSKM genes [40]. The advantage of this system, in our opinion, is the fact that all four reprogramming factors are delivered into a cell by the same viral particle.

Cells isolated from hybrid embryos of *M. levis* \times *M. arvalis*, which are used as a model object for study-

Table 3. Conditions tested in experiments on obtaining iPSC lines of interspecies *M. levis* × *M. arvalis* vole hybrids

Cell type (experiment №)	Cell number, 10 ³	Transfection efficiency, %	Culturing conditions	Number of primary colonies of AP ⁺ cells	Number of iPSC lines
Brain cells (I exp.)	37.5	75	KSR, mLIF	85	–
	37.5		KSR, mLIF, 3iR	200	–
Fibroblasts (I exp.)	25	86.5	KSR, mLIF	–	–
	25		KSR, mLIF, 3iR	–	–
Fibroblasts (II exp.)	12.5	32.9	KSR, mLIF, bFGF, AA	–	–
	12.5		KSR, mLIF, bFGF, AA, 3iR	–	–
	12.5		KSR, FBS, mLIF, bFGF, AA	70	2
	12.5		KSR, FBS, mLIF, bFGF, AA, 3iR	100	–

Note. AA – ascorbic acid.

ing the phenomenon of non-random inactivation of the X chromosome [27, 28], were transfected with lentiviral constructs. Two types of hybrid cells were used: skin fibroblasts and cells isolated from the brain. Transduction was performed according to the scheme depicted in *Fig. 1A*. Transduction efficiency was assessed using lentiviruses expressing a green fluorescent protein by flow cytometry, and fluorescent microscopy was 86.5% in fibroblasts (*Fig. 1B*) and 75% in brain cells (*Fig. 1C*).

Pluripotency in common vole cells was induced in two types of media that had previously been used for obtaining iPSCs from the fibroblasts of prairie vole *M. ochrogaster* [16]. The first type of media included Advanced DMEM/F12 with 15% KSR and mLIF, the second medium contained the same components, but it also included an inhibitor cocktail called 3iR. The cocktail 3iR contains inhibitors of kinases MEK/ERK and GSK3b (PD0325901 and CHIR99021, respectively), an antagonist of the IL-type receptor TGF-β (A83-01) used for obtaining mouse and rat PSCs and maintaining them in undifferentiated state, and a ROCK inhibitor that enhances the survival of single cells in culture [41–46].

Four days after transduction, *M. levis* × *M. arvalis* hybrid cells were transferred on mitotically inactivated mouse embryonic fibroblasts and the medium with or without 3iR for obtaining iPSCs was added. The first morphological changes appeared 24 h after cell transfer, and an increase in cell proliferation was noted. After 5–8 days, the transduced cells that were isolated from the brain began to merge into primary homogeneous colonies (*Fig. 1D*), whereas the transduced fibroblasts did not give rise to such PSC-like primary colonies (*Table 3*).

Primary colonies obtained during brain cell reprogramming were transferred into individual wells by disaggregation with a glass capillary on day 10–13 after the start of the experiment. Cell morphology remained the same as in the case of mESCs: three-dimensional dense colonies with tight intercellular contacts (*Fig. 1E, F*). The cells in these colonies had a high nuclear/cytoplasmic ratio, which is a characteristic of ESCs. After transfer of primary colonies, most cells underwent differentiation on the first passage. However, about 40% of the cells retained an ESC-like morphology. The difference in the level of endogenous AP was detected in selected colonies by histochemical staining (*Fig. 1G*). Cell lines with no detectable AP activity were excluded from further analysis. Thus, a total of 11 cell lines obtained in the presence of inhibitors (viBr; vole inhibitor Brain) and 10 lines obtained in their absence (vBr) were selected based on the results of an AP activity analysis at early passages. It should be noted that most of the cells did not express one of the major early markers of pluripotency specific to mouse PSCs, surface antigen SSEA1. The iPSC-like cells underwent rapid proliferation in the medium containing doxycycline without altering the morphology for approximately 40 passages (more than 120 days). The cells were passaged every 2–3 days at a dilution of 1:8 to 1:10: they resisted repeated freezing in liquid nitrogen with further thawing without changing their phenotypic characteristics.

We repeatedly tried to culture the cell lines in the absence of doxycycline, abolishing the expression of transgenic OSKM, but this led to the flattening of the plump colonies of iPSC-like cells and differentiation by the second day of growth.

Thus, culture conditions that allowed to successfully obtain *M. ochrogaster* iPSCs failed to induce pluripotency in *M. levis* × *M. arvalis* hybrid cells. When reprogramming in these media, the fibroblasts of common vole hybrids did not form even primary colonies and the cells isolated from the brain that nonetheless demonstrated primary colony formation, but yet did not show activation of self-renewal mechanisms, which would have enabled them to maintain an iPSC-like morphology in the absence of OSKM expression.

Further, we focused on the selection of culture medium components which would allow the induction and maintenance of pluripotency in common vole cells *in vitro*. In order to achieve this, we studied various combinations of such medium components as mLIF, PKCi, bFGF, and ascorbic acid, which are used for obtaining and maintaining pluripotent cell cultures of other mammalian species [7, 9, 13, 47-55]. Medium components were tested in two iPSC-like doxycycline-dependent lines, one of which had been obtained and grown in a medium supplemented with inhibitors (viBr3 line), and a medium without inhibitors (vBr3).

A peculiar effect was observed upon the addition of substance G66983, a protein kinase C inhibitor (PKCi), to the medium, which is able to trigger and maintain the self-renewal capability of rodent pluripotent cells without activating LIF/STAT3- or suppressing ERK/GSK3-signaling pathways [48, 50]. The appearance of 3-5% SSEA1-positive cells (Fig. 1H) was observed on the 9th day of culturing in both iPSCs-like cell lines in the presence of PKCi, and the percent of these cells increased to 80-90% by day 19 (Fig. 1I). However, despite the expression of SSEA1, one of the markers in mouse, rat and *M. ochrogaster* pluripotent cells cultured *in vitro* [16, 56, 57], iPSC-like cells of *M. levis* × *M. arvalis* hybrids still retained the ability to differentiate in the absence of doxycycline. A more significant for the reprogramming of common vole cells result has been obtained using a medium containing all of the three components: mLIF, bFGF, and ascorbic acid, which for the first time allowed us to grow iPSC-like cultures of *M. levis* × *M. arvalis* hybrids for six passages in the absence of doxycycline without visible changes in morphology. Activation of endogenous pluripotency genes *Oct4* and *Nanog* expression has been noted in cell cultures (Fig. 1J). No similar effect has been shown upon addition of mLIF, bFGF, or ascorbic acid alone or in pairwise combinations to the culture medium for viBr3 and vBr3. However, since viBr3 and vBr3, which were cultured without doxycycline but in the presence of mLIF, bFGF, and ascorbic acid, showed differentiation after the sixth passage, we decided to perform the second experiment on reprogramming common vole cells, applying from the first stages the same condi-

tions that allowed us to maintain for some period *M. levis* × *M. arvalis* hybrid iPSCs in the medium without doxycycline.

Successful derivation of doxycycline-independent vole iPSCs

New reprogramming of *M. levis* × *M. arvalis* hybrids was conducted according to the previously described scheme (Fig. 1A) using embryonic skin fibroblasts. The efficiency of transduction was 32.9% (Fig. 2A). Taking into account the experience gained in the previous experiment, fibroblasts were transferred to the medium for reprogramming containing mLIF, bFGF, and ascorbic acid after transduction with lentiviruses expressing OSKM and a tetracycline-dependent transactivator. As in the previous experiment, we used a medium with or without the 3iR cocktail. Furthermore, we varied the content of KSR and FBS in the media: some media contained 15% KSR as described above, whereas others included a mixture of 7% FBS and 7% KSR.

Primary colonies were observed in fibroblasts already by the third day of culturing (on the 7th day of transduction) in reprogramming media containing a mixture of FBS and KSR. The formation of primary colonies was not shown for the medium containing 15% KSR (Table 3). By day 10-14 from the start of transduction, six primary colonies from the plates with 3iR and eight colonies from the plates without 3iR had been transferred from the cultures reprogrammed in the presence of KSR and FBS. Doxycycline was canceled 14 days after the start of reprogramming, i.e. on the first passage after seeding primary colonies in individual wells. As a result of culturing without doxycycline, most of the colonies underwent differentiation or were excluded after the first analysis of AP activity. Thus, two clones expressing AP and growing in the medium without inhibitors, 14vf2 and 14vf7, were selected. Two cell lines obtained in this experiment were cultured in the medium containing 7% FBS, 7% KSR, mLIF, bFGF, and ascorbic acid without changes in morphology for at least 28 passages (more than 4 months) in the absence of doxycycline. Removal of any of the medium components, namely mLIF, bFGF, or ascorbic acid, and a reduction in the FBS concentration lower than 7% led to gradual differentiation of the obtained iPSC lines.

Properties of doxycycline-independent vole cell lines

The obtained cell lines grow as flatten colonies with tight intercellular contacts and clear colony edge resembling human PSCs (Fig. 2B). Cell colonies grew pushing out the feeder cells and attached to the plastic-like human ESCs/iPSCs, but not above the feeder as mESCs. The proliferation intensity of the common

vole iPSC lines was comparable to that of human PSCs. It was found that supplementation of the medium with a ROCK inhibitor significantly enhances cell survival both during mechanical transfer (by capillary) or upon using TrypLE.

More than 70% of the cells in both iPSC lines exhibited the expected number of autosomes, equal to 50; X chromosome from *M. levis* and Y chromosome from *M. arvalis* (Fig. 2C).

A histochemical analysis demonstrated AP activity in undifferentiated cell lines (Fig. 2D), which was not detected after differentiation. Unlike in doxycycline-dependent lines, transcription of the introduced construct carrying reprogramming factors was not detected in stable iPSC lines of *M. levis* × *M. arvalis* hybrids (Fig. 2E). Moreover, both cell lines exhibited demethylation of CpG dinucleotides in the promoter of vole *Oct4*, which is indicative of its reactivation (Fig. 3A).

An immunofluorescence analysis showed that in undifferentiated state in the early passages the obtained cell lines expressed one of the key markers of pluripotency – surface antigen SSEA1 (Fig. 2F), which is specific to mESCs/iPSCs. Figure 2F clearly shows that staining both for AP and SSEA1 is located in the most bulky, tight undifferentiated part of the colony, while the upper left edge of the colony, which contains spread differentiated cells, is not stained either for AP or for SSEA1. However, after a series of passages we failed to detect SSEA1-expressing cells. In *M. levis* × *M. arvalis* hybrid iPSCs, we also detected the expression of the key genes of pluripotent state *Oct4* and *Sox2*, which remained stable during cell growth (Fig. 2G).

In both lines of *M. levis* × *M. arvalis* hybrid iPSCs the method of RT-PCR revealed expression of the endogenous genes *Nanog*, *Oct4*, *Sox2*, *Sall4*, and *Esrrb*, which are essential for maintaining the pluripotent state in mammalian cells. The initial line of embryonic skin cells lacked transcripts of these genes (Fig. 3B). Thus, taking into account the PSC-like morphology of the cell colony, AP expression, demethylation of the *Oct4* promoter, expression of *Nanog*, *Oct4*, *Sox2*, *Sall4*, and *Esrrb*, and unlimited proliferation in the absence of doxycycline in the culture without change in cell morphology, we can claim that we managed to obtain pluripotent iPSCs of *M. levis* × *M. arvalis* interspecific hybrids.

In order to study the ability of the cell lines to differentiate *in vitro*, we obtained embryoid bodies, which had already formed in the suspension culture by day 2 (Fig. 4A). Analysis of differentiated derivatives by immunofluorescent staining with antibodies to markers of specific cell types revealed derivatives in all three primary germ layers: ectoderm (Nestin, β -III-tubulin),

endoderm (SOX17, KRT18), and mesoderm (α -SMA, CD90) (Fig. 4B).

Investigating the spontaneous differentiation of *M. levis* × *M. arvalis* hybrid iPSCs that have undergone successful reprogramming in culture using antibodies that detect only endogenous OCT4 and KLF4, we found that expression of KLF4 is absent in pluripotent cells, but it appears at the beginning of their differentiation, correlating with the loss of the transcription factor OCT4 (Fig. 4C).

DISCUSSION

We have attempted to obtain iPSCs of hybrids from crossing *M. levis* and *M. arvalis* species, which belong to a group of common voles, genus *Microtus*. The first experiments on obtaining common vole iPSCs were performed using the same culture media and factors that had previously allowed to induce pluripotency and obtain iPSCs of prairie vole *M. ochrogaster* [16]. Prairie vole iPSCs were obtained in the presence of a KSR substitute for fetal bovine serum, as well as LIF cytokine, which triggers a key signaling cascade in mouse and rat pluripotent cells [57-59]. Generation of *M. ochrogaster* iPSCs can be achieved in the presence of the inhibitors CHIR99021, PD0325901, and A83-01 in the medium. Nevertheless, the induction of pluripotency did not take place upon reprogramming of common vole cells using four OSKM factors and a medium containing LIF and KSR. After careful selection of conditions, hybrid *M. levis* × *M. arvalis* iPSCs were obtained in a medium containing mLIF, bFGF, ascorbic acid, and a mixture of 7% FBS and 7% KSR. It is interesting that in most cases could not we obtain even primary iPSC-like colonies during the reprogramming of *M. ochrogaster* cells in a medium containing bFGF and FBS.

Our experience shows that mouse OSKM, which is used for the induction of pluripotency in somatic cells of different types and in different species, can be effective in obtaining common vole iPSCs. However, the presence of LIF cytokine in the medium, an important factor in the formation and maintenance of undifferentiated state in ESCs and iPSCs in rodents (mouse, rat, as well as prairie vole) [16, 57-59], is not sufficient for the induction and maintenance of pluripotency in common vole cells *in vitro*. This result is consistent with the reports of unsuccessful attempts to obtain ESCs of common vole group species from the inner cell mass of blastocysts in the presence of mouse or *M. levis* LIF [31]. Taken together, these data allow one to suggest that the signaling pathway triggered by cytokine LIF, due to some species-specific features of the common vole, cannot independently provide pluripotency in *in vitro* conditions.

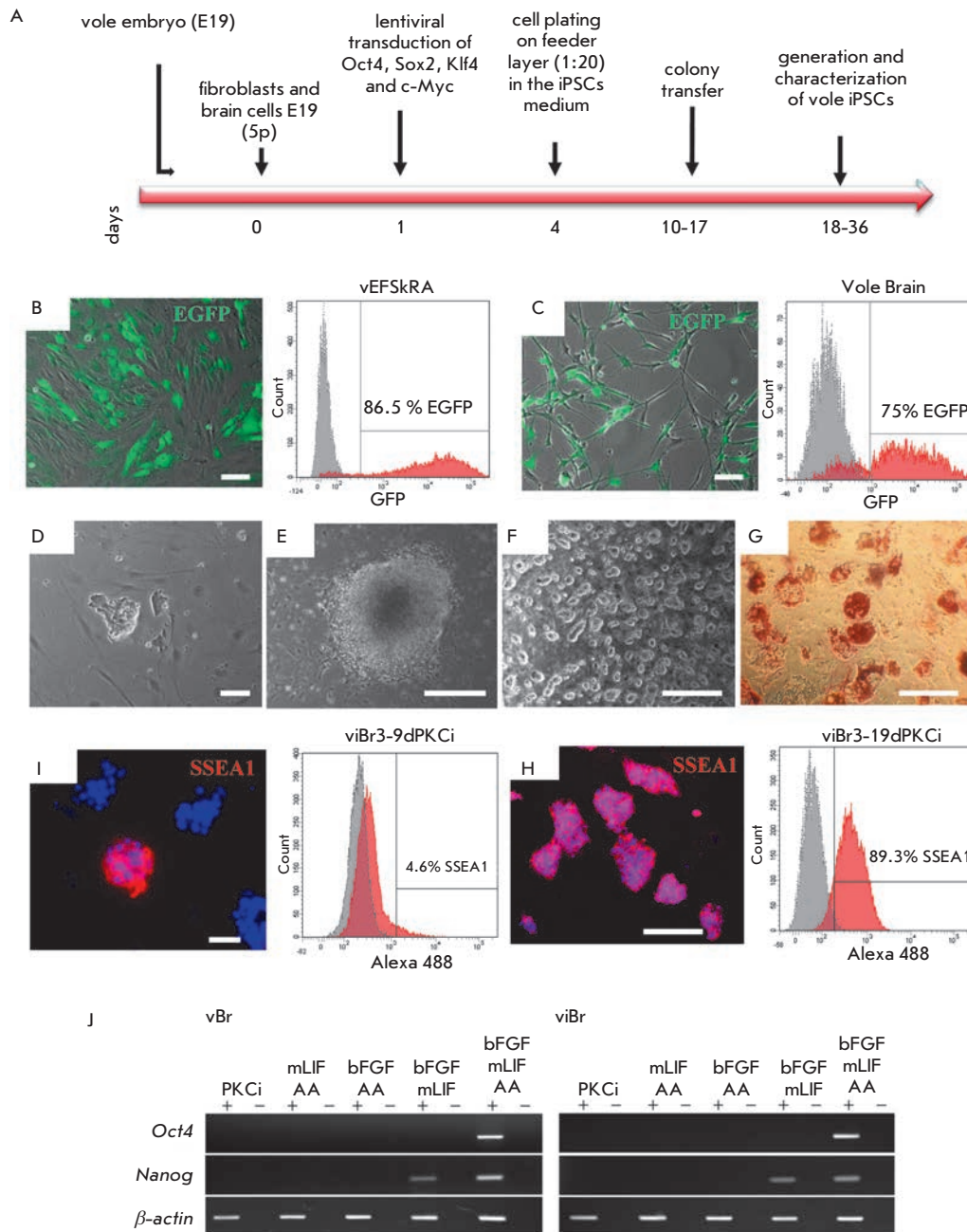


Fig. 1. Obtaining and characterizing doxycycline-dependent iPSC-like lines of *M. levis* × *M. arvalis* hybrid cells. **A** – scheme of the experiment on obtaining iPSCs of the common vole of genus *Microtus*. **B**, **C** – assessment of the efficiency of the transduction of the embryonic skin fibroblasts (vEFSkRA cell line) and brain cells (Vole Brain) of hybrid voles with lentivirus expressing EGFP (green signal) by fluorescence microscopy and flow cytometry. The percentage of GFP-positive cells among the fibroblasts and brain cells is 86.5% and 75%, respectively. **D** – primary colony morphology on the 8th day of the reprogramming of cells isolated from the brain. **E** – colony morphology on the 13th day of reprogramming of cells isolated from the brain. **F** – morphology of viBr3 cell line colonies grown in the medium supplemented with mLIF + 3iR, passage 3. **G** – histochemical assay of endogenous AP activity in the viBr3 cell line, passage 2. **H**, **I** – data of immunofluorescence analysis and flow cytometry on the presence of SSEA1-positive cells in the viBr3 line on days 9 and 19 of culturing in the presence of PKCi, respectively. Nuclei are stained with DAPI (blue signal). **J** – RT-PCR analysis of *Oct4* and *Nanog* expression in the vBr3 and viBr3 cell lines after culturing for 3 passages in the media supplemented with various components. AA – ascorbic acid. Scale bar **B–D**, **H** – 100 μm, **E–G**, **I** – 500 μm

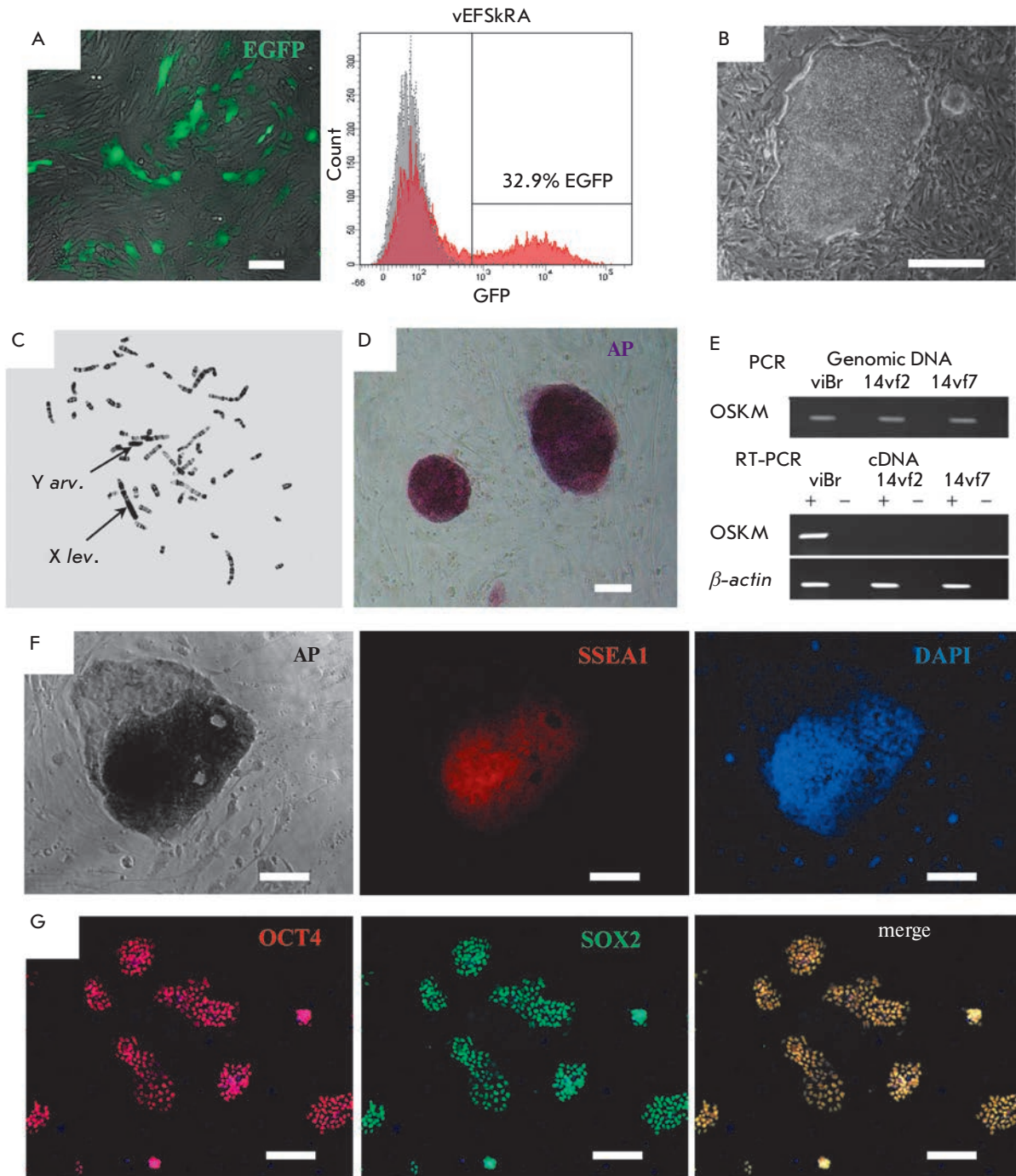


Fig. 2. Obtaining and characterizing iPSCs of common vole *M. levis* \times *M. arvalis* hybrids. **A** – efficiency of the transduction of vole embryonic fibroblasts (vEFSkRA) with a lentivirus expressing GFP (green signal), and assessment of the percentage of GFP-positive cells (32.9%) by fluorescence microscopy and flow cytometry. **B** – morphology of 14vf7 cell line colonies at passage 7. **C** – metaphase spread of 14vf7, passage 13. X lev. – X chromosome of *M. levis*, Y arv. – Y chromosome of *M. arvalis*. **D** – histochemical detection of endogenous AP activity, 14vf7 cell line, passage 6. **E** – RT-PCR analysis of the expression of the construct with exogenous factors of reprogramming (OSKM) in iPSC lines of common vole hybrids. **F** – immunofluorescence analysis of SSEA1 expression (red signal) and histochemical detection of AP activity, 14vf7 line, passage 4. Nuclei are stained with DAPI (blue signal). **G** – immunofluorescence analysis of pluripotency markers OCT4 (red signal) and SOX2 (green signal). Nuclei are stained with DAPI (blue signal). Scale bar **A**, **D**, **F**, **G** – 100 μ m, **B** – 500 μ m

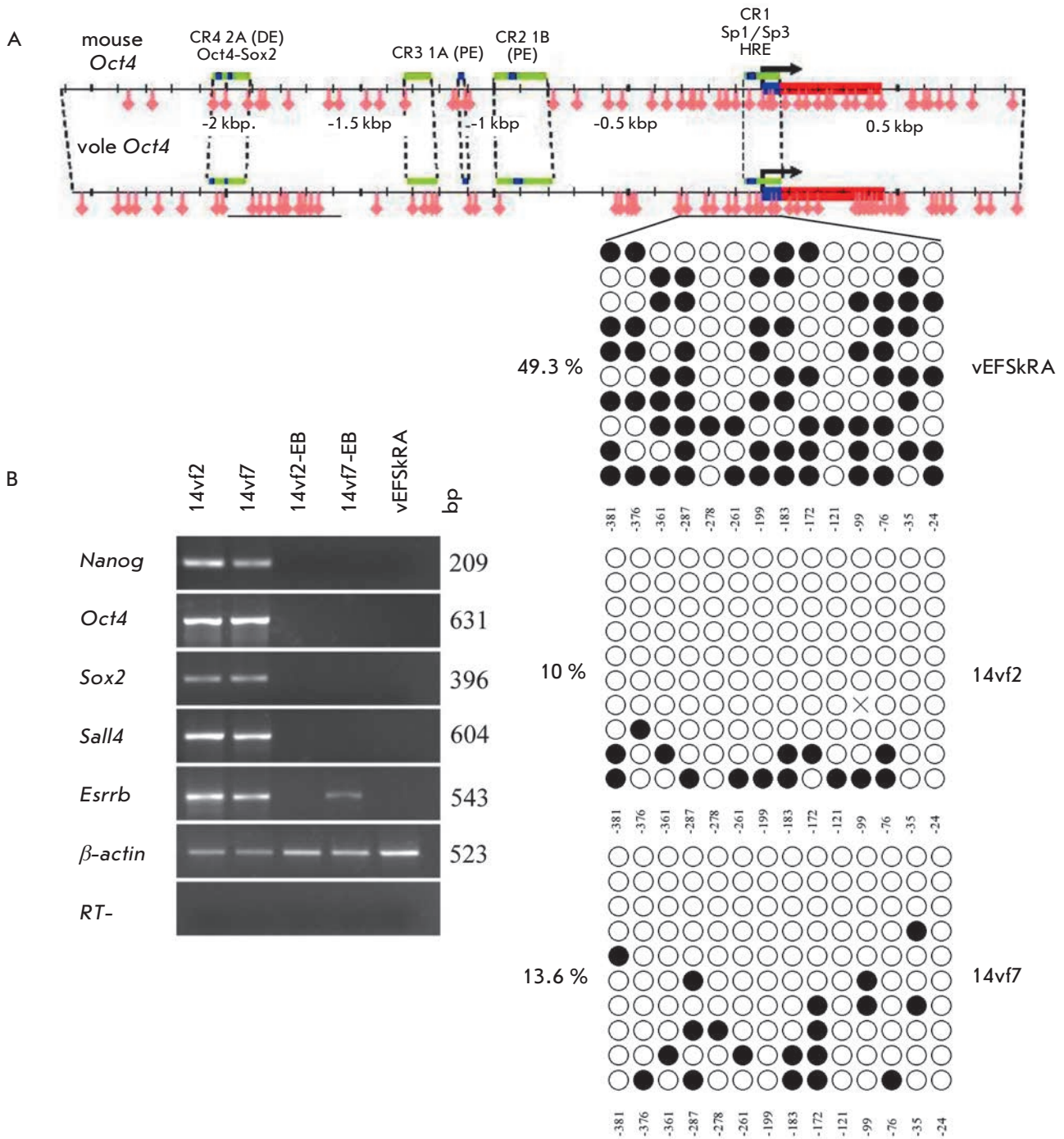


Fig. 3. Analysis of CpG dinucleotide methylation in the *Oct4* gene promoter and expression of genes specific for pluripotent cells in common vole iPSCs. **A** – comparison of CpG dinucleotide methylation of the *Oct4* promoter in the obtained iPSC lines (14vf2, 14vf7) and initial line (vEFSkRA) of embryonic skin fibroblasts. Schematic distribution of CpG dinucleotides in a mouse and a common vole *Oct4* promoter is presented at the top. Light and dark circles – unmethylated and methylated CpG dinucleotides, respectively. **B** – transcriptional activity of the genes responsible for the pluripotent state in the iPSC lines 14vf2 and 14vf7 and their differentiated derivatives (14vf2-EB and 14vf7-EB). Control – initial cell line vEFSkRA. (RT-) – negative control of reverse transcription reaction

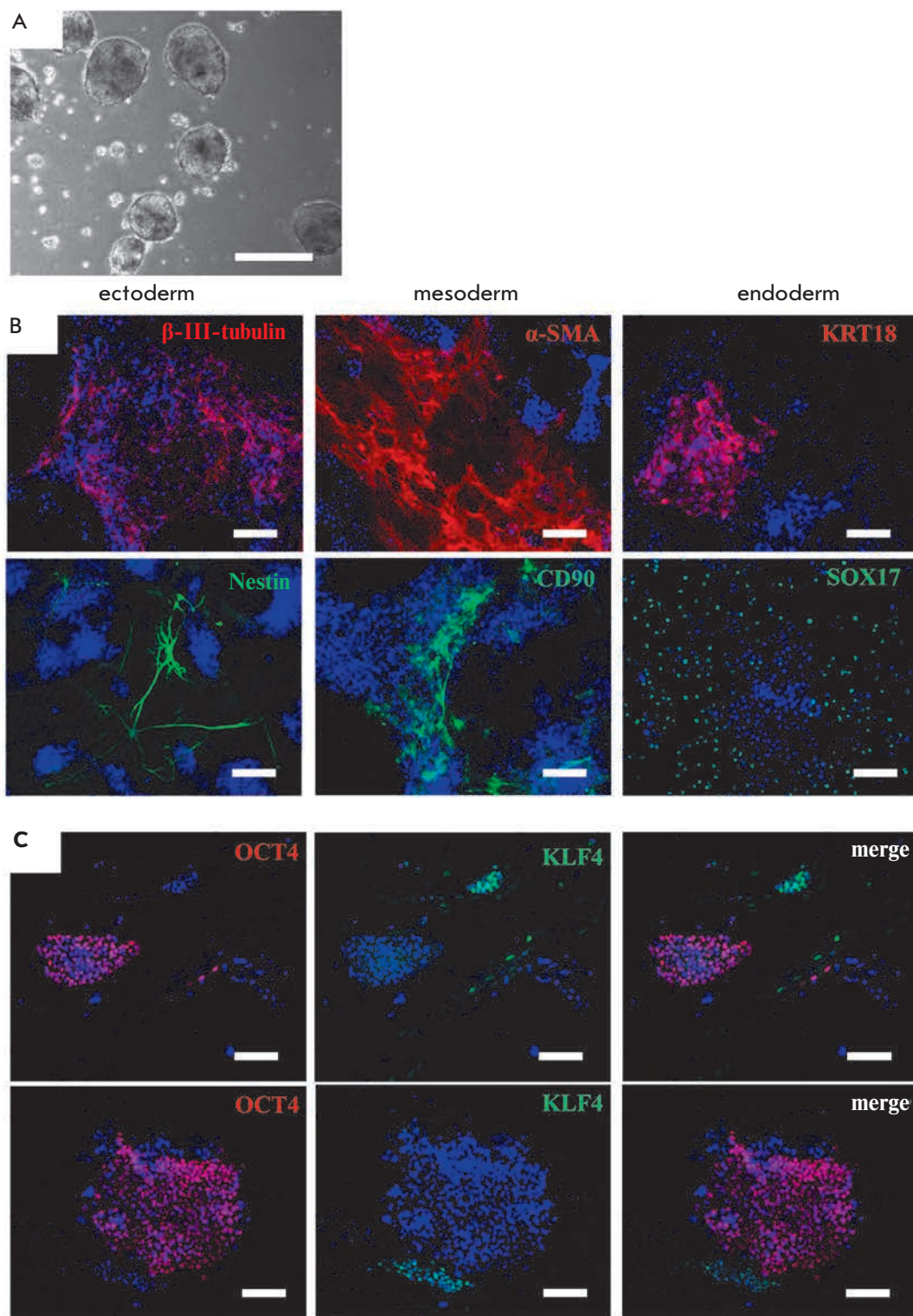


Fig. 4. Spontaneous differentiation of common vole iPSCs. **A** – morphology of embryoid bodies formed from 14v2 cells in suspension culture in 5 days. **B** – immunofluorescence analysis of differentiated derivatives of common vole iPSCs. Identification of ectoderm markers: β -III-tubulin (red signal), Nestin (green signal); mesoderm: α -SMA (red signal), CD90 (green signal); endoderm: KRT18 (red signal), SOX17 (green signal). **C** – immunofluorescent detection of the transcription factors OCT4 (red signal) and KLF4 (green signal). Nuclei are stained with DAPI (blue signal). Scale bar **A** – 500 μ m, **B**, **C** – 100 μ m

The factors bFGF and mLIF comprised the key combination for the induction and maintenance of pluripotency in common vole cells. Factor bFGF is known to participate in the triggering of the main signaling cascade of PSCs in such species as human, monkey, dog, cow, horse, and sheep [7, 9, 13, 15, 17, 51-55, 60-62]. A combination of LIF and bFGF is also used for the induction and maintenance of pluripotency in many mammalian species, including human [63], rabbit [13], dog [62], horse [7, 17], and sheep [9]. Comparison of human PSCs obtained and cultured in different conditions one of which includes bFGF only, and another bFGF, and LIF together showed that iPSCs and ESCs maintained using two factors are more similar in their characteristics of the transcriptome and epigenome to pluripotent cells of early embryos [63].

An important medium component that allowed us to conduct the reprogramming of differentiated cells of common voles to a pluripotent state is ascorbic acid. It has previously been shown that ascorbic acid possesses antioxidant properties, as well as activates histone demethylases and the TET proteins responsible for the most important epigenetic transformations upon pluripotency induction, which, in particular, is essential for the initiation of endogenous *Oct4* and *Nanog* expression in reprogrammed cells [47, 49].

It has been demonstrated on many mammalian species that it is preferable to use a medium supplemented with KSR, not FBS, when working with pluripotent cells [10, 13, 15, 16, 60]. However, during reprogramming of common vole fibroblasts to a pluripotent state, primary cell colonies were only obtained in media containing not less than 7% FBS. Moreover, substitution of FBS with KSR in the culture medium induced differentiation of the obtained common vole iPSCs even in the presence of growth factors.

The presence of mLIF, bFGF, ascorbic acid, and FBS in the culture medium is essential for maintaining the self-renewal and pluripotency capabilities of common vole iPSCs. Removal of any of these components triggers differentiation.

The obtained stable common vole iPSC lines share the properties of pluripotent cells. They exhibit AP activity and express the endogenous transcription factors *Nanog*, *Oct4*, *Sox2*, *Esrrb*, and *Sall4*, which are necessary for maintaining PSCs in the undifferentiated state [64-68]. Hereupon, the promoter region of *Oct4* is more hypomethylated in iPSCs than in embryonic fibroblasts. The obtained lines of common vole pluripotent cells are capable of unlimited self-renewal and differentiation *in vitro* into derivatives of the three primary germ layers. Nevertheless, unlike *M. ochrogaster* iPSCs, stable cell lines of common voles

with induced pluripotency do not maintain SSEA1 expression during culturing, and they also do not express KLF4. Probably, this may be due to the fact that various signaling cascades participate in maintaining the pluripotency of *M. ochrogaster* and common vole iPSCs. The expression of KLF4 and SSEA1 is specific to PSCs in which a key signaling cascade providing pluripotency is triggered by cytokine LIF and varies or absent in PSCs, the pluripotency of which is supported by a bFGF-activated signaling cascade [61, 69-71]. Thus, pluripotent cell lines derived from postimplantation mouse embryos in the presence of bFGF do not express KLF4, unlike ESCs and iPSCs cultured with LIF [71, 72]. Sheep, macaque, and human PSCs obtained in the presence of bFGF do not exhibit the SSEA1 surface antigen [8, 15, 61, 69]. It is noted that the pluripotent cells maintained with bFGF exhibit flatter morphology and proliferate slower [8, 72]. These two properties of bFGF-dependent pluripotent cells are specific to common vole iPSC lines. To speculate on whether the obtained iPSCs of *M. ochrogaster* and common voles fully reflect the properties of pluripotent cells of these species appears impossible: ESC lines that serve as a standard of pluripotency are absent both in prairie and common vole cultures. Nothing is known either about the properties of pluripotent cells of the pre- and postimplantation embryos of these species *in vivo*.

Thus, the technology of reprogramming by overexpression of four OSKM transcription factors allowed us to obtain for the first time common vole iPSCs, pluripotent cell lines that had been previously impossible to grow as a culture. These lines are to be used to study the processes of early development and pluripotency genes in common voles. We hope that the experience gained in the course of this work will allow us to further develop more effective approaches for the reprogramming of somatic cells of common voles and isolate ESCs from early preimplantation blastocysts and germinal cells of these rodents.

CONCLUSION

In this study, iPSCs of common vole *M. levis* × *M. arvalis* hybrids were obtained in a medium containing cytokine LIF, bFGF, ascorbic acid, and FBS. Common vole iPSCs obtained in these conditions are capable of self-renewal, they express the pluripotency genes *Oct4*, *Nanog*, *Sox2*, *Sall4*, and *Esrrb* and form derivatives of the three primary germ layers during differentiation *in vitro*. The results of our work will allow researchers to assess diversity and species-specific features in induction and maintenance of the pluripotent cells of various mammalian species.

The authors express their deep appreciation to I.S. Zaharova for her help in conducting experiments and processing the data of flow cytometry, E.A. Elisaphenko for assistance in designing primers for RT-PCR, E.A. Vaskova, A.V. Vyalkova and D.R. Bayzigitov for participation in several stages of the culturing work. The authors are also grateful to Rudolf Jaenisch, Konrad Hochedlinger and Didier Trono for

plasmids located in the Addgene depository. Work on the BD FACS Aria equipment was performed at the "Flow Cytometry" Center of Collective Use in ICG SB RAS.

This work was supported by RFBR grant № 15-04-03947 and budget project of the Institute of cytology and genetics SB RAS VI.60.1.2.

REFERENCES

- Evans M.J., Kaufman M.H. // *Nature*. 1981. V. 292. № 5819. P. 154–156.
- Martin G.R. // *Proc. Natl. Acad. Sci. USA*. 1981. V. 78. № 12. P. 7634–7638.
- Takahashi K., Yamanaka S. // *Cell*. 2006. V. 126. № 4. P. 663–676.
- Dutta D. // *Int. J. Dev. Biol.* 2013. V. 57. № 9–10. P. 667–675.
- Medvedev S.P., Shevchenko A.I., Elisaphenko E.A., Nesterova T.B., Brockdorff N., Zakian S.M. // *BMC Genomics*. 2008. V. 9. № 162. doi: 10.1186/1471-2164-9-162.
- Medvedev S.P., Elisaphenko E.A., Mazurok N.A., Zakian S.M. // *Dokl. Biochem. Biophys.* 2009. V. 425. № 1. P. 102–105.
- Breton A., Sharma R., Diaz A.C., Parham A.G., Graham A., Neil C., Whitelaw C.B., Milne E., Donadeu F.X. // *Stem Cells Dev.* 2013. V. 22. № 4. P. 611–621.
- Li Y., Cang M., Lee A.S., Zhang K., Liu D. // *PLoS One*. 2011. V. 6. № 1. e15947.
- Liu J., Balehosur D., Murray B., Kelly J.M., Sumer H., Verma P.J. // *Theriogenology*. 2012. V. 77. № 2. P. 338–346.
- Wu Z., Chen J., Ren J., Bao L., Liao J., Cui C., Rao L., Li H., Gu Y., Dai H., et al. // *J. Mol. Cell Biol.* 2009. V. 1. № 1. P. 46–54.
- Bao L., He L., Chen J., Wu Z., Liao J., Rao L., Ren J., Li H., Zhu H., Qian L., et al. // *Cell Res.* 2011. V. 21. № 4. P. 600–608.
- Ezashi T., Telugu B.P., Alexenko A.P., Sachdev S., Sinha S., Roberts R.M. // *Proc. Natl. Acad. Sci. USA*. 2009. V. 106. № 27. P. 10993–10998.
- Honda A., Hirose M., Hatori M., Matoba S., Miyoshi H., Inoue K., Ogura A. // *J. Biol. Chem.* 2010. V. 285. № 41. P. 31362–31369.
- Liao J., Cui C., Chen S., Ren J., Chen J., Gao Y., Li H., Jia N., Cheng L., Xiao H., et al. // *Cell Stem Cell*. 2009. V. 4. № 1. P. 11–15.
- Liu H., Zhu F., Yong J., Zhang P., Hou P., Li H., Jiang W., Cai J., Liu M., Cui K., et al. // *Cell Stem Cell*. 2008. V. 3. № 6. P. 587–590.
- Manoli D.S., Subramanyam D., Carey C., Sudin E., van Westerhuyzen J.A., Bales K.L., Blueloch R., Shah N.M. // *PLoS One*. 2012. V. 7. № 5. e38119.
- Nagy K., Sung H.K., Zhang P., Laflamme S., Vincent P., Agha-Mohammadi S., Woltjen K., Monetti C., Michael I.P., Smith L.C., et al. // *Stem Cell Rev.* 2011. V. 7. № 3. P. 693–702.
- Shimada H., Nakada A., Hashimoto Y., Shigeno K., Shionoya Y., Nakamura T. // *Mol. Reprod. Dev.* 2010. V. 77. № 1. P. 2.
- Takahashi K., Tanabe K., Ohnuki M., Narita M., Ichisaka T., Tomoda K., Yamanaka S. // *Cell*. 2007. V. 131. № 5. P. 861–872.
- Mazurok N.A., Rubtsova N.V., Isaenko A.A., Pavlova M.E., Slobodyanyuk S.Y., Nesterova T.B., Zakian S.M. // *Chromosome Res.* 2001. V. 9. № 2. P. 107–120.
- Dementyeva E.V., Shevchenko A.I., Anopriyenko O.V., Mazurok N.A., Elisaphenko E.A., Nesterova T.B., Brockdorff N., Zakian S.M. // *Chromosoma*. 2010. V. 119. № 5. P. 541–552.
- Rubtsov N.B., Rubtsova N.V., Anopriyenko O.V., Karamysheva T.V., Shevchenko A.I., Mazurok N.A., Nesterova T.B., Zakian S.M. // *Cytogenet. Genome Res.* 2002. V. 99. P. 323–329.
- Shevchenko A.I., Malakhova A.A., Elisaphenko E.A., Mazurok N.A., Nesterova T.B., Brockdorff N., Zakian S.M. // *PLoS One*. 2011. V. 6. № 8. e22771.
- Shevchenko A.T., Mazurok N.A., Slobodyanyuk S.Y., Zakian S.M. // *Chromosome Res.* 2002. V. 10. № 2. P. 117–126.
- Vaskova E.A., Dementyeva E.V., Shevchenko A.I., Pavlova S.V., Grigor'eva E.V., Zhelezova A.I., Vandeberg J.L., Zakian S.M. // *PLoS One*. 2014. V. 9. № 2. e88256.
- Sherstyuk V.V., Shevchenko A.I., Mazurok N.A., Zakian S.M. // *Dokl. Biochem. Biophys.* 2013. V. 450. № 5. P. 164–166.
- Nesterova T.B., Slobodyanyuk S.Y., Elisaphenko E.A., Shevchenko A.I., Johnston C., Pavlova M.E., Rogozin I.B., Kolesnikov N.N., Brockdorff N., Zakian S.M. // *Genome Res.* 2001. V. 11. № 5. P. 833–849.
- Zakian S.M., Kulbakina N.A., Meyer M.N., Semenova L.A., Bochkarev M.N., Radjabli S.I., Serov O.L. // *Genet. Res.* 1987. V. 50. № 1. P. 23–27.
- Sorokin M.A., Medvedev S.P., Shevchenko A.I., Slyn'ko N.M., Zakian S.M. // *Rus. J. Genet.* 2010. V. 46. № 2. P. 249–252.
- Mazurok N.A., Rubtsova N.V., Grigor'eva E.V., Matveeva N.M., Zhelezova A.I., Shilov A.G., Slobodianiuk S., Zakian S.M. // *Rus. J. Dev. Biol.* 2003. V. 34. № 3. P. 154–163.
- Grigor'eva E.V., Shevchenko A.I., Mazurok N.A., Elisaphenko E.A., Zhelezova A.I., Shilov A.G., Dyban P.A., Dyban A.P., Noniashvili E.M., Slobodyanyuk S.Y., et al. // *PLoS One*. 2009. V. 4. № 9. e7161.
- Shevchenko A.I., Demina V.V., Mazurok N.A., Zhelezova A.I., Efremov I.A., Shilov A.G., Shevela A.I., Belevantseva A.V., Vlasov V.V., Zakian S.M. // *Rus. J. Genet.* 2008. V. 44. № 11. P. 1280–1289.
- Grigor'eva E.V., Shevchenko A.I., Zhelezova A.I., Shilov A.G., Mazurok N.A., Dyban P.A., Dyban A.P., Zakian S.M. // *Bull. Exp. Biol. Med.* 2011. V. 150. № 4. P. 504–514.
- Carey B.W., Markoulaki S., Hanna J., Saha K., Gao Q., Mitalipova M., Jaenisch R. // *Proc. Natl. Acad. Sci. USA*. 2009. V. 106. № 1. P. 157–162.
- Maherali N., Ahfeldt T., Rigamonti A., Utikal J., Cowan C., Hochedlinger K. // *Cell Stem Cell*. 2008. V. 3. № 3. P. 340–345.
- Kingston R.E., Chen C.A., Okayama H. // *Curr. Protoc. Cell Biol.* 2003. Chapter 9. Unit 9.1.
- Pain B., Clark M.E., Shen M., Nakazawa H., Sakurai M., Samarut J., Etches R.J. // *Development*. 1996. V. 122. № 8. P. 2339–2348.
- Kumaki Y., Oda M., Okano M. // *Nucl. Acids Res.* 2008. V. 36. Web Server issue. W170–W175.
- Nesterova T.B., Duthie S.M., Mazurok N.A., Isaenko A.A., Rubtsova N.V., Zakian S.M., Brockdorff N. // *Chromosome Res.* 1998. V. 6. № 1. P. 41–48.

RESEARCH ARTICLES

40. Sorokin M.A., Elisafenko E.A., Mazurok N.A., Zakian S.M. // *Dokl. Biochem. Biophys.* 2013. V. 452. № 1. P. 229–233.
41. Bechard M., Dalton S. // *Mol. Cell. Biol.* 2009. V. 29. № 8. P. 2092–2104.
42. Buehr M., Meek S., Blair K., Yang J., Ure J., Silva J., McLay R., Hall J., Ying Q.L., Smith A. // *Cell.* 2008. V. 135. № 7. P. 1287–1298.
43. Chang M.Y., Kim D., Kim C.H., Kang H.C., Yang E., Moon J.I., Ko S., Park J., Park K.S., Lee K.A., et al. // *PLoS One.* 2010. V. 5. № 3. e9838.
44. Chen Y., Blair K., Smith A. // *Stem Cell Reports.* 2013. V. 1. № 3. P. 209–217.
45. Li P., Tong C., Mehrian-Shai R., Jia L., Wu N., Yan Y., Maxson R.E., Schulze E.N., Song H., Hsieh C.L., et al. // *Cell.* 2008. V. 135. № 7. P. 1299–1310.
46. Ying Q.L., Wray J., Nichols J., Battle-Morera L., Doble B., Woodgett J., Cohen P., Smith A. // *Nature.* 2008. V. 453. № 7194. P. 519–523.
47. Blaschke K., Ebata K.T., Karimi M.M., Zepeda-Martinez J.A., Goyal P., Mahapatra S., Tam A., Laird D.J., Hirst M., Rao A., et al. // *Nature.* 2013. V. 500. № 7461. P. 222–226.
48. Dutta D., Ray S., Home P., Larson M., Wolfe M.W., Paul S. // *Stem Cells.* 2011. V. 29. № 4. P. 618–628.
49. Esteban M.A., Wang T., Qin B., Yang J., Qin D., Cai J., Li W., Weng Z., Chen J., Ni S., et al. // *Cell Stem Cell.* 2010. V. 6. № 1. P. 71–79.
50. Rajendran G., Dutta D., Hong J., Paul A., Saha B., Mahato B., Ray S., Home P., Ganguly A., Weiss M.L., et al. // *J. Biol. Chem.* 2013. V. 288. № 34. P. 24351–24362.
51. Medvedev S.P., Grigor'eva E.V., Shevchenko A.I., Malakhova A.A., Dementyeva E.V., Shilov A.A., Pokushalov E.A., Zaidman A.M., Aleksandrova M.A., Plotnikov E.Y., et al. // *Stem Cells Dev.* 2011. V. 20. № 6. P. 1099–1112.
52. Medvedev S.P., Malakhova A.A., Grigor'eva E.V., Shevchenko A.I., Dementyeva E.V., Sobolev I.A., Lebedev I.N., Shilov A.G., Zhimulev I.F., Zakian S.M. // *Acta Naturae.* 2010. V. 2. № 2. P. 102–106.
53. Medvedev S.P., Shevchenko A.I., Zakian S.M. // *Acta Naturae.* 2010. V. 2. № 3. P. 30–46.
54. Medvedev S.P., Shevchenko A.I., Zakian S.M. // *Acta Naturae.* 2010. V. 2. № 2. P. 18–28.
55. Shutova M.V., Bogomazova A.N., Lagarkova M.A., Kiselev S.L. // *Acta Naturae.* 2009. V. 1. № 2. P. 91–92.
56. Ginis I., Luo Y., Miura T., Thies S., Brandenberger R., Gerecht-Nir S., Amit M., Hoke A., Carpenter M.K., Itskovitz-Eldor J., et al. // *Dev. Biol.* 2004. V. 269. № 2. P. 360–380.
57. Vassilieva S., Guan K., Pich U., Wobus A.M. // *Exp. Cell Res.* 2000. V. 258. № 2. P. 361–373.
58. Smith A.G., Heath J.K., Donaldson D.D., Wong G.G., Moreau J., Stahl M., Rogers D. // *Nature.* 1988. V. 336. № 6200. P. 688–690.
59. Williams R.L., Hilton D.J., Pease S., Willson T.A., Stewart C.L., Gearing D.P., Wagner E.F., Metcalf D., Nicola N.A., Gough N.M. // *Nature.* 1988. V. 336. № 6200. P. 684–687.
60. Han X., Han J., Ding F., Cao S., Lim S.S., Dai Y., Zhang R., Zhang Y., Lim B., Li N. // *Cell Res.* 2011. V. 21. № 10. P. 1509–1512.
61. Thomson J.A., Itskovitz-Eldor J., Shapiro S.S., Waknitz M.A., Swiergiel J.J., Marshall V.S., Jones J.M. // *Science.* 1998. V. 282. № 5391. P. 1145–1147.
62. Vaags A.K., Rosic-Kablar S., Gartley C.J., Zheng Y.Z., Chesney A., Villagomez D.A., Kruth S.A., Hough M.R. // *Stem Cells.* 2009. V. 27. № 2. P. 329–340.
63. Gafni O., Weinberger L., Mansour A.A., Manor Y.S., Chomsky E., Ben-Yosef D., Kalma Y., Viukov S., Maza I., Zviran A., et al. // *Nature.* 2013. V. 504. № 7479. P. 282–286.
64. Chambers I., Colby D., Robertson M., Nichols J., Lee S., Tweedie S., Smith A. // *Cell.* 2003. V. 113. № 5. P. 643–655.
65. Chambers I., Silva J., Colby D., Nichols J., Nijmeijer B., Robertson M., Vrana J., Jones K., Grotewold L., Smith A. // *Nature.* 2007. V. 450. № 7173. P. 1230–1234.
66. Masui S., Nakatake Y., Toyooka Y., Shimosato D., Yagi R., Takahashi K., Okochi H., Okuda A., Matoba R., Sharov A.A., et al. // *Nat. Cell. Biol.* 2007. V. 9. № 6. P. 625–635.
67. Mitsui K., Tokuzawa Y., Itoh H., Segawa K., Murakami M., Takahashi K., Maruyama M., Maeda M., Yamanaka S. // *Cell.* 2003. V. 113. № 5. P. 631–642.
68. Niwa H., Miyazaki J., Smith A.G. // *Nat. Genet.* 2000. V. 24. № 4. P. 372–376.
69. Adewumi O., Aflatoonian B., Ahrlund-Richter L., Amit M., Andrews P.W., Beighton G., Bello P.A., Benvenisty N., Berry L.S., Bevan S., et al. // *Nat. Biotechnol.* 2007. V. 25. № 7. P. 803–816.
70. Osorno R., Chambers I. // *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2011. V. 366. № 1575. P. 2230–2237.
71. Posfai E., Tam O.H., and Rossant J. // *Curr. Top. Dev. Biol.* 2014. V. 107. P. 1–37.
72. Tesar P.J., Chenoweth J.G., Brook F.A., Davies T.J., Evans E.P., Mack D.L., Gardner R.L., McKay R.D. // *Nature.* 2007. V. 448. № 7150. P. 196–199.

Modifiers of the Dipole Potential of Lipid Bilayers

S. S. Efimova*, O. S. Ostroumova

Institute of Cytology of the Russian Academy of Sciences, Tikhoretsky av. 4, St. Petersburg 194064, Russia

*E-mail: ssefimova@mail.ru

Copyright © 2015 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 This paper assesses the magnitude of change in the dipole potential (φ_d) of membranes caused by the adsorption of modifiers on lipid bilayers of various compositions. We tested flavonoids, muscle relaxants, thyroid hormones, and xanthene and styrylpyridinium dyes in order to assess their dipole-modifying properties. A quantitative description of the modifying action of flavonoids, muscle relaxants, thyroid hormones, and xanthene dyes is shown as the ratio of the maximum change in the bilayer dipole potential upon saturation and the absolute φ_d value of the unmodified membrane. The slopes of the linear relationship between the increase in the dipole potential of phospholipid bilayers and the concentration of styrylpyridinium dyes in membrane-bathing solutions were found. We described the relationships between the change in φ_d and the chemical structure of modifiers, as well as the charge and spontaneous curvature of lipid monolayers.

KEYWORDS Dipole modifiers, membrane dipole potential, xanthene and styrylpyridinium dyes, muscle relaxants, planar lipid bilayers, spontaneous curvature, thyroid hormones, flavonoids.

ABBREVIATIONS DPhPC – 1,2-diphytanoyl-sn-glycero-3-phosphocholine; DOPC – 1,2-dioleoyl-sn-glycero-3-phosphocholine; DOPS – 1,2-dioleoyl-sn-glycero-3-phosphoserine; DOPE – 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine.

INTRODUCTION

The finding of agents that are able to affect the value of the membrane dipole potential, φ_d , and hence to regulate the processes of transport through the plasma membrane both in health and disease is one of the most relevant problems of modern molecular pharmacology. This potential jump at the bilayer–solution interface occurs as a result of a certain relative orientation of the dipoles of membrane lipids and water molecules adsorbed on the bilayer surface [1–4]. The dipole potential of a membrane depends on its lipid composition. The essential role is played by the unsaturation, length, and number of hydrocarbon chains in phospholipid molecules [5–7]. The most common dipole modifiers are amphiphilic substances, whose molecules have a significant dipole moment and are characterized by a specific orientation on the phase interface. There exist published data of the successful use of dipole modifiers to study the molecular mechanisms of formation and functioning of ion channels formed by various toxins and antimicrobial agents [8–25]. It was found that dipole-modifying properties are characteristic of some flavonoids, steroids, thyroid hormones, and xanthene and styrylpyridinium dyes [1–3, 26–30].

Flavonoids are the most common phytochemical compounds. They are derivatives of benzo-gam-

ma-pyrone, whose structure is based on the scaffold consisting of two benzene rings (A and B) interconnected by a three-carbon fragment (C₂-C₃-C₄). Classification of flavonoids is based on the degree of oxidation of pyran (2-phenylchromane or C-ring). The following groups are recognized: chalcones, flavanones, flavonols, flavanones, flavanonols, iso-flavonoids, and others. Until recently, it was believed that the magnitude of membrane dipole potential can be affected only by chalcones, phloretin, and its glycoside, phlorizin [1, 3, 31].

Muscle relaxants are used to reduce the tone of skeletal muscles, including complete immobilization. Ammonium steroids (vecuronium, pancuronium, and rocuronium) are non-depolarizing relaxants. The structure of muscle relaxants is based on the steroid nucleus. There exist published data on the impact of some steroids on φ_d . Thus, it has been shown [32] that the introduction of cholesterol, 6-ketocholestanol, or a coprostanol in the membrane-forming solution of dimyristoylphosphocholine results in increased membrane dipole potential. Extraction of 5 α -androstan-3 β -ol from the lipid bilayer leads to an increase in membrane conductance induced by K⁺-nonactin [20]. This result indicates that 5 α -androstan-3 β -ol enhances the dipole potential of the bilayer. It was found [33] that the steroid

hormone pregnenolone reduces the dipole potential of liposome membranes formed of a mixture of dipalmitoylphosphatidylcholine and cholesterol.

Thyroid hormones thyroxine and triiodothyronine play a key role in metabolism regulation. They are iodinated tyrosine derivatives and differ from each other in the number and location of iodine atoms. It was also shown that iodine-containing thyroid hormones, as well as flavonoid phloretin, reduce the dipole potential of cholesterol-containing lipid bilayers [27].

Xanthene dyes are represented by two groups: fluoresceins and rhodamines. The first group includes fluorescein and its halogen derivatives (erythrosin, eosin, Rose Bengal, and phloxine B). The second group includes xanthene dyes that belong to the rhodamine family. They are fluorescein derivatives with both hydroxyl groups replaced by alkylated amino groups. It was shown that adsorption of Rose Bengal on the surface of the diphytanoylphosphatidylcholine membrane leads to a decrease in the membrane dipole potential [29], similarly to the effect of phloretin.

Styrylpyridinium dyes that belong to the AN-EPPS and RH series are potential-sensitive fluorochromes based on styrylhemicyanines. They differ in the length of their hydrocarbon tails and (or) polyene fragment. These fluorescent dyes have a high dipole moment due to a delocalized positive charge in the pyridine complex and negative charge of the sulfo-group at the other end of the molecule. RH dyes enhance the ϕ_d of phosphocholine membranes, and this ability decreases in the order RH 421, RH 237, and RH 160 [28].

The aim of our study was assessing and quantifying the effect of certain flavonoids, muscle relaxants, thyroid hormones, and fluorescent dyes on the dipole potential of lipid bilayers of various compositions. Particular attention was paid to the relationship between the chemical structure of modifiers and the dipole-modifying effectiveness of these compounds.

EXPERIMENTAL

Materials

The following reagents were used: KCl, HEPES, pentane, ethanol, chloroform, dimethylsulfoxide (DMSO), hexadecane, and squalene (Sigma, USA); phloretin, phlorizin, rutin, genistin, genistein, quercetin, myricetin, biochanin A, (\pm) catechin hydrate, (\pm) taxifolin hydrate, daidzein, 2',4',6'-trihydroxyacetophenone monohydrate (THAP), 2'-hydroxy-4',6'-dimethoxyacetophenone (DHAP), RH 421, di-8-ANEPPS, L-thyroxine, 3,3',5'-triiodo-L-thyronine, Rose Bengal, phloxineB, erythrosin, eosin Y, fluorescein, rhodamine 6G, rhodamine 101, pancuronium bromide, vecuronium

bromide, and rocuronium bromide (Sigma, USA); RH 160 and RH 237 (Molecular Probes, USA); 1,2-diphytanoyl-sn-glycero-3-phosphocholine (DPhPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dioleoyl-sn-glycero-3-phosphoserine (DOPS), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) (Avanti Polar Lipids, USA). The chemical structures of the used modifiers are shown in Table 1.

Measurements of the change in the dipole potential of lipid bilayers

The bilayer lipid membranes were formed using the Montal and Mueller method [34] by combining condensed lipid monolayers in a hole in the Teflon film that separated the experimental chamber into two (*cis*- and *trans*-) compartments. The volume of each compartment was 1.5 ml, Teflon film thickness was 10 μm , and hole diameter was about 50 μm . Before the beginning of the process, the hole in the Teflon membrane was pretreated with hexadecane. Monolayers were formed on the water-air interface using a 1 or 2 mg/ml lipid solution in pentane. DPhPC, DOPC, DOPE, DOPS, and an equimolar mixture of DOPE and DOPS (DOPS/DOPE) were used to obtain the monolayers. The experiments were performed using the same ionic composition of aqueous solutions of electrolyte in both compartments of the chamber (0.1 M KCl). Solution acidity (pH 7.4) was maintained with a 5 mM HEPES-KOH buffer.

Ionophores nonactin or valinomycin in the form of an ethanol (7 mg/ml) or methanol (0.8 mg/ml) solution, respectively, was added to the aqueous phase of the two compartments of the chamber to a final concentration of 10^{-7} - 10^{-5} M. It is known that phloretin in phospholipid membranes is less effective with respect to the transmembrane current induced by K^+ -valinomycin compared to that induced by K^+ -nonactin [1]. Similar results were obtained in preliminary experiments with other flavonoids, as well as muscle relaxants, xanthene dyes, and thyroid hormones. For this reason, changes in the membrane dipole potential caused by the introduction of these modifiers were measured using nonactin. It was found that styrylpyridinium dyes in the DPhPC bilayer are less effective with respect to the transmembrane current induced by K^+ -nonactin compared to that induced by K^+ -valinomycin. For this reason, valinomycin was used in the experiments measuring the increase in the membrane dipole potential caused by adsorption of these dyes.

Modifiers were added to both compartments of the chamber from millimolar solutions in ethanol, DMSO, or water to final concentrations in the membrane bathing solutions of 2.5 to 150 μM for flavonoids, 1 μM to 1 mM for muscle relaxants, 0.25 to 10 μM for xanthene

dyes, 1 to 50 μM for thyroid hormones, and 1 to 10 μM for styrylpyridinium dyes.

The final concentration of the solvent (ethanol, methanol, or DMSO) in the chamber did not exceed 0.1%. This concentration of the above-mentioned solvents did not cause the integrity and stability of the lipid bilayers. In the absence of ionophores, dipole modifiers at maximum concentrations likewise did not affect the conductance of the model membranes.

Transmembrane currents were measured and digitized in the voltage clamp mode, using Axopatch 200B and Digidata 1440A (Axon Instruments, USA). Silver-silver chloride electrodes (Ag/AgCl) connected to the solutions in the chamber through bridges with 1.5% agarose in the 2 M solution of KCl were used to apply the transmembrane potential (V) and record signals from the membrane. Measurements were performed at room temperature.

The data were processed using an 8-pole Bessel filter (Model 9002, Frequency Devices) and a filtering frequency of 1 kHz. Transmembrane current recordings were processed using the Clampfit 9.0 software package (Axon Instruments, USA). Statistical analysis of data was performed using the Origin 8.0 program (OriginLab, USA).

Membrane conductance (G) was determined as the ratio of the steady-state transmembrane current flowing through the lipid bilayer membrane (I) to the transmembrane potential (V), which was 50 mV. Change in the membrane dipole potential ($\Delta\phi_d$) caused by the introduction of the modifiers was assessed using Boltzmann statistics:

$$\Delta\phi_d = \frac{kT}{e} \ln\left(\frac{G_m}{G_m^0}\right), \quad (1)$$

where G_m^0 and G_m are the values of the steady-state K^+ -conductance of the bilayer induced by ionophore before and after the introduction of the modifier, e – electron charge, k – Boltzmann constant ($1,38 \times 10^{-23}$ J/K), and T – thermodynamic temperature ($T = 294$ K) [1].

The average values of the change in the membrane dipole potential were calculated as the arithmetic mean values of $\Delta\phi_d$ in each experimental system, measuring three to five bilayers (mean \pm SD).

Adsorption of flavonoids, muscle relaxants, thyroid hormones, and xanthene dyes on the surface of the lipid bilayers was described using the Langmuir isotherm:

$$\Delta\phi_d(C) = \frac{\Delta\phi_d(\infty)C}{C + K}, \quad (2)$$

where $\Delta\phi_d(C)$ – change in the membrane dipole potential at the concentration (C) of the modifier in the membrane bathing solution; $\Delta\phi_d(\infty)$ – maximum change in

the membrane dipole potential at $C \rightarrow \infty$; and K – dissociation constant of the modifier that characterizes its affinity to the lipid phase [3, 26]. The $\Delta\phi_d(\infty)$ value was determined using the plot of $\Delta\phi_d(C)$ function as the mean value corresponding to saturation, i.e. constant of the membrane dipole potential upon further increase in the concentration of the modifier. The K value was found as the slope of the linear approximation of the relationship

$$\frac{\Delta\phi_d(\infty)}{\Delta\phi_d(C)} \left(\frac{1}{C}\right)$$

The error of $\Delta\phi_d(\infty)$ was found as the maximum experimental error of the measurement of $\Delta\phi_d(C)$. The error of K was calculated as the error of the ratio:

$$\left(\frac{\Delta\phi_d(\infty)}{\Delta\phi_d(C)}\right).$$

No saturation effect was observed within the measured concentrations of styrylpyridinium dyes (10 μM). Further increase of the concentration of the dye results in the destruction of the lipid bilayer. For these reasons, we used the expression resulting from the linearization of the equation (2) at low concentrations of the dipole modifier ($C \ll K$) to describe the adsorption of styrylpyridinium dyes on the bilayer:

$$\Delta\phi_d(C) = \beta C, \quad (3)$$

where

$$\beta = \frac{\Delta\phi_d(\infty)}{K}$$

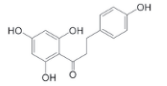
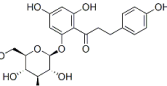
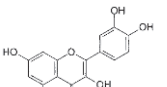
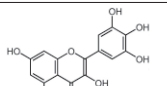
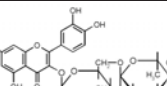
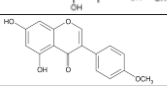
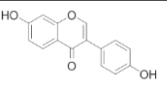
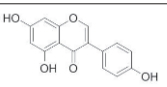
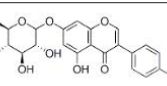
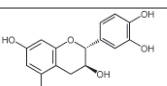
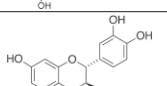
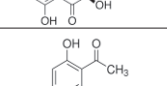
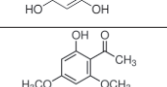
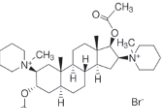
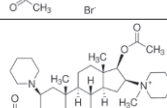
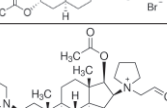
is the slope of the linear relationship between the increase in the bilayer dipole potential and dye concentration in the bathing solution [28].

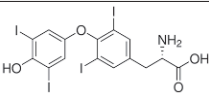
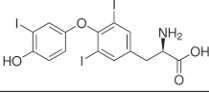
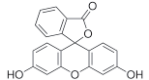
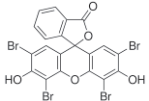
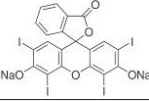
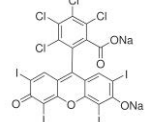
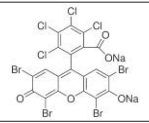
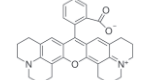
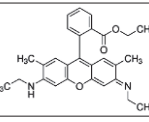
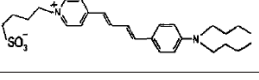
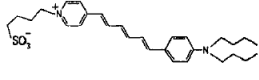
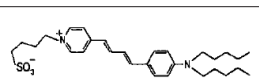
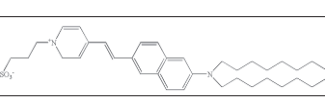
The relative value of the change in the dipole potential (γ) was used to compare the effectiveness of the dipole-modifying action of various modifiers. It was calculated as follows:

$$\gamma = \frac{\Delta\phi_d(\infty)}{\phi_{d_{nm}}} \cdot 100\%, \quad (4)$$

where $\phi_{d_{nm}}$ is the value of the dipole potential of the unmodified membrane, which can be found in the literature. In the absence of dipole modifiers, the dipole potential of DPhPC, DOPC, DOPS, and DOPE membranes was equal to 250 ± 40 [5, 35, 36], 225 ± 20 [5], 240 ± 20 mV [37, 38], and 220 ± 5 mV [5, 35, 36], respectively. $\phi_{d_{nm}}$ of DOPS/DOPE-bilayers was calculated as the average of $\phi_{d_{nm}}$ values for DOPS and DOPE membranes. The error of γ was calculated as the error of the ratio:

Table 1. The relative changes in the dipole potential of DPhPC membranes in the presence of various modifiers (γ) and their dissociation constants (K)

Class	Dipole modifier	Chemical structure	γ , %	K, μM
Flavonoids	Phloretin		-59 ± 12	$2.0 \pm 0.5^*$
	Phlorizin		-37 ± 7	$5.1 \pm 0.2^*$
	Quercetin		-42 ± 9	$3.3 \pm 0.5^*$
	Myricetin		-44 ± 12	$3.3 \pm 0.2^*$
	Ruthin		-17 ± 5	10.8 ± 0.5
	Biochanin A		-37 ± 10	$2.1 \pm 0.3^*$
	Daidzein		-8 ± 4	8.8 ± 0.2
	Genistein		-28 ± 8	$1.3 \pm 0.2^*$
	Genistin		-2 ± 2	9.6 ± 0.5
	Catechin		-1 ± 1	0.7 ± 0.2
	Taxifolin		-3 ± 1	5.8 ± 0.6
	THAP		-6 ± 3	$26.4 \pm 5.6^*$
DHAP		-20 ± 5	10.2 ± 0.4	
Muscle relaxants	Pancuronium		2 ± 2	0.1 ± 0.1
	Vecuronium		1 ± 1	0.1 ± 0.1
	Rocuronium		2 ± 2	0.1 ± 0.1

Thyroid hormones	Thyroxine		-24 ± 7	$3.5 \pm 0.1^{\text{®}}$
	Triiodothyronine		-23 ± 7	$5.3 \pm 0.2^{\text{®}}$
Xanthene dyes	Fluorescein		-2 ± 1	$0.4 \pm 0.1^{\text{®}}$
	Eosin Y		-2 ± 1	$0.4 \pm 0.1^{\text{®}}$
	Erythrosin		-26 ± 7	$0.8 \pm 0.1^{\text{®}}$
	Rose Bengal		-48 ± 11	$0.2 \pm 0.1^{\text{®}}$
	Phloxine B		-33 ± 7	$0.2 \pm 0.1^{\text{®}}$
	Rhodamine 101		-9 ± 4	0.3 ± 0.1
	Rhodamine 6G		-4 ± 1	0.4 ± 0.1
Styrylpyridinium dyes	RH 160		15 ± 6	-
	RH 237		19 ± 4	-
	RH 421		47 ± 9	-
	di-8-ANEPPS		1 ± 1	-

[®] Results are taken from [44].

* Results are taken from [46].

$$\frac{\Delta\phi_d(\infty)}{\phi_{d_nm}}$$

In the case of styrylpyridinium dyes, the γ value was calculated as the ratio of the change in the bilayer dipole potential at $5 \mu\text{M}$ concentration of the modifier to ϕ_{d_nm} . It was assumed that the agents having “weak” dipole-modifying properties are characterized by a γ value ranging from 0 to 10 %, “average”– 10 to 30 %, and “strong”– 30 to 60 %.

RESULTS AND DISCUSSION

Flavonoids

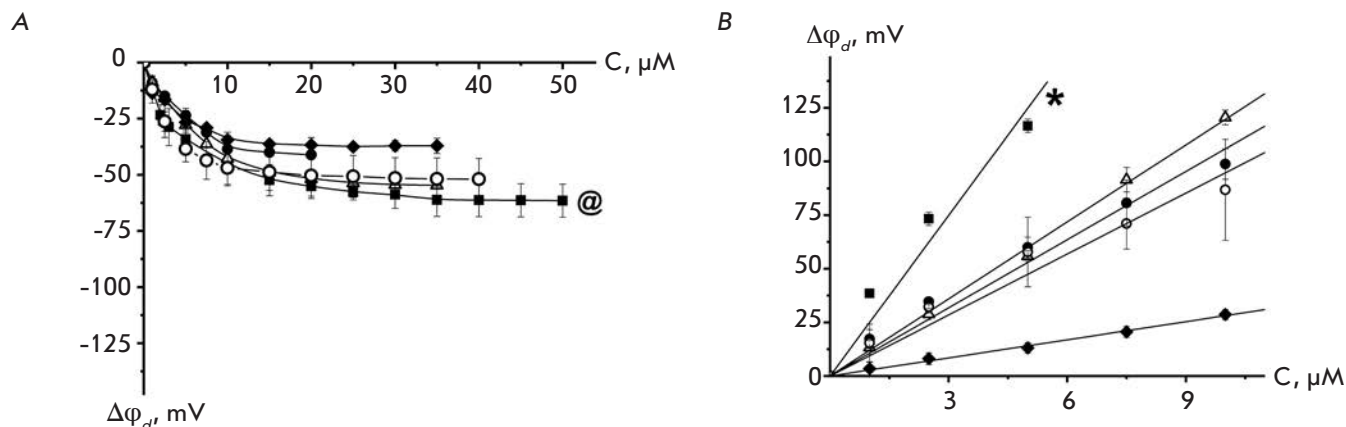
It is known that adsorption of phloretin on the membrane in a first approximation can be described by the Langmuir adsorption isotherm (1), which is characterized by the following: maximum change in the dipole potential upon saturation ($\Delta\phi_d(\infty)$) and dissociation constant of the flavonoid (K) [3, 39]. Table 1 shows γ

values (4) that characterize the relative change in the membrane dipole potential at the introduction of various modifiers. As shown from *Table 1*, all flavonoids reduce the membrane dipole potential. All the agents under study can be conveniently divided into three groups according to the intensity of the dipole-modifying properties. The first group includes modifiers with “low” effectiveness, which have a γ value ranging from 0 to 10 %. These are isoflavonoids daidzein and genistin, flavanol catechin, flavanonol taxifolin, and synthetic phloroglucinol THAP. The second group includes compounds having more pronounced dipole-modifying properties, so-called agents with “average”-effectiveness (γ ranges from 10 to 30 %). They are flavonol rutin, isoflavone genistein, and phloroglucinol DHAP. The third group includes the “strongest” flavonoid dipole modifiers characterized by a γ value of 30 to 60 %. They are chalcones phloretin and phlorizin, flavonols quercetin and myricetin, and isoflavone biochanin A.

When comparing the chemical structures of the flavonoids shown in *Table 1*, one can conclude that the dipole-modifying properties of modifiers are associated with the presence of a double bond in the C-ring of flavonoid molecules. There is no double bond in the C-rings of taxifolin or catechin as opposed to that of quercetin. As a result, bent-shaped flavanol and flavanonol have practically no influence on the ϕ_d value, while adsorption of planar flavonol on the membrane leads to a significant decrease in the dipole potential of the latter. When analyzing *table 1*, it can also be seen that a higher γ value correlates with a lower number of hydroxyl groups in the flavonoid molecule. Thus, the decrease in ϕ_d caused by the adsorption of phlorizin (phloretin glycoside) is less pronounced compared to that in the case of the more hydrophobic aglycone phloretin. A similar situation is observed for flavonols (quercetin/myricetin and rutin), isoflavones (biochanin A, genistein, and genistin), and phloroglucinols (DHAP and THAP). Unlike biochanin A, isoflavone daidzein has practically no effect on the value of the dipole potential of DPhPC membranes, despite its small number of hydroxyl groups. Since the dissociation constant of daidzein is higher compared to that of biochanin A, it can be assumed that the former has lower affinity to the lipid phase compared to the latter. The observed differences can also be caused by the different orientations of daidzein and biochanin A in the membrane due to the fact that the daidzein molecule has two hydroxyl groups located on opposite sides of the molecule, while in the biochanin A molecule they are located on the same side. It should also be noted that the dissociation constants of glycosides (phlorizin, rutin, and genistin) surpass this parameter in the corresponding aglycones (phloretin, quercetin, and genistein). It is likely that

this fact, as well as the less pronounced variation of ϕ_d in the presence of glycosides compared to that in the presence of aglycones, is due to the greater hydrophilicity of glycosides, and hence their lower affinity to the lipid phase. Phloroglucinols THAP and DHAP have the highest desorption constants among the studied flavonoids. The latter observation is consistent with the results of [40], which showed that the distribution coefficient of THAP between lecithin and water is 8 times lower than that of phloretin.

Table 2 shows the characteristic parameters of the adsorption isotherm of the “strongest” flavonoid modifier, phloretin, on lipid bilayers of different compositions. As *Table 2* shows, the ability of phloretin to reduce ϕ_d depends on the type of membrane-forming lipid. DOPE (because it is unsaturated) and DPhPC (because it has branched hydrocarbon chains) tend to form non-lamellar structures; so the bilayers formed of these lipids are characterized by elastic tension due to the deformation of monolayers having a negative spontaneous curvature. This tension can be detected when investigating the profile of the lateral pressure in the bilayer [41, 42]. DOPC forms a lamellar structure, and monolayers that contain DOPC have a very low spontaneous curvature. The maximum decrease in the dipole potential caused by the adsorption of phloretin is practically the same for DOPC, DOPE, and DPhPC membranes (*Tables 1* and *2*). These results indicate that the plane of adsorption of the modifier in the membrane does not match the plane of the lateral pressure jump in DOPE and DPhPC bilayers. Furthermore, phloretin is about 1.5 times less effective with respect to bilayers comprising negatively charged DOPS phospholipid (DOPS or DOPS/DOPE) compared to membranes formed of uncharged phospholipids (DPhPC, DOPC, or DOPE). A similar result was obtained [43] in the study of phloretin adsorption on neutral and negatively charged monolayers consisting of dimyristoylphosphatidylcholine and dimyristoylphosphatidylglycerol, respectively. Taking into account that the uncharged form of phloretin [1] is responsible for the decrease in the dipole potential, the observed differences cannot be determined by the decreased adsorption of the charged form of the modifier on DOPS-containing membranes. This is also evidenced by the close values of phloretin dissociation constants for uncharged DPhPC, DOPC, and DOPE and charged DOPS membranes (*Tables 1* and *2*). This suggests that the distribution ratio of the modifier is practically independent of the phospholipid composition of the membrane. This effect may stem from the positive spontaneous curvature of DOPS monolayers, arising from the “repulsion” of negatively charged lipid heads. As a result, phloretin molecules adsorp-



The dependence of change in the membrane dipole potential ($\Delta\phi_d$) on concentration of thyroxine (A) and RH 421 (B) in the bathing solution. Membranes were made formed of DPhPC[⊙] (■), DOPC (●), DOPS (◆), DOPE (Δ), and DOPS/DOPE (50/50 mol. %) (○) and bathed in 0.1M solution of KCl at pH 7.4. $V = 50$ mV.

⊙ Results are taken from [44]. * Results are taken from [46].

Table 2. The relative changes in the dipole potential of phospholipid bilayers in the presence of various modifiers (γ) and their dissociation constants (K)

Dipole modifier	Parameter	DOPC	DOPE	DOPS	DOPS/DOPE (50/50 mol.%)
Phloretin	γ , %	-62 ± 9	-58 ± 5	-38 ± 6	-41 ± 8
	K , μM	$0.7 \pm 0.2^*$	$2.2 \pm 0.4^*$	$2.7 \pm 0.8^*$	2.8 ± 0.2
Thyroxine	γ , %	-18 ± 5	-25 ± 4	-16 ± 5	-22 ± 6
	K , μM	3.8 ± 0.3	5.9 ± 0.5	2.8 ± 0.3	3.4 ± 0.2
RH 421	γ , %	27 ± 5	25 ± 2	5 ± 2	25 ± 9

* Results are taken from [46].

tion in this region may acquire an orientation different from that in the uncharged membranes and/or have a greater number of possible conformations.

Muscle relaxants

As Table 1 shows, addition of pancuronium bromide, vecuronium bromide, or rocuronium bromide to solutions bathing the DPhPC membrane has practically no effect on their dipole potential (γ value does not exceed 2 %). We can conclude that all the muscle relaxants under study have weak dipole-modifying properties. Given that a saturated steroid 5α -androstane- 3β -ol that has only one hydroxyl group enhances the dipole potential of the lipid bilayers [20], it can be suggested that the lack of impact by steroid relaxants on ϕ_d is due to modifications that enhance the hydrophilicity of the steroid molecule (additional acetate groups and nitrogen-containing heterocycles). High hydrophilicity and the presence of functional groups at different ends of the

pancuronium, vecuronium, and rocuronium molecules suggest that muscle relaxants adsorbed on the membrane surface are only slightly buried into the bilayer. A small depth of immersion is indirectly evidenced by the lack of effect by pancuronium bromide on the dipole potential of DOPC membranes (γ value is equal 3 ± 1), which, unlike DPhPC bilayers, do not have a lateral pressure jump on the hydrocarbon area. At the same time, small values of the dissociation constants of muscle relaxants (Table 1) are indicative of a high coefficient of distribution of these compounds between the bilayer and aqueous solution.

The surface charge of the membrane significantly affects the absorption of steroid muscle relaxants. Pancuronium bromide and vecuronium bromide enhance the dipole potential of negatively charged DOPS/DOPE membranes (γ value is equal 17 ± 3 %), while more hydrophobic rocuronium bromide has almost no effect on the ϕ_d value of DOPS/DOPE-bilayers

(γ value is equal about 1 %). The dependence of the effects on the charge of membrane-forming lipids suggests that charged forms of modifiers are responsible for the change in the dipole potential caused by the introduction of pancuronium bromide and vecuronium bromide. However, the dissociation constant of these muscle relaxants is two orders of magnitude higher and, therefore, affinity is lower in DOPS/DOPE bilayers compared to those in DPhPC membranes, which is indicative of the opposite. It is likely that the observed differences are not due to electrostatic interaction between the modifiers and the DOPS-containing membrane, but rather the positive spontaneous monolayer curvature, as in the case of phloretin. This suggests that in the DOPS/DOPE bilayer pancuronium and vecuronium are located near the repulsing, negatively charged serine heads.

Thyroid hormones

Comparing the γ value of thyroid hormones showed that thyroxine and triiodothyronine are dipole modifiers with “average” effectiveness that enhance the dipole potential of DPhPC membranes in a similar manner. Similar results were obtained previously [27]. This finding suggests that the presence of an additional iodine atom in the thyroxine molecule (compared to the triiodothyronine molecule) has little effect on the dipole moment of the modifier and its orientation in the bilayer.

Figure A shows the decrease in the dipole potential of DPhPC, DOPC, DOPS, DOPE, and DOPS/DOPE bilayers versus thyroxine concentration in the bathing solutions. *Table 2* shows γ values that characterize the relative changes in the membrane dipole potential caused by adsorption of thyroxine on lipid bilayers of different compositions. *Figure A* and *Table 2* show that the effectiveness of thyroxine weakly depends on the charge of membrane-forming lipids. Similar results were obtained when comparing DPhPC and diphytanoylphosphoserine bilayers [44]. The lack of a relationship between the modifier effects (both the γ and K values) and membrane surface charge indicates that the decrease in the membrane dipole potential is caused by the adsorption of the uncharged form of iodine-containing thyroid hormones. This finding is also evidenced by data reported in [27]. The closeness of γ values for the DPhPC, DOPC, DOPE, and DOPS membranes suggests that the spontaneous curvature of monolayers does not affect thyroxin adsorption. It can be assumed that the adsorption plane of the modifier is located between the areas corresponding to the lateral pressure jump in the DOPE and DPhPC membranes and location of charged serine residues in DOPS bilayers.

Xanthene dyes

Table 1 shows that the xanthene dyes discussed in this paper reduce the dipole potential of DPhPC membranes. Evaluation of the effectiveness of the dipole-modifying action of these dyes suggests that Rose Bengal and phloxine B belong to the most effective modifiers; erythrosin belongs to substances with “average” effectiveness; and fluorescein, eosin Y, rhodamine 6G, and rhodamine 101 belong to substances with “low” effectiveness.

When comparing the structures shown in *Table 1*, one can conclude that the type and location of halogen substituents in the dye molecule are the main factors determining the decrease in the membrane dipole potential caused by the introduction of these modifiers. It can be assumed that the pronounced decrease in the dipole potential of membranes caused by the introduction of erythrosine is due to the presence of iodine atoms in its molecule. The lack of the latter in fluorescein or replacement of iodine with bromine in eosin Y results in a loss of the dipole-modifying properties of these compounds. Similarly, replacement of iodine with bromine in phloxine B reduces the effectiveness of this modifier compared to that of Rose Bengal. In this case, the strong dipole-modifying properties of Rose Bengal should be attributed to the presence of chlorine in its structure. The presence of both the iodine and chlorine atoms in the Rose Bengal molecule makes it the most effective dipole modifier among the known xanthene dyes. Replacement of the hydroxyl group in the fluorescein molecule with the amino group in rhodamine molecules has no effect on the modifier’s capability to change the bilayer dipole potential.

Previously, we have shown that the anionic form of the xanthene dye is responsible for the reduction in the membrane dipole potential [44].

Comparing K values shows that xanthene dyes are characterized by an order of magnitude higher affinity for phospholipid membranes compared to that of flavonoids and thyroid hormones (*Table 1*).

Styrylpyridinium dyes

According to [28], the increase in the dipole potential of DPhPC membranes is directly proportional to the concentration of styrylpyridinium dyes in bathing solutions in a range from 0 to 15 μM . *Table 1* shows the γ value that characterizes the relative change in the membrane dipole potential caused by the introduction of 5 μM of the dye. Based on these results, an efficacy scale of styrylpyridinium dyes can be composed as follows: di-8-ANEPPS have no dipole-modifying effect, RH 160 and RH 237 are characterized by “average” effectiveness, and RH 421 is characterized by the highest effectiveness with respect to the dipole potential

of DPhPC membranes. These results for RH dyes are consistent with data presented in [28].

The ability to enhance the φ_d depends on the orientation and depth of immersion of the dye into the membrane. According to [45], the depth of immersion into the DPhPC bilayer increases in the order RH 160 < RH 421 < RH 237. RH 160 demonstrates minimum immersion into the membrane among the tested dyes, which is probably due to its lowest hydrophobicity. The highest effectiveness of RH 421 with respect to DPhPC membranes given its intermediate adsorption plane can be due to the closest orientation of the long axis of this dye in the membrane to the surface normal [45]. RH 421 and di-8-ANEPPS should have a similar dipole moment, as they have the same length of pyridine complexes. For this reason, the lack of influence of di-8-ANEPPS on the dipole potential of DPhPC membranes can be associated with longer hydrocarbon “tails” compared to those in RH 421, which determine the immersion and orientation of the dye in the bilayer, rather than the structural differences in the pyridine complex.

Figure B shows the dependence of increase in the dipole potential of DPhPC, DOPC, DOPE, DOPS, and DOPS/DOPE bilayers on RH 421 concentration in bathing solutions. Table 2 shows γ values that characterize the relative increase in the membrane dipole potential caused by the introduction of 5 μ M of RH 421 in membrane bathing solutions. The dependence of the dipole-modifying action of RH 421 on the type of membrane-forming lipid (DPhPC, DOPC, and DOPS) may indicate the influence of the lateral pressure profile on dye orientation in the bilayer. Furthermore, RH 421 has a low effectiveness with respect to negatively charged DOPS membranes. This may result from the repulsion of negatively charged sulfonate groups of the modifier and serine moieties. In

all probability, this contributes to the increase in the positive spontaneous curvature of the monolayer upon adsorption of RH 421 and change in the orientation of the dipole moment of the dye compared to that in DPhPC, DOPC, DOPE, and DOPS/DOPE membranes. The same conclusion was reached in the study of the channel-forming activity of antimicrobial peptides in the presence of RH 421 [15].

CONCLUSION

The dipole-modifying effect of certain flavonoids, steroid muscle relaxants, thyroid hormones, and xanthene and styrylpyridinium dyes on phospholipid bilayers of different compositions was quantitatively characterized. The structural features of the modifiers responsible for their ability to change the value of the membrane dipole potential were identified. Typically, more hydrophobic compounds have more pronounced dipole-modifying properties. In the case of flavonoids, the conformation of the molecule and position of hydroxyl groups are also important, while in the case of xanthene dyes, the important factors are the type and position of halogen substituents. Variation in the phospholipid composition of membranes allowed us to predict the plane of adsorption of the most effective compounds in each group of modifiers. Changing the lateral pressure profile of the bilayer affects the adsorption of phloretin, pancuronium bromide, vecuronium bromide, and RH 421.

This work was supported by the Russian Foundation of Science (No 14-14-00565). The experiments with muscle relaxants was partly supported by Russian Foundation for Basic Research (No 15-34-20356), Program of the Presidium of the Russian Academy of Sciences “Molecular and Cellular Biology”, SP-69.2015.4 and SS-1721.2014.4.

REFERENCES

- Andersen O.S., Finkelstein A., Katz I., Cass A. // *J. Gen. Physiol.* 1976. V. 67. P. 749–771.
- Franklin J.C., Cafiso D.S. // *Biophys. J.* 1993. V. 65. P. 289–299.
- Cseh R., Hetzer M., Wolf K., Kraus J., Bringmann G., Benz R. // *Eur. Biophys. J.* 2000. V. 29. P. 172–183.
- Wang L. // *Annu. Rev. Biochim.* 2012. V. 81. P. 615–635.
- Pickar A.D., Benz R. // *J. Membr. Biol.* 1978. V. 44. P. 353–376.
- Clarke R.J. // *Biochim. Biophys. Acta.* 1997. V. 1327. P. 269–278.
- Starke-Peterkovic T., Clarke R.J. // *Eur. Biophys. J.* 2009. V. 39. P. 103–110.
- Sun X., Garlid K.D. // *J. Biol. Chem.* 1992. V. 267. P. 19147–19154.
- Rokitskaya T.I., Kotova E.A., Antonenko Y.N. // *Biophys. J.* 2002. V. 82. P. 865–873.
- Hwang T.C., Koeppe R.E., Andersen O.S. // *Biochemistry.* 2003. V. 42. P. 13646–13658.
- Luchian T., Mereuta L. // *Langmuir.* 2006. V. 22. P. 8452–8457.
- Ostroumova O.S., Kaulin Y.A., Gurnev A.P., Schagina L.V. // *Langmuir.* 2007. V. 23. P. 6889–6892.
- Asandei A., Mereuta L., Luchian T. // *Biophys. Chem.* 2008. V. 135. P. 32–40.
- Mereuta L., Luchian T., Park Y., Hahm K.S. // *Biochem. Biophys. Res. Commun.* 2008. V. 373. P. 467–472.
- Apetrei A., Mereuta L., Luchian T. // *Biochim. Biophys. Acta.* 2009. V. 1790. P. 809–816.
- Ostroumova O.S., Malev V.V., Ilin M.G., Schagina L.V. // *Langmuir.* 2010. V. 26. P. 15092–15097.
- Lundbaek J.A., Koeppe R.E., Andersen O.S. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. P. 15427–15430.
- Ostroumova O.S., Efimova S.S., Schagina L.V. // *Biochim.*

RESEARCH ARTICLES

- Biophys. Acta. 2011. V. 1808. P. 2051–2058.
19. Mereuta L., Asandei A., Luchian T. // PLoS One. 2011. V. 6. P. e25276.
20. Ostroumova O.S., Efimova S.S., Schagina L.V. // PLoS One. 2012. V. 7. P. e30261.
21. Ostroumova O.S., Efimova S.S., Chulkov E.G., Schagina L.V. // PLoS One. 2012. V. 7. P. e45135.
22. Ostroumova O.S., Efimova S.S., Mikhailova E.V., Schagina L.V. // Eur. Biophys. J. 2014. V. 43. P. 207–215.
23. Efimova S.S., Schagina L.V., Ostroumova O.S. // Langmuir. 2014. V. 30. P. 7884–7892.
24. Chulkov E.G., Schagina L.V., Ostroumova O.S. // Biochim. Biophys. Acta. 2015. V. 1848. P. 192–199.
25. Efimova S.S., Zakharov V.V., Ostroumova O.S. // Cell and Tissue Biology. 2015. V. 9. P. 250–259.
26. Reyes J., Greco F., Motais R., Latorre R. // J. Membr. Biol. 1983. V. 72. P. 93–103.
27. Tsybul'skaya M.V., Antonenko Yu.N., Tropsha A.E., Yaguzhinsky L.S. Biofizika. 1984. V. 29. P. 801–805 (in Russian).
28. Malkov D.Y., Sokolov V.S. // Biochim. Biophys. Acta. 1996. V. 1278. P. 197–204.
29. Kotova E.A., Rokitskaya T.I., Antonenko Yu.N. // Membr. Cell Biol. 2000. V. 13. P. 411–420.
30. Issé B.A., Yunes Quartino P., Fidelio G.D., Fariás R.N. // Chem. Phys. Lipids. 2013. V. 175–176. P. 131–137.
31. Tarahovsky Y.S., Muzafarov E.N., Kim Y.A. // Mol. Cell Biochem. 2008. V. 314. P. 65–71.
32. Starke-Peterkovic T., Turner N., Vitha M.F., Waller M.P., Hibbs D.E., Clarke R.J. // Biophys. J. 2006. V. 90. P. 4060–4070.
33. Alakoskela J.M., Söderlund T., Holopainen J.M., Kinnunen P.K. // Mol. Pharmacol. 2004. V. 66. P. 161–168.
34. Montal M., Muller P. // Proc.Nat.Acad.Sci.USA. 1972. V. 65. P. 3561–3566.
35. Cseh R., Benz R. // Biophys. J. 1998. V. 74. P. 1399–1408.
36. Peterson U., Mannock D. A., Lewis R.N., Pohl P., McElhane R.N., Pohl E.E. // Chem. Phys. Lipids. 2002. V. 117. P. 19–27.
37. Flewelling R.F., Hubbell W.L. // Biophys. J. 1986. V. 49. P. 531–540.
38. Flewelling R.F., Hubbell W.L. // Biophys. J. 1986. V. 49. P. 541–552.
39. de Levie R., Rangarajan S.K., Seelig P.F., Andersen O.S. // Biophys. J. 1979. V. 25. P. 295–300.
40. Awiszus R., Stark G. // Eur. Biophys. J. 1998. V. 15. P. 299–310.
41. Cantor R.S. // Biophys. J. 1999. V. 76. P. 2625–2639.
42. Bezrukov S.M. // Current Opinion in Colloid Interface Sci. 2000. V. 5. P. 237–243.
43. Lairion F., Disalvo E.A. // Langmuir. 2004. V. 20. P. 9151–9155.
44. Efimova S.S., Schagina L.V., Ostroumova O.S. // J. Membr. Biol. 2014. V. 247. P. 739–745.
45. Passechnik V.I., Sokolov V.S. // Bioelectrochemistry. 2002. V. 55. P. 47–51.
46. Efimova S.S., Ostroumova O.S. // Langmuir. 2012. V. 28. P. 9908–9914.

Valproic Acid Increases the Hepatic Differentiation Potential of Salivary Gland Cells

O. S. Petrakova^{1, 2, 5, *}, V. V. Ashapkin³, V. Y. Shtratnikova⁴, L. I. Kutueva³, E. A. Vorotelyak^{1, 2, 5}, M. A. Borisov², V. V. Terskikh⁵, I. G. Gvazava^{2, 5}, A. V. Vasiliev⁵

¹Lomonosov Moscow State University, Faculty of Biology, Leninskie Gory, Moscow State University, 1, bld. 12, 119991, Moscow, Russia

²Pirogov Russian National Research Medical University, Ostrovitianov str., 1, 117997, Moscow, Russia

³Belozersky Institute, Moscow State University, Leninskie Gory, 1/40, 119991, Moscow, Russia

⁴Department of bioengineering and bioinformatics, Lomonosov Moscow State University, Leninskie Gory, 1/73, 119991, Moscow, Russia

⁵Koltsov Institute of Developmental Biology, Russian Academy of Sciences, Vavilova str., 26, 119334, Moscow, Russia

*E-mail: PetrakovaOl@yandex.ru

Copyright © 2015 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 studies of cell plasticity and differentiation abilities are important problems in modern cellular biology. The use of histone deacetylase inhibitor - valproic acid is a promising approach to increasing the differentiation efficiency of various cell types. In this paper we investigate the ability of mouse submandibular salivary gland cells to differentiate into the hepatic direction and the effect of valproic acid on the efficiency of this differentiation. It was shown that the gene expression levels of hepatocyte markers (*Aat*, *Afp*, *G6p*, *Pepck*, *Tat*, *Cyp3a13*) and liver-enriched transcription factors (*Hnf-3α*, *Hnf-3β*, *Hnf-4α*, *Hnf-6*) were increased after differentiation in salivary gland cells. Valproic acid increases the specificity of hepatic differentiation, reducing the expression levels of the ductal (*Krt19*, *Hhex1*, *Cyp7a1*) and acinar (*Ptf1a*) markers. After valproic acid exposure, the efficiency of hepatic differentiation also increases, as evidenced by the increase in the gene expression level of *Alb* and *Tdo*, and increase in urea production by differentiated cells. No change was found in DNA methylation of the promoter regions of the genes; however, valproic acid treatment and subsequent hepatic differentiation largely affected the histone H3 methylation of liver-enriched genes. Thus, mouse submandibular salivary gland cells are capable of effective differentiation in the hepatic direction. Valproic acid increases the specificity and efficiency of the hepatic differentiation of these cells.

KEYWORDS gene expression, hepatic differentiation, submandibular salivary gland cells, valproic acid.

ABBREVIATIONS DAPI - 4',6-diamidino-2-phenylindole, EGF - Epidermal Growth Factor, ITS - Insulin-Transferrin-Selenium, LPC - liver progenitor cells, qRT-PCR - quantitative real-time polymerase chain reaction, SGC - salivary gland cells, SGC-diff - differentiated salivary gland cells, SGC-VPA-diff - differentiated salivary gland cells, pretreated with valproic acid, VPA - valproic acid.

INTRODUCTION

The investigation of cell differentiation abilities is one of the most important problems in modern cell biology. *In vitro* expansion and subsequent differentiation can provide cells for regenerative medicine, as well as help in the study of development regulation and drug discovery. One of the promising approaches in this field is the use of small molecules possessing the ability to change the epigenetic status of cells [1]. Using small molecules facilitates cell reprogramming and increases the differentiation efficiency of various cell types [2-5].

Valproic acid (valproate, 2-*n*-propylpentanoic acid, VPA) has been used for decades as an effective antiepileptic drug with a broad spectrum of action [6]. Valproic acid acts as an inhibitor of histone deacetylases thereby causing an increase in gene expression [7]. The idea of histone deacetylase inhibitors application is based on the fact that histone acetylation causes the activation of various gene expressions, resulting in an increase in the transcription pool of cells and, hence, increases the cell differentiation ability. Some researchers believe that this effect on cell epi-

genetic regulation may also cause cells dedifferentiation [8].

With regard to the endoderm, liver and pancreas are two crucial organs of this germ layer. The ability of valproic acid to improve the efficiency of hepatic differentiation both in pluripotent and differentiated cells has been shown [9, 10]. The use of valproic acid during a standard hepatic differentiation protocol increases the hepatocyte differentiation of mouse embryonic stem (ES) cells and reduces the differentiation into ductal structures [9]. Human bone marrow cells much more efficiently differentiate in the hepatic direction after 5mM valproic acid exposure. These cells express albumin and produce urea more efficiently than cells not exposed to valproic acid treatment [10]. Human umbilical-cord-derived mesenchymal stem cells more efficiently undergo hepatic differentiation after 10mM valproic acid exposure in a concentration-dependent manner [11].

Cellular therapies of liver and pancreas disorders are hampered by a limited source of cells with the ability to differentiate within endoderm with high efficiency. The search for cells capable of differentiation in the hepatic and pancreatic directions with high efficiency is a challenge. Submandibular salivary gland cells have high differentiation abilities within endoderm and are an attractive source of adult cells for the cellular therapy of liver and pancreas disorders [12-14]. Ductal epithelial cells of submandibular salivary glands are easily accessible from patients and are easy to culture. As it was shown previously, submandibular salivary gland cells (SGC) isolated from humans and different animals (mouse, rat and swine) represent an active proliferating culture *in vitro* and possess high differentiation ability in the hepatic and pancreatic directions [13, 15-17]. Cultured submandibular salivary gland cells possess phenotypic convergence with liver progenitor cells (LPS) in mice [18]. A comparative *in vitro* analysis of mouse submandibular salivary gland and liver progenitor cells revealed similarities in cell markers gene expression: EpCAM, CD29, c-Kit, Sca-1, c-Met, cytokeratins 8, 18, 19, Afp, as well as in regulatory factors gene expression [18]. Under certain conditions, SGC acquire the ability of insulin or albumin expression [14-16] but the hepatic and pancreatic differentiation of SGC is incomplete [16]. The treatment of mouse submandibular salivary gland cells with 5mM valproic acid for 5 days causes an increase in hepatic (*G6p*, *Alb*, *Tdo*) and pancreatic (*Ngn3*, *Pax4*, *Ins1*) markers expression levels [19]. The effect of histone deacetylase inhibitors on the differentiation potential of cells is reversible and not specific; thus, small molecules treatment is usually combined with specific differentiation cytokines.

In this study, we investigate the effect of valproic acid treatment on the hepatic differentiation of mouse

submandibular salivary gland cells. For SGC hepatic differentiation, we chose a standard protocol [20, 21] including the main stages occurring during the differentiation of liver cells. Before performing the cell differentiation procedure, we treated SGC with 5mM valproic acid for 5 days and compared the effectiveness of hepatic differentiation of pretreated and intact cells. The first-passage SGC and LPC were used as controls. This approach will help to estimate the influence of valproic acid on the efficiency of salivary gland cells hepatic differentiation and will help assess the phenotypic plasticity of salivary gland cells.

MATERIALS AND METHODS

Animals

C57BL/6 male mice (aged 8-15 weeks) were used. All animal experiments were performed in accordance with the Ethics Committee for Animal Research of the Koltsov Institute of Developmental Biology, Russian Academy of Sciences, as approved by the Guidelines for Humane Endpoints for Animals Used in Biomedical Research, Regulations for Laboratory Practice in the Russian Federation.

Cell culture

The mice were anaesthetized by injecting 300 mg/kg of chloral hydrate (Sigma) intraperitoneally. The submandibular salivary glands and the liver were excised under aseptic conditions. The organs were dissected in phosphate-buffered saline (PBS), washed twice with PBS containing 40 µg/ml gentamicin and incubated in DMEM/F12 (1:1) medium (Gibco) with 0.1% type IV collagenase (Sigma) for 30-40 min at 37°C. The cells were pipetted and passed through a filter with 40 µm pores (Corning). The cell suspensions were washed twice in DMEM/F12 culture medium using "gentle" centrifugation (2 min, 100g). The cells were plated into culture dishes (Corning) coated with collagen type I at a density of $5 \cdot 10^3$ cells per cm^2 . The cells were cultured in DMEM/F-12 (1:1) supplemented with 10% fetal bovine serum (FBS) (HyClone), 2 mM glutamine (Gibco), 1% insulin-transferrin-selenium supplement (ITS) (Invitrogen), and 10 ng/ml epidermal growth factor (EGF) (Gibco). This standard culture medium was changed every 3 days. After the monolayer formation (on day 5-7), the cells were harvested using 0.25% trypsin/EDTA (Gibco) and plated onto collagen type I-coated culture dishes at a density of $1 \cdot 10^4$.

Culture conditions for hepatic differentiation of salivary gland cells

A first-passage monolayer culture of SGC was used for the hepatic differentiation (approximately 15-20 day

after cell isolation). SGC were divided into two groups: one group of SGC was treated for 5 days with 1mM valproic acid in a standard culture medium (the media with VPA was changed every day). Another group of SGC was incubated in a standard culture medium for 5 days without VPA treatment (the media was changed every day). Then, VPA was removed and the standard media was substituted with hepatic differentiation media: DMEM/F-12 (1:1) supplemented with 10% FBS (HyClone), 2 mM glutamine (Gibco), 1% ITS (Invitrogen), 0.03 mM Nicotinamide (Sigma), 20 ng/ml EGF (Gibco), and 20 ng/ml hepatocyte growth factor (HGF) (Invitrogen). This day was considered as the first day of differentiation. The same differentiation procedure was performed in both groups of SGC: cells were cultured in hepatic differentiation media with 20 ng/ml BMP2 (Invitrogen) for 5 days, then with 10 ng/ml oncostatin M (Invitrogen) and 0.1 μ M dexamethasone (Sigma) for 5 days, then with 1% N2 and 1% B27 (Invitrogen) for 5 days. Media changes were performed every 3 days during the differentiation procedure.

On the 15th day of cell differentiation, the cells were analyzed under 2D and 3D culture conditions. The undifferentiated first-passage SGC and LPC were used as controls.

Immunocytochemistry

For immunocytochemical staining, the cells were grown as a monolayer on collagen I-coated dishes for 3 days. The cells were fixed with 4% paraformaldehyde (Sigma) for 10 min and permeabilized in PBS containing 0.1% Triton-X (Sigma) for 30 min at room temperature. Non-specific binding was blocked for 30 min with PBS containing 1% bovine serum albumin (Sigma) at room temperature. The cells were incubated with primary antibodies diluted in PBS for 1 h at 37°C. The antibodies and dilutions used were as follows: anti-alpha-fetoprotein (Afp), 1:200 (R&D; MAB1368); anti-albumin (Alb), 1:200 (R&D; MAB1455); anti-cytokeratin 19 (Krt19), 1:100 (AbCam; ab15463-1). Then, the cells were washed three times in PBS and incubated with secondary antibodies diluted in PBS (1:1000) for 40 min at 37°C. The secondary antibodies used: Alexa Fluor® 488 donkey anti-rabbit IgG (H+L) (Invitrogen; A-21206), and Alexa Fluor® 488 goat anti-mouse IgG (H+L) (Invitrogen; A-11029). The cells were washed three times for 10 min in PBS (during the last wash DAPI (Sigma) was added). The cells were viewed using the fluorescence microscope Olympus IX51. For negative controls, the secondary antibodies were used.

Extraction of total ribonucleic acid (RNA)

The extraction of total RNA was performed using the AllPrep DNA/RNA Mini Kit (Qiagen) according to the

manufacturer's protocol. RNA was quantified using minifluorimeter Qubit and the RNA Assay Kit (Invitrogen). The RNA was converted into complementary deoxyribonucleic acid (cDNA) using the Superscript II kit (Invitrogen) and random primers according to the manufacturer's instructions. Five hundred nanograms of total RNA were used in the reaction.

Quantitative RT-PCR

Quantitative real-time PCR (qRT-PCR) was performed using the EVA Green kit (Syntol) and CFX96 system (BioRad). qRT-PCR reactions were performed for 40 cycles with preliminary incubation at 95°C for DNA-polymerase activation. Each cycle consisted of the following steps: denaturation at 95°C (30s), annealing at 57-59°C (30 s), and elongation at 72°C (45 s). The annealing temperature varied depending on the primer's melting temperature (primers are listed in *table 1* of supplemental materials). Fluorescence detection in a Fam channel and primary processing of the results were performed automatically using the system software. Samples were run in triplicate and normalized to *Gapdh*.

DNA methylation study by bisulphite sequencing

DNA was isolated and purified from frozen cells with the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich G1N70) by the method recommended by the supplier. After DNA quantity measurement with a Qubit fluorometer (Invitrogen), it was sodium bisulphite modified with the EpiTect Bisulfite Kit (Qiagen) by the method recommended by the supplier. The primers for PCR amplification of modified DNA were constructed with the aid of the online service BiSearch (<http://bisearch.enzim.hu/?m=search>). The list of the primers used is in *table 2* of supplemental materials. The target regions were chosen to contain a transcription initiation site proximal to CpG islands or immediately preceding the transcription initiation site for genes devoid of CpG islands. Two-step PCR amplification was carried out using 2 μ l of the primary PCR mix as a matrix for the second PCR step and changing one of the primers used in primer PCR for a more internal one. The real-time detection system DT-322 (DNA-Technologiya, Moscow, Russia) and qPCRmix-HS SYBR+ROX kit (Evrogen, Moscow, Russia) were used. Amplification conditions were 95°-5 min for DNA initial denaturation and Taq DNA polymerase, 40 cycles of [95°-30 sec - 52-56° (T_m minus 3°)-30 sec - 72°-45 sec], final elongation - 72°-2 min. The final PCR mixes were fractionated by 2% agarose gel electrophoresis, discrete bands of the expected length were quickly cut out under "long" (312 nm) UV light, and DNA was extracted with the GenElute Gel Extraction Kit

(Sigma-Aldrich NA1111-1KT). Sequencing of the PCR fragments obtained was done with an ABI PRISM® BigDye™ Terminator v. 3.1 kit on an Applied Biosystems 3730 DNA Analyzer. The results of the visualization and their export to a fasts format were done with the Sequence Scanner Software v1.0 (<http://www.appliedbiosystems.com/absite/us/en/home/support/software-community/free-ab-software.html>). DNA methylation patterns were obtained with the aid of the online service Meth Tools 2.0 (http://194.167.139.26/methtools/MethTools2_submit.html) [22].

Histone H3 methylation analysis

ChIP grade Abcam antibodies (H3K4me3 – ab1012, H3K9me3 – ab8898, H3K27me3 – ab6002) and the Imprint Chromatin Immunoprecipitation Kit (Sigma-Aldrich CHP1) were used for chromatin immunoprecipitation and subsequent fractionation by the method recommended by the supplier of the kit. The resulting methylated H3-associated DNA fractions were whole-genome amplified with the REPLI-g UltaFast Mini kit (Qiagen 150033) by the method recommended by the supplier. The relative content of the target gene sequences was estimated by quantitative PCR using the real-time detection system DT-322 (DNA-Technologiya, Moscow, Russia) and the qPCR-mix-HS SYBR+ROX kit (Evrogen, Moscow, Russia). 50 ng of each methylated H3-associated DNA fraction was used as a matrix for PCR amplification. The primers were chosen to amplify the target gene sequences immediately preceding the transcription initiation site (the list of primers is in *table 3* of supplemental materials). All the data obtained were normalized to those of the DNA samples obtained from the aliquots of input unfractionated chromatin.

Urea production analysis

Analysis of the cell's ability to produce urea was performed under 3D culture conditions in a collagen gel. The collagen gel was prepared using collagen of rat tails dissolved in 0.1% sterile acetic acid to a concentration of 5 mg/ml.

Before the gel preparation, all materials were cooled down to +4°C. The components were mixed in the following order: 0.34 M NaOH (Sigma) to a concentration of 0.023 mM, 7.5% Na₂CO₃ (PanEco) to a concentration of 0.26%, 10× DMEM (Sigma) to a concentration of 1×, glutamine (Gibco) to a concentration of 2 mM, HEPES (Gibco) to a concentration of 1%, fetal bovine serum (HyClone) to a concentration of 10%, and then the collagen in acetic acid to a concentration of collagen of 2%. The cell suspensions in a small volume of PBS were added in the last to a concentration of cells of 1×10⁶ per 1 ml of the gel. The mixture was stirred and placed into

35 mm Petri dishes (2 ml per dish). The Petri dishes were kept in a CO₂ incubator at 37°C for 30 min until complete polymerization of the gel. After the gel polymerization, 2 ml of the culture medium was added into each dish. The gel was placed then into a CO₂ incubator (zero hour of the gel preparation) and kept under standard culture conditions.

The amount of produced cell urea was determined in the culture medium using the Urea Assay Kit (Bio-Vision) in accordance with the manufacturer's recommendations. The samples of the medium were collected on days 1, 5, and 10 of cell incubation in the gel. The medium was changed to fresh 24 hours prior to sampling. The amount of urea was determined by the intensity of the chromogenic reaction on the Start Fax 2100 microplate reader (Awareness Technology Inc).

Statistical analysis

All experiments were performed using cell cultures obtained from three animals with at least three repeats in each culture. Data were analyzed using Student's t-test. A p value of 0.05 was considered significant.

RESULTS

Immunocytochemical and qRT-PCR analyses confirmed the increased expression of hepatocyte markers and decreased expression of ductal markers in SGC after VPA exposure

On the 15th day of hepatic differentiation, the cells were analyzed by immunostaining with antibodies to alpha-fetoprotein (a marker of undifferentiated liver), albumin (a marker of hepatocytes), and cytokeratin 19 (a marker of ductal cells), as well as by qRT-PCR with the primers listed in *Table 1*. The differentiated SGC pretreated with VPA were named SGC-VPA-diff, and differentiated SGC without VPA exposure were named SGC-diff. For the control of cell differentiation, we used undifferentiated first-passage SGC and LPC.

In the first-passage SGC and LPC cultures, expression of alpha-fetoprotein, albumin (weakly), and cytokeratin 19 were observed (*Fig. 1*). In LPC, cytokeratin 19 was localized near the nucleus and in SGC it was also detected near the cell membrane.

In differentiated SGC, an increase in alpha-fetoprotein and albumin expression was observed; what is more, in SGC-VPA-diff, the albumin expression level was higher (*Fig. 1*). The change of cellular localization of the ductal marker cytokeratin 19 occurred during hepatic differentiation: in SGC-diff, as well as in SGC-VPA-diff, the loss of cytokeratin 19 localization near the cell membrane was observed. Furthermore, in the SGC-VPA-diff culture a decrease in the cytokeratin 19 expression level was detected.

The qRT-PCR analysis confirmed the results obtained using immunocytochemistry. For a more convenient interpretation of the results, we considered the gene expression levels in first-passage SGC to be “one” (after the normalization to *Gapdh* in each culture) and the expression levels of corresponding genes in other cultures were measured relative to those determined in the first-passage SGC (*Fig. 2, 3*).

In SGC-diff, a significant increase in the expression of early differentiation markers (*Aat*, *Afp*) was observed (*Fig. 2*). The expression of the hepatocyte markers *G6p*, *Pepck*, *Tat*, and *Cyp3a13* also increased, but there was not significant change in *Alb* and *Tdo* ex-

pression. In SGC-VPA-diff, the expression increase in early differentiation markers (*Aat*, *Afp*) is lower than in SGC-diff; however, the expression level of *Alb* and *Tdo* increased significantly compared to SGC-diff (*Fig. 2*). The expression of ductal marker *Krt19* is 4.6-fold lower in SGC-VPA-diff than in the SGC culture. The expression level of ductal cytochrome *P450 7a1* also increased during hepatic differentiation; but in SGC-VPA-diff, to a lesser degree (*Fig. 2*).

The expression of liver-enriched transcription factors was also analyzed. In the first-passage LPC culture, the mRNA expression levels of the hepatocyte nuclear factors *Hnf-3 β* , *Hnf-4 α* , and *Hnf-6* are about

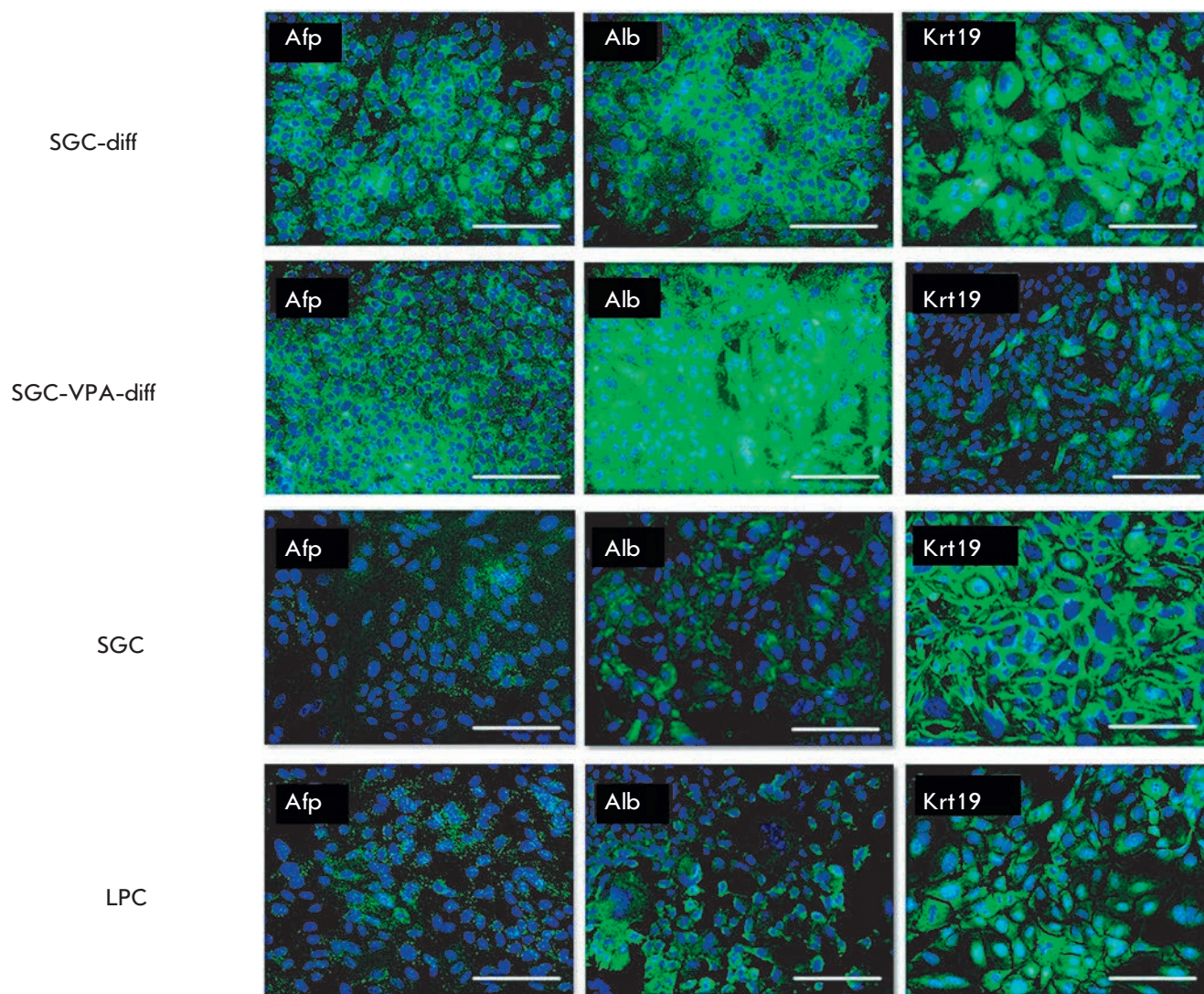


Fig. 1. Immunocytochemical analysis of hepatic differentiation, fluorescent microscopy. Cell nuclei were stained with DAPI (blue), the antigens were detected with Alexa Fluor 488-conjugated antibodies (green), scale bars = 100 μ m

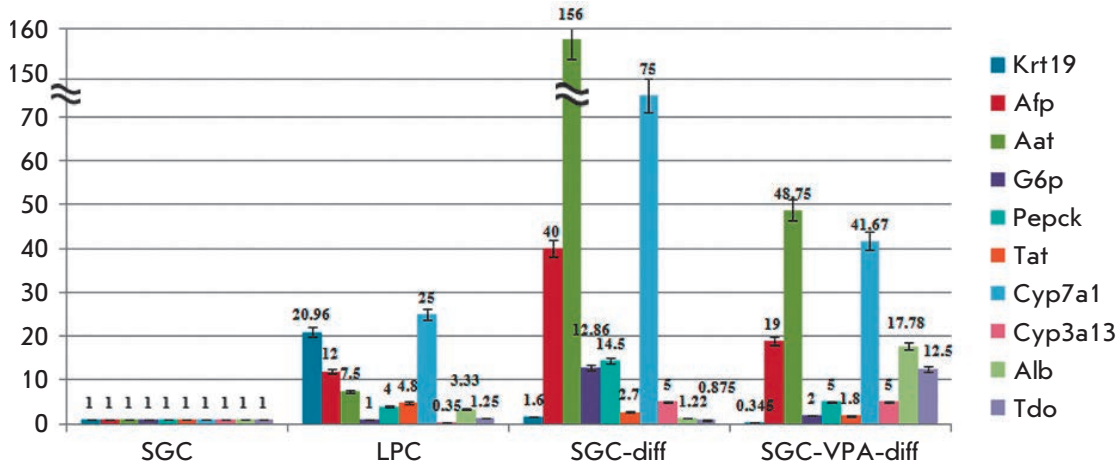


Fig. 2. qRT-PCR analysis of hepatic cell markers gene expression

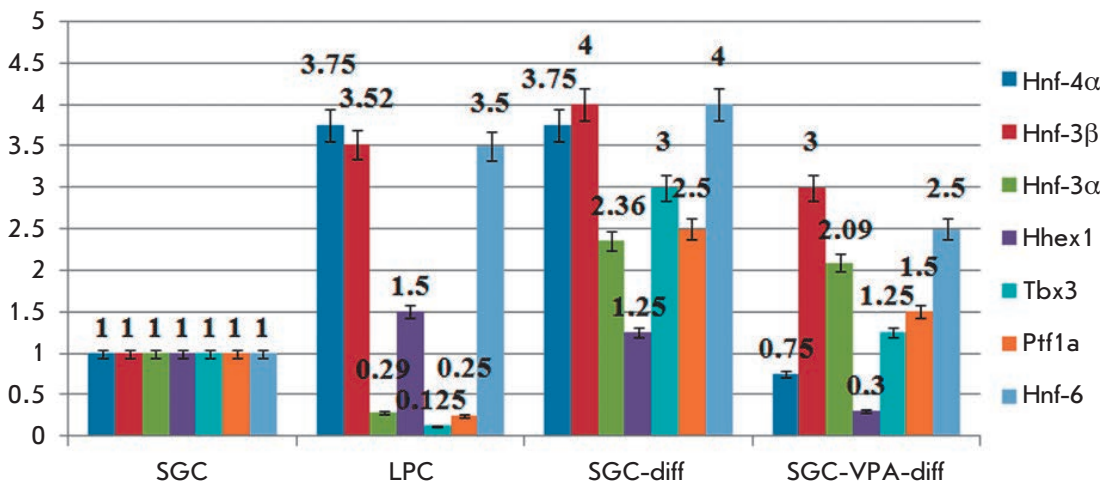


Fig. 3. qRT-PCR analysis of liver-enriched transcription factors gene expression

3.5-fold higher than in SGC, but *Hnf-3α* expression is about 3.4-fold higher in the SGC culture (Fig. 3). After hepatic differentiation of SGC, the gene expression levels of hepatocyte nuclear factors (*Hnf-3β*, *Hnf-4α* and *Hnf-6*) increased to values observed in the LPC culture. However, the expression levels of the early hepatic differentiation gene (*Tbx3*), as well as the transcription factors involved in ductal (*Hhex1*) and acinar (*Ptf1a*) formation, increased too in SGC-diff (Fig. 3). Valproic acid had an ambiguous effect on the transcription factors' gene expression: on the one hand, there was a decrease in the *Hhex1* expression level; furthermore, *Tbx3* and *Ptf1a* expression increased slightly. On the other hand, *Hnf-4α* expression did not increase in SGC-VPA-diff (Fig. 3).

In general, expression of the transcription factors required for hepatic differentiation and early hepatocyte markers (*Afp* and *Aat*) is significantly increased after differentiation of mouse submandibular salivary gland

cells. Increase in later differentiation markers expression (*G6p*, *Pepck*, *Tat*, *Alb*, and *Tdo*) is less pronounced. Thus, under 2D cultivation conditions, initiation and initial stages of hepatic differentiation occur. The effect of VPA treatment on the SGC differentiation is ambiguous: on the one hand, reduction of the ductal markers expression (*Krt19* and cytochrome *P450 7a1*) and increase in the hepatocyte markers expression – *Alb* and *Tdo* is observed, which may indicate an increase in differentiation specificity. On the other hand, the expression levels of the early hepatic markers in SGC-VPA-diff are usually lower than in SGC-diff. It is possible that SGC-VPA-diff is at a later stage of hepatic differentiation, which is characterized by a low level of early differentiation markers expression. Furthermore, it is known that the *Afp* and *Aat* genes normally express in salivary glands cells, and that the expression of *Aat* in ducts of the salivary glands increases during their development and differentiation [23, 24]. Initially,

the SGC culture consists of undifferentiated cells. During hepatic differentiation later differentiation stages typical of the salivary gland may emerge. Thus, the observed difference in *Afp* and *Aat* gene expression for SGC-diff and SGC-VPA-diff may be a result of differentiation, characteristic for salivary glands. In this case, the decrease in the expression of these genes in SGC-VPA-diff may indicate increased specificity of hepatic differentiation and decreased lineage commitment inherent to salivary glands.

DNA methylation analysis revealed no significant changes in the methylation of the promoter regions of liver-enriched genes in salivary gland cells

The DNA methylation pattern of liver-enriched genes was analyzed for SGC, LPC, and SGC-VPA-diff cultures at the first passage by bisulfite sequencing. CpG-islands near the transcription initiation point were analyzed; and in the case of island absence – in the area immediately preceding the transcription initiation point. In most cases, the methylation of these areas is most important for gene expression. In a graphic form, generated by the online service Meth Tools 2.0, the DNA methylation pattern is shown in *Figure 4*.

It was shown that the DNA methylation pattern is similar in all three cultures. Part of the *Gata4* gene CpG-island, located in the promoter region, is virtually not methylated. Methylated cytosines were located below the transcription initiation point. It is possible that the *Gata4* gene is in active or preactivated state in all three cultures. *Gata6* is hardly methylated, and it is possible that this gene is active or is in the preactivated state in all cultures. The promoter region of the *Hnf-1 α* gene is strongly methylated in all cultures, but this area is not a CpG island but sporadic CpG-site. So it is un-

clear how methylation affects *Hnf-1 α* expression. The *Hnf-3 β* gene is strongly methylated in the same manner in all three cultures; therefore, we can talk about its stable repression. But the CpG-island located at the beginning of the coding sequence is not in the promoter region. There are cases when such methylation does not prevent transcription. The *Alb* gene is methylated quite and almost to the same degree in all cultures, which may indicate its stable repression in all three cell cultures. Another possibility is that methylation of this area may be unessential for gene transcription.

Thus, for the most investigated genes no significant differences in DNA methylation patterns were observed in all cell cultures. Apparently, specific transcription control of these genes is carried out in the cells at the expense of other epigenetic modifications (possibly, histone modifications).

Valproic acid treatment and subsequent differentiation of salivary gland cells changes the histone H3 methylation in the chromatin areas associated with liver-enriched genes

Histone methylation plays one of the most important roles in the epigenetic regulation of transcription. Analyses of histone H3 methylation were performed on the following positions: H3K4me3 – signal most clearly correlating with the promoter transcriptional activity; H3K9me3 – signal correlating with inactivation of genes by a heterochromatinization mechanism; and H3K27me3 – signal by which the polycomb repressive complex 2 (PRC2) produces an inactivating effect on the gene activity.

It was shown that histone methylation of the early endoderm genes *Gata4* and *Gata6* is generally similar in SGC cultures (*Fig. 5*). Compared to the control LPC, in first-passage salivary gland cells, H3 histone of these genes is methylated higher on the H3K9me3 position, which indicates their heterochromatin inactivation. At the same time, in SGC, histone methylation is present on the H3K4me3 position for both genes. The low expression levels of *Gata4* and *Gata6* in SGC detected by the analysis of gene expression across the transcriptome [18] suggest that inactivating methylation in the H3K9me3 position in these cells is dominant. The histone methylation level of the *Gata4* and *Gata6* genes in the H3K9me3 position is much lower in LPC, and methylation in the H3K27me3 position is virtually absent. These results correlate with the relatively higher expression of these genes in liver progenitor cells, as shown by gene expression analysis across the transcriptome. The H3K9me3 histone methylation for *Gata4* and *Gata6* genes is greatly reduced in SGC-diff and SGC-VPA-diff. At the same time, H3K27me3 histone methylation of these genes slightly increases in

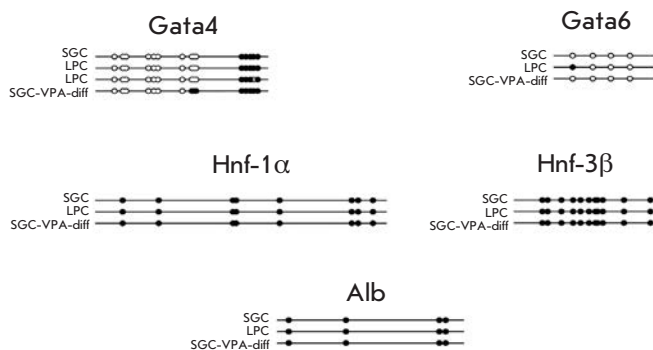


Fig. 4. Analysis of promoter region DNA methylation pattern of liver-enriched genes. Empty circles correspond to nonmethylated CpG dinucleotides, solid circles – to methylated

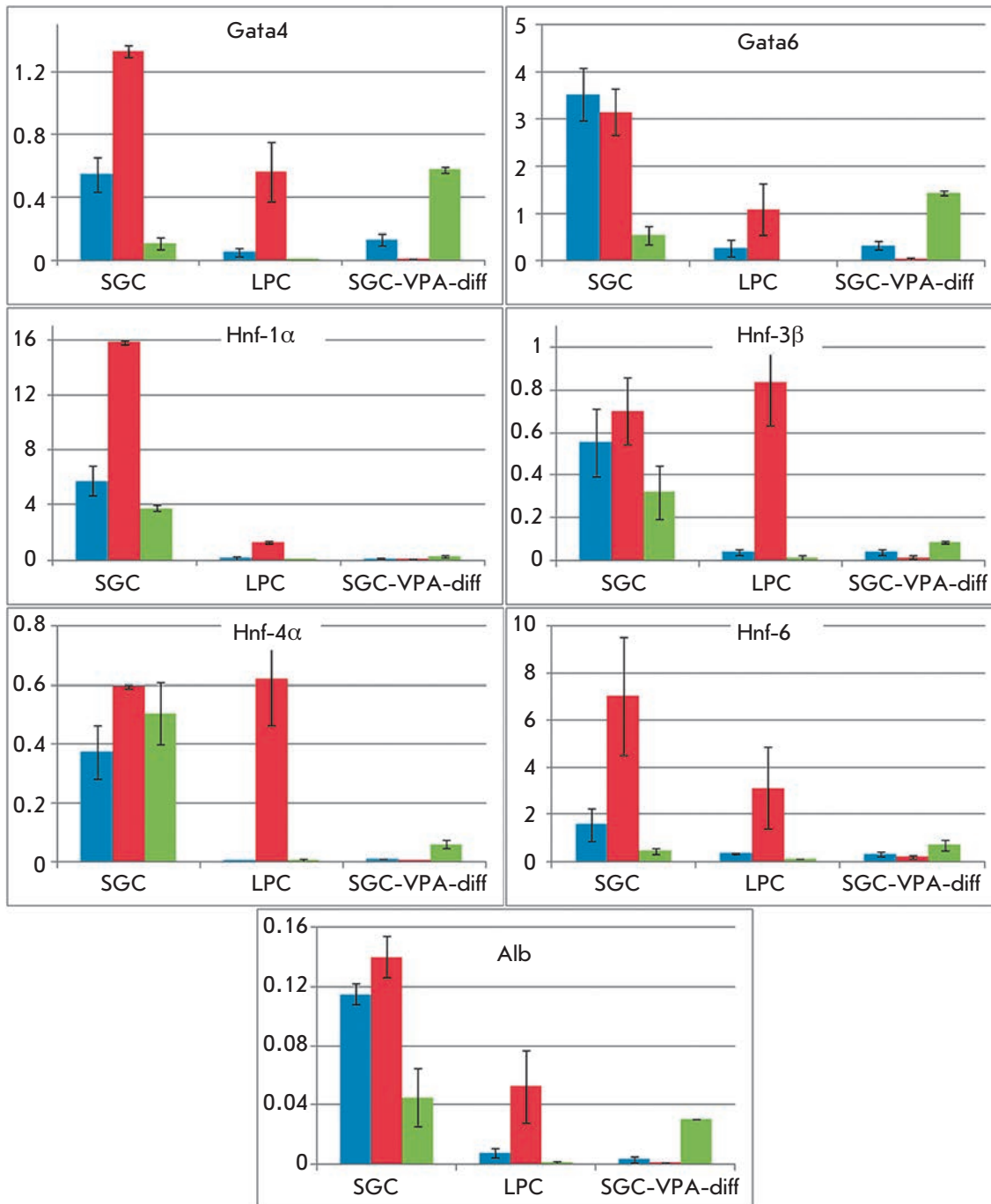


Fig. 5. Analysis of the histone H3 methylation in the cell cultures at the first passage. Blue color shows H3K4me3, red color – H3K9me3, green color – H3K27me3 position

SGC-VPA-diff. This can indicate the transcriptional activation of these genes after differentiation. However, it is also possible that in salivary gland cells secondary inactivation of these genes appears after differentiation.

The *Hnf-1α* transcription factor is strongly inactivated in SGC at the first passage by the heterochromatinization mechanism. In addition, there is increased H3K27me3 methylation compared to LPC. H3K9me3 and H3K27me3 methylation are almost completely ab-

sent for this gene after differentiation. The H3K4me3 methylation level in SGC-VPA-diff is equivalent to the level in the LPC culture at the first passage. This may indicate transcription activation of this gene after hepatic differentiation.

The hepatocyte nuclear factors *Hnf-3β* and *Hnf-4α* demonstrate similar histone methylation features. In the first-passage SGC, the H3 histone methylated in the H3K9me3 and H3K27me3 positions; however, the level of H3K9me3 methylation for the *Hnf-3β* gene is

slightly lower compared to LPC. These data correlate with the results obtained by RT-PCR and gene expression analysis across the transcriptome, according to which in the first-passage SGC *Hnf-3 β* expression is present. The H3K9me3 methylation level for the *Hnf-4 α* gene in SGC is equivalent to the level in LPC; however, the H3K27me3 methylation level is much higher in SGC. H3K9me3 methylation for the *Hnf-3 β* and *Hnf-4 α* genes is almost absent after differentiation, whereas H3K27me3 methylation persists at a low level. This could mean activation of these genes, which correlates with the RT-PCR results, according to which the *Hnf-3 β* expression level in SGC-VPA-diff increases 3-fold in comparison to SGC.

For the *Hnf-6* gene in SGC a high level of H3K9me3 methylation is typical. Methylation at this position is largely absent after differentiation, but it slightly increases the H3K27me3 methylation level, which could mean a secondary inactivation of this gene.

In first-passage SGC, the H3 histone of the *Alb* gene is methylated in the H3K9me3 and H3K27me3 positions. H3K9me3 methylation is almost absent, and the H3K27me3 methylation level decreases after differentiation. The H3K4me3 methylation level in SGC-VPA-diff is equivalent to that in LPC. These results could indicate transcriptional activation of this gene. According to RT-PCR results, the *Alb* mRNA expression level is low in first-passage SGC and increases 18-fold after differentiation.

Thus, these results indicate that the VPA treatment and following differentiation procedure affect the mechanisms of genome epigenetic regulation. DNA methylation of gene promoter regions differs very little in the studied cell cultures. However, the H3 histone methylation shows significant differences. The early endodermal markers *Gata4* and *Gata6* and hepatocyte nuclear factors *Hnf-1 α* and *Hnf-6* are inactivated by a heterochromatization mechanism in first-passage SGC. The *Hnf-3 β* , *Hnf-4 α* , and *Alb* genes in addition could be inactivated by the polycomb repressive complex 2 (PRC2). After differentiation in SGC, in almost all cases the removal of H3K9me3 methylation occurs. However, for the *Gata4*, *Gata6*, *Hnf-6*, and *Alb* genes secondary inactivation by H3K27me3 methylation is possible. In general, the histone methylation results correlate well with the gene expression data obtained by RT-PCR.

Valproic acid increases urea production in mouse salivary gland cells under 3D cultivation conditions

One of the detoxification functions of the liver is urea synthesis from ammonia, carried out by hepatocytes. Determination of the cell's ability to produce urea is widely used as a test for hepatic differentiation efficiency estimation [25]. To assess the hepatic differenti-

ation efficiency of SGC *in vitro*, we analyzed urea production by cells under 3D cultivation conditions (in the collagen gel). The undifferentiated first-passage SGC and LPC were used as a control.

The first-passage SGC and LPC synthesize almost no urea under 2D cultivation conditions but acquire the ability to produce urea under 3D cultivation conditions. Increased urea production level by the studied cells in the collagen gel during the whole observation period indicates that 3D cultivation conditions promote cell differentiation.

By the 15th day of cell incubation in collagen gel, the level of urea production reaches 24 mM per 1×10^6 cells per 24 hours (Fig. 6). For comparison, freshly isolated mouse hepatocytes produce about 350 mM of urea per 1×10^6 cells per 24 hours. The high level of urea production by SGC is evidence of their considerable potential for hepatic differentiation under certain culture conditions. The LPCs actively synthesize urea under 3D conditions: by the 15th day, the urea production level by PKP is 7.6 fold lower than that of the primary culture of hepatocytes. This indicates the high hepatic differentiation ability of LPC under 3D cultivation conditions.

After hepatic differentiation, the ability of SGC to produce urea increases, and, in the case of VPA treatment, SGCs produce urea at a comparable rate with LPC. Thus, hepatic differentiation affects not only the gene expression of submandibular salivary gland cells, but also the functional characteristics of these cells. Hepatic cell differentiation of SGC yields cells capable of performing some of the functions of hepatocytes. VPA can increase cell differentiation efficiency.

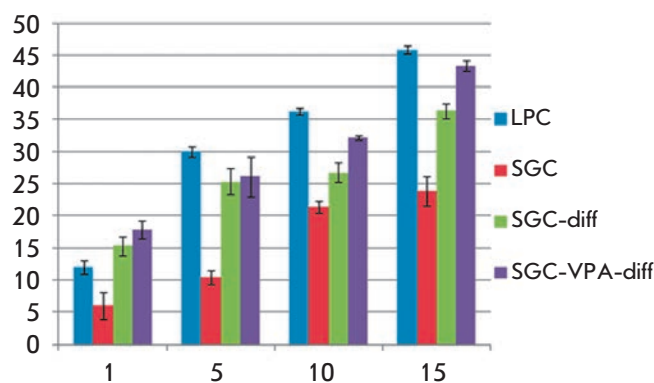


Fig. 6. The urea production analysis by differentiated cells in comparison with first-passage SGC and LPC under 3D cultivation conditions. The X-axis value is the days of cell cultivation in the gel, the Y-axis value is the amount of urea in the medium (mM) per 1×10^6 cells per 24 hours

DISCUSSION

Using small molecules that can affect the epigenetic regulation of gene expression and cause an increase in differentiation efficiency is a promising approach in cell biology. It has been shown that valproic acid increases the differentiation efficiency of various cell types. Dong *et al.* [9] showed that VPA treatment significantly increases the efficiency of hepatic differentiation of mouse ES cells and decreases the extent of their spontaneous differentiation into bile duct structures. The authors suggest that the possible mechanism accelerating the ES cells differentiation could be the transition of the cell cycle in the G0/G1 phase, which occurs after VPA treatment. Deceleration of the cell cycle contributes to the loss of pluripotency and differentiation of ES cells.

VPA can also increase the differentiation efficiency of committed cells. VPA exposure for 72 hours at 5 mM in human bone marrow cells increases H3 and H4 histone acetylation, which promotes DNA demethylation of these cells [10]. During the subsequent hepatic differentiation, human bone marrow cells express albumin and store glycogen more efficiently than cells not exposed to VPA. Differentiated bone marrow MSCs were able to produce urea; what is more, after VPA treatment urea synthesis was about 1.5 fold higher [10]. The authors suggested that the increased differentiation efficiency of bone marrow cells is due to demethylation of the genes involved in hepatic differentiation.

In addition, the histone deacetylase inhibitor - valproic acid can cause active DNA demethylation in a DNA-replication-independent manner [7]. This effect may also increase the transcriptional activity of genes. Furthermore, VPA activates the *Wnt* family genes [26]. It has been shown that the *Wnt*/ β -catenin signaling pathway is required for the formation and differentiation of endodermal cells of the pancreas and liver [27, 28].

Our study of submandibular salivary gland cells in comparison to liver progenitor cells in mice provides insight into the ability of SGC for hepatic differentiation and the effect of valproic acid treatment on the differentiation efficiency. We have shown that the initiation and initial stages of SGC hepatic differentiation is implemented effectively, affecting a wide range of transcription factors and various liver-enriched markers. However, under 2D cultivation conditions, it is quite difficult to achieve terminal differentiation stages. For this reason, we evaluated the efficacy of differentiation under 3D cultivation conditions. Cultivation of cells in a collagen gel promotes an expansion of their morphogenetic and differentiation potential [29, 30]. Under 3D cultivation conditions, SGC-VPA-diff produce urea at a high level, comparable to the control LPC. This indi-

cates an effective hepatic differentiation of SGC, which affects the functional features of these cells.

Hepatic differentiation has almost no effect on the DNA methylation of the genes' promoter regions, which is not surprising given the conservative mechanism of DNA methylation. However, VPA treatment and subsequent differentiation to a large extent influence H3 histone methylation. It was shown that histone methylation in most cases is higher in control SGC and usually lower in differentiated SGC. In general, this can be explained by the progenitor nature of SGC: many genes in progenitor cells have a bivalent configuration, i.e., enriched both in activating and inhibiting histone modifications. This allows them to differentiate into different directions. The histone methylation results correlate well with the results of gene expression in the studied cells.

Most likely, there are several mechanisms of VPA influence on the cell differentiation potential. The efficiency of cell differentiation depends not only on the target cell markers acquisition, but also on the loss of the differentiation features of the initial cell line. As is known from the experience of cell reprogramming, the committed cells acquire the differentiation markers of target cells easier, while the loss of parental cell line markers occurs more slowly [31]. According to our results, VPA can reduce the expression of a number of markers characteristic of initial SGC. One of the possible mechanisms of VPA influence on the differentiation efficiency is that VPA promotes the erasing of the parental cell line epigenetic program that accelerates differentiation. In addition, the possible mechanism of VPA influence on the differentiation efficiency could be H3 histone modifications of target genes and increased accessibility of these genes for the growth factors and cytokines used in the differentiation protocol.

CONCLUSION

The obtained results allow us to conclude that mouse submandibular salivary gland cells show significant phenotypic plasticity and are able to differentiate in the hepatic direction. Valproic acid affects the epigenetic regulation of gene expression by histone modifications and can increase the specificity and efficiency of hepatic differentiation for these cells. The possible mechanism of valproic acid influence on the differentiation efficiency could consist in erasing the parental cell line differentiation features or/and in facilitating the accessibility of target genes for the cytokines and growth factors used during hepatic differentiation.

This work was supported by a grant of the Russian Science Foundation (project #14-50-00029).

Table 1. Primers used in qRT-PCR

Primer	Gene	Nucleotide sequence	Amplicon, bp	Melting temperature, °C
Control				
<i>Gapdh</i>	Glyceraldehyde-3-phosphate dehydrogenase	AGGTCCGGTGTGAACGGATTTG GGGGTCCGTTGATGGCAACA	95	62.6 62.6
Liver-enriched mercers				
<i>Aat</i>	Alpha-1-antitrypsin	CTCGTCCGCTCACTAAACAAG GCTGTCTGAGAGTCAAGGTCTT	248	60.7 61.3
<i>Afp</i>	Alpha-fetoprotein	CCATCACCTTTACCCAGTTTGT CCCATCGCCAGAGTTTTTCTT	101	60.2 60.6
<i>Alb</i>	Albumin	TGCTTTTTCCAGGGGTGTGTT TACTTCCCTGCACTAATTTGGCA	167	62.4 60.2
<i>Krt19</i>	Cytokeratin 19	GGGGGTTTCAGTACGCATTGG GAGGACGAGGTCACGAAGC	113	62,9 62,1
<i>Cyp7a1</i>	Cytochrome P450, family 7, subfamily a, polypeptide 1	AACGGGTTGATTCCATACCTGG GTGGACATATTTCCCATCAGTT	126	62.0 60.0
<i>Cyp3a13</i>	Cytochrome P450, family 3, subfamily a, polypeptide 13	GATTCTTGCTTACCAGAAGGGC GCCGGTTTGTGAAGGTAGAGTA	170	61,0 61,7
<i>G6p</i>	Glucose-6-phosphatase	CGACTCGCTATCTCCAAGTGA GGGCGTTGTCCAAACAGAAT	208	61.0 60.9
<i>Pepck</i>	Phosphoenolpyruvate carboxykinase 1	TGACAGACTCGCCCTATGTG CCCAGTTGTTGACCAAAGGC	153	61.0 61.4
<i>Tat</i>	Tyrosine aminotransferase	AGCCGAATCCGAACAAAACC GCCGATAGATGGGGCATAGC	146	60.9 61.3
<i>Tdo</i>	Tryptophan 2,3-dioxygenase	AATCCATGACGAGCACCTATTCA TCACCTTGAGCATGTTCTCT	140	61.4 60.8
Liver-enriched transcription factors				
<i>Hhex1</i>	Hematopoietically expressed homeobox 1	CGAGACTCAGAAATACCTCTCCC CTGTCCAACGCATCCTTTTTTG	162	61.2 60.0
<i>Hnf-3α</i>	Hepatocyte nuclear factor 3α (Foxa1)	GGAGTTGAAGTCTCCAGCGTC GGGGTGATTAAGGAGTAGTGGG	157	62.4 61.7
<i>Hnf-3β</i>	Hepatocyte nuclear factor 3β (Foxa2)	TCCGACTGGAGCAGCTACTAC GCGCCACATAGGATGACA	176	62.8 61.8
<i>Hnf-4α</i>	Hepatocyte nuclear factor 4α	ATGCGACTCTCTAAAACCCTTG ACCTTCAGATGGGGACGTGT	135	60,0 62,7
<i>Hnf-6</i>	Hepatocyte nuclear factor 6	GCCCTGGAGCAAACCTCAAGT TTGGACGGACGCTTATTTTCC	231	62,4 60,6
<i>Tbx3</i>	T-box transcription factor 3	TGGAACCCGAAGAAGACGTAG TACCCGCTTGTGAAACTGG	84	61.2 62.1
Acinar marker				
<i>Ptf1a</i>	Pancreas specific transcription factor 1a	GCTACACGAATACTGCTACCG CGCAGCAATAGCTGACGTTG	134	60.3 62.0

Table 2. Primers used in bisulphite sequencing

Primer	Nucleotide sequence	Amplicon, bp	Melting temperature, °C
<i>AlbF1</i>	TTGGTAAAGATGGTATGATTTTG	397	58,4
<i>AlbF2</i>	ATTTTGTAATGGGGTAGGAAT	381	58,7
<i>AlbR</i>	ACCACCTAAAAATTCTCAAA		57,3
<i>Hnf-3βF</i>	AATGTGTATTAAAAGGGAGGAAA		60,0
<i>Hnf-3βR1</i>	CCRAACAACCCATTTAAATAATC	378	59,2
<i>Hnf-3βR2</i>	CCCAAAAACCTAAAATCAAA	180	57,9
<i>Gata4F1</i>	TATTGAGAGTAGGGAGGAAAGA	261	60,0
<i>Gata4F2</i>	AGGAAAGAGAAGGAGAATAAATA	247	58,8
<i>Gata4R</i>	CTAACTAACCTAAAAAAATCAC		57,2
<i>Gata6F1</i>	ATTTAGTAGTTTGTAGAGAGTAG	405	57,1
<i>Gata6F2</i>	TTYGATTTATAGTTTGGTATTTT	381	57,5
<i>Gata6R</i>	AATCCCTACAATCTTCTAAA		55,7
<i>Hnf-1αF1</i>	ATAGGGGTTTTTTTTTTTTGGG	373	62,3
<i>Hnf-1αF2</i>	GGGTGTAGTGATTTATTTTA	325	55,3
<i>Hnf-1αR</i>	ACTTTAACTTCAACCTTAC		56,7
<i>Hnf-4αF</i>	TTTGGTTTTTATAGGTATTAGGT		58,9
<i>Hnf-4αR1</i>	CTCTTTCTTTCTTTCTTTCTTTC	399	59,2
<i>Hnf-4αR2</i>	CTTTCTTTCTTTCTTTCTTTCTTTC	365	60,8
<i>Hnf-6F</i>	TTTTTTYGGTTTATTTGTGTTGG		60,1
<i>Hnf-6R1</i>	ATATCTTACCTTCTCTTACT	390	56,8
<i>Hnf-6R2</i>	TTCCCCTCTATCTTTTTTTTTTTC	363	60,6

Table 3. Primers used in histone H3 methylation analysis

Primer	Nucleotide sequence
<i>AlbF</i>	GGGGTAGGAACCAATGAAATG
<i>AlbR</i>	GAGGAGGAGGAGAAAGGTTA
<i>Hnf-3βF</i>	CACCTGCTTGTTGTTTTGAC
<i>Hnf-3βR</i>	AGTCCCTTCCTTTACGTCCA
<i>Gata4F</i>	TTGGGGGAGCTTTGGGAAGA
<i>Gata4R</i>	GGAAAAGAGCAGGGACTCGG
<i>Gata6F</i>	TACCACCACCACCATCACCAT
<i>Gata6R</i>	TCTGATCTTTACCTGTGCTG
<i>Hnf-1αF</i>	TGATGTTGGGCTAGGACTGA
<i>Hnf-1αR</i>	CAATTGGGAGTGAGCAGAAG
<i>Hnf-4αF</i>	AGACAGGGTGGATAGATAGC
<i>Hnf-4αR</i>	GACAGTGTGAGTATGTGTGCAG
<i>Hnf-6F</i>	CCACCACCTACACTACCTTA
<i>Hnf-6R</i>	GGTTATTCATAGAGGCCAGC

REFERENCES

1. Snykers S., Henkens T., De Rop E., Vinken M., Fraczek J., De Kock J., De Prins E., Geerts A., Rogiers V., Vanhaecke T. // *J. Hepatol.* 2009. V. 51. № 1. P. 187–211.
2. Haumaitre C., Lenoir O., Scharfmann R. // *Mol. Cell Biol.* 2008. V. 28. № 20. P. 6373–6383.
3. Vaes B.L., Lute C., van der Woning S.P., Piek E., Vermeer J., Blom H.J., Mathers J.C., Müller M., de Groot L.C., Steegenga W.T. // *Bone.* 2010. V. 46. № 2. P. 514–523.
4. Jeong S.G., Ohn T., Kim S.H., Cho G.W. // *Neurosci. Lett.* 2013. V. 554. P. 22–27.
5. Mike A.K., Koenig X., Koley M., Heher P., Wahl G., Rubi L., Schnürch M., Mihovilovic M.D., Weitzer G., Hilber K. // *Cell Physiol. Biochem.* 2014. V. 33. № 1. P. 205–221.
6. Perucca E. // *CNS Drugs.* 2002. V. 16. P. 695–714.
7. Detich N., Bovenzi V., Szyf M. // *J. Biol. Chem.* 2003. V. 278. № 30. P. 27586–27592.
8. Liu J., Liu Y., Wang H., Hao H., Han Q., Shen J., Shi J., Li C., Mu Y., Han W. // *Sci. Rep.* 2013. DOI: 10.1038/srep01185.
9. Dong X.J., Zhang G.R., Zhou Q.J., Pan R.L., Chen Y., Xiang L.X., Shao J.Z. // *World J. Gastroenterol.*, 2009. V. 15. № 41. P. 5165–5175.
10. Dong X., Pan R., Zhang H., Yang C., Shao J., Xiang L. // *PLoS One.* 2013. DOI: 10.1371/journal.pone.0063405.
11. An S.Y., Han J., Lim H.J., Park S.Y., Kim J.H., Do B.R., Kim J.H. // *Tissue Cell.* 2013. DOI: 10.1016/j.tice.2013.12.006.
12. Shubnikova E.A., Pogodina L.S. // *Ontogenez.* 2000. V. 31. № 6. P. 476–480.
13. Sato A., Okumura K., Matsumoto S., Hattori K., Hattori S., Shinohara M., Endo F. // *Cloning Stem Cells.* 2007. V. 9. № 2. P. 191–205.
14. Baek H., Noh Y.H., Lee J.H., Yeon S.I., Jeong J., Kwon H. // *J. Tissue Eng. Regen. Med.* 2012. DOI: 10.1002/term.1572.
15. Okumura K., Nakamura K., Hisatomi Y., Nagano K., Tanaka Y., Terada K., Sugiyama T., Umeyama K., Matsumoto K., Yamamoto T., Endo F. // *Hepatology.* 2003. V. 38. P. 104–113.
16. Hisatomi Y., Okumura K., Nakamura K., Matsumoto S., Satoh A., Nagano K., Yamamoto T., Endo F. // *Hepatology.* 2004. V. 39. № 3. P. 667–675.
17. Schwarz S., Rotter N. // *Methods Mol. Biol.* 2012. V. 879. P. 403–442.
18. Petrakova O.S., Terskikh V.V., EChernioglo E.S., Ashapkin V.V., Bragin E.Y., Shtratnikova V.Y., Gvazava I.G., Sukhanov Y.V., Vasiliev A.V. // *SpringerPlus.* 2014. DOI: 10.1186/2193-1801-3-183.
19. Petrakova O.S., Gvazava I.G., Ashapkin V.V., Shtratnikova V.Y., Terskih V.V., Sukhanov Y.V., Vasiliev A.V. // *Doklady Biological Sciences.* 2013. V. 453. P. 397–400.
20. Soto-Gutierrez A., Navarro-Alvarez N., Caballero-Corbalan J., Tanaka N., Kobayashi N. // *Acta Med Okayama.* 2008. V. 62. № 2. P. 63–68.
21. Yi F., Liu G.H., Izipisua Belmonte J.C. // *Cell Research.* 2012. V. 22. P. 616–619.
22. Grunau C., Schattevoy R., Mache N., Rosenthal A. // *Nucleic Acids Res.* 2000. V. 28. № 5. P. 1053–1058.
23. Tsuji T., Nagai N. // *Int. J. Dev. Biol.* 1993. V. 37. № 3. P. 497–498.
24. Chi J.G. // *J. Korean Med. Sci.* 1996. V. 11. № 3. P. 203–216.
25. You J., Shin D.S., Patel D., Gao Y., Revzin A. // *Adv. Healthc Mater.* 2014. V. 3. № 1. P. 126–132.
26. Hrebackova J., Hrabeta J., Eckschlager T. *Curr. // Drug Targets.* 2010. V. 11. № 3. P. 361–379.
27. Zaret K.S., Grompe M. // *Science.* 2008. V. 322. № 5907. P. 1490–1494.
28. Engert S., Burtscher I., Liao W.P., Dulev S., Schotta G., Lickert H. // *Development.* 2013. V. 140. № 15. P. 3128–3138.
29. Davydova D.A., Voroteliak E.A., Bragina E.E., Terskikh V.V., Vasiliev A.V. // *Tsitologiya.* 2011. V. 53. № 4. P. 325–331.
30. Chermnykh E.S., Vorotelyak E.A., Gnedeva K.Y., Moldaver M.V., Yegorov Y.E., Vasiliev A.V., Terskikh V.V. // *Histochem. Cell Biol.* 2010. V. 133. P. 567–576.
31. Sekiya S., Suzuki A. // *Nature.* 2011. V. 475. № 7356. P. 390–393.

Recombinant Immunotoxin 4D5scFv-PE40 for Targeted Therapy of HER2-Positive Tumors

E. A. Sokolova^{1,2}, O. A. Stremovskiy², T. A. Zdobnova¹, I. V. Balalaeva^{1*}, S. M. Deyev^{1,2}

¹Lobachevsky State University of Nizhny Novgorod, pr. Gagarina 23, 603950, Nizhny Novgorod, Russia

²Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, 117997, Moscow, Russia

*E-mail: irin-b@mail.ru

Received 21.08.2015

Copyright © 2015 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 Recombinant immunotoxins are extremely promising agents for the targeted therapy of tumors with a certain molecular profile. In this work, we studied the properties of a new recombinant HER2-specific immunotoxin composed of the scFv antibody and a fragment of *Pseudomonas* exotoxin A (4D5scFv-PE40). High affinity of the immunotoxin for the HER2 tumor marker, its selective cytotoxicity against HER2-overexpressing cells, and its storage stability were demonstrated. The 50% inhibitory concentration (IC₅₀) of the 4D5scFv-PE40 immunotoxin for HER2-overexpressing cancer cells was 2.5–3 orders of magnitude lower compared to that for CHO cells not expressing this tumor marker and was 2.5–3 orders of magnitude lower than IC₅₀ of free PE40 for HER2-overexpressing cancer cells. These findings provide a basis for expecting in the long run high therapeutic index values of the 4D5scFv-PE40 immunotoxin for its use *in vivo*.

KEYWORDS recombinant immunotoxin, 4D5scFv, *Pseudomonas* exotoxin A, HER2 tumor marker, targeted therapy. **ABBREVIATIONS** HER2 – human epidermal growth factor receptor 2; scFv – single-chain variable fragment; PE40 – fragment of *Pseudomonas* exotoxin A; MTT – 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide.

INTRODUCTION

The progress achieved in studying the molecular basis of carcinogenesis has revealed subtle biochemical differences between tumor and normal cells and, thus, provided opportunities for developing therapies based on these differences. The targeted therapy concept involves the development of drugs specifically interacting with target molecules that are expressed in tumor cells but are not present in normal tissues. This approach enables elimination of tumor cells with minimum negative impact on other tissues and organs.

Targeted agents include, primarily, monoclonal antibodies specifically interacting with surface receptor tumor markers (including anti-angiogenic antibodies) [1, 2] and low-molecular-weight inhibitors of enzymes [3]. An additional toxic component can be introduced into antibody-based drugs to enhance their tumor-specific effect [4]. The drug Kadcyla® [5], which was introduced into clinical practice for the treatment of metastatic breast cancer at the end of 2013, became the first bifunctional agent. This drug is a chemical conjugate of a full-length humanized antibody specific for the HER2 tumor marker and a toxic compound which

inhibits association of tubulin subunits during microtubule assembly. In the case where both the targeting and effector (toxic) modules are protein molecules, it becomes possible in principle to combine them into a single polypeptide chain by genetic engineering techniques. Recombinant bifunctional proteins, known as immunotoxins, are exclusively promising molecules for further development of targeted cancer treatment due to their strictly controlled composition, the possibility of biotechnological production in bacterial producers, the possibility of optimization of their properties by genetic engineering techniques, etc. [6].

In this study, we investigated the physico-chemical and functional properties of a new HER2-specific recombinant immunotoxin produced on the basis of a scFv format antibody and *Pseudomonas* exotoxin A.

MATERIALS AND METHODS

Preparation and characterization of proteins

Production of the recombinant immunotoxin 4D5scFv-PE40 and free polypeptides 4D5scFv and PE40 (ETA) was carried out in *E.coli* culture that was preliminarily

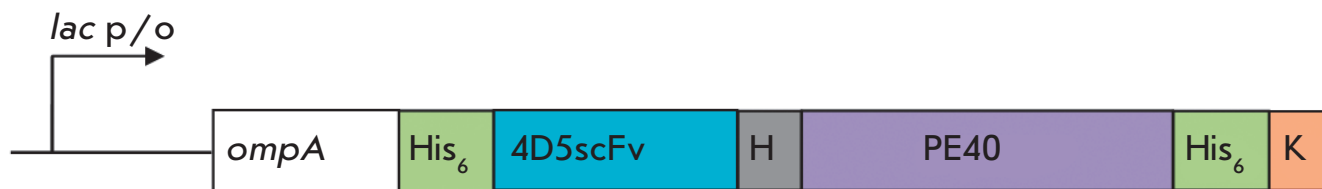


Fig. 1. Scheme of recombinant immunotoxin 4D5scFv-PE40. The following encoding regions are shown: *ompA* (white) is a signal peptide ensuring secretion of the desired recombinant protein to the periplasmic space; His₆ (green) is an oligohistidine peptide; 4D5scFv (blue) is the anti-HER2-antibody 4D5scFv; H (gray) is a flexible hydrophilic linker of the hinge region of mouse IgG (16 a.a.); PE40 (lilac) is a fragment of wild-type exotoxin A from *Pseudomonas aeruginosa* (domains II, Ib, and III); K (orange) is the KDEL oligopeptide

transformed with the plasmids pSD-4D5scFv-PE40, pSD-4D5scFv, and pSD-PE40 containing genes of the 4D5scFv-PE40, 4D5scFv, and PE40 proteins, respectively, under control of the *lac*-promoter. The plasmids were constructed from the vectors pIG6-4D5 and pIG6-4D5MOCB-ETA [7, 8].

Protein purification was performed in two steps by Ni²⁺-chelate affinity chromatography using a 1 mL HisTrap FF column (GE Healthcare, USA) and ion exchange chromatography on a 1 mL Q Sepharose FF column (GE Healthcare, USA).

Fractions containing the desired protein were analyzed by electrophoresis in 12% PAGE under denaturing conditions according to the standard protocol [9].

The dissociation constant of the complex between 4D5scFv-PE40 and the HER2 receptor was determined by surface plasmon resonance on a BIAcore 3000 optical biosensor (GE Healthcare, USA) using a recombinant extracellular domain of the HER2 receptor p185^{HER2-ECD} (Sino Biological, Inc., China).

Analysis of cytotoxicity

In the study, we used the SKOV-3 human ovarian adenocarcinoma cell line (catalog number ATCC-HTB-77), which is characterized by overexpression of the HER2 receptor, and HER2-negative Chinese hamster ovary (CHO) cells (ATCC-CCL-61). To generate the SKOV-kat fluorescent tumor cell line, SKOV-3 cells were transfected with the TurboFP635 red fluorescent protein gene using the pTurboFP635-C vector (Evrogen, Russia) [10].

Cells were grown in a RPMI-1640 medium (HyClone, USA) with 10% fetal calf serum (HyClone, USA) and 2 mM glutamine (PanEco, Russia) at 37 °C under 5% CO₂.

The cytotoxicity analysis of the studied proteins was performed using the standard MTT assay [11]. In this case, treatment of the cells was conducted in two ways. To evaluate the effect of short-term exposure, the cells were incubated in the presence of the studied proteins for 40 min at 4 °C. Then, the unbound proteins were washed out and the cells were added with the growth

Cytotoxicity of proteins 4D5scFv-PE40, PE40, and 4D5scFv

Cell line	IC ₅₀ , nM*					
	Short-term exposure (40 min)			Long-term exposure (72 h)		
	4D5scFv- PE40	PE40	4D5scFv	4D5scFv- PE40	PE40	4D5scFv
CHO	> 1000	> 1000	> 1000	8.7 (5.6–13.6)	2.9 (1.8–4.6)	> 100
SKOV-kat	22** (5.7–85.3)	> 1000	> 1000	0.008** (0.006–0.013)	4.9 (1.3–18.4)	> 100
SKOV-3	–	–	–	0.017** (0.011–0.025)	6.6 (3.1–14.0)	> 100

* The mean and 95% confidence interval are presented.

** Statistically significant difference from CHO.

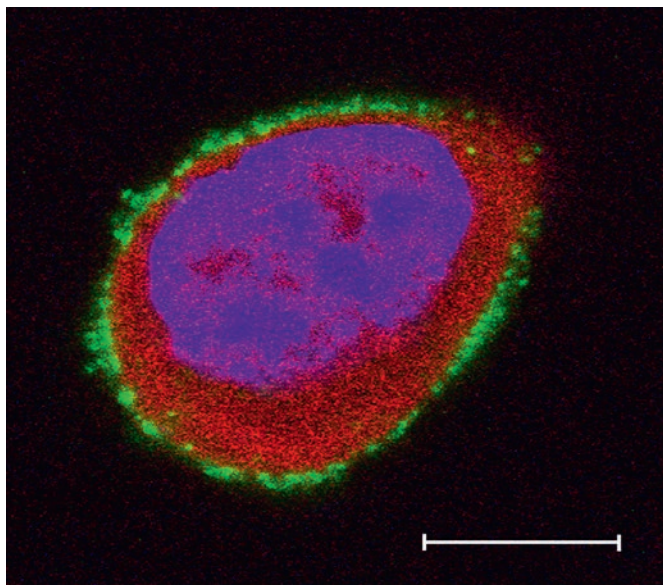


Fig. 2. The SKOV-kat cell expressing the red fluorescent protein TurboFP635 (red). The cell nucleus is counter-stained with Hoechst 33342 (blue). Expression of the HER2 receptor on the cell surface is confirmed by staining with complexes of quantum dots and the anti-HER2 antibody 4D5scFv [21] (green). The bar is 10 μm

medium and grown for 48 h. To analyze the effect of long-term exposure, the cells were grown in the presence of the studied proteins in the medium for 72 h.

The mean value and 95% confidence interval of the protein concentration leading to a 2-fold decrease in the culture viability (IC_{50}) were calculated using the GraphPad Prism 6 software.

RESULTS AND DISCUSSION

The tested immunotoxin 4D5scFv-PE40 is a single polypeptide chain that combines the targeting and toxic modules (*Fig. 1*).

The recombinant protein 4D5scFv-PE40 contains the 40 kDa PE40 fragment of exotoxin A from *Pseudomonas aeruginosa*, lacking the wild-type receptor-recognizing domain. The PE40 fragment is linked to the C-terminus of the anti-HER2-antibody of scFv format (4D5scFv) via a flexible hydrophilic 16-amino acid linker [12]. Because of this linker, the distance between connected fragments of the protein molecule, 4D5scFv and PE40, amounts to 2.5–2.7 nm, which allows the two protein domains to avoid steric hindrance and to retain their functional properties. The 4D5scFv antibody is a recombinant polypeptide composed of fused variable domains of light and heavy chains of the full-length

antibody 4D5 specific for the HER2 tumor marker. The antibody 4D5scFv proved to be an effective targeting agent for generating bifunctional cytotoxic proteins [13–18].

At the C-terminus, the immunotoxin molecule contains the oligopeptide KDEL, which is an endoplasmic reticulum translocation signal. Oligohistidine sequences are fused to both ends of the target protein for its purification by metal-chelate affinity chromatography. The signal peptide ompA provides secretion of the desired recombinant protein to the periplasmic space both for reducing its toxic effect on the cell and for increasing the level of a soluble fraction of the desired protein during biotechnological production in bacterial producers [19].

Successive purification of the recombinant protein by metal-chelate and ion exchange chromatography yielded the immunotoxin 4D5scFv-PE40 ($M_r=71$ kDa) with purity of more than 96%, the stability of which was confirmed by storage at +4 $^{\circ}\text{C}$ for 3 months; this can be regarded as a very good indicator for protein samples. For this period, preservation of homogeneity and high affinity of the immunotoxin 4D5scFv-PE40 for the HER2 receptor extracellular domain ($K_d \sim 7$ nM) was demonstrated. For comparison, K_d of the free antibody 4D5scFv determined also by plasmon resonance is 5.2 nM [20].

Investigation of the functional properties of 4D5scFv-PE40 was performed in SKOV-kat cell line [10] generated by transfection of the SKOV-3 parent cell line (human ovarian adenocarcinoma), which is characterized by overexpression of the HER2 tumor marker, with the TurboFP635 red fluorescent protein gene. Given the possible impact of transfection on the cells, special attention was paid to preservation of cell morphology and phenotype. Preservation of the HER2 receptor overexpression on the surface of the transfected cells was preliminarily confirmed by immunofluorescence analysis using HER2-targeted semiconductor quantum dots [21] (*Fig. 2*).

Analysis of the 4D5scFv-PE40 cytotoxicity under conditions of short-term (40 min) incubation in the cold demonstrated a highly selective toxic effect on SKOV-kat cells (Table). Since these conditions prevent protein internalization, it is obvious that the specific binding to the HER2 receptor and retention on the membrane affect the cell metabolism during subsequent culturing of the cells after immunotoxin removal from the medium.

In the body, long-term presence of a drug in the blood and especially in the intercellular matrix is typical. *In vitro* experiments demonstrated that the selectivity of 4D5scFv-PE40 cytotoxic action on SKOV-kat cells persists even for a 72-h incubation of the drug in the growth medium. The 4D5scFv-PE40 concentration

causing a twofold decrease in the viability of SKOV-kat cells (IC_{50}) is 3 orders of magnitude lower than IC_{50} of 4D5scFv-PE40 for CHO cells not expressing the HER2 receptor. It should be noted that PE40 also exhibits a toxic effect under these conditions; however, IC_{50} of the PE40 polypeptide for SKOV-3 and SKOV-kat cells is also 2.5–3 orders of magnitude higher than IC_{50} of the 4D5scFv-PE40 immunotoxin. These findings provide the basis for expecting in the long run high therapeutic index values of the immunotoxin for its use *in vivo*.

The resistance of SKOV-3, the parent cell line for SKOV-kat, to the action of cytotoxic agents is well-known [22]. Nevertheless, 4D5scFv-PE40 also exhibited toxicity against SKOV-3 in the picomolar concentration range, which was comparable to that against SKOV-kat (Table).

Fluorescent tumor cell lines are unique research tools. Expression of a fluorescent protein by SKOV-kat tumor cells provides an opportunity to continue studying the developed immunotoxin 4D5scFv-PE40 on experimental SKOV-kat-based tumor models using the highly informative methods of intravital whole-body optical imaging.

CONCLUSION

The success of the first targeted anticancer drugs has radically changed the approach to the development of new anticancer agents and led to a change in the standard treatment of many cancers. The biggest success was achieved in oncohematology and the treatment of disseminated tumors; however, the advantage of the targeted approach was also demonstrated in the treatment of solid tumors of a certain molecular profile.

The immunotoxin 4D5scFv-PE40 investigated in this study is designated for targeted therapy of tumors expressing the HER2 tumor marker. The specificity of this immunotoxin is identical to that of the known anticancer drug Herceptin®. At the same time, the presence of a toxic module in 4D5scFv-PE40 multiplies the specific toxicity of the protein against HER2-expressing cells and provides an opportunity to anticipate its effectiveness in further research into the therapeutic potential *in vivo*.

This work was supported by the Ministry of Education and Science of the Russian Federation (project RFMEFI57814X0051).

REFERENCES

- Deyev S.M., Lebedenko E.N., Petrovskaya L.E., Dolgikh D.A., Gabibov A.G., Kirpichnikov M.P. // *Russ. Chem. Rev.* 2015. V. 84 (1). P. 1–26.
- Hojjat-Farsangi M. // *Tumour Biol.* 2015. V. 36. P. 543–556.
- Roskoski R. Jr. // *Pharmacol. Res.* 2015. V. 100. P. 1–23.
- Weiner G.J. // *Nat. Rev. Cancer.* 2015. V. 15. № 6. P. 361–370.
- Lambert J.M., Chari R.V. // *J Med Chem.* 2014. V. 57. P. 6949–6964.
- Alewine C., Hassan R., Pastan I. // *Oncologist.* 2015. V. 20. № 2. P. 176–185.
- Deyev S.M., Waibel R., Lebedenko E.N., Schubiger A.P., Plückthun A. // *Nat. Biotechnol.* 2003. V. 21. P. 1486–1492.
- Di Paolo C., Willuda J., Kubetzko S., Lauffer I., Tschudi D., Waibel R., Plückthun A., Stahel R.A., Zangemeister-Wittke U. // *Clin. Cancer Res.* 2003. V. 9. P. 2837–2848.
- Laemmler U.K., Beguin F., Gujer-Kellenberger G. // *J. Mol. Biol.* 1970. V. 47. P. 69–85.
- Zdobnova T., Sokolova E., Stremovskiy O., Karpenko D., Telford W., Turchin I., Balalaeva I., Deyev S. // *Oncotarget.* 2015. V. 6. № 31. P. 30919–30928.
- Mosmann T. // *J. Immunol. Methods.* 1983. V. 65. P. 55–63.
- Muller K.M., Arndt K.M., Strittmatter W., Pluckthun A. // *FEBS Lett.* 1998. V. 422. P. 259–264.
- Edelweiss E., Balandin T.G., Ivanova J.L., Lutsenko G.V., Leonova O.G., Popenko V.I., Sapozhnikov A.M., Deyev S.M. // *PLOS ONE.* 2008. V. 3. P. e2434.
- Balandin T.G., Edelweiss E., Andronova N.V., Treshalina E.M., Sapozhnikov A.M., Deyev S.M. // *Invest. New Drugs.* 2011. V. 29. № 1. P. 22–32.
- Serebrovskaya E.O., Edelweiss E.F., Stremovskiy O.A., Lukyanov K.A., Chudakov D.M., Deyev S.M. // *Proc. Natl. Acad. Sci. USA.* 2009. V. 106. P. 9221–9225.
- Mironova K.E., Proshkina G.M., Ryabova A.V., Stremovskiy O.A., Lukyanov S.A., Petrov R.V., Deyev S.M. // *Theranostics.* 2013. V. 3. P. 831–840.
- Cao Y., Marks J.W., Liu Z., Cheung L.H., Hittelman W.N., Rosenblum M.G. // *Oncogene.* 2014. V. 33. P. 429–439.
- Zhang M., Qiu Z., Li Y., Yang Y., Zhang Q., Xiang Q., Su Z., Huang Y. // *Appl Microbiol Biotechnol.* 2013. V. 97. P. 3913–3923.
- Pines O., Inouye M. // *Mol Biotechnol.* 1999. V. 12. P. 25–34.
- Sokolova E.A., Zdobnova T.A., Stremovskiy O.A., Balalaeva I.V., Deyev S.M. // *Biochemistry (Mosc).* 2014. V. 79. № 12. P. 1376–1381.
- Balalaeva I.V., Zdobnova T.A., Krutova I.V., Brilkina A.A., Lebedenko E.N., Deyev S.M. // *J. Biophotonics.* 2012. V. 5. № 11–12. P. 860–867.
- Morimoto H., Safrin J.T., Bonavida B. // *J. Immunol.* 1991. V. 147. P. 2609–2616.

A Model System in S2 Cells to Test the Functional Activities of *Drosophila* Insulators

M. Tikhonov, N. B. Gasanov, P. Georgiev, O. Maksimenko

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

E-mail: mog@genebiology.ru

Received 19.03.2015

Copyright © 2015 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 Insulators are a special class of regulatory elements that can regulate interactions between enhancers and promoters in the genome of high eukaryotes. To date, the mechanisms of insulator action remain unknown, which is primarily related to the lack of convenient model systems. We suggested studying a model system which is based on transient expression of a plasmid with an enhancer of the *cop* transposable element, in *Drosophila* embryonic cell lines. We demonstrated that during transient transfection of circle plasmids with a well-known *Drosophila* insulator from the *gypsy* retrotransposon, the insulator exhibits in an enhancer-blocking assay the same properties as in *Drosophila* stable transgenic lines. Therefore, the *Drosophila* cell line is suitable for studying the main activities of insulators, which provides additional opportunities for investigating the functional role of certain insulator proteins.

KEYWORDS insulator, *cop* enhancer, Su(Hw), enhancer transcription, *hsp70* promoter.

ABBREVIATIONS S2 – *Drosophila* embryonic cell line; Sg4 – S2-derived cell line; bp – base pairs; *hsp70* promoter – *hsp70* gene promoter; SV40 – simian virus 40 polyadenylation signal.

INTRODUCTION

In cells of higher eukaryotes, an enhancer can activate a promoter at a distance of up to several hundred kilobase pairs [1–3]. The investigation of insulators may make a significant contribution to the understanding of the mechanisms of long-range interactions between regulatory elements. Insulators are regulatory elements capable of blocking the interaction between an enhancer and a promoter when located between them [4–7]. However, insulators do not directly affect the activity of the enhancer and promoter; i.e., the promoter can be activated by another enhancer, and the enhancer can activate another promoter. Recently, it became obvious that many insulator proteins provide specific interactions between distant regulatory elements and the structural domains of chromosomes [1].

Model systems derived from mammalian [8] and *Drosophila* [9–11] cell lines play an important role in the study of transcription factors acting as part of insulators. One of the problems in developing a convenient model system for the investigation of insulators is the relatively small number of described enhancers that are able to function effectively in *Drosophila* cell cultures.

An enhancer from the *cop* retrotransposon was previously shown to activate a promoter of the heat shock protein 70 gene in S2 cells from *Drosophila melanogaster*, having embryonic origin [10]. The 150 bp enhancer is located immediately after a 5'-long terminal repeat of the *cop* retrotransposon (Fig. 1A) and contains a 28 bp duplication at the 3'-end [12, 13]. The duplicated sequence comprises two copies of a TTGTGAAA octanucleotide in the inverted orientation. Three similar octanucleotides are located in the 5'-region of the enhancer. *Copia*-elements are known that contain an enhancer with only one 28 bp sequence and have a significantly reduced transcriptional activity. It is assumed that the TTGTGAAA sequence binds to a transcription factor, which determines the enhancer activity. Several transcription factors were also isolated that preferentially bind to the 5'-region of the enhancer and can both activate and inhibit transcription [13–15].

This work provides a detailed analysis of the *cop* enhancer in a model system that is used to test insulators in the *Drosophila* cell culture. The adequacy of the model system based on transient expression of circular plasmids in *Drosophila* cell cultures was studied using

an insulator localized in the regulatory region of the *Drosophila gypsy* retrotransposon [4–7]. Previously, the basic properties of regulatory elements of this class were described with an example of this insulator using model systems based on *Drosophila* stable transgenic lines. The present work demonstrates that all basic properties of the *gypsy* insulator are reproduced during transient expression of a circular plasmid in a *Drosophila* cell culture.

MATERIALS AND METHODS

Development of constructs

Plasmids pGL3basic and pGL3enhancer (Promega) were used as initial vectors. The *hsp70* gene promoter (–203... + 253 bp relative to the transcription start) was amplified with *D. melanogaster* genomic DNA and inserted at the restriction sites HindIII and EcoRI into the pGL3basic and pGL3enhancer vectors. The 168 bp *copia* enhancer was amplified with *D. melanogaster* genomic DNA and inserted into the pGL3basic and pGL3enhancer vectors (*he* construct) downstream of the polyadenylation signal at the BamHI restriction site. Constructs e_d and e_r were prepared by inserting the amplified *copia* enhancer upstream of the luciferase gene coding region. In the constructs e_dh and e_rh , the *copia* enhancer was cloned into the *h* vector upstream of the promoter, at the SmaI restriction site. In the case of the constructs $g_d e_d h$, $g_r e_d h$, $g_d e_r h$, $g_r e_r h$, $e_d g_d h$, $e_d g_r h$, $g_d e_d g_d h$, $g_r e_d g_d h$, $e_d s_d h$, $e_d s_d g_d h$, $e_d s_d g_r h$, and $e_d g_d s_d h$, a sequence of regulatory elements was first constructed on the basis of the pBluescript vector, and then the sequence was transferred to the *h* vector at the SmaI restriction site, upstream of the promoter. The *gypsy* insulator (from MDG4 retrotransposon) was a 450-bp fragment previously amplified in our laboratory. The SV40 virus polyadenylation signal was cut out from the pAc5.1hisB vector (Invitrogen) at restriction sites BamHI and SalI. In the case of constructs $he_d g_d$, $he_d g_r$, $hg_d e_d$, $hg_d e_d g_d$, and $hg_d e_d g_r$, a set of regulatory elements was also assembled in the pBluescript vector and transferred to the *h* vector at the BamHI restriction site, upstream of the polyadenylation signal. In the $g_d hg_d e_d$ construct, regulatory elements were inserted at the restriction sites SmaI and BamHI, upstream and downstream of the transcription unit, respectively.

Cell culturing and transfection

The *Drosophila* S2 cell culture was grown in a SFX medium (HyClone) at 25 °C. Cells were transfected with a Cellfectin II reagent (Invitrogen) according to the manufacturer's recommendations (about 8×10^5 cells per transfection). Two hours before transfection, the cells

were put into wells of a 12-well plate. 0.5 µg of DNA was used for one transfection. In all cases, co-transfection of the tested constructs (the firefly luciferase gene was used as a reporter gene) and a control construct (the jellyfish luciferase gene was under control of the actin gene promoter at a 1 : 19 ratio) was performed. The cells were harvested 48 h after transfection.

RNA extraction and reverse transcription

RNA was isolated from S2 cells using a TRI-reagent (Ambion) according to the manufacturer's recommendations. The isolated total RNA was purified from genomic DNA using a Turbo DNA-free reagent kit (Ambion). 1–5 µg of a RNA sample was mixed with a hexamer randomized primer (with a final concentration of 1–5 µM), heated to 70 °C, incubated for 5 min, and rapidly cooled in ice. Then, dNTPs at a concentration of 0.5 mM, buffer for reverse transcriptase, 5 units of the SUPERase-In RNase inhibitor (Ambion), and 60 units of ArrayScript Reverse Transcriptase (Ambion) were added. The reaction mixture was incubated at 42 °C for 2 h, then the enzymes were inactivated by heating to 95 °C for 5 min.

Quantitative real-time PCR

Quantitative real-time PCR was carried out in cDNA samples. Simultaneously, at least three independent reactions with each primer pair for each of three independently collected samples were conducted. Relative amounts of DNA were determined by $\Delta\Delta C_t$. Fragments of the *γTub37C* and *rpl32* genes were used as an endogenous control. The following primer pairs were used in the study:

tub (gctttccaagaagctcataca and gggttcagtcggtattatccag),
 rpl32 (gttcgatccgtaaccgatgt and ccagtcggatcgatgctaa),
 Fluc (ttgtccaacacccaacat and ttccgtgctccaaaacaaca),
 Rluc (cagtggtgggccagatgtaaacaa and taatacaccgcgctactggctcaa).

Dual luciferase assay

The dual luciferase assay was performed using a Firefly & the Renilla Luciferase Assay Kit (Biotium) according to the manufacturer's protocol. The measurement was conducted on a microplate analyzer with a sensitivity of 100 and 1 s exposure time.

RESULTS AND DISCUSSION

Activity of an enhancer of the *copia* mobile element depends on the *Drosophila* cell line

It was previously shown [10] that the *copia* enhancer (Fig. 1A) could cause a more than 100-fold increase in

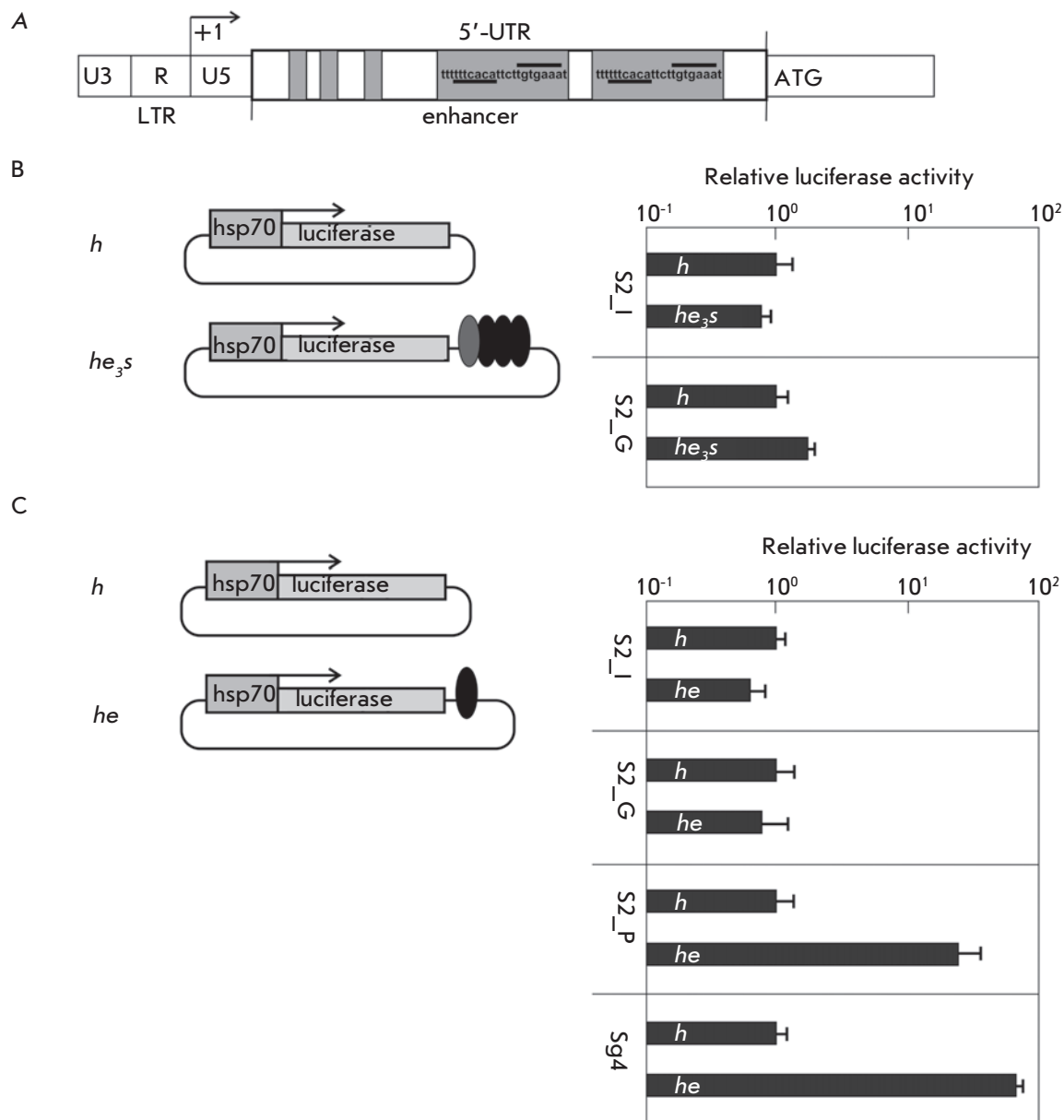


Fig. 1. A – schematic diagram of an enhancer from the *copia* retrotransposon. The enhancer is located in the 5'-untranslated region (5'-UTR). LTR – long terminal repeat. +1 – transcription start. ATG – start codon. Gray rectangles denote octanucleotide repeats. B – results of an analysis of the activity of an element consisting of three copies of the *copia* enhancer (black ovals) and one copy of the SV40 enhancer (gray oval) located at the 3'-end of the firefly luciferase reporter gene (gray box), which is under control of the *hsp70*-promoter (gray rectangle with an arrow). The control construct *h* and tested construct *he_{3s}* were transfected into two variants of S2 cells (S2_I and S2_G). The histogram presents, in a logarithmic scale, the firefly luciferase to jellyfish luciferase activity ratio. All data were normalized relative to the control construct *h*. The standard deviations were calculated on the basis of measurements of four biological replicates. C – analysis of the activity of one copy of the *copia* enhancer (black oval) located at the 3'-end of the firefly luciferase reporter gene. The control *h* and tested construct *he* were transfected into four variants of S2 cells (S2_I, S2_G, S2_P, and Sg4). The histogram presents, in a logarithmic scale, the firefly luciferase to jellyfish luciferase activity ratio. All data were normalized relative to the control construct *h*. The standard deviations were calculated on the basis of measurements of four biological replicates

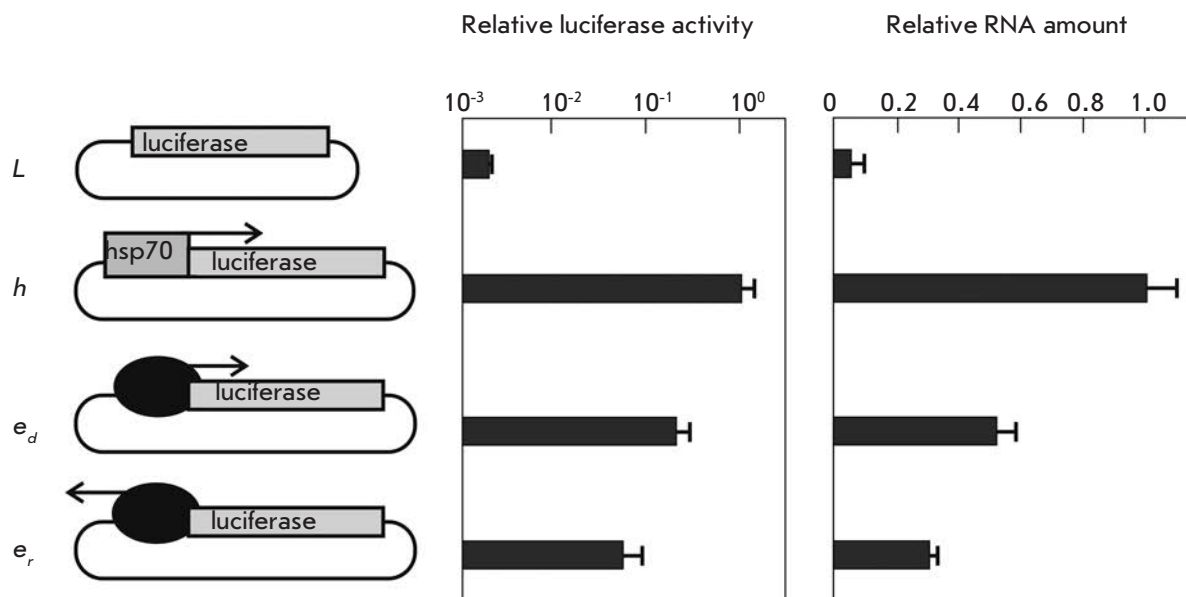


Fig. 2. Results of the analysis of the promoter activity of the *copia* enhancer (black oval). The arrow shows the enhancer orientation. Plasmids *L* (negative control, no promoter) and *h* (positive control with the *hsp70*-promoter) were used as a control. The left histogram presents, in a logarithmic scale, the firefly luciferase to jellyfish luciferase activity ratio. All data were normalized relative to the control construct *h*. The standard deviations were calculated on the basis of measurements of four biological replicates. The right histogram presents, in a linear scale, the relative amount of RNA transcribed from the firefly luciferase gene. All data were normalized relative to the expression levels of the *rpl32*, *tub*, and luciferase jellyfish genes. The standard deviations were calculated on the basis of measurements of four biological replicates

transcription from the promoter of the heat shock protein 70 (*hsp70*) gene of a plasmid transfected into S2 cells. However, according to [16], the *copia* enhancer does not stimulate transcription in S2 cells and its activity is detected only in the DH-33 cell line derived from *Drosophila hydei*.

The first possible explanation for these contradictory results was an assumption that the construct used in [10] contained additional regulatory elements that might have increased the *copia* enhancer activity in S2 cells. Indeed, an expression vector contained three copies of the *copia* enhancer at the 3'-side of the firefly luciferase reporter gene, which was controlled by the minimal *hsp70*-promoter (Fig. 1B). The SV40 (s) enhancer was located near *copia* enhancer copies (e₃) and could also participate in the stimulation of transcription [10].

To study the role of the complex organization of the enhancer region in the stimulation of transcription, we compared the activity of this construct (*he₃s*) and that of a construct containing only a promoter (*h*) in S2 cells from two different sources (Fig. 1B). One cell line was maintained in our laboratory (2S_G), and the second line was received from Invitrogen (2S_I). Surprisingly,

a complex element consisting of three *copia* enhancers and the SV40 enhancer was found not to stimulate the *hsp70*-promoter in any of the tested cell lines (Fig. 1B). Thus, the complex enhancer does not stimulate transcription in S2 cells.

These results may be explained by the differences in the set of transcription factors that are expressed in S2 cell lines independently cultivated for a long time. To test this assumption, a new vector was generated that contained only one copy of the *copia* enhancer downstream the reporter gene (Fig. 1C). We used two additional cell lines: S2_P (line used in the MODEncode project) and Sg4 (received from Pirrotta's laboratory, Rutgers University, USA). The Sg4 line is derived from the S2 line and differs from S2 in the expression profile of several genes.

On the basis of transfection of the control and tested plasmids into four cell lines, the *copia* enhancer was found to cause an approximately 80- to 100-fold increase in *hsp70*-promoter transcription in the Sg4 and S2_P lines, but not to have a stimulatory potential in two previously used S2 lines. Thus, one copy of the *copia* enhancer can efficiently stimulate transcription only in certain types of S2 cells.

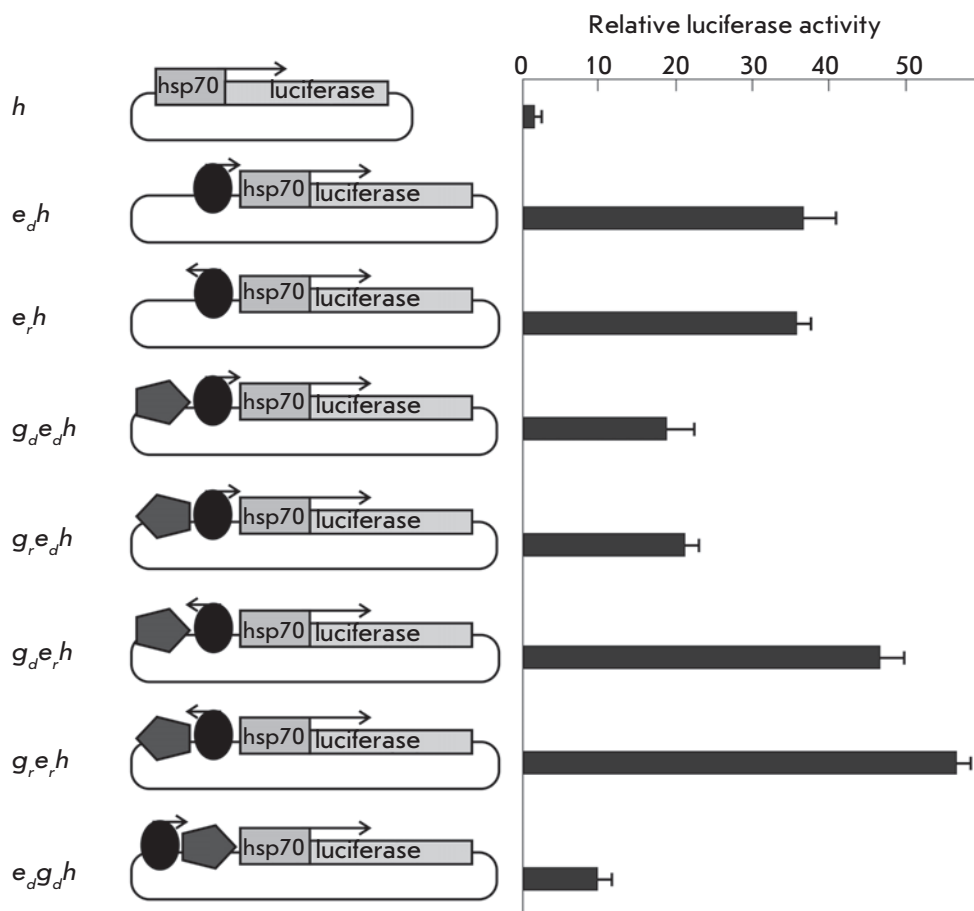


Fig. 3. Effect of the *gypsy* insulator on the *copia* enhancer activity. The results of the analysis of the activity of combinations of the enhancer (oval) and the insulator (pentagon) located upstream of the *hsp70* promoter are shown. The *copia* enhancer orientation is indicated by an arrow, and the insulator orientation is indicated by pentagon pointing. The histogram presents the firefly luciferase to jellyfish luciferase activity ratio. All data were normalized relative to the control construct *h*. The standard deviations were calculated on the basis of measurements of four biological replicates

The *copia* enhancer induces bidirectional transcription with an efficiency comparable to the *hsp70* promoter baseline activity

In study [10], it was shown that a complex regulatory element consisting of SV40 and *copia* enhancers induces bidirectional transcription. Currently, there is abundant data showing that transcription initiation occurs on most enhancers [2, 3]. Short unstable non-polyadenylated transcripts are most often transcribed from enhancers. Usually, the transcripts are not transported into the cytoplasm and not translated. So, we decided to test the *copia* enhancer ability to induce transcription. Some enhancers were previously shown to be capable of producing full-length mRNAs [2, 3]. Therefore, the *copia* enhancer ability to produce polyadenylated and translated RNA was studied.

For this purpose, a construct was generated where the *copia* enhancer was introduced in the direct or reverse orientation instead of the *hsp70*-promoter, upstream of the firefly luciferase reporter gene (Fig. 2). Plasmids with/without the *hsp70*-promoter upstream of the reporter gene were used as a control. These plasmids were used to transfect Sg4 cells. It was shown that the *copia* enhancer was able to start bidirectional

transcription and expression of the luciferase, but 5–20 times less intensively compared to the construct with the *hsp70*-promoter. In the direct orientation, the *copia* enhancer acts as a promoter which is about 3 times stronger than in the reverse orientation. Thus, the *copia* enhancer can act as a weak bidirectional promoter inducing the formation of functional mRNA, which is used as a template for luciferase synthesis. The level of transcripts synthesized from the *copia* enhancer and *hsp70*-promoter were compared by reverse RNA transcription, followed by quantitative PCR. Transcription from the promoter was found to be only 2–3 times more efficient than transcription from the enhancer. Thus, one copy of the *copia* enhancer can trigger bidirectional synthesis of RNA molecules suitable for passing through translation stages, and at a level comparable to the baseline activity of the *hsp70*-promoter.

An insulator from the *gypsy* retrotransposon has little effect on the activity of the *copia* enhancer located prior to a promoter

The strongest *Drosophila* insulator consisting of 12 binding sites of the Su(Hw) protein is located in the regulatory region of the *gypsy* retrotransposon [17–19].

The activity of the insulator in *Drosophila* transgenic lines depends on tested enhancers and promoters. For example, one copy of the insulator completely blocks the activity of *yellow* gene enhancers, but it has almost no influence on the *white* gene enhancer activity [20, 21]. By means of transfection of a circular plasmid into S2 *Drosophila* cells, it was shown [9] that one copy of the *gypsy* insulator placed before a reporter gene promoter causes a two-fold reduction in the activity of the *cop* enhancer introduced into the 3'-side of the gene. The two-fold reduction may be explained by the insulator influence on both the enhancer activity and the promoter located nearby. For example, the Su(Hw) protein is detected not only on an insulator, but also on the sequences of the *cop* enhancer and *hsp70*-promoter in the transfected constructs [22].

To determine the element whose activity is affected by the insulator, we used a series of constructs with the enhancer at position -233 bp relative to the transcription start from the *hsp70*-promoter (Fig. 3). The enhancer was placed in two orientations: direct (e_d) and reverse (e_r). The reporter gene expression level in transfected Sg4 cells was not dependent on the enhancer orientation. The insulator *gypsy* (*g*) was located immediately before the enhancer, in either direct or reverse orientation. Thus, four constructs were prepared in which the enhancer and insulator were placed in different orientations relative to each other and to the promoter. All the constructs were used to transfect Sg4 cells (Fig. 3). Determining the luciferase expression level demonstrated that the insulator orientation in constructs where the enhancer and the promoter had opposite orientations relative to each other did not affect, or slightly increased, the reporter gene expression level. In cases where the enhancer had a direct orientation, the reporter gene expression level was reduced approximately 2 times in the presence of the insulator in either orientation. Therefore, the insulator can affect the activity of the neighbor enhancer, which is located in close proximity to the promoter. In this case, the mechanism of influence is not associated with inhibition of interaction between the enhancer and the promoter. It is most likely that this orientation-dependent transcription inhibition is due to the direct interaction of proteins associated with the insulator and enhancer, which is consistent with the data on the distribution of insulator proteins [10].

The next task was to study the influence of the insulator on the expression level of the reporter gene at a position between the enhancer and promoter. For this purpose, we prepared a construct with the insulator inserted in position -233 bp relative to the transcription start of the *hsp70*-promoter (Fig. 3). The enhancer was located immediately before the insulator, in the direct

orientation; i.e., the insulator was located between the enhancer and the promoter. In this case, the insulator reduced the enhancer activity by about 4 times. Thus, the insulator interposed between the enhancer and the promoter causes stronger inhibition of reporter gene transcription compared to the case where the insulator is upstream of the enhancer. This result is consistent with the basic property of insulators – the ability to block an enhancer – which is implemented when an insulator is interposed between an enhancer and a promoter.

Two insulator copies surrounding an enhancer completely inactivate the enhancer activity

The obtained results demonstrate that one copy of an insulator is only capable of partially blocking the enhancer activity in a transient model based on circular plasmids. Previously, we demonstrated that only two *gypsy* insulator copies surrounding either the enhancer or the *white* reporter gene are able to completely block the enhancer activity in *Drosophila* transgenic lines [21]. According to the model, the interaction between insulators leads to the formation of a chromatin loop, which greatly complicates interactions among the protein complexes associated with enhancers and promoters. To determine whether this rule of functioning of insulators works in the transient model based on a circular plasmid, two additional constructs were generated in which the enhancer located before the *hsp70*-promoter was surrounded by two insulators arranged in one or opposite directions (Fig. 4A). The reporter gene expression level in both variants was found to be close to the level of a control plasmid which contained only the *hsp70*-promoter. Thus, two insulators surrounding an enhancer lead to complete inactivation of its activity, which is consistent with the results obtained previously in transgenic *Drosophila* lines.

In the above-mentioned experiments, insulators surrounding the enhancer were located near the promoter. A question arises as to whether the effect of full enhancer inhibition is retained if an enhancer surrounded by insulators located at a considerable distance from the promoter. To answer this question, a number of constructs were generated in which the *cop* enhancer was inserted in the direct orientation at position +2,230 bp relative to the transcription start of the firefly luciferase reporter gene (Fig. 4B). At this position, the enhancer stimulated reporter gene transcription about twice more efficiently than at the position before the promoter. The reporter gene expression level in two plasmids in which the insulator was in the direct/reverse orientation relative to the 3'-side of the enhancer was close to the expression of a plasmid containing an enhancer only. Thus, the insulator located after the

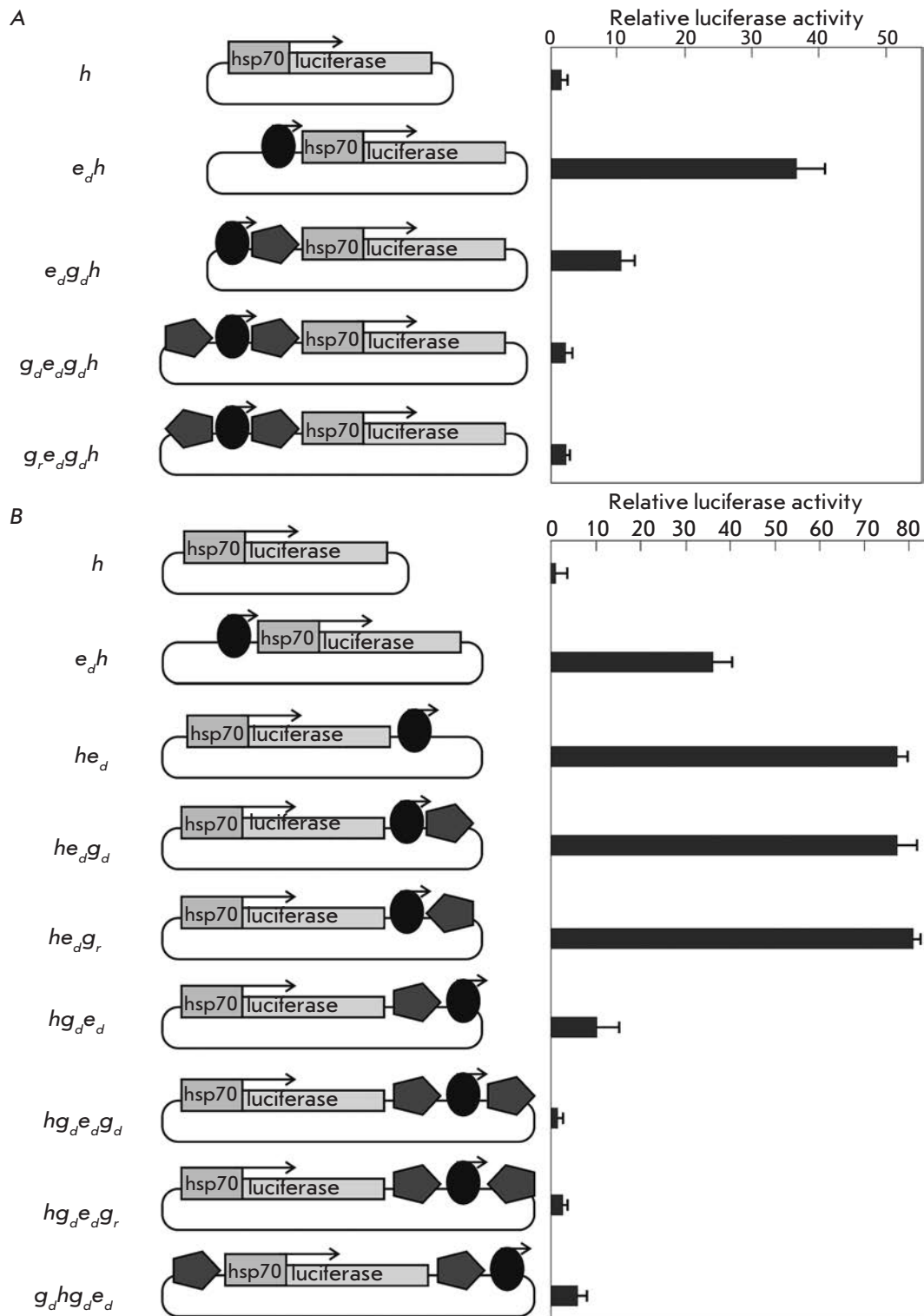


Fig. 4. Effect of two *gypsy* insulator copies surrounding an enhancer or a reporter gene. Results of the analysis of combinations of two insulator copies (pentagon) in different orientations, which surround the enhancer (oval) and are located upstream of the *hsp70* promoter (A) or at the 3'-end of the reporter gene, are shown (B). The *copia* enhancer orientation is indicated by an arrow, the insulator orientation is indicated by pentagon pointing. The histogram presents the firefly luciferase to jellyfish luciferase activity ratio. All data were normalized relative to the control construct *h*. The standard deviations were calculated on the basis of measurements of four biological replicates

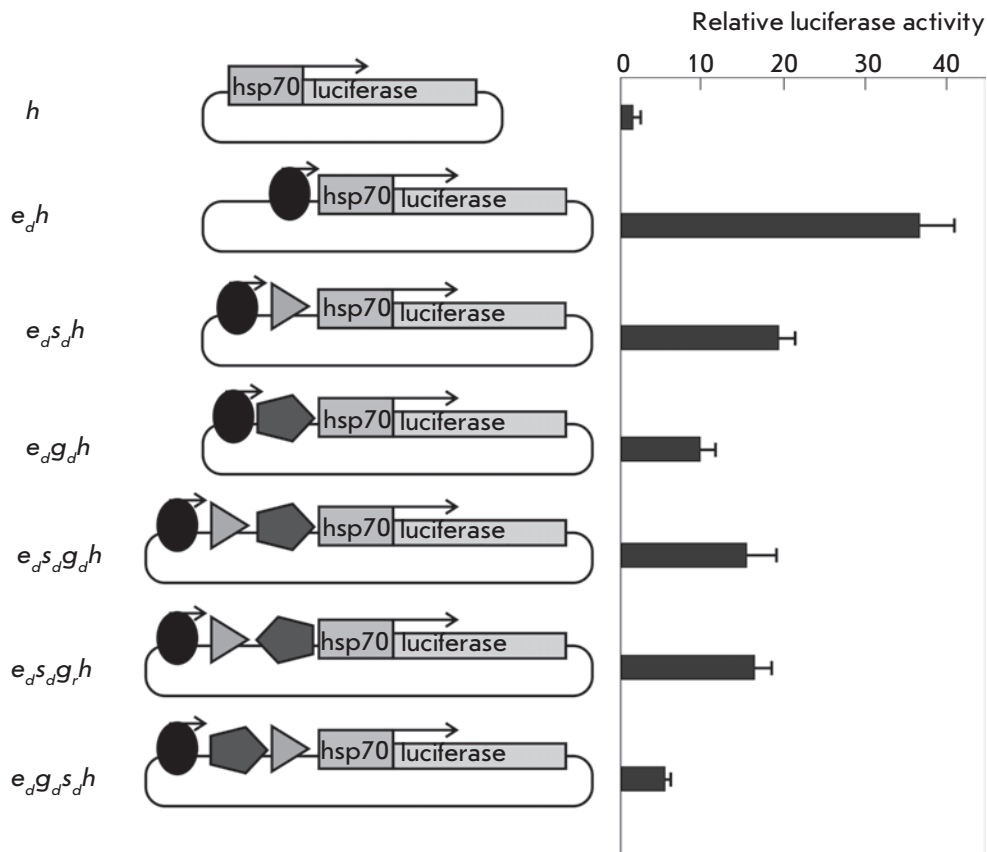


Fig. 5. Effect of transcription from an enhancer on the *gypsy* insulator activity. Results of the analysis of combinations of the insulator (pentagon) in different orientations, enhancer (oval), and SV40 virus transcription terminator (triangle) located upstream of the *hsp70* promoter are shown. The *copia* enhancer orientation is indicated by an arrow; the insulator orientation is indicated by pentagon pointing; the SV40 terminator orientation is indicated by triangle pointing. The histogram presents the firefly luciferase to jellyfish luciferase activity ratio. All data were normalized relative to the control construct *h*. The standard deviations were calculated on the basis of measurements of four biological replicates

enhancer did not affect its activity. However, when the insulator was located between the reporter gene and the enhancer, a six-fold reduction in the reporter gene expression level occurred. Thus, the mutual arrangement of the insulator and the enhancer relative to the promoter even in a circular plasmid determines the efficiency of transcription inhibition. In the next series of constructs, the enhancer was inserted between two unidirectional or bidirectional insulators (Fig. 4B). Transient transfection of these plasmids into Sg4 cells caused reporter gene expression at the level of a control plasmid containing the *hsp70*-promoter only. Thus, two insulator copies surrounding the enhancer completely block its activity. Therefore, the distance between the enhancer and the promoter does not affect efficiency in blocking the enhancer interposed between a pair of insulators.

In transgenic *Drosophila* lines, two insulator copies surrounding the reporter gene caused weaker inhibition of the enhancer activity than two insulator copies surrounding the enhancer [21]. To further test the degree of correlation of the results obtained in circular plasmids and transgenic *Drosophila* lines, we used a construct in which insulators surrounded the reporter gene, and the enhancer was located immediately after

the insulator, on the 3'-side of the gene. Sg4 cells transfected with this plasmid were detected with a weak enhancer activity, which is consistent with the assumption that insulators in this configuration are unable to completely block the enhancer. Complete enhancer inactivation was observed only when two insulators were located immediately next to the enhancer. Therefore, a complete correlation between the results obtained in transgenic *Drosophila* lines and in a transient model in Sg4 cells was found.

Transcription from an enhancer regulates the *gypsy* insulator activity

Previously, it was assumed [10, 23, 24] that transcription helps an enhancer move along chromatin in search of a promoter. According to this model, an insulator blocks promotion of the enhancer together with RNA polymerase II towards the promoter. Transcription, which is initiated on the enhancer, may also directly affect the activity of the promoter and the insulator.

To investigate the functional role of transcription initiated on an enhancer, we generated a number of plasmids with the 220 bp SV40 virus universal polyadenylation signal used to terminate transcription. In the first plasmid, the SV40 terminator was inserted

between the *copia* enhancer, which was positioned in direct orientation, and the *hsp70*-promoter (Fig. 5). Transfection of Sg4 cells with the plasmid caused a 2-fold reduction in the reporter gene expression level compared to a plasmid containing the enhancer only. This result may be partly explained by the fact that transcription initiated from the enhancer contributes to the reporter gene expression. The SV40 terminator stops this transcription and, thus, reduces the reporter gene expression level. However, we demonstrated above (Fig. 2) that luciferase expression from the *copia* enhancer is about 5 times lower than that from the *hsp70*-promoter. Therefore, the main possible explanation is related to the fact that the SV40 terminator is able to partially block the interaction between the enhancer and the promoter by stopping the movement of RNA polymerase II from the enhancer to the promoter. This interpretation is consistent with a model in which RNA polymerase II plays a certain role in signal transmission from the enhancer to the promoter [10, 24, 25].

In the other two plasmids, the enhancer was inserted in direct orientation relative to the promoter and was separated from the promoter by the SV40 terminator and the insulator, which was inserted in direct or reverse orientation (Fig. 5). When Sg4 cells were transfected with any of these plasmids, the reporter gene expression remained at the same level as that of a plasmid containing the SV40 terminator only. Interestingly, transcription in the presence of a combination of the insulator and the terminator reached a higher level compared to a plasmid containing the insulator only. Thus, the SV40 terminator partially suppresses the inhibitory activity of the insulator instead of the expected additive negative effect of the insulator and SV40 terminator on the reporter gene expression. When Sg4 cells were transfected with a plasmid with a reversed order of the insulator and the terminator, whereby the insulator occurred between the enhancer and the terminator, a reduction in the reporter gene expression

level was observed (Fig. 5). These data suggest that transcription from the enhancer increases the insulator activity, which leads to more effective inhibition of the enhancer.

CONCLUSION

The data obtained in this work suggest that embryonic *Drosophila* cell lines with a common origin differ in their expression levels of the transcription factors necessary for the functioning of the *copia* enhancer. Apparently, expression of other genes encoding transcription factors not essential for maintaining the cell line can vary in embryonic cell lines. Thus, cell lines, even with a common origin, can greatly vary in their sets of transcription factors and, as a consequence, in the functional activity of regulatory elements.

We developed a model system that makes it possible to study the activity of insulators in *Drosophila* embryonic cell lines. In the circular plasmid-based transient model, the most well-known insulator *gypsy* retains its basic properties described using model systems based on transgenic *Drosophila* lines [25]. One copy of the insulator blocks only partially the enhancer activity, whereas two copies surrounding either an enhancer or a reporter gene cause complete inactivation of the enhancer.

Recently, our laboratory demonstrated that transcription via an enhancer inhibits its activity [26]. In the present study, we found that the *copia* enhancer has the properties of a weak bidirectional promoter, and transcription from the enhancer can increase the enhancer-blocking activity of the MDG4 insulator. Indeed, there is data according to which binding of transcripts to the Su(Hw) complex can regulate insulator activity [27, 28].

This work was supported by the Russian Science Foundation (project № 14-24-00166).

REFERENCES

- Maksimenko O., Georgiev P. // *Front. Genet.* 2014. V. 5. P. 28.
- Spitz F., Furlong E.E. // *Nat. Rev. Genet.* 2012. V. 13. P. 613–626.
- Erokhin M., Vassetzky Y., Georgiev P., Chetverina D. // *Cell. Mol. Life Sci.* 2015. V. 72. P. 2361–2375. doi: 10.1007/s00018-015-1871-9.
- Chetverina D., Aoki T., Erokhin M., Georgiev P., Schedl P. // *BioEssays.* 2014. V. 36. P. 163–172.
- Kyrchanova O., Georgiev P. // *FEBS Lett.* 2014. V. 588. P. 8–14.
- Herold M., Bartkuhn M., Renkawitz R. // *Development.* 2012. V. 139. P. 1045–1057.
- Matzat L.H., Lei E.P. // *Biochim. Biophys. Acta.* 2014. V. 1839. P. 203–214.
- Fedoseeva D.M., Tchurikov N.A., // *Dokl. Acad. Nauk.* 2013. V. 451. P. 339–343.
- Bohla D., Herold M., Panzer I., Buxa M.K., Ali T., Demmers J., Krüger M., Scharfe M., Jarek M., Bartkuhn M., Renkawitz R. // *PLoS One.* 2014. V. 9. P. e107765.
- Tchurikov N.A., Kretova O.V., Moiseeva E.D., Sosin D.V. // *Nucl. Acids Res.* 2009. V. 37. P. 111–122.
- Dai Q., Ren A., Westholm J.O., Serganov A.A., Patel D.J., Lai E.C. // *Genes Dev.* 2013. V. 27. P. 602–614.
- McDonald J.F., Matyunina L.V., Wilson S., Jordan I.K., Bowen N.J., Miller W.J. // *Genetica.* 1997. V. 100. P. 3–13.
- Wilson S., Matyunina L.V., McDonald J.F. // *Gene.* 1998. V. 209. P. 239–246.
- Cavarec L., Jensen S., Casella J.F., Cristescu S.A., Heid-

RESEARCH ARTICLES

- mann T. // *Mol. Cell. Biol.* 1997. V. 17. P. 482–494.
15. Cavarec L., Heidmann T. // *Nucl. Acids Res.* 1993. V. 21. P. 5041–5049.
16. Cavarec L., Jensen S., Heidmann T. // *Biochem. Biophys. Res. Commun.* 1994. V. 203. P. 392–399.
17. Holdridge C., Dorsett D. // *Mol. Cell Biol.* 1991. V. 11. P. 1894–1900.
18. Geyer P.K., Corces V.G. // *Genes Dev.* 1992. V. 6. P. 1865–1873.
19. Mazo A.M., Mizrokhi L.J., Karavanov A.A., Sedkov Y.A., Krichevskaja A.A., Ilyin Y.V. // *EMBO J.* 1989. V. 8. P. 903–911.
20. Maksimenko O., Golovnin A., Georgiev P. // *Mol. Cell. Biol.* 2008. V. 28. P. 5469–5477.
21. Kyrchanova O., Maksimenko O., Stakhov V., Ivlieva T., Parshikov A., Studitsky V.M., Georgiev P. // *PLoS Genet.* 2013. V. 9. P. e1003606.
22. Smirnov N.A., Didych D.A., Akopov S.B., Nikolaev L.G., Sverdlov E.D. // *Biochemistry (Mosc)*. 2013. V. 78. P. 1141–1150.
23. Zhao H., Dean A. // *Nucl. Acids Res.* 2004. V. 32. P. 4903–4919.
24. Zhu X., Ling J., Zhang L., Pi W., Wu M., Tuan D. // *Nucl. Acids Res.* 2007. V. 35. P. 5532–5544.
25. Ling J., Ainol L., Zhang L., Yu X., Pi W., Tuan D. // *J. Biol. Chem.* 2004. V. 279. P. 51704–51713.
26. Erokhin M., Davydova A., Parshikov A., Studitsky V.M., Georgiev P., Chetverina D. // *Epigenetics Chromatin.* 2013. V. 6. P. 31.
27. King M.R., Matzat L.H., Dale R.K., Lim S.J., Lei E.P. // *J. Cell Sci.* 2014. V. 127. P. 2956–2966.
28. Nazer E., Lei E.P. // *Curr. Opin. Genet. Dev.* 2014. V. 25. P. 68–73.

Profiling of *Mycoplasma gallisepticum* Ribosomes

G. Y. Fisunov, D. V. Evsyutina, A. A. Arzamasov, I. O. Butenko, V. M. Govorun

Scientific research institute of physical-chemical medicine, Malaya Pirogovskaya, 1a, 119435, Moscow, Russia

E-mail: herr.romanoff@gmail.com

Received 16.06.2015

Copyright © 2015 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 development of high-throughput technologies is increasingly resulting in identification of numerous cases of low correlation between mRNA and the protein level in cells. These controversial observations were made on various bacteria, such as *E. coli*, *Desulfovibrio vulgaris*, and *Lactococcus lactis*. Thus, it is important to develop technologies, including high-throughput techniques, aimed at studying gene expression regulation at the level of translation. In the current study, we performed proteomic profiling of *M. gallisepticum* ribosomes and identified high abundant noncanonical proteins. We found that binding of mRNAs to ribosomes is mainly determined by two parameters: (1) abundance of mRNA itself and (2) complimentary interactions between the 3' end of 16S rRNA and the ribosome binding site in the 5'-untranslated region of mRNA.

KEYWORDS mycoplasma, ribosome, ribosome profiling.

INTRODUCTION

System research using the so-called “omics technologies” increasingly reveal unexpected phenomena and new regulated events that simply cannot be determined using either an omics technology alone or traditional methods of analysis. However, the joint analysis of data obtained from the quantitative measurement of RNA or protein and peptide levels inevitably creates a significant number of artifacts due to mistakes in each of the methods used. This fact requires careful cross analysis and additional confirmation of the obtained data using alternative methods or orthogonal processing. For this reason, prokaryotes are used as subjects for testing a common methodology for the analysis of the joint behavior of macromolecules in living systems and their reciprocal impact.

The European Molecular Biology Laboratory has established a project dedicated to studying the causative agent of human respiratory diseases, a representative of the class *Mollicutes* – *Mycoplasma pneumoniae* [1, 2]. Then, our group selected another representative of this class (*M. gallisepticum*) to carry out a similar study. Soon afterwards, American researchers developed a computer model of the metabolism and adaptive responses of the smallest self-replicating bacteria – *M. genitalium* [3]. Despite these successes, a large number of issues remains unexplored, requiring additional methods for accessing dynamic processes during expression of the minimum set of genetic information encoded in the genome of *Mycoplasma*.

Mollicutes, which include *M. gallisepticum*, are characterized by a significant genome reduction. The average size of the genome of *Mycoplasma* typically ranges from 800 thousand to 1 million bp (1 million in *M. gallisepticum*) [4]. Due to genome reduction, *Mycoplasma* lost the well-known mechanisms of regulation of gene expression [5].

As we have showed earlier, *M. gallisepticum* responds to stress at the transcriptional level [6]. At the same time, these changes are generally slightly reflected at the level of translation [6]. This phenomenon may be caused by two reasons: (1) the rate of translation in *M. gallisepticum* is not enough to reveal changes in the protein level during the experiment (30 min), (2) mRNA binds selectively to ribosomes during the stress response. The mechanism of selective attaching of mRNA to ribosomes can be realized through interaction with antisense RNA that blocks the ribosome binding site [7]. Moreover, even within the same cell, ribosomes may differ from each other both in the nucleotide sequence of rRNA [8] and in protein composition. For example, ribosomes in *Escherichia coli*, which do not include S1 protein, translate mainly leaderless transcripts [9]. Ribosomes may bind regulatory proteins modulating translation in particular transcripts [10]. Ribosomes may bind non-canonical proteins whose primary function is not related to translation. For example, glycogen synthase in *Sascharomyces cerevisiae* can affect translation of various RNAs [11].

The development of high-throughput technologies has led to the accumulation of a large database on transcription and translation in the entire cell. According to the classical view, protein level is generally determined by the level of the corresponding mRNA; however, in some cases, it is not. High-throughput technologies have increased the number of cases when the protein level does not correlate with the level of mRNA. Such data were obtained for a variety of bacteria, including *E. coli* [12], *Desulfovibrio vulgaris* [13], and *Lactococcus lactis* [14]. The Pearson correlation coefficient between the level of mRNA and protein level, according to the published data, can vary from 0.53 to 0.19, depending on the type and state of the bacteria. Significant progress in studying the regulation of gene expression at the translational level was achieved with a technology of ribosome profiling [15] which enables observing translation almost in real time. Thus, the stage of binding mRNA to the ribosome and the process of translation is an extremely significant part in the regulation of gene expression in bacteria. Since *Mollicutes* in general and *M. gallisepticum* in particular are characterized by the lack of transcriptional regulatory mechanisms, regulation of gene expression at the translational level can be perhaps the most significant part in determining protein abundance in the cell.

In the present study, a high-throughput proteomic profiling of *M. gallisepticum* ribosomes to determine the composition of ribosomes and transcriptional profiling of ribosome-bound mRNA using real-time PCR were performed.

EXPERIMENTAL SECTION

M. gallisepticum S6 culturing

M. gallisepticum S6 was cultured in a liquid medium (20 g/l Tryptose, 5 g/l NaCl, 1.3 g/l KCl, 3 g/l Tris, 5% yeast dialysate, 6% horse serum, and 1% glucose; pH 7.4) to mid-logarithm growth phase as described in [16].

Ribosome purification

Chloramphenicol was added to 12 ml of *M. gallisepticum* cell culture to a final concentration of the solution of 100 µg/ml, which was thoroughly mixed and incubated for 5 min on ice. Then, the cells were pelleted by centrifugation at 4,500 g for 20 min (4°C). The supernatant was removed, and the cell pellet, which was obtained from 50 ml of culture, was resuspended in 500 µl of lysis buffer containing 20 mM HEPES, 100 mM NaCl, 6 mM MgCl₂, 2 mM spermidine, 100 µg/ml chloramphenicol, 5 µl protease inhibitor (GE Healthcare), and 200 units of RNase inhibitor (Thermo Scientific) (pH 7.5). Following the resuspension, 15 µl of NP-40 was added to the buffer and the composition was mixed

Table 1. Spearman correlation between biological replicates of ribosome-bound mRNA sample. Levels of mRNAs were measured by real-time PCR

	Rep1	Rep2	Rep3
Rep1	1	0.87	0.92
Rep2	0.87	1	0.90
Rep3	0.92	0.90	1

thoroughly. Then, the cell lysate was frozen for at least 1 hour at -75°C. The cell lysate was purified by centrifugation at 20,000 g for 20 min (4°C). The supernatant was collected and fractionated by centrifugation on a sucrose step gradient.

Sucrose step gradient was created in a 5 ml polycarbonate tube by layering sucrose solutions of different densities using a pipette. The volume of each layer was 750 µl, and the difference in density was 10%. In this study, we used 10–50% sucrose gradients (a total of 5 layers). Sucrose solution was prepared using the same buffer as for cell lysis (without adding NP-40, chloramphenicol, and inhibitors of proteases and RNases). The mixture was centrifugated at 50,000 rpm (200,620 g on average) for 1 h at 4°C, using the Optima centrifuge (Beckman Coulter) and the MLS 50 swinging bucket rotor (Beckman Coulter). 200 µl aliquots of fractions were collected using a pipette.

RNA isolation from fractions

Each fraction was added to 400 µl of Trizol LS reagent (Life Technologies). The content was thoroughly mixed, and 200 µl of chloroform was added. The composition was then mixed again and centrifuged for 10 minutes at 16,000 g (4°C). The supernatant was collected and resuspended in an equal volume of isopropanol. The sample was incubated for at least 1 h at -20°C. RNA was pelleted by centrifugation at 16,000 g for 20 min (4°C). The pellet was washed with 80% (v/v) ethanol. The RNA sample was then dissolved in 10 µl of water (Panreac). RNA abundance in the fractions was measured by quantitative real-time PCR. The experiment was conducted in 3 biological replicates.

Protein extraction from fractions and trypsin digestion

For protein precipitation, each fraction was diluted 10-fold with deionized water and trichloroacetic acid (Sigma-Aldrich) was added to a final concentration of 10% (v/v). The mixture was left at 4°C overnight and

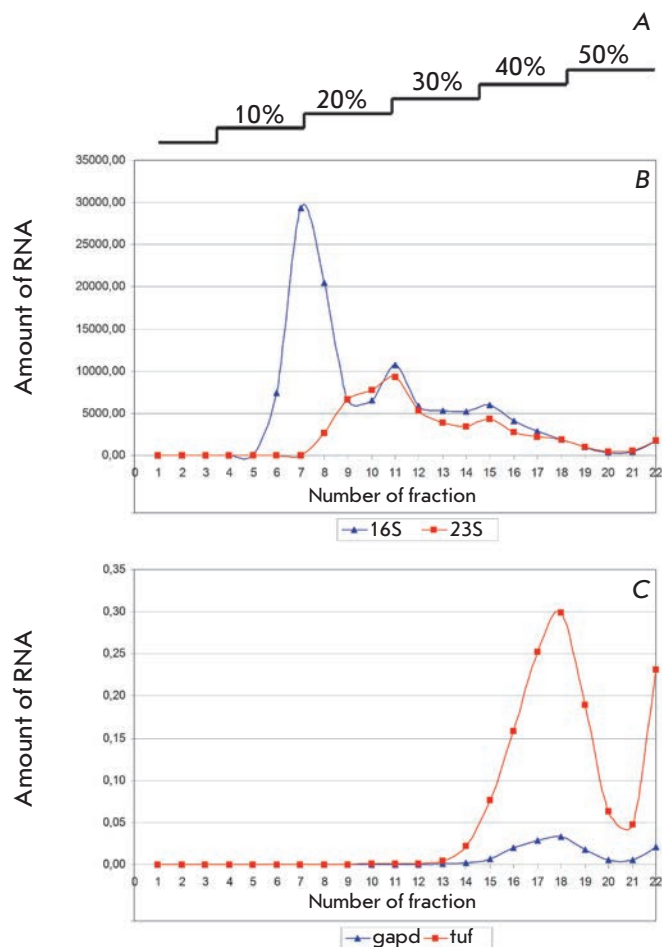


Fig. 1. Fractionation of cytoplasm of *M. gallisepticum* in a sucrose gradient. Abundance of 16S rRNA, 23S rRNA, *gapd* mRNA, and *tuf* mRNA was measured. All results (for rRNA and mRNA) were normalized to the 16S mRNA level in fraction 1. A – step gradient of sucrose (regarding obtained fractions); B – abundance of rRNA; C – abundance of mRNA

then centrifugated for 15 min at 16,000 g. The pellet was washed twice with 1 ml of cold acetone (Pancreac) to remove residual trifluoroacetic acid.

Protein pellets were redissolved in 25–35 μ l of 50 mM ammonium bicarbonate solution (Pancreac) containing 0.5% RapiGest SF (Waters) and 1 μ l of Nuclease Mix (GE Healthcare). Then, the mixture was left for 30 min at 4°C, incubated for 5 min at 100°C, and centrifuged for 10 min at 16,000 g. The supernatant was collected, and the protein content was determined in each sample using bicinchoninic acid (Bicinchoninic Acid Protein Assay Kit, Sigma-Aldrich). In order to reduce disulfide bonds, dithiothreitol (Bio-Rad) was added to the protein solution to a final concentration of 10 mM (the reaction was conducted on a shaker (600

rpm) for 30 minutes at 60°C). The subsequent alkylation of cysteine residues by iodoacetamide (final concentration of 30 mM; Bio-Rad) was performed for 30 min at room temperature in the dark. Then, trypsin (Trypsin Gold, Mass Spectrometry Grade, Promega) was added to protein samples; trypsin:protein ratio (w/w) was 1:50. Trypsin digestion was performed during 16 hours at 37°C. The reaction was stopped by adding 10% trifluoroacetic acid (Sigma-Aldrich) (pH after trifluoroacetic acid addition should be 2.0). Then, the sample was incubated for 45 min at 37°C and centrifuged (15 min at 16,000 g) to remove RapiGest SF. The mixture of tryptic peptides was additionally purified by solid phase extraction using Discovery DSC-18 mini columns (Supelco) according to the manufacturer's recommendations. For further mass spectrometry analysis, the eluate was dried in the CentriVap vacuum concentrator (Labconco) and dissolved in 10 μ l of 3% acetonitrile solution containing 0.1% formic acid.

RNA isolation from cell culture

RNA was extracted from the cell culture according to [16]. Triple volume of Trizol LS reagent (Thermo Scientific) was added to aliquotes of cell culture. Phase separation was induced by adding chloroform (80 μ l per 100 μ l of cell culture). Samples were centrifuged at 10,000 g for 15 min (4°C). Then, the RNA samples were reprecipitated with isopropanol (1:1 v/v).

cDNA synthesis and qPCR

cDNA synthesis and real-time PCR were performed as described in [16]. RNA samples were treated with DNase I (Thermo Scientific). Then, cDNA was synthesized with reverse transcriptase (H minus Reverse Transcriptase, Thermo Scientific) and random hexamers. RiboLock RNase inhibitor (Thermo Scientific) was used to enhance the stability of RNA.

Quantitative real-time PCR (qRT-PCR) was performed on the C1000 Touch thermal cycler (Bio-Rad) with the CFX96 optical module (Bio-Rad). For PCR, 10X PCR buffer (Lytech) (1.5-fold final dilution), 10X dNTP mixture (Lytech), Taq-polymerase (Lytech), SYBR Green I dye (Life Technologies), 5 pmol primers, and 2% formamide were used. Data normalization was carried out according to the average mRNA abundance in 21 housekeeping genes (*eno*, *gaphd*, *tpiA*, *tuf*, *tsf*, *acoA*, *acoB*, *aceF*, *ldh*, *ackA*, *pgk*, *fba*, *pgi*, *pfkA*, *gpmI*, *pykF*, *tktA1*, *rpiB*, *eutD*, *prsA*, and *lpd*). The same as in [6] primers were used for qRT-PCR.

Identification of Proteins

Chromatography-mass spectrometry analysis of peptide extracts was performed using the Q-Exactive HF mass spectrometer (Thermo Fisher Scientific) coupled

with the Ultimate 3000 RSLCnano LC system (Dionex) through the Nanospray Flex ion source (Thermo Fisher Scientific).

Peptides were separated by inverted-phase chromatography using Acclaim PepMap precolumn (C18 stationary phase, length of 2 cm, 75 mm i.d., particle size of 3 μm , and pore size of 100 \AA ; Dionex) and Zorbax column (Zorbax 300SB-C18 stationary phase, length of 15 cm, 75 mm i.d., particle size of 3.5 μm , and pore size of 100 \AA ; Agilent Technologies). Each sample was applied to the precolumn in water for high-pressure liquid chromatography (HPLC) with 0.1% formic acid (v/v) for 5 minutes at a flow rate of 2 $\mu\text{l}/\text{min}$. Then, the precolumn was placed in front of the column. Peptides were eluted with a mixture of solvent A (water for HPLC with 0.1% formic acid (v/v)) and solvent B (79.9% acetonitrile for HPLC (v/v), 20% water for HPLC, and 0.1% formic acid (v/v)), increasing the density gradient of solvent B from 5 to 40% (v/v) for 120 minutes at a flow rate of 300 nl/min . Then, the system was washed for 10 minutes with a mixture of 99% solvent B (v/v) and for 10 minutes with a mixture of 5% solvent B.

The voltage was 2000 V; the temperature of the capillary was 200°C. The mass spectrometer operated in a data-dependent acquisition mode; in each cycle, a panoramic spectrum was obtained; the most intense 20 peaks in the panoramic spectrum were selected in turn for fragmentation and recording of product ion spectra and then excluded from the assessment for 10 seconds. A panoramic spectrum was recorded at a resolution of 70,000 in a mass-to-charge ratio value of 400 to 1200 m/z with automatic gain control (AGC) setting of 106 and time limits for filling of 50 ms. Product ion spectra were recorded at the resolution of 17,500 and AGC of 105 with time limits for filling of 100 ms. Collision energy was 30 V; the width of the ion isolation window was 2 m/z .

On the basis of mass chromatograms (.raw file format), a list of centroid spectra in Mascot Generic Format using the MSConvert utility of the ProteoWizard package (version 3.0.7.414, 64 bits) was compiled, which was then interpreted by the Mascot search engine (Matrix Science Inc.). Protein identification was carried out using the CP006916.2 protein sequence database of *M. gallisepticum* S6, which was supplemented by sequences for common protein contaminants. The following parameters of identification were used: tryptic peptides; maximum of one missed cleavage; precursor-ion charge was +2 or +3; the allowable error in the mass of parent ions was 10 ppm; the accuracy of the mass peak measurement of fragments was 0.5 Da; ESI-TRAP instrument; no permanent modification; cysteine carbamidomethylation and oxidation of methionine residues were variable modifications.

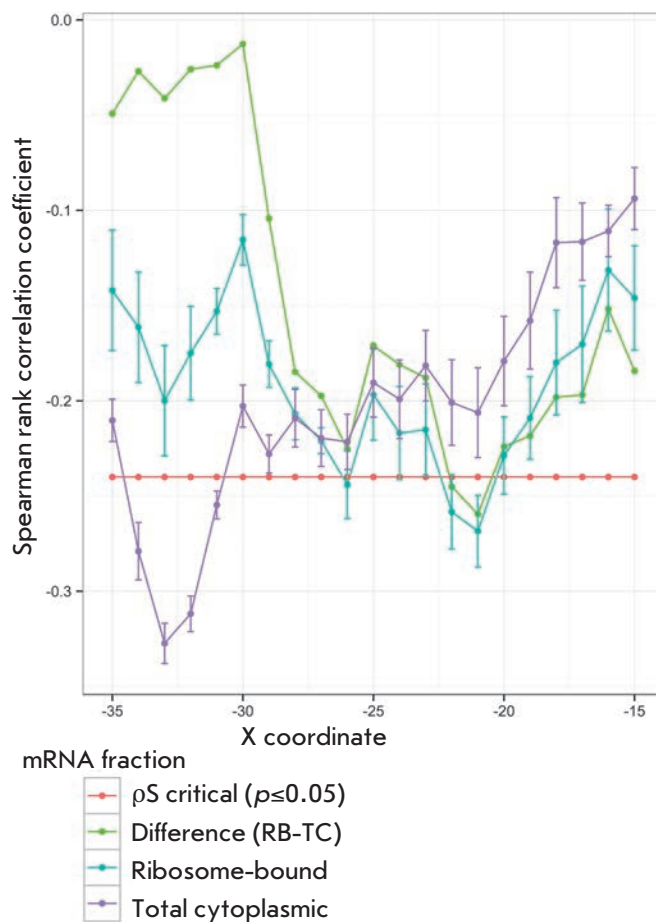


Fig. 2. Correlation between mRNA abundance in a given fraction and dG of the secondary structure near the start codon in a sliding window of 30 nt ($p < 0.05$ for $\rho > 0.25$)

The list of statistically significant identities was defined as a list of proteins with 2 or more identified peptides with $p < 0.05$.

RESULTS AND DISCUSSION

Distribution of RNA in isolated fractions

After the cytoplasm of *M. gallisepticum* was fractionated, distribution of RNA in the fractions was measured by RT-qPCR (see Experimental section). The results are presented in Fig. 1. Small ribosomal subunits and large ones cause peaks in fractions 7 and 11, respectively. *M. gallisepticum* is characterized by an approximately 4-fold abundance of 16S rRNA compared to 23S rRNA [16], which is consistent with the observed picture. Fractions after 12–22 reveal an equimolar ratio of 16S and 23S rRNA. Relatively high levels of mRNA was detected only in polysomes (fractions 15

and above). The highest amount of mRNA was detected in fractions 17 and 18, comprising about 1 microgram. This makes the corresponding fractions most suitable for further analysis, especially when using high-throughput sequencing technologies.

Suitability of the technique for the quantitative analysis of the abundance of transcripts associated with ribosomes was tested for fraction 18. Transcriptional profiling of 67 genes in three biological replicates using RT-qPCR was conducted. The reproducibility of the data was evaluated using Spearman correlation (see table). In each pair of samples, the correlation coefficient ranged from 0.87 to 0.92, indicating good reproducibility of the method. The correlation between the abundance of transcripts in the ribosome-bound mRNA fraction and total cytoplasmic mRNA fraction was 0.78.

Proteomic profiling of *M. gallisepticum* ribosomes

In order to validate the methods of ribosome purification from *M. gallisepticum*, proteomic profiling of fractions 7 (30S subunit), 17, and 18 (70S ribosomes associated with the mRNA) was conducted and a semi-quantitative evaluation of protein abundance according to emPAI was made.

In fraction 7, 18 out of 20 proteins of the small ribosome subunit and only 8 out of 33 proteins of the large subunit were detected, which agrees with the data on the distribution of rRNA in fractions. Thus, fraction 7 was mainly composed of small ribosome subunits. It should be noted that fraction 7 included a significant amount of additional cellular proteins.

Ribosomal proteins were mostly abundant, as expected, in fractions 17–18. A total of 47 out of 53 ribosomal proteins (19 out of 20 proteins of 30S subunit and 28 out of 33 proteins of 50S subunit) were detected. All proteins that were not identified were small in size (less than 100 amino acids); the latter probably impeded identification of these proteins. The abundance of ribosomal proteins in this fraction was regarded as equimolar. Proteins of the small and large subunits had the same emPAI value. Moreover, in fraction 18, a high content of ribosome-associated proteins (EF-Tu and EF-Ts translation factors and Tig and DnaK chaperones) and HU protein was detected. It is known that HU protein, which is a bacterial histone-like protein, can bind both DNA and RNA [17]. It is possible that this protein can bind mRNA or rRNA within the ribosomes.

In fraction 18, we identified a GCW_03230 protein with a high emPAI value. This conservative protein of unknown function is common in many mycoplasmas. Considering the small size (74 amino acids), the feature of this protein is the extreme pI value (11.0), which makes the protein similar to ribosomal proteins. GCW_03230 is likely a new ribosomal protein. In frac-

tion 18, a number of proteins with a relatively high emPAI value, which were not directly related to the process of translation (e.g., triosephosphate isomerase, thioredoxin, a number of proteins of unknown function), were also detected. On the one hand, their presence can be explained by nonspecific interactions with ribosomes after cell lysis. On the other hand, it has been shown recently that such proteins can modulate ribosome activity *in vivo* [11].

Influence of secondary structures and ribosome binding site on mRNA abundance in the ribosome-bound pool

Our results indicate that the mRNA abundance in the ribosome-bound fraction generally corresponds to the mRNA abundance in the total RNA fraction. However, some mRNAs were significantly more or less abundant in this fraction. Binding efficiency of mRNA to the ribosome was determined also by complementary interactions between the 3' end of 16S rRNA and the ribosome binding site in the 5'-untranslated region (5'-UTR) of mRNA and by the presence of secondary structures in this region that mediate or block binding to ribosomes.

Using the RNA duplex program, we modeled *in silico* the interaction between the 3'-terminal region of 16S rRNA (UUACCUCCUUUCU; underlined is the canonical ribosome binding site in *E. coli*) and the 25-nucleotide region upstream of the start codon of each gene. Thus, we obtained results of the binding force of the 16S rRNA with the 5'-UTR of the corresponding mRNA. Spearman correlation between our evaluation of the capacity of the ribosome binding site and the abundance of the corresponding mRNA in the ribosome-bound RNA fraction was 0.39 ($p < 0.01$). We selected mRNA with more than 2-times up-abundance (19) and down-abundance (25) in the ribosome-bound RNA fraction compared to the total mRNA. The energy of duplex formation with the 3' end of 16S rRNA in 5'-UTR of the up-abundant mRNAs was on average half that of down-abundant mRNAs (dG was -4.96 and -2.52 kcal/mol, respectively).

Despite the expected low efficiency of binding to ribosomes (dG > 0), some mRNAs (e.g., GCW_02495 and *putA*) were more abundant in the ribosome-bound fraction than in the total RNA fraction. In the case of GCW_02495, this paradox can be explained by the fact that the GCW_02495 gene is expressed with polycistronic mRNA, together with the adjacent GCW_02490 gene that has a very effective ribosome binding site (dG = -11.8 kcal/mol). Thus, the corresponding mRNA generally binds well with the ribosome.

Several mRNAs were less abundant in the ribosome-bound fraction than in the total RNA fraction despite

the predicted efficiency of ribosome binding. These mRNAs are *GCW_00085*, *glpF*, *gyrA*, *gyrB*, *ruvA*, *potD*, and *hrcA*. This behavior can be explained by the presence of certain secondary structures in the 5'-UTR of mRNA which prevent ribosome binding. Using the quickfold program, we calculated dG values of hairpin structure formation in the region of ribosome binding site and start codon using a sliding window of 30 nucleotides. As a result, we found that the dG value of the secondary structure near the start codon correlates with mRNA abundance in the ribosome-bound fraction (Fig. 2). The best correlation was identified in the range of -21 ... +9 nucleotides upstream of the start codon both for mRNA abundance in the ribosome-bound fraction and for the relative mRNA abundance in the ribosome-bound fraction regarding the total RNA level. Thus, the mRNA abundance in the ribosome-bound

fraction in *M. gallisepticum* can be modulated by secondary structures in the start codon region.

CONCLUSIONS

Our results suggest that the amount of ribosome-bound mRNA in *M. gallisepticum* is largely determined by two parameters: (1) the level of gene transcription and (2) the efficacy of the complementary interaction between the 3'-end of 16S rRNA and the ribosome binding site in the 5'-UTR of mRNA. We have developed a quantitative and reproducible method for obtaining the ribosome-bound fraction of mRNA from *M. gallisepticum*, which can be used for studying the process of translation in this bacterium.

The work was supported by the grant from the Russian Science Foundation (No 14-24-00159).

REFERENCES

- Güell M. et al. Transcriptome complexity in a genome-reduced bacterium. // *Science*. 2009. V. 326. № 5957. P. 1268–1271.
- Kühner S. et al. Proteome organization in a genome-reduced bacterium. // *Science*. 2009. V. 326. № 5957. P. 1235–1240.
- Karr J.R. et al. A whole-cell computational model predicts phenotype from genotype // *Cell*. 2012. V. 150. № 2. P. 389–401.
- Fisunov G. et al. Core proteome of the minimal cell: comparative proteomics of three mollicute species // *PLoS One*. 2011. V. 6. № 7. P. e21964.
- Moreno-Campuzano S., Janga S.C., Pérez-Rueda E. Identification and analysis of DNA-binding transcription factors in *Bacillus subtilis* and other Firmicutes—a genomic approach. // *BMC Genomics*. 2006. V. 7. P. 147.
- Mazin P. V. et al. Transcriptome analysis reveals novel regulatory mechanisms in a genome-reduced bacterium // *Nucleic Acids Res.* 2014. V. 42. № 21. P. 13254–13268.
- Faner M.A., Feigh A.L. Identifying and Characterizing Hfq-RNA Interactions // *Methods*. 2013. V. 63. № 2. P. 144–159.
- Hillebrand A. et al. The seven *E. coli* ribosomal RNA operon upstream regulatory regions differ in structure and transcription factor binding efficiencies // *Biol. Chem.* 2005. V. 386. № 6. P. 523–534.
- Moll I., Resch A., Bläsi U. Discrimination of 5'-terminal start codons by translation initiation factor 3 is mediated by ribosomal protein S1 // *FEBS Lett.* 1998. V. 436. № 2. P. 213–217.
- Shi J. et al. SuhB is a novel ribosome associated protein that regulates expression of MexXY by modulating ribosome stalling in *Pseudomonas aeruginosa*. // *Mol. Microbiol.* 2015. V. 98. № 2. P. 370–383.
- Fuchs G. et al. Proteomic Analysis of Ribosomes: Translational Control of mRNA populations by Glycogen Synthase GYS1 // *J Mol Biol.* 2011. V. 410. № 1. P. 118–130.
- Lu P. et al. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. // *Nat. Biotechnol.* 2007. V. 25. № 1. P. 117–124.
- Nie L., Wu G., Zhang W. Correlation of mRNA expression and protein abundance affected by multiple sequence features related to translational efficiency in *Desulfovibrio vulgaris*: A quantitative analysis // *Genetics*. 2006. V. 174. № 4. P. 2229–2243.
- Picard F. et al. Bacterial translational regulations: high diversity between all mRNAs and major role in gene expression. // *BMC Genomics*. 2012. V. 13. P. 528.
- Ingolia N.T. Ribosome profiling: new views of translation, from single codons to genome scale. // *Nat. Rev. Genet.* Nature Publishing Group, 2014. V. 15. № 3. P. 205–213.
- Gorbachev A.Y. et al. DNA repair in *Mycoplasma gallisepticum* // *BMC Genomics*. 2013. V. 14. P. 726.
- Balandina A., Kamashev D., Rouviere-Yaniv J. The bacterial histone-like protein HU specifically recognizes similar structures in all nucleic acids. DNA, RNA, and their hybrids // *J. Biol. Chem.* 2002. V. 277. № 31. P. 27622–27628.

Attenuation of Vaccinia Virus

S. N. Yakubitskiy¹, I. V. Kolosova¹, R. A. Maksyutov¹, S. N. Shchelkunov^{1,2}

¹State Research Center of Virology and Biotechnology "Vector", Koltsovo, Novosibirsk region, Russia

²Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia

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

Copyright © 2015 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 Since 1980, in the post-smallpox vaccination era the human population has become increasingly susceptible compared to a generation ago to not only the variola (smallpox) virus, but also other zoonotic orthopoxviruses. The need for safer vaccines against orthopoxviruses is even greater now. The Lister vaccine strain (LIVP) of vaccinia virus was used as a parental virus for generating a recombinant 1421ABJCN clone defective in five virulence genes encoding hemagglutinin (*A56R*), the IFN- γ -binding protein (*B8R*), thymidine kinase (*J2R*), the complement-binding protein (*C3L*), and the Bcl-2-like inhibitor of apoptosis (*N1L*). We found that disruption of these loci does not affect replication in mammalian cell cultures. The isogenic recombinant strain 1421ABJCN exhibits a reduced inflammatory response and attenuated neurovirulence relative to LIVP. Virus titers of 1421ABJCN were 3 lg lower versus the parent VACV LIVP when administered by the intracerebral route in new-born mice. In a subcutaneous mouse model, 1421ABJCN displayed levels of VACV-neutralizing antibodies comparable to those of LIVP and conferred protective immunity against lethal challenge by the ectromelia virus. The VACV mutant holds promise as a safe live vaccine strain for preventing smallpox and other orthopoxvirus infections.

KEYWORDS vaccinia virus, virulence genes, vaccine, attenuation, protection.

ABBREVIATIONS PFU – plaque forming units; VARV – variola virus; VACV – vaccinia virus; CPXV – cowpox virus; ECTV – ectromelia virus; IFN – interferon; CBP – complement binding protein; LIVP – a strain L-IVP of vaccinia virus; MCS – multiple cloning site; MPA – mycophenolic acid; PCR – polymerase chain reaction; *gpt* – xanthine-guanine phosphoribosyl transferase gene; CPE – cytopathic effect.

INTRODUCTION

Smallpox was declared eradicated in 1980 by the World Health Organization, after which human vaccination against smallpox was halted [1]. This decision was mainly due to a variety of adverse events, ranging in severity from benign to lethal [1, 2].

Following its termination, smallpox immunity in the human population has been decreasing. Naïve individuals become susceptible to not only variola (smallpox) virus (VARV), but also related orthopoxviruses, which are mainly maintained in the rodent population as an animal reservoir. [3]. Humans and animals can be infected by the monkeypox virus (MPXV) and cowpox virus (CPXV). A wide distribution in the human population is likely to allow orthopoxviruses to adapt to the human immune system and lead to new virus strains infectious for humans [3, 4]. Recently, an increased rate of zoonotic poxvirus outbreaks has been reported worldwide [3], the concerns about potential bioterrorist attacks with VARV have been discussed [5].

The only way to protect against the growing threat posed by orthopoxviruses is vaccination [1, 2]. Immune

deficiency in humans for the last decades has led to a situation whereby classic live vaccinia-virus-based vaccines cannot be efficiently used for severe, associated complications that rarely ensued during the smallpox vaccination era. Consequently, there is an urgent need for orthopox vaccine candidates with a perfect safety record and high immunogenicity, as compared to those used in the past.

First, attenuated VACV strains were produced through multiple passages on chicken embryo fibroblasts (strain MVA) [6] or cell cultures of rabbit kidney (strain LC16m8) [7]. Attenuation was accompanied by spontaneous, extended deletions and other mutations in viral genes implicated in virulence, replication capacity, and host-range [5].

With the advent of genetic engineering tools, live-attenuated vaccine strains can be tailored to the particular virus variant by inserting, deleting, or disrupting genes of choice [8]. Attenuated strains generated by the deletion of virulence genes retain replication competence but are no longer pathogenic. Promising vaccine candidates are recombinant attenuated VACV strains,

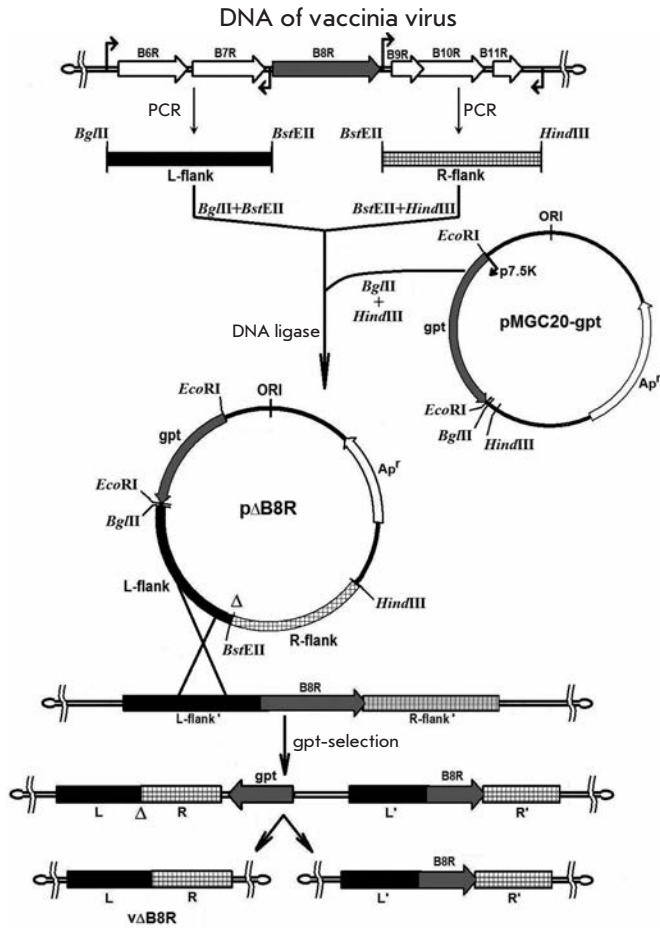


Fig. 1. Schematic diagram illustrating the generation of VACV B8R mutants (see text for details)

capable of eliciting protective antibodies comparable with the smallpox vaccine, but with much reduced post-vaccination side-effects.

Our expertise in complete genome sequencing of orthopoxvirus species and strains pathogenic for man [9–12], targeted mutagenesis [15, 16], and knowledge of the functions of the genes of these viruses [2, 13, 14] allowed us to elaborate and implement a novel approach towards generating attenuated VACVs. This approach involves the targeted excision or disruption of a virulence gene, without affecting viral replication and host-range.

The objective of this work was twofold: (i) to engineer a VACV variant with a knockout in 5 virulence genes, which could be used as a promising live-attenuated vaccine candidate against smallpox and other orthopoxvirus-related diseases in humans; and (ii) to examine its biological characteristics. The isogenic VACV strain, designated 1421ABJCN, was compared

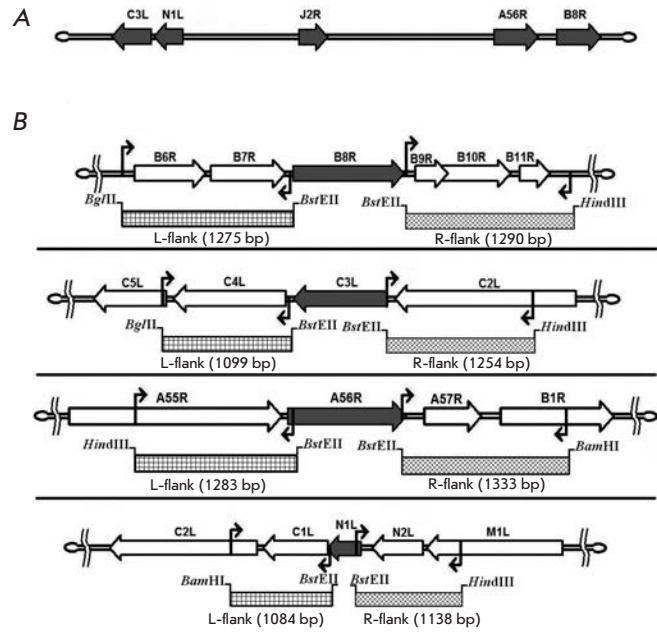


Fig. 2. Schematic diagram illustrating the amplification of PCR fragments flanking the deleted sequences in the VACV genome. A. Map of the VACV genome with the location of the C3L, N1L, J2R, A56R, and B8R virulence genes (indicated with arrows); B. Schematic of amplified regions (shaded area) flanking a sequence stretch upstream (L-flank) and downstream (R-flank) of the deleted genes (grey block arrows). The zigzag arrows stand for the position of primer pairs so that a deletion in one gene does not affect the adjacent genes and amplified DNA fragments are flanked by the restriction sites

to its parental strain LIVP (clonal variant) in terms of viral propagation in cell cultures, pathogenicity in mouse and rabbit models, protective immunogenicity, and protection of experimental animals against a lethal challenge with a virulent orthopoxvirus. These characteristics are carefully considered in pre-clinical trials of new vaccines against orthopoxvirus infection [17].

EXPERIMENTAL SECTION

Bacteria, viruses, cell cultures.

In this study, we used the *Escherichia coli* strains JM109 and XL2-blue, the VACV strain LIVP (derived from the Lister strain background obtained at the Institute for antivirals, Moscow) and the ectromelia virus (ECTV) strain K-1 from the collection of SRC VB Vector, passaged on cultures of African green monkey kidney cells CV-1, and Vero and 4647 from the cell culture collection of SRC VB Vector.

Integrating plasmids

Integrating plasmids used for the deletion of target genes in the VACV genome were derived from the backbone vector pMGC20-gpt carrying the selective marker *E. coli gpt*-gene under the control of the VACV 7.5K promoter [16] (Fig. 1). Upstream (L-flank) and downstream (R-flank) DNA fragments flanking the gene to be deleted were amplified by PCR from the vaccinia virus strain LIVP, using primer pairs with restriction sites (Fig. 1, 2). Both flanking sequences were digested and inserted into the plasmid pMGC20-gpt upstream and downstream of each gene (Fig. 2). Correct construction of pΔB8R, pΔC3L, pΔA56R, and pΔN1L was confirmed by restriction profiling and sequencing.

The plasmid pΔTK carries a MCS sequence which inactivates the thymidine kinase gene (*J2R*) through to an in-frame translation disruption [18].

Generation of VACV strains with deleted/disrupted virulence genes

VACV was propagated on a CV-1 cell culture in a DMEM medium supplemented with 2% fetal bovine serum. All mutant VACV variants were engineered using the transient dominant selection strategy [16]. Following 4-5 passages in the selective medium (before the onset of cytopathic effects), the virus was plaque-purified using an agarose overlay [19] and replated in a non-selective medium for secondary cloning. DNA of plaque-purified clones of recombinant VACVs was extracted as described [15]. After PCR verification, one clone of the 2-3 plaque-purified clones was titered by the plaque assay method on CV-1 [19].

PCR analysis of recombinant VACV variants

Obtained VACV clones were verified for the desired insertions/deletions by PCR with the corresponding primers:

ΔB8R:

5'-TCACAAATATGATGGTGATGAGCGA-3'
5'-CGTGATATACCCTAGCCATAGGCAT-3'.

ΔC3L:

5'-TCGCGCTTTACATTCTCGAATCT-3'
5'-TGTTCTGTGTTCTTGCGGTGA-3'.

ΔA56R:

5'-GTGGTATGGGACACCACAAATCCAA-3'
5'-ATTAAACATTCTAGAAATTAATCCCGCTC-3'.

ΔN1L:

5'-GGGTTGGATCCTTTACACATAGATCTACTA-CAGGCGGAACA-3'

5'-GGGAAAGCTTAATTTGTGAAGATGCCATG-TACTACGCT-3'.

J2R-MCS:

5'-ATATGTTCTTCATGCCTAAACGA-3'
5'-ATGAAGGAGCAAAAGGTTGTAAC-3'.

PCR was carried out in 0.2 ml tubes (Applied Biosystems, USA) on a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems, USA). The reaction mix contained SE buffer for Taq-DNA polymerase, NTPs and Taq-DNA polymerase (SibEnzim, Russia), and sterile deionized water. The thermal profile consisted of an initial denaturation step at 94°C for 1.5 min, followed by 20 cycles of 94°C for 20 s, 55°C for 30 s, 72°C for 1 min, and ending with a final elongation step at 72°C for 5 min. PCR fragments were stored at 4°C until needed.

Growth curves

To evaluate viral replication kinetics, the parental clone 14 of VACV LIVP and mutant VACV strains with inactivated virulence genes were seeded onto 90–100% confluent monolayers of CV-1 or Vero in 6 well plates at a multiplicity of infection (m.o.i) of 0.1 pfu/cell. Viral titers were determined by plaque assay at time points 0, 24, 48, and 72 h post infection.

Harvest and purification of mutant VACVs

A monolayer of cell line 4647, which is recommended for variola virus production [20], grown on tissue-culture flasks with a growth area of 175 cm² (volume 650 ml), was infected with VACV at a multiplicity of infection of 1.0 pfu/cell. Cells were incubated for 48 h at 37°C until the cytopathic effect was evident throughout the culture. A total of 80% of the maintenance medium was removed, and the remaining 20% medium (10 ml) was lysed by three cycles of freeze-thawing, followed by ultrasonic treatment with a MSE 500 disintegrator (22 kHz, 2-3 times for 10-15 s). Cell debris was removed by centrifugation (4,000 g for 10 min). The supernatant was centrifuged at 30,000 g for 1.5 h. The pellet of virus was re-suspended in 4 ml isotonic saline. Infection titers were determined by plaque assay on 4647 monolayer cells.

Animals

Depending of the experiment, Balb/c female mice weighing 14–16 g (aged 4–5 wk) or 1-2-day-old suckling mice of the same strain weighing 5–6 g were used. The animals were placed in groups of ten. We also used Chinchilla rabbits weighing 2.5–3 kg. All animals were obtained from the animal facility of SRC VB Vector. The use and care of animals followed the guidelines for the experimental use of animals. Animal care proce-

dures were approved by the Animal Ethical Committee of SRC VB Vector.

Assessment of mutant VACV virulence

To assess the virulence properties of the genetically engineered VACV strains, 1-2-day-old suckling mice received an intracerebral inoculation of recombinant VACV 1421ABJCN or the parental VACV LIVP clone 14, diluted in isotonic saline at a dose of 10^2 pfu/0.01 ml per mice. Mock-infected controls were inoculated with isotonic saline. Infected animals were examined for death on a daily basis over an 8-day period.

To determine if recombinant VACV strains can replicate and propagate in the brain cells of suckling mice inoculated by the same route, mice were euthanized by cervical dislocation on day 3 post infection and brain tissue was aseptically sampled. Samples from each group were pooled to prepare a 10% suspension in a DMEM medium as previously described [21]. Infection titers were determined by plaque assay on 4647 monolayer cells.

Testing for virulence in rabbits was performed by intradermal injection in a shaved area on the lateral aspect of each flank. Virus suspensions were serially 10-fold diluted in isotonic saline to give 10^2 – 10^7 pfu/0.05 ml, followed by intradermal inoculation in the volume as described above. Each dilution was injected at two separate sites across a shaved area of each rabbit: one flank for VACV 1421ABJCN and the other flank for the parental VACV LIVP. The animals were monitored daily for up to 14 days for the occurrence and disappearance of skin lesions, depending on titer and virus strain.

Immunogenicity

The presence of specific antibodies was evaluated by the neutralization activity of the serum of 4- to 5-week old mice infected with 1421ABJCN or LIVP in a dose range of 10^6 , 10^7 or 10^8 pfu/0.1 ml per mice. Control mice were sham-inoculated with isotonic saline. The immunization was carried out twice at a 28-day interval. On day 28 post second immunization (day 56 post first immunization), ether-anesthetized mice were bled via the retrobulbar venous plexus, blood was allowed to clot overnight at 4°C overnight, and the sera separated from cellular constituents by centrifugation at 5,000 g for 10 min. The sera from each group were pooled and heated at 56°C for 30 min for the inactivation of complement. Virus neutralization titers were determined on a 4647 cell culture as described [22].

Protective immunity

The ability of engineered VACV strains to protect against a lethal orthopoxvirus challenge was assessed in 4- to 5-week old mice following the immunization

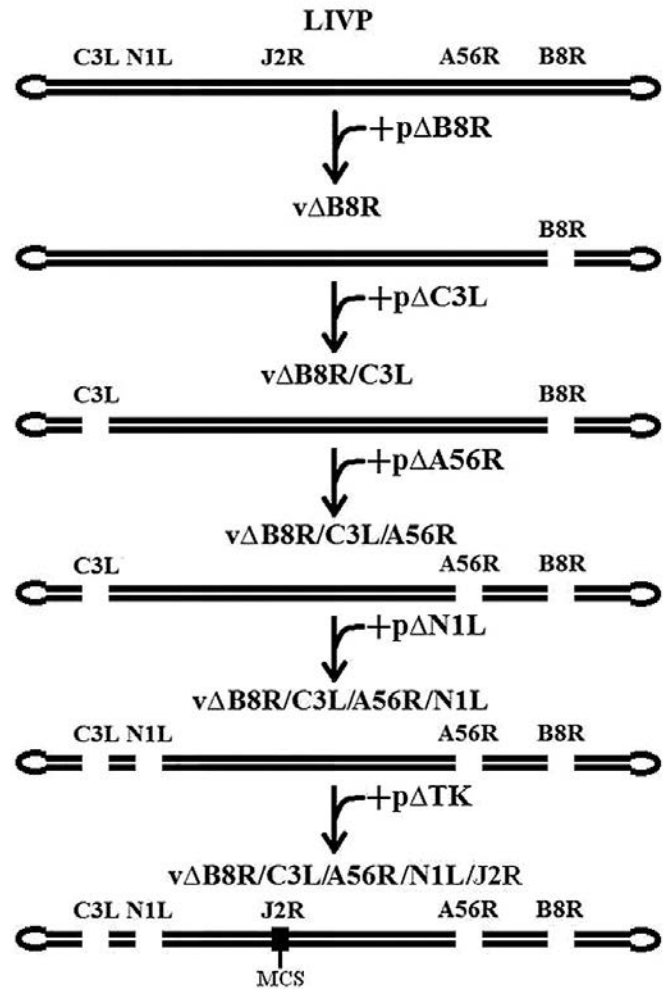


Fig. 3. The strategy to introduce mutations in each of the VACV virulence genes. The order of sequential inactivation with plasmids targeting the genomic sites of choice is shown

procedure as described (see Immunogenicity section). On day 28 post second immunization, lightly anesthetized mice were intranasally infected with highly pathogenic ECTV at a dose of $10 \text{ LD}_{50}/0.02 \text{ ml}$ per mice as previously reported [23]. Survival and mortality of the inoculated mice were monitored for 14 days.

Statistical analysis

Experimental data were analyzed by Student's t-test using the Origin Professional 8.1.10.86 software. $P < 0.05$ was considered statistically significant [24].

RESULTS

Generation of vaccinia virus clones

Before plasmid integration, individual clones of LIVP VACV were isolated by serial plaque-purifications us-

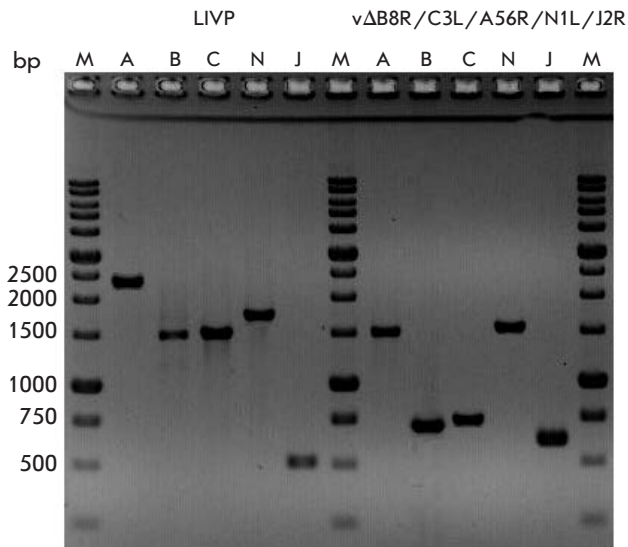


Fig. 4. Verification of deletions/insertions by PCR. PCR products formed from DNA of the parent clone 14 VACV L1VLP and 1421ABJCN with disrupted virulence genes (see text for details). M – molecular size marker, the lengths are given on the left. Lanes A, B, C, N, J contain VACV DNA fragments amplified from the genes *A56R*, *B8R*, *C3L*, *N1L*, and *J2R*, respectively

ing agarose overlay [19] to ensure strain homogeneity. Five candidate VACV clones were picked for subsequent genomic DNA extraction and *Hind*III restriction profiling. Clone 14, which was consistent with the restriction pattern of the parental L1VLP VACV, was selected.

VACV clones with targeted disruption of virulence genes

The genes required for virulence, the interferon- γ (IFN γ) binding protein (*B8R*), secreted complement-binding protein (*C3L*), hemagglutinin (*A56R*), endogenous apoptotic inhibitor (*N1L*), and thymidine kinase (*J2R*) were sequentially inactivated (Fig. 2A). The targeting plasmids p Δ B8R, p Δ C3L, p Δ A56R, p Δ N1L, and p Δ TK were used to attempt plasmid integration and gene disruption.

CV-1 confluent cells were infected with VACV L1VLP clone 14 and transfected with recombinant p Δ B8R plasmid under *gpt*-selection. Single-crossover integration of the targeting plasmid into the viral genome produced a viral DNA genome carrying a *gpt* selection marker, a continuous stretch of viral DNA with targeted deletion of the gene, and the same stretch of viral DNA without deletion. This recombinant DNA segment can only be retained under selection conditions [15, 16].

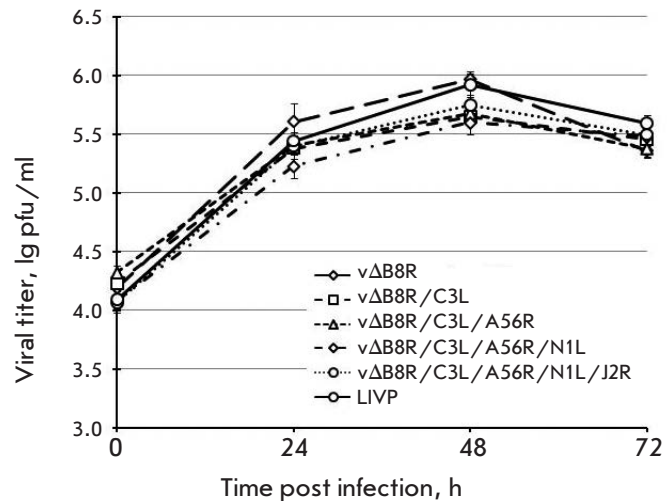


Fig. 5. Growth curves for VACV mutants in CV-1 cells

Due to rapid deletion of *gpt* gene upon removal of selection and high frequency of intramolecular recombination within the R-R' region, recombinant v Δ B8R was obtained (Fig. 1). Candidate clones of this variant were identified by PCR (see Experimental section, data not shown).

One B8R-negative clone was further subjected to deletion of the *C3L* gene to produce a VACV strain negative for *B8R* and *C3L*, followed by sequential disruption of the other three loci *A56R*, *N1L*, and *J2R* (Fig. 3). The final strain with disrupted *B8R*, *C3L*, *A56R*, *N1L*, and *J2R* genes was designated 1421ABJCN.

The expected deletions/insertions were verified by PCR protocols targeting the mutated genes (see Experimental section). Figure 4 shows the PCR analysis of the 1421ABJCN strain with disrupted genes. The full-size PCR products of the genes *A56R*, *B8R*, *C3L*, *N1L*, and *J2R*, amplified from the parental clone L1VLP, were 2366, 1555, 1542, 1784, and 512 bp, respectively. PCR fragments appropriate to the deletion variant genes of Δ A56R, Δ B8R, Δ C3L, and Δ N1L and the insertional variant *J2R*-MCS of 1421ABJCN were 1425, 737, 751, 1431, and 617 bp, respectively, which was consistent with the expected sizes.

Growth curves for mutant VACV in mammalian cell cultures

The highly attenuated vaccine strain must have permissive replication in mammalian culture cells. For this reason, the engineered VACV recombinants were characterized for growth kinetics in CV-1, Vero, and 4647 cells. To analyze viral replication titers, the cell

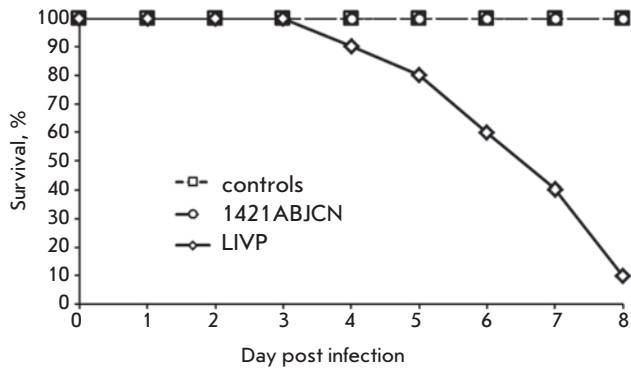


Fig. 6. Time-course of mortality following intracerebral infection of newborn mice with 1421ABJCN and LIVP VACV

cultures were infected with VACV variants with a different combination of disrupted genes. Growth curves of the recombinant viruses in CV-1 cells (Fig. 5) showed no significant differences in replication properties, as well as when compared to the parental VACV LIVP. Similar findings were observed for Vero and 4647 cells.

We found that recombinant VACV variants displayed robust growth and retained a stable genotype over 10 passages in CV-1 and 4647 cells.

Neurovirulence

Lethality after intracerebral challenge was studied in newborn mice observed for 8 days after inoculation. Mice infected with VACV LIVP at a dose of 10² pfu/mice started dying from day 4, with mortality reaching 90% by day 8. Mice inoculated with VACV 1421ABJCN at the same dose showed no mortality (Fig. 6).

In addition, viral titers were determined in the brain of suckling mice on day 3 post infection. The growth characteristics of 1421ABJCN (virus titer ± SD: 2.78 ± 0.66 lg pfu/g of organ) were significantly different from those of the parental VACV LIVP (6.12 ± 0.20 lg pfu/g of organ).

Pathogenicity for rabbits

Pathogenicity studies were performed by bilateral shaving of both flanks, followed by intradermal inoculation of virus for edematous and necrotic lesions to occur. VACV 1421ABJCN caused an inflammatory response at a dose higher than 10⁵ pfu per injection, whereas the parental VACV LIVP induced swelling at a dose of 10² pfu/injection (Table). Swelling and necrotic plaques of 1421ABJCN were less severe than those of VACV LIVP. The lesions induced by 1421ABJCN completely disappeared by day 9 post inoculation versus 14 days for LIVP lesions to heal.

Edematous and necrotic lesions in a rabbit model of intradermal infection with VACV strains

Day post infection	Viral titer, pfu/injection											
	1421ABJCN						LIVP					
	10 ²	10 ³	10 ⁴	10 ⁵	10 ⁶	10 ⁷	10 ²	10 ³	10 ⁴	10 ⁵	10 ⁶	10 ⁷
1	-	-	-	-	S	S	-	-	-	S	S	S
	-	-	-	S	S	S	-	-	-	S	S	S
2	-	-	-	-	S	S	-	-	S	S	S	S/N
	-	-	-	S	S	S	-	-	S	S	S	S/N
3	-	-	-	-	S	S/N	S	S	S/N	S/N	S/N	S/N
	-	-	-	S	S/N	S/N	S	S	S/N	S/N	S/N	S/N
6	-	-	-	-	N	N	-	-	S/N	S/N	S/N	S/N
	-	-	-	-	N	N	-	-	-	S/N	S/N	S/N
7	-	-	-	-	-	N	-	-	S/N	S/N	S/N	S/N
	-	-	-	-	-	N	-	-	-	-	S/N	S/N
9	-	-	-	-	-	N	-	-	N	S/N	S/N	S/N
	-	-	-	-	-	N	-	-	-	-	S/N	S/N
11	-	-	-	-	-	-	-	-	-	N	N	S/N
	-	-	-	-	-	-	-	-	-	-	N	S/N
13	-	-	-	-	-	-	-	-	-	N	N	N
	-	-	-	-	-	-	-	-	-	-	-	N
14	-	-	-	-	-	-	-	-	-	-	-	N
	-	-	-	-	-	-	-	-	-	-	-	-

S – swelling, N – necrosis. Skin lesions were observed in two replicates for each virus dose

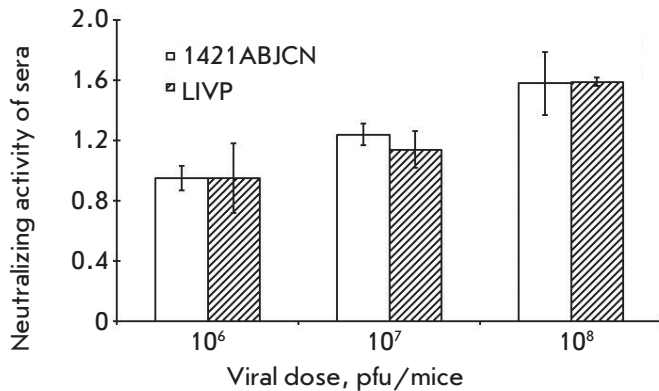


Fig. 7. Levels of serum-neutralizing activity to VACV following double subcutaneous immunization of mice with 1421ABJCN and L1VP at different doses. Neutralizing titers were expressed in $-\log_{10}$ of the highest dilution which gives 50% neutralization of VACV \pm standard deviation

Immunogenicity

The ability of VACV variants to elicit virus-neutralizing antibodies was examined in mice on day 28 post second immunization. Mice were infected at doses of 10^6 , 10^7 or 10^8 pfu/mice. *Figure 7* demonstrates that recombinant 1421ABJCN induces serum levels of virus-neutralizing antibodies comparable with the parental VACV L1VP. These findings suggest that the attenuated VACV variant has the same immunogenic properties as VACV L1VP. Sera of mock-infected controls had no detectable neutralizing antibodies to VACV.

Protective properties of VACV variants

To examine the protective efficacy of VACV variants against lethal challenge, mice were intradermally immunized twice at different doses (10^6 , 10^7 or 10^8 pfu/mice), followed by intranasal ECTV administration at a dose of 10 LD_{50} /mice. Infected mice were examined for death over a 2-week period. Immunization with both 1421ABJCN and VACV L1VP at all doses conferred 100% protection (*Fig. 8*).

DISCUSSION

The genus *Orthopoxvirus* of the family *Poxviridae* includes species pathogenic to humans such as VARV, MPXV, CPXV, and VACV. VARV is a dangerous human pathogen that has caused millions of human casualties in the past [1]. VARV has evolved to become an exclusive anthroponotic agent. Owing to the absence of other natural reservoirs (susceptible hosts) and the development of a smallpox vaccine with vaccinia virus as the active constituent, smallpox was globally eradicated [1, 2].

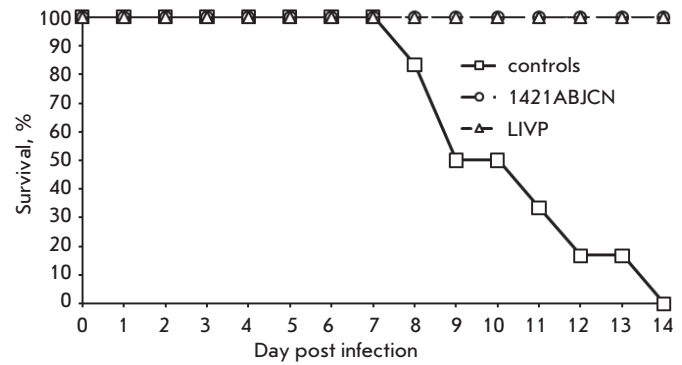


Fig. 8. Time-course of mortality after double subcutaneous immunization of mice with 1421ABJCN or L1VP VACV at a dose of 10^6 pfu/mouse, followed by challenge with ECTV at a dose of 10 LD_{50} /mouse

Following the termination of the program of routine vaccination since 1980, the number of humans susceptible to smallpox and other orthopoxviruses has increased due to lack of seroprevalence [4]. This situation increases the risk of infection and further transmission of orthopoxviruses in the human population, which could re-emerge as novel pathogens, spreading efficiently from human-to-human [3].

To construct a highly attenuated VACV variant for use as an orthopoxvirus vaccine and/or oncolytic virus, we sequentially deleted/disrupted the genes responsible for viral virulence. A literature search allowed us to identify putative virulence genes of VACV suitable for targeting: *B8R*, *C3L*, *A56R*, *N1L*, and *J2R*.

The *B8R* gene of VACV encodes a secreted glycoprotein in the form of a homodimer in infected cells at the early stage of infection. This protein is homologue to the extracellular domain of the IFN- γ receptor, capable of inhibiting IFN- γ activity in various mammalian models [25]. Deleting *B8R* gene results in attenuation with regard to the wild-type strain in mouse infection studies [26] and mild disease even at high doses [27].

The *C3L* gene codes for a secreted complement-binding protein (CBP) in infected cells at the early stage of infection [28] and inactivates complement through interaction with C3b and C4b [14, 29]. CBP suppresses the inflammatory response [30] and antibody-dependent neutralization of VACV virions aided by complement proteins [31]. In animal experiments, mutant VACV strains negative for CBP displayed reduced virulence [14, 31].

The early/late gene *A56R* encodes a surface hemagglutinin protein that mediates viral attachment to host cells, inhibits fusion of infected cells, and promotes pro-

teolytic activation of infectivity. Deletion of the *A56R* gene from the NYCBH VACV strain showed a 40-fold decrease in LD₅₀ versus the parental strain when given to mice by the intracerebral and intranasal routes [32]. Inactivation of the HA gene of the vaccinia virus strain WR leads to significant attenuation [33].

The VACV *N1L* gene encodes an intracellular homodimer with early/late expression [34]. It is non-essential for replication *in vitro*, but it plays an important role in virulence *in vivo*. The protein N1L belongs to the family of Bcl-2-like proteins [35], inhibits apoptosis, and regulates the function of the nuclear factor kappa B pathway [36].

The early VACV *J2R* gene expresses a viral thymidine kinase. *J2R* mutant viruses exhibit reduced virulence *in vivo* [37, 38].

Figures 1 and 2 show a schematic diagram for target inactivation of the virulence genes to generate the expected VACV recombinants. Deletions/insertions were sequentially introduced into the target genes a VACV L1VP clone 14 (Fig. 3) and verified by PCR analysis (Fig. 4) and sequencing.

Because we were concerned about a live, highly attenuated strain without replication deficiency in mammalian cell cultures, we sought to determine replication properties for mutant viruses on CV-1, Vero, and 4647 cells. Growth curves were performed, and the results (Fig. 5) indicate that VACV variants defective in one or more of the virulence genes replicated to titers comparable to those of VACV L1VP in all cell lines. These findings demonstrate that mutations in these genomic loci did not alter viral replication in mammalian cell cultures, which are needed for the manufacturing of live vaccines.

The neurological complications following vaccinia virus vaccination seem to be associated with viral infection in the brain, followed by encephalitis. The hallmark of the VACV vaccine strain is neurovirulence observed in intracerebrally infected mice [39, 40]. The vaccine strain VACV L1VP at a dose of 10² pfu/mice

caused mortality in 90% of the mice by day 8 post inoculation, whereas the descendant strain 1421ABJCN had no virulence in an intracerebral mice model at the same dose (Fig. 6). The titer of the recombinant VACV 1421ABJCN was a 1,000-fold lower than that of the parental strain VACV L1VP in the brain of suckling mice on day 3 post infection.

In a rabbit model of intradermal infection, VACV 1421ABJCN was at least 2 log₁₀ less virulent than the parent strain VACV L1VP (Table).

Both viruses elicited comparable levels of neutralizing antibodies when administered subcutaneously to mice (Fig. 7) and conferred complete protection against challenge with highly virulent ECTV (10 LD₅₀/mice) even at the lowest used immunization dose (10⁶ pfu/mice) (Fig. 8).

CONCLUSION

We constructed a genetically engineered, replication-competent in mammalian cultured cells, VACV 1421ABJCN with targeted inactivation of 5 virulence genes. Our results show that VACV 1421ABJCN vaccination leads to reduced reactogenic properties and neurovirulence versus the parent VACV L1VP that is currently used for human vaccination in the Russian Federation. Besides the attenuated phenotype, VACV 1421ABJCN is immunogenic and protective as L1VP. This virus holds promise for use as a next-generation orthopoxvirus vaccine strain and could serve as a safe vector for recombinant polyvalent vaccines and/or oncolytic viruses.

The authors thank P.F. Safronov and I.N. Babkina for assistance in plasmid production.

The work was supported by the Federal Target Program "Chemical and Biological Security of the Russian Federation (2009–2014)" and grant № 15-04-01326a from the Russian Foundation for Basic Research.

REFERENCES

1. Fenner F, Henderson D.A., Arita I., Jezek Z., Ladnyi I.D. Smallpox and Its Eradication. Geneva: World Health Organization, 1988. 1460 p.
2. Shchelkunov S.N., Marennikova S.S., Moyer R.W. Orthopoxviruses Pathogenic for Humans. Berlin, Heidelberg, New York: Springer, 2005. 425 p.
3. Shchelkunov S.N. // PLoS Path. 2013. V. 9. № 12. e1003756.
4. Shchelkunov S.N. // Vaccine. 2011. V. 29. Suppl. 4. D49–D53.
5. Arntstein A.W., Grabenstein J.D. // Expert Rev. Vaccines. 2008. V. 7. P. 1225–1237.
6. Mayr A., Stickl H., Muller H.K., Danner K., Singer H. // Zentralbl. Bakteriologie. 1978. V. 167. P. 375–390.
7. Kenner J., Cameron F., Empig C., Jobs D.V., Gurwith M. // Vaccine. 2006. V. 24. P. 7009–7022.
8. Mos B. // Vaccine. 2013. V. 31. P. 4220–4222.
9. Shchelkunov S.N., Massung R.F., Esposito J.J. // Virus Res. 1995. V. 36. P. 107–118.
10. 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.
11. 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.
12. Shchelkunov S.N., Totmenin A.V., Safronov P.F., Mikheev M.V., Gutorov V.V., Ryazankina O.I., Petrov N.A., Babkin

RESEARCH ARTICLES

- I.V., Uvarova E.A., Sandakhchiev L.S., et al. // *Virology*. 2002. V. 297. P. 172–194.
13. Shchelkunov S.N. // *Virus Genes*. 2010. V. 41. P. 309–318.
14. Shchelkunov S.N. // *Adv. Virol.* 2012. V. 2012, Article ID 524743, 17 pages.
15. Falkner F.G., Moss B. // *J. Virol.* 1990. V. 64. P. 3108–3111.
16. Kochneva G., Kolosova I., Maksyutova T., Ryabchikova E., Shchelkunov S. // *Arch. Virol.* 2005. V. 150. P. 1857–1870.
17. Guidelines for preclinical studies of drugs (Immuno biological drugs), part two. Edited by Mironova A.N. 2012. M.: Grif and K. 536 p.
18. Maksyutov R.A., Tregubchak T.V., Denisova N.I., Maksyutov A.Z., Gavrilova E.V. // *Russ. J. Immunol.* 2013. V. 4. P. 456–459.
19. *Virology. A practical approach*. Edited by W.J. Mahy. 1985. IRL Press Limited. 344 p.
20. Scarnovich M.O., Radaeva I.F., Vdovichenko G.V., Nechaeva E.A., Sergeev A.A., Petrishenko V.A., Plyasunov I.V., Shishkina L.N., Ternovoy V.A., Smetannikova M.A., et al. // *Voprosy Virusologii*. 2007. V. 2. P. 37–40.
21. Vijaysri S., Jentarra G., Heck M.C., Mercer A.A., McInnes C.J., Jacobs B.L. // *Vaccine*. 2008. V. 26. P. 664–676.
22. Leparc-Goffart I., Poirier B., Garin D., Tissier M-H, Fuchs F., Crance J.-M. // *J. Clin. Virol.* 2005. V. 32. P. 47–52.
23. Martinez M.J., Bray M.P., Huggins J.W. // *Arch. Pathol. Lab. Med.* 2000. V. 124. P. 362–377.
24. Ashmarin I.P., Vorobyev A.A. *Statistical methods in microbiological investigations*. L.: State Publishing House of medical literature. 1962, 186 p.
25. Mossman K., Upton C., Buller R.M., McFadden G. // *Virology*. 1995. V. 208. P. 762–769.
26. Verardi P.H., Jones L.A., Aziz F.H., Ahmad S., Yilma T.D. // *J. Virol.* 2001. V. 75. P. 11–18.
27. Denes B., Gridley D.S., Fodor N., Takatsy Z., Timiryasova T.M., Fodor I. // *J. Gene Med.* 2006. V. 7. P. 814–823.
28. Kotwal G.J., Isaacs S.N., McKenzie R., Frank M.M., Moss B. // *Science*. 1990. V. 250. P. 827–830.
29. Sahu A., Isaacs S.N., Soulika A.M., Lambris J.D. // *J. Immunol.* 1998. V. 160. P. 5596–5604.
30. Miller C.G., Shchelkunov S.N., Kotwal G.J. // *Virology*. 1997. V. 229. P. 126–133.
31. Isaacs S.N., Kotwal G.J., Moss B. // *Proc. Natl. Acad. Sci.* 1992. V. 89. P. 628–632.
32. Lee M.S., Roos J.M., McGuigan L.C., Smith K.A., Cormier N., Cohen L.K., Roberts B.E., Paynet L.G. // *J. Virol.* 1992. V. 66. P. 2617–2630.
33. Shida H., Hinuma Y., Hatanaka M., Morita M., Kidokoro M., Suzuki K., Maruyama T., Takahashi-Nishimaki F., Sugimoto M., Kitamura R., et al. // *J. Virol.* 1988. V. 62. P. 4474–4480.
34. Bartlett N.W., Symons J.A., Tschärke D.C., Smith G.L. // *J. Gen. Virol.* 2002. V. 83. P. 1965–1976.
35. Cooray S., Bahar M.W., Abrescia N.G., McVey C.E., Bartlett N.W., Chen R.A., Stuart D.I., Grimes J.M., Smith G.L. // *J. Gen. Virol.* 2007. V. 88. P. 1656–1666.
36. Maluquer de Motes C., Cooray S., Ren H., Almeida G.M., McGourty K., Bahar M.W., Stuart D.I., Grimes J.M., Graham S.C., Smith G.L. // *PLoS Path.* 2011. V. 7. e1002430.
37. Buller R.M., Smith G.L., Cremer K., Notkins A.L., Moss B. // *Nature*. 1985. V. 317. P. 813–815.
38. Taylor G., Stott E.J., Wertz G., Ball A. // *J. Gen. Virol.* 1991. V. 72. P. 125–130.
39. Li Z., Rubin S.A., Taff R.E., Merchlinsky M., Ye Z., Carbone K.M. // *Vaccine*. 2004. V. 22. P. 1486–1493.
40. Zhang C.X., Sauder C., Malik T., Rubin S.A. // *Biologicals*. 2010. V. 38. P. 278–283.

Role of the Lipid Environment in the Dimerization of Transmembrane Domains of Glycophorin A

A. S. Kuznetsov¹, P. E. Volynsky¹, R. G. Efremov^{1-3*}

¹ M.M. Shemyakin and Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, GSP-7, Miklukho-Maklaya Str., 16/10, 117997, Moscow, Russia

² Higher School of Economics, Myasnitskaya Str., 20, 101000, Moscow, Russia

³ Joint Supercomputer Center of Russian Academy of Sciences, Leninskiy Pr., 32a, 119991, Moscow, Russia

* E-mail: r-efremov@yandex.ru

Received 01.09.2015

Copyright © 2015 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 An efficient computational approach is developed to quantify the free energy of a spontaneous association of the α -helices of proteins in the membrane environment. The approach is based on the numerical decomposition of the free energy profiles of the transmembrane (TM) helices into components corresponding to protein-protein, protein-lipid, and protein-water interactions. The method was tested for the TM segments of human glycophorin A (GpA) and two mutant forms, Gly83Ala and Thr87Val. It was shown that lipids make a significant negative contribution to the free energy of dimerization, while amino acid residues forming the interface of the helix-helix contact may be unfavorable in terms of free energy. The detailed balance between different energy contributions is highly dependent on the amino acid sequence of the TM protein segment. The results show the dominant role of the environment in the interaction of membrane proteins that is changing our notion of the driving force behind the spontaneous association of TM α -helices. Adequate estimation of the contribution of the water-lipid environment thus becomes an extremely urgent task for a rational design of new molecules targeting bitopic membrane proteins, including receptor tyrosine kinases.

KEYWORDS transmembrane domain, glycophorin A, molecular dynamics, protein-protein interactions, role of the lipid membrane, free energy of intermolecular interactions.

ABBREVIATIONS GpA – human glycophorin A; MP – membrane proteins; MD – molecular dynamics; POPC – palmitoyloleoylphosphatidylcholine; RMSD – root-mean-square deviation; RTK – receptor tyrosine kinase; TM – transmembrane.

INTRODUCTION

Protein-protein interactions play an important role in the formation of the supramolecular complexes that perform the essential functions in the cell. The study of these interactions is particularly difficult in the case of membrane proteins (MPs), as they lose their native properties in a nonmembranous environment. Since the α -helix is the most common structural element in the transmembrane (TM) domains of MPs, in some cases, the study of protein-protein interactions is reduced to the consideration of the behavior of TM α -helices in the membrane. This behavior determines the spatial structure of the membrane-bound fragments of ion channels [1–4] and functioning of biotopic (i.e., having only one TM-helix) membrane proteins: in particular, receptor tyrosine kinases (RTKs) [5–9]. It was shown that activation of RTKs requires the formation of dimers or oligomers, wherein the association process involves TM

domains interaction [6, 10–12]. Disruption of MPs leads to severe diseases, such as diabetes mellitus or cancer; so, their study is particularly important. It was shown that a number of disorders in RTK functions result from mutations in the TM domains [13–16]. Designing therapeutic agents selectively acting on the TM-segments of the target protein referred to as “interceptor” peptides is a challenging endeavor. However, resolving this problem requires a detailed understanding of all the molecular mechanisms of signal transmission through these domains of target RTKs [17–20]. Thus, the interaction of α -helices in the membrane is a key process which requires a detailed study.

Experimental methods for studying the interaction of TM domains may include the use of hybrid molecular biological structures with marker-proteins, for example, FRET [21] and TOXCAT [22], as well as the determination of the spatial structure of dimers by nu-

clear magnetic resonance spectroscopy in media that mimic the membrane environment [5, 23–25]. Both types of approaches yield good results; however, they are associated with complex and long-term expression and production of target proteins, as well as with difficulties in stabilizing oligomeric states in membrane-like environments. Therefore, computational simulation techniques which efficiently evaluate the parameters of protein complexes are increasingly used to deal with the problem. In particular, molecular dynamics (MD) is used to quantify the free energy association in studying the role of the medium and the effect of mutations on the dimerization of the TM domains of membrane proteins [11, 12, 26, 27].

Based on the results of an analysis of the TM domains amino acid sequences of several MPs, in particular glycoporphin A (GpA), the concept of “dimerization motifs,” i.e., specific residues that are located in the contact area and determine the interaction between α -helices, has been proposed [28–31]. The so-called glycoporphin dimerization motif includes two glycine residues separated in the sequence by the other three residues and is designated as GG4. This motif is also found in other proteins [5], but in some cases it can be non-functional [9]. Glycoporphin A is still being considered a good model for the theoretical and experimental study of the influence of point mutations and the properties of the medium on the dimerization of TM helices [27, 29, 32–34]. Despite the fact that in these works a key role in dimer formation is assigned to protein-protein interactions, it is shown that the parameters of the lipid membrane influence dimerization [35–37]. These differences were attributed to the condition of hydrophobic mismatch, wherein the most efficient incorporation of a protein into the membrane is ensured by the fact that the length of the membrane-spanning segment of a protein must be equal to the hydrophobic thickness of the lipid bilayer [38].

It is known that the lipid membrane is not a homogeneous medium and tend to form clusters of lipids even in the simplest model systems [39, 40]. Proteins, in turn, often contain binding sites on their surface for the molecules of phospholipids and cholesterol which can modulate their activity [4, 21, 41, 42]. Incorporation of any protein into the membrane changes the properties of the latter one [43, 44], and the dimerization process may induce more complex effects [40, 45]. Thus, the interaction of the TM domains of MP is not limited only to the search for the most favorable protein-protein contacts, but also represents a complex combination of the contributions and interactions of the proteins and the medium.

Therefore, the question arises as to how to calculate the contribution of membrane environment effects to

the free energy of TM domains dimerization of a protein. Moreover, it is necessary to identify the role of each amino acid residue in the process. Important information can be gleaned by studying the effects of point mutations in the amino acid sequences on the distribution of contributions to the energy of the protein and the environment. In the present study, we have investigated the effects of two mutations in the TM domain of GpA on the formation of a dimer. We studied substitutions that influence different factors of dimerization: the mutation Gly83Ala disrupting compact folding of helices and Thr87Val mutation that interferes with the formation of the intermolecular hydrogen bond. The molecular dynamics method was used for the study at the atomic level in an explicit zwitterionic lipid bilayer. The results allow us to draw a conclusion about the dominant role of the membrane in the initial stage of dimer formation and the different molecular mechanisms of disruption of TM complexes association in mutant proteins.

MATERIALS AND METHODS

The systems

For the study, we chose two mutations affecting amino acid residues, which are the most important for glycoporphin A TM domain dimerization: Gly83Ala and Thr87Val [46, 47]. The amino acid sequences of the peptides are shown in *Table 1*. Monomers and dimers of the TM domain were studied in the hydrated lipid bilayers of palmitoyloleoylphosphatidylcholine (POPC). In the monomers, the protein was represented as an ideal α -helix; the initial conformation of the dimer was built on the basis of the experimental structure of the wild-type GpA dimer (PDB ID: 2KPF [48]). Models of dimers of mutant GpA forms were built similarly, substituting the corresponding residues, followed by energy relaxation of the structure. The models of the protein were

Table 1. Amino acid sequences of the studied peptides

Peptide	Amino acid sequence
GpA	<i>SEPEITLIIFGVMAGVIGTILLISYGIRR</i>
GpA Thr87Val	<i>SEPEITLIIFGVMAGVIGVILLISYGIRR</i>
GpA Gly83Ala	<i>SEPEITLIIFGVMA<u>A</u>VIGTILLISYGIRR</i>

Note. Amino acid substitutions are underlined. Italics show terminal residues; their influence was not considered in calculation of the energy profiles for protein-protein interaction.

placed into a lipid bilayer (128 molecules of POPC), and water molecules were added using *genbox* utility. The size of the calculated cell was $6.5 \times 6.5 \times 7.5 \text{ nm}^3$.

Molecular dynamics

MD trajectories were calculated using the GROMACS software package version 4.6.7 and GROMOS96 43a2 force field [49]. The integration time step was 2 fs. Periodic boundary conditions applied in all directions. The calculations were carried out at a constant pressure of 1 atm and a temperature of 315 K. The system pressure was controlled using a Berendsen barostat algorithm [50]. The V-rescale thermostat algorithm was employed for the protein, lipids, and water [51]. Electrostatic interactions were treated using Ewald summation and van der Waals interactions using the Lennard-Jones potential truncated at a cut-off distance of 1.0/1.2 nm.

Before calculating the MD trajectory, energy minimization of the system was performed using the steepest descent method (50,000 steps) and then during the first 50 ps the MD temperature in the system was increased linearly from 5 to 315 K. For relaxation of the membrane environment, a trajectory of 5 ns with a fixed protein molecule was first calculated, followed by 50 ns MD run without any restrictions. The stability of the dimer was analyzed from the change of the crossing angle formed between α -helical axes and the inclination relative to the plane of the bilayer, change of the secondary structure, as well as the value of the root-mean-square deviation (RMSD) of the coordinates of protein backbone atoms from their initial values.

Calculations of the free energy of α -helices dimerization

Profiles of the free energy of TM-domains association of the receptor were obtained by integrating the mean force acting between the monomers. The distance be-

tween the mass centers of the peptides was used as the reaction coordinate. A range of dimer states characterized by different distances between the monomers was generated for subsequent MD simulations (32 points, from 0.75 to 2.10 nm with 0.05 nm increments). These initial states were obtained by parallel transfer of monomers at a predetermined distance along the line passing through their mass centers. The membrane was relaxed *via* 50-ns MD, and the length of the production run was also 50 ns. For each of the states, the value of the mean force acting between the protein monomers was calculated and integrated. In the resulting energy profiles, we evaluated the standard error by defining five independent non-overlapping fragments of each MD trajectory. Two independent calculations were performed for each system; the total length of the accumulated MD trajectories was about 10 μs .

Decomposition of the interaction energy of α -helices

The total energy profile was decomposed into components according to the following approach: at each MD step forces were recalculated using the coordinates of the atoms, taking into account only the sub-systems of interest. Further, these forces were averaged over the length of the MD trajectory, projected on the direction corresponding to the reaction coordinate, and integrated. For plotting the distribution of individual amino acid residues contributions, we selected energy values corresponding to the global minimum in the total profile of the potential of mean force.

RESULTS AND DISCUSSION

Stability of the peptides in the lipid bilayer

As shown by MD, all models of the TM-domains of GpA are stable in terms of the investigated parameters. Thus, the α -helices in the central part were folded and

Table 2. Stability of GpA dimers in MD simulations. The root-mean-square deviation (RMSD) of the resulting structure from the initial one, crossing angle and secondary structure change

Structure	RMSD from the initial structure, nm*	Crossing angle of α -helices axes, degrees	Content of α -helix conformation, %
Initial	0.0	-40 ± 2	84 ± 2
GpA	2.9 ± 0.2 (1.5 ± 0.1)	-39 ± 3	74 ± 2
GpA Thr87Val	2.4 ± 0.3 (1.7 ± 0.3)	-57 ± 4	73 ± 2
GpA Gly83Ala	1.9 ± 0.4 (1.3 ± 0.5)	-47 ± 7	77 ± 2

*RMSD was calculated based on the backbone atomic coordinates; the value for Ile73-Ile95 residues is given in parenthesis.

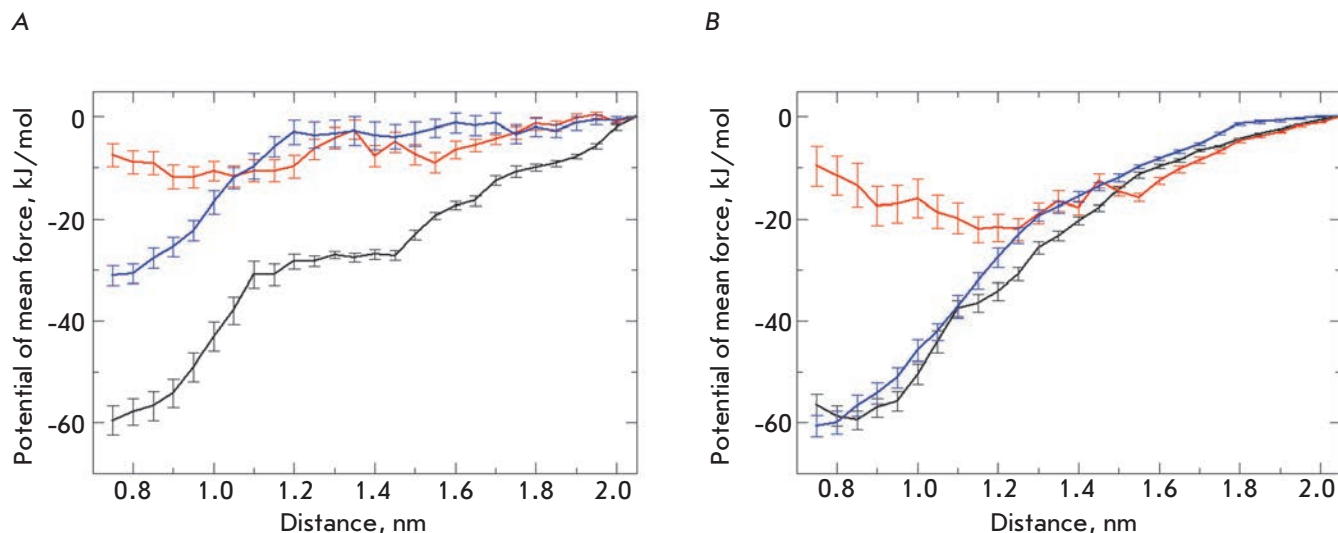


Fig. 1. Profiles of the free energy of TM-domains dimerization of wild-type GpA (black) and its mutant forms, Thr87Val (red) and Gly83Ala (blue). (A): total free energy profiles; (B): contribution of protein-protein interactions for the central part of α -helices (residues Ile73-Ile95)

the crossing angles between the axes of α -helices in dimers were very close to the initial values (*Table 2*). In the case of monomers, we observed a helix inclination relative to the membrane normal caused by the adjusting of the peptide to the environment, referred to as a hydrophobic mismatch. In all systems, there was a partial unfolding of the helices ends at the contact areas with water. The intermolecular hydrogen bond, which is regarded as one of the important factors that stabilize the dimer, was formed by Thr87 residues in the wild-type dimer GpA and its Gly83Ala mutant. However, the dimer with the mutation Thr87Val, which makes this bond impossible, was stable in MD and did not differ significantly by other parameters.

The free energy of GpA α -helices dimerization

The obtained profiles of the association energies of all three systems under study indicate the presence of a stable dimeric state of each protein, but the minimum energy depths are different (*Fig. 1A*). Thus, mutant Thr87Val has a very small free energy of dimerization (in absolute value) (-16 ± 3 kJ/mol) compared to the wild-type dimer (-60 ± 3 kJ/mol). Mutation Gly83Ala also weakens the dimerization, but not too much (-30 ± 5 kJ/mol). Thus, the behavior of the studied peptides varies. For a more detailed study of the differences, we quantified similar energy profiles of a direct protein-protein interaction without the contributions of lipids and water (*Fig. 1B*). It was found that the mutant Gly83Ala drastically differs in energy from the wild-type dimer and reveals an energy profile of the same

shape and depth, while the mutant Thr87Val is also considerably weaker in this case. Thus, the mutation Gly83Ala affects the interaction of the GpA TM-domain with the lipid environment rather than the direct contact of the monomers. When comparing the curves, it is noticeable that in the case of total energy profiles, minima are shifted toward shorter distances compared with the profiles characterizing protein-protein interactions. Thus, the membrane “brings” the monomers in closer contact compared to their equilibrium position, without including the medium effects.

Contribution of amino acid residues

A detailed study of the distribution of energy contributions from the residues showed that the residues which lie at the dimerization interface can form energetically unfavorable contacts (Gly79, Val80, Gly83, Thr87), while the main contribution that promotes the formation of the complex is made by residues that interact with the membrane (Phe78, Ala82, Ile89, Tyr93) (*Fig. 2A*). The influence of the Thr87Val mutation is seen in protein-protein interactions, while the Gly83Ala mutation has in this case a compensation effect (*Fig. 2B*). Thus, the influences of the two mutations differ: substitution of Thr87Val disrupts protein-protein interactions, making it impossible for a hydrogen bond to be formed, while Gly83Ala leads to a minor rearrangement in the structure destabilizing the interaction of the dimer with the membrane environment. It should also be noted that the Gly83 residue in the dimer does not interact with lipids and indirectly affects the entire

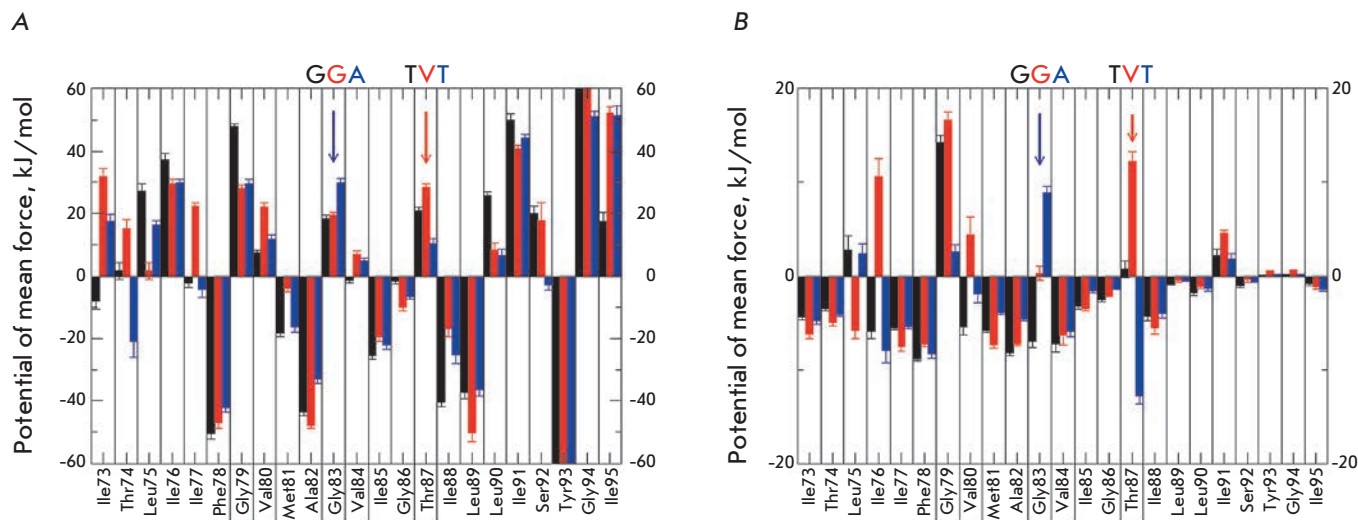


Fig. 2. Diagrams of amino acid residues contributions to the free energy of dimerization for wild-type GpA (black) and its mutant forms, Thr87Val (red) and Gly83Ala (blue). (A): contributions to the total free energy; (B): contributions of protein-protein interactions. Color arrows point to the mutated residues. Frames show amino acid residues that form helix-helix contacts in the dimer

structure, improves overall packing in protein-protein contacts (Table 2, Fig. 1B), but decreases the interaction with the membrane environment. Thus, in describing the interaction of α -helices in the membrane, it is insufficient to consider just the protein-protein interactions and packing density in the protein structure. The lipid environment can make an equal contribution to the stabilization of the dimeric form: so, it is important to study the effects of TM-peptides of different nature on the membrane.

CONCLUSIONS

Results of atomistic molecular dynamics simulations of the interaction between glycoporphin A TM domains and two mutant forms show that the lipid membrane plays an important role in dimer formation, together with the direct contact between the monomers. One of the main conclusions in this work is that the interactions between residues lying on the TM helix-helix interface are often energetically unfavorable. This is, however, compensated by favorable contributions of

the protein-environment contacts to the total free energy of the system. We propose two different scenarios of the disruption of GpA TM-helices association caused by point mutations. Thus, the mutation Thr87Val directly disrupts the protein-protein interactions, and the substitution Gly83Ala has an indirect effect, affecting the membrane environment of the receptor. Thus, the water-lipid environment actively participates in the functioning of the receptor systems of the cell, and its role should be taken into account when considering the functions of membrane proteins, as well as for a rational design of new molecules modulating the function of signaling systems, primarily receptor tyrosine kinases, and other membrane proteins in a goal-oriented manner.

Simulations were carried out using the computational resources of the Joint Supercomputer Center of the Russian Academy of Sciences.

This study was supported by the Russian Scientific Foundation, grant № 14-14-00871.

REFERENCES

- Moreau A., Gosselin-Badaroudine P., Chahine M. // Q. Rev. Biophys. 2014. V. 47. No 4. P. 364–388
- Raja M. // Arch. Biochem. Biophys. 2011. V. 510, No 1. P. 1–10
- Williamson I.M., Alvis S.J., East J.M., Lee A.G. // Cell. Mol. Life Sci. 2003. V. 60. No 8. P. 1581–1590
- Prasanna X., Chattopadhyay A., Sengupta A. // Biophys. J. 2014. V. 106. No 6. P. 1290–1300
- Bocharov E.V., Mineev K.S., Goncharuk M.V., Arseniev A.S. // Biochim. Biophys. Acta – Biomembr. 2012. V. 1818. No 9. P. 2158–2170
- Bocharov E.V., Lesovoy D.M., Goncharuk S.A., Goncharuk M.V., Hristova K., Arseniev A.S. // Structure. 2013. V. 21. No 11. P. 2087–2093
- Jura N., Endres N.F., Engel K., Deindl S., Das R., Lamers M.H., Wemmer D.E., Zhang X., Kuriyan J. // Cell. 2009.

- V. 137. No 7. P. 1293–1307
8. Endres N.F., Das R., Smith A.W., Arkhipov A., Kovacs E., Huang Y., Pelton J.G., Shan Y., Shaw D.E., Wemmer D.E., et al. // *Cell*. 2013. V. 152. No 3. P. 543–556
 9. Bocharov E.V., Mayzel M.L., Volynsky P.E., Mineev K.S., Tkach E.N., Ermolyuk Y.S., Schulga A.A., Efremov R.G., Arseniev A.S. // *Biophys. J.* 2010. V. 98. No 5. P. 881–889
 10. Fleming K.G., Engelman D.M. // *Proc. Natl. Acad. Sci. U. S. A.* 2001. V. 98. No 25. P. 14340–14344
 11. Efremov R.G., Vereshaga Y.A., Volynsky P.E., Nolde D.E., Arseniev A.S. // *J. Comput. Aided Mol. Des.* 2006. V. 20. No 1. P. 27–45
 12. Polyansky A.A., Volynsky P.E., Efremov R.G. // *J. Am. Chem. Soc.* 2012. V. 134. No 35. P. 14390–14400
 13. Volynsky P.E., Polyansky A.A., Fakhrutdinova G.N., Bocharov E.V., Efremov R.G. // *J. Am. Chem. Soc.* 2013. V. 135. No 22. P. 8105–8108
 14. Roskoski R. // *Pharmacol. Res. Off. J. Ital. Pharmacol. Soc.* 2014. V. 79. P. 34–74
 15. Li E., Hristova K. // *Biochemistry*. 2006. V. 45. No 20. P. 6241–6251
 16. Hubert P., Sawma P., Duneau J.-P., Khao J., Hénin J., Bagnard D., Sturgis J. // *Cell Adhes. Migr.* 2010. V. 4. No 2. P. 313–324
 17. Najumudeen A.K. // *Sci. Signal*. 2010. V. 3. № 138. P. jc6
 18. Bennisroune A., Fickova M., Gardin A., Dirrig-Grosch S., Aunis D., Crémel G., Hubert P. // *Mol. Biol. Cell*. 2004. V. 15. No 7. P. 3464–3474
 19. Arpel A., Sawma P., Spenlé C., Fritz J., Meyer L., Garnier N., Velázquez-Quesada I., Hussenet T., Aci-Sèche S., Baumlin N., et al. // *Cell Rep*. 2014. V. 8. No 6. P. 1714–1721
 20. Yin H., Slusky J.S., Berger B.W., Walters R.S., Vilaire G., Litvinov R.I., Lear J.D., Caputo G.A., Bennett J.S., DeGrado W.F. // *Science*. 2007. V. 315. No 5820. P. 1817–1822
 21. Yano Y., Kondo K., Kitani R., Yamamoto A., Matsuzaki K. // *Biochemistry*. 2015. V. 54. No 6. P. 1371–1379
 22. Russ W.P., Engelman D.M. // *Proc. Natl. Acad. Sci. U. S. A.* 1999. V. 96. No 3. P. 863–868
 23. Bocharov E.V., Volynsky P.E., Pavlov K.V., Efremov R.G., Arseniev A.S. // *Cell Adhes. Migr.* 2010. V. 4. No 2. P. 284–298
 24. Mineev K.S., Lesovoy D.M., Usmanova D.R., Goncharuk S.A., Shulepko M.A., Lyukmanova E.N., Kirpichnikov M.P., Bocharov E.V., Arseniev A.S. // *Biochim. Biophys. Acta – Biomembr.* 2014. V. 1838. No 1. P. 164–172
 25. Smith S.O., Bormann B.J. // *Proc. Natl. Acad. Sci. U. S. A.* 1995. V. 92. No 2. P. 488–491
 26. Pohorille A., Jarzynski C., Chipot C. // *J. Phys. Chem. B*. 2010. V. 114. No 32. P. 10235–10253
 27. Hénin J., Pohorille A., Chipot C. // *J. Am. Chem. Soc.* 2005. V. 127. No 23. P. 8478–8484
 28. Engelman D.M., Adair B.D., Brünger A., Flanagan J.M., Hunt J.F., Lemmon M.A., Treutlein H., Zhang J. // *Soc. Gen. Physiol. Ser.* 1993. V. 48. P. 11–21
 29. Melnyk R.A., Kim S., Curran A.R., Engelman D.M., Bowie J.U., Deber C.M. // *J. Biol. Chem.* 2004. V. 279. No 16. P. 16591–16597
 30. Doura A.K., Kobus F.J., Dubrovsky L., Hibbard E., Fleming K.G. J. // *Mol. Biol.* 2004. V. 341. No 4. P. 991–998
 31. Prakash A., Janosi L., Doxastakis M. // *Biophys. J.* 2011. V. 101. No 8. P. 1949–1958
 32. Cuthbertson J.M., Bond P.J., Sansom M.S.P. // *Biochemistry*. 2006. V. 45. No 48. P. 14298–14310
 33. Duong M.T., Jaszewski T.M., Fleming K.G., MacKenzie K.R. // *J. Mol. Biol.* 2007. V. 371. No 2. P. 422–434
 34. Mueller B.K., Subramaniam S., Senes A. // *Proc. Natl. Acad. Sci. U. S. A.* 2014. V. 111. No 10. P. E888–E895
 35. Janosi L., Prakash A., Doxastakis M. // *Biophys. J.* 2010. V. 99. No 1. P. 284–292
 36. Johnson R.M., Rath A., Melnyk R.A., Deber C.M. // *Biochemistry*. 2006. V. 45. No 28. P. 8507–8515
 37. Mehrbod M., Mofrad M.R.K. // *PLoS Comput. Biol.* 2013. V. 9. No 3. P. e1002948
 38. Caputo G.A. // *Methods Mol. Biol.* 2013. V. 1063. P. 95–116
 39. Pyrkova D.V., Tarasova N.K., Pyrkov T.V., Krylov N.A., Efremov R.G. // *Soft Matter*. 2011. V. 7. No 6. P. 2569–2579
 40. Parton D.L., Tek A., Baaden M., Sansom M.S.P. // *PLoS Comput. Biol.* 2013. V. 9. No 4. P. e1003034
 41. Lensink M.F., Govaerts C., Ruyschaert J.-M. // *J. Biol. Chem.* 2010. V. 285. No 14. P. 10519–10526
 42. Poveda J.A., Giudici A.M., Renart M.L., Molina M.L., Montoya E., Fernández-Carvajal A., Fernández-Ballester G., Encinar J.A., González-Ros J.M. // *Biochim. Biophys. Acta – Biomembranes*. 2014. V. 1838. No 6. P. 1560–1567
 43. Dave P.C., Tiburu E.K., Damodaran K., Lorigan G.A. // *Biophys. J.* 2004. V. 86. No 3. P. 1564–1573
 44. Stangl M., Schneider D. // *Biochim. Biophys. Acta – Biomembr.* 2015. V. 1848. No 9. P. 1886–1896
 45. Kuznetsov A.S., Polyansky A.A., Fleck M., Volynsky P.E., Efremov R.G. // *J. Chem. Theory Comput.* 2015. V. 11. No 9. P. 4415–4426
 46. Lemmon M.A., Flanagan J.M., Treutlein H.R., Zhang J., Engelman D.M. // *Biochemistry*. 1992. V. 31. No 51. P. 12719–12725
 47. Lee J., Im W. // *J. Am. Chem. Soc.* 2008. V. 130. No 20. P. 6456–6462
 48. Mineev K.S., Bocharov E.V., Volynsky P.E., Goncharuk M.V., Tkach E.N., Ermolyuk Y.S., Schulga A.A., Chupin V.V., Maslennikov I.V., Efremov R.G., et al. // *Acta Naturae*. 2011. V. 3. No 2. P. 90–98
 49. Van Der Spoel D., Lindahl E., Hess B., Groenhof G., Mark A.E., Berendsen H.J.C. // *J. Comput. Chem.* 2005. V. 26. No 16. P. 1701–1718
 50. Berendsen H.J.C., Postma J.P.M., van Gunsteren W.F., DiNola A., Haak J.R. // *J. Chem. Phys.* 1984. V. 81. No 8. P. 3684
 51. Bussi G., Donadio D., Parrinello M. // *J. Chem. Phys.* 2007. V. 126. No 1. P. 014101

Sulfoxides, Analogues of L-Methionine and L-Cysteine As Pro-Drugs against Gram-Positive and Gram-Negative Bacteria

N. V. Anufrieva¹, E. A. Morozova¹, V. V. Kulikova¹, N. P. Bazhulina¹, I. V. Manukhov², D. I. Degtev², E. Yu. Gnuchikh², A. N. Rodionov¹, G. B. Zavilgelsky², T. V. Demidkina^{1*}

¹Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilova Str., 32, Moscow, 119991, Russia

²State Research Institute of Genetics and Selection of Industrial Microorganisms, 1-st Dorozhniy pr., 1, Moscow, 117545, Russia

*E-mail: tvd@eimb.ru, tvdemidkina@yandex.ru

Copyright © 2015 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 problem of resistance to antibiotics requires the development of new classes of broad-spectrum antimicrobial drugs. The concept of pro-drugs allows researchers to look for new approaches to obtain effective drugs with improved pharmacokinetic and pharmacodynamic properties. Thiosulfinates, formed enzymatically from amino acid sulfoxides upon crushing cells of genus *Allium* plants, are known as antimicrobial compounds. The instability and high reactivity of thiosulfinates complicate their use as individual antimicrobial compounds. We propose a pharmacologically complementary pair: an amino acid sulfoxide pro-drug and vitamin B6 – dependent methionine γ -lyase, which metabolizes it in the patient's body. The enzyme catalyzes the γ - and β -elimination reactions of sulfoxides, analogues of L-methionine and L-cysteine, which leads to the formation of thiosulfinates. In the present work, we cloned the enzyme gene from *Clostridium sporogenes*. Ionic and tautomeric forms of the internal aldimine were determined by lognormal deconvolution of the holoenzyme spectrum and the catalytic parameters of the recombinant enzyme in the γ - and β -elimination reactions of amino acids, and some sulfoxides of amino acids were obtained. For the first time, the possibility of usage of the enzyme for effective conversion of sulfoxides was established and the antimicrobial activity of thiosulfinates against Gram-negative and Gram-positive bacteria *in situ* was shown.

KEYWORDS Pro-drugs, vitamin B6-dependent enzymes, cloning of *Clostridium sporogenes* methionine γ -lyase gene, alliin, allicin, sulfoxides of amino acids, Gram-positive and Gram-negative bacteria.

ABBREVIATIONS PLP – pyridoxal 5'-phosphate, MGL – methionine γ -lyase, His-tag – poly-histidine fragment, His-tag MGL – methionine γ -lyase with poly-histidine fragment, *megL* – gene encoding of MGL in *Clostridium sporogenes*, DTT – dithiothreitol, NADH – reduced form of β -nicotinamide adenine dinucleotide, EDTA – ethylenediaminetetraacetic acid.

INTRODUCTION

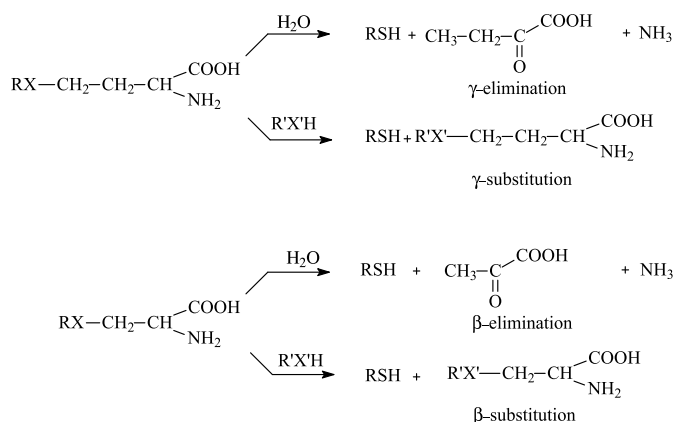
The development of new antimicrobial agents with a minimal inherent risk of inducing rapid resistance to antibiotics is one of the most pressing issues nowadays. Many potentially effective antimicrobial agents are rapidly degraded in the human body and have high toxicity, preventing their use in the concentrations necessary for treatment. This issue can be resolved through the concept of pro-drugs, compounds that must be metabolized in the body of a patient. This concept has been successfully used in tumor therapy [1].

In the present study, we propose using this approach to create effective antimicrobial therapy using a phar-

macological pair of a pro-drug and a biocatalyst metabolizing it. Recently, we have demonstrated that methionine γ -lyase (MGL) [EC 4.4.1.11] from *Citrobacter freundii* catalyzes the β -elimination reaction of a non-protein amino acid, (\pm)-S-(2-propenyl)-L-cysteinesulfoxide ((\pm)-alliin), resulting in 2-propene thiosulfinate (allicin), a natural antibiotic [2].

MGL catalyzes the γ -elimination reaction of L-methionine to produce methylmercaptan, α -ketobutyric acid, and ammonia. The enzyme catalyzes the β -elimination reactions of L-cysteine and its S-substituted derivatives to the corresponding mercaptans, pyruvic acid and ammonia, and the substitution reactions at the

C_β- and C_γ-atoms of L-cysteine and L-methionine and their analogues [3, 4]:



X = S, O or Se;
X' = S or Se

MGL is present in fungi [5], *Arabidopsis thaliana* [6], in various bacteria, including pathogenic *Aeromonas* spp. [7], *Clostridium sporogenes* [8], *Porphyromonas gingivalis* [9], and pathogenic protozoa *Entamoeba histolytica* [10] and *Trichomonas vaginalis* [11]. The enzyme has no counterpart in mammals and, therefore, may be considered as a target in pathogens. This approach was implemented using a suicide substrate of the enzyme. Catalysis of the γ -elimination reaction of trifluoromethionine led to the formation of trifluoromethanethiol, which spontaneously decomposes to thiocarbonyl difluoride, which has an antimicrobial effect on MGL-containing *T. vaginalis* [12], *P. gingivalis* [13], and *E. histolytica* [14]. However, the high toxicity of thiocarbonyl difluoride prohibits the use of trifluoromethionine as an antimicrobial agent.

Allicin, the most well-known antimicrobial and antitumor component of garlic, accounts for about 70% of all thiosulfinates [15] formed by the β -elimination reaction of alliin, which is catalyzed by PLP-dependent alliinase [EC 4.4.1.4] [16] upon crushing of garlic. The antimicrobial action of allicin and other thiosulfinates formed enzymatically during the crushing of plant cells of genus *Allium* is largely due to their ability to oxidize the sulfhydryl groups of proteins/enzymes of bacterial cells, whereas animal cells are partially protected by the presence of glutathione [17]. The antimicrobial, anti-inflammatory, antioxidant, and anticarcinogenic effects of organic sulfo-compounds, cell extracts of garlic and onions [18, 19], have been known since ancient times. However, isolated thiosulfinates are not used in medicine due to their high reactivity and, hence, instability. Only allicin has been studied extensively as an individual biologically active compound, and its antitumor, antioxidant, antibacterial, and antifungal properties have been identified [20–22].

MGL ability to catalyze the γ - and β -elimination reactions of methionine sulfoxide [23] and alliin [2] to produce thiosulfinates allows one to use the concept of pro-drugs to develop a new antimicrobial agent, using the substrates of the enzyme, alliin and other sulfoxides, as pro-drugs *in situ* generating thiosulfinates.

Previously, we cloned the *C. sporogenes* gene (*megL*) encoding MGL with a polyhistidine fragment (His-tag) at the N-terminus of the polypeptide chain and determined some kinetic characteristics of the recombinant enzyme (His-tag MGL). *C. sporogenes* MGL catalyzed the γ -elimination reaction of L-methionine at a faster rate than the enzyme from *C. freundii* [24] and showed higher cytotoxic activity against a number of tumor cells [25].

The cleavage of His-tag by thrombin increases the rate of the physiological substrate cleavage by *C. sporogenes* MGL by 1.5 times. In this study, we cloned the *C. sporogenes* MGL gene without His-tag. The steady-state kinetic parameters of the γ - and β -elimination reactions of a number of well-known substrates and sulfoxides, analogues of cysteine and methionine, and the spectral characteristics of *C. sporogenes* MGL have been determined. The antibacterial activity of mixtures containing MGLs from *C. sporogenes* and *C. freundii* and the sulfoxides of amino acids has been demonstrated in a solid medium. It has been shown that the kinetic parameters of the recombinant PLP-dependent MGL make it possible, in principle, to use the enzyme to convert pro-drugs, sulfoxides of amino acids, to thiosulfinates.

MATERIALS AND METHODS

Reagents, enzymes

The following compounds were used in the study: pyridoxal 5'-phosphate, L-methionine, L-cysteine, L-homocysteine, L-norvaline, L-norleucine, L- α -aminobutyric acid, alliin, S-ethyl-L-cysteine, S-ethyl-L-homocysteine, L-alanine, O-acetyl-L-serine, lactate dehydrogenase from rabbit muscle, DTT, NADH, sodium periodate, ethyl bromide (all Sigma, USA); EDTA, protamine sulfate (Serva, USA); lactose (Panreac, Spain); glucose, glycerol, magnesium sulfate, ammonium sulfate, monopotassium phosphate, disodium phosphate ("Reakhim," Russia); yeast extract, tryptone (Difco, USA); DEAE-Sephacrose (GE Healthcare, Sweden); O-acetyl-L-homoserine was produced by L-homoserine acetylation as described previously [26]. 2-Nitro-5-thiobenzoic acid was obtained according to [27]. (\pm)-L-methionine sulfoxide was obtained according to the standard procedure [28]. Synthesis of (\pm)-S-ethyl-L-cysteine and (\pm)-S-ethyl-L-homocysteine sulfoxides was performed according to [29–31].

Restriction and ligation reactions were carried out with enzymes from Promega (USA). A “working buffer” with pH 8.0, containing 100 mM potassium phosphate, 0.1 mM PLP, 1 mM DTT, and 1 mM EDTA was used.

Escherichia coli strain BL21 (DE3) F⁻ *ompT hsdS_B gal dcm* (DE3) (Novagen) was used to express the *C. sporogenes* MGL gene. *E. coli* strain K12 AB2463 - *arecA* derivative of *E. coli* K12, has a F⁻, *thr-1 leu-6 proA2 his-4 thi-1 argE3 lacY1 galK2 ara-14 xyl-5 mtl-1 tsx-33 rpsL31 supE44, recA13* genotype. It was used for cloning, production, and storage of the plasmid. *C. freundii* strain ATCC 21434 from the American Type Culture Collection (USA) was kindly provided by R. S. Phillips. The *Staphylococcus aureus* strain 015 was kindly provided by Yu. F. Belyi. The plasmid with D-2-hydroxyisocaproate dehydrogenase was kindly provided by K. Muratore.

Cloning of the *C. sporogenes* MGL gene

The pET28a-*megL*_sporog plasmid was constructed based on the pET28a plasmid, containing the *C. sporogenes megL* gene with a polyhistidine fragment (His-tag) and designated as pET28a::*megL_s*_HT [24]. The amplicon (*megL*_sporog), containing the *megL* gene without His-tag, was obtained by PCR. pET28a plasmid carrying *megL* with His-tag was used as a template. The primers included the NcoI restriction site (underlines): *megL*_sporog:5'-CGCG-CGGCAGCCCCATGGAGAA-3' (forward), *megL*_sporog:5'-CCGGATCTCAGTGGTGGTGGTG-3' (reverse).

*MegL*_sporog amplicon was cloned in the pET28a vector by the NcoI and EcoRI sites in the *recA*- *E. coli* strain AB2463. The cloning was controlled by sequencing the insert. Transformation was carried out using the *E. coli* strain BL21 (DE3).

Biomass growth and enzyme purification

Cells of *E. coli* BL21 (DE3), containing the MGL gene without His-tag in the pET28a *megL*_sporog plasmid, were grown in the “inducing” medium [32] at 37 °C with stirring (180 rpm) for 24 hours. The cells were collected by centrifugation and stored at -80 °C. The cells were destroyed and purified from nucleic acids as described previously [33]. Further purification was carried out by ion exchange chromatography on a column with DEAE-Sepharose equilibrated with the working buffer. The column was pre-washed with the working buffer containing 100 mM KCl. The enzyme was eluted with the working buffer containing 500 mM KCl, concentrated and dialyzed against the working buffer. The purity of the preparation was checked by polyacrylamide gel electrophoresis under denaturing con-

ditions according to Laemmli [34]. The concentration of the purified preparations was determined using a $A_{1\%}^{278}$ coefficient of 0.8 [23].

Assay of the enzyme activity and steady-state kinetics parameters

MGL activity during the purification was assayed in the γ - and β -elimination reactions by measuring the reduction of NADH absorption at 340 nm ($\epsilon = 6220 \text{ M}^{-1}\text{cm}^{-1}$) at 30 °C to estimate the rate of keto acids formation in the conjugation reaction with D-2-hydroxyisocaproate dehydrogenase (the γ -elimination reaction) or lactate dehydrogenase (the β -elimination reaction). The reaction mixtures contained the working buffer, 0.2 mM NADH, 10 units of lactate dehydrogenase or 70 μg of D-2-hydroxyisocaproate dehydrogenase, 30 mM S-ethyl-L-cysteine, or 30 mM L-methionine. One unit of enzyme activity was defined as the amount of the enzyme that catalyzes the formation of 1.0 $\mu\text{M}/\text{min}$ of pyruvate (or α -ketobutyrate). The specific activity of 95% pure enzyme preparations was 26.8 units/mg for the γ -elimination reaction of L-methionine and 8.32 unit/mg for the β -elimination reaction of S-ethyl-L-cysteine.

Steady-state kinetic parameters for the γ - and β -elimination reactions were measured in the same manner by varying the substrates concentrations. The obtained data were processed according to the Michaelis-Menten equation using the EnzFitter software. Calculations were based on the molecular weight of an enzyme subunit of 43 kDa. Inhibition of the γ -elimination reaction of L-methionine by various amino acids was studied under the conditions described above by varying the concentrations of substrates and inhibitors in the reaction mixture. The values of inhibition constants were determined using the EnzFitter software. The data were processed in Dixon coordinates [35].

Spectral studies

The absorption spectrum of holoenzyme was recorded at 25 °C on a Cary-50 spectrophotometer (Varian, USA) in the working buffer without PLP. The enzyme concentration was 1.036 mg/mL.

Antimicrobial activity of drugs

Overnight cultures of *C. freundii* and *S. aureus* grown in a Luria-Bertani medium (LB-medium) at 37 °C were diluted 100-fold in a LB-medium and grown at 37 °C with constant stirring to an optical density of 0.2–0.3 at 600 nm. The bacterial cultures were plated on solid-medium dishes (LB-agar). Mixtures of MGLs from different sources and sulfoxides of amino acids pre-incubated at room temperature for 1 hour were applied to 12 mm filter paper disks placed on the dishes. The

Table 1. Kinetic parameters of the γ - and β -elimination reactions*

Substrate	<i>C. sporogenes</i> MGL			<i>C. sporogenes</i> His-tag MGL**			<i>C. freundii</i> MGL***			<i>P. putida</i> MGL****		
	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , M ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , M ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , M ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , M ⁻¹ s ⁻¹
L-Met	21.61	0.60	3.60×10^4	9.86	0.43	2.28×10^4	6.2	0.7	8.85×10^3	48.6	0.90	5.4×10^4
(±)-L-MetO	21.66	11.39	1.90×10^3	8.59	7.89	1.09×10^3	8.12	4.65	1.75×10^3	-	-	-
S-Et-L-Hcy	21.31	0.24	8.87×10^4	7.05	0.27	2.54×10^4	6.78	0.54	1.25×10^4	33.4	0.27	1.23×10^5
(±)-S-Et-L-HcyO	0.48	0.60	8.0×10^2	-	-	-	-	-	-	-	-	-
O-Ac-L-Hse	37.26	3.18	1.17×10^4	-	-	-	2.1	2.91	7.21×10^2	78.0	2.22	3.51×10^4
S-Et-L-Cys	6.53	0.43	1.52×10^4	6.3	0.358	1.76×10^4	5.03	0.17	2.96×10^4	5.79	0.48	1.21×10^4
(±)-S-Et-L-CysO	1.39	0.33	4.21×10^3	-	-	-	-	-	-	-	-	-
O-Ac-L-Ser	5.31	8.01	6.6×10^2	-	-	-	2.13	4.28	4.98×10^2	-	-	-
(±)-Alliin	11.43	1.43	7.99×10^3	-	-	-	5.9	4.7	1.26×10^3	-	-	-

*The error did not exceed 10%. **Data from [25]. ***Data from [2, 23, 33]. ****Data from [37].

concentrations of MGLs from *C. sporogenes* and *C. freundii* and sulfoxides were 10 and 2.5 mg/mL, respectively. The dishes were incubated for 24 hours at 37 °C, and inhibition zones were then measured. The control solutions of the enzymes and the sulfoxides mixtures retained their antibacterial activity for 2 weeks.

Determination of allicin

Allicin, produced in the mixtures containing MGL and alliin, was determined in a reaction with 2-nitro-5-thiobenzoic acid. The mixture of MGL and alliin was added to 1 mL of 0.1 mM 2-nitro-5-thiobenzoic acid in a 100 mM potassium-phosphate buffer containing 0.2 mM PLP, pH 8.0. The mixture was incubated for 30 min at room temperature. Allicin molar concentration was calculated by the decrease in absorbance at 412 nm using a molar absorption coefficient of 2-nitro-5-thiobenzoic acid at 412 nm of $28,300 \text{ M}^{-1}\text{cm}^{-1}$ [27].

RESULTS AND DISCUSSION

Kinetic parameters of the β - and γ -elimination reactions

Previously [25], we showed that cleavage of His-tag from *C. sporogenes* MGL by thrombin leads to a 1.5-fold increase in the activity of the enzyme in the physiological reaction with L-methionine. In this work, we have determined the parameters of steady-state kinetics of *C. sporogenes* MGL without His-tag in the γ -elimination reactions of five substrates (L-methionine, L-methionine sulfoxide, S-ethyl-L-homocysteine, S-ethyl-L-homocysteine sulfoxide and O-acetyl-L-homoserine) and in the β -elimination reactions of four substrates (S-ethyl-L-cysteine, S-ethyl-L-cysteine sulfoxide, O-acetyl-L-serine and alliin). Table 1 summa-

rizes the parameters for MGL from *C. sporogenes*, for MGLs derived from two other bacterial sources, and *C. sporogenes* His-tag MGL.

The k_{cat} values for *C. sporogenes* MGL in the γ -elimination reactions of three substrates, L-methionine, S-ethyl-L-homocysteine, and L-methionine sulfoxide, were 2–3 times higher than for *C. sporogenes* His-tag MGL. K_M values for the first two substrates were close, and the K_M value for L-methionine sulfoxide was slightly higher than that for His-tag MGL.

The presence of the His-tag fragment does not affect the kinetic parameters of the β -elimination reaction of S-ethyl-L-cysteine, and K_M and k_{cat} values for MGL are almost identical to those for His-tag MGL. In the γ - and β -elimination reactions, the elimination of the side-chain groups of the substrates is catalyzed by different acid groups of the enzyme. Presumably, in the case of the β -elimination reaction catalyzed by PLP-dependent lyases, this group is the side group of the lysine residue (Lys210 in *C. freundii* MGL) which binds the coenzyme [36]. In PLP-dependent γ -elimination and γ -replacement reactions, this role is attributed to the conservative tyrosine residue (Tyr113 in *C. freundii* MGL) involved in the stacking interaction with the coenzyme ring [36]. This assumption is confirmed by the data obtained for the mutant form of *Pseudomonas putida* MGL, in which Tyr114 is replaced with Phe [37]. It has also been shown that the acid/base properties of Tyr113 in *C. freundii* MGL are regulated by the Cys115/Tyr113/Arg60 triad [2]. Arg60 is located in the mobile N-terminal loop of the enzyme, and the nitrogen atom of the guanidine group is positioned within a hydrogen-bond distance from the hydroxyl group of Tyr113 in the three-dimensional structure of the holoenzyme [38], the structures of MGL com-

plexes with amino acids modeling the Michaelis complex [39], and in the spatial structure of the external aldimine of the enzyme with glycine [40]. The His-tag fragment may affect the conformation of the N-terminal loop and, therefore, the relative arrangement of the hydroxyl group of Tyr113 and the guanidine group of Arg60, which, in turn, may affect the pK_a value of the hydroxyl group of Tyr113. That may explain the increase in the γ -eliminating activity of *C. sporogenes* MGL compared with His-tag MGL.

Comparison of the enzymes from three bacterial sources (Table 1), *P. putida*, *C. freundii*, and *C. sporogenes*, showed that their affinity for both the physiological substrate and its analogues are almost equal. The efficiency of catalysis in the reaction γ -elimination of L-methionine for *C. sporogene* and *P. putida* MGLs is close, and the k_{cat}/K_M value for *C. freundii* MGL is somewhat lower. The kinetic parameters of the β -elimination reaction of S-ethyl-L-cysteine are very similar for the three enzymes.

C. sporogenes MGL catalyzes the γ -elimination reaction of L-methionine sulfoxide with a catalytic efficiency which is an order of magnitude higher than that in the γ -elimination reaction of S-ethyl-L-homocysteine sulfoxide. The rate of the β -elimination reaction of S-ethyl-L-cysteine sulfoxide, catalyzed by the enzyme, is 15 times lower than the rate of the γ -elimination reaction of L-methionine sulfoxide, but due to the greater affinity of *C. sporogenes* MGL to this substrate, the overall catalytic efficiency is virtually the same. Among the reactions with amino acids sulfoxides, the enzyme most effectively catalyzes the β -elimination reaction of alliin.

The enzyme from *C. sporogenes* catalyzes the γ -elimination reaction of L-methionine sulfoxide more effectively than *C. freundii* MGL (k_{cat} value is 2.5 times higher). The rate of alliin cleavage by *C. sporogenes* MGL is almost 2 times higher than that of the enzyme from *C. freundii*, the substrate affinity is 3 times higher, and the efficiency of catalysis is 6.3 times higher.

Amino acids with a linear side chain inhibited the γ -elimination reaction of L-methionine competitively. Table 2 shows the inhibition constants for *C. sporogenes*, *C. freundii*, and *P. putida* MGLs. All of these enzymes demonstrate an increase in binding with an increase in the number of methylene groups in amino acids with linear side chains, which can be attributed to the hydrophobic nature of the active site of the enzyme from *P. putida* [41] and *C. freundii* [38]. The significant increase in the affinity of the enzyme from the three sources then switching from L-norvaline to L-norleucine and close values of K_i for L-norleucine and K_M for L-methionine and S-ethyl-L-cysteine may be attributed to the presence of a “pocket” for the amino acid methyl group in the MGL active site.

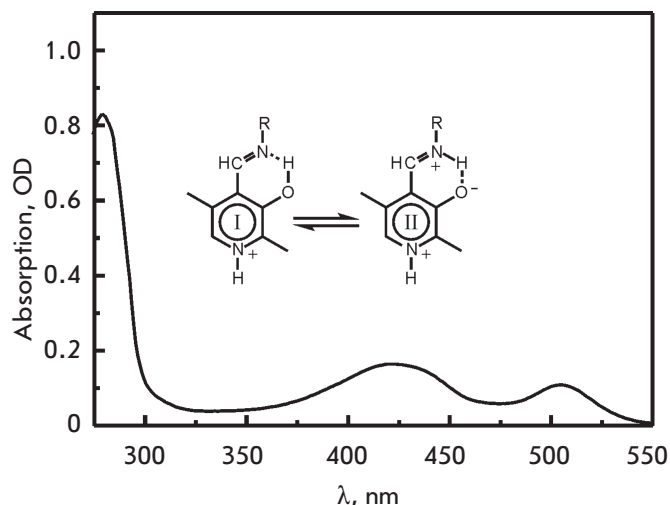


Fig. 1. Absorption spectrum of the holoenzyme *C. sporogenes* MGL

Table 2. Inhibition of the γ -elimination reaction of L-methionine*

Amino acid	K_i , mM		
	<i>C. freundii</i> **	<i>C. sporogenes</i>	<i>P. putida</i> ***
L-Ala	3.4	1.5	5.1
L-Abu	8.3	2.0	8.4
L-Nva	4.7	1.9	3.0
L-Nle	0.6	0.37	0.5

*The error did not exceed 10%.

**Data from [23].

***Data from [43].

Spectral characteristics of the enzyme

The absorption spectrum of *C. sporogenes* MGL holoenzyme (Fig. 1) at pH 8.0 is similar to the spectrum of *C. freundii* MGL [23], with a predominant absorption band of the ketoenamine form of the internal aldimine in the region 422–425 nm (Fig. 1, structure II). Just like *C. sporogenes* His-tag MGL [24], the spectrum contains an intense absorption band with a maximum in the region 502–505 nm, which is attributed to a quinonoid intermediate in the spectra of PLP-dependent enzyme complexes with amino acids and model compounds [42].

Deconvolution of the holoenzyme spectrum in the region 300–500 nm using lognormal curves was performed according to [23]. Table 3 shows the parameters of the absorption bands obtained after deconvolution. In addition to the ketoenamine form, the

Table 3. Parameters of the absorption spectrum bands of the internal aldimine *C. sporogenes* MGL

Structure	E , eV	$\nu \times 10^{-3}$, cm^{-1}	λ , nm	$\epsilon \times 10^{-3}$, $\text{M}^{-1}\text{cm}^{-1}$	$W \times 10^{-3}$, cm^{-1}	ρ	f	n , %
II ¹	2.92	23.53	425.0	10.46	3.58	1.58	0.22	64.7
II ²	3.24	26.15	382.4	7.76	4.00	1.37	0.02	7.5
I	3.63	29.28	341.5	9.44	3.65	1.23	0.03	10.0
II ^{1,2}	3.79	30.56	327.2	10.27	3.47	1.29	0.01	5.6
II ^{2*}	4.28	34.55	289.4	5.98	5.06	1.20	0.18	
*	4.46	35.99	277.9	6.70	4.70	1.50	0.26	

E , electron transition energy; ν , wave number; λ , wavelength; ϵ , molar absorption coefficient; W , half-width; ρ , asymmetry; f , oscillator force; n , contents of tautomers and conformers. The content of PLP in the enzyme is 87.8%.

* Experimental information about these bands is insufficient.

Above-line indices (1, 2) correspond to the first and second electron transitions of structure II. Above-line indices (1,2) correspond to two conformers of structure II (the conformer with the aldimine group in the plane perpendicular to the pyridine cycle plane and the conformer with the aldimine bond released from the coenzyme ring plane but with retained coupling and a hydrogen bond between the aldimine nitrogen atom and the coenzyme 3'-oxygroup).

Table 4. Inhibition of cell culture by mixtures containing MGL and sulfoxides of amino acids

Amino acid sulfoxide	Inhibition zone, mm ²			
	<i>C. freundii</i> MGL		<i>C. sporogenes</i> MGL	
	<i>C. freundii</i>	<i>S. aureus</i>	<i>C. freundii</i>	<i>S. aureus</i>
(±)-Alliin	380	754	254	754
(±)-L-MetO	452	491	177	227
(±)-S-Et-L-CysO	314	491	254	314
(±)-S-Et-L-HcyO	254	415	227	227

internal aldimine (Fig. 1, structure II, $\epsilon = 10410 \text{ M}^{-1}\text{s}^{-1}$) is represented by minor structures, enol tautomer (Fig. 1, structure I), and two ketoenamine conformers with the aldimine bond perpendicular to the plane of the coenzyme ring (absorption in the region of 380 nm) and with the aldimine bond partly removed from the plane of the ring but retaining its coupling with π -electrons of the cofactor and a hydrogen bond between aldimine nitrogen and the 3'-oxygroup of PLP (absorption in the region 327–328 nm). The ionic form of the internal aldimine and tautomeric equilibrium are almost the same as those for *C. freundii* MGL. The absorption in the region 502–505 nm requires further investigation.

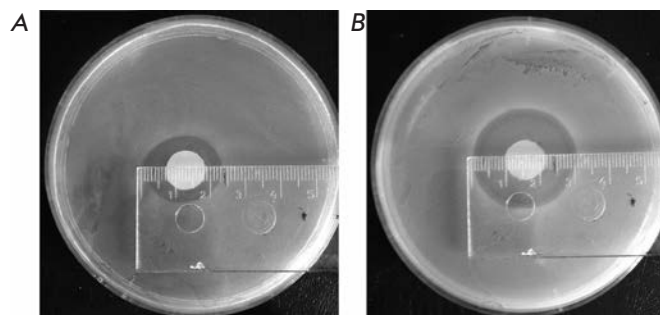


Fig. 2. Diffusion in agar by the Kirby-Bauer method [44]. The mixture of *C. sporogenes* MGL (10 mg/ml) and alliin (2.5 mg/ml) in 100 mM potassium phosphate-buffer was applied on the – A) cell culture of *C. freundii*, B) cell culture of *S. aureus*

Antimicrobial activity of mixtures of *C. freundii* and *C. sporogenes* MGLs with sulfoxides of amino acids

The antibacterial activity of mixtures of MGLs from two sources and sulfoxides of amino acids was assessed using bacterial cultures of Gram-positive *S. aureus* and Gram-negative *C. freundii* (Table 4). All mixtures showed a bacteriostatic effect against Gram-positive and Gram-negative bacteria. The most significant effect was observed for the culture of *S. aureus* (Fig. 2). The bacteriostatic effect was comparable to the inhibition of bacterial cell growth by kanamycin. The inhibition zones of kanamycin (0.05 mg) and a mixture comprising 0.04 mg of allicin in the *C. freundii* culture amounted to 314 and 346 mm², respectively.

Therefore, the data obtained show that the recombinant enzyme effectively catalyzes the conversion of amino acids sulfoxides into thiosulfinates. This suggests that a pharmacological pair of MGL and a sulfoxide can ensure production of thiosulfinates in the amounts necessary for therapeutic purposes.

CONCLUSIONS

MGL catalyzes the γ - and β -elimination reactions of sulfoxides, analogues of methionine and cysteine, with a catalytic efficiency comparable to the efficiency of the γ - and β -elimination reactions of these amino acids.

Using a solid medium, we have demonstrated that mixtures of sulfoxides and MGL are promising as antimicrobial agents against Gram-positive and Gram-negative bacteria *in situ*.

The strongest bacteriostatic effect for the mixture of amino acids sulfoxides and MGL have been observed for Gram-positive bacteria *S. aureus*, and the bacteriostatic effect of allicin produced *in situ* is comparable with the effect of kanamycin.

The authors thank State Research Institute of Genetics and Selection of Industrial Microorganisms for allowing the use of its facilities to clone Clostridium sporogenes methionine γ -lyase gene.

This work was supported by the Russian Science Foundation (project № 15-14-00009).

REFERENCES

- Wentworth P., Datta A., Blakey D., Boyle T., Partridge L.J., Blackburn G.M. // Proc. Natl. Acad. Sci. U.S.A. 1996. V. 93. P. 799–803.
- Morozova E.A., Revtovich S.V., Anufrieva N.V., Kulikova V.V., Nikulin A.D., Demidkina T.V. // Acta Crystallogr. D. Biol. Crystallogr. 2014. V. 70 (11). P. 3034–3042.
- Tanaka H., Esaki N., Soda, K. // Enzyme Microb. Technol. 1985. V. 7. P. 530–537.
- Faleev N.G., Troitskaya M.V., Ivoilov V.S., Karpova V.V., Belikov V.M. // Prikladnaya biokhimiya i microbiologiya. 1994. V. 30 (3). P. 458–463.
- El-Sayed A.S. // Appl. Microbiol. Biotechnol. 2010. V. 86. P. 445–467.
- Goyer A., Collakova E., Shachar-Hill Y., Hanson A.D. // Plant Cell Physiol. 2007. V. 48. P. 232–242.
- Nakayama T., Esaki N., Lee W.-J., Tanaka I., Tanaka H., Soda K. // Agric. Biol. Chem. 1984. V. 48. P. 2367–2369.
- Kreis W., Hession C. // Cancer Res. 1973. V. 33. P. 1862–1865.
- Yoshimura M., Nakano Y., Yamashita Y., Oho T., Saito T., Koga T. // Infection Immunity. 2000. V. 68. P. 6912–6916.
- Tokoro M., Asai T., Kobayashi S., Takeuchi T., Nozaki T. // J. Biol. Chem. 2003. V. 278. P. 42717–42727.
- Lockwood B., Coombs G. // Biochem. J. 1991. V. 279. P. 675–682.
- Coombs G.H., Mottram J.C. // Antimicrob. Agents and Chemother. 2001. V. 45. P. 1743–1745.
- Yoshimura M., Nakano Y., Koga, T. // Biochem. Biophys. Res. Commun. 2002. V. 292. P. 964–968.
- Sato D., Kobayashi S., Yasui H., Shibata N., Toru T., Yamamoto M., Tokoro G., Ali V., Soga T., Takeuchi T., Suematsu M., Nozaki T. // Int. J. Antimicrob. Agents. 2010. V. 35 (1). P. 56–61.
- Han J., Lawson L., Han G., Han P. // Anal. Biochem. 1995. V. 225. P. 157–160.
- Stoll A., Seebeck E. // Adv. Enzymol. 1951. V. 11. P. 377–400.
- Rabinkov A., Miron T., Konstantinovskii L., Wilchek M., Mirelman D., Weiner L. // Biochim. Biophys. Acta. 1998. V. 1379. P. 233–244.
- Rose P., Whiteman M., Moore P.K., Zhu Y.Z. // Nat. Prod. Rep. 2005. V. 22. P. 351–368.
- Lynett P.T., Butts K., Vaidya V., Garrett G.E., Pratt D.A. // Org. Biomol. Chem. 2011. V. 9. P. 3320–3330.
- Hirsch K., Danilenko M., Giat J., Miron T., Rabinkov A., Wilchek M., Mirelman D., Levy J., Sharoni Y. // Nutr. Cancer. 2000. V. 38. P. 245–254.
- Shadkhan Y., Shemesh E., Mirelman D., Miron T., Rabinkov A., Wilchek M., Osherov N. // J. Antimicrob. Chemother. 2004. V. 53. P. 832–836.
- Curtis H., Noll U., Störmann J., Slusarenko A.J. // Physiol. Mol. Plant Pathol. 2004. V. 65. P. 79–89.
- Morozova E.A., Bazhulina N.P., Anufrieva N.V., Mamaeva D.V., Tkachev Y.V., Streltsov S.A., Timofeev V.P., Faleev N.G., Demidkina T.V. // Biochemistry (Mosc.). 2010. V. 75. P. 1272–1280.
- Revtovich S.V., Morozova E.A., Anufrieva N.V., Kotlov M.I., Belyi Yu.F., Demidkina T.V. // Doklady Biochemistry and Biophysics. 2012. V. 445. P. 187–193.
- Morozova E.A., Kulikova V.V., Yashin D.V., Anufrieva N.V., Anisimova N.Y., Revtovich S.V., Kotlov M.I., Belyi Y.F., Pokrovsky V.S., Demidkina T.V. // Acta Naturae. 2013. V. 5. P. 96–102.
- Nagai S., Flavin M. // J. Biol. Chem. 1967. V. 242. P. 3884–3895.
- Miron T., Rabinkov A., Mirelman D., Weiner L., Wilchek M. // Anal. Biochem. 1998. V. 265. P. 317–332.
- Mitsudome T., Takahashi Y., Mizugaki T., Jitsukawa K., Kaneda K. // Angew. Chem. Int. Ed. Engl. 2014. V. 53 (32). P. 8348–8351.
- Frankel M., Gertner D., Jacobson H., Zilkha A. // Journal of the Chemical Society. 1960. P. 1390–1393.

RESEARCH ARTICLES

30. Briggs W.H., Xiao H., Parkin K.L., Shen C., Goldman I.L. // *Journal of Agricultural and Food Chemistry*. 2000. V. 48 (11). P. 5731–5735.
31. Waelsch H., Owades P., Miller H.K., Borek E. // *Journal of Biological Chemistry*. 1946. V. 166. P. 273–281.
32. Studier F.W. // *Protein Expr. Purif.* 2005. V. 41. P. 207–234.
33. Manukhov I.V., Mamaeva D.V., Morozova E.A., Rastorguev S.M., Faleev N.G., Demidkina T.V., Zavilgelsky G.B. // *Biochemistry (Mosc.)*. 2006. V. 71. P. 454–463.
34. Laemmli U.K. // *Nature*. 1970. V. 227. P. 680–685.
35. Dixon M. // *Biochem. J.* 1953, V. 55. P. 170–171.
36. Clausen T., Huber R., Laber B., Pohlentz H.D., Messerschmidt A. // *J. Mol. Biol.* 1996. V. 262. P. 202–224.
37. Inoe H., Inagaki K., Adachi N., Tamura T., Esaki N., Soda K., Tanaka H. // *Biosci. Biotechnol. Biochem.* 2000. V. 64. P. 2336–2343.
38. Nikulin A., Revtovich S., Morozova E., Nevskaya N., Nikonov S., Garber M., Demidkina T. // *Acta Crystallography, Section D*. 2008. V. 64. P. 211–218.
39. Revtovich S.V., Morozova E.A., Khurs E.N., Zakomirdina L.N., Nikulin A.D., Demidkina T.V., Khomutov R.M. // *Biochemistry (Mosc.)*. 2011. V. 76. P. 690–698.
40. Revtovich S.V., Faleev N.G., Morozova E.A., Anufrieva N.V., Nikulin A.D., Demidkina T.V. // *Biochimie*. 2014. V. 101. P. 161–167.
41. Motoshima H., Inagaki K., Kumasaka T., Furuichi M., Inoue H., Tamura T., Esaki N., Soda K., Tanaka N., Yamamoto M. et al. // *J. Biochem.* 2000. V. 128. P. 349–354.
42. Metzler C.M., Harris A.G., Metzler D.E. // *Biochemistry*. 1988. V. 27. P. 4923–4933.
43. Esaki N., Nakayama T., Sawada S., Tanaka H., Soda K. // *Biochemistry*. 1985. V. 24. P. 3857–3862.
44. Bauer A.W., Kirb W.M.M., Sherris J.C., Turck M. // *Am. J. Clin. Pathol.* 1966. V. 36. P. 493–496.

Chemical Polysialylation and *In Vivo* Tetramerization Improve Pharmacokinetic Characteristics of Recombinant Human Butyrylcholinesterase-Based Bioscavengers

S. S. Terekhov^{1*}, I. V. Smirnov^{1,3}, O. G. Shamborant¹, T. V. Bobik¹, D. G. Ilyushin¹, A. N. Murashev², I. A. Dyachenko², V. A. Palikov², V. D. Knorre¹, A. A. Belogurov Jr.^{1,3,4}, N. A. Ponomarenko¹, E. S. Kuzina¹, D. D. Genkin⁵, P. Masson³, A. G. Gabibov^{1,3,4}

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

²Branch of the Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences, Prospekt Nauki, 6, Pushchino, Moscow Region, 142290, Russia

³Kazan' Federal University, Kremlevskaya Str., 18, Kazan', Republic of Tatarstan, 420000, Russia

⁴Institute of Gene Biology of the Russian Academy of Sciences, Vavilova Str., 34/5, Moscow, 119334, Russia

⁵Pharmsynthez, Krasnogo Kursanta Str., 25th, Saint Petersburg, 197110, Russia

*E-mail: sterekhoff@gmail.com

Received 01.09.2015

Copyright © 2015 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 Organophosphate toxins (OPs) are the most toxic low-molecular compounds. The extremely potent toxicity of OPs is determined by their specificity toward the nerve system. Human butyrylcholinesterase (hBChE) is a natural bioscavenger against a broad spectrum of OPs, which makes it a promising candidate for the development of DNA-encoded bioscavengers. The high values of the protective index observed for recombinant hBChE (rhBChE) make it appropriate for therapy against OP poisoning, especially in the case of highly toxic warfare nerve agents. Nevertheless, large-scale application of biopharmaceuticals based on hBChE is restricted due to its high cost and extremely rapid elimination from the bloodstream. In the present study, we examine two approaches for long-acting rhBChE production: I) chemical polysialylation and II) *in-vivo* tetramerization. We demonstrate that both approaches significantly improve the pharmacokinetic characteristics of rhBChE (more than 5 and 10 times, respectively), which makes it possible to use rhBChE conjugated with polysialic acids (rhBChE-CAO) and tetrameric rhBChE (4rhBChE) in the treatment of OP poisonings.

KEYWORDS bioscavenger, biopharmaceutical, biodistribution, butyrylcholinesterase, polysialylation, pharmacokinetics, *in vivo* tetramerization.

ABBREVIATIONS OPs – organophosphates; hAChE – human acetylcholinesterase; 4rhBChE – tetrameric recombinant human butyrylcholinesterase; 4rhBChE-CAO – chemically polysialylated 4rhBChE; MRT – mean residence time; PRAD – proline-rich attachment domain; CAO – Colominic (polysialic) Acid Oxidized.

INTRODUCTION

Organophosphate toxins (OPs), despite their more than 150-year history, remain some of the most topical objects in modern toxicology. OP toxins represent several classes of organophosphorus compounds that irreversibly inhibit human acetylcholinesterase (hAChE). The inhibition of acetylcholinesterase, in turn, leads to the development of the salivation, lacrimation, urination, diaphoresis, gastrointestinal upset, and emesis (miosis) syndrome (SLUDGE(M)). Acute poisoning leads to con-

vulsions, permanent brain damage, respiratory arrest, and death. Currently, OP victims (about 260,000 per year) are mainly suicides. This is especially true for the Western Pacific Region, which accounts for approximately 50% of the total number of suicide attempts [1]. Also, poisonings by organophosphate pesticides often occur among farmers. In addition, there is a potential threat in military use of neuroparalytic warfare poisonous agents or their use in terrorist attacks. The convenient management of OP poisoning [2] includes

combined therapy with muscarinic receptor antagonists (usually atropine) and acetylcholinesterase reactivators (pralidoxime or obidoxime). Unfortunately, this therapy is not a panacea; it does not increase survival in poisoning by organophosphorus pesticides [3] and also does not prevent permanent brain damage.

An alternative approach for the treatment of OP poisonings is the application of bioscavengers – biomolecules binding and inactivating OPs [4–7]. Human butyrylcholinesterase is a natural bioscavenger (suicide inactivator) in OP poisoning [8]. hBChE that has a large active site cavity and a unique similarity to hAChE inactivates a wide range of OPs, often more effectively than hAChE [9]. The use of hBChE in the therapy of OP poisonings not only improves the survival rate, but also obviates the side effects of long-term OP poisoning, including permanent brain damage [10]. Despite the obvious advantages, the application of hBChE in the treatment of OP poisonings is very limited by the high cost of hBChE-based drugs and the extremely rapid elimination ($\tau_{1/2} \approx 2$ min) of monomeric and dimeric recombinant hBChE (rhBChE) forms from the circulation [11]. Thus, the main efforts in the development of an effective therapeutic drug have focused on enhancing rhBChE production [12] and improving the pharmacokinetics of rhBChE-based drugs through chemical conjugation with polyethylene glycol [13–16] and polysialic acids (CAO) [17], or fusing rhBChE with human serum albumin [18]. Recently, we showed [19] that a high production level and, at the same time, a significant improvement in the pharmacokinetic characteristics of rhBChE can be achieved by *in vivo* rhBChE tetramerization. We demonstrated that simulation of natural rhBChE tetramerization [20] in a conventional expression system of the CHO cell line provides effective biotechnological production of a tetrameric rhBChE-based (4rhBChE) biopharmaceutical. 4rhBChE produced by *in vivo* tetramerization had pharmacokinetic characteristics ($\tau_{1/2}$ 32 ± 1.2 h, MRT 43 ± 2 h) similar to those of a tetrameric hBChE drug obtained from human blood plasma [21].

The purpose of this study was to investigate the possibility of further enhancement in the pharmacokinetic characteristics of the 4rhBChE drug by means of chemical polysialylation and to determine the effect of polysialylation on the biodistribution profile of rhBChE-based biopharmaceuticals in mouse models.

MATERIALS AND METHODS

4rhBChE-based biopharmaceuticals used in the study

4rhBChE was produced in CHO-K1 cells transfected with the construct pFUSE PRAD-F2A-BChE. The cells simultaneously expressed genes of the tetramerization

peptide (PRAD-peptide) and human butyrylcholinesterase [19]. rhBChE was obtained as a mixture of oligomers [17] with a predominant content of the dimeric form. BChE was successively purified by affinity chromatography on a XK10/50 column (GE Healthcare, USA) packed with a procainamide-Sepharose sorbent and ion exchange chromatography on a MonoQ 5/50 column (GE Healthcare, USA). According to polyacrylamide gel electrophoresis with Coomassie staining and staining for the specific butyrylcholinesterase activity by the Karnovsky and Roots method [22], the protein purity was greater than 95%.

Chemical polysialylation of rhBChE biopharmaceuticals

rhBChE samples were chemically conjugated to oxidized polysialic acids with a mean molecular weight of 24 kDa (Xenetic Biosciences) by reductive amination according to [17, 23]. The conjugation was performed in 0.1 M potassium phosphate buffer, pH 6.9, with the molar rhBChE : CAO ratio being 1 : 50 per the rhBChE monomer. The final NaBH_3CN concentration was 3 mg/mL. The reaction was conducted at 25 °C for 48 h. The resulting rhBChE-CAO conjugate was purified from reaction by-products by repeated dialysis using Amicon Ultra-15 30K concentrators (Millipore, USA). The efficiency of modification was determined by electrophoresis in 8% polyacrylamide gel (with SDS, but without β -mercaptoethanol). The concentration of active rhBChE was determined by the Ellman's method [23] using 1 mM butyrylthiocholine iodide (Sigma) and 0.5 mM 5,5-dithiobis(2-nitrobenzoic acid) (Sigma) in 0.1 M potassium phosphate buffer, pH 7.0, at 25 °C. The formation of a reaction product, 5-thio-2-nitrobenzoic acid, was detected spectrophotometrically at a wavelength of 412 nm using a product molar absorption coefficient of $13,600 \text{ M}^{-1} \text{ cm}^{-1}$. The BChE concentration was evaluated based on the specific activity of 720 units of activity per 1 mg of pure BChE.

Determination of the pharmacokinetic parameters of rhBChE biopharmaceuticals and rhBChE-CAO conjugates

The concentration of rhBChE, rhBChE-CAO, 4rhBChE, and 4rhBChE-CAO in blood plasma was determined using four groups of BALB/c mice, 18 animals each. Each group consisted of three subgroups, six animals each, for time intervals of 2 min–3 h (subgroup I), 1 h–3 days (subgroup II), and 1–8 days (subgroup III). BChE biopharmaceuticals were administered intravenously at a dose of 200 $\mu\text{g}/\text{mouse}$ (subgroups I and II) and 500 $\mu\text{g}/\text{mouse}$ (subgroup III). Blood samples were collected from the orbital sinus after 2, 5, 10, 15, and 30 min, 1, 2, 3, 6, 9, and 24 h, and 2, 3, 4, 5, 6, and 7 days

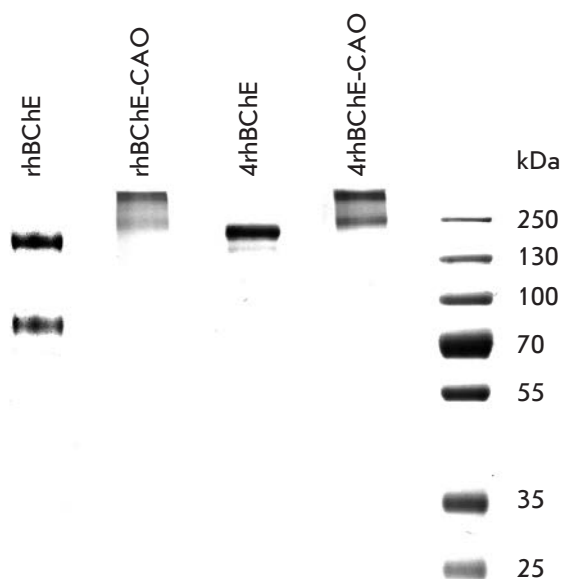


Fig. 1. Electrophoretic analysis of the studied BChE-based biopharmaceuticals and their conjugates with polysialic acids (CAO). Separation was carried out in 8% PAGE under non-reducing conditions, followed by Coomassie R-250 staining. An rhBChE sample is present as a mixture of monomeric, dimeric, and tetrameric forms; a 4rhBChE sample is present exclusively in the tetrameric form. Polysialylated BChE samples obviously have higher molecular weights, but broad bands; this effect was described for chemically polysialylated proteins [17]

after administration. The BChE concentration in mouse blood serum was determined based on the BChE activity according to the Ellman's method [24]. The pharmacokinetic characteristics of the samples were obtained based on fitting the BChE elimination curve using the two-compartment model [17] with the SigmaPlot 12.5 software (Systat software).

Profiling of the rhBChE and rhBChE-CAO conjugate biodistribution

rhBChE and rhBChE-CAO samples were radiolabeled with ^{125}I using chloramine-T at a dose of 10^6 cpm/mg. Labeled rhBChE and rhBChE-CAO samples were intravenously administered to BALB/c line mice (three groups of six animals for each drug) at a dose of 10^5 cpm/mouse. Mice were sacrificed after 0.5, 3, and 48 h, and samples of their blood and tissues were collected and weighed. Collected samples were measured using a WIZARD automated gamma counter (PerkinElmer). Accumulation in tissue was defined as a ratio of the organ specific radioactivity (cpm/g) to the blood specific radioactivity (cpm/mL) at a given time.

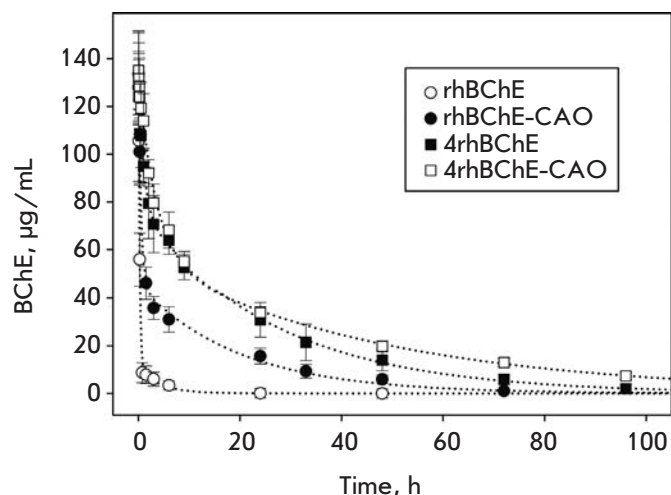


Fig. 2. Pharmacokinetics of the elimination of BChE-based biopharmaceuticals from the circulation after an intravenous injection. To evaluate BChE biopharmaceutical concentrations in blood plasma, 4 groups of BALB/c mice, 18 animals each, were used. The animals were intravenously administered with rhBChE, rhBChE-CAO, 4rhBChE, and 4rhBChE-CAO at doses of 200 and 500 $\mu\text{g}/\text{mouse}$. The BChE concentration in mouse blood serum was determined by the Ellman's method. The pharmacokinetic characteristics of the biopharmaceuticals were obtained by fitting the drug elimination curve to a two-compartment model

RESULTS AND DISCUSSION

Major advances associated with the therapeutic application of rhBChE biopharmaceuticals were achieved for rhBChE chemically modified with polyethylene glycol. Previously, we demonstrated [17] that chemical polysialylation can be used as an alternative modification that repeatedly improves the pharmacokinetic characteristics of rhBChE. The pharmacokinetic characteristics of rhBChE-polysialic acid (rhBChE-CAO) conjugates are inferior to those of rhBChE-polyethylene glycol conjugates [12, 13, 15], but the former have a significant advantage in biodegradability over the latter. In this study, we compared the pharmacokinetic characteristics of rhBChE-CAO and 4rhBChE biopharmaceuticals without chemical modification [19], as well as evaluated the effect of chemical polysialylation on the 4rhBChE-CAO conjugate pharmacokinetics.

Polysialylation of rhBChE and 4rhBChE proceeds with efficiency of over 95% and leads to the formation of high-molecular-weight products with a modification degree of about six CAO molecules per BChE monomer (Fig. 1). The resulting rhBChE-CAO and

Pharmacokinetic characteristics of BChE biopharmaceuticals

Biopharmaceutical	Pharmacokinetic parameters		
	$\tau_{1/2distr.}$, h	$\tau_{1/2el.}$, h	MRT, h
rhBChE	0.2±0.1	3±1	3±1.6
rhBChE-CAO	0.3±0.1	14±2	19±3
4rhBChE	2.4±0.3	33±2	43±4
4rhBChE-CAO	0.8±0.2	19±2	27±3

4rhBChE-CAO conjugates had low toxicity and did not cause the death of the experimental animals after an intravenous injection at a dose of up to 1,500 mg/kg, which, in turn, may indicate a potential for increasing the protective index by more than an order of magnitude relative to the data obtained previously for the warfare agent VR [17].

To evaluate the pharmacokinetic characteristics of the produced BChE drugs, a mouse model of intravenous drug administration and determination of the residual butyrylcholinesterase activity in blood serum were used. The endogenous BChE activity level in mouse blood serum was $2.0 \pm 0.5 \mu\text{g/mL}$, which enabled a highly accurate estimation of the administered drug's concentration. *Figure 2* shows the elimination curves for the studied biopharmaceuticals. It is obvious that practical application of rhBChE without modification is largely complicated by an extremely rapid elimination of rhBChE from the circulation. The modification with polysialic acids provides a more than 5-fold enhancement of the rhBChE pharmacokinetic characteristics (*Table*), which significantly extends the range of its therapeutic applications and enables its use for the prevention of OP poisoning. At the same time, 4rhBChE has characteristics that are more than twice better than those of the rhBChE-CAO conjugate; thereby, 4rhBChE has the longest clearance time from circulation among the studied drugs. The biotechnological production of 4rhBChE is similar to that of rhBChE and is much more economically feasible than the production of the rhBChE-CAO conjugate due to the absence of modification (which uses a 50-fold excess of CAO) and purification stages. At the same time, one could expect that polysialylation of 4rhBChE would lead to a further enhancement of the pharmacokinetic characteristics of 4rhBChE, but this does not occur. The elimination pharmacokinetics of 4rhBChE-CAO and 4rhBChE is almost identical on the first day; at lengthier times, 4rhBChE-CAO is eliminated faster than unmodified 4rhBChE. Therefore, chemical polysialylation repeatedly enhances the pharmacokinetic characteristics of

monomeric and dimeric rhBChE forms but does not improve the 4rhBChE pharmacokinetics. Since hBChE is present in human blood plasma only in the tetrameric form, which ensures its long-term circulation, and chemical polysialylation of 4rhBChE does not lead to an improvement in the pharmacokinetic characteristics of 4rhBChE-CAO compared to those of 4rhBChE, we may assume that the longer circulation of 4rhBChE is primarily associated with no increase in the hydrodynamic radius of 4rhBChE. Apparently, 4rhBChE complex formation leads to masking protein domains responsible for the rapid elimination of rhBChE.

To study the impact of chemical polysialylation on the profile of biodistribution and accumulation of 4rhBChE drugs, experiments with the drugs labeled with the ^{125}I radioisotope were conducted. 4rhBChE and 4rhBChE-CAO were administered intravenously, and their accumulation in different compartments was analyzed after 0.5, 3, and 48 h relative to the appropriate radioactivity of the blood samples (*Fig. 3*). No specific accumulation of 4rhBChE and 4rhBChE-CAO in organs occurs within the first 3 h, but pronounced urinary excretion is observed, which is apparently associated with the biodegradation products. Accumulation of the drugs in the kidneys and liver occurs 48 h after and is significantly more pronounced for 4rhBChE. As mentioned earlier, the elimination pharmacokinetics of 4rhBChE and 4rhBChE-CAO are very similar within the first 24 h, which also manifests itself in the similarity of biodistribution profiles. At the same time, after 48 h, the pharmacokinetic properties of 4rhBChE are better than those of 4rhBChE-CAO. Apparently, this is related to the more pronounced 4rhBChE accumulation in the kidneys, which leads to a reduction in its excretion rate. Along with this, an extremely low rhBChE level in the brain, as well as adipose and muscle tissue, should be noted. Residual radioactivity in these compartments is apparently associated with vascularization, which indicates a limited penetration capability typical of rhBChE.

CONCLUSION

The purpose of this work was to study the impact of alternative approaches for increasing the duration of rhBChE circulation on rhBChE-based drugs pharmacokinetics. The improvement in the pharmacokinetic characteristics of rhBChE-CAO compared to those of rhBChE is apparently associated with an increase in the hydrodynamic radius of rhBChE-CAO and masking of the rhBChE domains (in particular, the C-terminal domain of hBChE) responsible for tetramerization by CAO molecules. A similar effect can be achieved through the production of 4rhBChE. The use of 4rhBChE is an economically attractive alternative to bi-

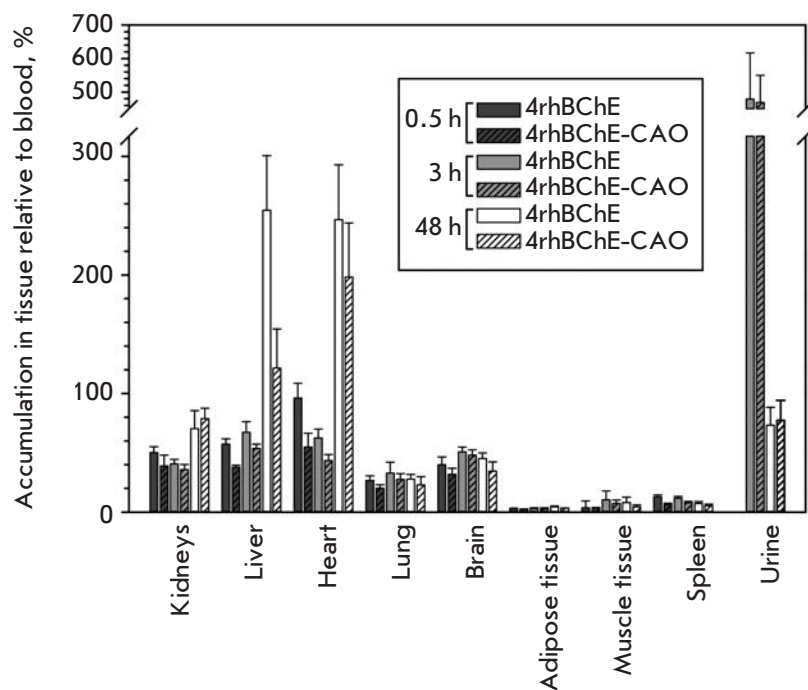


Fig. 3. Biodistribution profiles for ^{125}I -radiolabeled 4rhBChE and 4rhBChE-CAO 0.5, 3, and 48 h after their intravenous administration. ^{125}I -labeled samples were obtained according to the standard protocol using chloramine T and purified by gel filtration. BALB/c mice (3 groups of 6 animals each per protein) were used in the experiments. Samples were injected into the tail vein at a dose of 10^5 cpm/mouse. The animals were euthanized after 0.5, 3, or 48 h. The appropriate organs were isolated, weighed, and used for the radiological analysis on a WIZARD automatic gamma counter (PerkinElmer). Accumulation in tissue was defined as a ratio of the organ specific radioactivity (cpm/g) to the blood specific radioactivity (cpm/mL) at a given time

opharmaceuticals on the basis of modified rhBChE, because 4rhBChE allows one to achieve better pharmacokinetic parameters compared to those of rhBChE and rhBChE-CAO. Chemical modification of 4rhBChE by polysialic acids, in turn, does not lead to further improvement in the 4rhBChE-CAO pharmacokinetics, which may indicate the existence of additional natural mechanisms of 4rhBChE stabilization. At the same time, it should be recognized that further optimization of the polysialylation reaction, complete standardization of the chemical modification process, and use of recently proposed genetic expression constructs [19] may again bring rhBChE-CAO to the fore among potential OP bioscavengers.

The low toxicity of 4rhBChE-based biopharmaceuticals extends the opportunity for bioscavenger application in the treatment of OP poisonings. At the same time, it should be noted that application of 4rhBChE biopharmaceuticals is limited by the need to administer stoichiometric amounts of the enzyme with respect to OP. This fact, in turn, leads to the possibility of the protective index of 4rhBChE therapy (ratio of LD_{50} for animals after treatment to LD_{50} for animals without treatment) being high only in the case of warfare agents (i.e., highly toxic agents with low LD_{50}). Further

improvement of 4rhBChE-based biopharmaceuticals should be obviously associated with the creation of catalytic bioscavengers – enzymes that catalytically inactivate OPs. This will repeatedly reduce the therapeutic dose and extend the capabilities of this therapy for pesticide poisoning, because the high LD_{50} value in this case leads to the necessity of introducing an excessively high amount of the 4rhBChE drug. At the same time, the transition to catalytic bioscavengers should be associated with rapid and effective ($k_2/K_M \approx 10^7 \text{ M}^{-1} \text{ min}^{-1}$) OP elimination [13], which is particularly important in the treatment of OP poisonings [25, 26]. Of particular importance for developing a catalytic bioscavenger will be the issue of standardizing a highly productive clone which is economically viable for launching production, as well as its possible certification to FDA requirements. In the case of the stoichiometric rhBChE-based bioscavengers discussed in this article, the latter condition is absolutely feasible.

This work was supported by a grant under contract with the Russian Ministry of Education and Science RFMEFI60414X0069 and partially was supported by RFBR (grant number 14-04-0067).

REFERENCES

- Gunnell D., Eddleston M., Phillips M.R., Konradsen F. // BMC Public Health. 2007. V. 7. P. 357.
- Eddleston M., Buckley N.A., Eyer P., Dawson A.H. // Lancet. 2008. V. 371. P. 597–607.

- Eddleston M., Eyer P., Worek F., Juszczak E., Alder N., Mohamed F., Senarathna L., Hittarage A., Azher S., Jegannathan K., et al. // PLoS Med. 2009. V. 6. e1000104.
- Nachon F., Brazzolotto X., Trovaslet M., Masson P. // Chem. Biol. Interact. 2013. V. 206. P. 536–544.

RESEARCH ARTICLES

5. Masson P., Lockridge O. // *Arch. Biochem. Biophys.* 2010. V. 494. P. 107–120.
6. Masson P., Rochu D. // *Acta Naturae.* 2009. V. 1. № 1. P. 68–79.
7. Radic Z., Dale T., Kovarik Z., Berend S., Garcia E., Zhang L., Amitai G., Green C., Radic B., Duggan B.M., et al. // *Biochem. J.* 2013. V. 450. P. 231–242.
8. Wille T., Thiermann H., Worek F. // *Arch. Toxicol.* 2014. V. 88. P. 301–307.
9. Shenouda J., Green P., Sultatos L. // *Toxicol. Appl. Pharmacol.* 2009. V. 241. P. 135–142.
10. Sun W., Doctor B.P., Lenz D.E., Saxena A. // *Chem. Biol. Interact.* 2008. V. 175. P. 428–430.
11. Duysen E.G., Bartels C.F., Lockridge O. // *J. Pharmacol. Exp. Ther.* 2002. V. 302. P. 751–758.
12. Huang Y.J., Huang Y., Baldassarre H., Wang B., Lazaris A., Leduc M., Bilodeau A.S., Bellemare A., Cote M., Herskovits P., et al. // *Proc. Natl. Acad. Sci. USA.* 2007. V. 104. P. 13603–13608.
13. Geyer B.C., Kannan L., Garnaud P.-E., Broomfield C.A., Cadieux C.L., Cherni I., Hodgins S.M., Kasten S.A., Kelley K., Kilbourne J., et al. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. P. 20251–20256.
14. Rosenberg Y.J., Gearhart J., Mao L., Jiang X., Hernandez-Abanto S. // *Chem. Biol. Interact.* 2014. V. 210. P. 20–25.
15. Sun W., Luo C., Tipparaju P., Doctor B.P., Saxena A. // *Chem. Biol. Interact.* 2013. V. 203. P. 172–176.
16. Chilukuri N., Sun W., Naik R.S., Parikh K., Tang L., Doctor B.P., Saxena A. // *Chem. Biol. Interact.* 2008. V. 175. P. 255–260.
17. Ilyushin D.G., Smirnov I.V., Belogurov A.A., Dyachenko I.A., Zharmukhamedova T., Novozhilova T.I., Bychikhin E.A., Serebryakova M.V., Kharybin O.N., Murashev A.N., et al. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. P. 1243–1248.
18. Huang Y.J., Lundy P.M., Lazaris A., Huang Y., Baldassarre H., Wang B., Turcotte C., Cote M., Bellemare A., Bilodeau A.S., et al. // *BMC Biotechnol.* 2008. V. 8. P. 50.
19. Terekhov S., Smirnov I., Bobik T., Shamborant O., Zenkova M., Chernolovskaya E., Gladkikh D., Murashev A., Dyachenko I., Palikov V., et al. // *Biochimie.* 2015. V. 118. P. 51–59.
20. Li H., Schopfer L.M., Masson P., Lockridge O. // *Biochem. J.* 2008. V. 411. P. 425–432.
21. Saxena A., Ashani Y., Raveh L., Stevenson D., Patel T., Doctor B.P. // *Mol. Pharmacol.* 1998. V. 53. P. 112–122.
22. Karnovsky M.J., Roots L. // *J. Histochem. Cytochem.* 1964. V. 12. P. 219–221.
23. Smirnov I.V., Vorobiev I.I., Belogurov A.A., Genkin D.D., Deyev S.M., Gabibov A.G. // *Methods in Molecular Biology.* 2015. V. 1321. P. 389–404.
24. Ellman G.L., Courtney K.D., Andres V.Jr., Feather-Stone R.M. // *Biochem. Pharmacol.* 1961. V. 7. P. 88–95.
25. Worek F., Seeger T., Reiter G., Goldsmith M., Ashani Y., Leader H., Sussman J.L., Aggarwal N., Thiermann H., Tawfik D.S. // *Toxicol. Lett.* 2014. V. 231. P. 45–54.
26. Worek F., Seeger T., Goldsmith M., Ashani Y., Leader H., Sussman J.S., Tawfik D., Thiermann H., Wille T. // *Arch. Toxicol.* 2014. V. 88. P. 1257–1266.

2-(2,4-Dioxy-1,2,3,4-Tetrahydropyrimidin-1-yl)-N-(4-Phenoxyphenyl)-Acetamides As a Novel Class of Cytomegalovirus Replication Inhibitors

D. A. Babkov¹, M. P. Paramonova¹, A. A. Ozerov¹, A. L. Khandazhinskaya², R. Snoeck³, G. Andrei³, M. S. Novikov^{1,*}

¹Volgograd State Medical University, Pavshikh Bortsov Sq., 1, Volgograd 400131, Russia

²Engelhardt Institute of Molecular Biology, Russian Academy of Science, Vavilov Str., 32, Moscow, 119991, Russia

³Rega Institute for Medical Research, KU Leuven, Minderbroedersstraat 10, Leuven B-3000, Belgium

*E-mail: m-novikov1@mail.ru

Received 25.05.2015

Copyright © 2015 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 A series of novel uracil derivatives, bearing *N*-(4-phenoxyphenyl)acetamide moiety at N3 of a pyrimidine ring, has been synthesized. Their antiviral activity has been evaluated. It has been found that the novel compounds possess high inhibitory activity against replication of human cytomegalovirus (AD-169 and Davis strains) in HEL cell cultures. In addition, some of the derivatives proved to be inhibitory against varicella zoster virus.

KEYWORDS uracil derivatives, synthesis, antiviral activity, human cytomegalovirus.

ABBREVIATIONS HIV, human immunodeficiency virus, CMV, cytomegalovirus, AIDS, acquired immunodeficiency syndrome, HMDS, hexamethyldisilazane, DMSO, dimethyl sulfoxide, DMF, *N,N*-dimethylformamide, VZV, varicella zoster virus.

Cytomegalovirus (CMV) is widespread in the human population and has been found in people of all geographical regions as well as in representatives of all socio-economic groups [1]. CMV causes a lifelong latent infection that can reactivate periodically. In healthy individuals, the infection is usually asymptomatic [2]; however in individuals with reduced immune status, particularly in AIDS patients [3] and those receiving immunosuppressive therapy after organ transplantation [4], CMV is associated with significant morbidity and mortality. CMV is considered to be the most dangerous cause of congenital diseases. The virus can be transmitted from the mother to the fetus, resulting in a stillbirth, birth defects, and developmental disorders [5].

Ganciclovir, foscarnet, cidofovir and their prodrugs valganciclovir, and cidofovir are used to treat CMV [6]. However, these drugs cause many adverse side effects [6]. Long-term therapy of a CMV infection can lead to the emergence of resistant variants of CMV [7], therefore the search for new highly effective anti-CMV agents is an urgent task.

We have recently synthesized a number of 1-cinamyl-3-benzyl-uracil derivatives which effectively blocked the replication of HIV-1 and CMV in cell cultures [8], and we describe the synthesis and properties of 1-[ω -(phenoxy)alkyl]uracil derivative as an anti-CMV agent [9]. In the continuation of the search for new inhibitors of CMV replication, we synthesized uracil derivatives bearing *N*-(4-phenoxyphenyl)acetamide moiety at N3 of a pyrimidine ring and studied their antiviral properties.

2-Chloro-*N*-(4-phenoxyphenyl)-acetamide (1)

The suspension of 3.9 g (21.06 mmol) of 4-(phenoxy)aniline (2) and 0.15 g of NH₄Cl in 25 mL of HMDS was refluxed for 12 hours until a clear solution was obtained. The excess of HMDS was removed under reduced pressure, and 50 mL of anhydrous 1,2-dichloroethane was added to the residue (dark-colored oily liquid); then, 1.7 ml (21.37 mmol) of chloroacetyl chloride was added dropwise to the solution at 0 °C. The resulting mixture was stirred at 0 °C for 2 hours and allowed to stand overnight at room temperature. The reaction mixture was

then evaporated under reduced pressure on a rotary evaporator and recrystallized from an ethyl acetate-hexane (1:1) mixture. The resulting product was a light purple fine crystalline substance (80% yield), m.p. 105–106 °C, R_f 0.62 (ethyl acetate-hexane, 1:1). $^1\text{H-NMR}$ spectrum (DMSO- D_6), δ , ppm, J (Hz): 4.24 (2H, s, COCH_2), 6.97 (2H, d, $J = 8.7$, H-3', H-5'), 7.00 (2H, d, $J = 9.0$, H-2'', H-6''), 7.10 (1H, t, $J = 7.4$, H-4''), 7.36 (2H, t, $J = 8.5$, H-3'', H-5''), 7.61 (2H, d, $J = 9.0$, H-2', H-6'), 10.30 (1H, s, NH). $^{13}\text{C-NMR}$ -spectrum (DMSO- D_6), δ , ppm: 47.7, 122.2, 123.6, 125.4, 134.2, 138.5, 156.6, 161.4, 168.7.

General procedure for the synthesis of 2-(2,6-dioxy-3,6-dihydropyrimidin-1(2H)-yl)-N-(4-phenoxyphenyl)acetamides (4)-(11)

A mixture of 1.42 mmol of the appropriate 1-substituted-uracil (**12**)–(**19**) and 0.29 g (2.10 mmol) K_2CO_3 in 10 mL of DMF solution was stirred at 80 °C for 1 h, cooled to room temperature, and 2.12 mmol of 2-chloro-N-(4-phenoxyphenyl)acetamide (**1**) was added to the mixture; the reaction mixture was stirred at the same temperature for 24 hours. Then the reaction mixture was filtered, evaporated *in vacuo*, and purified by flash chromatography, followed by recrystallization of the product from ethyl acetate-hexane (1:1).

2-(3-Benzyl-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl)-N-(4-phenoxyphenyl)acetamides (4)

Yield 85%, m.p. 186–187 °C, R_f 0.60 (1,2-dichloroethane-ethyl acetate, 1:1). $^1\text{H-NMR}$ -spectrum (DMSO- D_6) δ , ppm, J (Hz): 4.21 (2H, s, CH_2), 4.53 (2H, s, CH_2), 5.38 (1H, d, $J = 7.8$, H-5), 6.51–6.56 (4H, m, H-4', H-3', H-5', H-4''), 6.65 (2H, d, $J = 8.5$, H-2', H-6'), 6.84–6.94 (6H, m, H-3'', H-5'', H-2''', H-3''', H-5''', H-6'''), 7.15 (2H, d, $J = 8.9$, H-2'', H-6''), 7.43 (1H, d, $J = 7.8$, H-6), 9.88 (1H, s, NH). $^{13}\text{C-NMR}$ -spectrum (DMSO- D_6), δ , ppm: 42.9, 51.0, 100.1, 117.5, 119.1, 120.3, 122.6, 127.1, 127.4, 128.3, 129.5, 134.3, 136.1, 144.1, 150.8, 151.4, 156.9, 161.8, 164.7.

2-[3-(4-Methylbenzyl)-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl]-N-(4-phenoxyphenyl)acetamide (5)

Yield 83%, m.p. 193–194 °C, R_f 0.54 (1,2-dichloroethane-ethyl acetate, 1:1). $^1\text{H-NMR}$ -spectrum (DMSO- D_6) δ , ppm, J (Hz): 2.29 (3H, s, CH_3), 4.65 (2H, s, CH_2), 4.92 (2H, s, CH_2), 5.80 (1H, d, $J = 7.8$, H-5), 6.98 (2H, d, $J = 8.5$, H-3', H-5'), 7.00 (2H, d, $J = 8.9$, H-2', H-6'), 7.10 (1H, dt, $J = 7.3$ and 1.0, H-4''), 7.19 (2H, d, $J = 7.9$, H-2'', H-6''), 7.23 (2H, d, $J = 7.8$, H-3'', H-5''), 7.36 (2H, dt, $J = 7.4$ and 1.2, H-3''', H-5'''), 7.60 (2H, d, $J = 8.8$, H-2''', H-6'''), 7.84 (1H, d, $J = 8.0$, H-6), 10.25 (1H, s, NH). $^{13}\text{C-NMR}$ -spectrum (DMSO- D_6), δ , ppm: 24.9, 47.5, 55.4, 104.7, 122.2, 123.7, 125.0, 127.2, 131.8, 133.4, 134.2, 137.7, 138.9, 141.3, 148.6, 155.4, 156.1, 161.6, 166.4, 169.3.

2-[3-(3,5-Dimethylbenzyl)-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl]-N-(4-phenoxyphenyl)acetamide (6)

Yield 79%, m.p. 99–101 °C, R_f 0.53 (1,2-dichloroethane-ethyl acetate, 1:1). $^1\text{H-NMR}$ -spectrum (DMSO- D_6) δ , ppm, J (Hz): 2.24 c (6H, CH_3), 4.64 c (2H, CH_2), 4.87 c (2H, CH_2), 5.80 d (1H, $J = 7.9$, H-5), 6.92 c (3H, H-2', H-4', H-6'), 6.96 d (2H, $J = 8.0$, H-2'', H-6''), 6.98 d (2H, $J = 8.9$, H-3'', H-5''), 7.09 t (1H, $J = 7.3$, H-4''), 7.35 t (2H, $J = 7.8$, H-3''', H-5'''), 7.58 d (2H, $J = 8.8$, H-2''', H-6'''), 7.82 d (1H, $J = 7.9$, H-6), 10.29 c (1H, NH). $^{13}\text{C-NMR}$ -spectrum (DMSO- D_6), δ , ppm: 25.1, 47.5, 55.6, 104.7, 122.1, 123.7, 124.9, 127.2, 129.5, 133.4, 134.2, 138.9, 140.5, 142.0, 148.7, 155.4, 156.1, 161.5, 166.4, 169.3.

2-(3-Cinnamyl-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl)-N-(4-phenoxyphenyl)acetamide (7)

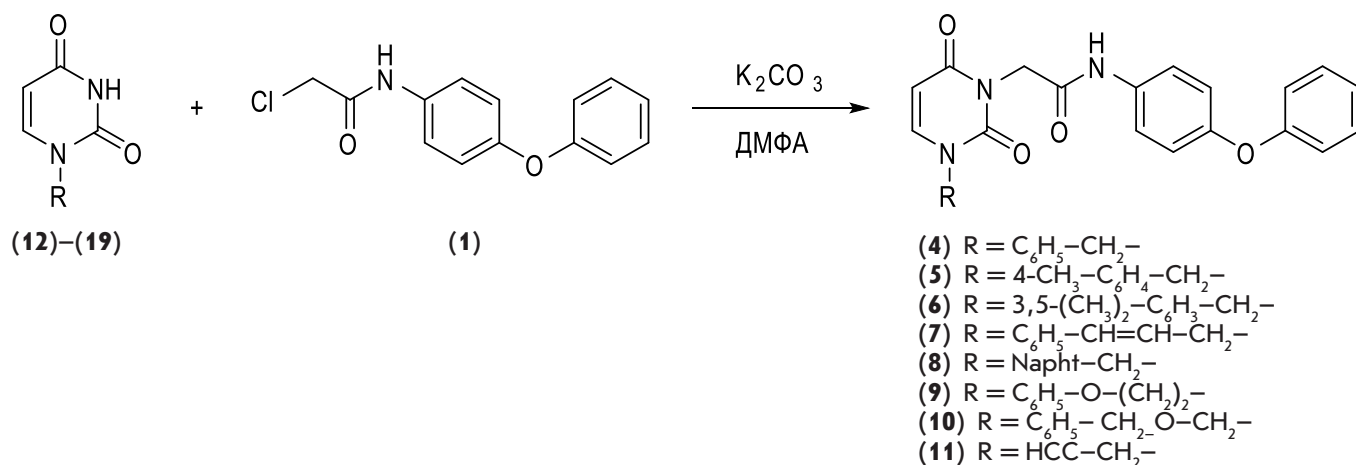
Yield 88%, m.p. 184–185 °C, R_f 0.41 (1,2-dichloroethane-ethyl acetate, 1:1). $^1\text{H-NMR}$ -spectrum (DMSO- D_6) δ , ppm, J (Hz): 4.54 (2H, d, $J = 5.6$, CH_2), 4.64 (2H, s, CH_2), 5.81 (1H, d, $J = 8.0$, H-5), 6.34 (1H, dt, $J = 6.0$, =CH-), 6.60 (1H, d, $J = 16.0$, PhCH=), 6.96 (2H, d, $J = 7.8$, H-3'', H-5''), 6.98 (2H, d, $J = 8.9$, H-2'', H-6''), 7.09 (1H, t, $J = 7.3$, H-4''), 7.25 (1H, t, $J = 7.4$, H-4'), 7.33 (2H, t, $J = 7.8$, H-3''', H-5'''), 7.35 (2H, t, $J = 8.4$, H-3', H-5'), 7.43 (2H, d, $J = 7.5$, H-2', H-6'), 7.58 (2H, d, $J = 8.9$, H-2''', H-6'''), 7.77 (1H, d, $J = 7.8$, H-6), 10.28 (1H, s, NH). $^{13}\text{C-NMR}$ -spectrum (DMSO- D_6), δ , ppm: 47.4, 54.2, 104.7, 122.1, 123.7, 125.0, 127.2, 128.0, 130.7, 132.2, 132.9, 134.2, 137.0, 138.9, 140.1, 148.4, 155.2, 156.1, 161.5, 166.5, 169.3.

2-[3-(Napht-1-ylmethyl)-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl]-N-(4-phenoxyphenyl)acetamide (8)

Yield 85%, m.p. 197–198.5 °C, R_f 0.57 (1,2-dichloroethane-ethyl acetate, 1:1). $^1\text{H-NMR}$ -spectrum (DMSO- D_6) δ , ppm, J (Hz): 4.72 (2H, s, CH_2), 5.49 (2H, s, CH_2), 5.85 (1H, d, $J = 8.0$, H-5), 6.99 (2H, d, $J = 7.8$, H-3'', H-5''), 7.02 (2H, d, $J = 9.0$, H-2'', H-6''), 7.11 (1H, t, $J = 7.3$, H-4''), 7.36 (1H, d, $J = 7.1$, H-4'), 7.37 (2H, dt, $J = 8.6$ and 0.9, H-3''', H-5'''), 7.51 (1H, t, $J = 7.9$, H-6'), 7.57–7.63 (4H, m, H-3', H-7', H-2''', H-6'''), 7.76 (1H, d, $J = 7.8$, H-6), 7.92 (1H, d, $J = 8.3$, H-8'), 7.99 (1H, d, $J = 7.7$, H-4'), 8.12 (1H, d, $J = 8.2$, H-5'), 10.28 (1H, s, NH). $^{13}\text{C-NMR}$ -spectrum (DMSO- D_6), δ , ppm: 47.6, 53.2, 105.1, 122.2, 123.7, 125.0, 127.2, 129.1, 129.7, 130.4, 130.9, 132.5, 132.9, 134.2, 134.6, 136.1, 137.6, 138.9, 148.3, 155.6, 156.2, 161.6, 166.4, 169.3.

2-[3-[2-(Phenoxy)ethyl]-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl]-N-(4-phenoxyphenyl)acetamide (9)

Yield 90%, m.p. 154–155 °C, R_f 0.46 (1,2-dichloroethane-ethyl acetate, 1 : 1). $^1\text{H-NMR}$ -spectrum (DMSO- D_6) δ ,



ppm, J (Hz): 4.16 (2H, d, $J = 5.4$, NCH₂), 4.21 (2H, d, $J = 5.4$, OCH₂), 4.64 (2H, s, CH₂), 5.78 (1H, d, $J = 8.0$, H-5), 6.94–7.00 (7H, m, H-2', H-4', H-6', H-2'', H-3'', H-5'', H-6''), 7.10 (1H, t, $J = 7.3$, H-4'''), 7.29 (2H, t, $J = 7.9$, H-3''', H-5'''), 7.36 (2H, t, $J = 7.6$ and 1.2, H-3', H-5'), 7.59 (2H, d, $J = 8.9$, H-2''', H-6'''), 7.78 (1H, d, $J = 7.9$, H-6), 10.24 (1H, s, NH). ¹³C-NMR-spectrum (DMSO-D₆), δ , ppm: 29.7, 47.4, 52.4, 69.4, 104.0, 118.9, 122.2, 123.7, 125.0, 125.3, 127.2, 133.8, 134.1, 138.9, 149.3, 155.4, 156.2, 161.6, 162.2, 166.4, 169.3.

2-[3-(Benzyloxymethyl)-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl]-N-(4-phenoxyphenyl) acetamide (10)

Yield 84%, m.p. 163–164°C, R_f 0.47 (1,2-dichloroethane–ethyl acetate, 1 : 1). ¹H-NMR-spectrum (DMSO-D₆) δ , ppm, J (Hz): 4.59 (2H, s, CH₂), 4.63 (2H, s, CH₂), 5.26 (2H, s, CH₂), 5.83 (1H, d, $J = 7.8$, H-5), 6.96 (2H, d, $J = 7.9$, H-3'', H-5''), 6.99 (2H, d, $J = 8.9$, H-2'', H-6''), 7.09 (1H, dt, $J = 7.6$ and 1.0, H-4'''), 7.26–7.38 (7H, m, C₆H₅, H-3''', H-5'''), 7.59 (2H, d, $J = 9.1$, H-2''', H-6'''), 7.84 (1H, d, $J = 8.0$, H-6), 10.32 (1H, s, NH). ¹³C-NMR-spectrum (DMSO-D₆), δ , ppm: 43.2, 70.4, 77.2, 100.9, 117.9, 119.5, 120.7, 123.0, 127.69, 127.74, 128.3, 130.0, 134.7, 137.4, 143.8, 151.3, 151.9, 157.4, 162.1, 165.1.

2-(3-Propargyl-2,6-dioxo-3,6-dihydropyrimidin-1(2H)-yl)-N-(4-phenoxyphenyl)acetamide (11)

Yield 81%, m.p. 226–228°C, R_f 0.68 (1,2-dichloroethane–ethyl acetate, 1:1). ¹H-NMR-spectrum (DMSO-D₆) δ , ppm, J (Hz): 3.42 (1H, s, =CH), 4.59 (2H, s, CH₂), 4.60 (2H, d, $J = 8.0$, CH₂), 5.81 (1H, d, $J = 8.0$, H-5), 6.95 (2H, d, $J = 7.7$, H-3'', H-5''), 6.98 (2H, d, $J = 8.9$, H-2'', H-6''), 7.08 (1H, t, $J = 7.6$, H-4'''), 7.35 (2H, dt, $J = 8.5$ and 1.1, H-3''', H-5'''), 7.56 (2H, d, $J = 8.9$, H-2''', H-6'''), 7.80 (1H, d, $J = 7.9$, H-6), 10.33 (1H, s, NH). ¹³C-NMR-spectrum (DMSO-D₆), δ , ppm: 42.0, 47.4, 80.3, 82.4, 105.1, 122.1, 123.7, 125.0, 127.3, 134.2, 138.8, 147.6, 154.8, 156.1, 161.5, 166.3, 169.2.

Antiviral research

Activity of the compounds was evaluated against the following viruses: thymidine kinase deficient (TK-) herpes simplex virus type 1 (HSV-1) KOS strain, HSV-1 KOS strain resistant to acyclovir (ACVr), herpes simplex virus type 2 Lyons and G strains, CMV (AD-169 and Davis strains), varicella-zoster (VZV, OKA and YS strains), vaccinia virus Lederle strain, respiratory syncytial virus (Long strain), vesicular stomatitis virus, Coxsackie virus B4, parainfluenza virus 3, influenza A (sub-types H1N1, H3N2), influenza virus B, reovirus-1 virus, Sindbis virus, and Punta Toro virus. Investigations were carried out as described in [9].

2-Chloro-N-(4-phenoxyphenyl)acetamide (1) was synthesized as described previously [10]. The uracil derivatives substituted at N¹ (12)–(19) were obtained by condensation of equimolar amounts of 2,4-bis-(trimethylsilyloxy)-pyrimidine and arylmethylchloride/bromide as described in [8]. The treatment with equimolar amount of the chloride (1) in DMF in the presence of K₂CO₃, as shown, resulted in the target 4-phenoxyacetanilides (4)–(11) with 79–90% yields.

Anti-CMV properties of the uracil derivatives (4)–(11) were studied in HEL cell culture against CMV (AD-169 and Davis strains). It has been found that certain compounds of this series exhibit strong inhibitory activity against CMV, which is comparable with the effect of ganciclovir. The uracil derivatives substituted at position 1 of the pyrimidine ring with benzyl (compound (4)) and 3,5-dimethylbenzyl (compound (6)) were the most active. They inhibited CMV replication with EC₅₀ = 3.06–8.9 μ M. Other modifications of the structure resulted in complete loss of inhibitory activity.

It has also been found that the compounds (4) and (6) exhibit significant activity against VZV. They blocked VZV replication (Oka strain) in a HEL cell culture with EC₅₀ = 8.18 μ M (compound (4)) and 17.0 μ M (compound (6)), which is inferior to the protective action of acyclo-

vir ($EC_{50} = 1.33 \mu\text{M}$) and brivudine ($EC_{50} = 0.026 \mu\text{M}$), currently used to treat infections caused by this virus [11]. However, the thymidine kinase deficient VZV mutant strain (07-1), which is resistant to acyclovir and brivudine, was susceptible to 1-benzyl-3-acetanilide uracil derivatives with $EC_{50} = 6.68 \mu\text{M}$ (compound (4)) and $16.1 \mu\text{M}$ (compound (6)).

Therefore, the uracil derivatives whose synthesis is described in this work represent a new class of inhibitors of CMV reproduction whose effect is comparable to that of ganciclovir. Furthermore, some compounds of

this series have pronounced inhibitory effect on VZV, both the wild-type strain (OKA) and the strain (07-1) resistant to the action of acyclovir. The data demonstrate that it is a promising direction for the development of new effective antiviral agents.

This work was supported by RFBR (grant number 13-04-01391_A), the biological part of the work supported by the grant GOA 10/014.

REFERENCES

1. Cannon M.J., Schmid D.S., Hyde T.B. // Rev. Med. Virol. 2010. V. 20. P. 202–213.
2. Gandhi M.K., Khanna R. // Lancet Infect. Dis. 2004. V. 4. P. 725–738.
3. Baroco A.L., Oldfield E.C. // Curr. Gastroenterol. Rep. 2008. V. 10. P. 409–416.
4. Nashan B., Gaston R., Emery V., Säemann M.D., Mueller N.J., Couzi L., Dantal J., Shihab F., Mulgaonkar S., Seun K.Y., et al. // Transplantation. 2012. V. 93. P. 1075–1085.
5. Dollard S.C., Grosse S.D., Ross D.S. // Rev. Med. Virol. 2007. V. 17. P. 355–363.
6. Ahmed A. // Infect Disord Drug Targets. 2011. V. 5. P. 475–503.
7. Lurain N.S., Chou S. // Clin. Microb. Rev. 2010. V. 23. P. 689–712.
8. Novikov M.S., Valuev-Elliston V.T., Babkov D.A., Paramonova M.P., Ivanov A.V., Gavryushov S.A., Khandazhinskaya A.L., Kochetkov S.N., Pannecouque C., Andrei G., et al. // Bioorg. Med. Chem. 2013. V. 21. P. 1150–1158.
9. Novikov M.S., Babkov D.A., Paramonova M.P., Khandazhinskaya A.L., Ozerov A.A., Chizhov A.O., Andrei G., Snoeck R., Balzarini J., Seley-Radtke K.L. // Bioorg. Med. Chem. 2013. V. 21. P. 4151–4157.
10. Novikov M.S., Babkov D.A., Paramonova M.P., Chizhov A.O., Khandazhinskaya A.L., Seley-Radtke K.L. // Tetrahedron Lett. 2013. V. 54. P. 576–578.
11. De Clercq E. // Med. Res. Rev. 2005. V. 25. P. 1–20.

Mutations in Parkinson's Disease-Associated *PARK2* Gene Are Accompanied by Imbalance in Programmed Cell Death Systems

E. V. Konovalova¹, O. M. Lopacheva^{1,2}, I. A. Grivennikov³, O. S. Lebedeva³, E. B. Dashinimaev⁴, L. G. Khaspekov¹, E. Yu. Fedotova¹, S. N. Illarionov^{1*}

¹Research Center of Neurology, Volokolamskoe Shosse, 80, Moscow, 125367, Russia

²International Biotechnological Center, Lomonosov Moscow State University, Leninskie Gory, 1/12, Moscow, 119991, Russia

³Institute of Molecular Genetics, Russian Academy of Sciences, Kurchatov Square, 2, Moscow, 123182, Russia

⁴Koltsov Institute of Developmental Biology, Russian Academy of Sciences, Vavilova Str., 26, Moscow, 119334, Russia

*E-mail: snillario@gmail.com

Received 04.06.2015

Copyright © 2015 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 Parkinson's disease is caused by the degeneration of midbrain dopaminergic neurons. A rare recessive form of the disease may be caused by a mutation in the *PARK2* gene, whose product, Parkin, controls mitophagy and programmed cell death. The level of pro- and anti-apoptotic factors of the Bcl-2 family was determined in dopaminergic neurons derived from the induced pluripotent stem cells of a healthy donor and a Parkinson's disease patient bearing *PARK2* mutations. Western blotting was used to study the ratios of Bax, Bak, Bcl-2, Bcl-XL, and Bcl-W proteins. The pro-apoptotic Bak protein level in *PARK2*-neurons was shown to be two times lower than that in healthy cells. In contrast, the expression of the anti-apoptotic factors Bcl-XL, Bcl-W, and Bcl-2 was statistically significantly higher in the mutant cells compared to healthy dopaminergic neurons. These results indicate that *PARK2* mutations are accompanied by an imbalance in programmed cell death systems in which non-apoptotic molecular mechanisms play the leading role.

KEYWORDS Parkinson's disease, dopaminergic neurons, induced pluripotent stem cells, *PARK2*, mutation, programmed cell death.

Parkinson's disease (PD) is a common neurodegenerative disorder caused by lesions to pigmented neurons in the midbrain substantia nigra and degeneration of the dopaminergic nigrostriatal pathway. Treatment of PD is symptomatic and does not prevent further loss of nigral cells, which necessitates identification of the key elements in the disease pathogenetic cascade [1].

Genetics plays an important role in the development of PD. Monogenic forms account for approximately 10% of all PD cases [1, 2], while the other (sporadic) cases are of multifactorial nature. A total of 20 genetic PD loci have been identified to date [2]. One of the involved genes, *PARK2*, is located on the 6q25.2-27 chromosome and associated with the development of a particular form of the disease: an autosomal recessive PD, which is characterized by an early onset [3]. Mutations in *PARK2* cause about 15% of familial and 4% of spo-

radic PD cases with the onset before the age of 40 years [3–5]. The *PARK2* gene encodes cytosolic ubiquitin-E3-ligase, the Parkin protein. The main Parkin function is to regulate mitophagy. It acts in tandem with the PINK1 mitochondrial protein, which is a product of another gene of autosomal recessive Parkinson's disease [6]. The sequence of molecular events occurs as follows: the dysfunctional mitochondrion is depolarized, thereby stabilizing PINK1; the latter recruits Parkin from the cytosol and activates it during its delivery to the mitochondrion, using the PINK1-kinase activity; then, the activated Parkin initiates selective autophagy of the damaged organelle [7].

Therefore, the interaction between Parkin and PINK1 controls the state of mitochondria. Currently, Parkin is regarded as a polyvalent neuroprotective agent that plays the key role in the survival of dopaminergic neurons upon exposure to various neurotoxins [8].

Studying *PARK2*/Parkin cell biology and neurochemistry is an extremely important area of modern neuroscience. The subject of our research is mutant Parkin expressed in cells of a patient with a rare *PARK2*-associated PD form. The dopaminergic neuron-enriched culture of nerve cells produced by reprogramming of dermal fibroblasts into induced pluripotent stem cells (iPSCs), followed by their directed differentiation, was used as a model system [9, 10]. In order to elucidate the disease pathogenesis, we examined the dynamics of several pro- and anti-apoptotic factors in the mutant cell culture in comparison to the neuron culture derived from a healthy donor.

We used cell cultures of a healthy donor and a patient with autosomal recessive juvenile-onset PD, who carried compound heterozygous mutations in the *PARK2* gene (del202-203AG and IVS1+1G/A). Previously, we derived fibroblasts from skin biopsies of the donor and the patient. To generate iPSC clones, the fibroblasts were re-programmed using original lentiviral vectors with pluripotency factors (LeGO-hOCT4, LeGO-hSOX2-IRES-GFP, LeGO-hc-Myc, and LeGO-hKLF4). The cells were then differentiated into the neuronal type using the differentiation factors. The techniques for producing and re-programming fibroblasts, as well as inducing differentiation of iPSCs into specialized dopaminergic neurons, were described previously [10].

Differentiation of cells into neurons was confirmed by immunohistochemical staining with antibodies to β -III-tubulin and tyrosine hydroxylase (TH). Western blotting, which was performed according to standard procedures, was used to determine the protein levels in cell cultures.

The results were processed using Microsoft Excel and the GraphPad Prism 6 software.

During differentiation, Po2 and Tr5 cell lines expressed neuronal markers that were detected by immunohistochemical staining (*Fig. 1*). The β -III-tubulin protein is a classic marker of early neuronal differentiation, and TH is traditionally considered to be a specific marker of neuronal dopaminergic differentiation. As can be seen from *Fig. 1*, all iPSCs undergo differentiation since, all DAPI-stained cell nuclei are located within neurons (β -III-tubulin and TH-positive). The number of dopaminergic-differentiated neurons is approximately 80% of the total number of β -III-tubulin-positive cells.

The ratio of pro- and anti-apoptotic factors is different in Po2 and Tr5 cell cultures derived from the healthy donor and the carrier of *PARK2* mutations, respectively (*Fig. 2*).

The pro-apoptotic Bak protein level in Tr5 mutant dopaminergic neurons was statistically significantly

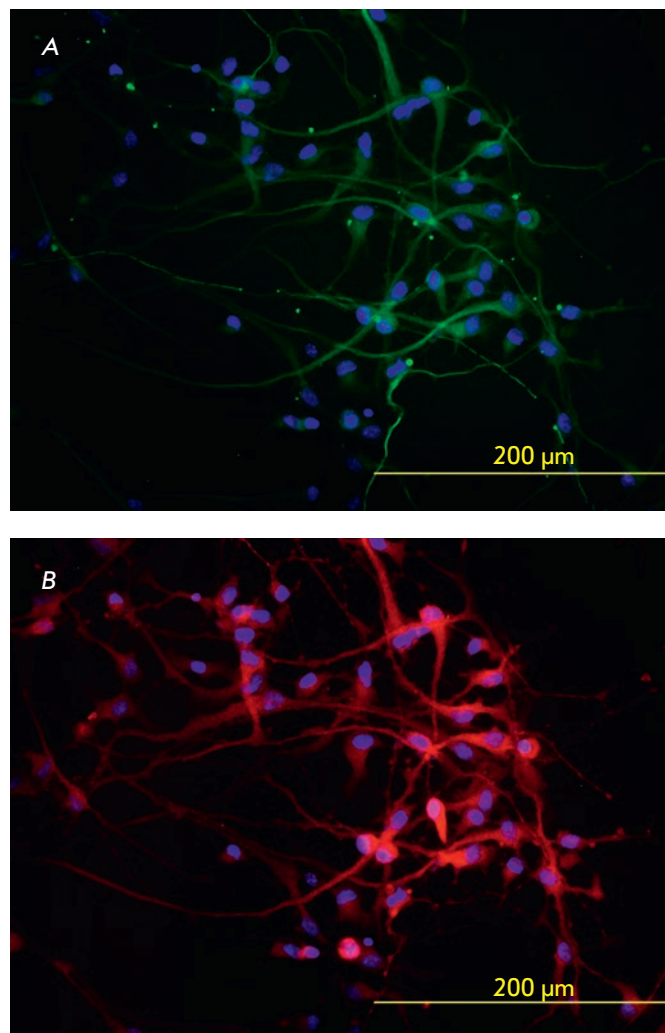


Fig. 1. Immunohistochemical staining of iPSC-differentiated neurons. *A* – staining for β -III-tubulin (green); *B* – staining for TH (red); cell nuclei are shown in blue (DAPI staining)

lower (by 58%) than that in healthy donor cells in the Po2 culture ($p < 0.001$). No significant differences in the pro-apoptotic Bax protein level were observed in the cell cultures under study.

The levels of the anti-apoptotic proteins Bcl-2, Bcl-W, and Bcl-XL were statistically significantly higher in Tr5 cells expressing mutant *PARK2*. For example, the levels of the Bcl-2 protein, Bcl-W protein, and Bcl-XL protein in mutant cells compared to those in wild-type cells were 222% ($p < 0.01$), 150.5% ($p < 0.01$), and 55% ($p < 0.05$) higher, respectively.

Each sample contained 20 μ g of the total protein. Uniformity of protein loading onto gel was additionally monitored using the GAPDH protein, whose signal level was similar in all samples (*Fig. 2B*).

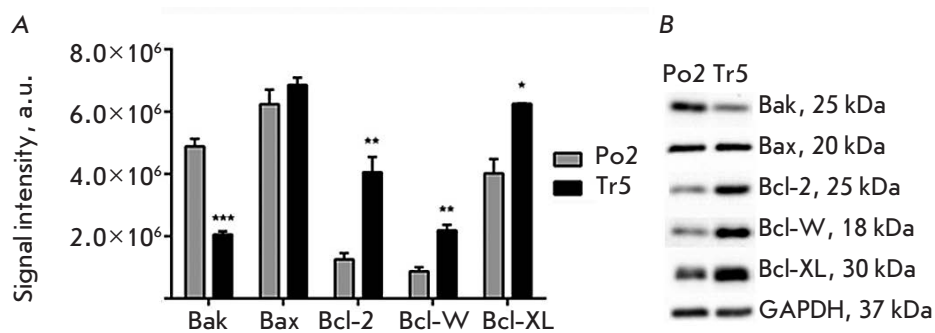


Fig. 2. Quantitative analysis of pro- and anti-apoptotic proteins in the Po2 cell culture derived from a healthy donor and in the Tr5 culture derived from a PD patient carrying *PARK2* mutations (Western blotting). **A** – signal intensity (arbitrary units) of Bak, Bax, Bcl-2, Bcl-W, and Bcl-XL bands ($n = 3$). $p < 0.001$ (***) , $p < 0.01$ (**), and $p < 0.05$ (*) for comparison of Po2 and Tr5 cultures. **B** – representative pictures of Western blotbands

Several variants of programmed cell death have been described for PD and other neurodegenerative diseases: classical apoptosis, autophagic pathway, AIF/PARP-dependent pathway, paraptosis, etc [11, 12]. In some cases, however, cell death patterns do not correspond to a single mechanism and have a complex nature. The release of cytochrome *c* from mitochondria plays an important role in cell death [13]. Cytochrome *c* activates the cytosolic protein Apaf-1 and pro-caspase-9 to form the apoptosome, and its release is strictly regulated in the cell by the pro- and anti-apoptotic factor ratio. The imbalance in the system, which we found in *PARK2*-compromised neurons, may be associated with the functional properties of defective Parkin, but this hypothesis needs further confirmation in cultures with normal and mutant genotypes.

The Parkin protein acts as a mitochondrial neuroprotector. In particular, a specific autonomic influence of Parkin on the mitochondrial mechanisms governing cytochrome *c* release and triggering apoptosis reactions was found [14]. Recently, a Parkin-like cytoplasmic protein associated with p53 (PARC) was identified; like Parkin, PARC is a E3-ligase and initiates proteasomal degradation of cytochrome *c* [15]. Proteasomal degradation of ARTS, which is a mitochondrial protein initiating an early (prior to the release of cytochrome *c*) step of the caspase pathway, may be an important stage in the protective action of Parkin [16]. However, the detailed molecular mechanisms of Parkin-associated neuroprotection still remain unclear.

A feature of the Tr5 culture is the lack of normal Parkin, since the neurons of this culture have two recessive inactivating mutations in the *PARK2* gene (frameshift deletion del202-203A and IVS1+1G/A splicing mutation). Thus, the Parkin-mediated selective processing of mitochondrial proteins and cytochrome *c* is disrupted in the mutant culture. In this

case, in the Tr5 culture, we found a statistically significant decrease in the pro-apoptotic Bak protein concentration instead of the expected initiation of apoptosis and, conversely, an increase in the level of all tested anti-apoptotic proteins of the Bcl family. These results are similar to the data on an increased Bcl-XL level *in vivo* and *in vitro*, in a paraquat-induced Parkinsonism model [17]. Prevention of apoptotic Bax protein translocation to the mitochondria is considered to be one of the functions of Parkin [18]; however, dysregulation of this process in *PARK2*-mutant cells does not affect the expression of Bax and may be not associated with the changes in its level in the culture. In recent years, microglial activation and inflammation have been considered to be important in the cytotoxic mitochondrial cascade in PD [19]; these processes do not require the induction of mitochondrial apoptosis [20], which is indirectly confirmed by our data.

CONCLUSIONS

1. We characterized the Bax/Bak/Bcl system in the case of autosomal recessive PD with a complex *PARK2* mutation: *PARK2* mutations were found to lead to a complex imbalance in programmed cell death systems, in which the non-apoptotic molecular mechanisms likely play the leading role.
2. These preliminary data need to be confirmed in dopaminergic neurons derived from other homozygous and heterozygous carriers of *PARK2* mutations.
3. These findings demonstrate that approaches to the treatment of neurodegenerative diseases using inhibition of apoptosis [11, 12] may be ineffective in the case of *PARK2*-associated PD.

This work was supported by the Russian Science Foundation (grant № 14-15-01047).

REFERENCES

1. Jenner P, Morris H.R., Robbins T.W., Goedert M., Hardy J., Ben-Shlomo Y., Bolam P, Burn D, Hindle J.V., Brooks D. // *J. Parkinson's Dis.* 2013. V. 3. P. 1–11.
2. Bonifati V. // *Parkinsonism Relat. Disord.* 2014. V. 20 (Suppl. 1). P. S23–S28.
3. Zagorovskaya T.B., Illarioshkin S.N., Slominskii P.A., Ivanova-Smolenskaya I.A., Markova E.D. Limborskaya S.A., Levin O.S., Miloserdova O.V., Proskokova T.N., Bagyeva G.H., Bris A. // *S.S. Korsakov zhurn. nevrologii i psikiatrii.* 2004. V. 8. P. 66–72 [In Russian].
4. Periquet M., Lücking C.B., Vaughan J.R., Bonifati V., Dürr A., De Michele G., Horstink M.W., Farrer M., Illarioshkin S.N., Pollak P., et al. // *Am. J. Hum. Genet.* 2001. V. 68. P. 617–626.
5. Kilariski L.L., Pearson J.P., Newsday V., Majounie E., Knipe M.D., Misbahuddin A., Chinnery P.F., Burn D.J., Clarke C.E., Marion M.H., et al. // *Mov. Disord.* 2012. V. 27. P. 1522–1529.
6. Park J., Lee S.B., Lee S., Kim Y., Song S., Kim S., Bae E., Kim J., Shong M., Kim J.M., Chung J. // *Nature.* 2006. V. 441. P. 1157–1161.
7. Matsuda N., Sato S., Shiba K., Okatsu K., Saisho K., Gautier C.A., Sou Y.S., Saiki S., Kawajiri S., Sato F., et al. // *J. Cell Biol.* 2010. V. 189. P. 211–221.
8. Lim K.L., Ng X.H., Grace L.G., Yao T.P. // *Antioxid. Redox. Signal.* 2012. V. 16. P. 935–949.
9. Takahashi K., Yamanaka S. // *Cell.* 2006. V. 126. P. 663–676.
10. Lebedeva O.S., Lagarkova M.A., Kisilev S.L., Mukhina I.V., Vedunova M.V., Usova O.V., Stavrovskaya A.V., Yamschikova N.G., Fedotova E.Yu., Grivennikov I.A. et al. // *Neirokhimiya* 2013. V. 3. P. 233–241.
11. Esposito E., Cuzzocrea S. // *Curr. Med. Chem.* 2010. V. 17. P. 2764–2774.
12. Lockshin R.A., Zakerib Z. // *J. Cell Mol. Med.* 2007. V. 11. P. 1214–1224.
13. Kim J., Yang Y., Song S.S., Na J.H., Oh K.J., Jeong C., Yu Y.G., Shin Y.K. // *Biophys. J.* 2014. V. 107. P. 1601–1608.
14. Berger A.K., Cortese G.P., Amodeo K.D., Weihofen A., Letai A., LaVoie M.J. // *Hum. Mol. Genet.* 2009. V. 18. P. 4317–4328.
15. Gama V., Swahari V., Schafer J., Kole A.J., Evans A., Huang Y., Cliffe A., Golitz B., Sciaky N., Pei X.H., et al. // *Sci. Signal.* 2014. V. 7(334). P. ra67. doi: 10.1126/scisignal.2005309.
16. Kemeny S., Dery D., Loboda Y., Rovner M., Lev T., Zuri D., Finberg J.P., Larisch S. // *PLoS One.* 2012. V. 7. P. e38837.
17. Fei Q., McCormack A.L., Di Monte D.A., Ethell D.W. // *J. Biol. Chem.* 2008. V. 283. P. 3357–3364.
18. Charan R.A., Johnson B.N., Zaganelli S., Nardozi J.D., LaVoie M.J. // *Cell Death Dis.* 2014. V. 5. P. e1313. doi:10.1038/cddis.2014.278.
19. Yan J., Fu Q., Cheng L., Zhai M., Wu W., Huang L., Du G. // *Mol. Med. Rep.* 2014. V. 10. P. 2223–2233.
20. Allam R., Lawlor K.E., Yu E.C., Mildenhall A.L., Moujalled D.M., Lewis R.S., Ke F., Mason K.D., White M.J., Stacey K.J., et al. // *EMBO Rep.* 2014. V. 15. P. 982–990.

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 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 symbols (approximately 40 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 (20 pages in A4 format, including tables and references). They should contain no more than ten figures. Lengthier articles can only be accepted with the preliminary consent of the editors.

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 should be sent to the editors in electronic form: the text should be in Windows Microsoft Word 2003 format, and the figures should be in TIFF format with each image in a separate file. In a separate file there should be a translation in English of: the article's title, the names and initials of the authors, the full name of the scientific organization and its departmental affiliation, the abstract, the references, and figure captions.

MANUSCRIPT FORMATTING

The manuscript should be formatted in the following manner:

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

(1,500 characters).

- Keywords (3 – 6). These should include the field of research, methods, experimental subject, and the specifics of the work. List of abbreviations.

- INTRODUCTION
- EXPERIMENTAL PROCEDURES
- RESULTS AND DISCUSSION
- CONCLUSION

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

- REFERENCES

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

RECOMMENDATIONS ON THE TYPING AND FORMATTING OF THE TEXT

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

GUIDELINES FOR AUTHORS

- 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 publication order is established by the date of acceptance of the article. 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.

A manuscript which has been sent back to the authors for improvements requested by the editors and/or reviewers is reviewed again, after which the editorial board makes another decision on whether the article can be accepted for publication. The published article has the submission and publication acceptance dates set at the beginning.

The return of an article to the authors for improve-

ment does not mean that the article has been accepted for publication. After the revised text has been received, a decision is made by the editorial board. The author must return the improved text, together with the original text and 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

ActaNaturae

SUBSCRIPTION TO

Acta Naturae journal focuses upon interdisciplinary research and developments at the intersection of various spheres of biology, such as molecular biology, biochemistry, molecular genetics, and biological medicine.

Acta Naturae journal is published in Russian and English by Park Media company. It has been included in the list of scientific journals recommended by the State Commission for Academic Degrees and Titles of the Ministry of Education and Science of the Russian Federation and the Pubmed abstracts database.



SUBSCRIBE AT THE EDITORIAL OFFICE

Leninskie Gory, 1-75G, Moscow, 119234 Russia
Telephone: +7 (495) 930-87-07, 930-88-51
Bio-mail: podpiska@biorf.ru
Web site: www.actanaturae.ru

SUBSCRIBE USING THE CATALOGUES OR VIA THE INTERNET:

ROSPECHAT (The Russian Press)
Indices: 37283, 59881
www.pressa.rosp.ru

INFORMNAUKA
Index: 59881
www.informnauka.com

INTER-POCHTA
17510
www.interpochta.ru

INFORMATION FOR AUTHORS:

If you would like to get your research paper published in *Acta Naturae* journal, please contact us at journal@biorf.ru or call +7 (495) 930-87-07.

