Letter from the Editors

Dear readers,

We would like to bring you the fifteenth, and this year's last, issue of *Acta Naturae*. As always, the issue begins with the Forum section containing three articles. One of them–*Made in Russia*–is devoted to the State Program "Development of Pharmaceutical and Medical Industries" that is awaiting passage. According to the program, the share of domestically produced drugs and medicinal products in the total volume of drugs consumed should considerably increase. However, the term "domestically produced drugs" still remains to be defined. In the publication, actors in the Russian Pharmaceutical industry exchange opinions about the issue. The remaining two publications are comments made by corresponding members of the Russian Academy of Sciences V.I. Tsetlin and A.N. Tomilin on the awarding of the 2012 Nobel Prizes in chemistry to professors R. Lefkowitz and B. Kobilka for their research into G protein-conjugated receptors and the Nobel Prize in Medicine to J. Gurdon and S. Yamanaka for the discovery of the possibility of reprogramming differentiated cells to pluripotent cells. The findings of the laureates in these paramount and extremely interesting fields of study have been deservedly appreciated by the world scientific community.

The research section of the issue consists of reviews and research articles. The subject area of the reviews is rather broad: from physicochemical biology to population genetics. We believe these reviews are devoted to topical questions and will be interesting to many readers.

Six research articles also embrace a broad range of disciplines: bioorganic chemistry (E.S. Matyugina *et al.*), bionanotechnologies (Yu.M. Yevdokimov *et al.*), molecular biology (E.N. Lyukmanova *et al.*), and cell biology (A.V. Golovin *et al.*). The subject area of the study by A.M. Ziganshin *et al.* is relatively new (the scientific foundations of waste processing in the agricultural sector). We deem such a broad range of studies typical of *Acta Naturae* useful, since it allows readers to learn something new from research areas that usually go unnoticed because of lack of time.

We would also like to draw the conclusions of the year 2012. It is in this very year that our journal has achieved its first success in international recognition by appearing in PubMed. Let us jointly do our best to further increase the authority of *Acta Naturae*.

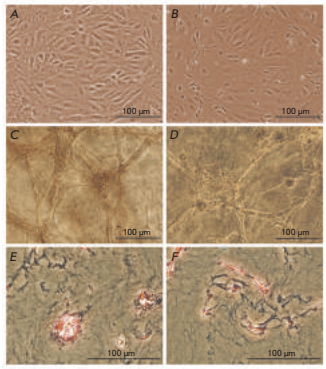
Happy New 2013 Year to our dear colleagues and readers! ●

Editorial Board

Effect of 3D Cultivation Conditions on the Differentiation of Endodermal Cells

O. S. Petrakova, V. V. Ashapkin, E. A. Voroteliak, E. Y. Bragin, V. Y. Shtratnikova, E. S. Chernioglo, Y. V. Sukhanov, V. V. Terskikh, A. V. Vasiliev

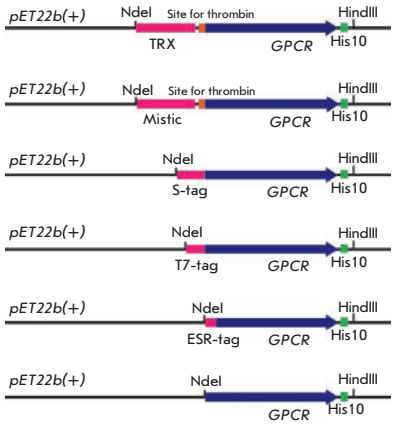
Cellular therapy of endodermal organs is one of the most important problems of modern cellular biology and biotechnology. One of the most promising directions in this field is studying the transdifferentiation abilities of cells within the same germ layer. The authors demonstrated that the liver progenitor cells also acquire the pancreatic differentiation capability under 3D cultivation conditions. Thus, postnatal salivary gland cells exhibit a considerable differentiation potential within the endodermal germ layer and can be used as a perspective source of endodermal cells for cellular therapy of liver pathologies.



Salivary gland and progenitor liver cell morphology under 2D and 3D cultivation conditions

N-Terminal Fusion Tags for Effective Production of G-Protein-Coupled Receptors in Bacterial Cell-Free Systems

Design of the vectors containing GPCR genes and additional 5-prime end sequences



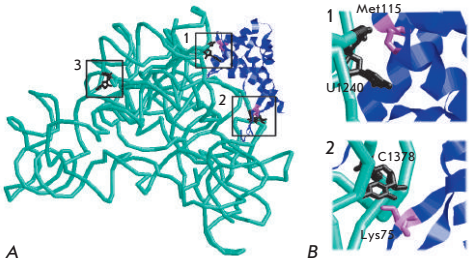
E.N. Lyukmanova, Z.O. Shenkarev, N.F. Khabibullina, D.S. Kulbatskiy, M.A. Shulepko, L.E. Petrovskaya, A.S. Arseniev, D.A. Dolgikh, M.P. Kirpichnikov

The family of G-protein-coupled receptors (GPCR) is one of the biggest families of membrane proteins. Despite their enormous importance to new developments in the fields of pharmacology and medicine, these receptors remain poorly studied, mostly due to the absence of high-efficiency systems for heterologous production. A methodology for the efficient production of GPCR in bacterial cell-free systems was proposed. The authors demonstrated that the yields of receptors can be increased (by 5-38 times) via the expression of fusion constructs containing additional N-terminal amino acid sequences.

Identification of Novel RNA-Protein Contact in Complex of Ribosomal Protein S7 and 3′-Terminal Fragment of 16S rRNA in *E. coli*

A. V. Golovin, G. A. Khayrullina, B. Kraal, A. M. Kopylov

For prokaryotes *in vitro*, 16S rRNA and 20 ribosomal proteins are capable of hierarchical self- assembly yielding a 30S ribosomal subunit. The self- assembly is initiated by interactions between 16S rRNA and three key ribosomal proteins: S4, S8, and S7. These proteins also have a regulatory function in translation of their polycistronic operons recognizing a specific region of mRNA. Therefore, a studying of RNA – protein interactions within the binary complexes is obligatory for understanding ribosome biogenesis. The non-conventional RNA – protein contact within the binary complex of recombinant ribosomal protein S7 and its 16S rRNA binding site (236 nucleotides) was identified. The structure of the binary RNA–protein complex formed at the initial steps of self-assembly of the small subunit has apparently to be rearranged during the formation of the final subunit structure.



The correlation between the X-ray 30S ribosomal subunit crystal data and cross-links of 30S ribosomal subunit in solution for *E.coli*

ActaNaturae
OCTOBER–DECEMBER 2012 VOL. 4 № 4 (15)

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

Address: 119234 Moscow, Russia, Leninskiye Gory, Nauchny
Park MGU, vlad.1, stroeniye 75G.
Phone/Fax: +7 (495) 930 88 50
E-mail: vera.knorre@gmail.com, mmorozova@strf.ru,
actanaturae@gmail.com

Номер подписан в печать 28 декабря 2012 г.
Тираж 200 экз. Цена свободная.
Отпечатано в типографии «МЕДИА-ГРАНД»

CONTENTS

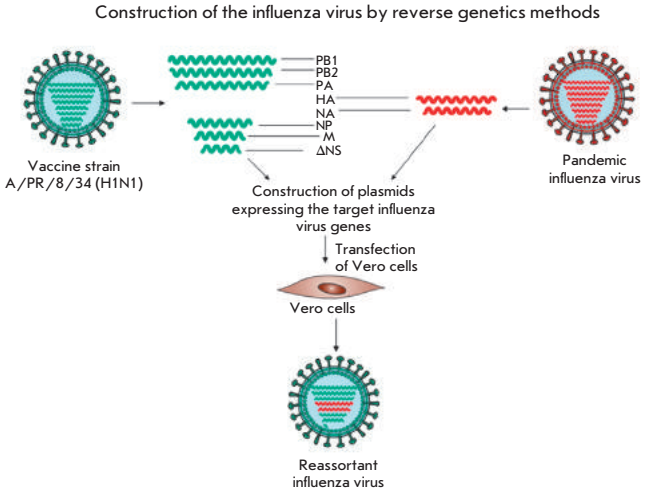


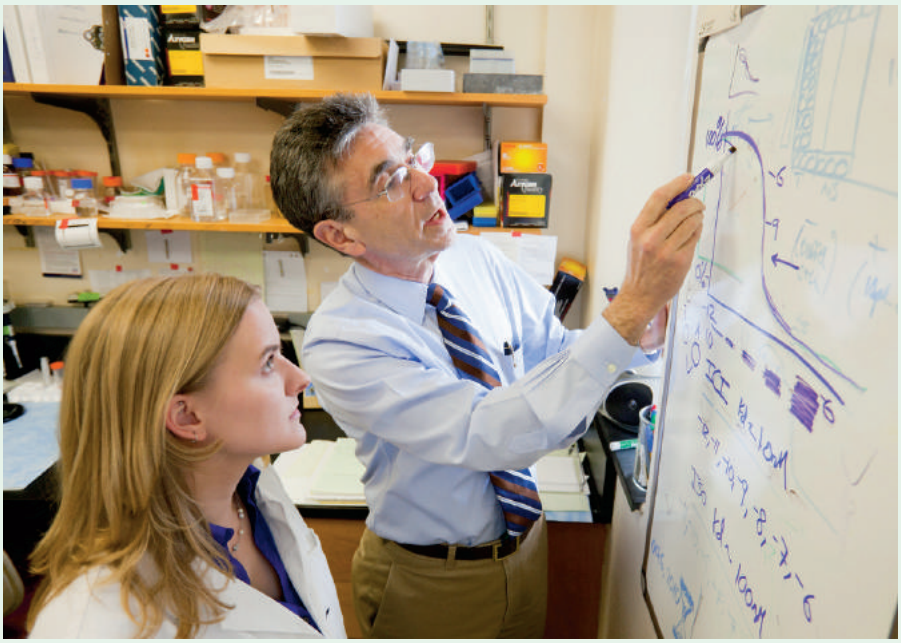
IMAGE ON THE COVER PAGE
Production of a new influenza reassortant strain by reverse genetics methods (E.S. Sedova *et al.*)

A Long Road to the Spatial Structure and Mechanisms of Action

The 2012 Nobel Prize in chemistry has been awarded to Robert Lefkowitz, professor at Duke University Medical Center, and Brian Kobilka, professor at Stanford University School of Medicine, for "research into G protein-coupled receptors." The wording "research," rather than "discovery, elucidation, or identification of something," speaks to the fact that such studies have been under way for nearly 40 years, from almost the zero level to the peaks conquered over the past few years.

In the early 1980s, G protein-coupled receptors (GPCR) and membrane receptors that belong to other classes had yet to be obtained as individual proteins and subsequently cloned. Here, one should do justice to R. Lefkowitz, who employed the synthesis of new ligands, affinity modifications, and other techniques to purify and isolate adrenergic receptors. In 1987, the clones were obtained and expression of the human β-adrenergic receptor was carried out at R. Lefkowitz's laboratory, with the active participation of B. Kobilka, at that time a young researcher working at the laboratory. Since then, R. Lefkowitz has led the field in the thorough investigation of the functions of adrenergic receptors and in the elucidation of the general principles of cellular signaling with the participation of GPCR. He has discovered β-adrenergic receptor kinase, elucidated various functions of β-arrestin, and thoroughly characterized the processes of phosphorylation and desensitization of GPCR. These functional studies performed by R. Lefkowitz have always been closely linked to medicine and were carried out with allowance for the diversity of the physiological functions of adrenergic receptors (in particular, regulation of cardiac activity).

During the past two decades, B. Kobilka has focused on determining the spatial structure of GPCR. One should not forget that all GPCRs are membrane proteins whose crystallization (similar to that of any other membrane protein) is still considered a challenging task. The first concept regarding the spatial structure of GPCR ("seven pillars") was obtained through the electron microscopic study of bacteriorhodopsin (Henderson, Unwin, 1975). Although this protein is a light-activated proton pump (not a GPCR), the presence of bound retinal and the homology of primary structures allowed to hypothesize that the actual GPCR, rhodopsin, has the same structure. In the late XX – early XXI century, the high-resolution electron microscopic structure of bacteriorhodopsin was obtained (Henderson, 1998) and the crystalline structure of rhodopsin was determined (Palczewski, 2000). It is worth noting here that the primary structures of bacteriorhodopsin and rhodopsin were determined in our country at Academician Ovchinnikov's school (1982), which required developing ways to isolate and analyze membrane proteins. Prior to the identification of the crystal structure of rhodopsin, bacteriorhodopsin had been used as a model for the analysis of the spatial organization of GPCRs. Rhodopsin was subsequently used for this purpose; however, it is obvious that this approach had certain limitations, and that determining the detailed spatial organization of individual GPCR was necessary. To do this, B. Kobilka and his colleagues set out to master the entire arsenal of techniques used for the solubilization, stabilization, and crystallization of membrane proteins, including the synthesis of new detergents, analysis of GPCR–antibody complexes, or chimerization with easily crystallizable foreign proteins that do not disturb the function of GPCR. As a result, a high-resolution crystal structure of the β2-adrenergic receptor was published in 2007. The structures of the complexes between this receptor and various ligands were subsequently determined. This has opened up possibilities for designing novel drugs. Huge experience has been accumulated, and the general methodology has been elaborated, since over the past few years B. Kobilka and his colleagues have determined the crystal structures of two individual subtypes of muscarinic acetylcholine receptors



Chris Hildreth/Duke Photography/Duke University

Linda A. Cicero / Stanford News Service

Winners of the 2012 Nobel Prize in chemistry Robert Lefkowitz and Brian Kobilka

and the μ-opioid receptor. Interestingly, the spatial structures of an additional three subtypes of opioid receptors have been determined in other American laboratories almost simultaneously. In all the cases, those included both the structures of ligand-free GPCR receptors and the structures of their complexes with pharmacologically important ligands, which opened up the way to elucidating the pharmacological differences between receptor subtypes belonging to the same family. B. Kobilka and his colleagues have also determined the structure of the β2-adrenergic receptor complex with G protein. The identified crystal structures not only facilitate the designing of novel drugs, but also open up the way to a thorough elucidation of the molecular mechanisms of cellular signaling with the participation of these receptors.

GPCR is the most common class of membrane receptors, which are considered to be the targets of over 30% of all commercially available drugs. However, this does not mean that other classes of membrane receptors deserve no attention or that their spatial structures are of lesser interest. Noteworthy that the 2003 Nobel Prize in chemistry was awarded to P. Agre and R. MacKinnon for the identification of the spatial structures of aquaporin and the potential-dependent K^+ channel. The crystal structures of a number of ligand-gated channels, which are involved in memory formation and transfer of pain signals and participate in other essential physiological processes, have been determined over the past few years. ●

V.I. Tsetlin,
Corresponding member
of the RAS

The 2012 Nobel Prize in Physiology or Medicine

The 2012 Nobel Prize in physiology or medicine has been awarded to John Bertrand Gurdon and Shinya Yamanaka "for the discovery that mature cells can be reprogrammed to become pluripotent."

The 50-year-long story started in 1962 (the year Yamanaka was born), when John Gurdon used nuclear transplantation to demonstrate that frog skin and intestinal cells could spawn a new organism. In this way, fundamental evidence to the fact that the stem characteristics, lost during the development of an amphibian, can be in principle regained (Gurdon, 1962). Of course, at that time nobody discussed any therapeutic applications of the finding. More than 40 years later, an article by Shinya Yamanaka was published, in which it was demonstrated for the first time that somatic mammalian cells can also be reprogrammed to the pluripotent[1] state (Takahashi and Yamanaka, 2006). During this period of the next 40 years, a number of events that were not as sound, but still very important, occurred, which ultimately led to Yamanaka's discovery. Martin Evans pioneered the derivation of mouse embryonic stem (ES) cells in 1981 (Evans and Kaufman, 1981), which not only paved the way for the numerous studies of gene functions via gene knockout techniques (the 2007 Nobel Prize), but has also brought into light the first known type of pluripotent cells with a considerable therapeutic potential. Despite significant effort, human ES cells were obtained much later, only in 1999 (Thomson *et al.*, 1998), which allowed a number of researchers to consider the possibility of tissue replacement therapy in humans. I should also mention such a significant event as the cloning of Oct4 (Okamoto *et al.*, 1990; Scholer *et al.*, 1990), one of the central genes essential not only for the maintenance of cellular pluripotency, but also for its induction (as was subsequently demonstrated by Yamanaka).

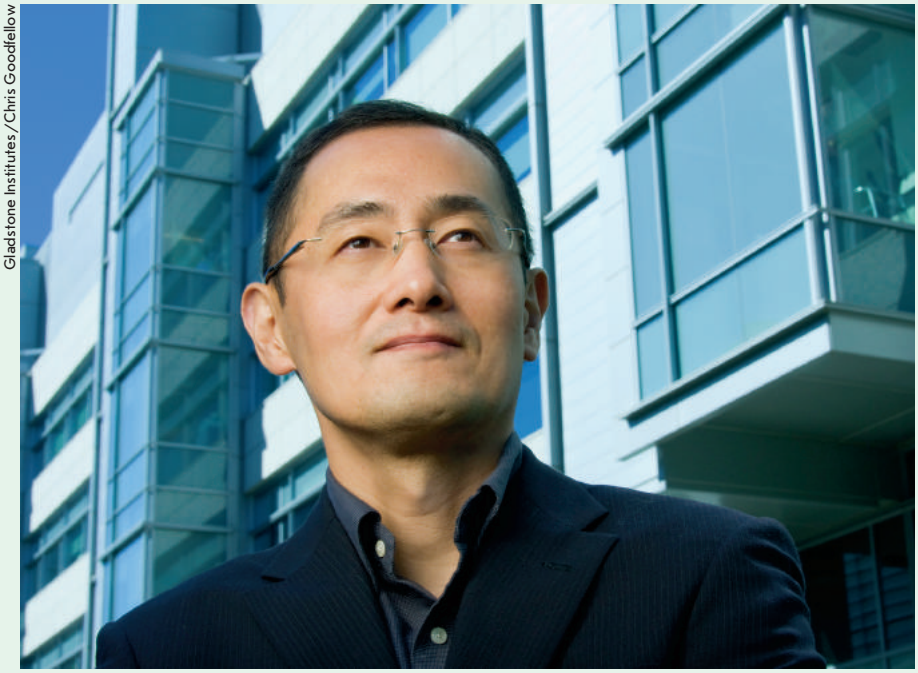
Gladstone Institutes/Chris Goodfellow

John Gurdon

Cloning of a sheep was the next milestone in cell reprogramming following Gurdon's studies; it provided the first evidence that converting mammalian somatic nuclei to the totipotent[2] state is possible.

[1] *The term "pluripotency" is used to denote the ability of cells to self-renew and differentiate to all cell types of adult mouse tissues (with the exception for two extraembryonic cell types, trophectoderm and primitive endoderm).*

[2] *In other words, the ability to give rise to all embryonic and extraembryonic cell lines; two mammalian pluripotent cell types have been known: zygote and early blastomeres.*



Gladstone Institutes/Chris Goodfellow

Shinya Yamanaka

The reprogramming was achieved by transplanting somatic nuclei into oocyte cytoplasm (Campbell *et al.*, 1996). A series of studies showing that reprogramming can be achieved by spontaneous or induced fusion of somatic cells with ES cells is also worth mentioning (Matveeva *et al.*, 1998; Tada *et al.*, 2001; Terada *et al.*, 2002; Ying *et al.*, 2002). It became obvious that specialization of cell types during mammalian development is a reversible process.

Meanwhile, the apparent disadvantages of human ES cells came to the forefront: they were associated both with ES cell production requiring the sacrifices of human embryos and with high risks of immune rejection of differentiated derivatives of ES cells in a recipient's body.[3] Thus, the task of obtaining pluripotent stem cells from somatic cells arose in the early 2000s; this task was bound to eliminate both the ethical and practical problems associated with ES cells. Several research groups (including the one I was heading at the Max Planck Institute in Freiburg) which believed that reprogramming of somatic cells to the pluripotent state by forced gene expression was a feasible aim were doing their best, developing sophisticated approaches for the screening of reprogramming factors and selection for pluripotency. Yamanaka has drawn the final line in this pursuit. He simply picked 24 transcription factors expressed in mouse ES cells and determined the minimum combination of them sufficient for the induction pluripotency in mouse fibroblasts: Oct4, Sox2, Klf4, and cMyc (the so-called Yamanaka cocktail) (Takahashi and Yamanaka, 2006). The cells obtained in this way, named induced pluripotent stem (iPS) cells, possess characteristics that are almost identical to those of ES cells.

Six years have passed since induced pluripotency was discovered, and almost 5,000 articles have been published. These articles describe the alternative combinations of transcription factors, new methods for obtaining iPS cells in various species (including humans), propose various delivery methods (viral, plasmid, transposon, protein, RNA), use various starting somatic cell lines to obtain iPS cells, etc. A number of studies that expose the caveats of iPS cells (low generation efficiency, uncertain epigenetic status, chromosome instability, increased rate of point mutations as compared to that in ES cells, etc.) have also been published. These drawbacks postpone the introduction of iPS cells to clinics. However, it is already clear that iPS cells are indispensable both for creating *in vitro* models of a broad range of human diseases (the so-called diseases in a Petri dish) and for *in vitro* screening of drugs to treat these diseases.

Thus, the well-expected and absolutely deserved 2012 Nobel Prize has marked the start (made by Gurdon) and the finish (crossed by Yamanaka) lines in the 50-year-long marathon race in the pursuit of pluripotency. Beyond all doubts, the results of this race promise enormous benefits for mankind's health. ●

[3] *No databanks of characterized ES cells that would embrace all possible histotypes have been launched thus far.*

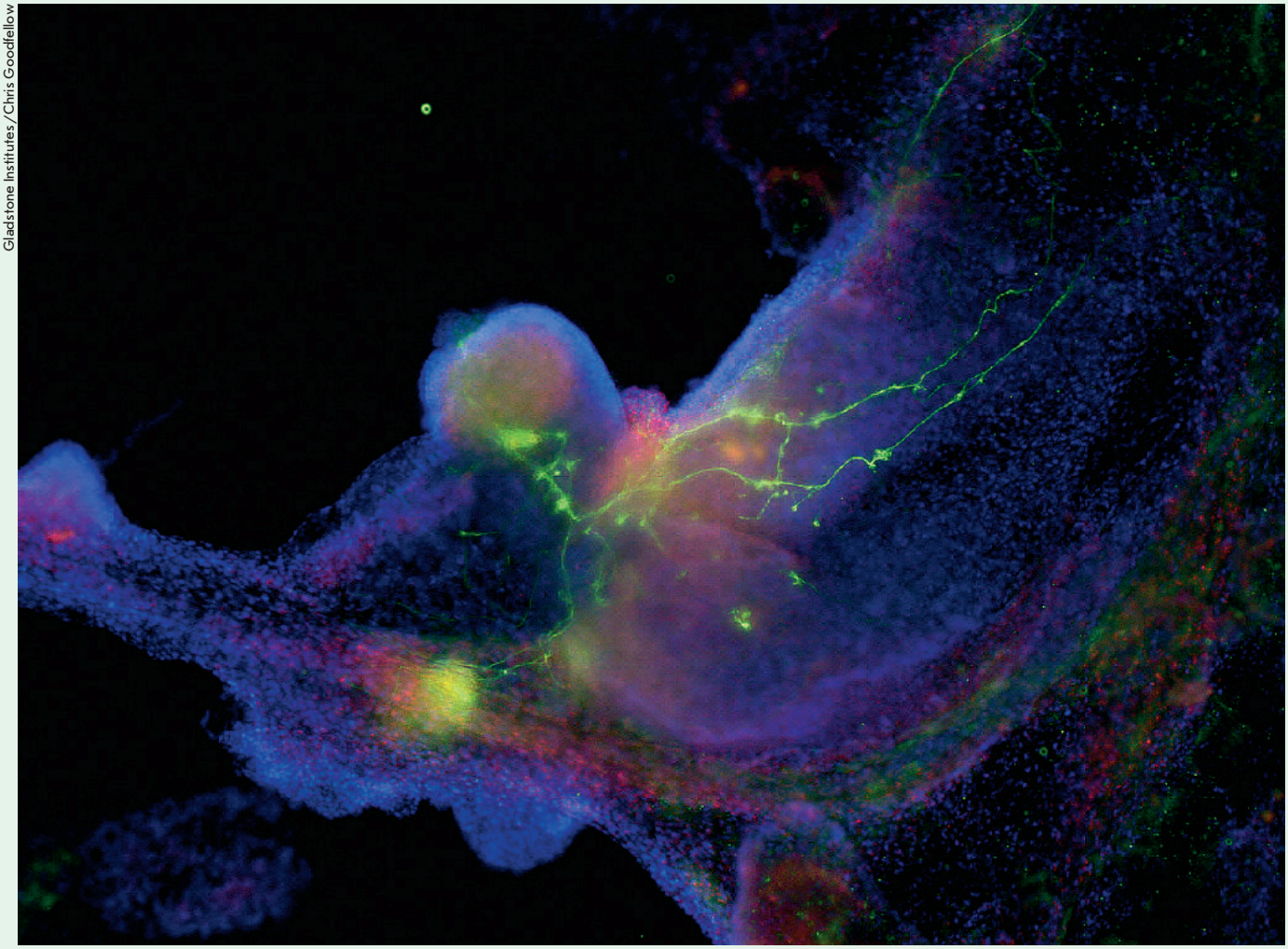
Campbell K.H., McWhir J., Ritchie W.A., Wilmut I. Sheep cloned by nuclear transfer from a cultured cell line // Nature. 1996. V. 380. P. 64–66.

Evans M.J., Kaufman M.H. Establishment in culture of pluripotential cells from mouse embryos // Nature. 1981. V. 292. P. 154–156.

Gurdon J.B. Adult frogs derived from the nuclei of single somatic cells // Developmental Biology. 1962. V. 4. P. 256–273.

Matveeva N.M., Shilov A.G., Kaftanovskaya E.M., Maximovsky

Gladstone Institutes/Chris Goodfellow



Induced pluripotent stem cells, which are known as iPS cells and act very much like embryonic stem cells, are here growing into heart cells (red) and nerve cells (green)

L.P., Zhelezova A.I., Golubitsa A.N., Bayborodin S.I., Fokina M.M., Serov O.L. *In vitro* and *in vivo* study of pluripotency in intraspecific hybrid cells obtained by fusion of murine embryonic stem cells with splenocytes // Molecular Reproduction and Development. 1998. V. 50. P. 128–138.

Okamoto K., Okazawa H., Okuda A., Sakai M., Muramatsu M., Hamada H. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells // Cell. 1990. V. 60. P. 461–472.

Scholer H.R., Dressler G.R., Balling R., Rohdewohld H., Gruss P. Oct-4: a germline-specific transcription factor mapping to the mouse t-complex // EMBO J. 1990. V. 9. P. 2185–2195.

Tada M., Takahama Y., Abe K., Nakatsuji N., Tada T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells // Curr. Biol. 2001. V. 11. P. 1553–1558.

Takahashi K., Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors // Cell. 2006. V. 126. P. 663–676.

Terada N., Hamazaki T., Oka M., Hoki M., Mastalerz D.M., Nakano Y., Meyer E.M., Morel L., Petersen B.E., Scott E.W. Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion // Nature. 2002. V. 416. P. 542–545.

Thomson J.A., Itskovitz-Eldor J., Shapiro S.S., Waknitz M.A., Swiergiel J.J., Marshall V.S., Jones J.M. Embryonic stem cell lines derived from human blastocysts // Science. 1998. V. 282. P. 1145–1147.

Ying Q.L., Nichols J., Evans E.P., Smith A.G. Changing potency by spontaneous fusion // Nature. 2002. V. 416. P. 545–548.

A.N. Tomilin,
corresponding
member of the RAS

Made in Russia

By the end of 2012, the Government of the Russian Federation is to approve the State Program "Development of the Pharmaceutical and Medical Industries" for 2013–2020, which includes the current Federal Target-Oriented Program "Pharma-2020." One of the objectives within the State Program prepared by the Ministry of Industry and Trade is to "increase the share of domestically produced drugs and medicinal products in overall consumption by the public healthcare services of the Russian Federation by 48%." However, the term "domestically produced drug" still remains to be legislatively defined. According to the draft resolution issued by the Ministry of Industry and Trade in May 2012, a "domestic drug" should mean a drug whose production cycle in the territory of the Russian Federation starts from a substance or a ready-to-consume formulation. Until 2014, the Ministry was ready to regard even those drugs whose packaging was made in Russia as Russian ones. However, no further steps followed. Therefore, the question pertaining to which drugs and which produced by which pharmaceutical companies should be regarded as domestic drugs remains open. Actors of the Russian pharmaceutical industry share their opinions.

Andrey Ivashchenko, Chairman of the Board of Directors of the Chem-Rar High-Tech Center.

In your opinion, to what extent are the pharmacological production facilities to be localized in Russia so that a company could be regarded as a Russian manufacturer?

There are two aspects in localization of production facilities. What does the Russian government want? When buying pills for the public healthcare system costing 1billion USD, it wants a possibly higher share of this amount to be in rubles. In this case, the budget risks to the state are lower. Let us say that 30 billion rubles are to be spent to buy pills. If the drugs are imported and the rate of the US dollar to the Russian ruble moves by 10%, the state will have to look for an additional 3 billion rubles. Let us assume that a drug is being sold in Russia. What share of its total cost is in USD? If

Andrey Ivashchenko: One should see the interests of the State in the term "domestically produced drug"; in this case, the situation becomes clearer

it has been completely produced in Russia, its entire cost will be in rubles. If the substance was imported from China, 20–30% of the drug's cost will be in USD. Finally, if a drug has simply been bought from a distributor, 90% of its cost will be in USD.

The second aspect consists in the technologies used in our country. This aspect is important in terms of drug safety. If only pelletizing and bottle-filling are located within the country, the technology is very simple. It is better if chemical substances are synthesized in the country. Finally, if Russian manufacturers can produce both synthetic and biotechnological substances, this means that there is a complete set of technologies in the country.

Therefore, various combinations of these two aspects are possible. For instance, in the case where the patent is foreign but the drug is produced in Russia, the biggest share of the drug's cost will be in USD, but there will be production facilities in our country. The opposite case is when the patent is Russian but the drug is produced in China: the biggest share of the drug cost is in rubles, but we do not have the technology for its production. An ideal variant is when both technology and production facilities are Russian. A poor variant (which is nowadays most commonly observed in Russia) is when drugs are simply imported. Hence, the state policy is oriented towards proceeding from the worse alternative to the best one.

Another aspect of state policy is the protection of its industry under conditions of elimination of state boundaries (e.g., after joining the World Trade Organization, WTO). There is such a protection method as technical regulations. Let us assume that most pharmaceutical manufacturing companies in Russia today can only produce drugs in their finished dosage form (FDF). The Government, hence, decides that a manufacturer producing FDF in the country will be regarded as a domestic producer. When most pharmaceutical companies learn how to produce substances, those who produce them within the country will be regarded as Russian producers. I deem the situation will develop in this very direction. Until recently, a company could have been regarded as a Russian manufacturer by simply packaging the ready-to-consume drugs into boxes. Now the situation is different: the company has to be able to bottle and pelletize drugs under sterile conditions, etc. I am positive that with such a policy, in five years, when most manufacturers will start to produce substances, they will lobby for preferences.

One should see the interests of the State in the term "domestically produced drug"; in this case, the situation becomes clearer. It is a rather flexible system. It has been specified in the Federal Target-Oriented Program "Pharma–2020" that drugs produced in Russia should make up 50% of the pharmaceutical market by 2020. Let us assume that Russian companies open plants in China, where the cost of production is low. The largest share of the pharmaceutical market will be in rubles. However, where will we produce vaccines and pills in case of war? Some technologies need to be localized in Russia. This is exactly what the Government is doing right now. There is a list of 57 strategic medications that are to be fully produced in the territory of Russia. It is clear from this list that production of these medications requires one to master the major pharmaceutical technologies.

I deem the Ministry of Industry and Trade holds a rather reasonable position. It is a different matter that it runs counter to the position held by the Ministry of Healthcare and Social Development and doctors. It does not matter for the doctors who

produces the pills. Moreover, they consider imported medications to be of better quality (and they are often right). A compromise, which is not subject to market regulation, has to be reached. There should be a system of interdepartmental regulations. Instead, the Ministry of Industry and Trade is now slamming on the gas pedal, whereas the Ministry of Healthcare and Social Development does not actually care. It steps on the brake, which eventually can damage the entire engine.

There have not been interdepartmental mechanisms for balancing between different interests thus far. However, the number of companies investing funds into import substitutive production is on the increase. Hence, the issue is on the agenda, and some solution is needed. Some kind of mechanism will emerge, but it is unclear yet whether it will be an interdepartmental committee or a self-regulatory organization.

The question pertaining to production localization deals with the same issue. The Government makes foreign companies localize production facilities without giving any guarantees that procurement will be made. Foreign companies have been looking attentively at the situation and eventually have started building facilities. But it is obvious that if rivals invested funds into the construction, they will ask the State for preferences. Russian manufacturers will do so as well. Hence, falling behind is inadmissible. The avalanche-like process has started. The first to construct their facilities were those who had something to lose – various East European players who are being driven out of the market. They produce branded generics, and Russia is the last reserve where these products can be sold. Therefore, such companies as Polpharma, Stada, Gedeon Richter were the first ones to localize their production facilities in Russia. Next were the European innovative firms: Novartis and Sanofi-Aventis. They are followed by American companies falling behind by a year or two. Japanese firms are at the very tail.

The contests held by the Ministry of Industry and Trade include two types of events. First-priority events include everything associated with import substitution. There is intense rivalry in all the group 1 contests. There are several players for each lot, which pushes down prices. It means that there are a lot of companies in Russia dealing with import substitutive production. In turn, this fact means that all of them will have their drugs certified in 3–4 years and will compete further during the stage of Government procurement. And all these medications will be manufactured in Russia.

How will it affect the demand for developments made by our scientists?

Of course there is no demand for developments that fall behind, whereas cutting – edge developments are needed. Global competition still exists; there is no way to escape it. Identically to the USA or Europe, Russia will learn to build virtual regulatory barriers to protect industries that are important to us. However, there will be no rigid wall. No one can protect us against competition from foreign manufacturers; hence, we need to reach for cooperation models, incorporate into a high-profit unit of the added value chain, together with the foreign producers (e.g., into research & development). Of course, it is important to make the lives of the local developing companies easier; however, we also should not isolate ourselves from the world market. Balance is required. We are moving towards it; although the imbalance between the State policies hinders this move. But this is presumably a stage that we need to go through.

Aleksandr Bykov, Director of Government and Public Relations, Novo Nordisk.

According to one of the viewpoints, only the pharmaceutical manufacturers that have full-cycle production facilities (from substance to the drug in its finished dosage form) on the territory of Russia should be regarded as domestic manufacturers. Do you think this approach is reasonable?

I think it is not quite reasonable to insist on substance production. The development of the pharmaceutical industry should be guided not by drug safety, but by intensive scientific activity at pharmaceutical production facilities, modernization of the industry, and development of its intellectual potential. The safety thesis is applicable to a very narrow group of medicators only.

On the other hand, let us assume that we start producing substances. It is a very complex chemical production process during which we are bound to encounter a lack of some additional ingredients, catalysts, or equipment. According to this logic, we will have to relocate the production of these ingredients, catalysts, and equipment providing functioning of our chemical production facility to Russia, as well. But this is a way towards North Korea's Juche ideology, which insists on relying on domestic resources only.

In your opinion, what should state policy in this area be?

I think that the investment attractiveness of the market should be developed. Preference should be given to companies that localize their production facilities in Russia, thus contributing to the Russian economy's modernization. Among these companies, there can also be Russian enterprises that reorganize their technological process in accordance with GMP standards or participate in joint cutting-edge developments.

Aleksandr Bykov: The development of the pharmaceutical industry should be guided not by drug safety, but by intensive scientific activity at pharmaceutical production facilities, modernization of the industry, and development of its intellectual potential

Biotechnology and new directions in drug synthesis should come under focus when developing the pharmaceutical industry. If we do not get integrated into the global industry, we can find ourselves left far behind. Of course, we will be supplied with medications. The process of producing mass medications via chemical synthesis is relatively simple. But if foreign manufacturers are forced to produce biodrugs here (while the volume of domestic consumption is small), the costs of production organization will be higher than the potential profits. Hence, investing here will be less attractive, and investments will go into the production facilities of India and China.

Do you mean that the best definition for the term "Russian producer" is "companies producing drugs in their finished dosage form in Russia"?

Exactly! By 2018, the volume of medications produced in Russia (according to the lists of strategically important drugs and vitally essential and most important medicines) is scheduled to increase to 90%. However, joining the WTO assumes that the excessive trade barriers will be eliminated: by then there probably will be no need at all in establishing the nationality of manufacturers or products. The requirements to the products manufactured in Russia and imported ones should be identical. In this case, it will be reasonable to align the principles of price formation for drugs produced in Russia with global ones. Of course, Government procurement in the frame of the WTO is not regulated. However, it is quite possible that in five years Russia will have signed some additional agreements, which will determine this sector of the market, as well. In this case, the competitive advantage will disappear.

What can be the response of foreign manufacturers to the requirement to produce substances in Russia so that a drug is regarded as domestically produced?

There will be no response. The companies that have already localized their production facilities in Russia will not close them down. This measure will stimulate further localization of substance production only for companies whose drugs are consumed in huge amounts. In this case, they may consider substance production in Russia. However, if only a few thousand people need a drug, it is not reasonable at all to build a plant that would operate for only a day or two. Production of small amounts of substances will be economically a nonstarter. India and China are producing substances for the entire world: for their domestic markets, the foreign market, the USA, and Europe. Large multinationals subsequently purify the substances produced in India and China and produce drugs in their finished dosage form. Hence, is it possible to compete with India in substance production? And do we actually need to compete? Substance production is not a highly intellectual process. It is simply the chemical production and is also associated with certain environmental costs.

Which ways of stimulating the development of new drugs in Russia should be used?

I deem it necessary to develop pharmaceutical clusters. They are the link between science and technology and production. The research organizations (institutes, centers, laboratories) do not structurally belong to the pharmaceutical companies working in the pharmaceutical cluster. Meanwhile, these institutions can solve the urgent problems of the industry and be additionally supported via the scheme of public – private partnership. However, they need to cooperate with the international pharmaceutical industry so that their developments can reach markets. There are very few cases when a drug was recognized only in the local market and successfully sold.

Victor Dmitriev: The current Federal Law FZ-94 with price being the major driving force is against both patients and Russian manufacturers

Viktor Dmitriev, General Director, Association of Russian Pharmaceutical Manufacturers

Your organization has recently proposed to elucidate the question pertaining to the definition given to a domestically produced drug. Why is that important?

Russian products participate in government tenders; hence, it is important that local products be given some preference, since government procurement makes up to 30% of the pharmaceutical market. They have indeed been given for 3 years already according to the Resolution of the Ministry of Economic Development on 15% preference points when making public procurement in accordance to the Federal Law FZ-94. However, in reality these preference points appear to be inefficient, since the technical documentation for the tender is drawn in such a way that domestic manufacturers cannot take advantage of their preference points.

What should be done to make the preference point system work?

First, political will is required. Second, the current Federal Law FZ-94 with price being the major driving force is against both patients and Russian manufacturers. We would propose to provide a separate article in the Law (or even a separate Law) to regulate drug procurement. The reason for that is that drugs are goods that cannot be regulated based on general criteria. What do I mean? The Federal Law on technical regulation assumes that quality is determined by a buyer. If he liked a certain type of sausages, he would buy them again. If not, he simply would not buy them. This cannot be said about drugs, since their effect or quality can be tested only under laboratory conditions. I can provide the following example: the effect of statins used to prevent cardiovascular diseases cannot be felt; it can only be assessed according to analysis results.

Another point is that drugs possess a special property that other products lack. It is the so-called placebo effect, when a person expects amelioration of his condition and even feels it, but the amelioration is not necessarily caused by the drug. The waiting process can also contribute to it.

Hence, I deem it necessary to adopt a separate law in which the following provisions will be made. First, if several domestic manufacturers (e.g., two, as is in Belarus) participate in a public tender, participation of drugs produced by foreign companies should be prohibited. Second, the participants in the tender should be obligated to hold a GMP certificate. The reason for that is that the pharmaceutical industry is very heterogeneous now. There are companies that are in compliance with world standards. The net cost of the drugs produced by them is higher than that of drugs produced by companies that are not in compliance with GMP standards. However, the quality of the drug is guaranteed in the former case. According to the third provision, supply volume and product prices should be guaranteed.

Vladimir Shipkov, Executive Director, Association of International Pharmaceutical Manufacturers

The Ministry of Industry and Trade has offered different criteria for defining a domestically produced drug. What is your sense as to how the question should be solved?

Indeed, the Ministry of Industry and Trade has not resolved this question yet. I am not sure whether that is a good or bad thing. The last project prepared by the Ministry has provided additional motivation depending on the extent of production localization. On one hand, I would appreciate this approach, although it has not been implemented thus far in the form of normative documents in force. On the other hand, the ef-

Vladimir Shipkov: The idea that only the medications whose full-cycle production was carried out in Russia should b regarded as Russian ones makes no sense

forts at defining a local product or a local manufacturer under the conditions of joining the WTO make little sense. Since the pharmaceutical industry is among the most globalized ones, it would be more reasonable to think about supporting manufacturers who produce goods in accordance with the generally accepted international quality standards. The potential customer should not care where a certain drug was produced: in Russia, Ukraine, Kazakhstan, France, the USA, or somewhere else. What really matters is that the drug corresponds to strict international requirements.

In your opinion, what preference points should be given to Russian manufacturers?

The preference points that are being given right now look a disservice to me. It is wrong to give 15% preference points for products manufactured in the Russian Federation (starting in 2012, in Belarus as well), since it is given to any medications regardless of whether they were produced in compliance with GMP standards or under dubious conditions. The manufacturers should be motivated to invest into modernization by encouraging GMP implementation. Hence, the preference points should be given only to those manufacturers who work in compliance with these standards. However, I think that, much higher preference points (about 30–35%) should be given in our situation. In this case there will be real motivation. Do you want to gain preference points? Do you want to enter the markets of third-world countries instead of hiding behind the "iron curtain?" In this case, implement GMP at your production facilities. Furthermore, a differentiated approach should also be used: a deeper localization process means higher preference points (but provided that a manufacturer is in compliance with GMP standards). After a while, when all enterprises implement GMP standards, this requirement can be dropped.

What requirements to the degree of localization do you think there should be?

All localization forms should be encouraged. The idea of regarding only those medications whose full-cycle production took place in Russia as Russian ones makes no sense. Among Russian manufacturers very few deal with full-cycle production (including substance synthesis). There is no way inept requirements to full-cycle production can attract investments. Instead, a differentiated preference point system for all localization degrees is needed (e.g., secondary package – 5%; a drug in its final dosage form – 15%, full-cycle production – more points). If an investor today wants to invest into packaging and use the minimum preference points, this should be encouraged as well. After working under these conditions for a certain period of time and seeing that the other manufacturers are awarded more preference points because of a deeper degree of localization, he eventually will implement deeper localization as well, increase investments, hire workers, and establish better production conditions. ●

Prepared by Ekaterina Borovikova

Recombinant Influenza Vaccines

E. S. Sedova[1*], D. N. Shcherbinin [1], A. I. Migunov[2], Iu. A. Smirnov [1,3], D. Iu. Logunov[1], M. M. Shmarov[1], L. M. Tsybalova [2], B. S. Naroditskiĭ[1], O. I. Kiselev[2], A. L. Gintsburg[1]
[1]Gamaleya Research Institute of Epidemiology and Microbiology, Gamaleya Str., 18, Moscow, Russia, 123098
[2]Research Institute of Influenza, prof. Popov Str., 15/17, Saint Petersburg, Russia, 197376
[3]Ivanovsky Research Institute of Virology, Gamaleya Str., 16, Moscow, Russia, 123098
*E-mail: sedova-es@yandex.ru
Received 27.07.2012

ABSTRACT **This review covers the problems encountered in the construction and production of new recombinant influenza vaccines. New approaches to the development of influenza vaccines are investigated; they include reverse genetics methods, production of virus-like particles, and DNA- and viral vector-based vaccines. Such approaches as the delivery of foreign genes by DNA- and viral vector-based vaccines can preserve the native structure of antigens. Adenoviral vectors are a promising gene-delivery platform for a variety of genetic vaccines. Adenoviruses can efficiently penetrate the human organism through mucosal epithelium, thus providing long-term antigen persistence and induction of the innate immune response. This review provides an overview of the practicability of the production of new recombinant influenza cross-protective vaccines on the basis of adenoviral vectors expressing hemagglutinin genes of different influenza strains.**
KEYWORDS **Recombinant vaccine; influenza; immunization.**
ABBREVIATIONS **WHO – World Health Organization; DNA – deoxyribonucleic acid; HA – hemagglutinin; NA – neuraminidase; VLP – virus-like particles; RNA – ribonucleic acid; NP – nucleoprotein; DC – dendritic cell; APC – antigen-presenting cell; Ad – adenovirus; NK – natural killers.**

INTRODUCTION

Influenza is the most common infectious disease. According to the WHO, 20–30% of children and 5–10% of adults are infected with influenza annually; the severe complications caused by it result in the death of 250, 000-500, 000 people. The economic burden inflicted by influenza epidemics is estimated at 1–6 million USD per 100, 000 population [1]. The burden and mortality increase significantly during pandemics. Thus, according to different sources, the influenza pandemic that struck in 1918–1919 caused 50-100 million deaths [2].

Prevention through vaccination is the most sensible measure to protect people against influenza and to contain its spread [3]. Modern influenza vaccines typically induce the formation of antibodies against the influenza virus' surface antigens: hemagglutinin (HA) and neuraminidase (NA). These vaccines include both live and inactivated (whole-virion, split, subunit) vaccine types. The efficiency of seasonal vaccines directly depends on the degree of correspondence between the antigenic structure of the influenza virus strains within the vaccine and the strains circulating among the population during a given epidemic season. The influenza virus surface proteins undergo progressive antigenic variation (antigenic drift), thus requiring annual renewal of the strain composition of vaccines [4].

The development of highly immunogenic and safe vaccines inducing the immune response of a broad spectrum of action is currently one of the major problems encountered in efficient influenza prevention. The 2009–2010 pandemic caused by the influenza A(H1N1) pdm09 virus and the existing pandemic threat of avian influenza A(H5N1) sustain the interest in designing new vaccines capable of inducing broad protective immunity [5].

The use of reverse genetics techniques to design influenza vaccines

The existing influenza vaccines can be subdivided into two types: the attenuated (live) and inactivated (including subunit) types. All of those are rather widely used for population immunization and have shown themselves to perform well. Attenuated vaccines are influenza viruses with attenuated virulence [6]. The epidemically topical virus strains are also used to produce inactivated subunit vaccines, although the use of high pathogenic strains is limited by strict requirements imposed on the biological safety of the production process [7].

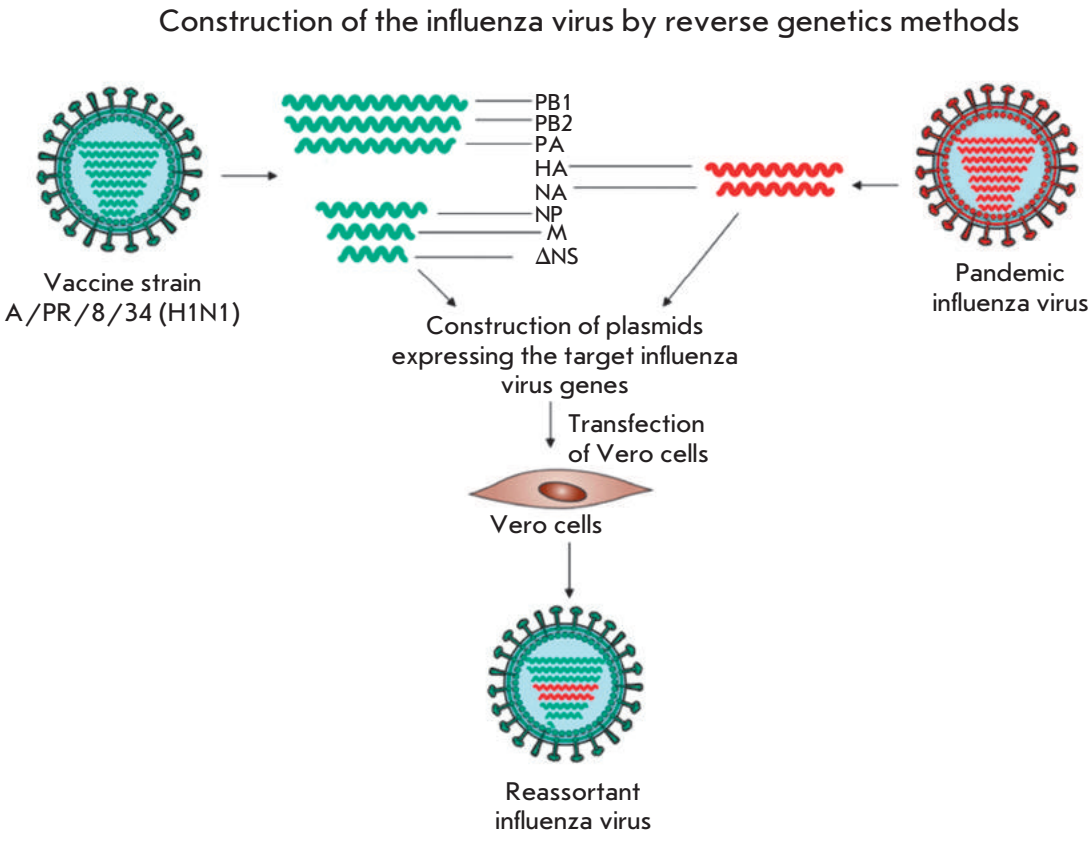


Fig. 1. Production of a new influenza reassortant strain by reverse genetics methods

The conventional methods for producing vaccine strains of the influenza A virus have a number of drawbacks. The use of both attenuation based on viral adaptation to the organism of the heterologous host [8] and reassortment after coinfection with epidemic strains and attenuation donors [9] does not always make it possible to preserve the equilibrium between the virulence level of the original virus and its immunogenicity. Excessive attenuation may result in the production of strains that have lost all ability to reproduce in the cells of the human respiratory tract.

The use of reverse genetics techniques is an alternative method for obtaining vaccine strains. Reverse genetics techniques allow to reconstruct a biologically active viral particle by coinfecting permissive cell lines with plasmids containing the genes that encode viral proteins. The virulence and antigenic properties of the influenza virus can be manipulated by altering these genes [10].

The reverse genetics techniques can be used to obtain reassortant influenza viruses. Thus, the plasmids encoding the segments in the genomes of pandemic or circulating seasonal strains and the attenuated vaccine strain of the influenza A virus are used to transfect permissive eukaryotic cells. As a result, the assembly of whole virions of the virus carrying a combination of proteins of both the vaccine and pathogenic strains occurs (*Fig. 1*). The influenza A(H1N1) virus that caused the 1918 pandemic (the so-called Spanish Influenza) was successfully obtained and examined using this very technique [11].

Reverse genetics techniques allow to reduce the viral virulence by introducing mutations into various viral genes. Thus, mutations in two genes encoding the polymerase proteins PB1 and PB2 of the avian influenza A/guinea fowl/Hong Kong/WF10199 (H9N2) virus caused the loss of viral pathogenicity for chickens [12]. The deletion of the nonstructural protein NS1 resulted in attenuation of the influenza A virus. The vaccine obtained using this method has successfully undergone phase I clinical trials [13, 14]. The introduction of mutations to the M2 protein, which is essential for ion channel formation, also results in virus attenuation [15]. After variation of the amino acid sequence at the fragmentation sites of HA of the highly pathologic influenza A H5 virus by targeted mutagenesis, the virus acquires the characteristics of low pathogenic viruses [16].

The reverse genetics techniques have shown good results in obtaining attenuated strains of the influenza virus [17]. However, the use of reassortment in the case of vaccine strains brings to the fore the question of biosafety because of possible mutations that can re-

cover or increase the viral virulence [18]. Furthermore, the extensive use of live attenuated influenza vaccines casts suspicion because of the possible reassortment of a live vaccine with the circulating strains of human influenza viruses [19, 20]. Vaccine strains of the influenza virus in preparative amounts are most commonly produced in chicken embryos, which makes vaccination of individuals allergic to chicken protein impossible. Another drawback of the vaccines produced using chicken embryos is the dependence of the technological process on fertility in the chicken flock.

Recombinant subunit vaccines

The problems associated with the use of chicken embryos and the necessity to attenuate pathogenic strains of the influenza virus can be solved using recombinant subunit vaccines. The use of various expression systems for rapid production of individual viral proteins in preparative amounts is one of the new approaches to the production of subunit influenza vaccines [21].

In one of the popular expression systems, influenza antigens are produced in insect cells using baculoviral vectors carrying the genes of the target antigens. The autographa californica multiple nucleopolyhedrovirus (AcMNPV) is the most commonly used. Sf9 cell lines obtained from *Spodoptera frugiperda* ovarian tissue are typically used for work with AcMNPV. This system can be used to produce various antigens of the influenza A virus. Immunization of mice with the recombinant HA of the influenza H5N1 virus obtained in the baculovirus expression system resulted in the induction of a high level of virus-neutralizing antibodies. However, either an adjuvant or prime-boost immunization using an inactivated influenza H5N1 virus or the recombinant adenovirus carrying the HA gene of the influenza virus was required in order to attain any significant antibody level [22].

The ion channel-forming protein M2 is considered the most promising candidate for influenza subunit vaccines. M2 is one of the three influenza A virus proteins that are expressed on the virion's surface; as opposed to HA and NA, this protein is highly conserved. In viruses circulating in the human population, the M2 protein ectodomain (M2e) has undergone virtually no changes since 1933 [23]; hence, the M2e protein is regarded as a candidate for designing broad-spectrum vaccines. Thus, the possibility of using the cucumber mosaic virus to express the M2e protein of the A H5N1 influenza virus in plants has been demonstrated [24].

The low immunogenicity and, therefore, the need for repeated vaccination and use of adjuvants are the drawbacks of recombinant subunit vaccines, as well as of conventional subunit vaccines. One of the ways of solving this problem consists in including molecular adjuvants (ligands of various receptors of the innate immunity system) in the composition of subunit vaccines. The recombinant protein STF2.4×M2e, which is produced in *Escherichia coli* cells and includes flagellin (the toll-like receptor 5 (TLR-5) ligand), has protected immunized mice against a lethal dose of the influenza virus [25]. The safety and efficacy of a vaccine based on this construct was demonstrated in adult volunteers [26].

Intramuscular immunization of mice with the recombinant fusion protein 4×M2e·HSP70c produced in *E. coli* and consisting of sequential repeats of the M2e and HSP70 proteins of *Mycobacterium tuberculosis* resulted in a significant decrease in weight loss, a reduced viral titer in the lungs, and a less pronounced manifestation of the symptoms of the disease after the mice had been infected with a lethal dose of the influenza A H1N1, H3N2, or H9N2 viruses [27].

Virus-like particles (VLP)

Virus-like particles (VLP) are antigenic determinants of virions without genomic RNA fragments. Due to the absence of genetic material, VLP are incapable of infecting human and animal cells, which makes them safe [18]. The surface proteins of influenza VLP can be the conformational epitopes of the cells of the immune system as native virions.

It has been demonstrated in a number of studies that participation of the internal protein of the influenza virus M1 plays the key role in the formation of influenza VLP. This protein is bound to the lipid site of the apical plasma membrane domain, interacts with the surface glycoproteins of the influenza virus, and initiates assembly and budding of VLP containing the lipid membrane of the host cell, with three transmembrane proteins of the influenza virus incorporated into it [28].

Influenza VLP have been obtained in various expression systems. Either simultaneous expression of NA or addition of exogenous NA is required to provide efficient release of influenza VLP containing HA from mammalian cells. This fact can be attributed to the ability of active NA to cleave the sialic acids on the surface of the cell membrane [29]. Influenza VLP containing HA can be produced in insect cells even in the absence of NA expression, since the sialic acids in these cells are not bound to N-glycans during post-translational modification [30].

One of the approaches in producing influenza VLP in insect cells assumes the use of recombinant baculoviruses (*Fig. 2*) [1]. It has been demonstrated on animal models that the influenza surface antigens within VLP, which have been obtained using recombinant bacoloviruses, induce the production of both antihemagglutinating and virus-neutralizing antibodies and of the effectors of the cellular immune response. Furthermore,

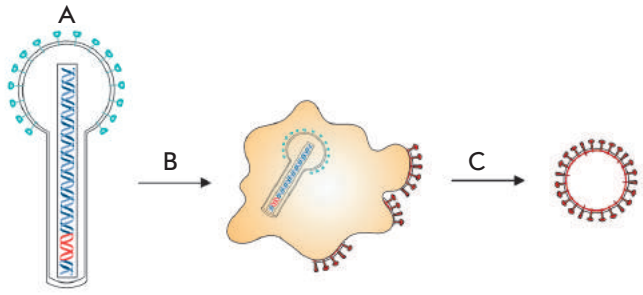


Fig. 2. Production of virus-like particles in the baculovirus expression system A – contraction of recombinant baculovirus expressing the gene of influenza antigen, B – transduction of insect cells, C – budding of the virus-like particles

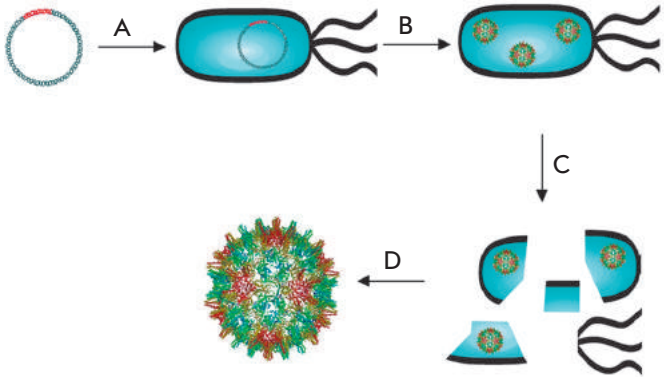


Fig. 3. Production of proteasomes. A – transformation of a producer strain by plasmid with the target gene, B – assembly and expression of proteasomes in cells, C – separation of proteasomes from the producer cells, D – purification and production of proteasomes

the influenza VLP vaccine induces protective immunity against homologous and heterologous strains of the influenza A virus [31].

A vaccine based on VLP carrying antigens of the pandemic influenza A H1N1(2009) virus has undergone phase II clinical trials in 4,563 healthy adult volunteers and has demonstrated that it's safe and has immunogenicity [32].

The use of recombinant baculoviruses for the expression of influenza virus proteins in insect cells results in the accumulation of baculoviruses, along with VLP, in the culture fluid. Since these structures are of similar sizes, it is difficult to isolate VLP from baculovirus particles. Influenza VLP can be generated in mammalian cells using other DNA- and viral vectors. Thus, a system for producing influenza VLP in Vero cells using DNA vectors carrying the genes of the HA, NA, M1 and M2 proteins of the influenza virus has been designed. The use of the modified vaccinia virus Ankara to generate VLP containing proteins of the influenza H5N1 (HA, NA, M1) virus in mammalian cells has been described. These VLP are capable of inducing a protective immune response in mice [33].

Thus, production of VLP is a promising direction in the efforts to design new types of influenza vaccines. In order to enhance the immunogenicity, attempts at introducing immune-stimulating components into the structure of influenza VLP have been made. For this purpose, recombinant baculoviruses carrying the flagellin (TLR-5 ligand) gene have been produced. The presence of recombinant flagellin within influenza VLP containing the HA of the influenza A/PR8 (H1N1) virus considerably enhanced the immunogenicity and protective properties of VLP after immunized mice had been infected with the heterologous strain of the influenza virus [34].

Proteasomes

Nano-sized structures containing the target antigen bound to a carrier consisting of biological macromolecules can be produced using genetic engineering techniques. The so-called proteasomes (a complex of proteins approximately 30-60 nm in diameter, which carries the target antigen on its surface) can be obtained by self-assembly of these macromolecules. Despite the fact that many authors refer to these structures as virus-like particles, opposite to VLP, proteasomes are formed on the basis of a carrier protein.

Proteasomes are most frequently based on virus coat proteins (e.g., the adenovirus penton [35], human papillomavirus L1 protein [36], hepatitis B virus HBc antigen [37] (*Fig. 3*).

Proteasomes containing the M1 protein of the influenza A virus bound to the structure comprising adenovirus surface proteins (dodecahedron) via the WW domain. The dodecahedron – antigen complex is capable of activating human dendritic cells, which introduce the antigen into cytotoxic T lymphocytes after activation [38]. The human papillomavirus L1 protein [36], the coat protein of Qβ bacteriophage [39], the papaya mosaic virus capsid protein [40], and the woodchuck hepatitis virus antigen have been used as a carrier of the influenza virus M2e protein or various epitopes of the M2 protein.

The hepatitis B virus HBc antigen, whose monomers can assemble into nano-sized particles, arouses the greatest interest as a carrier protein. These chimeric particles have been used as a carrier protein of the influenza virus M2e protein. The fusion protein M2e-HBc has been produced in *E. coli* cells. Immunization with recombinant M2e-HBc proteasomes has protected mice

against a lethal influenza infection even in the presence of preexisting antibodies against the HBc antigen [37]. The system for producing M2e-HBc proteasomes in *Nicotiana benthamiana* cells using the recombinant viral vector based on the potato virus X has been described [42].

The ability of proteasomes to carry a large number of antigenic determinants on their surface is an undoubted advantage [36]. However, the immunogenicity of the antigens represented in such a way is not always sufficient. The drawbacks of proteasomes also include their ability to carry small peptides only.

Genetic vaccines

The principle in designing any genetic vaccine consists in that a certain gene or region of the pathogen genome is incorporated into the carrier vector, which is subsequently used for vaccination. These vaccines provide the delivery of genetic material into the host cells and expression of the genes of the pathogen proteins in them. As a result, the pathogen antigens expressed by the cells in the organism are recognized by the immune system, which causes the induction of both the humoral and cell-mediated immune responses. The structure of the target antigens is very similar to that formed upon viral infection. Production of genetic vaccines does not require isolation and purification of antigens and, hence, handling pathogens. Furthermore, the use of various recombinant virus-based vectors can have an additional immunostimulating effect due to the presence of molecular pathogen-associated structures inducing innate immunity in them [43].

Among a great variety of genetic vaccines, three major groups can be distinguished: DNA vaccines, bacterial vector-based vaccines, and viral vector-based vaccines.

DNA vaccines

DNA vaccines are bacterial plasmids with the incorporated target gene and regulatory elements providing gene expression after this construct is introduced into the organism [44].

The levels of cell-mediated and humoral response induced by the introduction of a DNA vaccine are often insufficient for the development of immunity against pathogens. Therefore, DNA vaccines are typically used along with adjuvants to enhance immunogenicity and together with electroporation and gen-gun procedures (the latter method is delivery using a "gene gun," a device that injects microscopic DNA-coated particles) to provide better penetration of the genetic material into the cells [10].

Phase I clinical trials of the DNA vaccine expressing HA of the avian influenza virus, A/Vietnam/1203/04 (H5N1) where an adjuvant was used, has demonstrated the formation of hemagglutinin-inhibiting antibodies in 47-67% and induction of the T cell response in 75-100% of immunized volunteers. A 3-valent vaccine containing plasmids expressing NP, M2, and HA of the same influenza virus induced the T cell response in 72% of immunized individuals [45]. The use of a DNA vaccine for priming of the immune systems in combination with various other types of vaccines (VLP [46], an attenuated vaccine [47], a recombinant adenovirus [48]) appears rather promising.

Recombinant bacterial vector-based vaccines

Attenuated strains of bacteria, such as BCG, *Listeria monocytogenes, Salmonella typhi, S. typhimurium, Shigella flexneri*, etc., are used as bacterial vectors in designing genetic vaccines. Bacterial vectors are characterized by the ability to deliver an antigen to the antigen-presenting cells and the possibility of producing vaccines for intramucosal introduction. The use of bacterial vectors activates the innate immunity as a result of the interaction between the bacterial components and the receptors of the innate immunity system [49].

Immunization of mice with *L.monocytogenes*-based bacterial vectors carrying the *NP* gene of the influenza A virus reduced the influenza virus titer in the lungs of infected mice [50]. The safeness and immunogenicity of this vaccine has been demonstrated in volunteers [51]. The use of the *Bordetella pertussis*-based vaccine vector BPZE1 carrying the M2e protein gene of the influenza A virus induced the formation of anti-M2e antibodies in mice and reduced the influenza virus titers in the lungs after the animals had been infected with A/PR8 (H1N1). However, this vaccine failed to provide complete protection when the animals had been infected with a lethal dose of the virus [52].

When using bacterial vectors, the resulting immune response is not always sufficient to provide protection; therefore, additional means to enhance the vaccine's immunogenicity should be employed. The possibility of transferring the plasmid carrying the transgene to other bacteria is a serious downside in the case of bacterial vectors. What's more, there is a possibility of insertional mutagenesis [53].

Recombinant viral vector-based vaccines

Viral vectors are recombinant viruses with the target gene and a combination of regulatory elements incorporated into their genome. Viral vectors hold a special position among the existing antigen delivery systems due to the fact that they possess the following properties: a natural mechanism of interaction with cells and penetration into them; they deliver foreign genetic material to the cell nuclei; are capable of providing long-

term antigen expression; and their capsid protects the antigen-encoding genetic material [54].

Viral vector-based vaccines efficiently activate cytotoxic T lymphocytes, which play a particularly significant role when performing vaccination against intracellular pathogens. These vaccines can have a broad range of activities due to the induction of the T cell response to conserved epitopes that are potentially capable of ensuring protection against various pathogenic strains (including the influenza virus) [55].

Viral vectors are capable of activating the innate immunity by binding the genetic material or their capsid proteins to pattern-recognition receptors (TLR, RIG-1, etc.) [56]. Viral vectors are recognized by TLR, such as TLR2, TLR3, TLR4, TLR7, TLR8, and TLR9. The interaction between these receptors and ligands results in the activation of various transcription factors, which leads to the formation of an inflammation locus and rapid activation of the defense reactions of the organism [57].

One needs to be guided by the following criteria when choosing a viral vector for genetic immunization: the vaccine should not cause any symptoms of the disease; it needs to be safe for immune-deprived individuals, as well as for elderly people and children; the intrinsic proteins of the recombinant virus should not cause a strong immune response; the viral vector needs to be simple for genetic manipulations and be capable of incorporating large fragments of foreign DNA; the resulting vectors need to have a high viral titer and provide a high expression level of the target antigens; the DNA of a viral vector should not be integrated into the host cell genome after the immunization; and the vector needs to be completely eliminated from the organism after the immune response is induced. Furthermore, the presence of a preexisting immune response to the proteins of the viral vector in immunized individuals is undesirable, since it can considerably reduce the level of the immune response to the target antigen [58].

Not all the viruses possess the properties required for the construction of efficient vectors. Poxviruses [59], the recombinant Newcastle disease virus [60], and adenoviruses [61] are those most frequently used today to design viral vector-based influenza vaccines.

Recombinant poxviruses

Poxviruses (Poxviridae) are DNA-containing viruses with a large genome. The vaccinia virus is a poxvirus that is most commonly used as a viral vector; its advantages include simple and inexpensive production, as well as high packaging capacity (up to 25 thousand nucleotide pairs) [59]. Attenuated vaccinia viruses (such as the modified vaccinia Ankara virus and the attenuated NYVAC strain based on the Copenhagen strain) are used for vaccine production. MVA was attenuated by repeated passivation in chicken embryo fibroblasts, which resulted in the loss of a number of genes that are not essential for replication in avian cells and in reduced reproduction in human cells. Attenuation of the NYVAC strain was achieved via deletion of 18 genes; as a result, the virus became replicatively defective for human cells [62].

It has been demonstrated that immunization of mice with MVA expressing the HA genes of the highly pathogenic avian influenza H5N1 virus protects mice against both the homologous and heterologous strains of the influenza H5N1 virus, as well as induces virus-neutralizing antibodies and HA-specific $CD4^+$- and $CD8^+$ T cells [63]. The MVA-based vaccine expressing the HA gene of the influenza A/California/07/09 (H1N1) virus proved efficient in the double immunization of mice, macaques, and polecats [64]. The efficiency of the vaccine based on the NYVAC strain expressing the HA gene of the avian influenza A(H5N1) virus was demonstrated for pigs [65].

A serious drawback of vaccinia virus-based vectors is the preexisting immunity to this virus, which formed in the human population as a result of immunization against smallpox. Therefore, it is reasonable to use vectors based on the canarypox and fowlpox viruses, against which there are no preexisting antibodies in the human population. Immunization of chickens and ducks with the recombinant fowlpox virus with the HA gene of the avian influenza A virus incorporated into its genome has protected birds against infection with lethal doses of homologous influenza viruses [66]. The high packaging capacity of poxviruses allows one to simultaneously introduce several transgenes (e.g., the HA and NP genes of the influenza A virus) into the genome [66]. However, the canarypox and fowlpox viruses induce a weaker immune response to the target antigens, as compared to that induced by the vaccinia virus, and require repeated immunization or the use of adjuvants [66].

Recombinant Newcastle disease virus

The Newcastle disease virus (NDV) belongs to the Paramyxoviridae family. This virus has a nonsegmented single-stranded RNA genome containing six genes that encode seven proteins: the NP, P and V proteins, the M protein, the fusion protein or F protein, HA–NA, and the large polymerase protein L. Since the expression level of each viral protein decreases in the direction from the 3' to the 5' terminus of the genome, when NDV is used as a vector, the expression level of a foreign gene can be controlled based on its position in the viral genome. The virulence level and tropicity of NDV depend on the site of the fragmentation of the F protein

by nucleases, which is required to provide fusion of the viral coat and the cell membrane. Thus, the virus virulence can be altered via amino acid replacements in the F protein, which can be regarded as a convenient basis for constructing vaccine vectors [60].

NDV expressing the HA gene of the influenza A/WSN/3 (H1N1) virus has been constructed using the reverse genetics method. Mice has been successfully protected against infection with the influenza A/WSN/3 (H1N1) virus using this construct [60]. NDV expressing the HA genes of the highly pathogenic avian influenza A H5 and H7 viruses has protected immunized birds against infection with lethal doses of homologous influenza A viruses. The efficiency of immunization with the recombinant NDV expressing the HA gene of the highly pathogenic avian influenza A H5 virus was demonstrated in mice [67].

In nature, only birds are infected by NDV; hence, humans have no antibodies against this virus. Therefore, there is no problem of preexisting immune response for this viral vector. However, a significant drawback of the vaccine vector NDV is that the consequences of the introduction of recombinant NDV have not been sufficiently studied and it remains unclear whether NDV-based influenza vaccines are safe for humans. Furthermore, NDV is characterized by a low packaging capacity and a complex procedure in constructing vectors carrying several target antigens. Preparative amounts of NDV are produced in chicken embryos, a method which has a number of drawbacks, as shown above [68].

Recombinant adenoviruses

Recombinant adenoviruses (Adenoviridae) are the best studied and most frequently used recombinant viral vectors. Adenovirus virions consist of a double-stranded DNA molecule surrounded by a protein capsid.

A number of adenovirus types have been thoroughly characterized (at the genetic level as well). The genomes of most of them have been fully sequenced. Detailed data on the structure, physicochemical, and biological properties of adenoviruses enable their use in designing recombinant vaccines and gene therapeutic agents [61]. Approximately 24% of the genetic vaccines that are currently undergoing clinical trials are vaccines based on recombinant adenoviruses (clinicaltrials.gov) (*Fig. 4*).

Adenoviruses possess significant properties for vaccine vectors: they are capable of providing high levels of expression of the target transgene in the target cell and of transducing both dividing and postmitotic cells. Adenovirus DNA remains in the extrachromosomal form. Adenoviruses can be accumulated to high titers in cell culture. The process of designing a new recombinant adenovirus takes several weeks, which allows a prompt response to a changing epidemiological situation [61].

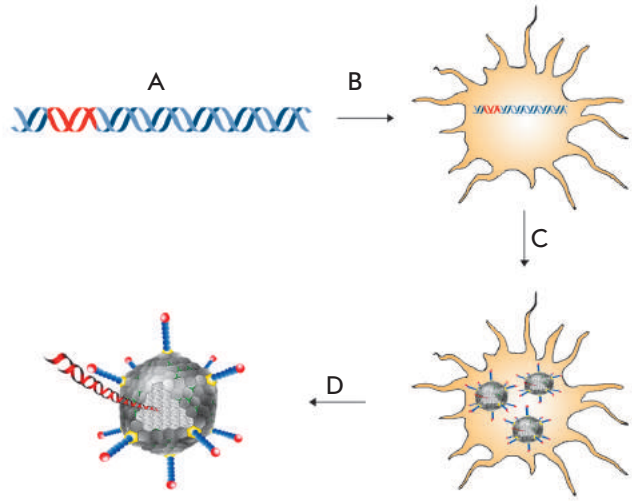


Fig. 4. Production of recombinant adenovirus. A – production of genomic DNA of human recombinant adenovirus expressing the genes of influenza antigens, B – transfection of permissive cell line with recombinant virus DNA, C – expression of virus genes in eukaryotic cells and assembly of recombinant virus particles, D – purification of adenovirus virions from a cell suspension

Vaccines based on recombinant adenoviruses against a number of pathogens causing such diseases as malaria, tuberculosis, brucellosis, etc. [69, 70] and various viruses (influenza A virus, human immunodeficiency virus, human papillomavirus, rabies virus, Ebola virus, etc. [71–74]) are currently under development.

The best studied representative of adenoviruses, human adenovirus serotype 5 (Ad5), is the most commonly used among the adenoviruses used to construct recombinant viral vectors [75, 76]. Replication-defective Ad5 are used to produce vaccines and gene therapeutic agents. In these Ad5, various genome regions (E1, E2, E3, E4) essential for virus replication are deleted. Cell lines complementing the functions of the removed regions *in trans* have been designed to produce and accumulate these viruses. The vectors enable inserting up to 10,000 bp [77].

When injected into the organism, adenoviruses are capable of activating TLR-9 and RIG-1 receptors. The innate immunity is simultaneously activated as a result of adenovirus penetration of antigen-presenting cells [78].

Adenovirus-transduced dendritic cells expressing the target antigen or the activated dendritic cells that have captured the antigen produced by epithelial cells

act as an interlink between the innate and adaptive immunity systems. Upon mucosal immunization, priming of dendritic cells occurs in mucosal tissues; hence, activated T and B lymphocytes (as well as the memory cells originating from them) acquire the ability to express α4β7 integrin. This molecule allows T and B lymphocytes to migrate through the endothelium layer to submucous tissues to a spot where it is possible to enter in contact with a pathogen [79].

It has been demonstrated in experiments with laboratory animals that cross-immunity is developed after mucosal immunization with vaccines of various types [80]. The major component of the adaptive immunity of mucosal tunics is antibodies, which mostly refer to secretory immunoglobulin A (sIgA), to secretory immunoglobulin M (sIgM) to a smaller extent, and to IgG of both plasmatic and local origin. Expression of sIgA presumably determines the cross-protectivity of the vaccine [81]. The other advantages of mucosal vaccines over injective ones include the absence of skin damage during the immunization and lower reactogenicity [82].

Thanks to the activation of the innate immune response, intranasal introduction of recombinant Ad5 carrying no transgenes into mice can also protect against the influenza A virus, since it induces production of a broad range of anti-inflammatory cytokines and chemokines (including type I interferons) and nitrogen oxide activates natural killer (NK) cells. The protective effects caused by the introduction of Ad5 carrying no transgenes are retained for at least 3 weeks after a single intranasal immunization. The introduction of recombinant adenovirus protects nonspecifically against low doses of the influenza virus during the period of time that is required for the formation of the adaptive immune response to the target antigens. Thus, the protective effect of Ad5-based vaccines starts almost immediately after the immunization [83]; due to the specific immune answer to the transgene, the effect lasts for over 6 months.

Vaccines based on recombinant adenoviruses against various influenza A serotypes are currently under development in different countries. One such vaccine has successfully passed phase I clinical trials in the USA and proved to be safe for humans and highly immunogenic for the influenza A H5N2 virus [84].

The problem related to the design of influenza vaccines triggering the heterosubtypic immunity that can protect against various strands of the influenza virus is urgent; new approaches in solving the problem have appeared recently. The conformational epitopes of HA have been identified for various influenza A subtypes; broad-spectrum antibodies against these epitopes can be secreted both after the infection and after live virus vaccination [85]. Vaccination with recombinant adenoviral vectors imitates an infection of mucosal cells of the upper air passages, thus providing expression of antigens with a native tertiary structure, which allows to trigger the formation of these cross-reactive antibodies. Recombinant adenoviral vaccines can also induce a strong T cell immune response characterized by a broader spectrum of action compared to that of the humoral immune response.

The possibility of using recombinant Ad5 to induce a heterosubtypical immune response against the influenza A virus has been studied at the Molecular Biotechnology Laboratory of the Gamaleya Research Institute of Epidemiology and Microbiology. It has been demonstrated that twice-repeated intranasal immunization of mice with the recombinant adenovirus carrying the HA gene of the influenza A H5N2 virus provides high-level induction of specific antibodies against this virus and ensures complete protection of mice against infection with a lethal dose of the H5N2 virus (50 LD_{50}) [86]. The mice immunized with this recombinant adenovirus were also protected against infection with the influenza H1N1 and H2N3 viruses, which belong to the H1 group (H1, H2, H5, H6, H11, H13, and H16); however, they were not protected against infection with the influenza H3N2 virus, which belongs to the H3 group (H3, H14, and H4) [87, 80].

The data obtained allow to assume that a panel of adenoviral vectors carrying the HA genes of influenza A viruses belonging to different groups can be used to design a vaccine that would protect against most epidemic strains of the influenza A virus.

A serious drawback of using Ad5 as a vector for designing vaccines is that most people have anti-Ad5 antibodies. The presence of these antibodies can significantly reduce immunization efficacy. However, it has been demonstrated in a number of studies that Ad5-based vaccines can avoid the effect of the preexisting immune response upon intranasal immunization (as opposed to parenteral introduction) [88–90]. Upon intranasal introduction of Ad5, the transgene is efficiently delivered through the mucosal barrier. Even a single intranasal administration of Ad5-based vaccines results in prolonged *ex vivo* expression of the transgene despite the preexisting immunity both in laboratory animals and in primates [89].

A single intranasal immunization of mice with a recombinant adenoviral vector carrying the HA gene of the avian influenza A H5N2 virus has protected immunized animals against infection with this virus. No differences between the levels of protection were observed in mice with virus-neutralizing anti-Ad5 antibodies present in their blood and in Ad5-naive mice [80].

The intranasal immunization of mice with a preexisting immune response to Ad4, recombinant Ad4 carrying the HA gene of the influenza virus, resulted in a lower level of production of antibodies against the influenza virus as compared to that in nonprimed mice and a decrease in the cell-mediated immune response by over 2 orders of magnitude (depending on the dose of recombinant Ad4). However, despite the preexisting immunity, the animals remained fully protected against infection with a lethal dose of the influenza virus [74]. Analogous data were obtained for Ad5 carrying the HA and NP genes of the influenza virus. An increase in the dose of the recombinant adenovirus leveled the reduction of the immune response [90].

Incorporation of elements that allow an optimization of the expression level of the transgene into the vector and selection of the optimal dose of Ad5 made it possible to achieve a significant level of transgene-specific $CD8^+$ cells in immunized animals even at high levels of Ad5-neutralizing preexisting antibodies [91].

Thus, recombinant Ad5 carrying genes of various antigens of the influenza A virus are rather promising as candidates for influenza vaccines. They are safe, efficacious, and can be used to design a universal intranasal influenza vaccine.

CONCLUSIONS

The data presented in this review attest to a vigorous research effort aimed at constructing influenza vaccines using new approaches that employ the promise of reverse genetics methods and recombinant technologies, as well as the production of VLP, proteasomes, and subunit vaccines in various expression systems. The new approaches have enabled to achieve significant progress in the design of new influenza vaccines. Some of these vaccines are currently undergoing either preclinical or clinical trials. Among the vectors used to design genetic vaccines, adenoviral vectors hold a special position. They are capable of efficiently penetrating the respiratory mucosal tunic, which makes it possible to achieve mucosal immunization, thus ensuring a lasting presence of the antigen in the organism and activation of the innate immunity. Human recombinant adenoviruses serotype 5 carrying the genes of various antigens of the influenza A virus have the potential of being used as influenza vaccines. They are safe, efficacious, and could allow to design a universal influenza vaccine delivered intranasally. Upon immunization, the recombinant adenovirus acts as an adjuvant; it is capable of boosting immunity with respect to the transgene. Producing this vaccine takes several weeks, which would allow to respond promptly to a changing epidemiological situation. Recombinant adenoviral vectors carrying the HA genes of various subtypes of influenza A viruses can be used to form a heterosubtypic immune response against most epidemic variants of the influenza A virus. Thus, adenoviruses can be used to design a universal recombinant influenza vaccine. ●

REFERENCES

1. Girard M., Cherian T., Pervikov Y., Kieny M. // Vaccine. 2005. V. 23. № 50. P. 5708–5724.
2. Taubenberger J., Morens D. // Emerg. Infect. Dis. 2006. V. 12. № 1. P. 15–22.
3. Gendon Iu. // Vopr. Virusol. 2007. V. 52. № 1. P. 4–10.
4. Kiselev O.I. // Biotechnology in Russia. 2010. № 2. P. 1–18.
5. Bodewes R., Osterhaus A., Rimmelzwaan G. // Viruses. 2010. V. 2. P. 166–188.
6. Nichol K. // Vaccine. 2001. V. 19. № 31. P. 4373–4377.
7. Swayne D. // Ann. N.Y. Acad. Sci. 2006. V. 1081. P. 226–227.
8. Mahmud M., Maassab H., Jennings R., Potter C. // Arch. Virol. 1979. V. 61. № 3. P. 207–216.
9. Snyder M., Clements M., Betts R., Dolin R., Buckler-White A., Tierney E., Murphy B. // J. Clin. Microbiol. 1986. V. 23. № 5. P. 852–857.
10. Plotkin S. // Clin. Vaccine Immunol. 2009. V. 16. № 12. P. 1709–1719.
11. Tumpey T., Basler C., Aguilar P., Zeng H., Solórzano A., Swayne D., Cox N., Katz J., Taubenberger J., Palese P., et al. // Science. 2005. V. 310. № 5745. P. 77–80.
12. Song H., Nieto G., Perez D. // J. Virol. 2007. V. 81. № 17. P. 9238–9248.
13. Li Z., Jiang Y., Jiao P., Wang A., Zhao F., Tian G., Wang X., Yu K., Bu Z., Chen H. // J. Virol. 2006. V. 80. № 22. P. 11115–11123.
14. Wacheck V., Egorov A., Groiss F., Pfeiffer A., Fuereder T., Hoeflmayer D., Kundi M., Popow-Kraupp T., Redlberger-Fritz M., Mueller C., et al. // J. Infect. Dis. 2010. V. 201. № 3. P. 354–362.
15. Watanabe T., Watanabe S., Kim J., Hatta M., Kawaoka Y. // J. Virol. 2007. V. 82. № 5. P. 2486–2492.
16. Middleton D., Bingham J., Selleck P., Lowther S., Gleeson L., Lehrbach P., Robinson S., Rodenberg J., Kumar M., Andrew M. // Virology. 2007. V. 359. № 1. P. 66–71.
17. Harvey R., Wheeler J., Wallis C., Robertson J., Engelhardt O. // Vaccine. 2008. V. 26. № 51. P. 6550–6554.
18. Vasil'ev Iu. // Vopr Virusol. 2008. V. 6. № 1. P. 4–15.
19. Kiselev O., Tsybalova L., Pokrovskiĭ V. // Zh. Mikrobiol. Epidemiol. Immunobiol. 2006. V. 5. P. 28–38.
20. Wang C., Luo Y., Chen Y., Li S., Lin C., Hsieh Y., Liu H. // J. Virol. Methods. 2007. V. 146. № 1–2. P. 293–297.
21. Deroo T., Jou W., Fiers W. // Vaccine. 1996. V. 14. № 6. P. 561–569.
22. Lin S., Huang M., Tsou P., Huang L., Chong P., Wu S. // PLoS One. 2011. V. 6. № 5. e20052.
23. Tsybalova L., Kiselev O. // Vopr. Virusol. 2012. V. 57. № 1. P. 9–14.
24. Nemchinov L., Natilla A. // Protein Expr. Purif. 2007. V. 56. P. 153–159.
25. Huleatt J., Nakaar V., Desai P., Huang Y., Hewitt D., Jacobs A., Tang J., McDonald W., Song L., Evans R., et al. //

Vaccine. 2008. V. 26. № 2. P. 201–214.
26. Turley C., Rupp R., Johnson C., Taylor D., Wolfson J., Tussey L., Kavita U., Stanberry L., Shaw A. // Vaccine. 2011. V. 29. № 32. P. 5145–5152.
27. Ebrahimi S., Dabaghian M., Tebianian M., Zabeh Jazi M. // Virology. 2012. V. 430. № 1. P. 63–72.
28. Latha T., Galarza J. // J. Virol. 2001. V. 75. № 13. P. 6154–6165.
29. Chen B., Leser G., Morita E., Lamb R. // J. Virol. 2007. V. 81. № 13. P. 7111–7123.
30. Galarza J., Latham T., Cupo A. // Viral. Immunol. 2005. V. 18. № 2. P. 365–367.
31. Bright R., Carter D., Daniluk S., Toapanta F., Ahmad A., Gavrilov V., Massare M., Pushko P., Mytle N., Rowe T., et al. // Vaccine. 2007. V. 25. № 19. P. 3871–3878.
32. López-Macías C. // Hum. Vaccin. Immunother. 2012. V. 8. № 3. P. 411–414.
33. Tang X., Lu H., Ross T. // Viral. Immunol. 2011. V. 24. № 4. P. 311–319.
34. Wang B., Quan F., Kang S., Bozja J., Skountzou I., Compans R. // J. Virol. 2008. V. 82. № 23. P. 11813–11823.
35. Fender P., Ruigrok R., Gout E., Buffet S., Chroboczek J. // Nat. Biotechnol. 1997. V. 15. № 1. P. 52–56.
36. Matic S., Rinaldi R., Masenga V., Noris E. // BMC Biotechnol. 2011. V. 11. № 1. P. 106.
37. De Filette M., Martens W., Smet A., Schotsaert M., Birkett A., Londoño-Arcila P., Fiers W., Saelens X. // Vaccine. 2008. V. 26. № 51. P. 6503–6507.
38. Naskalska A., Szolajska E., Chaperot L., Angel J., Plumas J., Chroboczek J. // Vaccine. 2009. V. 27. № 52. P. 7385–7393.
39. Bessa J., Schmitz N., Hinton H., Schwarz K., Jegerlehner A., Bachmann M. // Eur. J. Immunol. 2008. V. 38. № 1. P. 114–126.
40. Denis J., Acosta-Ramirez E., Zhao Y., Hamelin M., Koukavica I., Baz M., Abed Y., Savard C., Pare C., Lopez Macias C., et al. // Vaccine. 2008. V. 26. № 27–28. P. 3395–3403.
41. Ameiss K., Ashraf S., Kong W., Pekosz A., Wu W.H., Milich D., Billaud J., Curtiss R. 3rd. // Vaccine. 2010. V. 28. № 41. P. 6704–6713.
42. Ravin N.V., Kotlyarov R.Y., Mardanova E.S., Kuprianov V.V., Migunov A.I., Stepanova L.A., Tsybalova L.M., Kiselev O.I., Skryabin K.G. // Biochemistry (Mosc.). 2012. V. 77. № 1. P. 33–40.
43. Bessis N., GarciaCozar F., Boissier M. // Gene Ther. 2004. V. 11. Suppl 1. S. 10–17.
44. Abdulhaqq S., Weiner D. // Immunol. Res. 2008. V. 42. P. 219–232.
45. Smith L., Wloch M., Ye M., Reyes L., Boutsaboualoy S., Dunne C., Chaplin J., Rusalov D., Rolland A., Fisher C., et al. // Vaccine. 2010. V. 28. № 13. P. 2565–2572.
46. Ding H., Tsai C., Gutiérrez R.A., Zhou F., Buchy P., Deubel V., Zhou P. // PLoS One. 2011. V. 6. № 1. e16563.
47. Suguitan A.L. Jr., Cheng X., Wang W., Wang S., Jin H., Lu S. // PLoS One. 2011. V. 6. № 7. e21942.
48. Guo J., Yao L., Chen A., Liu X., Fu J., Xu P., Zhang Z. // Sheng Wu Gong Cheng Xue Bao. 2011. V. 27. № 6. P. 876–883.
49. Stoeker L., Nordone S., Gunderson S., Zhang L., Kajikawa A., LaVoy A., Miller M., Klaenhammer T., Dean G. // Clin. Vaccine Immunol. 2011. V. 18. № 11. P. 1834–1844.
50. Ikonomidis G., Portnoy D.A., Gerhard W., Paterson Y. // Vaccine. 1997. V. 15. № 4. P. 433–440.
51. Johnson P., Blair B., Zeller S., Kotton C., Hohmann E. // Microbiol. Immunol. 2011. V. 55. № 5. P. 304–317.
52. Li R., Lim A., Ow S., Phoon M., Locht C., Chow V., Alonso S. // Vaccine. 2011. V. 29. № 33. P. 5502–5511.
53. Hense M., Domann E., Krusch S., Wachholz P., Dittmar K., Rohde M., Wehland J., Chakraborty T., Weiss S. // Cell Microbiol. 2001. V. 3. № 9. P. 599–609.
54. Draper S., Heeney J. // Nat. Rev. Microbiol. 2010. V. 8. P. 62–73.
55. Wei C., Boyington J., McTamney P., Kong W., Pearce M., Xu L., Andersen H., Rao S., Tumpey T., Yang Z., et al. // Science. 2010. V. 329. P. 1060–1064.
56. Tukhvatulin A. , Shcherbinin D., Logunov D., Shmarov M., Naroditskiĭ B.. // Vestn. RAMN. 2011. № 10. P. 47–54.
57. Tutykhina I.L., Logunov D.Y., Shcherbinin D.N., Shmarov M.M., Tukhvatulin A.I., Naroditsky B.S., Gintsburg A.L. // J. Mol. Med. (Berl.). 2011. V. 89. № 4. P. 331–341.
58. Kopecky-Bromberg S., Palese P. // Curr. Top. Microbiol. Immunol. 2009. V. 333. P. 243–267.
59. Gherardi M., Esteban M. // J. Gen. Virol. 2005. V. 86. № 11. P. 2925–2936.
60. Nakaya T., Cros J., Park M., Nakaya Y., Zheng H., Sagrera A., Villar E., García-Sastre A., Palese P. // J. Virol. 2001. V. 75. № 23. P. 11868–11873.
61. Karpov A. P., Tutykhina I.L., Logunov D.Y., Verkhovskaya L.V., Shmarov M.M., Valichov A.F., Shul'pin M.I ., Drygin V.V., Naroditskii B.S. // Biotechnology in Russia. 2007. V. 5. P. 38–44.
62. Pantaleo G., Esteban M., Jacobs B., Tartaglia J. // Curr. Opin. HIV AIDS. 2010. V. 5. № 5. P. 391–396.
63. Hessel A., Schwendinger M., Holzer G., Orlinger K., Coulibaly S., Savidis-Dacho H., Zips M., Crowe B., Kreil T., Ehrlich H., et al. // PLoS One. 2011 V. 6. № 1. e16247.
64. Kreijtz J., Süzer Y., Bodewes R., Schwantes A., van Amerongen G., Verburgh R., de Mutsert G., van den Brand J., van Trierum S., Kuiken T., et al. // J. Gen. Virol. 2010. V. 91.(Pt 11). P. 2745–2752.
65. Kyriakis C., De Vleeschauwer A., Barbé F., Bublot M., van Reeth K. // Vaccine. 2009. V. 27. № 16. P. 2258–2264.
66. Webster R., Kawaoka Y., Taylor J., Weinberg R., Paoletti E. // Vaccine. 1991. V. 9. № 5. P. 303–308.
67. Ge J., Deng G., Wen Z., Tian G., Wang Y., Shi J., Wang X., Li Y., Hu S., Jiang Y., et al. // J. Virol. 2007. V. 81. № 1. P. 150–158.
68. Xie Y., Sun H., Li D. // Chem. Biodivers. 2010. V. 7. № 3. P. 677–689.
69. Radosevic K., Rodriguez A., Lemckert A., van der Meer M., Gillissen G., Warnar C., von Eyben R., Pau M., Goudsmit J. // Clin. Vaccine. Immunol. 2010. V. 17. № 11. P. 16876–16894.
70. Ronan E., Lee L., Beverley P., Tchilian E. // PLoS One. 2009. V. 4. № 12. e8235.
71. Vemula S., Mittal S. // Expert. Opin. Biol. Ther. 2010. V. 10. № 10. P. 1469–1487.
72. Paris R., Kim J., Robb M., Michael N. // Expert. Rev. Vaccines. 2010. V. 9. № 9. P. 1055–1069.
73. Knowles M., Roberts D., Craig S., Sheen M., Nadin-Davis S., Wandeler A. // Vaccine. 2009. V. 27. № 20. P. 2662–2668.
74. Alexander J., Ward S., Mendy J., Manayani D., Farness P., Avanzini J., Guenther B., Garduno F., Jow L., Snarsky V., et al. // PLoS One. 2012. V. 7. № 2. e31177.
75. Rogozhin V.N., Beloysova R.V., Logunov D.Y., Shmarov M.M., Lunin V.G., Naroditskii B.S. // Veterinary medicine. 2011. №2.
76. Gribova I.Y., Tillib S.V., Tutykhina I.L., Shmarov M. M., Logunov D.Yu., Verkhovskaya L.V., Naroditskii B.S., Gintsburg A.L. // Acta Naturae. 2011. V. 3. № 3 (10). P. 66–72.

77. Gao G., Yang Y., Wilson J. // J. Virol. 1996. V. 70. № 12. P. 8934–8943.
78. Tutykhina I.L., Shcherbinin D.N., Shmarov M.M., Logunov D.Iu., Naroditskiĭ B.S. // Vestn. RAMN. 2011. № 10. P. 37–49.
79. Tamura S., Asanuma H., Ito Y., Hirabayashi Y., Suzuki Y., Nagamine T., Aizawa C., Kurata T., Oya A. // Eur. J. Immunol. 1992. V. 22. № 2. P. 477–481.
80. Shmarov M.M., Sedova E.S., Verkhovskaya L.V., Bogacheva E.A., Barykova Y.A., Shcherbinin D.N., Lisenko A.A., Tutykhina I.L., Neugodova G.L., Logunov D.Iu., et al. // Biopreparati. 2011. №1. P.31–35.
81. Roy C., Ault A., Sivasubramani S., Gorres J., Wei C., Andersen H., Gall J., Roederer M., Rao S. // Respir. Res. 2011. V. 12. P. 153.
82. Ugwoke M., Agu R., Verbeke N., Kinget R. // Adv. Drug Deliv. Rev. 2005. V. 57. № 11. P. 1640–1665.
83. Zhang J., Tarbet E., Toro H., Tang D.C. // Expert. Rev. Vaccines. 2011. V. 10. № 11. P. 1539–1552.
84. van Kampen K., Shi Z., Gao P., Zhang J., Foster K., Chen D., Marks D., Elmets C., Tang D.C. // Vaccine. 2005. V. 23. № 8. P. 1029–1036.
85. Sahini L., Tempczyk-Russell A., Agarwal R. // PLoS One. 2010. V. 5. № 2. e9268.
86. Sedova E.S., Shmarov M.M., Tutykhina I.L., Barykova Iu.A., Verkhovskaia L.V., Logunov D.Iu., Naroditskiĭ B.S., Gintsburg A.L. // Zh. Mikrobiol. Epidemiol. Immunobiol. 2010. № 3. P. 44–8.
87. Shmarov M.M., Sedova E.S., Verkhovskaya L.V., Rudneva I.A., Bogacheva E.A., Barykova Y.A., Shcherbinin D.N., Lysenko A.A., Tutykhina I.L., Logunov D.Y., et al. // Acta Naturae. 2010. V. 2. № 1(4). P. 119–26.
88. Croyle M., Patel A., Tran K., Gray M., Zhang Y., Strong J., Feldmann H., Kobinger G. // PLoS One. 2008. V. 3. № 10. e3548.
89. Zabner J., Petersen D., Puga A., Graham S.M., Couture L., Keyes L., Lukason M., St. George J., Gregory R., Smith A. // Nat. Genet. 1994. V. 6. № 1. P. 75–83.
90. Pandey A., Singh N., Vemula S., Couëtil L., Katz J., Donis R., Sambhara S., Mittal S. // PLoS One. 2012. V. 7. № 3. e33428.
91. Steffensen M., Jensen B., Holst P., Bassi M., Christensen J., Thomsen A. // PLoS One. 2012. V. 7. № 4. e34884.

Epigenetics of Pluripotent Cells

S. P. Medvedev[1,2,3], E. A. Pokushalov[1,2,3], S. M. Zakian[1,2,3*]
[1]Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Prospekt Lavrentyeva, 10, Novosibirsk, Russia, 630090
[2]Meshalkin Novosibirsk State Research Institute of Circulation Pathology, Rechkunovskaja Str., 15, Novosibirsk, Russia, 630055
[3]Institute of Chemical Biology and Fundamental Medicine, Siberian Branch, Russian Academy of Sciences, Prospekt Lavrentyeva, 8, Novosibirsk, Russia, 630090
*E-mail: zakian@bionet.nsc.ru
Received 08.08.2012

ABSTRACT **Pluripotency is maintained by a complex system that includes the genetic and epigenetic levels. Recent studies have shown that the genetic level (transcription factors, signal pathways, and microRNAs) closely interacts with the enzymes and other specific proteins that participate in the formation of the chromatin structure. The interaction between the two systems results in the unique chromatin state observed in pluripotent cells. In this review, the epigenetic features of embryonic stem cells and induced pluripotent stem cells are considered. Special attention is paid to the interplay of the transcription factors OCT4, SOX2, and NANOG with the Polycomb group proteins and other molecules involved in the regulation of the chromatin structure. The participation of the transcription factors of the pluripotency system in the inactivation of the X chromosome is discussed. In addition, the epigenetic events taking place during reprogramming of somatic cells to the pluripotent state and the problem of "epigenetic memory" are considered.**
KEYWORDS **embryonic stem cells; induced pluripotent stem cells; pluripotency; covalent histone modifications; DNA methylation.**
ABBREVIATIONS **ESC – embryonic stem cells; iPSC – induced pluripotent stem cells; DMR – differentially methylated regions; ICM – inner cell mass.**

INTRODUCTION

Pluripotency is the property of cells to differentiate into the derivatives of all three primary germ layers – ectoderm, endoderm, and mesoderm - as well as to form precursor cells of functional gametes during embryonic development. Inner cell mass (ICM) cells and epiblast cells from pre-implantation mammalian embryos are pluripotent [1]. The adult organism is formed from pluripotent cells during ontogenesis. However, these cells cannot give rise to extraembryonic organs and tissues.

Embryonic stem cells (ESCs) are obtained from the inner cell mass of pre-implantation embryos [2–4]. Under optimal cultivation conditions, ESCs can retain a number of properties intrinsic to the inner cell mass and embryonic epiblast cells (including pluripotency) for a long period [2–4]. The pluripotency makes ESCs promising for use in fundamental and applied research. ESCs are used as model systems to study the processes occurring during early embryogenesis in mammals and to *in vitro* simulate various diseases. Furthermore, pluripotent cells are a promising source of material for substitutive cellular therapy [5–7].

After the first mouse and human ESC lines were obtained, research into the molecular genetic basis involved in maintenance of the undifferentiated pluripotent state of ESCs started. It is known today that the pluripotent state of the cells of pre-implantation embryos and ESCs is maintained via a complex system of cell surface proteins, their molecular signal pathways, and the transcription factors that initiate the transcription of the target genes. The subsystem of the so-called "external regulators of pluripotency" includes several signaling pathways, among which the cascades triggered by the proteins LIF, BMP4, TGFβ, activin A, NODAL, and bFGF (FGF2) are the major ones [1].

ESC pluripotency is also controlled by the subsystem of "internal regulators of pluripotency" – transcription factors functioning in cell nuclei. The factors OCT4, NANOG, and SOX2 are among the key regulators in this subsystem [8, 9].

In 2006, the data on reprogramming of mouse somatic cells into the pluripotent state were published in *Cell* by a group of Japanese researchers [10]. This was one of the most outstanding discoveries of the past decade in the field of cell biology. Cells obtained by the repro-

gramming of somatic cells were called induced pluripotent stem cells (iPSCs) [10].

The development of the technology for obtaining animal and human induced pluripotent stem cells has opened up a broad range of possibilities for studying the dynamics of the epigenetic events occurring upon reprogramming and the features of the epigenomes of pluripotent cells. A large number of well-reproducible methods for obtaining iPSCs from a broad range of somatic cells are known today. Most researchers use a certain gene combination for reprogramming; many of these genes encode transcription factors (e.g., *Oct4*, *Sox2*, *Klf4*, *c-Myc*, *Nanog*, and *Lin28* genes) [10–13]. Furthermore, it has been demonstrated that mouse and human iPSCs can be obtained using miRNA [14, 15]. iPSCs have been successfully derived from various types of somatic cells. iPSCs were first obtained from fibroblasts of different origins, and subsequently from keratinocytes, melanocytes, blood cells, neural stem cells, pancreatic β-cells, B lymphocytes, and other cells [16–22]. Thus, it can be concluded that iPSCs can be derived from cells originating from all three primary germ layers (ectoderm, mesoderm, and endoderm), although efficiency in and the dynamics of the derivation of stable iPSC lines considerably depends on the method used and type of somatic cells [14, 23]. iPSCs obtained as a result of direct reprogramming have a number of common properties, which makes them such promising models for studies in the field of the biology of pluripotent cells and enables to use them to simulate human diseases and in regenerative medicine [6, 7]. In terms of their properties, induced pluripotent stem cells are very close to embryonic stem cells, which are derived from mouse and human pre-implantation embryos. Both cell types possess a similar morphology, sensitivity to growth factors and signaling molecules, and patterns of gene expression and differentiation [24]. In particular, during *in vitro* differentiation, iPSCs can form embryoid bodies consisting of the derivatives of all three germ layers. Furthermore, human iPSCs can form teratomes, whereas mouse iPSCs give rise to chimeras and are even capable of forming an entire organism when injected into tetraploid blastocysts [25–27]. It is obvious that all these properties typical of pluripotent cells are determined by the special state of epigenome, which is "inherited" by ESCs from the inner cell mass cells or is formed during reprogramming in the case of iPSCs.

Recent studies have demonstrated that transcription factors, signaling pathways, and miRNA closely interact with the system of enzymes and other specific proteins that participate in the formation of the chromatin structure. The unique state of chromatin in pluripotent cells is formed by this interplay.

The features of the epigenomes of embryonic stem cells and induced pluripotent stem cells are considered in this review. Special attention is focused on the interaction of the transcription factors OCT4, SOX2 and NANOG with Polycomb group proteins and the other molecules that participate in the regulation of the chromatin structure. The participation of the transcription factors of the system of pluripotency maintenance during the process of X chromosome inactivation is discussed. Moreover, the epigenetic events occurring upon reprogramming of somatic cells to a pluripotent state and the problems associated with the "epigenetic memory" are considered.

BIVALENT CHROMATIN DOMAINS IN PLURIPOTENT CELLS

Chromatin regions simultaneously enriched in marks of active and inactive chromatin (H3K4me3 and H3K27me3) are known as bivalent domains [28]. These domains were found in mouse and human ESCs [28–31]. The genes whose transcription start sites are associated with the bivalent domains are characterized by a low transcription level regardless of the presence of the active chromatin mark, which attests to the fact that H3K27me3 prevails over H3K4me3. A high level of the histone H2A variant H2AZ was detected in bivalent domains [32]. Most bivalent domains are connected with the transcription start sites of the genes associated with development, e.g., transcription factors of the families HOX, SOX, FOX, PAX, IRX, and POU [28]. During the differentiation, most bivalent domains become monovalent and contain either H3K27me3 or H3K4me3 depending on the type of differentiated derivatives [28, 33]. However, some domains remain in their bivalent state and are present in the epigenomes of precursor cells [33, 34]. In general, the existence of bivalent domains and preservation of active chromatin marks in the promoter regions of the genes involved in maintaining an undifferentiated state allows one to quickly switch between programs of gene transcription upon differentiation to certain derivatives.

INTERPLAY OF TRANSCRIPTION FACTORS OF THE SYSTEM OF PLURIPOTENCY MAINTENANCE WITH POLYCOMB GROUP PROTEINS AND CHROMATIN REMODELING FACTORS

The existence of the so-called open chromatin in ESCs and simultaneous reliable repression of the differentiation genes are provided by the interaction system both at the protein – DNA and protein – protein levels. The investigation into the proteome of pluripotent cells and, in particular, the proteins forming the main system of pluripotency maintenance (OCT4, NANOG, SOX2) has shown that proteins not only interact with one another,

thus regulating the transcription of a number of genes, but that they also form a complex interaction network with the other transcription factors and proteins that participate in chromatin modification and remodeling. The proteins involved in pluripotency maintenance interact with the components of protein complexes, such as PRC1 and 2, BAF, NuRD, etc. [35–38].

POLYCOMB GROUP PROTEINS. COMPLEXES PRC1 AND 2

Polycomb group proteins are an evolutionary-conserved family of regulators of the chromatin structure. The role of these proteins is to achieve and maintain the transcriptional silencing of homeotic genes [39–41].

Two complexes belonging to the Polycomb family are known in mammals: PRC1 (Polycomb Repressive Complex 1) and PRC2, which play an essential role in embryonic development and in the maintenance of stem cell self-renewal and normal differentiation.

The mammalian PRC1complex consists of several subunits; homologues of these were found in Drosophila: CBX1, 2 and 3; MEL18, BMI1, RING1A (RING1), RING1B (RNF2) and PHC1, 2 and 3. It is considered that the role of PRC1 is to maintain the genes in their repressed state, which is originally achieved by the PRC2 complex. This function is activated through the activity of the subunits RING1A and 1B, which belong to the E3 ligase family and perform monoubiquitination of H2A histone at K119 (H2AK119Ub1) [42–44]. Mice having mutations in the PRC1 subunit genes (except for RING1B) remain alive, which may attest either to the existence of alternative mechanisms or to redundancy of the PRC1 function for the normal regulation of embryonic development [45]. However, it has been ascertained that the components of the PRC1 complex (e.g., BMI1) are required to ensure the functioning of several types of regional stem cells (hematopoietic, neural, lung and intestinal stem cells) [46–49]. It is of interest that the function of BMI1 and PRC1 in regional stem cells is presumed to be confined to control over the system regulating the level of reactive oxygen in mitochondria [50]. Furthermore, the absence of RING1A and 1B causes spontaneous differentiation of mouse ESCs and activates the genes associated with differentiation and development. Interestingly, the promoters of a large number of genes repressed by PRC1 are bound to the OCT4 transcription factor, which also participates in the transcriptional repression of these genes. Binding of PRC1 to the target genes depends on OCT4, whereas binding of OCT4 is PRC1-independent [51]. Proteomic studies have demonstrated that RING1B (RNF2) physically interacts with the NANOG transcription factor in ESCs [37]. These facts indicate that there is a close relationship between the system of transcription factors that maintain pluripotency and the system of regulators of the chromatin structure (in particular, PRC1).

A new function of CBX proteins, components of the PRC1 complex, in the regulation of the self-renewal and differentiation of mouse ESCs has been recently detected [52, 53] (*Fig. 1*). Five CBX proteins associated with PRC1 – CBX2, CBX4, CBX6, CBX7, and CBX8 are known in mammals [54]. The methods of ChIP-Seq (chromatin immunoprecipitation followed by sequencing of enriched DNA) and co-immunoprecipitation were used to demonstrate that in undifferentiated mouse ESCs, 97% of the CBX7 binding sites contain the complexes PRC1 and PRC2; 86% of them are also marked by H3K27me3. Several sites are located within the development-associated genes (e.g., sites in the HOX gene cluster [52]).

It has also been demonstrated using a quantitative proteomic analysis that only CBX7 co-localizes with H3K27me3 in undifferentiated mouse ESCs, whereas CBX2 and CBX8 interact with this histone modification in differentiated cells and fibroblasts [53]. Furthermore, it has been established via chromatin immunoprecipitation that CBX7 in a complex with PRC1 interacts with the *Cbx2, Cbx4* and *Cbx8* gene promoters, suppressing their transcription in ESCs [52]. Contrariwise, CBX2, CBX4 and CBX8, which can participate in *Cbx7* expression, bind to PRC1 during the differentiation process [52, 53]. Suppression of *Cbx7* expression in ESCs results in enhanced expression of the *Cbx2, Cbx4*, and *Cbx8* genes, disruption of the ESC morphology, and spontaneous differentiation. Ectopic increased *Cbx7* expression suppresses differentiation and X chromosome inactivation in female cells and enhances their self-renewal [53]. In addition, *miR-125* and *miR-181* participate in the suppression of *Cbx7* transcription, which supports the fact that miRNAs play a significant role in the regulation of the Polycomb protein function [53]. Thus, the dynamic system of PRC1 and CBX proteins, which are mutually regulated, participates in the regulation of the self-renewal and differentiation of ESCs. The function of these complexes is regulated by PRC2 (H3K27me3); their combinations change depending on cell status (*Fig. 1*).

The mammalian protein complex PRC2 contains the EED (embryonic ectoderm development), SUZ12 (suppressor of zeste 12), and EZH1 (enhancer of zeste 1) or EZH2 (enhancer of zeste 2) subunits. EZH2 is the protein with the SET domain, which is attributable to the proteins functioning as histone methyltransferases and performs di- and trimethylation of histone H3 at K27 (H3K27me2/3). As opposed to PRC1, gene mutations of PRC2 subunits cause significant disruptions in embryonic development and embryonic death [45, 56, 57]. Disrupted gastrulation (the pattern of the germi-

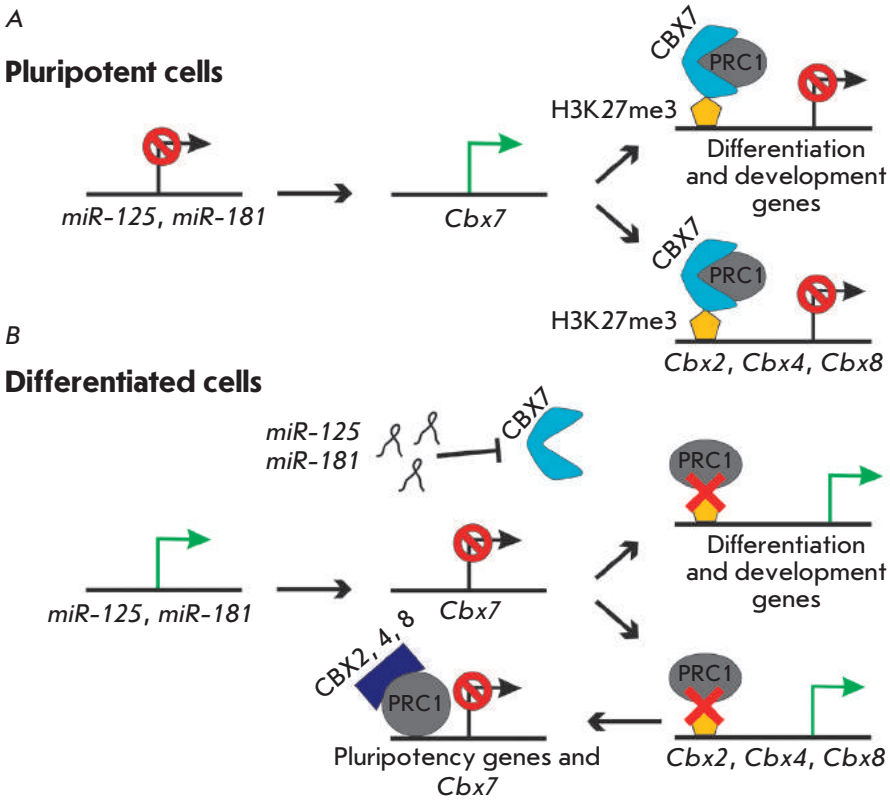


Fig. 1. Model illustrating the role of CBX proteins in PRC1 regulation in pluripotent cells and during differentiation. (A) In pluripotent cells, the CBX7/PRC1 complex binds to the regulatory regions of the genes involved in development and differentiation and the genes encoding the CBX2, 4, and 8 proteins, which represses their transcription. This binding depends on H3K27me3 established by PRC2. (B) Expression of microRNAs *miR-125* and *miR-181*, which repress CBX7 expression, is activated during differentiation. The absence of CBX7/PRC1 results in activation of the differentiation genes, as well as *Cbx2*, *4* and *8*. PRC1, together with CBX2, CBX4, and CBX8, represses the transcription of the genes responsible for pluripotency maintenance and *Cbx7* [55]

nal streak is changed in the anteroposterior direction), hypertrophied extraembryonic mesoderm, and undeveloped embryonic mesoderm are observed in embryos with the mutant *Eed* gene and the absence of H3K27 methylation [57, 58]. However, *Eed* mutant blastocysts can be used to obtain ESCs possessing pluripotency but tending towards spontaneous differentiation [59]. A similar situation has also been observed for *Suz12* mutants. Despite the fact that the death of *Suz 12* mutant mouse embryos is observed, ESCs can be successfully obtained. Although ESCs obtained from the mutant embryos possess a high level of transcription of the differentiation genes, they do not form neurons during *in vitro* differentiation and slightly differentiate to endoderm derivatives when embryoid bodies are formed [56]. Deletion of the *Ezh2* gene results in no changes in the properties of ESCs obtained from mutant embryos; this fact can be attributed to the effect of the EZH1 subunit, which also has histone–methyltransferase activity and mediates the setting of the mark of inactive chromatin within the PRC2 target genes [60].

It has been recently found that the JARID2 protein from the family JUMONJI C (JMJ C) is one of the subunits of the PRC2 complex. JUMONJI proteins belong to histone demethylases; however, JARID2 lacks such activity. It has been shown that JARID2 is required to provide efficient binding of PRC1 and PRC2 to the promoters of the target genes; the binding pattern of PRC2 and JARID2 to the DNA of the target genes in the scale of the genome of mouse ESCs coincides to over 90% [61–66]. There is abundant controversial experimental data on the effect of *Jarid2* knockout or knockdown on the H3K27me3 level in promoters of the PRC2 target genes. A decrease in the H3K27me3 level has been observed in some studies [63, 65, 66], whereas both the absence of changes [61] and an increase in the H3K27me3 level [62] have also been reported. However, the differentiation process has been shown to be disrupted or decelerated in JARID2-deficient ESCs; i.e., JARID2 affects pluripotency in some way [62, 63, 66]. In addition, JARID2, jointly with MTF2 and esPRC2p48 proteins, is capable of enhancing efficiency in obtaining induced pluripotent stem cells from mouse embryonic fibroblasts via overexpression of the *Oct4*, *Sox2*, and *Klf4* genes. On the contrary, knockout of the genes encoding JARID2, MTF2, and esPRC2p48 considerably suppresses reprogramming [67].

There are several hypotheses on the molecular basis of the effect of JARID2 on cell pluripotency; none of those has been sufficiently supported through experimental data [68]. The major role of JARID2 is believed to be not in the modulation of the histone-methyltransferase activity of PRC2, but in recruiting a special

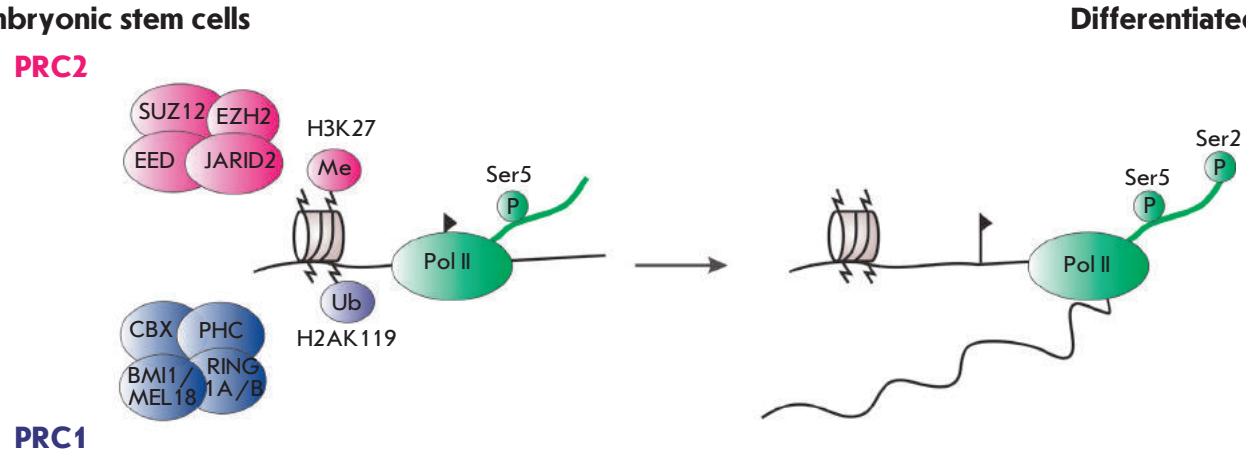


Fig. 2. JARID2 is necessary to recruit the Ser5-phosphorilated form of RNA-polymerase II in the bivalent domains of mouse ESC epigenome. In pluripotent cells, the formation of "bivalent domains" requires RNA-polymerase II phosphorylated at Ser5 (green oval) in the promoter regions of the genes whose expression is repressed by PRC2. PRC2 subunits mediate H3K27 methylation; in turn, it recruits PRC1 responsible for H2AK119 monoubiquitylation [68]

initiating form of RNA polymerase II [66, 68] (*Fig. 2*). This form of RNA polymerase has a phosphorylated serine residue at position 5 (Ser5P-RNAP) (while in the elongating form, the serine residue at position 2 is also phosphorylated); the presence of this form is typical of bivalent epigenome domains, which are formed with the participation of PRC1 and PRC2 [69, 70]. The presence of this polymerase form within promoters of the genes participating in cell differentiation appears to be required for rapid and reliable switching between transcription programs when the differentiation process is initiated.

Thus, it can be concluded that PRC2 plays an essential role in the regulation of mammalian development and ESC differentiation; however, this complex does not affect the process of obtaining ESCs and their self-renewal. Abundant experimental evidence of joint regulation of the target genes by PRC2 and transcription factors OCT4, SOX2, and NANOG, which are the central part of the system of gene transcription regulation in mouse and human ESCs, has been obtained. Whole-genome studies have demonstrated that OCT4, SOX2, NANOG, and the PRC2 subunits colocalize in the genes responsible for development, intracellular signal transfer, morphogenesis and organogenesis; hence, they can function jointly [8, 28, 71, 72].

TRITHORAX (TRXG) COMPLEX

The proteins of the Trithorax complex are among the major regulators of the embryonic development of both invertebrates and vertebrates [73]. The role of Trithorax during their development is opposite to that of Polycomb proteins; they mediate H3K4me3, which is mostly associated with transcription activation. Unlike the PRC1 and PRC2 complexes, the role of Trithorax proteins in maintaining cell pluripotency remains poorly studied [74]. Identically to that in invertebrates, Trithorax in mammals is a multi-subunit complex containing histone methyltransferases SET/MLL. The main subunits of the complex, WDR5, ASH2L and RBBP5, are needed to activate the SET/MLL enzymes [75]. The ASH2L/RBBP5 heterodimer is known to directly contribute to the histone methyltransferase activity of the MLL1 complex [76]. Furthermore, experimental evidence of the fact that ASH2L is needed for normal embryogenesis and X chromosome inactivation in female mice has been obtained [77]. The WDR5 subunit is also the major component of the mammalian Trithorax complex. Its function consists in "providing" H3K4 residues and in carrying out an efficient interaction between the entire Trithorax complex and H3K4, and thus in implementation of its histone-methyltransferase activity [75]. Furthermore, WDR5 is known to recognize H3K4me2 and mediate the transition of H3K4 into a trimethylated state (H3K4me3) [78]. It has recently been established that WDR5 is required not only to ensure the normal development of vertebrates, but also plays a crucial role in maintaining ESC pluripotency and cell reprogramming to a pluripotent state [74]. The inhibition of WDR5 expression was ascertained to abruptly reduce the self-renewal of mouse ESCs. Proteomic studies have allowed to establish that WDR5 physically interacts with OCT4 in undifferentiated ESCs, so that the targets of these two proteins overlap

to a significant extent. Thus, it has been demonstrated that the Trithorax complex, along with OCT4, SOX2 and NANOG, positively regulates gene transcription in mouse ESCs. Furthermore, it has been shown in experiments on the reprogramming of somatic cells that the Trithorax complex (WDR5) is required to provide efficient formation of iPSC clones [74].

BAF COMPLEX

Numerous studies have shown that ATP-dependent chromatin remodeling protein complexes play an essential role in the embryonic development of mammals in general and in maintaining cell pluripotency in particular [79–84]. Nearly 30 proteins possessing ATP-dependent chromatin-remodeling activity have been known in mammals. These proteins have been grouped into several families in accordance with the structure of the ATPase domain. In mammalian cells, chromatin-remodeling ATPases interact with each other and the other proteins and act within protein complexes consisting of several subunits. BAF, NuRD, ISWI can be mentioned as examples of such complexes. The BAF protein complex participates in the redistribution of nucleosomes and is present in all cell types. However, the subunit composition of this complex may vary for different cell types, and chromatin structure is controlled in a fashion specific to each cell type. ESCs contain the BAF complex (known as esBAF), which in turn consists of a specific combination of the subunits BRG, BAF155, and BAF60a but contains no subunits BRM, BAF170, or BAF60c [83, 84]. It has been experimentally demonstrated that the inactivation of most subunits of the BAF complex causes the death of mouse embryos at the early stages of development and also leads to cellular oncotransformation [79, 85–88]. Furthermore, embryonic death in case of loss of the subunits BRG, BAF47, and BAF155 has been caused by the disturbance of the formation of pluripotent cells. Careful screening of libraries of interfering RNAs has demonstrated that such subunits as BRG and BAF155 are required to maintain the morphology of ESC colonies and expression of *Nanog* [89, 90]. According to proteomic data, several subunits of chromatin-remodeling complexes physically interact with OCT4 and NANOG in ESCs [35–38, 83, 84].

The transcription factors OCT4 and NANOG can interact with chromatin remodeling complexes via specific proteins [37, 90]. Thus, it has been demonstrated that the chromosome scaffold protein TIF1b (Transcription Intermediary Factor-1b) is required for maintaining the activity of the *GFP* transgene in ESCs under the control of the *Oct4* promoter [92]. Interestingly, TIF1b used to be known as a protein that participates in transcriptional silencing and heterochromatin formation through recruitment of the heterochromatin protein HP1 and histone methyltransferase SETDB1 and NuRD. However, the phosphorylated form of TIF1b can interact with the ESC-specific form of the BAF complex, localized in euchromatin, and is capable of affecting the efficiency of the induced pluripotent stem cell derivation [91]. Furthermore, overexpression of ESC-specific components of this complex, BRG1 and BAF155, increases the efficiency of the reprogramming of somatic cells in the absence of *c-Myc* overexpression [93, 94].

It has recently been demonstrated that the esBAF complex is directly associated with the activity of the LIF-STAT3 signaling pathway, which is necessary for maintaining pluripotency of mouse ESCs [95, 96]. The transcription factor STAT3 is known to activate gene groups in various cell types containing specific BAF complexes; however, it is only in ESCs that it contributes to the regulation of the target genes required to maintain an undifferentiated status of ESCs. However, the mechanism of such a specific effect of STAT3 remained unclear for a long time.

L. Ho *et al*. [96] have ascertained that binding of STAT3 to the target sites in the genome of mouse ESCs depends on BRG1, the ATPase subunit of the ESC-specific esBAF complex. The effect of BRG1 within STAT3-binding sites forms the chromatin structure, which is required for gene activation by interleukin LIF. BRG1 deletion induces PRC2-mediated transcriptional silencing of a number of genes at the level of the entire genome via H3K27me3. STAT3 targeted genes undergo transcriptional silencing as well. Based on these facts, a conclusion has been drawn that the major role of BRG1 in mouse ESCs is enhancing the action of the LIF-STAT3 signaling pathway and counteracting the repression of this pathway by the Polycomb proteins (PRC2). It is an interesting fact that BRG1 can act jointly with Polycomb to enhance repression of the differentiation genes (e.g., HOX family genes). Thus, the esBAF complex can act both antagonistically and synergically with PRC2; however, both types of actions work towards the maintenance of pluripotency [96] (*Fig. 3*).

NuRD COMPLEX

The mammalian protein complex NuRD (Nucleosome Remodeling Deacetylase), which exhibits ATP-dependent remodeling and histone deacetylase activity, consists of at least six subunits [97, 98]. NuRD contains histone acetylases HDAC1 and HDAC2, whose activity is dependent on the chromodomain-containing ATPase subunits Mi2a and Mi2b. In addition, the complex contains proteins binding methylated cytosine MBD 1, 2 and 3 (methyl-CpG-binding proteins), proteins MTA1,

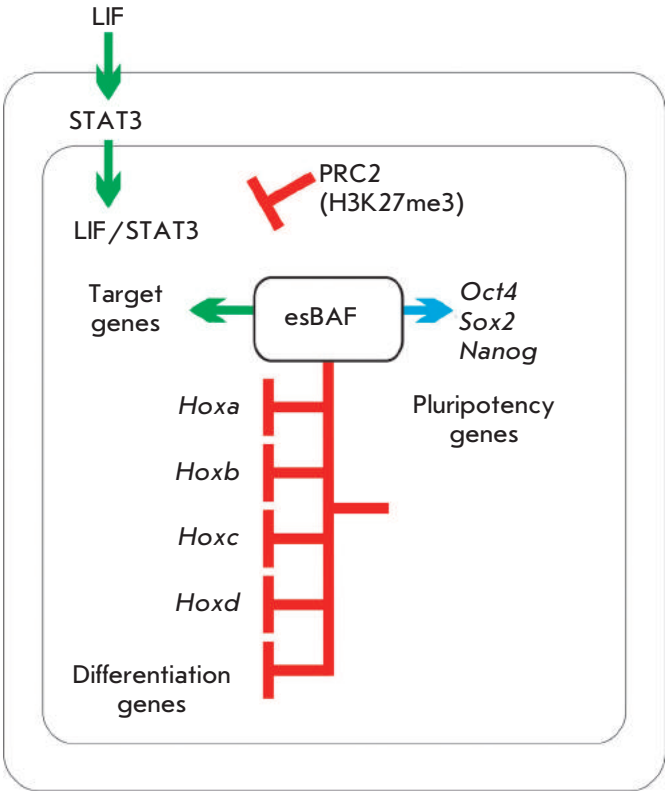


Fig. 3. Cooperation of esBAF and PRC2 during pluripotency maintenance. esBAF and PRC2 can act both synergically and antagonistically. esBAF antagonizes PRC2 when regulating the target genes of the LIF-STAT3 signaling pathway, thus preparing the chromatin structure for activation by the phosphorylated STAT3 form (green arrow). Meanwhile, esBAF acts together with PRC2 to repress transcription of HOX genes (red lines with stoppers). However, expression of pluripotency genes can be activated or repressed by esBAF (blue arrow) [96]

2 and 3 (Metastasis-associated proteins), WD40-containing proteins RbAP46 and RbAP48, as well as two proteins containing zinc finger domains (p66a and p66b). Several subunits of the NuRD complex have been demonstrated to be required to maintain the pluripotency and differentiation of ESCs. Embryonic stem cells with a deletion of the gene encoding MBD3 retain their viability and expression of pluripotency markers; however, they cannot differentiate both *in vitro* and *in vivo* upon formation of chimeric animals [99]. Nevertheless, it has been demonstrated in one of the later studies that *Mbd3* knockout in mouse ESCs enhances the transcription level of such trophoblast markers as *Cdx2*, *Eomesodermin*, and *Hand1* and the acetylation level of histone H3 in the promoter regions of these genes. Furthermore, knockout cells grown in media for trophoblast stem cells have differentiated into trophoblast cells expressing CDX2 and CADHERIN 3 [100]. It has been shown in *in vivo* experiments that MBD3 is required for the development of epiblast from ICM cells after implantation. In MBD3-deficient embryos, the pluripotency genes *Oct4*, *Nanog* and *Sox2*, as well as their target genes, are expressed at the normal level; however, their normal transcriptional silencing can be disturbed after the implantation. On the contrary, the cultured ICM of MDB3-deficient embryos cannot give rise to pluripotent ESC lines, although they form a significant number of endodermal derivatives [101].

Mouse ESCs contain a specific subfamily of NuRD complexes known as NODE (NANOG and OCT4 associated deacetylase). The NODE complex consists of histone deacetylases HDAC1 and HDAC2, as well as MTA1 and 2. However, this complex contains almost no MBD3 or RBBP7 subunits (they are detected in substoichiometric amounts). NODE is of interest because it physically interacts with the OCT4 and NANOG transcription factors in mouse ESCs [35]. NODE exhibits deacetylase activity, which is MBD3-independent. A knockout of the genes encoding the NODE subunits results in an enhancement of the expression of the genes responsible for differentiation; hence, it causes differentiation of ESCs into various cell derivatives. It has been demonstrated in experiments aimed at translation inhibition that, unlike MBD3 which is required to repress transcription of the genes maintaining the undifferentiated state, MTA1 participates in the inhibition of the differentiation genes, such as *Gata6* and *FoxA2* [35]. Thus, ESCs contain at least two subfamilies of NuRD complexes that act in opposite directions: 1) MBD3-containing complexes that regulate (inhibit) the transcription of the pluripotency genes (*Oct4*, *Nanog* and *Sox2*, etc.) and are required for ESC differentiation into various cellular derivatives and for cell differentiation during early embryonic development; 2) HDAC1-, HDAC2-, and MTA1-containing complexes interacting with OCT4 and NANOG and participating in the transcription activation of the genes responsible for the maintenance of an undifferentiated state.

It has recently been established that the MBD3-containing NuRD complex is required to modify H3K27me3 by the PRC2 complex within the promoters of the genes participating in development and differentiation processes. Thus, NuRD does not simply repress gene transcription; it is also responsible for the equilibrium between H3K27 acetylation and methylation in embryonic stem cells [102]. However, that is not the only example of interaction between chromatin-remodeling complexes in ESCs. The NuRD complex, namely,

its MBD3 subunit, closely interacts with esBAF (BRG1) in mouse ESCs [103]. The MBD3 and BRG1 subunits colocalize within the transcription start sites and multidirectionally regulate the transcription of a vast gene set. Moreover, MBD3 and BRG1 play a significant role in transcription regulation through hydroxymethylation of cytosine residues. The MBD3 subunit colocalizes with the TET1 protein and 5-hydroxymethylcytosines (5hmC) *in vivo*; the binding of MBD3 to promoters is TET1-dependent. *In vitro* experiments have shown that MBD3 binds to 5-hydroxymethylcytosines more efficiently as compared to 5-methylcytosines; knockout of the *Mbd3* gene mostly affects the transcription of 5hmC-marked genes, whereas MBD3 and BRG1 are required to maintain the 5-hydroxymethylation level [103].

5-Hydroxymethylation of DNA was believed to be just an intermediate stage during the demethylation of 5-methylcytosines [104]. However, it turns out that knockout of the genes belonging to the *Tet* family, which encode the proteins performing hydroxylation of 5-methylcytosines, disrupts the differentiation (*Tet1, Tet2*) and self-renewal (*Tet1*) of ESCs [105–107]. In addition, DNA 5-hydroxymethylation can be retained for a long time during early embryonic development and seems to perform regulatory functions [108, 109]. All these facts indicate that 5-hydroxymethylation can be an independent regulatory state of the epigenome and that NuRD and esBAF play a crucial role in its regulatory potential. Meanwhile, DNA 5-hydroxymethylation directly affects the joint regulatory action of NuRD and esBAF.

Tip60-p400 COMPLEX

The Tip60-p400 complex exhibits histone acetyltransferase and remodeling activity; it can act both as an activator and a repressor of transcription [110, 111]. In addition, Tip60-p400 participates in the replacement of forms of H2AZ-H2B histones [112, 113]. The embryos with a knockout of the *Tip60* and *Trrap* genes that encode the Tip60-p400 subunits die at the preimplantation stage [114, 115]. The inhibition of the translation of several Tip60-p400 subunits in ESCs via RNA interference has demonstrated that Tip60-p400 is important for the normal self-renewal and differentiation of cells. It has been demonstrated using chromatin immunoprecipitation that p400 colocalizes with NANOG and H3K4me3 (the active chromatin mark) in undifferentiated mouse ESCs. The spectra of the NANOG and Tip60-p400 target genes overlap to a significant extent. Furthermore, NANOG and H3K4me3 are required to provide binding of Tip60-p400 to the target genes. In turn, Tip60-p400 acetylates histone H4 [89].

DIRECT REGULATION OF THE GENES ENCODING THE PROTEINS MODULATING THE CHROMATIN STRUCTURE USING THE TRANSCRIPTION FACTORS PARTICIPATING IN THE MAJOR SYSTEMS OF PLURIPOTENCY MAINTENANCE

The transcription factors that make up the system of pluripotency maintenance, in addition to their interaction with protein complexes, can directly regulate the genes of chromatin-modifying enzymes. In ESCs, OCT4 activates the demethylase genes JMJD1A/KDM2A and JMJD2C/KDM4B, which demethylate H3K9me2 and H3K9me3, respectively, whereas KDM2A and KDM4B, in turn, perform the demethylation of the promoter region of *Tcl1* and *Nanog*, respectively [116].

The transcription factors regulating pluripotency interact with the promoters of the genes whose products participate in the global regulation of the chromatin structure. Thus, such factors as OCT4, SOX2, NANOG, SMAD1, ZFX, and E2F1 are associated with the *Chd1* gene promoter [117]. This gene encodes the enzyme participating in chromatin remodeling. CHD1 binds to histone H3 di- or trimethylated at K4, which is a mark of active chromatin and the genes being transcribed, via two chromodomains [118]. The *Chd1* repression in mouse ESCs has no effect on the self-renewal of ESCs; however, it tilts the cells towards neural differentiation [119].

Factor UTF1 (undifferentiated embryonic cell transcription factor 1), which is transcribed at a high level in undifferentiated mouse ESCs, can participate in the formation of the global chromatin structure. This protein is bound to chromatin; it colocalizes in the regulatory regions of over 1,700 genes, most of which overlap with the previously identified target genes of the transcription factors NANOG, OCT4, KLF4, C-MYC, and REX1. Reduced synthesis of UTF1 increases the level of expression of most of its target genes and disrupts ESC differentiation. This fact indicates that UTF1 mainly represses the transcription of the genes involved in cell differentiation [120]. It has been demonstrated that the enhancer element localized in the 3' untranslated region of *Utf1* binds selectively to OCT4 and SOX2 [121].

Thus, regulators of the chromatin structure (CHD1 and UTF1), whose gene expression is directly regulated by the transcription factors that are components of the main internal system of pluripotency maintenance, have been found in ESCs.

PLURIPOTENCY AND DNA METHYLATION

In addition to covalent modifications of histones, DNA methylation is the major mechanism that regulates cellular processes in mammals [122]. Today DNA methylation is known to participate in fundamental phenomena and processes, such as embryogenesis, cell differentiation, genomic imprinting, cancerogenesis, regulation

of the transcription of mobile genetic elements, and X chromosome inactivation in female mammals [123–128].

DNA methylation indisputably plays a crucial role in the regulation of the self-renewal and pluripotency of cells [129]. Promoters of the major genes associated with the pluripotency maintenance and self-renewal of ESCs (*Oct4* and *Nanog*) are hypomethylated in undifferentiated cells and hypermethylated in stem and somatic trophoblast cells [130, 131]. During cell differentiation in a culture or in the embryonic development, promoters of the genes maintaining self-renewal undergo methylation with the participation of the DNA methyltransferases DNMT1, DNMT3A, and DNMT3B [132]. Knockout of the genes encoding the DNA methyltransferases DNMT1, DNMT3A, and DNMT3B causes a disruption of embryonic development and ESC differentiation *in vitro* [132–135]. However, mouse ESCs with simultaneously knocked out genes *Dnmt1*, *Dnmt3a*, and *Dnmt3b* retain their self-renewal ability [136]. DNA methylation performed by DNMT3A and DNMT3B participates in reliable repression of pluripotency genes in embryonic development. Histone methyltransferase G9a, which establishes H3K9me3 within the *Oct4* promoter, recruits the heterochromatin protein HP1 and DNA methyltransferases into this region [137].

Cytosine residues in CpG dinucleotides undergo methylation in mammalian genomes [138]. Pluripotent cells are characterized by a reduced methylation level of CpG-rich promoters (containing the so-called GpG islands) and an increased methylation level of CpG-deficient promoters [129, 139]. Most of the CpG-deficient promoters contain H3K4me3, the active chromatin mark. H3K4me3 appears to be established as a result of binding of nonmethylated CpG islands to CPF1 associated with histone methyltransferase SETD1 [140]. In turn, H3K4 methylation can "protect" gene promoters against the impact of DNA methyltransferases [141].

It has recently been demonstrated that a significant fraction (up to 25% in human ESCs) of methylated cytosine residues in ESC and iPSC genomes localizes outside CpG [142–144]; non-CpG methylation is predominantly observed in exons rather than in the regulatory gene regions [143, 144]. The pattern of non-CpG methylation in different pluripotent cell lines is very diverse, whereas non-CpG methylation is almost absent in some differentiated cells. Furthermore, knockout of the *DNMT3A* and *DNMT3B* genes in human ESCs drastically reduces the non-CpG methylation level [145].

Numerous experimental data indicate that the reprogramming of somatic cells to the pluripotent state (obtaining iPSCs) is accompanied by a global change in methylome towards the state characteristic of pluripotent cells [144, 146, 147]. Promoters of the genes participating in self-renewal maintenance (e.g., *Oct4* and *Nanog*) undergo demethylation [11, 12, 148]. Such DNA demethylases as TET1 and AID can participate in the reprogramming. Demethylase TET1, which catalyzes the conversion of 5-methylcytosine into 5-hydroxymethylcytosine, is essential for the maintenance of self-renewal of mouse ESCs; it regulates DNA methylation in the *Nanog* promoter [106]. Furthermore, it has been demonstrated using a reprogramming model with mouse embryonic stem/human fibroblast hybrid cells that demethylase AID is required for demethylation of promoters of the human genes *OCT4* and *NANOG* [149]. The fact that the use of inhibitors of DNA methyltransferases allows one to enhance the efficiency of iPSC derivation also lends support to the idea of the significance of methylation for cell reprogramming [146, 150].

PLURIPOTENCY FACTORS IN THE REGULATION OF X CHROMOSOME INACTIVATION

X chromosome inactivation is a complex process occurring during early mammalian embryogenesis. In mice, imprinted inactivation of the X chromosome inherited from the male parent takes place during the first series of zygote divisions. At the blastocyst stage, the X chromosome is reactivated in ICM cells. Random inactivation of one of the two X chromosomes occurs during gastrulation and differentiation of ICM cells [151–153]. X-inactivation is regulated by a certain locus at the X chromosome, which is known as the X-inactivation center [154]. This locus comprises several genes; however, the *Xist* and *Tsix* genes, which are anti-parallel-transcribed and encode nuclear untranslated RNAs, are considered to be the major regulators [155, 156]. *Xist* RNA was shown to be transcribed monoallelically from the inactive X chromosome, to coat it, and to induce modifications corresponding to inactive chromatin [155]. On the contrary, the *Tsix* gene is a negative regulator of the *Xist* gene; it is transcribed from the active X chromosome [157]. Since X-inactivation takes place during early embryogenesis, an investigation into its dynamics and molecular basis is rather complicated, almost infeasible when humans are used as the objects. Hence, pluripotent cell lines obtained from pre-implantation embryos (ESCs) or by reprogramming mouse or human somatic cells (iPSCs) are currently the most suitable and commonly used models to study X-inactivation. However, studies of the X chromosome status and molecular genetic studies of the regulation of the X-inactivation have revealed a number of differences between mice and humans.

Embryonic stem cells of female mice derived from pre-implantation blastocysts (3.5 days post coitum) retain a number of the properties of ICM cells; in particu-

lar, they can maintain two active X chromosomes in a series of mitotic divisions [152]. Random inactivation of one of the two X chromosomes takes place during the differentiation of mouse ESCs. This property of mouse ESCs is reproducible and stable [152, 153].

The situation is more complex for human ESCs, which are also derived from blastocysts (5–9 days post coitum) [4]. A large-scale analysis of a number of human ESC lines has shown that they can be divided into three classes [158]. The first class comprises ESCs with two active X chromosomes, which undergo random inactivation during differentiation; this class corresponds to mouse ESCs. The second class comprises ESC lines in which one of the chromosomes is inactive and the *XIST* gene is transcribed; however, cells retain all of their pluripotency features. The third class contains lines with one X chromosome being inactive; however, the *XIST* gene is not transcribed even after cell differentiation. The inactive X chromosomes in the lines of the second class carry inactive chromatin marks, such as H3K27me3, H4K20me, and the histone variant macroH2A. Interestingly, the lines that belong to the third class carry almost no inactive chromatin marks. Meanwhile, a molecular genetic analysis shows that transcription of most of the genes of the inactive X chromosomes is repressed [158].

The fact that pluripotency is not associated with the epigenetic status of X chromosomes in human pluripotent stem cells has also been demonstrated for iPSCs. Mouse iPSCs, similar to ESCs, have two active X chromosomes (in cells derived from females); one of those undergoes random inactivation after the differentiation is induced [159]. However, human iPSCs can have all the features of pluripotent cells and contain an inactive X chromosome; i.e., they can fall into the second class of ESCs [160]. The status of the X chromosome can be changed during reprogramming, resulting in the emergence of subclones corresponding to the first and third classes of ESCs. It has been mentioned that reactivation of the inactive X chromosome can occur during the reprogramming of human somatic cells [147]. In all likelihood, the isolation of clones of ESCs and iPSCs carrying two active X chromosomes can be achieved by varying cell culture conditions. Thus, it has recently been demonstrated that cell culturing under conditions of physiological oxygen concentration (5%) can considerably enhance efficiency in obtaining human ESCs of the first class. On the contrary, transition of the cells to the second and third classes, according to the status of the X chromosome, can be caused by various physiological stress factors [161]. Furthermore, overexpression of *KLF4* in the presence of a combination of inhibitors of signaling pathways in human ESCs and iPSCs can also cause reactivation of the inactive X chromosome [162]. This fact attests to the instability of the status of the X chromosome in human pluripotent cells.

Despite the fact that the association between the pluripotency of mouse cells and X chromosome status during embryogenesis and in culture is rather obvious, no direct evidence of association between these phenomena at the molecular level had existed until recently. However, the association between transcription factors and regulation of the *Xist* and *Tsix* genes has been revealed. Thus, the transcription factors NANOG, OCT4, and SOX2 have potential binding sites in the first intron of the *Xist* gene and are bound to it in undifferentiated mouse ESCs [163] (*Fig. 4*). Knockout of *Oct4* and *Nanog* induces activation of *Xist* transcription. Thus, pluripotency factors can inhibit *Xist* expression via the *Tsix*-independent mechanism [163]. It was established later that the factors NANOG, OCT4, and SOX2 can inhibit *Xist* transcription by repressing the expression of its activator, *Rnf12*. However, the removal of the first intron of *Xist* does not result in X-inactivation [164, 165] (*Fig. 4*).

The factors associated with maintenance of the pluripotency and repression of *Xist* can participate in the activation of *Tsix* transcription [167] (*Fig. 4*). Thus, binding of OCT4, SOX2, and KLF4 has been detected within the *Xite* enhancer; although this interaction has not been confirmed in other works [168]. The REX1, c-MYC, and KLF4 binding sites have been detected in the *DXPas34* regulatory element. It has been established that REX1 is required mostly for the elongation of *Tsix* RNA, rather than for the assembly of the transcription complex. Thus, all the aforementioned studies support the fact that the system of pluripotency maintenance is associated with an active status of both X chromosomes in undifferentiated mouse ESCs. Human ESCs do not exhibit these regularities. Human *XIST* transgenes remain active in mouse ESCs despite the presence of pluripotency-maintaining factors. Different mechanisms (e.g., DNA methylation) seem to participate in the regulation of human *XIST* and *TSIX* genes. In mouse ESCs, the *Xist* promoter is only partially methylated even at the active X chromosome; thus, gene transcription is presumably repressed by transcription factors. In human ESCs of the first type, the *XIST* promoter is almost completely methylated (100%). In addition, the differences can be attributed to the fact that the properties of human ESCs (gene expression pattern, sensitivity to signaling molecules) are similar to those of mouse epiblast stem cells, where one X chromosome is inactivated, despite the expression of pluripotency factors [169].

In all likelihood, investigations into the status of the X chromosome in human iPSCs should be used in standard tests carried out for newly obtained lines, to-

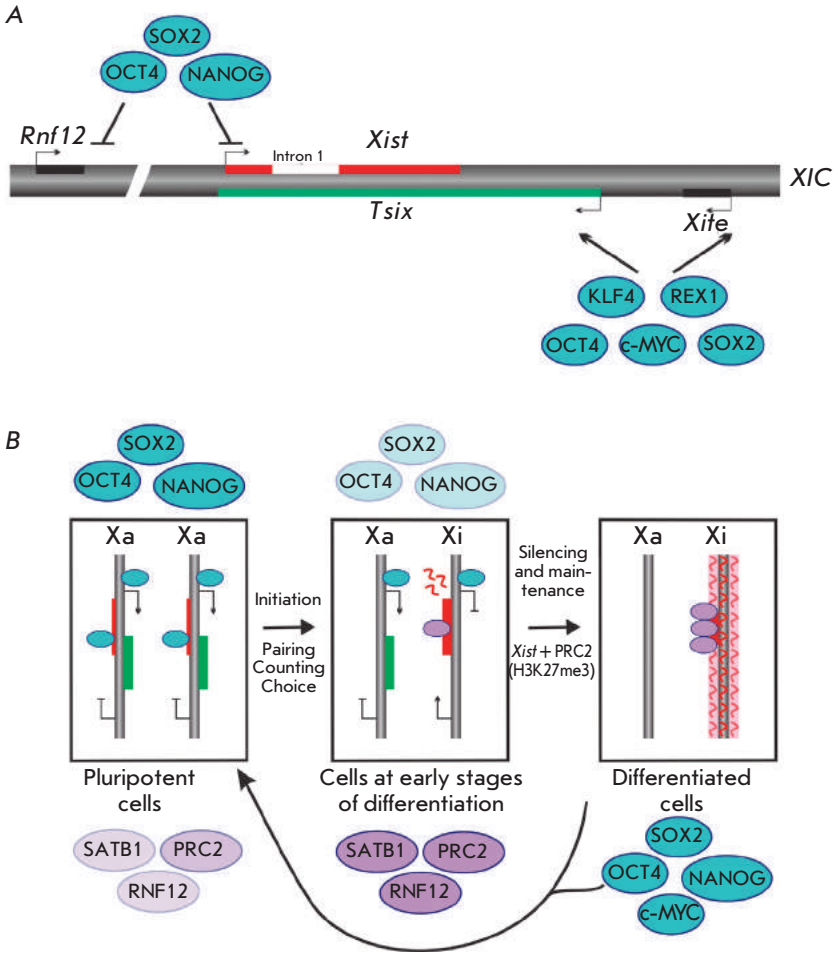


Fig. 4. Transcription factors of pluripotency in the regulation of X chromosome inactivation. (A) Scheme of the mouse X chromosome inactivation center (XIC). *Xist*, *Tsix*, and their activators – *Rnf12* and *Xite* – are shown in red, green, and black, respectively. In undifferentiated female mouse ESCs, the transcription factors OCT4, SOX2, and NANOG bind to the first intron of *Xist* and *Rnf12*, repressing their transcription. Meanwhile, OCT4, SOX2, KLF4, REX1, and c-MYC bind to the regulatory regions of *Tsix* and *Xite*, activating their transcription. (B) In female mouse ESCs, *Tsix* is activated and *Xist* is repressed by the proteins involved in pluripotency maintenance. During the differentiation, one of the X chromosomes is inactivated. X-inactivation is a multistage process including the initiation of inactivation, establishment, and maintenance of transcriptional silencing. Initiation of inactivation occurs due to the decrease in pluripotency factor expression and involvement of chromatin structure regulators (such as SATB1 and PRC2) in the process. Overexpression of OCT4, SOX2, NANOG, and c-MYC in somatic cells induces reprogramming to the pluripotent state, which is accompanied by reactivation of the inactive X chromosome [166]

gether with an analysis of the expression of the pluripotency markers that determine the patterns of gene transcription and differentiation. By choosing clones of cells with an inactivated paternal or maternal X chromosome, one can selectively obtain lines with inactive mutant alleles and, hence, cells that can be used to treat X-linked diseases.

EPIGENETIC EVENTS OCCURRING DURING CELL REPROGRAMMING TO A PLURIPOTENT STATE. "EPIGENETIC MEMORY"

Reprogramming of somatic cells to the pluripotent state is accompanied by a global change in their epigenomes [146, 159, 170]. A number of chemical inhibitors of the enzymes participating in the formation of the chromatin structure are currently used to enhance efficiency in generating human and mouse iPSCs. In particular, the use of the histone methyltransferase G9a inhibitor (BIX-01294) and inhibitors of DNA methyltransferases (5'-azacytidine, RG108) and histone deacetylases (valproic acid, TSA, SAHA, sodium butyrate) allows one to increase the reprogramming efficiency tens of times [18, 20, 150, 171–173]. Furthermore, the mechanism of the effect of ascorbic acid (vitamin C) on efficiency in iPSC isolation was recently elucidated [174].

Ascorbic acid is known to considerably enhance (from 3.8 to 8.75%) the efficiency of reprogramming of fibroblasts and stem cells from adipose tissue; however, its mechanism of action remained unclear [175]. Histone demethylases JHDM1A and 1B turn out to be the major effectors of ascorbic acid. Ascorbic acid induces JHDM1A/1B-mediated demethylation of histone H3 at K36 (H3K36me2/3) in a culture of embryonic mouse fibroblasts and during the reprogramming process (*Fig. 5*). It has been proven that JHDM1A/1B are needed for the reprogramming and participate in the acceleration of the cell cycle and inhibition of cell aging via repression of the *Ink4/Arf* locus (*Fig. 5*). A high cell division rate and inhibition of the mechanisms of aging and apoptosis are required to provide complete and efficient reprogramming of somatic cells [176–180]. Furthermore, JHDM1A/1Bs, together with OCT4, activate the

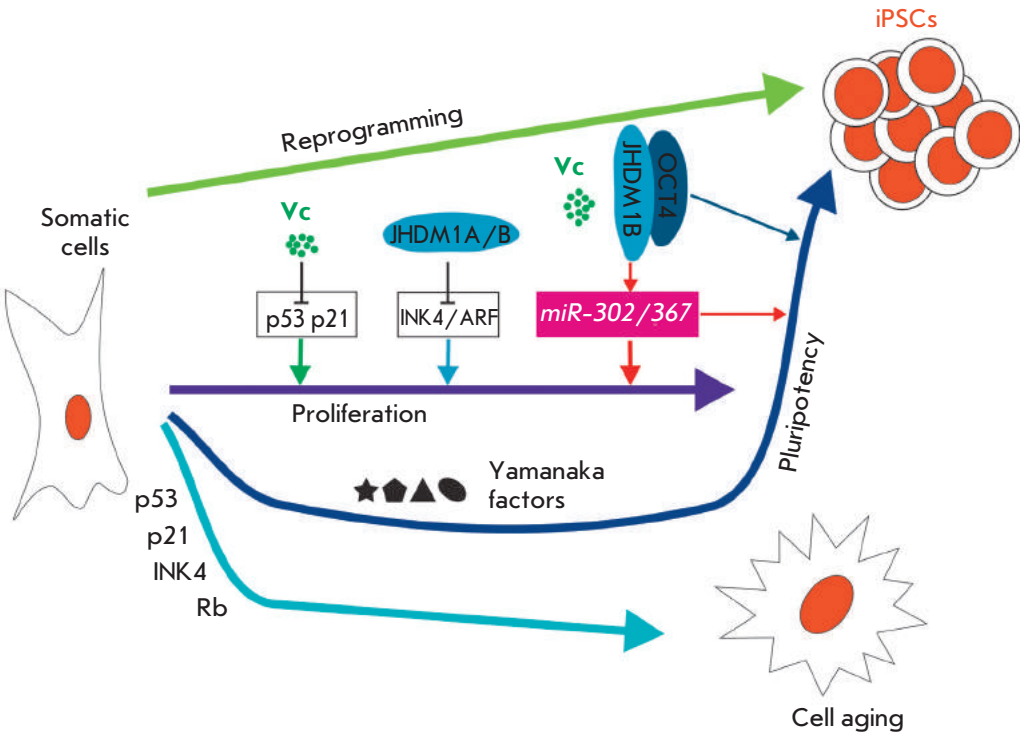


Fig. 5. The joint effect of ascorbic acid (vitamin C) and Jhdm1a/1b on the reprogramming of somatic cells to the pluripotent state. Vitamin C and Jhdm1a/1b counteract cell aging by repression of the p53/p21 and Ink4/Arf proteins. In addition, ascorbic acid and the Jhdm1b/Oct4 complex activate expression of the microRNA 302/367 cluster, thus increasing reprogramming efficiency [174]

expression of the miRNA 302/367 cluster, which is also involved in cell reprogramming [14, 15, 174] (*Fig. 5*).

T. Onder *et al.* [181] have screened a set of interfering RNAs inhibiting the translation of 22 genes whose products participate in DNA and histone methylation. The inhibition of the translation of mRNA of the genes encoding the components of the complexes PRC1 (BMI1, RING1) and PRC2 (EZH2, EED, SUZ12) considerably reduces the efficiency of reprogramming in human fibroblasts. Reduced efficiency was also observed during the inhibition of *EHMT1* and *SETDB1* encoding H3K9 histone methyltransferases. *YY1*, *SUV39H1*, and *DOT1L* were among the genes in which inhibition of mRNA translation considerably enhanced the reprogramming efficiency. The *YY1* gene encodes a protein acting both as a transcription activator and repressor, depending on the specific context. *SUV39H1* encodes H3K9 histone methyltransferase; *DOT1L* encodes H3K79 methyltransferase. More attention has been given to *DOT1L*. It turns out that repression of *DOT1L* via RNA interference or chemical inhibition of DOT1L can substitute the functions of *KLF4* and *c-MYC* in experiments for generating iPSCs from human fibroblasts. In addition, inhibition of DOT1L at the early stages of reprogramming results in the activation of *NANOG* and *LIN28*, which are also used in the case of human cells. A genome-wide analysis the H3K79me2 distribution has demonstrated that the genes associated with epithelial-mesenchymal transition, whose expression is specific to fibroblasts, lose this histone modification at the early stages of reprogramming. DOT1L inhibition accelerates deletion of H3K79me2 within the genes subjected to transcriptional silencing in iPSCs [181].

All these facts attest to a crucial role played by the system of epigenetic regulators in the reprogramming process.

High-performance analysis methods were used to reveal a high degree of similarity between iPSCs and ESCs in terms of the gene expression pattern and epigenomic state both at the level of DNA methylation and distribution the covalent histone modifications H3K27me3 and H3K4me3 [147, 182].

Despite significant similarity between iPSCs and ESCs at the molecular level, it has been demonstrated in a series of studies that the transcriptomes and epigenomes of individual iPSC lines can possess certain common features and retain a number of characteristics that are intrinsic to the original somatic cells [183–186]. The phenomenon of retaining certain features of the epigenomes of somatic precursors is known as epigenetic memory [187, 188].

The modern methods of molecular genetic analysis allow one to carry out high-resolution genome-wide studies of DNA methylation and distribution of cova-

lent histone modification. The study by R. Lister *et al.* [144], which employed the Methyl-C-Seq technique, can be given as an example. This technique allows one to carry out genome-wide studies of cytosine methylation at single-nucleotide resolution. The authors have tried to avoid the possible effect of the method of obtaining iPSCs and types of somatic cells on the results obtained. Five iPSC lines were used in this study: one line was obtained via retroviral transduction of adipose tissue cells with *OCT4*, *SOX2*, *KLF4*, and *c-MYC*; the second line was obtained via lentiviral transduction of lung fibroblasts IMR90 with *OCT4*, *SOX2*, *NANOG*, and *LIN28*; and three lines were obtained from foreskin fibroblasts using non-integrating episomal vectors. Furthermore, two ESC lines and trophoblast derivatives of iPSCs and ESCs differentiated using BMP4 were included in the study. The methylation status of 75.7–94.5% of all cytosine residues in the genomes of 11 cell lines has been determined. It is of interest to note that the authors have focused on not only methylation of cytosines within CpG dinucleotides but on non-CpG-methylation as well (CpH, where H = A, C or T). It has been shown that at the genomic scale, human iPSCs and ESCs have similar methylation patterns. The genomes of pluripotent cells tend to be more methylated (on average) than those of somatic cells. Serious differences at the level of CpH-methylation of DNA have been revealed. Somatic cells, including adipose tissue stem cells, are characterized by an extremely low level of such type of methylation; whereas the share of methylated cytosines within CpH dinucleotides in DNA in iPSCs and ESCs is 20–30% of the total amount of methylated cytosine residues in the genome. Moreover, enrichment of exons and introns in methylated CpH is observed both in ESCs and iPSCs.

It is interesting to mention that despite the general similarities between the methylomes of ESCs and iPSCs, a number of differences between them have been revealed, including 1,175 differentially methylated regions (DMRs) with a length varying from 1 to 11 thousand base pairs (the total length being 1.68 million bp). No DMRs have been detected between two ESC lines analyzed under the same conditions. Differentially methylated regions in ESCs and iPSCs can be subdivided into two groups. The first group contains DMRs whose emergence can be attributed to inheritance of the methylation pattern of the somatic precursor cells of iPSCs (44–49% of the total number of DMR). The second group contains DMRs whose methylation pattern is specific to iPSCs (i.e., differs from the DMR pattern both in somatic cells and in ESCs). DMRs of this kind make up 51–56% of the total number of detected DMRs. DMR distribution varies in five of the iPSC lines that have been analyzed: 62% occur in two lines out of five; 16% occur in all five lines. These regions can be regarded as "hotspots" of epigenetic reprogramming, which require increased attention when obtaining iPSCs. A significant number of DMRs (80%) are associated with CpG islands; 62% localize near the genes or in the genes; 29 and 19% lie within 2 thousand bp from the transcription start or termination sites, respectively. A bioinformatic analysis of the function of the genes localized near DMRs and occurring in all the iPSCs under analysis showed no marked predominance of the genes involved in certain cellular processes. This attests to the fact that methylation disturbance during the reprogramming can affect a large number of cellular functions. Another important regularity is the predominance of hypomethylation in DMRs (109 out of 130, 92%) in all five lines. The disturbances in methylome reprogramming when obtaining iPSCs can be attributed to insufficient methylation.

DMRs have also been detected by an attentive analysis and comparison of CpH methylation in ESCs and iPSCs. A total of 29 regions have been found; they are characterized by the extensive length (half of these regions is over 1 million bp long; the longest one is 4.8 million bp); the total length of CpH-DMR is 32.4 million bp. Most CpH-DMRs in iPSCs are hypomethylated as compared with ESCs; they localize near centromeres and telomeres. These regions are enriched in histone H3 trimethylated at K9 (H3K9me3) and colocalized with hypermethylated CpG-DMRs. Most genes localized in these regions are characterized by an increased level of methylation of promoter regions and, therefore, by a reduced transcription level. It is interesting that the level of the inactive chromatin mark (H3K27me3) is reduced in these regions. Thus, extensive domains associated with the near-centromeric and near-telomeric regions with aberrant distribution of histone modifications, disturbed patterns of CpG and CpH methylation, and a disturbed level of gene transcription, have been revealed in human iPSCs. These "hotspots" of epigenomes undoubtedly need to be subjected to a thorough investigation when obtaining new human iPSC lines [144].

An investigation into CpG methylation in 22 human iPSC lines derived from various somatic cells (endometrial cells, umbilical vein epithelial cells, amnion cells, fetal lung fibroblasts, and menstrual blood cells) has also revealed differences from ESCs [186]. 1,459 differentially methylated CpG sites corresponding to 1,260 genes were detected when comparing all iPSC and ESC lines using a DNA microchip containing probes for 24,273 CpG sites within 13,728 genes. However, the number and distribution of these sites in different iPSC lines varied considerably. The reason may be that the lines were obtained from somatic

cells of different types. In more than 15 lines out of 22, only 20 sites were shared. It is worth noting that the number of these sites was increased in XX iPSCs. The comparison of these data with the results obtained by R. Lister *et al.* [144] has revealed 72 differentially methylated promoters in both studies. However, according to [186], most DMRs in iPSCs were hypermethylated as compared to ESCs; hence, it was postulated that the iPSC genome is methylated to a higher extent. On the contrary, R. Lister *et al.* [144] have reported hypomethylation of CpG dinucleotides in iPSCs. However, these differences can be attributed to the features of the experimental approaches used. In particular, K. Nishino [186] has analyzed the CpG localized mostly within the CpG islands in the promoter regions of the genes, whereas R. Lister *et al.* determined cytosine methylation in the entire genome. Furthermore, it has been clearly demonstrated [186] that the level of aberrant hypermethylation at later passages (30–40) is considerably lower than that at earlier ones (4–6), whereas R. Lister *et al.* [144] used iPSC lines which had undergone tens of passages.

The surprising similarity between the transcriptomes and epigenomes of these cells and those of ESCs was emphasized in the early studies devoted to obtaining mouse and human iPSCs. Furthermore, it has been demonstrated that the gene transcription pattern in somatic cells at the genome-wide level changes to the maximum extent. However, it was established later that iPSCs retain certain (often rather insignificant) features of somatic transcriptomes and epigenomes [187, 188]. Despite its apparent unimportance, incomplete reprogramming of particular loci can considerably affect the properties of pluripotent cells by changing their differentiation ability. Thus, a significant similarity between mouse ESCs and iPSCs at the level of mRNA and miRNA transcription (with the exception of several transcripts) has been detected [189]. In particular, aberrant silencing of the imprinted *Dlk1-Dio3* locus has been observed in certain iPSC clones, including those derived from hematopoietic precursor cells, which are also characterized by a low transcription level of this locus. This effect is assumed to be caused by the "epigenetic memory." Due to the transcriptional disturbance in the *Dlk1-Dio3* locus iPSCs become incapable of efficient formation of chimera and cannot form the mouse organism via tetraploid complementation. It is interesting to mention that treatment with valproic acid, the histone deacetylase inhibitor, leads to transcription activation in the *Dlk1-Dio3* locus and restores the iPSC capability of tetraploid complementation and efficient formation of chimeric animals [189].

A number of interesting studies have been devoted to the investigation of the effect of the origin of iPSCs on their differentiation pattern [185, 186, 189, 190]. Thus, the properties of iPSCs derived from mouse hematopoietic, neuronal precursors and fibroblasts were compared with those of ESCs. Embryonic stem cells originated either from blastocysts obtained by nuclear transfer from somatic cells or from those obtained by natural fertilization. First, it turned out that the type of somatic cells strongly affects efficiency and quality in reprogramming. The molecular genetic parameters of iPSCs derived from hematopoietic cells were much closer to those of ESCs, whereas fibroblast-derived iPSCs gave rise only to partially reprogrammed clones. iPSCs derived from neuronal precursors were the closest to ESCs. Second, the differences between iPSCs and embryo-derived pluripotent cells have been revealed via the analysis of DNA methylation. Similar to the earlier studies, it has been established that iPSCs and embryo-derived pluripotent cells differed by a large number of DMRs. iPSCs obtained from neural precursors and fibroblasts are characterized by residual methylation of the loci responsible for the formation of the hematopoietic line, which causes a decreased differentiation level of these iPSCs in the corresponding direction. Third, the limitations on the directions of differentiation of iPSCs of a certain origin can be eliminated. If iPSCs derived from neuronal precursors are differentiated into hematopoietic cell lines and secondary iPSCs are subsequently obtained from these derivatives, these secondary iPSCs will have a higher potential towards differentiation into blood cells. Furthermore, the impact of the inhibitors of histone deacetylases and DNA methylation (such as trichostatin A and 5-azacytidine) on the epigenome can considerably reduce the effect of the cell origin on their differentiation [187]. It should be mentioned that iPSCs at very early passages were used in [187]. Aberrant cytosine methylation at early passages and, therefore, disruption of the pattern of gene expression and iPSC differentiation have also been revealed in other studies. Thus, it has been demonstrated that mouse iPSCs derived from fibroblasts, B lymphocytes, bone marrow granulocytes, and precursor cells of skeletal muscles possess "epigenetic memory," which is manifested at the transcriptional level and results in differentiation predominantly into the cell types from which they had been obtained [190]. It has been established that the genes that are markers of certain somatic cells can continue being expressed at a high level in pluripotent cells, with the inactive chromatin marks (H3K27me3) in their promoter regions being reduced and active chromatin marks (H3Ac and H3K4me3) being increased. No differences in the methylation of the promoters of these genes have been observed [190]. It is significant that these transcription disturbances and shifts in cell differentiation are elimi-

nated after long-term cultivation of the iPSCs clones. These data, along with the results of other studies, attest to the fact that reprogramming is a gradual process; the establishment of the completely reprogrammed state of epigenome and cells in general requires a large number of rounds of genome replication.

In addition to the studies focused on the disturbance of epigenome reprogramming and "epigenetic memory" in mouse iPSCs, several papers have already been published, which have confirmed the fact that a similar phenomenon exists in the reprogramming of human cells. It has been demonstrated that iPSCs derived from neuronal precursors and β cells of the pancreatic gland and human retinal pigment epithelium can have a non-random differentiation pattern; i.e., the direction of differentiation is strongly tilted towards the precursor type of somatic cells [188, 191, 192]. Aberrantly methylated regions have also been detected in iPSCs derived from umbilical cord blood cells and neonatal keratinocytes, and the existence of the "epigenetic memory" has been established, which consists in predominant differentiation into parent-type cells and is retained even after a large number of passages [193].

Thus, the problem of "epigenetic memory" today remains among the major hurdles in the derivation and application of induced pluripotent stem cells. It is a pressing problem, especially due to the fact that iPSCs display great potential for use in regenerative medicine and as models for human diseases. Resolution of this problem will not only enable efficient usage of human and animal iPSCs for biomedical purposes, but can also provide new fundamental knowledge on the organization and role of cell epigenomes in culture and during the embryonic development of organisms. ●

This work was supported by the Russian Foundation for Basic Research (grants № 11-04-00847-a and 12-04-00208-a) and the Program of the Russian Academy of Sciences "Molecular and Cell Biology".

REFERENCES

1. Boiani M., Scholer H.R. // Nat. Rev. Mol. Cell Biol. 2005. V. 6. № 11. P. 872–884.
2. Evans M.J., Kaufman M.H. // Nature. 1981. V. 292. № 5819. P. 154–156.
3. Martin G.R. // Proc. Natl. Acad. Sci. USA. 1981. V. 78. № 12. P. 7634–7638.
4. 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.
5. Cohen D.E., Melton D. // Nat. Rev Genet. 2011. V. 12. № 4. P. 243–252.
6. Grskovic M., Javaherian A., Strulovici B., Daley G.Q. // Nat. Rev. Drug. Discov. 2011. V. 10. № 12. P. 915–929.
7. Tiscornia G., Vivas E.L., Belmonte J.C. // Nat. Med. 2011. V. 17. № 12. P. 1570–1576.
8. Boyer L.A., Lee T.I., Cole M.F., Johnstone S.E., Levine S.S., Zucker J.P., Guenther M.G., Kumar R.M., Murray H.L., Jenner R.G., et al. // Cell. 2005. V. 122. № 6. P. 947–956.
9. Loh Y.H., Wu Q., Chew J.L., Vega V.B., Zhang W., Chen X., Bourque G., George J., Leong B., Liu J., et al. // Nat. Genet. 2006. V. 38. № 4. P. 431–440.
10. Takahashi K., Yamanaka S. // Cell. 2006. V. 126. № 4. P. 663-676.
11. Takahashi K., Tanabe K., Ohnuki M., Narita M., Ichisaka T., Tomoda K., Yamanaka S. // Cell. 2007. V. 131. № 5. P. 861–872.
12. Yu J., Vodyanik M.A., Smuga-Otto K., Antosiewicz-Bourget J., Frane J.L., Tian S., Nie J., Jonsdottir G.A., Ruotti V., Stewart R., et al. // Science. 2007. V. 318. № 5858. P. 1917–1920.
13. Maherali N., Hochedlinger K. // Cell Stem. Cell. 2008. V. 3. № 6. P. 595–605.
14. Anokye-Danso F., Trivedi C.M., Juhr D., Gupta M., Cui Z., Tian Y., Zhang Y., Yang W., Gruber P.J., Epstein J.A., et al. // Cell Stem. Cell. 2011. V. 8. № 4. P. 376–388.
15. Miyoshi N., Ishii H., Nagano H., Haraguchi N., Dewi D.L., Kano Y., Nishikawa S., Tanemura M., Mimori K., Tanaka F., et al. // Cell Stem. Cell. 2011. V. 8. № 6. P. 633–638.
16. Aasen T., Raya A., Barrero M.J., Garreta E., Consiglio A., Gonzalez F., Vassena R., Bilic J., Pekarik V., Tiscornia G., et al. // Nat. Biotechnol. 2008. V. 26. № 11. P. 1276–1284.
17. Utikal J., Maherali N., Kulalert W., Hochedlinger K. // J. Cell Sci. 2009. V. 122. Pt 19. P. 3502–3510.
18. Huangfu D., Osafune K., Maehr R., Guo W., Eijkelenboom A., Chen S., Muhlestein W., Melton D.A. // Nat. Biotechnol. 2008. V. 26. № 11. P. 1269–1275.
19. Loh Y.H., Hartung O., Li H., Guo C., Sahalie J.M., Manos P.D., Urbach A., Heffner G.C., Grskovic M., Vigneault F., et al. // Cell Stem. Cell. 2010. V. 7. № 1. P. 15–19.
20. 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.
21. Stadtfeld M., Brennand K., Hochedlinger K. // Curr Biol. 2008. V. 18. № 12. P. 890–894.
22. Hanna J., Markoulaki S., Schorderet P., Carey B.W., Beard C., Wernig M., Creyghton M.P., Steine E.J., Cassady J.P., Foreman R., et al. // Cell. 2008. V. 133. № 2. P. 250–264.
23. Maherali N., Ahfeldt T., Rigamonti A., Utikal J., Cowan C., Hochedlinger K. // Cell Stem. Cell. 2008. V. 3. № 3. P. 340–345.
24. Yamanaka S. // Cell. 2009. V. 137. № 1. P. 13–17.
25. Boland M.J., Hazen J.L., Nazor K.L., Rodriguez A.R., Gifford W., Martin G., Kupriyanov S., Baldwin K.K. // Nature. 2009. V. 461. № 7260. P. 91–94.
26. Kang L., Wang J., Zhang Y., Kou Z., Gao S. // Cell Stem. Cell. 2009. V. 5. № 2. P. 135–138.
27. Zhao X.Y., Li W., Lv Z., Liu L., Tong M., Hai T., Hao J., Guo C.L., Ma Q.W., Wang L., et al. // Nature. 2009. V. 461. № 7260. P. 86–90.
28. Bernstein B.E., Mikkelsen T.S., Xie X., Kamal M., Huebert D.J., Cuff J., Fry B., Meissner A., Wernig M., Plath K., et al. // Cell. 2006. V. 125. № 2. P. 315–326.

29. Azuara V., Perry P., Sauer S., Spivakov M., Jorgensen H.F., John R.M., Gouti M., Casanova M., Warnes G., Merkenschlager M., et al. // Nat. Cell Biol. 2006. V. 8. № 5. P. 532–538.
30. Pan G., Tian S., Nie J., Yang C., Ruotti V., Wei H., Jonsdottir G.A., Stewart R., Thomson J.A. // Cell Stem. Cell. 2007. V. 1. № 3. P. 299–312.
31. Zhao X.D., Han X., Chew J.L., Liu J., Chiu K.P., Choo A., Orlov Y.L., Sung W.K., Shahab A., Kuznetsov V.A., et al. // Cell Stem. Cell. 2007. V. 1. № 3. P. 286–298.
32. Creyghton M.P., Markoulaki S., Levine S.S., Hanna J., Lodato M.A., Sha K., Young R.A., Jaenisch R., Boyer L.A. // Cell. 2008. V. 135. № 4. P. 649–661.
33. Mikkelsen T.S., Ku M., Jaffe D.B., Issac B., Lieberman E., Giannoukos G., Alvarez P., Brockman W., Kim T.K., Koche R.P., et al. // Nature. 2007. V. 448. № 7153. P. 553–560.
34. Mohn F., Weber M., Rebhan M., Roloff T.C., Richter J., Stadler M.B., Bibel M., Schubeler D. // Mol. Cell. 2008. V. 30. № 6. P. 755–766.
35. Liang J., Wan M., Zhang Y., Gu P., Xin H., Jung S.Y., Qin J., Wong J., Cooney A.J., Liu D., et al. // Nat. Cell. Biol. 2008. V. 10. № 6. P. 731–739.
36. Pardo M., Lang B., Yu L., Prosser H., Bradley A., Babu M.M., Choudhary J. // Cell Stem. Cell. 2010. V. 6. № 4. P. 382–395.
37. Wang J., Rao S., Chu J., Shen X., Levasseur D.N., Theunissen T.W., Orkin S.H. // Nature. 2006. V. 444. № 7117. P. 364–368.
38. van den Berg D.L., Snoek T., Mullin N.P., Yates A., Bezstarosti K., Demmers J., Chambers I., Poot R.A. // Cell Stem. Cell. 2010. V. 6. № 4. P. 369–381.
39. Ringrose L. // Bioessays. 2006. V. 28. № 4. P. 330–334.
40. Ringrose L. // Curr. Opin. Cell Biol. 2007. V. 19. № 3. P. 290–297.
41. Hauenschild A., Ringrose L., Altmutter C., Paro R., Rehmsmeier M. // PLoS Biol. 2008. V. 6. № 10. P. e261.
42. de Napoles M., Mermoud J.E., Wakao R., Tang Y.A., Endoh M., Appanah R., Nesterova T.B., Silva J., Otte A.P., Vidal M., et al. // Dev. Cell. 2004. V. 7. № 5. P. 663–676.
43. Wang H., Wang L., Erdjument-Bromage H., Vidal M., Tempst P., Jones R.S., Zhang Y. // Nature. 2004. V. 431. № 7010. P. 873–878.
44. Elderkin S., Maertens G.N., Endoh M., Mallery D.L., Morrice N., Koseki H., Peters G., Brockdorff N., Hiom K. // Mol. Cell. 2007. V. 28. № 1. P. 107–120.
45. Voncken J.W., Roelen B.A., Roefs M., de Vries S., Verhoeven E., Marino S., Deschamps J., van Lohuizen M. // Proc. Natl. Acad. Sci. USA. 2003. V. 100. № 5. P. 2468–2473.
46. Lessard J., Sauvageau G. // Nature. 2003. V. 423. № 6937. P. 255–260.
47. Park I.K., Qian D., Kiel M., Becker M.W., Pihalja M., Weissman I.L., Morrison S.J., Clarke M.F. // Nature. 2003. V. 423. № 6937. P. 302–305.
48. Dovey J.S., Zacharek S.J., Kim C.F., Lees J.A. // Proc. Natl. Acad. Sci. USA. 2008. V. 105. № 33. P. 11857–11862.
49. Molofsky A.V., Pardal R., Iwashita T., Park I.K., Clarke M.F., Morrison S.J. // Nature. 2003. V. 425. № 6961. P. 962–967.
50. Liu J., Cao L., Chen J., Song S., Lee I.H., Quijano C., Liu H., Keyvanfar K., Chen H., Cao L.Y., et al. // Nature. 2009. V. 459. № 7245. P. 387–392.
51. Endoh M., Endo T.A., Endoh T., Fujimura Y., Ohara O., Toyoda T., Otte A.P., Okano M., Brockdorff N., Vidal M., et al. // Development. 2008. V. 135. № 8. P. 1513–1524.
52. Morey L., Pascual G., Cozzuto L., Roma G., Wutz A., Benitah S.A., Di Croce L. // Cell Stem. Cell. 2012. V. 10. № 1. P. 47–62.
53. O'Loghlen A., Munoz-Cabello A.M., Gaspar-Maia A., Wu H.A., Banito A., Kunowska N., Racek T., Pemberton H.N., Beolchi P., Lavial F., et al. // Cell Stem. Cell. 2012. V. 10. № 1. P. 33–46.
54. Morey L., Helin K. // Trends Biochem. Sci. 2010. V. 35. № 6. P. 323–332.
55. Camahort R., Cowan C.A. // Cell Stem. Cell. 2012. V. 10. № 1. P. 4–6.
56. Pasini D., Bracken A.P., Hansen J.B., Capillo M., Helin K. // Mol. Cell Biol. 2007. V. 27. № 10. P. 3769–3779.
57. Shumacher A., Faust C., Magnuson T. // Nature. 1996. V. 383. № 6597. P. 250–253.
58. Faust C., Schumacher A., Holdener B., Magnuson T. // Development. 1995. V. 121. № 2. P. 273–285.
59. Chamberlain S.J., Yee D., Magnuson T. // Stem. Cells. 2008. V. 26. № 6. P. 1496–1505.
60. Shen X., Liu Y., Hsu Y.J., Fujiwara Y., Kim J., Mao X., Yuan G.C., Orkin S.H. // Mol. Cell. 2008. V. 32. № 4. P. 491–502.
61. Peng J.C., Valouev A., Swigut T., Zhang J., Zhao Y., Sidow A., Wysocka J. // Cell. 2009. V. 139. № 7. P. 1290–1302.
62. Shen X., Kim W., Fujiwara Y., Simon M.D., Liu Y., Mysliwiec M.R., Yuan G.C., Lee Y., Orkin S.H. // Cell. 2009. V. 139. № 7. P. 1303–1314.
63. Pasini D., Cloos P.A., Walfridsson J., Olsson L., Bukowski J.P., Johansen J.V., Bak M., Tommerup N., Rappsilber J., Helin K. // Nature. 2010. V. 464. № 7286. P. 306–310.
64. Zhang Z., Jones A., Sun C.W., Li C., Chang C.W., Joo H.Y., Dai Q., Mysliwiec M.R., Wu L.C., Guo Y., et al. // Stem. Cells. 2011. V. 29. № 2. P 229–240.
65. Li G., Margueron R., Ku M., Chambon P., Bernstein B.E., Reinberg D. // Genes Dev. 2010. V. 24. № 4. P. 368–380.
66. Landeira D., Sauer S., Poot R., Dvorkina M., Mazzarella L., Jorgensen H.F., Pereira C.F., Leleu M., Piccolo F.M., Spivakov M., et al. // Nat. Cell Biol. 2010. V. 12. № 6. P. 618–624.
67. Zhang Z., Jones A., Sun C.W., Li C., Chang C.W., Joo H.Y., Dai Q., Mysliwiec M.R., Wu L.C., Guo Y., et al. // Stem. Cells. 2011. V. 29. № 2. P. 229–240.
68. Landeira D., Fisher A.G. // Trends Cell Biol. 2011. V. 21. № 2. P. 74–80.
69. Stock J.K., Giadrossi S., Casanova M., Brookes E., Vidal M., Koseki H., Brockdorff N., Fisher A.G., Pombo A. // Nat. Cell Biol. 2007. V. 9. № 12. P. 1428–1435.
70. Guenther M.G., Levine S.S., Boyer L.A., Jaenisch R., Young R.A. // Cell. 2007. V. 130. № 1. P. 77–88.
71. Boyer L.A., Plath K., Zeitlinger J., Brambrink T., Medeiros L.A., Lee T.I., Levine S.S., Wernig M., Tajonar A., Ray M.K., et al. // Nature. 2006. V. 441. № 7091. P. 349–353.
72. Lee T.I., Jenner R.G., Boyer L.A., Guenther M.G., Levine S.S., Kumar R.M., Chevalier B., Johnstone S.E., Cole M.F., Isono K., et al. // Cell. 2006. V. 125. № 2. P. 301–313.
73. Ringrose L., Paro R. // Annu Rev. Genet. 2004. V. 38. P. 413–443.
74. Ang Y.S., Tsai S.Y., Lee D.F., Monk J., Su J., Ratnakumar K., Ding J., Ge Y., Darr H., Chang B., et al. // Cell. 2011. V. 145. № 2. P. 183–197.
75. Dou Y., Milne T.A., Ruthenburg A.J., Lee S., Lee J.W., Verdine G.L., Allis C.D., Roeder R.G. // Nat. Struct. Mol. Biol. 2006. V. 13. № 8. P. 713–719.
76. Cao F., Chen Y., Cierpicki T., Liu Y., Basrur V., Lei M., Dou Y. // PLoS One. 2010. V. 5. № 11. P. e14102.
77. Pullirsch D., Hartel R., Kishimoto H., Leeb M., Steiner G., Wutz A. // Development. 2010. V. 137. № 6. P. 935–943.

78. Wysocka J., Swigut T., Milne T.A., Dou Y., Zhang X., Burlingame A.L., Roeder R.G., Brivanlou A.H., Allis C.D. // Cell. 2005. V. 121. № 6. P. 859–872.
79. Kim J.K., Huh S.O., Choi H., Lee K.S., Shin D., Lee C., Nam J.S., Kim H., Chung H., Lee H.W., et al. // Mol. Cell Biol. 2001. V. 21. № 22. P. 7787–7795.
80. Wu J.I., Lessard J., Olave I.A., Qiu Z., Ghosh A., Graef I.A., Crabtree G.R. // Neuron. 2007. V. 56. № 1. P. 94–108.
81. Wu J.I., Lessard J., Crabtree G.R. // Cell. 2009. V. 136. № 2. P. 200–206.
82. Lessard J., Wu J.I., Ranish J.A., Wan M., Winslow M.M., Staahl B.T., Wu H., Aebersold R., Graef I.A., Crabtree G.R. // Neuron. 2007. V. 55. № 2. P. 201–215.
83. Ho L., Jothi R., Ronan J.L., Cui K., Zhao K., Crabtree G.R. // Proc. Natl. Acad. Sci. USA. 2009. V. 106. № 13. P. 5187–5191.
84. Ho L., Ronan J.L., Wu J., Staahl B.T., Chen L., Kuo A., Lessard J., Nesvizhskii A.I., Ranish J., Crabtree G.R. // Proc. Natl. Acad. Sci. USA. 2009. V. 106. № 13. P. 5181–5186.
85. Roberts C.W., Galusha S.A., McMenamin M.E., Fletcher C.D., Orkin S.H. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. № 25. P. 13796–13800.
86. Bultman S.J., Gebuhr T.C., Pan H., Svoboda P., Schultz R.M., Magnuson T. // Genes Dev. 2006. V. 20. № 13. P. 1744–1754.
87. Gao X., Tate P., Hu P., Tjian R., Skarnes W.C., Wang Z. // Proc. Natl. Acad. Sci. USA. 2008. V. 105. № 18. P. 6656–6661.
88. Guidi C.J., Sands A.T., Zambrowicz B.P., Turner T.K., Demers D.A., Webster W., Smith T.W., Imbalzano A.N., Jones S.N. // Mol. Cell Biol. 2001. V. 21. № 10. P. 3598–3603.
89. Fazzio T.G., Huff J.T., Panning B. // Cell. 2008. V. 134. № 1. P. 162–174.
90. Schaniel C., Ang Y.S., Ratnakumar K., Cormier C., James T., Bernstein E., Lemischka I.R., Paddison P.J. // Stem. Cells. 2009. V. 27. № 12. P. 2979–2991.
91. Seki Y., Kurisaki A., Watanabe-Susaki K., Nakajima Y., Nakanishi M., Arai Y., Shiota K., Sugino H., Asashima M. // Proc. Natl. Acad. Sci. USA. 2010. V. 107. № 24. P. 10926–10931.
92. Hu G., Kim J., Xu Q., Leng Y., Orkin S.H., Elledge S.J. // Genes Dev. 2009. V. 23. № 7. P. 837–848.
93. Singhal N., Graumann J., Wu G., Arauzo-Bravo M.J., Han D.W., Greber B., Gentile L., Mann M., Scholer H.R. // Cell. 2010. V. 141. № 6. P. 943–955.
94. He L., Liu H., Tang L. // Stem. Cell Rev. 2012. V. 8. № 1. P. 128–136.
95. Matsuda T., Nakamura T., Nakao K., Arai T., Katsuki M., Heike T., Yokota T. // EMBO J. 1999. V. 18. № 15. P. 4261–4269.
96. Ho L., Miller E.L., Ronan J.L., Ho W.Q., Jothi R., Crabtree G.R. // Nat. Cell Biol. 2011. V. 13. № 8. P. 903–913.
97. Denslow S.A., Wade P.A. // Oncogene. 2007. V. 26. № 37. P. 5433–5438.
98. McDonel P., Costello I., Hendrich B. // Int. J. Biochem. Cell Biol. 2009. V. 41. № 1. P. 108–116.
99. Kaji K., Caballero I.M., MacLeod R., Nichols J., Wilson V.A., Hendrich B. // Nat. Cell Biol. 2006. V. 8. № 3. P. 285–292.
100. Zhu D., Fang J., Li Y., Zhang J. // PLoS One. 2009. V. 4. № 11. P. e7684.
101. Kaji K., Nichols J., Hendrich B. // Development. 2007. V. 134. № 6. P. 1123–1132.
102. Reynolds N., Salmon-Divon M., Dvinge H., Hynes-Allen A., Balasooriya G., Leaford D., Behrens A., Bertone P., Hendrich B. // EMBO J. 2011. V. 31. № 3. P. 593–605.
103. Yildirim O., Li R., Hung J.H., Chen P.B., Dong X., Ee L.S., Weng Z., Rando O.J., Fazzio T.G. // Cell. 2011. V. 147. № 7. P. 1498–1510.
104. Bhutani N., Burns D.M., Blau H.M. // Cell. 2011. V. 146. № 6. P. 866–872.
105. Tahiliani M., Koh K.P., Shen Y., Pastor W.A., Bandukwala H., Brudno Y., Agarwal S., Iyer L.M., Liu D.R., Aravind L., et al. // Science. 2009. V. 324. № 5929. P. 930–935.
106. Ito S., D'Alessio A.C., Taranova O.V., Hong K., Sowers L.C., Zhang Y. // Nature. 2010. V. 466. № 7310. P. 1129–1133.
107. Koh K.P., Yabuuchi A., Rao S., Huang Y., Cunniff K., Nardone J., Laiho A., Tahiliani M., Sommer C.A., Mostoslavsky G., et al. // Cell Stem. Cell. 2011. V. 8. № 2. P. 200–213.
108. Inoue A., Zhang Y. // Science. 2011. V. 334. № 6053. P. 194.
109. Iqbal K., Jin S.G., Pfeifer G.P., Szabo P.E. // Proc. Natl. Acad. Sci. USA. 2011. V. 108. № 9. P. 3642–3647.
110. Ikura T., Ogryzko V.V., Grigoriev M., Groisman R., Wang J., Horikoshi M., Scully R., Qin J., Nakatani Y. // Cell. 2000. V. 102. № 4. P. 463–473.
111. Cai Y., Jin J., Tomomori-Sato C., Sato S., Sorokina I., Parmely T.J., Conaway R.C., Conaway J.W. // J. Biol. Chem. 2003. V. 278. № 44. P. 42733–42736.
112. Sapountzi V., Logan I.R., Robson C.N. // Int. J. Biochem. Cell Biol. 2006. V. 38. № 9. P. 1496–1509.
113. Squatrito M., Gorrini C., Amati B. // Trends Cell Biol. 2006. V. 16. № 9. P. 433–442.
114. Herceg Z., Hulla W., Gell D., Cuenin C., Lleonart M., Jackson S., Wang Z.Q. // Nat. Genet. 2001. V. 29. № 2. P. 206–211.
115. Gorrini C., Squatrito M., Luise C., Syed N., Perna D., Wark L., Martinato F., Sardella D., Verrecchia A., Bennett S., et al. // Nature. 2007. V. 448. № 7157. P. 1063–1067.
116. Loh Y.H., Zhang W., Chen X., George J., Ng H.H. // Genes Dev. 2007. V. 21. № 20. P. 2545–2557.
117. Chen X., Xu H., Yuan P., Fang F., Huss M., Vega V.B., Wong E., Orlov Y.L., Zhang W., Jiang J., et al. // Cell. 2008. V. 133. № 6. P. 1106–1117.
118. Sims R.J., 3rd, Chen C.F., Santos-Rosa H., Kouzarides T., Patel S.S., Reinberg D. // J. Biol. Chem. 2005. V. 280. № 51. P. 41789–41792.
119. Gaspar-Maia A., Alajem A., Polesso F., Sridharan R., Mason M.J., Heidersbach A., Ramalho-Santos J., McManus M.T., Plath K., Meshorer E., et al. // Nature. 2009. V. 460. № 7257. P. 863–868.
120. Kooistra S.M., van den Boom V., Thummer R.P., Johannes F., Wardenaar R., Tesson B.M., Veenhoff L.M., Fusetti F., O'Neill L.P., Turner B.M., et al. // Stem. Cells. 2010. V. 28. № 10. P. 1703–1714.
121. Nishimoto M., Fukushima A., Okuda A., Muramatsu M. // Mol. Cell Biol. 1999. V. 19. № 8. P. 5453–5465.
122. Cedar H., Bergman Y. // Nat. Rev. Genet. 2009. V. 10. № 5. P. 295–304.
123. Bestor T.H. // Hum. Mol. Genet. 2000. V. 9. № 16. P. 2395–2402.
124. Li E., Bestor T.H., Jaenisch R. // Cell. 1992. V. 69. № 6. P. 915–926.
125. Lippman Z., Gendrel A.V., Black M., Vaughn M.W., Dedhia N., McCombie W.R., Lavine K., Mittal V., May B., Kasschau K.D., et al. // Nature. 2004. V. 430. № 6998. P. 471–476.
126. Okano M., Bell D.W., Haber D.A., Li E. // Cell. 1999. V. 99. № 3. P. 247–257.
127. Reik W. // Nature. 2007. V. 447. № 7143. P. 425–432.
128. Straussman R., Nejman D., Roberts D., Steinfeld I., Blum

B., Benvenisty N., Simon I., Yakhini Z., Cedar H. // Nat. Struct. Mol. Biol. 2009. V. 16. № 5. P. 564–571.
129. Fouse S.D., Shen Y., Pellegrini M., Cole S., Meissner A., van Neste L., Jaenisch R., Fan G. // Cell Stem. Cell. 2008. V. 2. № 2. P. 160–169.
130. Hattori N., Nishino K., Ko Y.G., Ohgane J., Tanaka S., Shiota K. // J. Biol. Chem. 2004. V. 279. № 17. P. 17063–17069.
131. Hattori N., Imao Y., Nishino K., Ohgane J., Yagi S., Tanaka S., Shiota K. // Genes Cells. 2007. V. 12. № 3. P. 387–396.
132. Li J.Y., Pu M.T., Hirasawa R., Li B.Z., Huang Y.N., Zeng R., Jing N.H., Chen T., Li E., Sasaki H., et al. // Mol. Cell Biol. 2007. V. 27. № 24. P. 8748–8759.
133. Li E., Beard C., Jaenisch R. // Nature. 1993. V. 366. № 6453. P. 362–365.
134. Kaneda M., Okano M., Hata K., Sado T., Tsujimoto N., Li E., Sasaki H. // Nature. 2004. V. 429. № 6994. P. 900–903.
135. Ueda Y., Okano M., Williams C., Chen T., Georgopoulos K., Li E. // Development. 2006. V. 133. № 6. P. 1183–1192.
136. Tsumura A., Hayakawa T., Kumaki Y., Takebayashi S., Sakaue M., Matsuoka C., Shimotohno K., Ishikawa F., Li E., Ueda H.R., et al. // Genes Cells. 2006. V. 11. № 7. P. 805–814.
137. Feldman N., Gerson A., Fang J., Li E., Zhang Y., Shinkai Y., Cedar H., Bergman Y. // Nat. Cell Biol. 2006. V. 8. № 2. P. 188–194.
138. Bird A. // Genes Dev. 2002. V. 16. № 1. P. 6–21.
139. Meissner A., Mikkelsen T.S., Gu H., Wernig M., Hanna J., Sivachenko A., Zhang X., Bernstein B.E., Nusbaum C., Jaffe D.B., et al. // Nature. 2008. V. 454. № 7205. P. 766–770.
140. Thomson J.P., Skene P.J., Selfridge J., Clouaire T., Guy J., Webb S., Kerr A.R., Deaton A., Andrews R., James K.D., et al. // Nature. 2010. V. 464. № 7291. P. 1082–1086.
141. Ooi S.K., Qiu C., Bernstein E., Li K., Jia D., Yang Z., Erdjument-Bromage H., Tempst P., Lin S.P., Allis C.D., et al. // Nature. 2007. V. 448. № 7154. P. 714–717.
142. Ramsahoye B.H., Biniszkiewicz D., Lyko F., Clark V., Bird A.P., Jaenisch R. // Proc. Natl. Acad. Sci. USA. 2000. V. 97. № 10. P. 5237–5242.
143. Lister R., Pelizzola M., Dowen R.H., Hawkins R.D., Hon G., Tonti-Filippini J., Nery J.R., Lee L., Ye Z., Ngo Q.M., et al. // Nature. 2009. V. 462. № 7271. P. 315–322.
144. Lister R., Pelizzola M., Kida Y.S., Hawkins R.D., Nery J.R., Hon G., Antosiewicz-Bourget J., O'Malley R., Castanon R., Klugman S., et al. // Nature. 2011. V. 471. № 7336. P. 68–73.
145. Ziller M.J., Muller F., Liao J., Zhang Y., Gu H., Bock C., Boyle P., Epstein C.B., Bernstein B.E., Lengauer T., et al. // PLoS Genet. 2011. V. 7. № 12. P. e1002389.
146. Mikkelsen T.S., Hanna J., Zhang X., Ku M., Wernig M., Schorderet P., Bernstein B.E., Jaenisch R., Lander E.S., Meissner A. // Nature. 2008. V. 454. № 7200. P. 49–55.
147. Lagarkova M.A., Shutova M.V., Bogomazova A.N., Vassina E.M., Glazov E.A., Zhang P., Rizvanov A.A., Chestkov I.V., Kiselev S.L. // Cell Cycle. 2010. V. 9. № 5. P. 937–946.
148. Okita K., Ichisaka T., Yamanaka S. // Nature. 2007. V. 448. № 7151. P. 313–317.
149. Bhutani N., Brady J.J., Damian M., Sacco A., Corbel S.Y., Blau H.M. // Nature. 2010. V. 463. № 7284. P. 1042–1047.
150. Shi Y., Desponts C., Do J.T., Hahm H.S., Scholer H.R., Ding S. // Cell Stem. Cell. 2008. V. 3. № 5. P. 568–574.
151. Sado T., Wang Z., Sasaki H., Li E. // Development. 2001. V. 128. № 8. P. 1275–1286.
152. Mak W., Nesterova T.B., de Napoles M., Appanah R., Yamanaka S., Otte A.P., Brockdorff N. // Science. 2004. V. 303. № 5658. P. 666–669.
153. Okamoto I., Otte A.P., Allis C.D., Reinberg D., Heard E. // Science. 2004. V. 303. № 5658. P. 644–649.
154. Chureau C., Prissette M., Bourdet A., Barbe V., Cattolico L., Jones L., Eggen A., Avner P., Duret L. // Genome Res. 2002. V. 12. № 6. P. 894–908.
155. Brockdorff N., Ashworth A., Kay G.F., McCabe V.M., Norris D.P., Cooper P.J., Swift S., Rastan S. // Cell. 1992. V. 71. № 3. P. 515–526.
156. Lee J.T., Davidow L.S., Warshawsky D. // Nat. Genet. 1999. V. 21. № 4. P. 400–404.
157. Lee J.T., Lu N. // Cell. 1999. V. 99. № 1. P. 47–57.
158. Silva S.S., Rowntree R.K., Mekhoubad S., Lee J.T. // Proc. Natl. Acad. Sci. USA. 2008. V. 105. № 12. P. 4820–4825.
159. Maherali N., Sridharan R., Xie W., Utikal J., Eminli S., Arnold K., Stadtfeld M., Yachechko R., Tchieu J., Jaenisch R., et al. // Cell Stem. Cell. 2007. V. 1. № 1. P. 55–70.
160. Tchieu J., Kuoy E., Chin M.H., Trinh H., Patterson M., Sherman S.P., Aimiuwu O., Lindgren A., Hakimian S., Zack J.A., et al. // Cell Stem. Cell. 2010. V. 7. № 3. P. 329–342.
161. Lengner C.J., Gimelbrant A.A., Erwin J.A., Cheng A.W., Guenther M.G., Welstead G.G., Alagappan R., Frampton G.M., Xu P., Muffat J., et al. // Cell. 2010. V. 141. № 5. P. 872–883.
162. Hanna J., Cheng A.W., Saha K., Kim J., Lengner C.J., Soldner F., Cassady J.P., Muffat J., Carey B.W., Jaenisch R. // Proc. Natl. Acad. Sci. USA. 2010. V. 107. № 20. P. 9222–9227.
163. Navarro P., Chambers I., Karwacki-Neisius V., Chureau C., Morey C., Rougeulle C., Avner P. // Science. 2008. V. 321. № 5896. P. 1693–1695.
164. Navarro P., Moffat M., Mullin N.P., Chambers I. // Hum. Genet. 2011. V. 130. № 2. P. 255–264.
165. Barakat T.S., Gunhanlar N., Pardo C.G., Achame E.M., Ghazvini M., Boers R., Kenter A., Rentmeester E., Grootegoed J.A., Gribnau J. // PLoS Genet. 2011. V. 7. № 1. P. e1002001.
166. Orkin S.H., Hochedlinger K. // Cell. 2011. V. 145. № 6. P. 835–850.
167. Navarro P., Oldfield A., Legoupi J., Festuccia N., Dubois A., Attia M., Schoorlemmer J., Rougeulle C., Chambers I., Avner P. // Nature. 2010. V. 468. № 7322. P. 457–460.
168. Donohoe M.E., Silva S.S., Pinter S.F., Xu N., Lee J.T. // Nature. 2009. V. 460. № 7251. P. 128–132.
169. 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.
170. Mattout A., Biran A., Meshorer E. // J. Mol. Cell Biol. 2011. V. 3. № 6. P. 341–350.
171. Shi Y., Do J.T., Desponts C., Hahm H.S., Scholer H.R., Ding S. // Cell Stem. Cell. 2008. V. 2. № 6. P. 525–528.
172. Huangfu D., Maehr R., Guo W., Eijkelenboom A., Snitow M., Chen A.E., Melton D.A. // Nat. Biotechnol. 2008. V. 26. № 7. P. 795–797.
173. Mali P., Chou B.K., Yen J., Ye Z., Zou J., Dowey S., Brodsky R.A., Ohm J.E., Yu W., Baylin S.B., et al. // Stem. Cells. 2010. V. 28. № 4. P. 713–720.
174. Wang T., Chen K., Zeng X., Yang J., Wu Y., Shi X., Qin B., Zeng L., Esteban M.A., Pan G., et al. // Cell Stem. Cell. 2011. V. 9. № 6. P. 575–587.
175. 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.
176. Hong H., Takahashi K., Ichisaka T., Aoi T., Kanagawa O., Nakagawa M., Okita K., Yamanaka S. // Nature. 2009. V. 460. № 7259. P. 1132–1135.
177. Kawamura T., Suzuki J., Wang Y.V., Menendez S., Mor-

era L.B., Raya A., Wahl G.M., Belmonte J.C. // Nature. 2009. V. 460. № 7259. P. 1140–1144.
178. Marion R.M., Strati K., Li H., Murga M., Blanco R., Ortega S., Fernandez-Capetillo O., Serrano M., Blasco M.A. // Nature. 2009. V. 460. № 7259. P. 1149–1153.
179. Li H., Collado M., Villasante A., Strati K., Ortega S., Canamero M., Blasco M.A., Serrano M. // Nature. 2009. V. 460. № 7259. P. 1136–1139.
180. Ruiz S., Panopoulos A.D., Herrerias A., Bissig K.D., Lutz M., Berggren W.T., Verma I.M., Izpisua Belmonte J.C. // Curr. Biol. 2011. V. 21. № 1. P. 45–52.
181. Onder T.T., Kara N., Cherry A., Sinha A.U., Zhu N., Bernt K.M., Cahan P., Mancarci O.B., Unternaehrer J., Gupta P.B., et al. // Nature. 2012. V. 483. № 7391. P. 598–602.
182. Guenther M.G., Frampton G.M., Soldner F., Hockemeyer D., Mitalipova M., Jaenisch R., Young R.A. // Cell Stem. Cell. 2010. V. 7. № 2. P. 249–257.
183. Chin M.H., Mason M.J., Xie W., Volinia S., Singer M., Peterson C., Ambartsumyan G., Aimiuwu O., Richter L., Zhang J., et al. // Cell Stem. Cell. 2009. V. 5. № 1. P. 111–123.
184. Newman A.M., Cooper J.B. // Cell Stem. Cell. 2010. V. 7. № 2. P. 258–262.
185. Bock C., Kiskinis E., Verstappen G., Gu H., Boulting G., Smith Z.D., Ziller M., Croft G.F., Amoroso M.W., Oakley D.H., et al. // Cell. 2011. V. 144. № 3. P. 439–452.
186. Nishino K., Toyoda M., Yamazaki-Inoue M., Fukawatase Y., Chikazawa E., Sakaguchi H., Akutsu H., Umezawa A. // PLoS Genet. 2011. V. 7. № 5. P. e1002085.
187. Kim K., Doi A., Wen B., Ng K., Zhao R., Cahan P., Kim J., Aryee M.J., Ji H., Ehrlich L.I., et al. // Nature. 2010. V. 467. № 7313. P. 285–290.
188. Bar-Nur O., Russ H.A., Efrat S., Benvenisty N. // Cell Stem. Cell. 2011. V. 9. № 1. P. 17–23.
189. Stadtfeld M., Apostolou E., Akutsu H., Fukuda A., Follett P., Natesan S., Kono T., Shioda T., Hochedlinger K. // Nature. 2010. V. 465. № 7295. P. 175–181.
190. Polo J.M., Liu S., Figueroa M.E., Kulalert W., Eminli S., Tan K.Y., Apostolou E., Stadtfeld M., Li Y., Shioda T., et al. // Nat. Biotechnol. 2010. V. 28. № 8. P. 848–855.
191. Marchetto M.C., Yeo G.W., Kainohana O., Marsala M., Gage F.H., Muotri A.R. // PLoS One. 2009. V. 4. № 9. P. e7076.
192. Hu Q., Friedrich A.M., Johnson L.V., Clegg D.O. // Stem. Cells. 2010. V. 28. № 11. P. 1981–1991.
193. Kim K., Zhao R., Doi A., Ng K., Unternaehrer J., Cahan P., Hongguang H., Loh Y.H., Aryee M.J., Lensch M.W., et al. // Nat. Biotechnol. 2011. V. 29. № 12. P. 1117–1119.

Effect of 3D Cultivation Conditions on the Differentiation of Endodermal Cells

O. S. Petrakova*[1,2], V. V. Ashapkin[3,4], E. A. Voroteliak[1], E. Y. Bragin[4], V. Y. Shtratnikova[4], E. S. Chernioglo[1], Y. V. Sukhanov[1], V. V. Terskikh[1], A. V. Vasiliev[1]

[1]Koltzov Institute of Developmental Biology, Russian Academy of Sciences, Vavilova Str. 26, Moscow, Russia, 119334

[2]Lomonosov Moscow State University, Faculty of Biology, Leninskie Gory 1/12, Moscow, Russia, 119991

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

[4]Center of Innovation and Technology of Biologically Active Compounds and Their Applications, Russian Academy of Sciences, Gubkina Str. 3/2, Moscow, Russia, 117312

*E-mail: PetrakovaOl@yandex.ru

Received 02.10.2012

ABSTRACT **Cellular therapy of endodermal organs is one of the most important issues in modern cellular biology and biotechnology. One of the most promising directions in this field is the study of the transdifferentiation abilities of cells within the same germ layer. A method for an *in vitro* investigation of the cell differentiation potential (the cell culture in a three-dimensional matrix) is described in this article. Cell cultures of postnatal salivary gland cells and postnatal liver progenitor cells were obtained; their comparative analysis under 2D and 3D cultivation conditions was carried out. Both cell types have high proliferative abilities and can be cultivated for more than 20 passages. Under 2D cultivation conditions, the cells remain in an undifferentiated state. Under 3D conditions, they undergo differentiation, which was confirmed by a lower cell proliferation and by an increase in the differentiation marker expression. Salivary gland cells can undergo hepatic and pancreatic differentiation under 3D cultivation conditions. Liver progenitor cells also acquire a pancreatic differentiation capability under conditions of 3D cultivation. Thus, postnatal salivary gland cells exhibit a considerable differentiation potential within the endodermal germ layer and can be used as a promising source of endodermal cells for the cellular therapy of liver pathologies. Cultivation of cells under 3D conditions is a useful model for the *in vitro* analysis of the cell differentiation potential.**

KEYWORDS **3D conditions; collagen gel; differentiation; endoderm; submandibular salivary gland cells; liver progenitor cells.**

ABBREVIATIONS **2D conditions – two-dimensional conditions; 3D conditions – three-dimensional conditions; BrdU - 5-bromo-2'-deoxyuridine; cDNA – complementary deoxyribonucleic acid; DAPI - 4',6-diamidino-2-phenylindole; DNA – deoxyribonucleic acid; EGF – epidermal growth factor; ITS – insulin–transferrin–selenium; mRNA – messenger ribonucleic acid; PLPC – postnatal liver progenitor cells; PSGC – postnatal salivary gland cells; RNA – ribonucleic acid; Real-time PCR – real-time polymerase chain reaction; RT-PCR – reverse transcription polymerase chain reaction.**

INTRODUCTION

Investigating the plasticity of the cellular phenotype and the ability of cells to undergo transdifferentiation within a single germ layer remain a pressing issue in contemporary cell biology. Such research may be helpful not only in finding solutions to fundamental problems, such as the elucidation of the differentiation pathways in the embryogenesis process, establishment of the histogenetic relationships between different cell types, but also in outlining new approaches in regenerative medicine.

The potential of cellular therapy for treating liver pathologies is being actively researched. Despite the relative success that has been achieved in investigations carried out on laboratory animals, no safe and sufficiently efficient approach has been found thus far [1, 2]. Today, the main task is to search for easily obtainable cells that can undergo hepatic differentiation with sufficient efficiency. In terms of experimental and clinical studies on embryonic stem cells, bone marrow- [7, 9] and adipose-derived [10–12] mesenchymal cells, as well as amniotic fluid cells [13, 14], are the ones

that have been best studied [3, 6]. However, only partial transdifferentiation was demonstrated in all studies; no functionally active state has been achieved: hence, the search for an optimal source of cells to treat liver pathologies still continues.

The salivary gland remains a relatively poorly studied source of endodermal cells. However, the availability of cellular material from the salivary gland, the possibilities of using these cells in autologic and allogenic variants, and the relatively low invasiveness of the biopsy procedure makes this source of endodermal cells promising in research.

A sufficient amount of data on the *in vitro* cultivation of salivary gland cells of human and animal origin has been accumulated today. The *in vitro* cultivated salivary gland cells are an actively proliferating culture that can undergo a significant number of passages [15]. Salivary gland cells of human and animal origin (mouse, rat, pig) are characterized by the expression of cytokeratins 18 and 19 and, frequently, of α-fetoprotein [16, 17]. Under certain cultivation conditions, these cells acquire the ability to synthesize glucagon, albumin, or insulin [18].

One of the approaches used in the *in vitro* investigation of the phenotypic plasticity of cells, cell cultivation in a 3D matrix (under 3D conditions) using a type I collagen gel, is discussed in the present work. Cell cultivation in the collagen gel is used to investigate the morphogenetic potential of cells [19, 20], cellular migration [21], and to assess the cell differential potential [22]. Moreover, cultivation of pancreatic β cells in a type I 4% collagen gel contributes to higher rates of survival for these cells and increases their functional activity [23]. As for salivary gland cells, mouse postnatal salivary gland cells (PSGCs) cultured under 3D conditions (matrigel) acquire the ability to express α-fetoprotein and albumin, which is typical of hepatic differentiation [16]. Type I collagen gel, along with fibronectin, is known to be the main component of the hepatic extracellular matrix. Thus, the present study allows not only to shed some light on the differentiation potential of endodermal cells, but also assist in assessing the abilities of a collagen matrix to initiate and maintain *in vitro* hepatic differentiation of cells.

The present study was aimed at investigating the capacity of mouse salivary gland cells to undergo hepatic differentiation during cultivation in a collagen gel.

The comparative analysis of the properties of mouse submandibular salivary gland cells and progenitor cells isolated from the liver was carried out. The morphological, immunophenotypic, and biochemical characteristics of the cell cultures were compared under 2D and 3D cultivation conditions; the gene expression profile of these cells was also analyzed by PCR.

EXPERIMENTAL

Animals

8–20-week-old male *C57BL/6* mice were used in the present work. The animals were kept under standard conditions and had food and water available ad libitum. All the procedures were carried out in accordance with the rules established by the bioethics committee at the Koltzov Institute of Developmental Biology, Russian Academy of Sciences.

Isolation and cultivation of mouse postnatal submandibular salivary gland cells and mouse liver progenitor cells

To obtain the mouse liver and submandibular salivary gland cell cultures, the animals were anesthetized by inhaled chloroform vapor and dissected. The neck and belly regions were cleaned with ethanol, skin was cut using sterile scissors, the liver and both submandibular salivary glands were extracted using tweezers. The organs were transferred into sterile test tubes containing DMEM/F12 1 :1 medium (Gibco) and 40 μg/ml of gentamicin. After the organs were minced and blood vessels and mesenchymal tissues were removed, the salivary gland and liver homogenate was washed twice with PBS, followed by treatment with a type IV collagenase solution (4 mg/ml, Sigma) in DMEM/F12 1 :1 medium for 30–40 minutes at 37°C. Cell suspensions were pipetted and passed through a 40-μm-pore-size filter to separate small cells from larger polyploid ones. The cells were washed twice with the cultivation medium. "Soft" centrifugation at 100 *g* was carried out for 2 min. The removal of erythrocytes and precipitation of primarily small cells with a higher specific density occurred under these conditions. After the supernatant was removed, the cells were re-suspended in a complete growth medium containing DMEM/F12 1:1, a 10% embryonic bovine serum (HyClone), 2 mM glutamine (Gibco), 1× ITS (Invitrogen), and 10 ng/ml EGF (Invitrogen). The cells were plated into culture dishes (Corning) coated with type I collagen at a density of 5×10^3 cells/cm^2 and cultivated under standard conditions at 37°C and 5% CO_2. The medium was replaced daily during the first 5 days and subsequently replaced every 3 days. During the splitting procedure cells were washed twice with PBS and incubated in the presence of 0.25% trypsin for 5 min at 37°C. The cells were split at a ratio of 1:3 and plated into type I collagen coated culture dishes.

Preparation of collagen gel, cultivation of cells under 3D conditions, collagen gel contraction

Collagen gel was prepared using the conventional procedure: type I collagen was extracted from rat tails as

described previously [20] and dissolved in sterile 0.1% acetic acid (5 mg/ml). First-passage cells were collected using trypsin and diluted with PBS, with allowance for the fact that the final concentration of cells be 1×10^6 cells/ml of gel. All materials were cooled to +4°C prior to gel preparation; the subsequent operations were carried out under cold conditions. Sterile components were added into a separate test tube in the following sequence: 0.34 M NaOH (Sigma) until final concentration of 0.023 mM, 7.5% Na_2CO_3 (PanEco) until final concentration of 0.26%, 10× DMEM (Sigma) until final concentrations of 1×, 100× glutamine (Gibco) until final concentration of 2 mM, 100× HEPES (Gibco) until final concentration of 1×, embryonic bovine serum (HyClone) until final concentration of 10%, followed by the addition of the collagen solution in acetic acid until a final concentration of collagen of 4%. The components were subsequently mixed 2–3 times, cells in PBS were added, the resulting mixture was additionally stirred once or twice, and the gel was poured into 35 mm Petri dishes (2 ml per plate). The gel was incubated in a CO_2 incubator at 37°C for 30 min until complete polymerization. After the gel had polymerized, 2 ml of complete growth medium was added to each dish. The gel was separated from the dish walls using a pipette tip. The gel was then placed into a CO_2 incubator; this very instant was assumed to be the zero hour of gel preparation. The cells were subsequently cultivated in the gel in the CO_2 incubator under the standard conditions; the medium was replaced every 2 days. In order to determine the extent of gel contraction, its diameter was measured every 24 h counting from the zero hour of gel preparation. Gel containing no cells was used as the negative control of contraction.

Immunohistochemistry

Collagen gel was incubated in the presence of 4% paraformaldehyde at room temperature for 30 min, followed by paraffin embedding in accordance with the standard methodology to prepare 40-μm-thick paraffin sections. The sections were stained with azure-eosin.

Immunocytochemistry

Cells were plated onto the Petri dishes coated with the type I collagen gel 48 h prior to fixation in order to carry out immunocytochemical staining. Paraformaldehyde (4%) was used for fixation (10 min, room temperature). The dishes were subsequently washed with PBS containing 0.1% Triton X-100, followed by blocking using 1% bovine serum albumin in PBS for 30 min at room temperature. The incubation in the presence of the primary antibodies in PBS was carried out for 60 min at 37°C (or at +4°C overnight) using the dilution recommended by the manufacturer (typically, 1:200–1:500).

Table 1. Antibodies used in the present work

Antibody	Antigen	Company, catalog number
Primary antibodies		
CK19	Cytokeratin 19	AbCam, # ab15463-1
ALB	Albumin	R & D, # MAB1455
CYP P450	Cytochrome P450 1A1	Millipore, # AB1258
BrdU	Bromodeoxyuridine	AbCam, # ab8152
Secondary antibodies		
Alexa Fluor® 488 donkey anti-rabbit IgG (H + L)		Invitrogen, # A-21206
Alexa Fluor® 488 goat anti-mouse IgG (H + L)		Invitrogen, # A-11029

The plates were washed with PBS 3 times for 10 min at 37°C, followed by incubation in the presence of the secondary antibodies in PBS (1:1000 dilution) for 40 min at 37°C. Washing with PBS was carried out 3 times for 10 min at 37°C with DAPI (Sigma) added during the last washing phase. The analysis was carried out on a fluorescent microscope. The antibodies used in the experiments are listed in *Table 1*.

During the analysis of the cells after cultivation under 3D conditions for 10 days, the gel was washed with PBS, minced, and incubated in the presence of 0.075% type II collagenase (Sigma) for 60 min at 37°C. After the gel had been digested, the cells were washed using PBS, plated onto dishes coated with type I collagen, and cultivated under the standard conditions (CO_2 incubator, 48 h). The conventional staining procedure was subsequently carried out (see above).

Determination of the proliferative activity of cells using bromodeoxyuridine

The proliferative activity of cells during cultivation under 2D and 3D conditions was determined by their ability to incorporate bromodeoxyuridine (BrdU). Fifteen hours prior to fixation, BrdU (Sigma) was added to the cells until a final concentration of 10 μM. The cells were subsequently washed using PBS during cultivation under 2D conditions and fixed in 70% ethanol (30 min, +4°C). An equivalent volume of 4N HCl was then added, and incubation at room temperature for 30 min was carried out. The cells were washed with PBS until neu-

Table 2. Primers used in RT-PCR

Primer	Gene	Nucleotide sequence	Amplicon, bp	Melting temperature, ˚C
mGAPDH	Glyceraldehyde-3-phosphate	5'-AGG TCG GTG TGA ACG GAT TTG-3' 5'-GGG GTC GTT GAT GGC AAC A-3'	95	62.6 62.6
mKRT8	Keratin 8	5'-TCC ATC AGG GTG ACT CAG AAA-3' 5'-AAG GGG CTC AAC AGG CTC T-3'	242	60.1 60.0
mKRT14	Keratin 14	5'-GGC TGG AGC AGG AGA TCG CCA-3' 5'-AGG ACC TGC TCG TGG GTG GAG ACCA-3'	90	61.0 62.0
mKRT19	Keratin 19	5'-GGG GGT TCA GTA CGC ATT GG-3' 5'-GAG GAC GAG GTC ACG AAG C-3'	113	62.9 62.1
mAFP	Alpha fetoprotein	5'-CCA TCA CCT TTA CCC AGT TTG T-3' 5'-CCC ATC GCC AGA GTT TTT CTT-3'	101	60.2 60.6
m1AAT	Alpha-1-antitrypsin	5'-CTC GTC CGC TCA CTA AAC AAG-3' 5'-GCT GTC TGA GAG TCA AGG TCT T-3'	248	60.7 61.3
mTAT	Tyrosine aminotransferase	5'-AGC CGA ATC CGA ACA AAA CC-3' 5'-GCC GAT AGA TGG GGC ATA GC-3'	146	60.9 61.3
mPEPCK	Phosphoenolpyruvate carboxykinase 1	5'-TGA CAG ACT CGC CCT ATG TG-3' 5'-CCC AGT TGT TGA CCA AAG GC-3'	153	61.0 61.4
mALB	Albumin	5'-TGC TTT TTC CAG GGG TGT GTT-3' 5'-TTA CTT CCT GCA CTA ATT TGG CA-3'	167	62.4 60.2
mCYP3A13	Cytochrome P450, family 3, subfamily a, polypeptide 13	5'-ATG AGG CAG GGA TTA GGA GAA G-3' 5'-TGA GAG GAA CAG TGG ATC AAA GA-3'	189	60.7 60.7
mIns2	Insulin-2 preproprotein	5'-GCT TCT TCT ACA CAC CCA TGT C-3' 5'-AGC ACT GAT CTA CAA TGC CAC-3'	147	60.6 60.1
mAmy	Amylase	5'-AAC GAA AGA GAA ATT GAA ACC-3' 5'-GCC CCC ACT CCA CAC ATG TGG-3'	213	60.0 62.0

tral pH values, incubated in the presence of the primary anti-BrdU antibodies in PBS (1:1000 dilution) for 60 min at 37˚C, and subsequently washed 3 times for 10 min using PBS at 37˚C, followed by incubation of the cells in the presence of secondary antibodies (1:1000 dilution) for 60 min at 37˚C. The cells were then washed again 3 times for 10 min in PBS at 37˚C; DAPI (Sigma) was added during the final washing phase. The cells were analyzed using a fluorescent microscope, 5,000 cells were counted for the statistical analysis.

In order to determine the fraction of proliferating cells after 10 days of cultivation under 3D conditions, the cells in gel were incubated in the presence of 10 μM BrdU for 15 h, followed by gel washing with PBS, minced, and incubated in the presence of type II collagenase (Sigma) for 60 min at 37˚C. Following the gel digestion , the cells were washed with PBS, plated onto dishes coated with type I collagen, and cultivated under the standard conditions in the CO_2 incubator. After the cells were attached, a staining procedure identical to that carried out under 2D conditions was performed (see above).

The gel was stained with anti-BrdU antibodies to analyze the features of cell growth and to reveal morphogenetic features under 3D conditions. After the cells were cultivated in the gel for 10 days, 10 μM BrdU was added to the medium for 15 h; the gel was subsequently fixed with 4% paraformaldehyde for 10 min at room temperature. The gel was then incubated in 70% ethanol (30 min, +4˚C), an equivalent volume of 4N HCl was added, and incubation at room temperature for 15 min was performed. The gel was washed using PBS to obtain neutral pH values and incubated in the presence of the primary anti-BrdU antibodies in PBS (1:1000 dilution) for 16 h on a shaker at room temperature. The gel was subsequently washed 3 times for 10 min using PBS on a shaker at room temperature, followed by incubation of the cells in the presence of the secondary antibodies (1:1000 dilution) for 2 h on a shaker at room temperature. The cells were then washed 3 times for 10 min in PBS at 37˚C, and DAPI (Sigma) was added during the last wash. The cells were analyzed on a fluorescent microscope.

Isolation of total RNA from cells

Total RNA was isolated from the cells during the first passage, when cultivation occurred under 2D conditions and after 10 days of incubation in gel, when cells were cultivated under 3D conditions. The gel was digested using type II collagenase to isolate RNA from

Table 3. Primers used in real-time PCR

Primer	Gene	Nucleotide sequence	Amplicon, bp	Melting temperature, °C
mGAPDH	Glyceraldehyde-3-phosphate	5'-AGG TCG GTG TGA ACG GAT TTG-3' 5'-GGG GTC GTT GAT GGC AAC A-3'	95	62.6 62.6
mKRT19	Keratin 19	5'-GGG GGT TCA GTA CGC ATT GG-3' 5'-GAG GAC GAG GTC ACG AAG C-3'	113	62.9 62.1
mAFP	Alpha fetoprotein	5'-CCA TCA CCT TTA CCC AGT TTG T-3' 5'-CCC ATC GCC AGA GTT TTT CTT-3'	101	60.2 60.6
m1AAT	Alpha-1-antitrypsin	5'-CTC GTC CGC TCA CTA AAC AAG-3' 5'-GCT GTC TGA GAG TCA AGG TCT T-3'	248	60.7 61.3
mTAT	Tyrosine aminotransferase	5'-AGC CGA ATC CGA ACA AAA CC-3' 5'-GCC GAT AGA TGG GGC ATA GC-3'	146	60.9 61.3
mPEPCK	Phosphoenolpyruvate carboxykinase 1	5'-TGA CAG ACT CGC CCT ATG TG-3' 5'-CCC AGT TGT TGA CCA AAG GC-3'	153	61.0 61.4

the cells cultivated under 3D conditions; the cells were precipitated using centrifugation. RNA was isolated using AllPrep DNA/RNA Mini Kit (Qiagen) in accordance with the manufacturer's recommendations. RNA concentration was determined using a Qubit mini-fluorometer and the RNA Assay Kit (Invitrogen). Reverse transcriptase Superscript II (Invitrogen) and random primers were used for reverse transcription. 500 ng of total RNA was used for the reaction.

RT-PCR analysis of mouse salivary gland cells and mouse progenitor liver cells

The RT-PCR was carried out using the ScreenMix kit (Eurogen) in accordance with the manufacturer's recommendations. The reaction conditions were as follows: pre-incubation at 95°C for 5 min to activate DNA polymerase, followed by 25–30 cycles: denaturation at 95°C for 15 s; annealing at 57–59°C for 15 s; and elongation at 72°C for 1 min. *Table 2* lists the markers and the melting temperatures of the primers used in the experiment.

Agarose gel electrophoresis

Electrophoresis was carried out in 1.5% agarose gel (Helicon) and a TAE buffer (PanEco) at a voltage of 80 V. DNA Ladder (Promega) was used as molecular weight markers in 1 kb and 100 bp increments. The probe volume was 6 μl per well. The gel was analyzed under UV light (360 nm) after staining with ethidium bromide (Sigma).

Quantitative PCR

Quantitative real-time PCR was carried out using a real-time PCR kit and an EVA Green stain (Sintol) on a CFX96 real-time PCR instrument (BioRad). The reaction conditions were as follows: pre-incubation at 95°C for 5 min to activate DNA polymerase, followed by 40 cycles: denaturation at 95°C for 30 s; annealing at 57–59°C for 30 s; and elongation at 72°C for 45 s. The annealing temperature varied slightly for different genes, depending on the melting temperatures of the primers (*Table 3*). Fluorescence was determined in the Fam channel; the initial analysis of the results was carried out automatically using software supplied with the device. *GAPDH* mRNA was used as an internal standard relative to which the concentrations of the other mRNAs were determined. The cDNA samples of each investigated gene were (where possible) analyzed simultaneously and in parallel in the adjacent wells of the device under strictly identical conditions.

Determination of the rate of urea production by cells

The rate of urea production by cells was determined using the Urea Assay Kit (BioVision) in accordance with the manufacturer's recommendations. The amount of urea was measured in the culture medium, and the old medium was replaced with a fresh one 24 h prior to sampling. The cells cultivated under 2D conditions were analyzed during the first passage; the media samples from cells cultivated under 3D conditions were collected on the 1st, 5th and 10th days of incubation in gel.

Statistical analysis of the data

All experiments were carried out in three repeats using cell cultures obtained from three different animals. Each procedure was performed under identical conditions in three technical repeats. The statistical analysis was performed using a Student's t-test at a 95% confidence interval for the biological repeats and 99% confidence interval for the technical repeats.

RESULTS AND DISCUSSION

Morphological characteristics of mouse liver and mouse salivary gland cells cultivated under 2D and 3D conditions

Following the isolation, the PLPCs were attached to the plastic coated with type I collagen for 1-2 days; the PSGCs were attached for 2-3 days. The cells were polygonal and mononuclear; they were characterized by a small size and high nucleocytoplasmic ratio. The formation of dense colonies followed. The PSGC monolayer was formed on the 5th day, and the PLPC monolayer was formed on the 7th day (*Fig. 1A,B*). At this stage, the morphology of the cells isolated from the liver and the salivary gland was almost identical. The cell population doubling time was maximal during the 0 passage: approximately 35 h for PSGCs and 50 h for PLPCs. Following the monolayer formation and during the subsequent cultivation, the cell population doubling time was stabilized around 42 h for PSGCs and 63 h for PLPCs. Both cell cultures could undergo over 20 passages, which indicates that they have a high proliferative potential and generally consist of undifferentiated cells.

During the first passage, PSGCs and PLPCs were incorporated into the 4% collagen gel at a concentration of 1 × 10^6 cells/ml of gel. Morphological changes were observed during the subsequent 10 days of cell incubation in the gel. PSGCs became elongated and formed clusters increasing in size with the lapse of time. Bundles consisting of several dozen cells radiated from the clusters (*Fig. 1C*). The bundles of the PSGC cells in the gel had a tubular structure as can be seen in the paraffin sections (*Fig. 1E*). The PLPCs also formed clusters in the form of small bundles, although the emergence of large clusters was less pronounced (*Fig. 1D*). These structures are not hollow as can be seen in the sections (*Fig. 1F*). Some PLPCs retained their round shape, became larger, and contained many granules under 3D conditions.

A decrease in the collagen gel size (contraction) occurs as the cellular bundles grow. It was established that the degree of contraction depends on the cytoskeleton of the cells and reflects their contractility [19]. The investigated cells acquired contractile capabilities during the differentiation in myoepithelium. Hence, the myoepithelial differentiation potential of the PSGCs and PLPCs can be determined from the degree of contraction of the collagen gel.

Contraction of the collagen gel was observed during the entire period of incubation of PSGCs and PLPCs under 3D conditions. PSGCs cause a significant contraction, as soon as after 5 days of cell incubation in the gel, its area decreases to 14% of the initial size (*Fig. 2*). Meanwhile, PLPCs contract the gel to a lesser extent:

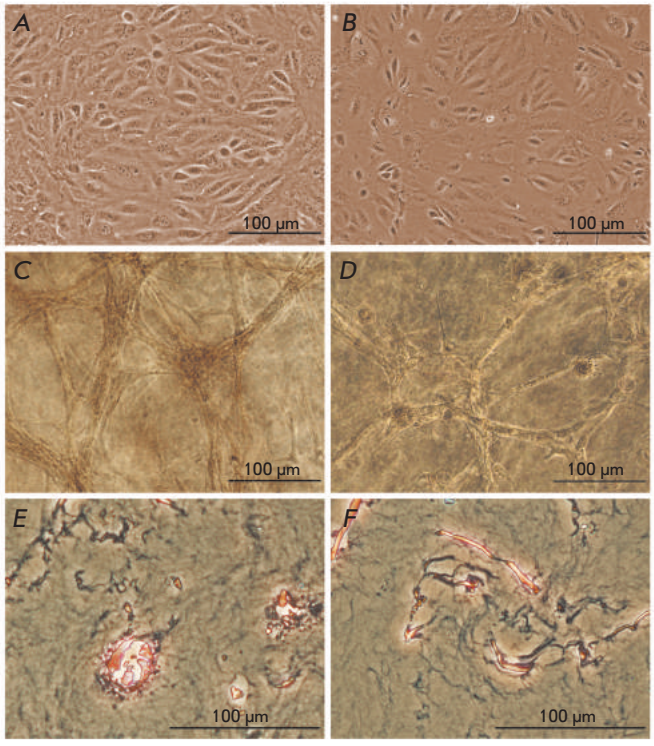


Fig. 1. The morphology of salivary gland and liver progenitor cells under 2D and 3D cultivation conditions, phase contrast microscopy. A) PSGC, monolayer culture, 0 passage; B) PLPC, monolayer culture, 0 passage; C) PSGC, 10-th day of cultivation in 4% collagen gel, 1 passage; D) PLPC, 10th day of cultivation in 4% collagen gel, 1 passage; E) PSGC, histological section of collagen gel, 10th day of cultivation, azure-eosin staining; F) PLPC, histological section of collagen gel, 10th day of cultivation, azure-eosin staining

the gel area was more than 30% of its initial size on the 10th day of incubation (*Fig. 2*). Therefore, the PSGCs have a higher myoepithelial differentiation potential, which is in agreement with published data [24].

Our data indicate that there is a similarity between the morphological characteristics of PSGCs and PLPCs during cultivation under 2D conditions. However, the morphogenetic characteristics of cells differ significantly under 3D conditions.

Immunocytochemical analysis of mouse liver and mouse salivary gland cells cultures under 2D conditions and after cultivation in collagen gel

In order to determine the effect of 3D cultivation conditions on the expression of hepatic markers, immunophenotyping of PSGCs and PLPCs cultured on

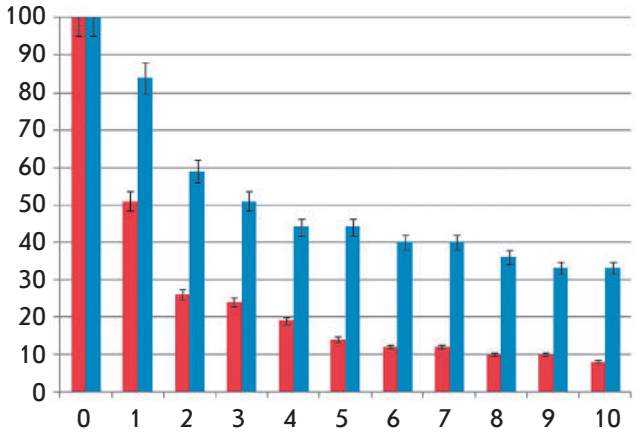


Fig. 2. 4% collagen gel contraction by salivary gland and liver progenitor cells, cell concentrations are 1 x 10^6 per ml of gel, 1 passage. X-axis value is days of contraction, Y-axis value is gel area in relation to the initial gel area (%). Red columns relate to PSGC; blue columns, to PLPC

plastic and after incubation for 10 days in a collagen gel was carried out. Immunophenotyping of endodermal cells was performed during the first passage using the markers listed in *Table 1*.

During cultivation under 2D conditions, the cells in both cultures are weakly positive with respect to albumin and hepatocyte-specific cytochrome P450 1A1 (*Fig. 3A, B, C, D*). Staining for cytokeratin 19 is typical of both cultures (*Fig. 3C,F*); however, cytokeratin localization differs for PSGCs and PLPCs. Cytokeratin 19 localizes in the perinuclear space and under the plasma membrane in PSGCs, which can presumably be attributed to the developed system of tight junctions in these cells as they have a barrier function. In PLPCs, cytokeratin 19 localizes predominantly in the perinuclear space. Both cultures contain a small number of cells in which cytokeratin 19 is distributed throughout the entire cytoplasm.

During the analysis of the cells cultivated for 10 days under 3D conditions, the gel was digested using type II collagenase. The cells were then plated onto dishes coated with type I collagen and cultivated under the standard conditions (CO_2 incubator for 48 h), followed by cellular immunophenotyping. The expression level of cytochrome P450 1A1 increased due to cultivation of PSGCs in the collagen gel (*Fig. 3H*), whereas the expression level of albumin remained intact. The expression of cytokeratin 19 decreased, and its localization changed: in most cells, cytokeratin 19 localized in the perinuclear space (*Fig. 3I*). Following the incubation in the collagen gel, the expression of albumin and cytochrome P450 increased, while the expression of cytokeratin 19 decreased in PLPCs (*Fig. 3J–L*).

Hence, the investigated cell cultures are characterized by a similar expression pattern of hepatic markers during cultivation under 2D conditions. The immunocytochemical analysis data indicate that hepatic markers are expressed in both cultures. However, their level of expression in the investigated endodermal cells is low; which attests to the fact that PSGCs and PLPCs are in an undifferentiated state. Both PSGCs and PLPCs express markers typical of the ductal (cytokeratin 19) and the hepatic (albumin, cytochrome P450 1A1) lineage; i.e., they have a bipotent differentiation potential, which is usually observed in oval cells (e.g., [25]). Following cultivation under 3D conditions, the expression of the hepatic differentiation markers increases, which is more evident in the case of PLPCs. The expression of ductal differentiation markers (cytokeratin 19) decreases in both cultures.

Determination of the proliferative capacity of mouse salivary gland and mouse liver progenitor cells under 2D and 3D conditions

The analysis of the proliferative capacity of cells based on the determination of BrdU incorporation demonstrated that PSGCs have a high proliferative potential. During the first passage of cultivation under 2D conditions, BrdU was incorporated into more than 90% of the salivary gland cells (*Fig. 4A*) and 30% of PLPCs (*Fig. 4D*). The PLPC population appeared to be more heterogeneous and to contain cells of variable levels of differentiation.

On the 10^{th} day of cultivation in the collagen gel, BrdU was incorporated only into 52% of PSGCs and 11.5% of PLPCs (*Fig. 4B,E*). Hence, the cellular proliferation of both cultures under 3D conditions slows down approximately twofold. In order to determine the features of cellular growth in gel and the morphogenetic features under 3D conditions, BrdU staining was carried out without isolation of cells from the gel. As a result, no specific patterns in the distribution of proliferating cells were identified: BrdU-positive cells localized both in the outer and inner layers along the entire length of the cellular bundles (*Fig. 4C,F*).

RT-PCR for mouse salivary gland and mouse progenitor liver cells cultures under 2D and 3D cultivation conditions

The RT-PCR analysis of the PSGC and PLPC was carried out during the first passage under 2D cultivation conditions and after cell incubation in gel under 3D conditions for 10 days. RT-PCR was performed for a wide range of markers specific to endoderm cells. The

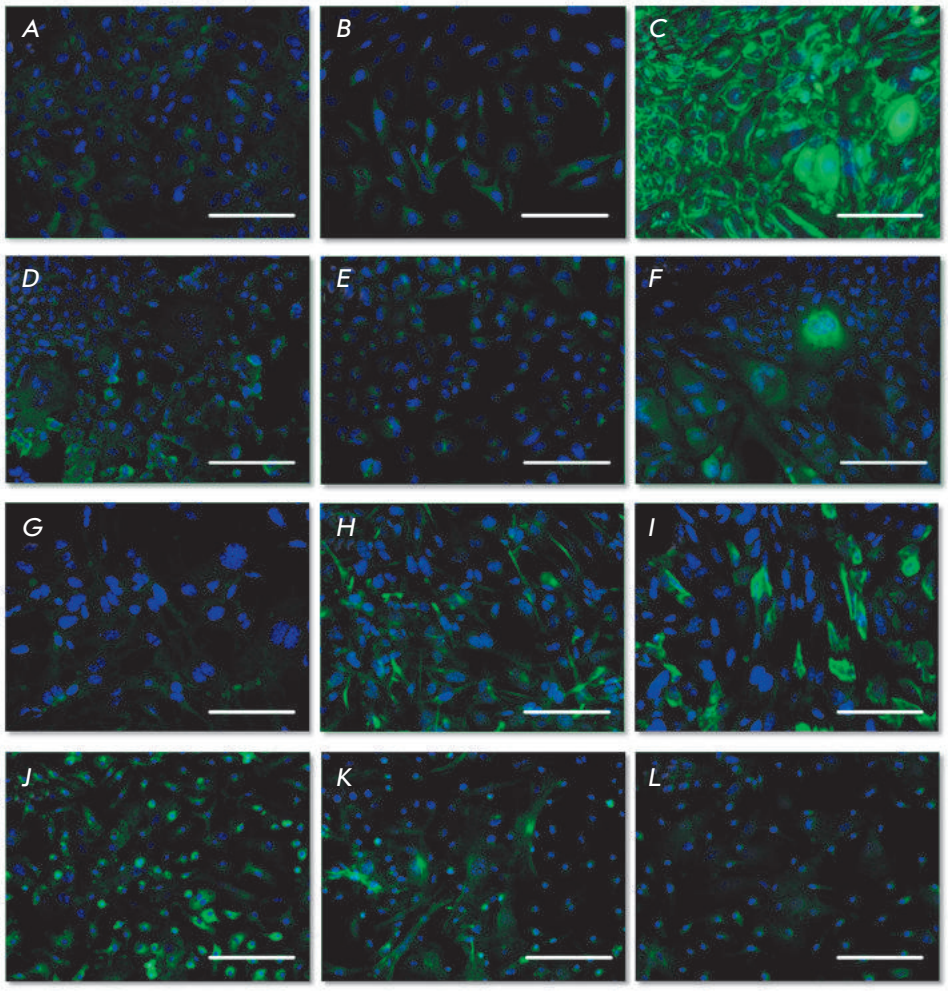


Fig. 3. Immunocytochemistry of salivary gland and liver progenitor cells on the 1st passage, fluorescent microscopy. Cell nuclei stained by DAPI (blue color), the antigens stained by Alexa Fluor 488 (green color), bar equals to 100 microns. A) PSGC, albumin, 2D conditions; B) PSGC, cytochrome P450 1A1, 2D conditions; C) PSGC, cytokeratin 19, 2D conditions; D) PLPC, albumin, 2D conditions; E) PLPC, cytochrome P450 1A1, 2D conditions; F) PLPC, cytokeratin 19, 2D conditions; G) PSGC, albumin, after cell cultivation for 10 days in the collagen gel; H) PSGC, cytochrome P450 1A1, after cell cultivation for 10 days in the collagen gel; I) PSGC, cytokeratin 19, after cell cultivation for 10 days in the collagen gel; J) PLPC, albumin, after cell cultivation for 10 days in the collagen gel; K) PLPC, cytochrome P450 1A1, after cell cultivation for 10 days in the collagen gel; L) PLPC, cytokeratin 19, after cell cultivation for 10 days in the collagen gel

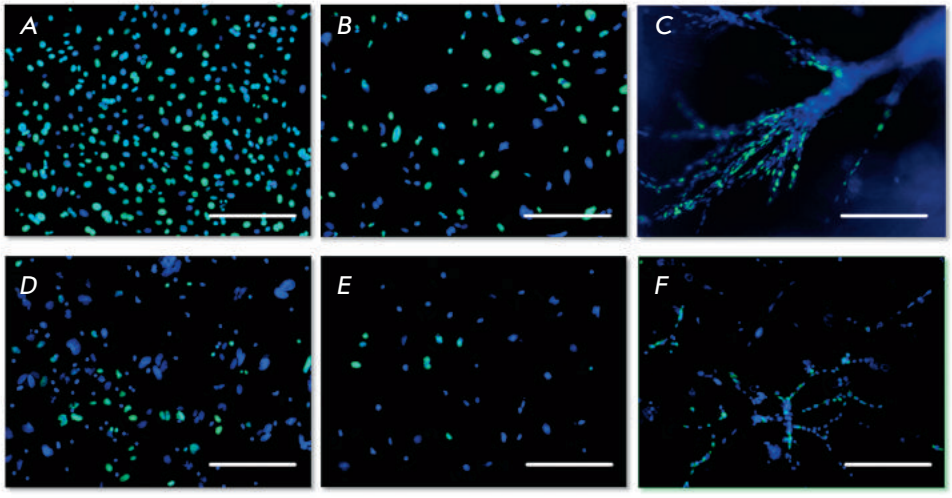


Fig. 4. Analysis of the proliferative activity of salivary gland and liver progenitor cells under 2D and 3D cultivation conditions, 1st passage, fluorescent microscopy. Cell nuclei stained by DAPI (blue color), BrdU stained by Alexa Fluor 488 (green color), bar is equal to 100 μm. A) PSGC, 2D conditions; B) PSGC isolated from the collagen gel after cultivation for 10 days under 3D conditions; C) PSGC, 10th day of cultivation in the collagen gel (without isolation); D) PLPC, 2D conditions; E) PLPC isolated from the collagen gel after cultivation for 10 days under 3D conditions; F) PLPC, 10th day of cultivation in the collagen gel (without isolation)

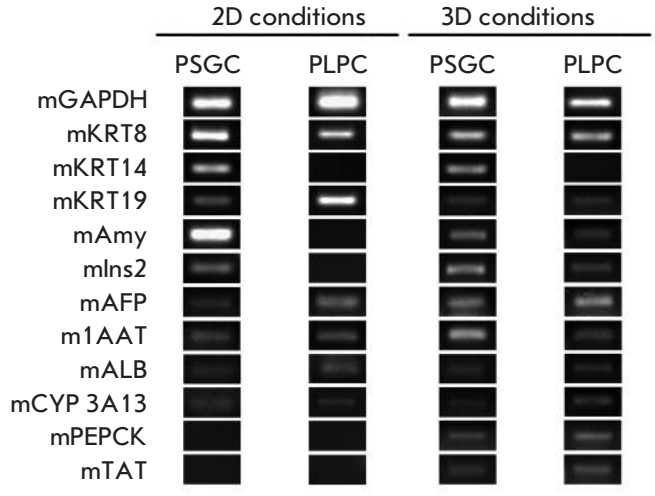


Fig. 5. RT-PCR analysis of salivary gland and liver progenitor cells on the 1st passage under 2D and 3D cultivation conditions

Table 4. RT-PCR for PSGCs and PLPCs during the 1st passage under 2D and 3D cultivation conditions*

Primer	2D conditions		3D conditions, 10th day
	PSGC	PLPC	PSGC
mGAPDH	1	1	1
mKRT19	14.1	295.59	7.9
mAFP	0.01	0.12	1.19
m1AAT	0.04	0.30	3.19
mTAT	0.09	0.43	0.32
mPEPCK	0.02	0.08	0.18

* The data have been standardized with respect to GAPDH

housekeeping gene *GAPDH* was used as a positive control during PCR.

According to the RT-PCR data, both PSGCs and PLPCs express cytokeratins 8, 18, 19, specific to endodermal epithelial cells during the first passage under 2D conditions. Moreover, PSGCs express cytokeratin 14, specific to the cells associated with the basement membrane, and also express amylase and insulin (*Fig. 5*), whereas PLPCs are negative with respect to these markers. As for the hepatocyte-specific markers, both PSGCs and PLPCs express α-fetoprotein and α-1-antitrypsin, albumin and cytochrome P450 3A13. The markers of the later stages of hepatocyte differentiation (*PEPCK, TAT*) were not detected under 2D cultivation conditions.

The expression of *PEPCK* and *TAT,* which are typical of the later stages of hepatic differentiation, was detected in PSGCs and PLPCs after cultivation in the collagen gel for 10 days. Furthermore, PLPCs are characterized by expression of amylase and insulin, which are the cell markers of pancreatic differentiation. This phenotypic plasticity of liver progenitor cells is in close agreement with the published data, according to which liver and pancreatic stem cells can transdifferentiate into each other under certain *in vitro* cultivation conditions [2, 26].

Quantitative Real-Time PCR for salivary gland cells cultures compared to liver progenitor cells under 2D and 3D cultivation conditions

The comparative analysis of the PSGC and PLPC cells cultures was carried out during the first passage using quantitative real-time PCR for the markers specific to liver cells (*Table 4*). In addition, the expression of these genes in PSGCs was analyzed after cultivation in the collagen gel for 10 days. The data for each culture was normalized with respect to *GAPDH*, the level of expression of which was assumed to be equal to 1.

During the first passage under 2D conditions, PSGCs express cytokeratin 19 at a relatively high level; however, its expression level in PLPCs was over 18 times higher. After the incubation in the gel, the expression of cytokeratin 19 in salivary gland cells dropped twofold. The level of expression of α-fetoprotein specific to the oval cells in both cultures was relatively low, but its expression was 10 times higher in PLPCs. The hepatic markers were expressed in the investigated cells at a relatively low level; the expression of these markers was several times higher in PLPCs. The level of α-fetoprotein expression in PSGCs increased by over 100 times after the cell incubation in gel for 10 days and significantly exceeded its expression level in PLPCs under 2D cultivation conditions. The level of expression of α-1-antitrypsin in PSGCs specific to the initial stages of hepatic differentiation increased 80 times under 3D conditions. The level of expression of the markers specific for the later stages of hepatic differentiation significantly increased as well.

Determination of the rate of urea production by mouse salivary gland cells and mouse progenitor liver cells under 2D and 3D cultivation conditions

The rate of urea production by PSGCs and PLPCs under 2D and 3D cultivation conditions was analyzed to

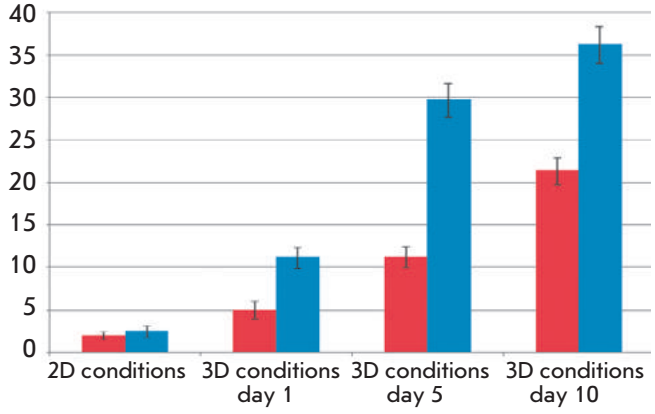


Fig. 6. Analysis of urea synthesis by salivary gland and progenitor liver cells under 2D and 3D cultivation conditions, 1st passage. The Y-axis value is amount of urea (mM) per 1 x 10^6 cells per 24 h. Red columns relate to PSGC; blue columns, to PLPC

determine the functional activity of the cells under investigation.

Both cultures of endodermal cells producted virtually no urea under 2D conditions: ~2 mM of urea per 1 × 10^6 of PSGCs during 24 h was detected during the first passage, and ~2.5 mM of urea was detected for PLPCs (*Fig. 6*). This value slightly decreases as the number of passages increases (data not shown). However, the level of urea production by both PLPCs and PSGCs gradually increases during cultivation under 3D conditions. On the 10^{th} day of cultivation, the rate of urea production was 15 times higher for PLPCs and 10 times higher for PSGCs as compared to the initial levels. Hence, the level of urea production by PSGCs was equal to ~60% of that by PLPCs after cultivation for 10 days under 3D conditions.

Liver and pancreatic stem cells are known to be capable of undergoing transdifferentiation within the endodermal germ layer. Thus, the ductal cells isolated from the pancreas and transplanted into the liver differentiate into hepatocytes [27]. Oval cells can differentiate into endocrine and exocrine pancreatic cells [26]. Pancreatic islet cells can differentiate into hepatocytes in culture with increased plating density. It has been demonstrated that pancreatic acinar cells can differentiate into hepatocytes when exposed to dexamethasone [2]. However, very little data is available about the differentiation potential of salivary gland cells. A fivefold increase in albumin expression is observed in cell cultures derived from pig salivary glands after such cells are stimulated with nicotinamide [18]. Mouse salivary gland cells transplanted into the liver through the portal vein are capable of engrafting the liver and producing albumin and α-1-antitrypsin [16].

Generally speaking, our results have confirmed the high phenotypic flexibility of mouse salivary gland cells within the endodermal germ layer. PSGCs undergo significant and specific hepatic differentiation in the collagen gel without any additional stimulation with growth factors and cytokines.

CONCLUSIONS

A mouse postnatal submandibular salivary gland cell culture was obtained and compared to mouse postnatal progenitor liver cells under 2D and 3D cultivation conditions.

Mouse postnatal salivary gland cells and mouse progenitor liver cells are endodermal epithelial cells. PSGCs are ductal cells characterized by a high proliferative potential. The PLPC population is heterogeneous and contains cells at various differentiation stages. Both cultures possess the bipotent potential of hepatic and ductal differentiation. Moreover, PSGCs are capable of pancreatic differentiation, while PLPCs acquire this ability during cultivation in a collagen gel.

PSGCs and PLPCs are considerably similar in terms of the expression of various cellular markers and seem to possess similar differential potentials. In general, PSGCs and PLPCs are characterized by a significant phenotypic flexibility and the ability to undergo transdifferentiation within the endodermal germ layer.

Under 3D cultivation conditions, PSGCs and PLPCs undergo differentiation, which is characterized by a slowdown of cellular proliferation and an increase in the expression level of differentiation markers. Under 3D conditions, PSGCs are characterized by a decrease of the expression of ductal markers and an increase in the expression of hepatic markers in a similar degree with PLPCs.

Cell culturing in a collagen gel is a convenient model of *in vitro* analysis of the cellular differential potential. During cultivation in a collagen gel, postnatal salivary gland cells undergo hepatic differentiation in the absence of any additional stimulation by cytokines and growth factors. Hence, the investigated postnatal salivary gland cells can undergo hepatic transdifferentiation and become a convenient source of cells for the cellular therapy of liver pathologies. ●

This work was supported by the Russian Foundation for Basic Research (project № 11-04-12061-ofi-M-2011).

REFERENCES

1. Uryvaeva I.V. Stem cells in liver regeneration. The stem cells biology and cellular technologies. M.: Medicine, 2009. V. 2. 456 P.
2. Zaret, K.S. // Nat. Rev. Genet. 2008. V. 9. P. 329–340.
3. Rambhatla L., Chiu C.P., Kundu P., Peng Y., Carpenter M.K. // Cell Transplant. 2003. V. 12. № 1. P. 1–11.
4. Hay D.C., Fletcher J., Payne C., Terrace J.D., Gallagher R.C., Snoeys J., Black J.R., Wojtacha D., Samuel K., Hannoun Z., *et al.* // Proc. Natl. Acad. Sci. USA. 2008. V. 26. № 105. P. 12301–12306.
5. Soto-Gutierrez A., Navarro-Alvarez N., Caballero-Corbalan J., Tanaka N., Kobayashi N. // Acta Med. Okayama. 2008. V. 62. P. 63–68.
6. Mizumoto H., Aoki K., Nakazawa K., Ijima H., Funatsu K., Kajiwara T. // Transplant. Proc. 2008. V. 40. № 2. P. 611–613.
7. Snykers S., Vanhaecke T., De Becker A., Papeleu P., Vinken M., Van Riet I., Rogiers V. // BMC Dev. Biol. 2007. V. 2. № 7. P. 24.
8. De Kock J., Vanhaecke T., Rogiers V., Snykers S. // Aatex. 2008. V. 14. P. 605–611.
9. Gumerova A.A., Shafigullina A.K., Trondin A.A., Gazizov I.M., Andreeva D.I., Kaligin M.S., Rizvanov A.A., Kiasov A.P. // Cell Transplantation and Tissue Engineering. 2011. V. 6(4). P. 72–81.
10. Sgodda M., Aurich H., Kleist S., Aurich I., Konig S., Dollinger M.M., Fleig W.E., Christ B. // Exp. Cell Res. 2007. V. 313. P. 2875–2886.
11. Seo M.J., Suh S.Y., Bae Y.C., Jung J.S. // Biochem. Biophys. Res. Commun. 2005. V. 328. P. 258–264.
12. Stock P., Staege M.S., Muller L.P., Sgodda M., Volker A., Volkmer I., Lutzkendorf J., Christ B. // Transplant. Proc. 2008. V. 40. P. 620–623.
13. Dawn M.D., De Coppi P., Bartsch G., Atala A. // Meth. Enzymology. 2006. V. 419. P. 426–438.
14. Zheng Y.B., Gao Z.L., Xie C., Zhu H.P., Peng L., Chen J.H., Chong Y.T. // Cell Biol. Internat. 2008. V. 32. № 11. P. 1439–1448.
15. Gvazava I.G., Vasilev A.V., Balan O.V., Terskikh V.V. // Tsitologiia. 2011. V. 53. № 2. P.129–134.
16. Hisatomi Y., Okumura K., Nakamura K., Matsumoto S., Satoh A., Nagano K., Yamamoto T., Endo F. // Hepatology. 2004. V. 39. P. 667–675.
17. Sato A., Okumura K., Matsumoto S., Hattori K., Hattori S., Shinohara M., Endo F. // Cloning Stem. Cells. 2007. V. 9. P. 191–205.
18. Matsumoto S., Okumura K., Ogata A., Hisatomi Y., Sato A., Hattori K., Matsumoto M., Kaji Y., Takahashi M., Yamamoto T., *et al.* // Cloning Stem Cells. 2007. V. 9. P. 176–190.
19. Davydova D.A., Voroteliak E.A., Bragina E.E., Terskikh V.V., Vasil'ev A.V. // Tsitologiia. 2011. V. 53(4). P. 325–31.
20. Shinin V.V., Chernaia O.G., Terskikh V.V. // Ontogenez. 2002. V. 33(3). P. 176–81.
21. Voroteliak E.A., Leonova O.G., Shinin V.V., Vasil'ev A.V., Terskikh V.V. // Dokl. Akad. Nauk. 1999. V. 369(5). P. 695–697.
22. 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.
23. Zhang Y., Jalili R.B., Warnock G.L., Ao Z., Marzban L., Ghahary A. // The Am. J. Pathol. 2012. V. 181. № 4. P. 1296–1305.
24. Babaeva A.G., Shubnikova E.A. Structure, function and adaptive growth of salivary glands. M.: MSU. 1979. P 192.
25. Duncan A.W., Dorrell C., Grompe M. // Gastroenterology. 2009. V. 137(2). P. 466–481.
26. Reddy J.K., Rao M.S., Yeldandi A.V., Tan X.D., Dwivedi R.S. // Digestive Diseases Sci. 1991. V. 36(4). P. 502–509.
27. Dabeva M.D., Hwang S.G., Vasa S.R., Hurston E., Novikoff P.M., Hixson D.C., Gupta S., Shafritz D.A. // Proc. Natl. Acad. Sci. USA. 1997. V. 8. № 94. P. 7356–7361.

N-Terminal Fusion Tags for Effective Production of G-Protein-Coupled Receptors in Bacterial Cell-Free Systems

E.N. Lyukmanova[1,*], Z.O. Shenkarev[1], N.F. Khabibullina[1,2], D.S. Kulbatskiy[1], M.A. Shulepko[1,2], L.E. Petrovskaya[1], A.S. Arseniev[1], D.A. Dolgikh[1,2], M.P. Kirpichnikov[1,2]

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

[2]Faculty of Biology, Lomonosov Moscow State University, Leninskie Gory 1/12, Moscow, Russia, 119991

*E-mail: ekaterina-lyukmanova@yandex.ru

ABSTRACT **G-protein-coupled receptors (GPCR) constitute one of the biggest families of membrane proteins. In spite of the fact that they are highly relevant to pharmacy, they have remained poorly explored. One of the main bottlenecks encountered in structural-functional studies of GPCRs is the difficulty to produce sufficient amounts of the proteins. Cell-free systems based on bacterial extracts from *E. coli* cells attract much attention as an effective tool for recombinant production of membrane proteins. GPCR production in bacterial cell-free expression systems is often inefficient because of the problems associated with the low efficiency of the translation initiation process. This problem could be resolved if GPCRs were expressed in the form of hybrid proteins with N-terminal polypeptide fusion tags. In the present work, three new N-terminal fusion tags are proposed for cell-free production of the human β2-adrenergic receptor, human M1 muscarinic acetylcholine receptor, and human somatostatin receptor type 5. It is demonstrated that the application of an N-terminal fragment (6 a.a.) of bacteriorhodopsin from *Exiguobacterium sibiricum* (ESR-tag), N-terminal fragment (16 a.o.) of RNAse A (S-tag), and Mistic protein from *B. subtilis* allows to increase the CF synthesis of the target GPCRs by 5–38 times, resulting in yields of 0.6–3.8 mg from 1 ml of the reaction mixture, which is sufficient for structural-functional studies.**

KEYWORDS **Cell-free expression; GPCR; translation initiation.**

ABBREVIATIONS **a.a. – amino acid residue; β2AR – human β2-adrenergic receptor; CF system – cell-free expression system; ESR – bacteriorhodopsin from *Exiguobacterium sibiricum*; ESR-tag – 6 a.a. N-terminal fragment of the ESR; FM – feeding mixture; GPCR – G-protein-coupled receptor; M1-mAChR – human muscarinic acetylcholine receptor M1; MP – membrane protein; RM – reaction mixture; SSTR5 – human somatostatin receptor type 5; S-tag – 16 a.a. N-terminal fragment of ribonuclease A; T7-tag – 11 a.a. N-terminal fragment of the leader peptide of the protein 10 of the bacteriophage T7; TM – transmembrane; TRX – thioredoxin from *Escherichia coli*.**

INTRODUCTION

Integrated membrane proteins (MPs) participate in a number of processes essential for single-cell and metazoan organisms. These proteins are responsible for cellular energetics, intercellular recognition, signal transduction, and transport of various substances through the cell membrane [1]. Recent data indicate that MPs make up over 25% of all amino acid sequences in the genomes of higher organisms, including the human genome. G-protein-coupled receptors (GPCR) are among the most pharmacologically important MP classes. Over 800 GPCR genes have been identified in the human genome [3], and membrane receptors of this class are the targets of ~30% of modern drugs [4]. GPCRs are characterized by homological spatial organization and contain seven transmembrane (TM) helices, as well as the extracellular N- and intracellular C-terminal regions [5]. The binding sites of low-molecular-weight ligands localize in the TM domain of the receptor, whereas peptide hormones and regulatory proteins interact with the N-terminal region and extracellular loops [5].

GPCRs are of particular interest for pharmacological research; however, structural and functional investigations of these receptors are complicated [5] because of the infeasibility of isolating a sufficient amount of the protein from natural sources and the problems concerned with designing high-performance systems to heterologously produce these MPs [6]. Over the last decade, the joint use of expression systems based on eukaryotic cells and new methods of X-ray structure analysis has enabled to determine the spatial structure of a series of GPCRs [5], including the human β2-

adrenoreceptor (β2AR) [7] and human muscarinic M2 and M3 cholinoreceptors (mAChR) [8, 9]. These studies have led to a better understanding of the principles of the spatial organization of GPCR. However, a thorough investigation into the functional dynamics and mechanisms of membrane receptor functioning requires the use of high-resolution spectroscopic methods, such as heteronuclear NMR spectroscopy [10]. The current NMR spectroscopy methods require milligram amounts of protein samples labeled with stable isotopes (^{2}H, ^{13}C, ^{15}N) [10], which are expensive when eukaryotic systems are used. Meanwhile, the use of conventional bacterial expression systems for GPCR production often does not allow to achieve high yields of the target protein and is complicated due to the necessity to develop re-naturation protocols [11].

Cell-free (CF) expression systems [12], and in particular those based on bacterial extracts, have recently gained increasing popularity as an alternative tool for the recombinant production of MPs [13]. As compared with the systems based on cell production, CF systems have a number of advantages, including exclusive production of the target protein, the possibility to synthesize toxic proteins, simple procedure for synthesizing selectively isotope-labeled samples, and the possibility of direct introduction of various agents and cofactors to the reaction mixture to stabilize the native spatial structure of the synthesized protein in the solution [12, 13]. Thus, the components of membrane-mimicking media, such as detergent micelles, lipid/detergent bicelles, liposomes, and lipid–protein nanodiscs, can be added to the reaction mixture to produce soluble MPs [13–15].

According to the published data, direct expression of GPCR genes in CF systems is inefficient [14, 16–18]. The low efficiency of the translation initiation process [18], due to the formation of a secondary structure of the 5-prime end mRNA fragment, is among the possible reasons [19, 20]. In most cases, this problem can be solved and the desired level of GPCR production can be attained by inserting additional nucleotide sequences encoding N-terminal polypeptide fusion tags, such as the fragment of the protein 10 leader sequence of bacteriophage T7 (T7-tag, 11 a.a.; hereinafter, the sequence length is given with allowance for the N-terminal Met residue) [14,16], the thioredoxin protein from *E. coli* (TRX) [17], or 1–6 a.a. long synthetic sequences [18] at the 5-prime end of the target protein gene.

Three novel N-terminal fusion tags are proposed in this work in order to increase the efficiency of cell-free production of human GPCR by the example of β2AR, M1-mAChR, and somatostatin receptor type 5 (SSTR5). It is shown that the use of nucleotide sequences encoding the N-terminal fragment (6 a.a.) of bacteriorhodopsin from Gram-positive bacteria *Exiguobacterium sibiricum* (ESR-tag), the N-terminal fragment (16 a.a.) of ribonuclease A (N-terminal fragment of S-peptide, S-tag), and Mistic protein from *Bacillus subtilis* allows to increase the receptor yield by 5–38 times, providing a sufficient level of target protein production for further structural and functional studies.

EXPERIMENTAL

Design and cloning of the *GPCR* genes with additional 5-prime end sequences

Truncated human β2AR, M1-mAChR, and SSTR5 receptor genes with additional substitutions, and 3-prime end sequences encoding the 10 His residues (His10-tag) (see the Results and Discussion section) are used in this study. The molecular weights of the target proteins were 38.2, 32.6, and 32.7 kDa, respectively. Nucleotide sequences encoding the T7-tag (11 a.a., MASMTGGQQMG), S-tag (16 a.a., MKETAAAKFERQHMDS), TRX (11.8 kDa), and Mistic protein (12.8 kDa) were introduced in one reading frame to the 5-prime end of the truncated GPCR genes (*Fig. 1*) using conventional genetic engineering techniques. The nucleotide sequence encoding the ESR-tag (6 a.a., MEEVNL) was introduced to the 5-prime end of the truncated *GPCR* genes to replace the regions encoding N-terminal extracellular fragments of the receptors (see the Results and Discussion section) via single-stage PCR. All these gene constructs were cloned in the pET22b(+) vector (Novagen, USA) under the control of the T7 promoter. The resulting vectors were named *pET22b(+)/GPCR*, *pET22b(+)/T7-tag-GPCR*, *pET22b(+)/S-tag-GPCR*, *pET22b(+)/TRX-GPCR*, *pET22b(+)/Mistic-GPCR*, and *pET22b(+)/ESR-GPCR (Fig. 1)*.

Cell-free production of GPCR

GPCRs were synthesized in the continuous cell-free system based on the *E. coli* S30 extract using protocols [15, 21]. The final concentrations of the components of the reaction mixture were as follows: 100 mM HEPES-KOH (Fluka, USA), pH 8.0; 8 mM $Mg(OAc)_2$, 90 mM KOAc, 20 mM potassium acetyl phosphate (Sigma, USA), 20 mM potassium phosphoenolpyruvate (Aldrich, USA), 1.3 mM of each amino acid, except for Arg, Cys, Met, Trp, Asp, Glu, whose concentrations were 2.3 mM; 0.15 mg/ml folic acid (Sigma), each of four ribonucleoside triphosphates at a concentration of 1 mM; proteinase inhibitor (X1 Complete protease inhibitor®, Roche Diagnostics, Germany); 0.05% of NaN_3; 2% of polyethylene glycol 8000 (Sigma); 0.3 U/μl of ribonuclease inhibitor RiboLock (Fermentas, Lithuania); 0.04 mg/ml of pyruvate kinase (Fermentas, Lithuania); 5.5 μg/ml of T7 polymerase; 0.3 mg/ml of plasmid DNA, 0.5 mg/ml of total tRNA (from *E. coli* MRE 600)

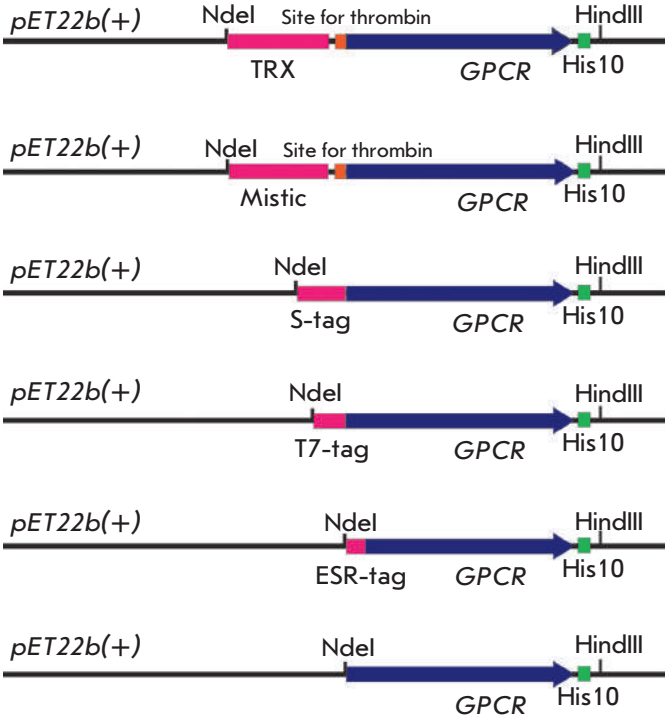


Fig. 1. Design of the vectors containing the GPCR genes and additional 5-prime end sequences. *N*-terminal fusion tags coding sequences, GPCR genes, thrombin sites, and polyhistidine sequences are shown in pink, blue, orange, and green, respectively. Restriction endonuclease sites at which the GPCR genes were cloned into the *pET22b(+)* vector are shown

(Roche Diagnostics, Switzerland), 30% of the total volume of the reaction mixture of the *E. coli* S30 extract. The feeding mixture (FM) had the same composition, except for the high-molecular-weight components: S30 extract, plasmid, enzymes, and ribonuclease inhibitor. The synthesis was carried out without the addition of any membrane-mimicking media in RM and FM. The RM and FM volumes were 50 and 750 μl, respectively. The RM was placed into the reactor separated from the FM solution with a dialysis membrane (pore size 12 kDa, Sigma, USA), followed by incubation for 20 h at 30°C under moderate stirring.

Isolation and purification of GPCR samples

The RMs containing synthesized GPCRs were centrifuged for 15 min at 14000 rpm. The resulting precipitates were solubilized in buffer A (20 mM Tris-HCl, 250 mM NaCl, 1 mM NaN_3, pH 8.0) containing 1% of sodium dodecyl sulfate (SDS), 1 mM dithiothreitol, and 8 M urea. The solubilized proteins were transferred to the column with Ni^{2+}-sepharose (GE Healthcare, Sweden), washed with 10 column volumes of buffer A containing 1% SDS, and eluted with 3 volumes of buffer A containing 1% SDS and 500 mM imidazole. The GPCR samples were dialyzed against buffer A containing 1% SDS.

The eluate fractions were analyzed by SDS-PAGE and Western blotting using mouse monoclonal antibodies against the hexahistidine sequence (His-tag® Monoclonal antibody, Novagen, USA). The amount of purified GPCR samples was determined spectrophotometrically at room temperature based on absorption at 280 nm. The CD spectra were recorded at room temperature on a J-810 spectrometer (Jasco, Japan).

RESULTS AND DISCUSSION

Design of the *GPCR* genes

The truncated variants of the receptors containing additional point substitutions were used to increase the stability of the GPCR samples and to reduce the aggregation tendency of the proteins. Genetic engineering methods were used to excise the N- and C-terminal extramembrane regions that do not participate in ligand binding [7–9, 22–24]. The deletion of the C-terminal regions of the receptors resulted in the removal of cysteine residues (241,435, and 320), which are presumably the sites of post-translational binding of palmitic acid residues in human β2AR, M1-mAChR and SSTR5 molecules, respectively [7, 23, 24]. In addition, the fragment of the third cytoplasmic loop (L3), which also does not participate in ligand binding, was deleted from the M1-mAChR molecule [8, 9, 25]. The genes obtained encoded the regions 25–340, 19–224/354–426, and 37–319 of human receptors β2AR, M1-mAChR, and SSTR5, respectively. Additional His_{10}-tag sequences were inserted at the 3-prime end of the genes in order to provide further purification of recombinant proteins by Ni^{2+} affinity chromatography.

The truncated genes of the β2AR, M1-mAChR, and SSTR5 receptors encoded 10, 9, and 10 cysteine residues, respectively. Among those, only the residues from the extracellular region presumably participate in the formation of disulfide bonds (Cys106–Cys191 and Cys184–Cys190 in β2AR; Cys98–Cys178 and Cys391–Cys394 in M1-mAChR; Cys112–Cys186 in SSTR5, the numeration is provided for the native sequence of the receptors). In order to reduce the aggregation of recombinant proteins due to the formation of "non-native" disulfide intermolecular bonds, transmembrane and cytoplasmatic Cys residues were substituted via site-directed mutagenesis. Thus, the data [26, 27] were used to substitute Cys77, Cys116 and Cys125 residues in β2AR for Val; and to substitute Cys285, Cys327, and Cys265 for Ser. In M1-mAChR, the Cys69, Cys205, Cys417, and Cys421 residues were substituted for Ser [28]. In SSTR5, the Cys129, Cys237, and Cys260 residues were substituted

for Ser; the Cys169, Cys218, and Cys220 residues were substituted for Val; and Cys51 and Cys298, for Gly. Furthermore, an additional stabilizing Glu122Trp substitution was introduced to the β2AR sequence [29].

Expression of the *GPCR* genes in cell-free system

The introduction of membrane-mimicking components to the RM allows to synthesize MPs in the soluble and functionally active forms [13–18]. However, most of these additives (e.g., detergent molecules) may reduce the productivity of the system via the partial or complete inhibition of the synthesis of the target protein [14–17]. For this reason, we did not use membrane-mimicking compounds for the synthesis when performing the comparative analysis of the efficiency of expression of the *GPCR* genes with additional 5-prime end regions. It should be mentioned that the target proteins accumulated as a precipitate in the RM. The precipitates were dissolved in a hard detergent (SDS) in the presence of urea and dithiothreitol as a reducing agent. The amount of synthesized proteins was determined spectrophotometrically after the dissolved precipitates had been purified via Ni^{2+} affinity chromatography. The synthesis of the target proteins was confirmed using monoclonal antibodies against the hexahystidine sequence.

As one would expect, the direct expression of the truncated *β2AR*, *M1-mAChR*, and *SSTR5* genes in CF systems based on the *E. coli* S30 extract was inefficient. The yield of the target proteins after the purification did not exceed 0.1 mg per 1ml of RM (*Fig. 2*). It should be noted that highly efficient production (with a yield of up to 1.6 mg/ml of RM) of bacteriorhodopsin from Gram-positive bacteria *Ex. sibiricum* (ESR) [30], a structural homolog of the GPCRs, which also contains seven TM helices, has been previously observed [15]. We supposed that the low yield of the model GPCRs could be attributed to the low efficiency of translation initiation due to the formation of a secondary mRNA structure at the beginning of the target gene. In order to confirm this assumption, the 5-prime end regions encoding the extracellular N-terminal amino acid residues preceding the first TM helix (25–33, 19–23, and 37–38 in β2AR, M1-mAChR, and SSTR5, respectively) in the truncated GPCR genes were substituted with the nucleotide sequence encoding the first 6 a.a. of bacteriorhodopsin ESR (ESR-tag, the sequence length is indicated with allowance for the N-terminal Met) (*Fig. 1*). This substitution allowed one to significantly increase efficiency in the production of the target protein (*Fig. 2*). The yield of the ESR-tag-β2AR hybrid protein was comparable to that of the ESR protein, whereas the level of synthesis of the remaining two hybrid proteins (ESR-tag-M1-mAChR and ESR-tag-SSTR5) was approximately three times lower (~0.5 mg/ml of RM).

Comparison of the efficiency in GPCR synthesis with various N-terminal fusion tags

The results obtained have confirmed that the 5-prime end sequence plays a significant role in efficient expression in a cell-free system. However, the yields of the target proteins attained using the ESR-tag presumably were not optimal. Thus, synthesis of recombinant MPs in continuous CF systems based on the *E. coli* S30 extract with yields of up to 4–6 mg/ml of RM has been described in the literature [14]. For further optimization of the synthesis for the model GPCRs, we tested four N-terminal fusion tags. Two of those, the T7-tag (11 a.a.) and TRX protein (11.8 kDa), have previously been used in cell-free production of GPCRs [14, 16, 17], whereas the Mistic protein (12.8 kDa) was used for GPCR production in E. coli [31, 32]. In addition, we tested the sequence encoding the N-terminal fragment (16 a.a.) of ribonuclease A (N-terminal fragment of S-peptide, S-tag), which is used to detect and purify recombinant proteins via affinity chromatography [33], but has never been used as an N-terminal fusion tag for the production of recombinant MPs. In contrast to the method used to design hybrid genes with the 5-prime end fragment encoding the ESR-tag, nucleotide sequences encoding the T7-tag, TRX, Mistic, and S-tag were added in a single reading frame to the 5-prime end of the genes of the truncated GPCR variants (*Fig. 1*).

In most cases, the use of N-terminal fusion tags increased the yield of model receptors, but the yield levels varied for different proteins. Thus, the use of T7-tag increased the yields of M1-mAChR and SSTR5 receptors to ~0.5 mg/ml of RM, whereas the β2AR level stayed low and was comparable to that observed during direct expression. The use of the TRX also provided a small increase in the synthesis of the target proteins to ~0.3–0.7 mg/ml of RM (hereinafter, the amounts of the target proteins are given without the protein-fusion tags part, *Fig. 2*). Meanwhile, the use of the N-terminal fusion tags Mistic and S-tag allowed one to considerably increase the production of β2AR and M1-mAChR (*Fig. 2*). The highest yield of β2AR (~ 1.9 mg/ml of RM) was observed when using the Mistic protein, and the highest yield of M1-mAChR (~ 3.6 mg/ml of RM) was attained for the S-tag hybrid protein (*Fig. 2*). However, none of the sequences used has enabled to attain a considerable increase in the SSTR5 yield. The yields of this receptor (0.4–0.7 mg/ml of RM) were very close when using various hybrid constructs (*Fig. 2*). It seems that the translation initiation for SSTR5 is not the only crucial factor for providing efficient cell-free synthesis. Further optimization of the nucleotide sequence of the gene (e.g., substitution of the codon variants uncommon for *E. coli*) is presumably required to increase the production level in a cell-free system. It should be noted

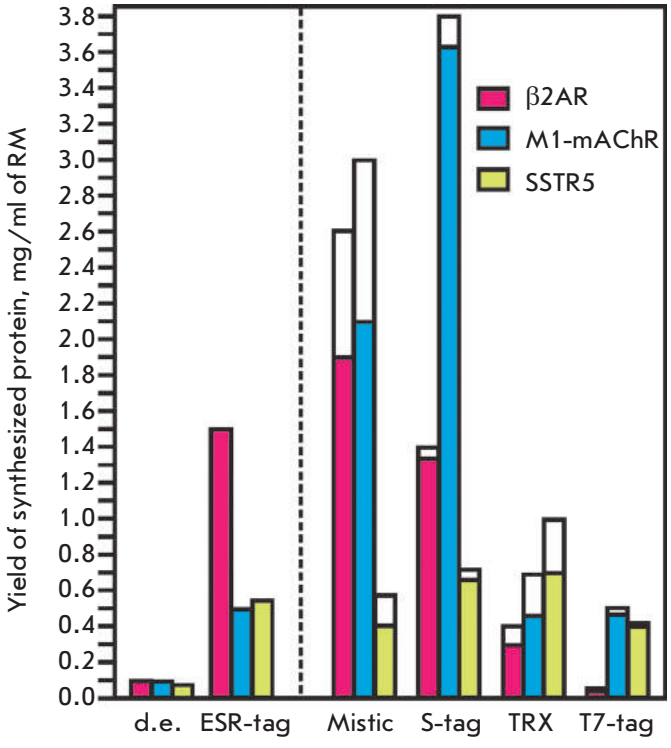


Fig. 2. Analysis of the CF synthesis of β2AR, M1-mAChR, and SSTR5 using different *N*-terminal fusion tags. The yields in CF expression of GPCRs without any *N*-terminal tags are designated as "d.e." (direct expression). Synthesis yields of hybrid proteins are shown by unfilled bars; the yields of the target proteins (colored bars) are shown after the deduction of a part of *N*-terminal partners. Each value represents the average of three experiments. The systematic error does not exceed 15%. The amounts of proteins produced were measured by UV-Vis absorption spectroscopy at 280 nm after purification by Ni^{2+}-chromatography

Free energy of formation of the secondary structure by the 5-prime end mRNA fragment (ΔG, kcal/mol)

Fusion tag/*GPCR*	β2AR	M1-mAChR	SSTR5
Direct expression	-5.6	-8.2	-19.3
ESR-tag	-3.1	-3.5	-6.4
Mistic	-1.3	-1.3	-1.3
S-tag	-3.3	-3.3	-3.3
TRX	-7.8	-7.8	-7.8
T7-tag	-5.5	-7.6	-7.3

Note. The free energy was calculated using the M-fold software [35] for mRNA fragments containing 4 nucleotides upstream of the start codon, the start codon, and 34 nucleotides of the target protein gene or fusion tag, downstream of the start codon

that a similar SSTR5 yield (~ 0.5 mg/ml of RM) was earlier observed in the bacterial continuous CF system when using a full-length (nontruncated) hybrid of the receptor with the N-terminal T7-tag sequence [34].

As previously mentioned, the increase in efficiency in protein synthesis when using additional 5-prime end sequences can presumably be attributed to the reduction in the ability of the 5-prime end mRNA fragment to form a secondary structure. To confirm this assumption, the formation of a secondary structure of the 5-prime end mRNA fragment used for GPCR production was modeled. The modeling was performed using the M-fold software to analyze the free energy of formation of the secondary structure of RNA [35]. The free energy of secondary structure formation was calculated for the mRNA fragments containing four nucleotides upstream of the start codon, the start codon, and 34 nucleotides of the target protein gene or the fusion tag downstream of the start codon, as was described in [20]. The computation (*Table*) has shown that the native sequences of the truncated receptors can form stable secondary structures (ΔG ~ −5.6, −8.2, and −19.3 kcal/mol for β2AR, M1-mAChR, and SSTR5, respectively). The use of T7-tag and TRX slightly reduces the stability of the secondary structure of the 5-prime end mRNA fragment (ΔG ~ −5.5−7.8 kcal/mol). Meanwhile, the use of the N-terminal sequence of bacteriorhodopsin ESR considerably reduces the stability of the secondary structure of the 5-prime end mRNA fragment in β2AR and M1-mAChR (ΔG ~ −3.1 and −3.5 kcal/mol, respectively). Secondary structures of mRNA characterized by the lowest stability were obtained for Mistic and S-tag sequences (ΔG ~ −1.3 and −3.3, respectively). The qualitative correlation between the calculated energies and the yields of GPCRs indirectly supports the important role of the formation of an 5-prime end mRNA secondary structure in the decrease in the efficiency of translation initiation and, as a consequence, in the total efficiency of the cell-free synthesis.

Modification of the 5-prime end region of the target protein gene is not the only way to prevent the formation of a secondary mRNA structure and increase efficiency in translation initiation. Nucleotide sequences from the 5-prime end untranslated regions of mRNA can also affect these processes. In this study, we used genetic constructs based on a pET22b(+) vector (Novagene) containing the *lac*-operator sequence inserted between the T7 promoter and the ribosome-binding site (RBS). According to published data, the use of *pIVEX* vectors (Roche Applied Science, USA) lacking the *lac*-

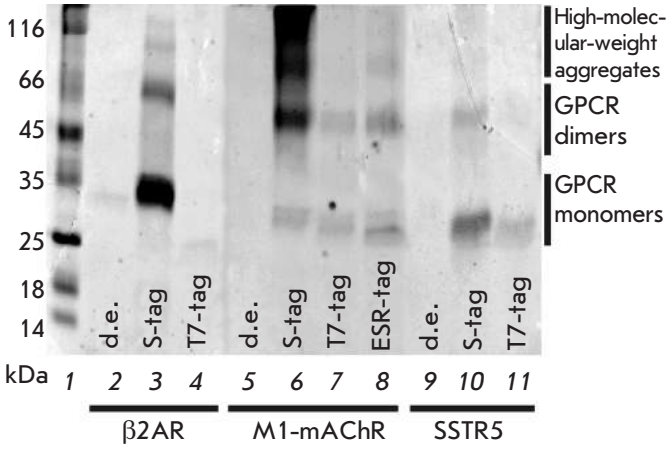


Fig. 3. SDS-PAGE analysis of the synthesized GPCRs with different N-terminal fusion tags after purification by Ni^{2+}-chromatography. Monomers, dimers, trimers and high-molecular-weight aggregates are present in the GPCR samples. 1 – molecular weight markers; 2, 5, 9 – receptors synthesized without N- terminal fusion tags ("d.e.," direct expression); 3, 6, 10 – receptors with S-tag; 4, 7, 11 – receptors with T7-tag; 8 – ESR-tag-M1-mAChR

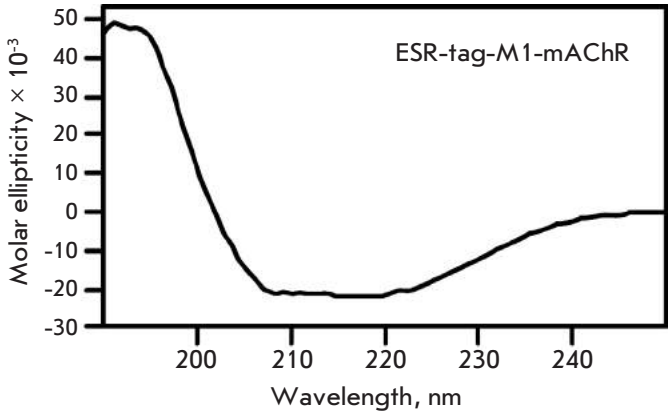


Fig. 4. CD spectrum of ESR-tag-M1-mAChR in 1% SDS

operator can increase efficiency in the direct expression of the *GPCR* genes in bacterial CF systems [34]. In order to verify this assumption, we tested efficiency in the direct expression of the truncated *M1-mAChR* gene using the *pIVEX2.3* vector. The yield of the target protein (~ 0.1 mg/ml of RM) in this case was no higher than that obtained via direct expression of the *M1-mAChR* gene cloned in the *pET22b*(+) vector. The data obtained were in close agreement with the results of the investigation of olfactory GPCRs, whose production in a bacterial cell-free system using *pIVEX* vectors was characterized by low efficiency [36]. Moreover, the use of N-terminal fusion tags was also required to provide highly efficient expression of human protein genes cloned into the *pIVEX* vectors [37].

Another method to solve the problem of low efficiency in translation initiation in CF systems can include the rational design of the 5-prime end sequence of the target protein gene using synonymous substitutions (without any changes in the encoded amino acid sequence), which is aimed at reducing the mRNA ability to form a secondary structure [20]. This approach was used to produce mammal cytokines when the presence of the fusion tag sequence (N-terminal fragment of chloramphenicol aminotransferase, 5a.a.) hindered the formation of the spatial structure [38].

Analysis of recombinant GPCRs

Purified GPCR samples solubilized in a hard SDS detergent (1%) were analyzed by SDS-PAGE. The representative gels are shown in *Fig. 3*. The resulting samples, as well as the other MP samples, possess an anomalous electrophoretic mobility, which is presumably caused by incomplete denaturation of MP molecules in SDS [39]. Separate bands corresponding to receptor monomers, dimers, trimers, and higher order aggregates were detected on the gels (*Fig. 3*). This behavior is typical for GPCRs, which tend to form dimers and trimers in biological membranes and are prone to spontaneous aggregation due to hydrophobic interactions between TM helices even in hard detergent solutions [31]. The aggregation level of GPCR samples depends on the receptor type, the sequence of N-terminal fusion tags, and presumably on the protein concentration in the sample. Thus, the highest amount of high-molecular-weight aggregates was observed in S-tag-M1-mAChR samples characterized by the most efficient synthesis. The secondary structure of the ESR-tag-M1-mAChR hybrid, which exhibited the lowest degree of aggregation in solution, was analyzed by CD spectroscopy (*Fig. 4*). The analysis of the resulting data revealed that the α-helical structure was the predominant one (α-helix – 65%, β-sheet – 4%, β-turn – 9%, and irregular regions – 22%), which attests to the fact that the secondary structure of the receptor is partially formed in the environment of SDS micelles. It should be noted that the content of α-helical elements in a molecule of the truncated M1-mAChR receptor calculated similarly to that in the known crystal structures of M2 and M3-mAChR [8, 9] is supposed to be equal to ~72%.

Further investigation of recombinant GPCRs requires either an optimization of the procedure of target-protein solubilization from the RM precipitate, followed by the development of renaturation methods for the obtained samples, or the use of membrane-mimicking media during the CF synthesis, which allows to synthesize MPs in the functionally active form in some cases [13,15, 34, 35].

CONCLUSIONS

The data obtained have demonstrated that the use of amino acid sequences of the ESR-tag, S-tag, and Mistic protein as N-terminal fusion tags allows to achieve a highly efficient production of human GPCRs in a cell-free system based on the *E. coli* S30 extract. Utilization of these sequences provides yields of target protein production (0.6 – 3.8 mg/ml of RM) that are sufficient for further structural and functional studies. The present work is the first to demonstrate the possibility of using the ESR-tag and S-tag to increase the level of heterologous production of MPs. ●

This work was supported by the Federal Target-Oriented Program "Scientific and Scientific-Pedagogical Personnel of the Innovative Russia" for 2009–2013, Program of the Russian Academy of Science "Molecular and Cell Biology", Russian Foundation for Basic Research (grant № 11-04-01864-a), the U.M.N.I.K. Program, and the Ministry of Education and Science (contracts № 8268 and 8789).

REFERENCES

1. Lundstrom K.H. Structural Genomics on Membrane Proteins. CRC Press, 2006.
2. Wallin E., von Heijne G. // Protein Sci. 1998. V. 7. № 4. P. 1029–1038.
3. Fredriksson R., Lagerström M.C., Lundin L.G., Schiöth H.B. // Mol. Pharmacol. 2003. V. 63. № 6. P. 1256–1272.
4. Overington J.P., Al-Lazikani B., Hopkins A.L. // Nat. Rev. Drug. Discov. 2006. V. 5. № 12. P. 993–996.
5. Rosenbaum D.M., Rasmussen S.G., Kobilka B.K. // Nature. 2009. V. 459. № 7245. P. 356–363.
6. McCusker E.C., Bane S.E., O'Malley M.A., Robinson A.S. // Biotechnol. Prog. 2007. V. 23. № 3. 540–547.
7. Cherezov V., Rosenbaum D.M., Hanson M.A., Rasmussen S.G., Thian F.S., Kobilka T.S., Choi H.J., Kuhn P., Weis W.I., Kobilka B.K., et al. // Science. 2007. V. 318. № 5854. P. 1258–1265.
8. Haga K., Kruse A.C., Asada H., Yurugi-Kobayashi T., Shiroishi M., Zhang C., Weis W.I., Okada T., Kobilka B.K., Haga T., et al. // Nature. 2012. V. 482. № 7386. P. 547–551.
9. Kruse A.C., Hu J., Pan A.C., Arlow D.H., Rosenbaum D.M., Rosemond E., Green H.F., Liu T., Chae P.S., Dror R.O., et al. // Nature. 2012. V. 482. № 7386. P. 552–556.
10. Cavanagh, J., Fairbrother, W. J., Palmer, A.G., III, Skelton, N.J., and Rance, M. Protein NMR Spectroscopy Principles and Practice. 2nd ed. N.Y.: Academic Press.
11. Kiefer H. // Biochim. Biophys. Acta. 2003. V. 1610. № 1. P. 57–62.
12. Shirokov V.A., Kommer A., Kolb V.A., Spirin A.S. // Methods Mol. Biol. 2007. V. 375. P. 19–55.
13. Schneider B., Junge F., Shirokov V.A., Durst F., Schwarz D., Dötsch V., Bernhard F. // Methods Mol. Biol. 2010. V. 601. P. 165–186.
14. Klammt C., Schwarz D., Fendler K., Haase W., Dötsch V., Bernhard F. // FEBS J. 2005. V. 272. P. 6024–6038.
15. Lyukmanova E.N., Shenkarev Z.O., Khabibullina N.F., Kopeina G.S., Shulepko M.A., Paramonov A.S., Mineev K.S., Tikhonov R.V., Shingarova L.N., Petrovskaya L.E., et al. // Biochim. Biophys. Acta. 2012. V. 1818. № 3. P. 349–358.
16. Klammt C., Schwarz D., Eifler N., Engel A., Piehler J., Haase W., Hahn S., Dötsch V., Bernhard F. // J Struct. Biol. 2007. V. 158. № 3. P. 482–493.
17. Ishihara G., Goto M., Saeki M., Ito K., Hori T., Kigawa T., Shirouzu M., Yokoyama S. // Protein Expr. Purif. 2005. V. 41. P. 27–37.
18. Haberstock S., Roos C., Hoevels Y., Dötsch V., Schnapp G., Pautsch A., Bernhard F. // Protein Expr. Purif. 2012. V. 82. № 2. P. 308–316.
19. Hall M.N., Gabay J., Debarbouille M., Schwartz M. // Nature. 1982. V. 295. P. 616–618.
20. Kudla G., Murray A.W., Tollervey D., Plotkin J.B. // Science. 2009. V. 324. № 5924. P. 255–258.
21. Khabibullina N.F., Lyukmanova E.N., Kopeina G.S., Shenkarev Z.O., Arseniev A.S., Dolgikh D.A., Kirpichnikov M.P. // Russian J. Bioorg. Chem. 2010. V.36. № 5. P. 603–609.
22. Greenwood M.T., Hukovic N., Kumar U., Panetta R., Hjorth S.A., Srikant C.B., Patel Y.C. // Mol. Pharmacol. 1997. V. 52. № 5. P. 807–814.
23. Hukovic N., Panetta R., Kumar U., Rocheville M., Patel Y.C. // J Biol Chem. 1998. V. 273. № 33. P. 21416–21422.
24. Kaye R.G., Saldanha J.W., Lu Z.L., Hulme E.C. // Mol. Pharmacol. 2011. V. 79. № 4. P. 701–709.
25. Shapiro R.A., Nathanson N.M. // Biochemistry. 1989. V. 28. № 22. P. 8946–8950.
26. Gether U., Lin S., Ghanouni P., Ballesteros J.A., Weinstein H., Kobilka B.K. // EMBO J. 1997. V. 16. № 22. P. 6737–6747.
27. Fraser C.M. // J Biol Chem. 1989. V. 264. № 16. P. 9266–9270.
28. Savarese T.M., Wang C.D., Fraser C.M. // J Biol Chem. 1992. V. 267. № 16. P. 11439–11448.
29. Roth C.B., Hanson M.A., Stevens R.C. // J Mol Biol. 2008. V. 376. № 5. P. 1305–1319.
30. Petrovskaya L.E., Lukashev E.P., Chupin V.V., Sychev S.V., Lyukmanova E.N., Kryukova E.A., Ziganshin R.H., Khatypov R.A., Erokhina L.G., Spirina E.V., et al. // FEBS lett. 2010. V. 584. № 19. P. 4193–4196.
31. Petrovskaya L.E., Shulga A.A., Bocharova O.V., Ermolyuk Y.S., Kryukova E.A., Chupin V.V., Blommers M.J., Arseniev A.S., Kirpichnikov M.P. // Biochemistry (Mosc.). 2010. V. 75. № 7. P. 881–891.
32. Chowdhury A., Feng R., Tong Q., Zhang Y., Xie X.Q. // Protein Expr. Purif. 2012. V. 83. № 2. P. 128–134.
33. Terpe K. // Appl. Microbiol. Biotechnol. 2003. V. 60. № 5. P. 523–533.
34. Klammt C., Perrin M.H., Maslennikov I., Renault L., Krupa M., Kwiatkowski W., Stahlberg H., Vale W., Choe S. // Protein Sci. 2011. V. 20. P. 1030–1041.
35. Zuker M. // Nucleic Acids Res. 2003. V. 31. № 13. P. 3406–3415.
36. Kaiser L., Graveland-Bikker J., Steuerwald D., Vanberghem M., Herlihy K., Zhang S. // Proc. Natl. Acad. Sci. USA. 2008. V. 105. № 41. P. 15726–15731.
37. Michel E., Wüthrich K. // J. Biomol. NMR. 2012. V. 53. № 1. P. 43–51.
38. Goerke A.R., Swartz J.R. // Biotechnol. Bioeng. 2008. V. 99. № 2. P. 351–367.
39. Rath A., Glibowicka M., Nadeau V.G., Chen G., Deber C.M. // Proc. Natl. Acad. Sci. USA. 2009. V. 106. № 6. P. 1760–1765.

Identification of Novel RNA-Protein Contact in Complex of Ribosomal Protein *S7* and 3′-Terminal Fragment of 16S rRNA in *E. coli*

A. V. Golovin[1], G. A. Khayrullina[1], B. Kraal[3], A. M. Kopylov[2*]

[1]Department of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Leninskie Gory, 1/73, Moscow, Russia, 119991

[2]Chemistry Department, Lomonosov Moscow State University, Moscow, Leninskie Gory, 1/3, Moscow, Russia, 119992

[3]Leiden Institute of Chemistry, Leiden University, 2300 RA Leiden, the Netherlands

*E-mail: kopylov.alex@gmail.com

Received 02.07.2012

ABSTRACT **For prokaryotes in vitro, 16S rRNA and 20 ribosomal proteins are capable of hierarchical self- assembly yielding a 30S ribosomal subunit. The self-assembly is initiated by interactions between 16S rRNA and three key ribosomal proteins: S4, S8, and S7. These proteins also have a regulatory function in the translation of their polycistronic operons recognizing a specific region of mRNA. Therefore, studying the RNA–protein interactions within binary complexes is obligatory for understanding ribosome biogenesis. The non-conventional RNA–protein contact within the binary complex of recombinant ribosomal protein S7 and its 16S rRNA binding site (236 nucleotides) was identified. UV–induced RNA–protein cross-links revealed that S7 cross-links to nucleotide U1321 of 16S rRNA. The careful consideration of the published RNA– protein cross-links for protein S7 within the 30S subunit and their correlation with the X-ray data for the 30S subunit have been performed. The RNA – protein cross–link within the binary complex identified in this study is not the same as the previously found cross-links for a subunit both in a solution, and in acrystal. The structure of the binary RNA–protein complex formed at the initial steps of self-assembly of the small subunit appears to be rearranged during the formation of the final structure of the subunit.**

KEYWORDS **ribosome; initiation; self-assembly; ribosomal protein S7; UV– induced cross-linking.**

ABBREVIATIONS **XRD – X-ray diffraction analysis; RNP – ribonucleoprotein; EcoS7, TthS7, BstS7 –proteins S7 isolated from *E. coli*, *T. thermophilus* and *B. stearothermophilus*, respectively; Tth30S and Eco30S – small ribosomal subunits isolated from *T. thermophilus* and *E. coli*, respectively.**

INTRODUCTION

The *in vitro* self-assembly of bacterial ribosomes has been relatively well described [1–5]. The phenomenology of the events resulting in the formation of individual ribosomal subunits has been well established. However, the thorough analysis of the interaction between rRNA and proteins is just being started.

The individual assembly of the small 30S ribosomal subunit and the large 50S ribosomal subunit occurs during the formation of the prokaryotic 70S ribosomes. The small ribosomal subunit of *Escherichia coli* consists of 1542-nucleotides-long 16S rRNA and 20 different medium-size proteins. The proteins that are the first to bind to the 16S rRNA (S4, S7, S8, S15) resulting in the formation of the so-called "structural core" of a small subunit play a crucial role during the self-assembly of the 30S subunit [6, 7]. The ribosomal assembly process has come into the focus of researchers again now that the structure of the small subunit of thermophilic and mesophilic ribosomes has been identified using X-ray diffraction (XRD) analysis [8–10]. The possibility of describing the sequence of events during the self-assembly using a specific structural terminology has become real [3, 4, 11]. Moreover, the potential opportunity for interfering in the ribosome biogenesis process may stimulate the designing of fundamentally different and powerful antibacterial agents.

As opposed to the 50S subunit, the expressive discrete character of the structure of the 30S subunit makes the experimental study of its self-assembly much easier: it consists of 4 domains (*Fig. 1*) [8, 9]. Three RNP domains are capable of assembling independently [12–17]. The minimal fragment of 16S rRNA (236 nucleotides long) isolated from *E. coli* (D3LH, Eco16S), which is capable of specific binding to protein S7, the key participant in the subunit assembly process, has been found for the major 3'-terminal domain, [18].

The present work is devoted to the investigation of the rRNA–protein contacts in the binary complex of the recombinant ribosomal protein S7 with a binding site on a fragment of 16S rRNA isolated from *E. coli* using UV-induced RNA-protein cross-linking. A non-conventional rRNA-protein contact has been identified: protein S7 is cross-linked to the nucleotide U1321. The annotation of the previously published rRNA–protein cross-links of protein S7 in the 30S subunit in solution and the XRD data obtained for a small subunit crystal has been carried out. The newly identified rRNA–protein cross-link in the binary complex does not match any of the annotated cross-links found in the intact subunit. It can be hypothesized that the structure of the binary rRNA–protein complex that is formed during the initial stages of the small ribosomal subunit assembly must undergo rearrangement during the formation of an intact subunit.

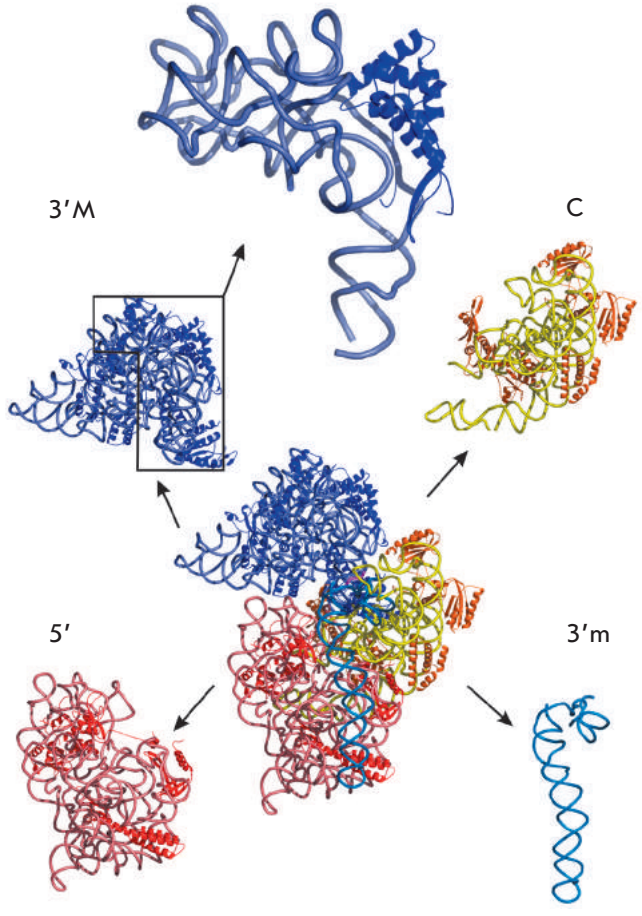


Fig. 1. X-ray structure of the 30S small ribosomal subunit isolated from *T. thermophilus* (PDB 1FJF [8]). Domains are specified as follows: (5') – 5'-end, (C) – central, (3'M) – major 3'-end, and 3'm – minor 3'-end domains. Proteins are shown as dark blue, orange, and red ribbons. 16S rRNA is shown as light blue, cyan, yellow, and pink ribbons. Top: scaled-up 16S RNA (Eco16S)-ribosomal protein S7 complex extracted *in silico*

EXPERIMENTAL

T4 polynucleotide kinase (PNK) and PNK buffer (New England Biolabs, USA), reverse transcriptase of the avian myeloblastosis virus (RT-AMV), Taq DNA polymerase, RNase inhibitor, proteinase K, nucleoside triphosphate and its dideoxy derivatives (Roche, Germany), [γ-^{32}P]ATP (Amersham, Germany), bovine serum albumin (BSA, MBI, Fermentas, Lithuania), 0.45 μm nitrocellulose filters (Millipore HA, USA; Schleicher & Schuell BA85, Germany), Ni-NTA-agarose (QIAGEN, Germany), phenylmethylsulfonyl fluoride (PMSF, Merk, Germany) were used. The pFD3LH plasmid was kindly provided by L. Brakier-Gingras (University of Montreal, Canada).

Buffer A: 50 mM Tris-HCl (pH 9.5), 1.5 mM $MgCl_2$, 20 mM $(NH_4)_2SO_4$, 1 mM dithiothreitol (DTT), 0.005% NP-40, 5% dimethyl sulfoxide (DMSO), 1 mM betaine. Buffer B: 40 mM Tris-HCl (pH 7.9), 12 mM $MgCl_2$, 10 mM NaCl, 10 mM DTT, 2 mM spermidine. Buffer C: 0.3 M NaAc (pH 5.2), 1 mM EDTA, 0.2% phenol. Buffer D: 50 mM Hepes-KOH (pH 7.0), 100 mM KCl. Buffer E: 50 mM Tris-HCl (pH 8.5), 10 mM $MgCl_2$, 60 mM KCl, 10 mM DTT, 0.5 mM dNTP. Buffer F: 20 mM Tris-Ac (pH 7.8); 7 mM $MgAc_2$, 300 mM NH_4Cl, 0.2% BSA.

Isolation of the recombinant protein S7 of *E.coli* (EcoS7) and protein S7 of *Thermus thermophilus* (TthS7) from the superproducer strain of *E. coli*

The EcoS7 was isolated from the superproducer strain of *E. coli* in accordance with the QIAGEN protocols as was briefly mentioned previously [19]. Cells were collected by centrifugation, suspended in 50mM Tris-HCl (pH 8.0) containing 500 mM NaCl and lysozyme. After the incubation, glycerin was added to 10%; mercaptoethanol, to 5 mM; PMSF, to 0.5 mM; and Triton X-100, to 1%. Following the subsequent ultrasonication, inclusion bodies were dissolved in a buffer containing 8 M urea, applied to Ni-NTA-agarose; after rinsing, the urea concentration in the eluent was re-

duced to 0. The protein was eluted using a 0–0.5 M gradient of an imidazole solution in a buffer of 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 1 mM mercaptoethanol, 5% glycerol, and 0.5 mM PMSF. The protein was transferred into the buffer of 20 mM Hepes-KOH (pH 7.5), 100 mM NaCl, 0.2 mM DTT, 5% glycerin, 0.5 mM PMSF by dialysis, and kept at –70°C. Prior to complex formation, the protein was transferred into a buffer of 20 mM Tris-HCl (pH 7.6), 4 mM $MgAc_2$, 400 mM NH_4Cl, 0.2 mM EDTA, and 4 mM mercaptoethanol. Protein S7 was isolated in a similar fashion [19, 20].

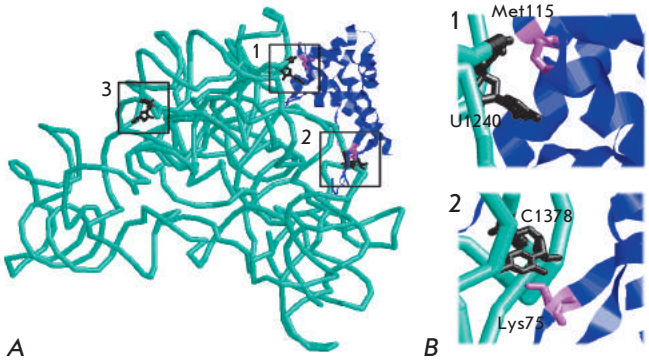


Fig. 2. The correlation between the XRD data obtained for the 30S ribosomal subunit isolated from *E.coli* in crystal and data regarding the cross-links of this subunit in solution. A – EcoS7–Eco16S complex structure (*in silico* extraction from Eco30S). 16S rRNA – cyan ribbon, protein S7– blue ribbon. RNA–protein cross-links are shown in brackets: 1 – U1240-Met115; 2 – C1378-Lys75; 3 - U1321–protein S7 within the binary complex (*Table*). B – Details of the RNA–protein contacts are shown in Fig. 2A

DNA amplification by PCR

The matrix DNA fragment was amplified using the pFD3LH plasmid containing cDNA of the minimal fragment of 16S rRNA under the control of the T7 phage promoter. PCR was carried out in 50-400 μl of buffer A containing 200 mM of dNTP, 20 pmol of primers, 50–500 ng of pFD3LH, and 2–5 AU Taq DNA polymerase. The 5'-terminal primer AGGGATCCTAATACGACTCACTATAGGG corresponds to the promoter sequence of the T7 phage RNA-polymerase and is complementary to the vector sequence; the 3'-end primer GTAAGCTTACAA GGCCCGGGAACGTATTCACC is complementary to the fragment G1370-U1393 of the Eco16S (non-complementary sequence is underlined). The primers were synthesized by MWG-Biotech AG company (Germany). PCR was carried out on a thermal cycler (BioRad, USA) under the following conditions: preincubation at –95°C for 2 min; cycle at 95°C for 45 s; at 60°C for 30 s; at 72°C for 30–60 s. After 25 cycles, additional incubation at 72°C for 7 min was carried out. DNA was purified through electrophoresis in 1–2% agarose gel, 3 volume extraction (according to gel weight) with 6 M NaI (56°C, 5 min) with subsequent purification using the PCR Purification Kit (Roche, Germany).

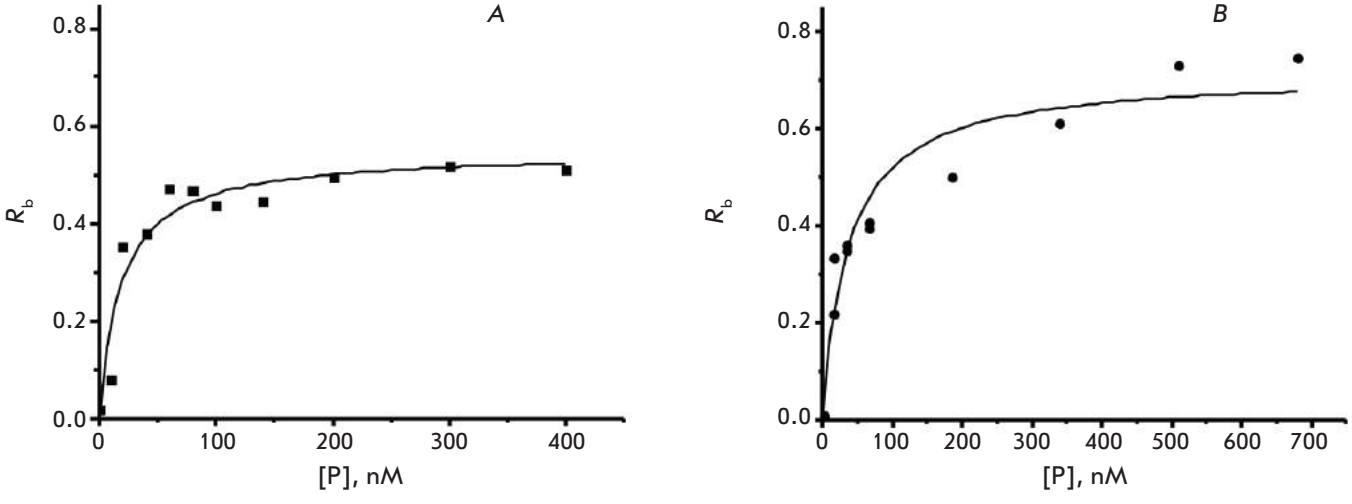


Fig. 3. Binding isotherms for the EcoS7–Eco16S complexes (A) and TthS7–Eco16S (B). aK_d = 21.5 ± 1.9 nM, and 35.8 ± 9.3 nM, respectively. The initial concentration of the Eco16S – 20 nM, [P] – protein concentration, R_b – fraction of the protein bound Eco16S

Analysis of the correlation between the XRD data obtained for the 30S ribosomal subunit of *E. coli* and the data for the cross-links between the 16S rRNA and protein S7 within the 30S subunit in solution

№	Cross-link with the 16S rRNA	Cross-link with protein S7	Distance in the Eco30S, Å	Reagent	Size of the reagent, Å	Reference
1.1	A1238–U1240	S7	3.0	API	8.6	[21]
1.2	A1238–U1240	S7	3.0	IT	5	[22]
1.3*	U1240	M115**	2.7	IT	5	[23–25]
1.4	U1240***	S7	2.7	UV	0	[26]
1.5	16S rRNA	M115**	2.7	UV	0	[27]
2.1	A1377–C1378	S7	3.8	IT	5	[22]
2.2	C1378	K75	3.8	IT	5	[25]

Note. Numeration in the first column: the first number denotes the contact number, 1 (1238–1240) or 2 (1377–1378), the second number is the order number of the cross-link: 1–5 for the first contact, 1–2 for the second contact. API – Methyl-*p*-azidophenylacetimidate; IT – 2-iminothiolane.
* Analogous cross-link was identified in the small subunit of *Bacillus stearothermophilus* (Met115 Bst7) [24, 27].
** In studies [23-25, 27], Met115 was denoted as Met114 (an error in sequencing of protein EcoS7 [28] (R91 was missing [29])).
*** Until 1979, incorrect numeration of the 16S rRNA [30] was used (U1239 instead of U1240).
Data not included in the Table. A) 30S subunit of *E. coli*. 1. Identified cross-link C1265 between the 16S rRNA and protein S7 [30]. C1265 is located at a distance of 35 Å from the nearest amino acid residue of protein S7 in the Eco30S. 2. Identified cross-links 278-280, 1139-1144, 1155-1158, 1531-1542 between the 16S rRNA and protein S7 [31]. The minimal distance between the 1531-1542 segment of the 16S rRNA and protein S7 in the Eco30S is equal to 11 Å. B) The 30S subunit of *B. stearothermophilus:* identified cross-link between the 16S rRNA with the Lys8 residues of protein S7 [27].

Transcription of the 16S rRNA (Eco16S) segment *in vitro*

Transcription of Eco16S was carried out on a PCR-copy of matrix DNA containing the T7 phage RNA-polymerase promoter in 100 µl of a solution containing the following components: 2.5 mM NTP, 1000 AU T7 phage RNA polymerase, 60 AU RNase inhibitor, 1 µg/ml of pyrophosphatase, and 4 µg of matrix DNA in buffer B at 37^{0}C for 1 h. After the transcription, the solution was subjected to phenol deproteinization followed by chloroform extraction and ethanol precipitation. The RNA was purified in 8% polyacrylamide gel containing 7 M urea and eluted from the gel by diffusion in buffer B. After the elution, the RNA was treated with phenol, chloroform, and subsequently precipitated in ethanol. The precipitated RNA was dissolved in 50 µl of water; the RNA integrity was confirmed using 8% polyacrylamide (containing 7 M urea) gel electrophoresis. The RNA concentration was determined from the absorbance at 260 nm: 1 mg of RNA – 22 o.u.

Obtaining the EcoS7–Eco16S and the TthS7–Eco16S complexes

Complex formation was performed in 200 µl of buffer F. The RNA and the protein were renatured separately at 37˚C for 30 min then mixed and heated at 37˚C for 30 min. The degree of complex formation was determined using adsorption on nitrocellulose membranes at the filtration rate of 0.5 ml/min by titrating the constant quantity of the ^{32}P-labeled RNA with an increasing quantity of the protein [19]. The radioactivity of the filters was determined in 10 ml of water in accordance with Cherenkov's method using a Tracor Analytic counter (France). The apparent dissociation constant (aK_d) was determined using XMGRACE software, GNU (http://plasma-gate.weizmann.ac.il/Grace/), by the following equation:

$$\alpha = \frac{P_0}{K_d * R_0 + R_0^2 + P_0 R_0},$$

where P_0 is concentration of protein S7, R_o is the fixed concentration of Eco16S, K_d* (aK_d) is the apparent dissociation constant, α is the bound fraction in Eco16S complex.

UV-induced covalent RNA-protein cross-linking in the EcoS7–Eco16S and the TthS7–Eco16S complexes

Complex formation was performed in 200 μl of buffer F at the RNA concentration of 150 nM and a 10-fold molar excess of protein. The protein was renatured at 37°C, mixed with RNA, and kept at 37°C for 30 min. The complex was kept under UV light at 260 nm (Stratolinker, USA, power of 2400 μV) on ice for 10 min. The UV light source was located 15 cm away from the complex; the light intensity was controlled by measuring the uridine concentration.

Obtaining the oligodeoxyribonucleotide primers labeled with ^{32}P at the 5'-terminus

The labeled primer (3'-terminal primer for PCR) for reverse transcription was obtained using kination with PNK in the presence of [γ-^{32}P]ATP. PNK buffer (10μl) containing 20 pmol of the primer, 3 μl of the [γ-^{32}P] ATP (0.4 MBq/μl), and10 AU PNR, and subsequently incubated at 37^{0}C for 1 h. The reaction was halted by adding 90 μl of 0.3 M NaAc (pH 5.2) with subsequent phenol deproteinization and chloroform extraction. The primer was precipitated in ethanol and dissolved in 40 μl of water.

Mapping the Eco16S nucleotide cross-linked to protein S7 in the EcoS7–Eco16S and the TthS7–Eco16 complexes

After the irradiation, the complex was treated with proteinase K to remove protein S7. Mapping of the Eco16S nucleotide cross-linked to protein S7 was carried out by reverse transcription using the primer labeled at its 5'-terminus. The hybridization of the primer with RNA was carried out in 4.5 μl of buffer D containing 2–5 pmol of RNA and 0.5 pmol of primer. RNA was denatured at 95^{0}C for 1 min followed by slow cooling to 42.5^{0}C. Reverse transcription was carried out in 8.5 μl of buffer E containing 2.2 AU RT-AMV at the same temperature for 1 h. One of the ddNTPs (70–400 μM) was added during the control sequencing. Samples were analyzed using 8% polyacrylamide gel containing 7 M urea.

RESULTS AND DISCUSSION

The spatial structures of the 30S small ribosomal subunits isolated from the thermophilic bacterium *T. thermophilus* (Tth30S) [8, 9] and from *E. coli* [10] were determined using XRD. No biochemical data describing the assembly of the 30S subunit in *T. Thermophilus* in solutions have been obtained thus far; only the possibility of domain assembly has been identified [14, 15, 17]. Most biochemical data on the assembly of ribosomes were obtained using *E. coli* ribosomes. Hence, the analysis of the correlation of the biochemical data obtained for the Eco30S in solution and those obtained using XRD for the Eco30S and Tth30S is of particular interest.

The RNA–protein cross-links were widely used to investigate the contacts in the 30S bacterial ribosome subunit in solution. Several cross-links of the 16S rRNA and protein S7 in the structure of the 30S subunit isolated from *E. coli* (*Table*) have been described. Two of these typical cross-links have been reliably identified as U1240–Met115 and C1378–Lys75; this finding correlates well with the XRD data for crystals (*Table*, *Fig. 2*). Hence, we used UV-induced cross-linking in the present work to identify the possible rRNA-protein contacts in the Eco16S fragment–protein S7 binary complex.

It had been previously shown that complexes of protein S7 and the intact 16S rRNA create a cross-link un-

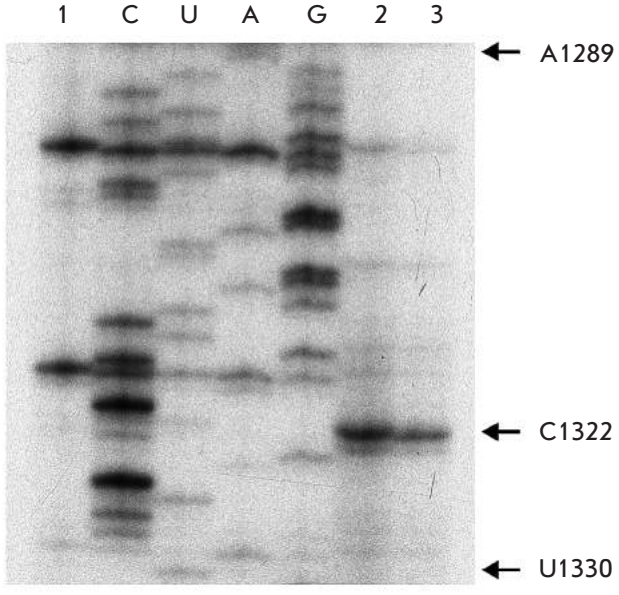


Fig. 4. Primer extension analysis of EcoS7-Eco16S, and TthS7 - Eco16S cross-links within the binary complexes. Radioautography of electrophoresis of reverse transcription products in 8% PAAG 8M urea. Lane *1* – cDNA from Eco16S rRNA after UV irradiation. Lanes C, U, A, G – Eco16S rRNA sequencing of the region A1289 – U1330. Lanes 2 and 3 – UV-induced cross-links of Eco16S – EcoS7 and Eco16S-TthS7, respectively. The arrow shows C1322 corresponding to the U1322 cross-link

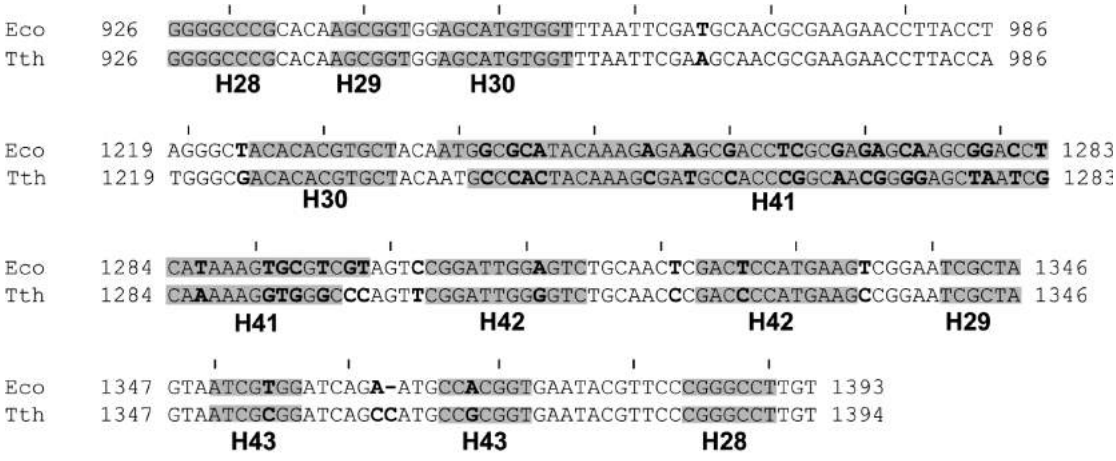


Fig. 5. Alignment of the primary structures of the Eco16S and Tth16S fragments. Conventional numbering of nucleotides in the Eco16S was used; the numbering for the Tth16S fragment was in accordance with the PDB 1FJF [12] for Tth30S. Non-identical nucleotides are shown in bold; double-stranded regions are shown in gray

der UV irradiation [32]; however, no cross-linked residues have been identified.

Brakier-Gingras *et al.* have demonstrated [18] that protein EcoS7 is capable of binding to a small fragment of 16s rRNA (236 nucleotides, D3LH, Eco16S), which is the key element in the structure of the major 3'-terminal domain of 16S rRNA. The EcoS7–Eco16S complex was obtained by the authors using EcoS7 isolated from an aggregated ribosomal protein in accordance with the standard methodology [33]. The apparent dissociation constant of this protein complex (aK_d) was relatively high (620 ± 80 nM) [18]. The recombinant protein containing 6 additional histidine residues (6 His) at the N terminus was subsequently used. The recombinant protein also bound to the Eco16S; its aK_d was considerably less, in the range of 110–210 nM [34, 35]. It is considered that the additional fragment containing 6 His residues does not affect the binding of the protein to 16S rRNA [35]; whereas the difference in the constants reflects the difference in isolation methods. A recombinant protein EcoS7 was used in the present work, which had 6 His residues at its N terminus [19].The EcoS7–Eco16S complex turned out to be more stable than it used to be considered [34, 35]; its aK_d was 21.5 ± 1.9 nM (*Fig. 3*), which attests to its high activity.

The EcoS7-Eco16S complex was irradiated with UV light; the cross-linking efficiency was determined using polyacrylamide gel electrophoresis under denaturing conditions from the ratio between the radioactivity in the RNP zone and the total radioactivity of rRNA. The duration of the irradiation was selected in such a way as to provide maximum yield of the cross-linked RNP. The position of the cross-linked heterocyclic bases in rRNA was determined to identify the Eco16S–EcoS7 contact using reverse transcription after protein hydrolysis with proteinase K; allowance was made for the fact that reverse transcriptase stops one nucleotide before the modified one. The analysis of the "cross-linked" Eco16S-EcoS7 complex (*Fig. 4*, lane *2*) definitively identifies the unique stop-signal corresponding to the C1322 nucleotide (cross-linked to U1321).The additional "stop" signals in the remaining locations have not been identified. The position of the cross-link is shown in the tertiary structure of the 30S ribosomal subunit isolated from *E. coli* (*Fig. 2*).

The identified contact between the Eco16S rRNA and protein EcoS7 differs from all the known contacts, which are formed during the cross-linking of the 16S rRNA with protein S7 in a small subunit of *E. coli* ribosomes in solution (*Table*). Moreover, the contact between the Eco16S and protein EcoS7 identified by us does not match the structure of the analogues RNP domain in the structure of the 30S *E. coli* subunit in the crystal (*Fig. 2*): the amino acid residue of protein S7 closest to U1321 is located at a distance of 35 Å (*Table*).

The difference identified can be attributed to the fact that during the interaction between protein S7 and the 16S rRNA at the initial stages of ribosome assembly the structure of the assembled binary complex differs from the final structure of the corresponding RNP domain within the subunit. Based on the analysis of the structure of the RNP domain in Eco16S and Tth30S, one can assume that the Eco16S in the binary complex with protein S7 is likely to be characterized by an uncoiled state of four-helix bundles (H30, H41, H42, H43), which are packed side-by-side in the crystal structure of Eco16S and Tth30S [19]. Some additional factors may presumably be required for sta-

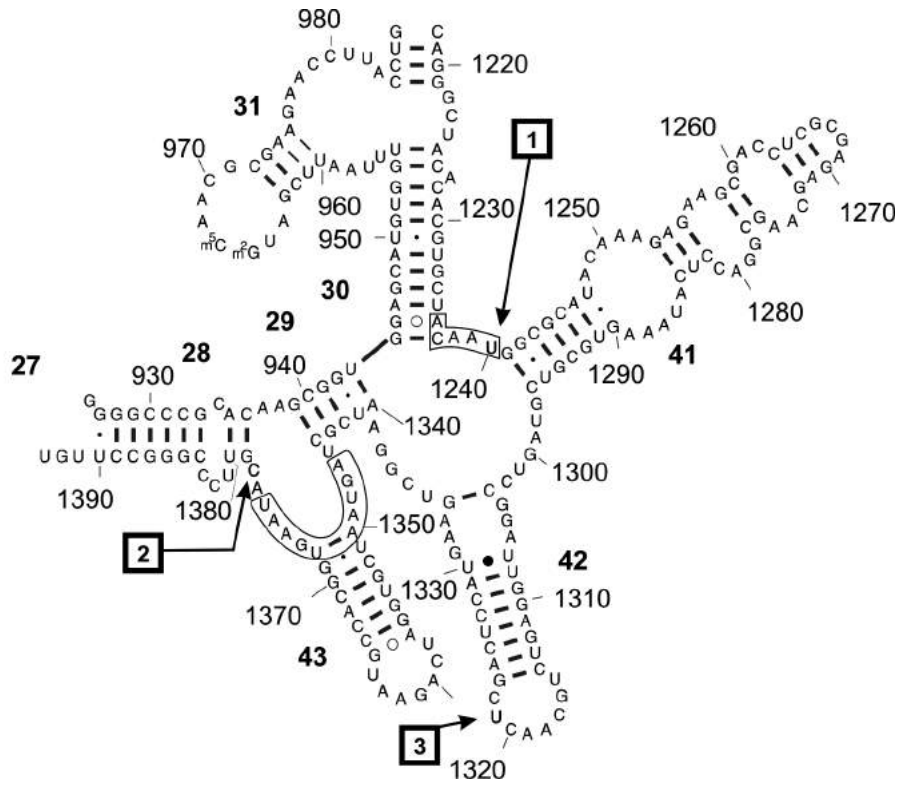


Fig. 6. Model of the secondary structure of the major 3′-terminal domain of the 16S rRNA (D3LH, Eco16S) used in this investigation [18]. RNA-protein cross-links are shown by arrows. Cross-links are taken from the *Table*. Cross-links for the 30S *E. coli* subunit: (1) U1240 - M115, (2) C1378 - K75; and for the binary complex: (3) U1321 – protein S7 identified in this work. 16S rRNA sites identical to the streptomycin mRNA binding site for protein S7 are shown in brackets [30]

bilization of the binary complex in its compact state during the self-assembly of ribosomes (e.g., high local concentration of Mg ions [19] or an interaction with the other proteins in the domain). This hypothesis is in agreement with the existence of an additional special Thx protein in thermophilic ribosomes; this protein is of a particularly strong basic character, and therefore, it can stabilize the compact structure of this RNP domain [8, 9].

The comparative analysis of the heterologous structure of the TthS7-Eco16S protein complex is of indisputable interest. In this case, protein TthS7 can be regarded as a "natural mutant" of protein EcoS7 [19]. We had previously shown [19, 36] that ThtS7 can form stable complexes with Eco16S. In the present work, the heterologous complex had a K_d of 35.8 ± 9.3 nM (*Fig. 3*), which was comparable to the a K_d of the homologous EcoS7–16S complex (a K_d = 21.5 ± 1.9 nM). The contact sites of the recombinant ThtS7 and the Eco16S fragment were also identified in a similar fashion to the Eco16S-EcoS7 complex: protein ThtS7 cross-links to U1321 (*Fig. 4*). It appears that a similar RNA–protein contact exists in the heterologous complex. Interestingly, the position 1321 in the 16S rRNA is phylogenetically conserved and the substitution was found only in thermophilic 16S rRNA (*Fig. 5*).

CONCLUSIONS

It has been demonstrated in the present study that binary complexes of the ribosomal protein S7 and its local binding site located at 16S rRNA can be obtained for the investigation of the initial stages of the assembly of small bacterial ribosomal subunits. This possibility is in close agreement with the previously shown possibility of assembly of the individual domain RNP complexes of bacterial ribosomes [5, 37]. The S7-containing complex cross-links to the residue of the U1321 under UV irradiation of binary complexes (260 nm) both in homologous (EcoS7-Eco16S) and heterologous (TthS7–Eco16S) complexes. As a result of searching for similar structures in the 16S rRNA and mRNA, Saito and Nomura [38] have proposed that the recombinant protein S7 recognizes a specific motif in the 16 rRNA structure, which is located next to the identified cross-link (*Fig. 6*). Moreover, the cross-link of protein S7 and the mRNA fragment next to the tentative motif was identified [39]. The combination of these data argues in favor of Saito and Nomuro's assumption [38] with regard to the possibility of the initial recognition of this RNA motif by protein S7.

It can be proposed that the formation of the intact small ribosomal subunit results in reorganization of the contacts in the initial binary complex. Such a rearrangement can also be observed in other RNA–pro-

tein complexes; for instance, in complexes of tRNA and phenylalanine-tRNA- synthetase [40]. Some interesting rearrangements have also been identified during the dissociation of binary RNA–protein complexes [41]. ●

The authors would like to thank A.A. Bogdanov, Y. Brozius, T.I. Rassohin, T.S. Rozhdestvenskiy, V.A. Spiridonova for their interest, helpful discussions, and support. This work was supported by the German Academic Exchange Service (DAAD), the Russian Foundation for Basic Research (grant № 11-04-01990-a), the Russian Foundation for Basic Research and the Netherlands Organization for Scientific Research (grants № 03-04-89001, 047.015.018).

REFERENCES

1. Spirin A.S. // Ribosomes / Ed. Siekevitz P. N.Y.: Kluwer Acad. / Plenum Publ., 1999. P. 98–101.
2. Nomura M. // J. Bacteriol. 1999. V. 181. P. 6857–6864.
3. Culver G.M. // Biopolymers. 2003. V. 68. P. 234–249.
4. Williamson J.R. // RNA. 2003. V. 9. P. 165–167.
5. Mulder A.M., Yoshioka C., Beck A.H., Bunner A.E., Milligan R.A., Potter C.S., Carragher B., Williamson J.R. // Science. 2010. V. 330. № 6004. P. 673–677.
6. Bogdanov A.A., Kopylov A.M., Shatsky I.N. // Subcell. Biochem. 1980. V. 7. P. 81–116.
7. Nowotny V., Nierhaus K.H. // Biochemistry. 1988. V. 27. P. 7051–7055.
8. Wimberly B.T., Brodersen D.E., Clemons W.M. Jr., Morgan-Warren R.J., Carter A.P., Vonrhein C., Hartsch T., Ramakrishnan V. // Nature. 2000. V. 407. P. 327–339.
9. Schluenzen F., Tocilj A., Zarivach R., Harms J., Gluehmann M., Janell D., Bashan A., Bartels H., Agmon I., Franceschi F., Yonath A. // Cell. 2000. V. 102. P. 615–623.
10. Dunkle J.A., Xiong L., Mankin A.S., Cate J.H.D. // Proc. Natl. Acad. Sci. USA. 2010. V. 107. P. 17152–17157.
11. Kopylov A.M. // Biochemistry. 2002. V. 67. P. 372–382.
Weitzmann C.J., Cunningham P.R., Nurse K., Ofengand J. // FASEB J. 1993. V. 7. P. 177–180.
12. Weitzmann C.J., Cunningham P.R., Nurse K., Ofengand J. // FASEB J. 1993. V. 7. P. 177–180.
13. Samaha R.R., O'Brien B., O'Brien T.W., Noller H.F. // Proc. Natl. Acad. Sci. USA. 1994. V. 91. P. 7884–7888.
14. Agalarov S.C., Selivanova O.M., Zheleznyakova E.N., Zheleznaya L.A., Matvienko N.I., Spirin A.S. // Eur. J. Biochem. 1999. V. 266. P. 533–537.
15. Agalarov S.C., Zheleznyakova E.N., Selivanova O.M., Zheleznaya L.A., Matvienko N.I., Vasiliev V.D., Spirin A.S. // Proc. Natl. Acad. Sci. USA. 1998. V. 95. P. 999–1003.
16. Jagannathan I., Culver G.M. // J. Mol. Biol. 2003. V. 330. P. 373–383.
17. Serdyuk I., Ulitin A., Kolesnikov I., Vasiliev V., Aksenov V., Zaccai G., Svergun D., Kozin M., Willumeit R. // J. Mol. Biol. 1999. V. 292. P. 633–639.
18. Dragon F., Brakier-Gingras L. // Nucl. Acids Res. 1993. V. 21. P. 1199–1203.
19. Rassohin T. I., Golovin A. V., Petrova E. V., Spiridonova V. A., Karginova O. A., Rozhdestvensky T. S., Brosius J., Kopylov A. M. // Molecular biology. 2001. № 35. P. 617–627.
20. Karginov A.V., Karginova O.A., Spiridonova V.A., Kopylov A.M. // FEBS Lett. 1995. V. 369. P. 158–160.
21. Osswald M., Greuer B., Brimacombe R., Stoffler G., Baumert H., Fasold H. // Nucl. Acids Res. 1987. V. 15. P. 3221–3240.
22. Wower I., Brimacombe R. // Nucl. Acids Res. 1983. V. 11. P. 1419–1437.
23. Urlaub H., Thiede B., Muller E.C., Wittmann-Liebold B. // J. Protein Chem. 1997. V. 16. P. 375–383.
24. Urlaub H., Kruft V., Wittmann-Liebold B. Methods in Protein Structure Analysis. N.Y.: Plenum Press, 1995. P. 275–282.
25. Urlaub H., Thiede B., Muller E.C., Brimacombe R., Wittmann-Liebold B. // J. Biol. Chem. 1997. V. 272. P. 14547–14555.
26. Zwieb C., Brimacombe R. // Nucl. Acids Res. 1979. V. 6. P. 1775–1790.
27. Urlaub H., Kruft V., Bischof O., Muller E.C., Wittmann-Liebold B. // EMBO J. 1995. V. 14. P. 4578–4588.
28. Reinbolt J., Tritsch D. // FEBS Lett. 1978. V. 91. P. 297–301.
29. Johanson U., Hughes D. // Gene. 1992. V. 120. P. 93–98.
30. Ehresmann B., Backendorf C., Ehresmann C., Millon R., Ebel J.P. // Eur. J. Biochem. 1980. V. 104. P. 255–262.
31. Greuer B., Osswald M., Brimacombe R., Stoffler G. // Nucl. Acids Res. 1987. V. 15. P. 3241–3255.
32. Ehresmann B., Reinbolt J., Backendorf C., Tritsch D., Ebel J. // FEBS Lett. 1976. V. 67. P. 316–319.
33. Wittmann H.G. Ribosomes. Cold Spring Harbor: CSHL, 1974. P. 93–114.
34. Robert F., Brakier-Gingras L. // Nucl. Acids Res. 2001. V. 29. P. 677–682.
35. Robert F., Gagnon M., Sans D., Michnick S., Brakier-Gingras L. // RNA. 2000. V. 6 P. 1649–1659.
36. Spiridonova V.A., Golovin A.V., Drygin D.Yu., Kopylov A.M. // Biochem. Mol. Biol. Int. 1998. V. 44. P. 1141–1146.
37. Agalarov S.C., Selivanova O.M., Zheleznyakova E.N., Zheleznaya L.A., Matvienko N.I., Spirin A.S. // Eur. J. Biochem. 1999. V. 266. № 2. P. 533–537.
38. Saito K., Nomura M. // J. Mol. Biol. 1994. V. 235. № 1. P. 125–139.
39. Golovin A., Spiridonova V., Kopylov A. // FEBS Lett. 2006. V. 580. № 25. P. 5858–5862.
40. Klipcan L., Moor N., Finarov I., Kessler N., Sukhanova N., Safro M.G. // J. Mol. Biol. 2012. V. 415. № 3. P. 527–537.
41. Anunciado D., Dhar A., Gruebele M., Baranger A.M. // J. Mol. Biol. 2011. V. 408. P. № 5. P. 896–908.

Carbocyclic Analogues of Inosine-5′-Monophosphate: Synthesis and Biological Activity

E.S. Matyugina[1], S.N. Andreevskaya[2], T.G. Smirnova[2], A.L. Khandazhinskaya[1*]
[1]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilova Str., 32, Moscow, Russia, 119991
[2]Central Tuberculosis Research Institute, Russian Academy of Medical Sciences, 2, Yauzskaya Alley, Moscow, Russia, 107564
*E-mail: khandazhinskaya@bk.ru
Received 03.08.2012

ABSTRACT **9-(4'-Phosphonomethoxy-2'-cyclopenten-1'-yl)hypoxanthine and 9-(4'-phosphonomethoxy-2',3'-dihydroxycyclopenten-1'-yl)hypoxanthine were synthesized as isosteric carbocyclic analogues of inosine-5'-monophosphate. The synthesized compounds were shown to be capable of inhibiting the activity of human type II inosine-5'-monophosphate dehydrogenase (IMPDH II) (IC_{50} = 500 μM) and to have no significant effects on the growth of *Mycobacterium tuberculosis*.**
KEYWORDS **Carbocyclic nucleosides; competitive inhibition; inosine-5'-monophosphate; human IMPDH II,** *Mycobacterium tuberculosis*.

INRODUCTION

Inosine monophosphate dehydrogenase (IMPDH, [EC 1.1.1.205]) is one of the key enzymes in a *de novo* biosynthesis of purine nucleotides (GTP and dGTP). Inosine-5'-monophosphate (IMP) is the natural substrate of IMPDH. IMPDH catalyzes NAD^+-dependent reactions leading to the formation of NADH and xanthosine-5'-monophosphate, which is then converted into guanosine-5'-monophosphate (GMP). The inhibition of IMPDH causes a decrease in the intracellular levels of guanine-containing nucleotides, leading to antimicrobial, antiparasitic, antiviral, anticancer, and immunodepressive effects [1, 2].

The existing inhibitors of IMPDH can be classified into 3 groups with respect to enzyme binding: analogues of IMP, analogues of NAD^+, and allosteric inhibitors. The modified nucleosides belonging to the first group undergo intracellular phosphorylation and, in the form of 5'-monophosphates, competitively interact with the IMP binding site. Both types of inhibitors exist among the analogues of IMP: reversible (ribavirin-5'-monophosphate, 3'-deazaguanosine, mizoribin) and irreversible (5'-monophosphates of 2-vinyl inosine, 6-chloropurine nucleoside, 5-ethinyl-1-ribofuranosyl-imidazole-4-carboxamide). The most widely known representatives of the second group of inhibitors include tiazofurine, selenazofurine, and mycophenolic acid.

Human IMPDH exists in 2 isoforms, types I and II, showing 84% homology. Type I enzyme is prevalent in normal lymphocytes and leucocytes; type II is found mostly in actively dividing and cancerous cells. Bacterial IMPDH molecules isolated from different sources significantly differ from the human enzyme, showing 30-41% homology. The affinity of IMPDH isolated from different sources may vary significantly for the same types of inhibitors [2]. Thus, human IMPDH type II is more sensitive to mycophenolic acid (K_i = 7nM) as compared to human IMPDH type I (K_i = 33nM). The sensitivity of bacterial enzymes to mycophenolic acid is considerably lower (K_i = 0.2–20 μM). The selectivity of IMPDH with respect to inhibitors makes this enzyme a rather attractive target for potential anticancer, antimicrobial, and antiparasitic compounds [2].

It has been shown recently that the inhibition of IMPDH isolated from *Mycobacterium tuberculosis* suppresses the growth of the bacterium [3]. The main objective in efforts to treat tuberculosis today is searching for new drugs that are effective against strains resistant to existing medicinal agents. New compounds should work via different mechanisms compared to those of the existing therapeutic agents. Hence, searching for new anti-tuberculosis agents not only among classes of antibiotics, but also among compounds of another nature seems reasonable. There are no analogues of nucleosides among the therapeutic agents currently used to treat tuberculosis;

in combination with their potential to inhibit IMPDH, it makes these analogues primary candidates for investigation as antimycobacterial agents.

The present article describes the synthesis of 9-(4'-phosphonomethoxy-2'-cyclopenten-1'-yl)hypoxanthine (1) and 9-(4'-phosphonomethoxy-2',3'-dihydroxycyclopenten-1'-yl)hypoxanthine (2) (Fig. 1), the isosteric carbocyclic analogues of IMP. The ability of these compounds to inhibit human IMPDH II and to suppress the growth of *M. tuberculosis* is also assessed.

EXPERIMENTAL

All the reagents and solvents used in the experiments are commercially available (Acros, Aldrich, and Fluka). Thin-layer chromatography (TLC) was performed on Kieselgel 60 F_{254} plates (Merck) in the following systems: 98:2 $CHCl_3$–MeOH (system A); 9:1 $CHCl_3$–MeOH (system B), 4:1 dioxane–NH_3 (system C), 7:2:1 isopropanol–NH_3–water (system D). Column chromatography was performed using a 40-63 μm Kieselgel (,Merck), a 25-40 μm Lichroprep RP-18 (Merck), and a DEAE-Toyopearl (Toysoda, Japan). The elution systems are specified below.

The UV spectra were recorded using a Shimadzu UV-1201 spectrophotometer (Japan). ^1H and ^{31}P NMR spectra were recorded on an AMX III-400 spectrometer (Bruker) with operating frequencies of 400 MHz for ^1H NMR (chemical shifts relative to the internal standards are provided: Me_4Si for organic solvents and sodium 3-(trimethylsilyl)-1-propansulfonate (DSS) for D_2O) and 162 MHz for ^{31}P NMR (with suppression of phosphorus-proton spin-spin coupling; chemical shifts relative to the external standards, 85% phosphoric acid, are provided). Chemical shifts are given in parts per million (ppm).

The starting 6-chloro-9-(4'-hydroxy-2'-cyclopenten-1'-yl)purine (3) was synthesized in accordance with the previously described methodology [4].

6-Ethoxy-9-(4'-hydroxy-2'-cyclopenten-1'-yl)purine (4).

Calcined K_2CO_3 (300 mg, 2.3 mmol) was added to a solution of 9-(4'-hydroxy-2'-cyclopent-1'-yl)-6-chloropurine (300 mg, 13 mmol) dissolved in 10 ml of ethanol; the resultant suspension was refluxed for 1 h. The course of the reaction was controlled using TLC (system A).

The solvent was removed under reduced pressure; the residue was applied onto a silica gel column; system B was used for elution; the target fractions were concentrated under reduced pressure. Product **4** (220 mg, 78%) was isolated as a white foamy substance. ^1H NMR (CD_3OD): 8.42 (1H, s, H_2), 7.95 (1H, s, H_8), 6.34–6.33 (1H, m, $H_{2'}$), 5.82 (1H, m, $H_{3'}$), 5.34–5.32 (1H, m, $H_{1'}$), 4.86 (1H, m, $H_{4'}$), 4.64–4.62 (2H, m, O-CH_2), 3.02–2.98 (1H, m, H_{a5}), 2.23–2.19 (1H, m, H_{b5}), 1.5–1.48 (3H, m, CH_3).

6-Ethoxy-9-(4'-ethylphosphonomethoxy-2'-cyclopenten-1'-yl)purine (5).

NaH (33.5 mg, 1.4 mmol) and Cs_2CO_3 (234 mg, 0.72 mmol) were added to the solution of compound **4** (230 mg, 0.93 mmol) in 5 ml of dimethylformamide (DMF) under stirring in an argon atmosphere. The reaction mixture was stirred for 1.5 h at room temperature, then ethyl ester of *p*-toluene sulfonyloxymethyl phosphonic acid (334 mg, 1.8 mmol) solution in DMF (2 ml) was added. The mixture was stirred for 12 h at room temperature. The course of the reaction was controlled using TLC (system B). After removing the solvent under reduced pressure, the residue was applied onto the DEAE-Toyopearl column and eluted with a linear gradient of NH_4HCO_3 (0–0.2 M). The target compound **5** was eluted using 0.1 M NH_4HCO_3; the fraction was concentrated, and the target product was isolated on a LiChroprep RP-18 and eluted using a linear gradient of aqueous ethanol (0–10%). The product was eluted using an 8% aqueous ethanol solution.

A total of 240 mg (67%) of compound **5** was obtained in the form of a colorless oil. ^1H NMR (D_2O): 8.14 (1H, s, H_2), 8.06 (1H, s, H_8), 6.34–6.32 (1H, m, H_2), 6.15 (1H, m, H_3), 5.35 (1H, m, H_1), 4.63 (1H, m, H_4), 4.38 (2H, m, O–CH_2), 3.76–3.72 (2H, m, O–CH_2), 3.58–3.56 (2H, m, O–CH_2–P), 2.89 (1H, m, H_{a5}), 1.80 (1H, m, H_{b5}), 1.33–1.29 (3H, m, CH_3), 1.15–1.11 (3H, m, CH_3). ^{31}P NMR (D_2O): 17.99 s.

6-Ethoxy-9-(4'-ethylphosphonomethoxy-2',3'-dihydroxycyclopent-1'-yl)purine (6).

Solutions of osmium tetroxide in dioxane (0.5 M) and N-methyl morpholine oxide (0.3 ml, 3 mmol) were added

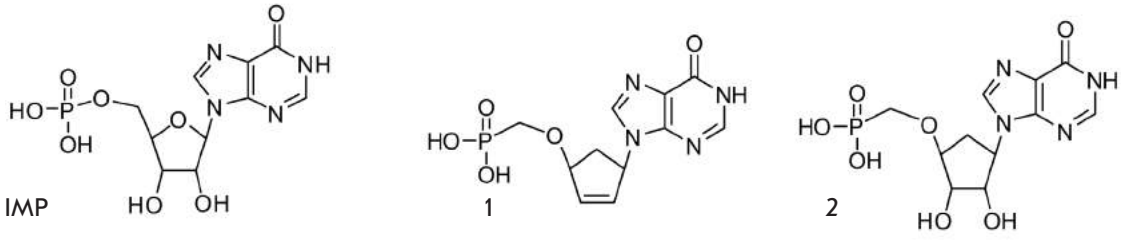


Fig. 1. Inosine-5'-monophosphate and its isosteric carbocyclic analogues

to the suspension of phosphonate **5** (200 mg, 0.54 mmol) in a 10:1 mixture of dioxane and water solvents. The solution was stirred for 3 h at room temperature. The course of the reaction was controlled using TLC (system D). After the solvent had been removed under reduced pressure, the residue was applied onto the DEAE-Toyopearl column and eluted with a linear gradient of NH_4HCO_3 (0–0.3 M), and subsequently repurified on a Lichroprep RP-18 column eluted with water. The yield of product **6** was 74%. The UV spectra (H_2O, pH 7) λ_{max} 252.0 nm (ε 9600), ^{1}H NMR (D_2O): 8.36 (1H, s, H_2), 8.29 (1H, s, H_8), 4.85 (1H, m, H_1), 4.23 (1H, m, H_4), 3.92–3.89 (2H, m, O–CH_2), 3.73 (2H, m, O–CH_2), 3.60–3.59 (2H, m, O–CH_2–P), 2.88–2.85 (1H, m, H_{a5}), 2.10 (1H, m, H_{b5}), 1.37 (3H, m, CH_3), 1.21 (3H, m, CH_3). ^{31}P NMR (D_2O): 18.23 s.

9-(4'-Phosphonomethoxy-2'-cyclopenten-1'-yl)hypoxanthine (1).

Trimethylbromosilane (0.65 ml, 5 mmol) was added to the suspension of phosphonate **5** (100 mg, 0.27 mmol) in DMF under argon atmosphere; the resulting mixture was stirred for 3 h at room temperature. The course of the reaction was controlled using TLC (system B). The reaction mixture was neutralized with 25% aqueous ammonia; the solvent was removed under reduced pressure. The residue was purified on a Lichroprep RP-18 column and eluted with water to give 70 mg (84%) of product **4** in the form of lyophilizate. ^{1}H NMR (D_2O): 8.39 (1H, s, H_2), 8.26 (1H, s, H_8), 6.44–6.42 (1H, m, $H_{2'}$), 6.18–6.17 (1H, m, H_3), 5.57–5.55 (1H, m, H_1), 4.81 (1H, m, H_4), 3.61 (2H, m, O–CH_2–P), 3.04 (1H, m, H_{a5}), 1.93 (1H, m, H_{b5}). ^{31}P NMR (D_2O): 16.66 s.

9-(4'-phosphonomethoxy-2',3'-dihydroxycyclopent-1'-yl)hypoxanthine (2).

The compound **2** was obtained in a similar fashion to compound **1,** obtained from compound **6** (140 mg, 0.35 mmol). A total of 105 mg (81%) of the product was isolated as lyophilizate. The UV spectra (H_2O, pH 7) λ_{max} 251.0 nm (ε 9300). ^{1}H NMR (D_2O): 8.27 (1H, s, H_2), 8.11 (1H, s, H_8), 4.20 (1H, m, H_1), 3.93 (1H, m, H_4), 3.53–3.51 (2H, m, O–CH_2–P), 2.81 (1H, m, H_{a5}), 2.07 (1H, m, H_{b5}). ^{31}P NMR (D_2O): 14.06 s.

BIOLOGICAL TESTS

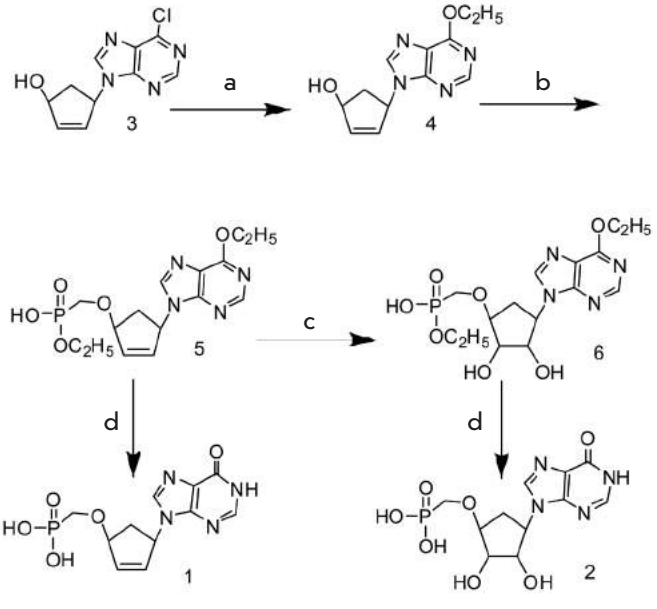
Experiments on the ability of the synthesized compounds to inhibit human IMPDH II were conducted by NovoCib company (France). Compounds **1** and **2** were tested on a human recombinant IMPDH II (~0.0003 activity units per well) at 37°C in 200 µl of a buffer solution (KH_2PO_4 0.1 M, pH 7.8, NAD 250 µM, DTT 2 mM) using a 96-well microplate. The reaction was initiated by the addition of a substrate, IMP, at a concentration of 250 µM. Prior to reaction initiation, the compounds were incubated in a buffer with IMPDH II for 5 minutes. The absorbance was measured at 340 nm using an iEMS Reader MF device (Labsystems, Finland). Ribavirin was used as a positive control. The influence of the synthesized compounds on human IMPDH II activity was simultaneously tested in two identical experiments.

Antituberculosis activity.

The compounds were tested on a laboratory strain of *M. tuberculosis* H37Rv sensitive to antituberculosis drugs. The mycobacteria was transferred into a single-cell suspension of single cells at the same growth phase and standardized with respect to CFU [5]. The enriched liquid growth medium Dubois (Difco, USA) was used.

Evaluation of compound efficacy.

The effect of the compounds on the growth of mycobacterial strains was estimated using a Bactec MGIT 960 automated growth detection system (BD, USA). The suspension of mycobacterial cells (500 µl) was inoculated in a nutrient medium (7.9 ml). The final concentration of *M. tuberculosis* in the sample was 10^6 CFU/ml. Three replicates for each sample (concentration) were analyzed, including the control samples. The antimicrobial activity of the compounds was evaluated using the proportion method with the TB Exist software [6]. The growth of mycobacterial cells cul-



Scheme a – Ethanol (pure), K_2CO_3, refluxing for 1 hour; b – NaH, $TosOCH_2P(O)OC_2H_5OH$, Cs_2CO_3, DMF; c – OsO_4, NMMO, dioxane; d - $(CH_3)_3SiBr$, DMF

tured in the presence of the compound and the growth of the control culture diluted 100 times as compared to the test sample is assessed in this analysis. The culture is considered to be sensitive to such a concentration of the compound at which the growth rates in the experiment do not exceed 100 growth units (GU), when 400 GU are recorded in the control sample; the compound is regarded as active in this case. Furthermore, the absolute concentration method was used to assess the effect of the compounds at concentrations lower than the minimum inhibitory concentration (MIC) on the viability of mycobacterial cells on the basis of the inhibition of bacterial growth as compared to the control. Bacterial growth was determined automatically at 1 h intervals and recorded using the Epicenter software (BD, USA).

RESULTS AND DISCUSSION

Over the past decades, carbocyclic nucleosides have been intensively studied. These compounds have been found to be biologically active; in particular, they turn out to have antiviral and anticancer properties [7]. These nucleosides are recognized by many enzymes and receptors, since their structure is similar to that of natural nucleosides. Meanwhile, they are highly resistant to C-N bond cleavage by phosphorylases and hydrolases.

Hydroxycyclopentene is used as a carbocyclic moiety in compounds **1** and **2**. The analogues with such moiety are known as 5'-norcarbocyclic nucleosides. The substitution of the primary 5'-hydroxyl residue for the secondary 4'-hydroxyl results in toxicity decrease due to the loss of substrate properties with respect to cellular kinases. Taking into account the fact that intracellular phosphorylation of 5'-norcarbocyclic nucleosides is infeasible, we synthesized the methylene phosphonates of 9-(4'-hydroxy-2'-cyclopenten-1'-yl)hypoxanthine and 9-(4',2',3'-trihydroxycyclopent-1'-yl)hypoxanthine (*scheme*). It was previously shown that such isosteric phosphonates imitate the corresponding monophosphates but are more stable to the action of hydrolyzing and dephosphorylizing enzymes [8].

9-(4'-Hydroxy-2'-cyclopenten-1'-yl)-6-chloropurine **3** was obtained via condensation of epoxycyclopentene and 6-chloropurine in accordance with the previously described procedure [4].Refluxing of compound **3** in ethanol led to the formation of ester **4**, which subsequently reacted with the ethyl esters of tosyloxymethylphosphonic or iodomethylphosphonic acid, to yield monophosphonate **5**.

Ester **5** was hydrolyzed to 9-(4'-phosphonylmethyloxy-2'-cyclopenten-1'-yl)hypoxanthine **1** using excess amounts of trimethylbromosilane. In order to obtain monophosphonate **2**, the double bond of compound **5** was oxidized with osmium tetroxide in the presence of N-methylmorpholine-N-oxide, and the ethyl groups were subsequently removed from compound **6** in the presence of trimethylbromosilane (*scheme*). The target compounds **1** and **2** were purified on a DEAE-Toyopearl column eluted with a linear concentration gradient of NH_4HCO_3. The subsequent purification and removal of salts was performed on a Lichroprep RP-18 column. The final yields were 84 and 81%, respectively.

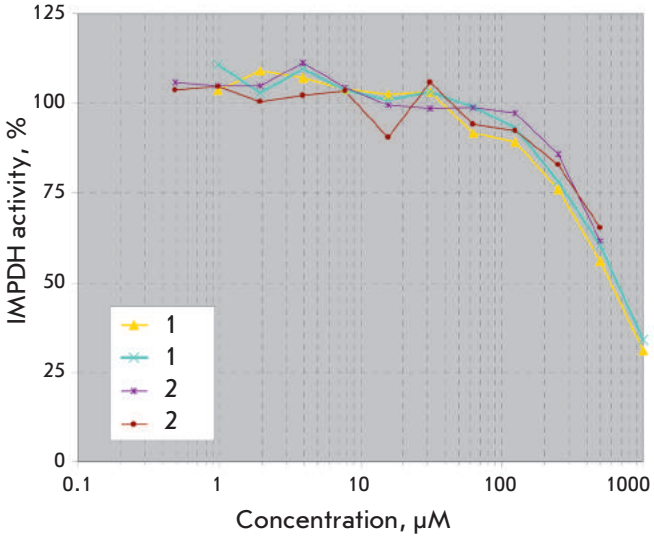


Fig. 2. Dose-dependent inhibition of IMPDH II by compounds **1** and **2**

Compounds **1** and **2** were tested as human IMPDH II inhibitors (Fig. 2).

It is clear from *Fig. 2* that the carbocylic analogue **1** at a concentration of 500 μM inhibited enzymatic activity by 50% (K_i 474 μM), whereas compound **2** did so by 35–39% (K_i 975 μM). Ribavirin monophosphate was used as the control and at a concentration of 2 μM inhibited the enzymatic activity by 50%; the K_m value of the IMP (natural substrate) in this system was 124.4 μM.

The ability of the monophosphonates **1** and **2** to inhibit the growth of a *M. tuberculosis* was also tested. The growth of *M. tuberculosis* culture H37Rv under the action of compounds **1** and **2** at concentrations of 2–100 μg/ml (5–320 μM) was identical to that observed in the control group: the initial phases of culture growth were detected on day 7; entry into stationary phase was detected on day 17. The duration of active bacterial growth is 10 days. Compound **2**, at concentrations of 200 μg/ml (578 μM), caused an insignificant delay (2 days) in bacterial growth as compared to the control group.

CONCLUSIONS

Synthesized 9-(4'-phosphonomethoxy-2'-cyclopenten-1'-yl)hypoxanthine and 9-(4'-phosphonomethoxy-2',3'-dihydroxycyclopent-1'-yl)hypoxanthine are weak inhibitors of human IMPDH II. These compounds at concentrations of 20–200 μg/ml do not affect the growth of *M. tuberculosis* H37Rv *in vitro*. This fact can be attributed both to the structural features of the mycobacterial cell wall and, hence, the difficulties associated with penetrating the membrane, or to the existence of alternative pathways for synthesizing essential compounds in mycobacteria. The hypothesis that IMPDH of *M. tuberculosis* could be less sensitive to the compounds under study compared with human IMPDH II should not be dismissed, either. ●

The authors would like to acknowledge the contribution from L. Balakireva and N. Godard from NovoCib company (Lyon, France) for biological tests of the synthesized compounds with the recombinant human IMPDH II.

This work was supported by the Russian Foundation for Fundamental Research (grants № 11-04-12035-ofi-m, 11-04-00603) and Russian Academy of Sciences Presidium Program on Molecular and Cellular Biology.

REFERENCES

1. Nair V., Shu Q. // Antiviral. Chem. Chemotherapy. 2007. V. 18. P. 245–258.
2. Shu Q., Nair V. // Med. Res. Rev. 2008. V. 28. № 2 P. 219–232.
3. Usha V., Hobrath J.V., Gurcha S.S., Reynolds R.C., Besra G.S. // PLoS One. 2012. V. 7(3):e33886. Epub 2012 Mar 29.
4. Khandazhinskaya A.L., Shirokova E.A., Shipitsin A.V., Karpenko I.L, Belanov E.F., Kukhanova M.K., YaskoM.V. // Collect. Chech. Chem. Commun. 2006. V. 71. P. 1107–1121.
5. Andreevskaya S.N., Chernousova L.N., Smirnova T.G., Larionova E.E., Kuzmin A.V. // Problems of Tuberculosis and Lung Disease (Russian). 2006. V. 12. P. 43–48.
6. Siddiqi S.H., Rusch–Gerdes S. 2006. MGIT procedure manual for Bactec MGIT 960 TB system. Foundation for Innovative New Diagnostics, Geneva, Switzerland.
7. Matyugina E.S., Khandazhinskaya A.L., Kochetkov S.N. // Russian Chemical Reviews. 2012. V. 81. № 8. P. 729–746.
8. Wu T, Froeyen M., Kempeneers V., Pannecouque C., Wang J., Busson R., De Clercq E., Herdewijn P. // J Am. Chem Soc. 2005 V. 127. №14. P. 5056–65.

Gold Nanoparticle Clusters in Quasinematic Layers of Liquid-Crystalline Dispersion Particles of Double-Stranded Nucleic Acids

Yu. M. Yevdokimov[1]*, V. I. Salyanov[1], E. I. Katz[2], S. G. Skuridin[1]
[1]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilova Str., 32, Moscow, Russia, 119991
[2]Landau Institute for Theoretical Physics, Russian Academy of Sciences, Kosygina Str. 2, Moscow, Russia, 119334
*E-mail: yevdokim@eimb.ru
Received 13.06.2012

ABSTRACT **The interaction between gold nanoparticles and particles of cholesteric liquid-crystalline dispersions formed by double-stranded DNA and poly(I)×poly(C) molecules is considered. It is shown that small-sized (~ 2 nm) gold nanoparticles induce two different structural processes. First, they facilitate the reorganization of the spatial cholesteric structure of the particles into a nematic one. This process is accompanied by a fast decrease in the amplitude of an abnormal band in the CD spectrum. Second, they induce cluster formation in a "free space" between neighboring nucleic acid molecules fixed in the structure of the quasinematic layers of liquid-crystalline particles. This process is accompanied by slow development of the surface plasmon resonance band in the visible region of the absorption spectrum. Various factors influencing these processes are outlined. Some assumptions concerning the possible mechanism(s) of fixation of gold nanoparticles between the neighboring double-stranded nucleic acid molecules in quasinematic layers are formulated.**
KEYWORDS **DNA; poly(I)×poly(C); liquid-crystalline dispersions of nucleic acids; gold nanoparticles; circular dichroism; absorption spectroscopy; abnormal optical activity; surface plasmon resonance; structure of biopolymer lyotropic liquid crystals; cytotoxicity of nanoparticles.**
ABBREVIATIONS **DAU – daunomycin; CD – circular dichroism; Au nanoparticles (nano-Au) – gold nanoparticles; SPR – surface plasmon resonance; PEG – poly(ethylene glycol); UV region – ultraviolet region; CLCD – cholesteric liquid-crystalline dispersion.**

INTRODUCTION

Metal and metal oxide nanoparticles are known to be characterized by their inherent ability to exhibit specific properties depending on the nanoparticle's size. These properties of nanoparticles differ substantially from those typical of a "bulky" sample of the initial material. Nano-sized gold (Au) nanoparticles that are used both for research and applied purposes [1] (in particular, for diagnosis and treatment of certain diseases [2, 3]) are among the most vivid examples of the existence of such differences. Although the *in vitro* and *in vivo* cytotoxicity of Au nanoparticles has been investigated by several research teams, the data pertaining to the biological effects induced by Au nanoparticles are rather controversial [4, 5]. It is quite possible that the reason for this is that different biological systems have been used to study the effect of nanoparticles; in this case, it is difficult to compare their action mechanisms.

The data [3, 6] provide a background to assume that the *in vitro* and *in vivo* action of Au nanoparticles on spatially arranged DNA structures is similar to that of molecules that possess mutagenic activity. Particles of DNA cholesteric liquid-crystalline dispersion (CLCD) are known to be among the structures that model certain spatial features of DNA within biological objects [7]. Indeed, the physicochemical features of DNA CLCD particles indicate some properties, which are characteristic of Protozoan chromosomes (e.g., chromosomes of Dinoflagellate, etc.) and DNA-containing bacteriophages [8–10].

Hence, DNA CLCD is a system of undoubted interest both in terms of nano- and biotechnologies.

When studying the effect of Au nanoparticles on various biological macromolecules and systems, several facts should be borne in mind. Au nanoparticles, especially the small-sized ones, tend to spontaneously

aggregate in water–salt solutions [1, 11, 12] and to form various complexes and aggregates with the solution components and dissolved macromolecules [13–16]. This process, accompanied by the approaching of neighboring Au nanoparticles, results not only in the enhancement of the so-called surface plasmon resonance (SPR) band typical of individual Au nanoparticles, but also in excitation of the collective vibrations of the electronic system and interaction between the neighboring "plasmons." The latter effect, known as plasmon overlapping, is accompanied [1, 17, 18] by a shift of the SPR band toward the shorter or longer wavelengths of the absorption spectrum depending on a number of parameters (interparticle distance, size and shape of the resulting aggregates, dielectric permittivity of the medium [19, 20], existence of "interlayers" between the neighboring Au nanoparticles [21, 22], etc.). It is obvious that the complex formation (and possible aggregation of neighboring Au nanoparticles) is dependent on the concentration and charge of Au nanoparticles, their size, and the properties of the solvent components. This means that when studying the interaction between Au nanoparticles and biopolymer molecules, control experiments are to be carried out which would prove the absence of "parasitic" optical effects induced by the formation of nonspecific aggregates between Au nanoparticles and the solvent components under the conditions used.

Hence, this work was aimed not only at proving the fact that there are no unnspecific aggregates between Au nanoparticles and the solvent components, but also at analyzing the interaction between Au nanoparticles and the double-stranded DNA molecules fixed in the spatial structure of the CLCD particles formed by phase exclusion of DNA molecules from water–salt solutions.

MATERIALS AND METHODS

Colloid gold solutions (hydrosols) containing spherical nanoparticles of different sizes were used in this study. Au nanoparticles were synthesized according to the previously described procedures [23–25]. The first hydrosol was obtained using the procedure [23] and contained Au nanoparticles with a mean diameter of ~15 nm. The second hydrosol containing Au nanoparticles 5.5 nm in diameter was synthesized according to [24]. Finally, the third hydrosol containing quasi-metallic Au nanoparticles 2–3 nm in diameter was obtained according to the procedure described in [25]. The mean size of the Au nanoparticles in the initial solutions was determined via dynamic light scattering and electron microscopy. The numerical concentration of Au nanoparticles in the first, second, and third hydrosols was 10^{12}, 10^{13}, and 10^{15} particles/cm^3, respectively.

The Au nanoparticles were negatively charged; their ξ-potentials were as follows: for 2–3 cm particles, −18 ± 7 mV (immediately after synthesis), −25 ± 5 mV (2 days after the synthesis) and −38 ± 5 mV (9 months after the synthesis); for 5 nm particles, −32 ± 4 mV; for 15 nm particles, −44 ± 3 mV.

The original solutions of Au nanoparticles were stored at 4°C in light-impermeable containers and used 2.5 months following the synthesis.

A calf thymus depolymerized DNA (Sigma, USA) with a molecular mass of (0.3–0.7) × 10^6 Da after additional purification was used. A synthetic double-stranded polyribonucleotide poly(I)×poly(C) (Sigma, USA; lot 023K4032) was used without additional purification. DNA and poly(I)×poly(C) concentrations in the water–salt solutions were determined spectrophotometrically using the known values of the molar extinction coefficients (ε_{max} = 6,600 $M^{-1}\times cm^{-1}$ for DNA and ε_{max} = 4,900 $M^{-1}\times cm^{-1}$ for poly(I)×poly(C)).

Poly(ethylene glycol) samples (PEG; Serva, Germany; molecular mass of 4,000 Da) were used without additional purification.

The absorption spectra were taken by Cary 100 Scan (Varian, USA) spectrophotometer. The circular dichroism (CD) spectra were recorded using an SKD-2 portable dichrometer. The CD spectra were represented as a dependence of the difference between the intensities of absorption of left- and right-handed polarized light (ΔA; $\Delta A = (A_L - A_R)$) on the wavelength (λ).

CLCD of DNA in PEG-containing water–salt solutions were prepared according to the previously described procedure [7].

A series of control experiments were carried rut to check the possible interaction between Au nanoparticles and biopolymer molecules (nucleic acids and proteins).

As has already been mentioned in Introduction, a number of questions pertaining to the behavior of negatively charged small-sized Au nanoparticles under the conditions used were to be answered. Are these Au nanoparticles capable of:

a) forming aggregates in solutions of low or high ionic strength;

b) interacting (form complexes) with a neutral polymer (PEG) used to form DNA CLCD particles;

c) affecting single-stranded nucleic acid molecules in low- or high-ionic-strength solutions; and

d) affecting double-stranded DNA molecules under conditions that prevent dispersion formation in a PEG-containing water–salt solution.

Absorption spectra

The absorption spectra of Au nanoparticles recorded at different times after PEG (C_{PEG} = 150 mg/ml) addition

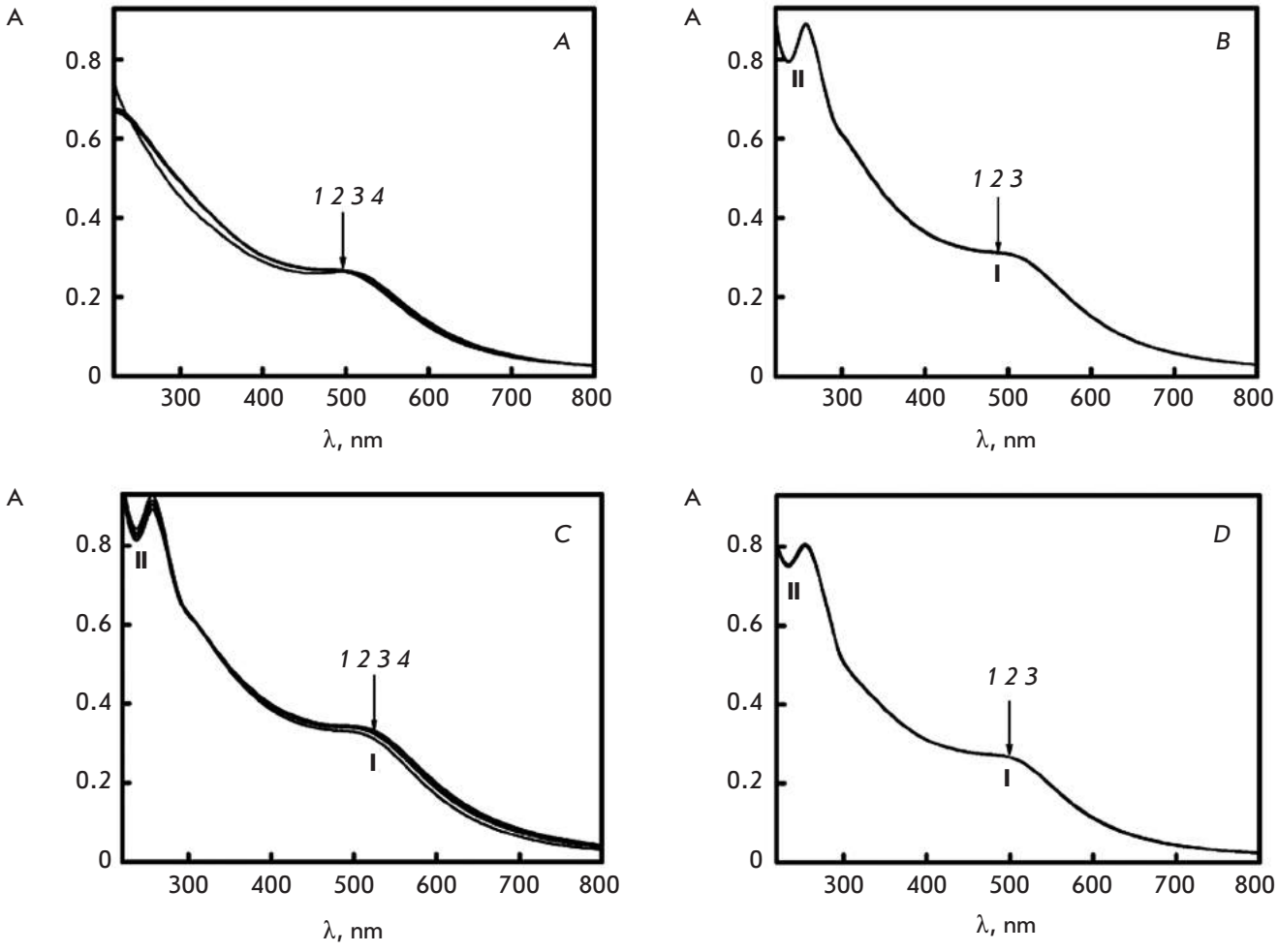


Fig. 1. The absorption spectra of Au nanoparticles under various conditions.
A. The absorption spectra of Au nanoparticles in a PEG-containing water–salt solution (curves *1–4*): the PEG-containing solution was treated with Au nanoparticles for: *1* – 1 min; *2* – 6 min; *3* – 17 min; *4* – 280 min; C_{PEG} = 150 mg/ml; 0.27 M NaCl + 1.78×10^{-3} M Na^+-phosphate buffer; $C_{Nano\text{-}Au} = 0.82\times10^{14}$ particles/ml.
B. The absorption spectra of Au nanoparticles in a water–salt solution of a single-stranded polynucleotide (polyA) (curves *1–3*): poly(A)-water–salt solution treated with Au nanoparticles for: *1* – 1 min; *2* – 17 min; *3* – 30 min; $C_{Poly(A)}$ = 9 μg/ml; refer to *Fig. 1A* for the other conditions.
C. The absorption spectra of Au nanoparticles in a PEG-containing water–salt solution of a single-stranded polynucleotide (polyA) (curves *1–4*): the poly(A)-PEG-containing solution was treated with Au nanoparticles for: *1* – 1 min; *2* – 7 min; *3* – 15 min; *4* – 30 min; $C_{Poly(A)}$ = 9 μg/ml; refer to *Fig. 1A* for the other conditions.
D. The absorption spectra of Au nanoparticles in a water-salt-DNA-PEG-containing solution of low ionic strength (curves *1–3*): the DNA-PEG-containing solution was treated with Au nanoparticles for: 1 – 1 min; 2 – 17 min; 3 – 180 min; C_{DNA} = 9 μg/ml; 0.0009 M NaCl; refer to *Fig. 1A* for the other conditions

to the solution are compared in *Fig. 1A*. It is clear that the absorption spectrum is characterized by a poorly pronounced band (I) at λ ~ 500 nm and a broadband in the short wave spectral region, which is caused by electron transitions both between the d orbitals and the sp hybridized orbitals of Au [26]. The amplitude constancy of the band at λ ~ 500 nm in the absorption spectrum and the absence of either red or blue shifts in this band unambiguously attest to the fact that negatively charged small-sized Au nanoparticles do not tend to aggregate near the surface of PEG molecules under the conditions used.

Figure. 1B shows the absorption spectra recorded at different time intervals after Au nanoparticles addition to the water–salt solution of synthetic single-stranded polynucleotide poly(A). *Figure. 1C* shows the absorption

spectra recorded after Au nanoparticles were added to a PEG-containing (C_{PEG} = 150 mg/ml) water–salt solution of the same biopolymer. There are two bands in the absorption spectra in *Figs. 1B,C*: the band in the UV region of spectrum (**I**) corresponds to Au nanoparticles; the band in the UV region of spectrum (**II**) contains the contribution of the absorption of chromophores in the polynucleotide. The position of these bands and their maxima do not change over time after Au nanoparticles are added to the solutions.

The absorption spectra of Au nanoparticles recorded at different time intervals after the Au nanoparticles were added to the water–polymer solution (C_{PEG} = 150 mg/ml) of low ionic strength containing double-stranded DNA molecules are shown in *Fig. 1D*. The absorption spectrum contains two bands; the band in the visible region of spectrum (**I**) corresponds to Au nanoparticles, whereas that in the UV region of spectrum (**II**) corresponds to absorption of DNA chromophores. Phase separation of double-stranded DNA molecules does not happen under the conditions used (ionic strength 0.001 and C_{PEG} = 150 mg/ml); thus, no DNA CLCD are formed. No changes in the amplitudes of both bands are observed under these conditions.

Circular dichroism spectra

The CD spectra of water–salt solutions containing linear double-stranded DNA or poly(I)×poly(C) molecules attest to the fact that treatment of these molecule with Au nanoparticles causes no optical changes in them (spectra are not shown).

Thus, the absence of any noticeable changes in the amplitude and position of the 500 nm band in the absorption spectra shown in *Fig. 1A* and in the CD spectra indicates that small-sized negatively charged Au nanoparticles neither undergo aggregation in aqueous solutions of low or high ionic strength nor form aggregates near PEG molecules under the selected conditions. Moreover, no changes in the amplitudes of the bands characterizing the optical properties of nitrogen bases or small-sized Au nanoparticles are observed under conditions when there is no phase separation of single-stranded polynucleotide molecules (*Fig. 1C*) or double-stranded DNA (*Fig. 1D*) and a biopolymer molecule dispersion is not formed [7].

The influence of small-sized Au nanoparticles on double-stranded DNA and the poly(I)×poly(C) molecules fixed in the spatial structure of CLCD particles has been investigated with allowance for the results of control experiments.

RESULTS AND DISCUSSION

Before analyzing the effect of Au nanoparticles on double-stranded DNA and the poly(I)×poly(C) molecules fixed in the spatial structure of CLCD particles, let's provide some illustrations of the structure of the initial liquid-crystalline dispersion particles. In physicochemical terms, each particle in the dispersion is a "droplet" of a concentrated DNA solution, whose structure and properties are determined by the osmotic pressure of the solution [7]. A "droplet" cannot be held in one's hands or immobilized on a substrate, since the "droplet" structure will change without the osmotic pressure of the solution, and DNA molecules will be converted from their condensed into an isotropic state. Each CLCD particle consists of double-stranded nucleic acid molecules forming its neighboring (so-called quasinematic) layers [7]. *Fiure. 2* illustrates certain features of the quasinematic layer consisting of ordered neighboring double-stranded molecules of nuclear acids (in particular, DNA). In the case of phase separation, the dispersion particles (hence, the quasinematic layer as well) do not contain molecules of a water-soluble polymer (PEG) molecule. There is "free space" both between the neighboring DNA molecules in the same layer and between the DNA molecules in the neighboring layers. The distance between two neighboring DNA molecules in a layer (d) can vary within the 2.5–5.0 nm range, depending on the osmotic pressure of the solution. Under the conditions used (C_{PEG} = 150 and 170 mg/ml), the distance between two DNA molecules determined via an X-ray diffraction analysis of the phases obtained by low-speed precipitation of the initial DNA CLCD particles [7] was 3.6 and 3.2 nm, respectively. DNA molecules ordered in layers retain almost all their diffusion degrees of freedom. Due to the anisotropic properties of DNA molecules, each subsequent quasinematic layer is rotated by a certain angle (approximately 0.1° [7]) with respect to the previous one. The rotation gives rise to the helical (cholesteric) structure of a liquid-crystalline dispersion particle. The emergence of this structure can be easily detected according to the abnormal optical activity manifested as a characteristic intense band in the CD spectrum in the region of absorption of DNA chromophores (nitrogen bases). High local concentration of DNA and the ordered arrangement of these macromolecules in a layer provide conditions for a rapid interaction between molecules of various low-molecular-mass compounds ("guests") with DNA molecules (intercalation between base pairs, fixation in the grooves on the molecule surface, etc.). The distortion of the secondary DNA structure accompanying this interaction affects not only the properties of all quasinematic layers, but also the character of the interaction between them (hence, the structural features of any CLCD particle and its properties as well). Since the properties of the quasinematic layer(s) are determined by the physicochemical properties of DNA CLCD par-

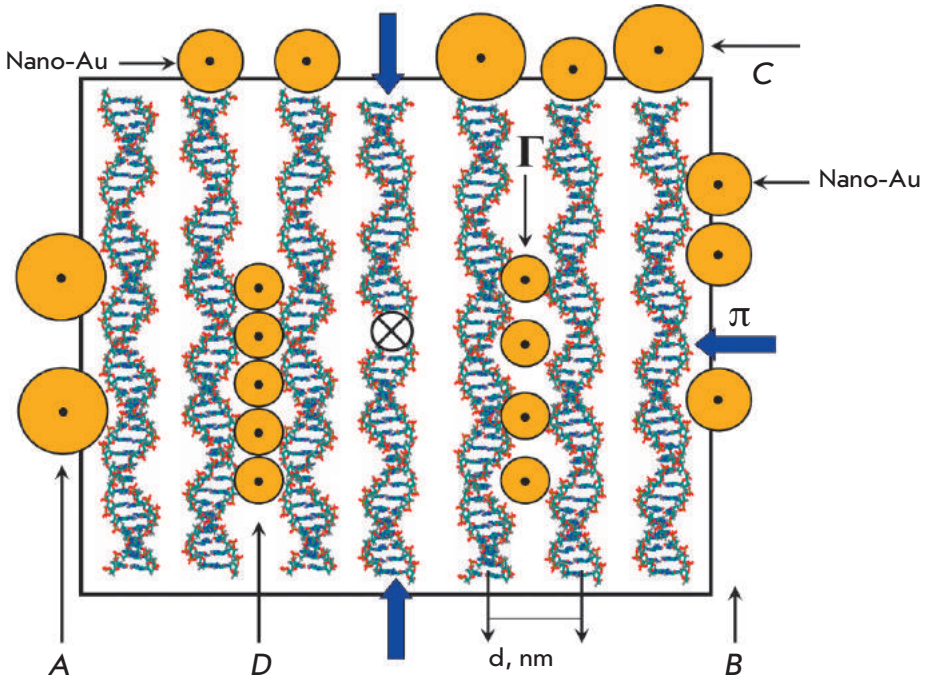


Fig. 2. A hypothetical scheme of the arrangement of Au nanoparticles of different sizes (*A*–*D*) near the DNA molecules forming the quasinematic layer. The frame and wide arrows indicate the presence of osmotic pressure (π) in the PEG-containing solution; *d* – distance between the axes of the neighboring DNA molecules

ticles, we will use this very term when reporting further results. Finally, complete separation of the chains of double-stranded DNA molecules in a quasinematic layer and their folding into individual random coils is infeasible for steric reasons [27, 28].

These features of the quasinematic layer allow to hypothesize about the possible mechanisms of fixation of Au nanoparticles ("guests") near the double-stranded DNA molecules of the quasinematic layer (*Fig. 2*).

First, Au nanoparticles of any size (*Figs. 2A–C*) can interact both with the "surface" DNA molecules and with the base pairs of the terminal groups of DNA molecules in the quasinematic layers, thus forming complexes (ensembles) with them [13, 29–31].

Second, it is quite possible that Au nanoparticles, whose size is comparable to the distance between the DNA molecules in the quasinematic layer, can diffuse inside the layers (*Fig. 2D*), interact with the neighboring DNA molecules within the same quasinematic layer or neighboring quasinematic layers, and form linear clusters.

One can assume that binding even of a small number of negatively charged Au nanoparticles to DNA molecules (in particular, to the terminal groups in these molecules) results in dipole formation (it should be mentioned there is no need for penetration of Au particles into the quasinematic layer). Dipoles from the neighboring (DNA–Au) complexes within a quasinematic layer, as well as the layers, will tend to be organized in parallel fashion, which can eventually induce a change in the helical twisting of the neighboring quasinematic layers made of DNA molecules. The twist angle between these layers (~ 0.1° [7]) can fall to zero, which is equivalent to untwisting of the cholesteric helical structure, and this process will manifest itself as the attenuation of the abnormal band in the CD spectrum of liquid-crystalline dispersion particles.

It is obvious that although it has no significant effect on the forces (sterical, etc.) that determine the tendency of the neighboring DNA molecules to organize in a parallel fashion, even a small number of negatively charged Au nanoparticles can induce changes in the contributions (in particular, anisotropic contribution to the van der Waals interaction) that control the helical twisting of the neighboring quasinematic layers of DNA molecules. In this case, the helical twisting of the neighboring quasinematic layers will be disturbed and the twist angle between these layers (~ 0.1° [7]) can be equal to zero, which is equivalent to untwisting of the cholesteric helical structure accompanied by attenuation of the abnormal band in the CD spectrum of liquid-crystalline dispersion particles.

Therefore, it can be expected that if negatively charged Au nanoparticles somehow interact with double-stranded DNA molecules in CLCD particles, this interaction will be accompanied by changes in the abnormal optical activity, which is characteristic for this dispersion.

It is also quite possible that when neighboring Au particles localize near DNA molecules in a certain fash-

ion, interaction between these nanoparticles can result in the emergence of a surface plasmon resonance band in the absorption spectrum [1, 13, 19].

Changes in circular dichroism spectra caused by the treatment of DNA CLCD particles with Au nanoparticles

Treatment of DNA CLCD particles with Au nanoparticles results in a decrease in the amplitude of the abnormal negative band in the CD spectrum (*Fig. 3*). The fact that the band has a negative sign indicates that right-handed helical double-stranded DNA molecules give rise to a left-handed helical structure of the CLCD particles [7].

Due to the effect of Au nanoparticles, the amplitude of the abnormal band in the CD spectrum of DNA CLCD decreases within a rather short period of time. The decrease in the amplitude of the abnormal band in the CD spectrum of DNA CLCD particles becomes pronouncedly stronger as the concentration of Au nanoparticles in the solution increases. It should be mentioned that noticeable changes in the amplitude of the abnormal band in the CD spectrum of DNA CLCD starts at some critical concentration of Au nanoparticles in a solution of approximately 1,000 Au nanoparticles per DNA CLCD particle (*Fig. 3, inset*).

Similar data characterizing the decrease in the abnormal band in the CD spectrum of CLCD formed by synthetic double-stranded poly(I)×poly(C) molecules caused by treatment with Au nanoparticles were presented in [6]. It should be mentioned that the emergence of a positive band in the CD spectrum of this CLCD attests to the fact that the right-handed helices of double-stranded poly(I)×poly(C) molecules form CLCD particles with right-handed twisting of their spatial helical structure.

The rapid decrease in the amplitude of the band in the CD spectrum of DNA CLCD depends on the size of Au nanoparticles. In particular, if Au nanoparticles are 2 nm in diameter, the amplitude of the abnormal band in the CD spectrum decreases by 75%, whereas when 15-nm diameter nanoparticles are used, it decreases by only 20% [32].

The decrease in the amplitude of the band in the CD spectrum of DNA CLCD is also dependent on the temperature of the solution where the dispersion particles are treated with Au nanoparticles [32].

In combination with the differences in the efficiency of the changes in the CD spectrum for nanoparticles of different sizes, the scheme shown in *Fig. 2* allows to assume that there are two reasons for the decrease in the abnormal optical activity of DNA CLCD or poly(I)×poly(C) CLCD particles. First, individual Au nanoparticles of any size (*Figs. 2A–C*) can interact with the "surface" DNA molecules to yield complexes or linear ensembles (clusters). In this case, small-sized Au nanoparticles can localize in the grooves of the "surface" DNA molecules [31, 33] or form complexes with pairs of DNA nitrogen bases (in particular, with N7 atoms of purines [34–37]). Second, Au nanoparticles whose sizes are comparable to the distance between the DNA molecules in quasinematic layers can diffuse inside the layers to interact with DNA molecules. It is important to mention two aspects here. 1) It was found as early as in the first experiments [13, 38, 39] that Au nanoparticles can form ensembles near the surface of linear single-stranded DNA molecules. Ensemble formation from Au nanoparticles was subsequently shown to be accompanied by the formation of planar suprastructures consisting of repeating double-stranded DNA molecules and Au nanoparticles. These results demonstrate unambiguously that, after they interact with Au nanoparticles, DNA molecules tend to form planar suprastructures [30, 39, 40], despite the fact that the original DNA molecules possess anisotropic properties [7]. 2) In case of CLCD particles of double-stranded DNA molecules, formation of an ensemble even of a small number of Au nanoparticles on "surface" DNA molecules or near DNA molecules in quasinematic layers will result in changes in the character of the interaction between neighboring quasinematic layers. This can result in the attenuation of the helical twisting of the neighboring layers; i.e., the spatial helical structure of CLCD particles will untwist.

With allowance for the formation of planar structures considered above, it can be stated that Au nanoparticles (in case of CLCD particles) initiate a parallel (rather than helical) arrangement of the neighboring quasinematic layers of DNA molecules.

Regardless of the aforementioned reasons, combination of the control experiments (*Fig. 1*) and the results obtained (*Fig. 3*) allows one to suggest that the action of Au nanoparticles is directed towards the double-stranded DNA molecules fixed in the CLCD particles. Meanwhile, the rapid decrease in the abnormal band in the CD spectrum can be attributed to binding of an appreciably small number of Au nanoparticles to the DNA molecules in CLCD particles. This process is accompanied by the disturbance of the helical mode of ordering in the neighboring quasinematic layers; i.e., Au nanoparticles induce a transition similar to the known cholesteric → nematic transition [7].

Thus, the changes in the CD spectra of DNA CLCD (or poly(I)×poly(C) CLCD) indicate that Au nanoparticles of different sizes can interact with the double-stranded molecules of nucleic acids or synthetic polynucleotides within CLCD particles (the efficiency of the interaction may vary), although most of the de-

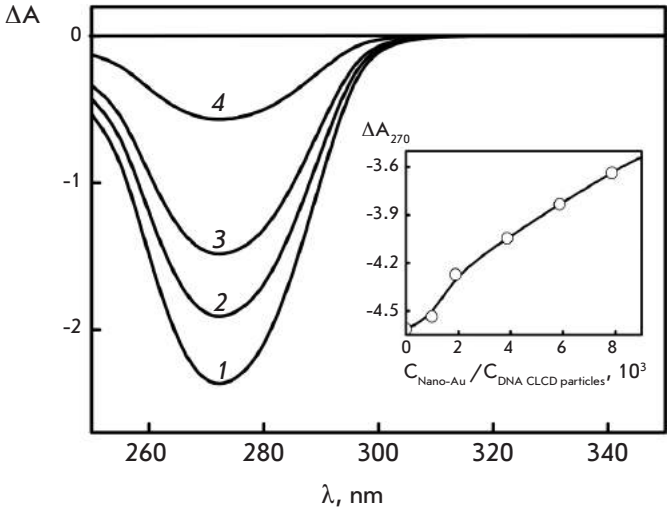


Fig. 3. The CD spectra of DNA CLCD treated with Au nanoparticles (2 nm): *1* – $C_{Nano-Au} = 0$; *2* – $C_{Nano-Au} = 0.07 \times 10^{14}$ particles/ml; *3* – $C_{Nano-Au} = 0.26 \times 10^{14}$ particles/ml; *4* – $C_{Nano-Au} = 0.82 \times 10^{14}$ particles/ml. (Treatment time – 3 h). $C_{DNA} = 9$ μg/ml; refer to *Fig. 1A* for the other conditions. $\Delta A = (A_L - A_R) \times 10^{-3}$ opt. units; L = 1 cm. Inset: the dependence of the ΔA_{270} value on the $C_{Nano-Au}/C_{DNA\ CLCD\ particles}$ ratio obtained for the solution ($C_{PEG} = 170$ mg/ml) providing maximum abnormal optical activity of DNA CLCD particles is shown as an example

tails of the mechanism underlying the interaction remain unclear.

Changes in the absorption spectra caused by the treatment of DNA CLCD particles with Au nanoparticles

The analysis of the absorption spectra of Au nanoparticles permits an assessment of the size of the ensembles formed by these particles under various conditions [41–44].

Noticeable changes both in the visible and in the UV spectral regions are observed after DNA CLCD particles are treated with small-sized Au nanoparticles (*Fig. 4A*). This treatment is primarily accompanied by changes in band (**I**) at 550 nm (SPR band) [41, 42]. *Figure 4B* shows the data obtained by treating CLCD formed by poly(I)×poly(C) molecules (their particles are characterized by left-handed twisting of the spatial structure) with Au nanoparticles. It is clear that treatment with Au nanoparticles in this case is also accompanied by the development of the plasmon effect.

The emergence of the SPR band is responsible for the pink-violet color of the solution containing DNA CLCD or poly(I)×poly(C) CLCD and treated with Au nanoparticles. The control experiments (*Fig. 1*) have demonstrated that the band at ~505 nm is poorly pronounced in the absorption spectrum of Au nanopar-

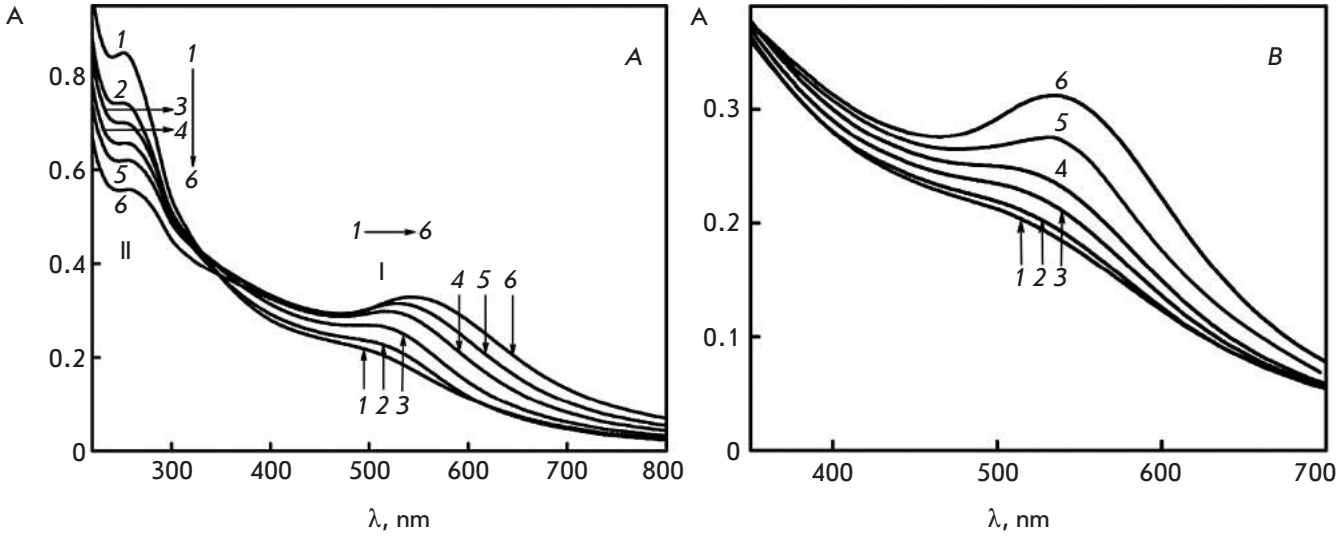


Fig. 4. The absorption spectra of DNA CLCD (A) and poly(I)×poly(C) CLCD (B) treated with Au nanoparticles during various time intervals. A. Treatment for: 1 – 0 min; 2 – 1 min; 3 – 8 min; 4 – 21 min; 5 – 36 min; 6 – 100 min. B. Treatment for: 1 – 2 min; 2 – 30 min; 3 – 50 min; 4 – 85 min; 5 – 115 min; 6 – 256 min. $C_{DNA} = 9$ μg/ml; $C_{Poly(I)\times poly(C)} = 9$ μg/ml; refer to *Fig. 1A* for the other conditions

ticles and remains almost unchanged when solvent properties are varied. The intensity of the SPR band gradually increased over time; its maximum shifted from λ ~ 505 to ~ 550 nm. Meanwhile, the amplitude of band (**II**) in the UV region of the spectrum corresponding to the absorption of DNA chromophores decreases over time. It should be also mentioned that according to theoretical calculations [45], similar changes in bands (**I**) and (**II**) in the absorption spectrum are responsible for the increase in the volume fraction of Au nanoparticles in the ensemble formed by these particles.

It is characteristic that the treatment of DNA CLCD particles with Au nanoparticles 5 and 15 nm in diameter does not result in any changes in the absorption spectra of these nanoparticles. This fact gives ground to hypothesize that there are noticeable differences in the mechanisms of action of small- and large-size Au nanoparticles on DNA CLCD particles. Indeed, it can be seen from the scheme shown in *Fig. 2* that Au nanoparticles of any size (*A*–*C*) can localize near the "surface" DNA molecules of the quasinematic layer and form linear ensembles. Formation of these ensembles even from a small number of Au nanoparticles can be accompanied by the enhancement of the SPR band [1].

It is important to note that the emergence of the plasmon effect does not require direct contact between neighboring Au nanoparticles, and the plasmon effect can be observed as long as the distance between the neighboring nanoparticles is shorter than the wavelength of the incident light [1].

The absence of changes in the absorption spectrum of CLCD particles after they are treated with Au nanoparticles 5 and 15 nm in diameter, in combination with the scheme given in *Fig. 2*, allows one to assume that in addition to the known fact that Au nanoparticles are ordered near single-stranded or linear double-stranded DNA molecules [29–31, 39, 40], there is a different mechanism of arrangement of small-sized Au nanoparticles in DNA CLCD particles.

The evolution of the SPR band during the treatment of DNA CLCD with Au nanoparticles lasts for ~100 min (*Fig. 5*); then, its saturation occurs. The direct proportional dependence between the amplitude of the SPR band (until the saturation point) and the $t^{0.5}$ value is retained. Under the assumption that the amplitude of the SPR band is associated with the concentration of Au nanoparticles in the resulting ensemble, the dependence shown in the inset (*Fig. 5*) represents the diffusion of Au nanoparticles [46] into the quasinematic layers of CLCD particles.

Figure 6 (inset) shows the dependence between the position of the SSR band maximum on the size of spherical Au nanoparticles, which was constructed by averaging the published data [40–43]. It was demon-

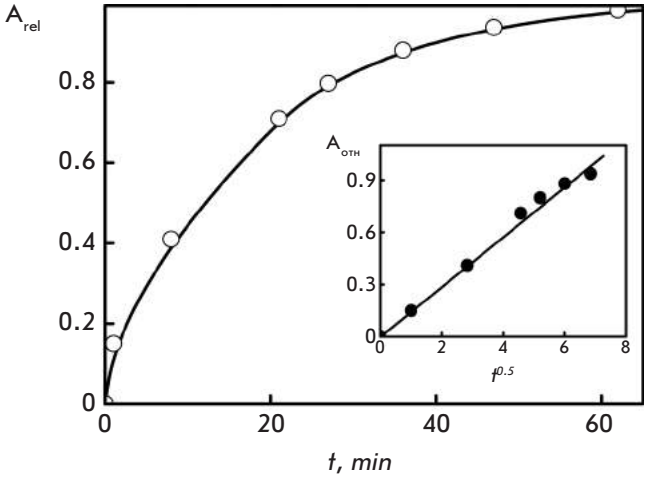


Fig. 5. Time dependence of the amplitude of the SPR band (λ = 550 nm) for Au nanoparticles interacting with DNA CLCD. C_{DNA} = 9 μg/ml; refer to *Fig. 1A* for the other conditions. Inset: dependence of the amplitude of the SPR band (λ = 550 nm) for Au nanoparticles interacting with DNA CLCD on $t^{0.5}$ value (**t** is given in min)

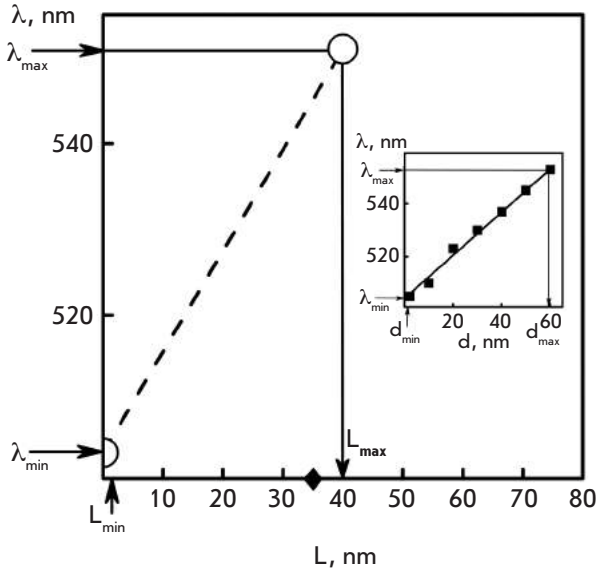


Fig. 6. Position of the surface plasmon resonance (SPR) peak as a function of the size of the linear clusters of Au nanoparticles, which are formed in the spatial structure of DNA CLCD particles. Symbol (♦) shows the data for the linear cluster of Au nanoparticles formed within the spatial structure of poly(I)×poly(C) CLCD particles. Inset: dependence of the position of the SPR peak on the diameter of spherical Au nanoparticles (the average data are taken from [40–43])

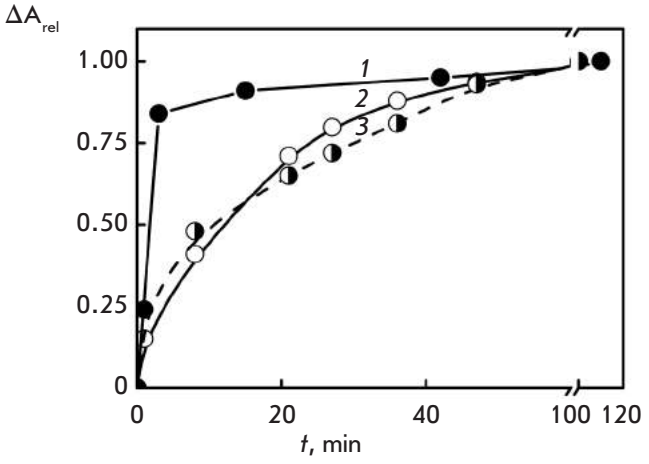


Fig. 7. Dependence of the CD band amplitude (λ = 270 nm, curve *1*), location of the SPR peak (λ = 550 nm, curve *2*), and band at λ = 270 nm (curve *3*) in the absorption spectrum of DNA CLCD on the time of treatment with Au nanoparticles. C_{DNA} = 9 μg/ml; refer to *Fig. 1A* for other conditions

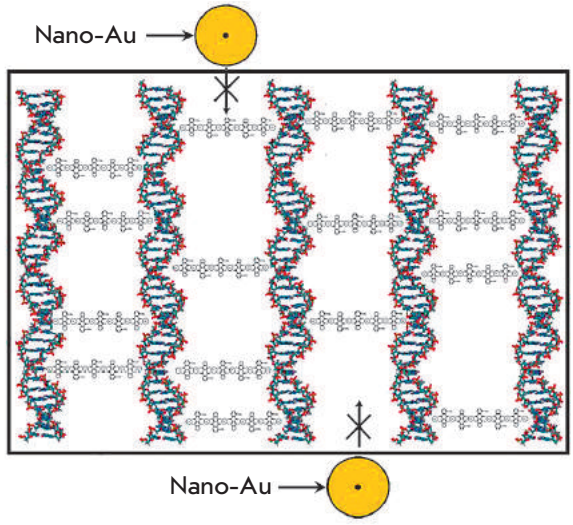


Fig. 8. A hypothetical structure of the quasinematic layer in a DNA nanoconstruction. The neighboring DNA molecules forming the quasinematic layer are "cross-linked" via nanobridges, which do not allow Au nanoparticles to penetrate into the layer and to form clusters in the «free space» between DNA molecules (the probability of their interaction with the "surface" DNA molecules remains unchanged). The frame means the presence of osmotic pressure in the PEG-containing solution

strated by comparing the results shown in *Fig. 4* with this dependence that the size of Au nanoparticles after their binding to DNA CLCD particles has the potential to increase from 2 to ~60 nm. Although this estimation is not consistent enough, since the dependence characterizes the properties of Au nanoparticles of spherical shape, it still can be used for comparative assessment of the size of Au nanoparticles formed under various conditions.

The results presented in [6] and characterizing low-angle X-ray scattering from the phases formed by DNA CLCD particles treated with Au nanoparticles allow one to make a more accurate estimation of the particle size. These results indicate that linear clusters of Au nanoparticles with a maximum size of 40 nm are formed within the structure. The SPR band is characterized by a maximum at λ ~ 550 nm [6]. The dependence of the position of the SPR peak on the linear size of Au clusters (*Fig. 6*) can be constructed using these findings (i.e., it directly describes Au nanoparticle clusters formed upon interaction between Au nanoparticles and particles of CLCD of various nucleic acids). It is clear that the actual size of the resulting ensemble (the linear cluster of Au nanoparticles) for DNA increases from 2 to 40 nm. Treatment of poly(I)×poly(C) CLCD with Au nanoparticles results in an increase in the size of Au nanoparticles up to 34 nm (these data are indicated by ♦ symbol on the X axis in *Fig. 6*).

It should also be mentioned that the size of the linear clusters of Au nanoparticles was never higher than 40 nm under the experimental conditions used (negatively charged Au particles, high ionic strength of solutions [47, 48], etc.).

The results presented in *Figs. 5* and *6* enable one to analyze more thoroughly the diffusion mechanism of formation of Au nanoparticle clusters. Since the concentration of Au nanoparticles "outside" DNA CLCD particles is higher than that "inside" (i.e., between the quasinematic layers), the concentration gradient induces the emergence of a diffusion flow of Au nanoparticles. The flow stops when the concentrations "outside" and "inside" DNA CLCD particles become equal. If the characteristic time of attaining this equilibrium is t, the size of a cluster formed by the diffused Au nanoparticles increases as the square root of time (i.e., as $t^{0.5}$). One can expect this process to be hindered by the lower translational entropy value of the Au nanoparticles concentrated inside a cluster (i.e., in the «free space» between the quasinematic layers) as compared to that of the Au nanoparticles which are freely distributed over the solution. Since the entropy factor is proportional to k_BT, the size of the Au nanoparticle clusters formed in nucleic acid CLCD particles will decrease with increasing temperature of the solution.

Thus, in our case the shift in the position of the SPR band is associated with the size of the linear Au nanoparticle clusters within nucleic acid CLCD particles

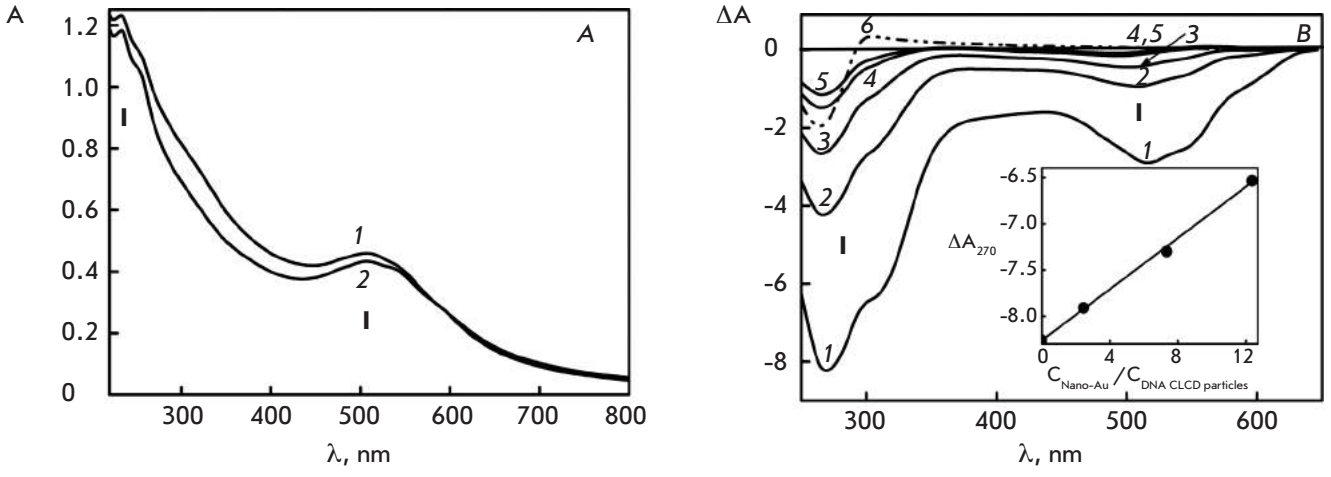


Fig. 9. The absorption (A) and CD (B) spectra of a DNA nanoconstruction treated with Au nanoparticles. A – Treatment for: 1 – 0 min; 2 – 100 min. B – Treatment for: 1 – 0 min; 2 – 10 min; 3 – 25 min; 4 – 55 min; 5 – 100 min. C_{DNA} = 5 μg/ml; C_{PEG} = 150 mg/ml; C_{DAU} = 3.2×10^{-5} M; C_{Cu} = 1×10^{-5} M; refer to *Fig. 1A* for the other conditions. Inset: ΔA_{270} value as a function of the $C_{Nano\text{-}Au}/C_{DNA\ CLCD\ particles}$ ratio in the solution

formed under various conditions rather than with an increase in the true size of individual Au nanoparticles. The problem of estimating Au nanoparticles in a cluster based on the results of optical changes remains unsolved, since the position of the SPR peak depends on the number and distance between the Au nanoparticles in a cluster, the dielectric permittivity of the medium, and other parameters [19].

With allowance for these results and the hypothetic scheme (*Fig. 2*) showing all possible ways for Au nanoparticles to bind to the DNA molecules fixed in the structure of CLCD particles, as well as for the changes in the amplitudes of the bands localizing in different regions of the absorption spectrum (*Fig. 4*), which was not observed in the control experiments with single-stranded polynucleotide or double-stranded DNA molecules under conditions impeding their condensation (*Fig. 1*), one can consider that small-sized Au nanoparticles (2 nm) can form linear clusters in CLCD particles.

Although capable of interacting with the "surface" DNA molecules (*Figs. 2A,B*) or terminal groups of DNA molecules (*Fig. 2C*) in quasinematic layers, Au nanoparticles 5 and 15 nm in diameter are too large to be incorporated between the DNA molecules in these layers.

Figure 7 shows the curves that characterize the rate of changes in the amplitude of the abnormal band in the CD spectrum of DNA CLCD, of the SPR band, and of the band located in the UV region of the absorption spectra after the dispersion is treated with small-sized Au nanoparticles. It is clear that the treatment of DNA CLCD with Au nanoparticles is accompanied by two simultaneous processes: a fast decrease in the abnormal optical activity of DNA CLCD and a slower evolution of the SPR band. The process recorded on the basis of the changes in the abnormal band in the CD spectrum lasts 10–15 min, whereas the evolution of the SPR band requires approximately 60 min.

Thus, in addition to the fast interaction between Au nanoparticles (of any size) and DNA CLCD particles (which is required to change their abnormal optical activity to a certain extent), incorporation of small-sized Au nanoparticles in the structure of CLCD particles yielding Au nanoparticle clusters is also possible.

Absorption and CD spectra obtained for CLCD particles with DNA molecules cross-linked by nanobridges treated with Au nanoparticles

An important issue is where the Au nanoparticle clusters localize. It can be assumed that small-sized Au nanoparticles diffuse into the "free space" between neighboring DNA molecules in the quasinematic layers of CLCD particles to cluster there. This process is accompanied by the emergence and evolution of the SPR band (*Fig. 4*).

In order to verify this assumption, the "free space" between the neighboring DNA molecules in CLCD particles was filled with appreciably strong nanobridges [49] consisting of alternating antibiotic molecules and copper ions (*Fig. 8*). This process resulted in the formation of a DNA nanoconstruction. In this case, the "free space" becomes inaccessible for diffusion and clustering of Au nanoparticles.

If the assumption about the localization of Au nanoparticle clusters is valid, treatment of the DNA nano-

construction with Au nanoparticles will not result in any changes in the bands located both in the UV and visible regions of the absorption spectrum. Indeed, it is clearly seen in *Fig. 9A* that no significant changes in the absorption spectrum of the nanoconstruction obtained from CLCD particles due to the formation of nanobridges between DNA molecules are observed and that SPR band (**I**) does not evolve in this case. Meanwhile, band (**II**) in the UV region of the spectrum remains virtually intact. This means that small-sized Au nanoparticles cannot insert themselves between the neighboring DNA molecules in quasinematic layers, since the "free space" is occupied by nanobridges [49].

One can focus on the fact that the nanobridges increase the rigidity of the spatial structure of the nanoconstructions [49]. Hence, although "surface" DNA molecules in particles of nanoconstructions are available for interacting with Au nanoparticles, the untwisting process (in the case when a nanoconstruction is treated with Au nanoparticles) accompanied by a decrease in the abnormal band in the CD spectrum of the nanoconstructions will require a longer period of time and can be terminated even at a smaller "depth" of this process. The CD spectra of the original DNA CLCD (dashed curve *6*), DNA nanoconstruction (i.e., CLCD with the neighboring DNA molecules cross-linked via nanobridges; curve *1*), and the same nanoconstruct treated with Au nanoparticles (curves *2–5*) are compared in *Fig. 9B*. It is clear that the formation of a DNA nanoconstruction from the original CLCD is accompanied by amplification of the band in the UV region and the emergence of an additional band in the visible region of the spectrum, which is caused by the formation of nanobridges containing chromophores absorbing within this wavelength range [49]. The amplification indicates that the twist angle of the neighboring quasinematic layers increases due to the formation of nanobridges [7]. After the nanoconstruct is treated with Au nanoparticles at a high concentration ($C_{Nano\text{-}Au} = 0.82 \times 10^{14}$ particles/ml), the amplitude of the bands in the UV and visible regions of the spectrum decreases despite the fact that the absorption spectrum does not contain the SPR band.

Figure 10 shows a comparison of the kinetic curves characterizing the changes in the abnormal optical activity caused by treatment of the original DNA CLCD and DNA nanoconstructions with Au nanoparticles. It is clear that the depth and rates of these processes are different for the original DNA CLCD and DNA nanoconstructions, which supports the thesis that the bridges play a stabilizing role.

The results shown in *Fig. 9* additionally demonstrate that small-sized Au nanoparticles can interact with the "surface" molecules of double-stranded DNA, thus inducing the cholesteric → nematic transition, even if

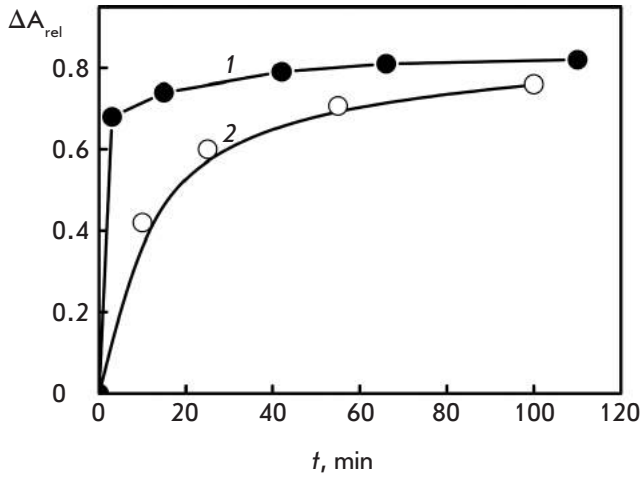


Fig. 10. Kinetic curves characterizing the change in the abnormal optical activity of the initial DNA CLCD (curve *1*) and DNA nanoconstruction (curve *2*) induced by treatment with Au nanoparticles (2 nm).
Curve 1- $C_{DNA} = 9$ μg/ml; $C_{DNA} = 5$ μg/ml; $C_{PEG} = 150$ mg/ml; $C_{DAU} = 3.2 \times 10^{-5}$ M; $C_{Cu} = 1 \times 10^{-5}$ M; refer to *Fig. 1A* for the other conditions

nanobridges form between the neighboring DNA molecules, but cannot diffuse between DNA molecules in the quasinematic layers, since the "free space" is filled with nanobridges.

Thus, the SPR band can emerge and evolve only if there is "free space" between DNA molecules in quasinematic layers. It is in this very space that Au nanoparticle clusters are formed.

We previously demonstrated that the interaction between Au nanoparticles and the "surface" DNA molecules in CLCD particles induces changes in the helical spatial distribution of neighboring quasinematic DNA layers (i.e., formation of the nematic structure). It is possible that the probability of one (or several) right-handed helical double-stranded DNA molecule rotating 180° with respect to its neighbor(s) due to rotational diffusion in the quasinematic layers located at nanodistances increases at this very moment. In this case, the reactive groups of a DNA molecule (1) localize in the "free space" facing the identical groups of its neighbor (2), which can be referred to as a type of face-to-face phasing of the reactive groups of DNA molecules. Therefore, clustering of negatively charged Au nanoparticles in the "free space" between DNA molecules (*Fig. 2*) may result from two processes. First, Au nanoparticles may diffuse into the "free space" between the neighboring "phased" DNA molecules (1 and 2) (in this case, it is a one-dimensional diffusion of Au nanoparticles between these DNA molecules). Second, the interaction between a DNA

particle in the quasinematic layer and a negatively charged small-sized Au nanoparticle can be conditionally regarded as the equivalent interaction between a plane and a spherical particle [50]. In this case, the interaction of the Au nanoparticle can be determined by the so-called Casimir effect [51–54].

For either version of the processes discussed above (provided that the experimental conditions are fixed), one can assume that Au nanoparticles can form linear clusters between DNA molecules (direct contact between neighboring Au nanoparticles in clusters can be absent) [55]. The clustering of Au nanoparticles is accompanied by the evolution of the SPR band.

Thus, different processes can determine "sliding" ("retraction") of Au nanoparticles into the "free space" between neighboring DNA molecules in quasinematic layers.

Thus, if one accepts the hypothesis of the ordering mechanism of negatively charged Au nanoparticles in quasinematic layers, it becomes clear why small-sized Au nanoparticles form clusters only in CLCD particles comprising double-stranded molecules of nucleic acids or synthetic polyribonucleotides (poly(I)×poly(C))., neighboring double-stranded DNA molecules fixed in the spatial structure of CLCD particles. This conclusion allows one to regard a DNA CLCD particle as a matrix that specifically adsorbs small-sized Au nanoparticles and provides conditions for the formation of linear clusters from these nanoparticles. The cytotoxicity of Au nanoparticles can presumably be attributed to their tendency to cluster. ●

CONCLUSIONS

These findings demonstrate that small-sized Au nanoparticles form clusters in the "free space" between the

The authors are sincerely grateful to V.M. Rudoi and O.V. Dement'eva (Frumkin Institute of Physical Chemistry and Electrochemistry, Russian Academy of Sciences) for synthesizing high-quality Au nanoparticle samples, for determining their parameters, and for providing critical comments in preparing the manuscript.

This work was supported by the Federal Target-Oriented Program "Research and Elaboration of Priority Directions of Science and Engineering Development of the Scientific-Engineering Complex of Russia for 2007–2013" (Government Contract № 14.527.12.0012 dated October 13, 2011; application code «2011-2.7-527-012-001») and by the Russian Foundation for Basic Research (project № 11-04-00118-a).

REFERENCES

1. Dykman, L.A., Bogatyrev V.A., Shchegolev S.Y., Khlebtsov N.G. Gold nanoparticles: synthesis, properties and biomedical applications. M.: Nauka, 2008. 318 p.
2. Wiwanitkit V., Sereemaspun A., Rojanathanes R. // Fertil. Steril. 2009. V. 91. № 1. P. e7–e8.
3. Zakhidov S.T., Marshak T.L., Malonina E.A., Kulibin A.Yu., Zelenina I.A., Pavlyuchenkova O.V., Rudoy V.M., Dement'eva O.V., Skuridin S.G., Yevdokimov Yu.M. // Biological membranes (Russian Edition). 2010. V. 27. № 4. P. 349–353.
4. Kang B., Mackey M.A., El-Sayed M.A. // J. Am. Chem. Soc. 2010. V. 132, № 5. P. 1517–1519.
5. Tsoli M., Kuhn H., Brandau W., Esche H., Schmid G. // *Small*. 2005. V. 1. № 8–9. P. 841–844.
6. Yevdokimov Yu.M., Skuridin S.G., Salyanov V.I., Popenko V.I., Rudoy V.M., Dement'eva O.V., Shtykova E.V. // J. Biomater. Nanobiotechnol. 2011. V. 2, № 4. P. 461–471.
7. Yevdokimov Yu.M., Salyanov V.I., Semenov S.V., Skuridin S.G. DNA liquid–crystalline dispersions and nanoconstructions (Russian Edition). Moscow: Radiotekhnika, 2008, 296 p.
8. Livolant F., Leforestier A. // Prog. Polym. Sci. 1996. V. 21, № 6. P. 1115–1164.
9. Leforestier A., Livolant F. // Proc. Natl. Acad. Sci. USA. 2009. V. 106, № 23. P. 9157–9162.
10. Leforestier A., Livolant F. // J. Mol. Biol. 2010. V. 396. № 2. P. 384–395.
11. Schmid G., Klein N., Korste L. // Polyhedron. 1988. V. 7. № 8. P. 605–608.
12. Li H., Rothberg L. // Proc. Natl. Acad. Sci. USA. 2004. V. 101. № 39. P. 14036–14039.
13. Storhoff J.J., Mucic R.C., Mirkin C.A. // J. Cluster Sci. 1997. V. 8. № 2. P. 179–216.
14. Keating C.D., Kovaleski K.M., Natan M.J. // J. Phys. Chem. B. 1998. V. 102. № 47. P. 9404–9413
15. Harada G., Sakurai H., Matsushita M.M., Izuoka A., Sugawara T. // Chem. Lett. 2002. V. 31. No 10. P. 1030–1031.
16. Tkachenko A.G., Xie H., Coleman D., Glomm W., Ryan J., Anderson M.F., Franzen S., Feldheim D.L. // J. Am. Chem. Soc. 2003. V. 125, № 16. P. 4700–4701.
17. Katz E., Willner I. // Angew. Chem. Int. Ed. Engl. 2004. V. 43. № 45. P. 6042–6108.
18. Khlebtsov N.G., Dykman L.A., Krasnov Ya.M., Melnikov A.G. // Colloid. J. 2000. V. 62. № 6. P. 765–779.
19. Westcott S. L., Oldenburg S. J., Lee T. R., Halas N. J. J. // Chem. Phys. Lett. 1999. V. 300. № 5–6. P. 651–655.
20. Kelly K.L., do Coronado E., Zhao L.L., Schatz G.C. // J. Phys. Chem. B. 2003. V. 107. № 3. P. 668–677.
21. Schmitt J., Decher G., Dressick W.J., Brandow S.L., Geer R.E., Shashidhar R., Calvert J.M. // Adv. Mater. 1991. V. 9. № 1. P. 61–65.
22. Grabar K.C., Freeman R.G., Hommer M.B., Natan M.J. // Anal. Chem. 1995. V. 67. № 4. P. 735–743.
23. Turkevich J., Stevenson P.C., Hillier J. // Discuss. Faraday Soc. 1951. V. 11. No 0. P. 55–75.
24. Brown K.R., Walter D.G., Natan M.J. // Chem. Mater. 2000. V. 12. № 2. P. 306–313.
25. Duff D.G., Baiker A., Edwards P.P. // Langmuir. 1993. V. 9.

№ 9. P. 2301–2309.
26. Noginov M. A., Zhu G., Belgrave A. M., Bakker R., Shalaev V. M., Narimanov E. E., Stout S., Herz E., Suteewong T., Wiesner U. // Nature. 2009. V. 460. № 7259. P. 1110–1113.
27. Grasso D., Fasone S., La Rosa C., Salyanov V. // Liq. Crystals. 1991. V. 9. № 2. P. 299–305.
28. Grasso D., Campisi R.G., La Rosa C. // Thermochim. Acta. 1992. V. 199, № 1. P. 239–245.
29. Loweth C.J., Caldwell W.B., Peng X., Alivisatos A.P., Schultz P.G. // Angew. Chem. Int. Ed. 1999. V. 38. № 12. P. 1808–1812.
30. Kumar A., Pattarkine M., Bhadbhade M., Mandale A.B., Ganesh K.N., Datar S.S., Dharmadhikari C.V., Sastry M. // Adv. Mater. 2001. V. 13. № 5. P. 341–344.
31. Liu Y., Meyer-Zaika W., Franzka S., Schmid G., Tsoli M., Kuhn H. // Angew. Chem., Int .Ed. 2003. V. 42. № 25. P. 2853–2857.
32. Skuridin S.G., Dubinskaya V.A., Rudoy V.M., Dement·eva O.V., Zakhidov S.T., Marshak T.L., Kuzmin V.A., Popenko V.I., Yevdokimov Yu.M. // Dokl. Acad. Nauk.(Russian Edition) 2010. V. 432. № 6. P. 838–841.
33. Zherenkova L.V., Komarov P.V., Khalatur P.G. // Colloid. J. (Russian Edition). 2007. V. 69. № 5. P. 753–765.
34. Herne T.M., Tarlov M.J. // J. Am. Chem. Soc. 1997. V. 119. № 38. P. 8916–8920.
35. Petrovykh D.Y., Kimura–Suda H., Whitman L.J., Tarlov M.J. // J. Am. Chem. Soc. 2003. V. 125. № 17. P. 5219–5226.
36. Parak W.J., Pellegrino T., Micheel C.M., Gerion D., Williams S.C.,
Alivisatos A. Paul. // Nano Lett. 2003. V. 3. № 1. P. 33–36.
37. Kira A., Kim H., Yasuda K. // Langmuir. 2009. V. 25. № 3. P. 1285–1288.
38. Mirkin C.A., Letsinger R.L., Mucic R.C., Storhoff J.J. // Nature. 1996. V. 382. № 6592. P. 607–609.
39. Sastry M., Kumar A., Datar S., Dharmadhikari C.V., Gahesh K.N. // Appl. Phys. Lett. 2001. V.78. № 19. P. 2943–2945
40. Warner M.G., Hutchison J.E. // Nature Materials. 2003. V. 2. P. 272–277. doi:10.1038/nmat853.
41. Link S., El–Sayed M.A. // J. Phys. Chem. B. 1999. V. 103. № 40. P. 8410–8426.
42. Su K.H., Wei Q.H., Zhang X., Mock J.J., Smith D.R., Schultz S. // Nano Lett. 2003. V. 3. № 8. P. 1087–1090.
43. Rechberger W., Hohenau A., Leitner A., Krenn J.R., Lamprecht B., Aussenegg F.R. // Opt. Commun. 2003. V. 220. № 1–3. P. 137–141.
44. Kamat P.V. // J. Phys. Chem. B. 2002. V. 106. № 32. P. 7729–7744.
45. Storhoff J.J., Lazarides A.A., Mucic R.C., Mirkin C.A., Letsinger R.L., Schatz G.C. // *J. Am. Chem. Soc*. 2000. V. 122. № 19. P. 4640–4550.
46. Grabar K.C., Smith P.C., Musick M.D., Davis J.A., Walter D.G., Jackson M.A., Guthrie A.P., Natan M.J. // J. Am. Chem. Soc. 1996. V. 118. № 5. P. 1148–1153.
47. Biggs S., Mulvaney P., Zukovski C.F., Grieser F. // J. Am. Chem. Soc. 1994. V. 116. № 20. P. 9150–9157.
48. Thompson D.W., Collins I.R. // J. Colloid. Interface Sci. 1992. V. 152. № 1. P. 197–204.
49. Yevdokimov Yu.M., Salyanov V.I., Skuridin S.G. Nanostructures and nanoconstructions based on DNA (Russion Edition). Moscow: Science–Press, 2010. 254 p.
50. Maia Neto P.A., Lambrecht A., Reynaud S. // Phys. Rev. 2008. V. A 78. № 1. 012115. [4 pages].
51. Casimir H.B.G. // Proc. K. Ned. Akad. Wet. 1948. V. 51. P. 793–795.
52. Mostepanenko. V.M., Trunov N.N. // Uspekhi Fiz. Nauk. 1988. V. 156. № 5. P. 385–426.
53. Rodriguez A.W., Capasso F., Johnson S.G. // Nature Photonics. 2011. V. 5. № 4. P. 211–221.
54. Dzyaloshinskii I.E., Lifshitz E.M., Pitaevskii L.P. // Adv. Phys. 1961. V. 10. № 38. P. 165–209.
55. Chan H.B., Bao Y., Zou J., Cirelli R.A., Klemens F., Mansfield W.M., Pai C.S. // Phys. Rev. Lett. 2008. V. 101. № 3. 030401. [4 pages].

Methanogenic Community Dynamics during Anaerobic Utilization of Agricultural Wastes

A. M. Ziganshin*[,1], E. E. Ziganshina[1], S. Kleinsteuber[2], J. Pröter[3], O. N. Ilinskaya[1]
[1]Kazan (Volga Region) Federal University, Kremlyovskaya Str., 18, Kazan, Russia, 420008
[2]UFZ-Helmholtz Centre for Environmental Research, Permoser Str., 15, Leipzig, Germany 04318
[3]German Biomass Research Centre, Torgauer Str., 116, Leipzig, Germany, 04347
*E-mail: a.ziganshin06@fulbrightmail.org
Received 08.08.2012

ABSTRACT **This work is devoted to the investigation of the methanogenic archaea involved in anaerobic digestion of cattle manure and maize straw on the basis of terminal restriction fragment length polymorphism (T-RFLP) analysis of archaeal 16S rRNA genes. The biological diversity and dynamics of methanogenic communities leading to anaerobic degradation of agricultural organic wastes with biogas production were evaluated in laboratory-scale digesters. T-RFLP analysis, along with the establishment of archaeal 16S rRNA gene clone libraries, showed that the methanogenic consortium consisted mainly of members of the genera *Methanosarcina* and *Methanoculleus*, with a predominance of *Methanosarcina* spp. throughout the experiment.**
KEY WORDS **archaeal 16S rRNA genes; T-RFLP analysis; biogas production; methanogens.**
ABBREVIATIONS **T-RFLP – terminal restriction fragment length polymorphism; OTU – operational taxonomic unit; oTS – organic total solids; OLR – organic loading rate; HRT – hydraulic retention time; VFA – volatile fatty acids.**

INTRODUCTION

One of the most effective methods for reducing the negative effects of the waste from the agricultural and processing industries on the environment is their anaerobic digestion. Anaerobic digestion of wastes is accompanied by the destruction of most organic components and production of biogas consisting of methane (50–75%) and carbon dioxide (25–50%), with trace amounts of other components. In contrast to bioethanol and biodiesel mostly produced from energy crops, biogas is obtained during utilization of residual biomass and various organic wastes [1–7], such as cattle manure. However, due to the low biodegradability of manure, its utilization in anaerobic reactors is characterized by an insignificant biogas yield. Anaerobic co-digestion of manure and plant biomass promotes substrate hydrolysis, optimizes the distribution of nutrients in the bioreactor, thus activating microbial growth and the biomethane yield [8, 9]. The co-digestion of several different substrates has been actively investigated over the past years [9–13].

The first three stages of anaerobic co-digestion (hydrolysis, acidogenesis, and acetogenesis) are performed by bacterial communities; the fourth stage is performed by aceticlastic and hydrogenotrophic methanogens, which consume acetate, molecular hydrogen, and carbon dioxide to produce methane [1, 6, 14].

Independently of the mode of digestion (psychro-, meso-, or thermophilic) and feedstock composition, the major participants in methanogenesis are the members of the orders of *Methanomicrobiales* and/or *Methanosarcinales* [2, 5, 7, 15–18]. However, there is a lack of information about the changes in microbial association during methanogenic fermentation.

The present study was devoted to the investigation of pathways for utilization of agricultural wastes (manure and maize straw) with biogas production in laboratory-scale biogas reactors and to studying the diversity, structure, and dynamics of the methanogenic communities involved in this process using modern methods of molecular biology. The determination of the composition and dynamics of the microbial communities in biogas reactors, jointly with the analysis of substrate destruction, is aimed at revealing the potential for intensification of the anaerobic process. The use of the universal phylogenic marker 16S rRNA and T-RFLP (terminal restriction fragment length polymorphism) will contribute to the study of the composition and temporal changes to the microbial consortium.

MATERIALS AND METHODS

Digester configurations

Table 1 lists the main technological parameters of the anaerobic processing of cattle manure and maize straw. All bioreactors were run under mesophilic conditions (38°C). The bioreactors R 4.13 and R 4.14 were loaded with cattle manure and maize straw; the bioreactors R 4.15 and R 4.16, with cattle manure and extruded maize straw. Feeding a new portion of substrate and unloading of the digested mixture were performed daily; the volume of the digesting mixture was maintained at the level of 30 L; the hydraulic retention time (HRT) was kept constant in the bioreactors (35 days). The biogas yield, composition and pH were analyzed daily, whereas the concentrations of organic acids and ammonium ions were measured twice a week.

Analytical methods

Biogas production was monitored using Ritter TG 05 drum-type gas meters (Bochum, Germany); biogas composition was measured by an infrared landfill gas analyzer, GA 94 (Ansyco, Germany). Ammonium concentration was analyzed by coloring of the liquid phase of the bioreactor contents with Nessler's reagent on a spectrophotometer DR/2000 (HachCompany, USA) at 425 nm.

The total acid capacity was determined by titration with 0.025–0.1 M H_2SO_4 in a pH range from 4.5 to 3.5 using a Titration Excellence T90 titrator (Mettler-Toledo, Switzerland). The concentration of volatile fatty acids (VFA) was analyzed by gas chromatography using a 5890 series II GC (Hewlett Packard, USA) equipped with an HS40 automatic headspace sampler (Perkin Elmer, USA) and an Agilent HP-FFAP column (30 m×0.32 mm×0.25 μm), as described previously [7].

DNA extraction and purification

Samples were collected from four reactors once a month and were immediately used for DNA extraction and purification. The digested biomass mixture was centrifuged at 20,000 g for 10 min. Total DNA was subsequently extracted and purified using a FastDNA Spin Kit for soil (Qbiogene, Germany) according to the manufacturer's recommendations. The total amount of extracted and purified DNA was measured on a NanoDrop ND-1000 UV–visible spectrophotometer (PeqLab, Germany).

Amplification, cloning and sequencing of archaeal 16S rRNA

All molecular manipulations were performed according to our previous work [7]. Archaeal 16S rRNA genes were amplified from the total DNA as a template in a DNA Engine Tetrad 2 Peltier Thermal Cycler (Bio-Rad) using a combination of universal primers UniArc21F (5'-TTCYGKTTGATCCYGSCRG-3') and UniArc931R (5'-CCCGCCAATTCCTTTHAG-3') and 2 × *Taq*MasterMix (Qiagen, Germany). The composition of the reaction mixture was as follows: 6 μL of 2 × *Taq*MasterMix, 0.5 μL of UniArc21F (5 pmol/μL), 0.5 μL of UniArc931R (5 pmol/μL), 4 μL of H_2O, and 1 μL of 100-fold diluted DNA template (equivalent of 1–3 ng). The amplification was started with denaturation at 95°C for 5 min, followed by 35 cycles: denaturation at 94° C for 1 min, annealing at 54° C for 1 min, and elongation at 72° C for 2 min. The final elongation was carried out at 72° C for 2 min.

The PCR products were purified using a QIAGEN PCR Cloning Kit (QIAGEN, Germany). The presence of inserts of archaeal 16S rRNA genes of the desired size in positive clones after cloning was analyzed using the vector-specific primers M13uni(–21) (5'-TGTAAAACGACGGCCAGT-3') and M13rev(–29) (5'-CAGGAAACAGCTATGACC-3'). 1 μL of the M13-amplicons were further treated with HaeIII endonuclease (New England Biolabs, Germany) and separated by Phor-agarose gel electrophoresis (Biozym, Germany). The lengths of restricted fragments were analyzed using the Phoretix™ 1D Database Version 2.00 and Phoretix™ 1D Advanced Version 5.20 (Nonlinear Dynamics, Great Britain) software; clones were grouped into clusters, and dendrograms were constructed. The representative clones from large clusters were selected to further determine their nucleotide sequences.

The PCR products of the representative clones were purified using a Promega PCR Purification Kit (Promega, USA). The nucleotide sequences of the 16S rRNA genes were determined using a BigDye™ Terminator Cycle Sequencing Ready Reaction Kit 1.1 on an ABI-PRISM 3100 Genetic Analyzer automated sequencer (Applied Biosystems). The POP-6™ polymer was used as a separation matrix. The BLAST tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi) [21] was employed to search for similar sequences in the GenBank database. The Ribosomal Database Project (http://rdp.cme.msu.edu) [22] was used for taxonomic assignment.

T-RFLP analysis

The T-RFLP analysis was performed in accordance with our previous work [7]. The archaeal 16S rRNA genes were amplified using a universal primer pair UniArc21F-FAM and UniArc931R and 2 × *Taq*MasterMix (Qiagen, Germany) with the PCR parameters as described above. The forward primer UniArc21F-FAM was marked with a FAM fluorophor (phosphoramidite fluorochrome-5-carboxyfluorescein) at the 5' end. The amplicons of the archaeal 16S rRNA genes containing

Table 1. Main configurations of anaerobic digestion of cattle manure and maize straw

Digester*	Organic loading rate**, g_{oTS} day^{-1}	Substrate composition, g day^{-1}			Biogas yield under standard conditions, L g^{-1}_{oTS}	Biogas composition			pH	Acid capacity, g L^{-1}	NH_4^+-N, g L^{-1}
		cattle manure	straw	total***		CH_4, %	CO_2, %	H_2S, ppm.			
R 4.13	74.1	723.6	28.2	857	0.40	58.7	40.2	3450	7.63	1.49	1.20
	71.2	518.7	26.3	857	0.36	59.8	38.7	2216	7.50	1.90	1.24
	71.7	694.6	26.3	857	0.33	55.6	42.9	2145	7.61	1.80	1.16
R 4.14	74.1	723.6	28.2	857	0.40	59.3	39.8	4183	7.66	1.42	1.22
	71.2	518.7	26.3	857	0.38	58.4	40.2	1928	7.53	1.66	1.28
	71.7	694.6	26.3	857	0.37	56.7	42.1	2092	7.58	1.43	1.31
R 4.15	72.1	723.6	83.7	857	0.39	58.1	41.1	~5000	7.75	1.54	1.47
	68.6	518.7	78.1	857	0.39	59.3	39.2	2234	7.56	1.28	1.39
	69.1	694.6	78.1	857	0.39	56.8	42.6	2373	7.74	1.37	1.26
R 4.16	72.1	723.6	83.7	857	0.41	58.6	40.6	4558	7.76	1.51	1.54
	68.6	518.7	78.1	857	0.38	59.0	40.1	2056	7.54	1.53	1.36
	69.1	694.6	78.1	857	0.39	57.2	41.5	3155	7.61	1.37	1.27

* Digester parameters are presented at three sampling times, when methanogenic communities were analyzed (except for biogas yield, biogas composition, and pH, with values averaged over 1 week).
** Organic total solids.
*** Water was added to final concentration of 857 ml day^{-1}.

FAM fluorophor were purified using a SureCleanPlus kit (Bioline, Germany) and treated with the MseI and HaeIII restrictases (New England Biolabs, Germany). After a 16-hour-long incubation at 37°C, DNA fragments were precipitated with 3 M sodium acetate (pH 5.5) and absolute ethanol. The supernatant was removed; the precipitate was dried in vacuum, and the resulting DNA fragments were resuspended in 10 μL of Hi-Di formamide containing 0.25 μL of GeneScan-500 ROX™ STANDARD or MapMarker® 1000 size standard. The samples were denatured at 95°C for 5 min, cooled on ice (approximately for 5 min), and analyzed on an ABIPRISM 3100 Genetic Analyzer (Applied Biosystems). The POP-6™ polymer was used as a separation matrix. The resulting T-RLFP patterns were analyzed using the GeneMapper V3.7 software (Applied Biosystems). The theoretical T-RF values of the representative phylotypes listed in the clone library were calculated using the NEB cutter Version 2.0 (http://tools.neb.com/NEBcutter2) and confirmed experimentally by T-RFLP analysis using the corresponding clones as templates.

RESULTS AND DISCUSSION

The use of renewable energy sources, in particular various types of organic waste, is an essential aspect of "green technologies" for biofuel production [1]. The aim of this work was to investigate the dynamics of methanogenic associations during the conversion of cattle manure and maize straw in model mesophilic digesters.

Table 1 lists the major operational parameters of the anaerobic digestion of agricultural waste as substrates. Anaerobic biomass destruction was carried out in four laboratory-scale digesters with an operating volume of 30 L at 38°C. In the reactors R 4.13 and R 4.14, cattle manure and maize straw were co-digested; the reactors

Table 2. Results of sequencing of archaeal 16S rRNA gene clones and experimentally determined terminal restriction fragments (T-RF)

Clone, bp	Nearest relative (GenBank accession No) / coincidence %	Taxonomic status in accordance with RDP 10	MseI-T-RF, bp	HaeIII-T-RF, bp
ar_B9 (863)	Uncultured archaeon clone: FA69 (AB494258) / 99%	*Methanoculleus* sp.	37	67
ar_A1 (864)	Uncultured *Methanoculleus* sp. Clone: DMMR219 (HM218939) / 99%	*Methanoculleus* sp.	36	67
OTU 1		*Methanoculleus sp. I*	36/37	67
ar_A2 (863)	Uncultured archaeon clone: MTSArc_G8 (EU591664) / 99%	*Methanoculleus* sp.	499	67
OTU 2		*Methanoculleus sp. II*	499	67
ar_E12 (864)	Uncultured archaeon clone: WA50 (AB494245) / 100%	*Methanocorpusculum* sp.	97	241
OTU 3		*Methanocorpusculum sp.*	97	241
ar_E10 (567)	Uncultured euryarchaeote clone: B35_F_A_A05 (EF552199) / 99%	*Methanosarcina* sp.	557	220
ar_H2 (873)	Uncultured euryarchaeote clone: B35_F_A_A05 (EF552199) / 99%	*Methanosarcina* sp.	557	220
OTU 4		*Methanosarcina sp. I*	557	220
ar_E6 (873)	Uncultured archaeon clone: SA42 (AB494252) / 99%	*Methanosarcina* sp.	859	220
ar_F10 (873)	Uncultured archaeon clone: SA42 (AB494252) / 99%	*Methanosarcina* sp.	858	220
OTU 5		*Methanosarcina sp. II*	858/859	220
ar_G8 (874)	Uncultured archaeon clone: SA42 (AB494252) / 99%	*Methanosarcina* sp.	877	220
OTU 6		*Methanosarcina sp.III*	877	220

R 4.15 and R 4.16 were loaded with manure and extruded maize straw. The organic loading rate (OLR) was varied from 71.2 to 74.1 g_{oTS} day^{-1} (organic total solids) in the reactors R 4.13 and R 4.14. In the reactors R 4.15 and R 4.16, the OLR was lower and varied from 68.6 to 72.1 g_{oTS} day^{-1}. Throughout the experiment the HRT was kept constant (35 days). Depending on the particular feedstock, the biogas yield varied from 0.33 to 0.41 L g^{-1}_{oTS} with a methane content of 56–60%. As can be seen in *Table 1,* pH in all bioreactors was maintained at approximately 7.5–7.8; acid capacity ranged between 1.3 and 1.9 g L^{-1}, and ammonium concentration varied from 1.2 to 1.5 g L^{-1}. These parameters are favorable for methanogenesis [23].

The biological diversity and dynamics of methanogenic communities digesting cattle manure and maize straw were investigated by amplification, cloning, restriction analysis, and sequencing of the archaeal 16S rRNA genes. The methanogenic association structure was determined at three sampling points with a 1-month interval.

Amplification, cloning, sequencing of archaeal 16S rRNA, and T-RLFP analysis revealed a relatively large diversity of archaeal species in the reactors. During the T-RLFP analysis of archaeal 16S rRNA gene amplicons containing FAM flurophor were treated with endonucleases MseI and HaeIII. Belonging of the peaks in T-RLFP patterns to certain phylogenic groups was determined by the length of the terminal restriction fragments (T-RF) of 16S rRNA gene clones. In total, 9 clones were selected from the clone library for sequencing. The clones were classified into 6 operational taxonomic units (OTUs) on the basis of their T-RF lengths (*Table* 2). Three phylotypes were attributed

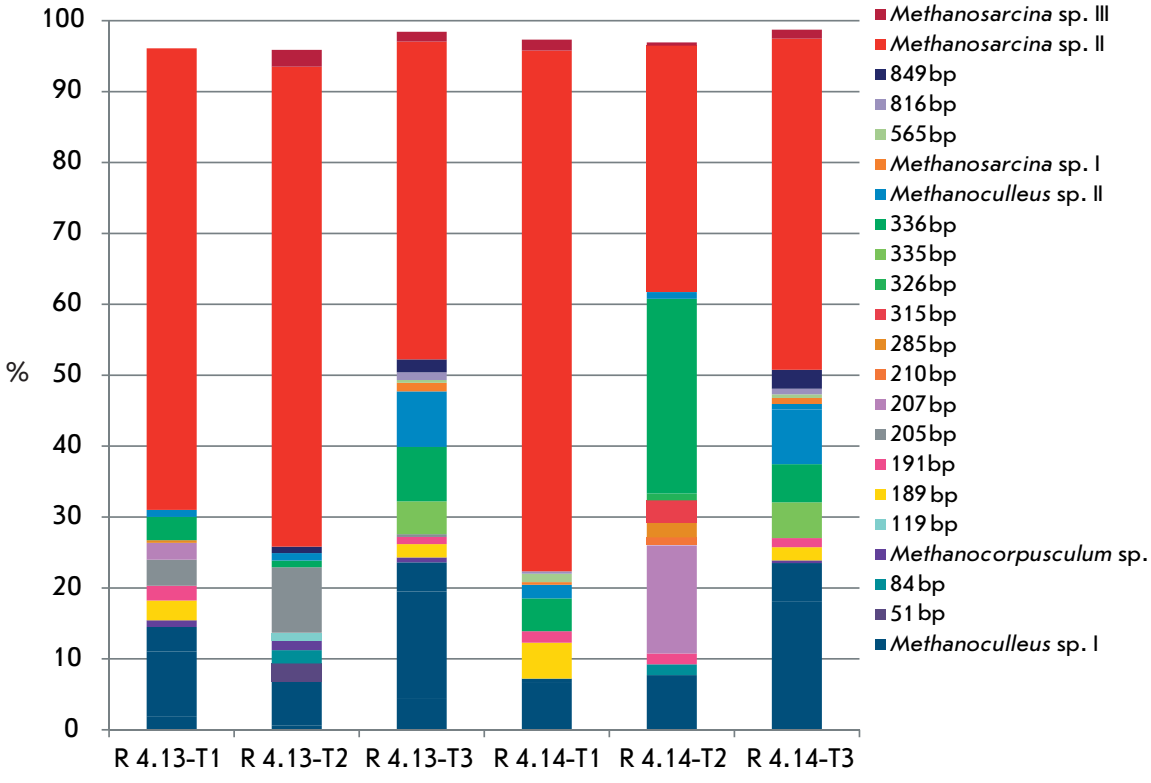


Fig. 1. Dynamics of methanogenic communities in the digesters R 4.13 and R 4.14 based on T-RFLP analysis (determined with the restriction enzyme MseI)

to the order *Methanomicrobiales* (OTU 1, OTU 2, OTU 3), and three were attributed to the order *Methanosarcinales* (OTU 4, OTU 5, OTU 6). Up to 22 different T-RFLP profiles (with abundance of more than 1%) were detected by T-RLFP analysis of 16S rRNA genes using MseI restrictase. Since the main T-RFs in the reactors were identified, we identified the methanogens playing the key role in biogas production.

Figure 1 shows the distribution of groups of methanogens (community dynamics) during anaerobic digestion of manure and straw (R 4.13 and R 4.14). This distribution was obtained based on MseI restriction profiles (results of HaeIII restriction are not shown). In the first sample, when the organic loading rate was 74.1 g_{oTS} day^{-1}, the T-RFLP analysis revealed the predominance of methanogens of the genus *Methanosarcina* and hydrogenotrophic methanogens of the genus *Methanoculleus* in the archaeal community of the bioreactors R 4.13 and R 4.14. Thus, the total ratio of representatives of *Methanosarcina* sp. (OTU 4, OTU 5, and OTU 6) and *Methanoculleus* sp. (OTU 1, OTU 2) was 65 and 15%, respectively, of the total T-RF peak areas in the reactor R 4.13. In the reactor R 4.14, methanogens of the genera *Methanosarcina* (75%) and *Methanoculleus* (9%) were detected. Other archaeal members with low abundance (1–3%) were classified into the minor groups. A decrease in OLR to 71.2 g_{oTS} day^{-1} with a subsequent increase to 71.7 g_{oTS} day^{-1} resulted in a change in the composition of the microbial community. Thus, the relative abundance of members of the genus *Methanosarcina* (OTU 4, OTU 5, OTU 6) in the two next sampling points reached 70/47% and 35/49% values for the reactors R 4.13 and R 4.14, respectively. The relative abundance of the species of the genus *Methanoculleus* (OTU 1, OTU 2) in the reactors R 4.13 and R 4.14 was 8/31 and 9/32%, respectively (two next sampling points).

Hydrogentrophic methanogens from the genus *Methanocorpusculum* were found among the minor associations and they comprised less than 2% of the total T-RF area. Furthermore, the major peak corresponding to 336 bp was detected in T-RLFP patterns; however, this phylotype was not present among the cloned archaeal 16S rRNA genes, and, hence, it was assigned to the unidentified group of the methanogenic community.

It is clear from *Fig. 2* that the composition of the methanogenic communities from the bioreactors R 4.15 and R 4.16 with manure and extruded straw as the used substrates was represented by similar groups as those detected in the reactors R 4.13 and R 4.14. The OLR at three sampling points for the R 4.15 and R 4.16 reactors

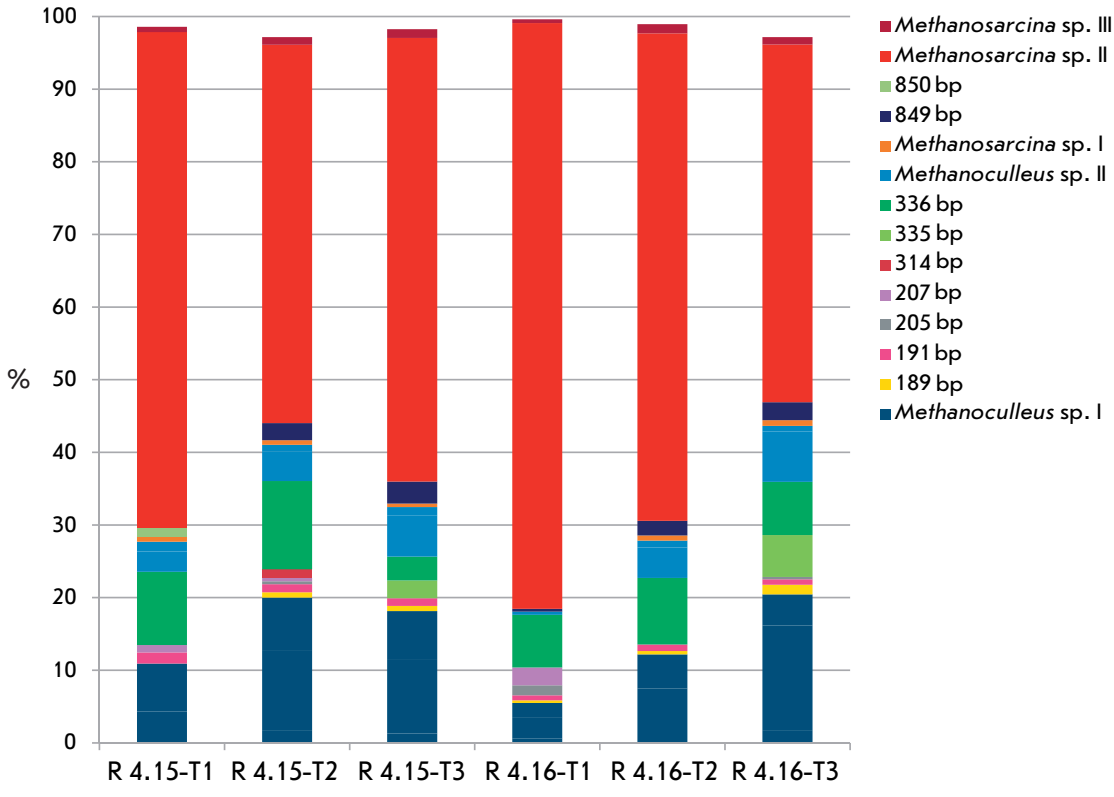


Fig. 2. Dynamics of methanogenic communities in the digesters R 4.15 and R 4.16 based on T-RFLP analysis (determined with the restriction enzyme MseI)

were 72.1, 68.6, and 69.1 g_{oTS} day^{-1}, respectively. The members of the genera *Methanosarcina* (70, 54, and 63% of the total abundance in three sampling points, respectively) and *Methanoculleus* (15, 25, and 25% of the total abundance in three sampling points, respectively) were the predominant taxons in the digester R 4.15. Reactor R 4.16, as well as R4.15, was dominated by members of the genera *Methanosarcina* (81, 69, and 51%) and *Methanoculleus* (6, 17, and 28%). Similar to that in the reactors R 4.13 and R 4.14, high abundance of the T-RF peak corresponding to 336 bp was detected; however, the taxonomic group of archaea corresponding to this restriction length profile was not determined.

These findings substantiate the possibility of effective co-digestion of manure and maize straw, yielding biogas. It has been demonstrated that members of the genera *Methanosarcina* and *Methanoculleus* prevail throughout the fermentation process. In addition, the methanogenic community dynamics during utilization of organic waste has been investigated for the first time. *Methanoculleus* species utilize hydrogen and carbon dioxide for methanogenesis [2], whereas the members of the genus *Methanosarcina* are likely to decompose acetate yielding methane and carbon dioxide or to utilize hydrogen, carbon dioxide, and methylated compounds yielding methane [24]. In all likelihood, the increased concentration of organic acids in the reactors inhibits representatives of the strictly aceticlastic genus *Methanosaeta* and stimulates the development of *Methanosarcina* spp. [14, 25]. ●

This work was supported by the grant "Alğarış" of the Republic of Tatarstan (Russia) (2010) and the joint scholarship of the DAAD Program and the Ministry of Education and Science of the Russian Federation ("Mikhail Lomonosov II" Program, 2011).

REFERENCES

1. Antoni D., Zverlov V.V., Schwarz W.H. // Appl. Microbiol. Biotechnol. 2007. V. 77. P. 23–35.
2. Krause L., Diaz N.N., Edwards R.A., Gartemann K.H., Krömeke H., Neuweger H., Pühler A., Runte K.J., Schlüter A., Stoye J., et al. // J. Biotechnol. 2008. V. 136. P. 91–101.
3. Ahn H.K., Smith M.C., Kondrad S.L., White J.W. // Appl. Biochem. Biotechnol. 2010. V. 160. P. 965–975.
4. Bedoya I.D., Arrieta A.A., Cadavid F.J. // Bioresour. Technol. 2009. V. 100. P. 6624–6629.
5. Goberna M., Insam H., Franke-Whittle I.H. // Appl. Environ. Microbiol. 2009. V. 75. P. 2566–2572.

6. Weiland P. // Appl. Microbiol. Biotechnol. 2010. V. 85. P. 849–860.
7. Ziganshin A.M., Schmidt T., Scholwin F., Il'inskaya O.N., Harms H., Kleinsteuber S. // Appl. Microbiol. Biotechnol. 2011. V. 89. P. 2039–2052.
8. Holm-Nielsen J.B., Seadi T.A., Oleskowicz-Popiel P. // Bioresour. Technol. 2009. V. 100. P. 5478–5484.
9. El-Mashad H.M., Zhang R. // Bioresour. Technol. 2010. V. 101. P. 4021–4028.
10. Gomez X., Moran A., Cuetos M.J., Sanchez M.E. // J. Power Sources. 2006. V. 157. P. 727–732.
11. Davidsson A., Lövstedt C., la Cour Jansen J., Gruvberger C., Aspegren H. // Waste Manage. 2008. V. 28. P. 986–992.
12. Fountoulakis M.S., Petousi I., Manios T. // Waste Manage. 2010. V. 10. P. 1849–1853.
13. Nayono S.E., Gallert C., Winter J. // Bioresour. Technol. 2010. V. 101. P. 6998–7004.
14. Demirel B., Scherer P. // Rev. Environ. Sci. Biotechnol. 2008. V. 7. P. 173–190.
15. O'Reilly J., Lee C., Collins G., Chinalia F., Mahony T., O'Flaherty V. // Water Res. 2009. V. 43. P. 3365–3374.
16. Kröber M., Bekel T., Diaz N.N., Goesmann A., Jaenicke S., Krause L., Miller D., Runte K.J., Viehöver P., Pühler A., Schlüter A. // J. Biotechnol. 2009. V. 142. P. 38–49.
17. Lee C., Kim J., Hwang K., O'Flaherty V., Hwang S. // Water Res. 2009. V. 43. P. 157–165.
18. Nettmann E., Bergmann I., Pramschüfer S., Mundt K., Plogsties V., Herrmann C., Klocke M. // Appl. Environ. Microbiol. 2010. V. 76. P. 2540–2548.
19. Abdo Z., Schüette U.M., Bent S.J., Williams C.J., Forney L.J., Joyce P. // Environ. Microbiol. 2006. V. 8. P. 929–938.
20. Culman S.W., Bukowski R., Gauch H.G., Cadillo-Quiroz H., Buckley D.H. // BMC Bioinformatics. 2009. V. 10. P. 171–180.
21. Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. // J. Mol. Biol. 1990. V. 215. P. 403–410.
22. Wang Q., Garrity G.M., Tiedje J.M., Cole J.R. // Appl. Environ. Microbiol. 2007. V. 73. P. 5261–5267.
23. Gerardi M.H. The microbiology of anaerobic digesters. Hoboken: Wiley-Interscience, 2003. 177 p.
24. Kendall M.M., Boone D.R. // The order Methanosarcinales. The Prokaryotes – a Handbook on the Biology of Bacteria. 3rd ed. / Eds Dworkin M., Falkow S., Rosenberg E., Schleifer K.H., Stackebrandt E. New York: Springer, 2006. P. 244–256.
25. Karakashev D., Batstone D.J., Angelidaki I. // Appl. Environ. Microbiol. 2005. V. 71. P. 331–338.

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.

- 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 improvement does not mean that the article has been accepted for publication. After the revised text has been received, a decision is made by the editorial board. The author must return the improved text, together with the 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

NANOTECHNOLOGIES *in Russia*

Peer-review scientific journal

Nanotechnologies in Russia (Rossiiskie Nanotekhnologii) focuses on self-organizing structures and nanoassemblages, nanostructures including nanotubes, functional nanomaterials, structural nanomaterials, devices and facilities on the basis of nanomaterials and nanotechnologies, metrology, standardization, and testing in nanotechnologies, nanophotonics, nanobiology.

→ Russian edition: http://nanoru.ru

→ English edition: http://www.springer.com/materials/nanotechnology/journal/12201

Issued with support from:

The Ministry of Education and Science of the Russian Federation