

Letter from the Editors

Dear readers of *Acta Naturae*! To open the current issue of the Journal, let's first congratulate all of us on a remarkable event: according to Thomson Reuters, our journal's impact factor has now reached 0.872. Therefore, we are now in second place among Russian life sciences journals, second only to *Biochemistry*. The journal's IF has almost doubled in the last year! This is not boasting, but we do believe it is a great success. The decisive factor, in our opinion, was our push to be included in international databases and, first of all, to insure free online access to our journal. It seems that we shouldn't, in any case, be content with what we have achieved, because 0.872 is a very modest figure at the international level. We are confident that continued coordinated work between the editorial board, publishers, and, above all, our authors will further increase *Acta Naturae's* rating.

Now let us turn to the new issue. It begins with a report on the third international conference "Genetics of Aging and Longevity" held in Sochi. Life is such that over time this issue is bound to become a concern for all of us. In the context of the conference, it was clear how deeply molecular methods have penetrated the field of medicine, and how much progress has been achieved in gerontology. Let's live longer and better lives!

The issue includes a large number of reviews on various topics. The review by P.V. Dubovskii and Yu.N. Utkin is devoted to cobra cardiotoxins and their antibacterial activity. A.A. Nemudryi *et al.* describe the newest methods of genome editing (the TALEN and CRISPR systems) and their application in the editing and functional screening of genomes, in generating cell models of human

hereditary diseases, studying epigenomes, and visualizing cellular processes. A review by Kazan microbiologists (V.M. Chernov *et al.*) focuses on the contamination of cell cultures by mycoplasmas. Any researcher who uses cell biology methods faces this problem, and the review will certainly be of interest to a very wide range of readers. A publication by V.V. Koval, D.G. Knorre, and O.S. Fedorova is devoted to the molecular basis of DNA repair. Finally, a review by I.M. Larina *et al.* analyzes the approaches used in OMICs disciplines (proteomics, transcriptomics, and metabolomics) for a deeper understanding of the physiological adaptation of a healthy individual to extreme conditions.

Three full-length research articles and a brief report are also offered to your attention. The article by I.V. Golubev *et al.* is dedicated to the protein engineering of D-amino acid oxidase, an enzyme which is essential in biotechnology. E.V. Chetina *et al.* studied the relationship between the expression of the genes of the structural proteins of the extracellular matrix, proteinases, and their inhibitors during fetal development. The publication by N.Yu. Glazova *et al.* is devoted to the influence of an antidepressant, fluvoxamine, on neonatal development. A brief report by A.P. Yakimov *et al.* concludes the issue by describing the cloning of the *ribT* gene, which is part of the riboflavin biosynthesis operon of *Bacillus subtilis*.

We hope that the articles in this issue of *Acta Naturae* will be interesting and useful to you. We look forward to your comments and new articles!

See you in the next issue! ●

Editorial Board

15-17 october 2014
**SAINT-PETERSBURG
EXPOFORUM**



ST. PETERSBURG INTERNATIONAL HEALTH FORUM



**MEDIZ
SAINT-PETERSBURG**
Medicine and Health
www.mediz-spb.ru



PHARMACY
www.pharma.primexpo.ru



BIOINDUSTRY
www.bio.expoforum.ru



HEALTHCARE TOURISM
www.healthtourism.primexpo.ru



VISUS-EXPO
www.kmc-med.com



AESTHETIC MEDICINE
www.primexpo.ru



www.pmfz.expoforum.ru

+7 812 240 40 40

EXPOFORUM



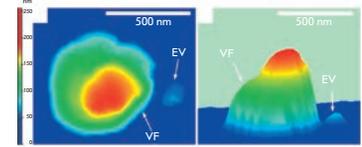
0+



Mycoplasma Contaminations of Cell Cultures: Vesicular Traffic in Bacteria and the Infectious Genes Control Problem

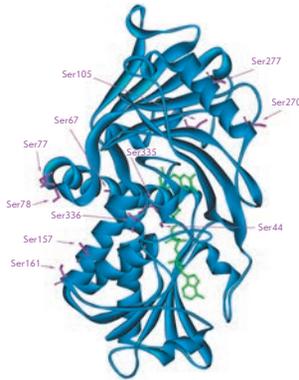
V.M. Chernov, O.A. Chernova, J. T. Sanchez-Vega, A.I. Kolpakov, O.I. Il'inskaya

Cell cultures are subjected to contamination both by cells of other cultures and by microorganisms. Contamination of cell cultures by mycoplasmas is of particular significance. This review describes features of mycoplasmas and their extracellular vesicles as well as the interaction of contaminants with eukaryotic cells, analyzes the problems of modern diagnostic methods and eradication of mycoplasma contamination of cell cultures and perspectives for their solution.



A. laidlawii PG8 cells

A Study of the Structure-Function-Stability Relationship in Yeast D-Amino Acid Oxidase: Hydrophobization of α -Helices



I.V. Golubev, N.V. Komarova, K.V. Ryzhenkova, T.A. Chubar, S.S. Savin, V.I. Tishkov

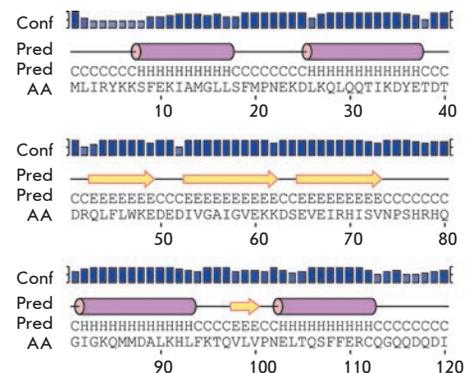
Within the approach to enhancement of the protein thermostability through hydrophobization of α -helices in D-amino acid oxidase from the *Trigonopsis variabilis* yeast (TvDAAO, [EC 1.4.3.3]), 8 mutants with Ser/Ala substitutions were produced. Mutations at positions 78, 270, 277, and 336 led to the enzyme destabilization, Ser77Ala and Ser335Ala substitutions had almost no effect on the TvDAAO stability, and Ser67Ala and Ser105Ala substitutions increased the thermostability by 1.5 and 2 times, respectively. The k_{cat}/K_M value in Ser105Ala TvDAAO with a number of D-amino acids was 1.2–3 times higher than that in the parent enzyme.

A general view of the TvDAAO subunit with Ser residues located in the α -helical regions

Possible Function of the *ribT* Gene of *Bacillus Subtilis*: Theoretical Prediction, Cloning, and Expression

A.P. Yakimov, T.A. Seregina, A.A. Kholodnyak, R.A. Kreneva, A.S. Mironov, D.A. Perumov, A.L. Timkovskii

The complete decipherment of functions and interaction of elements of the riboflavin biosynthesis operon (*rib* operon) of *Bacillus subtilis* is necessary for the development of superproducers of this important vitamin. The function of its terminal *ribT* gene has not been established up to date. In this work, the search for homologs of the hypothetical amino acid sequence of the gene product through databases as well as an analysis of the homologs were performed, the distribution of secondary structure elements was theoretically predicted, and the tertiary structure of the RibT protein was proposed. The real possibility of RibT protein production in quantities sufficient for further investigation of its structure and functional activity was demonstrated.



Secondary structure elements predicted for the *ribT* gene product

Founders

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

Editorial Council

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

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

Editorial Board

Managing Editor: V.D. Knorre
Publisher: K.V. Kiselev

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

Project Head: S.B. Nevskaya
Editor: N.Yu. Deeva

Strategic Development Director: E.L. Pustovalova

Designer: K.K. Oparin

Photo Editor: I.A. Solovey

Art and Layout: K. Shnaider

Copy Chief: Daniel M. Medjo

Address: 119234 Moscow, Russia, Leninskiye Gory, Nauchny
Park MGU, vlad.1, stroeniye 75G.

Phone/Fax: +7 (495) 930 88 50

E-mail: vera.knorre@gmail.com, mmorozova@strf.ru,
actanaturae@gmail.com

Reprinting is by permission only.

© ACTA NATURAE, 2014

Номер подписан в печать 28 августа 2014 г.

Тираж 200 экз. Цена свободная.

Отпечатано в типографии «МЕДИА-ГРАНД»

CONTENTS

Letter from the Editors 1

FORUM

A.V. Belikov, M.V. Shaposhnikov,
A.A. Moskalev

There Are Many Ways to Fight Aging. 6

REVIEWS

P.V. Dubovskii, Y.N. Utkin

Cobra cytotoxins: structural organization
and antibacterial activity 11

A.A. Nemudryi, K.R. Valetdinova,
S.P. Medvedev, S.M. Zakian

TALEN and CRISPR/Cas Genome Editing
Systems as Discovery Tools 19

V. M. Chernov, O. A. Chernova,
J. T. Sanchez-Vega, A. I. Kolpakov,
O. N. Ilinskaya

Mycoplasma Contamination Of Cell Cultures:
Vesicular Traffic In Bacteria And Control
Of Infectious Agents. 41

V. V. Koval, D. G. Knorre, O. S. Fedorova
**Structural Features of the Interaction
 between Human 8-Oxoguanine DNA
 Glycosylase hOGG1 and DNA**52

I. M. Larina, V. A. Ivanisenko, E. N. Nikolaev,
 A. I. Grigorev
**Proteome of healthy human during
 PHYSICAL activity in extreme
 environments**66

RESEARCH ARTICLES

I. V. Golubev, N. V. Komarova,
 K. V. Ryzhenkova, T. A. Chubar, S. S. Savin,
 V. I. Tishkov
**Study of structure-function-stability
 relationships in yeast D-amino acid oxidase:
 hydrophobization of alpha-helices**76

E. V. Tchetina, F. Mwale, A. R. Poole
**Changes in Gene Expression Associated with
 Matrix Turnover, Chondrocyte Proliferation and
 Hypertrophy in the Bovine Growth Plate**89

N. Yu. Glazova, S. A. Merchieva,
 M. A. Volodina, E. A. Sebentsova,
 D. M. Manchenko, V. S. Kudrun,
 N. G. Levitskaya
**Effects of neonatal fluvoxamine administration
 on physical development and serotonergic
 system activity in white rats.**98

SHORT REPORTS

A.P. Yakimov, T.A. Seregina, A.A. Kholodnyak,
 R.A. Kreneva, A.S. Mironov, D.A. Perumov,
 A.L. Timkovskii

**Possible Function of the *ribT* Gene of *Bacillus
 subtilis*: Theoretical Prediction, Cloning, and
 Expression**106

Guidelines for Authors.....110

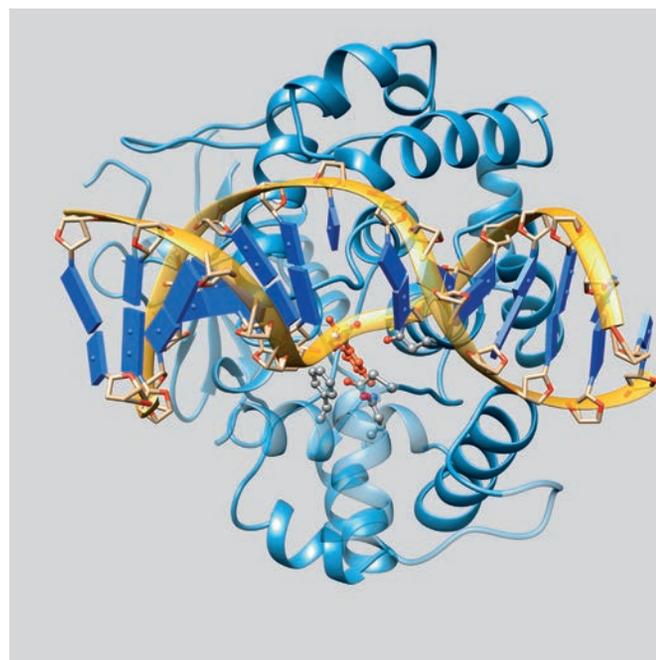


IMAGE ON THE COVER PAGE

See the article by Koval et al.

There Are Many Ways to Fight Aging. Results of the Third International Conference “Genetics of Aging and Longevity”

A. V. Belikov¹, M. V. Shaposhnikov^{2,3}, A. A. Moskalev^{1,2,3*}

¹Moscow Institute of Physics and Technology, Institutskiy Pereulok, 9, Dolgoprudny, Moscow Region, Russia, 141700

²Institute of Biology, Komi Science Centre, Ural Division of the Russian Academy of Sciences, Kommunisticheskaya Str., 28, Syktyvkar, Russia, 167982

³Syktyvkar State University, Oktyabrsky Prosp., 55, Syktyvkar, Komi Republik, Russia, 167001

*E-mail: amoskalev@list.ru

The 3rd International Conference “Genetics of Aging and Longevity” was held in Sochi, Russia, April 6 to 10, 2014, which was attended by over 300 delegates from 18 countries. Apart from 50 oral presentations given by world-leading gerontologists and an extensive poster session, four round tables dedicated to the development and integration of aging theories, the development of personalized medicine, and attracting venture capital to research were held. The conference participants signed an open letter to the World Health Organization with the appeal to organize the worldwide collection and integration of data on age-related diseases. The conference clearly demonstrated the progress in investigation of aging and the development of technologies for intervention in this process. In particular, identification of somatic mutations and DNA damages in individual cells has become possible; reliable markers of biological age have been characterized; many genes, alleles, processes, metabolites, intestinal bacteria, and external factors that influence the aging rate have been identified; online databases of age-related changes as well as the genomes of long-lived individuals and long-lived animal species have been generated. Preclinical trials of pharmacological agents that are likely capable of slowing aging, such as nicotinamide riboside, selective inhibitors of TORC1, and IGF-1 receptor-blocking antibodies, have been carried out. Methods of cultivation and transplantation of artificial organs, targeted delivery of drugs to individual cells and organelles, targeted genome editing, and introduction of artificial chromosomes have been suggested.

In April 2014, the 3rd International Conference “Genetics of Aging and Longevity” organized by the Institute of Biology, Komi Science Centre, the Ural Division of the Russian Academy of Sciences together with the “Science for Life Extension” foundation was held at Congress Centre of the Radisson Blu Resort hotel in the post-Olympic Sochi. Researchers from 18 countries (Russia, USA, UK, the Netherlands, Germany, Italy, China, Israel, Jordan, Ukraine, Belarus, Azerbaijan,

Poland, Uzbekistan, Kazakhstan, Canada, Estonia, and Sweden) working in different fields of biology, medicine, and computer science as well as entrepreneurs and investors gathered for five days, being joined together by the desire to understand the mechanism(s) of aging and to develop methods to prevent and possibly to suppress processes underlying all age-related diseases.

The key topics of the conference were the search for molecular targets of drugs slowing aging; mapping the genomes of long-lived

individuals and long-lived animal species; the relationship of genes, epigenetic regulation, metabolism, intestinal microflora, lifestyle, and environment in controlling life expectancy of the individual; the development of technologies for artificial growth, cryopreservation, and transplantation of organs; the use of nanocapsules and artificial chromosomes for genome editing.

One of the important trends in gerontology is the search for molecular targets for drugs to cause slowing or ceasing the aging proc-

ess. For example, Michael Petrascheck (The Scripps Research Institute, USA) presented the results of screening conducted in nematodes that identified 57 compounds with known mammalian pharmacology that extend the lifespan of *Caenorhabditis elegans*, with 16 of them providing a more than 30% increase in the lifespan. The main targets of these compounds were receptors for dopamine, serotonin, epinephrine, norepinephrine, and hormones. Most (33) of these compounds also increase resistance to oxidative stress. Brian Kennedy (Buck Institute for Research on Aging, USA) spoke about research aimed at separating the beneficial effects of rapamycin, primarily increasing the life span, from its side effects, in particular, induction of insulin resistance. Rapamycin appeared to inhibit the TORC1 (beneficial effects) and TORC2 (side effects) complexes via different mechanisms. This makes it possible to develop selective TORC1 inhibitors, which is the research area of Delos Pharmaceuticals. Preclinical trials revealed a decrease in severity of side effects of this new class of drugs. Nir Barzilai (Albert Einstein College of Medicine, USA) reported that injection of the insulin-like growth factor 1 (IGF-1), but not insulin, into the brain ventricles of old rats improves the peripheral insulin sensitivity via the IGF-1 receptor. He also pointed out the prospect of using IGF-1 receptor-blocking antibodies that are not able to cross the blood-brain barrier for reducing the effect of IGF-1 on the peripheral (carcinogenic), but not on the central nervous system (safety). The use of these antibodies in a combination with rapamycin would probably minimize rapamycin-induced insulin resistance. Furthermore, inhibition of the very IGF-1 homolog, DAF-2, extends the life span of nematodes. For example, David Gems (University College London,

UK) expanded our understanding of the mechanisms of this phenomenon. The DAF-16 transcription factor turned out to activate AAKG-4, the subunit of atypical AMP-independent AMP-kinase. This subunit, in turn, accelerates activation of DAF-16 by direct phosphorylation; it is necessary for nematode life span extension caused by inactivation of DAF-2. Another target of DAF-16, the MDL-1 transcription factor, also promotes the positive effect of DAF-2 inactivation on the lifespan of nematodes.

According to the report of Matt Kaeberlein (University of Washington, USA), rapamycin extends the life span and improves the condition of mice with a defect of the mitochondrial complex I. These mice are used as a model organism in studying subacute necrotizing encephalomyopathy (Leigh syndrome). The number of defective mitochondria is known to increase with age. Rapamycin is supposed to prevent glycolytic product accumulation and lactic acidosis. M. Kaeberlein also supposed that the cause of a mitochondrial disorder may be intracellular NAD depletion. Thus, the possibility to restore the NAD level, to correct the Leigh syndrome, and to extend the life span, when nicotinamide riboside is included in the diet of mice, is under study. Nicotinamide riboside is already on the US market, as a vitamin supplement, under the brand name NIAGEN™. Blanka Rogina (University of Connecticut Health Center, USA) reported that the dPGC-1 protein increases the lifespan of *Drosophila* caused by a mutation in the transmembrane citric acid transporter, INDY, via stimulating mitochondrial biogenesis and reducing the level of oxidative stress. A reduced intestinal INDY level supports homeostasis of stem cells and, as a consequence, the intestinal tract integrity. Identification of a human INDY homolog would help, per-

haps, to develop new drugs slowing aging. Robert Shmookler Reis (University of Arkansas for Medical Sciences, USA) spoke about the discovery of a new protein, CRAM-1, in intracellular protein aggregates of a mutant nematode used to study the Huntington's disease. Inactivation of this protein leads to a decrease in size of the aggregates, a delayed development of paralysis, and restoration of chemotaxis in two mutant nematodes, which are models of the Huntington's disease and Alzheimer's disease. CRAM-1 appeared to cause condensation of oligoubiquitin, which probably ceases proteasomal protein degradation and prevents autophagy. Pharmacological inhibitors of the protein homolog in humans could be useful in treatment of age-associated neurodegenerative diseases.

Another promising trend in gerontology is investigation of genes extending the lifespan of model organisms. For example, Vera Gorbunova (University of Rochester, USA) reported that SIRT6, a protein, overexpression of which extends the lifespan of mice, plays a role of the genome guardian. SIRT6 promotes repair of DNA damages via mono-ADP-ribosylation and activation of the PARP-1 protein. A regulator of the SIRT6 protein was found, and SIRT6 was demonstrated to maintain the genome stability also in other ways. Bill Orr (Southern Methodist University, USA) spoke how the redox-sensing functions of peroxiredoxins can regulate expression of longevity genes in *Drosophila*. PRX-5 controls the balance between innate immunity and aging. A reduction in the PRX-5 expression increases resistance to infections, whereas an excess of PRX-5 leads to decreased immunity along with an increase in the life span (30%) and resistance to oxidative stress.

Certainly, an indispensable tool in the search for new longevity genes

is studying long-lived individuals and long-lived animal species. For example, based on genome-wide population studies, Claudio Franceschi (University of Bologna UNIBO, Italy) found the protective allele *APOE* (T), the carriership of which correlates with low blood pressure, low probability of stroke, and large lifespan. Unexpectedly beneficial effects of mutations in the mitochondrial complex I were also revealed that are observed in the complexes III and V in the absence of mutations. Yousin Suh (Albert Einstein College of Medicine, USA) reported that certain variants of single nucleotide polymorphisms in promoter regions of the *sirt-1* gene inhibit binding of the transcriptional repressor CTCF and promote binding of the transcriptional activator ZFR. This prevents the *sirt-1* activation in oxidative stress and increases the risk of myocardial infarction. A longevity-associated variant of single nucleotide polymorphisms in the enhancer region of the *foxo-3* gene similarly enhances expression of this gene in response to oxidative stress. The SKAT analysis of single nucleotide polymorphisms in centenarians and control groups demonstrated that first 25 longevity-associated genes comprise mainly genes associated with identification and repair of double-strand DNA breaks.

Vadim Gladyshev (Harvard Medical School, USA) presented the results of genome and gene expression analyses in mammals. Sequencing the genomes of atypically long-lived mammals, the naked mole rat (underground rodent) and Brandt's bat, and subsequent comparison with the genomes of closely related species revealed a number of longevity-associated genes. Research of this group is aimed at comparing gene expression in species with low and high life span in order to identify new longevity genes. Andrei Seluanov (University of Rochester,

USA) reported that high molecular weight hyaluronic acid protects the naked mole rat from the cancer development, while low molecular weight hyaluronic acid, in contrast, acts as a carcinogen. Furthermore, naked mole rat ribosomes are capable of synthesizing proteins with the record low number of errors. Joao Pedro de Magalhaes (University of Liverpool, UK) provided some results of genome sequencing for the bowhead whale, which is the most long-lived mammal. Only five genes subjected to the recent action of selection, including those responsible for the immune response, were found.

DNA mutations and damages play a role in aging also at the level of individual somatic cells. Jan Vijg (Albert Einstein College of Medicine, USA) presented new methodologies to identify and analyze somatic mutations and epigenetic changes, such as single cell whole genome sequencing and single cell transcripto-genomics that allows a researcher to identify whether a studied mutation is transcribed or it remains silent. Studies in mice and *Drosophila* conducted using these methodologies demonstrated that the number of somatic mutations and the fraction of silent mutations increase with age. Alex Maslov (Albert Einstein College of Medicine, USA) presented a new method to assess the level of DNA damages using quantitative PCR amplification of long fragments. A slight age-related increase in the level of DNA damages in the liver, but not in the brain, of laboratory animals was found. Another method to determine structural somatic DNA variations (translocations, inversions, duplications) using high-throughput sequencing helped to identify an increase in the number of these changes in the liver and brain of aging mice, but not in the intestines or heart. Somatic cells affected by DNA damage change to

the state of senescence, an irreversible delay of the cell division. Judy Campisi (Buck Institute for Research on Aging, USA) presented a transgenic mouse that is suitable for visualization and destruction of senescent cells identified by the p16-INK4a overexpression. Senescent cells appeared to accumulate after exposure to ionizing radiation or introduction of doxorubicin, to promote metastasis of tumors, and to be responsible for side effects of chemotherapy. At the same time, a short-term (but not persistent) presence of senescent cells in skin lesions facilitates accelerating their healing by secretion of a potent growth factor.

Currently, increasing support is given to the theory about the relationship of genes, epigenetic regulation, metabolism, intestinal microflora, lifestyle, and environment in controlling the individual's life span. For example, Claudio Franceschi (University of Bologna UNIBO, Italy) demonstrated that the metabolic profile of centenarians differs drastically from that of seniors, but is similar to the profile of a young population. Furthermore, according to microbiome shotgun sequencing, relative abundance of *Proteobacteria*, such as the *Escherichia* and *Ruminococcus* genera, is increased in the intestines of centenarians. Daniel Promislow (University of Washington, USA) presented a new methodology for studying the metabolome that combines highly sensitive orbitrap mass spectrometry with the WGCNA analysis of data to identify metabolomic modules of correlating metabolites. He demonstrated that caloric intake restriction, which is a repeatedly confirmed method to increase the life span, has a strong effect not only on the levels of certain metabolites in *Drosophila*, but also on the interaction among them. Alexei Moskalev (Institute of Biology, Komi Scientific Center,

Ural Division of the Russian Academy of Sciences; Moscow Institute of Physics and Technology, Russia) discussed the mechanisms for extending the life span after moderate stress, such as exposure to ionizing radiation (hormesis phenomenon), discovered in mutant *Drosophila* lines. In this case, destruction of sensitive cells, stimulation of the cell stress response, activation of the immune system, and acceleration of organism growth occur. RNA sequencing revealed an increase in the activity of four genes (including *sugarbabe*, *tram-track*, and *fat*) and a reduced activity of 48 age-related genes (e.g., *keap-1* and *relish*) after exposure to a low dose of the γ -radiation.

Investigation of the aging process, with allowance for its complexity, is impossible without involvement of computer technologies. For example, Joao Pedro de Magalhaes (University of Liverpool, UK) spoke about online databases: Digital Aging Atlas, Human Aging Genomic Resources, and the Naked Mole-Rat Genome Resource. His university colleague, Daniel Wuttke (University of Liverpool, UK), presented a new open-source online platform, Denigma, for collaboration of scientists to study the aging mechanisms. This platform is based on machine learning using databases, data sets as well as individually input data and on developing a model for aging on the basis of these data. The system uses ontologies, decomposition of problems into subproblems, and machine logic.

The search for reliable biomarkers of aging is the urgent need of gerontology. For example, Claudio Franceschi (University of Bologna UNIBO, Italy) reported on several biomarkers of biological age, which were confirmed in population studies, such as blood levels of N-glycans, particularly the NGA2F/NA2F ratio, hypermethylation of the *elovl-2* and *fhl-2* loci, metabolic

signatures in blood and urine, and circulatory mitochondrial DNA. Ancha Baranova (Research Centre for Medical Genetics, Russia; George Mason University, Fairfax VA, USA) presented a simple, but reliable method to identify stages of various diseases or aging that is based on mRNA profiling using oligonucleotide microarrays or low coverage high-throughput sequencing followed by calculating the profile distance from the norm in false coordinates. This method is suitable to predict the progression of malignant tumors.

A separate section of the conference was devoted to bioengineered approaches to healthy aging. For example, Shay Soker (Wake Forest School of Medicine, USA) reported on the latest advances in growing organs in bioreactors. Simple organs, such as cornea, blood vessels, and bladder, are grown relatively easily, while complex ones, such as liver, kidney, and pancreas, require scaffolds. Paolo Macchiarini (Karolinska Institutet, Sweden) presented a technology for growing the non-immunogenic trachea and other thoracic organs and their clinical application. Mesenchymal stem cells and mononuclear leukocytes were experimentally found to be equally well suitable for cellularization of bioscaffolds or bioartificial three-dimensional nanomaterials. In addition, G-CSF (granulocyte-colony stimulating factor) appeared to accelerate cellularization and survival of bioscaffolds in the human body, while erythropoietin reduces the fraction of apoptotic cells. Greg Fahy (21st Century Medicine Inc., USA) spoke about the advances and problems in organ vitrification. Cryoprotective mixtures, protocols for organ perfusion with these mixtures under high pressure, and methods for rapid organ defrosting were developed. Nevertheless, the differences in an optimal freezing rate of dif-

ferent cell types and organ zones remain a serious problem.

Andre Watson (Ligandal Technology, USA) presented a commercial technology for targeted delivery of drugs and non-viral genome editing tools (such as CRISPR and TALEN) to individual cells and organelles by means of nanocapsules. The efficiency of this technology has been confirmed in a number of experiments *in vitro* and *in vivo*. Ksenia Yurieva (Human Stem Cells Institute, Russia) reported on the emergence of technologies for generating artificial human chromosomes. Currently, there are two approaches: “top-down”, in which a normal chromosome is depleted of all genes, with telomere and centromere regions only being left, and “bottom-up”, when a chromosome is synthesized from scratch. To copy and introduce artificial chromosomes, a microcell transfer technology is used, in which carrier cells are used for multiplication of chromosomes, then they are fragmented, and the resultant single chromosome fragments are fused to target cells. The opportunities offered by selective genome editing and implementation of artificial chromosomes to extend the human life span are truly endless.

Therefore, pharmacological and genetic experiments on model organisms as well as comprehensive investigation of long-lived individuals and long-lived animal species will identify the relationship of genes, epigenetic regulation, metabolism, intestinal microflora, lifestyle, and environment in controlling the individual’s life span as well as discover new genes, alleles, processes, metabolites, strains of intestinal bacteria, and external factors affecting the aging rate. The progress of research is facilitated by the introduction of new techniques and technologies, such as identification of somatic mutations and quantification of DNA damages

in individual cells, determination of complete metabolic profiles or identification of biological age using biochemical or genetic markers. Due to the extraordinary complexity of the aging process and a huge amount of accumulated knowledge, the introduction and extensive use of online databases of age-related changes as well as online depositories of the genomes of long-lived individuals and long-lived animal species, including resources using artificial intelligence, are becoming indispensa-

ble. All these efforts have enabled the development of drugs that are likely capable of slowing aging and are at the stage of preclinical trials: nicotinamide riboside, selective TORC1 inhibitors, and IGF-1 receptor-blocking antibodies. In the near future, the emergence of new classes of anti-aging drugs is possible, for example, those that inhibit inflammatory responses, oxidative stress and the formation of protein conglomerates, activate DNA repair, destroy senescent cells, or

even affect neurotransmission. In parallel, bioengineering approaches to healthy aging are developed, including growing and transplantation of artificial organs, targeted delivery of drugs to cells and organelles, directed genome editing, and introduction of artificial chromosomes.

The conference participants addressed the World Health Organization with the appeal about the need to monitor and integrate data related to age-dependent diseases. ●

Cobra Cytotoxins: Structural Organization and Antibacterial Activity

P. V. Dubovskii, Y. N. Utkin*

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

*E-mail: utkin@mx.ibch.ru

Received: 08.04.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT Cardiotoxins (cytotoxins, CT) are β -structured proteins isolated from the venom of cobra. They consist of 59–61 amino acid residues, whose antiparallel chains form three ‘fingers’. In contrast to neurotoxins with an overall similar fold, CTs are amphiphilic. The amphiphilicity is caused by positively charged lysine and arginine residues flanking the tips of the loops that consist primarily of hydrophobic amino acids. A similar distribution of amino acid residues is typical for linear (without disulfide bonds) cationic cytolytic peptides from the venoms of other snakes and insects. Many of them are now considered to be lead compounds in combatting bacterial infections and cancer. In the present review, we summarize the data on the antibacterial activity of CTs and compare it to the activity of linear peptides.

KEYWORDS antibacterial activity, lipopolysaccharide, peptidoglycan, plasma membrane, three-finger cardiotoxins (cytotoxins), cytolytic cationic peptides.

ABBREVIATIONS AMP – antimicrobial peptide; GAG – glucosaminoglycan; CL – cardiolipin; LPS – lipopolysaccharide, LTA – lipoteichoic acid; XRA – X-ray analysis; PG – phosphatidylglycerol; PE – phosphatidylethanolamine; CT – cytotoxin (cardiotoxin) from snake venom, NMR – nuclear magnetic resonance.

INTRODUCTION

Cytolytic peptides are present in the venom of snakes and insects. They contain extended hydrophobic regions flanked by positively charged lysine and arginine residues as a structural motif [1]. Cytolytic peptides can be either linear [2–7] or contain disulfide bonds [8]. In the latter case, they can only be β -sheets [9–12] or contain both β -sheet and α -helical regions [13, 14]. The interest in cytolytic peptides stems from the fact that some of them display both antibacterial and anti-proliferative activities [15–18]. They are widely used to design peptides with improved therapeutic indices [19–23]. The design process is typically based on the principle of combining various motifs in one peptide (e.g., the cytolytic motif, the motif inducing membrane fusion, and the one promoting cell penetration). However, the systematic use of such peptides is undermined by their susceptibility to proteolysis in the bloodstream [24, 25]. Therefore, we believe that peptides with compact structures stabilized by one or several disulfide bonds would be of great interest.

Three-finger toxins from cobra venom belong to the family of cytotoxins (cardiotoxins, CT) [12, 26–28] and can kill various types of cells by disrupting their plasma membranes. Studies of CT interaction with model lipid membranes have demonstrated that its mechanism depends on the cytotoxin type: either P or S [29, 30]. The

P-type includes CTs with a Pro30 residue at the tip of the second loop; the S-type, those with a Ser28 residue (*Table*). Data on CT interactions with model phospholipid membranes suggest that these toxins destabilize the lipid bilayer of anionic phospholipid-containing membranes [30, 31]. Evidently, in a living cell CTs target the plasma membrane (or the membranes of intracellular organelles) that contain such phospholipids. The interaction of CTs with the components of the surface membrane of eukaryotic cells disrupts its barrier properties and/or leads to the penetration of CT into the cell and its subsequent interaction with organelles, resulting in cell death [32–35]. Presumably, this scenario requires anionic glycolipid sulfatide to be present in the membrane [36]. On the other hand, bacterial cell membranes are almost entirely composed of anionic phospholipids [37] and therefore should be substantially more vulnerable to CTs. The aim of this review is to verify this claim.

All CT molecules contain such structural and functional motifs as the membrane-binding motif and a ‘belt’ of charged residues surrounding it, as well as clusters of conserved polar residues [12]. We can expect toxin activity to be defined by the efficiency of these motifs at certain stages of the CT penetration into a bacterial cell, as well as by the succession of their involvement in the interactions with the cell. Let us first examine the spatial structure of a CT molecule.

Cardiotoxins: properties and conformational characteristics

Cobra species, <i>Naja</i>	Abbreviation	Alternative names	ID ¹	I/II ²	S/P ³	HTL ⁴	Net positive charge (neutral pH)	Method	PDB code ⁵
<i>N. mossambica</i>	M1	CTX: IIb, VIII CT-1	P01467	I	P	3.4	8	NMR	2CCX
	M3	CTX VII4 CT-4	P01470	II	S	1.0	10	XRA	1CDT
<i>N. atra</i>	A1	CTX:-1, I CT:-1(CX1)	P60304	II	S	12.5	7	NMR	2CDX
	A2	CTX:-2, II CT:1A, -2(CX2)	P01442	II	S	12.9	8	NMR	1CRF 1CRE
	A3	CTX:-3, III CT-3	P60301	II	P	11.7	9	NMR	2CRT, 2CRS
								NMR	1I02
								XRA	1H0J
								XRA	1XT3
	A4	CTX:-4, IV CT:-4	P01443	II	S	12.9	9	NMR	1KBT 1KBS
A4b	CTX:-A4b; -T CT: D-1; -5	P07525	II	S	9.8	9	NMR	1CHV	
A6	CTX:6, N CT:-6, N	P80245	I	P	9.3	8	XRA	1UG4	
<i>N. oxiana</i>	CII (CTII)	CT-2	P01441	II	P	16.3	10	NMR	1CB9, 1CCQ
	CTI (CTI)	CT-1	P01451	II	S	8.9	6	NMR	1FFJ
								NMR	1RL5
<i>N. pallida</i>	Tg	CTX: gamma CT-1	P01468	I	P	3.4	9	NMR	1CXO
								XRA	1TGX

¹Code of the amino acid sequence in the Swiss-Prot database of protein structures (www.uniprot.org).

²Classification into CT Group I and II is based on the presence of either two Pro (Group I) or a single Pro (group II) residues in the loop I sequence.

³Classification of CT into S- and P-type is based on the presence of the S28 and P30 residues, respectively, at the end of loop II.

⁴Residues 5-11, 24-37, 46-50 and the Kyte-Doolittle hydrophobicity scale were used for calculations, a higher value corresponds to a higher hydrophobicity of the HTL.

⁵Protein structure database PDB (www.rcsb.org/pdb/home/home.do).

SPATIAL STRUCTURE OF A CT MOLECULE

The researchers working at the Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry of the Russian Academy of Sciences have made a significant contribution to the elucidation of the spatial structure of CT molecules, which has been discussed in several papers [30, 38–40] and reviews [12, 23, 27]. Herein, we present a brief overview. CTs are characterized by a high degree of homology of their amino acid sequences. *Figure 1* shows the alignment of CT amino acid sequences, whose spatial architecture was determined by X-ray diffraction or NMR. *Table* lists the supplementary data (references, short names, charges of these toxins, etc.).

All CTs are β -sheet proteins with three-finger folding [41] (*Fig. 2*). Four disulfide bonds formed by eight cysteine residues are the conserved elements of their spatial architecture (*Fig. 1*). It should be noted that the Asn60 residue, located in the immediate vicinity of the last cysteine residue that is conserved in all CTs, plays an important structural role. The side chain of this residue participates in three hydrogen bonds in the hydrophobic core of the molecule.

CT loops are formed by the antiparallel strands of the β -sheets (*Fig. 2*). The sizes of the β -sheets, small (formed by two strands of loop I) and large (formed by both strands of loops II and III), and their twist are sim-

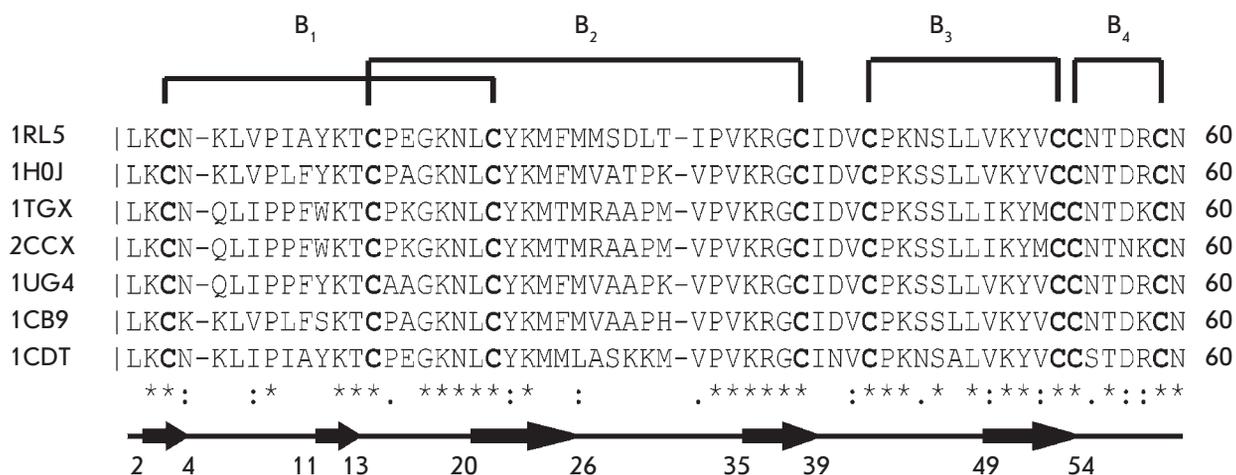


Fig. 1. Alignment of the amino acid sequences (cysteine residues are shown in bold) of CTs whose spatial structures have been determined by NMR or X-ray analysis (Table). Codes of the corresponding structures in the PDB bank are shown on the left. Sequences of asterisks, colons, and dots indicate the following residues: conservative, close and more distant in terms of properties, respectively. Disulfide bonds (B1 to B4) are designated by brackets above the sequence. The secondary structure is given below: sections of antiparallel chains are indicated by arrows under which the boundaries (residue number) of their formation are marked

ilar in different CTs. Dissimilarities in the structures of different CTs are observed in regions with irregular structures, i.e. at the tips of loops I and II. In Group I CTs (this group includes CTs with two Pro residues at the tip of the loop I), this loop is bent (Fig. 2A) and has a ‘banana-twist’ shape [42]. It is achieved through two Pro residues at the tip of loop I (the first of which (Pro8) is present in the *cis*-configuration) and type VIa rotation (stabilized by hydrogen bonds between 10 HN ... O = C 7 and side chain NH of the Gln5 residue and C=O of residue 7). In Group II CTs (this group includes CTs with one Pro8 residue at the tip of loop I), this loop is more extended (Fig. 2A).

An interesting feature of CTs is the Ω-shape of the loop II tip (Fig. 2A). NMR spectroscopy revealed that this region binds a water molecule with a long lifetime of the bound state [43]. This molecule may participate in up to three hydrogen bonds: with one of the amide protons of loop II and with two carbonyl groups of the polypeptide backbone of this portion of the molecule.

The loop III structure is the most conserved one in all CTs. It starts with residues 40–45 (numbers are given for a CT built of 60 amino acid residues) that form a cross-turn with a right twist (Fig. 2A). It connects the external polypeptide chains of the three-stranded β-sheet. Residues 46–49 at the tip of loop III form a type I β-turn. Residues 49–54 form an antiparallel strand structure with residues 20–26. It is interesting to note that the length of this final strand is strictly identical in all known structures of CTs [12].

The tips of the loops play the crucial role in the interactions of CTs with detergent micelles and lipid membranes [38, 44]. They form a membrane-binding CT motif (Fig. 2A). The degree of hydrophobicity of the residues in this motif can serve as a more subtle basis for CT classification (Table, column HTL) [23] than the previously proposed subdivision of CTs into the S- and P-types [29].

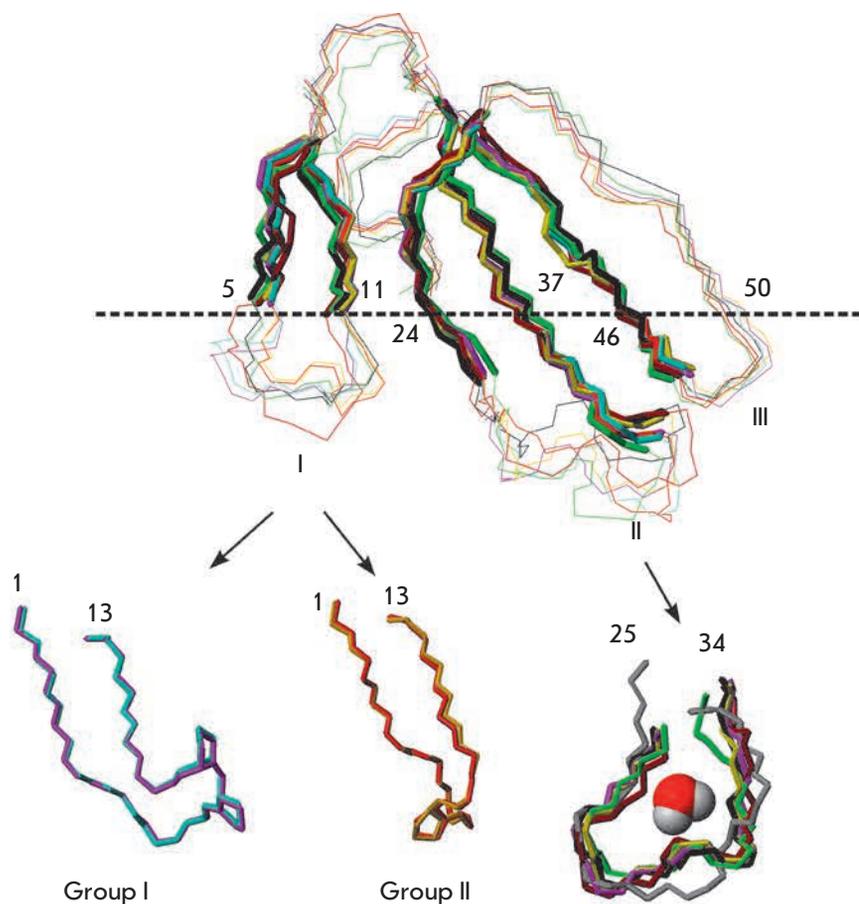
In general, the overall positive charge of a CT molecule ranges from 4 to 12 [23]. The difference is due to variations in the ratio between the negatively charged amino acids residues (aspartic and glutamic acids) and the positively charged lysine and arginine residues (Fig. 2B). The latter mediate the interactions of CTs with the cell surface polyanionic glycopolymers of animal cells, glycosaminoglycans (GAGs) [45]. Charge distribution in a CT molecule defines the corresponding association constant [46].

The biological effect of CTs on various types of cells is mediated by: 1) the interaction with the components of the cell wall (if any) and the plasma membrane; 2) penetration into a cell; and 3) the subsequent interaction with cellular organelles.

ANTIBACTERIAL ACTIVITY OF CTs

The venoms of snakes and insects have long been considered to be a source of various biologically active compounds [47–55], including antibacterial ones. The low incidence of infections in snake bite wounds is a clear indication that the venom includes antibacterial compounds [56]. It was assumed that such activity is

A



B

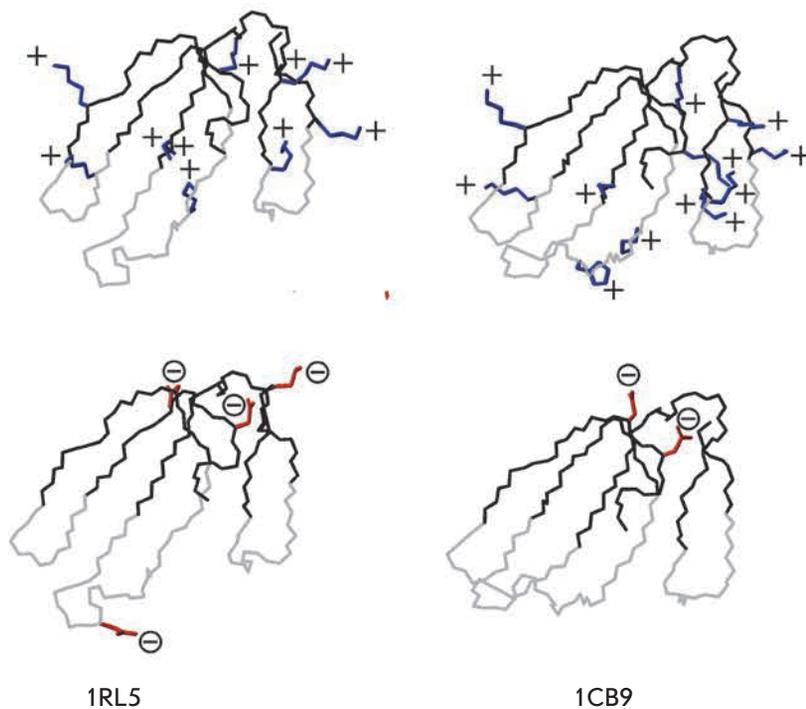


Fig. 2. Detailed spatial structure of CTs. A – superposition of structures of cytotoxins from Fig. 1 (details are also given in Table) based on the elements of their secondary structure (strands of antiparallel β -structure are shown in bold). The loops are numbered with Roman numerals. The dotted line shows the boundary between the membrane and water established by NMR in the model system [38]. The numbers of amino acid residues located at this boundary (according to [38]) are given next to the line. Superposition of residues from the first and the second loops is shown at the bottom of the set (only for the crystal structures from the set). For loop I, the superposition of Group I CTs (characterized by the presence of two Pro residues at the tip of the loop) and Group II CTs (single residue at the tip of the loop Pro) is given. Heavy atoms of the side chain Pro residues are displayed. A water molecule is shown in the center of loop II (formation of its hydrogen bonds: see text for details). B – distribution of positive (Lys, Arg, His residues, *top*, end groups of residues are marked with a "+", in blue) and negative (Asp, Glu, residues, *bottom*, end groups of the side chains of these residues are marked with a "-" in a circle, red) charges, exemplified by the structures of CT2 of *N. oxiana* (*left*) and CT1 of *N. oxiana* (*right*). The set of PDB codes is shown below (representing the structure #1 from the set of 20 deposited structures). Residues of the polypeptide backbone, forming a membrane motif, are shown in gray

required to protect snakes from bacteria of their preys [57].

Studies of whole venoms of several snake species revealed their antibacterial activity [58, 59]. For example, the venoms of some African and Asian cobras (genus *Naja*) and some Australian elapids (*Notechis scutatus*, *Pseudechis australis*) display a very prominent antibacterial effect, especially against *Aeromonas hydrophila* [59]. The venoms of one Asian (*N. oxiana*) and one African (*N. melanoleuca*) cobras are an exception and do not display this activity. Gram-negative bacteria of genus *Escherichia coli* show the highest resistance to the effects of all venoms. Gram-negative *Pseudomonas aeruginosa* and Gram-positive *Bacillus subtilis* are less resistant. Gram-positive cocci *Staphylococcus aureus* and Gram-negative bacteria *A. hydrophila* are the most resistant ones. These data show that the venoms affect both Gram-negative and Gram-positive bacteria. In an earlier study [57], it was suggested that L-amino acid oxidase, a protein with a molecular weight of ~ 140 kDa present in the venom, is responsible for its antibacterial activity. Later, a number of oxidases from the venoms of different snakes were found to have antibacterial activity (e.g., [60, 61]).

Studies of the antibacterial activity of whole insect and snake venoms against relatively stable *E. coli* bacteria demonstrated increasing efficiency in the series of *Crotalus adamanteus* < *Vipera russellii* << *N. naja sputatrix* < *Apis mellifera* (honeybee) [62]. According to electron spectroscopy data, the plasma membrane is the main target. Studies of antibacterial activity in a number of snake, scorpion, and bee venoms against Gram-negative bacterium *Burkholderia pseudomallei* showed that the venoms of snakes *C. adamanteus*, *Daboia russelli russelli*, *Agkistrodon halys*, *P. australis*, *Bungarus candidus*, and *Pseudechis guttatus* display a high activity comparable with that of chloramphenicol and ceftazidime [63]. This remarkably high activity is attributed to the presence of proteins with enzymatic activity, L-amino acid oxidase and phospholipase A2, in the venom. It is believed that the oxidative activity of L-amino acid oxidase produces hydrogen peroxide that kills bacteria. Introduction of hydrogen peroxide interceptors, such as catalase, abolished the antibacterial activity of the enzyme [64]. Phospholipase A2 cleaves phospholipids, causing membrane permeabilization [65].

The first paper on the antibacterial activity of CT was published in 1968 [66]. It has been reported that a CT extracted from the venom of ringhals *Hemachatus haemachatus* (family *Elapidae*) inhibited *S. aureus* at a concentration of 50 µg/mL. The amino acid sequence of CT was not established at the time. It was known only that the protein has a molecular weight of ~ 7 kD and contains four disulfide bonds.

More detailed information on the antibacterial activity of CTs has been obtained later. In particular, it was found that the CT P4 (amino acid sequence unknown) from *N. nigricollis* is active against several Gram-positive bacteria: *B. subtilis*, *Micrococcus flavus*, *Sarcina lutea* [67]. The minimal inhibitory concentrations were in the range of 1.6–6.25 µg/mL. The CT, however, was inactive against Gram-negative bacteria and other microorganisms (yeast, fungi). It can be assumed that CT targets bacterial membranes that contain substantial amounts of anionic phospholipids. Many cytolytic peptides, such as melittin [68] laticins [69, 70], vaprines [71, 72], and cathelicidin [73], can disrupt the membrane at similar concentrations. Another CT, namely CT3 from *N. atra* (otherwise known as A3, *Table*), has been active not only against Gram-positive (*S. aureus*), but also against Gram-negative (*E. coli*) microorganisms [74], even though previous reports indicated no activity for whole *N. atra* venom against *E. coli* [61]. These differences may be explained by the peculiarities of the individual *E. coli* strains that were used in the cited papers. These can only be the differences in the lipopolysaccharide layer of these bacteria (its outer O-antigen portion consisting of branched polysaccharides). Chen *et al.* [74] presented electron microscopy images of bacteria before and after interaction with CT3. It can be seen that the toxin causes characteristic damage to the plasma membrane (protrusions, bubbles and cracks) and, therefore, penetrates through the lipopolysaccharide (LPS) layer. This may occur by substitution of Ca²⁺ ions in the phosphate groups of lipid A via the interaction between the charged side chains of the lysine residues of the toxin and the phosphate groups of lipid A of the LPS, followed by loosening of this layer [75]. An alternative mechanism of cell penetration for antimicrobial peptides through LPS is self-promoted uptake, typical of linear (containing no disulfide bonds) AMPs, such as cecropins [76, 77]. Binding of these peptides to LPS facilitates their penetration into the plasma membrane and increases their membrane-permeabilizing ability. In the case of CT3, some of the molecules remain bound to the LPS, providing other ones with a passage into the plasma membrane. This was demonstrated in experiments with a fluorescent dye leakage from the liposomes formed by phospholipids, whose composition corresponded to that of the phospholipids of the plasma membrane of the bacteria under study [78]. Pre-incubation of CT3 with LPS reduced dye leakage. Thus, the cell wall of Gram-negative bacteria is the major obstacle to the penetration of CTs into the plasma membrane. The high proportion of anionic phospholipids in the plasma membrane enables its destruction by CT molecules. Since the plasma mem-

brane is associated with important cellular functions such as respiration, transportation, osmoregulation, lipid synthesis and others, the loss of its integrity results in cell death [74, 78].

The interaction of CT3 with the cell walls of Gram-positive bacteria (mainly with lipoteichoic acid (LTA), which has no polysaccharide moiety) was also examined by Chen *et al.* [74]. Pre-incubation of CT3 with LTA reduced the dye leakage from the liposomes formed from anionic phospholipids (phosphatidylglycerol (PG) : cardiolipin (CL), 6:4) mimicking the plasma membrane of *B. subtilis*. The effective concentration of CT3 (the concentration causing the death of 50% of bacteria) is approximately an order of magnitude smaller (~ 0.9 μM) than that against *E. coli*. This fact most likely indicates that most of the CT3 molecules are not bound to the plasma membranes of these bacteria and are present in the aqueous solution and/or on the outer membrane (LPS), which is a major obstacle for CT molecules. Therefore, CT molecules are too large and conformationally rigid to penetrate this barrier.

As discussed above, the antibacterial action of CTs may be attributed to their membrane activity. To elucidate the mechanism of the destructive action of CT on membranes, Cao *et al.* [79] analyzed the interaction of toxin CT3 from *N. atra* and toxin gamma from *N. nigricollis* with model membranes of *E. coli* (phosphatidylethanolamine (PE)/PG, 75/25 mol/mol) and *S. aureus* (PG/CL, 60/40 mol/mol). Toxin gamma was equally effective in destroying both PE/PG and PG/CL vesicles. However, CT3 was more effective against PG/CL vesicles. The fusogenic activity of the toxins correlated with their ability to disrupt membranes. For example, CT3, in contrast to toxin gamma, induced a more pronounced membrane fusion with an increase in the cardiolipin content. These data demonstrate that the fusogenic and antibacterial activities of CTs are related.

The attempts to use CT amino acid sequences to design antimicrobial agents that would be smaller but more active than their parent peptides deserve special mention. It was reported earlier that 7- to 12-residue long peptides from loop I of the CT4 of *N. mossambica* display *in vivo* toxicity, albeit lesser than that of the original toxin [80]. The 14-membered cyclic peptide (with one disulfide bond) L1AD3 has the amino acid sequence of loop I of the CT3 of *N. atra* and can induce apoptosis in leukemic T-cells [81, 82] when used at micromolar concentrations. The peptide has a β-hairpin

conformation in aqueous solutions, similar to the corresponding moiety within the original CT. Although these short analogs have not been reported to exhibit antibacterial activity, it can be assumed that the β-structural analog possesses this activity. There are several β-structural antimicrobial peptides with one disulfide bond that display a broad range of activities (e.g., [83]). The compact size of these cationic peptides allows them to penetrate through the LPS of Gram-negative bacteria and to destabilize the plasma membrane because of their favorable charge/hydrophobicity ratio. We believe that the emergence of interest in the antibacterial activity of CTs will soon be followed by the development of antimicrobial peptides based on their amino acid sequences.

Notably, the application of computer methods of analysis have allowed researchers to elucidate the evolutionary relationship between animal venom toxins and antimicrobial proteins [84]. It is likely that animal toxins retain their antibacterial function during evolution.

CONCLUSIONS

The antibacterial activity of CTs varies widely between different members of this family of peptides. The data presented in this review clearly show that penetration through the peptidoglycan layer, bacterial lipopolysaccharide, plays the crucial role in the manifestation of peptide activity. This has been confirmed in a recent study of the comparative activity of five different CTs against some Gram-positive and Gram-negative bacteria [85], which demonstrated that the activity is determined by amino acid residues outside of the CT membrane-binding motif. It might be easier to understand the rules governing the interaction of CTs with the polymers forming the outer membrane and the peptidoglycan layer of a bacterial cell for the peptides whose spatial structure depends on the abundance of disulfide bonds than for mobile linear peptides that have no disulfide bonds. We believe that the next step will be designing peptides based on CTs amino acid sequences. Some steps have already been taken in this direction, and one of the peptides, L1AD3 [81, 82], can be used to treat leukemia. It is likely that there will be more examples in the future.

This work was supported by the Russian Foundation for Basic Research (grant 13-04-02128).

REFERENCES

1. Kini R.M., Evans H.J. // *Int. J. Pept. Prot. Res.* 1989. V. 34. P. 277–286.
2. Landon C., Meudal H., Boulanger N., Bulet P., Vovelle F. // *Biopolymers.* 2006. V. 81. P. 92–103.
3. Dubovskii P.V., Volynsky P.E., Polyansky A.A., Chupin V.V., Efremov R.G., Arseniev A.S. // *Biochemistry.* 2006. V. 45. P. 10759–10767.
4. Bhattacharjya S., Domadia P.N., Bhunia A., Malladi S., David S.A. // *Biochemistry.* 2007. V. 46. P. 5864–5874.

REVIEWS

5. Dubovskii P.V., Volynsky P.E., Polyansky A.A., Karpunin D.V., Chupin V.V., Efremov R.G., Arseniev A.S. // *Biochemistry*. 2008. V. 47. P. 3525–3533.
6. Abbassi F., Lequin O., Piesse C., Goasdoué N., Foulon T., Nicolas P., Ladram A. // *J. Biol. Chem.* 2010. V. 285. P. 16880–16892.
7. Legrand B., Laurencin M., Sarkis J., Duval E., Moret L., Hubert J.F., Cohen M., Vie V., Zatyiny-Gaudin C., Henry J., Baudi-Floc'h M., Bondon A. // *Biochim. Biophys. Acta.* 2011. V. 1808. P. 106–116.
8. Wang G., Li X., Wang Z. // *Nucleic Acids Res.* 2009. V. 37. D933–937.
9. Bauer F., Schweimer K., Klüber E., Conejo-García J.R., Forssmann W.G., Rosch P., Adermann K., Sticht H. // *Protein Sci.* 2001. V. 10. P. 2470–2479.
10. Powers J.P.S., Rozek A., Hancock R.E.W. // *Biochim. Biophys. Acta.* 2004. V. 1698. P. 239–250.
11. Sayyed-Ahmad A., Kaznessis Y.N. // *PloS One.* 2009. V. 4. № 3. e4799.
12. Konshina A.G., Dubovskii P.V., Efremov R.G. // *Curr. Protein Pept. Sci.* 2012. V. 13. P. 570–584.
13. Yount N.Y., Kupferwasser D., Spisni A., Dutz S.M., Ramjan Z.H., Sharma S., Waring A.J., Yeaman M.R. // *Proc. Natl. Acad. Sci. USA.* 2009. V. 106. P. 14972–14977.
14. Dubovskii P.V., Vassilevski A.A., Samsonova O.V., Egorova N.S., Kozlov S.A., Feofanov A.V., Arseniev A.S., Grishin E.V. // *FEBS J.* 2011. V. 278. P. 4382–4393.
15. Mader J.S., Hoskin D.W. // *Expert Opin. Investig. Drugs.* 2006. V. 15. P. 933–946.
16. Hoskin D.W., Ramamoorthy A. // *Biochim. Biophys. Acta.* 2008. V. 1778. P. 357–375.
17. Schweizer F. // *Eur. J. Pharmacol.* 2009. V. 625. P. 190–194.
18. Riedl S., Zweytick D., Lohner K. // *Chem. Phys. Lipids.* 2011. V. 164. P. 766–781.
19. Fazio M.A., Jouvencal L., Vovelle F., Bulet P., Miranda M.T., Daffre S., Miranda A. // *Biopolymers.* 2007. V. 88. P. 386–400.
20. Haney E.F., Hunter H.N., Matsuzaki K., Vogel H.J. // *Biochim. Biophys. Acta.* 2009. V. 1788. P. 1639–1655.
21. Feliu L., Oliveras G., Cirac A.D., Besalu E., Roses C., Colomer R., Bardaji E., Pianas M., Puig T. // *Peptides.* 2010. V. 31. P. 2017–2026.
22. Fadnes B., Uhlin-Hansen L., Lindin I., Rekdal O. // *BMC Cancer.* 2011. V. 11. P. 116.
23. Dubovskii P.V., Konshina A.G., Efremov R.G. // *Curr. Med. Chem.* 2014. V. 21. P. 270–287.
24. Nguyen L.T., Chau J.K., Perry N.A., de Boer L., Zaat S.A., Vogel H.J. // *PloS One.* 2010. V. 5. e12684.
25. Verardi R., Traaseth N.J., Shi L., Porcelli F., Monfregola L., De Luca S., Amodeo P., Veglia G., Scaloni A. // *Biochim. Biophys. Acta.* 2011. V. 1808. P. 34–40.
26. Dufton M.J., Hider R.C. // *Pharmacol. Ther.* 1988. V. 36. P. 1–40.
27. Kumar T.K., Jayaraman G., Lee C.S., Arunkumar A.I., Sivaraman T., Samuel D., Yu C. // *J. Biomol. Struct. Dyn.* 1997. V. 15. P. 431–463.
28. Kini R.M., Doley R. // *Toxicol.* 2010. V. 56. P. 855–867.
29. Chien K.Y., Chiang C.M., Hseu Y.C., Vyas A.A., Rule G.S., Wu W. // *J. Biol. Chem.* 1994. V. 269. P. 14473–14483.
30. Dubovskii P.V., Lesovoy D.M., Dubinnyi M.A., Konshina A.G., Utkin Y.N., Efremov R.G., Arseniev A.S. // *Biochem. J.* 2005. V. 387. P. 807–815.
31. Dubinnyi M.A., Lesovoy D.M., Dubovskii P.V., Chupin V.V., Arseniev A.S. // *Solid State Nucl. Magn. Reson.* 2006. V. 29. P. 305–311.
32. Feofanov A.V., Sharonov G.V., Dubinnyi M.A., Astapova M.V., Kudelina I.A., Dubovskii P.V., Rodionov D.I., Utkin Y.N., Arseniev A.S. // *Biochemistry (Mosc.).* 2004. V. 69. P. 1148–1157.
33. Feofanov A.V., Sharonov G.V., Astapova M.V., Rodionov D.I., Utkin Y.N., Arseniev A.S. // *Biochem. J.* 2005. V. 390. P. 11–18.
34. Wang C.H., Wu W.G. // *FEBS Lett.* 2005. V. 579. P. 3169–3174.
35. Wu M., Ming W., Tang Y., Zhou S., Kong T., Dong W. // *Am. J. Chin. Med.* 2013. V. 41. P. 643–663.
36. Wu P.L., Chiu C.R., Huang W.N., Wu W.G. // *Biochim. Biophys. Acta.* 2012. V. 1818. P. 1378–1385.
37. Epand R.M., Rotem S., Mor A., Berno B., Epand R.F. // *J. Am. Chem. Soc.* 2008. V. 130. P. 14346–14352.
38. Dubovskii P.V., Dementieva D.V., Bocharov E.V., Utkin Y.N., Arseniev A.S. // *J. Mol. Biol.* 2001. V. 305. P. 137–149.
39. Dubovskii P.V., Lesovoy D.M., Dubinnyi M.A., Utkin Y.N., Arseniev A.S. // *Eur. J. Biochem.* 2003. V. 270. P. 2038–2046.
40. Tjong S.C., Chen T.S., Huang W.N., Wu W.G. // *Biochemistry.* 2007. V. 46. P. 9941–9952.
41. Galat A., Gross G., Drevet P., Sato A., Menez A. // *FEBS J.* 2008. V. 275. P. 3207–3225.
42. Chen T.S., Chung F.Y., Tjong S.C., Goh K.S., Huang W.N., Chien K.Y., Wu P.L., Lin H.C., Chen C.J., Wu W.G. // *Biochemistry.* 2005. V. 44. P. 7414–7426.
43. Dementieva D.V., Bocharov E.V., Arseniev A.S. // *Eur. J. Biochem.* 1999. V. 263. P. 152–162.
44. Dauplais M., Neumann J.M., Pinkasfeld S., Menez A., Roumestand C. // *Eur. J. Biochem.* 1995. V. 230. P. 213–220.
45. Vyas A.A., Pan J.J., Patel H.V., Vyas K.A., Chiang C.M., Sheu Y.C., Hwang J.K., Wu W.G. // *J. Biol. Chem.* 1997. V. 272. P. 9661–9670.
46. Vyas K.A., Patel H.V., Vyas A.A., Wu W.G. // *Biochemistry.* 1998. V. 37. P. 4527–4534.
47. Pal S.K., Gomes A., Dasgupta S.C., Gomes A. // *Indian J. Exp. Biol.* 2002. V. 40. P. 1353–1358.
48. de Lima D.C., Abreu A.P., de Freitas C.C., Santos D.O., Borges R.O., Dos Santos T.C., Cabral M.R., Rodrigues C.R., Castro H.C. // *Evid. Based Complement. Alternat. Med.* 2005. V. 2. P. 39–47.
49. Koh D.C., Armugam A., Jeyaseelan K. // *Cell. Mol. Life Sci.* 2006. V. 63. P. 3030–3041.
50. Samy P.R., Gopalakrishnakone P., Thwin M.M., Chow T.K., Bow H., Yap E.H., Thong T.W. // *J. Appl. Microbiol.* 2007. V. 102. P. 650–659.
51. Gomes A., Bhattacharjee P., Mishra R., Biswas A.K., Dasgupta S.C., Giri B. // *Indian J. Exp. Biol.* 2010. V. 48. P. 93–103.
52. Kapoor V.K. // *Indian J. Exp. Biol.* 2010. V. 48. P. 228–237.
53. King G.F. // *Expert Opin. Biol. Ther.* 2011. V. 11. P. 1469–1484.
54. Lazarev V.N., Govorun V.M. // *Appl. Biochem. Microbiol.* 2010. V. 46. P. 803–814.
55. Koh C.Y., Kini R.M. // *Toxicol.* 2012. V. 59. № 4. P. 497–506.
56. Talan D.A., Citron D.M., Overturf G.D., Singer B., Froman P., Goldstein E.J. // *J. Infect. Dis.* 1991. V. 164. P. 195–198.
57. Thomas R.G., Pough F.H. // *Toxicol.* 1979. V. 17. P. 221–228.
58. Glaser H.S.R. // *Copeia.* 1948. P. 245–247.
59. Stiles B.G., Sexton F.W., Weinstein S.A. // *Toxicol.* 1991. V. 29. P. 1129–1141.
60. Okubo B.M., Silva O.N., Migliolo L., Gomes D.G., Porto W.F., Batista C.L., Ramos C.S., Holanda H.H., Dias S.C.,

REVIEWS

- Franco O.L., Moreno S.E. // *PLoS One*. 2012. V. 7. e33639.
61. Rima M., Accary C., Haddad K., Sadek R., Hraoui-Bloquet S., Desfontis J.C., Fajloun Z. // *Infect. Disord. Drug. Targets*. 2013. V. 13. P. 337–343.
62. Stocker J.F., Traynor J.R. // *J. Appl. Microbiol.* 1986. V. 61. P. 383–388.
63. Samy P.R., Pachiappan A., Gopalakrishnakone P., Thwin M.M., Hian Y.E., Chow V.T., Bow H., Weng J.T. // *BMC Infect. Dis.* 2006. V. 6. P. 100
64. Lee M.L., Tan, N.H., Fung S.Y., Sekaran S.D. // *Comp. Biochem. Physiol. C. Toxicol. Pharmacol.* 2011. V. 153. P. 237–242.
65. Samy P.R., Gopalakrishnakone P., Ho B., Chow V.T. // *Biochimie*. 2008. V. 90. P. 1372–1388.
66. Aloof-Hirsch S., de Vries A., Berger A. // *Biochim. Biophys. Acta*. 1968. V. 154. P. 53–60.
67. Mollmann U., Gutsche W., Maltz L., Ovadia M. // *Toxicon*. 1997. V. 35. P. 487–487.
68. Raghuraman H., Chattopadhyay A. // *Biosci. Rep.* 2007. V. 27. P. 189–223.
69. Kozlov S.A., Vassilevski A.A., Feofanov A.V., Surovov A.Y., Karpunin D.V., Grishin E.V. // *J. Biol. Chem.* 2006. V. 281. P. 20983–20992.
70. Vassilevski A.A., Kozlov S.A., Zhmak M.N., Kudelina I.A., Dubovskii P.V., Shatursky O.Y., Arseniev A.S., Grishin E.V. // *Rus. J. Bioorgan. Chem.* 2007. V. 33. P. 376–382.
71. Torres A.M., Wong H.Y., Desai M., Mochhala S., Kuchel P.W., Kini R.M. // *J. Biol. Chem.* 2003. V. 278. P. 40097–40104.
72. Nair D.G., Fry B.G., Alewood P., Kumar P.P., Kini R.M. // *Biochem. J.* 2007. V. 402. P. 93–104.
73. Wang Y., Hong J., Liu X., Yang H., Liu R., Wu J. Wang A., Lin D., Lai R. // *PLoS One*. 2008. V. 3. e3217.
74. Chen L.W., Kao P.H., Fu Y.S., Lin S.R., Chang L.S. // *Toxicon*. 2011. V. 58. № 1. P. 46–53.
75. Bhunia A., Domadia P.N., Torres J., Hallock K.J., Ramamoorthy A., Bhattacharjya S. // *J. Biol. Chem.* 2010. V. 285. P. 3883–3895.
76. Piers K.L., Brown M.H., Hancock R.E. // *Antimicrob. Agents Chemother.* 1994. V. 38. P. 2311–2316.
77. Arcidiacono S., Soares J.W., Meehan A.M., Marek P., Kirby R. // *J. Pept. Sci.* 2009. V. 15. P. 398–403.
78. Chen L.W., Kao P.H., Fu Y.S., Hu W.P., Chang L.S. // *Peptides*. 2011. V. 32. № 8. P. 1755–1763.
79. Kao P.H., Lin S.R., Hu W.P., Chang L.S. // *Toxicon*. 2012. V. 60. P. 367–377.
80. Marchot P., Bougis P.E., Ceard B., van Rietschoten J., Rochat H. // *Biochem. Biophys. Res. Commun.* 1988. V. 153. P. 642–647.
81. Smith C.A., Hinman C.L. // *J. Biochem. Mol. Toxicol.* 2004. V. 18. P. 204–220.
82. Smith C.A., Hinman C.L. // *Arch. Biochem. Biophys.* 2004. V. 432. P. 88–101.
83. Shenkarev Z.O., Balandin S.V., Trunov K.I., Paramonov A.S., Sukhanov S.V., Barsukov L.I., Arseniev A.S., Ovchinnikova T.V. // *Biochemistry*. 2011. V. 50. P. 6255–6265.
84. Kaplan N., Morpurgo N., Linial M. // *J. Mol. Biol.* 2007. V. 369. P. 553–566.
85. Dubovskii P.V., Vorontsova O.V., Utkin Y.N., Arseniev A.S., Efremov R.G., Feofanov A.V. // *Mendeleev Communications*. 2015. in press.

TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery

A. A. Nemudryi^{1,2,3,§}, K. R. Valetdinova^{1,2,3,4,§}, S. P. Medvedev^{1,2,3,4}, S. M. Zakian^{1,2,3,4*}

¹Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, Lavrentyev Prosp., 10, Novosibirsk, Russia, 630090

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

³Meshalkin Novosibirsk State Research Institute of Circulation Pathology, Ministry of Health of the Russian Federation, Rechkunovskaya Str., 15, Novosibirsk, Russia, 630055

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

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

§These authors contributed equally to the writing of this review.

Received 03.03.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT Precise studies of plant, animal and human genomes enable remarkable opportunities of obtained data application in biotechnology and medicine. However, knowing nucleotide sequences isn't enough for understanding of particular genomic elements functional relationship and their role in phenotype formation and disease pathogenesis. In post-genomic era methods allowing genomic DNA sequences manipulation, visualization and regulation of gene expression are rapidly evolving. Though, there are few methods, that meet high standards of efficiency, safety and accessibility for a wide range of researchers. In 2011 and 2013 novel methods of genome editing appeared – this are TALEN (Transcription Activator-Like Effector Nucleases) and CRISPR (Clustered Regulatory Interspaced Short Palindromic Repeats)/Cas9 systems. Although TALEN and CRISPR/Cas9 appeared recently, these systems have proved to be effective and reliable tools for genome engineering. Here we generally review application of these systems for genome editing in conventional model objects of current biology, functional genome screening, cell-based human hereditary disease modeling, epigenome studies and visualization of cellular processes. Additionally, we review general strategies for designing TALEN and CRISPR/Cas9 and analyzing their activity. We also discuss some obstacles researcher can face using these genome editing tools.

KEYWORDS TALEN, CRISPR/Cas9, genome editing.

ABBREVIATIONS TALENs – transcription activator-like effector nucleases; CRISPR – clustered regulatory interspaced short palindromic repeats/Cas9; PAM – protospacer adjacent motif; sgRNA – single guide RNA; crRNA – CRISPR RNA; tcacrRNA – trans-activating CRISPR RNA; SpCas9 – *Streptococcus pyogenes* Cas9; pre-crRNA – poly-spacer precursor crRNA.

INTRODUCTION

Genetic engineering emerged in the laboratory of Paul Berg in 1972 in the form of a recombinant DNA technology, when scientists combined the *E. coli* genome with the genes of a bacteriophage and the SV40 virus. Since then, this science has achieved tremendous success; the molecular genetic mechanisms and phenomena that can now be reproduced *in vitro* have been discovered and studied in detail. Studies in the field of molecular genetics and biochemistry of bacteria and viruses have allowed the development of methods to manipulate DNA, generate various vector systems and methods for their delivery to the cell. All of this has enabled not only transgenic microorganisms production, but also genetically modified plants and animals. The

application area of genetic engineering has experienced rapid development, which provided the impetus for progress in selection and biotechnology. However, the conventional genetic engineering strategy has several drawbacks and limitations, one of which is the complexity of manipulations with large animal and human genomes.

From 1990 to 2003, the nucleotide sequence of human nuclear DNA was determined and about 20.5 thousand genes were identified within the “Human Genome” International Project. Similar projects are also currently under implementation; the genome nucleotide sequences of the main model biological objects (*E. coli*, nematode, drosophila, mouse, and others) have been deciphered. However, these projects provide data

on the DNA nucleotide sequence only, but they yield no information about function of individual genome elements, or how they interrelate in an entire system. Understanding the functional relationships in the human genome will make it possible not only to identify the cause-and-effect relations in the pathology of hereditary as well as multifactorial diseases, but also to find targets for their treatment.

In 2003, the U.S. National Human Genome Research Institute launched a new international project, ENCODE (Encyclopedia Of DNA Elements), which aim was to join the efforts of scientists and obtain a complete list of the functional elements of the human genome, including the elements that act at the protein and RNA level, as well as the regulatory elements that control the fundamental genetic processes (transcription, translation, and replication). To establish these functional relationships, two strategies are used: switching off a gene (knockout or knockdown) and enhancing the gene activity or its ectopic expression. The traditional methods – transgenesis using homologous recombination in mice [1] and also the use of various viral, including lentiviral vectors – are not only expensive but also quite labor-intensive; they do not allow one to introduce precise changes into a strictly defined genome locus.

Currently, researchers have several tools that allow them to solve the problems of precise plant's, animal's, and human's genome editing.

As early as 1996, a zinc finger protein domain coupled with the FokI endonuclease domain was demonstrated for the first time to act as a site-specific nuclease cutting DNA at strictly defined sites *in vitro* [2]. This chimeric protein has a modular structure, because each zinc finger domain recognizes one nucleotide triplet (zinc finger nuclease, ZFN). This method became the basis for editing cultured cells, including pluripotent stem cells, plant and animal models [3–8]. However, the ZFN-based technology has a number of disadvantages, including the complexity and high cost of protein domains construction for each particular genome locus and the probability of inaccurate cleavage of target DNA due to single nucleotide substitutions or inappropriate interaction between domains. Therefore, an active search for new methods for genome editing was continued. In recent years, this search has led to the development of new tools for genome editing: TALENs (transcription activator-like effector nucleases) and CRISPR/Cas (clustered regulatory interspaced short palindromic repeats). These systems are characterized by a relative construction simplicity and a high functional efficiency in human, animal, and plant cells. These systems, which are extensively used for various genome manipulations, allow one to solve complex

problems, including the mutant and transgenic plants and animals generation, development and investigation of disease models based on cultured human pluripotent cells. Furthermore, chimeric proteins based on the TALE and inactivated Cas9 nuclease DNA-binding domains were used in experiments on the regulation of gene transcription and for studying the epigenomes and behavior of chromosomal loci in the cell cycle.

This review describes in detail the possibilities in the construction, implementation, and analysis of the TALEN and CRISPR/Cas9 function using examples of various model systems, as well as the complexities and problems associated with the use of these genome editing tools.

NATURAL BACTERIAL TALE AND CRISPR/CAS SYSTEMS AS THE BASIS FOR THE DEVELOPMENT OF NEW TOOLS FOR EUKARYOTIC GENOME EDITING

TALEN

In 2011, *Nature Methods* named the methods of precise genome editing, including the TALEN system, method of the year [9]. The history of this system's development is associated with the study of bacteria of the *Xanthomonas* genus. These bacteria are pathogens of crop plants, such as rice, pepper, and tomato; and they cause significant economic damage to agriculture, which was the motivate for their thorough study. The bacteria were found to secrete effector proteins (transcription activator-like effectors, TALEs) to the cytoplasm of plant cells, which affect processes in the plant cell and increase its susceptibility to the pathogen. Further investigation of the effector protein action mechanisms revealed that they are capable of DNA binding and activating the expression of their target genes via mimicking the eukaryotic transcription factors.

TALE proteins are composed of a central domain responsible for DNA binding, a nuclear localization signal, and a domain that activates the target gene transcription [10]. The capability of these proteins to bind to DNA was first described in 2007 [11], and just a year later two groups of researchers deciphered the code for recognition of the target DNA by TALE proteins [12, 13]. The DNA-binding domain was demonstrated to consist of monomers, each of them binds one nucleotide in the target nucleotide sequence. Monomers are tandem repeats of 34 amino acid residues, two of which are located at positions 12 and 13 and are highly variable (repeat variable diresidue, RVD), and it is they that are responsible for the recognition of a specific nucleotide. This code is degenerate; some RVDs can bind to several nucleotides with different efficiencies. Before the 5'-end of a sequence bound by a TALE monomer, the target DNA molecule always contains the same nucleo-

tide, thymidine, that affects the binding efficiency [14]. The last tandem repeat that binds a nucleotide at the 3'-end of the recognition site consists only of 20 amino acid residues and therefore is called a half-repeat.

After deciphering the code of DNA recognition by TALE proteins, which attracted the attention of researchers across the world due to its simplicity (one monomer – one nucleotide), the first studies on the construction of chimeric TALEN nucleases were launched. For that purpose, the sequence encoding the DNA-binding domain of TALE was inserted into a plasmid vector previously used for creating ZFN [15]. This resulted in the generation of genetic constructs expressing artificial chimeric nucleases that contain the DNA-binding domain and the catalytic domain of restriction endonuclease FokI. This system allow ones, by combining monomers of the DNA-binding domain with different RVDs, to construct artificial nucleases, the target of which can be any nucleotide sequence. Most studies use monomers containing RVDs such as Asn and Ile (NI), Asn and Gly (NG), two Asn (NN), and His and Asp (HD) for binding the nucleotides A, T, G, and C, respectively. Since the NN RVD can bind both G and A, a number of studies was performed to find monomers that will be more specific. It has been shown that the use of NH or NK monomers for more specific binding of guanine reduces the risk of off-target effects [16, 17]. The first amino acid residue in the RVD (H and N) was found not to be directly involved in the binding

of a nucleotide, but to be responsible for stabilizing the spatial conformation. The second amino acid residue interacts with a nucleotide, with the nature of this interaction being different: D and N form hydrogen bonds with nitrogenous bases, and I and G bind target nucleotides through van der Waals forces [18].

An artificial DNA-binding domain is inserted into a genetic construct comprising a nuclear localization signal, half-repeat, N-terminal domain, and the FokI catalytic domain. TALENs work as pairs and their bindings sites are chosen so that they are located on opposite DNA strands and are separated by a small fragment (12–25 bp), a spacer sequence. Once in the nucleus, artificial nucleases bind to target sites: the FokI domains located at the C-termini of a chimeric protein dimerize to cause a double-strand break in a spacer sequence (Fig. 1).

In theory, a double-strand break can be introduced in any region of the genome with known recognition sites of the DNA-binding domains using artificial TALEN nucleases. The only limitation to the selection of TALEN nuclease sites is the need for T before the 5'-end of the target sequence. However, site selection may be made in most cases by varying the spacer sequence length. The W232 residue in the N-terminal region of the DNA-binding domain was demonstrated to interact with 5'-T, affecting the efficiency of TALEN binding to the target site [19]. However, this limitation can also be overcome by selection of mutant variants of the

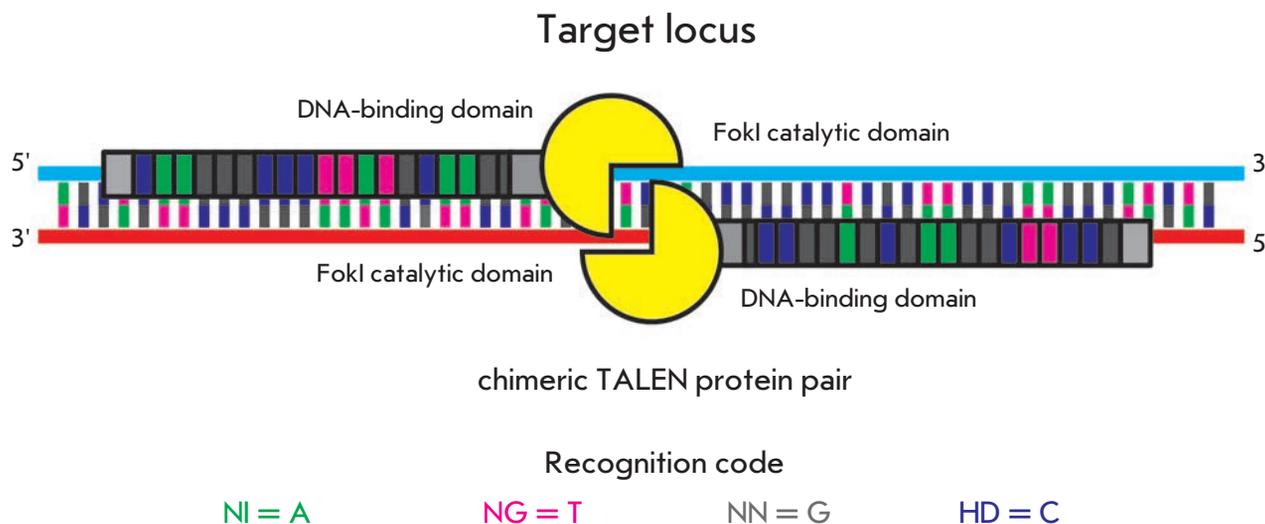


Fig. 1. A scheme for introducing a double-strand break using chimeric TALEN proteins. One monomer of the DNA-binding protein domain recognizes one nucleotide of a target DNA sequence. Two amino acid residues in the monomer are responsible for binding. The recognition code (single-letter notation is used to designate amino acid residues) is provided. Recognition sites are located on the opposite DNA strands at a distance sufficient for dimerization of the FokI catalytic domains. Dimerized FokI introduces a double-strand break into DNA

TALEN N-terminal domain that are capable of binding to A, G, or C [14].

CRISPR/Cas

About two years later after the discovery of the chimeric TALEN proteins, another genome editing system, CRISPR, elements of which are non-coding RNAs and Cas proteins (CRISPR associated), was developed and started to be extensively used. In contrast to the chimeric TALEN proteins, recognition by the CRISPR/Cas system is carried out via the complementary interaction between a non-coding RNA and the target site DNA. In this case, a complex of non-coding RNA and Cas proteins, which have nuclease activity, is formed. As early as 1987, mysterious repeats were discovered in some bacterial genes [20], the functions of which remained unknown for nearly 20 years. Sequencing of bacterial genomes revealed similar nucleotide sequences in the genome of many microorganisms that have the characteristic structure: short regions of the unique DNA (spacers) are separated from each other by short palindromic repeats (Fig. 2). Due to this feature, they received the name CRISPR (see abbreviations). Furthermore, these CRISPR cassettes are located in close proximity to the *cas* genes (CRISPR associated), the protein products of which have helicase and nuclease activity [21]. In 2005, three independent groups of bioinformaticians reported that the spacer DNA is often homologous to the DNA of many phages and plasmids [22–24]. Furthermore, in 2007, it was shown that *Streptococcus thermophilus* cells bearing in the CRISPR locus a spacer that is complementary to a bacteriophage genomic DNA fragment become resistant to the phage [25]. Thus, it became apparent that the CRISPR/Cas system is the unique mechanism providing microorganisms protection against foreign DNA penetration and acting along with the restriction-modification system as a limiter of the horizontal transfer of genetic information.

The CRISPR systems are widespread in prokaryotes: they are found in 87% of archaea and 48% of eubacteria [26]. This is why different species are widely varied both in the number of CRISPR cassettes in the genome (1–18) and in the number (60, on average) and size of repeats (23–37 bp, on average), as well as in the number and size of spacers (17–84 bp). Yet, the length of spacers and repeats in one cassette is constant and repeat sequences are almost identical [27].

The protection mechanism includes three main stages (Fig. 2). At the first stage, adaptation, a small fragment of foreign DNA that entered a bacterial cell is inserted into the CRISPR locus of the host genome, forming a new spacer. In the viral genome, this fragment is present as a protospacer that is complemen-

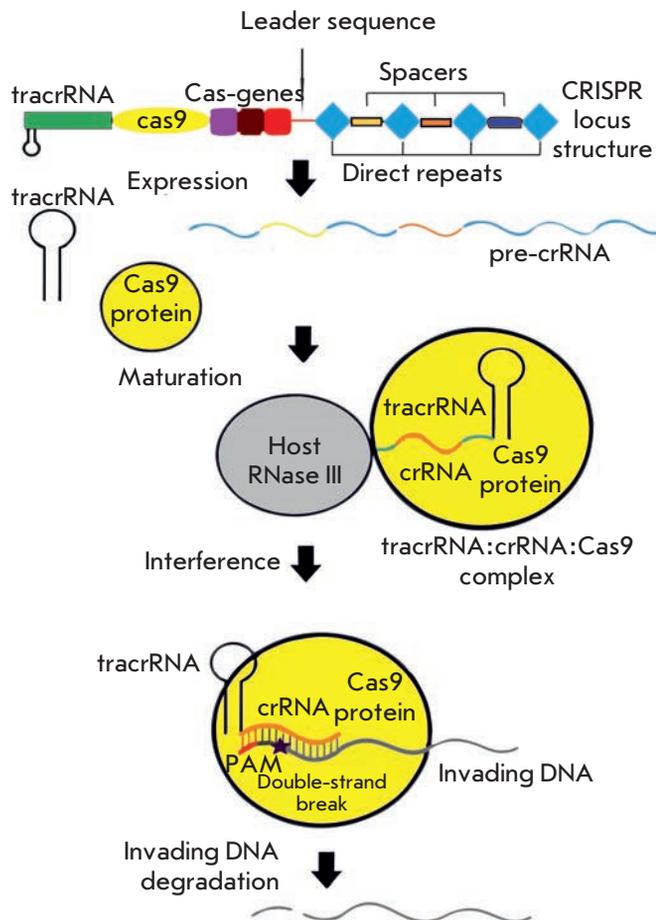


Fig. 2. A mechanism of CRISPR/Cas9 action in bacterial cells (see the text for details)

tary to the spacer and flanked by a short (2–5 bp), conserved sequence called PAM (protospacer adjacent motif) [28, 29]. The new spacer is always inserted on the AT-rich side of the leader sequence located before a CRISPR cassette that also contains promoter elements and landing sites for regulatory proteins [30, 31]. Apparently, this is the way the targets of most of the CRISPR/Cas systems are formed.

At the second stage, transcription, the entire CRISPR locus is transcribed into a long pre-crRNA (poly-spacer precursor crRNA) (Fig. 2). The processing of an immature transcript into mature crRNA in most of the CRISPR/Cas systems is implemented by Cas6 endonuclease [32–36]. Short crRNAs (CRISPR RNA) of 39–45 nucleotides contain one spacer sequence, and their ends contain repeats involved in the formation of the stem loop structure: the last eight nucleotides of the repeat with a hydroxyl group at the 5'-end form the stem, and the hairpin structure with 2', 3'-cyclic phosphate forms the loop at the 3'-end [37, 38].

Bioinformatic analysis

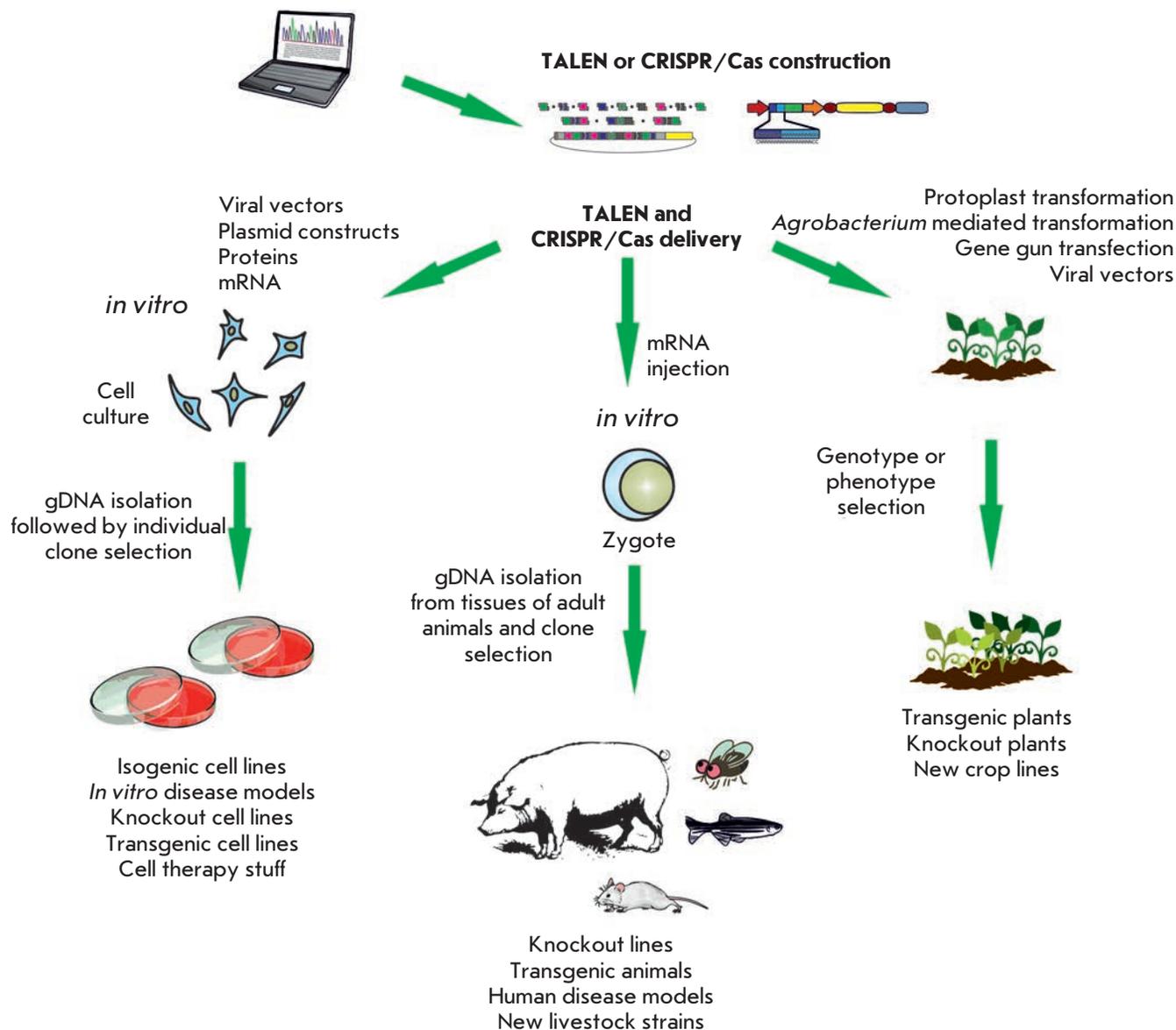


Fig. 3. A general scheme of the strategy for using the TALEN and CRISPR/Cas systems in genomic engineering

The third stage, the interference of foreign DNA or RNA, is provided by the interaction between crRNA and a complex of Cas proteins; crRNA recognizes complementarily the protospacer sequence, and Cas proteins provide its degradation (Fig. 2).

For target DNA degradation by the effector complex, any interaction between the complementary nucleotides of crRNA and target DNA at positions -2, -3, and -4 (if the first protospacer base is taken as +1) should be avoided [39]. Apparently, complementary interactions between crRNA and the target DNA at these positions disrupt the effector complex formation, which

prevents cleavage of genomic DNA and its subsequent degradation.

Long-term co-evolution of viruses and their hosts has led to the formation of viral protection mechanisms against the CRISPR interference [40], which explains a wide variety of the CRISPR/Cas systems in bacteria and archaea. Bioinformatic studies subdivide all CRISPR/Cas systems into three main types (I–III) and, at least, 10 subtypes [21, 27, 41]. Among these, the type II-A CRISPR/Cas system isolated from the *S. pyogenes* pathogen is currently the one used most widely in genomic engineering. A minimum set of the *cas* genes

was found in this bacterium [27, 41]. One polyfunctional Cas9 protein performs both the processing of pre-crRNA and the interference of foreign DNA [42]. The crRNA processing also depends on a small non-coding RNA, tracrRNA (trans-activating crRNA). tracrRNA molecules bind complementarily to repeat sequences in pre-crRNA, forming a duplex, while one of the ribonucleases of the host cell, RNase III, cuts the duplex in the presence of Cas9 to form mature crRNA containing a 20 nucleotide spacer sequence at the 5'-end. Cas9 makes a double-strand break in the target locus in the presence of Mg²⁺ ions, with the HNH nuclease domain of the enzyme cutting the DNA strand complementary to crRNA, and the RuvC domain cutting the non-complementary strand [43]. The target DNA for Cas9 of *S. pyogenes* should necessarily contain 5'-NGG-3' PAM [43, 44], three nucleotides from which cleavage occurs. In *S. thermophilus* and *Neisseria meningitidis*, targets for type II Cas9 have a different consensus (5'-NG-GNG-3' and 5'-NNNNGATT-3', respectively).

GENOMIC ENGINEERING USING TALENS AND CRISPR/CAS9

The general strategy in genomic engineering using site-specific nucleases comprises four main stages (Fig. 3):

1. Selection of a target nucleotide sequence in the genome;
2. Generation of a nuclease construct directed at the selected target;
3. Delivery of this construct to the cell nucleus; and
4. Analysis of produced mutations.

Selection of a target nucleotide sequence in the genome

An important aspect of working with the TALEN and CRISPR/Cas9 systems is careful selection of sites for the specific introduction of a double-strand break. The need for a preliminary bioinformatic analysis is explained by the possibility for off-target effects – introducing non-specific double-strand breaks into the genome. When selecting desired sites, regions of repeated sequences, as well as regions with a high homology to other genome regions, should be avoided.

The off-target effects, when using the chimeric TALEN protein system, arise for several reasons. First, these are differences in the binding efficiency of RVD and specific nucleotides. HD and NN monomers form strong hydrogen bonds with nucleotides, while NG and NI form weak hydrogen bonds. This causes a possible binding of the DNA recognition domain to sites that differ from the target sites in a few nucleotides. Second, the degeneracy of the code for the binding of nucleotides by monomers may lead to, for example,

interaction between NG and A. Third, dimerization of the FokI domains of two nucleases with identical DNA-binding domains (formation of homodimers) is possible. This issue has been resolved in a number of studies by producing TALENs that contain the FokI domains acting as obligate heterodimers. Finally, the possible off-target effects may result from the fact that the size of the spacer DNA between the nuclease recognition sites is not fixed. This property makes it possible to introduce double-strand breaks during the binding of nucleases to off-target sites located at a distance sufficient for the dimerization of the FokI domains [45].

Since Cas9 nuclease of *S. pyogenes* needs the obligatory presence of the PAM with the 5'-NGG-3' consensus, though it is not much, but it limits selection of a target. In particular, target sites in the human genome are located in every 8–12 bp [46, 47]. One of the main drawbacks of the CRISPR/Cas9 system is a relatively high probability of off-target mutations. A number of studies carried out *in vitro* [43], in bacteria [48], and in human cells [46] have demonstrated that some single nucleotide substitutions in the 20-nucleotide spacer region of sgRNA (single-guide RNA) may lead to a significant reduction in the activity of CRISPR/Cas9, especially if these substitutions are located in the last 10–12 nucleotides of the 3'-end of this region of sgRNA [49]. At the same time, substitutions at the 5'-end of sgRNA have actually no effect on the system's activity [43, 46, 48]. However, cases are known when some single- and dinucleotide substitutions at the 3'-end of sgRNA do not affect the CRISPR/Cas9 system's activity and, instead, inhibit its action, if they are located at the 5'-end [49]. In general, the off-target effect is determined by the position of substitutions, when 8–12 bp at the 3'-end of the guide sequence are less important for Cas9 than the 5'-end nucleotides; by the number of substitutions, which should not be more than three; by features of the very target site; by the concentration of introduced Cas9 and sgRNA [46–49]. The search for and development of methods based on the use of Cas9 orthologs, the activity of which needs a PAM with a more complex consensus sequence, will overcome these drawbacks. For example, type II CRISPR/Cas of *N. meningitidis* recognizes the PAM with the 5'-NNNNGATT-3' consensus, which certainly limits the choice of a target but may increase the specificity.

In order to increase the specificity of genome editing based on the CRISPR/Cas system, two Cas9 nucleases with a pair of sgRNAs are used [50, 51] by analogy with pairs of ZFNs and TALENs, which cause breaks in DNA only under the action of two independent proteins with the FokI domains. Mutations in one of the catalytically active domains (D10A in HNH and H840A in

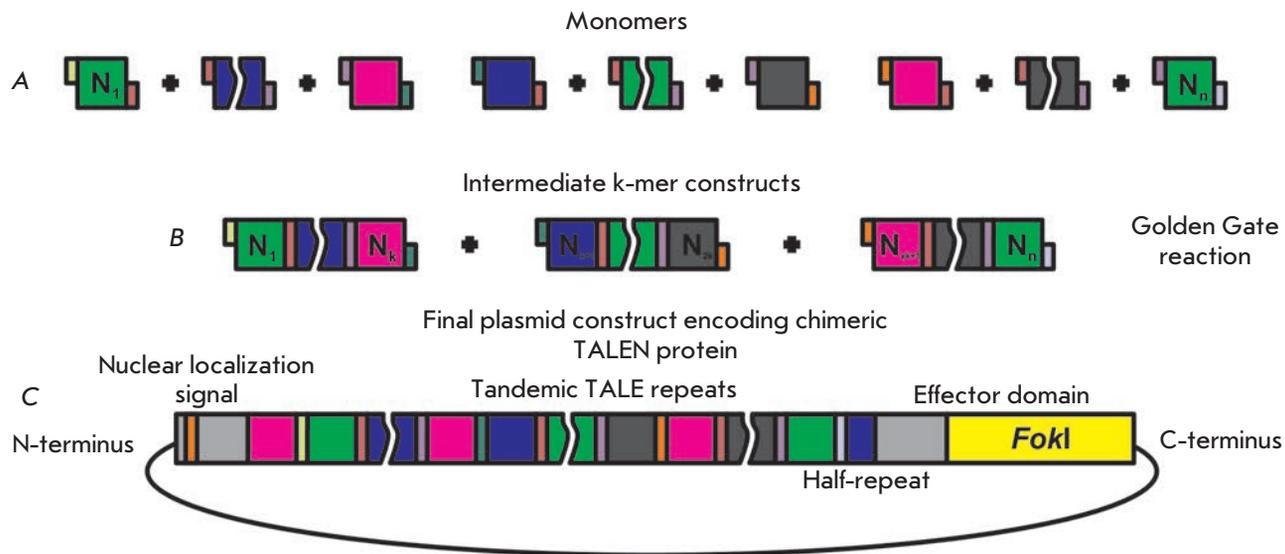


Fig. 4. A scheme of the modular hierarchical ligation strategy based on the Golden Gate cloning system to generate genetic constructs expressing chimeric TALEN proteins. **A** – at the first stage, a library of monomers is generated, which is a kind of a “construction kit” comprising a set of parts. These parts are amplified sequences of monomers with specific oligonucleotide primers. The primers are selected in such a way that hydrolysis by type IIS restriction endonucleases results in the formation of sticky ends, which define the monomer position in a final construct. **B** – a single Golden Gate reaction enables simultaneous ligation of multiple monomers, which results in intermediate k-mer constructs. **C** – at the last stage, the Golden Gate reaction is carried out, resulting in restriction and ligation of several intermediate k-mer constructs and the backbone plasmid containing the remaining TALEN elements

RuvC) convert Cas9 nuclease into DNA nickase [43, 46, 52]. If cleavage of both DNA strands by a pair of Cas9 nickases leads to the formation of site-specific double-strand breaks that are repaired via non-homologous end joining (NHEJ), individual single-strand damages are primarily repaired via highly accurate base excision repair (BER) [53]. The use of two Cas9 nickases with a sgRNA couple was demonstrated to provide a significant reduction in the production of off-target mutations, with the yield of target mutations generally corresponding to that for the use of nuclease [50, 51].

The mentioned properties of target site recognition by the CRISPR/Cas9 and TALEN systems were taken into account when developing computer algorithms that search for these sites. Currently, on-line software is available that was developed by different teams and designated for the selection of potential sites for the TALEN [54–59] and CRISPR/Cas9 [47, 60–62] systems, as well as for the detection of possible off-target effects.

Generation of genetic constructs expressing CRISPR and TALEN

TALEN. The DNA-binding domain consists of almost identical repeats: so there are certain technical difficulties in creating genetic constructs expressing TALENs. A number of methods have been suggested that enable

to construct TALE DNA-binding domains consisting of 20–30 or even more monomers. One of the strategies is based on standard DNA cloning using hydrolysis by type II DNA restriction endonucleases and ligation: REAL (REstriction and Ligation [63]). At the first stage, a library of monomers is generated that are introduced with endonuclease restriction sites at the 5’- and 3’-termini. After DNA hydrolysis, pair-wise ligation is carried out resulting in the formation of dimers ($N_1N_2, N_3N_4, N_{2k-1}N_{2k}$) that are then combined in tetramers and so forth. In this case, the correct sequence is achieved by the use of various restriction endonucleases. This technique is rather difficult and time consuming, because at each stage, the reaction products should be purified, and the correct orientation should be confirmed. To accelerate this process, a library of 376 members was generated that consists of mono-, di-, tri-, and tetramers (REAL-Fast, [64]).

To increase the efficiency and to accelerate the assembly process, the Golden Gate reaction is used [65, 66], which is simultaneous ligation and hydrolysis by restriction endonucleases in the same reaction mixture (Fig. 4). In the Golden Gate reaction, type IIS restriction endonucleases are used that hydrolyze DNA at a fixed distance from the recognition site, for example, BsmBI or BsaI. Therefore, the “scarless” assembly oc-

curs during ligation, because restriction endonucleases “cut off” their own site from a monomer, and the ligation product is not subjected to restriction. A library that contains different variants of the all four monomers, corresponding to the various positions (e.g., 1 to 20) in the future DNA-binding domain, is generated by amplifying sequences of monomers (NI, HD, NG, and NN) with different oligonucleotide primers. Treatment of these monomers with IIS restriction endonucleases produces sticky ends complementary to the sticky ends of neighbor monomers. In a single reaction, several monomers can be simultaneously ligated; for example, four [67] and six [68]. Next, using the Golden Gate reaction again, it is possible to ligate several tetra- or hexamers and to clone a complete sequence into a plasmid vector containing the 3'-half-repeat and FokI catalytic domain.

In order to reduce the time needed to develop genetic constructs expressing TALEN, a method was proposed that enables the exclusion of DNA ligation and, accordingly, steps related to verification of its results. The selected DNA-binding domain is assembled from monomers with long, specific, single-strand ends (10–30 nucleotides). Upon mixing several monomers, annealing of complementary single-strand ends occurs, whereby the monomers are arranged into a desired sequence. Then, *E. coli* cells are transformed with the resulting mixture and ligation occurs already in bacteria with involvement of their own enzymes [69].

These methods for developing genetic constructs expressing TALENs are relatively simple, and, according to various estimates, their implementation takes 1–2 weeks, if appropriate reagents are available. In addition to the simplicity and efficiency, this technology is also easily accessible; currently at the Addgene Depository (<http://www.addgene.org/TALEN/>), kits for the construction of TALENs developed by different groups of authors [64, 68–71] can be purchased and used in the laboratory.

Also, there are available systems for the automated high-performance production of constructs expressing TALENs nucleases. For example, a commercial platform from Cellectis Bioresearch enables one to generate up to 7,200 of these constructs annually. Three methods based on the use of solid phase surfaces have been described in the scientific literature [72–74]. These methods avoid an analysis of intermediate constructs, their purification by extraction from the gel, and other stages, which makes these methods suitable for automated production and accelerates the process. The idea behind these methods is to use streptavidin-coated magnetic particles with attached biotinylated double-strand DNA adapters. Sequential alternation of the phases of DNA hydrolysis by restriction endo-

nucleases and ligation is used to extend a sequence of monomers that is connected via an adapter with the magnetic particle. The reaction products are purified by means of washing buffers on a magnetic plate. In this case, by-products and reaction components are washed away and the target product is retained in a test-tube (or well) due to the attraction between the magnetic particles and the plate. At the end, restriction endonucleases are used to cleave the links between the biotinylated adapter and the synthesized sequence of monomers of the DNA-binding domain of TALEN. The sequence is then cloned into a plasmid vector by means of DNA ligation. This method allows one to quickly and efficiently synthesize in parallel genetic constructs in 96-well plates using multichannel pipettes or robotic pipetting stations.

CRISPR/Cas9. It was demonstrated that for cleavage of DNA *in vitro* [43, 52] and in bacterial cells [42] using CRISPR/Cas9, the following components are necessary and sufficient: non-coding RNAs (tracrRNA and pre-crRNA), RNase III, and the Cas9 protein. The use of this system in mammalian cells exhibits several features.

First, SpCas9 nuclease (Cas9 of *S. pyogenes*) should be adapted for adequate transcription in high eukaryotic cells, in particular codon-optimized, and attachment of nuclear localization signals (NLS) is necessary to provide a nuclear compartmentalization; two NLS are sufficient for effective guiding of Cas9 to the nucleus [46].

Second, maturation of pre-crRNA in eukaryotic cells does not require the introduction of exogenous RNase III, since this function is successfully performed by its own cellular RNases [75–77].

Third, instead of two non-coding RNAs, single chimeric sgRNA is often introduced, in which mature crRNA is fused with a part of the tracrRNA through the synthetic “stem-loop” structure to simulate the natural crRNA-tracrRNA duplex [43] (Fig. 5). To transcribe sgRNA, an appropriate promoter is required: for example, the RNA polymerase III U6 promoter.

Basic plasmid constructs containing the elements necessary for CRISPR/Cas9 activity were produced in the Feng Zhang’s laboratory. The pX260/pX334 plasmids contain three expression cassettes: Cas9 nuclease/nickase, CRISPR mRNA, and tracrRNA (Fig. 6). To change the target sequence, this construct only needs cutting off the original 30 nucleotide guide sequence flanked by BbsI sites and replacing it with an artificially synthesized one. To this effect, 30-mer oligonucleotides complementary to the target sequence and containing the appropriate sticky ends are melted together and ligated to the plasmid.

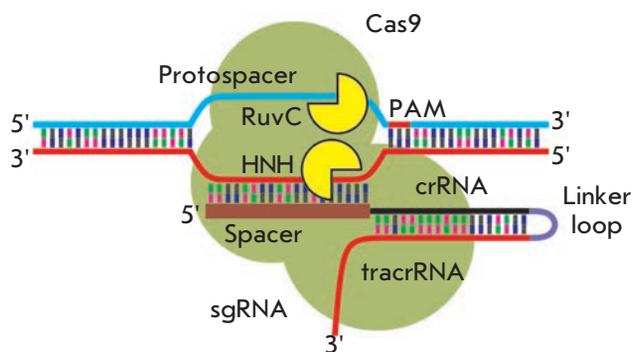


Fig. 5. Single chimeric sgRNA to introduce double-strand breaks into the target loci. A complex of sgRNA and Cas9 is capable of introducing double-strand breaks into selected DNA sites. SgRNA is an artificial construct consisting of elements of the CRISPR/Cas9 system (crRNA and tracrRNA) combined into a single RNA molecule. A protospacer is a site that is recognized by the CRISPR/Cas9 system. A spacer is a sequence in sgRNA that is responsible for complementary binding to the target site. RuvC and HNH are catalytic domains causing breaks at the target site of the DNA chain. PAM is a short motif (NGG in the case of CRISPR/Cas9) whose presence at the 3'-end of the protospacer is required for introducing a break

pX330/pX335 plasmids contain two expression cassettes: Cas9 nuclease/nickase and chimeric sgRNA comprising 85-nucleotide tracrRNA. The principle of changing the guide sequence is the same, but the sequence length is shorter – 20 nucleotides – and the 20th position should be occupied by guanine, because the U6 promoter used in this case comprises this base at the transcription start point. Furthermore, these plasmids can be inserted with additional elements, such as the 2A-GFP or 2A-Puro sites, for subsequent selection of cells bearing the plasmid.

Delivery of constructs expressing CRISPR/Cas9 components

To transform human, mouse, and other cell cultures, plasmids providing extensive production of Cas9 nuclease and sgRNA *in vitro* are more often used [46, 78–80]. To transform the whole organism, a method based on microinjection of *cas9* mRNA and sgRNA into single-celled embryos was developed [81–83]. This method is widely used in mouse, zebrafish (*Danio rerio*), and drosophila. For large-scale genome-wide knockout using large sgRNA libraries, lentiviral vectors are employed [84, 85]. In plants, which have cells with a thick cellular wall, the method of protoplast-plasmid transformation in cell cultures [86, 87], as well as agroinfil-

tration using *Agrobacterium tumefaciens* [88, 89], is widely used.

Analysis of mutations caused by CRISPR/Cas9 and TALEN

Due to the activity of CRISPR/Cas9 or TALENs systems, a double-strand break is introduced into eukaryotic DNA in the region of the CRISPR/Cas9 protospacer or spacer sequence separating the TALEN recognition sites (Fig. 7). In the absence of a homologous donor DNA, the double-strand break is repaired by nonhomologous end joining. During this process, errors occur and small insertions or deletions happen at a high frequency in the joining region [90]. A number of techniques based on the detection of such changes in a target DNA have been developed to study the activity of artificial nucleases in eukaryotic cells (Fig. 7).

A method based on TOPO cloning allows one to study the nucleotide sequences of mutant alleles resulting from nonhomologous DNA end joining as well as highly accurate quantification of the efficiency of artificial nucleases (Fig. 7). Eukaryotic cells are treated with artificial nucleases, then genomic DNA is isolated, and the DNA segment containing a nuclease recognition site is amplified by PCR. The PCR products are cloned in a plasmid vector, followed by sequencing of the clones produced after the transformation to *E. coli* cells [72]. Based on this, the variety of the generated mutations and their frequency are determined. Furthermore, if the cells treated with artificial nucleases are used to produce clonal populations, then lines carrying certain mutations may be selected after sequencing. For example, based on a selection of clones with a deletion of a certain size, cell lines were produced in which the reading frame impaired by the Duchenne muscular dystrophy mutation was restored [91].

The artificial nuclease activity is analyzed using enzymes that cleave the phosphodiester bonds in unpaired DNA segments (Fig. 7). Amplification of a segment selected as a target for artificial nucleases produces a mixture of DNA molecules, the nucleotide sequences of which are different due to the insertions or deletions that occurred during nonhomologous end joining. Denaturation followed by re-hybridization of a PCR product results in the formation of heteroduplexes containing loops in unpaired segments. After re-hybridization, PCR products are treated with enzymes such as phage T7 endonuclease I [92] or nucleases of the CELI family [93], and then the resultant fragments are separated by electrophoresis. Detection of hydrolysis products indicates that a PCR product mixture contains fragments with insertions or deletions resulting from nonhomologous end joining. The efficiency of artificial nucleases may be estimated by the ratio of

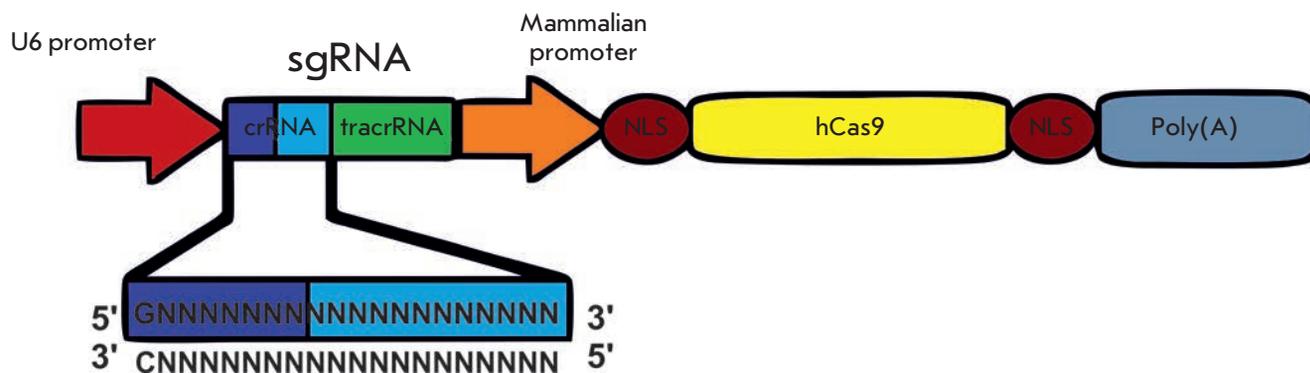


Fig. 6. A scheme of a genetic construct expressing CRISPR/Cas system elements. hCas9 is the Cas9 protein sequence optimized for expression in eukaryotic cells. sgRNA is a single chimeric RNA containing the parts of crRNA and tracrRNA necessary for the activity. NLS is the nuclear localization signal, which provides penetration of the construct into the nucleus. Poly (A) is the polyadenylation signal

intensity of the main product and the fragments produced during the hydrolysis, but this is an inaccurate estimate [92].

The properties of the resulting heteroduplexes are different from those of homoduplexes. One of these differences is a change in the melting curve profile that can be detected using the high resolution melting analysis (HRMA) (Fig. 7). A short segment (100–300 bp) containing a double-strand break site is amplified using real-time PCR with fluorescent intercalating dyes. Then, after denaturation and rehybridization, HRMA is performed. Based on a comparison of the control and test samples, changes in the melting curve profile and, hence, changes in the nucleotide sequences resulting from nonhomologous end joining can be determined [94]. This analysis is sensitive and simple, but this is a qualitative method that does not allow one to estimate accurately the efficiency of artificial nucleases, as well as the nature of DNA changes.

Another method to determine whether a double-strand break was introduced into a target site is the analysis of the electrophoretic mobility of heteroduplexes. Unpaired segments of the single-strand DNA that form loops in heteroduplexes reduce their mobility in a 15% polyacrylamide gel compared to that of homoduplexes. This property makes it possible not only to determine whether a double-strand break occurred, but also to evaluate the variety of generated mutations as well as to genotype different clones, because different size deletions or insertions change the heteroduplex mobility in different ways. In this case, the mobility profile for lines containing the same mutation is also identical [95].

The efficiency of artificial nucleases can be quantified and compared using methods based on genetic

reporter constructs containing genes of luminescent proteins. In this case, single-strand annealing (SSA) is used, which is one of the ways used to repair double-strand breaks in the genome of eukaryotes. If a double-strand break occurs between two direct repeats, then annealing of the complementary sequences flanking the break occurs via SSA. Then, the nonhomologous regions are hydrolyzed by specific nucleases and the synthesis and ligation of new DNA occur in single-strand segments. The sequence between direct repeats where a double-strand break occurred is always deleted, and one sequence remains instead of two repeated sequences. This process is used to restore a reporter gene; e.g., the luciferase gene. After a double-strand break introduced into the target sequence cloned into a plasmid vector between two repeat elements of the reporter gene, the reporter function is restored by means of SSA. Therefore, the efficiency of artificial nucleases can be quantified by the level of luminescence. In this case, reporter constructs are transfected into eukaryotic cells such as HEK293 lines or some yeast strains. The disadvantage of this method is that it does not take into account the genomic environment in which the target site is located; so, its results may not correlate with the results obtained when working with target sites in the genome [96].

Japanese scientists have developed a method of analysis based on the impairment/restoration of the *lacZα* gene function (Fig. 7). For this purpose, the site designated for introducing a double-strand break is cloned into the *lacZα* gene. In this case, oligonucleotide primers are selected in such a way that the wild type target site impairs (1) or preserves (2) the reading frame. If a double-strand break occurred in the site that was repaired by nonhomologous end joining, then

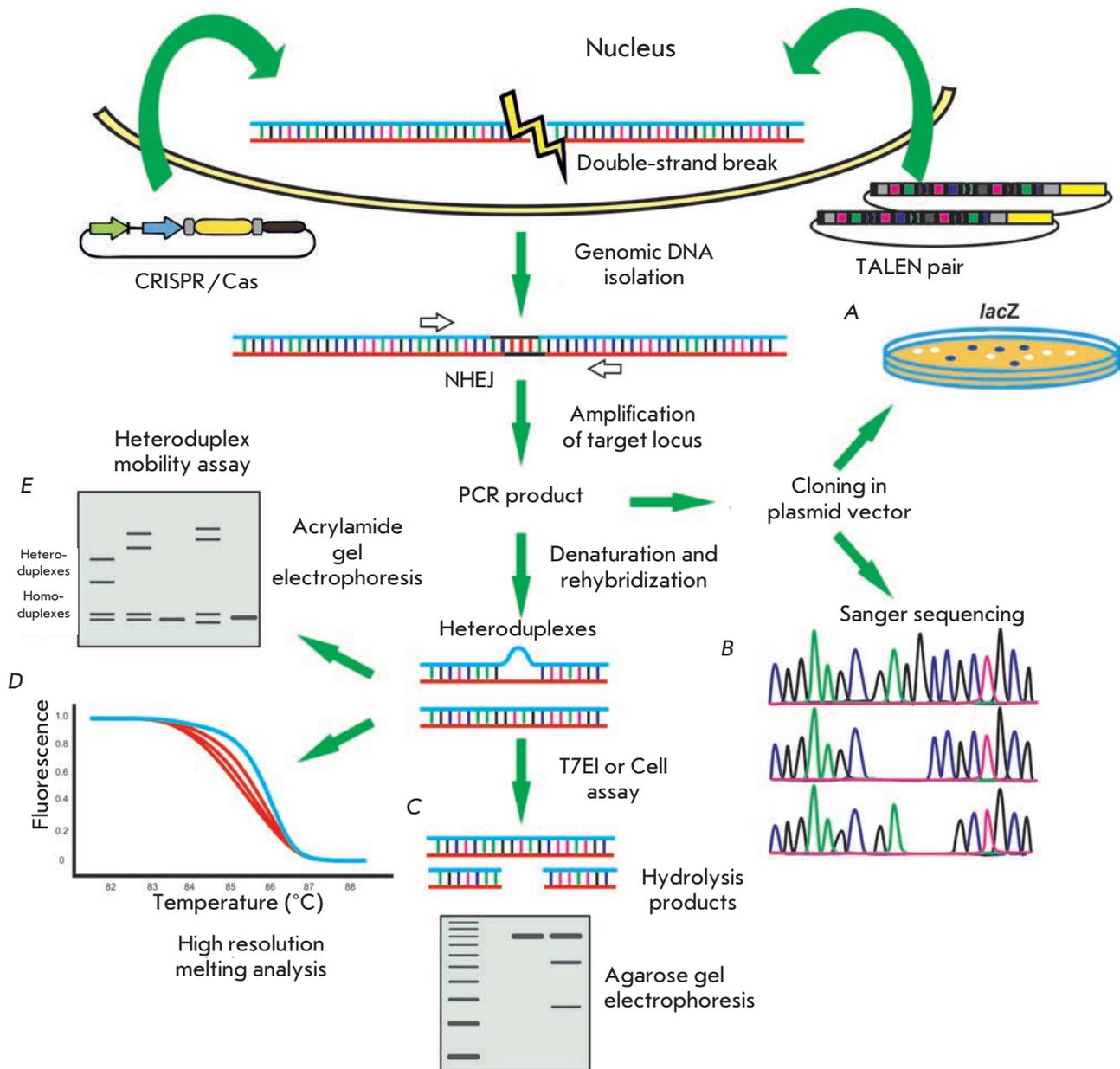


Fig. 7. A scheme of various analyses to identify and determine the efficiency of double-strand break introduction caused by the TALEN and CRISPR/Cas systems. First, constructs encoding CRISPR/Cas9 or TALEN are delivered into cells. In cells, double-strand breaks occur in the target loci that are repaired by nonhomologous end joining (NHEJ). This results in the formation of insertions or deletions. Next, the target locus is amplified by PCR. PCR products are analyzed by the following methods. **A** – a target segment is cloned into a plasmid vector. Impairment or, instead, recovery of the reading frame of the *lacZ* gene occurs due to the insertions or deletions. Based on the count of blue and white colonies after the transformation of *E. coli*, the efficiency of the CRISPR/Cas9 or TALEN systems is determined; **B** – after cloning into a plasmid vector and *E. coli* transformation, Sanger sequencing is performed. Clones containing insertions/deletions are counted, the efficiency is determined; **C** – after denaturation and re-hybridization of the PCR product, DNA heteroduplexes are formed; e.g., one strand is “wild type,” and the other contains a deletion. After treatment with enzymes that cut DNA in unpaired segments, samples are loaded onto a gel and electrophoresis is carried out. The hydrolysis products mean that the sample contained heteroduplexes; hence, a break appeared in double-strand genomic DNA under the action of CRISPR/Cas9 or TALEN; **D** – a high resolution melting analysis enables heteroduplex detection. Blue is the control samples, red is the samples containing heteroduplexes; **E** – unpaired DNA regions reduce the heteroduplex mobility in a 15% polyacrylamide gel. After gel electrophoresis, bands corresponding to homo- and heteroduplexes can be observed

in the first case, after cloning the reading frame will be restored in the one-third constructions due to deletions or insertions. Accordingly, after the transformation of *E. coli* cells with the produced constructs, a fraction of the colonies will be blue in color. In the second case, the reading frame will be impaired in the two-third constructions due to mutations caused by artificial nucleases. The colonies with these genetic constructs will be white in color. The efficiency of artificial nucleases can then be determined by simply counting the fraction of blue or white colonies in the first and second cases, respectively [97].

APPLICATION OF CRISPR/CAS9 AND TALEN SYSTEMS

Nuclease makes double-strand breaks in a target site that are repaired by the cell through one of two possible mechanisms:

Nonhomologous end joining, when errors occur that result in indel type (insertions, deletions) mutations in the target locus.

Homologous recombination, in which an intact homolog serves as a template to restore the original DNA structure; this is quite a rare event in the cell, but the use of CRISPR/Cas9 and TALENs increases the probability of homologous recombination by several orders of magnitude. If CRISPR/Cas9 components are added with artificially synthesized DNA showing homology with a nucleotide sequence at the break, then it may serve as a template for another way to repair DNA, homology-directed repair (HDR), in which a small piece of an artificial template is introduced into the target locus. As such a template, two types of constructs are most often used: single-strand oligonucleotides and plasmid vectors. In the first case, oligonucleotides homologous to the site for double strand break introduction are artificially synthesized; the optimum oligonucleotide length is about 90 nucleotides [98]. These oligonucleotides may be slightly different from the target site. When plasmid vectors are used as donor molecules for recombination, sufficiently long homology arms are cloned in them (500 to several thousand base pairs). These homology arms can flank additional elements such as reporter genes, antibiotic resistance genes, and so forth. Besides transgenesis HDR can be used to alter the genome via original nucleotide sequence replacing: synonymous substitutions can be generated to provide a new restriction site or a mutant allele, for example, which causes some hereditary disease, can be replaced with wild type allele. (genetic correction) However, HDR occurs vigorously only in dividing cells and its efficiency is highly depends on the cell type, stage of life, as well as the target locus of the genome and template itself [99].

Therefore, the following mutations can be produced

using site-specific nucleases:

- non-homologous end joining in the absence of a donor plasmid mediates deletions or insertions of several nucleotides in the target site and, as one of the results, knockout due to reading frame mutations and stop codon formation [100];
- in the presence of double-strand oligonucleotides or a donor plasmid, DNA fragments of more than 14 kb can be inserted through nonhomologous end-joining-mediated ligation [101, 102];
- simultaneous introduction of several double-strand breaks may lead to deletions, inversions, or translocations of the DNA regions located between these breaks [46, 103];
- homologous recombination in the presence of a donor plasmid with homology arms flanking the inserted fragment [104], a linear donor sequence with homology of less than 50 bp [105], or an oligonucleotide [103] leads to insertion of one or more transgenes for the correction or replacement of existing genes.

Currently, these methods are extensively used in basic and applied research. In this case, genome editing is possible both *in vitro* upon delivery of TALEN or CRISPR/Cas elements to cell cultures and *in vivo* by mRNA injections into zygotes (Fig. 3).

In vitro genome editing

The HEK 293T/HEK 293FT cell lines are commonly used to test the efficiency of the TALEN and CRISPR/Cas systems in a human *in vitro* model, because they can be transfected easily by plasmids and are relatively simple to maintain, [46, 50, 68, 78, 106]. According to different authors, the level of targeted mutations and also homologous recombination with donor plasmids/oligonucleotides varies widely, which probably depends not only on the method, but also on the cell line and the genomic target itself (Table). Cultured lines of induced pluripotent stem cells and human embryonic stem cells are of particular interest for regenerative medicine, the investigation of the structure and functioning of complex gene networks, the development of drug search systems, and a variety of other basic and biomedical studies.

Using the TALEN system, Ding *et al.* [71] introduced double-strand breaks and obtained human stem cell lines with mutations in various disease genes. In total, 15 genes were mutated and a comprehensive phenotype analysis of differentiated derivatives of stem cells with mutations in four of them (*APOB*, *SORT1*, *AKT2*, and *PLIN1*) was performed. New data on the role of these genes in the pathogenesis of diseases were obtained due to these cell models. For example, the *APOB* gene product was demonstrated to be necessary for the replication of the hepatitis C virus in human hepato-

cytes. Viral replication is greatly reduced in cells with a homozygous mutation in this gene. What is more, the E17K mutation in the *AKT2* gene leads to a decrease in the glucose synthesis in human hepatocytes and an increase in the level of triglycerides in adipocytes.

In addition to the generation of the models required for developing approaches to the treatment of diseases, artificial nucleases may be used directly for therapeutic purposes. One of such trends is the therapy of chronic viral infections. TALENs may be constructed to allow the introduction of mutations into the open reading frames of viruses such as HIV, hepatitis B, and herpes, which may be present in the body in a latent state and not be affected by therapy against replicating viruses [107, 108]. For example, the C-C chemokine receptor type 5 gene of T lymphocytes, whose mutations render a person resistant to HIV, can be modified using TALENs [100, 109].

With the use of a CRISPR/Cas9-based technology, isogenic human stem cells were generated [110], methods to correct a mutant cell phenotype are developed [111], and studies on the gene expression regulation [112–116], functional relationships between large groups of genes [84, 85], and imaging regions of the active genome regions in living cells [117] are conducted.

The development of panels of human isogenic pluripotent stem cells will implement modeling of hereditary and multifactorial diseases, screening of large drug libraries, as well as searching for new mutations involved in the pathological process. Currently, all these areas are under extensive study. For example, the CRISPR/Cas9 system was effectively used to generate a ICF syndrome model (ICF is the immunodeficiency, centromeric region instability and facial anomalies syndrome) using human-induced pluripotent stem cells. Homozygous mutations in the *DNMT3B* gene with a frequency of 63% were generated, with cells having the centromeric instability phenotype [110]. It seems particularly relevant to study severe neurodegenerative diseases such as Alzheimer's, Parkinson's, and various muscle atrophies.

Since Cas9 recognizes a particular target in the genome with the participation of a short guide sequence in sgRNA, currently, it is relatively easy to generate a large genome-wide library of oligonucleotides and, accordingly, sgRNAs. Furthermore, the use of lentiviruses, which are stably maintained in the genome and replicated together with genomic DNA, as a vector for the delivery of CRISPR/Cas9 components, has allowed one to develop a new GeCKO (Genome-scale CRISPR/Cas9 knockout) technology [84]. A large sgRNA library allows researchers to turn off the transcription of many genes simultaneously and thereby identify the functional relationships among them, their role in cer-

tain life processes, or their involvement in the pathological process. For example, the genes necessary for the life activity of cancer cells (A375 cell line of human melanoma) and pluripotent stem cells (HUES62 line) were identified using a lentivirus library comprising 18,080 genes (three or four sgRNAs for each gene) [84]. The development of resistance to vemurafenib (PLX), which is a BRAF inhibitor of protein kinases in melanoma, was demonstrated to involve not only the *NF1* and *MED12* genes, but the *CUL3* gene as well as the genes of the STAGA complex of histone-specific acetyltransferases: *TADA1* and *TADA2* [84]. Based on a lentiviral library comprising about 73,000 sgRNAs, the genes involved in the proliferation and cell cycle were studied using the HL60 and KBM7 tumor cells [83]. It was demonstrated that mutations resulting in the formation of nonfunctional products of four DNA mismatch repair (MMR) genes (*MSH2*, *MSH6*, *MLH1*, and *PMS2*) cause resistance to a nucleotide analog, 6-thioguanine, and therefore provide the cell proliferation. The activity of the genes *TOP2A*, *CDK6*, *BCR*, and *ABL1* and the genes that encode ribosomal proteins was also studied.

Therefore, the use of CRISPR/Cas9 libraries allows one to perform functional screening of genomes, which may yield important information about the physiology and biochemistry of different types of cells and could help reveal the molecular mechanisms of disease development and identify potential targets for drug and gene therapy.

CRISPR/Cas9-based methods can be effectively used to edit the genomes of cultured stem cells. In particular, the use of genome-editing systems enables one to correct point mutations in the cells obtained from patients. The object of research in this case may be induced pluripotent stem cells and regional stem cell. In this case, both complex genetic constructs and single-strand DNA oligonucleotides can be used as donor molecules [98].

An interesting example of this approach is a study in which correction of the *CFTR* locus (cystic fibrosis transmembrane conductor regulator) was performed in cultured intestinal stem cells derived from cystic fibrosis (CF) patients [111]. This approach enables the so-called organoids obtaining: functional multicellular structures with a corrected genome that are autologous with respect to the cell donor, which may be administered back to the patient. Certainly, this trend offers great opportunities for cell therapy of human diseases.

The controlled introduction of transgenes into the genome may be used in the case of functional correction of the genetic abnormalities associated with gene deletions or expression impairments that manifest themselves in a significant reduction in the level of gene products (protein or RNA). There are genome re-

gions, the introduction of transgenes into which is considered safe. These are sites like AAVS1 that provide stable expression of the introduced transgene [118]. Thus, the TALEN and CRISPR/Cas systems can effectively be used in the functional genomics of cells for the generation of cell models of human diseases and for cell therapy.

***In vivo* genome editing**

In genetics, for many years of its existence, a number of model objects have been formed that are studied in most detail and used in most basic and applied research. Model organisms include, for example, yeast, nematode, drosophila, arabidopsis, zebrafish, laboratory mice, and rat. These and a number of other model organisms are extensively used to perform experiments on genomic engineering using the CRISPR/Cas9 and TALEN systems.

Various applications of CRISPR/Cas and modifications of the genome editing technology in the nematode *Caenorhabditis elegans* are presented in a number of studies [119–126]. By injection of a mRNA/Cas9 protein and *in vitro/in vivo* produced sgRNA in germline cells, stable targeted genome modifications were produced in adult animals in the next generation, including small insertions/deletions, larger chromosome deletions and rearrangements [119], and transgene introduction by homologous recombination with donor molecules [121, 123]. This method is widely used to study the processes of dosage compensation in nematode and to compare gene functions in related species of *C. elegans* and *C. briggsae* [122].

The fruit fly, *Drosophila melanogaster*, is among the most studied model objects. However, the production of new mutant alleles by homologous recombination still remains a very labor-intensive procedure [127–129]. Injection of *cas9* mRNA and sgRNA into drosophila embryos provides double-strand breaks in the target loci of the genome, repair of which leads to the generation of insertion/deletion type mutations at a very high level (Table). Embryo injection produces mutations in both alleles of the target gene in all cells of a developing, and adult afterwards, insect; however, a certain percentage of mosaics emerges in this case [130–132]. These mutations are stably transmitted from generation to generation, which provides the possibility to generate new lines of flies [133]. Recently, an application was developed (<http://www.flyrnai.org/crispr>) that enables effective experiments planning for editing the drosophila genome. Therefore, the CRISPR/Cas9 technology allows quick and efficient generation of mutations to further study the gene activity in *Drosophila*.

The zebrafish is currently a very popular object

not only for basic research of the structural and functional relationships in the genome, but also for modeling of metabolic and neurodegenerative diseases in humans *in vivo* [134]. Various target and stably inheritable modifications were generated by injection of CRISPR/Cas9 components into zebrafish embryos (Table). In 2011, the international Zebrafish Mutation Project was launched to generate mutant alleles in each zebrafish protein-coding gene. All data are analyzed on the website http://www.sanger.ac.uk/Projects/D_rerio/zmp. As of June 2013, mutant models of 46% of all protein-coding zebrafish genes have been generated.

Laboratory animals, such as mouse and rat, are considered the most important model objects for the investigation of human diseases, basic research of the structure and function of genes and regulation of their expression, as well as in pharmacology and toxicology. Previously, mouse lines with knockout of specific genes were produced by homologous recombination in embryonic stem cells [1, 83], as well as by insertional mutagenesis [135, 136]. These are very time- and labor-consuming experiments, and generation of double knockout animals is a more difficult task. The CRISPR-Cas9-based genome editing technology is a faster and less labor-intensive way to do the job in a single step. Targeted injection of site specific nucleases into a single cell zygote causes double-strand DNA breaks at the target locus [137–139]. These breaks are repaired via the nonhomologous end joining mechanism that leads to the generation of mutant rats and mice carrying deletions or insertions at the cleaved site [140, 141]. Upon addition of a donor plasmid or oligonucleotide, the breaks can be repaired through the high precision homologous recombination mechanism that enables the production of animals carrying target DNA inserts [83, 142, 143].

Genome editing using CRISPR/Cas9 makes possible the introduction of mutations both into one gene and into a few genes at once. It was demonstrated that CRISPR-Cas9 generates, with a high efficiency, mutations in five genes simultaneously in mouse embryonic stem cells, and injection of *cas9* mRNA and sgRNAs targeting the *Tet1* and *Tet2* genes into a mouse zygote generates animals with biallelic mutations in both genes with an efficiency of 80% [83]. Similar results were obtained in experiments in rats, with both mice and rats stably inheriting identifiable mutations [144, 145]. Furthermore, effective correction of a mutation in the *Crygc* gene in mice with a dominant form of cataract induced by this mutation was performed [146]. Generation of model rodents carrying specific mutations in several loci makes it possible to analyze the functions of genes belonging to gene families with redundant functions as well as epistatic interactions

REVIEWS

Genomic engineering using TALEN and CRISPR/Cas

Nuclease	Object	Gene	Objective	Reference
TALEN	Human cells (<i>Homo sapiens</i>)	<i>ccr5, akt2, e17k, angptl3, apob, atgl, c6orf106, celsr2, cftr, ciita, foxo1, foxo3, gli1, glut4, hbb, hdac1, hdac2, hdac6, hmga2, hoxa13, hoxa9, hoxc13, hpvt, il2rg, jak2, kras, linc00116, maoa, map2k4, mdm2, met, mlh1, msh2, mutyh, myc, mycl1, mycn, nbn, ncor1, ncor2, nlrc5, ntf3, pdgfra, pdgfrb, phf8, plin1, pms2, ppp1r12c (aavs1), ptch1, pten, rara, rbbp5, recq4, ret, runx1, sdhb, sdhc, sdhd, setdb1, sirt6, smad2, sort1, sox2, klf4ss18, suz12, tfe3, tp53, trib1, tsc2, ttn, vhl, xpa, xpc, abl1, alk, apc, atm, axin2, bax, bcl6, bmpr1a, brca1, brca2, cbx3, cbx8, ccnd1, cdc73, cdk4, cdh4, chd7, cttnb1, cyld, ddb2, ercc2, ewsr1, ext1, ext2, ezh2, fanca, fancb, fancf, fancg, fes, fgfr1, fh, flcn, flt4, mstn, aavs2, oct4, pitx3</i>	Knockout, insertion	[67, 68, 70–72, 74, 92, 176–179, 180]
	Yeast (<i>Saccharomyces cerevisiae</i>)	URA3, ADE2, LYS3	Knockout, insertion	[181]
	Nematode (<i>Caenorhabditis elegans</i>)	<i>ben-1, tex-1, sdc-2</i>	Knockout	[182]
	Drosophila (<i>Drosophila melanogaster</i>)	<i>yellow, crhdr1, ponzr1, bmil, cdh5, dip2a, elmo1, epas1b, fh, golden, gria3, hey2, hif1ab, ikzf1, jak3, moesina, myod, phf6, ppp1cab, ryr1a, ryr3, scl6a3, tbc6, tnkb, th, fam46c, smad5</i>	Knockout, insertion	[94, 183–187]
	Silkworm (<i>Bombyx mori</i>)	<i>blos2</i>	Knockout	[188]
	Cricket (<i>Gryllus bimaculatus</i>)	<i>lac2</i>	Knockout	[189]
	Western clawed frog (<i>Xenopus tropicalis</i>)	<i>ets1, foxd3, grp78/bip, hhex, noggin, ptf1a/p48, sox9, vpp1</i>	Knockout	[190]
	Mouse (<i>Mus musculus</i>)	<i>c9orf72, fus, lepr, pak1ip1, gpr55, rprm, fbxo6, smurf1, tmem74, wdr20a, dcaf13, fam73a, mlkl, mstn, pibf1, sepw1, rab38, zic2</i>	Knockout, insertion	[179, 191–196]
	Rat (<i>Rattus norvegicus</i>)	<i>bmpr2, IgM</i>	Knockout	[197, 198]
	Pig (<i>Sus scrofa</i>)	<i>amely, dmd, gdf8, ggta, ghdrhdr, il2rg, ldlr, raq2, rela (p65), sry</i>	Knockout	[199]
	Cattle (<i>Bos taurus</i>)	<i>acan, gdf8, ggta, mstn, prnp</i>	Knockout	[179, 199]
	Arabidopsis (<i>Arabidopsis thaliana</i>)	<i>adh1</i>	Knockout	[70]
	Tobacco (<i>Nicotiana benthamiana</i>)	<i>surA, surB, hax3</i>	Knockout, insertion	[156, 157]
	False brome grass (<i>Brachypodium distachyon</i>)	<i>aba1, cck2, coi1, hta1, rht, sbp, smc6, spl</i>	Knockout	[154]
	Rice (<i>Oryza sativa</i>)	<i>avrxa7, pthxo3, badh2, cck2, dep1, sd1</i>	Knockout	[154, 155]
CRISPR/Cas	Yeast (<i>Saccharomyces cerevisiae</i>)	CAN1, ADE2	Knockout, insertion	[200]
	Human cells (<i>Homo sapiens</i>)	<i>dnmt3b-tdTomato, pou5f1(oct4), emx1, dyrk1a, grin2b, egfp, ccr5, c4bpb, pvalb, aavs, akt2, celsr2, ciita, glut4, linc00116, sort1, ldlr</i>	Knockout, insertion	[46, 51, 78, 80, 201, 202]
	Nematode (<i>Caenorhabditis elegans</i>)	<i>dpy-11, unc-4, ben-1, unc-36, daf-2, klp-12, lab-1, egfp, dpy-11, lin-5, rol-1, dpy-3, unc-1, dpy-13, unc-119, klp-12</i>	Knockout, insertion	[119–124, 126]
	Drosophila (<i>Drosophila melanogaster</i>)	<i>yellow, white, rosy, cg14251 (k81), cg3708cg17629 (kl-3), light</i>	Knockout, insertion	[130–133]
	Danio rerio (<i>Danio rerio</i>)	<i>etsrp, gata5, etsrp, gsk3b, apoea, fh, fh1, th1, rgs4, tia1l, tph1a, drd3, egfp, tyr, gol, mitfa, ddx19, sema3fb, dre-mir-126a, dre-mir-126b, dre-mir-17a-1-dre-mir-92a-1, dre-mir-17a-2-dre-mir-92a-2, fgd5, ensdarg00000070653, ensdarg00000076787, psmf1, dre-mir-126a, dre-mir-17a-2, dre-mir-92a-2, tardbp, tardbpl, c13h9orf72</i>	Knockout, insertion, chromosomal rearrangements	[81, 82, 203–206]
	Frog (<i>Xenopus tropicalis</i>)	<i>tyr, six3</i>	Knockout	[207]
	Pig (<i>Sus scrofa</i>)	<i>gdf8, p65</i>	Knockout, insertion	[208]
	Mouse (<i>Mus musculus</i>)	<i>tet1, tet2, tet3, sry, uty, rosa26, hpvt, egfp, th, rheb, uhrf2</i>	Knockout, insertion	[83, 144, 209, 210]
	Rat (<i>Rattus norvegicus</i>)	<i>dnmt1, dnmt3a, dnmt3b, tet1, tet2, tet3, mc3r, mc4r</i>	Knockout, insertion	[144, 145, 211]
	Arabidopsis (<i>Arabidopsis thaliana</i>)	<i>pds33, fls2, bri1, jaz1, gaj, chl, chl2, 5g13930</i>	Knockout, insertion	[87, 88, 149]
	Tabacco (<i>Nicotiana benthamiana</i>)	<i>Pds</i>	Knockout, insertion	[88, 89]
	Rice (<i>Oryza sativa</i>)	<i>ods, badh2, mrk2, 02g23823, roc5, spp, ysa, myb1, cao1, lazy1, sweet11, sweet14</i>	Knockout, insertion	[86, 150, 152]
	Wheat (<i>Triticum aestivum</i>)	<i>Mlo</i>	Knockout	[86]

of genes. Data combining information on knockout of a certain mouse gene are available on the IMPC site (International Mouse Phenotyping Consortium, <https://www.mousephenotype.org/>).

Genome editing using TALENs and the CRISPR/Cas9 system is extensively used in plants. Targeted editing of plant genomes may be used to solve problems of both fundamental (investigation of gene function) and applied science (production of plants with new properties such as resistance to pathogens and herbicides, changes in metabolism, productivity, etc.) [147]. In this case, the protoplast transformation or *in planta* expression with *Agrobacterium tumefaciens* (agroinfiltration) is primarily used for the delivery of genetically engineered constructs [148]. Gene knockouts and precise modification have been produced in plants, such as arabidopsis, wheat, rice, and tobacco [86, 88, 89, 149–153].

Editing of plant genomes using the TALEN system has, to date, been carried out in four model objects [70, 154–157]. Rice resistant to the pathogen of *Xanthomonas oryzae pv* serves as an example of a plant that acquired new properties due to genome editing using the TALEN system. A double-strand break was introduced into the wild-type pathogen TAL effector recognition site at the locus of the *Os11N3* gene using artificial TALENs. In this way, plants resistant to infection by *X. oryzae pv* were produced [155].

ALTERNATIVE WAYS TO USE TALE AND CRISPR/CAS9

Deciphering the recognition code between TALE proteins and target nucleotide sequences, as well as developing methods to generate artificial DNA-binding domains based on this code, has allowed scientists to construct chimeric proteins capable of acting directly on the genome. These proteins are composed of DNA-binding and effector domains. Nuclease domains are mainly used as the effector domain; however, in a number of studies, chimeric proteins were generated that contained, besides the DNA-binding domain, recombinase, histone methyltransferase, and histone deacetylase domains and domains that activate or suppress gene expression. These chimeric proteins have enormous prospects for application both in applied and in fundamental science. The CRISPR/Cas9 system is modified similarly: a certain effector domain, e.g., a transcriptional activator or repressor, the GFP fluorescent protein, etc, is attached to the catalytically inactive Cas9 protein.

Regulation of gene expression using the TALE and CRISPR/Cas9 systems

For targeted activation of gene expression, constructs containing the TALE DNA-binding domain and the

synthetic VP64 domain [158], TALE-TF, are used. Once in the nucleus, a chimeric protein binds to a target nucleotide sequence and the VP64 domain attracts endogenous activators of gene expression [159]. In this case, the target gene expression is statistically significantly increased, which is usually confirmed by real time PCR. Activation of noncoding genes is also possible, e.g., the genes of miRNAs [160]. Suppression of the target gene expression can be achieved using chimeric proteins containing the KRAB [161] or SRDX [162] domains.

A possible therapeutic application of TALE-TF is the targeted regulation of the expression of the genes associated with human diseases. To test this approach, a strategy was used to increase the expression level of the *FXN* gene that encodes the frataxin protein. Expansion of GAA trinucleotide repeats in this gene leads to the development of Friedreich's ataxia. In this case, the protein structure does not change but its expression is reduced. It was demonstrated that the *FXN* gene expression in human fibroblasts could be increased using TALE-TF, despite an increased number of the trinucleotide repeats [163].

Activation of endogenous gene expression avoids the use of ectopic overexpression of the reprogramming factors Oct4, Sox2, Klf4, and c-Myc (OSKM) in producing induced pluripotent stem cells. As a result, induced pluripotent stem cells can be produced that do not contain transgenes and, respectively, the risk of insertional mutagenesis, which arises when using lentiviral vectors expressing OSKM, can be reduced. For example, reprogramming of mouse embryonic fibroblasts to a pluripotent state was achieved through targeted activation of the expression of the *Oct4* and *Nanog* genes under the influence of TALE-TFs containing the VP64 domain [164].

More recently, transcription factors were generated for the targeted regulation of gene expression in response to an external chemical stimulus. These factors consist of the TALE DNA-binding domain and ligand-binding domain of the steroid hormone receptor. When a ligand (ecdysone) enters the cell, dimerization of the ligand-binding domain and, respectively, activation of the target gene expression occur [165].

A recently developed system of light-inducible transcriptional effectors (LITEs) is a combination of two, very promising trends in modern biotechnology: optogenetics and genomic engineering. This system consists of two parts. The first is the TALE DNA-binding domain connected to a light-sensitive domain, cryptochrome 2 (CRY2), isolated from *Arabidopsis thaliana*. The second is the VP64 transcriptional activator coupled with CIB1, which is able to interact with CRY2. CRY2 alters its conformation under the blue light irradiation and binds to CIB1, thereby attracting VP64 to

the target site [166]. A study by Konermann *et al.* [166], who developed the LITE system, demonstrated a statistically significant increase in the expression of some genes both in mouse neurons *in vitro* and in the brain *in vivo*. They also proposed a system in which the VP64 domain is replaced by methyltransferase or deacetylase capable of modifying histones.

An interesting application of targeted transcriptional regulation by TALE-TF is the development of genetic logic circuits inside the cell based on the interaction of several TALE-TFs with each other's promoters and with a reporter gene and the promoters of the factors that regulate expression. Based on this approach, logical NOT-OR [167] and AND [168] circuits were produced inside cells.

Catalytically inactive dCas9 or dCas9 coupled with factors regulating gene expression also allows one to activate or repress transcription in human, bacterial, and yeast cells [112–116]. For this purpose, the *E. coli* omega-subunit of RNA polymerase [113], tandem copies of the viral VP64 protein, and the KRAB domain can be used [112, 115]. For example, highly specific silencing of the *CD71* and *CXCR4* genes (at the level of 60–80%) as well as effective knockdown of the *TEF1* locus in yeast were achieved [112]. Furthermore, multiplex activation/repression of the promoters of several genes was achieved, with the regulation type (positive or negative) being controlled by the target position in the gene promoter [114, 115]. Therefore, the CRISPR/Cas9 system can be used as a modular platform that binds a given nucleotide sequence and attracts protein factors to it, thereby opening up opportunities of using this system as the main method for a precise regulation of gene expression in eukaryotic cells.

Imaging of internal genomic loci using the TALE and CRISPR/Cas9 systems

Chromatin organization and dynamics are known to play a decisive role in the regulation of genome activity. However, it is extremely difficult to obtain images of functional genomic loci in living cells. The use of the TALE and CRISPR/Cas9 systems opens up new possibilities for solving this problem.

Target DNAs in dynamics were visualized using constructs containing the TALE DNA-binding domain and a fluorescent protein [169–171]. This approach allows one to study the spatial and temporal organization of repeated genomic elements, including centromeric and telomeric repeats.

A method for imaging repetitive elements in the telomeres and the coding genes in living cells was developed using the endonuclease-deficient Cas9 protein labeled with EGFP and structurally optimized sgRNAs [117]. The repetitive and nonrepetitive elements in the

MUC4 and *MUC1* genes responsible for the production of various forms of mucin, which is a component of the protective mucus in various epithelial tissues and important in malignancy, were visualized in RPE, HeLa, and UMUC3 tumor cell lines [117]. Therefore, a possibility emerges to monitor the number of gene copies in living cells. The dynamics of telomere elongation and degradation, subnuclear localization of the *MUC4* loci, and cohesion of the replicated *MUC4* loci on sister chromatids and their changing behavior during mitosis were observed using this method [117]. This strategy has significant potential for the study of the conformation and dynamics of native chromosomes in living human cells.

Chimeric recombinases and transposases as an alternative to TALEN

Recombinases and transposases are an alternative to TALEN in genome editing. Their advantages include the lack of dependence on the intracellular repair mechanisms. These enzymes also perform cleavage and ligation at target sites, and respectively in this case, no accumulation of double-strand breaks, which may lead to cell death, occurs. In addition, recombinases and transposases insert donor DNA into the genome, which simplifies detection of their activity. The disadvantage of these chimeric enzymes is a fairly high level of off-target effects [172]. A catalytic domain of Gin recombinase [173, 174] or piggyBack transposase [175] is used as an effector domain. The TALE recombinase activity was demonstrated using a reporter gene, the promoter of which was specifically cut out by Gin recombinase. The possibility to edit the genome using transposase was demonstrated in the case of the *CCR5* locus.

CONCLUSION

The development of the TALEN and CRISPR/Cas9 systems is an important step in the progress achieved in modern genomic engineering. The emergence of these systems, due to their low cost and simplicity, has become a powerful impetus to the development of both fundamental and applied science. Prospects for the use of these systems in a variety of areas ranging from the food industry to personalized medicine are really amazing. However, until now, some problems have remained unresolved that are related to specificity and safety (due to possible off-target effects), delivery methods in therapeutic applications, and there is no answer to the question as to which of these systems combines the highest efficiency and safety?

The use of the CRISPR/Cas9 system has a number of advantages over the ZFN and TALEN based methods: it is much easier to produce, it is more efficient,

and is suitable for high-performance and multiplex genome editing in a variety of cell lines and in living organisms. To refocus it on a new target needs only replacing the 20-nucleotide guide sequence of sgRNA. Also, Cas9 causes a break strictly between the 17th and 18th nucleotides in the target sequence (counting from the 5'-end of the spacer), i.e. at a distance of three nucleotides from the PAM. Moreover, simultaneous editing of several genes is greatly simplified by introducing a combination of sgRNAs. The use of nickase and modification of the sgRNA construction for a more accurate target recognition in the genome allow researchers to avoid undesired off-target effects.

The TALEN system is more labor-consuming, it takes more time to construct compared to CRISPR/Cas9. However, there are now methods of automated design of TALEN-expressing constructs, which allows their efficient production on a commercial scale. Also,

the fact that TALENs cause breaks only upon dimerization of the FokI domain, i.e. in pairs, increases the specificity and reduces the risk of off-target effects.

To date, there is no definitive answer to the question of which of the systems should be used. A detailed comparison of the two systems, with each having its own features, is required. It is quite conceivable that a universal answer to this question will never be found, and for each particular case, it will be necessary to test different variants and to choose among them those that are most appropriate to the research goals.

This work was supported by the interdisciplinary integration project of the Siberian Branch of the Russian Academy of Sciences №55 and the Russian Foundation for Basic Research (grant № 12-04-00208-a).

REFERENCES

1. Capecchi M.R. // Nat. Rev. Genet. 2005. V. 6. № 6. P. 507–512.
2. Kim Y.G., Cha J., Chandrasegaran S. // Proc. Natl. Acad. Sci. USA. 1996. V. 93. № 3. P. 1156–1160.
3. Bibikova M., Golic M., Golic K.G., Carroll D. // Genetics. 2002. V. 161. № 3. P. 1169–1175.
4. Townsend J.A., Wright D.A., Winfrei R.J., Fu F., Maeder M.L., Joung J.K., Voytas D.F. // Nature. 2009. V. 459. № 7245. P. 442–445.
5. Zhang F., Maeder M.L., Unger-Wallace E., Hoshaw J.P., Reyon D., Christian M., Li X., Pierick C.J., Dobbs D., Peterson T., et al. // Proc. Natl. Acad. Sci. USA. 2010. V. 107. № 26. P. 12028–12033.
6. Torikai H., Reik A., Liu P.Q., Zhou Y., Zhang L., Maiti S., Huls H., Miller J.C., Kebriaei P., Rabinovitch B., et al. // Blood. 2012. V. 119. № 24. P. 5697–5705.
7. Provasi E., Genovese P., Lombardo A., Magnani Z., Liu P.Q., Reik A., Chu V., Paschon D.E., Zhang L., Kuball J., et al. // Nat. Med. 2012. V. 18. № 5. P. 807–815.
8. Lombardo A., Cesana D., Genovese P., Di Stefano B., Provasi E., Colombo D.F., Neri M., Magnani Z., Cantore A., Lo Riso P., et al. // Nat. Methods. 2011. V. 8. № 10. P. 861–869.
9. Becker M. // Nat. Methods. 2012. V. 9. № 1. P. 1.
10. Schornack S., Meyer A., Römer P., Jordan T., Lahaye T. // J. Plant Physiol. 2006. V. 163. № 3. P. 256–272.
11. Romer P., Hahn S., Jordan T., Strauss T., Bonas U., Lahaye T. // Science. 2007. V. 318. № 5850. P. 645–648.
12. Boch J., Scholze H., Schornack S., Landgraf A., Hahn S., Kay S., Lahaye T., Nickstandt A., Bonas U., et al. // Science. 2009. V. 326. № 5959. P. 1509–1512.
13. Moscou M.J., Bogdanove A.J. // Science. 2009. V. 326. № 5959. P. 1501.
14. Lamb B.M., Mercer A.C., Barbas C.F. // Nucleic Acids Res. 2013. V. 41. № 21. P. 9779–9785.
15. Christian M., Cermak T., Doyle E.L., Schmidt C., Zhang F., Hummel A., Bogdanove A.J., Voytas D.F. // Genetics. 2010. V. 186. № 2. P. 757–761.
16. Cong L., Zhou R., Kuo Y.C., Cunniff M.M., Zhang F. // Nat. Commun. 2012. V. 3. P. 968.
17. Christian M.L., Demorest Z.L., Starker C.G., Osborn M.J., Nyquist M.D., Zhang Y., Carlson D.F., Bradley P., Bogdanove A.J., Voytas D.F. // PLoS One. 2012. V. 7. № 9. e45383.
18. Streubel J., Blucher C., Landgraf A., Boch J. // Nat. Biotechnol. 2012. V. 30. № 7. P. 593–595.
19. Mak A.N., Bradley P., Cernadas R.A., Bogdanove A.J., Stoddard B.L. // Science. 2012. V. 335. № 6069. P. 716–719.
20. Ishino Y., Shinagawa H., Makino K., Amemura M., Nakata A. // J. Bacteriol. 1987. V. 169. № 2. P. 5429–5433.
21. Haft D.H., Selengut J., Mongodin E.F., Nelson K.E. // PLoS Comput. Biol. 2005. V. 1. № 6. e60.
22. Bolotin A., Quinguis B., Sorokin A., Ehrlich S.D. // Microbiology. 2005. V. 151. P. 2551–2561.
23. Mojica F.J., Diez-Villasenor C., Garcia-Martinez J., Soria E. // J. Mol. Evol. 2005. V. 60. № 2. P. 174–182.
24. Pourcel C., Salvignol G., Vergnaud G. // Microbiology. 2005. V. 151. P. 653–663.
25. Barrangou R., Flemaux C., Deveau H., Richards M., Boyaval P., Moineau S., Romero D.A., Horvath P. // Science. 2007. V. 315. № 5819. P. 1709–1712.
26. Grissa I., Vergnaud G., Pourcel C. // BMC Bioinformatics. 2007. V. 8. P. 172.
27. Makarova K.S., Haft D.H., Barrangou R., Brouns S.J., Charpentier E., Horvath P., Moineau S., Mojica F.J., Wolf Y.I., Yakunin A.F., et al. // Nat. Rev. Microbiol. 2011. V. 9. № 6. P. 467–477.
28. Mojica F.J., Diez-Villasenor C., Garcia-Martinez J., Almendros C. // Microbiology. 2009. V. 155. P. 733–740.
29. Swarts D.C., Mosterd C., van Passel M.W., Brouns S.J. // PLoS One. 2012. V. 7. № 4. e35888.
30. Hale C.R., Majumdar S., Elmore J., Pfister N., Compton M., Olson S., Resch A.M., Glover C.V., Graveley B.R., Tern R.M. // Mol. Cell. 2012. V. 45. № 3. P. 292–302.
31. Lillestol R.K., Shah S.A., Brugger K., Redder P., Phan H., Christiansen J., Garrett R.A. // Mol. Microbiol. 2009. V. 72. № 1. P. 259–272.
32. Carte J., Wang R., Li H., Terns R.M., Terns M.P. // Genes. Dev. 2008. V. 22. № 24. P. 3489–3496.
33. Haurwitz R.E., Jinek M., Wiedenheft B., Zhou K., Doudna J.A. // Science. 2010. V. 329. № 5997. P. 1355–1358.

REVIEWS

34. Gesner E.M., Schellenberg M.J., Garside E.L., George M.M., Macmillan A.M. // *Nat. Struct. Mol. Biol.* 2011. V. 18. № 6. P. 688–692.
35. Richter H., Zoepfel J., Schermuly J., Maticzka D., Backofen R., Randau L. // *Nucleic Acids Res.* 2012. V. 40. № 19. P. 9887–9896.
36. Sashital D.G., Jinek M., Doudna J.A. // *Nat. Struct. Mol. Biol.* 2011. V. 18. № 6. P. 680–687.
37. Hale C.R., Zhao P., Olson S., Duff M.O., Graveley B.R., Wells L., Terns R.M., Terns M.P. // *Cell.* 2009. V. 139. № 5. P. 945–956.
38. Makarova K.S., Grishin N.V., Shabalina S.A., Wolf Y.I., Koonin E.V. // *Biol. Direct.* 2006. V. 1. P. 7.
39. Marraffini L.A., Sontheimer E.J. // *Nature.* 2010. V. 463. № 7280. P. 568–571.
40. Bondy-Denomy J., Pawluk A., Maxwell K.L., Davidson A.R. // *Nature.* 2013. V. 493. № 7432. P. 429–432.
41. Makarova K.S., Aravind L., Wolf Y.I., Koonin E.V. // *Biol. Direct.* 2011. V. 6. P. 38.
42. Sapranaukas R., Gasiunas G., Fremaux C., Barrangou R., Horvath P., Siksnys V. // *Nucleic Acids Res.* 2011. V. 39. № 21. P. 9275–9282.
43. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. // *Science.* 2012. V. 337. № 6096. P. 816–821.
44. Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E. // *Nature.* 2011. V. 471. № 7340. P. 602–607.
45. Mussolino C., Cathomen T. // *Curr. Opin. Biotechnol.* 2012. V. 23. № 5. P. 644–650.
46. Cong L., Ran F.A., Cox D., Lin S., Barretto R., Habib N., Hsu P.D., Wu X., Jiang W., Marraffini L.A., et al. // *Science.* 2013. V. 339. № 6121. P. 819–823.
47. Hsu P.D., Scott D.A., Weinstein J.A., Ran F.A., Konermann S., Agarwala V., Li Y., Fine E.J., Wu X., Shalem O., et al. // *Nat. Biotechnol.* 2013. V. 31. № 9. P. 827–832.
48. Jiang W., Bikard D., Cox D., Zhang F., Marraffini L.A. // *Nat. Biotechnol.* 2013. V. 31. № 3. P. 233–239.
49. Fu Y., Foden J.A., Khayter C., Maeder M.L., Reyon D., Joung J.K., Sander J.D. // *Nat. Biotechnol.* 2013. V. 31. № 9. P. 822–826.
50. Mali P., Aach J., Stranges P.B., Esvelt K.M., Moosburner M., Kosuri S., Yang L., Church G.M. // *Nat. Biotechnol.* 2013. V. 31. № 9. P. 833–838.
51. Ran F.A., Hsu P.D., Lin C., Gootenberg J.S., Konermann S., Trevino A.E., Scott D.A., Inoue A., Matoba S., Zhang Y. // *Cell.* 2013. V. 154. № 6. P. 1380–1389.
52. Gasiunas G., Barrangou R., Horvath P., Siksnys V. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 39. P. 2579–2586.
53. Dianov G.L., Hubscher U. // *Nucleic Acids Res.* 2013. V. 41. № 6. P. 3483–3490.
54. Doyle E.L., Booher N.J., Standage D.S., Voytas D.F., Brendel V.P., Vandyk J.K., Bogdanove A.J. // *Nucleic Acids Res.* 2012. V. 40. P. W117–122.
55. Fine E.J., Cradick T.J., Zhao C.L., Lin Y., Bao G. // *Nucleic Acids Res.* 2013. V. 42. № 6. e42.
56. Grau J., Boch J., Posch S. // *Bioinformatics.* 2013. V. 29. № 22. P. 2931–2932.
57. Grau J., Wolf A., Reschke M., Bonas U., Posch S. // *PLoS Comput. Biol.* 2013. V. 9. № 3. e1002962.
58. Heigwer F., Kerr G., Walther N., Glaeser K., Pelz O., Breinig M., Boutros M. // *Nucleic Acids Res.* 2013. V. 41. № 20. e190.
59. Neff K.L., Argue D.P., Ma A.C., Lee H.B., Clark K.J., Ecker S.C. // *BMC Bioinformatics.* 2013. V. 14. P. 1.
60. Grissa I., Vergnaud G., Pourcel C. // *Nucleic Acids Res.* 2007. V. 35. P. W52–57.
61. Rousseau C., Gonnet M., Le Romancer M., Nicolas J. // *Bioinformatics.* 2009. V. 25. № 24. P. 3317–3318.
62. Xiao A., Cheng Z., Kong L., Zhu Z., Lin S., Gao G., Zhang B. // *Bioinformatics.* 2014. V. 30. № 8. P. 1180–1182.
63. Sander J.D., Cade L., Khayter C., Reyon D., Peterson R.T., Joung J.K., Yeh J.R. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 697–698.
64. Reyon D., Khayter C., Regan M.R., Joung J.K., Sander J.D. // *Curr. Protoc. Mol. Biol.* 2012. C.12. U. 12.15.
65. Weber E., Engler C., Gruetzner R., Werner S., Marillonnet S. // *PLoS One.* 2011. V. 6. № 2. e16765.
66. Engler C., Gruetzner R., Kandzia R., Marillonnet S. // *PLoS One.* 2009. V. 4. № 5. e5553.
67. Zhang F., Cong L., Lodato S., Kosuri S., Church G.M., Arlotta P. // *Nat. Biotechnol.* 2011. V. 29. № 2. P. 149–153.
68. Sanjana N.E., Cong L., Zhou Y., Cunniff M.M., Feng G., Zhang F. // *Nat. Protoc.* 2012. V. 7. № 1. P. 171–192.
69. Schmid-Burgk J.L., Schmidt T., Kaiser V., Honing K., Hornung V. // *Nat. Biotechnol.* 2013. V. 31. № 1. P. 76–81.
70. Cermak T., Doyle E.L., Christian M., Wang L., Zhang Y., Schmidt C., Baller J.A., Somia N.V., Bogdanove A.J., Voytas D.F. // *Nucleic Acids Res.* 2011. V. 39. № 12. e82.
71. Ding Q., Lee Y.K., Schaefer E.A., Peters D.T., Veres A., Kim K., Kuperwasser N., Motola D.L., Meissner T.B., Hendriks W.T. // *Cell Stem Cell.* 2013. V. 12. № 2. P. 238–251.
72. Reyon D., Tsai S.Q., Khayter C., Foden J.A., Sander J.D., Joung J.K. // *Nat. Biotechnol.* 2012. V. 30. № 5. P. 460–465.
73. Briggs A.W., Rios X., Chari R., Yang L., Zhang F., Mali P., Church G.M. // *Nucleic Acids Res.* 2012. V. 40. № 15. e117.
74. Wang Z., Li J., Huang H., Wang G., Jiang M., Yin S., Sun C., Zhang H., Zhuang F., Xi J.J. // *Angew. Chem. Int. Ed. Engl.* 2012. V. 51. № 34. P. 8505–8508.
75. Jinek M., Doudna J.A. // *Nature.* 2009. V. 457. № 7228. P. 405–412.
76. Malone C.D., Hannon G.J. // *Cell.* 2009. V. 136. № 4. P. 656–668.
77. Meister G., Tuschl T. // *Nature.* 2004. V. 431. № 7006. P. 343–349.
78. Cho S.W., Kim S., Kim J.M., Kim J.S. // *Nat. Biotechnol.* 2013. V. 31. № 3. P. 230–232.
79. Jinek M., East A., Cheng A., Lin S., Ma E., Doudna J. // *Elife.* 2013. V. 2. e00471.
80. Mali P., Yang L., Esvelt K.M., Aach J., Guell M., DiCarlo J.E., Norville J.E., Church G.M. // *Science.* 2013. V. 339. № 6121. P. 823–826.
81. Chang N., Sun C., Gao L., Zhu D., Xu X., Zhu X., Xiong J.W., Xi J.J. // *Cell Res.* 2013. V. 23. № 4. P. 465–472.
82. Hwang W.Y., Fu Y., Reyon D., Maeder M.L., Tsai S.Q., Sander J.D., Peterson R.T., Yeh J.R., Joung J.K. // *Nat. Biotechnol.* 2013. V. 31. № 3. P. 227–229.
83. Wang H., Yang H., Shivalia S.C., Dawlaty M.M., Cheng W.A., Zhang F., Jaenisch R. // *Cell.* 2013. V. 153. № 4. P. 910–918.
84. Shalem O., Sanjana N.E., Hartenian E., Shi X., Scott D.A., Mikkelsen T.S., Heckl D., Ebert B.L., Root D.E., Doench J.G. // *Science.* 2014. V. 343. № 6166. P. 84–87.
85. Wang T., Wei J.J., Sabatini D.M., Lander E.S. // *Science.* 2014. V. 343. № 6166. P. 80–84.
86. Shan Q., Wang Y., Li J., Zhang Y., Chen K., Liang Z., Zhang K., Liu J., Xi J.J., Qiu J.L. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 686–688.
87. Mao Y., Zhang H., Xu N., Zhang B., Gao F., Zhu J.K. // *Mol. Plant.* 2013. V. 6. № 6. P. 2008–2011.
88. Li J.F., Norville J.E., Aach J., McCormack M., Zhang D., Bush J., Church G.M., Sheen J. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 688–691.

89. Nekrasov V., Staskawicz B., Weigel D., Jones J.D., Kammoun S. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 691–369.
90. Lieber M.R. // *Annu. Rev. Biochem.* 2010. V. 79. P. 181–211.
91. Ousterout D.G., Perez-Pinera P., Thakore P.I., Kabadi A.M., Brown M.T., Qin X., Fedrigo O., Mouly V., Tremblay J.P., Gersbach C.A. // *Mol. Ther.* 2013. V. 21. № 9. P. 1718–1726.
92. Kim H., Um E., Cho S.R., Jung C., Kim H., Kim J.S. // *Nat. Methods.* 2011. V. 8. № 11. P. 941–943.
93. Guschin D.Y., Waite A.J., Katibah G.E., Miller J.C., Holmes M.C., Rebar E.J. // *Methods Mol. Biol.* 2010. V. 649. P. 247–256.
94. Dahlem T.J., Hoshijima K., Jurynek M.J., Gunther D., Starker C.G., Locke A.S., Weis A.M., Voytas D.F., Grunwald D.J. // *PLoS Genet.* 2012. V. 8. № 8. e1002861.
95. Ota S., Hisano Y., Muraki M., Hoshijima K., Dahlem T.J., Grunwald D.J., Kawahara A. // *Genes Cells.* 2013. V. 18. № 6. P. 450–458.
96. Kim H.J., Lee H.J., Kim H., Cho S.W., Kim J.S. // *Genome Res.* 2009. V. 19. № 7. P. 1279–1288.
97. Hisano Y., Ota S., Arakawa K., Muraki M., Kono N., Oshita K., Sakuma T., Tomita M., Yamamoto T., Okada Y. // *Biol. Open.* 2013. V. 2. № 4. P. 363–367.
98. Yang L., Guell M., Byrne S., Yang J.L., De Los Angeles A., Mali P., Aach J., Kim-Kiselak C., Briggs A.W., Rios X. // *Nucleic Acids Res.* 2013. V. 41. № 19. P. 9049–9061.
99. Saleh-Gohari N., Helleday T. // *Nucleic Acids Res.* 2004. V. 32. № 12. P. 3683–3688.
100. Perez E.E., Wang J., Miller J.C., Jouvenot Y., Kim K.A., Liu O., Wang N., Lee G., Bartsevich V.V., Lee Y.L. // *Nat. Biotechnol.* 2008. V. 26. № 7. P. 808–816.
101. Cristea S., Freyvert Y., Santiago Y., Holmes M.C., Urnov F.D., Gregory P.D., Cost G.J. // *Biotechnol. Bioeng.* 2013. V. 110. № 3. P. 871–880.
102. Maresca M., Lin V.G., Guo N., Yang Y. // *Genome Res.* 2013. V. 23. № 3. P. 539–546.
103. Chen F., Pruett-Miller S.M., Huang Y., Gjoka M., Taunton J., Collinwood T.N., Frodin M., Davis G.D. // *Nat. Methods.* 2011. V. 8. № 9. P. 753–755.
104. Moehle E.A., Rock J.M., Lee Y.L., Jouvenot Y., DeKolver R.C., Gregory P.D., Urnov F.D., Holmes M.C. // *Proc. Natl. Acad. Sci. USA.* 2007. V. 104. № 9. P. 3055–3060.
105. Orlando S.J., Santiago Y., DeKolver R.C., Freyvert Y., Boydston E.A., Moehle E.A., Choi V.M., Gopalan S.M., Lou J.F., Li J. // *Nucleic Acids Res.* 2010. V. 38. № 15. e152.
106. Hu R., Wallace J., Dahlem T.J., Grunwald D.J., O'Connell R.M. // *PLoS One.* 2013. V. 8. № 5. e63074.
107. Bloom K., Ely A., Mussolino C., Cathomen T., Arbuthnot P. // *Mol. Ther.* 2013. V. 21. № 10. P. 1889–1897.
108. Schiffer J.T., Aubert M., Weber N.D., Mintzer E., Stone D., Jerome K.R. // *J. Virol.* 2012. V. 86. № 17. P. 8920–8936.
109. Holt N., Wang J., Kim K., Friedman G., Wang X., Taupin V., Crooks G.M., Kohn D.B., Gregory P.D., Holmes M.C. // *Nat. Biotechnol.* 2010. V. 28. № 8. P. 839–847.
110. Horii T., Tamura D., Morita S., Kimura M., Hatada I. // *Int. J. Mol. Sci.* 2013. V. 14. № 10. P. 19774–19781.
111. Schwank G., Koo B.K., Sasselli V., Dekkers J.F., Heo I., Demircan T., Sasaki N., Boymans S., Cuppen E., van der Ent C.K. // *Cell Stem Cell.* 2013. V. 13. № 6. P. 653–658.
112. Gilbert L.A., Larson M.H., Morsut L., Liu Z., Brar G.A., Torres S.E., Stern-Ginossar N., Brandman O., Whitehead E.H., Doudna J.A. // *Cell.* 2013. V. 154. № 2. P. 442–451.
113. Bikard D., Jiang W., Samai P., Hochschild A., Zhang F., Marraffini L.A. // *Nucleic Acids Res.* 2013. V. 41. № 15. P. 7429–7437.
114. Perez-Pinera P., Kocak D.D., Vockley C.M., Adler A.F., Kabadi A.M., Polstein L.R., Thakore C.A. // *Nat. Methods.* 2013. V. 10. № 10. P. 973–976.
115. Farzadfard F., Perli S.D., Lu T.K. // *ACS Synth. Biol.* 2013. V. 2. № 10. P. 604–613.
116. Heintze J., Luft C., Ketteler R. // *Front Genet.* 2013. V. 4. P. 193.
117. Chen B., Gilbert L.A., Cimini B.A., Schnitzbauer J., Zhang W., Li G.W., Park J., Blackburn E.H., Weissman J.S., Qi L.S., et al. // *Cell.* 2013. V. 155. № 7. P. 1479–1491.
118. Smith J.R., Maguire S., Davis L.A., Alexander M., Yang F., Chandran S., French-Constant C., Pedersen R.A. // *Stem Cells.* 2008. V. 26. № 2. P. 496–504.
119. Chiu H., Schwartz H.T., Antoshechkin I., Sternberg P.W. // *Genetics.* 2013. V. 195. № 3. P. 1167–1171.
120. Cho S.W., Lee J., Carroll D., Kim J., Lee J. // *Genetics.* 2013. V. 195. № 3. P. 1177–1180.
121. Katic I., Grosshans H. // *Genetics.* 2013. V. 195. № 3. P. 1173–1176.
122. Lo T.W., Pickle C.S., Lin S., Ralston E.J., Gurling M., Schartner C.M., Bian Q., Doudna J.A., Meyer B.J. // *Genetics.* 2013. V. 195. № 2. P. 331–348.
123. Tzur Y.B., Friedland A.E., Nadarajan S., Church G.M., Calarco J.A., Colaiacovo M.P. // *Genetics.* 2013. V. 195. № 3. P. 1181–1185.
124. Waaijers S., Portegijs V., Kerver J., Lemmens B.B., Tijsterman M., van den Heuvel S., Boxem M. // *Genetics.* 2013. V. 195. № 3. P. 11887–11891.
125. Dickinson D.J., Ward J.D., Reiner D.J., Goldstein B. // *Nat. Methods.* 2013. V. 10. № 10. P. 1028–1034.
126. Friedland A.E., Tzur Y.B., Esvelt K.M., Colaiacovo M.P., Church G.M., Calarco J.A. // *Nat. Methods.* 2013. V. 10. № 8. P. 741–743.
127. Maggert K.A., Gong W.J., Golic K.G. // *Methods Mol. Biol.* 2008. V. 420. P. 155–174.
128. Rong Y.S., Golic K.G. // *Science.* 2000. V. 288. № 5473. P. 2013–2018.
129. Venken K.J., Bellen H.J. // *Nat. Rev. Genet.* 2005. V. 6. № 3. P. 167–178.
130. Bassett A.R., Tibbit C., Ponting C.P., Liu J. // *Cell Rep.* 2013. V. 4. № 1. P. 220–228.
131. Gratz S.J., Cummings A.M., Nguyen J.N., Hamm D.C., Donohue L.K., Harrison M.M., Wildonger J., O'Connor-Giles K.M. // *Genetics.* 2013. V. 194. № 4. P. 1029–1035.
132. Yu Z., Ren M., Wang Z., Zhang B., Rong Y.S., Jiao R., Gao G. // *Genetics.* 2013. V. 195. № 1. P. 289–291.
133. Ren X., Sun J., Housden B.E., Hu Y., Roesel C., Liu L.P., Yang Z., Mao D., Sun L., Wu Q., et al. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 47. P. 19012–19017.
134. Seth A., Stemple D.L., Barroso I. // *Dis. Model. Mech.* 2013. V. 6. № 5. P. 1080–1088.
135. Copeland N.G., Jenkins N.A. // *Nat. Rev. Cancer.* 2010. V. 10. № 10. P. 696–706.
136. Kool J., Berns A. // *Nat. Rev. Cancer.* 2009. V. 9. № 6. P. 389–399.
137. Bogdanove A.J., Voytas D.F. // *Science.* 2011. V. 333. № 6051. P. 1843–1846.
138. Carroll D., Beumer K.J., Morton J.J., Bozas A., Trautman J.K. // *Methods Mol. Biol.* 2008. V. 435. P. 63–77.
139. Urnov F.D., Rebar E.J., Holmes M.C., Zhang H.S., Gregory P.D. // *Nat. Rev. Genet.* 2010. V. 11. № 9. P. 636–646.
140. Carbery I.D., Ji D., Harrington A., Brown V., Weinstein E.J., Liaw L., Cui X. // *Genetics.* 2010. V. 186. № 2. P. 451–459.
141. Geurts A.M., Cost G.J., Freyvert Y., Zeitler B., Miller J.C., Choi V.M., Jenkins S.S., Wood A., Cui X., Meng X., et al.

- // Science. 2009. V. 325. № 5939. P. 433.
142. Cui X., Ji D., Fisher D.A., Wu Y., Briner D.M., Weinstein E.J. // *Nat. Biotechnol.* 2011. V. 29. № 1. P. 64–67.
143. Meyer M., de Angelis M.H., Wurst W., Kuhn R. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. № 34. P. 15022–15026.
144. Li D., Qiu Z., Shao Y., Guan Y., Liu M., Li Y., Gao N., Wang L., Lu X., Zao Y., et al. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 681–683.
145. Li W., Teng F., Li T., Zhou Q. // *Nat. Biotechnol.* 2013. V. 31. № 8. P. 684–686.
146. Wu Y., Liang D., Wang Y., Bai M., Tang W., Bao S., Yan Z., Li D., Li J. // *Cell Stem Cell.* 2013. V. 13. № 6. P. 659–662.
147. Chen K., Gao C. // *J. Genet. Genomics.* 2013. V. 40. № 6. P. 271–279.
148. van der Hoorn R.A., Laurent F., Roth R., De Wit P.J. // *Mol. Plant. Microbe Interact.* 2000. V. 13. № 4. P. 439–446.
149. Feng Z., Zhang B., Ding W., Liu X., Yang D., Wei P., Cao F., Zhu S., Zhang F., Mao Y., et al. // *Cell Res.* 2013. V. 23. № 10. P. 1229–1232.
150. Xie K., Yang Y. // *Mol. Plant.* 2013. V. 6. № 6. P. 1975–1983.
151. Miao J., Guo D., Zhang J., Huang Q., Qin G., Zhang X., Wan J., Gu H., Qu L. // *Cell Res.* 2013. V. 23. № 10. P. 1233–1236.
152. Jiang W., Zhou H., Bi H., Fromm M., Yang B., Weeks D.P. // *Nucleic Acids Res.* 2013. V. 41. № 20. e188.
153. Upadhyay S.K., Kumar J., Alok A., Tuli R. // *G3 (Bethesda)*. 2013. V. 3. № 12. P. 2233–2238.
154. Shan Q., Wang Y., Chen K., Liang Z., Li J., Zhang Y., Zhang K., Liu J., Voytas D.F., Zheng X. // *Mol. Plant.* 2013. V. 6. № 4. P. 1365–1368.
155. Li T., Liu B., Spalding M.H., Weeks D.P., Yang B. // *Nat. Biotechnol.* 2012. V. 30. № 5. P. 390–392.
156. Zhang Y., Zhang F., Li X., Baller J.A., Starker C.G., Bogdanove A.J., Voytas D.F. // *Plant Physiol.* 2013. V. 161. № 1. P. 20–27.
157. Mahfouz M.M., Li L., Shamimuzzaman M., Wibowo A., Fang X., Zhu J.K. // *Proc. Natl. Acad. Sci. USA.* 2011. V. 108. № 6. P. 2623–2628.
158. Beerli R.R., Segal D.J., Dreier B., Barbas C.F. 3rd. // *Proc. Natl. Acad. Sci. USA.* 1998. V. 95. № 25. P. 14628–14633.
159. Perez-Pinera P., Ousterout D.G., Brunger J.M., Farin A.M., Guilak F., Crawford G.E., Hartemink A.J., Gersbach C.A. // *Nat. Methods.* 2013. V. 10. № 3. P. 239–242.
160. Maeder M.L., Linder S.J., Reyon D., Angstman J.F., Fu Y., Sander J.D., Joung J.K. // *Nat. Methods.* 2013. V. 10. № 3. P. 243–245.
161. Li Y., Moore R., Guinn M., Bleris L. // *Sci. Rep.* 2012. V. 2. P. 897.
162. Mahfouz M.M., Li L., Piatek M., Fang X., Mansour H., Bangarusamy D.K., Zhu J.K. // *Plant Mol. Biol.* 2012. V. 78. № 3. P. 311–321.
163. Chapdelaine P., Coulombe Z., Chikh A., Gerard C., Tremblay J.P. // *Mol. Ther. Nucleic Acids.* 2013. V. 2. e119.
164. Gao X., Yang J., Tsang J.C., Ooi J., Wu D., Liu P. // *Stem Cell Reports.* 2013. V. 1. № 2. P. 183–197.
165. Mercer A.C., Gaj T., Sirk S.J., Lamb B.M., Barbas C.F. 3rd. // *ACS Synth. Biol.* 2013.
166. Koneremann S., Brigham M.D., Trevino A.E., Hsu P.D., Heidenreich M., Cong L., Platt R.J., Scott D.A., Church G.M., Zhang F. // *Nature.* 2013. V. 500. № 7463. P. 472–476.
167. Gaber R., Lebar T., Majerle A., Ster B., Dobnikar A., Bencina M., Jerala R. // *Nat. Chem. Biol.* 2014. V. 10. № 3. P. 203–208.
168. Lienert F., Torella J.P., Chen J.H., Norsworthy M., Richardson R.R., Silver P.A. // *Nucleic Acids Res.* 2013. V. 41. № 21. P. 9967–9975.
169. Miyanari Y., Ziegler-Birling C., Torres-Padilla M.E. // *Nat. Struct. Mol. Biol.* 2013. V. 20. № 11. P. 1321–1324.
170. Thanisch K., Schneider K., Morbitzer R., Solovei I., Lahaye T., Bultmann S., Leonhardt H. // *Nucleic Acids Res.* 2013. V. 42. № 6. e38.
171. Ma H., Reyes-Gutierrez P., Pederson T. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 52. P. 21048–21053.
172. Owens J.B., Urschitz J., Stoytchev I., Dang N.C., Stoytcheva Z., Belcaid M., Maragathavally K.J., Coates C.J., Segal D.J., Moisyadi S. // *Nucleic Acids Res.* 2012. V. 40. № 14. P. 6978–6991.
173. Mercer A.C., Gaj T., Fuller R.P., Barbas C.F. 3rd. // *Nucleic Acids Res.* 2012. V. 40. № 21. P. 11163–11172.
174. Lamb B.M., Mercer A.C., Barbas C.F. 3rd. // *Nucleic Acids Res.* 2013. V. 41. № 21. P. 9779–9785.
175. Owens J.B., Mauro D., Stoytchev I., Bhakta M.S., Kim M.S., Segal D.J., Moisyadi S. // *Nucleic Acids Res.* 2013. V. 41. № 19. P. 9197–9207.
176. Miller J.C., Tan S., Qiao G., Barlow K.A., Wang J., Xia D.F., Meng X., Paschon D.E., Leung E., Hinkley S.J., et al. // *Nat. Biotechnol.* 2011. V. 29. № 2. P. 143–148.
177. Mussolino C., Morbitzer R., Lutge F., Dannemann N., Lahaye T., Cathomen T. // *Nucleic Acids Res.* 2011. V. 39. № 21. P. 9283–9293.
178. Sun N., Liang J., Abil Z., Zhao H. // *Mol. Biosyst.* 2012. V. 8. № 4. P. 1255–1263.
179. Xu L., Zhao P., Mariano A., Han R. // *Mol. Ther. Nucleic Acids.* 2013. V. 2. e112.
180. Hockemeyer D., Wang H., Kiani S., Lai C.S., Gao Q., Casady J.P., Cost G.J., Zhang L., Santiago Y., Miller J.C., et al. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 731–734.
181. Li T., Huang S., Zhao X., Wright D.A., Carpenter S., Spalding M.H., Weeks D.P., Yang B. // *Nucleic Acids Res.* 2011. V. 39. № 14. P. 6315–6325.
182. Wood A.J., Lo T.W., Zeitler B., Pickle C.S., Ralston E.J., Lee A.H., Amora R., Miller J.C., Leung E., Meng X., et al. // *Science.* 2011. V. 333. № 6040. P. 307.
183. Liu J., Li C., Yu Z., Huang P., Wu H., Wei C., Zhu N., Chen Y., Zhang B., Deng W.M., et al. // *J. Genet. Genomics.* 2012. V. 39. № 5. P. 209–215.
184. Bedell V.M., Wang Y., Campbell J.M., Poshusta T.L., Starker C.G., Krug R.G. 2nd, Tan W., Penheiter S.G., Ma A.C., Leung A.Y., et al. // *Nature.* 2012. V. 491. № 7422. P. 114–118.
185. Moore F.E., Reyon D., Sander J.D., Martinez S.A., Blackburn J.S., Khayter C., Ramirez C.L., Joung J.K., Langenau D.M. // *PLoS One.* 2012. P. 7. № 5. e37877.
186. Huang P., Xiao A., Zhou M., Zhu Z., Lin S., Zhang B. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 699–700.
187. Cade L., Reyon D., Hwang W.Y., Tsai S.Q., Patel S., Khayter C., Joung J.K., Sander J.D., Peterson R.T., Yeh J.R. // *Nucleic Acids Res.* 2012. V. 40. № 16. P. 8001–8010.
188. Ma S., Zhang S., Wang F., Liu Y., Liu Y., Xu H., Liu C., Lin Y., Zhao P., Xia Q. // *PLoS One.* 2012. V. 7. № 9. e45035.
189. Watanabe T., Ochiai H., Sakuma T., Horch H.W., Hamaguchi N., Nakamura T., Bando T., Ohuchi H., Yamamoto T., Noji S., et al. // *Nat. Commun.* 2012. V. 3. P. 1017.
190. Lei Y., Guo X., Liu Y., Cao Y., Deng Y., Chen X., Cheng X., Cheng C.H., Dawid I.B., Chen Y., et al. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 43. P. 17484–17489.
191. Panda S.K., Wefers B., Ortiz O., Floss T., Schmid B., Haass C., Wurst W., Kuhn R. // *Genetics.* 2013. V. 195. № 3. P. 703–713.

REVIEWS

192. Qiu Z., Liu M., Chen Z., Shao Y., Pan H., Wei G., Yu C., Zhang L., Li X., Wang P., et al. // *Nucleic Acids Res.* 2013. V. 41. № 11. e120.
193. Wu J., Huang Z., Ren J., Zhang Z., He P., Li Y., Ma J., Chen W., Zhang Y., Zhou X., et al. // *Cell Res.* 2013. V. 23. № 8. P. 994–1006.
194. Sung Y.H., Baek I.J., Kim D.H., Jeon J., Lee J., Lee K., Jeong D., Kim J.S., Lee H.W. // *Nat. Biotechnol.* 2013. V. 31. № 1. P. 23–24.
195. Wefers B., Meyer M., Ortiz O., Hrabe de Angelis M., Hansen J., Wurst W., Kuhn R. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 10. P. 3782–3787.
196. Davies B., Davies G., Preece C., Puliyadi R., Szumska D., Bhattacharya S. // *PLoS One.* 2013. V. 8. № 3. e60216.
197. Tong C., Huang G., Ashton C., Wu H., Yan H., Ying Q.L. // *J. Genet. Genomics.* 2012. V. 39. № 6. P. 275–280.
198. Tesson L., Usal C., Menoret S., Leung E., Niles B.J., Remy S., Santiago Y., Vincent A.I., Meng X., Zhang L., et al. // *Nat. Biotechnol.* 2011. V. 29. № 8. P. 695–696.
199. Carlson D.F., Tan W., Lillico S.G., Stverakova D., Proudfoot C., Christian M., Voytas D.F., Long C.R., Whitelaw C.B., Fahrenkrug S.C. // *Proc. Natl. Acad. Sci. USA.* 2012. V. 109. № 43. P. 17382–17387.
200. DiCarlo J.E., Norville J.E., Mali P., Rios X., Aach J., Church G.M. // *Nucleic Acids Res.* 2013. V. 41. № 7. P. 4336–4343.
201. Hou Z., Zhang Y., Propson N.E., Howden S.E., Chu L., Sontheimer E.J., Thomson J.A. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 39. P. 15644–15649.
202. Ding Q., Regan S.N., Xia Y., Oostrom L.A., Cowan C.A., Musunuru K. // *Cell Stem Cell.* 2013. V. 12. № 4. P. 393–394.
203. Hwang W.Y., Fu Y., Reyon D., Maeder M.L., Kaini P., Sander J.D., Joung J.K., Peterson R.T., Yeh J.J. // *PLoS One.* 2013. V. 8. № 7. e68708.
204. Jao L.E., Wente S.R., Chen W. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 34. P. 13904–13909.
205. Xiao A., Wang Z., Hu Y., Wu Y., Luo Z., Yang Z., Zu Y., Li W., Huang P., Tong X., et al. // *Nucleic Acids Res.* 2013. V. 41. № 14. e141.
206. Hruscha A., Krawitz P., Rechenberg A., Heinrich V., Hecht J., Haass C., Schmid B. // *Development.* 2013. V. 140. № 24. P. 4982–4987.
207. Nakayama T., Fish M.B., Fisher M., Oomen-Hajagos J., Thomsen G.H., Grainger R.M. // *Genesis.* 2013. V. 51. № 12. P. 835–843.
208. Tan W., Carlson D.F., Lancto C.A., Garbe J.R., Webster D.A., Hackett P.B., Fahrenkrug S.C. // *Proc. Natl. Acad. Sci. USA.* 2013. V. 110. № 41. P. 16526–16531.
209. Fujii W., Kawasaki K., Sugiura K., Naito K. // *Nucleic Acids Res.* 2013. V. 41. № 20. e187.
210. Shen B., Zhang J., Wu H., Wang J., Ma K., Li Z., Zhang X., Zhang P., Huang X. // *Cell Res.* 2013. V. 23. № 5. P. 720–723.
211. Ma Y., Zhang X., Shen B., Lu Y., Chen W., Ma J., Bai L., Huang X., Zhang L. // *Cell Res.* 2014. V. 24. № 1. P. 122–125.

Mycoplasma Contamination of Cell Cultures: Vesicular Traffic in Bacteria and Control over Infectious Agents

V. M. Chernov^{1,2}, O. A. Chernova^{1,2}, J. T. Sanchez-Vega³, A. I. Kolpakov^{2*}, O. N. Ilinskaya²

¹Kazan Institute of Biochemistry and Biophysics, Kazan Scientific Center, Russian Academy of Sciences, Lobachevskogo Str., 2/3, 1420111, Kazan, Russia

²Kazan (Volga Region) Federal University, Kremlyovskaya Str., 18, 420008, Kazan, Russia

³National Autonomous University of Mexico, Coyoacán, 04510, Mexico

E-mail: ljoscha@mail.ru

Received 16.04.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT Cell cultures are subject to contamination either with cells of other cultures or with microorganisms, including fungi, viruses, and bacteria. Mycoplasma contamination of cell cultures is of particular importance. Since cell cultures are used for the production of vaccines and physiologically active compounds, designing a system for controlling contaminants becomes topical for fundamental science and biotechnological production. The discovery of extracellular membrane vesicles in mycoplasmas makes it necessary to take into consideration the bacterial vesicular traffic in systems designed for controlling infectious agents. The extracellular vesicles of bacteria mediate the traffic of proteins and genes, participate in cell-to-cell interactions, as well as in the pathogenesis and development of resistance to antibiotics. The present review discusses the features of mycoplasmas, their extracellular vesicles, and the interaction between contaminants and eukaryotic cells. Furthermore, it provides an analysis of the problems associated with modern methods of diagnosis and eradication of mycoplasma contamination from cell cultures and prospects for their solution.

KEYWORDS diagnosis and eradication, cell cultures, mycoplasma contamination.

INTRODUCTION

With the use of cell cultures expanding in fundamental and practical studies, it is utterly important to elaborate a system for rigorous testing of any contamination of the material. Working with cell cultures always presents a risk of contamination either with eukaryotic cells from other cultures or with microorganisms, including fungi, viruses and bacteria. Mycoplasma contamination is of particular preoccupation as it does not manifest itself conspicuously [1–3].

In 1956, for the purpose of investigating the effects of mycoplasma on eukaryotic cells, Robinson *et al.* infected cell cultures with these microorganisms. They found that the original cell culture had already been contaminated with mycoplasma. This was the first report on the detection of mycoplasma in cell cultures [4]. Subsequently, it became clear that mycoplasma contamination is the scourge of cell cultures. It turns out that all cell cultures originating from various eukaryotic organisms (mammals, birds, reptiles, fishes, insects and plants) are subject to mycoplasma contamination. Experimental studies conducted in various countries have shown the mycoplasma infection rate among cul-

tures in different laboratories to vary from 15% to 80% and, in some, to even reach 100% [3, 5].

Mycoplasma is an umbrella term for representatives of the Mollicutes class, the smallest bacteria lacking a cell wall and capable of self-reproduction. The small genome size limits the biosynthetic abilities of these microorganisms and defines their parasitic way of life. The great attention to mycoplasma is nowadays dictated, on the one hand, by the study of the molecular patterns of minimal cellular function and, on the other hand, by practical necessity. Mycoplasmas parasitize humans, animals, and plants, where some of them are agents of socially significant diseases, and the main contaminants of cell cultures and vaccines. Control over mycoplasma infection is a serious problem, the solution to which can probably be found in the molecular mechanisms of adaptation that allow mycoplasma to survive under various conditions and to overcome the protection barrier of higher eukaryotes and their persistence [1–3, 6–8].

MYCOPLASMAS ARE THE MAIN CONTAMINANT OF CELL CULTURES

The significant amount of theoretical and practical data accumulated recently has dramatically changed our no-

tion of mycoplasma pathogenicity. It has become clear that bacteria have elaborated sophisticated mechanisms to survive under severe conditions and remain virulent [9–18], whereas the conditions of *in vitro* cultivation of eukaryotic cells favor mycoplasma growth [13, 19]. Together with cells from the original organisms, whose tissues are used to create an *in vitro* culture, researchers themselves, as well as components of the medium and laboratory facilities, can act as a source of mycoplasma contamination. In this context, all representatives of Mollicutes are considered to be potential contaminants of cell cultures. At the moment, there are almost as many as 30 types of mycoplasmas that have been identified in cell cultures, whereas 95% of cases are caused by the following 6 mycoplasmas: *Mycoplasma arginini*, *M. fermentans*, *M. hominis*, *M. hyorhina*, *M. orale* and *Acholeplasma laidlawii* [2, 3]. This knowledge allows one to assume that these bacteria possess special features that define their prevalence in their ecological niche, and, consequently, that contamination can be controlled through the adaptation mechanisms of mycoplasmas.

A. laidlawii is a mycoplasma species that appears to have unique adaptation abilities. This widely spread

type is the agent of phytomycoplasmosis [1, 20, 21]. Although it is present in humans and animals in various pathological processes, there has been no reliable evidence of its pathogenicity so far [1, 3, 5]. Mapping of the *A. laidlawii* genome carried out in Russia [22] have made it possible to establish the adaptation mechanisms of this mycoplasma using post-genomic technologies. Genomic, transcript, and proteomic profiling, along with the nanoscopic analysis, have allowed researchers to identify the stress-reactive proteins and genes of *A. laidlawii*. It has been demonstrated that the mechanisms of mycoplasma survival under severe conditions, as well as the mechanisms of formation of host-parasite relationships and virulence, are connected to the secretion of extracellular vesicles by this bacteria [16, 20, 21, 23, 24].

Extracellular membrane vesicles mediate the common secretion mechanism in prokaryotes and eukaryotes and constitute an important part of the bacterial secretome [25]. Along with the membrane components, they may contain cytoplasmic proteins, toxins, DNA, and RNA [26, 27]. Discovered in gram-negative bacteria several decades ago, extracellular vesicles were

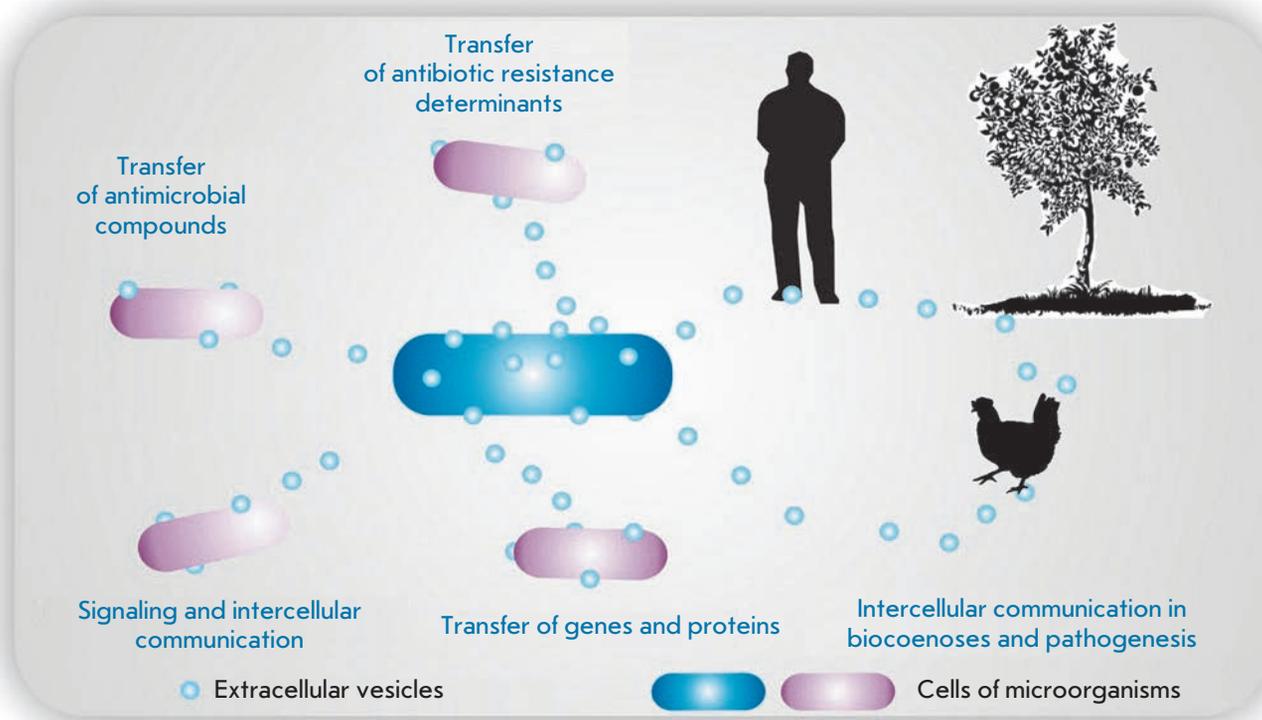


Fig. 1. Extracellular vesicles mediate the traffic of a broad range of components, transfer of virulence determinants, and development of resistance to antimicrobial agents; they participate in signaling, intercellular communication, and pathogenesis [41]

recently been found in archaea [28], gram-positive bacteria [29], and in the smallest wall-less prokaryotes; namely mycoplasma [16, 24]. Vesicles were shown to play an important role in cell-to-cell communication as carriers of essential cell-specific information [25, 30–32]. The internalization of these nanostructures triggers cell-target reprogramming, which can be detected by proteomic and transcript analyses [33, 34]. Bacteria-secreted extracellular vesicles mediate the protein traffic and transfer of virulence determinants, participate in the formation of the host–parasite system and that of the resistance to antibacterials and, respectively, in the adaptation to different environmental conditions (*Fig. 1*) [25, 27]. In accordance with the virulence criteria, the extracellular vesicles of pathogenic bacteria belong to a new type of infectious agents, which makes it necessary to adjust current approaches to the control of bacterial infections [21, 31, 35].

A. laidlawii cells have been shown to secrete vesicles (20–120 in diameter) into the intracellular space under different growth conditions; however, the vesicle generation rate considerably increases under stress (*Fig. 2*). Vesicles determine such virulent properties of mycoplasma as infectivity, invasiveness, and toxigenicity; they also induce the clastogenic effect in eukaryotic cells *in vitro* (*Fig. 3*). Vesicle penetration precedes mycoplasma invasion of plant tissues, destroys their ultrastructure, induces modulation of gene expression and protein synthesis in infected organisms, and mediates the development of mycoplasma resistance to antibacterials [16, 20, 21, 24, 36]. Global proteomic profiling has allowed researchers to “make an inventory” of the proteins of *A. laidlawii* extracellular vesicles PG8 secreted in an axenic culture [37]. It turns out that most polypeptides exported from mycoplasma cells with vesicles are virulence factors including adhesins, enzymes of a protein, polysaccharide, and nucleic acid degradation (*Fig. 4*).

In addition to membrane components and cytoplasmic proteins, the extracellular vesicles of *A. laidlawii* PG8 contain a specific set of nucleotide sequences that can be used as markers of bacterial vesicles in analyzed species [20, 24, 36]. Similar data on the structure and composition of extracellular vesicles were obtained for *M. gallisepticum* (*Fig. 2*), a widespread agent of avian diseases and the main contaminant of viral chick embryo vaccines [24]. The results indicate that vesicular traffic associated with extracellular membrane vesicles in archaea, classic gram-positive, and gram-negative bacteria was also found in the smallest wall-less prokaryotes. This fact makes it necessary to reconsider our understanding of the interaction between the mycoplasma and the cells of higher organisms and to design a strategy for controlling infectious agents.

MYCOPLASMA CONTAMINATION CONTROL

Since mycoplasmas do not have a rigid cell wall, close contact between the cytoplasmic membrane of the host and that of the parasite is possible; under certain conditions, this may cause cell fusion [1, 38]. Some mycoplasmas have specific organelles at their poles (the so-called tips or blebs) that mediate gliding motility and adhesion between bacteria and the eukaryotic cell membrane [1, 39]. Adhesion can be accompanied by invasion of the cell [3]. However, even when staying on the surface and thus in close contact with the host cell membrane, mycoplasmas induce modulation of the genome expression and cause considerable changes in the metabolism in eukaryotic cells [3, 38]. A series of studies aimed at determining the patterns of transcription profile modulation in cell cultures upon mycoplasma contamination show that the latter triggers changes in the expression of a broad range of genes in the host cell (*Table 1*). The genes whose expression changes include a significant portion of the most important ones encoding regulatory proteins, such as oncogenes, tumor suppressor genes [40], cytokines [41], receptors, and components of signaling pathways [42]. Changes in the expression may become overt as soon as several hours after inoculation [42], whereas prolonged cultivation of inoculated cells (18 weeks) may lead to their irreversible transformation to the extent of malignant degeneration [40]. The nature of transcript profile modulation in inoculated cells varies substantially depending on the mycoplasma type, cell culture type, multiplicity of infection, and cultivation period. Thus, contamination with mycoplasmas makes it impossible to adequately evaluate the results obtained using an inoculated culture. In particular, the effect of compounds suggested as promising pharmaceutical agents cannot be studied.

Despite the fact the hundreds of genes whose expression changes upon contamination of eukaryotic cells with mycoplasma have been identified [41–45], no common markers of mycoplasma contamination have been found. Mycoplasmas may trigger the activation of macrophages cultivated *in vitro*, suppression of antigen presentation, modification of the immune reactivity, signal transduction, viral proliferation, and apoptosis [40, 46–54]. Mycoplasma contamination may remain unnoticed for a rather long time; visible changes appear only at high multiplicity of the infection [1, 3]. The most serious effect of contamination is the loss of the cell culture due to the growth of microorganisms and, respectively, the irreversible worsening of the condition of the cells. Depending on the mycoplasma species, cell line and cultivation conditions, one may observe various cytopathic reactions, including, for instance, chromatin condensation, leopard cells, chromosome aberrations, suppression of cell division, and deprivation

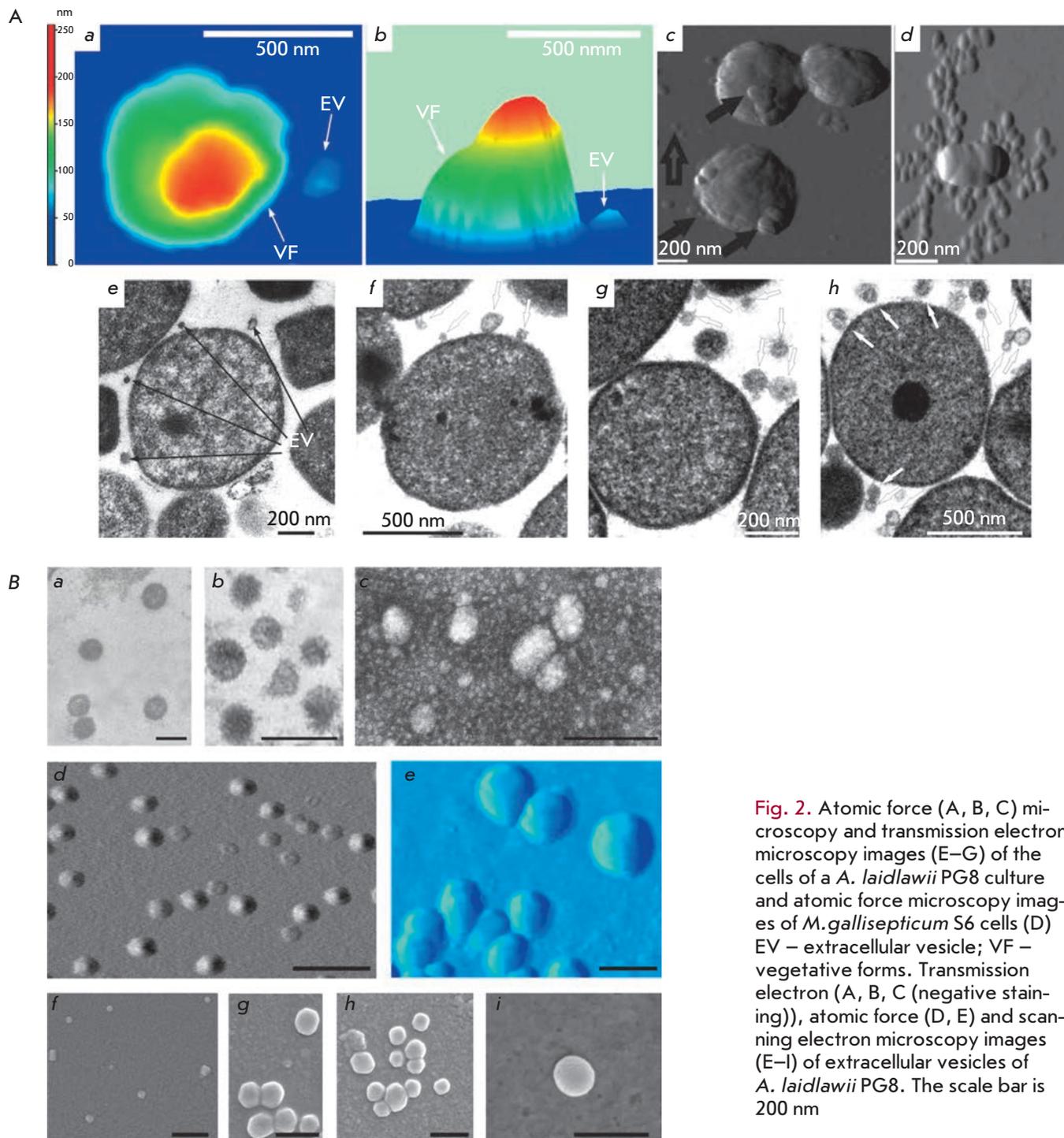


Fig. 2. Atomic force (A, B, C) microscopy and transmission electron microscopy images (E–G) of the cells of a *A. laidlawii* PG8 culture and atomic force microscopy images of *M. gallisepticum* S6 cells (D) EV – extracellular vesicle; VF – vegetative forms. Transmission electron (A, B, C (negative staining)), atomic force (D, E) and scanning electron microscopy images (E–I) of extracellular vesicles of *A. laidlawii* PG8. The scale bar is 200 nm

of cell culture growth [3, 5]. The main reason for these reactions is mycoplasma interference with cell metabolism, competitive absorption of nutrients and release of bacterial toxins, enzymes of protein, and DNA and RNA degradation [1, 38]. The extracellular vesicles of mycoplasma may actively participate in these processes. We have demonstrated in a series of special ex-

periments that the RNA activity of *A. laidlawii* PG8 and that of *M. hominis* PQ37 account for 86% and 89%, respectively, of the overall activity of the cellular and extracellular RNases of these bacteria [55]. The ribonucleic activity of the secreted vesicles may to a large extent determine the genotoxic properties of these contaminants revealed earlier [56–58]. Taking into account

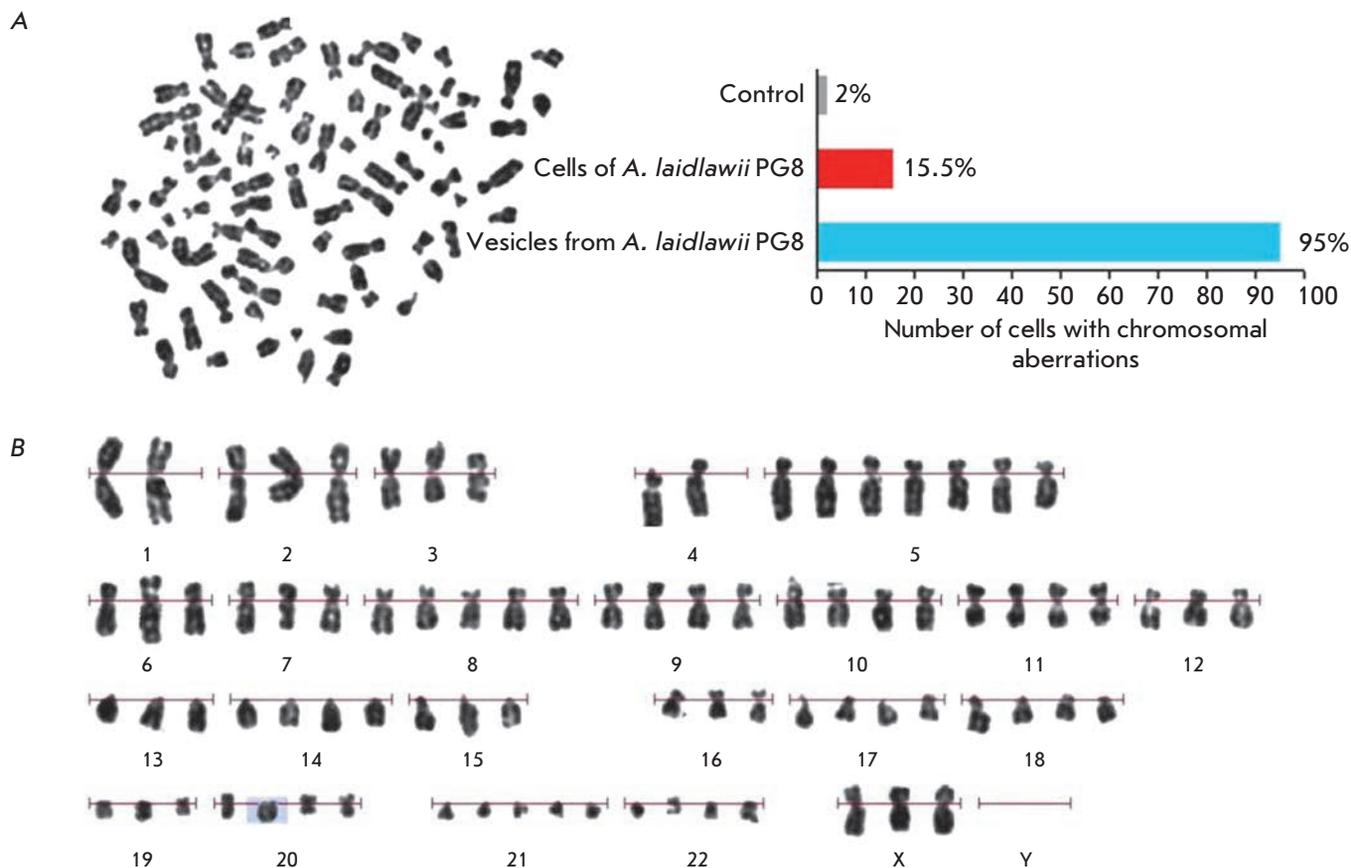


Fig. 3. Metaphase plate (A) and karyogram (B) of human peripheral blood lymphocytes after the cells were incubated with vesicles of *A. laidlawii* PG8

Table 1. Change of mRNA expression of a number of genes in cells inoculated with mycoplasma in 3-7 days after contamination

Mycoplasma	Cell culture	Induction of mRNA expression	Suppression of mRNA expression	References
<i>M. fermentans</i>	Epithelial cells of prostate HPV E7	14 cytokines	TGFβ1, TGFβ3	[41]
<i>M. genitalium</i>		12 cytokines	GM-CSF, IL-1Ra, M-CSF	
<i>M. hominis</i>		12 cytokines	TGFβ2	
<i>M. penetrans</i>		14 cytokines	TGFβ2	
<i>M. fermentans</i>	Epithelial cells of cervical canal HPV E6	17 cytokines	0	[41]
<i>M. genitalium</i>		13 cytokines	G-CSF, IL-1Ra	
<i>M. hominis</i>		13 cytokines	IL-1α, IL-1β	
<i>M. penetrans</i>		15 cytokines	TGFβ2, TGF-β3	
<i>M. synoviae</i>	Chicken macrophages MDM	Cytokines, lysozyme, apoptosis inhibitor, 11 enzymes, 4 types of receptors, 10 proteins of the signaling system	ovotransferrin, glutathione S-transferase, guanylate-binding protein	[42]
<i>M. fermentans incognitas</i>	Mice embryoblast C3H	92 genes encoding oncogenes and tumor suppressors	43 genes encoding oncogenes and tumor suppressors	[40]
<i>Phytoplasma</i>	Paulownia culture	769 genes	437 genes	[45]

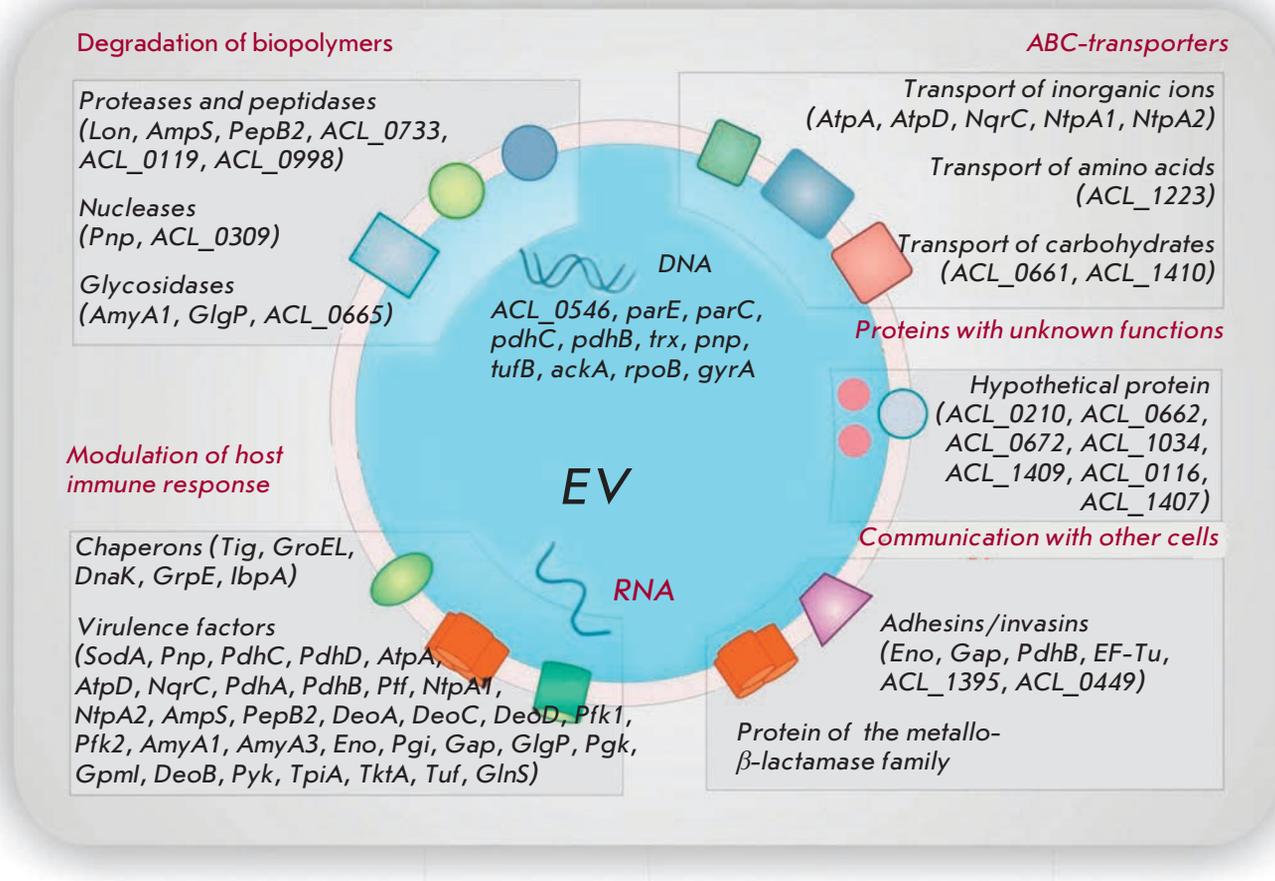


Fig. 4. Extracellular vesicles of *A. laidlawii* PG8 contain a specific set of DNA and RNA nucleotide sequences

the cytotoxic potential of numerous bacterial RNases [59–61], one may assume that the cytopathic reactions of contaminated cell cultures are substantially determined by the activity of the vesicular RNases of their mycoplasmas. The revealed high RNase activity of mycoplasma vesicles determines the apoptotic effect of these enzymes on the target cells of the mycoplasma vesicular traffic.

Since mycoplasmas may influence almost all the parameters of eukaryotic cells, the results obtained with infected cells should be treated with suspicion. Due to this fact, the editors of journals suggest that authors provide results of the verification of the experimental data (in particular, cell lines) for mycoplasma contamination. Since many viral vaccines are created using a primary cell culture, the problem of their contamination with mycoplasma is of special importance as vaccine contamination poses a potential risk to human health [1, 3, 5]. In this regard, many countries demand that products created using primary cell cultures, such as viral vaccines against measles, rubella, poliomyelitis,

rabies, mumps and some others, be thoroughly checked for mycoplasma contamination [3].

Thus, mycoplasma contamination of cell cultures is a serious problem both for fundamental studies and applied research. It is clear that all cell lines being purchased should undergo strict control for mycoplasma contamination before they reach a laboratory, whereas the cultures that are already in use should be regularly checked. The discovery of extracellular vesicle traffic in mycoplasma makes it necessary to control new-type infectious agents as well.

Methods for mycoplasma detection

There are no common markers of cell contamination with mycoplasma. Among specific diagnostic tools (Table 2), there are three approaches recommended by international expert organizations.

Microbiological cultivation is the main approach to detect mycoplasma [3, 62]. In this analysis, an aliquot of the cell culture supernatant is added to a liquid medium to cultivate mycoplasmas. After several days of

Table 2. Methods used to detect mycoplasma in cell cultures

Microbiological cultivation
Electronic microscopy
Biochemical assays
Detection of adenosine phosphorylase activity (6-MPDR)
Enzymatic conversion ATP → ADP detected by luciferase
Chromatographic detection of the transformation of radioactively labeled uridine to uracil with the uridine phosphorylase of mycoplasma
Immunoassays
Immunofluorescence
ELISA
Molecular biology tests
Hybridization analysis
Dot-blot hybridization with specific probes
PCR, RT-PCR
Microscopic detection
Direct staining of DNA with fluorescent dye (DAPI, Hoechst 33258)
Fluorescent in situ hybridization (FISH) using probes labeled with fluorescent dyes

■ – officially approved by a number of international expert organizations:
 FDA Points to Consider (May 1993), Regularien 21CFR610.30;
 USDA federal code #9CFR113.28;
 United States Pharmacopoeia, (USP 33/NF 28 <63> and <1226>, Mycoplasma tests, 2010); European Pharmacopoeia (EP 2.6.7., Mycoplasmas, 7th ed.; 2012);
 Japanese Pharmacopoeia (JP);
 ICH Guideline for biotechnological/biological products.

incubation, the culture is transferred to an agar plate containing the same components as the medium. The plates are then incubated for some time (up to 2 weeks) under aerobic conditions at 37 °C. The emergence of two-phase “fried-egg” colonies indicates that mycoplasmas are present in the test samples. This test is theoretically highly sensitive, but it requires a lot of time (up to 4 weeks) and expensive media. Furthermore, many types of mycoplasmas poorly grow on cell-free media, whereas some of them are impossible to grow *in vitro* [1, 62]. In this test, the medium can also become infected from the outside: either from a researcher, medium components, or laboratory facilities. Thus, this detection method includes the risk of obtaining false-positive and false-negative results. Moreover, the cultivation procedure does not allow one to reveal the extracellular vesicles of bacteria.

The second recommended approach to detect mycoplasma contamination is staining DNA with fluorescent DAPI or Hoechst 33258 [3, 62, 63]. This test is very simple and does not require much time; the result can be obtained in as early as 2–3 hours. However, certain parameters of the condition of cell culture may lead researchers to a wrong decision about whether the culture is contaminated with mycoplasma

or not. For instance, extracellular vesicles secreted by eukaryotic cells in a mycoplasma-free culture contain DNA and RNA, which significantly complicates the interpretation of the results, whereas administration of antibiotics makes it impossible to use the proper test. Nevertheless, this approach is very popular due to its simplicity and the possibility to use it for detecting uncultivable mycoplasmas or those growing poorly on cell-free media. In this analysis, the test culture supernatant is added to a mycoplasma-free indicator cell culture (lines Vero B4, NIH 3T3 or 3T6) [64]. Cells are grown in flasks containing sterile slips, which are washed and stained with fluorescents after several days of cell culture growth. In this case, prolonged duration of the test poses a risk that contaminants would spread in the laboratory.

Polymerase chain reaction (PCR) is nowadays the most effective way to detect mycoplasma [1, 3, 62, 65, 66]. PCR variants allow one to detect mycoplasma DNA and RNA. Oligonucleotides for the amplification of variable regions of 16S rDNA or rRNA and sequences of 16–23S intragenic regions are usually used as primers. PCR can include either a single amplification cycle or the nested PCR with two pairs of primers. The latter variant increases test sensitivity and specificity, but

at the same time it poses a risk of obtaining false results due to the possible contamination of target DNA. In addition, the medium components can be Taq polymerase inhibitors: so the test should be carried out using extracted DNA rather than the raw lysate of the cell culture supernatant. Administration of antibiotics may lead to false results: so the culture should be grown without antibiotics for at least 2 weeks before performing the test.

The use of reverse transcription PCR (RT-PCR) to detect rRNA increases test sensitivity; however, this variant is labor-intensive. Taking into account the fact that mycoplasma titer in cell cultures is sufficient to register bacterial DNA, a simple one-step PCR is acceptable. It meets the requirements for a short-term test: it is easy to perform, highly sensitive, specific, and cost-effective. Meanwhile, positive PCR results do not necessarily mean that the sample has the living cells of a contaminant (which is important to keep in mind while analyzing the material after the measures aimed at mycoplasma eradication). In certain cases, when PCR testing of the culture analyzed for mycoplasma contamination yields positive results, amplicon sequencing is needed to draw the final conclusions. Nevertheless, PCR has been approved by international expert organizations, and nowadays there are enough commercial sets for testing cultures for mycoplasma contamination available on the market [3, 62]. The primers used in these sets are ineffective in detecting extracellular vesicles; however, the discovery of mycoplasma-specific nucleic acid sequences in vesicles [20, 21, 24, 36] presents a challenge for developing PCR tests that would detect the corresponding infectious agents.

In addition to the officially approved approaches that have been listed above, there also are other methods: immunoassays and hybridization tests that in addition to using antisera, monoclonal antibodies, and DNA-RNA hybridization employ radioactive or fluorescent tags; biochemical and microscopic methods, etc. (Table 2) [1, 3, 43, 62, 67, 68]. All these methods are characterized by different sensitivities and are not free of the disadvantages typical of the aforementioned approaches.

The data presented above is evidence that the problem of detecting mycoplasma contamination has yet to be solved. All the available methods have disadvantages and limitations, so it is recommended that a cell culture be simultaneously tested using several techniques [1, 3, 62]. It is clear that in order to test the medium components for the presence of such infectious agents as extracellular bacterial vesicles, special tests against the markers of these organelles need to be elaborated. Detection of common marker sequences to reveal the respective infectious agents implies a complex study

of extracellular vesicles in various Mollicutes species. Only the first steps have been made in this direction so far [16, 20, 36, 37].

METHODS FOR MYCOPLASMA ERADICATION

Elimination of the infected cell culture and obtainment of a new, clear one is believed to be the best way to solve the problem of mycoplasma contamination [1, 3, 69]. If this is impossible, then one is faced with the decontamination issue, which means mycoplasma eradication without damaging eukaryotic cells. However, despite the fact that numerous approaches for the elimination of mycoplasma have been suggested and tried over several decades, an effective one has not been found yet. Nevertheless, researchers have remained persistent, and successful cases of cell culture decontamination by virtue of either new or modified approaches are reported from time to time [1, 3, 69–71]. The most popular one is the use of antibiotics.

Specific features of mycoplasma biology define the pattern of their susceptibility to antibiotics. Many of those turn out to be inefficient as mycoplasmas lack targets they are aimed at. For instance, they lack cell wall peptidoglycan whose synthesis is inhibited by penicillin [1, 3, 72]. On the other hand, some antibiotics do not cause mycoplasma death, but they slow down its growth and thus disguise the presence of a contaminant [2]. This fact is the reason why antibiotics are not recommended for prophylactic use upon *in vitro* cultivation [2, 5, 69]. Nonetheless, researchers continue to look for agents for cell culture decontamination among antibiotics [2, 3, 67, 69].

Three groups of antibiotics exhibiting some activity against mycoplasma are known thus far: macrolides, quinolones, and tetracyclines [3, 69, 72]. It has been reported in a number of publications that serial treatment of cell cultures with certain combinations of antibacterial agents belonging to these groups effectively removes mycoplasmas [3, 67, 69]. However, experimental attempts to decontaminate cell cultures according to the reported protocols often fail [1, 71, 73]. Taking into account this fact, together with the negative impact of antibiotics on cell cultures, most researchers remain skeptical of attempts to eradicate mycoplasma with antibiotics, while commercial companies continue to actively advertise these products.

A significant problem of antibiotic therapy against mycoplasma infections is that mycoplasmas quickly develop resistance [1, 19, 74]. The mechanisms of rapid development of resistance to antibiotics are not clear. It is assumed that, alongside the known mechanisms of developing resistance to such antibiotics as quinolones, the mycoplasmas use other mechanisms that have not been identified yet [75–77]. Extracellular vesicles

have recently been reported to potentially mediate the mechanisms of developing resistance to antibiotics in bacteria [78, 79], including mycoplasma [36]. Involvement of extracellular vesicles in the formation of mycoplasma resistance to antibiotics has been proved for *A. laidlawii*. To prove it, we used mycoplasma strains characterized by different susceptibilities to ciprofloxacin: the laboratory (PG8) and PG8R, which was derived from it in a stepwise manner and showed high resistance to the antibiotic. It turned out that these strains also had different clearance mechanisms and different vesicle generation rates. It was found that the high resistance of a PG8R strain is associated with a high vesicle generation rate and that vesicles, in turn, participate in the ciprofloxacin traffic exhibiting a bacteriostatic effect towards *Staphylococcus aureus* (a strain sensitive to the antibiotic). The strain with high resistance to ciprofloxacin was found to have a C® T transition at the 272 position (causing a serin to leucin transition – Ser (91) Leu in the target protein molecule) in *parC* locus (determining resistance to fluoroquinolone) of the target gene (topoisomerase IV). It turned out that the vesicles of this mycoplasma strain export the mutant gene of the target protein. Export of the antibiotic target genes mediated by extracellular vesicles favors a quick distribution of the mutant target of quinolones over the microbiocenosis by horizontal transfer [80]. Performance of this pattern has been recently demonstrated in model systems of *Escherichia coli* and *Pseudomonas aerogenosa* [81, 82]. The study of these processes in mycoplasma has not been completed yet, although it is already clear that extracellular vesicles are the important component of the mechanisms of quick adaptation to antibacterial products. Considering the fact that vesicle secretion is the process that allows microorganisms to survive under various conditions

[27, 32], searching for effective antibiotic means of cell culture decontamination does not appear promising.

Thus, mycoplasma contamination of cell cultures and mycoplasma diagnosis and elimination remain serious problems [1, 3, 7, 69, 83, 84]. It is absolutely clear that reliable methods for detecting infectious agents and decontamination methods are needed, which would be based first and foremost on a thorough investigation of mycoplasma genetics and physiology. The discovery of the extracellular vesicular traffic in mycoplasmas mediating cell-to-cell interactions and pathogenesis makes it necessary to take into account new infectious agents. Since cell cultures are used to produce vaccines and physiologically active compounds, quickly solving the discussed issue is topical both for fundamental science and the biotechnological production of pure, next-generation products.

The authors sincerely thank the personnel of the Laboratory of Molecular Basics of Pathogenesis at the Kazan Institute of Biochemistry and Biophysics of the Russian Academy of Sciences, who participated in the experimental work: A.A. Mouzykantov, N.B. Baranova, E.S. Medvedeva, G.F. Shaymardanova and M.V. Trushina.

This work was performed within the scope of the Program for Increasing the Competiveness of the Kazan (Volga region) Federal University of the Ministry of Education and Science of the Russian Federation and supported by the Russian Foundation for Basic Research (grants No 14-04-00883a, 12-04-01052a, 12-04-01226a), the Presidential grant (MK-3823.2023.4) and a grant for government support of leading scientific schools of the Russian Federation (No NSh-825.2012.4).

REFERENCES

1. Borkhsenius S.N., Chernova O.A., Chernov V.M., Vonsky M.S. Mycoplasma. St. Petersburg.: Science. 2002. 320 p. (In Russian)
2. Uphoff C., Drexler H. // Meth. Mol. Biol. 2011. V. 731. P. 105–114.
3. Rottem S., Kosower N.S., Kornspan J.D. // Contamination of Tissue Cultures by Mycoplasmas, Biomedical Tissue Culture / Ed. Dr. Luca Ceccherini-Nelli. 2012.
4. Robinson L.B., Wichelhausen R.H., Roizman B. // Science. 1956. V. 124. P. 1147–1148.
5. Drexler H.G., Uphoff C.C. // Cytotechnology. 2002. V. 39. № 2. P. 75–90.
6. Razin Sh. // Prokaryotes. 2006. V. 4. P. 836–904.
7. Folmsbee M., Howard G., McAlister M. // Biologicals. 2010. V. 38. P. 214–217.
8. Nikfarjam L., Farzaneh P. // Cell J. 2012. V. 13. № 4. P. 203–212.
9. Kühner S., van Noort V., Betts M.J., Leo-Macias A., Batisse C., Rode M., Yamada T., Maier T., Bader S., Beltran-Alvarez P., et al. // Science. 2009. V. 326. № 5957. P. 1235–1240.
10. Yus E., Maier T., Michalodimitrakis K., van Noort V., Yamada T., Chen W.-H., Wodke J.A.H., Gell M., Mart nez S., Bourgeois R., et al. // Science. 2009. V. 326. № 5957. P. 1263–1268.
11. Maier T., Schmidt A., Gueell M., Kuehner S., Gavin A.C., Aebersold R., Serrano L. // Mol. Syst. Biol. 2011. V. 7. P. 511. doi: 10.1038/msb.2011.38.
12. Oshima K., Ishii Y., Kakizawa S., Sugawara K., Neriya Y., Himeno M., Minato N., Miura C., Shiraishi T., Yamaji Y., et al. // PLoS One. 2011. V. 6. e23242.
13. van Noort V., Seebacher J., Bader S., Mohammed S., Vonkova I., Betts M.J., Kuhner S., Kumar R., Maier T., O'Flaherty M., et al. // Mol. Syst. Biol. 2012. V. 8. P. 571.
14. Lluch-Senar M., Luong K., Lloréns-Rico V., Delgado J., Fang G., Spittle K., Clark T.A., Schadt E., Turner S.W., Korlach J., et al. // PLoS Genet. 2013. V. 9. № 1. e1003191.

15. Hopfe M., Deenen R., Degrandi D., Köhrer K., Henrich B. // *PLoS One*. 2013. V. 8 (Suppl 1). e54219.
16. Chernov V.M., Chernova O.A., Medvedeva E.S., Mouzykantov A.A., Ponomareva A.A., Shaymardanova G.F., Gorshkov O.V., Trushin M.V. // *J. Proteomics*. 2011. V. 74. № 12. P. 2920–2936.
17. Parraga-Nino N., Colome-Calls N., Canals F., Querol E., Ferrer-Navarro M. // *J. Proteome Res*. 2012. V. 11. № 6. P. 3305–3316.
18. Liu W., Fang L., Li M., Li S., Guo S., Luo R., Feng Z., Li B., Zhou Z., Shao G., Chen H., Xiao S. // *PLoS One*. 2012. V. 7. № 4. e35698.
19. Razin S., Hayflick L. // *Biological*. 2010. V. 38. № 2. P. 183–190.
20. Chernov V.M., Chernova O.A., Gorshkov O.V., Baranova N.B., Mouzykantov A.A., Nesterova T.N., Ponomareva A.A. // *Doklady Biochemistry and Biophysics*. 2013. №450. P.483–487. (In Russian)
21. Chernov V.M., Chernova O.A., Mouzykantov A.A., Baranova N.B., Gorshkov O.V., Trushin M.V., Nesterova T.N., Ponomareva A.A. // *Scientific World J*. 2012. V. 2012. Article 315474. doi :10.1100/2012/315474.
22. Lazarev V.N., Levitskii S.A., Basovskii Y.I., Chukin M.M., Akopian T.A., Vereshchagin V.V., Kostjukova E.S., Kovalova G.Y., Kazanov M.D., Malko D.B., et al. // *J. Bacteriol*. 2011. V. 193. № 18. P. 4943–4953.
23. Chernov V.M., Chernova O.A., Medvedeva E.S., Davydova M.N. // *AS Rep*. 2011. № 438. P. 562–565. (In Russian)
24. Chernov V.M., Chernova O.A., Mouzykantov A.A., Efimova I.R., Shaymardanova G.F., Medvedeva E.S., Trushin M.V. // *Scientific World J*. 2011. V. 11. P. 1120–1130.
25. Kulp A., Kuehn M.J. // *Annu. Rev. Microbiol*. 2010. V. 64. P. 163–184.
26. Lee E.Y., Choi D.S., Kim K.P., Gho Y.S. // *Mass Spectrom*. 2008. V. 27. № 6. P. 535–555.
27. Deatherage B.L., Cookson B.T. // *Infect. Immun*. 2012. V. 80. № 6. P. 1948–1957.
28. Gaudin M., Gauliard E., Schouten S., Houel-Renault L., Lenormand P., Marguet E., Forterre P. // *Environmental Microbiol. Rept*. 2013. V. 5. P. 109–116.
29. Lee E.Y., Choi D.Y., Kim D.K., Kim J.W., Park J.O., Kim S., Kim S.H., Desiderio D.M., Kim Y.K., Kim K.P., et al. // *Proteomics*. 2009. V. 9. № 24. P. 5425–5436.
30. Amano A., Takeuchi H., Furuta N. // *Microbes Infect*. 2010. № 12. P. 791–798.
31. Yoon H., Ansong C., Adkins J.N., Heffron F. // *Infect. Immun*. 2011. V. 79. № 6. P. 2182–2192.
32. Schertzer J.W., Whiteley M. // *J. Mol. Microbiol. Biotechnol*. 2013. V. 23. № 1–2. P. 118–130.
33. Kluge S., Hoffmann M., Bendorff D., Rapp E., Reichl U. // *Proteomics*. 2012. V. 12. № 12. P. 1893–1901.
34. Pollak C.N., Delpino M.V., Fossati C., Baldi P.C. // *PLoS One*. 2012. V. 7. № 11. e50214.
35. Pierson T., Matrakas D., Taylor Y.U., Manyam G., Morozov V.N., Zhou W., van Hoek M.L. // *J. Proteome Res*. 2011. V. 10. № 3. P. 954–967.
36. Medvedeva E.S., Baranova N.B., Mouzykantov A.A., Grigorieva T.Y., Davydova M.N., Trushin M.V., Chernova O.A., Chernov V.M. // *Scientific World J*. 2014. V. 2014. Article 150615. doi: 10.1155/2014/150615.
37. Mouzykantov A.A., Baranova N.B., Medvedeva E.S., Grigorieva T.Y., Chernova O.A., Chernov V.M. // *Doklady Biochemistry and Biophysics*. 2014. V.455. №1. P.99–104. (In Russian)
38. Rottem S. // *Physiol. Rev*. 2003. V. 83. № 2. P. 417–432.
39. Razin Sh., Herrmann R. *Molecular biology and pathogenicity of Mycoplasmas*. N.Y.: Kluwer Acad. Plenum Publ., 2002. 572 p.
40. Zhang S., Tsai S., Lo S. // *BMC Cancer*. 2006. № 6. P. 116–126.
41. Zhang S., Wear D.J., Lo S. // *FEMS Immunol. Med. Microbiol*. 2000. V. 27. № 1. P. 43–50.
42. Lavrič M., Maughan M.N., Bliss T.W., Dohms J.E., Benčina D., Keeler Jr., Narat M. // *Veterinary Microbiology*. 2008. V. 126. № 1–3. P. 111–121.
43. Wang Z., Farmer K., Hill G.E., Edwards S.V. // *Mol. Ecol*. 2006. V. 15. № 5. P. 1263–1273.
44. Cecchini K.R., Gorton T.S., Geary S.J. // *J. Bacteriol*. 2007. № 189. P. 5803–5807.
45. Mou H.-Q., Lu J., Zhu S.-F., Lin C.-L., Tian G.-Z., Xu X., Zhao W.-J. // *PLoS One*. 2013. V. 8. № 10. e77217.
46. Harper D.R., Kangro H.O., Argent S., Heath R.B. // *J. Virol. Methods*. 1988. V. 20. № 1. P. 65–72.
47. Cole B., Mu H.H., Pennock M.D., Hasebe A., Chan F.V., Washburn L.R., Peltier M.R. // *Infect. Immun*. 2005. V. 73. № 9. P. 6039–6047.
48. Gerlic M., Horowitz J., Farkash S., Horowitz S. // *Cell Microbiol*. 2007. V. 9. P. 142–153.
49. Quah B., O'Neill H. // *J. Leukoc. Biol*. 2007. V. 82. № 5. P. 1070–1082.
50. Zhang S., Lo S. // *Curr. Microbiol*. 2007. V. 54. № 5. P. 3888–3895.
51. Gong M., Meng L., Jiang B., Zhang J., Yang H., Wu J., Shou C. // *Mol. Cancer Ther*. 2008. V. 7. № 3. P. 530–537.
52. Kraft M., Adler K.B., Ingram J.L., Crews A.L., Atkinson T.P., Cairns C.B., Krause D.C., Chu H.W. // *Eur. Respir. J*. 2003. V. 31. № 1. P. 43–46.
53. Dusanic D., Bencina D., Oven I., Cizelj I., Bencina M., Narat M. // *Veterinary Res*. 2012. V. 43. P. 7–20.
54. Elkind E., Vaisid T., Kornspan J.D., Barnoy S., Rottem S., Kosower N.S. // *Cell. Microbiol*. 2012. V. 14. № 6. P. 840–851.
55. Ilinskaya O.N., Sokurenko Yu.V., Ulyanova V.V., Vershinina V.I., Zelenikhin P.V., Kolkakov A.I., Medvedeva E.S., Baranova N.B., Davidova M.N., Mouzykantov A.A. et al. // *Mikrobiologiya (Microbiology)*. 2014. V.83. №3. P. 320–327. (In Russian)
56. Chernov V.M., Chernova O.A., Margulis A.B., Mouzykantov A.A., Baranova N.B., Medvedeva E.S., Kolkakov A.I., Ilinskaya O.N. // *American-Eurasian J. Agric. & Environ. Sci*. 2009. V. 6. № 1. P. 104–107.
57. Ilinskaya O.N., Ivanchenko O.B., Karamova N.S. // *Mutagenesis*. 1995. V. 10. № 3. P. 165–170.
58. Ilinskaya O.N., Ivanchenko O.B., Karamova N.S., Kipenskaya L.V. // *Mut. Res*. 1996. V. 354. № 2. P. 203–209.
59. Makarov A.A., Ilinskaya O.N. // *FEBS Lett*. 2003. V. 540. P. 15–20.
60. Makarov A.A., Kolchinsky A., Ilinskaya O.N. // *BioEssays*. 2008. V. 30. № 8. P. 781–790.
61. Ulyanova V., Vershinina V., Ilinskaya O. // *FEBS J*. 2011. V. 278. № 19. P. 3633–3643.
62. Uphoff C.C., Drexler H.G. // *Meth. Mol. Biol*. 2013. V. 946. P. 1–13.
63. Chen T.R. // *Exp. Cell Res*. 1977. V. 104. № 2. P. 255–262.
64. Spierenburg G.T., Polak-Vogelzang A.A., Bast B.J. // *J. Immunol. Meth*. 1988. V. 114. № 1–2. P. 115–119.
65. Harasawa R., Mizusawa H., Fujii M., Yamamoto J., Mukai H., Uemori T., Asada K., Kato I. // *Microbiol. Immunol*. 2005. V. 49. № 9. P. 859–863.
66. Sung H., Kang S.H., Bae Y.J., Hong J.T., Chung Y.B., Lee C.K., Song S. // *J. Microbiol*. 2006. V. 44. № 1. P. 42–49.

REVIEWS

67. Mariotti E., D'Alessio F., Mirabelli P., Di Noto R., Fortunato G., Del Vecchio L. // *Biologicals*. 2012. V. 40. № 1. P. 88–91.
68. Degeling M.H., Maguire C.A., Bovenberg M.S., Tannous B.A. // *Anal. Chem.* 2012. V. 84. № 9. P. 4227–4232.
69. Uphoff C.C., Drexler H.G. // *Meth. Mol. Biol.* 2013. V. 946. P. 15–26.
70. Nir-Paz R., Prévost M.C., Nicolas P., Blanchard A., Wróblewski H. // *Antimicrob. Agents Chemother.* 2002. V. 46. № 5. P. 1218–1225.
71. McGarrity G.J., Kotani H., Butler H. // *Mycoplasmas: molecular biology and pathogenesis* / Eds Maniloff J., McElhaney R.N., Finch L., Baseman J.B. Washington: ACM, 1992. 906 p.
72. Bébéar C., Pereyre S., Peuchant O. // *Future Microbiol.* 2011. № 4. P. 423–431.
73. Singh S., Puri S.K., Srivastava K. // *Parasitol. Res.* 2008. V. 104. № 1. P. 181–184.
74. Couldwell D.L., Tagg K.A., Jeoffreys N.J., Gilbert G.L. // *Int. J. STD AIDS.* 2013. V. 24. № 10. P. 822–828.
75. Raherison S., Gonzalez P., Renaudin H., Charron A., Bébéar C., Bébéar C.M. // *Antimicrob. Agents Chemother.* 2005. V. 49. № 1. P. 421–424.
76. Ruiz J., Pons M.J., Gomes C. // *Int. J. Antimicrob. Agents.* 2012. V. 40. № 3. P. 196–203.
77. Yamaguchi Y., Takei M., Kishii R., Yasuda M., Deguchi T. // *Agents Chemother.* 2013. V. 57. № 4. P. 1772–1776.
78. Ciofu O., Beveridge T.J., Kadurugamuwa J., Walther-Rasmussen J., Hoiby N. // *J. Antimicrob. Chemother.* 2000. V. 45. № 1. P. 9–13.
79. Lee J., Lee E.Y., Kim S.H., Kim D.K., Park K.S., Kim K.P., Kim Y.K., Roh T.Y., Gho Y.S. // *Antimicrob. Agents Chemother.* 2013. V. 57. № 6. P. 2589–2595.
80. Stanhope M.J., Walsh S.L., Becker J.A., Italia M.J., Ingraham K.A., Gwynn M.N., Mathie T., Poupard J.A., Miller L.A., Brown J.R., et al. // *Antimicrob. Agents Chemother.* 2005. V. 49. № 10. P. 4315–4326.
81. Manning A.J., Kuehn M.J. // *BMC Microbiol.* 2011. V. 11. P. 258. doi: 10.1186/1471-2180-11-258.
82. Maredia R., Devineni N., Lentz P., Dallo S.F., Yu J., Guentzel N., Chambers J., Arulanandam B., Haskins W.E., Weitao T. // *Scientific World J.* 2012. V. 2012. Article 402919. doi: 10.1100/2012/402919.
83. David S.A., Volokhov D.V., Ye Z., Chizhikov V. // *Appl. Environ. Microbiol.* 2010. V. 76. № 9. P. 2718–2728.
84. Windsor H.M., Windsor G.D., Noordergraaf J.H. // *Biologicals.* 2010. V. 38. № 2. P. 204–210.

Structural Features of the Interaction between Human 8-Oxoguanine DNA Glycosylase hOGG1 and DNA

V. V. Koval^{1,2}, D. G. Knorre^{1,2}, O. S. Fedorova^{1,2*}

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

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

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

Received 27.02.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT The purpose of the present review is to summarize the data related with the structural features of interaction between the human repair enzyme 8-oxoguanine DNA glycosylase (hOGG1) and DNA. The review covers the questions concerning the role of individual amino acids of hOGG1 in the specific recognition of the oxidized DNA bases, formation of the enzyme–substrate complex, and excision of the lesion bases from DNA. Attention is also focused upon conformational changes in the enzyme active site and disruption of enzyme activity as a result of amino acid mutations. The mechanism of damaged bases release from DNA induced by hOGG1 is discussed in the context of structural dynamics.

KEYWORDS protein-nucleic acid recognition, human 8-oxoguanine DNA glycosylase, repair enzymes, loss-of-function mutants, structural analysis of hOGG1.

ABBREVIATIONS AP – apurinic/aprimidinic site; 8-oxoG – 7,8-dihydro-8-oxoguanine.

INTRODUCTION

The vast majority of the tertiary structures of proteins and their complexes elucidated and deposited in databases over the past half-century (as of April 2014, there are about 100,000 entries in the archive for the three-dimensional structural data, the Protein Data Bank, or PDB) have naturally turned the attention to the relationship between proteins' structures and their functions. This paper reviews the structural features of hOGG1 enzyme interaction with DNA damage sites.

The genomes of all living organisms are continuously exposed to a heavy load of exogenous and endogenous mutagens, some of which are reactive oxygen species (ROS), highly reactive cellular by-products, xenobiotics, UV light, ionizing radiation, etc. Among the wide range of agents that attack DNA [1–4] and cause numerous diseases [5–10], an important role is attributed to the ROS, such as $O_2^{\cdot-}$, H_2O_2 , and OH^{\cdot} [11, 12], produced under aerobic conditions. The mutagenic lesions occurring due to oxidative damage of purine DNA residues involve the 7,8-dihydro-8-oxoguanine (8-oxoguanine, 8-oxoG) and 5-formamidopyrimidine derivatives of adenine, 4,6-diamino-5-formamidopyrimidine (Fapy A), and guanine, 2,6-diamino-4-oxy-5-formamidopyrimidine (Fapy G) [13, 14].

A buildup of DNA lesions causes structural damage to the DNA molecule. For example, 8-oxoguanine can

mismatch with adenine in the Hoogsteen mode, which results in a C to A substitution in the first replication cycle (a 8-oxoG/A mismatched pair is formed), followed by the incorporation of T opposite A in the second round of replication, thereby producing a G/C → T/A transversion [15, 16].

To prevent 8-oxoG accumulation in DNA, there is a pathway to protect cells against the mutagenic effect (the GO-system) [17, 18]. This pathway has been described in details for *Escherichia coli*. It consists of three enzymes: Fpg (MutM), a specific N-glycosylase/AP-lyase that catalyzes the excision of 8-oxoG; MutY, a specific N-glycosylase that removes adenine opposite 8-oxoG; and MutT, a nucleotide hydrolase involved in the cleavage of the pyrophosphate bond in 8-oxo-dGTP. Eukaryotic cells have structural or functional analogs of these bacterial enzymes [19, 20]. The excision of 8-oxoguanine from DNA in eukaryotes is carried out by the 8-oxoguanine glycosylase (OGG1) [21]. Each human cell has been shown to contain approximately 50 thousand copies of OGG1 that protect the genomic DNA from the accumulation of oxidized purine nucleotides [22].

In human cells, the *OGG1* gene is located on the short arm of Chromosome 3 (3p25/26). The OGG1 primary transcript gives rise to two distinct mRNAs encoding proteins of the 345 and 424 amino acids; α -hOGG1 and

β -hOGG1, respectively [21, 23–26]. Both isoforms share the first 316 amino acid residues, with the C-termini varying [21, 23–27]. Studies aimed at addressing their cellular location show that α -hOGG1 localizes in the nucleus, while β -hOGG1 localizes in mitochondria [27]. The nuclear isoform of α -hOGG1 is highly conserved and is well studied in humans, the yeast *Saccharomyces cerevisiae*, the plant *Arabidopsis thaliana* cells, the fruit fly *Drosophila melanogaster* cells, and mammals [28]. The α -hOGG1 of *Saccharomyces cerevisiae* and humans share up to a 38% homology. The β -hOGG1 isoform localizes only in mitochondria [28]. Catalytic and structural characteristics have been addressed for only α -hOGG1.

THE CATALYTIC MECHANISM OF THE OGG1 PROTEIN

The hOGG1 enzyme acts both as a DNA-glycosylase and a β -lyase, hydrolysing the N-glycosidic bond of the damaged base to release 8-oxoG, followed by catalytic cleavage of the 3'-phosphodiester bond. The chemical mechanism of action was first suggested by Wallace *et al.* for endonuclease III of *E. coli* [29]. The main ideas proposed in that study were experimentally proved by Lloyd *et al.* [30] using the analysis of cross-links formed between the enzyme and DNA molecule. According to

the proposed mechanism, upon incubation of DNA and the enzyme, followed by sodium borohydride treatment, a covalent bond was formed between the two molecules evidencing the occurrence of Schiff base intermediate.

The catalytic action of hOGG1 involves a mechanism by which the ϵ -amino group of Lys249 participates in the removal of the 8-oxoG base from the C1' of the ribose moiety and promotes the elimination of the 3'-phosphodiester bond through the formation of a Schiff base intermediate with C1' of the deoxyribose moiety (*Fig. 1*) [31–33]. Verdine *et al.* earlier showed that K249Q mutant hOGG1 exhibited no catalytic activity but retained the ability for the recognition of oxidatively damaged DNA [21]. The second chemical event is the scission of phosphodiester bond at the C3 of the 2'- deoxyribose moiety via β -elimination (AP-lyase activity).

Based on numerous structural and biochemical studies, Verdine *et al.* [34] proposed a role of 8-oxoG being displaced in acid/base interactions with the N ϵ , (Lys249 amino group), C2', and O4' atoms. The structure of the intermediate reduced by sodium borohydride, produced through the reaction of hOGG1 with 8-oxoG-containing DNA, has been characterized. The

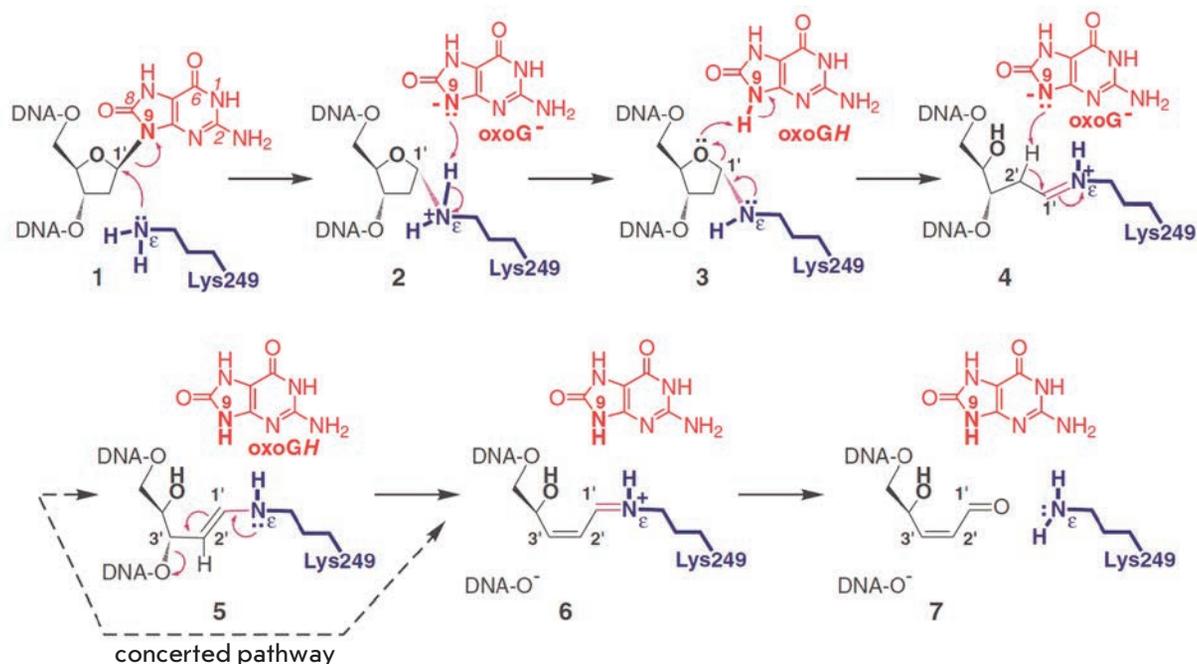


Fig. 1. Detailed stepwise mechanistic proposal for the entire cascade of reactions catalyzed by hOGG1 in which the reaction product serves to catalyze further processing of the substrate. (Reprinted by permission from Macmillan Publishers Ltd: [Nat. Struct. Biol.] Fromme J.C., Bruner S.D., Yang W., Karplus M., Verdine G.L. *Nat. Struct. Biol.* 2003. V. 10. № 3. P. 204–211, copyright 2003)

8-oxoG base to be removed from the DNA strand is not released but retained in the active site and acts as a cofactor at the β -elimination stage. The N9 of the enzyme in the 8-oxoG base is positioned near N ϵ and O4', which permits the transfer of a proton between these two atoms.

Figure 1 illustrates the enzymatic mechanism of hOGG1 as described [34]. The initial attack of the side-chain amino group of Lys249 on the C1' of the deoxyribose sugar results in cleavage of the glycosidic bond to produce an 8-oxoG⁻ anion. The anion deprotonates the ϵ -NH₂ of Lys249 to form the aminated intermediate **3**. The protonated 8-oxoGH transfers the proton to the O4' of the ribose sugar, with the aminated intermediate **3** rearranging to a Schiff base **4** with a sugar ring opening. The Schiff base **4** with a proton on the N ϵ of Lys249 donates it back to 8-oxoG⁻ to yield 8-oxoGH and an uncharged Schiff base **5** promoting the loss of the 3' phosphate via β -elimination. DNA with a 5' phosphate group and a positively charged intermediate **6** arise. The intermediate carries an α,β -unsaturated Schiff base at the 3' end with a positive charge. Upon hydrolysis, the intermediate **6** releases the enzyme and a DNA molecule with 4-hydroxy-2,3-pentalen-1 at the 3' end.

The Protein Data Bank currently holds information on 27 structures of hOGG1. The crystal structures of native hOGG1 [35] and DNA-bound hOGG1 have been solved: catalytically inactive K249Q mutant hOGG1 with 8-oxoG-containing DNA [35, 36], N149C with 8-oxoG-containing DNA and intact DNA [37], D268N with 8-oxoG-containing DNA and DNA with a tetrahydrofuran residue (F-ligand) – a “stop” substrate for hOGG1 [33, 38]; and the WT hOGG1 complex with the F-ligand [33]. In addition, the irreversibly linked adduct of hOGG1 with AP substrate has been described [34], formed through a borohydride-trapped Schiff bases for hOGG1 variant forms (H270A, Q315A, Q315F, G42A) containing mutation of amino acids that bind 7,8-dihydro-8-oxoguanine in the native structure [32]. The structural analysis has also been conducted for a late-stage intermediate wherein 8-oxoG is almost completely inserted in the active site; however, catalytically active conformation has not been achieved yet [32].

STRUCTURAL CHARACTERISTICS OF K249Q HOGG1 LACKING ENZYMATIC ACTIVITY

The first structure of complex of hOGG1 with 8-oxoG-containing substrate was obtained for hOGG1 with a K149Q mutation [36]. It was shown earlier [21] that the mutant form, in which Lys249 is replaced with Glu, lacks catalytic activity but retains substrate recognition. Since the authors [36] failed to yield high-quality crystals for full-length hOGG1 K249Q complexed with the DNA duplex containing 8-oxoG/C, limited diges-

tion by trypsin allowed to remove unstructured amino- and carboxyl-termini as well as amino acids at positions 80–82, facilitating crystallization and the analysis of the hOGG1 core domain comprising residues 12–325 (Fig. 2).

It has been demonstrated that hOGG1 shares a common fold with the members of the superfamily of DNA repair enzymes involved in base excision repair (BER), such as endonuclease III and alkyl-DNA-glycosylase AlkA from *E. coli* [39]. The repair proteins of this family occur in numerous organisms, from bacteria to mammals, repairing DNA bases damaged by oxidation, alkylation, and deamination. These enzymes possess a unique structural motif “helix-hairpin-helix” (HhH) [40], followed by a Gly/Pro-rich loop and conserved residues of Gly, Pro and Asp (HhH-GPD). The hOGG1 structure also contains two α -helix domains shared by all members of the HhH-GPD superfamily and antiparallel β -sheet present in the alkyl-DNA-glycosylase AlkA only.

The protein has high affinity to 8-oxoG-containing DNA (Fig. 2). The 8-oxoG base flips out of the DNA helix and fits into the pocket of the enzyme's active site, which is in agreement with similar structures in other members of the HhH-GPD superfamily [31, 40, 41]. In the case of the 8-oxo-dG base, the heterocyclic compound exists in the *syn*-conformation in relation to the glycosidic bond; although, upon binding to the active site of hOGG1, it takes the *anti*-conformation; i.e., as is the case for a normal duplex DNA. The flipped-out conformation of the glycosidic moiety and DNA backbone leads to the extrusion of 8-oxoG out of the DNA helix and insertion deeply into the active site of hOGG1.

The interaction of hOGG1 with the phosphate groups of 8-oxoG-containing DNA, 8-oxoG itself, and the complementary cytosine contribute to a contact surface of 2.268 Å² [36]. Though in most DNA-binding proteins the contact region contains many lysine and arginine residues, for the interaction with the phosphates groups of hOGG1, the DNA backbone is bound to a nearly uncharged groove lined by a single basic His270 residue. The resulting complex is unique in a way that hOGG1 contains many α -helices with the N-termini oriented towards the DNA (Fig. 2). This disposition of α -helices enhances the helix-dipole interaction, promoting dipole electrostatic contacts rather than salt bridges while binding DNA substrates. Only one of the helices, α L, is involved in direct contact with the DNA backbone. The α L helix with a loop and the α K helix form the conserved motif HhH. In addition to the interaction through the α L helix (Val250 and Gln249) and the phosphate group p⁻², the highly conserved glycine residue (Gly245), located within the loop, forms a hydrogen bond with the phosphate p⁻³ (Fig. 3). The structural motif HhH is brought into con-

tact with a DNA substrate from the 3'-side of the oxidized base; at this site, the structure of duplex DNA is very similar to the B-form. Consequently, the HhH motif is mainly responsible for positioning the mutated base in the DNA duplex towards the active site pocket.

The phosphate groups p^{-1} , p^0 and p^1 play an important role in the stabilization of the uncommon DNA backbone conformation at the mutagenic site. The rotation required for flipping 8-oxoG out of the helix causes inward rotation of oxygen atoms of p^{-1} and imposes extra strain on the ribose phosphate DNA backbone. To decrease electrostatic repulsion between the closely positioned p^{-1} and p^1 , a partially hydrated Ca^{2+} ion, present in the crystal microenvironment, is brought in between and secured through a direct bond with p^1 and the water bridge with p^{-1} (Fig. 3B). Although Ca^{2+} could be displaced by Mg^{2+} under physiological conditions, it does not come in direct contact with the protein, but its ligand, the water molecule, forms a hydrogen bond with DNA, thereby stabilizing its flipped-out and bent conformation.

The complementary cytosine locates inside the helix; however, it hardly stacks with neighboring nucleobases from the 5' side due to a kink in the chain that orients the duplex away from the enzyme. Beyond the active site, the DNA conformation is similar to that of canonical B-DNA form (Fig. 2).

The void remaining in the duplex after extrusion of 8-oxoG, is filled in by the amino acid residue of the conserved NNN-element (a motif of three Asn residues); namely, Asn149 that forms a hydrogen bond via its side-chain amide carbonyl with the exocyclic NH_2 of the estranged cytosine (Fig. 3B). In addition, hOGG1 plunges the indole ring of Tyr203 into the space between C^0 (estranged cytosine) and the base on the 5'-side (Fig. 3B), thus unstacking the two bases and creating a sharp kink in the DNA molecule, which significantly improves access to the Watson-Crick edge of base from the minor-groove side. C^0 gets unstacked from the base on its 3'-side (Fig. 3B T¹ base). The residues Arg154 and Arg204 of hOGG1 move together toward C^0 from the minor groove; one arginine above

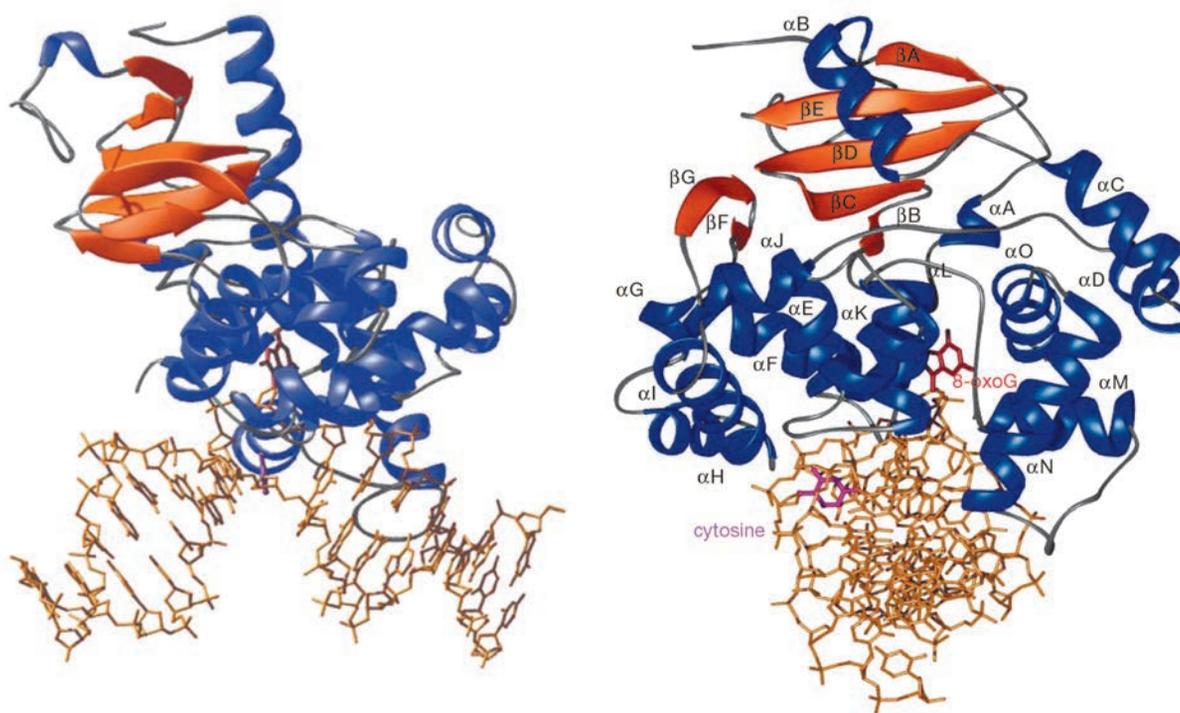


Fig. 2. The overall structure of the hOGG1-DNA complex. Two orthogonal views of the hOGG1-DNA complex with the protein shown as ribbon (blue, α -helices; orange, β -sheets; gray, elements with no secondary structure) and 15-base pair DNA oligonucleotide, as sticks (gold). The substrate oxoG base (red) is completely extruded from the DNA helix and inserted into an active site pocket. The complementary or "estranged" cytosine (purple) remains stacked in the helix. The enzyme bends the DNA ($\sim 70^\circ$) at the plane of the oxoG-C base pair. The bend in the DNA exposes the edge of the estranged cytosine to protein side-chains that make specific contacts. (Reprinted by permission from Macmillan Publishers Ltd: [Nature] Bruner S.D., Norman D.P.G., Verdine G.L. (2000) *Nature*. 403. 859–866, copyright 2000)

and the other below the plane of the pyrimidine ring, simultaneously creating hydrogen bonds with the acceptor N3 and O2 atoms of the estranged C⁰. These bonds seem to be exceptionally strong and occur in the presence of adjacent acceptor atoms, which are unique to cytosine among the other DNA bases. Along with the interaction of Asn149 and C⁰ amide carbonyl of the enzyme and the estranged cytosine, respectively, up to five hydrogen bonds could be created.

The role of Asn149 has been elucidated in the work [35] that demonstrated that Asn149 formed a hydrogen bond with N4 of the exocyclic NH₂ of the cytosine pairing with 8-oxoG. The hydrogen bonds made by the guanidine group of Arg204 with the N1 and O2 form a unique recognition site for the estranged cytosine and appear to play an essential role in the specific recognition of the estranged cytosine typical of hOGG1.

Recognition of 8-oxoG at the active site is accomplished by specific contacts between the lesion base and the amino acids. The enzyme recognizes the urea fragment in 8-oxoG containing the C8-carbonyl group, the N7 and N9 atoms, with N7 forming a hydrogen bond with the carbonyl of Gly42. Of all contacts associated with 8-oxoG, only the one made by Gly42 that would be different with oxoG versus guanine. Therefore, the authors of [36] conclude that 8-oxoG is discriminated from G by a single hydrogen bond. Significantly, the essential amino acid residue Gly42 is the only residue that the β -sheet domain contributes to the hOGG1–DNA interface.

In addition to the aforementioned residues responsible for 8-oxoG recognition, other residues of the hOGG1 active site also contribute to recognition. Phe319 and Cys253 stack toward opposite π -faces of the 8-oxoG, sandwiching the base in the active site (Fig. 4). The Gln315 amide NH₂- group, in cooperation with a tightly bound water molecule, interacts with O⁶ of 8-oxoG, and the Gln315 side-chain carbonyl forms two hydrogen bonds with N1 and N²H of 8-oxoG. Another tightly bound water molecule is hydrogen bonded to O⁶. Gln315 and Gly42, including the water molecules trapped in the active site, are not chemically able to form hydrogen bonds with A, C, and T through donor/acceptor interactions.

The structural analysis in [36] provides insight into the role of the catalytically important residues Lys249 and Asp268. Lys249 is located ~2.5 Å off the C1' of 8-oxoG close to the space in the active site into which the lesion base is extruded, with the Asp268 is suitably disposed to assist the protonation/deprotonation of Lys249. The intermediate, formed by an attack of the deoxyribose sugar by Lys249, rearranges to a Schiff base (Fig. 1). This rearrangement requires deprotonation of the Lys249 side-chain amino group, presuma-

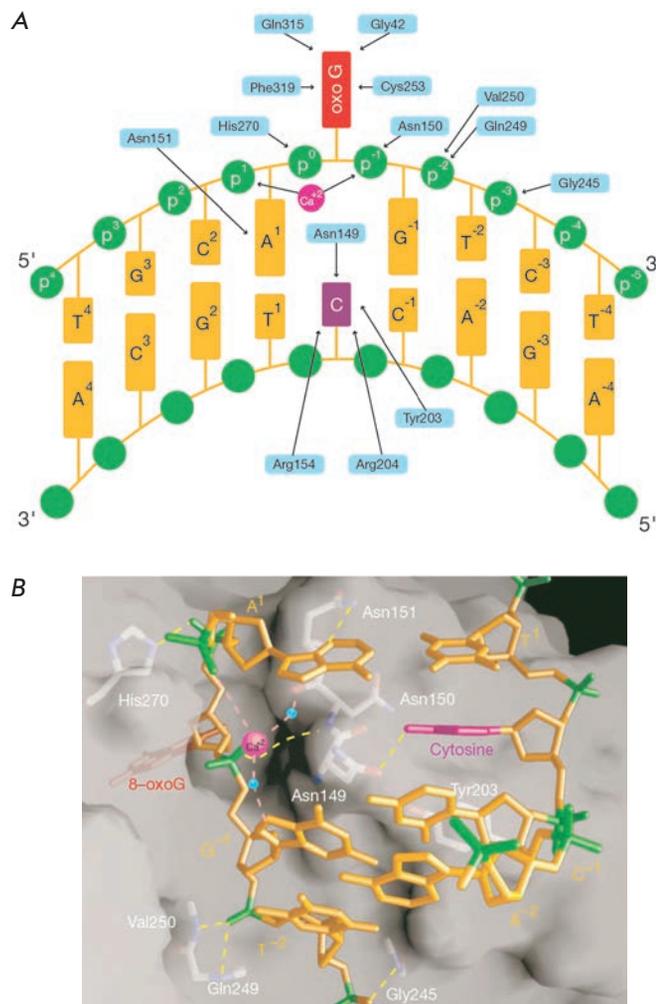


Fig. 3. Contact interface between hOGG1 and oxoG-C-containing DNA. A – The DNA-protein interface. Arrows indicate interactions between amino-acid residues (blue) and DNA bases (yellow). B – Surface of the DNA-protein interface. Hydrogen bonds are shown in yellow, and coordinative interactions with a calcium ion (magenta) are shown in pink. (Reprinted by permission from Macmillan Publishers Ltd: [Nature] Bruner S.D., Norman D.P.G., Verdine G.L. (2000) *Nature*. 403. 859–866, copyright 2000)

bly by Asp268, and O1' protonation, for which His270 is suitably positioned. This role of His270 sheds light on why this residue is invariant in members of the HhH-GPD superfamily, since it is involved in the catalyzing of the formation of Schiff base intermediates.

CATALYTICALLY ACTIVE HOGG1 COMPLEX WITH THE “STOP” SUBSTRATE

The structure of hOGG1 complex with DNA containing 2-oxomethyl-2-oxo-tetrahydrofuran (F) instead of

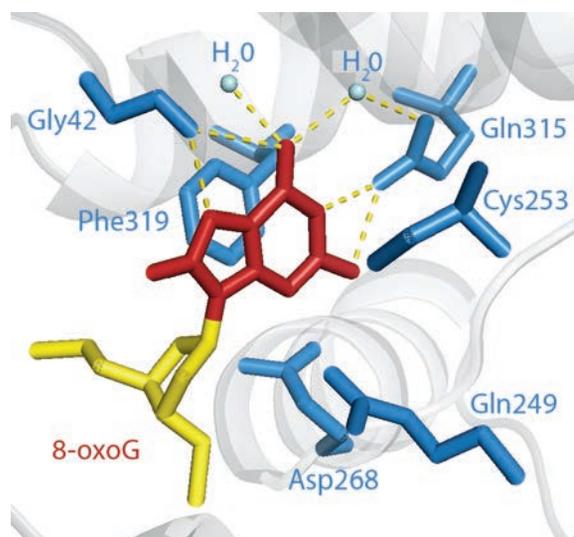


Fig. 4. Amino acids of the hOGG1 active site involved in the recognition of the 8-oxoG base. Visualization of the 3D structure (PDB ID: 1EBM [36]) using the software package PyMOL [43]

8-oxoG has been reported in [33]. This DNA substrate acts as a stop substrate for hOGG1. It is shown that in the produced structure (PDB ID: 1FN7) Asp268 is positioned rather far from Lys249 to initiate deprotonation of its side-chain amino group; the N–O distance is 3.7 Å. In addition, Asp268 does not contact with the His270 that is likely needed for hydrogen bonding to O1' of the ribose sugar as it was found in K249Q hOGG1 structure [36]. Taken together, the absence of the lesion base being flipped out of the helix and into the active site pocket causes conformational changes in the enzyme.

Based on the findings of [33], it was concluded that, firstly, the 8-oxoG-recognizing pocket in hOGG1 well fits the structure of the lesion, positioning Phe319, His270, and Asp268 in the appropriate arrangement. Secondly, the role of Asp268 in the deprotonating of Lys249 has been neither confirmed nor excluded. A hypothesis is proposed that this residue creates an electrostatic field that stabilizes the positive charge developing in the transition state, especially at O1' and C1' of the 8-oxoG deoxyribose.

hOGG1 CONFORMATIONAL CHANGES UPON DNA BINDING

The crystal structures of native hOGG1 and that bound to 8-oxoG-containing DNA were obtained and described by Bjørås *et al.* [35] at a resolution of 2.15 Å. It was found that the hOGG1 structure significantly differs from those of the free form and the one com-

plexed with 8-oxoG-containing DNA (*Fig. 5*). The Phe319, Cys253, Gly42, Gln43, Phe45, and Gln315 residues are responsible for recognition, which is consistent with other findings [36]. Phe319 and Cys253 sandwich 8-oxoG, whereas Gly42, Gln43, and Phe45 interact with the major-groove edge, recognizing the protonated N7 of 8-oxoG. The Gln315 amide oxygen is hydrogen-bonded to the N1-imino- and N2-amino groups of the ring involved in base pairing. The Phe319 residue takes different conformations in the free form of hOGG1 and that bound to DNA (*Fig. 5A*). When bound to DNA, the aromatic ring of Phe319 is oriented almost perpendicular to its position in the free form. The side-chain of Gln315 in the free enzyme is positioned under the aromatic ring of Phe319 (*Fig. 5A*). In complex with DNA, the nitrogen atom of the amide moiety of Gln315 and the carbonyl oxygen of Pro266 located on the opposite side of the binding site are involved in hydrogen bonding. This causes a significant conformational change in the enzyme, thereby creating a tight pocket for binding the damaged base.

In comparison with [36] a slightly different point of view for the role of His270 has been proposed [35]. It was demonstrated that conformational changes in the lesion-recognition pocket are accompanied by a change in the orientation of His270, which forms two hydrogen bonds when bound to DNA: one between the N ϵ 2 of the imidazole ring and the 5'-phosphate of 8-oxoG, the other between N δ 1 and the carboxyl group of Asp322 (*Fig. 5A*). In the free enzyme, one hydrogen bond with Asp322 is retained, although the imidazole ring of His270 rotates by more than 90° as compared to the DNA-bound conformation; packs against the phenyl ring of Phe319 and forms two layers of a sandwich completed by Gln315. The His270 rotamer is not compatible with the Phe319 and Gln315 conformations required for specific binding of the 8-oxoG in DNA. The side-chain conformations of Phe319 and Gln315 depend on the His270 conformation, which is itself governed in the 8-oxoG-DNA complex by interaction with the 5'-phosphate of 8-oxoG nucleotide. Overall, binding of the ribose-phosphate backbone to DNA influences the conformation of His270, which in turn causes a conformational alteration to Phe319 and Gln315, thus allowing entry of the aberrant base into the pocket.

Hence, the side-chains in Phe319, Gln315, and His270 behave as an entity, flipping between the closed and open states when binding the 8-oxoG-base.

In the free form, the amino acid region of 146 to 151 adopts a conformation different from that in a DNA-bound state wherein the atoms could shift away from their positions by up to 4–9 Å. Most pronounced changes occur within the center of the motif, with the side-chain of Asn149 in the free form extending back to-

ward the enzyme and hydrogen bonding to the amide oxygen and ϵ -amino group of the catalytic residue Lys249 (Fig. 5B). However, during complex formation with 8-oxoG-DNA, this oxygen atom retracts by $\approx 9 \text{ \AA}$ to pack against the estranged cytosine. The remaining residues of Asn in the NNN triplet, 150 and 151, whose hydrogen bonds in a DNA-bound structure stabilize the cytosine-recognizing protein site, are exposed to a solvent in the free form. The conformation of this inter-helical peptide seen in the free enzyme is unable to bind to 8-oxoG in its the extrahelical conformation. Flipping of this enzyme segment into target DNA, as well as flipping of the scissile base into the enzyme pocket, requires conformational alterations of the protein to allow the chemistry to take place.

An important sequel of His270 re-orientation in the hOGG1-8-oxoG-DNA complex is an incremental shift of the hinge and strand between Pro266 and Trp272, which moves the side-chain of Asp268 by over 1.5 \AA as compared to its position in the free enzyme (Fig. 5C).

In the DNA-bound hOGG1 complex with F-ligand [33], the Lys249 residue is located away from the carboxyl group of Asp268, which challenges its role in the deprotonating of the nucleophile, thus indicating an alternative function for this residue in the transition state charge stabilization of the deoxyribose ring. In addition, the carboxyl group of Asp268 and the imidazole group of His270 in the DNA-bound structure are sterically close for weak hydrogen bonding which would favor proton abstraction by Asp268, promoting its indirect catalytic contribution. On the other hand, the side-chains of Asp268 and His270 in the free hOGG1 are positioned $>4 \text{ \AA}$ apart, which essentially prevents their hydrogen bonding (Fig. 5C). By contrast, the side-chains of Lys249 and Asp268 sterically allow a hydrogen bond formation between the ϵ -amino- and carboxyl groups, respectively, as well as another hydrogen bond between the ϵ -amino group of Lys249 and the side-chain oxygen of Asn149 (Fig. 5C). The side-chain configuration of Lys249 in the active site corresponds to the protonated ϵ -aminogroup stabilized by the neutral hydrogen bond with the amide oxygen atom of Asn149 and the hydrogen-bonded deprotonated side-chain carboxyl of Asp268.

According to the catalytic mechanism for OGG1 [29–31], Lys249 provides a nucleophile attack on the C1' of the deoxyribose moiety, displacing 8-oxoG and generating Schiff's base covalent intermediate [39]. The protonated side-chain of Lys249 in the free enzyme lacks this function that would require a lone pair of electrons as in the neutral $\epsilon\text{-NH}_2$. Therefore, the Lys249 residue should be deprotonated for the reaction to proceed. This would only be possible if the interaction with Asp268 stabilizing the protonated Lys249 is cleared upon binding to the substrate.

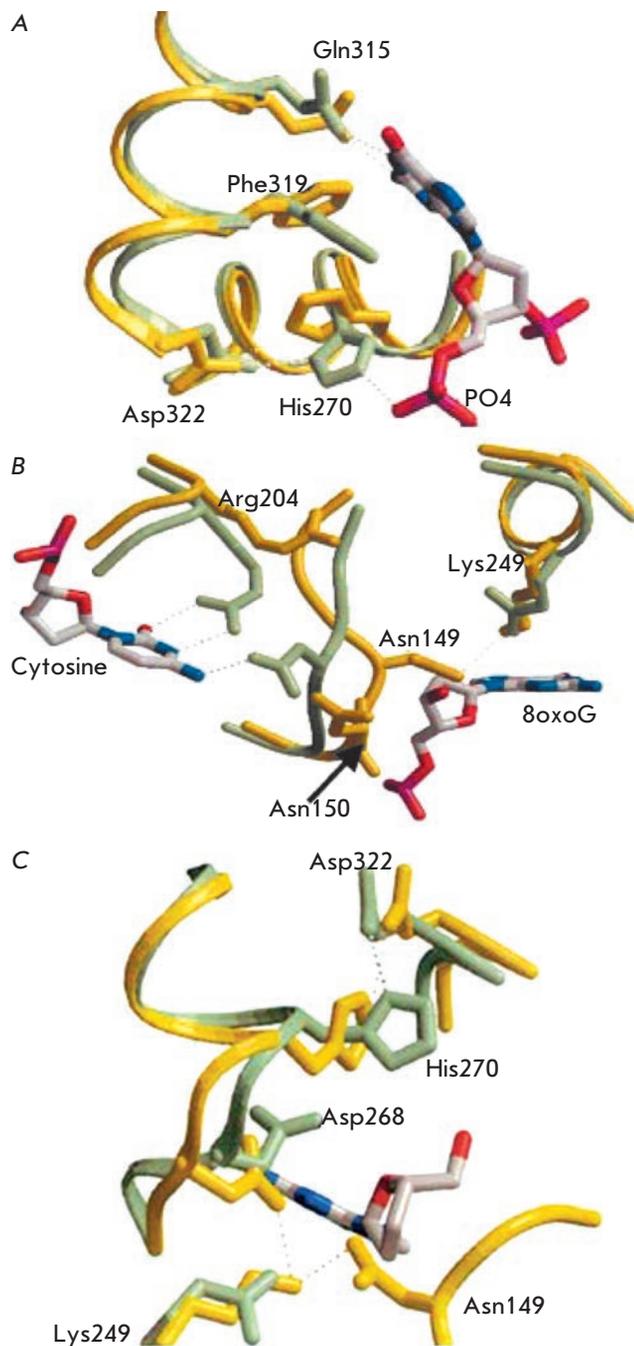


Fig. 5. Recognition of the 8-oxoG base by hOGG1 (the free enzyme is shown in yellow; hOGG1 in complex with DNA is shown in green). **A** – The conformation of His270 in native and DNA-bound structures together with interactions involving Gln315, Phe319, and Asp322 which form a trigger that switches between the closed and open states of the 8-oxoG binding pocket. **B** – Involvement of Asn149 in the recognition and binding of complementary cytosine. **C** – Conformational changes in the localization of Asp268 and His270 during binding of 8-oxoG. (Reprinted by permission from Elsevier Science Ltd.: [*J. Mol. Biol.*] Bjørås M., Seeberg E., Luna L., Pearl L.H., Barrett T.E. *J. Mol. Biol.* 2002. 317. 171–177, copyright 2002)

It was shown in [35] that binding of 8-oxo-dG promotes a change in the position of Asp268 (through movements in Phe319 and Phe270), which breaks its ion-pair/hydrogen bond to Lys249. At the same time, to promote the interaction between the enzyme and the estranged cytosine, the inter-helical hinge is extruded, accompanied by a loss of hydrogen bonding between Lys249 and the side-chain carbonyl group of Asn149. Following the removal of these neutralizing interactions, the protonated ϵ -amino group of Lys249 should be disfavored with regard to the neutral state generated by proton abstraction from the carboxyl group of Asp268, in parallel to its movement and loss of the hydrogen bond with Lys249.

Overall, the findings suggest [35] that the hydrogen bonds established by Asp268 and Asn149 with a protonated nitrogen of the ϵ -amino group of Lys249 act as trigger locks in the free enzyme. One hydrogen bond is involved in 8-oxoG recognition; the other, in the recognition of cytosine. Both should be removed for the enzyme to trigger a nucleophilic attack on the C1' of the deoxyribose.

THE ROLE OF Asp268 IN THE CATALYTIC ACTION OF hOGG1

Asp268 is catalytically important for both hOGG1 and other members of the structural family to which it belongs [26, 44]. Substitution of residue 268 for Ala or Asn abrogates the glycosylase and AP-lyase activities; however, substrate recognition is retained.

Initially, it was suggested that Asp268 could be important for oxidized base excision through deprotonation of Lys249, thereby converting the catalytically inactive cation into a nucleophilic neutral amine [36]. However, this is inconsistent with the fact that Asp located at the end of the α -protein helix needs to be less basic by several orders of magnitude. The fixed position should prevent the rotation of Asp268 around C $^{\alpha}$ -C $^{\beta}$ to interact with Lys249. The X-ray crystal structure of the native enzyme bound to AP-DNA [33] showed that Asp268 and Lys249 have no contact. Indeed, only the free form of the protein permits hydrogen bonding between the two, with Asp268 preserving its position in the α -helix, whereas Lys249 swivels about to enable contact [35]. Importantly, even though in the case of weakly hydrogen-bonded Glu268 and Lys249 (in mutant D268E) and a longer side-chain, the side-chain of Glu268 lacks contact with Lys249. It is likely that if the pK_a of Lys249 declines slightly due to the sequence context to produce a certain amount of neutral amine, this could be sufficient to break the glycosidic bond and excise the aberrant base.

To elucidate the role of Asp268 in the hOGG1-dependent catalysis, this residue was substituted for as-

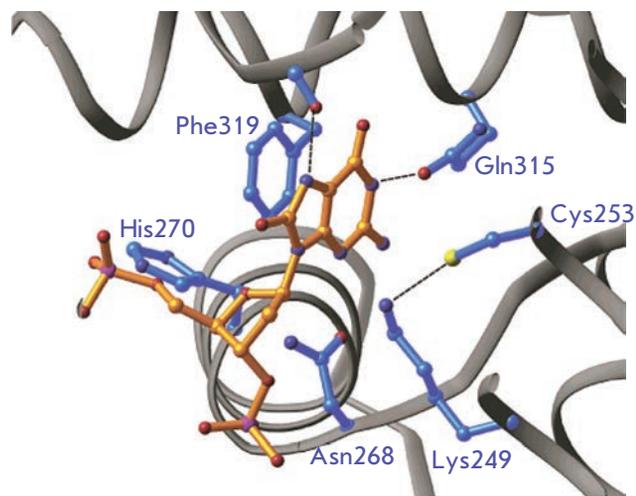


Fig. 6. Active-site structure of D268N hOGG1 in a complex with 8-oxoG-containing DNA. Visualization of PDB ID: 1N3C [38] (Reprinted by permission from American Chemical Society: [Biochemistry] Norman D.P.G., Chung S.J., Verdine G.L. *Biochemistry*. 2003. 42. 1564–1572, copyright 2003)

paragine (D268N), glutamate (D268E), and glutamine (D268Q) [38]. It was demonstrated that Asp268, located at the N-terminus of the α -helix (Fig. 6), plays a dual role in the catalysis of base excision and DNA strand scission. The mutation of this residue to asparagine led to a significant decline in the enzyme activity of D268N hOGG1. The D268N crystal structure first revealed the contribution of this nucleophilic Lys residue (Lys249) in the recognition of an aberrant guanine (8-oxoG or F).

Analysis of this structure allowed to suggest that 8-oxoG excision was performed through dissociative mechanism consisting of cleavage of the glycosidic bond, deprotonation of Lys249 (perhaps, by 8-oxoG anion) followed by subsequent linkage of Lys249, forming a Schiff base between the ϵ -NH $_2$ -group and C1' of the deoxyribose sugar. On the other hand, a change from aspartic acid to glutamine (D268Q hOGG1) and glutamate (D268E hOGG1), although it led to a protein fold disruption but did not abrogate catalytic activity.

It is shown [38] that complexes of Asp268 mutated hOGG1 with DNA are structurally similar to that of K249Q mutant hOGG1 [36]. The root mean square deviation of the protein backbone coordinates between the D268N and K249Q hOGG1 structures is 0.32 Å, which confidently lends credence to the identity of the structures. Comparison between the structure of wild-type and D268E mutant hOGG1 bound to F-analog yields root mean square deviations of 0.45 Å and 0.59 Å for WT hOGG1/F and hOGG1 D268Q/F structures,

respectively. This is indicative of the fact that no global conformational changes occur following mutations of active site amino acid. In all three structures, considerable alterations were only observed in the active site, especially in the first three residues of helix α -M (residues 269, 270 and 271). In all three structures, the nucleophilic ϵ -NH₂-group of Lys249 has a relatively fixed position, even though it is located at the end of a long alkyl chain. In this conformation, the ϵ -NH₂-group of Lys249 is located near the C1' of the deoxyribose moiety (3.4 Å), but in an orientation unfavorable to a straightforward nucleophilic attack on the glycosidic bond (Fig. 6). The ϵ -NH₂-group of Lys249 seems to be hydrogen-bonded to the sulfhydryl of Cys253, which is, by contrast, attached to the π -system of the 8-oxoG through a van der Waals interaction. Since Cys253 is located on the top face of the ribose moiety, Lys249 cannot form contact with Cys253 when the sugar is attacked from the underside. In the wild-type enzyme, the carbonyl group of Asn268 is hydrogen-bonded to the His270 side-chain, whereas in the D268N mutant, the NH₂-amide group of Asn268 is positioned too far off His270 for hydrogen bonding to occur. The absence of bonding clarifies the relatively weak binding of DNA to the D268N mutant variant, since His270 is in direct contact with the phosphate backbone on the 5'-side of the lesion. The remaining part of the active site and recognition pocket in D268N and K249Q mutant hOGG1 bound to 8-oxoG-DNA are structurally identical.

An alternative view of the role of Asp268 is the electrostatic stabilization of the positive charge on the O4' of the deoxyribose sugar in the transition state during base excision [38]. This is in agreement with the fact that in all hOGG1 structures with Asp at position 268, the carboxyl oxygen is positioned near the O4' of the sugar phosphate backbone DNA, at a distance of 3.2 Å. Due to the strong interaction with its α -helix, the orientation of the Asn268 side-chain in D268N is nearly the same as that of Asp268 in the wild-type enzyme. Consequently, a change of Asp to Asn at position 268 leads to the substitution of a negatively charged oxygen for a neutral NH₂ group, maintaining the position in the active site. The imide nitrogen of Asn268 and O4' of the deoxyribose moiety (3.4 Å) are positioned at a longer distance than would be expected the hydrogen bonding distance between these atoms; however, this is sufficient for O4' to experience the partially positive electrostatic field of the amide protons on Asn268. It is concluded [38] that the charge change from Asp (δ^-) to Asn (δ^+) at position 268 increases the transition state energy for base excision due to the enhanced positive charge on O4', which results in a significant decline in reaction rates.

Taken together, it was obtained [38] that hOGG1 containing Asn in place of Asp268 (D268N) exhibits no catalytic activity. The amino acid mutations D268Q or D268E, even though they confer a sterically disfavored conformation, do not diminish the catalytic activity. These findings argue against the role of Asp268 as an acid/basic catalyst in hOGG1 [36]; however, they support its involvement in the charge appearance on the O4' of the deoxyribose sugar in 8-oxoG.

RECOGNITION OF 8-oxoG

8-oxoG and the G base structurally differ by two positions: C8 carries O or H; and N7 contains H or a lone pair of electrons, respectively. For this reason, the H atom adjacent to N7 in 8-oxoG is capable of hydrogen bonding with the carboxyl of the Gly42 main chain, whereas G is not. To clarify the structural features of hOGG1 in complex with 8-oxoG and intact G, N149C mutant hOGG1, lacking catalytic activity, was probed tested [37]. Cys149 was connected with a linker to the C4 of cytosine complementary to 8-oxoG through a disulfide bond, thereby preventing the dissociation of the enzyme-DNA complex. The X-ray structures for N149C hOGG1 complexed with 8-oxoG-, G-, and 7-deaza-G-containing DNA duplexes have been solved.

The global structures of hOGG1 enzyme complexed with 8-oxoG- and G-containing DNA reveal a drastic helical axis kinks of $\sim 70^\circ$ and $\sim 80^\circ$, respectively, at the plane of the extruded base. The 8-oxoG base is located deeply in the active site pocket, whereas G lies against the enzyme surface at an exo-site positioned about 5 Å away from the pocket (Fig. 7A,B). The G base interacts with two active site residues Phe319 and His270 but these contacts differ from those made by 8-oxoG. In the 8-oxoG-containing complex, His270 is not in contact with the damaged base but forms hydrogen bond with its 5'-phosphate. In a G-containing DNA, His270 interacts with the π -system of the base, without hydrogen bonding to its phosphate.

Free energy calculations made for the interactions of 8-oxoG and G with the hOGG1 active site (ΔA_1 and ΔA_2 respectively) using quantum mechanics/molecular dynamics simulation techniques [37] showed that the value for free energy discrimination $\Delta\Delta A = \Delta A_1 - \Delta A_2$ is -6.8 kcal/mol, which indicates a 10^5 -fold preference for binding 8-oxoG in relation to G into the active site.

The hydrogen bond between the carbonyl oxygen of Gly42 and the proton at N7 in 8-oxoG strongly stabilizes the complex conformation. In the case of G, this will be replaced by Coulomb repulsion between the carbonyl oxygen of Gly42 and the lone pair of electrons at N7, if G and Gly42 are positioned similarly as in the 8-oxoG complex.

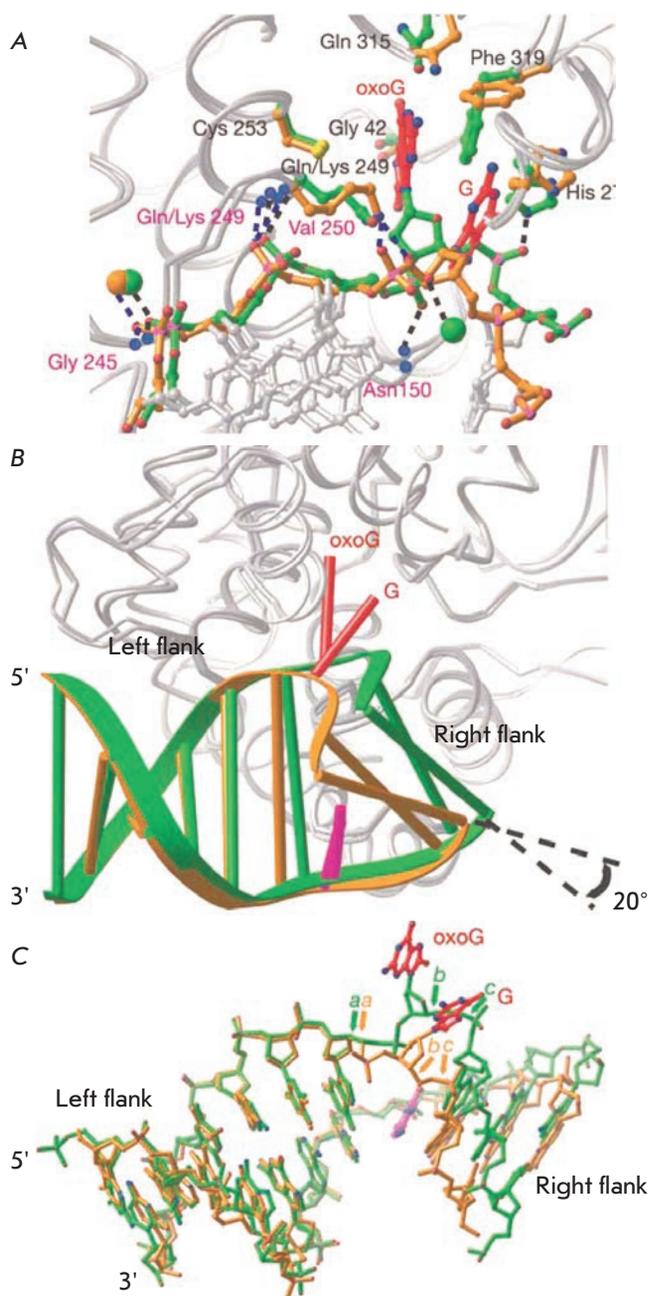


Fig. 7. Superposition of the N149C hOGG1 complexes with 8-oxoG-containing DNA (shown in green) and with the G-containing DNA (shown in yellow) in the region around the protein–DNA interface. **A** – Active site residues localization during interaction with the 8-oxoG and G bases. **B** – Changes in the structure of the DNA duplex depending on binding of damaged or undamaged bases. **C** – Comparison of the DNA in the two complexes, using the left flank for superposition. Arrows labeled **a**, **b**, and **c** indicate bonds that have undergone significant rotations: +110° for **a** (C4'–C5' bond of the residue 3' to 8-oxoG/G), +119° for **b** (C4'–C5' bond of 8-oxoG/G), and –151° for **c** (P–O5' bond of 8-oxoG/G). (Reprinted by permission from Macmillan Publishers Ltd: [Nature] Banerjee A., Yang W., Karplus M., Verdine G.L. *Nature*. 2005. 434. 612–618, copyright 2005)

The study [37] provides insight into the recognition mechanism for the lesion base. It is likely that base extrusion involves more than one step and cannot occur as a single process but rather progresses through multiple discrete states. This conclusion is supported by the fact that G-bound hOGG1 could mimic the intermediate produced upon binding to 8-oxoG, immediately before flipping out into the enzyme pocket. The multi-stage recognition mechanism for an aberrant base by hOGG1 is also corroborated in the study whereby 2-aminopurine and tryptophan fluorescence emissions were used for recording the dynamics of conformation changes [45, 46].

When compared, the structures of G- and 8-oxoG-DNA bound to the enzyme reveal events taking place at the final extrusion step. At the 3'-side (*Fig. 7*, left flank) of the damaged base the structures are quite similar. They retain the hydrogen bond contact with the HhH, containing Gly245, Gln249, and Val250, as well as electrostatic interaction with the divalent metal ion (*Fig. 7A*). The only difference is the 3'-phosphate of the extra-helical nucleoside hydrogen bonded to Lys249 in the G-containing complex; whereas in the 8-oxoG-bound structure, Lys249 is unable to approach the 3'-phosphate and must rotate to face the active site to promote catalysis. It has been suggested [37] that the contacts on the 3'-side of the lesion, even more likely that of the 3' phosphate of the DNA molecule with Lys249, are made before the base extrusion is complete. From the 5'-side of the extra-helical nucleoside, the helix conformation for the G- and 8-oxoG-structure varies. Consequently, when bound to G-DNA, the helix has a more pronounced bend (~80° versus ~70°); the duplex from the 5'-terminus is also over-rotated by ~20° about the vertical axis (*Fig. 7B*). This discrepancy is due to the loss of hydrogen bonds established between the 3'- and 5'-phosphates and the main chain NH-group of Asn150 and the side-chain NH-group of His270 in the complex with 8-oxoG-containing DNA. A divalent cation Ca²⁺ that coordinates the 3'-phosphate and stabilizes the bend by inner- and outer-sphere contacts to the bases flanking the extra-helical nucleoside is also missing in the G complex (*Fig. 7A*). These contacts in the 8-oxoG complex are formed only after the target base has been inserted into the lesion recognition pocket. Despite the apparent advances in understanding the structural features underlying the high specificity towards damaged base, it still remains unclear how hOGG1 recognizes 8-oxoG inside the DNA helix. According to [47] this issue could be addressed by identifying hOGG1 variants that recognize the intrahelical lesion base but with a diminished capacity for binding the lesion base outside the helix.

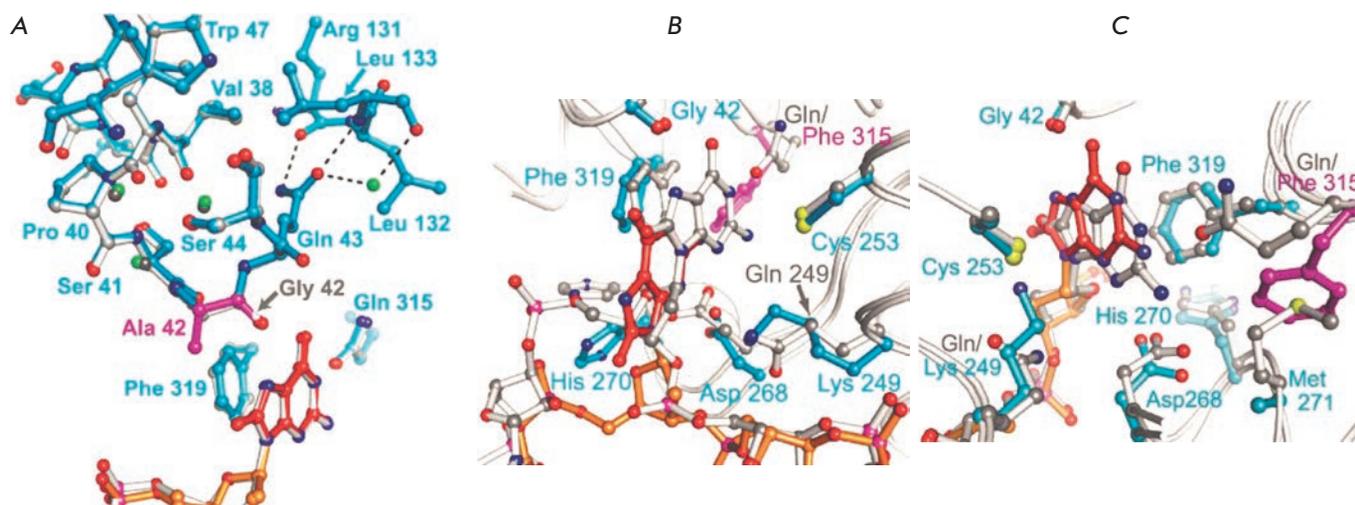


Fig. 8. Changes in the orientation of the active site amino acids in the complex of hOGG1 with 8-oxoG-containing DNA as a result of substitutions at the amino acids involved in formation of contacts with 8-oxoG. **A** – Superposition of amino acid residues in the hOGG1 active sites of the WT enzyme (gray backbone) and of the G42A hOGG1 (blue backbone). **B** – Active site amino acids in the Q315F*149 complex; 8-oxoG residue in the WT hOGG1 is shown in gray, in Q315F hOGG1 it is shown in red. **C** – Active-site amino acids in the Q315F*292 complex. (Reprinted by permission from the American Society for Biochemistry and Molecular Biology: [J. Biol. Chem.] Radom C.T., Banerjee A., Verdine G.L. *J. Biol. Chem.* 2007. 282. 9182–9194, copyright 2007)

A structural analysis has been carried out [47] for hOGG1 variants with mutated amino acids that are in contact with 8-oxoG. As in similar studies [37], the authors employed the disulfide cross-linking strategy for irreversibly linking cysteine in hOGG1 with the C4 of cytosine complementary to the oxidized guanine. Mutations at position His270, which interacts with the 5'-phosphate (H270A mutant hOGG1), and at position Gln315 contacting with the outer face of the DNA molecule (Q315A mutant hOGG1), do not affect the structure but eliminate its functionality.

On the other hand, an Ala substitution at position Gly42 (G42A mutant hOGG1), removing a specific contact with 8-oxoG [36], disfavors binding of hOGG1 to DNA. As mentioned above, Gly42 is the only residue in hOGG1 that directly distinguishes between G and 8-oxoG: N7-H of 8-oxoG forms hydrogen bond with the carbonyl oxygen of Gly42 [37]. A substitution of the hydrogen atom at C_α of Gly42 for a bulkier methyl group of Ala sterically impedes the binding of 8-oxoG in the active site pocket (Fig. 8A), followed by a conformational rearrangement in hOGG1.

A mutation to Q315F, which sterically disfavored the entry of 8-oxoG into the active site pocket, was studied for two variants: Q315F*149 and Q315F*292, wherein the cytosine of complementary strand was covalently linked to Cys149 or Cys292, respectively. The Q315F*149 hOGG1 variant actually failed to insert the

8-oxoG base into the recognition pocket but only in the exo-site (Fig. 8B). However, Q315F*292 allowed a nearly complete insertion of 8-oxoG, thus promoting hydrogen bond formation between Gly42 and N7-H of 8-oxoG (Fig. 8C). The authors suggest that it is the covalent bond to the remote Cys292 residue that stabilizes the weak interaction of 8-oxoG with the recognition pocket. Notwithstanding that 8-oxoG is inserted into the pocket, no cleavage takes place. Importantly, a Gln to Phe mutation in Q315F mutant hOGG1 abrogated the specificity for both intact and 8-oxoG-DNA.

To clarify the role of certain amino acid residues in the hOGG1 active site, participating in 8-oxoG binding, a photocleavable analog of 8-oxoG carrying the C6 *o*-nitrophenylisopropyl group (PC) was synthesized [32]. The use of this analog in a DNA substrate coupled with flash-freezing shed light on the structure of a very late-stage intermediate in the base excision.

The structure of the PC-bound enzyme does not differ from that of the G-bound complex; i.e., the modified PC-base was positioned in the exo-site of the hOGG1 enzyme [37]. Upon irradiation of the crystal with 373-nm laser light for 30 s at 4°C, the PC group was cleaved, unmasking a 8-oxoG/C base pair (FC-complex). Subsequent cryotrapping in liquid nitrogen and analysis of the captured structure demonstrated that the 8-oxoG localizes in the active site pocket in the same position as in the hOGG1/8-oxoG complex [36, 47].

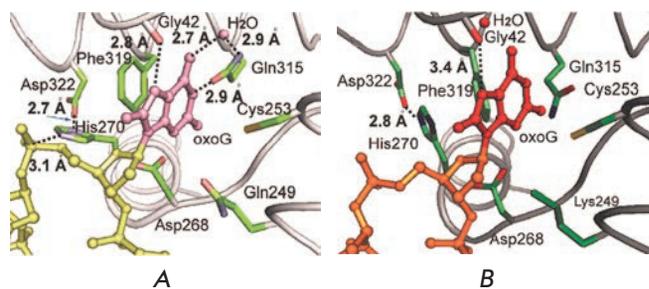


Fig. 9. Active site view of the hOGG1/8-oxoG in the LRC complex [36] (A) and in the FC complex (B). (Reprinted by permission from American Chemical Society: [J. Am. Chem. Soc.] Lee S., Radom C.T., Verdine G.L. *J. Am. Chem. Soc.* 2008. 130. 7784–7785, copyright 2008)

Notably, the hydrogen bond contact, recognizing the 8-oxoG base and formed by the N7-H of 8-oxoG and the carbonyl oxygen of Gly42 is retained, although the bond is longer than that in the 8-oxoG-bound substrate (LRC-complex, 3.4 versus 2.8 Å) (Fig. 9).

No other active site residues are in contact with 8-oxoG. The three key amino acids in hOGG1, known to establish contacts with 8-oxoG, in particular Phe319, Cys253, and Gln315, were shifted away from their positions in the FC-complex, as observed in earlier described structures [36, 47] (Fig. 4, Fig. 9A). In addition, the contact between His270 and the 5'-phosphate of 8-oxoG found in the LRC-complex is not observed in the FC-complex. This is replaced by His270 stacked with Phe319, whereas the hydrogen bond between His270 and Asp322 existing in the FC-complex is not seen in the hOGG1/8-oxoG lesion recognition complex (LRC). The catalytically important nucleophilic side-chain of Lys249 in the FC-complex is disordered and is not involved in the salt bridge with Cys253 (Lys249(NH₃⁺)/Cys253(S⁻)) predicted to play a role in 8-oxoG recognition [37]. The difference between the FC- and LRC-complexes is not limited with side-chain motions. The α -O helix, harboring the active site residues Gln315, Phe319, and Asp322, is retracted from the active site in the FC (Fig. 9B).

Overall, the structure studied in [32] is the very late-stage intermediate discovered to date in the DNA glycosylase reaction, wherein the 8-oxoG base has achieved nearly complete insertion, with the active site not yet accommodated for the base excision to occur (see also [46]). It is also evidenced [32] that the transition of the target base from the exo-site to the pocket proceeds much faster than the subsequent conformational rearrangements for the active site to achieve catalytic competence.

The strategy of disulfide cross-linking [37, 47] enabled the generation of DNA-enzyme adducts in crystal form for an X-ray analysis. The hOGG1 structure, in which the intact G was extruded out of the double-stranded DNA and inserted into the pocket, was obtained and analyzed [48]. No cleavage of the N-glycosidic bond was observed. The failure to break this bond was not due to the disulfide crosslink in the enzyme-DNA adduct, since the G to 8-oxoG replacement in this complex did promote base excision. These findings indicate a mechanism by which the G base is rejected late in the base excision pathway after it has been mistakenly inserted into the hOGG1 active site. This mechanism is triggered on those rare occasions when G overcomes the transition energy barrier from the exo-site to the active site pocket. The mechanism of the rejection of G remains unknown. It was earlier shown that the N-glycosidic bond in G nucleoside is more labile to hydrolysis at neutral pH than in 8-oxoG [49]. For this reason, the discrimination between the N-glycosidic bonds of G and 8-oxoG cannot be explained by a variation in bond stability.

There is also evidence in support of the existence of a mechanism by which G is rejected once it has been presented to the recognition pocket [46]. Using stopped-flow kinetics, the catalytic effects of hOGG1 with specific and non-specific DNA substrates were studied. The combination of tryptophan and 2-aminopurine fluorescence was used to follow the conformational dynamics of DNA, as well as the conformational transitions in the enzyme. The duplex DNA molecules used contained either 8-oxoG or a non-damaged G base. When binding to hOGG1, duplex DNA exhibited double helix disruption (judged by an increase in the 2-aminopurine fluorescence) at ~10 and 20 ms, respectively. This is due to DNA bending and flipping G or 8-oxoG out of the helix [46]. At > 20 ms in the case of 8-oxoG, a decline in the 2-aminopurine fluorescence was recorded, which was likely to correspond to the entry of amino acids into the DNA gap and complex stabilization. No similar changes in the 2-aminopurine fluorescence were observed with the G-base. This means that binding of hOGG1 to a non-specific DNA substrate could lead to DNA bending and eversion of the G base into the exo-site; however, G is rejected and the enzyme fails to achieve a catalytically competent state.

Verdine *et al.* [48] proposed that the energy barrier of the transition state for breaking the N-glycosidic bond of G could be higher than that for 8-oxoG because of the conformational changes in the active site or deprivation of the transition state stabilization (enzyme-G) through hydrogen bonding with Gly42. In addition, the G base in the active site is in a slightly different position relative to 8-oxoG as described recently [37]. This

could prevent the enzyme from adopting the optimal conformational state required for an attack on the C1' of the deoxyribose sugar by the Lys249 side-chain amino group. This is in good agreement with the finding indicating that introduction of active site mutations, that even slightly disturb 8-oxoG disposition, namely D268N [38] and Q315F [47], dramatically (but not completely) decreases the hOGG1 catalytic activity.

Lukina *et al.* [50] engineered C253L and C253I mutant hOGG1 forms with an occluded active site pocket by replacement of a Cys253 to a bulky leucine or isoleucine. Despite the perturbed active site geometry afforded by this mutation and dramatic decline in catalytic activity, the enzyme was still catalytically active. These findings argue for the concept of active site plasticity that postulates that the hOGG1 active site is flexible enough to mitigate the steric hindrance brought about by mutations [50].

The engineered complex [48] is structurally very similar to that produced with an 8-oxoG lesion. The differences in the interaction with G and 8-oxoG could be attributed to conformational adjustments to the lesion recognition pocket. It is suggested that while scanning for the target base, hOGG1 can occasionally flip the non-damaged G base into its active site [48]. The obtained data provide a basis to conclude that such erroneous insertions of G into the hOGG1 active site do not result in N-glycosidic bond cleavage due to the discrimination capacity between 8-oxoG and G at the catalytic stage.

There is a hypothesis that hOGG1 captures altered dynamic characteristics in the 8-oxoG-C base pair relative to the G-C base pair. This was tested by assessing the spontaneous opening of complementary pairs in double-stranded DNA using NMR spectroscopy and proton exchange [51]. It was demonstrated that the rate of spontaneous opening of 8-oxoguanine and the lifetime of the base in the extra-helical state are the same as those of a canonical guanine-cytosine base pair. This

finding does not support the role of the opening dynamics of 8-oxoguanine in the recognition of the lesion by DNA glycosylases.

CONCLUSIONS

In this review, we have focused on the structures of the DNA repair enzyme human 8-oxoguanine-DNA-glycosylase in free form and in complexes with DNA substrates. The currently available literature and data on the 3D structures of hOGG1 deposited in the Protein Data Bank (PDB) have been summarized. Lys249 and Asp268 are shown to be the key amino acids responsible for catalysis. Gly42, Asn149, Cys253, His270, Gln315, and Phe319 are the amino acids important for discrimination between 8-oxoG and G and for binding the target base in the active site pocket.

The hOGG1-mediated mechanism of oxidized base excision is reviewed in terms of structural dynamics. It is obvious that the eversion of damaged base from the DNA helix into active site of enzyme cannot occur as a concerted one-step process but, it seems to proceed through multiple discrete states. It is reasonable to hypothesize that base-specific cleavage by hOGG1 is controlled throughout the interaction: lesion recognition, base extrusion, binding of 8-oxoG into the active site pocket, and the catalytic hydrolysis of the N-glycosidic bond.

This work was supported by the Russian Foundation for Basic Research (Grants № 13-04-00013 (OSF), 14-04-00806 (VVK), 14-04-00531 (DGK)), the RAS Program Molecular and Cellular Biology (№ 6.11), the Grants Council under the President of the Russian Federation (NSh-1205.2014.4 (DGK)), and the Ministry of Education and Science of the Russian Federation (Joint Laboratory Award of Novosibirsk State University and the Novosibirsk Research Center, Siberian Branch of the Russian Academy of Sciences).

REFERENCES

- Wallace S.S. // *Free Radic. Biol. Med.* 2002. V. 33. № 1. P. 1–14.
- Marnett L.J. // *Carcinogenesis*. 2000. V. 21. № 3. P. 361–370.
- Dizdaroglu M., Jaruga P., Birincioglu M., Rodriguez H. // *Free Radic. Biol. Med.* 2002. V. 32. № 11. P. 1102–1115.
- Boiteux S., Guillet M. // *DNA Repair (Amst.)*. 2004. V. 3. № 1. P. 1–12.
- Cooke M.S., Evans M.D., Dizdaroglu M., Lunec J. // *FASEB J.* 2003. V. 17. № 10. P. 1195–1214.
- Evans M.D., Dizdaroglu M., Cooke M.S. // *Mutat. Res.* 2004. V. 567. № 1. P. 1–61.
- Xie Y., Yang H., Cunanan C., Okamoto K., Shibata D., Pan J., Barnes D.E., Lindahl T., McIlhatton M., Fishel R., Miller J.H. // *Cancer Res.* 2004. V. 64. № 9. P. 3096–3102.
- Wan J., Bae M.-A., Song B.-J. // *Exp. Mol. Med.* 2004. V. 36. № 1. P. 71–77.
- Gu Y., Desai T., Gutierrez P.L., Lu A.-L. // *Med. Sci. Monit.* 2001. V. 7. № 5. P. 861–868.
- Raha S., Robinson B.H. // *Trends Biochem. Sci.* 2000. V. 25. № 10. P. 502–508.
- Halliwell B., Gutteridge J.M.C. *Free Radicals in Biology and Medicine*. Oxford: Oxford Univ. Press, 2002.
- Jezeq P., Hlavata L. // *Int. J. Biochem. Cell. Biol.* 2005. V. 37. № 12. P. 2478–2503.
- Bernards A.S., Miller J.K., Bao K.K., Wong I. // *J. Biol. Chem.* 2002. V. 277. № 23. P. 20960–20964.
- Kasai H., Nishimura S. // *Nucleic Acids Res.* 1984. V. 12. № 4. P. 2137–2145.

REVIEWS

15. Shibutani S., Takeshita M., Grollman A.P. // *Nature*. 1991. V. 349. № 6308. P. 431–434.
16. Grollman A.P., Moriya M. // *Trends Genet.* 1993. V. 9. № 7. P. 246–249.
17. Michaels M.L., Miller J.H. // *J. Bacteriol.* 1992. V. 174. № 20. P. 6321–6325.
18. Fowler R.G., White S.J., Koyama C., Moore S.C., Dunn R.L., Schaaper R.M. // *DNA Repair (Amst.)*. 2003. V. 2. № 2. P. 159–173.
19. Sakumi K., Furuichi M., Tsuzuki T., Kakuma T., Kawabata S., Maki H., Sekiguchi M. // *J. Biol. Chem.* 1993. V. 268. № 31. P. 23524–23530.
20. Slupska M.M., Baikalov C., Luther W.M., Chiang J.-H., Wei Y.-F., Miller J.H. // *J. Bacteriol.* 1996. V. 178. № 13. P. 3885–3892.
21. Lu R., Nash H.M., Verdine G.L. // *Curr. Biol.* 1997. V. 7. № 6. P. 397–407.
22. Cappelli E., Hazra T., Hill J.W., Slupphaug G., Bogliolo M., Frosina G. // *Carcinogenesis*. 2001. V. 22. № 3. P. 387–393.
23. Radicella J.P., Dherin C., Desmaze C., Fox M.S., Boiteux S. // *Proc. Natl. Acad. Sci. USA*. 1997. V. 94. P. 8010–8015.
24. Roldan-Arjona T., Wei Y.F., Carter K.C., Klungland A., Anselmino C., Wang R.P., Augustus M., Lindahl T. // *Proc. Natl. Acad. Sci. USA*. 1997. V. 94. P. 8016–8020.
25. Aburatani H., Hippo Y., Ishida T., Takashima R., Matsuba C., Kodama T., Takao M., Yasui A., Yamamoto K., Asano M. // *Cancer Res.* 1997. V. 57. P. 2151–2156.
26. Rosenquist T.A., Zharkov D.O., Grollman A.P. // *Proc. Natl. Acad. Sci. USA*. 1997. V. 94. P. 7429–7434.
27. Nishioka K., Ohtsubo T., Oda H., Fujiwara T., Kang D., Sugimachi K., Nakabeppu Y. // *Mol. Biol. Cell.* 1999. V. 10. P. 1637–1652.
28. Boiteux S., Radicella J.P. // *Arch. Biochem. Biophys.* 2000. V. 377. P. 1–8.
29. Kow Y.W., Wallace S.S. // *Biochemistry*. 1987. V. 26. № 25. P. 8200–8206.
30. Dodson M.L., Michaels M.L., Lloyd R.S. // *J. Biol. Chem.* 1994. V. 269. № 52. P. 32709–32712.
31. Labahn J., Schäfer O.D., Long A., Ezaz-Nikpay K., Verdine G.L., Ellenberger T.E. // *Cell*. 1996. V. 86. № 2. P. 321–329.
32. Lee S., Radom C.T., Verdine G.L. // *J. Am. Chem. Soc.* 2008. V. 130. № 25. P. 7784–7785.
33. Norman D.P.G., Bruner S.D., Verdine G.L. // *J. Am. Chem. Soc.* 2001. V. 123. № 2. P. 359–360.
34. Fromme J.C., Bruner S.D., Yang W., Karplus M., Verdine G.L. // *Nat. Struct. Biol.* 2003. V. 10. № 3. P. 204–211.
35. Bjørås M., Seeberg E., Luna L., Pearl L.H., Barrett T.E. // *J. Mol. Biol.* 2002. V. 317. № 2. P. 171–177.
36. Bruner S.D., Norman D.P.G., Verdine G.L. // *Nature*. 2000. V. 403. № 6772. P. 859–866.
37. Banerjee A., Yang W., Karplus M., Verdine G.L. // *Nature*. 2005. V. 434. № 7033. P. 612–618.
38. Norman D.P.G., Chung S.J., Verdine G.L. // *Biochemistry*. 2003. V. 42. № 6. P. 1564–1572.
39. Nash H.M., Bruner S.D., Schärer O.D., Kawate T., Addona T., Sponner E., Lane W.S., Verdine G.L. // *Curr. Biol.* 1996. V. 6. № 8. P. 968–980.
40. Thayer M.M., Ahern H., Xing D., Cunningham R.P., Tainer J.A. // *EMBO J.* 1995. V. 14. № 16. P. 4108–4120.
41. Guan Y., Manuel R.C., Arvai A.S., Parikh S.S., Mol C.D., Miller J.H., Lloyd S., Tainer J.A. // *Nat. Struct. Biol.* 1998. V. 5. № 12. P. 1058–1064.
42. Bjørås M., Luna L., Johnsen B., Hoff E., Haug T., Rognes T., Seeberg E. // *EMBO J.* 1997. V. 16. № 20. P. 6314–6322.
43. The PyMOL Molecular Graphics System. Version 1.6.0.0 Schrödinger, LLC.
44. Krokan H.E., Standal R., Slupphaug G. // *Biochem. J.* 1997. V. 325. Pt. 1. P. 1–16.
45. Kuznetsov N.A., Koval V.V., Zharkov D.O., Nevinsky G.A., Douglas K.T., Fedorova O.S. // *Nucleic Acids Res.* 2005. V. 33. № 12. P. 3919–3931.
46. Kuznetsov N.A., Koval V.V., Nevinsky G.A., Douglas K.T., Zharkov D.O., Fedorova O.S. // *J. Biol. Chem.* 2007. V. 282. № 2. P. 1029–1038.
47. Radom C.T., Banerjee A., Verdine G.L. // *J. Biol. Chem.* 2007. V. 282. № 12. P. 9182–9194.
48. Crenshaw C.M., Nam K., Oo K., Kutchukian P.S., Bowman B.R., Karplus M., Verdine G.L. // *J. Biol. Chem.* 2012. V. 287. № 30. P. 24916–24928.
49. Bialkowski K., Cysewski P., Olinski R. // *Z. Naturforsch.* 1996. V. 51. № 1–2. P. 119–122.
50. Lukina M.V., Popov A.V., Koval V.V., Vorobjev Y.N., Fedorova O.S., Zharkov D.O. // *J. Biol. Chem.* 2013. V. 288. № 40. P. 28936–28947.
51. Every A.E., Russu I.M. // *J. Mol. Recognit.* 2013. V. 26. № 4. P. 175–180.

The Proteome of a Healthy Human during Physical Activity under Extreme Conditions

I. M. Larina^{1*}, V. A. Ivanisenko², E. N. Nikolaev³, A. I. Grigorev¹

¹SSC RF Institute for Biomedical Problems, Russian Academy of Sciences, Khoroshevskoye shosse, 76a, 123007, Moscow, Russia

²Institute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Akad. Lavrentiev Ave., 10, 630090, Novosibirsk, Russia

³Institute of Biochemical Physics, Russian Academy of Sciences, Kosygina Str., 4, 119334, Moscow, Russia

*E-mail: irina.larina@gmail.com

Received 08.04.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT The review examines the new approaches in modern systems biology, in terms of their use for a deeper understanding of the physiological adaptation of a healthy human in extreme environments. Human physiology under extreme conditions of life, or environmental physiology, and systems biology are natural partners. The similarities and differences between the object and methods in systems biology, the OMICs (proteomics, transcriptomics, metabolomics) disciplines, and other related sciences have been studied. The latest data on environmental human physiology obtained using systems biology methods are discussed. The independent achievements of systems biology in studying the adaptation of a healthy human to physical activity, including human presence at high altitude, to the effects of hypoxia and oxidative stress have been noted. A reasonable conclusion is drawn that the application of the methods and approaches used in systems biology to study the molecular pattern of the adaptive mechanisms that develop in the human body during space flight can provide valuable fundamental knowledge and fill the picture of human metabolic pathways.

KEYWORDS integrative physiology, space flight, proteomics, systems biology.

ABBREVIATIONS OMICs – biological disciplines integrated in the group of post-genomic technologies, with the names ending in *-omics*; MALDI – matrix-assisted laser desorption/ionization; ESI – electrospray ionization; PCR – polymerase chain reaction; HUPO – Human Proteome Organization; C-HPP – Chromosome-Centric Human Proteome Project; HLPP – Human Liver Proteome Project; KEGGDB – Kyoto Encyclopedia of Genes and Genomes database; PGC-1 α – peroxisome proliferator-activated receptor- γ coactivator 1 α ; HIF – hypoxia-inducible factor; HSP70 – heat shock protein 70 kDa; PDIA3 – protein disulfide isomerase family A, member 3; ROS – reactive oxygen species; CV – coefficient of variation.

INTRODUCTION

Proteomics¹ appeared in the late XXth century as a set of methods for the large-scale study of proteins [1]. Proteomics emerged as a result of a gradual development and sophistication of the classical methods used to study proteins, starting from gravimetry and photometry to disc electrophoresis, gradient, and 2D electrophoresis [2–4]. A considerable leap in the pace of development of this field of protein research took place after the possibility of using mass spectrometry to identify protein molecules was discovered. New methods of protein ionization without disturbing their

primary structure appeared in the late 1980s; namely, matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI) [5, 6]. The term “proteomics” recently appeared to identify the branch of systems biology that studies the protein composition of cells, tissues, body fluids, and organisms using primarily high-performance methods of mass spectrometry.

To date, tremendous progress has been made in technologies that enable one to identify proteins, measure their concentration in a sample, determine their abundance in cells, tissues and organisms, and also reveal post-translational modifications². The number of

¹ The term was first introduced by P. James in 1997, by analogy with genomics.

² Chemical modifications of amino acid residues after translation.

peptides and proteins that can be identified and quantified has been steadily increasing (www.SwissProt.com).

Yet, despite the impressive achievements in implementing mass spectrometry technologies in experimental biology, the results obtained using proteomic methods have had a lesser impact on medicine than those obtained with genomic research. In our opinion, this is mainly because proteomic problems are more complicated due to the lack of methods for the amplification of protein molecules similar to PCR, and also because of the fact that the amount of proteins is much greater than that of genes (by many orders of magnitude). There are also some other complicated reasons. No doubt, such a gap between theory and practical application impedes the rate at which such discoveries as the revealing of genetic functions etc. are implemented into practical knowledge and used in clinical medicine.

DEVELOPMENT OF PROTEOMICS: ACHIEVEMENTS AND COMPLICATIONS

The status of post-genome OMICs¹, including proteomics, has been summarized in detail in numerous reviews [7–13]. Thus, Lander [10] noted that the decade of the post-genomic era has been characterized by intensive accumulation and cataloging of data on full sets of cellular components as a result of the intensive development of the global study of the structures of genomes, proteomes, transcriptomes, metabolomes and other “-omes.” However, the existing gap between structural success and functional implementation – understanding the functions that characterize the vital activities and the disorders in pathological conditions – devalue such discoveries. The functional analysis remains critical today for advancing towards practical application [10]. This statement has also been supported by Alberts [8].

As for proteomics, Bensimon *et al.* [7] emphasize four reasons for the lag in practical implementation (these reasons are of conceptual and technical nature). First of all, many scientists find mass spectrometry technologies to be quite complicated and requiring expensive equipment that needs constant improvement. The same can be attributed to genomics and its high-performance technologies, but the process of obtaining results using a proteomic analysis is nonlinear and it utilizes several different protocols; therefore, proteomic methods in reality appear to be objectively more complex. In summary, one can conclude that the progress achieved in proteomics is closely related to

the improvement in mass spectrometry methods and increasing the accessibility to mass spectrometry-based instruments of proteomics for the wide range of scientists involved in the field of proteomics and the adjacent fields. Indeed, an analysis of the published data shows that a significant portion of high-quality results in the field of proteomics are generated by a relatively small number of laboratories. It has been noted [7] that, nowadays, from 7,000 to 10,000 human proteins can be reliably identified, and this without taking into account major proteins (present at high concentrations in samples).

When the analysis of cell line proteomes became possible, such an approach drew the attention of many researchers. However, the studies identified only approximately 100 high-abundance proteins [14–17].

Second, the studies that use high-performance methods of mass spectrometry conducted in order to identify markers show no significant advantages in the case of hypothesis-driven research, which remains the major method in life sciences. The repeating cycles of experiments with the generation and testing of hypotheses using proteomic data sets do not allow one to arrive at the expected benefit at the initial stage of the discovery of markers.

Third, it has become generally accepted that the cataloging of proteins in a sample or the predicting of their potential synthesis from a gene located on a specific chromosome (which is the main goal of the initiative of the Chromosome-Centric Human Proteome Project (C-HPP) by HUPO) are needed at the current stage but are insufficient for a biological understanding of the physical and functional interaction between proteins under conditions of dynamic molecular networks, where identifying the function of specific proteins is as important as determining the structure and function of individual proteins [2, 10, 17]. The understanding that biological processes should be studied using the dynamic networks of interacting molecules and changes in the network structure or topology determine the phenotype and underlies the new field of systems biology currently under development [7].

Fourth, the technical limitations of mass spectrometry (as the main “breakthrough” tool in proteomics) in terms of the data integrity and reproducibility of peptide identification and the protein correspondence decreases the value of the results of comparing the proteomic data obtained by various researchers and laboratories. As a result, specialized mass spectrometry teams generate large, high-quality datasets that are difficult or simply impossible to interpret and apply today, while the overwhelming majority of researchers in the field of life sciences perform analyses of small sets

¹ OMICs – collective name of proteomics, transcriptomics, peptidomics, metabolomics and other post-genomic disciplines.

of proteins using methods that were developed decades ago (e.g., Western blot [18] and enzyme-linked immunosorbent assay).

It is important to mention that the number of human proteins predicted using the genomic sequence and identified experimentally increases but not at such an impressive rate as it has been predicted. For example, 11 years of HLPP (Human Liver Protein Project) studies led to the discovery of 12,168 proteins of the hepatic tissue and organelles from four main types of the hepatic tissue, and the number is expected to reach 13,000 by 2015 [19]. Currently, a total of 20,128 nucleotide sequences encoding proteins in the human genome, among which the existence of 15,646 proteins has been experimentally confirmed (75%, assuming that 1 gene = 1 protein), have been found [20, 21]. The existence of the so-called “missing proteins,” as well as the degree of uncertainty arising from the lack of a strict correspondence between the number of genes and proteins and other, already known, molecular biological principles complicate the study of the human proteome.

The proteins in the organism do not function alone. They form multiprotein complexes on the one hand and complex functional and dynamic networks on the other hand [22–25]. The organization into functional modules demonstrates the complexity and diversity of the proteome at the subcellular, cellular, and tissue levels. Understanding the organization and function of protein networks, which describe the molecular mechanisms of biological processes, is necessary to clarify the regulation (and maintenance) of the degree of health and reserves, and well as the development of human diseases (oncologic, neurodegenerative, cardiovascular ones, etc.). The study of protein interactions, including their association with non-protein molecules, and analysis of the protein networks formed by protein–protein interactions are important tools for the diagnosis, determination of disease pathogenesis, and search for molecular targets for therapeutic interventions [26, 27].

In addition, since most eukaryotic proteins are modular and polyfunctional [28–30], a protein acquires the ability to accomplish a range of different functions participating in various pathways. In these circumstances, eukaryotic protein networks usually interfere with each other [22, 31–33]. The tasks of evaluating the interaction between the molecular components of a biological system and the integration of such information into systems of networks and pathways that can be used to develop models, to predict the behavior of the system, constitute a serious challenge for researchers who develop bioinformatic methods of analysis for proteomics [22, 25, 33–36].

PROTEOMICS AND SYSTEMS BIOLOGY

Proteomics and other OMICs disciplines (genomics, transcriptomics, and metabolomics) are not only new research tools and new measurement possibilities. Their emergence and development have brought new meaning to systems biology.

In their review, Edwards and Thiele [37] stated the following about the meaning of the term “systems biology”: “If then it is nothing new, why is systems biology suddenly so visible? Some have implicitly argued [38, 39] that systems biology is a mirage, no more than a rebranding of the type of holistic thinking that some biologists and integrative physiologists have been using for decades.” As it often happens to scientific terms, the meaning of the concept “systems biology” in its current form is different from the one used previously in the aforementioned sciences. Thus, although the term “systems biology” is not new, its meaning has now changed. This is attributed to the development of technologies, especially genome sequencing, and computational and analytical platforms, such as mass spectrometry and nuclear magnetic resonance. In order to truly study a large system in its entirety, one requires the ability to fully model and measure it. Prior to the sequencing of total genomes, this had been an insurmountable experimental challenge for biologists. With the enhancement of computational power during data processing and development of genomics, transcriptomics, proteomics, metabolomics, and fluxomics¹, it has become possible to study the “profiles” and models of a biological system or subsystem in its entirety. Systems biology and the so-called OMICs disciplines are not identical. Systems biology, using most of the OMICs data, goes beyond the scope of these methods [40–42]. At the same time, the term “systems biology” is understood rather narrowly by most scientists who use genomics and OMICs as a complex approach that utilizes experimental data obtained at different levels of life organization. This is mostly due to the specific understanding of the term, giving it a general meaning. A well-known physiologist, Noble *at al.* [43], in full agreement with other specialists, defines systems biology as an approach but not a field of science. Meanwhile, scientists who work in the field of systems biology consider it to be a scientific discipline that tries to study biological systems in a holistic, rather than a reductionist, way. This includes the collection of dynamic global datasets, along with phenotypic data from various levels of the biological information hierarchy, in order to identify and

¹ *Methods of mathematical description or prediction of metabolic reaction rates in biological systems that are considered to be a key novel computing technology.*

explain the mechanisms of the emerging characteristics¹ of the system [9, 25, 44].

MATHEMATICAL MODELING AND COMPUTER TECHNOLOGIES IN MODERN BIOLOGY

One has to admit that the most fundamental difference between systems biology and OMICs disciplines and other related sciences, such as integrative physiology, is the central role of mathematical modeling and computer technologies [45–47]. It is impossible today to consider all cellular processes and simultaneously study the molecular mechanisms of a process or phenomenon even when using high-performance experimental methods. Systems biology provides the tools for solving such problems, since it incorporates the methods of mathematical and computer modeling [48]. First of all, they is the method of molecular dynamics [49], interactome mapping² (including experimental methods) [22], and the development of special algorithms for the construction, design, and visualization of intra- and intercellular processes and phenomena [51–53], which are used in computer modeling. It is obvious that understanding of the term “systems biology” has significantly changed, and a number of authors note that a consensus is emerging on that subject [42, 45, 54]. Given the vast amounts of data generated by the methods of genomics, transcriptomics, proteomics, and metabolomics, their manual interpretation by a researcher without the help of bioinformatics is completely ineffective or mostly impossible, and inevitably this calls for a need to apply the methods and principles of systems biology.

The genome-scale construction of the metabolic processes in a human organism with further reconstruction, published in 2007, was named Recon 1 [55]. Recon 1 provided the description of 1,496 open reading frame functions for 2,766 metabolites and 3,311 reactions located in seven intercellular compartments (cytoplasm, mitochondria, nucleus, endoplasmic reticulum, Golgi apparatus, lysosome, and peroxisome) and in the extracellular environment. This first comprehensive, genome-scale human metabolic reconstruction covers most of the known central metabolic pathways occurring in any human cell [60]. Moreover, the reconstruction made in 2007 served as a starting point for further tissue- and cell-type specific metabolic reconstructions using the data generated

in OMICs studies (e. g., transcriptomic and proteomic data). Now these reconstructions are made for human macrophages [61], hepatocytes [62], myocytes, and adipocytes [63].

The major issue confronting systems biology as concerns human physiology is to fill the gaps existing in molecular networks up to their complete reconstruction.

One of the methods to pinpoint the missing reactions in the reconstructing network is to compare the results of model calculations with experimental data [64]. Numerous computational algorithms to apply in this method have been published [64–67]. Moreover, the metabolomics data of cells, tissues, and body fluids [68–70] can be used to reveal the missing links in the human metabolism. There are several different computational approaches [65, 67] that are used to find the missing candidate protein in a reaction and the corresponding genes [71, 72]. Such computational methods determine one or several reactions that take place in the organisms of other species collected in the universal protein interactions database, such as the ligand database KEGG [73], and they add them into a metabolic model, thus filling the gaps in potential missing knowledge. If experimental confirmation cannot be found in the scientific literature, one has to predict the missing genes and reactions and formulate hypotheses that require experimental testing.

EXPERIMENTAL METHODS IN HUMAN SYSTEMS BIOLOGY

How are such methods helpful to a researcher studying extreme conditions (and what is the value in collecting data on human physiology in extreme environments employing such methods for systems biology)? The concept of disturbances plays a key role in systems biology [45]. Systems biology is based on: (1) the ability to measure all variables of interest (OMICs); (2) the presence of a conceptual framework for data interpretation (there are models); (3) and the application of the disturbance method in the experiment. Such effects on the organism, which are capable of disturbing homeostasis, allow one to define the mechanisms that help maintain a constant internal environment and preserve health resources, and the adaptive potential of an organism. However, if the object of the research is a healthy human, then the list of methods (conditions), ethically appropriate and available for exposure that lead to the decline of his homeostasis, would be quite short and would include physical activity, the use of pharmaceuticals, nutrition manipulations (e. g., the use of lipid emulsions [70] or directive changes in salt consumption [74]), functional exercise testing [75], environmental stud-

¹ These are the characteristics that appear in complex systems as a result of the interaction between their components; they cannot be predicted based on the characteristics of individual components. If the organization of a complex system is hierarchical, the characteristics influence each other as “top to bottom” and “bottom to top” [13, 56–59].

² Interactome, as the whole set of protein–protein interactions in a cell or an organism, is more complex than the proteome, and has been recently used as the measure of organism complexity [50].

ies, including exposure to extreme temperatures, high barometric pressure and hypoxia, and, finally, space flights. Thus, extreme conditions are among the few ways to cause a decline in homeostasis in a healthy human and provide “experimental” data for systems biology. Therefore, we suggest that ecological and gravitational physiology and human systems biology are natural symbionts.

ENVIRONMENTAL STUDIES AND SYSTEMS BIOLOGY

Only a few scientists admit the benefit of experimental data on the influence of physical exercises on a human organism for systems biology [76, 77]. At the same time, the application of systems biology methods enables one to display the whole set of proteins and metabolites inherent to the stress phase (throughout a 1.060-km non-stop cycling event) [78], and also to reveal the mechanism underlying the phenotypic response to physical activity (during the process of adaptation of fish white and red myofibrils to the training load), with the activation of metabolic networks in white myofibrils (catabolism of carbohydrates, protein synthesis, muscle contraction, and detoxification) and insufficient expression of others in red myofibrils (responsible for energy production, muscle contraction, and maintenance of homeostasis) [79]. Application of the analytical capabilities of various OMICs provided the data that helped prove the role of PGC-1 α as a transcriptional coactivator that coordinates the activation of metabolic genes (responsible for mitochondrial biogenesis) in human skeletal muscle in response to physical activity [80]; allowed one to determine the connection between genome-mediated muscle plasticity and modulation of hypoxia-specific mitochondrial biogenesis [81]; and also to establish the metabolic pathways that are activated during physical exercise, revealing several dynamically regulated miRNA–mRNA networks [82].

Systems biology studies of environmental human physiology have been started [42, 47]. Some interesting studies in the field of high-altitude genetics and proteomics have been carried out. Several studies have convincingly shown that human populations living at high altitudes experience genetic divergence. Thus, the Tibetans whose ancestors have been living at high altitudes for more than 10,000 years have acquired and inherited new mutations of the gene encoding the oxygen-sensitive hypoxia-inducible factor (HIF) [83, 84]. Studies of the proteome of the skeleton muscle biopsy samples obtained from volunteers who had spent one week at an altitude of 4,500 m [85], carried out using 2D gel electrophoresis, revealed a much larger number of proteins (involved in iron transfer and oxidative metabolism), the number of which significantly

differs from that of experienced climbers after a stay at a much higher altitude. The changes in the urinary peptidome [86] and human plasma proteome [87] in response to high-altitude exposure have also been studied. In the latter case, special attention was paid to the identification of high-altitude pulmonary edema biomarkers. These and similar studies provide building blocks for coordinated efforts in systems biology and physiology in understanding the human physiological reaction to high altitudes. The extensive data on experimental hypoxia, including experiments using the methods of systems biology, have been reported in [88–95].

The ascent of a man to the highest mountain peaks initiated a surge of studies in the field of the physiological outcomes of physical activity at high altitudes. It has been established that the catabolic effects of chronic exposure to a hypoxic environment on muscles are a result of insufficient activation of hypoxia-sensitive signaling pathways and suppression of the energy-intensive processes of protein translation [96]. The study of the proteome modulation caused by hypobaric hypoxia allowed one to establish that efficient use of energy-generating pathways in conjugation with an abundance of antioxidant enzymes makes the cortex less vulnerable to hypoxia than the hippocampus [97]. The experimental study of pulmonary hypertension under hypobaric hypoxia conditions showed the characteristic structural remodeling in lungs, the mechanism of which involves isoforms of heat shock protein 70 (HSP70) and protein disulfide isomerases A3 (PDIA3) [98].

The studies of the mechanisms of oxidative stress and, more broadly, cell redox homeostasis have conclusively proved the dual role of reactive oxygen species (ROS) [99]. Uncontrolled overproduction of ROS damages cellular structures, including membrane lipids, proteins, and DNA [100]. At the same time, there is increasing evidence that ROS act as secondary messengers of intracellular signaling cascades, which can cause and maintain the oncogenic phenotype of cancer cells, but they are also involved in senescence and apoptosis [101]. Intensive studies in this field have even led to a change in the definition of the term “oxidative stress,” making the process dependable on the changes in the real post-translational thiol modification of proteins [102, 103]. The damage to ROS-induced signaling pathways has pathophysiological consequences that manifest themselves as disease progression (cardiovascular disorders, atherosclerosis, hypertension, ischemia-reperfusion injury, diabetes, neurodegenerative diseases – Alzheimer’s and Parkinson’s, rheumatoid arthritis). A positive role of ROS is its ability to protect

against infectious agents through the non-specific activation of T- and B lymphocytes, participate in the functioning of numerous signaling pathways, and induce mitogenesis [100].

HUMAN GRAVITATIONAL PHYSIOLOGY AND SYSTEMS BIOLOGY

Finally, space flight (its influence has been studied by physiologists and physicians for more than 50 years) can be regarded as an unprecedented in the history of human evolution experience of adaptation of a healthy person to extreme environmental conditions.

Paying tribute to the pioneers of this research in the Soviet Union (L.A. Orbeli, V.V. Parin, A.V. Lebedinsky, N.M. Sisakyan, O.G. Gazenco and many others), in the USA, and also in France, Germany, and Japan, we would like to refer the reader to the fundamental monographs that analyze the long-term results in this field [104–107]. Many effects observed in astronauts after a space flight have been well described at the physiological level. In general, it appears to be a complex pattern of adaptive reactions that involve all functional systems of the body. The need for astronauts to return to Earth, to their conventional habitat, and the obligations of physicians to keep them healthy led to efforts undertaken by specialists in all, without exception, space agencies, to develop measures to impede the onset of the phase of structural adaptation to the factors the human body is exposed to in space. Nevertheless, several tissues (e.g., bone tissue) exhibit slow re-adaptation to life on Earth. Space physiology, apparently, deals with the unique pattern of adaptation of human systems, tissues, and cells, thus demonstrating its possibilities. The phenomenology of the major changes induced by the conditions of space flight includes: a negative energy balance (more energy is spent than is received) that affects various processes in a human organism [107–110], a negative water and calcium balance [111, 112] but positive sodium balance [113, 114], demineralization and modification of bone tissue structure [115], ineffective thermoregulation [116–118], changes in the biorhythms of heat production, hormone secretion activity, cardiac function [118–121], reorganization of vasomotor reaction modulation [122], endothelial dysfunction [123], muscle hypotrophy [124–126], decreased muscle tone and speed-strength properties, functional deafferentation of sensor systems that leads to impairments in movement control [127, 128], modification of lung volume, breathing biomechanics and its regulation with chemoreceptors [129, 130], and space anemia [131]. Almost every field of knowledge still has unrevealed molecular mechanisms responsible for the formation of these new stages of physiological systems.

Adaptation of the human organism to any environ-

mental factor is performed with the help of proteins. For a long time, in accordance with analytical capabilities, working hypotheses were based on assumptions on the changes in the concentrations of working proteins or the efficiency of their performance (e.g., enzymatic activity) during adaptation. During the post-genomic era, it has become clear that such a level of study of the adaptation mechanisms will be followed by others: studies at the transcriptional level (i.e. the formation of a new set of functioning proteins) and studying the new protein complexes that are formed during adaptation, along with the protein interaction networks and new reaction cascades. These studies can be conducted with the help of systems biology, using its analytical and bio-information approaches. There are some needs that have been acknowledged by the community of gravitational physiologists but that have not been satisfied so far, such as reaching the level of OMICs. Glass [132], Jackman and Kandarian [133], Ventadour and Attaix [134] and Blottner [135] have noted that the biological effects of microgravity on the genome, proteome, transcriptome, and metabolome remain almost completely unknown.

We suggest that using the methods and approaches of systems biology to study the molecular pattern of adaptive mechanisms that is the most complicated (among all possible variants) at the present stage would both yield valuable knowledge and help to fill in the gaps in the picture of human metabolic pathways; many of these gaps have not even been considered to exist. The community of systems biologists has only begun to realize this exciting perspective. The first papers devoted to changes in the proteome of body fluids (urine and blood) in astronauts after flight [136, 137] have been published, arousing great interest, according to the number of visits on PLoSOne pages of open access (700 retrieves per week). We have studied the influence of overloads in a large-radius centrifuge [137] and breathing of the oxygen-nitrogen-argon mixture under hyperbaric conditions [138]. The characteristics of the blood and urine proteomes of a healthy human under model conditions of “dry immersion” [139] and long-term isolation [140] have been studied. Since the variability of the protein composition of human biological fluids might mask the effects of external impacts, we identified the indices of individual and group variability [141, 142]. It is necessary to take into account the parameters of group variability and the rates of manifestation of individual plasticity in order to determine functional shifts in the protein composition of body fluids during changes under conditions of the living environment, as well as disease progression. The use of direct mass spectrometric profiling of blood serum has allowed researchers to determine which proteins de-

termine the significant group variability (CV = 42.6%) and the dependence of this parameter on age. The individual variability indices turn out to depend linearly on the length of the periods between repeated surveys, increasing from 16 to 42% for periods from 1 day to 1 year. The common changes in the blood proteome observed for space flights and model experiments were modifications of the peaks of acute phase proteins (β -microglobulin, cystatin C) and lipid exchange (apolipoproteins CI, CIII, AII), as well as the shifts in the activity of blood proteolytic systems that can cause changes in the pattern of protein fragments.

High group and individual variability in the urinary protein profile has been noted by many scientists. We have showed that this is maintained even under the strict conditions of model experiments (with control of the intake levels of essential nutrients, fluids, level of locomotor activity, composition of the atmosphere, and sleep-wake rhythms). We observed the modification of the urinary proteome in healthy young men for 520 days with isolation in a hermoobject and managed to identify and characterize both the most plastic part of the low-molecular urinary subproteome and its constant component. Moreover, the proteins whose level in urine depended on the salt intake were discovered.

The study of the urine proteome of astronauts allowed one to identify the stable portion of the subproteome represented by 21 proteins with different tissue specificities and subcellular localizations. Three proteins (afamin, aminopeptidase A and aquaporin 2) appear in the urine of astronauts after long-term flights aboard the International Space Station; the frequency of their detection in samples is most likely related to the impact of space flight factors. The overloads experienced by astronauts at the initial and final stages of the flight can also affect the protein composition of extracellular fluid.

In the dry immersion model, the development of polyurea through a mechanism close to saluresis leads to the development of physiological proteinuria and competitively dependent sodium reabsorption in the proximal tubule of the nephron.

It is obvious that the proteins whose levels change under extreme conditions cannot be regarded as potential biomarkers of diseases, since they participate both in the natural molecular response of an organism during adaptation to a living environment and in the nonspecific component of a disease's pathogenesis.

CONCLUSIONS

The collaboration between physiologists and systems biologists in studying the adaptation of healthy humans to extreme environmental conditions is deepening and

mutually beneficial. Researchers note that the application of systems biology methods in the field of physiological adaptation to extreme environments enables one to move away from the reductionist approaches and avoid paradoxes (e. g., the so-called "lactate paradox in hypoxia"¹) when interpreting data [76, 143].

Now there is growing worldwide interest in collaboration between life science researchers and their colleagues (physicists, computer scientists, chemists, and mathematicians), which has been included in the agenda of the major organizations that fund science. Thus, partnerships between specialists working in the fields of systems biology/bioengineering and human physiology will become increasingly common. The new generation of scientists who are called to work in this field will become more transdisciplinary. We agree with the statement by Edwards [37] that "no longer can biology be considered a science for those who 'cannot do maths'." Consequently, psychologists and scientists working in life sciences should be prepared for modern challenges by expanding their knowledge in computational methods and mathematics in general to a level that will allow them to become productive systems biologists and interact with scientists from other fields. Oncoming advance of physicists and mathematicians is a more complicated process. We are not alone in this view. Paraphrasing Ideker *et al.* [45], we can say that "the contributions of cross-disciplinary scientists will be proportional to their understanding of biology."

Thus, new approaches inherent to modern systems biology can be used for a deeper understanding of the physiological adaptation of a healthy human to ex-

¹ This term refers to the phenomenon associated with the suppression of glycolysis during acclimatization to chronic hypoxia. It was shown that the acute phase of high-altitude adaptation is accompanied by a higher blood level of lactate at any period of submaximal load than under normoxia load, although the peak level of lactate remains unchanged. However, in individuals who have acclimatized to the altitude for more than 3 weeks, a load of the same absolute value and a maximum load cause a smaller increase in the blood lactate level compared to the same physical load in individuals in the un-acclimatized state. This phenomenon, initially regarded as a paradox (i.e., that does not correspond to a logical inference), suggested that ATP production in chronic hypoxia, apparently, does not depend on an increase in anaerobic glycolysis, but the production of mitochondrial ATP becomes better tuned to the hypoxic condition of the organism. Recent studies, however, have shown that the "lactate paradox" can only be a transitional feature of hypoxic adaptation to altitude, disappearing after more than 6 weeks, during the descent to the plains after a climb to altitudes above 5,000 m. Moreover, the decrease in the muscle ability to produce lactate during the period following acclimatization has not been shown in the studies. The question remains open as to whether the "lactate paradox" is caused by the decrease in lactate production in muscles due to the changes in the substrate preference or changes in lactate processing through the mitochondrial enzyme complexes MCT1 and MCT4 (monocarboxylate transporters 1 and 4) in muscle, or for better coupling of pyruvate synthesis with oxidation taking place in mitochondria. The question remains to be solved, along with defining a clear profile of the conditions under which it occurs. Several authors have suggested that a phenomenon analogous to the so-called "lactate paradox" can also occur in tissues other than muscles, in response to acute metabolic stress in chronic hypoxia.

treme environments. One can certainly agree with the opinion that environmental physiology and systems biology are natural partners [144]. Studies of human adaptation to various environmental factors, as well as the study of the response of the human body to a space flight, provide a unique platform for understanding human physiology from the systems' perspective, allowing scientists to approach homeostasis in an ethical and evolutionarily sound way. Finally, there is hope that the relationship between integrative physiology

and systems biology will develop, and that the fields will be thus better understood, leading us, in turn, to a more mature and deeper understanding of a healthy person's biology.

The work was partially supported by the Russian President's grant "Leading Scientific Schools" (NSH-1207.2012.4) and the Russian Foundation for Basic Research (grant № 13-04-01894).

REFERENCES

1. James P. // Quarterly Rev.Biophys. 1997. V. 30 № 4. P. 279–331. doi:10.1017/S0033583597003399. PMID 9634650.
2. Murray R.F., Harper H.W., Granner D.K., Mayes P.A., Rodwell V.W. Harper's Illustrated Biochemistry. New York: Lange Medical Books/McGraw-Hill, 2006.
3. Hey J., Posch A., Cohen A. // Meth. Mol. Biol. 2008. № 424. P. 225–239. doi:10.1007/978-1-60327-064-9_19. PMID 18369866.
4. O'Farrell P.H. // J. Biol. Chem. 1975. V. 250. № 10. P. 4007–4021.
5. Karas M., Bachmann D., Bahr D., Hillenkamp F. // Int. J. Mass Spectrom. Ion Proc. 1987. № 78. P. 53–68.
6. Fenn J.B., Mann M., Meng C.K., Wong S.F., Whitehouse C.M. // Science. 1989. V. 246. P. 64–71.
7. Bensimon A., Heck A.J.R., Aebersold R. // Annu. Rev. Biochem. 2012. V. 81. № 1. P. 379–405.
8. Alberts B. // Science. 2012. V. 337. № 6102. P. 1583.
9. Bard J. // Cell. 2013. V. 2. P. 414–431.
10. Lander E.S. // Nature. 2011. V. 470. № 7333. P. 187–197.
11. Sverdlov E.D. // Patol. Fiziol. Eksp. Ter. 2010. № 3. P. 3–23.
12. Sverdlov E.D. // Biochemistry (Mosc.). 2009. V. 74. P. 939–944.
13. Noble D. // Intersace Focus. 2012. V. 2. № 1. P. 55064.
14. Nagaraj N., Wisniewski J.R., Geiger T., Cox J., Kircher M., Kelso J., Pääbo S., Mann M. // Mol. Syst. Biol. 2011. № 7. P. 548.
15. Beck M., Schmidt A., Malmstroem J., Claassen M., Ori A., Szymborska A., Herzog F., Rinner O., Ellenberg J., Aebersold R. // Mol. Syst. Biol. 2011. № 7. P. 549.
16. Munoz J., Low T.Y., Kok Y.J., Chin A., Frese C.K., Ding V., Choo A., Heck A.J. // Mol. Syst. Biol. 2011. № 7. P. 550.
17. Vidal M., Chan D.W., Gerstein M., Mann M., Omenn G.S., Tagle D., Sechi S. // Clin. Proteomics. 2012. V. 9. № 1. P. 6.
18. Burnette W.N. // Anal. Biochem. 1981. V. 112. № 2. P. 195–203.
19. Zhang Y., Yang C., Wang S., Chen T., Li M., Wang X., Li D., Wang K., Ma J., Wu S., et al. // Liver Int. 2013. V. 33. № 8. P. 1239–1248. doi: 10.1111/liv.12173.
20. Legrain P. // Annual International Congress HUPO-2013, Yokohama, Japan. 12–18 September 2013. Book of Program & Keynote Lectures, 2013.
21. Radivojac P., Clark W.T., Oron T.R., Schnoes A.M., Wittkop T., Sokolov A., Graim K., Funk C., Verspoor K., Ben-Hur A., et al. // Nat. Methods. 2013. V. 10. № 3. P. 221–227.
22. Terentiev A.A., Moldogazieva N.T., Shaitan K.V. // Uspechi Biologicheskoi Khimii. 2009. V. 49. P.427–480.
23. Bose B. // Prog. Biophys. Mol. Biol. 2013. V. 113. № 3. P. 358–368.
24. Kitano H. // Science. 2002. V. 295. № 5560. P. 1662–1664.
25. Naylor S., Chen J.Y. // Per. Med. 2010. V. 7. № 3. P. 275–289.
26. Brenner S. // Philos Trans. R. Soc. Lond. B Biol. Sci. 2010. V. 365. № 1537. P. 201–212.
27. Brenner S., Sijnowski T.J. // Science. 2011. V. 334. № 6056. P. 567.
28. Carbo A., Hontecillas R., Kronsteiner B., Viladomiu M., Pedragosa M., Lu P., Philipson C.W., Hoops S., Marathe M., Eubank S., et al. // PLOs Comput. Biol. 2013. V.9. № 4. e1003027.
29. Chaufan C., Joseph J. // Int. J. Health Serv. 2013. V. 43. № 2. P. 281–303.
30. Cesareni G., Gimona M., Sudol M., Yaffe M. Modular protein domains. Weinheim, Germany: Wiley-VCH, 2004.
31. Tong A.H., Drees B., Nardelli G., Bader G.D., Branetti B., Castagnoli L., Evangelista M., Ferracuti S., Nelson B., Pauluzi S., et al. // Science. 2002. V. 295. № 5553. P. 21–24.
32. Sommer B., Kormeier B., Demenkov P.S., Arrigo P., Hippe K., Ates Ö., Kochetov A.V., Ivanisenko V.A., Kolchanov N.A., Hofestädt R. // J. Bioinform. Comput. Biol. 2013. V. 11. № 1. P. 1340005.
33. Demenkov P.S., Ivanisenko TV., Kolchanov N.A., Ivanisenko V.A. // In Silico Biol. 2011–2012. V. 11. № 3–4. P. 149–161.
34. Hood L., Heath J.R., Phelps M.E., Lin B. // Science. 2004. V. 306. № 5696. P. 640–643.
35. Park S., Yang J.S., Jang S.K., Kim S. // J. Proteome Res. 2009. V. 8. № 7. P. 3367–3376.
36. Ivanisenko V.A., Demenkov P.S., Pintus S.S., Ivanisenko TV., Podkolodny N.L., Ivanisenko L.N., Rozanov A.S., Bryanskaya A.V., Kostjukova E.S., Levizkiy S.A., et al. // Dokl. Biochem. Biophys. 2012. V. 443. P. 76–80.
37. Edwards L.M., Thiele I. // Extreme Physiol. Med. 2013. № 2. P. 1–8.
38. Hargreaves M. // J. Appl. Physiol. 2008. № 104. P. 1541–1542.
39. Greenhaff P.L., Hargreaves M. // J. Physiol. 2011. № 589. P. 1031–1036.
40. Carlson R.P., Oshota O.J., Taffs R.L. // Subcell. Biochem. 2012. № 64. P. 139–157.
41. Thiele I., Palsson B.O. // Nat. Protoc. 2010. № 5. P. 93–121.
42. Palsson B. Systems Biology. Cambridge: Cambridge Univ. Press, 2008.
43. Kohl P., Crampin E.J., Quinn T.F., Noble D. // Clin. Pharmacol. Ther. 2010. V. 88. № 1. P. 25–33.
44. Hood L., Rowen L., Galas D.J., Aitchison J.D. // Brief Funct. Genomic Proteomic. 2008. V. 7. № 4. P. 239–248.
45. Ideker T., Galitski T., Hood L. // Annu. Rev. Genomics Hum. Genet. 2001. № 2. P. 343–372.
46. Westerhoff H.V. // Meth. Enzymol. 2011. № 500. P. 3–11.

47. Shublaq N., Sansom C., Coveney P.V. // *Chem. Biol. Drug Des.* 2013. V. 81. № 1. P. 5–12. doi:10.1111/j.1747-0285.2012.01444.x.
48. You L. // *Cell Biochem. Biophys.* 2004. V. 40. P. 167–184.
49. Shaitan K.V., Turley E.V., Golik D.N., Tereshkina K.V., Levitsova O.V., Fedik I.V., Shaitan A.K., Li A., Kirpichnikov M.P. // *Rossiiskii khimicheskii zhurnal.* 2006. V.50. P.53–65.
50. Stumpf M.P., Thorne T., de Silva E., Stewart R., An H.J., Lappe M., Wiuf C. // *PNAS.* 2008. V.105. P. 6959–6964.
51. Visvanathan M., Breit M., Pfeifer B., Baumgartner C., Modre-Osprian R., Tilg B. // *Meth. Inf. Med.* 2007. V. 46. P. 381–391.
52. Suresh Babu C.V., Joo Song E., Yoo Y.S. // *Biochimie.* 2006. V. 88. №3–4. P. 277–283.
53. Conzelmann H., Saez-Rodriguez J., Sauter T., Bullinger E., Allgöwer F., Gilles E.D. // *Syst. Biol. (Stevenage).* 2004. V. 1. № 1. P. 159–169.
54. Wanjek C. // *The NIH Catalyst.* 2011. № 19. P. 1.
55. Duarte N.C., Becker S.A., Jamshidi N., Thiele I., Mo M.L., Vo T.D., Srivas R., Palsson B.O. // *Proc. Natl. Acad. Sci. USA.* 2007. № 104. P. 1777–1782.
56. Korn R. // *Biology and Philosophy.* 2005. V. 20. № 1. P. 137–151.
57. Noble D. // *Prog. Biophys. Mol. Biol.* 2013. V. 111. № 2–3. P. 59–65.
58. Noble D. // *Philos. Trans. A Math. Phys. Eng. Sci.* 2008. V. 366. № 1878. P. 3001–3015.
59. van Regenmortel M.H. // *EMBO Rep.* 2004. V. 5. № 11. P. 1016–1020.
60. Bordbar A., Palsson B.O. // *J. Intern. Med.* 2012. № 271. P. 131–141.
61. Bordbar A., Lewis N.E., Schellenberger J., Palsson B.O., Jamshidi N. // *Mol. Syst. Biol.* 2010. № 6. P. 422.
62. Gille C., Bolling C., Hoppe A., Bulik S., Hoffmann S., Hübner K., Karlstädt A., Ganeshan R., König M., Rother K., et al. // *Mol. Syst. Biol.* 2010. № 6. P. 411.
63. Bordbar A., Feist A.M., Usaite-Black R., Woodcock J., Palsson B.O., Famili I. // *BMC Syst. Biol.* 2011. № 5. P. 180.
64. Qu Z., Garfinkel A., Weiss J., Nivala M. // *Prog. Biophys. Mol. Biol.* 2011. № 107. P. 21–31.
65. Reed J.L., Patel T.R., Chen K.H., Joyce A.R., Applebee M.K., Herring C.D., Bui O.T., Knight E.M., Fong S.S., Palsson B.O. // *Proc. Natl. Acad. Sci. USA.* 2006. № 103. P. 17480–17484.
66. Satish Kumar V., Dasika M.S., Maranas S.D. // *BMC Bioinform.* 2007. № 8. P. 212.
67. Kumar V.S., Maranas C.D. // *PLoS Comput. Biol.* 2009. № 5. e1000308.
68. Illig T., Gieger C., Zhai G., Römisch-Margl W., Wang-Sattler R., Prehn C., Altmaier E., Kastenmüller G., Kato B.S., Mewes H.W., et al. // *Nat. Genet.* 2010. № 42. P. 137–141.
69. Suhre K., Wallaschofski H., Raffler J., Friedrich N., Harling R., Michael K., Wasner C., Krebs A., Kronenberg F., Chang D., et al. // *Nat. Genet.* 2011. № 43. P. 565–569.
70. Edwards L.M., Lawler N.G., Nikolic S.B., Peters J.M., Horne J., Wilson R., Davies N.W., Sharman J.E. // *J. Lipid Res.* 2012. V. 53. № 9. P. 1979–1986.
71. Orth J.D., Palsson B.O. // *Biotechnol. Bioeng.* 2010. № 107. P. 403–412.
72. Rolfsson O., Palsson B.O., Thiele I. // *BMC Syst. Biol.* 2011. № 5. P. 155.
73. Kanehisa M., Goto S., Hattori M., Aoki-Kinoshita K.F., Itoh M., Kawashima S., Katayama T., Araki M., Hirakawa M. // *Nucl. Acids Res.* 2006. № 34. P. 354–357.
74. Jens T., Maillet A., Lang R., Gunga H.C., Johannes B., Gauquelin-Koch G., Kihm E., Larina I., Gharib C., Kirsch K.A. // *Am. J. Kidney Dis.* 2002. V. 40. № 3. P. 508–516.
75. Grigoriev A.I., Arzamazov G.C., Dorohova B.R., Kozjyrevskaia G.I., Morukov B.V., Natochin J.V., Noskov V.B., Chmelkov V.P. Guidelines on the use of water and saline water load samples in the assessment of the functional state of human kidneys. Ministry of Health of the Russian Federation, Moscow, USSR. 1979. 31 p.
76. Vazquez A., Oltvai Z.N. // *PLoS One.* 2011. № 6. e19538.
77. van Beek J.H., Supandi F., Gavai A.K., de Graaf A.A., Binsl T.W., Hettling H. // *Philos. Trans. A Math. Phys. Eng. Sci.* 2011. № 369. P. 4295–4315.
78. Zauber H., Mosler S., von Heßberg A., Schulze W.X. // *Proteomics.* 2012.V. 12. № 13. P. 2221–2235. doi: 10.1002/pmic.201100228.
79. Martin-Perez M., Fernandez-Borras J., Ibarz A., Millan-Cubillo A., Felip O., de Oliveira E., Blasco J. // *J. Proteome Res.* 2012. V. 6. № 11(7). P. 3533–547. doi: 10.1021/pr3002832.
80. Pilegaard H., Saltin B., Neuffer P.D. // *J. Physiol.* 2003. V. 546 (Pt 3). P. 851–858.
81. Flueck M. // *Exp. Physiol.* 2010. V. 95. № 3. P. 451–462. doi: 10.1113/expphysiol.2009.047605.
82. Tonevitsky A.G., Maltseva D.V., Abbasi A., Samatov T.R., Sakharov D.A., Shkurnikov M.U., Lebedev A.E., Galatenko V.V., Grigoriev A.I., Northoff H. // *BMC Physiol.* 2013. V. 13. P. 9. doi: 10.1186/1472-6793-13-9.
83. Beall C.M., Cavalleri G., Deng L., Elston R.C., Gao Y., Knight J., Li C., Li J.C., Liang Y., McCormack M., et al. // *Proc. Natl. Acad. Sci. USA.* 2010. № 107. P. 11459–11464.
84. Yi X., Liang Y., Huerta-Sanchez E., Jin X., Cuo Z.X., Pool J.E., Xu X., Jiang H., Vinckenbosch N., Korneliussen T.S., et al. // *Science.* 2010. № 329. P. 75–78.
85. Viganò A., Ripamonti M., De Palma S., Capitanio D., Vasso M., Wait R., Lundby C., Cerretelli P., Gelfi C. // *Proteomics.* 2008. № 8. P. 4668–4679.
86. Mainini V., Gianazza E., Chinello C., Bilo G., Revera M., Giuliano A., Caldara G., Lombardi C., Piperno A., Magni F. // *Mol. Biosyst.* 2012. № 8. P. 959–966.
87. Ahmad Y., Shukla D., Garg I., Sharma N.K., Saxena S., Malhotra V.K., Bhargava K. // *Funct. Integr. Genomics.* 2011. № 11. P. 407–417.
88. Lai X., Nikolov S., Wolkenhauer O., Vera J. // *Comput. Biol. Chem.* 2009. № 3. P. 312–324.
89. Turan N., Kalko S., Stincone A., Clarke K., Sabah A., Howlett K., Curnow S.J., Rodriguez D.A., Cascante M., O’Neill L., et al. // *PLoS Comput. Biol.* 2011. № 7. e1002129.
90. Stefanini M.O., Qutub A.A., Mac Gabhann F., Popel A.S. // *Math. Med. Biol.* 2012. № 29. P. 85–94.
91. Baumann K., Carnicer M., Dragosits M., Graf A.B., Stadlmann J., Jouhten P., Maaheimo H., Gasser B., Albiol J., Matanovich D., et al. // *BMC Syst. Biol.* 2010. № 4. P. 141.
92. Schmierer B., Novak B., Schofield C.J. // *BMC Syst. Biol.* 2010. № 4. P. 139.
93. Mac Gabhann F., Qutub A.A., Annex B.H., Popel A.S. // *Wiley Interdiscip. Rev. Syst. Biol. Med.* 2010. № 2. P. 694–707.
94. Kojima T., Ueda Y., Adati N., Kitamoto A., Sato A., Huang M.C., Noor J., Sameshima H., Ikenoue T. // *J. Mol. Neurosci.* 2010. № 42. P. 154–161.
95. Selivanov V.A., Votyakova T.V., Zeak J.A., Trucco M., Roca J., Cascante M. // *PLoS Comput. Biol.* 2009. № 5. e1000619.
96. Flueck M. // *High Alt. Med. Biol.* 2009. V. 10. № 2. P. 183–193. doi: 10.1089/ham.2008.1104.
97. Sharma N.K., Sethy N.K., Bhargava K. // *J. Proteomics.* 2013. V. 21. № 79. P. 277–298. doi: 10.1016/j.jprot.2012.12.020.

98. Ohata Y., Ogata S., Nakanishi K., Kanazawa F., Uenoyama M., Hiroi S., Tominaga S., Toda T., Kawai T. // *Histol. Histopathol.* 2013. V. 28. № 7. P. 893–902.
99. Valko M., Leibfritz D., Moncol J., Cronin M.T., Mazur M., Telser J. // *Int. J. Biochem. Cell Biol.* 2007. V. 39. № 1. P. 44–84.
100. Blokhina O., Fagerstedt K.V. // *Plant Physiol. Biochem.* 2010. V. 48. № 5. P. 359–373. doi: 10.1016/j.plaphy.2010.01.007.
101. Jones D.P. // *Rejuvenation Res.* 2006. V. 9. № 2. P. 169–181.
102. Harris C., Shuster D.Z., Roman Gomez R., Sant K.E., Reed M.S., Pohl J., Hansen J.M. // *Free Radic. Biol. Med.* 2013. № 63. P. 325–337. doi:10.1016/j.freeradbiomed.2013.05.040.
103. Harris C., Hansen J.M. // *Methods Mol. Biol.* 2012. V. 889. P. 325–346. doi: 10.1007/978-1-61779-867-2_21.
104. Orbital station “Mir” / Edited by Grigorieva A.I. Moscow: LLC “Anicom”, 2002. V.1. P. 660.
105. Orbital station “Mir” / Edited by Grigorieva A.I. Moscow: LLC “Anicom”, 2002. V. 2. P. 623.
106. International space station / Edited by Grigorieva A.I. Moscow: IBMP RAS, 2011.
107. Leach-Huntoon C.S., Grigoriev A.I., Natochin V.Yu. // *Am. Astronautical Soc. Publ. SanDiego, California, USA.* 1998. V. 94. 220 p.
108. Stein T.P. // *Eur. J. Appl. Physiol.* 2013. V. 113. № 9. P. 2171–2181. doi: 10.1007/s00421-012-2548-9.
109. Larina I.M., Stein T.R., Leskiv M.G., Shluter M.D. // *Orbital station “Mir”* / Edited by Grigorieva A.I. Moscow: LLC “Anicom”, 2002. V. 2. P. 114–121.
110. Ajotto G.B., Himomura Y.S. // *J. Nutr. Sci. Vitaminol.* 2006. V. 52. P. 233–247.
111. Gazenko O.G., Grigoriev A.I., Natochin J.V. *Salt and water homeostasis and space flight.* Moscow: Nauka, 1986. P. 256.
112. Morukov B.V., Noskov V.B., Larina I.M., Natochin Iu.V. // *Ros. Fiziol. Zh. Im. I.M. Sechenova.* 2003. V.89. № 3. P. 356–367.
113. Gerzer R., Heer M. // *Curr. Pharm. Biotechnol.* 2005. V. 6. № 4. P. 299–304.
114. Drummer C., Norsk P., Heer M. // *Am. J. Kidney Dis.* 2001. V. 38. № 3. P. 684–690.
115. Oganov V.S., Bogomolov V.V., Bakulin A.V., Novikov V.E., Kabitskaia O.E., Murashko L.M., Morgun V.V., Kasparskii R.R. // *Fiziol. Cheloveka.* 2010. V. 36. № 3. P. 39–47.
116. Klimovitsky V.Y., Alpatov A.M., Hoban-Higgins T.M., Utekhina E.S., Fuller C.A. // *J. Gravit. Physiol.* 2000. V. 7. P. 149.
117. Lakota N.G., Larina I.M. // *Human Physiol.* 2002. V. 28. № 3. P. 82–92.
118. Robinson E.L., Fuller C.A. // *Pflügers Arch.* 2000. V. 441 (2–3 Suppl). P. R32–38.
119. Fuller P.M., Jones T.A., Jones S.M., Fuller C.A. // *Proc. Natl. Acad. Sci. USA.* 2002. V. 26. №99(24). P.15723–15728.
120. Larina I.M., Witson P., Smornova T.M., Chen J.-M. // *Physiologia cheloveka.* 2000. V. 26. № 4. P. 94–100.
121. Baevskiy R.M. // *Physiologia cheloveka.* 2002. V. 28. № 2. P. 70–82.
122. Fomina G.A., Kotovskaya A.R., Pochuev V.I., Zhernavkov A.F. // *Human Physiol.* 2008. V. 34. № 3. P. 92–97.
123. Navasiolava N.M., Dignat-George F., Sabatier F., Larina I.M., Demiot C., Fortrat J.-O., Gauquelin-Koch G., Kozlovskaya I.B., Custaud M.-A. // *Am. J. Physiol. Heart Circ. Physiol.* 2010. V. 299. № 2. P. H248–56. doi: 10.1152/ajp-heart.00152.2010.
124. Ohira Y., Yoshinaga T., Nomura T., Kawano F., Ishihara A., Nonaka I., Roy R.R., Edgerton V.R. // *Adv. Space Res.* 2002. V. 30. № 4. P. 777–781.
125. Edgerton V.R., Zhou M.Y., Ohira Y., Klitgaard H., Jiang B., Bell G., Harris B., Saltin B., Gollnick P.D., Roy R.R. // *J. Appl. Physiol.* 1995. V. 78. № 5. P. 1733–1739.
126. Shenkman B.S., Nemirovskaya T.L., Belozerova I.N., Cheglova I.A., Kozlovskaya I.B. // *Report of the Academy of Science.* 1999. V. 367. № 2. P. 279–281.
127. Tomilovskaya E.S., Berger M., Gerstenbrand F., Kozlovskaya I.B. // *J. Gravit. Physiol.* 2007. V. 14. № 1. P. 79–80.
128. Kornilova L.N., Naumov I.A., Azarov K.A., Sagalovitch V.N. // *Aviat Space Environ. Med.* 2012. V. 83. № 12. P. 1123–1134.
129. Baevsky R.M., Baranov V.M., Funtova I.I., Diedrich A., Pashenko A.V., Chernikova A.G., Drescher J., Jordan J., Tank J. // *J. Appl. Physiol.* 2007. V. 103. № 1. P. 156–161.
130. Baranov V.M., Popova Y.A., Suvorov A.V. *Space biology and medicine. V.2. Biomedical research on the Russian segment ISS.* Moscow, 2011. P. 72–92.
131. Grigor'ev A.I., Ivanova S.M., Morukov B.V., Maksimov G.V. // *Dokl. Biochem. Biophys.* 2008. V. 422. P. 308–311.
132. Glass D.J. // *Trends Mol. Med.* 2003. V. 9. № 8. P. 344–350.
133. Jackman R.W., Kandarian S.C. // *Am. Physiol. Cell Physiol.* 2004. V. 287. № 4. P. 834–843.
134. Ventadour S., Attaix D. // *Curr. Opin. Rheumatol.* 2006. V. 18. № 6. P. 631–635.
135. Blottner D., Serradj N., Salanova M., Touma C., Palme R., Silva M., Aerts J.M., Berckmans D., Vico L., Liu Y., et al. // *J. Comp. Physiol. B.* 2009. V. 179. № 4. P. 519–533. doi: 10.1007/s00360-008-0330-4.
136. Pastushkova L.Kh., Kireev K.S., Kononikhin A.S., Tiys E.S., Popov I.A., Starodubtseva N.L., Dobrokhotov I.V., Ivanisenko V.A., Larina I.M., Kolchanov N.A., et al. // *PLoS One.* 2013. V.8. №8. e71652.
137. Kireev K.S. *Kidneys and urinary tract proteins in healthy human urine proteome long-term space flight: Author's abstract. candidate of medical sciences.* Moscow: SSC RF – IBMP RAS, 2013. P. 25.
138. Paharukova N.A., Pastushkova L.H., Popova Y.A., Larina I.M. // *Bulletin of Experimental Biology and Medicine.* 2010. V. 148. № 2. P. 42–45.
139. Pastushkova L.H., Dobrokhotov I.V., Veselova O.M., Tiys E.S., Kononihin A.S., Novoselova A.M., Kupe M., Custo M.-A., Larina I.M. // *Physiologia cheloveka.* 2014. V.40. №3. P. 109–119.
140. Larina I.M., Kolchanov N.A., Dobrokhotov I.V., Ivanisenko V.A., Demenkov P.C., Tiys E.S., Valeeva O.A., Pastushkova L.H., Nikolaev E.N. // *Physiologia cheloveka.* 2012. V. 38. № 3. P. 107–115.
141. Trifonova O., Larina I., Grigoriev A., Lisitsa A., Moshkovskii S., Archakov A. // *Expert Rev. Proteomics.* 2010. V. 7. № 3. P. 431–438.
142. Paharukova N.A., Pastushkova L.H., Moshkovskiy S.A., Larina I.M. // *Biomedicina.* 2011. V. 5. № 3. P. 203–212.
143. Murray A.J. // *Genome Med.* 2009. № 1. P. 117.
144. Hood L.E., Omenn G.S., Moritz R.L., Aebersold R., Yamamoto K.R., Amos M., Hunter-Cevera J., Locascio L. // *Proteomics.* 2012. V. 12. № 18. P. 2773–2783.

Study of the Structure-Function-Stability Relationships in Yeast D-amino Acid Oxidase: Hydrophobization of Alpha-Helices

I. V. Golubev^{1,2}, N. V. Komarova^{2,3}, K. V. Ryzhenkova^{1,2}, T. A. Chubar^{1,2}, S. S. Savin^{2,3}, V. I. Tishkov^{1,2,3*}

¹Department of Chemistry, M.V. Lomonosov Moscow State University; Leninskie gory, 1/3, Moscow, 119991, Russia

²Innovations and High Technologies MSU Ltd, Tsymlyanskaya str., 16-96, Moscow, 109559, Russia

³A.N. Bach Institute of Biochemistry, Russian Academy of Sciences, Leninskiy ave., 33/2, Moscow, 119071, Russia

*E-mail: vitishkov@gmail.com

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT Hydrophobization of alpha-helices is one of the general approaches used for improving the thermal stability of enzymes. A total of 11 serine residues located in alpha-helices have been found based on multiple alignments of the amino acid sequences of D-amino acid oxidases from different organisms and the analysis of the 3D-structure of D-amino acid oxidase from yeast *Trigonopsis variabilis* (TvDAAO, EC 1.4.3.3). As a result of further structural analysis, eight Ser residues in 67, 77, 78, 105, 270, 277, 335, and 336 positions have been selected to be substituted with Ala. S78A and S270A substitutions have resulted in dramatic destabilization of the enzyme. Mutant enzymes were inactivated during isolation from cells. Another six mutant TvDAAOs have been highly purified and their properties have been characterized. The amino acid substitutions S277A and S336A destabilized the protein globule. The thermal stabilities of TvDAAO S77A and TvDAAO S335A mutants were close to that of the wild-type enzyme, while S67A and S105A substitutions resulted in approximately 1.5- and 2.0-fold increases in the TvDAAO mutant thermal stability, respectively. Furthermore, the TvDAAO S105A mutant showed on average a 1.2- to 3.0-fold higher catalytic efficiency with D-Asn, D-Tyr, D-Phe, and D-Leu as compared to the wild-type enzyme.

KEYWORDS D-amino acid oxidase from yeast *Trigonopsis variabilis*, protein engineering, hydrophobization of alpha-helices, site-directed mutagenesis, substrate specificity, thermal stability.

INTRODUCTION

D-amino acid oxidase (DAAO, [EC 1.4.3.3]) belongs to a class of FAD-containing oxidoreductases and catalyzes the oxidative deamination of D-amino acids to the corresponding α -keto acids [1]. DAAO is widespread in nature: the genes of this enzyme have been found in cells of molluscs, fishes, reptiles, amphibians, insects, birds, plants, mammals, as well as microorganisms, including fungi, yeasts and bacteria, where it performs important physiological functions [2, 3]. The processes of synthesis of optically active compounds, α -keto acids, and 7-aminocephalosporanic acid using DAAO have been designed. This enzyme is also used in biosensors to determine the D-amino acid content [2, 4, 5]. Two enzymes from yeasts *Rhodotorula gracilis* (RgDAAO) and *Trigonopsis variabilis* (TvDAAO) are the most widely

used ones. TvDAAO exhibits the highest activity with cephalosporin C (CephC) [6] and the best thermal stability [7] among all known D-amino acid oxidases. For example, TvDAAO retains 100% activity when incubated for 30 min at 45°C, while RgDAAO is completely inactivated under these conditions. The temperature stability of DAAO from *Arthrobacter protophormiae* and *Candida boidinii* was also studied. They are very similar to RgDAAO. Thus, at 50°C they completely lose their activity in 30 minutes [5, 8, 9].

In our laboratory, the D-amino acid oxidase gene from *T. variabilis* yeast has been cloned, the overexpression system of the recombinant enzyme in *Escherichia coli* cells in soluble and active form has been developed, and its properties have been studied [10]. Native TvDAAO is a homodimer [11], which has a 2-fold sym-

metry axis with the subunits mutually arranged in a “head-to-tail” manner. Each subunit contains one FAD cofactor molecule at the active site.

Increasing the thermal stability of practically important enzymes is both a fundamental and applied problem. Data produced in such experiments provide a more comprehensive and deeper understanding of the relationship between the structure, function, and stability of the protein being studied. At the same time, solving this problem allows one to reduce the loss of enzyme during its isolation and facilitates the purification process, which in turn reduces the cost of the final product. For example, mutant formate dehydrogenases from *Pseudomonas sp. 101* with increased temperature stability were obtained in our laboratory. This allowed us to introduce the thermal treatment step to the purification process of the recombinant enzyme. Heating the cell-free extract at 60°C for 20–30 minutes resulted in increased purity of the preparation ranging from 50 to 80–85% without a loss of the enzyme activity [12].

Very scarce data on increasing the thermal stability of TvDAAO using protein engineering have been published so far. Only two papers report data on the obtained TvDAAO mutants with point amino acid substitutions, which demonstrated a slight increase in thermal stability as compared to the wild-type enzyme [13, 14]. The covalent crosslinking of two TvDAAO subunits with Lys-Leu dipeptide [15] was also reported. It resulted in increased T_m by 2°C but worse catalytic properties with most substrates.

This paper presents the results of applying the general approach of increasing thermal stability based on hydrophobization to D-amino acid oxidase from the yeast *T. variabilis*. Hydrophobization was achieved by replacing the serine residues with the alanine residues in the α -helical segments of the TvDAAO structure. The effect of these substitutions on the catalytic properties of the enzyme was also studied.

EXPERIMENTAL

Molecular Biology Grade reagents were used for the genetic engineering experiments. Bacto tryptone, yeast extract and agar (Difco, USA), glycerol (99.9%) and calcium chloride (“ultra pure”), potassium hydrogen phosphate, sodium dihydrogen phosphate (“pure for analysis”), lysozyme (Fluka/BioChemika, Switzerland), isopropyl- β -D-thiogalactopyranoside (IPTG), 2,2'-azino-bis (3-ethylbenzothiazoline-6-sulfonate) (ABTS), kanamycin and chloramphenicol (Sigma, USA), and glucose and sodium chloride (“AR grade”, “Helicon”, Russia) were used in the microbiological experiments. Restriction endonucleases, DNA ligase of T4 phage, and Pfu-DNA polymerase (Thermo Scientific) were used for cloning DNA fragments and site-directed mutagen-

esis. Thermo Scientific reagent kits were used to isolate DNA from agarose gel and plasmids from *E. coli* cells. The oligonucleotides for the polymerase chain reaction (PCR) and sequencing were synthesized by Synthol (Russia). The MilliQ (Millipore, USA) purified water was used in these experiments.

We used the following *E. coli* bacterial strains in our study:

E. coli DH5 α : *fhuA2* Δ (*argF-lacZ*)U169 *phoA glnV44* Φ 80 Δ (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*.

E. coli BL21(DE3) pLysS Codon Plus: B⁻ *F⁻ ompT hsdS(r_B⁻ m_B⁻) dcm⁺ Tet^r gal λ (DE3) endA Hte [pLysS *argU ileY leuW Cam^r*].*

All reagents for the electrophoresis of proteins were manufactured by Bio-Rad (USA). Tris (tris (hydroxymethyl) aminomethane, “reagent grad”) from Merck (Germany), racemic amino acids from Dia-M (Russia), and Reanal (Hungary), 2,2'-azino-bis (3-ethylbenzothiazoline-6-sulfonate) (ABTS) (Sigma, USA), horseradish peroxidase (Dia-M, Russia) were used for purification and characterization of the enzyme.

Site-directed mutagenesis

Nucleotide substitutions were introduced using two-step PCR as described previously [13, 16]. The plasmid obtained based on pET-33b (+) with the *tvdaao* gene being under the control of a strong promoter of RNA polymerase of T7 phage was used as a template. The mutations were introduced using direct (T7_For) and reverse (T7_Rev) primers at the beginning and at the end of the gene, respectively, as well as direct (Mut_For) and reverse (Mut_Rev) primers carrying the desired replacements for the *tvdaao* gene. The primer sequences are shown below. The introduced mutations are highlighted in bold.

```
T7_For      5'-taatacgcactcactataggg-3'
T7_Rev      5'-gctagttattgctcagcgg-3'
Ser67Ala_For 5'-gactacgatgccgtcgcttattcctatctcttgcgagagctg-3'
Ser67Ala_Rev 5'-cgcaagataggataagcagcggcctcgtagtcggc-3'
Ser77Ala_For 5'-tggctcgaggcagccccgaggtggaa-3'
Ser77Ala_Rev 5'-gggctggctcgagccagctctcgaaag-3'
Ser78Ala_For 5'-ctcgaagcgccccgaggtggaattcg-3'
Ser78Ala_Rev 5'-tcgggggcgcttcgagccagctctcg-3'
Ser105Ala_For 5'-gaaggtgccatggcgccatctgtcaacgcaac-3'
Ser105Ala_Rev 5'-gacagatggcgcgcatggcaccctccagtttagg-3'
Ser270Ala_For 5'-ccgaaccgatcctgctctcaccatcgaatcctgt-3'
Ser270Ala_Rev 5'-gatgggtgagaggcaggatcgggttcggatgacc-3'
Ser277Ala_For 5'-ccatcgaatcctggctagagccctcgaccgatc-3'
Ser277Ala_Rev 5'-cgagggctctaggcagattcgatgggtgagaga-3'
Ser335Ala_For 5'-gtgctggttaccaggccctctaccgcatggctgat-3'
Ser335Ala_Rev 5'-gccgtaagaggctggtaccagcaccggc-3'
Ser336Ala_For 5'-ctggttaccaggcttacggcaggtgatgaaag-3'
Ser336Ala_Rev 5'-catgccgtaagggactggtaccagcaccg-3'
```

The reaction mixture for PCR contained 2.5 μ l of a 10x buffer for Pfu-DNA polymerase (200 mM Tris-HCl

(pH 8.8 at 25°C), 100 mM (NH₄)₂SO₄, 100 mM KCl, 1 mg/ml BSA, 1% (v/v) Triton X-100, 20 mM MgSO₄; 2.5 µl of a dNTP mixture (dATP, dGTP, dTTP, dCTP with the concentration of each component 2.5 mM); 1 µl of the DNA template (≈10 ng/µL); 2 µl of each primer (10 nM/ml); 0.5 µl of Pfu-DNA polymerase (2.5 U/µl); and deionized water to the total volume of the mixture of 25 µl. PCR was performed in a 0.5-ml thin-walled plastic tube (SSI, USA) using the Tertsik instrument (DNA-Technologies, Russia). A total of 30 µl of mineral oil was added to the tube before the PCR to prevent evaporation of the reaction mixture. The tube was heated for 5 min at 95°C, and the PCR reaction was then initiated by addition of the enzyme. The reaction was done according to the following scheme: the first step at 95°C, 30 sec; the second step at 54-58°C, 60s; and the third step at 72°C, 2 min, a total of 25-35 cycles. After the last cycle, the reaction mixture was further incubated for 10 min at 72°C. The temperature at the second step was 3-5°C below the melting temperature of the duplexes (T_m) formed by the primers. The empirical formula was used to evaluate T_m:

$$T_m = 2(nA + nT) + 4(nG + nC),$$

where nX is the number of X-type nucleotides (X = A, T, C, G) in the primer.

Two PCRs using T7_For/Mut_Rev (fragment 1) and Mut_For/T7_Rev (fragment 2) primer pairs were used to obtain fragments containing the desired substitution. The PCR products—fragment 1 and fragment 2—were purified using electrophoresis in a 1% agarose gel. The third uniting PCR was then performed with T7_For and T7_Rev primers, wherein the previously obtained fragments 1 and 2 were used as the DNA template. The product of the third PCR was purified in a similar way using a 1% agarose gel and treated with two restriction endonucleases. NcoI and Bsp119I were used in the case of substitutions at the 67, 77, 78, and 105 positions, while Bsp119I and XhoI were used for substitutions at the 270, 277, 335, and 336 positions. DNA fragments were purified using electrophoresis in a 1% agarose gel and ligated into the original vector, treated with the same restriction endonucleases. The ligation mixture was used to transform *E. coli* DH5α cells. The cells were then plated onto Petri dishes with an agar medium containing kanamycin (30 µg/ml) and incubated for 16 hours at 37°C. Three colonies of each mutant were taken from each plate and used to isolate plasmids. The presence of only desired mutations was proved by sequencing using the plasmid DNA at the center for collective use “Genome” (V.A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences).

Expression of TvDAAO mutants in *E. coli* cells.

TvDAAO and its mutants were expressed in *E. coli* cells BL21 (DE3) CodonPlus/pLysS. The cells were transformed using the appropriate plasmid and plated on Petri dishes with an agar medium containing kanamycin (30 µg/ml) to obtain the producer strain. A single colony was taken from the plate and cultured for 16 hours at 30°C in 10 ml of a 2YT medium (Bacto tryptone 16 g/l, yeast extract 10 g/l, sodium chloride 5 g/l, pH 7.5) in the presence of 30 µg/ml kanamycin and 25 µg/ml chloramphenicol to prepare the inoculum. In the morning, the cells were subcultured to a fresh medium (dilution 1: 100) and cultured at 30°C until the absorbance of A₆₀₀ ≈ 0.6-0.8 at 600 nm was reached. The inoculum was placed into the culture flasks in amounts of 10% of the total volume of a medium (LB modified medium – yeast extract 10 g/l, Bacto tryptone 5 g/l, glucose 5 g/l, sodium dihydrogen phosphate 1.5 g/l, dipotassium hydrogenphosphate 1 g/l, pH 7.5) containing 1 kanamycin 30 µg/m. Cultivation was carried out in 11 baffled conical flasks (the volume of the medium did not exceed 10-15% of the flask volume). The cultivation temperature ranged from 18 to 27°C, and the rotation rate of the shaker was 120-160 rpm. After reaching A₆₀₀ ≈ 0.6-0.8, enzyme expression was induced by adding IPTG to the medium to a final concentration of 0.1 mM. After induction, the cells were cultivated for 24 hours and then pelleted using the Eppendorf 5403 centrifuge (5 minutes, 5000 rpm, 4°C). The resulting pellet was resuspended in a 0.02-M Tris-HCl buffer (pH 8.0 at 25°C) in a ratio of 1: 4 (wt.). The resulting suspension was stored at -20°C.

Isolation and purification of TvDAAO mutants

Cell suspension in the 20 mM Tris-HCl buffer with pH 8.0 was twice frozen and thawed, and the cells were then disrupted using a Branson Sonifier 250 (Germany) under continuous cooling to isolate mutant TvDAAO. The precipitate was removed by centrifugation on an Eppendorf 5804 R centrifuge (11000 rpm, 30 min).

Purification of the enzyme included ion exchange chromatography on a MonoQ HR 10/10 column using the FPLC instrument manufactured by Pharmacia Biotech (Sweden) and desalting on a Sephadex G-25 carrier. [17] The purity of the preparations was monitored by analytical electrophoresis in a 12% polyacrylamide gel in the presence of 0.1% sodium dodecyl sulphate on a MiniProtean III instrument (BioRad, Austria) according to the manufacturer's protocols.

Kinetic assay

The activity of D-amino acid oxidase was determined using the bi-enzymatic system, including DAAO and

horseradish peroxidase. D-methionine was used as a substrate for the first enzyme, and ABTS was used as a substrate for the second enzyme. The activity was measured at 30°C based on the concentration of the ABTS oxidation product (absorbance at 414 nm, $\epsilon_{414} = 36600 \text{ l/mol/cm}$) on a Shimadzu UV-1800 spectrophotometer (Japan). A total of 770 μl of a 50-mM potassium phosphate buffer (PPB), pH 8.0, pre-saturated with air, 200 μl of a 100-mM sodium D-Met solution in 50 mM PPB, 20 μl of a ABTS water solution (16 mg/ml), and 10 μl of a peroxidase solution in 50 mM PPB (5 mg/ml) were added to the spectrophotometer's cuvette (working volume 1 ml, optical path 1 cm). After incubation for 10 min at 30°C, a sample of wild-type TvDAAO or the corresponding mutant was added to the cuvette (30 μl).

When determining the maximum reaction rate (V_m) and Michaelis constant (K_M), the concentration of the corresponding D-amino acid was varied from 0.5 to $5 K_M$. An approximate K_M value was determined in a separate experiment by measuring the reaction rate at concentrations of the corresponding D-amino acid of 0.1, 0.5, 1.0, 5.0, 10.0, and 50 mM. The kinetic parameters V_m and K_M were calculated by nonlinear regression using the OriginPro 8.5 SR1 (OriginLab) program. The catalytic constant k_{cat} was calculated based on the V_m value. The concentration of the active enzyme was determined spectrophotometrically based on absorbance at 455 nm using a FAD molar absorption coefficient of $10,800 \text{ m}^{-1}\text{cm}^{-1}$ [6].

Thermal inactivation study

The temperature stability of mutant TvDAAO and the wild-type enzyme was studied in a 0.1-M potassium phosphate buffer, pH 8.0. A series of 0.5 ml plastic test tubes containing 100 μl of the enzyme solution were prepared for each experiment. The tubes were placed to a preheated to the desired temperature water thermostat (temperature control accuracy $\pm 0.1^\circ\text{C}$). The test tubes were sampled one by one after fixed time intervals, rapidly cooled for 1-2 min in ice, and the enzyme activity was measured as described above. The sampling interval was adjusted to achieve a decrease in the enzyme activity to 10-15% of the baseline value during the experiment. The time dependence of the residual activity of the enzyme was plotted in semilogarithmic coordinates and processed using the OriginPro 8.5 SR1 (OriginLab) program as described in [18] to calculate the inactivation rate constant.

Computer simulation

Analysis of the TvDAAO structure, computer simulation of TvDAAO with amino acid mutations, and visualization of the protein globule was performed using the Accelrys Discovery Studio 2.1 software package.

RESULTS AND DISCUSSION

Selection of amino acid residues for site-directed mutagenesis

Rational protein design is a powerful method for studying the structure-function relationships and site-directed changes in an enzyme's properties. Comparison of the amino acid sequences of the enzymes of interest and the enzymes from thermophilic organisms, as well as analysis of the three-dimensional structure (if it is available for at least one enzyme in the family) is used for directional increase of thermal stability of the enzymes by means of identifying the amino acid residues that play an important role in the stability [19]. However, this approach is not applicable in the case of TvDAAO, since it is the most stable enzyme among the presently studied D-amino acid oxidases, and no amino acid sequences of DAAO from thermophilic microorganisms have been identified so far. Therefore, we decided to use one of the common approaches based on the hydrophobization of the α -helices in the enzyme's structure [19, 20] to improve the thermal stability of TvDAAO. This can be achieved using various substitutions, e.g., Ser \rightarrow Ala (most frequently used), Lys \rightarrow Arg, Gly \rightarrow Ala, Ser \rightarrow Thr, Lys \rightarrow Ala, Thr \rightarrow Ala, Lys \rightarrow Glu, Glu \rightarrow Arg, and Asp \rightarrow Arg [21]. Ser \rightarrow Ala substitution usually gives the highest stabilizing effect. For example, the hydrophobization of α -helices by means of Ser \rightarrow Ala substitution was used to increase the temperature stability of the formate dehydrogenase from *Pseudomonas sp. 101* [22].

We have analyzed the 3D-structure of TvDAAO to identify potential Ser residues in α -helices. The following points were taken into account when selecting the Ser residues that can be replaced by Ala: 1) the residues should be part of α -helices; 2) and they should not be conserved as well as located at the active site of the enzyme. The analysis of the TvDAAO structure has revealed a total of Ser 11 residues in the α -helices (*Fig. 1*). Comparison of DAAO amino acid sequences from different sources has shown that the Ser44 residue is conserved. It is located in the cofactor-binding domain of TvDAAO, and its side chain forms two hydrogen bonds with the FAD molecule, as established by computer analysis (*Fig. 2A*). Therefore, this residue was excluded from the list of potential replacement candidates. The Ser157 and Ser161 residues are located at the intersubunit area. Therefore, replacement of these residues is also undesirable, despite the fact that they do not participate in the formation of intersubunit hydrogen bonds [11]. Thus, eight Ser residues were selected to be replaced with Ala residues (positions 67, 77, 78, 105, 270, 277, 335, and 336). The Ser67, Ser105, Ser335, and Ser336

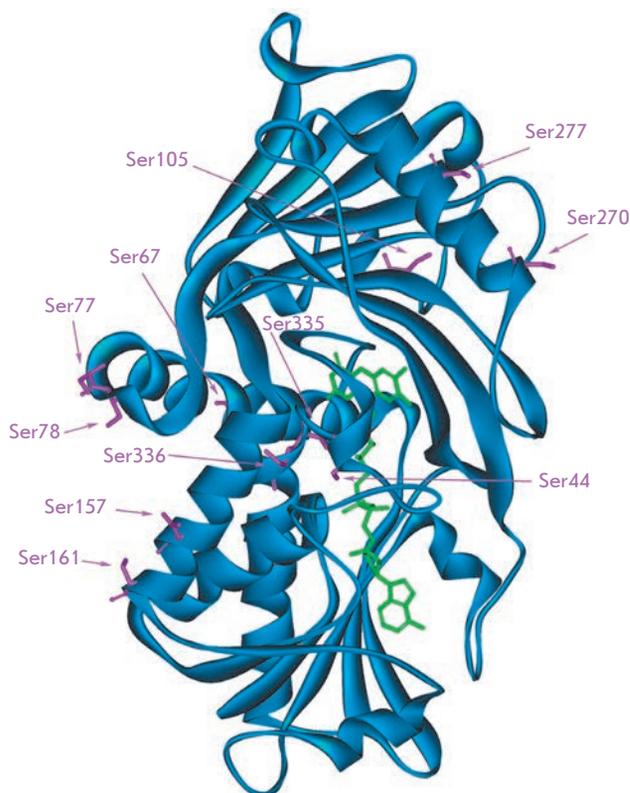


Fig. 1. General view of one subunit of TvDAAO with Ser residues located in alpha-helices

residues are located inside the protein globule, while Ser77, Ser78, Ser270, and Ser277 are exposed to the solution. *Figure 2B–F* shows the position of the selected residues in more detail. Ser67 is located in the middle, while the Ser77 and Ser78 residues are at the end of α 3-helix. Ser105 is located in the short α 4-helix, and Ser270 and 277 are located at the beginning and in the middle of the α 9-helix, respectively. The Ser335 and Ser336 residues are located at the beginning of the α 13-helix. All eight serine residues form two to six hydrogen bonds. Ser78, Ser105, and Ser270 form hydrogen bonds with other amino acid residues of the polypeptide chain including only the atoms involved in the peptide bonds. Since the side chains of these serines residues are not involved in hydrogen bonding with other amino acids, the replacement of these three residues with Ala should not result in losing of hydrogen bonds. The Ser residues at positions 67, 77, 277, 335, and 336 form hydrogen bonds both with the peptide bond atoms and with the side chain hydroxy-groups of other amino acids (see *Fig. 2*). On the one hand, the substitution of these five serine residues will result in a loss of the hydrogen bonds formed by the side chains, which can negatively affect the stability of the enzyme, but on the other hand, increased hydropho-

bicity of the α -helix can stabilize the protein globule, so that the total effect will be positive. Therefore, the replacement of these serine residues is of theoretical interest in terms of the influence of these two factors on the stability of TvDAAO.

Preparation of TvDAAO mutants with Ser / Ala substitutions

The nucleotide substitutions in the *tvdaao* gene that resulted in the desired mutation were introduced using PCR. Three plasmids were sequenced for each of the eight mutant *tvdaao* genes. It has been shown that in all cases only the desired mutations in the *tvdaao* gene were present and that there were no other nucleotide changes. Plasmids with mutated TvDAAO genes were used to transform *E. coli* BL21 (DE3) Codon Plus / pLysS cells. The resulting recombinant strains were cultivated as described in the Experimental section. All eight TvDAAO mutants were synthesized in a soluble form and demonstrated enzymatic activity. Two TvDAAO mutants with Ser78Ala and Ser270Ala substitutions could not be obtained in the purified form, as they were rapidly inactivated during cell disruption, which is indicative of strong destabilization of the protein globule. The remaining six TvDAAO mutants with Ser67Ala, Ser77Ala, Ser105Ala, Ser277Ala, Ser335Ala, and Ser336Ala substitutions were isolated and purified using anion exchange chromatography. Their purity was at least 99% according to the results of analytical electrophoresis in a polyacrylamide gel in the presence of sodium dodecyl sulfate (see *Fig. 3*, lanes 1–6).

Catalytic properties of TvDAAO mutants

The Michaelis constant (K_M) and catalytic constant (k_{cat}) with various D-amino acids were determined for the six TvDAAO mutants, including Ser67Ala, Ser77Ala, Ser105Ala, Ser277Ala, Ser335Ala, and Ser336Ala substitutions. The values of k_{cat} , K_M and catalytic efficiency k_{cat}/K_M of the TvDAAO mutants and the wild-type enzyme with various D-amino acids are shown in *Table 1*. The improvement of the kinetic parameter as compared to that of the wild-type enzyme is shown in bold on a gray background. For clarity, *Fig. 4* shows the catalytic efficiency values ($(k_{cat}/K_M)^{mut}/(k_{cat}/K_M)^{wt}$ 100%) of the TvDAAO mutants with respect to the values of the wild-type enzyme (100%). *Table 1* and *Fig. 4* show that the substitutions resulted in significant changes in the substrate specificity spectrum. The Ser335Ala TvDAAO mutant was the only enzyme that retained enzymatic activity with D-lysine, while Ser77Ala was the only mutant that retained activity with D-threonine. Furthermore, it should be mentioned that virtually all substitutions resulted in an increased catalytic efficiency with D-leucine.

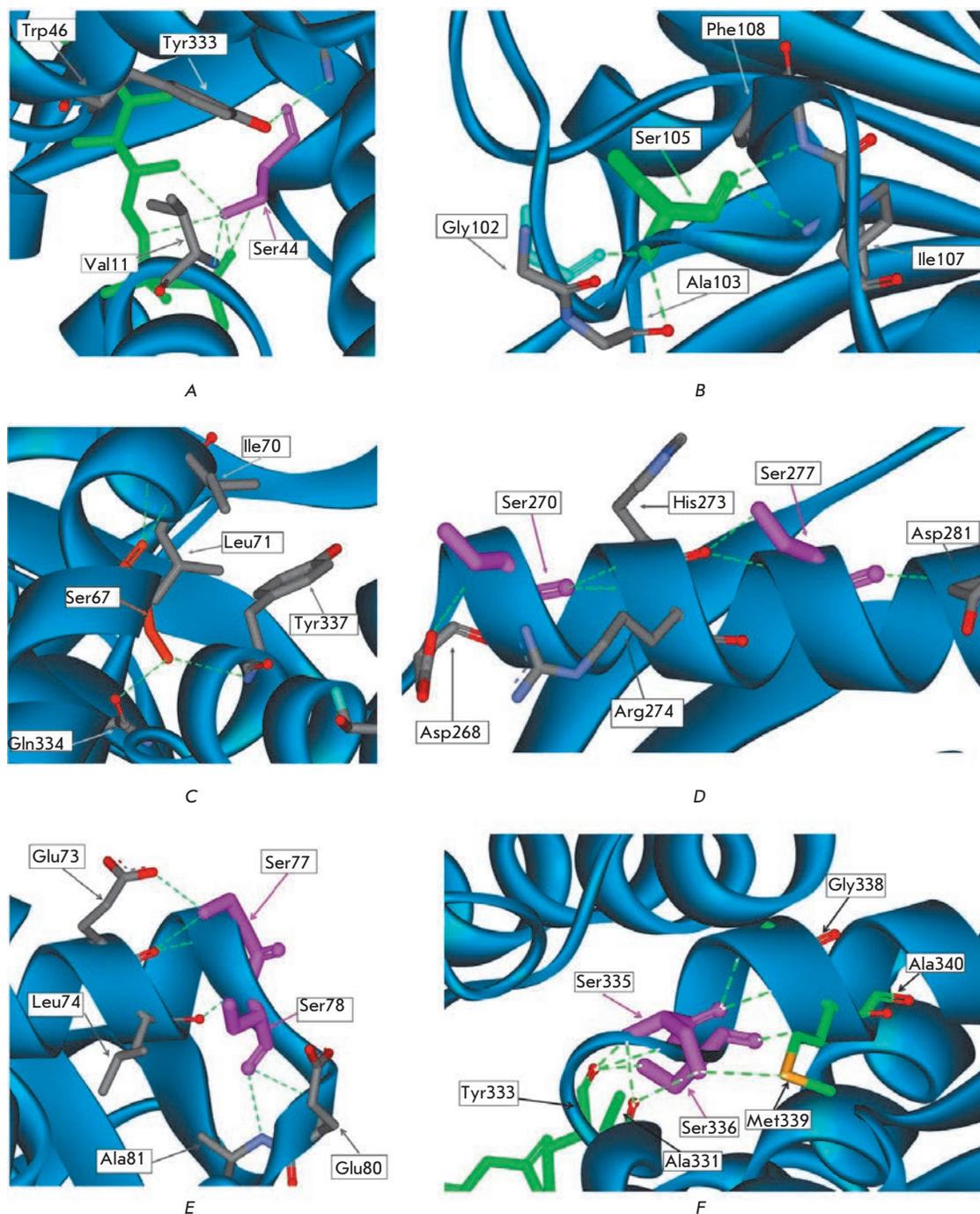


Fig. 2. Location and interactions of the Ser residues in alpha-helices in TvDAAO globule

The following points in relation to the individual mutants should be emphasized:

1. The properties of Ser67Ala TvDAAO are similar to those of the wild-type enzyme with many substrates. A significant increase in the catalytic efficiency (2.5 fold) was observed only with D-Leu. The enzyme was inactive with D-Thr and D-Lys.

2. Ser77Ala TvDAAO shows higher catalytic efficiency only with D-Asn. The enzyme is inactive with D-Thr, and the activity with D-Tyr, D-Met, and D-Val

is significantly decreased. Only this mutant retained its activity with D-Lys.

3. Ser105Ala TvDAAO has the best catalytic parameters between all mutant forms, except for the lack of activity with D-Thr and D-Lys. The catalytic efficiency decreased by 1.3-fold with D-Trp and 1.6-fold with D-Met, but it increased by 1.6-, 1.7-, and 3.0-fold with D-Tyr, D-Phe and D-Leu, respectively.

4. Ser335Ala TvDAAO has a higher catalytic activity with D-Ser as compared to that of the wild-type en-

Table 1. Catalytic properties of mutant TvDAAOs and wild-type on D-amino acids

D-amino acid	Enzyme											
	wt-TvDAAO			TvDAAO Ser67Ala			TvDAAO Ser77Ala			TvDAAO Ser105Ala		
	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹
D-Met	80.5 ± 0.8	0.46 ± 0.03	175	104 ± 3.0	1.35 ± 0.04	77.4	44.0 ± 2.0	0.85 ± 0.09	52.0	153 ± 2.0	1.38 ± 0.04	110
D-Ala	108.6 ± 2.0	16.7 ± 0.7	6.5	180 ± 13	31.0 ± 4.0	5.8	33.7 ± 1.0	11.1 ± 0.1	3.1	218 ± 8.0	31.3 ± 2.0	7.0
D-Ser	20.5 ± 0.9	36.6 ± 3.3	0.56	18.4 ± 0.7	36.0 ± 3.0	0.51	7.9 ± 0.2	18.3 ± 0.8	0.43	10.4 ± 0.4	21.4 ± 1.9	0.49
D-Val	85.3 ± 2.7	14.4 ± 1.2	5.9	133 ± 6.0	21.4 ± 1.6	6.2	19.1 ± 0.5	12.2 ± 0.8	1.56	154 ± 5.0	24.7 ± 1.3	6.3
D-Tyr	22.5 ± 1.9	0.45 ± 0.06	50.0	34.0 ± 4.0	0.87 ± 0.14	38.9	12.4 ± 0.8	6.5 ± 0.7	1.90	50.0 ± 7.0	0.63 ± 0.12	80.0
D-Trp	42.4 ± 1.4	0.49 ± 0.04	86.5	35.9 ± 1.3	0.53 ± 0.4	67.7	43.7 ± 0.8	0.68 ± 0.03	64.3	38.6 ± 1.6	0.60 ± 0.05	64.6
D-Leu	29.1 ± 0.3	0.78 ± 0.02	37.3	31.4 ± 0.5	0.34 ± 0.02	93.1	6.2 ± 0.5	0.20 ± 0.04	31.9	30.4 ± 0.4	0.28 ± 0.02	110
D-Phe	27.2 ± 0.8	0.37 ± 0.04	73.9	30.4 ± 1.2	0.41 ± 0.04	73.6	23.4 ± 0.5	0.32 ± 0.03	73.1	32.3 ± 0.6	0.26 ± 0.02	124.3
D-Asn	62.4 ± 2.0	22.6 ± 1.5	2.8	49.7 ± 1.7	25.7 ± 1.5	1.94	17.2 ± 1.0	4.7 ± 0.7	3.7	48.3 ± 1.2	14.5 ± 0.7	3.3
D-Thr	1.75 ± 0.04	11.1 ± 0.8	0.16	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction
D-Lys	3.54 ± 0.21	29.3 ± 3.4	0.12	no reaction	no reaction	no reaction	<1.2	>50	0.04	no reaction	no reaction	no reaction
D-amino acid	Enzyme											
	TvDAAO Ser277Ala			TvDAAO Ser335Ala			TvDAAO Ser336Ala					
	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹	k_{cat} , s ⁻¹	K_M , mM	k_{cat}/K_M , mM ⁻¹ s ⁻¹
D-Met	48.0 ± 1.0	0.73 ± 0.04	65.8	45.9 ± 1.1	4.6 ± 0.2	9.9	56.8 ± 1.5	2.69 ± 0.13	21.1	2.69 ± 0.13	21.1	21.1
D-Ala	81.0 ± 3.0	20.4 ± 1.4	4.0	8.2 ± 0.4	10.1 ± 1.2	0.82	101 ± 5.0	24.2 ± 1.9	4.2	24.2 ± 1.9	4.2	4.2
D-Ser	no reaction	no reaction	no reaction	34.1 ± 1.7	43.0 ± 5.0	0.79	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction
D-Val	71.0 ± 3.0	34.0 ± 3.0	2.1	20.1 ± 1.3	25.0 ± 3.0	0.82	50.0 ± 2.0	8.2 ± 0.9	6.1	8.2 ± 0.9	6.1	6.1
D-Tyr	0.45 ± 0.06	0.70 ± 0.14	0.65	5.5 ± 0.9	0.58 ± 0.13	9.5	15.8 ± 0.7	0.36 ± 0.03	44.2	0.36 ± 0.03	44.2	44.2
D-Trp	0.49 ± 0.04	0.26 ± 0.02	1.9	79.0 ± 5.0	1.23 ± 0.11	64.1	57.0 ± 1.0	0.69 ± 0.04	82.0	0.69 ± 0.04	82.0	82.0
D-Leu	12.0 ± 2.0	0.15 ± 0.02	80.0	5.9 ± 0.1	0.21 ± 0.01	27.5	5.0 ± 0.1	0.07 ± 0.01	76.4	0.07 ± 0.01	76.4	76.4
D-Phe	15.1 ± 0.4	0.18 ± 0.03	83.9	8.3 ± 0.2	0.38 ± 0.03	21.9	11.8 ± 0.3	0.24 ± 0.02	49.4	0.24 ± 0.02	49.4	49.4
D-Asn	11.8 ± 1.2	8.0 ± 1.5	1.48	46.0 ± 4.0	26.0 ± 3.0	1.78	32.0 ± 2.0	16.0 ± 3.0	2.0	16.0 ± 3.0	2.0	2.0
D-Thr	no reaction	no reaction	no reaction	1.8 ± 0.1	14.0 ± 2.0	0.13	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction
D-Lys	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction	no reaction

* - Improved catalytic parameters of mutant TvDAAOs in comparison to wild-type are marked with bold font and gray background

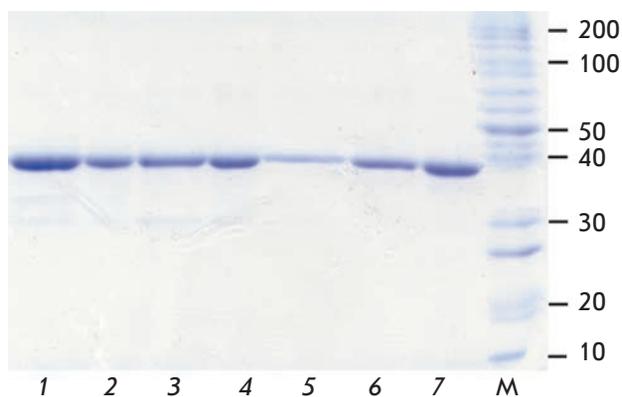


Fig. 3. Analysis by 12% SDS-PAGE of soluble mutant TvDAAOs with amino acid changes: 1 – Ser67Ala, 2 – Ser77Ala, 3 – Ser105Ala, 4 – Ser277Ala, 5 – Ser335Ala, 6 – Ser336Ala and 7 – wild-type enzyme. M – molecular-mass size marker

zyme. Moreover, only this mutant enzyme retained its activity with D-Thr.

Unlike the other enzymes, the Ser277Ala and Ser336Ala TvDAAO mutants were characterized by a complete loss of activity with D-Ser, along with retained activity with D-Ala, which can be used for the selective detection of D-Ala in biological samples in the presence of D-Ser. These enzymes are also inactive with D-Thr and D-Lys, but they exhibit higher catalytic efficiency with D-Leu.

TEMPERATURE STABILITY OF TVDAAO MUTANTS

Stability of Ser78Ala and Ser270Ala TvDAAO

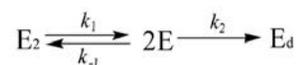
As noted above, Ser78Ala and Ser270Ala substitutions led to strong destabilization of the protein globule, so that the enzymes were inactivated during their isolation from the cells. Computer simulations have shown that the side chains of Ser78 and Ser270 do not form hydrogen bonds with neighboring residues. However, they are located in the immediate vicinity of the Glu80 and Asp268 residues, respectively, which can form hydrogen bonds in solution, both directly and through a water molecule, since in both cases the distance between the hydroxy group of serine and the carboxyl group is about 4 Å. Ser78 and Ser270 are located on the bends at the end of the $\alpha 3$ -helix and at the beginning of the $\alpha 9$ -helix, respectively, and therefore they appear to play an important role in maintaining the stability of TvDAAO secondary structure elements, as evidenced by the strong destabilization upon their replacement by alanine residues.

Stability of Ser67Ala, Ser77Ala, Ser105Ala, Ser277Ala, Ser335Ala, and Ser336Ala TvDAAO mutants

Figure 5A, B shows the time dependence of the residual activity of the TvDAAO mutants at the same concentration. As it can be seen from Fig. 5A, Ser77Ala and Ser335Ala substitutions result in slightly reduced stability. Ser67Ala substitution does not affect the stability of the enzyme, while Ser105Ala substitution results in noticeable stabilization. The most significant destabilization of the protein globule is observed in the case of Ser277Ala and Ser336Ala substitutions (Fig. 5B). The incubation temperature had to be reduced from 56 to 52°C to obtain inactivation curves comparable to those of the other TvDAAO mutants.

MECHANISM OF INACTIVATION OF TVDAAO MUTANTS

We showed [11, 13, 16] that inactivation of wild-type TvDAAO and its various mutants at elevated temperatures proceeds according to the following dissociative mechanism:



According to this mechanism, the first step includes reversible dissociation of the E_2 active dimer to form two inactive monomers E. Irreversible transition of E to the denatured monomer E_d then occurs. This mechanism was analyzed in details by O.I. Poltorak *et al.* [18]. The time dependence of the residual activity of the enzyme in this mechanism is described by a sum of two exponential functions, and the inactivation rate of the enzyme depends on its concentration [11, 13, 16]. The dissociative mechanism of the wild-type TvDAAO thermoinactivation is only observed within a temperature range of 50–60°C, when the rate constants k_1 and k_2 are comparable to each other. The rate constant k_1 increases more rapidly than the rate constant k_2 upon increasing temperature; therefore, the first and second steps become the limiting ones at temperatures below 50 and above 60°C, respectively, and the kinetics of inactivation is described by a single exponential function under these conditions, similarly to that of unimolecular reactions.

Analysis of the time dependence of the residual activity shows that the thermal inactivation mechanism of the Ser67Ala, Ser77Ala, Ser105Ala, Ser335Ala TvDAAO mutants (Fig. 5A) and Ser277Ala and Ser336Ala TvDAAO mutants (Fig. 5B) also does not differ from that of the wild-type enzyme. As an example, Fig. 6A, B shows the residual activity of Ser77Ala TvDAAO mutant vs incubation time in semilogarithmic coordinates at various temperatures and concentrations. Sim-

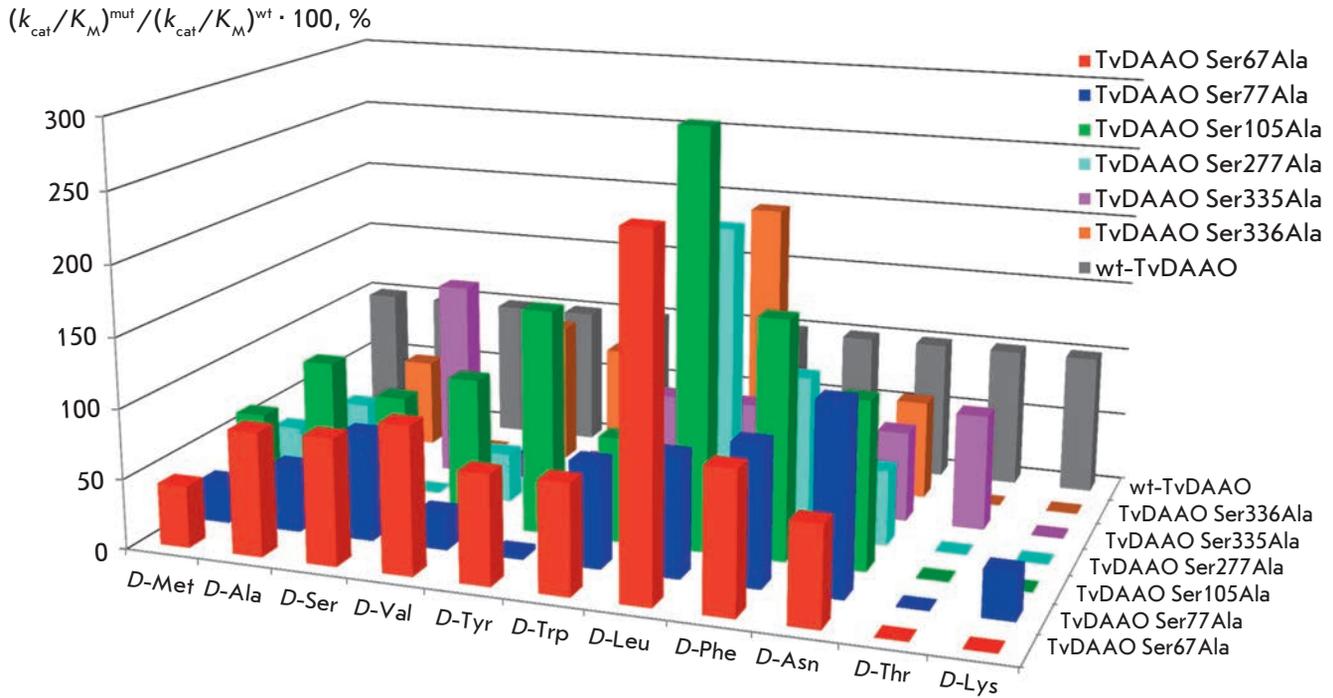


Fig. 4. Relative catalytic efficiencies ($(k_{cat}/K_M)^{mut}/(k_{cat}/K_M)^{wt} \cdot 100\%$) of mutant TvDAAOs with Ser67Ala, Ser77Ala, Ser105Ala, Ser277Ala, Ser335Ala, and Ser336Ala substitutions. The catalytic efficiency of the wild-type TvDAAO is taken as 100%

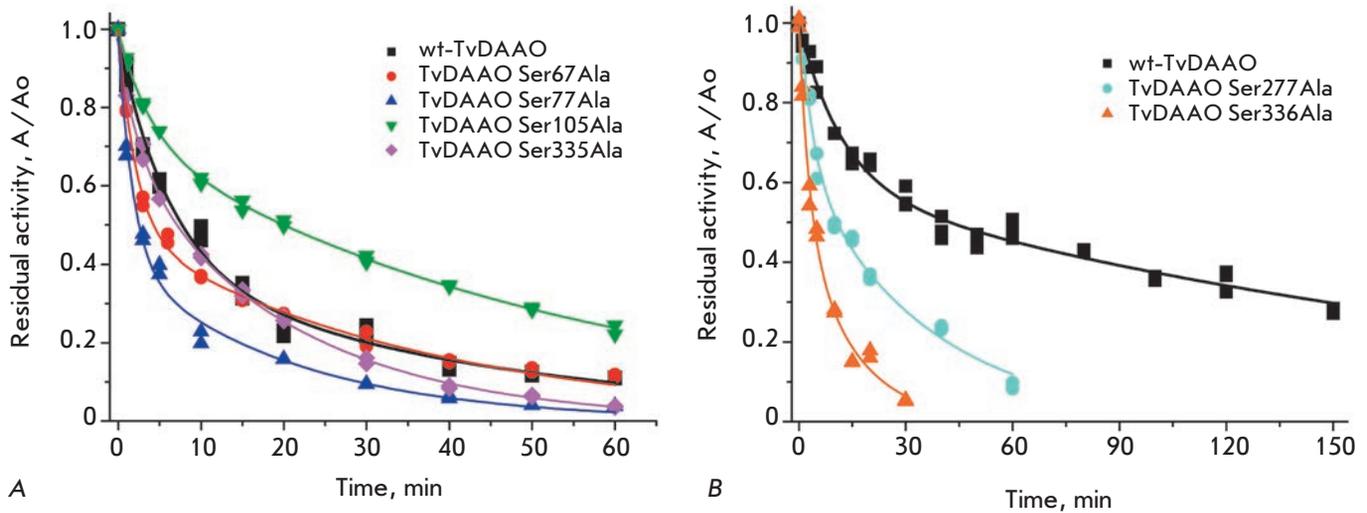
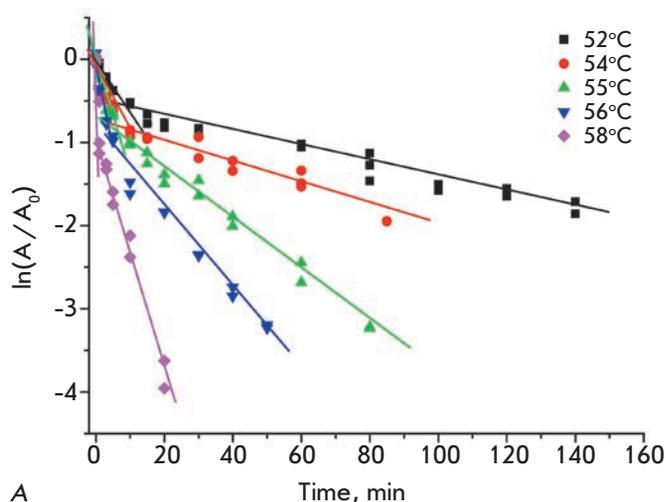


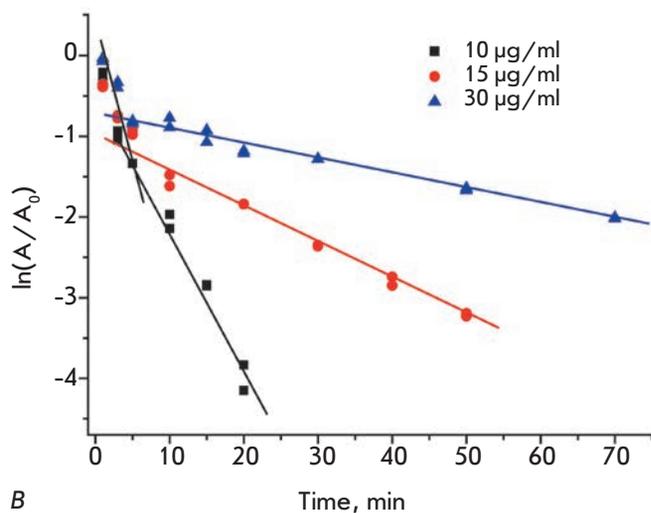
Fig. 5. A – Time dependence of the residual activity of the wild-type and mutant TvDAAOs S67A, S77A, S105A, and S335A at 56°C. 0.1 M potassium phosphate buffer, pH 8.0. Enzyme concentration is 10 µg/ml
 B – Time dependence of the residual activity of the wild-type and mutant TvDAAOs S277A and S336A at 52°C. 0.1 M potassium phosphate buffer, pH 8.0. Enzyme concentration is 10 µg/ml

ilar dependences were obtained for all other mutant enzymes. The presence of the break points on the thermal inactivation curves in semilogarithmic coordinates at different temperatures and increase in the slope of the second linear section along with a decrease in the

initial concentration of the enzyme provides evidence that thermal inactivation occurs through a dissociative mechanism [18]. We have calculated the rate constants of the thermal inactivation of the Ser67Ala, Ser77Ala, Ser105Ala, Ser277Ala, Ser335Ala, and Ser336Ala



A

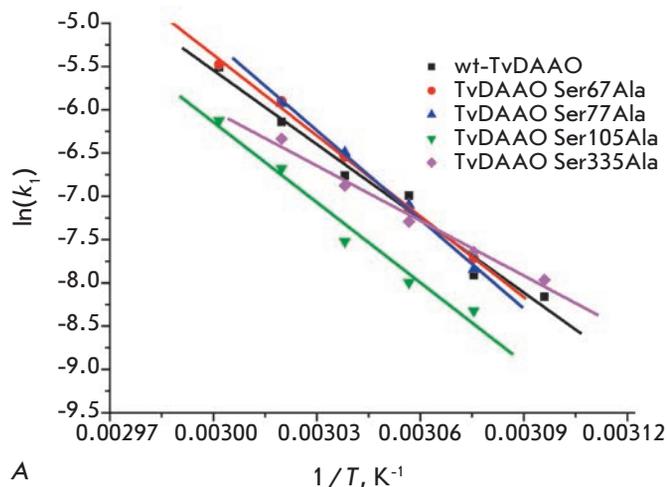


B

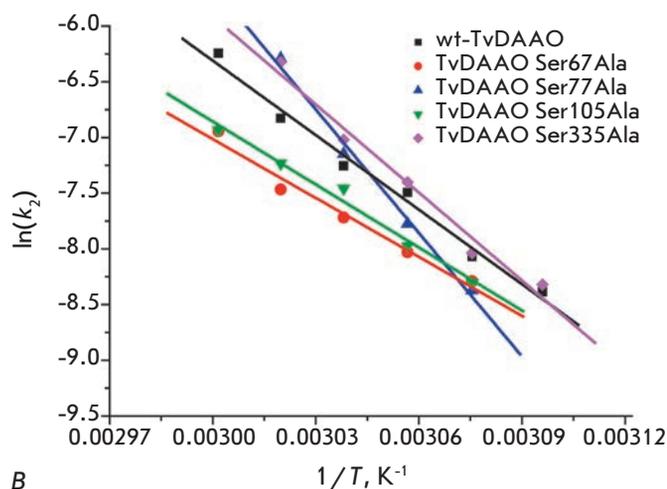
Fig. 6. Time dependence of the residual activity of mutant TvDAAO S77A in semi-logarithmic coordinates. A – thermal inactivation at different temperatures and fixed enzyme concentration of 15 $\mu\text{g/ml}$, B – thermal inactivation at different enzyme concentrations and temperature of 56°C. 0.1 M potassium phosphate buffer, pH 8.0

TvDAAO mutants for both stages of the process on the basis of the experimental dependence of the residual enzyme activity vs incubation time (Table 2).

Ser277Ala and Ser336Ala mutations in TvDAAO resulted in a shift in the temperature range associated with the dissociative mechanism by 4°C towards lower temperatures as compared to that of the wild-type enzyme. The Ser336Ala TvDAAO mutant was the least stable of all the mutants that were obtained and purified (see Fig. 5B, Tab. 2). At a temperature of 52°C (corresponding kinetic curves are shown in Fig. 5B), Ser277Ala and Ser336Ala mutations led to 3.9- and 7.7-fold increases in the first-step inactivation rate constants and 1.2- and 5.9-fold increases in the second-step in-



A



B

Fig. 7. Temperature dependences of the inactivation rate constants of the first (A) and second (B) stages for mutant TvDAAOs with S67A, S77A S105A, S335A mutations and wild-type TvDAAO in semi logarithmic coordinates $\ln(k_{in})$ vs $1/T$. 0.1 M potassium phosphate buffer, pH 8.0

activation rate constants, respectively (Tab. 2). Therefore, Ser336Ala mutation results in a destabilizing effect associated with both the first step of inactivation (enzyme dissociation into monomers), and the second step (denaturation of the protein globule). Ser277Ala substitution leads mainly to an increase in the first-step inactivation rate, but the effect is not as significant as that associated with Ser336Ala substitution. The higher effect of the enzyme destabilization associated with Ser336Ala substitution may be due to the fact that the Ser336 residue is located at the end of the $\alpha 13$ -helix and forms a hydrogen bond with the peptide bond of the Tyr333 residue, which, in turn, is in close contact with the *Si*-side of the isoalloxazine cycle of FAD and

Table 2. Kinetic parameters of dissociative thermal inactivation of mutant and wild-type TvDAAOs

Enzyme	Parameter	Temperature, °C							
		46	48	50	52	54	56	58	60
TvDAAO Ser67Ala	$k_1 \cdot 10^4, s^{-1}$	-	-	-	4.45	7.8	14.4	27.4	41.8
	$k_2 \cdot 10^4, s^{-1}$	-	-	-	2.51	3.25	4.44	5.7	9.6
TvDAAO Ser77Ala	$k_1 \cdot 10^4, s^{-1}$	-	-	-	3.92	8.3	15.1	26.7	-
	$k_2 \cdot 10^4, s^{-1}$	-	-	-	2.31	4.20	7.9	18.5	-
TvDAAO Ser105Ala	$k_1 \cdot 10^4, s^{-1}$	-	-	-	2.43	3.37	5.4	12.6	21.9
	$k_2 \cdot 10^4, s^{-1}$	-	-	-	2.48	3.44	5.8	7.2	9.7
TvDAAO Ser277Ala	$k_1 \cdot 10^4, s^{-1}$	3.02	5.4	10.2	14.4	25.9	-	-	-
	$k_2 \cdot 10^4, s^{-1}$	0.93	1.41	3.25	3.91	7.6	-	-	-
TvDAAO Ser335Ala	$k_1 \cdot 10^4, s^{-1}$	-	-	3.47	4.80	6.8	10.3	17.7	-
	$k_2 \cdot 10^4, s^{-1}$	-	-	2.43	3.23	6.1	8.9	17.9	-
TvDAAO Ser336Ala	$k_1 \cdot 10^4, s^{-1}$	2.88	6.8	13.8	28.4	-	-	-	-
	$k_2 \cdot 10^4, s^{-1}$	0.90	3.67	9.3	18.5	-	-	-	-
wt-TvDAAO	$k_1 \cdot 10^4, s^{-1}$	-	-	2.86	3.67	9.2	11.6	21.5	40.5
	$k_2 \cdot 10^4, s^{-1}$	-	-	2.28	3.13	5.6	7.1	10.8	19.4

* The decrease in inactivation rate constants of mutants as compared to the wild-type enzyme is marked with a green background; the increase is marked with a red background. Different shades show the extent of the effects – a greater effect corresponds to the darker color.

conserved residue Ser44 (*Fig. 2F*). In addition, Ser336 and Tyr333 occur in corresponding positions of the D-amino acid oxidases being the most homologous to TvDAAO, and apparently they play an important role in maintaining the conformation required for cofactor binding. The Ser277 residue is located in the middle of the $\alpha 9$ -helix on the surface of the enzyme (*Fig. 2D*) and forms a hydrogen bond with the carbonyl oxygen atom of the peptide bond of the His273 residue. The loss of this hydrogen bond could have a negative effect on the thermal stability of TvDAAO.

The temperature range, in which inactivation of the enzyme associated with substitutions of four Ser residues (67, 77, 105 and 335) occurs through the dissociative mechanism, remained the same, but the values of the thermal inactivation rate constants changed as compared to those of the wild-type enzyme (*Tab. 2*).

As mentioned above, the temperature dependences of the first- and second-stage rate constants differ. *Figure 7A, B* shows the dependence of $\ln(k)$ vs $1/T$ for the first- and second-stage rate constants, respectively.

Ser105Ala substitution results in a higher stability of the enzyme throughout the whole temperature range of 52 to 60°C (*Tab. 2, Fig. 5A*). The dependences of the

thermal inactivation rate constants for both stages are very close to those of the wild-type enzyme, but they are lower on the respective charts over the range from 52 to 60°C, as shown below (*Fig. 7A, B*). Ser105Ala substitution on average led to two-fold higher thermal stability of TvDAAO as compared to that of the wild-type enzyme at the first stage of thermal inactivation and 1.5-fold higher stability at the second stage, which is rather significant for this enzyme. A similar stabilizing effect was observed at the second stage of inactivation of the Ser67Ala TvDAAO mutant (about 1.6-fold), and the temperature dependence of the inactivation rate constant is also close to that of the wild-type enzyme, but the parameters of the first stage of the thermal inactivation of Ser67Ala TvDAAO are slightly inferior to those of the wild-type enzyme over the entire temperature range; i.e., the stability decreased by 20% on average. Nevertheless, this substitution also results in overall stabilization of the enzyme, although this stabilization is lower than in the case of Ser105Ala substitution. The Ser67 and Ser105 residues are located inside the protein globule in the middle of the $\alpha 3$ -helix and short $\alpha 4$ -helix. The hydroxy-group of Ser105 residue forms no hydrogen bonds with neighboring residues and is adjacent to

the hydrophobic residues Leu100, Ala103, and Ile107, while the hydroxy-group of Ser67 forms two hydrogen bonds with the polypeptide chain atoms of the Gln334 and Tyr337 residues, and it is also located in close vicinity to the hydrophobic residues Trp51, Leu70, Leu71, and the benzene ring of the Tyr337 residue. Ser105Ala substitution facilitates hydrophobic interactions within the protein globule without breaking any hydrogen bonds, which probably leads to an increase in the thermal stability of TvDAAO. Furthermore, the three dimensional structure of the enzyme may apparently undergo some conformational changes resulting in the stabilization of the dimer and improvement of its catalytic properties. By contrast, Ser67Ala substitution results in the loss of two hydrogen bonds with the Tyr337 and Gln334 residues, which are located in the spatially close α 13-helix, but could result in stronger hydrophobic interactions, which contributes to stabilization of the protein globule, as evidenced by the results of experiments.

Ser77Ala and Ser335Ala TvDAAO mutants differ from the rest of mutants in the temperature dependences of the rate constants of the first and the second stages of inactivation (Fig. 7A, B). The first- and the second-stage rate constants of the Ser77Ala mutant increase more rapidly with increasing temperature than those of the wild-type enzyme, and the dependence is more significant in the case of k_2 (Fig. 7B), which results in a lower overall stability of Ser77Ala TvDAAO at temperatures above 54°C, whereas at lower temperatures this mutant is more stable than the wild-type enzyme. In case of Ser335Ala TvDAAO, the temperature dependences of the inactivation rate constants k_1 and k_2 oppositely differ from those of the wild-type enzyme. Increase in k_1 with temperature is less pronounced, while k_2 is more strongly temperature-dependent. As a result, Ser335Ala TvDAAO is more stable than the wild-type enzyme at the first stage of inactivation at temperatures > 54°C; and at the second stage, at temperatures < 50°C. Thus, due to the complex temperature dependence of the constants of both inactivation stages, the stability of Ser335Ala TvDAAO at each temperature is given by the ratio of the constants of each inactivation stage. Nevertheless, the overall stability only slightly differs from that of the wild-type enzyme.

CONCLUSION

The effect of hydrophobization of α -helices in the structure of the D-amino acid oxidase from yeast *Trigonopsis variabilis* was studied by replacing eight serine residues with alanine residues. From the viewpoint of the structure-stability relationship, it is interesting that replacement of Ser residues on the surface of TvDAAO at positions 77, 78, 270, and 277 results in destabilization of the enzyme, while replacement of Ser 67, 105, 335, and 336 inside the protein globule leads to a reduced stability only in one case out of four. It should also be noted that replacement of the serine residues located at the ends of α -helices also negatively affects the thermal stability of the enzyme. These data are directly in contradiction to the results obtained for the formate dehydrogenase from the *Pseudomonas sp.* 101 bacteria [22]. The highest enzyme stabilization effect (1.6-fold) was observed upon replacing the Ser131 located on the surface of the protein globule, with the Ala residue. Furthermore, stabilization effect was also observed upon replacing Ser184, which is located at the end of the α 6-helix [22]. We therefore can conclude that, despite the generality of the approach based on hydrophobization of α -helices, the value and effect of stabilization depend directly on the structural features of the particular protein or enzyme.

In conclusion, we would like to mention that there are cases in protein engineering when a single amino acid substitution results in a significant stabilization of the enzyme [23, 24]. However, usually the improvement of thermal stability can be achieved by combining several successful point mutations. Each of these individual mutations has a moderate stabilization effect, while the temperature stability of a multipoint mutant enzyme becomes significant [12]. Thus, the method of hydrophobization of α -helices cannot be considered as the basic one, but rather as the additional approach to improving the stability of enzymes, due to the low stabilization effects of point amino acid mutations.

This work was supported by the Russian Foundation for Basic Research (grant № 14-04-00859-a).

REFERENCES

1. Tishkov V.I., Khoronenkova S. V. // Biochemistry (Moscow) 2005. V. 70. № 1. P. 40–54.
2. Khoronenkova S.V., Tishkov V.I. // Biochemistry (Moscow). 2008. V. 73. № 13. P. 1511–1518.
3. Pollegioni L., Piubelli L., Sacchi S., Pilone M.S., Molla G. // Cell. Mol. Life Sci. 2007. V. 64. № 11. P. 1373–1394.
4. Pollegioni L., Molla G. // Trends Biotechnol. 2011. V. 29. № 6. P. 276–283.
5. Pollegioni L., Molla G., Sacchi S., Rosini E., Verga R., Pilone M.S. // Appl. Microbiol. Biotechnol. 2008. V. 78. № 1. P. 1–16.
6. Pollegioni L., Caldinelli L., Molla G., Sacchi S., Pilone M.S. // Biotechnol. Prog. 2004. V. 20. № 2. P. 467–473.
7. Gabler M., Hensel M., Fischer L. // Enzyme Microb. Technol. 2000. V. 27. № 8. P. 605–611.
8. Yurimoto H., Hasegawa T., Sakai Y., Kato N. // Biosci. Biotechnol. Biochem. 2001. V. 65. P. 627–633.
9. Geueke B., Weckbecker A., Hummel W. // Appl. Microbiol.

RESEARCH ARTICLES

- Biotechnol. 2007. V. 74. № 6. P. 1240–1247.
10. Savin S.S., Chernyshev I.V., Tishkov V.I., Khoronenkova S.V. // Moscow Univ. Chem. Bul. 2006. V. 61. № 1. P. 13–19.
11. Cherskova N., Khoronenkova S., Tishkov V. // Rus. Chem. Bull. 2010. V. 59. № 1. P. 1–7.
12. Tishkov V.I., Popov V.O. // Biomol. Eng. 2006. V. 23. № 2–3. P. 89–110.
13. Komarova N.V., Golubev I.V., Khoronenkova S.V., Tishkov V.I. // Rus. Chem. Bull. 2012. V. 61. № 7. P. 1489–1496.
14. Wong K.-S., Fong W.-P., Tsang P.W.-K. // Nat. Biotechnol. 2010. V. 27. № 1. P. 78–84.
15. Wang S.-J., Yu C.-Y., Lee C.-K., Chern M.-K., Kuan I.-C. // Biotechnol. Lett. 2008. V. 30. № 8. P. 1415–1422.
16. Komarova N.V., Golubev I.V., Khoronenkova S.V., Chubar T.A., Tishkov V.I. // Biochemistry (Moscow). 2012. V. 77. № 10. P. 1181–1189.
17. Khoronenkova S.V., Tishkov V.I. // Anal. Biochem. 2008. V. 374. № 2. P. 405–410.
18. Poltorak O.M., Chukhray E.S., Torshin I.Y. // Biochemistry (Moscow). 1998. V. 63. № 3. P. 303–311.
19. Munoz V., Serrano L. // Nat. Struct. Biol. 1994. V. 1. P. 399–409.
20. Munoz V., Serrano L. // J. Mol. Biol. 1995. V. 245. P. 275–296.
21. Rose G., Geselowitz A., Lesser G. // Science. 1985. V. 229. № 7. P. 834–838.
22. Rojkova A., Galkin A., Kulakova L., Serov A., Savitsky P., Fedorchuk V., Tishkov V. // FEBS Lett. 1999. V. 445. № 1. P. 183–188.
23. Alekseeva A.A., Serenko A.A., Kargov I.S., Savin S.S., Kleymenov S.Y., Tishkov V.I. // Protein Eng. Des. Sel. 2012. V. 25. № 11. P. 781–788.
24. Alekseeva A.A., Savin S.S., Kleimenov S.Y., Uporov I.V., Pometun E.V., Tishkov V.I. // Biochemistry (Moscow). 2012. V. 77. P. 1199–1209.

Changes in Gene Expression Associated with Matrix Turnover, Chondrocyte Proliferation and Hypertrophy in the Bovine Growth Plate

E. V. Tchetina^{1*}, F. Mwale², A. R. Poole³

Joint Diseases Laboratory, Shriners Hospitals for Children and Departments of Surgery and Medicine, McGill University, Montreal, Quebec, Canada

Present Addresses:

¹Clinical Immunology Department, Research Institute of Rheumatology, Russian Academy of Sciences, Moscow, 115522, Russia

²Orthopaedics Research Laboratory, Jewish General Hospital, Lady Davis Institute for Medical Research, Division of Orthopedic Surgery, McGill University, Montreal, Quebec, H3T 1E2, Canada

³Department of Surgery, McGill University, McGill University, Montreal, Quebec, H3A 0G4, Canada

*Email: etchetina@mail.ru

Received 23.01.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT The aim of the study is to investigate the interrelationships between the expression of genes for structural extracellular matrix molecules, proteinases and their inhibitors in the bovine fetal growth plate. This was analyzed by RT-PCR in microsections of the proximal tibial growth plate of bovine fetuses in relationship to expression of genes associated with chondrocyte proliferation, apoptosis, and matrix vascularization. In the resting zone the genes for extracellular matrix molecule synthesis were expressed. Extracellular matrix degrading enzymes and their inhibitors were also expressed here. Onset of proliferation involved cyclic upregulation of cell division-associated activity and reduced expression of extracellular matrix molecules. Later in the proliferative zone we noted transient expression of proteinases and their inhibitors, extracellular matrix molecules, as well as activity associated with vascularization and apoptosis. With the onset of hypertrophy expression of proteinases and their inhibitors, extracellular matrix molecules, as well as activity associated with vascularization and apoptosis were significantly upregulated. Terminal differentiation was characterized by high expression of proteinases and their inhibitors, extracellular matrix molecules, as well as activity associated with apoptosis. This study reveals the complex interrelationships of gene expression in the physis that accompany matrix assembly, resorption, chondrocyte proliferation, hypertrophy, vascularization and cell death while principal zones of the growth plate are characterized by a distinct signature profile of gene expression.

KEYWORDS growth plate, gene expression, proteinases, chondrocyte differentiation.

ABBREVIATIONS ECM-extracellular matrix; MMP-metalloproteinase; TIMP- tissue inhibitor of metalloproteinases; HAS-hyaluronic acid synthase; COL- collagen; ADAMTS- A Disintegrin And Metalloproteinase with Thrombospondin Motifs; FGF- fibroblast growth factor; PTHrP- parathyroid hormone related peptide; Cbfa1-core-binding factor subunit alpha-1 (CBF-alpha-1); TGFb1- transforming growth factor beta 1; Ihh- Indian hedgehog; VEGF-vascular endothelial growth factor; PAI-1-plasminogen activator inhibitor-1; GAPDH-glyceraldehyde 3-phosphate dehydrogenase; RNA – ribonucleic acid; RT-PCR- Reverse Transcriptase Polymerase Chain Reaction; cDNA-complementary DNA.

INTRODUCTION

Endochondral ossification is a process involving chondrogenesis, chondrocyte hypertrophy, matrix mineralization, and vascularization followed by bone formation [1]. It begins during long bone formation in the embryo. After birth until adulthood, growth of the long bone is

centered in the cartilagenous growth plates, leading to an increase in bone length and epiphyseal growth. It is also an essential component of fracture repair.

In the growth plates distinct zones can be observed. Cells of the resting zone chondrocytes produce large amounts of extracellular matrix (ECM). In contrast cells

of the proliferative zone divide to give rise to columns of flattened cells that also secrete an ECM. At this time they express cell cycle-related genes such as cyclins. In the zone of maturation the cells round up and begin to enlarge into hypertrophic chondrocytes. The upper hypertrophic zone is characterized by cells that have enlarged 5- to 10-fold by a reduction in matrix volume per total tissue volume and which synthesize type X collagen [2-4]. In the lower hypertrophic zone calcification of the extracellular matrix occurs mainly in the longitudinal septa. The mineralization process, in combination with low oxygen tension, attracts blood vessels from the underlying primary spongiosum. Subsequently, the mineralized chondrocytes undergo apoptotic cell death [5].

The ECM of chondrocytes is a complex structure although 3 structural entities can be distinguished [6]. One of them is the complex of aggrecan molecules bound to hyaluronan and assembled into large aggregates. It is responsible for the cartilage compressive stiffness creating a highly hydrated matrix the expansion of which is constrained by a network of collagen fibrils composed of type II collagen, as well as a filamentous network of type VI collagen. Type II collagen fibrils contain a number of molecules at their surface, such as type IX collagen, decorin and fibromodulin. The key role of this network is to provide the tensile properties of this tissue. The non-fibrillar filaments of type VI collagen are involved both in cell-matrix and matrix-matrix interactions [6].

Changes in composition of the ECM occur as chondrocytes divide and mature. Metalloproteinases (MMPs) are generally considered to play a principal role in the cleavage of matrix macromolecules including type II collagen and aggrecan [3]. Only collagenases such as MMP-13, MMP-14 and cathepsin K, are capable of cleaving the triple helix of type II collagen [6]. This results in the unwinding (denaturation) of the triple helical domain which becomes susceptible to secondary cleavage by collagenases and other metalloproteinases such as stromelysin-1 (MMP-3) and gelatinases A and B (MMP-2 and MMP-9, respectively) [7]. MMP-13 is involved in the resorption of type II collagen that occurs during chondrocyte hypertrophy [8, 9]. Proteoglycan aggrecan can be cleaved by MMPs and by aggrecanases -1 and -2, ADAMTS-4 and ADAMTS-5, respectively [10]. In contrast the mechanism of type VI collagen degradation remains unclear. It is resistant to several extracellular matrix metalloproteinases in vitro including collagenases [11]. In cartilage MMP-2 or membrane-bound MMPs may be involved in its cleavage [12].

The activity of MMPs is further regulated by a family of specific inhibitors - tissue inhibitors of metallopro-

teinases, namely TIMPs -1, -2, -3 and -4 [13]. TIMP-1 and -2 inhibit the activity of all MMPs, whereas TIMP-3 only inhibits MMP-1, -2, -3, -9 and -13 [14]. Besides inhibiting MMPs, TIMPs also appear to perform other functions. TIMP-1 and -2 exhibit growth factor activity [15] and TIMP-3 is an active mitogen [16].

The complex coordinated regulation of chondrocyte maturation in the growth plate is exerted both by the systemic hormones and chondrocyte autocrine growth factors [5]. In our previous studies of the bovine growth plate we have shown two peaks of gene expression [17]. An increase in gene expression in the early proliferative zone was associated with the upregulation of the regulatory growth factors *FGF-2* and *PTHrP*. In contrast the second more pronounced peak of gene expression in the early hypertrophic zone was accompanied by the increase in *Cbfa1*, *TGFβ1* and *Indian hedgehog (Ihh)* expression. In the present study we extend the previous investigations to explore the relationships of gene expression patterns of matrix proteins to other proteinases and their inhibitors to the cellular and extracellular changes that occur in the physis of the bovine growth plate. These observations help provide more insight into the complex interrelationships of the expression of these molecules during this critical stage in endochondral ossification.

EXPERIMENTAL

Tissue Preparation

Bovine fetuses obtained from a local abattoir immediately after the slaughter of pregnant cows, were transported to the laboratory. Fetal age was determined by measurement of tibial length [18]. Fetuses ranged from 190 to 210 days old. Tissue preparation was essentially as described [2, 8]. Only blocks of growth plate with a flat fracture surface were used. Tissue blocks were trimmed to provide cross-sectional areas of approximately 25 mm². One hundred micrometer thick transverse sections were cut parallel to the fracture face (using a Vibratome; Ted Pella, Inc., California, USA), starting at the fracture face and extending through the hypertrophic zone into the upper proliferative zone of fetal bovine growth plate. They represented tissue labeled as A, B, C, and so on, from the fracture face. Their locations have been previously characterized [2, 8]. A series of sections of four growth plates was pooled (A with A and B with B, etc.) to permit collection of a sufficient amount of tissue for the analyses. Wet weights were determined immediately after sectioning: the weights ranged from 10 to 15 mg, depending on the sample. The weights of samples A and B were lower due to some irregularity of the fracture face.

Total RNA Isolation and Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)

Total RNA was isolated by a modification of the method of Chomczynski and Sacchi, which was described previously [9]. The RT-reaction was performed using total RNA isolated from the cartilage in a total volume of 20 ml using SuperScript TMII H-Reverse Transcriptase (as recommended by Invitrogen, Canada, Inc.).

Oligo sequences used for PCR are shown in Table 1. PCR was performed in a total volume of 25 μ l containing: 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 0.4 mM each of dATP, dGTP, dCTP, dTTP, 0.8 mM of each primer, 1ml of RT mixture and 2.5 units of AmpliTaq DNA polymerase (Perkin Elmer). The 30 cycles of PCR included denaturation (95°C, 1 min), annealing (50°C, 1 min) and extension (72°C, 5 min). After agarose (1.6%) gel electrophoresis, PCR products were visualized by ethidium bromide staining. GAPDH was used as reference for gel loading. The band intensities were determined to be below saturation by dilution analyses. Each analysis was performed at least 3 times at different dilutions of each sample of the original cDNA. The result of the single dilution for all the samples in a given set which showed most clearly differences in expression (e.g. COL2A1) is presented in Fig. 2 and 3. Results were analyzed using NIH 1.60 software to determine the pixel intensity for each band and autobackground subtraction was used to control for background signal (Fig. 3). These results were reproducible for growth plates from three different fetuses.

The isolated clones of each amplified cDNA fragment were sequenced (Sheldon Center, McGill Univer-

sity) to verify the identity of each cDNA product. To confirm the lack of chromosomal DNA contamination of RNA samples, PCR was also performed with RNA aliquots. To avoid variation in efficiency between experiments, all sections were simultaneously subjected to reverse transcription and all samples of cDNA were simultaneously amplified in PCR.

RESULTS AND DISCUSSION

Sequential transverse sections of the bovine tibial primary proximal growth plate (Fig. 1), which represent the hypertrophic (A-C), proliferative (D-J) and resting (K-L) zones [17], were generated. Using RT-PCR analyses of sequential transverse sections of the growth plate cartilage the expression of markers of chondrocyte proliferation and terminal differentiation has already been determined [17]. Here we present our analyses of gene expression of ECM proteins, HA synthase-2, and proteinases in the bovine fetal growth plate in the course of chondrocyte differentiation. We repeated these analyses several times on different fetuses. The data that is shown is representative of our repeated analyses. The data describing the expression of *GAPDH*, *cyclin B2*, *COL2A1*, *COL10A1*, *osteocalcin*, *MMP-13* and *MMP-9* is reproduced from our previous study [17] for reference.

In our RT-PCR analyses (Fig. 2) and its schematic presentation in relationship to *GAPDH* expression (Fig. 3) the onset of proliferation was defined at section J where the upregulation of the *cyclin B2* expression is observed. The onset of terminal differentiation was considered as sample C, where the expression of *COL10A1* the marker of hypertrophic chondrocytes is first detected.

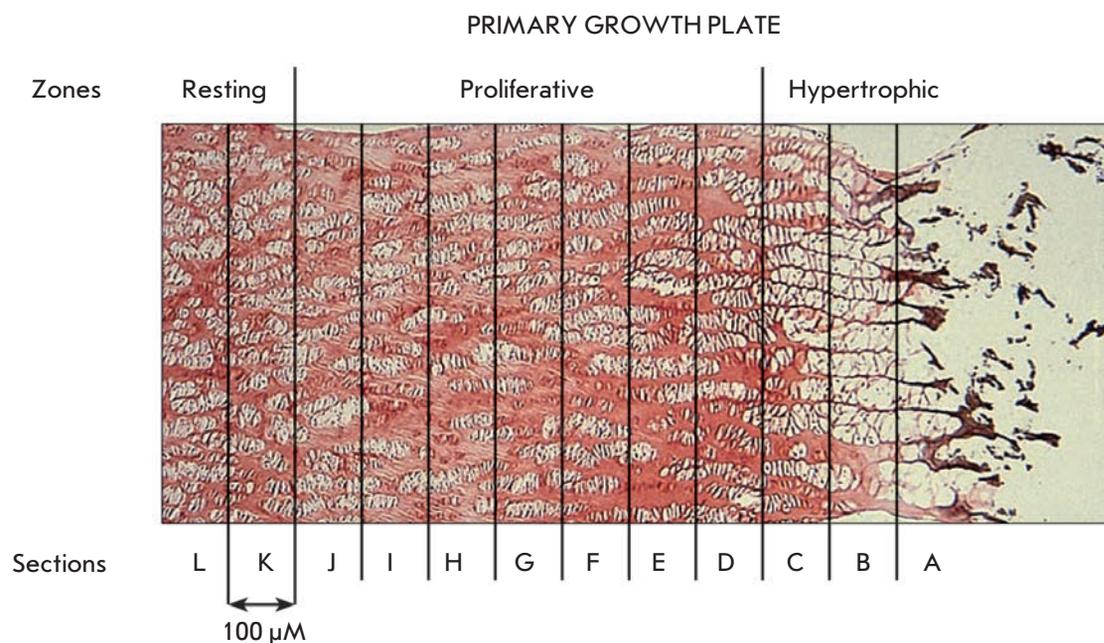


Fig. 1. Representation of the organization of the primary growth plate showing sampling sites extending from the hypertrophic zone through the proliferative zone into the resting zone

The resting zone

The resting zone (sections L and K) is characterized by the expression of extracellular matrix genes, namely *COL2A1*, *COL6A3*, *fibromodulin* and *decorin* being highest in section K. *HAS-2* expression is more pro-

nounced (in section L) than that of *aggrecan* in section K. *Osteocalcin* expression is also detected in section L. Of the proteinases tested *cathepsin K*, *MMP-14*, *MMP-13* (weakly) and *MMP-3* were all expressed in this zone. In contrast the expressions of the MMP in-

Table 1. Oligo sequences used for PCR

Genes	Forward primer	Reverse primer
Collagenase 3 (MMP-13)	GATAAAGACTATCCGAGAC	GAGTAACCGTATTGTTTCG
Membrane type 1-MMP (MMP-14)	GCATCCAGCAACTTTATG	CATCTGTGACGGAACTTTG
Stromelysin-1 (MMP-3)	TGCGTGGCAGTTTGCTCAGCC	GAGGTGACTCCACTCACATTC
Gelatinase A (MMP-2)	GCTACACACCTGATCTG	GACGGCAAGTATTGTTCTG
Gelatinase B (MMP-9)	GCAGAGGAATACCTGTAC	CACAACATCACCTACTG
Tissue inhibitor of metalloproteinase-1 (TIMP-1)	GAAAACCTGCAGGATGGAC	CACCAAGACCTACACTGTTG
Tissue inhibitor of metalloproteinase-2 (TIMP-2)	GGATATAGAGTTTATCTACAC	CATGATCCCGTGCTACATCTC
Tissue inhibitor of metalloproteinase-3 (TIMP-3)	CTTAGGCTGGAGGTCAACAAG	CAAGAACGAGTGTCTGGAC
Osteocalcin (bone Gla protein)	CTTTGTGTCCAAGCAGGA	CTATCGGCGCTTCTAC
Procollagen type II (COL2A1)	GAACCCAGAACAACAATCC	GTTTCGACTTTTCTCCCCTCT
Procollagen type X (COL10A1)	CTGAGCGATACCAAACACC	GTAAAGGTGTATCACTGAGAGG
Cyclin B2	GTTGACTATGACATGGTG	GTTTCGTGCACTTTGTCTTG
Procollagen type VI (COL6A3)	CATGCTTTGATTTACACTCG	CACTGCTGGTGTATATGTG
Fibromodulin	CAAGGCAATAGGATCAATG	GTTTGGCTTATGGAAGGTC
Decorin	TGAGTTTCAACAGCATCTCTGC	GTGAGCCTTTTCAGCAACC
Aggrecan	TGAGGAGGGCTGGAACAAGTACC	TGTTCCCTGCAATTACCACCTCC
Hyaluronan synthase 2 (HAS-2)	CTCATCAATAAGTGTGGCAG	CCTATATACCTCACTAGCC
Cathepsin K	GTGTGTCTGAGAATGATGGCTG	CAGCAAAGGTGTGTATTATG
Aggrecanase-1 (ADAMTS-4)	ACCACTTTGACACAGCCATTC	TGCTCTCGGACCTGTGGGGGT
Aggrecanase-2 (ADAMTS-5)	TGTGCTGTGATTGAAGACGAT	CCATCTACCGCTCCTGCAC
Caspase 3	CTGGTACAGATGTCGATGCAG	CATTGAGACAGACAGTGGTG
Vascular endothelial growth factor (VEGF)	GTTTCATGGATGTCTATCAG	GCACAACAAATGTGAATGCAG
Plasminogen activator inhibitor (PAI-1)	GATCCAAGAGGCAATGCAATTC	GATCAGCGACTTACTTGGTG
Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	GCTCTCCAGAACATCATCCCTGCC	AGCTCATTTCCTGGTATGACAACG

hibitors *TIMP-1* (strongly) and *TIMP-2* (weakly) and *TIMP-3* (strongly but only in section K) were detected. In contrast there was no expression of the gelatinases *MMP-2*, *MMP-9* nor of the aggrecanases *ADAMTS-4* and *-5* in sections L and K. *Caspase 3* was expressed only in section K where *MMP-3* and *TIMP-3* expression was strong.

Proliferative zone

The upregulation of *cyclin B2* in section J indicates the beginning of chondrocyte proliferation in the growth plate. This is associated with the downregulation of expression of all matrix proteins tested, namely *COL2A1*, *aggrecan*, *HAS-2*, *fibromodulin* and *decorin*, *osteocalcin* as well as *TIMPs* and proteinases previously upregulated. No expression of *COL6A3* and *caspase 3* was detected in this section and expression did not reappear until the lower proliferative zone in section E and D respectively. The expressions of *aggrecan*, *HAS-2*, *osteocalcin*, *TIMPs*, *MMP-13* and *MMP-14*, *ADAMTS-4* and *-5* were absent or markedly reduced in the central proliferative zone.

In section I *cyclin B2* expression was downregulated although its expression level was up again in section H dropping until section D when it rose again. Expression of *COL2A1* was maintained until section D when it rose with that of *COL6A3*, *fibromodulin*, *decorin* and *cathepsin K*, *MMP-13*, *MMP-3* and *caspase 3*. Chondrocytes in section E also expressed *COL6A3*, *fibromodulin*, *ADAMTS-4* and *MMP-9*, which also rise in section F. Expression of *TIMPs-1* and *-2* started to rise again, where *MMP-13* and *MMP-14* were weakly coexpressed with *MMP-3* in this region.

Caspase 3 expression in section D immediately preceded that of the hypertrophic chondrocytes phenotype identified by expression of *COL10A1* in section C. This is where *cathepsin K* expression reached a peak with *MMP-3* and *fibromodulin*, *COL2A1* and *COL6A3*. Here *cyclin B2* was again elevated.

Hypertrophic zone

The onset of hypertrophy in section C was accompanied by the strongest but transient expression of *VEGF* as well as *aggrecan* and *HAS-2*. *COL2A1* expression was maintained but that of *COL6A3* was not detected in this zone. In section C *MMP-13* continued to rise with *MMP-14*, *MMP-3*, *MMP-2* (latter only expressed here) and *MMP-9*. *ADAMTS-4* and *ADAMTS-5* were strongly expressed with *cathepsin K*. The inhibitors *TIMP-3* and *PAI-1* were markedly increased here.

In section B the proteinases continued to be expressed (except *MMP-2* which was only expressed in section C) but less so in the case of the aggrecanases. The inhibitors were also expressed but *PAI-1* expression was reduced.

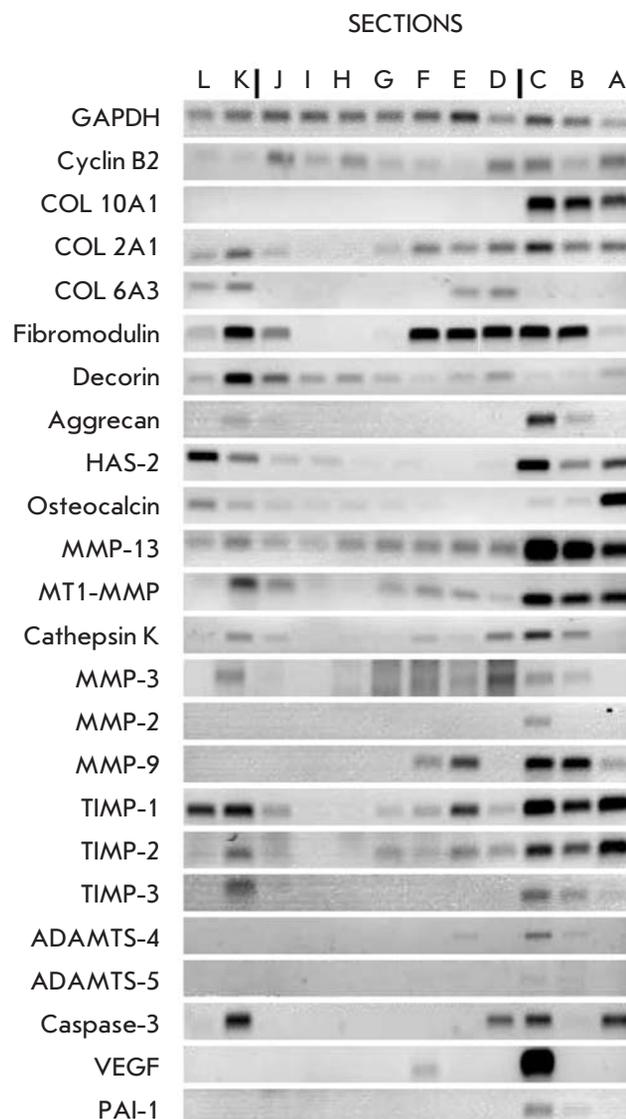


Fig. 2. A representative RT-PCR analyses of gene expression in the zones of the growth plate

The late hypertrophic zone of bovine growth plate is represented by section A. The gene expression analysis of this section revealed the highest expression of *cyclin B2*, *caspase 3*, *COL10A1*, *COL2A1*, *HAS-2* and *osteocalcin*. Collagenases *MMP-13* and *MMP-14* were also maximally expressed but *MMP-2*, *MMP-3*, *cathepsin K* and *ADAMTS-4* and *-5* were not expressed. *MMP-9* was weakly expressed. *Aggrecan* and *fibromodulin* expression was absent and weak, respectively. *TIMPs -1* and *-2* were further elevated, *TIMP-3* unchanged and *PAI-1* was absent.

The microanalytical methodology, involving RT-PCR analyses of sequential transverse sections of the primary proximal tibial growth plate, used in this and

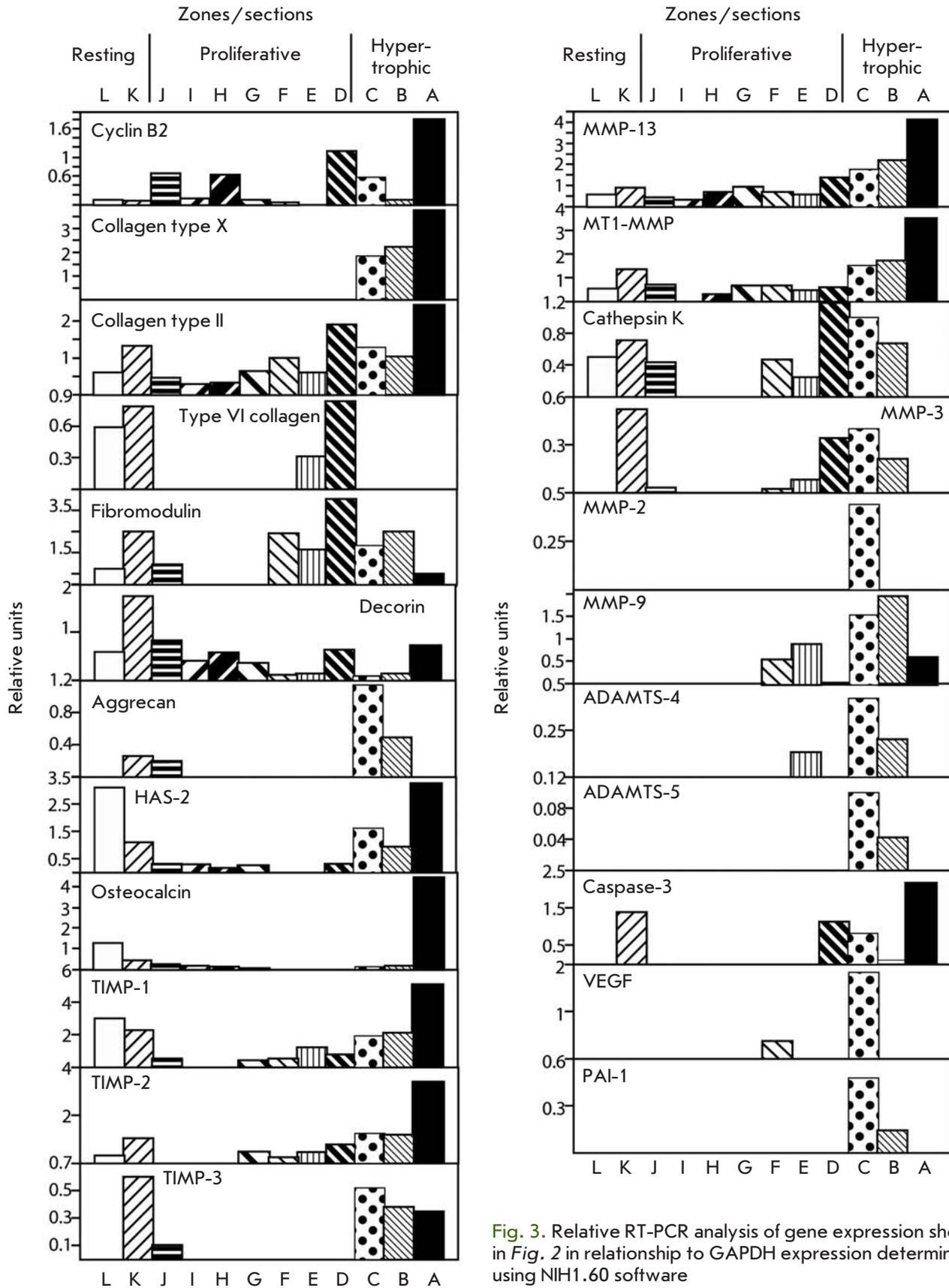


Fig. 3. Relative RT-PCR analysis of gene expression shown in Fig. 2 in relationship to GAPDH expression determined using NIH1.60 software

our previous study [17] permits expression analyses of the interrelationships of genes that have usually been studied individually concerning their involvement in chondrocyte differentiation, matrix assembly and remodeling in the growth plate.

From our present studies, combined with our earlier analyses, we can observe that each of the principal zones of the growth plate is characterized by a distinct signature profile of gene expression. Thus the resting zone (sections L and K) is characterized by the expression of matrix molecules that include the collagen fibrillar network of *COL2A1* and the associated proteoglycans *fibromodulin* and *decorin*, the filamentous collagenous network of *COL6A3*, the aggrecan network with *HAS-2* representing the synthesis of hyaluronan, a key component of aggregating proteoglycans. There is even a low level of expression of *osteocalcin* better known as a protein expressed by osteoblasts and terminally hypertrophic chondrocytes [19]. This matrix gene expression is associated with expression of *caspase 3* for reasons that are unclear. Moreover, a low expression level of the proteinases *MMP-3*, *-13*, *-14* is also seen, accompanied by expression of all three *TIMPs*. Although *MMPs* are regulated both at gene expression and protein level, the correspondent local increase of collagenase dependent collagen cleavage activity at (next to) this area of the growth plate has been also observed by us earlier [8]. This expression of matrix degradation genes is associated with expression of *caspase 3*, indicating the cell apoptosis which accompanies chondrocyte proliferation in animal growth plate [20]. However at this time matrix assembly dominates but is accompanied by limited matrix remodeling as was suggested by our previous direct analyses of matrix collagen and proteoglycan in this growth plate [8]. The increased expression of *cathepsin K* in this zone raises questions as to whether this is related to either extracellular and /or intracellular activity of this proteinase. At this stage it is worthy of note that *ADAMTS-4* and *-5* are not expressed until the hypertrophic zone and that evidence for their involvement in aggrecan degradation is not seen until hypertrophy is observed. Upregulation of matrix remodeling genes in the area adjacent to the beginning of chondrocyte proliferative activity is associated with the strong upregulation of proliferative zone related growth factors, namely *FGF-2*, *TGF β 2* and *PTHrP* [17] indicating their involvement in the regulation of matrix turnover.

The upper proliferating growth plate chondrocytes, delineated by the increased expression of *cyclin B2*, which is first observed in section J and then H. These early proliferative chondrocytes did not show any significant changes in relative expression of genes involved in matrix remodeling.

The downregulation of *cyclin B2* expression in section E, preceding hypertrophy and is associated with another expression maximum of matrix remodeling. In contrast to resting zone, at this time the upregulation of matrix proteins *COL2A1*, *COL6A3* and *fibromodulin* is not accompanied by significant increase in *decorin*, *aggrecan* and *HAS-2* expression. However, as in the resting zone, expression of matrix degrading genes *MMPs* and *ADAMTS-4* and their inhibitors *TIMPs* was detected. *Fibromodulin* has been shown to be strongly expressed only in the proliferative zone in the rat and mouse growth plates [6, 21]. In contrast previous studies using sequential transverse sections (200-400mm) of bovine growth plate revealed the presence of *fibromodulin* message in all the zones except the lower hypertrophic [22]. Our analyses reveal a 300mm region of the proliferating zone lacking significant expression of this protein. This study has also revealed that the distribution of *fibromodulin* expression in bovine growth plate is similar to that of *type II collagen* as was seen in mouse growth plate [23].

The gene expression of another collagen binding proteoglycan *decorin* progressively decreases in the proliferative zone confirming earlier data [21, 22]. This may be related to the ability of *decorin* to inhibit bone mineralization [23] which we know starts in section H in proliferating bovine growth plate chondrocytes [8]. *Decorin* expression is clearly greatest in the resting zone and decreases prior to hypertrophy.

In general the gene expression pattern in section D immediately preceding the hypertrophic zone is similar to that immediately preceding the onset of proliferation (section K) characterized by the expression of *COL2A1*, *COL6A3*, *fibromodulin*, *decorin*, *MMP-13*, *MMP-14*, *cathepsin K*, *MMP-3*, *TIMPs -1*, *-2* and *caspase 3*. However, at this time there is no *aggrecan*, little *HAS-2* expression and *TIMP-3* expression is also lacking. In spite of the similarity in gene expression pattern in sections K and D including that of *caspase 3*, the further fate of both groups of growth plate chondrocytes is not the same. Instead of chondrocyte progression to proliferation it is now to hypertrophy and is accompanied by the expression of different regulatory growth factors: namely *PTHrP* and *FGF-2* at the onset of proliferation, and *TGF β 1* and *Ihh* in the hypertrophic zone [17].

Immediately prior to hypertrophy there are some clear-cut changes in expression. *COL6A3* transiently peaks again as does *fibromodulin*. *Type II collagen* expression is also upregulated at this time. As we mentioned previously [17], *COL2A1* expression was detected throughout the growth plate. But when PCR was performed using equally diluted samples, three peaks of *COL2A1* expression were observed in samples K, D, and A. The highest level of *type II collagen* expression

in the lower proliferative and upper hypertrophic zones was also observed by others [24]. *MMP-9* is upregulated for the first time as is *ADAMTS-4*, although both transiently at this stage. *Cathepsin K* and *caspase 3* both rise again. The upregulation of the expression of these two genes in the proliferative and early hypertrophic chondrocytes were also observed in rodent and human growth plates [20, 25]. Clearly these changes reflect the cessation of proliferation and the beginning of hypertrophy.

The onset of hypertrophy is characterized by the sudden expression of *COL10A1*. This another gene expression maximum is characterized by the upregulation of *COL2A1*, *fibromodulin*, *aggrecan* and *HAS-2* expression and downregulation of *COL6A3*. The active process of ECM remodeling involving type II collagen loss mediated by collagenase [8] is accompanied by the upregulation of all the collagenases, gelatinases (*MMP-2* and *MMP-9*), *MMP-3*, *TIMPs* and expression of the aggrecanases *ADAMTS-4* and *-5*.

Growth plate vascularization is associated with the early transient hypertrophic upregulation of *VEGF* and persistent upregulation of *MMP-9* expression as observed by others [26, 27]. *MMP-9* expression clearly accompanies the expression of *VEGF* which is a chemoattractant and a mitogen for endothelial cells [28]. Active blood vessels ingrowth in the hypertrophic zone of the growth plate may account for upregulation of *cyclin B2* expression also seen in section D, C and A.

The final maximum of gene expression in section A is associated with the strong upregulation of collagenases *MMP-13* and *MMP-14*, the loss of expression of *cathepsin K*, and *ADAMTS-4* and *-5* and the maintenance or an increased in expression of the MMP inhibitors *TIMP-1*, *TIMP-2* and *TIMP-3* and is accompanied by an increase in the expression of *COL2A1*, *decorin* and *HAS-2*. No expression of *aggrecan* or *type VI collagen* is detected at that time but *osteocalcin* is again expressed. The downregulation of *fibromodulin* expression seen here in the late hypertrophic zone has previously been established [21, 22].

Overall, by using the enlarged bovine physis our study provides an original insight into the interrelationships of gene expression in chondrocyte proliferation and differentiation associated with extracellular matrix assembly, mineralization, and vascularization. Our approach is the first sequential presentation of various genes in one study that permits an analysis of individual gene expression changes associated both with respect to their alterations in the continuum of chondrocyte differentiation ending in cell death through the growth plate. It also allows for a comparison of the expression of various genes in each individual 100µm zone of the bovine physis.

In this respect upregulation of a gene in a distinct zone of the growth plate indicates its involvement in the processes associated with exact phase of chondrocyte differentiation. In contrast, downregulation of a gene indicates that its function is less important in that zone of the growth plate. In view of this the previously observed biphasic character of *MMP-13* expression in rodent growth plate [29] was supplemented by our original observation that that is not a case for *MMP-9* and *-2*, expressions of which were associated only with pre-hypertrophic and hypertrophic phases of chondrocyte differentiation. This further indicates the importance of collagenases *MMP-13*, *MT1-MMP*, *MMP-3*, and *cathepsin K* in extracellular matrix remodeling associated with further synthesis of chondrocyte-specific matrix supported by upregulation of extracellular matrix-related molecule expression here and in the following proliferative zone of the growth plate. In contrast, upregulation of *MMP-9*, *-2*, and both *aggrecanases* associated only with chondrocyte hypertrophy indicates their destructive activity in respect to chondrocyte-specific matrix. Moreover, the observed differences in matrix degrading molecule expression might be related also to differences in regulation of their expression as we previously reported [17] and differential growth factor profiles associated with early proliferative and hypertrophic zones in the bovine growth plate.

It is worth noting that early upregulation of genes involved in mineralization in the midst of proliferative zone in bovine growth plate observed in our previous studies [8] is also associated with upregulation of the genes related to extracellular matrix-related molecule expression, their inhibitors and vascularization markers: overt mineralization occurs later in the hypertrophic zone. This suggests that any alteration in chondrocyte metabolic activity is associated with specific extracellular matrix remodeling, which affects its properties and subsequent bone formation.

Therefore, our results indicating fluctuations in gene expression for extracellular matrix molecules, proteinases and their inhibitors in the bovine growth plate were expected. However, the exact profile of each gene pattern could not be predicted with accuracy prior to completion of this study.

CONCLUSIONS

The data presented here further define the complex changes and interrelationships in gene expression in the physis of the growth plate that occur in the course of chondrocyte maturation associated with matrix assembly, remodeling, cell proliferation, differentiation, vascular invasion and cell death. This investigation draws attention to distinct phases of expression of matrix molecules, proteinases and their inhibitors

and their relationships to the physiological events and regulatory molecules that are part of endochondral ossification.

This study was funded by Shriners Hospitals for Children and Canadian Institutes of Health (to A.R. Poole).

REFERENCES

1. Mackie E.J., Tatarczuch L., Mirams M. // *J. Endocrinol.* 2011. V. 211. P.109-121.
2. Alini M., Matsui Y., Dodge G.R., Poole A.R. // *Calcified Tissue International.* 1992. V. 50. № 4. P. 327-335.
3. Poole A.R. // *Cartilage in Health and Disease* / Ed. Koopman W. *Arthritis and Allied Conditions.* 15th ed. Chapter 11. Philadelphia, PA, USA: Lippincott, Williams and Wilkins, 2005. P. 223-269.
4. Svensson L., Oldberg A., Heinegard D. // *Osteoarthritis&-Cartilage.* 2001. V. 9. Suppl A. S23-S28.
5. Damron T.A., Zhang M., Pritchard M.R., Middleton F.A., Horton J.A., Margulies B.M., Strauss J.A., Farnum C.E., Spadaro J.A. // *Int. J. Radiat. Oncol. Biol. Phys.* 2009. V. 74. P. 949-956.
6. Simsa S., Hasdai A., Dan H., Ornan E.M. // *Am. J. Physiol. Regul. Integr. Comp. Physiol.* 2007. V. 292. P. R2216-2224.
7. Armstrong A.L., Barrach H.J., Ehrlich M.G. // *J. Orthop. Res.* 2002. V. 20. № 2. P. 289-294.
8. Mwale F., Tchetina E., Wu W., Poole A.R. // *J. Bone Miner. Res.* V. 17. № 2. P. 275-283.
9. Wu W., Tchetina E., Mwale F., Hasty K., Pidoux I., Reiner A., Chen J., van Wart H.E., Poole A.R. // *J. Bone Miner. Res.* 2002. V. 17. P. 639-651.
10. Mitani H., Takahashi I., Onodera K., Bae J.W., Sato T., Takahashi N., Sasano Y., Igarashi K., Mitani H. // *Histochem. Cell. Biol.* 2006. V. 126. P. 371-380.
11. Kielty C.M., Lees M., Shuttleworth C.A., Woolley D. // *Biochem. Biophys. Res. Commun.* 1993. V. 191. № 3. P. 1230-1236.
12. Myint E., Brown D.J., Ljubimov A.V., Kyaw M., Kenney M.C. // *Cornea.* 1996. V. 15. № 5. P. 490-496.
13. Joronen K., Salminen H., Glumoff V., Savontaus M., Vuorio E. // *Histochem. // Cell Biol.* 2000. V. 114. № 2. P. 157-165.
14. Kahari V.M., Saarialho-Kere U. // 1999. *Ann. Med.* V. 31. № 1. P. 34-45.
15. Bertaux B., Hornebeck W., Eisen A.Z., Dubertret L. // 1991. *J. Invest. Dermatol.* V. 97. № 4. P. 679-685.
16. Wick M., Burger C., Brusselbach S., Lucibello F.C., Muller R. // 1994. *J. Biol. Chem.* V. 269. № 29. P. 18953-18960.
17. Tchetina E., Mwale F., Poole A.R. // 2003. *J. Bone Miner. Res.* V. 5. № 5. P. 844-851.
18. Pal S., Tang L.H., Choi H., Haberman L.C., Ruoghley P.J., Poole A.R. // 1981. *Collagen Relat. Res.* V. 1 P. 151-176.
19. Pullig O., Weseloh G., Ronneberger D.-L., Kakonen S.M., Swoboda B. // 2000. *Calcif. Tissue Int.* V. 67. № 3. P. 230-240.
20. Chrysis D., Nilsson O., Ritzen E.M., Savendahl L. // 2002. *Endocrine.* V. 18. № 3. P. 271-278.
21. Wang Y., Middleton F., Horton J.A., Reichel L., Farnum C.E., Damron T.A. // 2004. *Bone.* V. 35. № 6. P. 1273-1293.
22. Alini M., Roughley P. // 2001. *Matrix Biol.* V. 19. № 8. P. 805-813.
23. Hoshi K., Kemmotsu S., Takeuchi Y., Amizuka N., Ozawa H. // 1999. *J. Bone Miner. Res.* V. 14. № 2. P. 273-280.
24. O'Keefe R.J., Puzas J.E., Loveys L., Hicks D.G., Rosier R.N. // 1994. *J. Bone Miner. Res.* V. 9, № 11. P. 1713-1722.
25. Soderstrom M., Salminen H., Glumoff V., Kirschke H., Aro H., Vuorio E. // 1999. *Biochim. Biophys. Acta.* V.1446. № 1-2. P.35-46.
26. Garcia-Ramirez M., Toran N., Andaluz P., Carrascosa A., Audi L. // 2000. *J. Bone Miner. Res.* V. 15. P. 534-540.
27. Takahara M., Naruse T., Takagi M., Orui H., Ogino T. // 2004. *J. Orthop. Res.* V. 22. № 5. P. 1050-1057.
28. Yang Y.Q., Tan Y.Y., Wong R., Wenden A., Zhang L.K., Rabie A.B. // 2012. *Int. J. Oral. Sci.* V. 4. P. 64-68.
29. Alvarez J., Balbin M., Santos F., Fernandez M., Ferrando M., Lopez J.M. // 2000. *J. Bone Miner. Res.* V. 15. P. 82-94.

Effects of Neonatal Fluvoxamine Administration on the Physical Development and Activity of the Serotonergic System in White Rats

N. Yu. Glazova¹, S. A. Merchieva², M. A. Volodina², E. A. Sebentsova¹, D. M. Manchenko², V. S. Kudrun³, N. G. Levitskaya^{1*}

¹ Institute of Molecular Genetics, Russian Academy of Sciences, Kurchatov Sq., 2, 123182 Moscow, Russia

² Biological Faculty, Lomonosov State University, Moscow, Leninskie Gory, 1/12, Russia

³ Institute of Pharmacology, Baltyiskaya str, 8, 125315, Moscow, Russia

E-mail: nglevitskaya@gmail.com

Received 16.06.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT Selective serotonin reuptake inhibitors (SSRIs), including fluvoxamine, are widely used to treat depressive disorders in pregnant women. These antidepressants effectively penetrate through the placental barrier, affecting the fetus during the critical phase of neurodevelopment. Some clinical studies have linked prenatal exposure to SSRIs with increased neonatal mortality, premature birth, decreased fetal growth and delay in psychomotor development. However, the effects of prenatal exposure to SSRIs remain unknown. The administration of SSRIs in rodents during the first postnatal weeks is considered as an model for studying the effects of prenatal SSRI exposure in human. The aim of this work was to study the acute effects of chronic fluvoxamine (FA) administration in white rat pups. The study was carried out in male and female rat pups treated with FA (10 mg/kg/day, intraperitoneally) from postnatal days 1 to 14. The lethality level, body weight, age of eye opening, and motor reflex maturation were recorded. The contents of biogenic amines and their metabolites in different brain structures were also determined. It was shown that neonatal FA administration led to increased lethality level, reduced body weight, and delayed maturation of motor reflexes. Furthermore, increased noradrenalin level in hypothalamus, serotonin level in hippocampus and serotonin metabolite 5-HIAA level in frontal cortex, hypothalamus, hippocampus, and striatum were observed in drug-treated animals compared to the control group. We can conclude that the altered activity of the serotonergic system induced by fluvoxamine administration at early developmental stages leads to a delay in physical and motor development.

KEYWORDS biogenic amines; fluvoxamine; neonatal administration; psychomotor development; selective serotonin reuptake inhibitors.

ABBREVIATIONS DA – dopamine; DOPAC – 3,4-dihydroxyphenylacetic acid; FA – fluvoxamine; 5-HIAA – 5-hydroxyindoleacetic acid; 5-HT – serotonin (5-hydroxytryptamine); HVA – homovanillic acid, IC – intact control, NA – noradrenalin, PND – postnatal day, SERT – serotonin transporter, SSRI – selective serotonin reuptake inhibitors.

INTRODUCTION

Depression is a common mental disease that occurs in more than 10% of the population. Women are much more susceptible to depression than men. Depressive symptoms are registered in 14–23% of women during the gestation period [1]. Selective serotonin reuptake inhibitors (SSRIs) have recently become the first-choice medications for depressive disorders, and their use is constantly on the rise. This group of drugs includes fluoxetine, citalopram, fluvoxamine (FA), paroxetine, sertraline, and others. All SSRIs have a similar

mechanism of action despite differences in their chemical structures [2]. The target of the action of SSRIs is the serotonin transporter (SERT) that is responsible for mediator reuptake from the synaptic cleft. SERT blockade increases serotonergic neurotransmission. SSRI antidepressants are used to treat depressive disorders in pregnant and nursing women. According to various sources, 6–13% of women in their gestation period are currently taking SSRIs [3, 4]. Moreover, duration of intake and daily doses of the prescribed drugs increase [5]. SSRIs effectively penetrate through the

placental barrier and are present in the amniotic fluid, umbilical blood, and fetal plasma [6]. The content of various antidepressants of this group in umbilical blood constitutes from 70 to 86% of the respective content in a mother's blood; thus, the fetus is exposed to physiologically active SSRI doses [7]. However, the consequences of SSRI influence on a developing organism have been poorly studied so far. The results of clinical studies are extremely contradictory. Some works report no effect of drug exposure on the pregnancy complications and conditions of newborns [8–10]. Other data indicate negative effects from SSRI exposure on pregnancy outcomes: an increase in the number of spontaneous miscarriages and neonatal mortality, higher risk of premature birth, and a decrease in the birth weight were observed [1, 3, 11]. Symptoms of poor neonatal adaptation occurred in 15–30% of newborns prenatally exposed to SSRIs (neonatal withdrawal syndrome). Respiratory impairment, hypoglycemia, unstable body temperature, sleep disturbances, hyperexcitability, and convulsions are observed in infants during the first days of life. The indicated symptoms disappear during 1–2 weeks [1, 7]. Moreover, SSRI exposure during pregnancy (especially the last trimester) leads to a lower Apgar score of the newborns, delayed psychomotor development, sleep disturbances, persistent pulmonary hypertension, and cardiovascular disorders [2, 3, 6, 12, 13]. All the listed effects are observed in the early neonatal period (from birth to 6 months). Information about the delayed effects of prenatal SSRI exposure is limited due to time-consuming nature and complexity of such studies [13, 14]. As it has been previously mentioned, the results of clinical studies of prenatal SSRI exposure effects are rather contradictory. The reason for that may be the large heterogeneity of the pregnant women sampling used for the study. Women with different severity of depression, taking different drugs of the SSRI group in different doses and at different pregnancy stages, could be included in one group [11].

Effects of exposure to SSRI drugs on the developing brain are being actively studied in experiments on animals, mostly rodents. The third trimester of pregnancy is the period of human CNS development most sensitive to the action of SSRI [12]. It is very difficult to compare the development of human and rodent brains in the right way; however, data on CNS maturation (including the serotonergic system) allow one to compare the last trimester of human pregnancy with the first weeks of life of rats [15, 16]. Therefore, SSRI effects during the first weeks of a rat's life can be regarded as a model for studying the prenatal effects of this group of drugs during the third trimester of human pregnancy [17]. It was experimentally shown that chronic administration of SSRIs in the neonatal period

causes long-term alterations in animal behavior. Adult rats and mice administered SSRIs during the first weeks of life displayed increased anxiety and depression, abnormalities in eating behavior, and alterations in the activity of the serotonergic system [17, 18].

Thus, clinical studies of the effects of prenatal SSRI exposure are mostly focused on neonatal abnormalities, with data on the delayed effects of the exposure being limited. On the contrary, animal experiments are mostly aimed at estimating the long-term effects of perinatal SSRI administration [15]; studies of the neonatal effects are scarce [4, 19, 20]. However, studies of the acute effects of neonatal SSRI administration in animals are required to prove the adequacy of the experimental models in use.

Fluvoxamine is a modern antidepressant belonging to the SSRI group. Fluvoxamine is similar to fluoxetine in its pharmacological properties, but it is highly effective and selective [21] and possesses anxiolytic activity. The effects of neonatal FA administration have not been studied so far. In the present work, we have studied the influence of chronic neonatal fluvoxamine administration on the physical development and the state of the serotonergic system in white rat pups.

EXPERIMENTAL

Experiments were performed using both male and female pups of outbred white rats. The animals were kept under standard vivarium conditions with free access to food and water and a 12-h light regimen. The pups' date of birth was considered postnatal day zero (PND 0). Two series of experiments were carried out.

The first series included 10 litters; the pups in each litter were divided into 3 groups: intact control ("IC"), control ("CON") and fluvoxamine ("FA"). The use of the "IC" group was necessary to estimate the influence of everyday experimental manipulations on the observed parameters. No differences were found between the "IC" and "CON" groups in the first series of experiments: therefore, for 10 litters used in the second series of experiments each litter was divided only into 2 groups ("CON" and "FA") in order to reduce the number of animals used. Rats from the "IC" group were subjected to everyday handling without drug administration from PND 1 to PND 14. Animals from the "CON" group were administered water for injections at 2 ml per kg of body weight by intraperitoneal (IP) injection, daily from PND 1 to PND 14. Rats from the "FA" group received IP injections of fluvoxamine (fluvoxamine maleate, Sigma) at 10 mg/kg of weight, daily from PND 1 to PND 14.

The age of eye opening and body weight of animals were registered in order to estimate the physical development of the pups. The level of psychomotor develop-

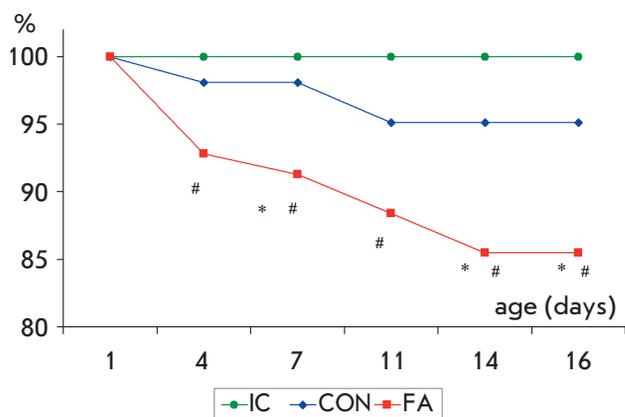


Fig. 1. Effect of neonatal fluvoxamine administration on the lethality level in rats. X-axis – the age of rats, Y-axis – the number of survived animals as a percentage of the initial number of rats in the group (“IC” n=34, “CON” n=90, “FA” n=88). * – significant difference from the control, # – significant difference from “IC” group (difference between two proportions) ($p < 0.05$)

ment was assessed in “righting reflex”, “gait reflex”, and “negative geotaxis reflex” tests. “Righting reflex”: a 6-day-old rat pup was placed in the supine position, noting the time needed for the animal to turn over (4 paws on the ground). “Gait reflex”: a 10-day-old rat pup was placed in the center of a circle 13 cm in diameter, noting the time needed for the animal to crawl out of the circle. “Negative geotaxis reflex”: a 12-day-old rat pup was placed on a 30-cm-long inclined surface (45°), with its head oriented towards the slope, noting the time needed for the animal to turn around, 180° [19, 20].

In order to study the effects of neonatal FA administration on the content of biogenic amines and their metabolites in a rat’s brain, some animals were decapitated at the age of PND 16 (48 h after the last injection). Brains were extracted with the following structures separated: frontal cortex, hypothalamus, hippocampus, and striatum. The specimens were rapidly frozen in liquid nitrogen and further stored at -70°C . Brain tissues were homogenized. High-performance liquid chromatography was used to determine the concentrations of biogenic amines and their metabolites, noradrenalin (NA), serotonin (5-HT), 5-hydroxyindoleacetic acid (5-HIAA), dopamine (DA), homovanillic acid (HVA), and 3,4-dihydroxyphenylacetic acid (DOPAC).

Statistical data analysis

The results were analyzed using the Statistica software package. The lethality levels in the groups were compared using the “difference between two proportions” test. Body weight data were analyzed by a

two-way ANOVA for repeated measurements, with gender and group as between subject factors. The two-way ANOVA (gender \times group) was used to compare the age of eye opening and psychomotor development of pups; post hoc testing was carried out by the LSD test. The content of biogenic amines in the brain was analyzed by a two-way ANOVA (gender \times group or litter \times group). The group means for normalized values of the contents of biogenic amines were compared using the Mann-Whitney test. The data in the figures are presented as the means \pm standard error of the means. Differences were considered to be statistically significant with $p < 0.05$.

RESULTS

The experiments were performed using both male and female animals. Application of the two-way ANOVA (factor 1 – group; factor 2 – gender) revealed no significant effect of the gender or interaction between the two factors in all the performed tests, allowing us to present the results obtained for the whole group of rats.

The lethality level in animal groups was estimated during the experiment. 100% of rats survived in the “IC” group by PND 16; 95.1%, in the “CON” group; and 85.5%, in the “FA” group (Fig. 1). Daily intraperitoneal injections of the solvent caused an increased lethality in the control group of rats; however, no statistically significant differences were observed compared to the intact control ($p > 0.20$). Chronic neonatal FA administration led to a significant increase in the lethality level compared to the control ($p < 0.03$).

Analysis of the eye opening age in pups with neonatal fluvoxamine administration indicated no significant effect of gender ($F_{1,193} = 2.73$, $p > 0.10$) and statistically significant effect of group ($F_{2,193} = 3.57$, $p < 0.03$). Post hoc analysis showed the presence of a slight but statistically significant decrease in the eye opening age in the “FA” group compared to the “IC” and “CON” groups (Fig. 2). 86.3% of the animals in the group of rats that received FA opened their eyes by the 16th day of life, as compared to 70.7% in the control group and 67.7% in the “IC” group ($p < 0.03$).

The measurement of the body weight revealed statistically significant differences between newborn pups in the control groups in the first and the second experimental series (6.14 ± 0.13 and 6.50 ± 0.08 g, $p < 0.01$), while no initial differences between the “FA” and “IC” groups and the respective control groups were detected. Figure 3 shows the change in the rats’ body weight in the “FA” and “IC” groups as compared to the respective control groups. In all experimental groups, the body weight increased from PND 1 to PND 16 ($F_{15,2325} = 4058.8$; $p < 0.001$ and $F_{15,930} = 1557$; $p < 0.001$, Fig. 3A and 3B, respectively). The effect of the gender

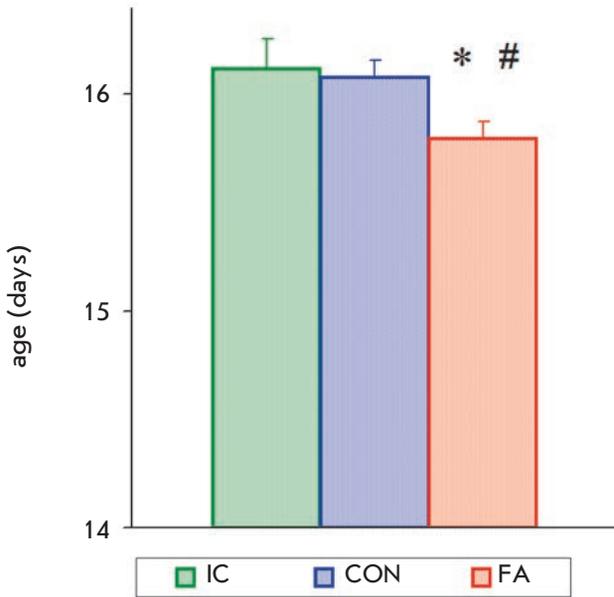


Fig. 2. Effect of neonatal fluvoxamine administration on the age of eye opening ("IC" n=33, "CON" n=86, "FA" n=75). * – significant difference from the control, # – significant difference from the "IC" group (LSD test) (p<0.05)

on the body weight was not statistically significant in the first and the second experimental series ($F_{1,62} = 0.70$; $p > 0.40$ and $F_{1,155} = 0.10$; $p > 0.80$, respectively). Comparison of the "IC" and "CON" groups revealed no significant effect of the group on the body weight change in rats ($F_{1,62} = 0.01$; $p < 0.98$). In the case of "CON" and "FA" groups, a significant effect of the group on the weight gain was revealed ($F_{1,155} = 4.1$; $p < 0.04$). Therefore, daily intraperitoneal injections of the solvent did not affect the weight gain, while FA administration decelerated the growth of the animals.

No statistically significant effects of gender on the development of motor reflexes in the rats were observed ($F_{1,68} = 0.17$; $p > 0.65$ in the "righting reflex" test and $F_{1,35} < 0.10$; $p > 0.80$ in the "negative geotaxis reflex" and "gait reflex" tests). At the same time, the group significantly affected the response time in the "righting reflex" ($F_{2,68} = 4.37$; $p < 0.04$) and "negative geotaxis reflex" ($F_{2,35} = 4.38$; $p < 0.05$) tests, while there was no effect on rat behavior in the "gait reflex" test ($F_{2,35} = 0.67$; $p = 0.52$). Post hoc analysis revealed no significant differences between the "IC" and "CON" groups. The values of the registered parameters were statistically significantly higher in the "FA" group

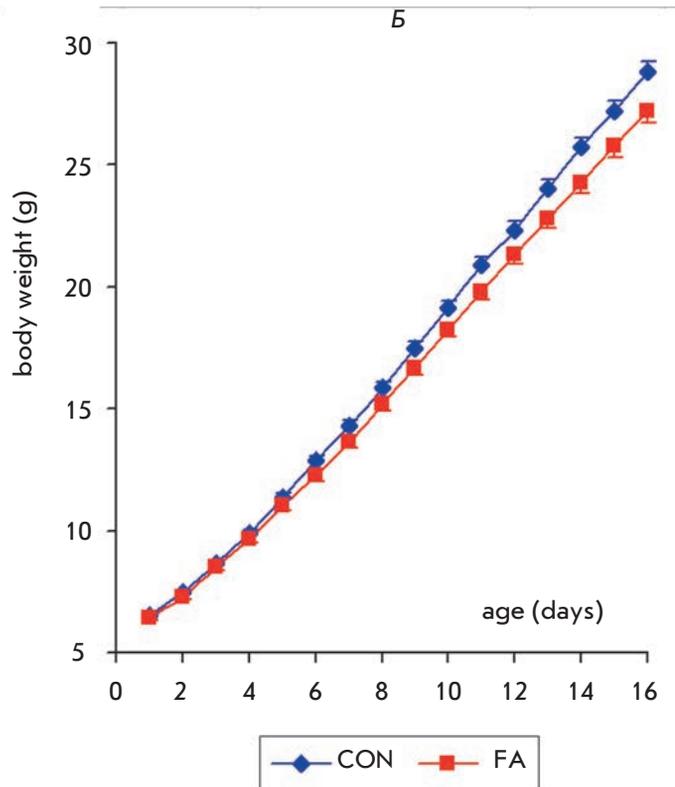
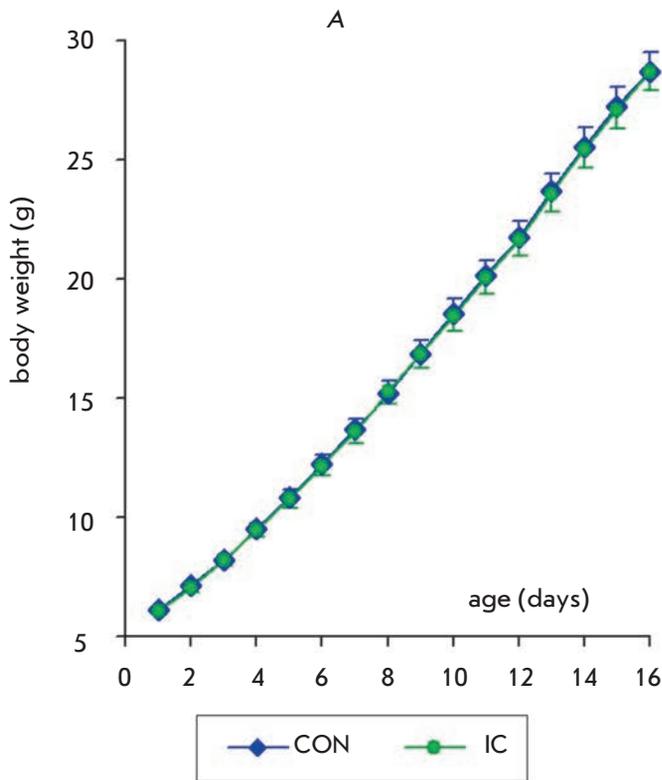


Fig. 3. Effect of neonatal fluvoxamine administration on the weight gain in rats. X-axis – the age of rats, Y-axis – body weight (a – "IC" n=33, "CON" n=33, b – "CON" n=86, "FA" n=75)

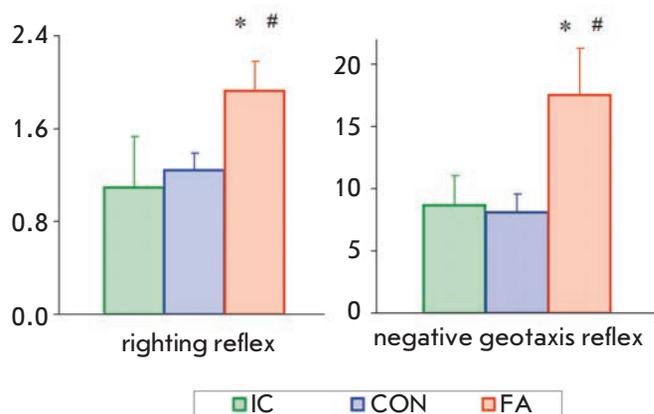


Fig. 4. Effect of neonatal fluvoxamine administration on the latency time in righting reflex (PND 6; "IC" n=16, "CON" n=30, "FA" n=28) and in negative geotaxis reflex (PND 10; "IC" n=13, "CON" n=14, "FA" n=14). * – significant difference from the control, # – significant difference from "IC" group (LSD test) ($p < 0.05$)

compared to the control and "IC" groups (Fig. 4).

The contents of biogenic amines and their metabolites in a rat's brain were measured on PND 16. The results are shown in Table. We observed no significant influence of gender on the registered parameters ($F < 3.0$; $p > 0.10$). Animals from five litters (3–4 rats from each group in every litter) were used for measurements. Application of the two-way ANOVA (factor

1 – group; factor 2 – litter) showed that the group significantly affected the following parameters: 5-HIAA content in hippocampus ($F_{2,38} = 4.36$; $p < 0.02$), frontal cortex ($F_{2,36} = 3.55$; $p < 0.04$) and striatum ($F_{2,40} = 12.13$; $p < 0.001$), as well as NA, 5-HT and 5-HIAA contents in hypothalamus ($F_{2,40} > 4.5$; $p < 0.02$). No significant influence of the group on the levels of DA and its metabolites was registered in any studied structure ($F < 2.60$; $p > 0.10$). Moreover, a significant influence of the litter on most parameters was noted ($F > 2.95$; $p < 0.05$), pointing to the variability of parameters in different litters. No significant interaction between the factors group and litter was observed ($F < 1.50$; $p > 0.20$). The values of the parameters for each litter were normalized to the respective controls to eliminate the influence of the litter. Further analysis revealed no statistically significant differences in the contents of biogenic amines and their metabolites in the studied brain structures between the "IC" and "CON" groups. In the "FA" group, a significant increase in NA content was observed in hypothalamus, as well as 5-HT content in the hippocampus and 5-HIAA content in all structures compared to the control (Fig. 5). 5-HIAA/5-HT ratio in the "FA" group in all structures was significantly higher than that in the "CON" group. A statistically significant increase in 5-HT contents in hippocampus and hypothalamus, NA content in the hypothalamus and 5-HIAA contents in hippocampus, hypothalamus and striatum, along with an increase in 5-HIAA/5-HT

Table. Contents of biogenic amines and their metabolites (nmol/g tissue) in different brain structures

Biogenic amines and their metabolites	Frontal cortex				Hippocampus				Hypothalamus				Striatum			
	IC	CON	FA	F (p)	IC	CON	FA	F (p)	IC	CON	FA	F (p)	IC	CON	FA	F (p)
NA	0.40 ± 0.07	0.36 ± 0.02	0.39 ± 0.04	0.96 (0.39)	0.72 ± 0.11	0.76 ± 0.10	0.86 ± 0.12	0.56 (0.57)	3.10 ± 0.21	3.37 ± 0.10	3.63 ± 0.10	4.50 (0.02)	0.50 ± 0.14	0.67 ± 0.17	0.55 ± 0.14	0.48 (0.62)
DA	0.11 ± 0.01	0.18 ± 0.03	0.16 ± 0.03	2.26 (0.12)	0.05 ± 0.01	0.08 ± 0.02	0.05 ± 0.01	1.93 (0.16)	0.96 ± 0.07	1.21 ± 0.10	1.09 ± 0.06	2.63 (0.09)	19.93 ± 0.63	19.72 ± 0.76	18.74 ± 0.70	1.02 (0.37)
DOPAC	0.11 ± 0.04	0.07 ± 0.02	0.05 ± 0.02	1.64 (0.21)	0.17 ± 0.02	0.17 ± 0.02	0.16 ± 0.02	1.92 (0.16)	0.38 ± 0.05	0.35 ± 0.05	0.44 ± 0.04	0.74 (0.48)	3.46 ± 0.15	3.46 ± 0.16	3.31 ± 0.17	0.21 (0.81)
HVA	0.21 ± 0.04	0.19 ± 0.04	0.10 ± 0.02	2.37 (0.12)	0.54 ± 0.17	0.75 ± 0.12	0.75 ± 0.16	0.65 (0.53)	0.37 ± 0.05	0.31 ± 0.04	0.35 ± 0.04	0.60 (0.55)	2.49 ± 0.20	2.56 ± 0.13	2.68 ± 0.11	0.25 (0.78)
5-HT	1.16 ± 0.09	1.16 ± 0.05	1.19 ± 0.08	0.40 (0.67)	1.22 ± 0.06	1.23 ± 0.03	1.30 ± 0.04	1.33 (0.28)	2.47 ± 0.15	2.92 ± 0.11	3.04 ± 0.11	8.33 (0.001)	1.15 ± 0.06	1.27 ± 0.06	1.36 ± 0.09	1.09 (0.35)
5-HIAA	0.44 ± 0.03	0.41 ± 0.03	0.47 ± 0.03	3.55 (0.04)	0.95 ± 0.07	0.96 ± 0.05	1.07 ± 0.06	4.36 (0.02)	1.49 ± 0.12	1.68 ± 0.09	1.89 ± 0.10	7.78 (0.002)	1.22 ± 0.08	1.32 ± 0.09	1.49 ± 0.09	12.13 (0.001)

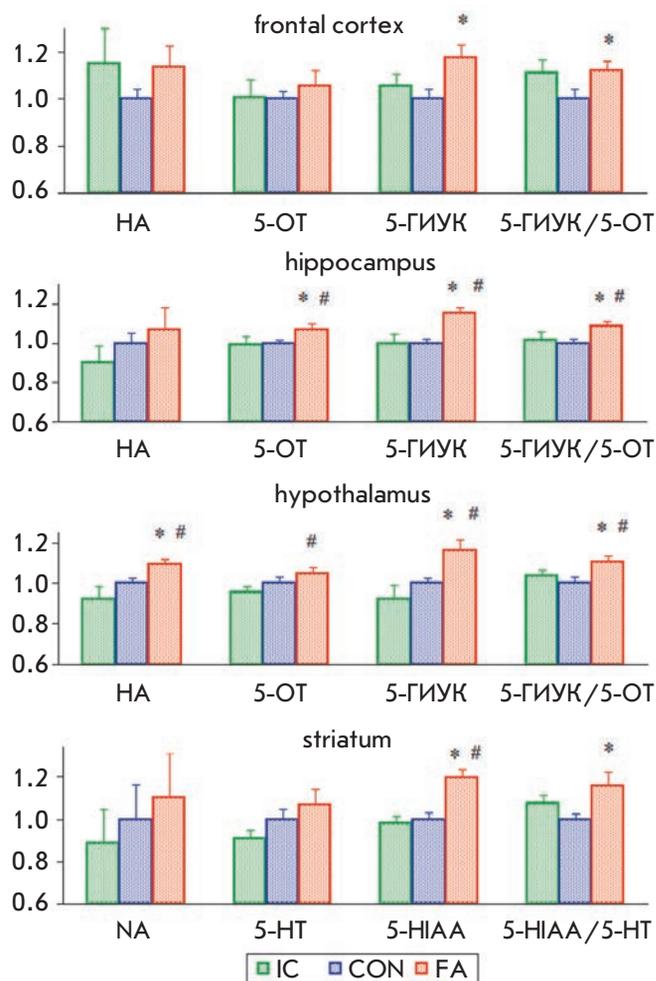


Fig. 5. Effect of neonatal fluvoxamine administration on the contents of noradrenalin, serotonin and its metabolite 5-HIAA in different brain structures. Y-axis – parameter values normalized to the control ("IC" n=18, "CON" n=19, "FA" n=18). * – significant difference from the control, # – significant difference from the "IC" group (Mann-Whitney U test) ($p < 0.05$)

ratio in the hippocampus compared to the intact control were registered in the group of rats received FA. Moreover, a tendency towards increasing 5-HIAA level in the frontal cortex and increasing 5-HIAA/5-HT ratio in the hypothalamus and striatum compared to the intact control was observed ($p < 0.10$).

DISCUSSION

Comparison of the data obtained for the intact control and the control groups allows one to conclude that daily intraperitoneal injections of the solvent during the first 14 days of life do not lead to any statistically significant increase in the lethality level and do not cause significant changes in the rate of somatic growth, time

of eye opening, and the development of motor reflexes, and therefore do not affect physical and sensomotor development of the animals. Moreover, daily injections of the solvent do not influence the state of the biogenic amine system in the brain of 16-day-old rats. Therefore, these experimental manipulations cause no significant changes in the physiological and neurochemical parameters registered in the present work.

A significant increase in the lethality level was noted in the group of rats that received daily FA injections compared to the control groups of animals. Moreover, a decelerated weight gain was observed in the "FA" group. Negative effects of neonatal administration of SSRI drugs on the change in body weight in the animals were registered in several studies. Thus, administration of citalopram [20], sertraline [19, 22], and fluoxetine [18] during the early development period in rats causes growth deceleration. The serotonergic system is known to play an important role in the regulation of appetite and food consumption. Drugs that increase the extracellular 5-HT content also display marked anorexigenic activity [23]; therefore, it cannot be excluded that the SSRI effect on the weight gain is linked to the anorexigenic effects of 5-HT. However, it was also shown that neonatal SSRI administration leads to the development of the hypermetabolic state in mice [22]. The increase in the metabolism level in animals that received FA may also be a reason behind the decelerated weight gain.

We have also shown that rat pups that received fluvoxamine injections opened their eyes earlier compared to the control animals. Monoamines are known to act as trophic factors during the active development of the nervous system. During the prenatal and early postnatal periods, serotonin serves as a signal factor in cell proliferation and differentiation in the nervous tissue; it also influences the development of the epithelial tissue [15, 24]. It was shown that neonatal administration of the serotonin precursor 5-hydroxytryptophan leads to earlier eye opening [25]. It can be assumed that the increase in the activity of the serotonergic system during this period accelerates the development of the visual analyzer. It is likely that this phenomenon explains the earlier eye opening in the group of rats receiving fluvoxamine injections.

We have registered an increase in the run-time for the righting and negative geotaxis reflexes in the group of rats receiving fluvoxamine injections, pointing to the deceleration in the maturation of motor reflexes. The delay in the development of motor reflexes was observed by Diero *et al.*, who showed a delayed maturation of reflexes in the rats that were neonatally administered sertraline [19] and citalopram [20]. Therefore, SSRI administration to rats during the early postnatal

period causes abnormalities in the development of motor functions. The changes in the content of biogenic amines caused by pharmacological or stress impacts during intensive brain development may lead to irreversible morphological and functional changes in the CNS [26]. Thus, neonatal fluoxetine administration decreases the number and size of 5-HT neurons in dorsal raphe nuclei and the quantity of 5-HT terminals in the hippocampus [27]. The increase in the serotonin content in the brain during the developmental period impairs axon myelination [28]. Neonatal SSRI impact causes morphological changes in neurons of the striatum and motor cortex: it reduces the length and branching of dendrites and the density of dendritic spines [17]. These changes may lead to the delay in the development of motor functions [28].

Our experiments showed that neonatal FA administration causes a deceleration of the somatic growth, decrease in eye opening age, and delay in the maturation of motor reflexes. The eye opening age and the changes in body weight reflect the level of animal's physical development, while dynamic tests for the execution of motor reflexes allow one to estimate the maturation of the vestibular system. The multidirectional effects of neonatal influences on physical and motor development of animals was demonstrated in several studies. Thus, neonatal stress caused by long-term maternal deprivation leads to a decrease in the eye opening age and a delay in the maturation of motor reflexes. This was accompanied by an increase in the activity of the serotonergic system in animals [29]. The influence of neonatal FA administration on the eye opening age is probably associated with its trophic function at the early stages of ontogenesis, since the acceleration of the development of nervous and epithelial cells may cause earlier maturation of the visual analyzer. The negative effect of FA on the maturation of motor reflexes can be determined by the morphological changes in CNS caused by neonatal SSRI administration. These changes impair the formation of connections between brain structures that may cause a delayed maturation of motor functions [28].

In our experiments, the levels of biogenic amines and their metabolites were measured 48 h after the last FA injection. Fluvoxamine has the shortest duration of action among all SSRI drugs; the half-life of this antidepressant is 15–17 h, while its metabolites exhibit no physiological activity [30]. Therefore, the withdrawal effects can be observed after 48 h. The experiments on adult animals showed that the content of serotonin metabolite 5-HIAA and 5-HIAA/5-HT ratio in various brain structures increased after discontinuation of

chronic SSRI administration to rats [30–32]. Depending on the duration of the drug's action, the effect develops 48–72 h after the last injection and persists for up to 2 weeks [33]. According to our data, discontinuation of FA administration to 14-day-old rats also increases 5-HIAA content and 5-HIAA/5-HT ratio in various brain structures. The 5-HIAA/5-HT ratio is an index of the serotonin turnover in the brain; an increase in this ratio indicates that the activity of the 5-HT system has increased.

According to the clinical data, an abrupt discontinuation of SSRI administration causes a withdrawal syndrome that includes the following symptoms: psychomotor agitation, anxiety, sleep disorders, vertigo, etc. The probable mechanism of this syndrome is an increase in the activity of the brain serotonergic system [30]. Impairment of neonatal adaptation was noted for 15–30% of newborns who received SSRI prenatally [1]. Most researchers also attribute these impairments to the termination of the drug's action [7, 11]. It can be assumed that the neonatal withdrawal syndrome is associated with the increased activity of the 5-HT system after termination of SSRI action, similarly to adult patients. Our data on the increase in the serotonin turnover rate in animals after the completion of the neonatal fluvoxamine administration course confirm this assumption.

Numerous clinical studies indicate that prenatal SSRI exposure (especially during the third trimester) negatively influence pregnancy outcomes and conditions of newborns. An increased number of spontaneous miscarriages and neonatal lethality, decreased birth weight, and further impairments of neonatal adaptation and delay in psychomotor development were noted [3, 11, 34]. The present work shows that chronic administration of the selective serotonin reuptake inhibitor fluvoxamine to white rat pups from the 1st to 14th days of life leads to an increase in the lethality level, deceleration of somatic growth, and delay in motor development. Moreover, increased activity of the serotonergic system in response to the discontinuation of drug administration is observed in various brain structures. Our data allow us to conclude that SSRI administration to rat pups during the first weeks of life can be considered as an adequate model for studying the prenatal effects of this drug group in humans.

This work was supported by the Program for Fundamental Studies of the Presidium of the Russian Academy of Sciences "Molecular and Cellular Biology" and the Russian Foundation for Basic Research (grant № 14-04-01913).

REFERENCES

1. Yonkers K.A., Wisner K.L., Stewart D.E., Oberlander T.F., Dell D.L., Stotland N., Ramin S., Chaudron L., Lockwood C. // *Gen. Hosp. Psychiatry*. 2009. V. 31. № 5. P. 403–413.
2. Diav-Citrin O., Ornoy A. // *Obstet. Gynecol. Int.* 2012. V. 2012. ID 698947.
3. Smith M.V., Sung A., Shah B., Mayes L., Klein D.S., Yonkers K.A. // *Early Hum. Dev.* 2013. V. 89. № 2. P. 81–86.
4. Haskell S.E., Hermann G.M., Reinking B.E., Volk K.A., Peotta V.A., Zhu V., Roghair R.D. // *Pediatr. Res.* 2013. V. 73. № 3. P. 286–293.
5. Tuccori M., Testi A., Antonioli L., Fornai M., Montagnani S., Ghisu N., Colucci R., Corona T., Blandizzi C., Del Tacca M. // *Clin. Ther.* 2009. V. 31. P. 1426–1453.
6. Casper R.C., Gilles A.A., Fleisher B.E., Baran J., Enns G., Lazzeroni L.C. // *Psychopharmacology*. 2011. V. 217. P. 211–219.
7. Rampono J., Simmer K., Ilett K.F., Hackett L.P., Doherty D.A., Elliot R., Kok C.H., Coenen A., Forman T. // *Pharmacopsychiatry*. 2009. V. 42. P. 95–100.
8. Gentile S., Galbally M. // *J. Affect. Disord.* 2011. V. 128. № 1–2. P. 1–9.
9. Altamura A.C., De Gaspari I.F., Rovera C., Colombo E.M., Mauri M.C., Fedele L. // *Hum. Psychopharmacol.* 2013. V. 28. № 1. P. 25–28.
10. Jimenez-Solem E., Andersen J.T., Petersen M., Broedbaek K., Lander A.R., Afzal S., Torp-Pedersen C., Poulsen H.E. // *Am. J. Psychiatry*. 2013. V. 170. № 3. P. 299–304.
11. Domar A.D., Moragianni V.A., Ryley D.A., Urato A.C. // *Hum. Reprod.* 2013. V. 28. № 1. P. 160–171.
12. Casper R.C., Fleisher B.E., Lee-Ancajas J.C., Gilles A.A., Gaylor E., Debattista A., Houme H. // *J. Pediatr.* 2003. V. 42. P. 402–408.
13. Olivier J.D.A., Blom T., Arentsen T., Homberg J.R. // *Progress Neuropsychopharmacol. Biol. Psychiatry*. 2011. V. 35. P. 1400–1408.
14. Harris S.S., Maciag D., Simpson K.L., Lin R.C.S., Paul I.A. // *Brain Res.* 2012. V. 1429. P. 52–60.
15. Homberg J.R., Schubert D., Gaspar P. // *Trends Pharmacol. Sci.* 2010. V. 31. № 2. P. 60–65.
16. Semple B.D., Blomgren K., Gimlin K., Ferriero D.M., Noble-Haesslein L.J. // *Prog. Neurobiol.* 2013. V. 106–107. P. 1–16.
17. Lee L.J., Lee L. J.-H. // *Developmental Neurobiol.* 2012. V. 72. № 8. P. 1122–1132.
18. Karpova N.N., Lindholm J., Pruunsild P., Timmusk T., Castrén E. // *Eur. Neuropsychopharmacology*. 2009. V. 19. P. 97–108.
19. Deiró T.C., Manhães-de-Castro R., Cabral-Filho J.E., Barreto-Medeiros J.M., Souza S.L., Marinho S.M., Castro F.M., Toscano A.E., Jesus-Deiró R.A., Barros K.M. // *Physiol. Behav.* 2006. V. 87. № 2. P. 338–344.
20. Deiró T.C., Carvalho J., Nascimento E., Medeiros J.M., Cajuhi F., Ferraz-Pereira K.N., Manhães-de-Castro R. // *Arq. Neuropsiquiatr.* 2008. V. 66. № 3-B. P. 736–740.
21. Hrdina P.D. // *J. Psychiatry Neurosci.* 1991. V. 16. № 2. P. 10–18.
22. Kummert G.J., Haskell S.E., Hermann G.M., Ni C., Volk K.A., Younes A.K., Miller A.K., Roghair R.D. // *J. Nutr. Metab.* 2012. V. 2012. ID 431574.
23. Halford J.C., Blundell J.E. // *Prog. Drug Res.* 2000. V. 54. P. 25–58.
24. Lauder J.M., Tamir H., Sadler T.W. // *Development*. 1988. V. 102. P. 709–720.
25. Bakke J.L., Lawrence N.L., Robinson S.A., Bennett J., Bowers C. // *Neuroendocrinology*. 1978. V. 25. № 5. P. 291–302.
26. Lopes de Souza S., Nogueira M.I., Bomfim de Jesus Deiro T.C., Manhaes de Castro F.M., Mendes da Silva C., Cesiana da Silva M., Oliveira de Lira L., Azmitia E.C., Manhaes de Castro R. // *Physiol. Behav.* 2004. V. 82. P. 375–379.
27. Mendes da Silva C., Goncalves L., Manhaes-de-Castro R., Nogueira M.I. // *Neurosci. Lett.* 2010. V. 483. P. 179–183.
28. Kinast K., Peeters D., Kolk S.M., Schubert D., Homberg J.R. // *Front. Cell. Neurosci.* 2013. V. 7. Article 72.
29. Mesquita A.R., Pego J.M., Summavielle T., Maciel P., Almeida O. F. X., Sousa N. // *Neurosci.* 2007. V. 147. P. 1022–1033.
30. Renoir T. // *Front. Pharmacol.* 2013. V. 4. № 45. P. 2–10.
31. Stenfors C., Ross S.B. // *Life Sci.* 2002. V. 71. № 24. P. 2867–2880.
32. Bosker F.J., Tanke M.A., Jongasma M.E., Cremers T.I., Jagtman E., Pietersen C.Y., van der Hart M.G., Gladkevich A.V., Kema I.P., Westerink B.H., et al. // *Neurochem. Int.* 2010. V. 57. № 8. P. 948–957.
33. Trouvin J.H., Gardier A.M., Chanut E., Pages N., Jacquot C. // *Life Sci.* 1993. V. 52. № 18. P. 187–192.
34. Stephansson O., Kieler H., Haglund B., Artama M., Engeland A., Furu K., Gissler M., Norgaard M., Nielsen R.B., Zoega H., et al. // *JAMA*. 2013. V. 309. № 1. P. 48–54.

Possible Function of the *ribT* Gene of *Bacillus subtilis*: Theoretical Prediction, Cloning, and Expression

A. P. Yakimov^{1,2}, T. A. Seregina³, A. A. Kholodnyak³, R. A. Kreneva¹, A. S. Mironov³,
D. A. Perumov¹, A. L. Timkovskii^{1,2,*}

¹B.P. Konstantinov Petersburg Nuclear Physics Institute, National Research Center "Kurchatov Institute", Orlova Roshcha, Gatchina, Leningrad Region, Russia, 188300

²St. Petersburg State Polytechnical University, Polytechnicheskaya Str., 29, St. Petersburg, Russia, 195251

³State Research Institute of Genetics and Selection of Industrial Microorganisms, 1st Dorozhnyi Proezd, 1, Moscow, Russia, 117545

*E-mail: alt@AT1660.spb.edu

Received 24.04.2014

Copyright © 2014 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT The complete decipherment of the functions and interactions of the elements of the riboflavin biosynthesis operon (*rib* operon) of *Bacillus subtilis* are necessary for the development of superproducers of this important vitamin. The function of its terminal *ribT* gene has not been established to date. In this work, a search for homologs of the hypothetical amino acid sequence of the gene product through databases, as well as an analysis of the homologs, was performed; the distribution of secondary structure elements was theoretically predicted; and the tertiary structure of the RibT protein was proposed. The *ribT* gene nucleotide sequence was amplified and cloned into the standard high-copy expression vector pET15b and then expressed after induction with IPTG in *E. coli* BL21 (DE3) strain cells containing the inducible phage T7 RNA polymerase gene. The *ribT* gene expression was confirmed by SDS-PAGE. The protein product of the expression was purified by affinity chromatography. Therefore, the real possibility of RibT protein production in quantities sufficient for further investigation of its structure and functional activity was demonstrated.

KEYWORDS proteomics; bioinformatics; homology search; theoretical protein structure; gene cloning; inducible expression.

The main stages in the riboflavin biosynthesis in *Bacillus subtilis* cells have been elucidated previously. This process turns out to be controlled by two regions of the genome: the *rib* operon and the bifunctional flavokinase/FAD synthase gene, *ribC*, which is part of the *truB-rpsO* operon [1, 2]. The *rib* operon, which controls the overall pathway of riboflavin production, starting with guanosine-5'-triphosphate (starting precursor), consists of five nonoverlapping genes. These are four consecutive structural genes: *ribG* (encodes bifunctional aminopyrimidine deaminase/uracil reductase), *ribB* (riboflavin synthase gene), *ribA* (GTP cyclohydrolase gene), *ribH* (lumazine synthase gene), as well as *ribT*, the operon's terminal gene, whose function has not been determined so far. Furthermore, this operon contains three regulatory elements: the *ribO* regulatory region with the major promoter P1 and two additional internal promoters P2 and P3.

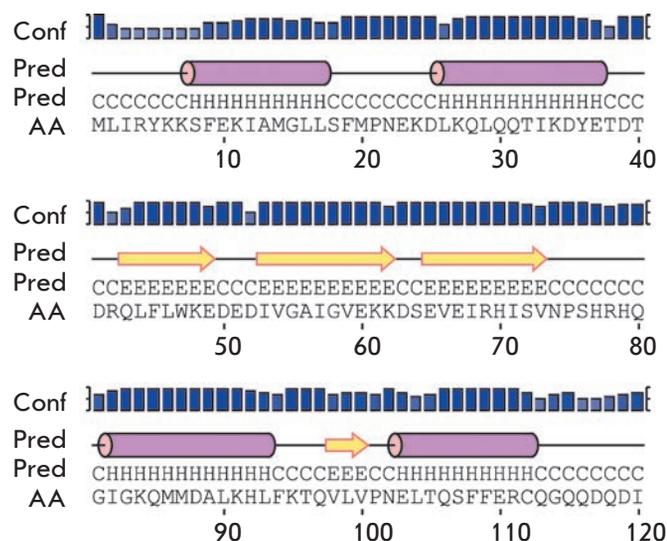


Fig. 1. Secondary structure elements predicted for the *ribT* gene product

SHORT REPORTS

Previously, we determined the relative functional activity of all three operon promoters [3]. In this case, a paradox came to light: the P2 and P3 promoters, when tested separately within corresponding fragments of the operon, differed considerably in their transcriptional activity. The activity of the P3 promoter, which regulates *ribT* gene transcription, exceeded several times the major P1 promoter activity. The P2 promoter, on the contrary, was tens of times weaker than P1. However, neither P2 nor P3 is regulated by flavins, but it is known that when the entire *rib* operon is transcribed under the major P1 promoter control, all elements of the operon are transcribed in concert to form a polycistronic mRNA [4]. And this is despite the presence of several promoters that differ in their transcriptional activity and regulatory mechanisms.

The *ribT* (*ypzK* in other nomenclatures) gene function still remains quite unclear. Indirect indications have been obtained that mutations in the *ribT* gene affect the *rib* operon activity and riboflavin accumulation. For example, Perkins *et al.* [5] demonstrated that the *ribT* gene inactivation does not lead to riboflavin auxotrophy, but it significantly reduces the riboflavin yield in producer strains. This suggests that the *ribT* gene function is important for the riboflavin biosynthesis, but it may become limiting at maximum intensity of riboflavin biosynthesis.

Therefore, elucidating the function of a *ribT* gene product may provide additional opportunities for the development of commercially promising superproducers of riboflavin, which is one of the most important vitamins.

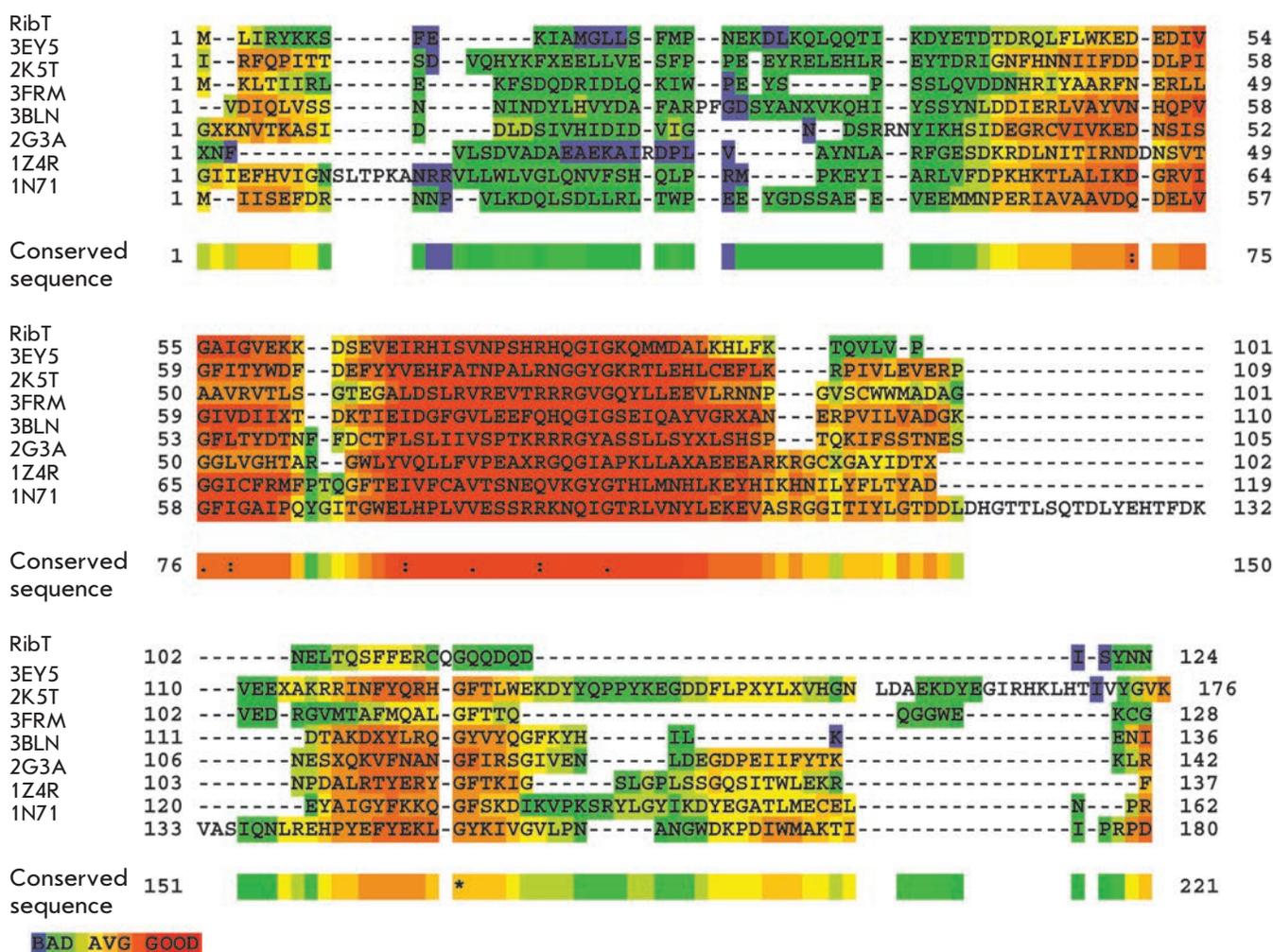


Fig. 2. Multiple alignment of the amino acid sequence of the *ribT* gene protein product to the assumed homologs (the description of the color scheme is provided at <http://www.jalview.org/help/html/colourSchemes/clustal.html>)

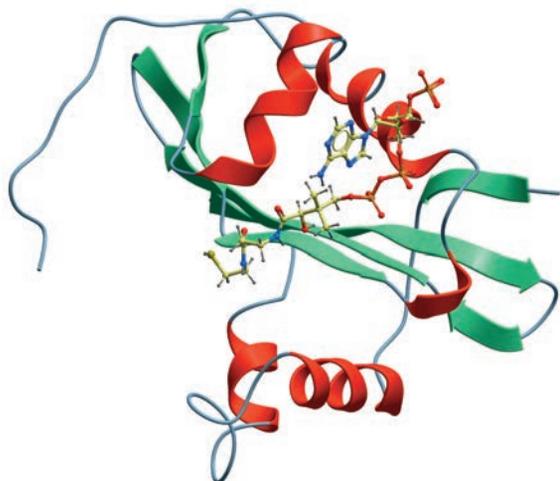


Fig. 3. Proposed tertiary structure for the *ribT* gene product in a complex with acetyl-CoA

Based on the known nucleotide sequence of the *ribT* gene, the amino acid sequence of its protein product was deduced. The hypothetical protein consists of 124 amino acids and has a molecular weight of 14.5 kDa. The prediction of the regions that form, with a high probability, the elements of the secondary structure (Fig. 1) was made using the PSIPRED service [6].

Then, based on this sequence, the search for homologs was carried out and a multiple alignment of the sequence of the *ribT* gene protein product was performed using the Clustal software [5] only among the proteins with structures deposited in the PDB [7] (Fig. 2).

Of them, 1N71, 3FRM, 2G3A, 1Z4R, 3EY5, 2K5T, and 3BLN were selected, and based on homology, the structure presented in Fig. 3 was built. Since acetyl CoA is present in the crystal structures of some homologous proteins, ligand docking to the hypothetical structure was carried out using the Molsoft ICM Pro software package [8].

Since most of the selected homologous proteins (except 3FRM, whose function is unknown) were acetyl transferases, the hypothetical *ribT* gene product could be assumed also to belong to this class of enzymes.

We assumed that the role of this gene's protein product might be the acetylation of the N(5) atom of flavins that results in the production of their reduced forms and maintains a high transcriptional level for the *rib* operon. Previously, with our participation, the mechanism of transcriptional inhibition through the direct interaction of flavins with the leader sequence of mRNA was established [9]. The consistency of our assumption is confirmed by the fact that the inhibitory transcriptional interaction with the leader sequence of a sensory RNA is performed by the oxidized form of FMN [10].

Therefore, the acetyl reduction produced by the *ribT* gene product may be important for maintaining a high level of riboflavin synthesis. This assumption requires direct experimental verification.

This defines the objective of further research that is to perform the full cycle of *ribT* gene expression in a preparative mode, to produce sufficient quantities of the purified native protein product, and to test directly its functional (enzymatic) activity *in vitro*.

To test the possibility of expression, the *ribT* gene was amplified from chromosomal DNA of *B. subtilis* using the primers RibT10, 5'-CGCCATATGTTA-ATTTCGTTATAAAAAATCGTTT-3', and RibT11, 5'-CGCCTCGAGTAATTATTGTATGAAATGTCTT-GATC-3', (the oligonucleotides used in this study were synthesized at Evrogen). The first oligonucleotide is complementary to the proximal region, and the second is complementary to the distal region of the *ribT* gene. PCR was performed on a MyCycler thermal cycler (BioRad) according to the following scheme: first, cells were disrupted at 95 °C for 3 min, then 25 cycles of amplification were carried out that included DNA denaturation at 95 °C for 30 s, primer annealing at 60 °C for 30 s, and completion of DNA at 72 °C for 30 s. At the last stage, completion of DNA was done at 72 °C for 2 min. As a result, a fragment of 372 bp was synthesized that contained the structural region of the *ribT* gene flanked by the restriction enzyme recognition sites, NdeI and XhoI. After electrophoretic separation of PCR products, the desired DNA fragment was eluted from the gel using a GeneClean kit (Fermentas). The *ribT* gene was cloned into the pET15b high-copy expression vector containing the T7 phage promoter, which is inducible by isopropyl β-D-1-thiogalactopyranoside (IPTG), at the restriction endonuclease sites NdeI and XhoI. The pET15b vector comprises nucleotide sequences for His-Tag before the NdeI restriction enzyme recognition site and for the target site of thrombin. *E. coli* TGI strain cells were transformed with the resulting ligase mixture. Selection of transformants was performed in a LB agar medium containing ampicillin as a selection marker. Screening for recombinant clones was performed by PCR using the plasmid primers pT7P, 5'-TAATACGACTCACTATAGGGG-3', and pT7T, 5'-GCTAGTTATTGCTCAGCGGT-3'. Plasmid DNA was isolated from the selected transformants, and the presence of insertion in hybrid plasmids was determined using a restriction analysis. These plasmids, designated as pET15b/*ribT*, were used to transform cells of the *E. coli* BL21 (DE3) strain containing the inducible T7 bacteriophage RNA polymerase gene.

The RibT protein synthesis was induced by adding IPTG to the growth medium at a final concentration of 1 mM. The *ribT* gene expression was determined by

polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAGE) for a total protein from *E. coli* BL21 (DE3) cells containing the pET15b/*ribT* plasmid. As a control, the lysate of BL21 (DE3) strain cells containing the pET15b plasmid without the insertion was used (Fig. 4A).

An additional fraction of a protein with a molecular weight of approximately 14.5 kDa is observed in the cell lysate of the BL21 (DE3) strain containing the pET15b/*ribT* plasmid after IPTG mediated induction, which is consistent with the predicted molecular weight of the RibT protein (see above).

The His-Tag-labeled recombinant RibT protein was isolated using TALON® Magnetic Beads (Clontech, USA). Cells of the *E. coli* BL21 (DE3) strain containing the pET15b/*ribT* plasmid, grown in the presence of 1 mM IPTG, were harvested by centrifugation. The biomass was re-suspended in a buffer of the following composition: 20 mM sodium phosphate, pH 7.0, 300 mM NaCl, and 20 mM imidazole. The cells were disrupted by sonication and centrifuged at 14,000 rpm for 20 min. The supernatant was incubated with TALON® Magnetic Beads at 4 °C for 1 h. The resin was then washed with four volumes of the same buffer. Elution of the protein was performed with a buffer: 20 mM sodium phosphate, pH 7.0, 300 mM NaCl, and 300 mM imidazole. The eluted protein's fractions were analyzed by SDS-PAGE electrophoresis (Fig. 4B). Lines 7–10 belonged to the target protein with a high degree of purity and a molecular weight of about 14.5 kDa, which is in good agreement with the theoretical prediction.

CONCLUSIONS

The theoretical amino acid sequence and analysis of *ribT* gene product homologs in the PDB allowed us to perform molecular modeling and predicting of the pos-

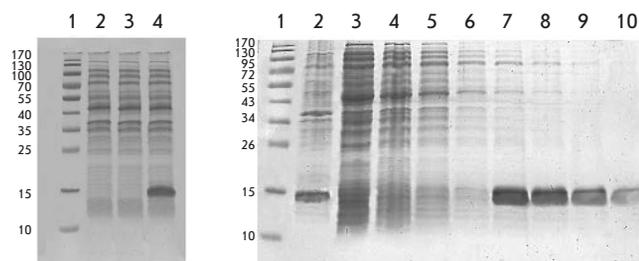


Fig. 4. SDS-PAGE analysis of the *ribT* gene expression and recombinant RibT protein purification.

A total protein fraction from *E. coli* BL21(DE3) cells.

Line 1: a protein marker; line 2: a protein fraction from *E. coli* BL21(DE3) cells with the pET15b plasmid without *ribT* gene insertion; line 3: a protein fraction from *E. coli* BL21(DE3) cells with the pET15b/*ribT* plasmid, grown without IPTG addition; line 4: protein fraction from *E. coli* BL21(DE3) cells with the pET15b/*ribT* plasmid, grown in the presence of 1 mM IPTG.

A protein fraction from *E. coli* BL21(DE3) pET15b/*ribT* cells before and after affinity sorption on TALON® Magnetic Beads. Line 1: a protein marker; line 2: a total protein fraction from *E. coli* BL21(DE3) pET15b/*ribT* cells after IPTG induction; lines 3–6: protein fractions with proteins unbound to the affinity sorbent; lines 7–10: fractions of the His-Tag-labeled RibT protein successively eluted from the affinity sorbent with 300 mM imidazole

sible three-dimensional structure of the protein. The *ribT* gene expression was experimentally implemented, and the possibility, in principle, to produce the RibT protein in quantities sufficient for further investigation of its structure and functional activity was proved.

This work was supported by the Russian Foundation for Basic Research (grant № 11-04-00476).

REFERENCES

- Chikindas M.L., Mironov V.N., Lukyanov E.V., Boretzky Yu.R., Arutyunova L.S., Rabinovich P.M., Stepanov A.I. // Mol. Gen. Mikrobiol. Virusol. (Russian). 1987. V. 4. № 1. P. 22–26.
- Mironov V.N., Perumov D.A., Kraev A.S., Stepanov A.I., Skryabin K.G. // Mol. Biol. (Russian). 1990. V. 24. № 3. P. 256–261.
- Sklyarova S.A., Kreneva R.A., Perumov D.A., Mironov A.S. // Genetika (Russian). 2012. V. 48. № 10. P. 1–9.
- Mironov V.N., Kraev A.S., Chikindas M.L., Chernov B.K., Stepanov A.L., Skryabin K.G. // Mol. Gen. Genet. 1994. V. 242. № 2. P. 201–208.
- Perkins J.B., Sloma A., Hermann et al. // J. Ind. Microbiol. Biotechnol. 1999. V. 22. P. 8–18.
- Jones D.T. // J. Mol. Biol. 1999. V. 292. P. 195–202.
- Bernstein F.C., Koetzle T.F., Williams G.J., Meyer Jr. E.E., Brice M.D., Rodgers J.R., Kennard O., Shimanouchi T., Tasumi M. // J. Mol. Biol. 1977. V. 112. № 3. P. 535–542.
- Abagyan R.A., Totrov M.M., Kuznetsov D.N. // J. Comp. Chem. 1994. V. 15. P. 488–506.
- Mironov A.S., Gusarov I., Rafikov R., Errais Lopez L., Shatalin K., Kreneva R.A., Perumov D.A., Nudler E. // Cell. 2002. V. 111. № 4. P. 747–756.
- Serganov A., Huang L., Patel D.J. // Nature. 2009. V. 458. P. 233–237.

GENERAL RULES

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

The journal is published by the Park Media publishing house in both Russian and English.

The journal *Acta Naturae* is on the list of the leading periodicals of the Higher Attestation Commission of the Russian Ministry of Education and Science

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

The maximum length of a review, together with tables and references, cannot exceed 60,000 symbols (approximately 40 pages, A4 format, 1.5 spacing, Times New Roman font, size 12) and cannot contain more than 16 figures.

Experimental articles should not exceed 30,000 symbols (20 pages in A4 format, including tables and references). They should contain no more than ten figures. Lengthier articles can only be accepted with the preliminary consent of the editors.

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

The manuscript should be sent to the editors in electronic form: the text should be in Windows Microsoft Word 2003 format, and the figures should be in TIFF format with each image in a separate file. In a separate file there should be a translation in English of: the article's title, the names and initials of the authors, the full name of the scientific organization and its departmental affiliation, the abstract, the references, and figure captions.

MANUSCRIPT FORMATTING

The manuscript should be formatted in the following manner:

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

(1,500 characters).

- Keywords (3 – 6). These should include the field of research, methods, experimental subject, and the specifics of the work. List of abbreviations.

- INTRODUCTION
- EXPERIMENTAL PROCEDURES
- RESULTS AND DISCUSSION
- CONCLUSION

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

- REFERENCES

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

RECOMMENDATIONS ON THE TYPING AND FORMATTING OF THE TEXT

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

GUIDELINES FOR AUTHORS

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

GUIDENESS FOR ILLUSTRATIONS

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

REVIEW AND PREPARATION OF THE MANUSCRIPT FOR PRINT AND PUBLICATION

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

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

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

A manuscript which has been sent back to the authors for improvements requested by the editors and/or reviewers is reviewed again, after which the editorial board makes another decision on whether the article can be accepted for publication. The published article has the submission and publication acceptance dates set at the beginning.

The return of an article to the authors for 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.

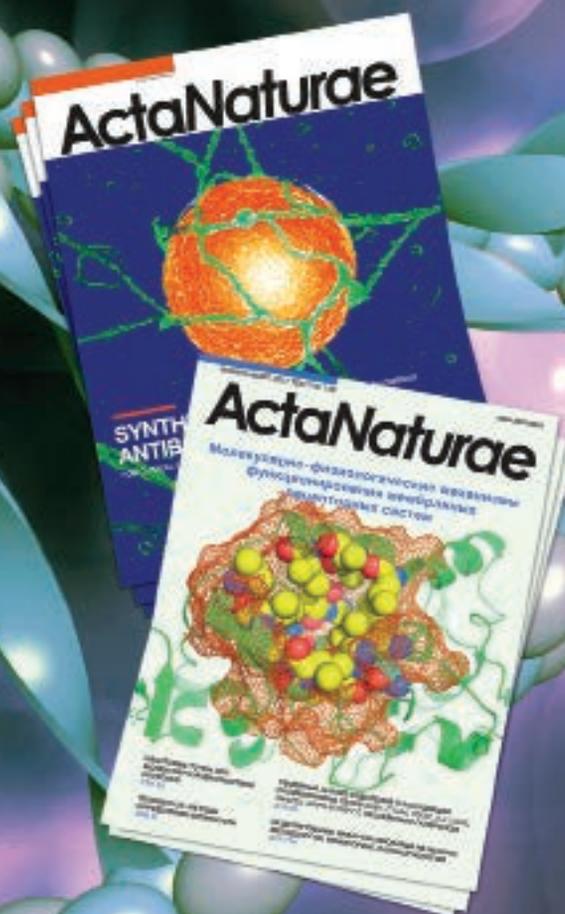
The following e-mail addresses can be used to contact the editorial staff: vera.knorre@gmail.com, actanaturae@gmail.com, tel.: (495) 727-38-60, (495) 930-87-07

ActaNaturae

SUBSCRIPTION TO

Acta Naturae journal focuses upon interdisciplinary research and developments at the intersection of various spheres of biology, such as molecular biology, biochemistry, molecular genetics, and biological medicine.

Acta Naturae journal is published in Russian and English by Park Media company. It has been included in the list of scientific journals recommended by the State Commission for Academic Degrees and Titles of the Ministry of Education and Science of the Russian Federation and the Pubmed abstracts database.



SUBSCRIBE AT THE EDITORIAL OFFICE

Leninskie Gory, 1-75G, Moscow, 119234 Russia
Telephone: +7 (495) 930-87-07, 930-88-51
Bio-mail: podpiska@biorf.ru
Web site: www.actanaturae.ru

SUBSCRIBE USING THE CATALOGUES OR VIA THE INTERNET:

ROSPECHAT (The Russian Press)
Indices: 37283, 59881
www.pressa.rosp.ru

INFORMNAUKA
Index: 59881
www.informnauka.com

INTER-POCHTA
17510
www.interpochta.ru

INFORMATION FOR AUTHORS:

If you would like to get your research paper published in *Acta Naturae* journal, please contact us at journal@biorf.ru or call +7 (495) 930-87-07.

