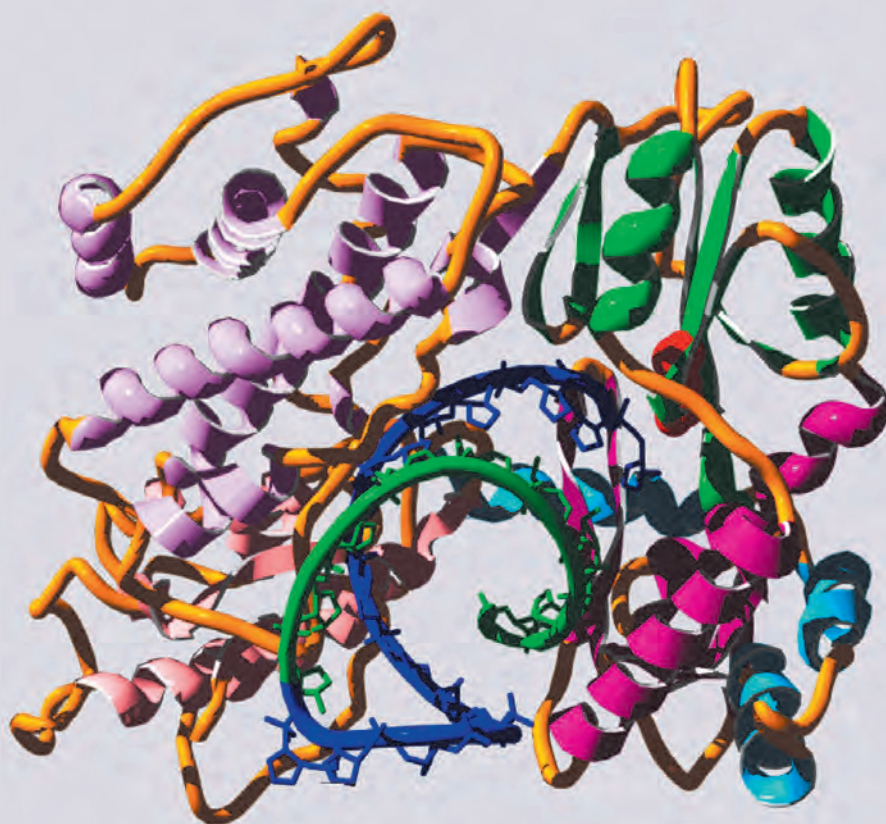


Acta Naturae

Coagulation Factor IX for Hemophilia B Therapy



PHYSICOCHEMICAL BIOLOGY: CONQUERED
BOUNDARIES AND NEW HORIZONS

P. 36

TELOMERE LENGTHENING AND OTHER
FUNCTIONS OF TELOMERASE

P. 44

Letter from the Editors

Dear readers of *Acta Naturae*,
We would like to bring to your attention the 13th issue of our journal. In spite of the popular superstition around this number, we actually believe it to be a fortunate one for us: our protracted struggle to secure a listing in the PubMed database is finally over: we are happy to report that now the contents, abstracts, and full-text articles of all *Acta Naturae* issues are available in this database. We congratulate all the authors, readers, and publishers of our journal on this momentous event – from now on *Acta Naturae* is a fully fledged scientific publication! We hope that our journal will acquire a respectable impact factor in the near future. Of course, our success depends primarily upon the continued high quality of our dear authors' work.

Let us return to the current issue of the journal. The “Forum” section is devoted to a topic that is significant for everyone: the publication activity of Russian scientists in the leading world journals and the quality criteria of these journals. It is no secret that the positive results of Russian researchers do not always result in publication in high-ranking international journals. We believe that the information given in the “Forum” section will provide an objective overview of the situation and will assist in the quest to find a solu-

tion to these appreciably difficult publishing problems.

As always, the research section of the journal opens with the reviews; this time there are three of them in total. First of all, it is a publication composed by the patriarch of Russian physicochemical biology, the Academician, D.G. Knorre. There is no need to embellish the title of the review with any further comments. We believe that this publication will be of significant interest to our dear readers.

The two other reviews are devoted to the topical problems of modern life science.

As always, a significant part of this issue is devoted to the medical application of the problems under consideration, which certainly reflects the global trends.

Referring back to the beginning of this Letter – the inclusion of *Acta Naturae* into PubMed – we would like to urge future authors of our journal to strive for high-quality work, in line with the high quality papers that we have received up until this moment. In return, the Editorial Board promises to always remain fair when considering the materials received. We express the sincere hope that our future collaboration will be fruitful. ●

*See you in the next issue of Acta Naturae,
Editorial Board*

Save 10% on Subscription for 2012

Details at www.actanaturae.ru

DISCUSSION

Modeling approaches can be further improved by implementing new algorithms of the conformational search and new scoring functions (methods to estimate the free energy of ligand binding). Scoring functions may include other components of molecular mechanics force fields [2] or empirical terms, e.g. hydrogen bonds described by their geometrical parameters [4]. In this work we studied stacking interactions, which usually are not properly taken into account in widely used scoring functions.

THE PARAMETERS OF STACKING INTERACTIONS

Of all the various types of interactions in biomolecular complexes (such as hydrogen bonds, salt bridges, etc.), the stacking of aromatic substances deserves special attention. Most drugs include aromatic fragments in their chemical structure, and stacking often plays a notable role in their recognition by protein targets. We have recently shown that an explicit account of stacking in scoring functions increases the efficiency of ATP docking [5]. The aromatic interactions were identified by the mutual orientation of two cycles described by geometrical parameters: the height h and displacement d of one cycle relative to the other, and the angle between their planes (Fig. 1).

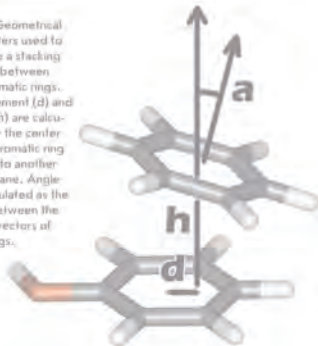
However, the range of these parameters, which corresponds to the presence or absence of a stacking contact, is still not very well defined and usually takes an arbitrary [6, 7]. Defining it more accurately would assist in developing more efficient scoring functions and should improve the prediction quality of the spatial structures of protein-ligand complexes by molecular modeling methods. With this aim in view, we performed an analysis of the spatial structures of protein-ligand complexes determined experimentally with atomic resolution where ligands contained aromatic moieties as a substructure.

The well-known example of stacking interactions is the parallel packing of purine and pyrimidine nucleosides in DNA [8, 9]. Some aromatic compounds tend to orient perpendicular to each other (T-shaped stacking), as has been shown for amino acids in proteins [7, 10] and for model systems of carbon aromatic cycles (benzene and naphthalene) [11–14]. Besides, such compounds participate in cation- π interactions, where a positively charged group interacts with the negatively charged cloud of aromatic π -electrons [15–17].

Taking all that into account, we analyzed the distribution of geometrical parameters h , d , and α for contacts of aromatic and guanine moieties of ligands with the aromatic side chains of receptor amino acids Phe, Tyr, Trp, and His, as well as with the positively charged guanidino group of Arg and amino group of Lys. The results obtained for guanine are presented in Fig. 2.

It can be seen that two distinct orientations are typical for Phe: parallel and perpendicular to the guanine plane (Fig. 3, shown in red and green, respectively). The displacement d lies in the same range (1–3 Å) for both types of contacts. Meanwhile, they differ in the value of height h , which is ≈ 3 Å for parallel and ≈ 1 Å for perpendicular orientation. Similar distributions were obtained for Tyr, Trp, and His, though the data are sparser in those cases. However, the T-shaped contact is not as typical for Tyr, Trp, and His as it is for Phe.

Fig. 1. Geometrical parameters used to describe a stacking contact between two aromatic rings. Displacement (d) and height (h) are calculated for the center of one aromatic ring relative to another ring's plane. Angle α is calculated as the angle between the normal vectors of both rings.



APRIL/JUNE 2009, No. 1

Acta Naturae



SYNTHETIC ANTIBODIES
FOR CLINICAL USE

REGULATING TELOMERASE IN ONCOGENESIS
P. 51

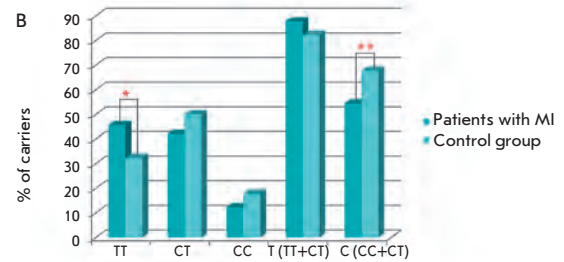
THE STRUCTURE OF THE MITOCHONDRIAL GENOME AS AN ACTIVATOR OF OPISTHORCHIASIS
P. 59

STACKING INTERACTIONS IN COMPLEXES OF FIBERS WITH ADENINE- AND GUANINE-CONTAINING LIGANDS

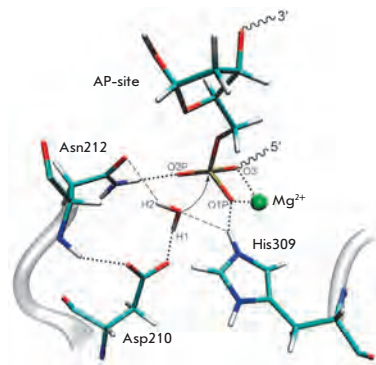
Contribution of the *TGFB1* Gene to Myocardial Infarction Susceptibility

R.M. Barsova, B.V. Titov, N.A. Matveeva, A.V. Favorov,
T.S. Sukhinina, R.M. Shahnovich, M.Ia. Ruda, O.O. Favorova

Carriage frequencies of alleles and genotypes of the *TGFB1* gene polymorphous loci -509C>T (rs1800469), 869T>C (rs1982073), 915G>C (rs1800471), which affect the level of cytokine TGF- β 1 production, were analyzed in the patients of Russian ethnic descent with myocardial infarction (MI) and in the control group of the same ethnic descent. The results obtained indicate the important role of the *TGFB1* gene in susceptibility to MI, including earlyonset and recurrent MI.



Carriage frequencies of alleles and genotypes of *TGFB1* gene polymorphous loci 869T>C in the patients with MI and in the control group.



The active site of the full-atomic model of APE1 enzyme-substrate complex.

Construction of a Full-Atomic Mechanistic Model of Human Apurinic/Apyrimidinic Endonuclease APE1 for Virtual Screening of Novel Inhibitors

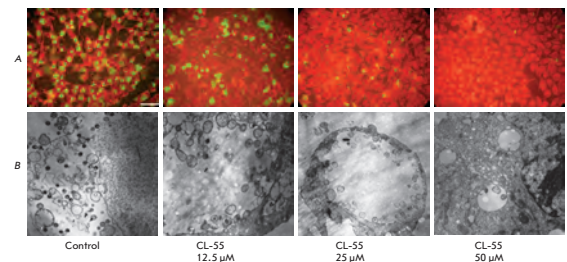
I.G. Khaliullin, D.K. Nilov, I.V. Shapovalova, V.K. Švedas

A full-atomic molecular model of human apurinic/apyrimidinic endonuclease APE1, an important enzyme in the DNA reparation system, has been constructed. The choice of the APE1 mechanism with an Asp210 residue as a proton acceptor was validated by means of a generalization of modelling and experimental data. Interactions were revealed in the active site that are of greatest significance for binding the substrate and potential APE1 inhibitors.

Development of Chlamydial Type III Secretion System Inhibitors for Suppression of Acute and Chronic Forms of Chlamydial Infection

N.A. Zigangirova, E.S. Zayakin, L.N. Kapotina, E.A. Kost,
L.V. Didenko, D.Y. Davydova, J.P. Romyanceva, A.L. Gintsburg

The Type III secretion system (T3SS) is currently considered to be one of the main pathogenicity factors in Gram-negative bacteria, which exhibit different types of parasitizing activity. T3TS is one of the most promising targets for the development of broad-spectrum antimicrobial drugs that do not develop resistance and are efficacious for the acute and chronic forms of infection. As a result of pilot screening using specific cellular and bacterial tests, followed by chemical optimization and detailed characterization of the biological activity, a new class of chlamydial T3SS inhibitors was obtained. The selected compounds have obvious advantages over the currently available inhibitors thanks to the high inhibitory activity with minimal damaging effects on eukaryotic cells.



Dose-dependent inhibition of the intracellular development cycle of *C. trachomatis* under the effect of CL-55.

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: A.I. Gordeyev

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: E.A. Novoselova

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: 119991 Moscow, Russia, Leninskiye Gory, Nauchny
Park MGU, vlad.1, stroeniye 75G.

Phone/Fax: +7 (495) 930 80 06

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

Reprinting is by permission only.

© ACTA NATURAE, 2012

Номер подписан в печать 28 июня 2012 г.

Тираж 200 экз. Цена свободная.

Отпечатано в типографии «МЕДИА-ГРАНД»

CONTENTS

Letter from the Editors 1

FORUM

V.A. Markusova

Quality of Scholarly Journals and Major
Selection Criteria for Coverage
by the Web of Science 6

M. N. Kotsemir

Publication Activity
of Russian Researches in Leading
International Scientific Journals 14

REVIEWS

D.G. Knorre

Physicochemical Biology:
Conquered Boundaries
and New Horizons 36

M.P. Rubtsova, D.P. Vasilkova,
A.N. Malyavko, Yu.V. Naraikina,
M.I. Zvereva, O.A. Dontsova

Telomere Lengthening
and Other Functions of Telomerase 44

N. A. Orlova, S. V. Kovnir, I. I. Vorobiev,
A.G. Gabibov

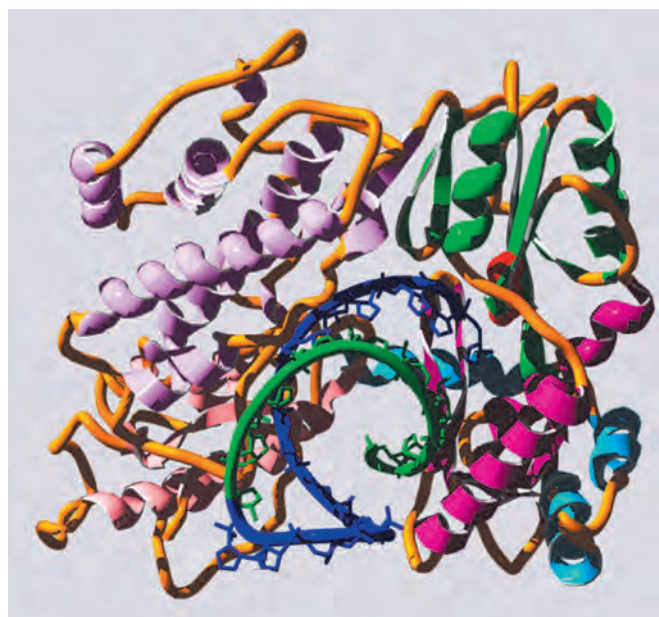
Coagulation Factor IX
for Hemophilia B Therapy 62

RESEARCH ARTICLES

- R.M. Barsova, B.V. Titov, N.A. Matveeva,
A.V. Favorov, T.S. Sukhinina,
R.M. Shahnovich, M.Ia. Ruda, O.O. Favorova
**Contribution of the *TGFB1* Gene
to Myocardial Infarction Susceptibility74**
- I.G. Khaliullin, D.K. Nilov, I.V. Shapovalova,
V.K. Švedas
**Construction of a Full-Atomic
Mechanistic Model of Human
Apurinic/Apyrimidinic Endonuclease APE1
for Virtual Screening of Novel Inhibitors.80**
- N.A. Zigangirova, E.S. Zayakin, L.N. Kapotina,
E.A. Kost, L.V. Didenko, D.Y. Davydova,
J.P. Rumyancheva, A.L. Gintsburg
**Development of Chlamydial Type III
Secretion System Inhibitors for Suppression
of Acute and Chronic Forms
of Chlamydial Infection87**

SHORT REPORTS

- Yu. A. Morozov, L. M. Khromykh,
I. I. Dementieva, M. A. Charnaia,
N. L. Kulikova, D.B. Kazansky
**Recombinant Human Cyclophilin A
in vitro Inhibits the Formation
of Fibrin Clot98**
- Guidelines for Authors.102**
- Erratum.104**

**IMAGE ON THE COVER PAGE**

Spatial structure of the *Tribolium castaneum* TERT complex and the RNA-DNA duplex (M.P. Rubtsova *et al.*)

Quality of Scholarly Journals and Major Selection Criteria for Coverage by the Web of Science

V.A. Markusova

All-Russian Institute for Scientific and Technical Information, Russian Academy of Sciences, Ul. Usievicha, 20, Moscow, 125190, Russia

E-mail: markusova@viniti.ru

Received 02.05.2012

ABSTRACT The major criteria for scholarly journals selection to be covered by the Web of Science are discussed. These criteria include the compliance of a journal with world standards, the international diversity, citation score of authors and editorial board members, the journal impact factor and journals self-citation. The demand for the unified transliteration of the authors name and the use of the unified English-language name and address of a research organization, as well as for including the information on funding organization (grant number and agency name) is emphasized.

KEYWORDS journals; impact factor; citation; self-citation; selection criteria; Journal Citation Report; Web of Science.

BIBLIOMETRIC INDICATORS AND THEIR ROLE IN THE EVALUATION OF THE RESEARCH ACTIVITY OF SCIENTIFIC INSTITUTIONS AND UNIVERSITIES

The significant geopolitical changes that have taken place during the past two decades, including the dissolution of the USSR, have considerably altered the landscape of world science. Starting with the 1990s, a wave of market liberalization resulted in the establishment of a new economy, which was accompanied by an unprecedented level of industrial activity, increased investments into the high-technology sector and the corresponding changes in the economic structure. The governments of a number of countries, including the developing ones, consider science and technology to be the key factor of economic growth and development and thus have confronted the challenge of establishing a knowledge-intensive economy. The detailed pattern of the resulting significant changes in the number of students, PhD students, postdoctoral fellows, and researchers both in the U.S. and

in other industrial and developing countries was presented in the report of the U.S. National Science Foundation titled Science & Engineering Indicators-2012 (SEI), published in January 2012, www.nsf.gov. Impressive statistics on the investments made in basic research in the U.S., the European Union, China, Russia, and the other leading world countries was presented there. Furthermore, the measures undertaken by the developing countries for the development of a science and technology (S&T) infrastructure, for the stimulation of industrial research and development (R&D), for the expansion of the higher education system and establishment of national systems for basic research, as well as for the establishment of their own markets for trade and foreign investments, were discussed. As a result of the recent financial crisis, investments into innovation research and development are experiencing downward economic pressure; in this context, increasing attention is being put on assessing the yield and efficiency of the national innovation systems.

The 5th section of the Report was devoted to the bibliometric statistics characterizing the levels of research and development (R&D) in these countries. This statistics is based on the data contained in the Science Citation Index-Expanded (SCI-E) and Social Science Citation Index-Expanded (SSCI-E), which are parts of the Web of Science (WoS) produced by the Thomson Reuters (hereinafter, TR).

The publication in a prestigious international journal covered by WoS is the most reliable method for bringing the results of the research to the attention of the world scientific community. Let us note that the experts who deal with assessing the output of scientific activity refer to all the journals covered by WoS as “international.” The journal’s prestige or its informational significance is assessed through the impact factor of a journal (IF).

In accordance with the Resolution of the Presidium of RAS № 212 dated October 12, 2009, “On the Establishment of the Committee for the Assessment of the Activity of Scientific Institutions,” the

institutes of the Russian Academy of Sciences are in the process of collecting a vast body of statistical data, including bibliometric indicators, such as the number of papers published by institute researchers in various databases (DB), citation score, and the mean weighted average impact factor of an institution. The hidden stones of using these indicators were discussed in the journal *Herald of the Russian Academy of Sciences* [1]; it was noted that the assessments of the output of scientific activity should be based on the decisions by experts in the corresponding field of science, whereas the bibliometric indicators are an additional tool for decision making.

The global financial crisis of 2008 has forced most foreign investment organizations supporting fundamental research to focus more on bibliometrics as an objective indicator of the quality of a scientific product. Thus, the UK Research Assessment Exercises has tied the distribution of financial support to universities to the indicators of scientific output. A trend towards the increased attention by university administrations to publications in prestigious international journals with high impact factors has also been observed. Unfortunately, this trend had a negative result: namely, the disappearance of a number of national journals on social sciences in the Netherlands [2].

The paper published by Dr. E. Garfield in *Science* [3] in 1955, where he discussed the idea of citation indexing, was a historical moment in the use of the informational flows of scientific literature consisting of the papers and citations in them. The experimental “Science Citation Index” (SCI) in natural science and engineering was published in 1963. The Institute for Scientific Information (ISI), which belonged to E. Garfield, started publishing SCI on a regular basis since 1964. Since 2001, ISI has been owned by the

Thomson Reuters company, the leader on the global market for information resources.

The scientific community rather rapidly welcomed the potential offered by the use of large collection of bibliographic information as a tool to assess the efficiency of the scientific activity of countries, universities, and research institutions. The bibliometric indicators of research output and its citedness in the U.S. and other countries were first published in 1972 in the report “Science Indicators” prepared by the U.S. National Science Foundation and since then have been published every other year. Since 1996, these reports have been known as Science & Engineering Indicators (SEI).

The compilation principle of all the information products launched by ISI is based on the fact that, beginning from the late XIX century, a paper published in a journal has been considered as a form of scientific communication. The well-known American expression “publish or perish” attests to the fact that the number of papers published is a significant factor of the professional acknowledgement of a researcher and facilitates his or her moving up the promotion ladder. The Nobel laureate and academician V.L. Ginzburg has mentioned that “The timely publication of the studies and support of the best ones is a necessary condition for success in research, namely, in providing the international acknowledgement of this success” [4].

Today, as it was 100 years ago, a scientific paper contains citations of predecessor research. If a paper contains no citations, this is considered as one of the signs of the low qualification of the beginning author and complicates the publication of the paper in a journal. When the author cites the works of other researchers, he or she demonstrates by so doing the conceptual relation-

ship between the subject of his or her paper and that of the works cited. Not being strictly formalized language, the citations enable to detect the internal conceptual links between publications [5].

The tens of thousands of articles, short notes, editorials, letters and reviews that are published daily in journals, as well as the millions of citations cited by these articles provide a way for penetrating knowledge communication, and its dissemination into science and the provide the empirical data on a significance of the research and scientific activity (6).

Along with the accumulation of large collections of bibliographic information in ISI and the development of computation facilities in the U.S., there emerged an opportunity for the establishment of a new information product based on the relationships between journals. The Journal Citation Reports (JCR) for natural sciences and engineering, which contained statistical data for 3,000 journals, was first published in 1975; since that time, it has been published annually. Its special edition, JCR-Social Sciences, has been published since 1978. According to E. Garfield, “Many researchers and editors often make the even more outrageous error of assuming that SCI was created just to produce its by-product called Journal Citation Report (JCR)”. The major purpose for these resources was not only to aid information retrieval but also to use it as an alerting tool, i.e. for selective dissemination of information” (7). The term “**impact factor**” (IF) was first proposed by E. Garfield jointly with Dr. Irving Sher in 1955 [8]. The introduction of this term promoted a more qualitative selection of scholarly journals (hereinafter, journal) by libraries and information services. It soon gained popularity as a symbol of the scientific prestige of a journal, although its values significantly vary depending

on the branch of knowledge and its relevant subject field.

The elaboration and development of methods for the analysis of bibliographic information has resulted in the emergence of a new scientific discipline: scientometrics. E. Garfield has noted that “We are witnessing the transformation of bibliometric research into a new field of industry – the assessment of the output of the research carried out by university and scientific teams” [7]. Although the discontent of the scientific community with the enthusiasm of bureaucrats from foundations and ministries for various ratings and evaluations is growing, the impact these indicators have on the financial support that global fundamental science can count on is becoming increasingly noticeable.

Over the past decade, the progress in the development of the databases (DB) and information technology of almost all social institutions and processes have resulted in the establishment of network technologies that enable one to work with very large collection of data. The information platform “Web of Knowledge” (WOK) is an example of a network technology of this type. One of its components, the information system Web of Science (WoS), comprises the expanded versions: SCI-Expanded, Social Science Citation Index-Expanded и Art & Humanities Citation Index. WoS is currently the largest citation DB with indexing of over 800 million citations in papers published over the period from 1900 until 2010. A total of 12,600 journals are covered by WoS. The attained level of development of network technologies allows one to use large data collection to resolve scientometric problems. In essence, we can speak about the establishment of “network” scientometrics. The network technology enables one to obtain more adequate evaluation of

the Russian science contribution to the global thesaurus of knowledge.

In 2005, the well-known scientific publisher Elsevier (The Netherlands) established the SCOPUS database (hereafter, SCOPUS) and posted it on the Internet; the system comprised articles and citations from 18,000 journals starting from 1996. A total of 230 Russian journals are used for this system. The number of Russian journals covered by the SCOPUS is higher compared to that for Web of Science; however, this system has a number of significant drawbacks. In particular, it is extremely difficult to obtain reliable results during the general search using the names of Russian research institutions and Russian Academy of Sciences. Still, Elsevier is very active in the Post-Soviet space; it organizes numerous workshops for potential users. The search for organizations in this system is more complicated compared to that in SCI. For a journal, it is easier to be covered into the SCOPUS system than into the Web of Science, since the “selection criteria” are considerably lower in the former there. **Let us to emphasize that all the comparative evaluations of the national science contribution to global science and university ranking are based on bibliometric indicators from the Web of Science.** The monitoring of these bibliometric indicators is carried out in all industrially developed countries. The data relating to research output assist in making strategic decisions on the directions of research development and evaluating the position of a research institution or a university with respect to global standards in a certain field of knowledge.

JOURNALS SELECTION CRITERIA FOR COVERAGE IN THE WEB OF SCIENCE (WoS)

When submission a journal for coverage into any foreign database, one needs to keep in mind the mission of the journal as the main channel of

scientific communication. The bibliographic data (including the correspondence address and e-mail of the author, working address of the organization, and source of financial support) and all citations within the journal are widely used to collect the bibliographic information and are regarded as important indicators of the researchers’ scientific activity.

A journal has different functions; in general, they provide ideas – on the research’ development of and its progress; competitive strength of science and the degree of its integration into the global scientific community; – on the publication activity of the authors; – on the publication activity and ranking of institutions; – on the degree of recognition and the level of publications in the global scientific community, according to the citations of these papers; and – on the quality of the national journals in comparison with the world flow of publications in the corresponding subject field, etc.

The research performed by English bibliographer S. Bradford was the theoretical foundation for the selection of journals in SCI. In 1930, he formulated one of the key regularities in the distribution of publications within scientific periodicals, which is known as the Bradford’s law of scattering. According to this law, three zones containing an equal number of articles devoted to a certain problem can be isolated in the list of journals belonging to a certain subject and arranged in such a manner that the number of articles devoted to this problem decreases. These zones differ in the number and quality of the journals comprising them: the first zone (the Bradford’s core) includes specialized journals (“core” journals) that are directly devoted to a certain subject; the second zone includes the journals that are par-

tially devoted to the adjacent fields of knowledge; the third zone, the largest one, comprises the journals with subject fields remote from the specified one. The number of journals in these three zones is roughly in proportion to $1 : n : n^2$, where n depends on the subject field [9]. One of the major principles of acquisition of library funds and activity of all information products and services is based on this law.

Thus, the analysis of 7,621 journals covered in the **Journal Citation Reports (JCR)** in 2008 demonstrated that less than 300 journals accounted for approximately 50% of the citations and contained roughly 30% of the papers published.

A group of “core” journals consisting of 3,000 journals publishes approximately 80% of all papers; among those, 90% were cited at least once. This group of “core” journals changes in accordance with the evolution of science; therefore, the task of the company’s personnel is to update the list of journals covered, as well as to identify and evaluate new journals. The journals that have become less useful are eliminated from the database.

During the first years after the SCI was established, the editors of journals were wary of their publication coverage by SCI; then, the situation changed dramatically with the years. The recognition of the significance of the SCI by scientific community and the commercial success of SCI resulted in the fact that E. Garfield’s office was filled with editors’ requests: “Do not perish us, include in the SCI” [10]. In the early 1970s, the caption “covered by SCI” first appeared on the covers of journals.

Three indicators connected both with the qualitative and quantitative characteristics of journals have been taken as a basis for the principles of journal selection process :

- citation data;
- compliance of a journal with cer-

tain scientific and publisher’s parameters (journal standard); and

- expert evaluation.

The TR company has used for several decades the technological and scientific achievements of the ISI, the unique information service that first started processing the references contained in scientific articles, letters to editors, editors’ columns, and book reviews published in journals. These data are invaluable sources of information pertaining to the qualitative characteristics used for the journals’ evaluation that have been in publication over a long period of time (11).

Selection of new journals is mostly based on expertise of the journal’s quality. The evaluation of journals and their exclusion from this system is a continuous process. Every other week, journals are added or excluded from the TR databases. About 2,000 new journals annually are examined by the company staff; only 10–12% of those are selected for processing.

Continual monitoring of journals ensures that they comply with high scientific standards and stay relevant to the products in which they are covered. All journals covered by Web of Science undergo selection process regardless of the database they are supposed to be covered in: Science Citation Index-Expanded, Social Science Citation Index-Expanded, or Art & Humanities Citation Index. Special attention is given to an evaluation of journals on social sciences or Art & Humanities due to the specificity of the citations patterns in these fields of knowledge (lower citations indicators compared to those for natural sciences).

TR editors responsible for the journal evaluation have a degree in the particular field of science they work in. Since these people daily monitor journals, they have become experts in journals in their corresponding fields of science

Scientific article titles

Scientific article titles need to be informative:

- the commonly accepted abbreviations only can be used in article titles;

- no transliterations should be used in the titles of articles and abstracts translated into English; the exception is for untranslatable names, such as proper names, names of instruments, etc.; no untranslatable slang can be used. This also refers to the abstracts (summary) and key words.

Geographic aspect

If a journal is not an outstanding exception and its subject field is of interest only to a small region, it is unlikely to find itself in TR products. Experience demonstrates that the best articles written by authors from Third World countries have been published in international journals; therefore, if a choice is to be made out of two journals with the same specialization, internationally oriented journals are preferred. Of course, this policy creates additional difficulties for authors from Third World countries with international ambition. TR has always been blamed (and this has now been generally acknowledged) for the fact that the service has biased American and western European journals. E. Garfield attributed this to the overabundance of research carried out in the U.S. and Western Europe, the results of this research being published in English, German, and French. Special attention has always been on the analysis of Soviet journals. If an ISI analysis has shown that a certain Soviet journal was not covered for preparation of the SCI but was highly cited by the other source journals, this journal was covered in the processing. For example, this was the case with *Teplofizika (Thermal Physics)*. At time of writing, the

Web of Science indexes a total of 166 Russian journals; among them, 157 are covered in the Science Citation Index ; 6, in the Social Science Citation Index (two of those are actually American journals focused on the politics and economics of CIS, Russia, and China), and 3 Russian journals (*Voprosy Istorii* (Questions of History), *Voprosy Filosofii* (Questions of Philosophy), and *Sotsiologicheskie Issledovaniia* (Sociological Studies)), in the Art & Humanities Citation Index.

Depth of coverage of the subject field

The depth of the subject field coverage is one of the evaluation criteria when deciding whether a new journal should be covered by TR products. Based on private experience between the TR editor and a professional scientific community or publishing company, a decision for a new journal selection of launched by this scientific community or publishing company can be made. Unfortunately, good quality journals published earlier do not always guarantee that a new journal will be of a high quality, as well. It is natural to expect that a new journal published by a respectable publishing company with solid experience will be of high quality. However, it is rather frequent that publishing companies, pushed by members of a particularly interested group of specialists, begins premature publishing of a new journal. Moreover, the community of publishing companies is not monolithic with respect to publishing standards; there are significant discrepancies in the quality and periodicity of the publications. The same observations apply for a number of journals that are subsidized by the government or partially subsidized by other organizations. The fate of such publications may hinge on annual budget fluctuations.

Compliance of the journal with major standards

The timeliness of a publication is one of the criteria in the evaluation process. At the first stage of consideration of candidates for inclusion in the TR, the journal needs to be in line with the claimed periodicity. The ability of a journal to come out strictly on schedule means that the editorial board has a large hoard of unpublished articles in its portfolio. It is unacceptable for a journal to violate the established publication timeliness. In order to confirm the timeliness of publication, the editorial board needs to successively forward three current issues of the journal as soon as they are published to the TR company.

The TR company also takes notice of compliance with the International Publishing Convention, which optimizes facilities for searching for the source articles. They include such items as the informative title of the journal, illustrative representation of article titles and abstract, full bibliographic description of all the references contained in the articles, and full address data of each author and organization.

In the modern world, English is the universal language of science. For this reason, the journals published in English or at least having bibliographic descriptions and abstracts in English are considered first by TR. The WoS includes a sufficient number of journals with only bibliographic data published in English. However, it is obvious that the journals that will play the leading role in the international scientific community in the future will be entirely published in English. This is particularly valid for the field of natural sciences.

The peer review institute ensuring the general quality of the materials submitted and completeness of the cited references is another important indicator of a journal's standard.

Contents of the journal

As mentioned above, the core of scientific literature is the basis for all scientific disciplines. This core is not static; it changes along with the evolution of science. The emergence of new fields of knowledge results in the appearance of new journals, while the materials published in the new field of science attain a certain critical value. The task of the editors is to determine whether a new journal will enrich the products in which it should be included or this subject area is represented sufficiently well without it.

The TR editors have access to a vast amount of cited literature and a long experience in daily work with new journals. Therefore, they are capable of evaluating emerging scientific trends and active fields of research that had been published in journals, as well as deciding whether it is necessary to cover a new journal by WoS.

International diversity of the authors and the editorial board

TR editors pay attention to the international diversity of the authors, editors, and members of the Editorial Board. It is a significant factor for journals that are aimed to reach the international scientific community. The task of the company staff is to update the list of covered journals and to identify and evaluate new journals. Contemporary research is carried out all over the world; the international diversity of the journal is likely to be important for the global scientific community.

However, there is a large number of excellent regional journals that are intended mostly for the local audience rather than the international community. They usually are not characterized by a great variety of materials from different countries, and TR does not impose this requirement on such journals.

All regional journals selected to be covered by TR products need to

provide complete bibliographic information in English (article title, abstract, key words, authors' addresses) and undergo the review process. **All citations within the articles need to be presented in Latin letters.** An appreciably large number of free software programs for creating bibliographic descriptions using Roman alphabets can be found on the Internet. A Google search "create citation" will offer you several free software programs allowing automatic creation of citations in accordance with the specified standards (e.g., <http://www.easybib.com/>). It should be noted that descriptions can be created for different publication types (a book, a journal article, an Internet resource, etc.).

Citation analysis

Evaluating a journal is a unique process, since TR editors have an enormous collection of citation data. Attention should be put on the significance of the interpretation of these data and their comprehension. The quantitative citation indicators should be used only in the context of a journal that belongs to the same subject field. For example, there are not so many articles in the field of crystallography, and they contain a smaller number of citations as compared to those in the field of biotechnology or genetics. Or let us take another field, Art & Humanities, which is covered in the Art & Humanities Index (A&HI). The articles relating to this discipline require a considerably longer period of time to accumulate a significant number of citations. However, there is nothing unusual if an article from another field of science (e.g., life sciences) reaches its citation peak 2–3 years after publication. These facts need to be taken into account for a fair evaluation.

The citation indicators are evaluated at least at two levels. First, the citation indicators of the journal

are determined on the basis of the impact factor or the total number of references to this journal. Then, the citation indicators of individual authors are evaluated. This type of analysis is always useful, in particular when a new journal has yet to build a citation history.

This type of evaluation is performed for the journals that have not been covered by TR products and need a re-evaluation. It refers to the journals for which an increase in the number of citations has been observed due to various reasons: translation of articles into English, changes in the editorial policy, changes of publishing companies, etc. **Free access of TR employees to all literature cited (over 12,500 journals) means that the company has access to data pertaining to the number of citations for the journals both covered and not covered in its products.**

Self-citation

Self-citation is the number of references to the articles published in the same journal. For instance, if journal X was cited by all journals 15,000 times, including 2,000 citations in journal X, the share of self-citation will be $2 \times 100/15 = 13.3\%$.

All journals tend to publish articles in a certain subject field. Since scientific progress is based on preliminary studies, there is nothing unusual about some extent of self-citation of the journal. However, it is appreciably difficult to detect deliberate self-citing.

A high level of self-citation is not unusual for the leading journals in a certain field of knowledge, since high-quality papers or the ones devoted to a new, rapidly developing scientific discipline are being constantly published in these journals. In the ideal situation, the authors cite their previous studies, since they are the most relevant for the work carried out, which does not depend on the journal they would

like to be published in. However, self-citation in these journals may become dominant in the total pool of citations. Self-citation can potentially distort the true value (position) of a journal in a particular subject field.

Among all the journals covered by the JCR-Science Edition, 80% have a share of self-citation below 20%.

A significant deviation of this indicator from the normal value forces TR to re-examine the journal to determine what effect the excess self-citation had on the growth of the impact factor. If it is detected that self-citation was used improperly, the impact factor of this journal will not be published, and it will be considered reasonable to exclude the journal from Web of Science. There has been a myth in the scientific community that the editorial boards can manipulate the impact factor. Let us provide a curious example. The journal *The World Journal of Gastroenterology* had an excellent start. In 2000, its first IF was 0.993. During subsequent years, it rose to 1.445, 2.532, and 3.318. However, the journal was excluded from the JCR in 2004. The permanent increase in the IF was caused by self-citation rather than by the recognition of this journal by other authors. When calculating the IF, it was discovered that over 90% of the citations were self-citations. The journal was re-covered in the JCR in 2008; its IF was 2.081, the share of self-citation was as low as 8%.

In 2003, Marie McVeigh, Director of the JCR and Bibliographic Policy, examined the factors that cause high self-citation. It was found that self-citation is independent of the volume or subject field of the journal and is determined by the pattern of behavior of individual journals. The elimination of self-citation upon calculation of IF will have no effect on the ranking of most journals. When M. McVeigh detects journals whose IFs are mostly based

on self-citation, she acts as a parent would do with naughty children: she punishes them and gives them some time to “reform.”

The journals are given a short period of time (usually, several years), during which the editors are required to make aware their policy consequences. If they follow common norms and standards, they are re-covered in the JCR [12].

Addresses of organizations and authors

It is necessary to follow the existing rules of transliteration and use unified spelling (with Latin letters) of the authors' last names and organization names. These data are of great significance for collection of analytical statistics on the activity of organizations and individual researchers. In WoS, there is a special “author finder” section which contains data on organizations where a researcher has worked and to his or her publications. The Russian Academy of Sciences wastes a large number of publications because Russian researchers forget to mention the fact that their organization is affiliated with the RAS.

If we would like to bolster the statistical presence for the RAS in global databases, one needs to ask the editors of journals published by the RAS to use the same English version of the name of a RAS institute. Of course, the authors need to use the same organization name; otherwise, it will be difficult to determine their real productivity.

Experience in working with the WoS has demonstrated that Russian scientists sometimes forget to specify the country (“Russia”) in the “organization address” field.

TR company has developed a researcher identifier (RESEARCHER ID) in open access for all Internet users: <http://researcherid.com>.

With a unique identifier assigned to each author in ResearcherID, you can

eliminate author misidentification and view an author's citation metrics. One needs to visit the aforementioned website, input his or her data, and write the commonly accepted English-language name of the Institute you are working. The data pertaining to your identifier will be subsequently sent to you via e-mail. This will allow to automatically add your publications into your profile and save up to 10,000 bibliographic records obtained upon searching in the WoS.

Sources of financial support

The data on the sources of financial support, provided in the last section of the article, including grant numbers, are very important information enabling one to search by “funding agency” in WoS. Therefore, journal editorial board members should pay attention to the presence of such data.

Social sciences and the Social Science Citation Index

All social science journals undergo the same evaluation process as the natural science and engineering journals. The indicators considered are as follows: compliance of the journal with conventional standards, including the timeliness of publication of the journal; journal content; international authors; and the number of citations. The citation statistics takes into account the fact that the general citation score in the field of social sciences generally are considerably lower than those in the field of natural sciences and engineering. Special attention is paid to the journals that focus on regional studies, since they play a special role in the field of social sciences. It is the local specialization that is of interest for research.

Art & Humanities and the Art & Humanities Citation Index

The compliance with conventional standards (including the timely

publication) is important for Art & Humanities, as well. However, the citation patterns in the field of Art & Humanities do not always correspond to certain predictable models used in social and natural sciences and engineering. Moreover, articles in the field of Art & Humanities frequently cite books, music compositions, as well as literature and art works. The presence of the English text is not obligatory in certain fields of Arts and Humanities, in which it is not needed for the national specificity of the study object (e.g., studies devoted to regional literature).

Journal Website

Journal website is one of the indicators of its visibility on the Internet. It is reasonable to specify on the website which foreign and Russian s process the journal. Prior to designing the website, we recommend having a look at foreign analogues of your subject area. The **links** should be provided on both Russian- and English-language websites. Guidelines for authors for a manuscript submission should be timely updated and correspond to the facilities provided by new information technologies. The website should be constantly updated.

PROCEDURE FOR SUBMISSION OF A JOURNAL TO THE THOMSON REUTERS

When requesting inclusion of a journal in TR, one needs to double-check the following:

- Titles of the articles need to be informative – this requirement is considered by the expert system to be a major one;
- Only commonly accepted abbreviations can be used in article titles;
- No transliterations from the Russian language should be used in the titles of articles and abstracts translated into English; exception is made for untranslatable names,

such as proper names, instruments, etc.; no untranslatable slang known only to Russian-speaking specialists can be used;

- Each paper in the journal should be peer reviewed;
- Timeliness of the publications is mandatory;
- Each printed and/or electronic material should have an ISSN;
- A list of references using Latin letters should be provided;
- The journal title should have an English-language name;
- Publication year is mandatory;
- Volume and issue of the publication are required;
- An English title of the article should be provided;
- Either page numbers or article number (article number is not the DOI number). If a journal uses the numeration of both pages and arti-

cles, they should be listed individually, similar to that indicated in parentheses (art. № 23, pp. 6–10, not 23.6.–23.10);

- An English-language abstract should be provided for each scientific article;
- Authors' last names and their addresses should be provided, including the e-mail of the author responsible for the distribution of reprints;
- All article identifiers, such as DOI, PII, and other article numbers are needed;
- Complete content list of each issue should include data pertaining to the pages/number of each article (except when only one article was published in the journal);
- Information on the source of financial support and grant number (if any) is needed;

– Assignment of these identifiers both in the original articles and in the references helps in the use of the articles containing citations, as well as the cited ones, and proper identification by the abstracting and indexing services.

To submit a journal to the Web of Science, one needs to

1. Establish an ongoing, complimentary subscription to the title for Thomson Reuters;
2. Send several most current issues of the journal;
3. Forward each subsequent issue of the journal to the following address: Thomson Reuters ATTN: PUBLICATION PROCESSING, 1500 Spring Garden Street, Fourth Floor, Philadelphia, PA 19130 USA. ●

REFERENCES

1. Ivanov V.V., Varshavskii A.E., Markusova V.A. // Herald of the RAS. 2011. № 7. P. 587–593.
2. Halfman W., Leydesdorff L. Is inequality among universities increasing? Gini coefficients and the elusive rise of elite universities. www.loet@leydesdorff.net
3. Garfield E. Citation Indexes for Science // Science. 1955. T. 122. № 3159. P. 108–111.
4. Ginzburg V.L. Sami vinovaty? Pochemu Rossiia poluchaet malo Nobelevskikh premii (Themselves to Blame? Why did Russia Get a Few Nobel Prizes) // Poisk. 2007. № 47. P. 4.
5. Price D.J. de S. Little Science, Big Science. New York: Columbia University Press. 1963.
6. Garfield E. How ISI selects journals for coverage: Quantitative and Qualitative Considerations // Current Contents. 1990. № 22. P. 2–10.
7. Garfield E. A Century of Citation Indexing. Key note address // 12th COLLNET Meeting, September 20–23, 2011. Istanbul: Istanbul Bilgi University, 2011.
8. Garfield E. The Agony and the Ecstasy – The History and Meaning of the Journal Impact Factor // J. Amer. Med. Association. 2006. V. 295. № 1. P.90–93.
9. Mikhailov A.I., Chernyi A.I., Giliarevskii R.S. Nauchnye kommunikatsii i informatika (Scientific Communications and Informatics). Moscow: Nauka, 1976. 435 p.
10. Garfield E. Errors – theirs, ours, yours // Current Contents. 1974. V. 25. P.5–6.
11. Testa J. The Thomson Reuters journal selection process http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/
12. Gaming the Impact Factor Puts Journal In Time-out. www.scholarlykitchen.sspnet.org

An extended session of the Scientific Publishing Council of the Russian Academy of Sciences with representatives of Thomson Reuters headed by Vice President J. Testa was held on May 14, 2012. Mr. Testa delivered his report on journals selection policy to be covered by the Web of Science, and answered questions from the audience.

The meeting between J. Testa and the Vice Presidents of the Russian Academy of Sciences, A.I. Grigor'ev and S.M. Aldoshin, was held to discuss the results of the extended session of the Scientific Publishing Council. During the meeting, the subsequent joint collaboration between the RAS and Thomson Reu-

ters company was discussed in terms of the following directions:

- Possible expansion of coverage of Russian journals to be covered in the Web of Science;
- The results of the free-of-charge test access of the RAS institutes to the Web of Science;
- The international conference on scientometrics issues that is planned to be held by the Institute for the Study of Science jointly with the Thomson Reuters company on October 10–12, 2013; and
- Establishment of a working group to solve problems of bibliometrics and collaboration between the Russian Academy of Sciences and the Thomson Reuters company.

Publication Activity of Russian Researches in Leading International Scientific Journals

M. N. Kotsemir

National Research University Higher School of Economics, Institute for Statistical Studies and Economics of Knowledge, ul. Miasnitskaia, Moscow, 101000, Russia

E-mail: mkotsemir@hse.ru

Received 12.04.2012

ABSTRACT The primary aim of this research report is to analyse the dynamics and structure of the publications of Russian authors, as well as to define the place of Russian science in the global scientific process. Bibliometric analysis methods are the main methods for quantitative analysis of scientific cooperation, efficiency, and other aspects of scientific activity. The information base for this research includes materials from science citation databases containing bibliographic descriptions of the articles published in scientific journals (mainly written in English) in a significant number of fields of science. Various parameters (e.g. dynamics of the number of publications, the number of citation, the level of co-authorship, the scientific specialization index, etc.) at various levels of aggregation (e.g. individual researchers, research organizations, countries and regions of the world) can be calculated based on these data. The results of bibliometric studies can be used in a number of ways:

- analysis of latest trends in the development of various scientific fields;
- evaluation of the effectiveness of research organizations;
- overall assessment of the scientific potential of Russia (its strengths and weaknesses);
- identification the most productive scientists in various fields of science;
- drawing the international comparisons of publications;
- analysis of collaboration networks of scientific teams.

The paper analyses the basic indicators of the publication activity of scientists in Russia and the leading countries over the period between 2001 and 2011. Publication activity of Russian scientists is analysed in the context of specific areas of science. This allows the identification of areas of specialization of Russian publications. The paper also examines the dynamics of highly-cited publications and the indicators of the international scientific collaboration of Russian researchers. In this paper, materials of Web of Science database were used for analysis of publication activity.

KEYWORDS publication activity; bibliometric indicators; publications of Russian authors; the level of citation; cross-country analysis.

INTRODUCTION

The international and national science citation indices contained in the bibliographic descriptions of articles and reference lists are used in bibliometric analysis. Evaluation of the publication activity in Russian statistics practice began fairly recently [1–5]. Despite its rapid development, the Russian science citation index (RSCI) still fails to reflect the publication activity of Russian scientists adequately. Against this background, the materials of the science citation databases Web of Science (WoS) and the electronic analytical database Essential Science Indicators based on Web of Science were chosen as the infor-

mation base for this work. Eugene Garfield established web of Science (WoS) in 1964. Web of Science is the first science citation database in the world owned by the Thompson Reuters information corporation. As of 2011, the Web of Science comprises approximately 48 million entries of scientific publications in over 15 thousand scientific journals. The database includes material presented in more than 148 thousand scientific conferences. The database includes science publications dating back to 1900.

The Essential Science Indicators (ESI) in cross-country terms contains data on three basic publication activity indicators (the number

of publications (“papers” column in ESI database); the number of citations received by these publications (“citations” column in ESI database), and the average number of citations per paper (“Citations per paper” column in ESI database) for 144 countries over the past 10 years. In turn, this ten-year period is divided into five-year sub-periods (at the time of writing, the data for 2001–2005, 2002–2006, 2003–2007, 2004–2008, 2005–2009, 2006–2010, and 2007–2011 were available in ESI database). Total number of publication was grouped into the 22 fields of science (according to the classification of the Essential Science Indicators).

Table 1. Language structure of Russian papers: 2001–2011 (%)

Language	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
English	94.1	91.3	93.0	93.0	93.9	94.2	94.0	95.0	94.7	94.4	94.8
Russian	4.6	6.3	4.5	5.0	4.1	3.8	4.4	3.5	3.1	1.9	3.9
Other languages	1.3	2.5	2.5	2.0	2.0	2.0	1.6	1.5	2.2	3.7	1.3

Notes: 1. The language structure of Russian papers was calculated as the dynamics of the share of Russian papers written in languages specified in the table (English, Russian, other languages) in the total number of publications by Russian authors for each year specified in the table.

2. In order to obtain the primary data from the Web of Science database, the option “Analyze results” was used (for the selected country “Russia” and the selected time period “2001–2011”); the initial data were then loaded from the category “Languages.” The following types of documents were selected for the analysis: article, proceedings paper, and review.

Source: author calculations on Web of Science database materials. All databases of the Web of Science portal were used.

Studied sample includes countries with more than 10 000 publications in ESI database for the years 2001 – 2011. This sample consists of 57 countries. In certain cases, all the countries of the world will be under analysis.

A publication is affiliated to a certain country (Russia for example) if its author (or at least one of the co-authors) specified this country (Russia) in his or her working address. The terms “publications by Russian authors”, “Russian publications”, “publications of Russia” are used as synonyms. The following types of documents are classified under the term “publication”: “article”, “proceedings paper”, and “review”.

When performing a cross-country analysis of publication activity, one should take into account the fact that English-language publications currently dominate the science citation databases. English-language publications accounted for 94.8% of the total number of publications presented in the WoS database between 1990 and 2011. At the same time publications in the French, German, Russian, Japanese, and Chinese languages accounted for only 1.2, 1.1, 0.6, 0.22,

and 0.1%, respectively. Moreover, the profiles of the publication activity in such fields of science as mathematics, computer science, natural science, and engineering are not sufficiently complete in these databases with area of social sciences and the humanities being the least complete. These limitations should be taken into account when interpreting bibliometric data.

GLOBAL TRENDS IN PUBLICATION ACTIVITY

The basic indicator of publication activity is the number of publications in peer-reviewed scientific journals. When comparing the publication activities of different countries, the share of publications attributed to the countries in the total world number of publications result in the countries rank in the global ranking of the number of publications. As mentioned, the international citation databases contain a relatively small number of non-English-language publications. Consequently, the analysis of the publications of non-English-speaking authors will focus primarily on their English-language papers, which actually constitute a relatively small share of the total number of pub-

lications in these countries. For example, only 10% of all Russian publications are indexed in the Web of Science database [1]. Therefore, the indicators of publication activity are artificially underestimated for such countries as Russia, Japan, China, India, as well as for Southeast Asia, Latin America, and the Middle East countries.

Web of Science database integrates five science citation databases:

- Science Citation Index Expanded (SCI-EXPANDED);
- Social Sciences Citation Index (SSCI);
- Arts & Humanities Citation Index (A&HCI);
- Conference Proceedings Citation Index-Science (CPCI-S); and
- Conference Proceedings Citation Index-Social Science & Humanities (CPCI-SSH).

Data from all the databases within the Web of Science database are used for all the calculations, tables, and figures in this paper.

Data on scientific publications can be presented in the WoS database in different ways using the option “Analyze results”. *Table 1* shows the language structure of publications by Russian authors in

Table 2. Dynamics of Russian publications: 2001–2011

Year	Number of Russian publications	Shares of Russian publications in the total world number of publications, %
2001	28 665	2.97
2002	29 612	3.00
2003	28 648	2.75
2004	28 835	2.64
2005	28 281	2.45
2006	27 462	2.24
2007	28 926	2.16
2008	30 673	2.16
2009	30 904	2.09
2010	29 224	2.06
2011	28 573	2.12

Notes: The option “Analyze results” was used to obtain the primary data in the Web of Science portal (for the selected country “Russia”); the initial data were then loaded from the category “Publication years.” The following types of documents were selected for the analysis: article, proceedings paper, and review.

Source: author calculations on Web of Science database materials. All databases of the Web of Science portal were used.

scientific journals indexed in the Web of Science. The overwhelming majority (93–95%) of papers are English-language publications. Throughout the studied period, the share of Russian-language publications in the total number of Russian publications in scientific journals indexed in the Web of Science varies between 1.9 and 6.3%.

In 2011, the number of Russian publications in WoS slightly decreased in comparison with 2001 (*Table 2*). The minimal number of Russian publications over the studied was reached in 2006 (27,462 publications), followed by slight growth. The share of Russian publications in the global number of publications has dropped from 2.97% in 2001 to 2.12% in 2011.

Over the past decade, the U.S. led the absolute world leader on the number of publications in WoS (*Table 3*). Over the period of 2001–2011, Russia has lost six positions in this ranking, dropping from 9th to 15th place, overtaken by Spain, Australia, South Korea, Brazil, and the

Netherlands. The positions of Canada and Western European countries remained almost unchanged.

Iran showed the highest growth of number of publications among the studied countries: by a factor of 11.5 over 2001–2011 (*Table 4*). Malaysia, Pakistan, China, Saudi Arabia, Thailand, and Turkey also showed high increase (more than 200% over the studied period) in the number of number of publications in WoS. Among European countries, Portugal showed the highest increase in the number of publications (149% over the 2011–2011). On the other hand some countries showed a decrease in the number of publications in scientific journals indexed in WoS in 2011 in comparison with 2001: Japan (by 7.4%), Venezuela (by 8.2%), and Belarus (by 15.9%).

Countries with a rapidly increasing publication activity have improved their positions in the global ranking of the number of publications. Iran’s position improved by 23 places (jump from the 42nd place in

2001 to the 19th place in 2011). Malaysia improved its position by 15 ranks (from the 45th place in 2001 to the 34th place in 2011). Some other countries (mainly from Asia and Latin America) also significantly improved their position in this ranking: Pakistan by 11 positions; Portugal by 9 positions; Turkey by 7 positions; Columbia by 6 positions; Saudi Arabia, Thailand, Brazil, and China by 4 positions each. In contrast, some countries (mainly from Eastern Europe) have lost many positions in the global ranking of number of publications. The following countries from the studied sample have lost six and more positions: Russia and Venezuela (6 positions), Hungary, Bulgaria and Slovakia (7 positions), Belarus (8 positions), and Ukraine (12 positions).

The dynamic growth of publication activity in developing countries led to a weakening of the dominance of the U.S. in the world science. In 2001, the share of the USA in the total world number of publications was 31.5%; in 2011, it decreased to the level of 27.1%. Meanwhile, the share of the second country in the rating increased from 8.9% in 2001 (in 2001, the second position was held by Japan) to 13.6% in 2011 (in 2011, the second position was held by China).

The data on the publication activity of different countries in the electronic analytical database Essential Science Indicators (ESI database) developed on Web of Science materials are presented in the “Countries/territories” subsection in the “Citation Rankings” section (option “View table of graph data”).

The following indicators are provided for a specified country for five-year periods as well as for the whole period of 2001–2011 (option “View table of graph data”):

- 1) Number of papers;
- 2) Number of citations;
- 3) Average citations per paper.

Table 3. Top 25 countries in the world rating for the number of publications

№	2001			2011		
	Country	Number of publications by the country	Share of the country in the total world number of publications, %	Country	Number of publications by the country	Share of the country in the total world number of publications, %
1	USA	303 917	31.48	USA	366 507	27.13
2	Japan	86 096	8.92	China	184 029	13.62
3	Great Britain	83 582	8.66	Great Britain	105 411	7.80
4	Germany	77 982	8.08	Germany	97 070	7.19
5	France	55 259	5.72	Japan	79 751	5.90
6	China	44 575	4.62	France	67 990	5.03
7	Canada	38 645	4.00	Canada	58 855	4.36
8	Italy	38 453	3.98	Italy	55 253	4.09
9	Russia	28 667	2.97	Spain	50 256	3.72
10	Spain	26 350	2.73	India	46 172	3.42
11	Australia	25 483	2.64	South Korea	45 971	3.40
12	The Netherlands	21 779	2.26	Australia	44 244	3.28
13	India	19 272	2.00	Brazil	34 122	2.53
14	South Korea	19 194	1.99	The Netherlands	33 523	2.48
15	Sweden	17 422	1.81	Russia	28 577	2.12
16	Switzerland	15 566	1.61	Taiwan	28 553	2.11
17	Brazil	13 324	1.38	Switzerland	24 655	1.83
18	Taiwan	13 018	1.35	Turkey	23 470	1.74
19	Poland	12 824	1.33	Iran	21 768	1.61
20	Belgium	11 964	1.24	Sweden	21 389	1.58
21	Israel	10 836	1.12	Poland	20 818	1.54
22	Finland	8 822	0.91	Belgium	18 686	1.38
23	Austria	8 779	0.91	Denmark	13 468	1.00
24	Denmark	8 754	0.91	Austria	12 852	0.95
25	Turkey	7 233	0.75	Israel	12 493	0.93

Notes: 1. In order to obtain the data on the total world number of publications in the Web of Science portal, the retrieval request "PY=2001 AND 2011" was input in the "Advanced search" insertion; the option "Analyze results" was then used to load the initial data from the category "Document types." The following types of documents were selected for analysis: article, proceedings paper, and review.

2. The option "Analyze results" was used to obtain the data on the number of publications in the countries studied on the Web of Science portal (for 2001 and 2011); the initial data were then loaded from the category "Countries/territories." The following types of documents were selected for analysis: article, proceedings paper, and review.

3. The sum of the shares for all countries is higher than 100%, since some papers were prepared in international co-authorship.

Source: author calculations on Web of Science database materials. All databases of the Web of Science portal were used.

Table 4. Country composition of the studied sample with the most dynamically increasing number of publications: 2001–2011

Country	Number of publications		The number of publications increased in 2001–2011 by a factor of
	2001	2011	
Iran	1 891	21 768	11.51
Malaysia	1 216	8 713	7.17
China	44 575	184 029	4.13
Pakistan	739	5 682	5.47
Saudi Arabia	1 451	5 651	3.89
Thailand	1 727	5 991	3.47
Turkey	7 233	23 470	3.24

Notes: In order to obtain the primary data from the Web of Science portal, the option "Analyze results" was used for the countries specified in the table; the initial data were then loaded from the category "publication years." The following types of documents were selected for the analysis: article, proceedings paper, and review.

Source: author calculations on Web of Science database materials. All databases of the Web of Science portal were used.

In essence, the latter indicator is the ratio between the second and the first indicators. All three indicators can be calculated both for each of the 22 fields of science specified by WoS and for the total number of publications ("All fields" category). The same indicators are available for the total world number of scientific papers. The data on global indicators of publication activity are presented in the "Baselines" subsection in the "Citation Analysis" section ("View field rankings table" option). Similar to the case of individual countries, the global indicators of publication activity can be calculated both for the 22 fields of science and for the total number of publications for the specified five-year periods and for the whole period of 2001–2011.

It should be noted here that the total number of publications of a certain country during a certain five-year period can be obtained

via the summing up the numbers of publications (and citations) of all 22 fields of science. In other words, the subjects of 22 fields of science specified in the Essential Science Indicators do not overlap. Meanwhile, the global number of publications (and citations) obtained via the summing up of publications of all countries exceeds the real total world number of publications (and citations), due to the fact that some papers result from international collaboration. Subsequently, these publications are assigned to all the countries whose authors participated in these publications. Therefore, by summing up the number of publications (and citations of these publications) of all countries, some papers (and citations) will actually be accounted several times.

The thematic structure of Russian publications in WoS differs from the global one (Table 5). Physics is the most significant field of

science in Russia. Publications on physics accounted for 28% of the total number of Russian publications in 2001–2011. The second and the third most significant fields of science are chemistry (22.2%) and engineering (7.6%).

Clinical medicine is the most significant field of science in the world structure of publications: its share in the total number of scientific publications was 21.0% in 2011–2011. The second and the third most significant fields of study are chemistry (12.0%) and physics (9.3%). The share of Russian publications on clinical medicine, computer and social sciences in total world number of publications was very small. No significant changes in the thematic structure of Russian publications were observed during the studied period. Nevertheless, it should be noted that in the structure of Russian publications the share of publications on pharmacology & toxicology, economics & business, multidisciplinary studies, and environment/ecology considerably increased.

By comparing the thematic structure of scientific publications in a specific country with the same global structure, it is possible to calculate the index of scientific specialisation of the country [1]. This index is calculated as the ratio between the shares of publications in the field of studies i , within the total number of publications of country j , and the same indicator for the global structure of publications. If this index is greater than one in a certain discipline, thus this discipline belongs to the sphere of scientific specialization sphere of that particular country.

The main areas of scientific specialization of Russian science are physics, space science, and geosciences (Fig. 1). Significant areas of specialization include mathematics, chemistry, and materials science. Publications on pharmacol-

Table 5. Distribution of publications over fields of science (%)

Field of science	World structure		Russian structure	
	2001–2005	2007–2011	2001–2005	2007–2011
Biology & biochemistry	6.08	5.17	4.01	3.48
Immunology	1.33	1.14	0.16	0.20
Clinical medicine	20.79	21.24	4.70	5.09
Computer science	2.97	2.09	1.20	0.84
Mathematics	2.47	2.81	4.42	5.46
Materials science	4.60	4.70	6.25	6.01
Microbiology	1.66	1.69	1.27	1.20
Molecular biology and genetics	2.83	2.78	2.12	2.24
Multidisciplinary	0.17	0.17	0.07	0.13
Geosciences	2.83	2.92	7.11	8.08
Space science	1.31	1.15	3.31	3.23
Plant & animal science	5.56	5.42	2.29	2.60
Environment/ecology	2.57	2.88	0.90	1.50
Neuroscience & behaviour	3.16	2.85	0.78	0.78
Social science	4.06	5.12	1.08	0.95
Psychiatry/psychology	2.38	2.49	0.50	0.44
Agricultural science	1.85	2.30	0.70	0.77
Engineering	8.02	8.66	7.95	7.27
Pharmacology & toxicology	1.79	1.96	0.19	0.46
Physics	9.71	8.98	28.29	27.34
Chemistry	12.31	11.62	22.58	21.76
Economics & business	1.54	1.84	0.10	0.18

Notes: The distribution of publications over fields of science was calculated as the dynamics of the shares of the publications in the specified fields of science in the total number of publications.

Source: author calculations on of Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories" country "Russia," option "View table of graph data" for Russian indicators ; section "Citation Analysis," subsection "Baselines," option "View field rankings table" for the total world indicators).

ogy and toxicology, economics and business, and multidisciplinary studies showed the most significant increase in the index of scientific specialization in 2007–2011 in comparison with 2001–2005. On the contrary, index of scientific specialization for publications on social sciences decreased dramatically.

Russian authors have contributed substantially to the total global number of scientific papers (*Fig. 2*) on physics (7.9% of the total world number), space science (7.1%), and geosciences (6.7%). Russian publications on chemistry, mathematics, and materials science were also significantly presented in the world science: they accounted to

3–5% of the total world number in 2001–2011. As previously mentioned, Russia has lost positions in world science over the period of 2001 – 2011. Let us analyse this phenomenon in the context of fields of science. In 2007–2011 in comparison with 2001–2005 Russian publications on social science, psychiatry & psychology, materi-

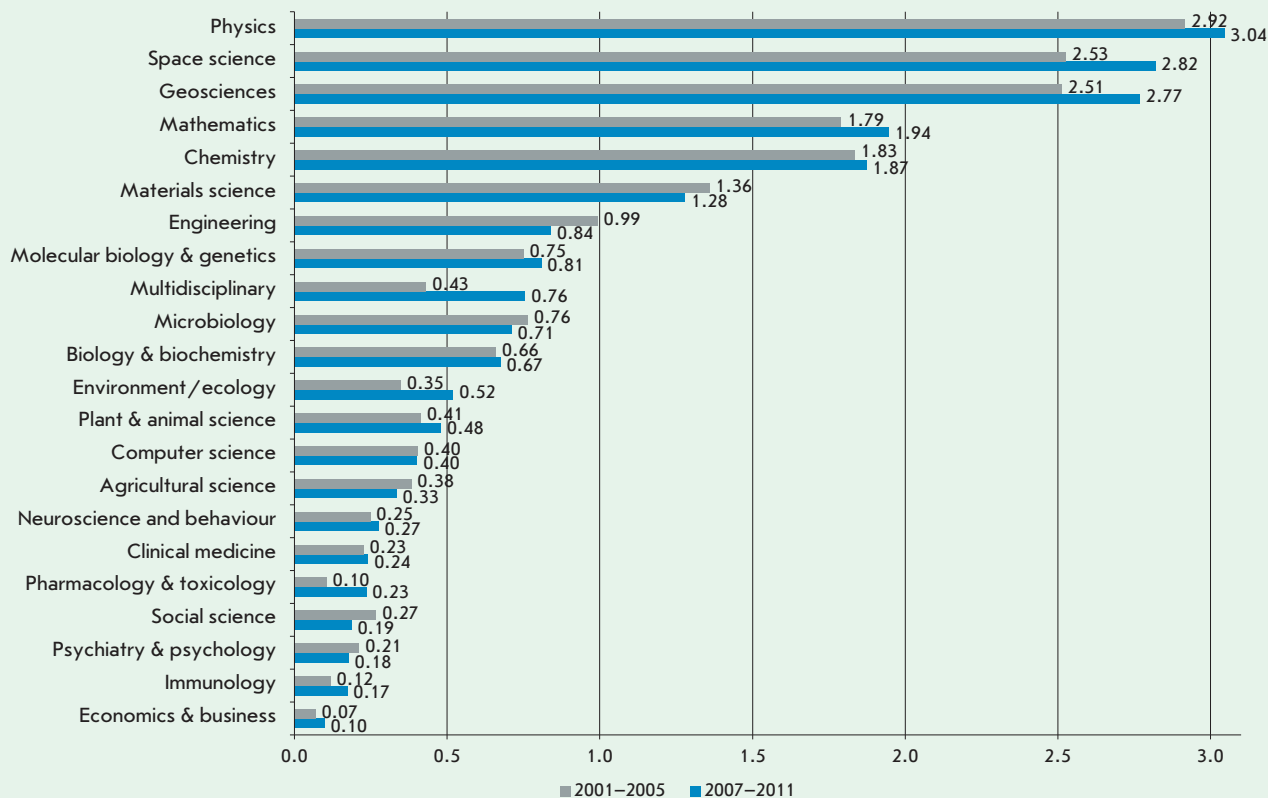


Fig. 1. Indices of the specialization of Russian publications in different fields of science.

Source: author calculations on Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories" country "Russia," option "View table of graph data" for indicators for Russia; section "Citation Analysis," subsection "Baselines," option "View field rankings table" for the total world indicators).

als science, and engineering have lost their positions in the world science more dramatically than publications on other field of science. The share of Russian publications in the other areas of its scientific specialization has also decreased. Meanwhile, Russia has strengthened its position in the fields of multidisciplinary studies, as well as pharmacology & toxicology, over the studied period.

MAJOR TENDENCIES IN PUBLICATION CITATION

Citation indices are used to evaluate the scientific impact of the publications of a certain researcher or a country in general. The average number of citations per publication is the basic indicator, which is de-

termined as the ratio between the number of citations on the publications published by a specific country and the number of these publications in scientific journals reviewed by the science citation database over a certain period. This indicator is typically assessed for a certain period (usually, 3–5 years) rather than for one year. The reason can be explained as follows: some period is necessary before a publication included in the international citation database can accumulate a significant number of citations. Let us note that in the context of this work indicator "average number of citations per publication" refer three kinds of scientific documents: article, proceedings paper and review.

The following procedure is used in the Essential Science Indicators database to calculate the average number of citations per paper. There are two basic indicators:

- total number of publications ("Number of papers" indicator in ESI database) of a specific country published during the specified 5-year period (2001–2005 for example) in scientific journals, indexed by Web of Science ("Number of papers" indicator in ESI database);
- number of citations on this publications during this (2001–2005) 5-year period ("Number of citations" indicator in ESI database).

Therefore, the average number of citations per paper (ACP further) indicator is simply the ration of B and A indicators. Such an approach

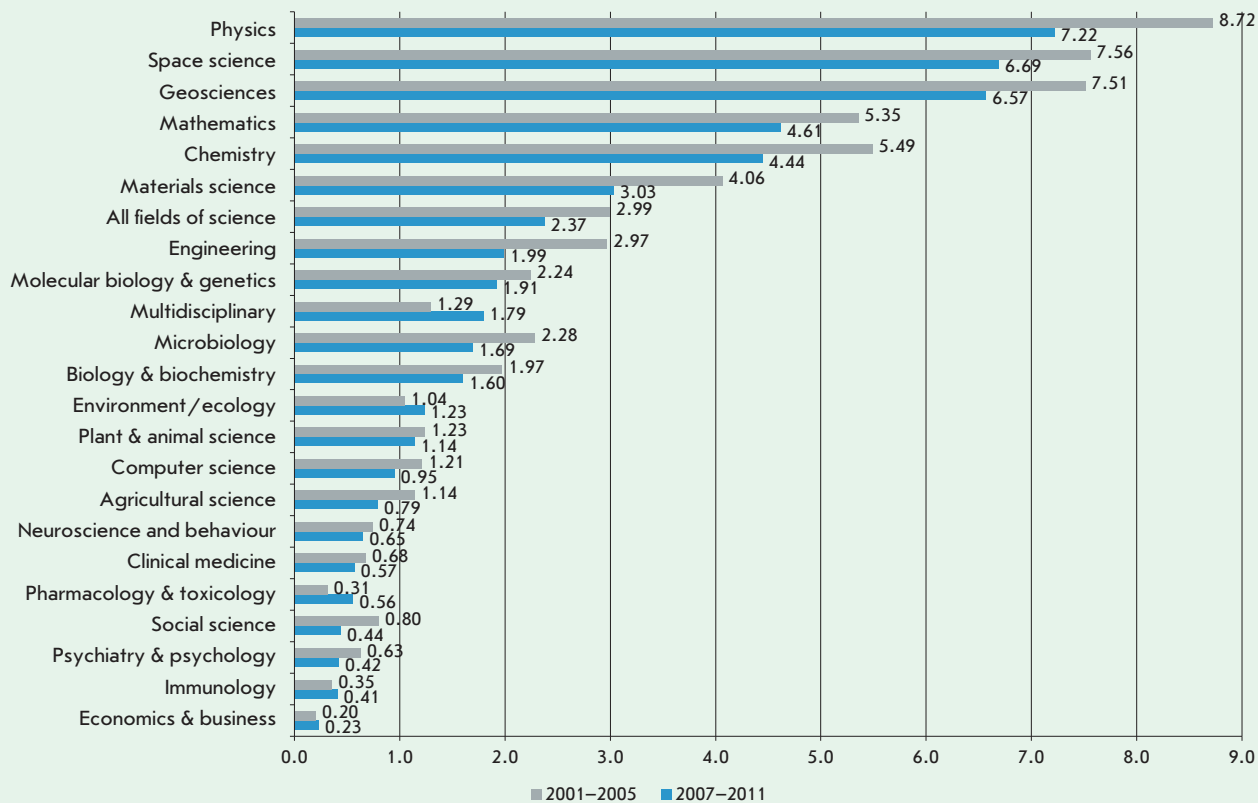


Fig. 2. Shares of Russian publications in the total world number of publications on fields of science (%).

Source: author calculations on Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories" country "Russia," option "View table of graph data" for Russian indicators; section "Citation Analysis," subsection "Baselines," option "View field rankings table" for the total world indicators).

to calculation the ACP indicator allows cross-time evaluation of average level of citation of publications.

Analysis of citation indicator should be performed with caution. The number of citations does not always adequately represent the quality of publications, since opportunities to be cited are not equal for authors of different countries. So-called "Matthew effect that has been repeatedly observed in different fields of science [6–10] should be mentioned here. The Matthew effect was first mentioned in 1968 in the article of the American sociologist Robert Merton in *Science* [6, 7]. The researcher analysed the psychosocial factors affecting the recognition and evaluation of scientific papers. The Matthew ef-

fect means that scientific research carried out by famous authors is somehow superior to that made by their less outstanding colleagues. The effect was named in light of the flowing quote from the Gospel According to St. Matthew: "For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken even that which he hath" (Matthew 25:29, King James Version). In context of countries Mathew effect means that publication from "prestigious" (e.g. from the USA) country has, other thing equal, (it is assumed that the article from the U.S. author and the article from Tajik author are devoted to the same problem, have the equal scientific level and are written in

the same language with the same level of the grammatical correctness), more opportunities than the publication from "unprestigious" (e.g. from Tajikistan) country.

The countries with a small number of publications in the international science citation databases often have extremely high level of the ACP indicator. Let us use the following example to illustrate. The Bermuda Islands, Panama, and Gambia held the first three positions in the ranking of the ACP indicator level over the period of 2001–2011, respectively. The value of the ACP indicator in these countries was 23.7, 17.9, and 17.0 citations per paper, respectively. Gabon and Guinea-Bissau were also among the top 20 countries of this

Table 6. The average number of citations per paper in the leading countries: 2001–2011

№	Страна	Average number of citations per paper for the country	Number of publications of the country	Position of the country in the ranking of the number of publications
1	Switzerland	16.61	194 618	16
2	USA	15.83	3 219 337	1
3	Denmark	15.83	104 212	23
4	The Netherlands	15.53	268 385	14
5	Sweden	14.82	189 413	18
6	Great Britain	14.79	912 495	2
7	Belgium	13.69	147 261	21
8	Finland	13.59	94 209	25
9	Germany	13.20	836 694	4
10	Canada	13.15	479 354	7
11	Austria	12.92	102 129	24
12	Israel	12.72	117 251	22
13	Norway	12.56	77 118	29
14	France	12.32	598 138	6
15	Australia	11.97	323 344	10
16	Italy	11.81	458 871	8
17	Ireland	11.51	49 358	37
18	New Zealand	10.84	61 205	34
19	Spain	10.42	364 197	9
20	Japan	10.35	815 789	5

Notes: 1. The top 20 countries from the studied sample in terms of the average number of citations per paper are considered.

2. All indicators were calculated for the total number of publications over the period of 2001–2011 in scientific journals indexed in the Web of Science.

3. The average number of citations per paper is determined as the ratio between the number of citations received in 2001–2011 by the publications of the country published in 2001–2011 to the number of these publications.

Source: author calculations on Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories," option "View table of graph data" for indicators of the countries listed in the table; section "Citation Analysis," subsection "Baselines," option "View field rankings table" for the total world indicators).

ranking. However, in all these countries with the exception of Panama, the total number of publications in scientific journals indexed by WoS over the period of 2001–2011 was below 1,000 (in Panama, the number of papers was 2,098).

According to ESI data, Russia held the 123rd position (out of 144) in the world rating on the ACP indicator level. The average number of citations received by a Russian paper published over the period of 2001–2011 was relatively low: 4.87,

much lower than the global average of 10.57 citations per paper. Despite this, Russia held the relatively high position of 21st in the ranking of the absolute number of citations received by papers.

Table 6 lists the first 20 countries for the level of the ACP indicator within the studied sample. English-speaking countries and Northern European countries hold the first 12 positions in this ranking, whereas Russia holds the 47th position out of 57 countries.

The value of the ACP indicator in all countries characterized with rapidly growing publication activity was lower compared to the global level. Among these countries, Thailand stood had the highest value of the ACP indicator: 7.57 citations per paper (the 72nd position in the world). The other countries characterized by a rapid increase in the number of scientific papers in the leading world journals held positions below the top 100 in the global

Table 7. Top 25 countries of the world ranking of the number of citations of publications

№	2001–2005			2007–2011		
	Country	Number of citations of the publications of the country	Share of the country in the total world number of publication citations, %	Страна	Number of citations of the publications of the country	Share of the country in the total world number of publication citations, %
1	USA	8 736 259	47.5	USA	11 542 290	42.2
2	Great Britain	2 231 223	12.1	Great Britain	3 330 285	12.2
3	Germany	1 900 402	10.3	Germany	2 788 268	10.2
4	Japan	1 576 262	8.6	China	2 219 953	8.1
5	France	1 244 048	6.8	France	1 852 765	6.8
6	Canada	971 332	5.3	Japan	1 840 922	6.7
7	Italy	870 611	4.7	Canada	1 641 349	6.0
8	The Netherlands	650 939	3.5	Italy	1 462 765	5.3
9	Australia	572 221	3.1	Spain	1 111 348	4.1
10	China	569 874	3.1	The Netherlands	1 086 107	4.0
11	Spain	549 353	3.0	Australia	1 071 029	3.9
12	Switzerland	529 890	2.9	Switzerland	857 170	3.1
13	Sweden	471 150	2.6	South Korea	677 451	2.5
14	Belgium	308 583	1.7	Sweden	666 464	2.4
15	South Korea	288 106	1.6	India	587 965	2.1
16	Denmark	265 646	1.4	Belgium	551 464	2.0
17	Israel	262 033	1.4	Brazil	436 681	1.6
18	Russia	258 172	1.4	Denmark	426 175	1.6
19	India	225 529	1.2	Taiwan	413 885	1.5
20	Finland	217 603	1.2	Austria	347 819	1.3
21	Austria	215 013	1.2	Israel	338 029	1.2
22	Brazil	185 243	1.0	Russia	317 770	1.2
23	Poland	178 917	1.0	Finland	309 117	1.1
24	Taiwan	173 626	0.9	Poland	302 810	1.1
25	Norway	140 394	0.8	Turkey	267 440	1.0

Notes: 1. The indicator "number of citations per publications of the country" is the number of citations received over the period of 2001–2005 (and 2007–2011) by the publications of the country in 2001–2005 (and 2007–2011).

2. The total world number of citations received by the publications is calculated for the actual number of the scientific publications of all the countries present in the Essential Science Indicators (section "Citation Analysis," subsection "Baselines," option "View field rankings table").

Source: author calculations on Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories," option "View table of graph data" for indicators for the countries listed in the table; section "Citation Analysis," subsection "Baselines," option "View field rankings table" for the total world indicators).

ranking of the value of the ACP indicator. Iran, the country with the most dynamic increase in the absolute number of papers among the countries of the studied sample, had one of the lowest indi-

cators of the average number of citations per paper in the world – 4.19 citations per paper (133rd position in the world rating). The value of the ACP indicator in European countries was higher than

in Asian countries. Japan had the highest average number of citations per paper among Asian countries: 10.35 citations per paper and the 35th position in the world rating for this indicator).

Among the BRIC countries, Brazil had the maximum level of the ACP indicator: 6.37 citations per paper (100th position in the global ranking). In China and India, the average number of citations per paper was also higher than in Russia: 103rd and 108th positions in the world, respectively (6.15 and 5.87). Estonia stands out among the countries of the former Soviet Union for the highest value of the ACP indicator: 9.35 citations per paper (45th position in the world). In all Baltic countries, the average number of citations per paper was higher than in Russia. On the contrary, the average number of citations per paper in Central Asian countries, as well as in Belarus and Ukraine, was lower than that in Russia. Azerbaijan had the lowest value of the ACP indicator among the former USSR countries: 2.71 citations per paper, 143rd position in the global ranking. Serbia held the lowest 144th position in the world rating on the value of the ACP indicator with 2.66 citations per paper.

As mentioned previously, the analytical database Essential Science Indicators enables one to dynamically compare the average number of citations per paper and the absolute number of citations received by publications. Among the countries of studied sample, the most significant growth (by more than 50%) in the average number of citations per paper was shown by such countries as Singapore (89.7%), Tunisia (68.5%), Iran (61.6%), China (60.4%), Egypt (57.9%), Algeria (57.4), Greece (55.8%) and Turkey (50.3%).

The U.S. was the stable global leader on absolute number of citations received by publications over the studied period. The dominance of the U.S. over the other countries in terms of this indicator is even more pronounced than that in terms of the number of publications (Table 7). The papers published by U.S. authors during the period of

2001–2011 received 45.3% of the total world number of citations. This indicator fell slightly from 47.5% to 42.2% over the studied period. The shares of Great Britain and Germany, which held the second and third positions, respectively, in the ranking of the number of citations over the period of 2001–2011, remained almost unchanged over the specified period. Russia held the 22nd position in this ranking. During the specified period, Russia lost by four positions dropping from 18th to 22nd.

China, which held the second position in the ranking of number of publications, ranked 7th in the ranking of the number of citations (5% of the total world number of citations). Over the period under analysis, China improved its rank by six positions: jumping from 10th to 4th place. In addition to China, some other countries with dynamic publication activity also improved their positions in this ranking: Iran (10 positions), Pakistan (8 positions), Malaysia (7 positions), China (6 positions), Brazil, Singapore, and Taiwan (5 positions each). Among the countries of the studied sample under more than threefold increase in the number of citations on publications issued over the period of 2001–2011 was observed in Iran (7.4 times), Pakistan (5.9 times), Malaysia (4.8 times), China (3.9 times), Algeria (3.7 times), Thailand (3.4 times), Nigeria (3.2 times), and Columbia (3.1 times). However, a dramatic increase in the absolute number of citations on publications did little to help these countries to score high citation indices.

The following countries lost positions in the ranking of the number of citations on publications: Russia, Estonia and Israel lost four positions each; New Zealand and Slovakia five positions each; Ukraine lost six positions, Hungary and Venezuela both lost eight positions. As in the case for the number of publications, the positions of the North European

and North American countries in the ranking of the number of citations remained almost unchanged over the period of 2001–2011. Despite the loss of position in the ranking of the number of citations on publications, Hungary held the relatively high 38th position in the rating for the level, in contrast to Venezuela, which held the 84th position. Israel, New Zealand, and Estonia also held relatively high positions 17th, 28th and 45th, respectively.

The distribution of citations on Russian publications over the fields of science for Russian papers, as well as the structure of papers, strongly differed from the world distribution (Table 8). 37.5% of all the citations over the period of 2001–2011 were received by Russian publications on physics. The share of the remaining fields of science in the total number of citations on publications by Russian authors was considerably lower. Nevertheless, the share of citations received by Russian publications on physics decreased in 2007–2011 as compared to that in 2001–2005. Meanwhile, the share of citations received by Russian publications on clinical medicine substantially increased from 3.8% to 7.8%.

The field of clinical medicine was dominant in the world structure of citations, accounting for 25.2% of the total number of citations. Chemistry and physics held the second and third positions in terms of the number of citations received (12.1% and 8% of the total number of citations, respectively). The share of citations received by the publications on biology & biochemistry decreased most significantly in the world structure of science over the period of 2001–2011.

The shares of citations on Russian publications in the total world number of citations for different fields of science are presented in Fig. 3. Russian publications on physics and astronomy received

Table 8. Distribution of the citations on publications over fields of science (%)

Field of science	World structure		Russian structure	
	2001–2005	2007–2011	2001–2005	2007–2011
Biology & biochemistry	10.47	7.63	6.66	5.54
Immunology	2.98	2.29	0.36	0.59
Clinical medicine	25.56	25.09	3.79	7.81
Computer science	0.82	0.91	0.32	0.27
Mathematics	0.68	0.94	1.36	1.86
Materials science	2.46	3.75	3.26	3.78
Multidisciplinary	0.07	0.12	0.01	0.20
Microbiology	2.64	2.35	1.59	1.52
Molecular biology and genetics	7.82	6.20	3.60	3.33
Geosciences	2.19	2.58	4.93	6.48
Space science	2.02	1.81	5.18	5.61
Plant & animal science	3.69	3.69	1.59	1.94
Environment/ecology	2.16	2.98	0.96	1.42
Neuroscience & behaviour	5.68	4.79	1.33	1.17
Social science	1.48	2.10	0.19	0.27
Psychiatry/psychology	2.00	2.36	0.25	0.21
Agricultural science	1.04	1.46	0.20	0.39
Engineering	2.92	4.37	4.90	4.39
Pharmacology & toxicology	2.05	2.29	0.31	0.52
Physics	8.60	7.67	43.49	35.49
Chemistry	12.09	13.71	15.65	17.15
Economics & business	0.58	0.91	0.08	0.09

Notes: The distribution of the publication citations over fields of science is calculated as the dynamics of the shares of citations for publications in the specified fields of science in the total number of citations on publications.

Source: author calculations on Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories" country "Russia," option "View table of graph data" for Russian indicators; section "Citation Analysis," subsection "Baselines," option "View field rankings table" for the total world indicators).

the largest share of the total world number of citations in comparison to the papers from other fields of science (the total being 6.1% over the period of 2001–2011). Russian publications on mathematics, geosciences, and space science received 2%–4% of the total world number of citations. The share of citations on Russian publications on physics and engineering in the total world number of such citations seriously decreased over the period of 2001–2011. On the other hand, this indicator for publications on multi-

disciplinary studies increased considerably (from 0.3 to 1.9%). *Table 9* lists the values of the ACP indicator for Russian publications from different fields of science over the period of 2001–2011. Russian publications on immunology had the highest average number of citations per paper. However, the world level of the ACP indicator for publications on immunology was higher than the Russian one.

Russian publications on mathematics, social and computer sciences had the lowest value of the ACP

indicator in comparison with publications from other fields of science over the period of 2001–2011. In 2001–2005, the value of the ACP indicator for Russian publications on economics & business was comparable to the corresponding world indicator. However, the level of ACP indicator of Russian publications from these fields, as well as those from the field of pharmacology & toxicology, decreased in 2007–2011 by 28% and 19%, respectively.

Russian publications on multidisciplinary studies showed quick

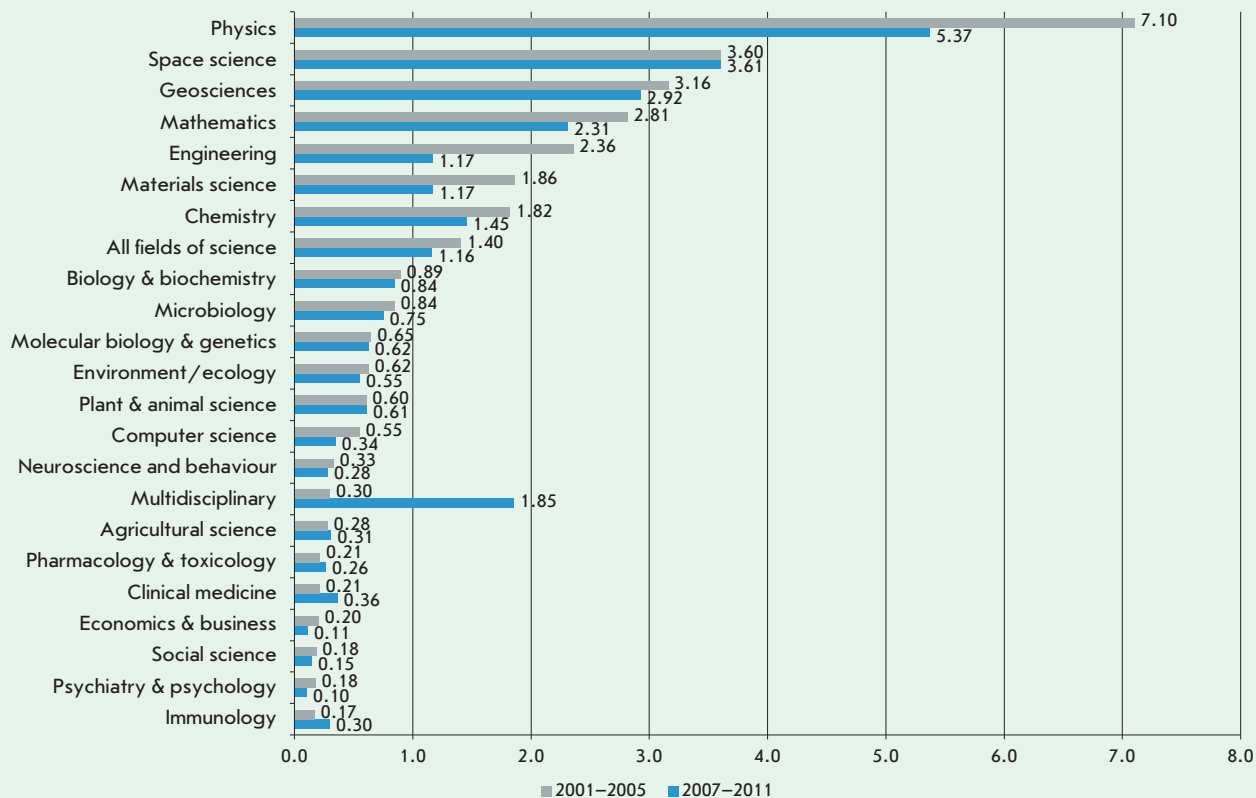


Fig. 3. Share of citations of Russian publications in the total world number of citations (%).

Source: author calculations on Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories" country "Russia," option "View table of graph data" for indicators for Russia; section "Citation Analysis," subsection "Baselines," option "View field rankings table" for the total world indicators).

growth of the value if the ACP indicator over the studied period: from 0.39 in 2001–2005 to 3.65 in 2007–2011. In 2007–2011, only the publications on multidisciplinary sciences had the level of the ACP indicator comparable to the world indicator. Russian publications on clinical medicine also showed relatively strong (by a factor of 2.28) increase of the ACP level over the studied period. In the world structure of science, publications on molecular biology & genetics had the highest value of the ACP indicator.

DYNAMICS OF HIGHLY-CITED PUBLICATIONS

Highly-cited publications are a relatively small group of the most

influential scientific publications. A publication in a certain field of science is considered to be highly cited if it is among the 1% of the most frequently cited publications in this field of science. Since the citation trends (frequency and time distribution of the citations) vary significantly in different fields of science, and the earlier published articles are cited more frequently than newly published ones, the distribution of citations over the years and individual fields of science are taken into account when determining the highly cited papers.

The highly-cited publications are frequently a result of international collaboration, which may bring together authors from different

countries. Many of them form the so-called research fronts, the most topical and rapidly developing areas of research and developments in the world. The highly cited publications characterize the science system of a certain country to some extent [11, 12].

The index of scientific specialization for the highly-cited publications is calculated using the same procedure as for "usual" publications.

The dynamics of highly-cited Russian publications is shown in Fig. 4. Their distribution over the fields of science is listed in Table 10. After a significant reduction in 2009, the number of highly cited papers in Russia increased to an even higher extent in 2010.

Table 9. The average level of citations of Russian publications: distribution over fields of science: 2001–2011

Field of science	2001–2005		2007–2011	
	Average number of citations per publication by Russian authors	The ratio between the average number of citations per publication by Russian authors and the world indicator	Average number of citations per publication by Russian authors	The ratio between the average number of citations per publication by Russian authors and the world indicator
Immunology	4.6	0.48	7.21	0.73
Space science	3.15	0.48	4.18	0.54
Biology & biochemistry	3.35	0.45	3.83	0.53
Clinical medicine	1.62	0.31	3.69	0.64
Multidisciplinary	0.39	0.23	3.65	1.04
Neuroscience & behavior	3.41	0.44	3.6	0.44
Molecular biology & genetics	3.42	0.29	3.58	0.33
Physics	3.1	0.81	3.12	0.74
Microbiology	2.52	0.37	3.03	0.44
Pharmacology & toxicology	3.35	0.68	2.71	0.47
Environmental science	2.16	0.60	2.28	0.45
Geosciences	1.39	0.42	1.93	0.44
Chemistry	1.39	0.33	1.9	0.33
Plant & animal science	1.39	0.49	1.8	0.54
Materials science	1.05	0.46	1.51	0.39
Engineering	1.24	0.79	1.45	0.58
Agricultural science	0.58	0.24	1.22	0.39
Economics & business	1.6	0.99	1.15	0.48
Psychiatry & psychology	1.01	0.28	1.13	0.24
Mathematics	0.62	0.53	0.82	0.50
Computer science	0.53	0.45	0.78	0.36
Social science	0.35	0.22	0.67	0.33
All fields of sciences	2.02	0.47	2.41	0.49

Notes: All indicators calculated for the total number of publications indexed in the Web of Science database over the period of 2001–2011.

Source: author calculations on Essential Science Indicators database materials (section "Citation Rankings, subsection "Countries/territories" country "Russia," option "View table of graph data" for Russian indicators; section "Citation Analysis," subsection "Baselines", option "View field rankings table" for the total world indicators).

In the total world structure of highly-cited publications the following fields of science were the most important over the period of 2001 – 2011: clinical medicine (22.0%), chemistry (11.8%), physics (8.7%), and engineering (8%). The shares of Russian highly-cited publications from different fields of science in their total world number

are given in *Table 10*). Russian highly-cited publications on physics accounted for almost half (48.6%) of their total number. Clinical medicine, engineering, chemistry, and geosciences in sum accounted for another 28.6% of Russian highly-cited publications in 2001 – 2011.

Highly-cited publications on physics had the highest index of

scientific specialization: Its share in the total number of the highly cited publications by Russian authors is more than fivefold higher than the identical world indicator. In Russia, the share of highly-cited publications on physics in the total number of highly cited papers is ~48%, whereas world indicator is ~8.6%. Hence, the index of scien-

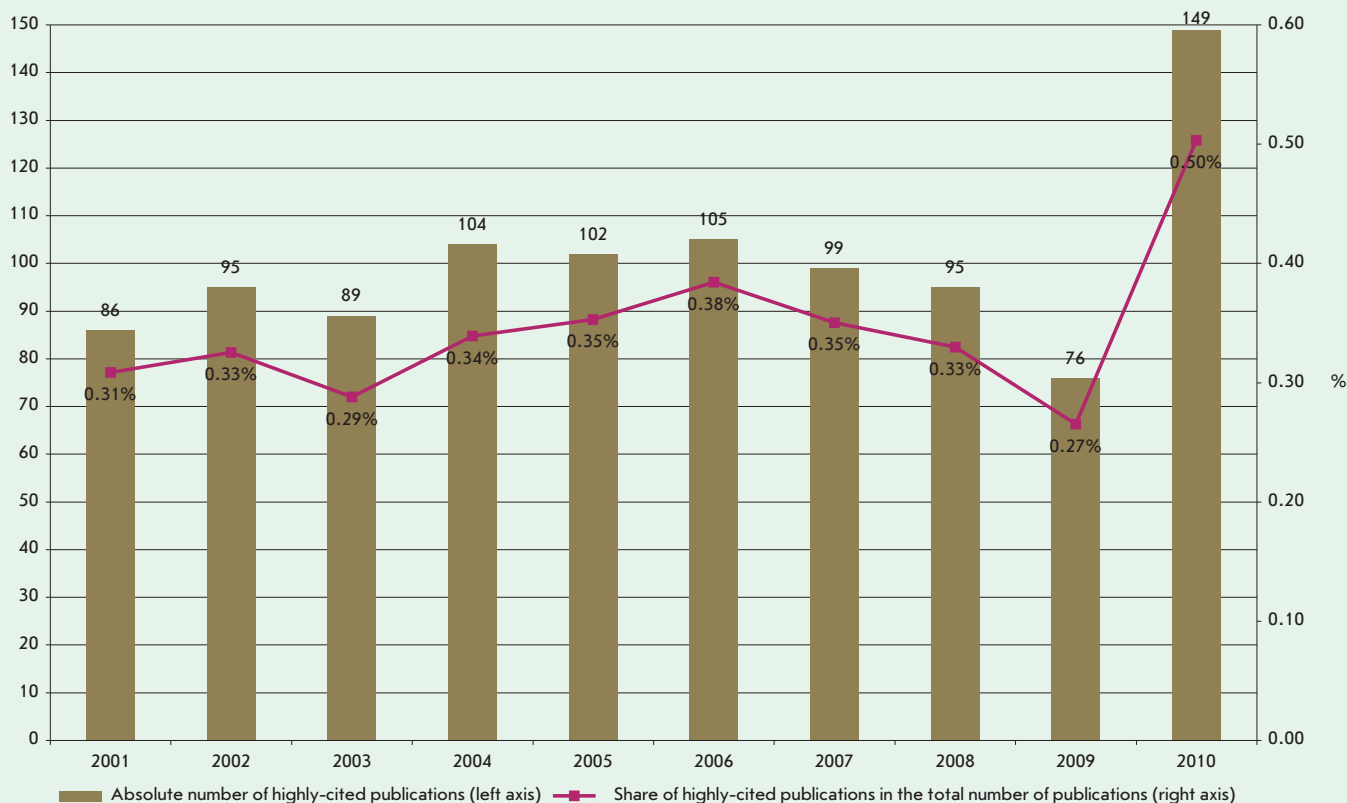


Fig. 4. Dynamics of the number of Russian highly-cited publications: 2001–2010.

Source: author calculations on Essential Science Indicators database materials (section "Most cited papers," subsection "Highly cited papers").

tific specialisation of Russian highly-cited publications on physics is $48\%/8.6\% = 5.6$. It is also should be noted that geosciences, space science, and mathematics were areas of specialisation of Russian highly-cited publications. These highly-cited Russian publications contribute most substantially to world science. On the contrary, Russian publications in the field of immunology, pharmacology & toxicology, neuroscience and behavior, psychiatry & psychology, economics & business, computer science, and social science prove the lowest potential of being highly cited.

Thus, the position of Russia in world science is strongest in the field of physics. Approximately half of highly-cited Russian publications and a significant share of the

total number of Russian publications in scientific journals indexed by Web of Science are publications on physics.

Table 11 allows for analysis of positions of highly-cited Russian publications in the world (Table 11). As previously mentioned for the total number of publication the USA was the once again the stable global leader in terms of highly-cited publications, with 34.1% of their total world number. The second and third positions in the global ranking of the number of highly-cited publications were held by Great Britain (8.8%) and Germany (7.2%), respectively. China with the largest number of highly cited publications among the Asian countries held the sixth position in this ranking. Russia held the 21st position with 0.64% of

total world number of highly-cited publications.

Over the period of 2001 – 2011, Switzerland had the highest share of highly cited publications in the total number of publications among the countries listed in Table 11 (2.3). On the contrary, in Russia this indicator was the lowest among the countries listed in Table 11 (0.39%).

INTERNATIONAL CO-AUTHORSHIP

The last section of this work is a brief review of the major directions of international co-authorship by Russian authors (methods for evaluating the international co-authorship were discussed in [13–15]). The intensity and direction of international collaboration are typically measured using the analysis of co-authorship indicators for re-

Table 10. Distribution of highly cited publications over fields of science: 2001–2011

Field of science	World	Russia	Share of Russian highly cited publications in the total number of highly cited papers, %	Specialization index of Russian publications
Biology & biochemistry	5501	46	0.84	0.79
Immunology	1252	2	0.16	0.15
Clinical medicine	21783	104	0.48	0.45
Computer science	2544	3	0.12	0.11
Mathematics	2377	27	1.14	1.08
Materials science	4572	27	0.59	0.56
Microbiology	1646	6	0.36	0.35
Molecular biology & genetics	2806	18	0.64	0.61
Multidisciplinary	180	4	2.22	2.11
Plant & animal science	2872	55	1.92	0.23
Geosciences	1240	27	2.18	1.82
Space science	5789	14	0.24	2.07
Environment/ecology	2762	16	0.58	0.55
Neuroscience & behavior	3012	4	0.13	0.13
Social science	4562	5	0.11	0.10
Psychiatry & psychology	2477	3	0.12	0.11
Agricultural science	2041	3	0.15	0.14
Engineering	7890	91	1.15	1.09
Pharmacology & toxicology	1848	4	0.22	0.21
Physics	8600	508	5.91	5.61
Chemistry	11699	77	0.66	0.62
Economics & business	1722	1	0.06	0.06
Total	99175	1045	1.05	

Notes: 1. All indicators were calculated for the total number of highly cited publications in the scientific journals indexed in the Web of Science database over the period of 2001–2011.

2. The data for December 2011 are listed in the table.

Source: author calculations on Essential Science Indicators database materials (section “Most cited papers,” subsection “Highly cited papers”).

searchers from different countries (e.g., [1–5, 14, 16–18]). Publication is considered to be written in international co-authorship if there are authors from at least two countries in the list of co-authors. In such publications, authorship is ascribed to all of the co-authors in an identical degree, irrespective of the share of contribution of the individual author to the paper. A high share of papers published in international co-authorship in the total number of publications of a country can be

indicative both of its central position in the international research networks and of a lack of self-sufficiency within the framework of national science. In line with the changes in scientific production in the world, which assume the formation of larger research groups, a tendency towards hyper- and mega-authorship bringing together tens of authors from different countries has appeared in certain fields of science (the methods and methodology of assessing the “mul-

tinationality” of publications written in international co-authorship were discussed in [10, 19, 20]).

The share of publications written in international co-authorship in the total number of publications is considered as a relative indicator of the level of integration of researchers of a particular country into the global scientific community. The basic indicators of the level of integration of Russian researchers into the global scientific community are shown in *Fig. 5*. The share of publications pre-

Table 11. The main indicators of highly cited publications of the leading countries: 2001–2011

Country	Number of highly cited publications	Position in the ranking of the number of highly cited publications	Share of highly cited publications of the country in the total world number of highly cited publications, %	Share of highly cited publications in the total number of publications of the country, %
USA	55953	1	34.10	1.83
Great Britain	14505	2	8.84	1.76
Germany	12649	3	7.72	1.61
France	7155	4	4.37	1.28
Canada	6717	5	4.09	1.49
China	5856	6	3.57	0.70
Japan	5659	7	3.45	0.73
Italy	5097	8	3.11	1.19
The Netherlands	4808	9	2.93	1.91
Australia	4210	10	2.57	1.38
Switzerland	4171	11	2.55	2.30
Spain	3584	12	2.18	1.06
Sweden	2747	13	1.68	1.53
Belgium	2310	14	1.41	1.68
Denmark	1940	15	1.18	1.98
South Korea	1773	16	1.08	0.63
Israel	1450	17	0.88	1.31
Austria	1438	18	0.87	1.50
India	1238	19	0.76	0.42
Finland	1172	20	0.72	1.32
Russia	1045	21	0.63	0.39

Notes: 1. All indicators were calculated for the total number of highly cited publications in the scientific journals indexed in the Web of Science over the period of 2001–2011.

2. The 2011 data correspond to the situation by the beginning of data for December 2011 is listed in the table.

Source: author calculations on Essential Science Indicators database materials (section "Most cited papers," subsection "Highly cited papers").

pared in international co-authorship in the total number of Russian publications over the period of 2001–2011 fluctuated around 29%–33%. It corresponded to the absolute number of approximately 9,000 publications. However, a decrease in the number of publications prepared in international co-authorship started in 2008.

Researchers from the USA and Germany were the main foreign partners of Russian researchers (Table 12): 26%–27% of the total number of Russian publications in international collaboration was prepared in co-authorship with

researches from these countries. French and British researchers were also significant partners of Russian scientists. The strengthening of scientific links between Russia and Asian countries should also be noted. The number of publications by Russian authors in co-authorship with their Indian, Chinese and Taiwanese, and South Korean colleagues increased by a factor of 3.6; 3.1, and 4.6, respectively, over the period of 2001–2011. The development of scientific cooperation between Russia and European countries was less dynamic. Among

the former USSR countries, only Ukraine and Belarus were included in the cohort of the 25 most significant scientific partners of Russia.

On the other hand, Russia was an important scientific partner for Central Asian and Caucasian countries, as well as for Mongolia, Ecuador, Ukraine, and Belarus (Table 13). Meanwhile, Russia was an insignificant scientific partner for such countries as the USA, Germany, France, and Great Britain. For these countries the share of publications prepared in co-authorship with Russia, in the total number

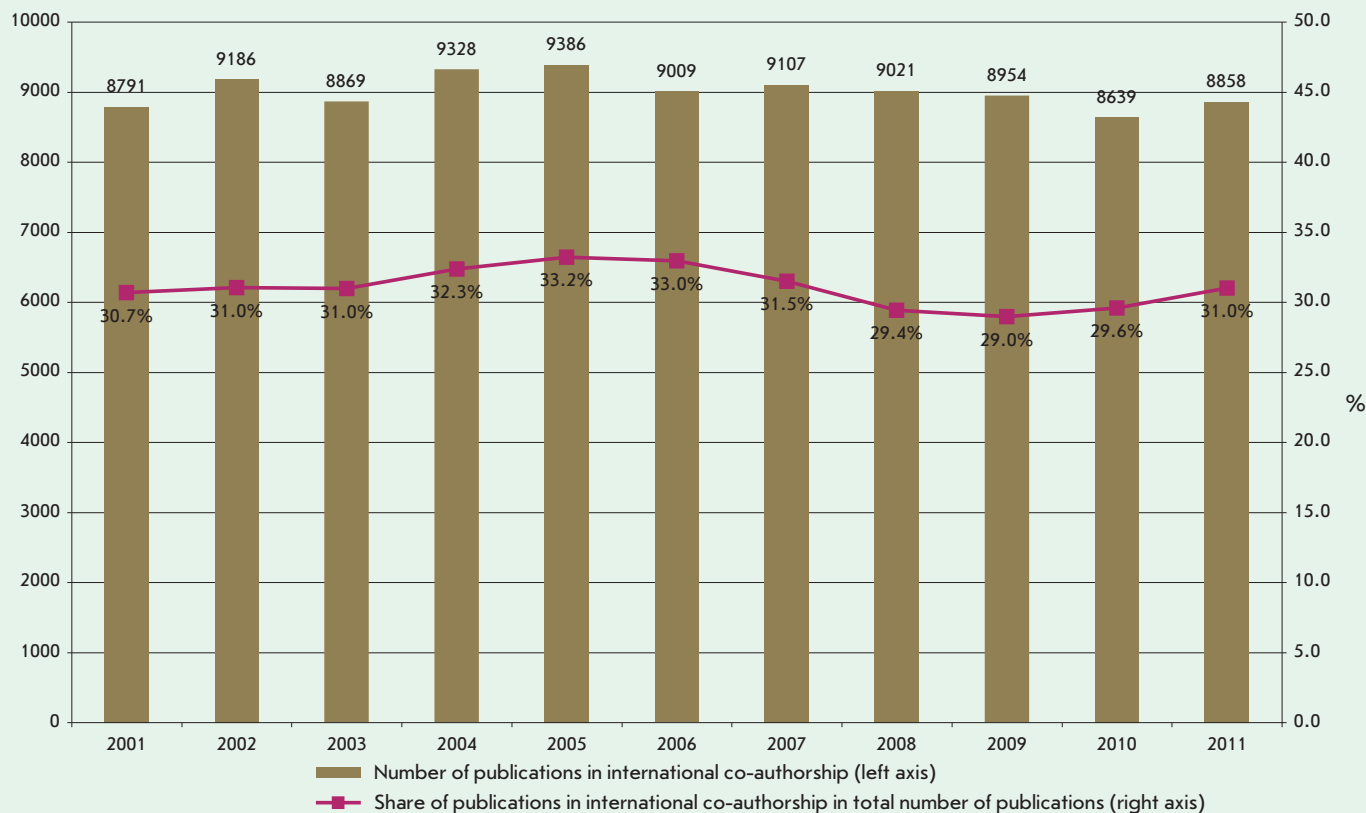


Fig. 5. Basic indicators of international collaboration by Russian authors: 2001–2011.

The number of publications of Russian authors prepared in international co-authorship was calculated using the following procedure: the publications were tallied for each year over the studied period and all countries, except for Russia, using the option "Analyze results" of the Web of Science database for the country "Russia" in the category "Countries/territories."

of publications in co-authorship amounted to less than 3.5%.

In Northern European countries, the level of integration of researchers of these countries into the global scientific community was considerably higher than in Russia (approximately 50%). This indicator was even higher in Indonesia, Cyprus, Tajikistan, Turkmenistan, Kyrgyzstan, and Uzbekistan: 60–65% in recent years [21–23]. This indicator was also high for countries with transition economies (e.g., Latvia, Estonia, and Belarus). The indicator has always been high for such countries as Germany, France, and Canada: almost half of all publications made in these countries were

the result of international collaboration. On the contrary, among the top 30 countries in number of publications, this indicator is the lowest for China, India, Turkey, and Iran. In general, the level of integration of researchers into the global scientific community in European countries was higher than the level observed in Asian countries.

Bibliometric analysis of the patterns of academic co-authorship at the world level have shown a significant increase in the number of publications prepared in international collaboration during the past 20–30 years. This phenomenon can be interpreted as the sign of the deepening of specialization and globali-

zation of knowledge production. In particular, the increasing role of the BRIC countries in international collaboration has been noted; with Russia as a leader on this indicator among the four BRIC countries. The number of Russian publication in Web of Science database prepared in international co-authorship over the period from 1980 to 2011 jumped from 3% to 31%, reaching parity with the level achieved by the USA.

The distribution of Russian publications in international co-authorship over various fields of science generally corresponds to the areas of the scientific specialization of Russia: physics holds a signifi-

Table 12. The main scientific partners of Russia

№	2001			2011		
	Country	Number of publications in co-authorship	Share of publications in the total number of Russian publications in international co-authorship, %	Country	Number of publications in co-authorship	Share of publications in the total number of Russian publications in international co-authorship, %
1	Germany	2389	27.18	Germany	2564	28.95
2	USA	2158	24.55	USA	2366	26.71
3	France	1076	12.24	France	1499	16.92
4	Great Britain	902	10.26	Great Britain	1390	15.69
5	Japan	681	7.75	Italy	924	10.43
6	Italy	651	7.41	Spain	730	8.24
7	Sweden	493	5.61	Japan	665	7.51
8	The Netherlands	432	4.91	China	635	7.17
9	Poland	423	4.81	Poland	621	7.01
10	Switzerland	346	3.94	Switzerland	582	6.57
11	Canada	319	3.63	Ukraine	540	6.10
12	Ukraine	318	3.62	Canada	514	5.80
13	Spain	315	3.58	Sweden	508	5.73
14	Finland	260	2.96	The Netherlands	494	5.58
15	Belgium	246	2.80	Finland	490	5.53
16	South Korea	220	2.50	Czech Republic	457	5.16
17	Czech Republic	216	2.46	Belgium	425	4.80
18	China	202	2.30	Austria	405	4.57
19	Israel	185	2.10	South Korea	398	4.49
20	Austria	151	1.72	Brazil	361	4.08
21	Brazil	150	1.71	India	331	3.74
22	Denmark	150	1.71	Taiwan	310	3.50
23	Mexico	149	1.69	Belarus	299	3.38
24	Norway	147	1.67	Australia	299	3.38
25	Australia	140	1.59	Greece	288	3.25

Notes: 1. The total of the shares for the countries is higher than 100%, since some papers were prepared in co-authorship with researchers from more than one country.

2. The procedure for calculating the total number of publications by Russian authors in international co-authorship is described in the legend to Fig. 5.

Source: author calculations on Web of Science database materials.

cant lead (in particular, solid body physics and sub-disciplines studying nuclear processes, fields and particles; optics and spectroscopy, and plasma physics). Astronomy & astrophysics, material science, physical chemistry, instrument engineering, biochemistry & mo-

lecular biology, geosciences, mathematics, and electronics follow these disciplines..

CONCLUSIONS

The publication activity of Russian scientists and the research efficiency of the leading scientific countries

were reviewed. The dynamics of the number of publications in scientific journals indexed in the Web of Science database and the basic citation indicators was analysed for the countries of studied sample (countries with more than 10000 publications in scientific journal, indexed

Table 13. Share of publications in co-authorship with Russian researchers in the total number of publications of the country: 2001–2011 (%)

Country	2001	Country	2011
Turkmenistan	40.00	Georgia	30.94
Tajikistan	24.00	Belarus	27.66
Armenia	19.83	Armenia	26.29
Kazakhstan	17.87	Kyrgyzstan	23.44
Georgia	17.56	Kazakhstan	23.18
Ecuador	16.35	Mongolia	19.29
Uzbekistan	13.02	Azerbaijan	18.04
Mongolia	12.50	Uzbekistan	15.45
Belarus	8.48	Tajikistan	14.81
Ukraine	6.40	Ecuador	13.61
Azerbaijan	5.44	Ukraine	10.34

Notes: 1. The indicators listed in the table are calculated as follows: for each country, the number of publications in co-authorship with Russian researchers is the indicator "Record count" in the box "Russia," which is shown after the function "Search within results for Countries/territories" in the option "Analyze results" in the Web of Science database is executed. The following types of documents were selected for the analysis: article, proceedings paper, and review.

2. The countries whose share of publications in co-authorship with Russian researchers in the total number of publications of the country was higher than 10% in 2001 or 2011 are listed.

Source: author calculations on Web of Science database materials. All databases of the Web of Science portal were used.

by Web of Science over the period of 2001 – 2011). The fields of scientific specialization for Russia were also identified. Moreover, the analysis has demonstrated that Russia lost positions in the world science over the period from 2001 to 2011. Countries with dynamic publication activity rose to overtake Russia in the world rating. The most significant loss of position for Russia occurred in the fields that are considered of traditional strength for Russia (physics, engineering, materials science, chemistry and mathematics). Furthermore, the average number of citations per paper in Russia was one of the lowest in the world.

The conclusions concerning the scientific specialization of Russia can be drawn through analysis of

the distribution of publications over various fields of science. The majority of Russian publications were in natural sciences and engineering, whereas the world structure of science was characterized by the predominance of publications on medical and biological sciences. The most significant contribution the Russian scientific output came from publications on physics. These publications accounted for almost 50% of Russian highly cited publications and for 20% of the total number of Russian publications. Russian publications on physics were much more broadly represented in the world science in comparison with publications in other fields of science. Moreover, Russian scientific papers on physics significantly contributed (in comparison with publications in

other fields of science) to the total number of Russian publications in international co-authorship.

The level of integration of Russian researchers into the world scientific community was relatively high. However, this indicator was higher in European countries (in particular, in Northern European countries). The main scientific partners of Russian researchers were scientists from the USA, Germany, Great Britain, and France. Scientific cooperation between Russian scientists and their colleagues from Asian countries (primarily from China, India and South Korea) considerably strengthened over the period of 2001 – 2011.

The following conclusions can be drawn from the analysis of the publication activity in the other countries. In 2001–2011, the U.S. clearly held the leading positions in the world science. The publication activity in Asian countries was much more dynamic than in European countries. Iran demonstrated the highest growth rate of the indicators of publication activity among the countries of the studied sample. However, the citation indicators and the level of international scientific collaboration were in European countries were higher than in Asian countries. Therefore, a significant rearrangement in world science has occurred over the past 10 years. Developing countries (primarily China, and also Iran, Thailand, Turkey, Malaysia, Brazil, Pakistan, and India) have begun closing the gap with the established leaders (North American countries, Northern European countries, and Japan). Because of the strengthening in this trend, Russia's lagging will manifest itself largely. By the next decade Russia may lose its status as one of the world's great scientific countries without timely reforms in scientific policy. ●

REFERENCES

1. Gokhberg L.M. *Statistika nauki (Statistics of Science)*. Moscow: TEIS, 2003. 478 p.
2. Gokhberg L.M., Sagieva G.S. // *Forsait (Foresight)*. 2007. V. 1. № 1. P. 44 – 53.
3. Kirchik O.I. // *Forsait (Foresight)*. 2011. V. 5. № 3. p. 34–42.
4. Rossijskij innovacionnyj indeks (Russian Innovation Index) / Ed. L. M. Gokhberg. Moscow: Higher School of Economics, 2011. 84 p.
5. Kotsemir M.N. // *Forsait (Foresight)*. V. 6. №1.
6. Merton R.K. // *Science*. 1968. V. 159. Iss. 3810. P. 56–63.
7. Merton R.K. // *ISIS*. 1998. V. 79. № 4. P. 606–623.
8. Gokhberg L., Pislyakov V. // 10th International Conference on Science and Technology Indicators. Book of Abstracts. 2008. Vienna. P. 400–403.
9. Batygin G.S. // *Vedomosti. Tyumen Oil and Gas University, Institute for Applied Ethics*. 2001. Iss. 18. P. 173–185.
10. Pislyakov V.V. // Preprint WP6/2010/01. 2010. Moscow. Higher School of Economics.
11. Aksnes D.W., Sivertsen G. // *Scientometrics*. 2004. V. 59. № 2. P. 213–224.
12. Tijssen R.J.W., Visser M.S., Van Leeuwen T.N. // *Scientometrics*. 2002. V. 54. P. 381–397.
13. Luukkonen T., Tijssen, R.J.W., Persson O., Silvertsen G // *Scientometrics*. 1993. V. 28. P. 15–36.
14. Glänzel W., Schubert A., Czerwon H.J. // *Scientometrics*. 1999. V. 45. P. 185 – 202.
15. Bookstein A., Moed H., Yitzhaki M. // *Information Processing and Management*. 2006. V. 42. P. 1422–1442.
16. Melin G., Persson O. // *Scientometrics*. 1996. V. 36. P. 363–377.
17. Arunachalam S., Doss M.J. // *Current Science*. 2000. V. 79. № 5.
18. Wang Y., Wu Y., Pan Y., Ma Z., Rousseau R. // *Scientometrics*. 2005. V. 62. № 2. P. 183 – 198.
19. Lange C., Glänzel W. // *Scientometrics*. 1997. V. 49. № 3. P. 593–604.
20. Nederhof A.J., Moed H.F. // *Scientometrics*. 1993. V. 27. P. 39 – 52.
21. *Indikatory nauki: 2010. Statisticheskij sbornik (Indicators of Science 2010: Statistics Digest)*. Moscow: HSE, 2010. 368 p.
22. *Indikatory nauki: 2011. Statisticheskij sbornik (Indicators of Science 2011: Statistics Digest)*. Moscow: HSE, 2011. 368 p.
23. *Indikatory nauki: 2012. Statisticheskij sbornik (Indicators of Science 2012: Statistics Digest)*. Moscow: HSE, 2012. 392 p.

Science and Technologies in Russia – STRF.RU



40% of scientists agree that the publication of research results helps the enlightenment of the society, leads to the growth of authority of scientific work

34% believe that wide communicating of research results helps to rise the foundation

12% hope that media communications helps them to stand out in public opinion...
...but

17% never speak to journalists*

Open your work to the world!

Physicochemical Biology: Conquered Boundaries and New Horizons

D.G. Knorre

Institute of Chemical Biology and Fundamental Medicine, Siberian Branch, Russian Academy of Sciences, prosp. Akademika Lavrent'eva, 8, Novosibirsk, 630090, Russia

E-mail: knorre@niboch.nsc.ru

Received 04.11.2011

Copyright © 2012 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT In this paper, we shall consider the main evolutionary stages that occurred within the field of physicochemical biology during the 20th century, following the determination of the tertiary structure of DNA by Watson and Crick and the subsequent successes in the X-ray structural analysis of biopolymers. The authors' ideas on the pre-emptive problems and the methods used in physicochemical biology in the 21st century are also presented, including an investigation of the dynamics of biochemical processes, studies of the functions of unstructured proteins, as well as single-molecule investigations of enzymatic processes and of biopolymer tertiary structure formation.

KEYWORDS DNA structure; enzyme active sites; unstructured proteins; dynamics of biochemical processes; single molecule studies of enzymatic processes and biopolymer tertiary structure formation.

The second half of the 20th century represented a period of tremendous achievement for mankind, witnessing an increase in the understanding of the essence of natural phenomena and heralding significant breakthroughs with regard to mankind's technical abilities. Progress achieved in the use of nuclear energy has enabled us to cover the earth with a network of nuclear power stations, while the appearance of powerful jet engines has opened the doors to space voyage; a process which began with Gagarin's flight, continued with the landing on the moon, and is now moving towards a flight to Mars. Advancements in electronics and materials science have enabled us to build computers that can perform trillions of computations per second and have facilitated the creation of devices smaller than a matchbox that are capable of storing gigabytes of information, as well as systems for ultrahigh-speed information transfer. These factors have resulted in the emergence of one of mankind's most impressive technical marvels: the Internet.

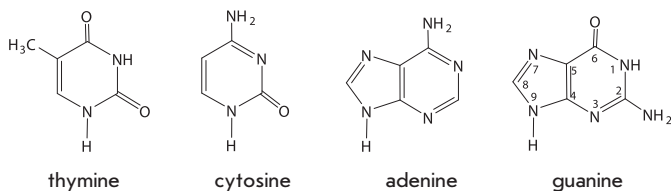
A less obvious, but no less significant, achievement was accomplished in our understanding of the chemical and physical foundation of the functioning of living organisms. It is fair to say that the introduction of physical and chemical approaches into the field of life sciences was a gradual process; this progress made a significant leap forward after the publication of the famous paper by James Watson and Francis Crick in 1953, which presented a three-dimensional structure of deoxyribonucleic acid (DNA) [1].

A plethora of mysteries exist within nature. Amongst these is the question of how an unthinkable huge amount of information is transferred accurately from one cell produced by the fertilization of an egg cell to an adult organism. Until 1944, the nature of the carrier of hereditary information remained unknown. Although Miescher had discovered nucleic acids in 1869, the prevailing view was that a protein played the role of such a carrier; indeed, it was not until 1944 that Avery demonstrated that DNA was the actual carrier of hereditary information [2].

It was evident that such a carrier had to have three main functions. First of all, it was necessary for the carrier to possess a huge storage capacity with which to store information relating to the manifold properties of a living organism, including its structure and the functions inherent to specific species and even individuals. Secondly, such a carrier should possess a mechanism for the realization of the information in to the definite structures of a living organism and its numerous functions (in order to express this information). Thirdly, the most important requirement was the existence of a mechanism for transferring the information to subsequent generations.

The main claim of DNA to be such a carrier is rooted in its chemical structure. DNA is a linear polymer that is comprised of four different monomers, i.e. nucleotides. Each monomer consists of three fragments: a carbohydrate residue (deoxyribose) bound to an orthophosphoric residue and one of the four

heterocyclic residues: adenine, guanine, thymine, and cytosine.



The nucleotides are bound by phosphodiester bonds between deoxyribose and phosphoric acid residues (Fig. 1).

Such a structural principle enables the existence of an innumerable quantity of various polymeric structures differing in the set and sequences of nucleotides. For a polymer built from n monomeric units, the number of combinations amounts to $4^n = 10^{0.6n}$. Even for a very short polymer of 200 monomeric units, this amount (10^{120}) exceeds the number of atoms in the observable part of the universe, which is estimated at (10^{80}). In fact, the DNA for even the simplest organisms is comprised of thousands, even millions, of nucleotides.

These calculations imply that most imaginable nucleotide sequences could not, in principle, appear in the universe and are subjected to natural selection. This means that the appearance or generation of organisms more fascinating than those that exist at the present time can not be excluded.

However, neither the light shed upon the role of the DNA as an information carrier, nor the huge information capacity of the DNA molecule, has enabled us to solve one of the most intriguing puzzles of this area of science: i.e., how this vast amount of information is transferred from one generation to the other. It has been established that the answer to this puzzle is rooted in the spatial structure of the DNA. According to the model proposed by Watson and Crick, which has been fully confirmed by numerous subsequent experimental studies, DNA is a construct of two polynucleotide chains which are bound together by hydrogen bonds. Within the structure proposed, adenine may interact only with thymine; and guanine with cytosine. Such sequences are considered complementary. The presence of such a correspondence means that any nucleotide sequence in one chain unambiguously corresponds to one nucleotide sequence in another chain, hereby following the mechanism of information transfer from maternal to two daughter cells at cell division. According to this mechanism, two polynucleotide chains separate prior to cell division, each of them governing the formation (synthesis) of a new complementary chain, thus double-stranded structures are formed that are

identical to the DNA of the maternal cell. The existence of such a process was confirmed by Meselson and Stahl [3] soon after the appearance of Watson-Crick's work. These authors prepared *Escherichia coli* cells grown on a medium containing $^{15}\text{NH}_4\text{Cl}$ as a single source of nitrogen. Thereafter, the cells were permitted to grow for several generations in a medium with the usual nitrogen isotope. In all subsequent cell generations, the presence of heavy DNA with the same amount of the heavy nitrogen isotope was observed; indicating that the ^{15}N -DNA formed in the first step of the experiment remained intact and was simply transferred to one of the daughter cells during each subsequent division of the daughter cells.

The defining feature of Watson-Crick's work was the fact that the structure of a biologically significant macromolecule was derived using the established geometrical parameters of definite chemical bonds; therefore, the elucidation of the biological phenomenon that begins with the physicochemical parameters of the molecule responsible for the phenomenon was achieved. Consequently, this work can be considered as having heralded the birth of physicochemical biology.

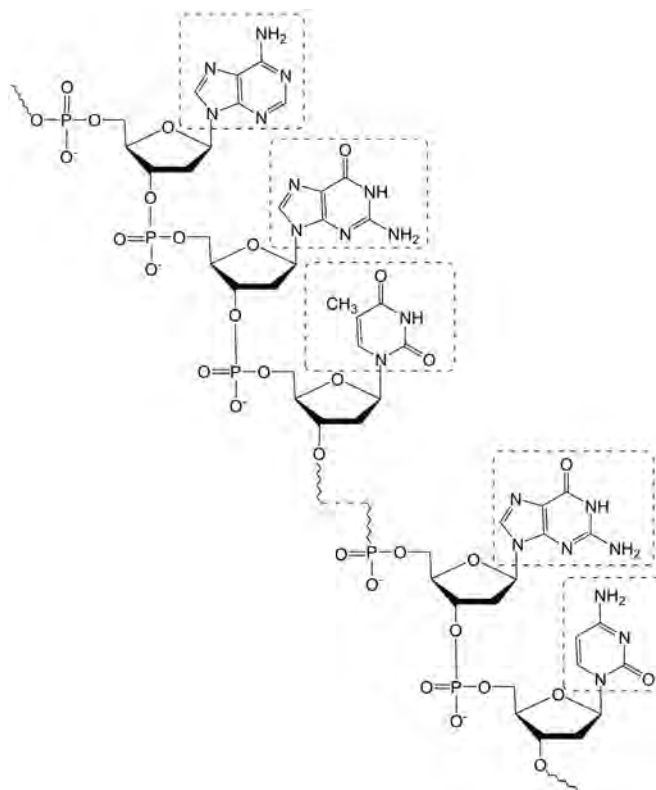


Fig. 1. Fragment of the structure of a DNA molecule.

Currently, physicochemical biology includes several scientific disciplines: biological chemistry, biophysics, bioorganic chemistry, and molecular biology. It can reasonably be argued that the traditional separation of these disciplines is not entirely appropriate. For example, molecular biology, according to Wikipedia, is defined as a science that deals with the molecular backgrounds of biological activity; however, biological chemistry has focussed for a considerable period of time upon molecular concepts which describe the most essential biochemical processes as being conversions of molecules with a commonly known chemical structure and has considered the catalysts of these processes to be individual compounds, i.e. as molecules. Therefore, the entire concept of modern biological chemistry refers to molecular science and could therefore reasonably lay claim to the appellation "molecular biology". For the remainder of this paper the term physicochemical biology will be used to refer to the science that studies biological phenomena on the basis of the physicochemical properties of separate atoms and chemical bonds.

The work of Watson and Crick inspired vigorous efforts which eventually resulted in the identification of the primary biochemical mechanisms that ensure the transfer and expression of genetic information. The concept of the matrix synthesis of biopolymers was the central element of these mechanisms; according to this, each step of elongation of a new biopolymer molecule is not only catalyzed by a specific enzyme, but is also checked by a special nucleic acid, indicating which monomer should be bound to the growing polymer chain at a given stage. These mechanisms are described in all contemporary international and Russian biological chemistry textbooks and manuals, such as [4, 5].

The discovery of the enzymes that catalyze the synthesis of complementary DNA molecules has led to the elaboration of the polymerase chain reaction (PCR) [6], which has found application in medical diagnosis, forensic science, and archaeology.

The elucidation of the mechanisms of DNA expression and the achievement of chemists in the synthesis of oligonucleotides of a desired sequence has led to the appearance of genetic engineering [7]. It has become possible to carve out definite genes, to modify them and then to subsequently insert them into the proper region of the genome, thus performing site-directed mutagenesis [8].

The greatest scientific effort in the field of physicochemical biology was launched in 1990 under the name "Human Genome," a program aimed at the mapping of the complete nucleotide sequence (sequencing) of human DNA [9]. As early as in 2001, Venter and 272 co-authors had published a complete nucleotide sequence

of the human genome [10]. The methods elaborated within the framework of the program and those that are still being improved have opened the doors to the obtaining of genetic maps for any individual; as well as for the obtaining of genome structures for all the animals, plants and microorganisms on Earth. Consequently, irrespective of the striking success that has been achieved in the elaboration of high-speed efficient sequencing methods, scientists dealing with molecular biology have had enough work on their plate to last for several decades.

The entire breadth of ground that physicochemical biology has covered, from Watson and Crick's effort to the determination of the structure of the human genome, can be viewed as an incremental effort with clearly formulated tasks and with the purpose of investigating and designing new, innovative methods. During that journey, new and unexpected advances were made along the way; among such advances is the discovery of ribozymes by Thomas Cech [11] and Sidney Altman [12], as well as the discovery of small interfering RNAs [13].

The appearance of new physicochemical peculiarities for living matter in the future is an eventuality which cannot be excluded *a priori*. The role of a significant portion of the human genome remains unclear, since only 1.5% of it is responsible for the 23,000 genes coding human proteins. A significant portion of the genome determines the synthesis of various non-coding RNAs: transfer and ribosomal RNAs, introns. However, this does not account for the remaining 98.5% of the genome, and thus a significant portion is considered junk DNA. Establishing the role of this DNA is one of the most challenging problems in the field of physicochemical biology. The functional importance of the extracellular nucleic acids present in appreciable amounts in the blood plasma still remains unclear [14].

Among the main achievements in physicochemical biology in the past century, it would be short-sighted not to mention the great progress achieved through X-ray crystallography and the NMR study of proteins in the understanding of the mechanisms of biological catalysis. A large body of material has been accumulated relating to the atomic structure of an enzyme's active sites and their complexes with specific ligands, which has enabled the formulation of reasonable hypotheses pertaining to the mechanisms of recognition and catalytic conversions. For a perspective on the degree of information obtained on the nature of an enzyme's active sites, a scheme of the arrangement of the reaction intermediate phenylalanine adenylate in the active site of the phenylalanine-tRNA-synthetase is presented in Fig. 2, with the bonds formed by enzyme active site

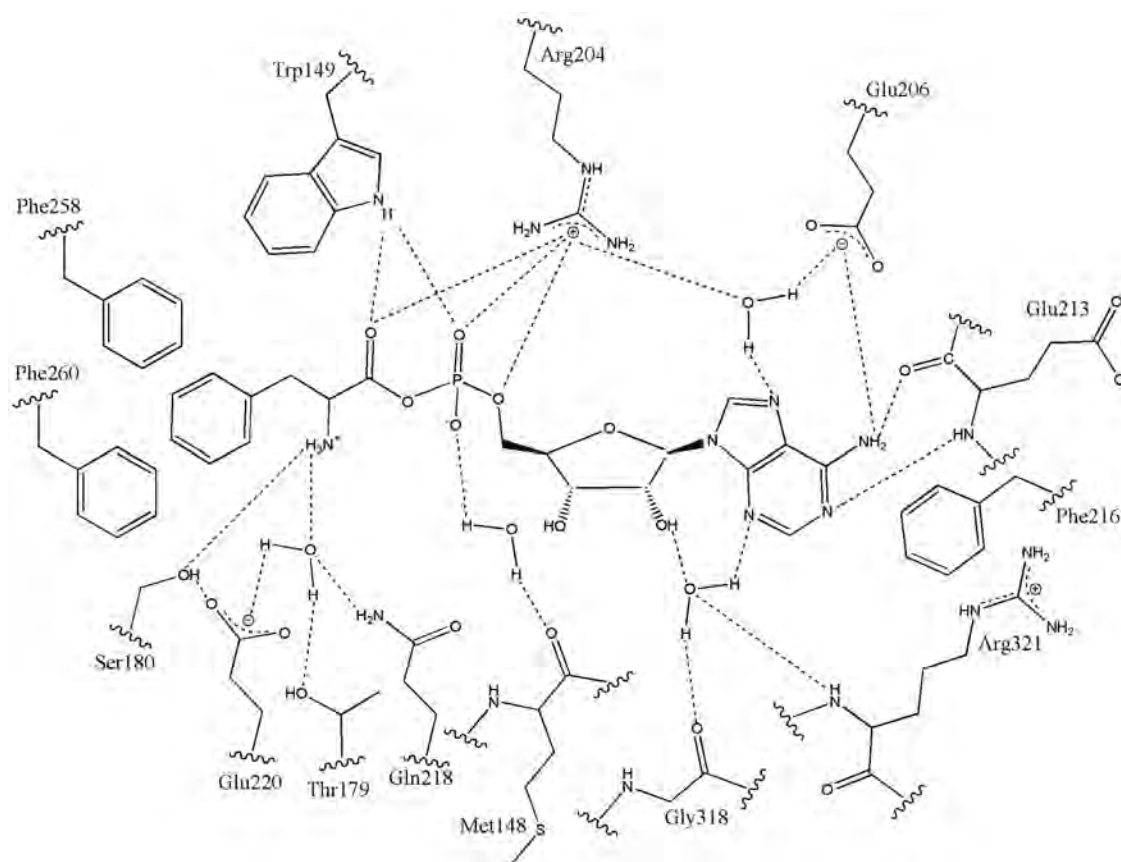


Fig. 2. Structure of the active site of phenylalanine-tRNA-synthetase in complex with the intermediate (phenylalanine adenylate). Dots indicate hydrogen bonds between the intermediate atoms and the enzyme with binding water molecules.

groups, including the water molecules participating in the interaction [15].

However, this author believes that the focus of physiochemical biology in the 21st century should shift to other matters. Several aspects which require primordial development both at the theoretical and experimental levels should receive more attention. First and foremost, the role of molecular dynamics requires significantly more attention.

Certainly, the dynamic character of the functioning of the biopolymer did not come as an unexpected revelation. However, organic chemistry, including bioorganic chemistry, has dealt predominantly with structures that are, in essence, static. Dynamic events have been considered as a transfer from the static structure of the reagents to the static structure of the reaction products. In the best cases intermediates were taken into account; however, these were also presented as static structures. It is absolutely common knowledge amongst chemists that molecules, including biopolymer molecules, are subjected to internal motions: the atomic vibrations proceeding at a subpicosecond time scale, fluctuations of a side radical at a pico-nanosecond time scale, conformational rearrangements in the millisecond range, and even slower intramolecular movements. The

problem with the molecular dynamics of biopolymers is not limited to simply stating and describing these motions but expands into establishing the role of these dynamic events in the recognition process, catalytic conversions, as well as intra- and extracellular signalling. At the time of writing, the most discussed topic is the role of dynamic factors in enzymatic catalysis [16].

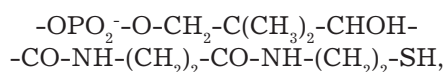
The role of dynamics in enzymatic catalysis was first brought under discussion in the induced fit hypothesis formulated by Koshland [17]. According to this hypothesis, no ideal compliance exists initially in the structure of the enzyme active site and the substrate which would enable procession to the execution of chemical conversion immediately after the formation of the enzyme-substrate complex. The process is supposed to be preceded by a conformational adjustment of the complex; i.e., a certain relocation of the atoms, which provides the necessary concordance of the chemical bonds subjected to conversion and a portion of the enzyme active site participating in the catalytic process.

The concept was confirmed by pre-stationary kinetic data. Such changes may be observed using rapid kinetic methods, such as stopped-flow in the millisecond range and relaxation methods (T-jump) in the microsecond range [18]. As an example, *Fig. 3* shows the curves ob-

tained by the stopped-flow method for the splitting of the base reaction (8-oxoguanine), which is catalyzed by 8-oxoguanine-DNA-glycosidase. The conversion was followed by fluorescence of tryptophan residues. At the first stage, the changes in conformation are clearly visible, whereas when several stages are recorded, the release of the reaction product (8-oxoguanine) is distinctly observed only at the final stage [19].

The time range in the use of relaxation methods is significantly expanded by the application of modern lasers capable of irradiating systems via femtosecond impulses, thus generating a T-jump within such a short time period [20, 21]. Moreover, if the solution contains a reagent with pK_a of the excited state significantly different from that of the ground state present in the solution, a pH jump may be performed via a laser impulse [22, 23].

The essential dynamics problem is the mechanism by which the system switches from one regime of functioning to another, significantly different, regime. The question already arises with regards to enzymes and enzymatic complexes that possess several catalytic functions manifesting themselves in a certain order. This relates to all the multifunctional enzymes which realize the sequential switching of different functions through a “swinging arm” that is capable of reaching various active sites. A number of such enzymes are known. For example, there is a fatty acid synthase which represents a complex of proteins that catalyze the sequential lengthening of the fatty acid carbon backbone by two-carbon fragments [24]. During the whole process, the growing chain of carbon residue is bound by a thioether bond to the SH-group of phosphopantothein



which is covalently bound by a phosphodiester bond to the serine residue of the acyl carrier protein (ACP). The arm contains a large number of σ -bonds and is therefore highly flexible. This allows the acyl residue to move alternately between the active sites catalyzing sequential stages of the biosynthesis of fatty acid from acetyl residues. The primary source of acetyl residues is the acetylated coenzyme A, $\text{CoAS}-\text{COCH}_3$, the main product of the catabolism of carbohydrates, fats, and a number of amino acids. The acetyl residue is carboxylated, and the malonyl residue formed is transferred from coenzyme A to ACP via the reaction

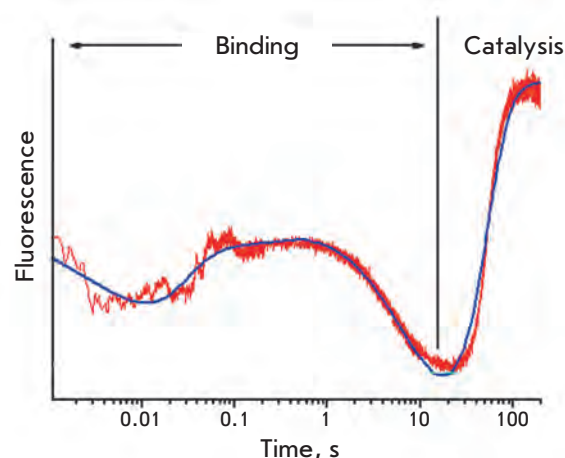


Fig. 3. Kinetic curve of the changes in fluorescence intensity (arbitrary units) during the initial phase of the reaction of 8-oxoguanine excision from the oligonucleotide containing an oxidized guanine residue.

Figure 4 represents a scheme of the processes that occur in all the two-carbon fragments introduced during the process. Malonyl-ACP is the direct donor of these fragments; it binds to the growing chain resulting in the detachment of CO_2 and the cleavage of the bond of the fragment with the protein, reduction of the fragment to $-\text{CHOHCH}_3$, its dehydration to $-\text{CH}=\text{CHCH}_3$, and reduction to $-\text{CH}_2\text{CH}_3$.

Clearly, each reaction proceeds with the participation of its own active site. The active sites may reside in different polypeptide chains (in eubacteria) or in one polyfunctional protein (in eukaryotes, including humans). The swinging arm must transport fragments of COCH_2R in a definite order to the four active sites for the procession of all sequential conversions.

The notion of intrinsically unordered proteins is a problem which has recently come to light and requires further study from the perspective of molecular dynamics [25–27]. Currently, there are a large number of such proteins which, in contrast to the commonly held view, function in the absence of a definite tertiary structure. Such proteins are unlikely to exist in the form of a statistically coiled polypeptide chain. In all likelihood, they represent an ensemble of rapidly, mutually transferring conformations in the solution. The predominance of proteins with an unordered conformation is typical of many neurodegenerative diseases, such as the Huntington disease and spinocerebellar ataxia (disorders of the gait and other types of movement coordination). However, many proteins with an unordered structure or at least containing rather ex-

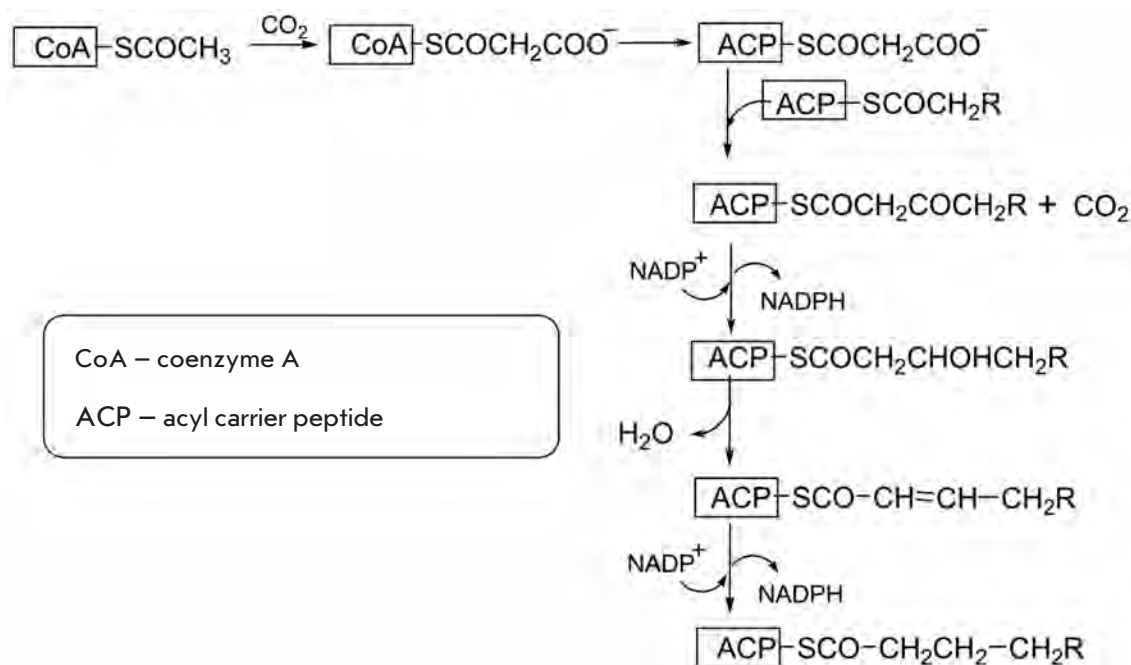


Fig. 4. Scheme representing the lengthening of two-carbon fragments catalyzed by fat acid synthase.

panded (more than 50 amino acid residues) unordered fragments are encountered within the established norm and more often in eukaryotes than in unicellular organisms. Among such proteins, the transcription factors and proteins responsible for chromatin remodeling and intracellular signalling occur more often. This certainly does not mean complete disorder. This can be supported by the fact that many of these proteins become structured after binding to their targets. The absence of order creates a serious problem for the elucidation of their spatial structure, since these proteins do not give rise to reflexes during the X-ray analysis. Meanwhile, more data has been accumulated pertaining to the fact that these unordered parts are most typical of poly-functional proteins. In all likelihood, the conformations with an affinity to different partners are also present among the conformations of these pseudo-unordered proteins. These proteins are typically characterized by a small content of bulky hydrophobic amino acid residues and an increased content of polar and charged residues.

At the time of writing, the theoretical investigation of biopolymer molecular dynamics is limited by the capabilities of computer techniques. The calculation of molecular dynamics assumes that a stepwise procedure is used, and it requires femtosecond time increments. Even for the modern supercomputers and software, the advance for several tens of nanosecond increment-scan only be attained for the biopolymers consisting of thousands atoms. Meanwhile, the most interesting con-

formational events occur in the micro- and even millisecond ranges.

Most studies focused on the physicochemical ground of the vital activity, in particular in the case of quantitative characteristics, were carried out *in vitro*. Most of the values obtained may be to a significant extent related to intracellular processes, especially to those in eukaryotic cells. A common eukaryotic cell may carry a large number of biopolymer molecules. The conditions in its cytosol do not significantly differ from those in a test tube. This may be demonstrated via a simple calculation for a spherical cell with a linear dimension of 20 μm, which is typical of a eukaryotic cell. A spherical cell of a 20-μm diameter was used to simplify the evaluation. For illustrative purposes, it is more convenient to perform calculations in Daltons (Da) as mass units and angstroms as length units (they can be qualified as the “cell” ones). Since 1g = 6 × 10²³Da and 1cm = 10⁸Å, the density is measured in 1g/cm³ = 0.6Da/Å³. The cell volume amounts approximately to 4 × 10¹⁵; the volume of a relatively large protein molecule (approximately 100kDa) is ~ 10⁵Å³. Assuming that the proteins occupy 10% of the cytosolic volume, 4 billions of such molecules can be accommodated in one cell. Therefore, it can be reckoned that cytosol conditions (with allowance made for the increased viscosity of the 10% protein solution) do not differ significantly from those in a tube. The arrangement of proteins on the cell surface can be estimated in a similar manner. Assuming that 10% of the plasma membrane surface is occupied by embed-

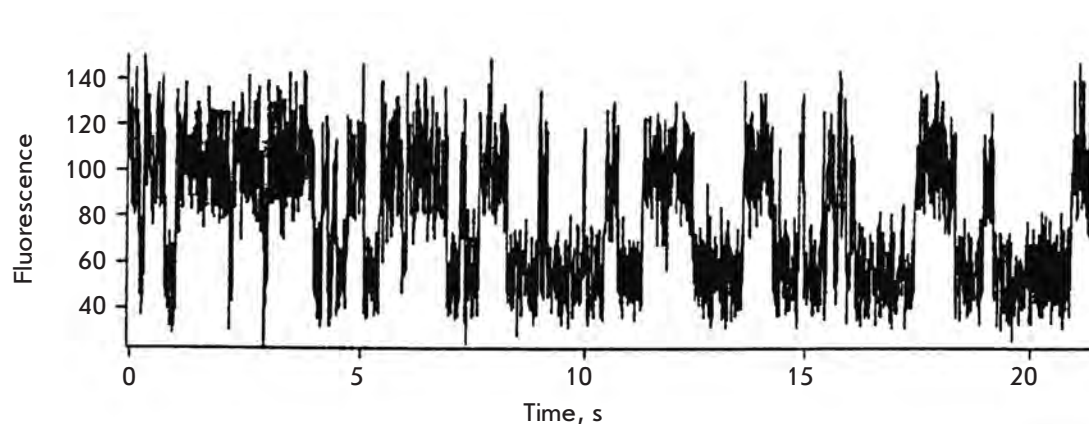
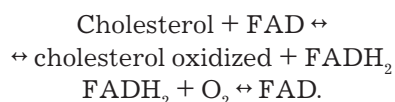


Fig. 5. Registration of cofactor fluorescence during catalytic oxidation of cholesterol.

ded proteins (functioning as receptors, transport and channel-forming proteins, etc.), it is simple to calculate that $\sim 4 \cdot 10^5$ proteins of 100 kDa can be accommodated therein.

Both *in vitro* and whole-cell studies provide data on the physicochemical characteristics of biochemical processes averaged on the entire set of the molecules involved in it. Therefore, the new possibilities that open up with the development of techniques for dealing with single molecules represent a new and important direction of research. On one hand, these investigations are aimed at elaborating methods for the sequencing of single DNA molecules; a considerable degree of progress has been made in such work in recent times [28]. The second direction is the investigation of reactions catalyzed by a single enzyme molecule. In this case, the reaction should be accompanied by a fluorescence change. Cholesterol oxidase (EC 1.1.3.6) [29], which catalyses cholesterol oxidation through molecular oxygen, can be used as an example. The reaction involves two stages:



Cholesterol is oxidized by fluorescent cofactor flavin adenine dinucleotide (FAD) bound to the protein moiety of the enzyme. During the process of cholesterol oxidation, FAD is transformed into the non-fluorescent reduced form FADH_2 . At the second stage of the reaction, FADH_2 is oxidized by molecular oxygen to the initial FAD. Each separate catalytic process is characterized by the attenuation and intensification of fluorescence, allowing one to follow each process of enzyme functioning. *Figure 5* shows the results of the registra-

tion of fluorescence upon the catalytic oxidation of cholesterol.

The investigation of macromolecule folding is another important aspect of the application of single molecule spectroscopy. Thus, single molecule fluorescence can be used to observe the dynamics of the formation of the spatial structure of RNA, which can also be recorded via the FRET (Forster resonance energy transfer) technique [30]. The intensity of the fluorescence energy transfer between the fluorophores bound to certain points of a fluorescence donor being irradiated and its acceptor is in inverse proportion to the sixth power of the distance between them. Any changes in the distance between the fluorophores during the formation of the spatial structure affect the fluorescence of an acceptor between them. The acceptor fluorescence will change with changes in the spatial structure.

The problems considered above, which arise in physicochemical biology, are related to proteins and nucleic acids, the investigation of which was a priority in research in the 20th century. When discussing new horizons in physicochemical biology, one should mention the demand for increasing the level of attention paid to other groups of compounds, with reference primarily directed at carbohydrates of an irregular structure, which play a significant role in the provision of a number of highly selective processes (i.e., the distribution of biochemical processes between cellular organelles). In addition to their cognitive significance, these directions will contribute significantly to the design of new drugs, the investigation of their interactions with living organisms, as well as their transformations and side effects. Therefore, these directions have the potential of becoming important elements of medicine in the 21st century. ●

REFERENCES

1. Watson G.D., Crick F.H. // *Nature*. 1953. V. 172. P. 737–738.
2. Avery O.T., MacLeod C., McCarty M. // *J. Exp. Med.* 1944. V. 79. P. 137–158.
3. Meselson M., Stahl F. // *Proc. Natl. Acad. Sci. USA*. 1958. V. 44. P. 671–682.
4. Lehninger A.L., Nelson D.L., Cox M.M. *Principles of Biochemistry*. N.Y.: Worth Publ., 2008. 1013 p.
5. Knorre D.G., Myzina S.D. *Biologicheskaya khimiya (Biological Chemistry)*. Moscow: “High School”, 1998. 479 p.
6. Saiki R.K., Scharf S., Faloona F., Mullis K.B., Horn G.T., Erlich H.A., Arnheim N. // *Science*. 1985. V. 230. P. 1350–1354.
7. Shchelkunov S.N. *Geneticheskaya inzheneriya (Genetic Engineering)*. Novosibirsk: Siberian University Publishment, 2004. 496 p.
8. Hutchison C.A. 3rd, Phillips S., Edgell M.H., Gillam S., Jahnke P., Smith M. // *J. Biol. Chem.* 1978. V. 253. P. 6551–6560.
9. Watson J.D. // *Science*. 1990. V. 248. P. 44–49.
10. Venter J.C., Adams M.D., Myers E.W., Li P.W., Mural R.J., Sutton G.G., Smith H.O., Yandell M., Evans Ch.A., Holt R.A., et al. // *Science*. 2001. V. 291. P. 1304–1351.
11. Kruger K., Grabowski P.J., Zaug A.J., Sands J., Gottschling D.E., Cech T.R. // *Cell*. 1982. V. 31. P. 147–157.
12. Guerrier-Takada C., Gardiner K., Marsh T., Pace N., Altman S. // *Cell*. 1983. V. 35. P. 849–857.
13. Hamilton A., Baulcomb D. // *Science*. 1999. V. 286. P. 950–952.
14. Vlassov V.V., Laktionov P.P., Rykova E.Y. // *BioEssays*. 2007. V. 29. P. 654–667.
15. Safo M., Moor N., Lavrik O. *The Aminoacyl-tRNA Synthetases*. Texas. Georgetown: Landes Biosciences, 2005. P. 250–270.
16. Callender R., Dyer R.B. // *Chem. Rev.* 2008. V. 106. P. 3031–3042.
17. Koshland D.E. // *Proc. Natl. Acad. Sci. USA*. 1958. V. 44. P. 98–104.
18. Hemmis G. *Metody issledovaniia bystrykh reaktsii (Methods of the Rapid Reactions Investigations)*. Moscow: Mir, 1977. 716 p.
19. Kuznetsov N.A., Zharkov D.O., Koval V.V., Buckle M., Fedorova O.S. // *Biochemistry*. 2009. V. 48. P. 11335–11343.
20. Anfirud P.A., Han C., Hochstrasser R.M. // *Proc. Natl. Acad. Sci. USA*. 1989. V. 86. P. 8387–8391.
21. Richard L., Genberg L., Deak J., Chiu H.-L., Miller R.J.D. // *Biochemistry*. 1992. V. 31. P. 10703–10715.
22. Gutman M., Huppert D., Pines E. // *J. Amer. Chem. Soc.* 1981. V. 103. P. 3709–3713.
23. Iu K.K., Kuczynski J., Fuernis S.J., Thomas J.K. // *J. Amer. Chem. Soc.* 1992. V. 114. P. 4871–4878.
24. Lehninger A.L., Nelson D.L., Cox M.M. *Principles of Biochemistry*. N.Y.: Worth Publ., 2008. P. 805–814.
25. Wright P.E., Dyson H.J. // *J. Mol. Biol.* 1999. V. 293. P. 321–331.
26. Dunker A.K., Silman I., Uversky V.N., Sussman J.L. // *Curr. Opin. Struct. Biol.* 2008. V. 18. P. 756–764.
27. Uverski V.N. // *Intern. J. Biochem. Cell. Biol.* 2011. V. 43. P. 1090–1103.
28. Pushkarev D., Neff N., Quake S. // *Nat. Biotechnol.* 2009. V. 27. P. 847–850.
29. Xie X.S., Lu H.P. // *J. Biol. Chem.* 1999. V. 274. P. 15967–15970.
30. Bokinsky G., Zhuang X. // *Acc. Chem. Res.* 2005. V. 38. P. 566–573.

Telomere Lengthening and Other Functions of Telomerase

M.P. Rubtsova^{1,2*}, D.P. Vasilkova¹, A.N. Malyavko¹, Yu.V. Naraikina³, M.I. Zvereva^{1,2}, O.A. Dontsova^{1,2}

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

²Belozersky Institute of Physicochemical Biology, Lomonosov Moscow State University, Leninskie gory, 1, str. 40, Moscow, 119992, Russia

³Lomonosov Moscow State University, Faculty of Bioengineering and Bioinformatics, Leninskie gory, 1, str. 73, Moscow, 119991, Russia

*E-mail: mprubtsova@gmail.com

Received 08.02.2012

Copyright © 2012 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 Telomerase is an enzyme that maintains the length of the telomere. The telomere length specifies the number of divisions a cell can undergo before it finally dies (i.e. the proliferative potential of cells). For example, telomerase is activated in embryonic cell lines and the telomere length is maintained at a constant level; therefore, these cells have an unlimited fission potential. Stem cells are characterized by a lower telomerase activity, which enables only partial compensation for the shortening of telomeres. Somatic cells are usually characterized by the absence of telomerase activity. Telomere shortening leads to the attainment of the Hayflick limit, the transition of cells to a state of senescence. The cells subsequently enter a state of crisis, accompanied by massive cell death. The surviving cells become cancer cells, which are capable both of dividing indefinitely and maintaining telomere length (usually with the aid of telomerase). Telomerase is a reverse transcriptase. It consists of two major components: telomerase RNA (TER) and reverse transcriptase (TERT). TER is a non-coding RNA, and it contains the region which serves as a template for telomere synthesis. An increasing number of articles focussing on the alternative functions of telomerase components have recently started appearing. The present review summarizes data on the structure, biogenesis, and functions of telomerase.

KEYWORDS telomerase; reverse transcriptase; telomeres; mitochondria; DNA damage; gene expression.

ABBREVIATIONS TER – telomerase RNA; hTR – human telomerase RNA; TERT – telomerase reverse transcriptase.

INTRODUCTION

The genetic information in eukaryotic cells is stored in linear DNA molecules known as chromosomes [1]. It was revealed as early as in the 1930s that the behavior of the whole chromosome and its fragments in cells varies. Torn chromosomes can fuse with each another, rearrange, and they are characterized by instability [2, 3]. An assumption was made back in the 1930s that these differences are caused by the presence of specific nucleotide sequences at the chromosome ends; these sequences were referred to as telomeres [3–5]. The telomeres consist of repeating sequences and a set of special proteins, which interact with these repeats and spatially organize them in a specific manner, resulting in the formation of the nucleoprotein complex known as telomeric heterochromatin [6, 7]. Shortening of the 5'-terminus of the daughter strand, caused by the removal of the terminal RNA-primer and the subsequent incomplete replication of linear DNA molecules, is observed during the genome replication occurring upon cell fission. The independent formulation of the so-

called “end-replication problem” was proposed in the 1970s by A.M. Olovnikov and J. Watson [8, 9]. Olovnikov hypothesised that there is a special enzyme, i.e. telomerase, which is capable of compensating for the “end-replication problem.” This enzyme was discovered in 1987 by C. Greider and E. Blackburn [10].

Telomerase consists of two major components: reverse transcriptase (TERT) and telomerase RNA (TER), which contains the template domain for the telomere repeat synthesis [9]. Furthermore, the telomerase complex contains numerous additional components that ensure the *in vivo* activity of the enzyme. Additional proteins participate in various processes. A number of these proteins are required in order for telomerase to attach to a telomere at a certain instant of the cell cycle [10], whereas the others serve to regulate the enzymatic activity [11]. It is already known that telomerase does not function in all higher eukaryotic cells; however, its components are almost always present in a cell. The data on the non-telomeric functions of telomerase components were recently reported.

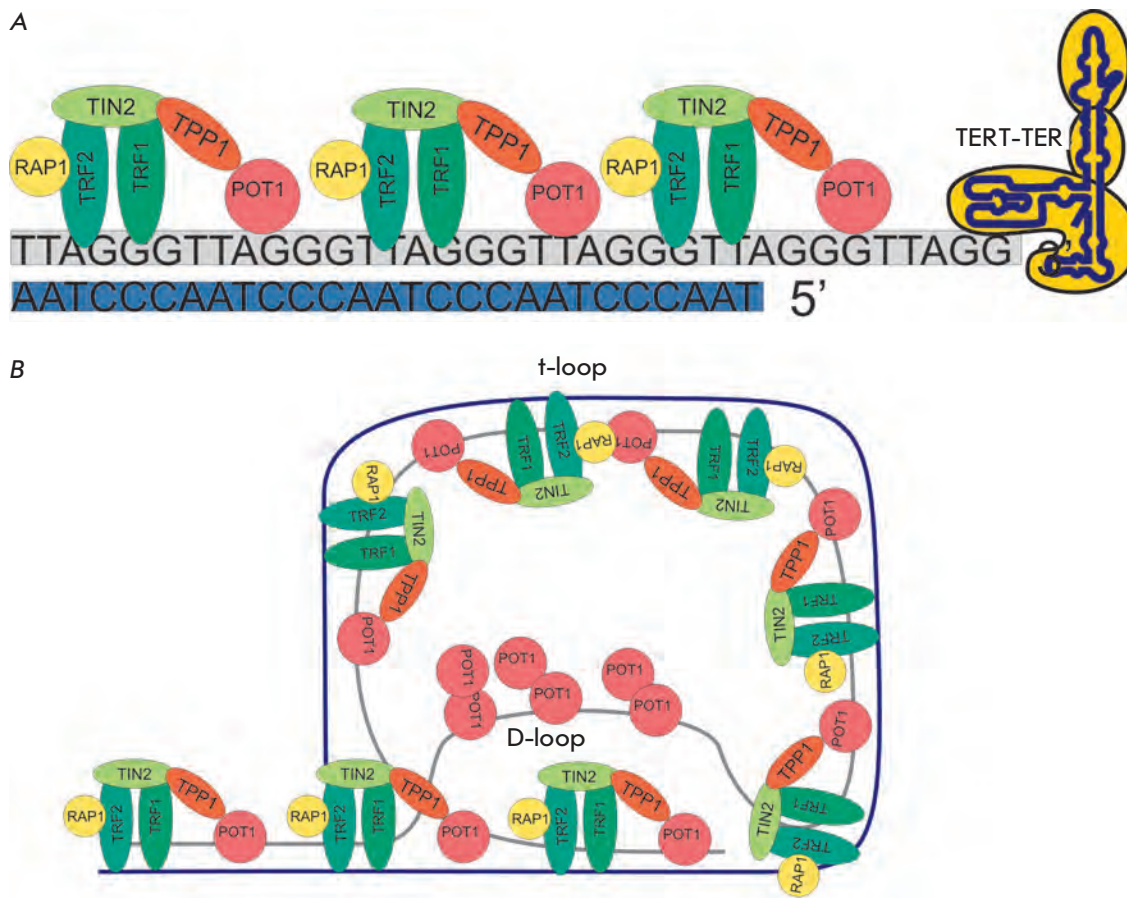


Fig. 1. Telomere structure. (A) – Schematic representation of the telomeric DNA complex, proteins of the shelterin complex and telomerase. (B) – Schematic representation of the shelterin complex bound to telomeric DNA, in the t-loop conformation.

TELOMERE STRUCTURE

Telomeres are the repeating nucleotide sequences bound to the specific proteins protecting chromosome ends against degradation and the double-strand break repair systems [12, 13]. As data accumulated, a hypothesis was postulated that telomeres consist of three distinct regions. Firstly, they contain the so-called cap, a terminal structure protecting the chromosome ends against degradation and the double-strand break repair systems (DDR – DNA damage response); they also regulate telomere elongation. The major component of a telomere is a double-stranded DNA (dsDNA) consisting of repeating and transcribed sequences. The third component of a telomere is represented by repeating telomere-associated sequences, the so-called subtelomeric regions [14, 15]. The telomere nucleotide sequence is enriched in thymidine and guanosine residues and is appreciably conserved. Mammalian telomeres are a double-stranded region consisting of TTAGGG repeats and the 150–200 nucleotide long 3' G-strand overhang. According to one of the hypotheses, the G-strand overhang is intertwined with the double-stranded telomeric region, thereby forming a t-loop. The so-called D-loop is formed at the site of the interaction between the pro-

truding 3'-terminus with the double-stranded region (Fig. 1). t-Loops were detected via electron microscopy after DNA was extracted and treated in a special manner. However, the existence of these structures in cells has as yet not been unequivocally proven; therefore, the D-loops are considered as tentative structures. Telomere functions depend on the minimal length of telomeric repeats and the activity of the protein complex associated with them. This complex is known as shelterin and consists of six proteins: TRF1, TRF2, POT1, TIN2, TPP1, and RAP1. The proteins TRF1, TRF2 (telomeric repeat binding factor 1 and 2) and POT1 (protection of telomeres protein 1) are bound to telomeric DNA. TRF1 and TRF2 are bound to the double-strand telomeric regions; POT1 can be bound to the 3'-protruding single-stranded region of the G-strand [16]. TRF1 and TRF2 bind telomeres independently; they do not interact with each other. Both proteins, which have the form of a homodimer and an oligomer, are capable of specifically binding the DNA duplex to the telomeric sequence 5'-YTAGGGTTR-3' [16–20]. POT1 binds highly specifically to the telomeric single-stranded DNA (ssDNA) 5'-TAGGGTTAG-3', attesting to a possible interaction both with the G-strand overhang and with the

sequence of the D-loop displaced by it [13, 21–23]. POT1 interacts with TRF1. It is believed that TRF1 facilitates the binding of POT1 to the single-stranded telomeric region in this manner. Via its independent domains, TIN2 (TRF1-interacting protein 2) simultaneously interacts with TRF1 and TRF2, as well as with the TPP1–POT1 complex, forming a bridge between the shelterin components [24, 25]. The C-terminal domain of TPP1 is bound to TIN2, the central domain is bound to POT1 [26–29]; thus, POT1 is attracted to the telomeres [30, 31]. Moreover, TPP1 contains a domain interacting with telomerase on its end. This fact supports the assumption that TPP1 attracts telomerase to the chromosome end [32]. Protein RAP1 forms a complex with TRF2 and the telomere [33, 34]. It has been demonstrated in studies undertaken by several research teams that RAP1 is not essential for telomere capping; however, it impedes recombination at telomeric regions and enhances their stability [35, 36]. Thus, RAP1 (unlike TRF1, TRF2, POT1, and TPP1) does not protect telomeres [32, 35, 36].

There is a hypothesis holding that G-quadruplex structures are formed in the telomeric regions of chromosomes. Four telomeric repeats can form a G-quadruplex, which inhibits telomerase activity [37–41]. The formation of these structures in ciliate cells has been clearly demonstrated using G-quadruplex-specific antibodies [42, 43]. The ability of the telomeric regions of higher eukaryotes to form G-quadruplex structures was indirectly supported by experimental data. According to [44], long 3'-protruding telomere ends form the G-quadruplex *in vitro*. The ligands binding the G-quadruplex structures are known to cause telomere shortening in cells. The telomerase activity remains intact, but the interaction between the shelterin complex and telomeric DNA is disrupted. Telomeres become unstable, and their binding to POT1 is disrupted, resulting in the activation of the DNA damage response system. This may be an indicator of the adverse effect of the stabilization of the G-quadruplexes in telomeric regions. These structures can presumably form in transition states; however, the telomeres cannot permanently maintain the structure of G-quadruplexes [45].

The chromosomes in eukaryotic cells are known to be packaged into chromatin by special proteins. It is believed that chromatin in a condensed state is untranscribed: euchromatin being associated with the cell transcription apparatus [46]. The telomeric regions of chromosomes form the so-called telomeric chromatin [47]. The assumption has been made that telomere elongation can depend upon the epigenetic status of telomeric chromatin [48]. Both the telomeric and subtelomeric regions are known to be enriched in histones that are typically bound to the repressed heterochromatic regions, such as histones H3 (H3K9m3) and H4 (H4K20)

trimethylated at lysine 9 and lysine 20 residues, respectively. Heterochromatin binding proteins 1 α , 1 β , and 1 γ (known as CBX5, CBX1, and CBX3, respectively) also bind to these regions [49–51]. Moreover, it has been ascertained that telomeric DNA is strongly methylated. The telomeres in the chromosomes in cells without N-methyltransferases (SUV420H1 and SUV39H1), which modify lysine residues in histones, are too long [49, 50]. The same was observed in cells with reduced methylation status of subtelomeric regions due to the deficiency in DICER1 or DNA-(cytosine-5)methyltransferases 1, 3A, and 3B (DNMT1, 3A and 3B) [52]. RNA containing telomeric repeats (TERRA – Telomeric Repeat containing RNA), or telomeric RNA – TelRNA, which is formed as a result of telomere transcription, was recently detected. These RNAs are capable of interacting with telomeric chromatin and of *in vitro* suppression of telomere elongation by acting as a potential telomerase inhibitor [48, 53, 54]. One can reasonably assume that the synthesis of TERRA cells is repressed in oncotransformed cells, rendering them incapable of suppressing telomerase activity.

Telomeric chromatin is dynamic, and its state may change. Differentiated somatic cells can be converted into induced pluripotent cells (iPS) via nuclear reprogramming [55]. The transition of the cells into a pluripotent state is accompanied by changes in the epigenetic status of telomeres: telomeric chromatin becomes less condensed; the histone content decreases, resulting in the subsequent formation of a large amount of TERRA; the level of telomere recombination becomes more frequent; and the telomere length becomes comparable with telomere length of embryonic stem cells [56]. Although no direct evidence exist so far to support the fact that the telomere length is regulated by changes in the chromatin state, the aforementioned observations lends credibility to the assumption that this theory is based on the truth.

TELOMERASE STRUCTURE

The assembly of telomerase, its existence in a cell, and its entry to the telomeres are processes that are similar in some aspects, yet differ in other aspects with regards to evolutionary distant organisms [57–59]. Properties common to all telomerase components have been revealed: reverse transcriptase (TERT), telomerase RNA (TER), and TER-binding proteins, which stabilize RNA and facilitate the assembly of the active enzyme. It should be noted that only TERT is a highly conserved telomerase component. The data obtained through the study of the components within telomerase are rather inconsistent [60–64]. Telomerase apparently interacts with various components throughout its vital activity and thus can be found in various complexes.

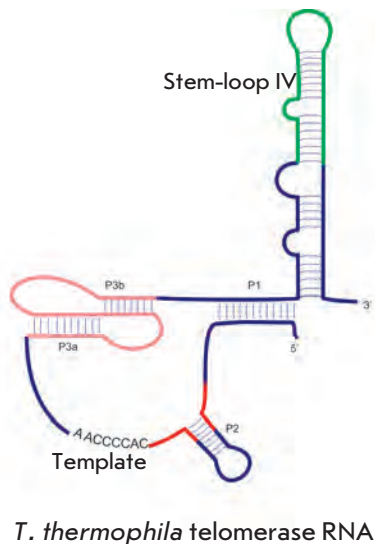
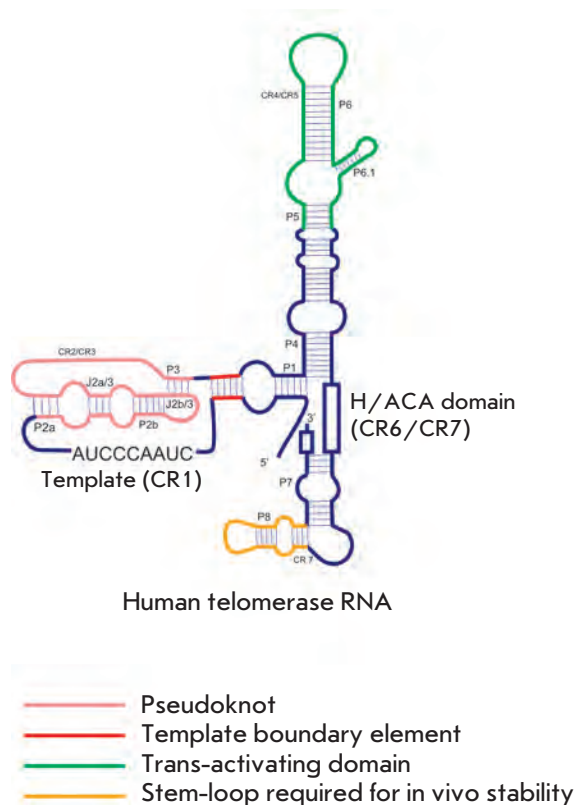


Fig. 2. Structure of telomerase RNA. The schematic representation of the secondary structures of human and *T. thermophila* telomerase RNA. The conserved elements are shown in color.

TER structure

Telomerase RNA is one of the major components of telomerase; it contains the region that acts as a template for telomere synthesis [65, 66]. Despite the differences in length and the nucleotide sequence of telomerase RNAs derived from different organisms, secondary structures of TER demonstrate high levels of similarity and contain similar structural elements [65, 67]. The template region, the pseudoknot, the trans-activating domain, and the domains required to ensure *in vivo* stability are the conserved elements of the TER structure (Fig. 2). Thus, TER contains the elements that are essential for telomerase activity and for the assembly, localization, and stability of RNA, but it is not required itself for enzymatic activity. The template domain of TER interacts with the 3' G-strand overhang of telomeres and guides DNA synthesis. This region can be single-stranded, although the differences in structure were detected via the analysis of the secondary structure of the transcript obtained *in vitro* and TER carried out in the *in vivo* experiments, attesting to its interaction with the other cell components [68–70]. Data indicating that a triplex structure is formed between the pseudoknot elements and the template domain was recently obtained by NMR spectroscopy. It is presumably the formation of this structure that can explain

the differences in the structure of the template domain of TER [71]. A hypothesis has also been put forth that in the absence of TERT and other necessary components, TER cannot form the correct structure. The template domain is flanked by two elements: the 5'-template boundary and the 3'-template recognition ones [72–75]. The 5'-element is a double-stranded region located immediately before the template domain; it regulates nucleotide addition during reverse transcription and, presumably, is the binding site with TERT. It has been demonstrated using mutagenesis that it is the secondary structure of this region, rather than the nucleotide sequence, that is of significance for the efficient functioning of telomerase. The 3'-recognition element is a single-stranded structure located after the template domain, which allows the 3'-terminus of the template to occupy the active site, stimulates telomerase activity and the ability to process after the repeats are added, and it contains the binding site of the N-terminus of TERT [76, 77].

Among the elements of the secondary structure of telomerase RNA, the pseudoknot has been the most intensely studied element. Changes in the stability of the pseudoknot result in a reduction in telomerase activity, which attests to the fact that this structural element plays an essential biological role [78, 79]. The recent

study of oligonucleotides representing the structural elements of the TER pseudoknot via NMR spectroscopy and molecular modeling have proved that it is the dynamics of the tertiary structure of the pseudoknot that plays a significant role in telomerase functioning [80–85]. The pseudoknot is formed due to the formation of the evolutionary conserved Hoogsteen triplet U*A*U between the U-rich loop 1 (J2b/3) and the major stem-loop (P3), which facilitates the maintenance of the structural integrity and is required for telomerase activity. Meanwhile, the A-rich loop 2 (part of J2a/3) enters the two other non-canonical triplet interactions, which facilitate the stabilization of the pseudoknot [86, 87]. Another Hoogsteen pair, A*U, is located between these two structural elements consisting of triplets. This pair is responsible for the stacking interaction between the two main stems, resulting in the formation of the final structure of the triple helix [85]. Nucleotide mutations inside the pseudoknot result in disintegration of the tertiary structure and considerably reduce telomerase activity, whereas compensatory mutations restore telomerase activity. These data confirm the fact that the tertiary structure has a more significant impact on the catalytic activity of the enzyme, compared to that of the nucleotide sequence [71, 87, 88]. It was assumed that the pseudoknot is necessary for correct orientation the template-primer duplex at the telomerase active site [71]. It is possible that the ability of this structure to exist in two conformations, the pseudoknot and the stem-loop, is of significance for telomerase functioning [83].

Unlike the pseudoknot, the structure of the trans-activating domain of TER has been subjected to less thorough study. The primary structures of this domain derived from different organisms are characterized by a high level of homology [7, 9, 87]. The trans-activating domain is a long stem-loop consisting of several, extremely stable helices broken apart by asymmetric loops and single-nucleotide bulges. This domain is required for the correct formation of the pseudoknot, nucleotide addition, and telomerase processivity upon repeat addition [88–90]. The P6.1 helix of the trans-activating domain of human TER is the one that has been best studied. This element is essential for enzyme functioning [90–93]. The role of the P6.1 in vertebrates has not been completely elucidated; however, it has been known that accurate structure of this helix is necessary for telomerase assembly, whereas specific sequences in the loops play a significant role in catalysis [93]. It is believed that the interaction of the P6.1 loop with the template domain yields the tertiary structure of TER, which thus explains the role of these enzymes in telomerase activity and ability to process [94].

The H/ACA domain, which is present in small nucleolar RNAs (snoRNAs) and in small Cajal body specific RNAs (scaRNAs), is located at the 3'-terminus of TER in vertebrates. The H/ACA domain is a single-stranded region containing the H-box (ANANNA, where N is a random nucleotide), the stem-loop that follows (containing the CAB-box), and the single-stranded 3'-terminus containing the ACA-box [94, 95]. The H/ACA domain is required to ensure the *in vivo* stability of telomerase RNA [96]. The CAB-box acting as a signal of localization in Cajal bodies is located inside this domain. The CAB-box does not participate in the 3'-terminal processing of telomerase RNA [97].

Data attesting to the fact that the first 17 nucleotides of human TER are essential for telomerase activity, and that absence of this region or mutation in it considerably reduces the enzymatic activity, have recently been reported. The ribooligonucleotide with this sequence was shown to form a G-quadruplex. It can be assumed that the structure of this element is likely to affect the structure of the P1 helix and positioning of the template domain of telomerase RNA [98–100].

Structure of telomerase reverse transcriptase

The telomerase catalytic subunit, TERT, is much more conserved in comparison with TER. It has a large number of motifs that are common to the other reverse transcriptases. Three domains can be distinguished in the structure of all known TERT: the RNA-binding domain (also known as TRBD and subdivided into RID1 and RID2), reverse transcriptase domain, and the poorly conserved C-terminal domain [57, 95, 101]. Certain TERT contain an additional N-terminal TEN domain, which is involved within the process of primer binding and facilitates the processive addition of telomere repeats [57, 102]. The primary structure of the reverse transcriptase domain is similar to the structures of other polymerases and contains seven conserved motifs (1, 2, A, B, C, D, and E). It is assumed that TERT originates from retrotransposons [103]. Intron-containing (the so-called Penelope-like) elements are the ones most similar to TERT.

Structures with high-resolution TEN and TRBD TERT domains and full-length TERT were recently derived from *Tetrahymena thermophila* [102, 103] and *Tribolium castaneum* [104], respectively; they showed new features of telomerase structure and function. As follows from the analysis of the structure of TERT derived from *T. castaneum*, high structural homology exists between TERT and other polymerases, including the p66 subunit of HIV reverse transcriptase [105]. The spatial arrangement of the major domains and of the key amino acid residues in them remains constant and corresponds to the right-hand structure that was first

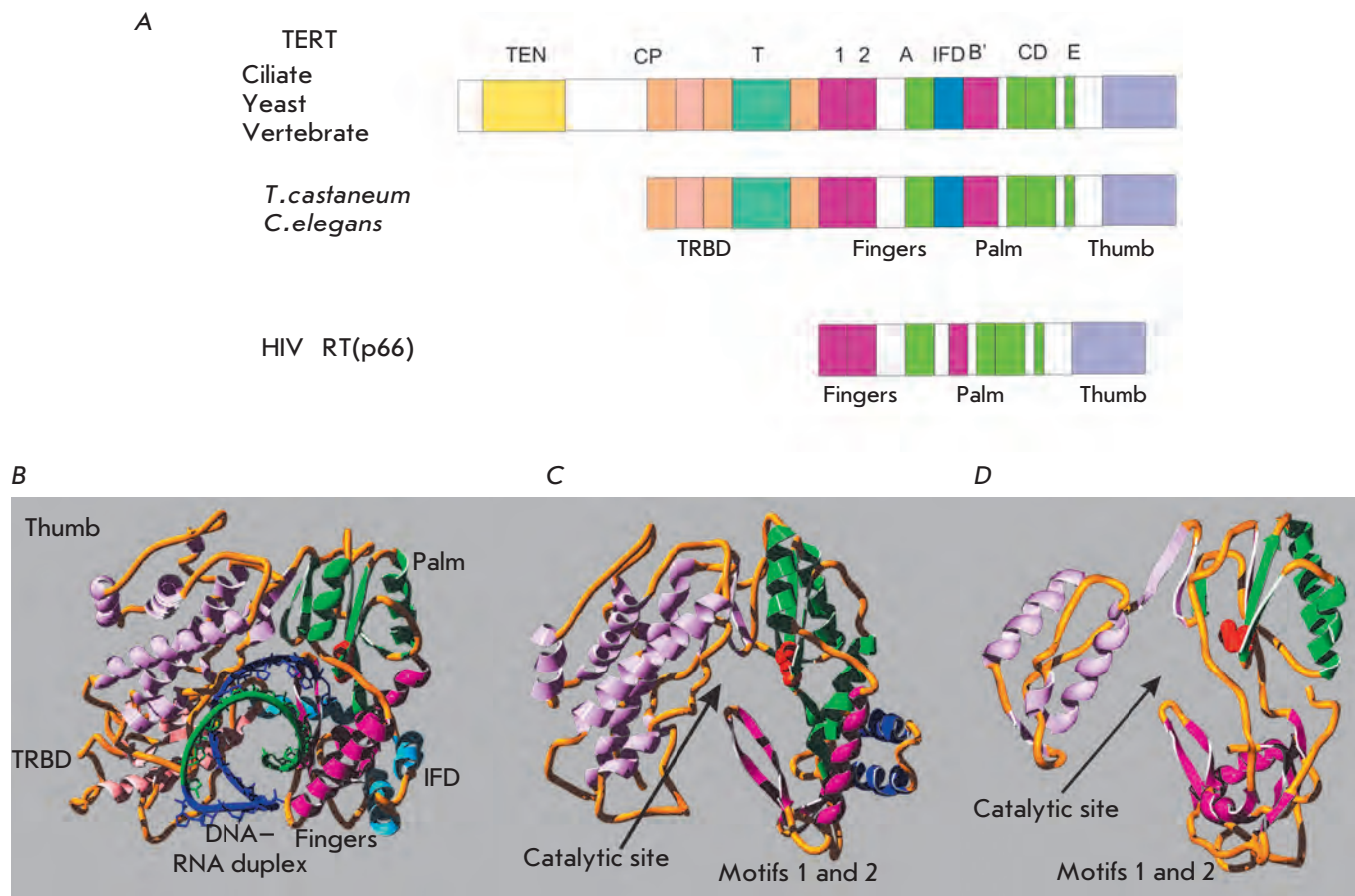


Fig. 3. Structure of telomerase reverse transcriptase. (A) – Schematic representation of the domain organization of TERT in different organisms and HIV reverse transcriptase. Homologous domains are shown in color. (B) – Spatial structure of the *T. castaneum* TERT complex and the RNA–DNA duplex. The image was obtained using the PDB file 3KY. (C) – Spatial structure of *T. castaneum* TERT. The image was obtained using the PDB file 3DU5. (D) – Spatial structure of HIV RT. The image was obtained using the PDB file 1N6Q. Asn residues of the enzyme catalytic sites are shown in red.

described for the Klenow fragment of DNA polymerase I from *Escherichia coli*. It is common to distinguish the so-called palm, fingers, and thumb subdomains in these structures [106]. It was revealed that the loop responsible for the binding and positioning of the template and nucleotide is located between the β -sheets of motifs 1 and 2 and oriented towards the active site [107, 108]. Meanwhile, there also are differences in the structure of TERT and other polymerases (Fig. 3). Thus, TERT contains an additional domain known as IFD (insertion in fingers domain). This domain is located outside the central ring between the “fingers” and the “palm.” It is clear from the structure of TERT derived from *T. castaneum* that three completely conserved domains form a ring-like structure. The conserved reverse transcriptase domain forms “fingers” and a “palm” (similar to the other polymerases) and occupies one side of

the ring, whereas the C-domain forms the “finger” and is located in direct proximity to the N-terminal RNA-binding domain and thus closes the ring. The inner diameter of this structure is 26 Å, and its depth is 21Å; approximately corresponding to the size of the A-helix consisting of 8 bp. The DNA–RNA duplex located in the polymerase active site has precisely this type of structure [109]. The surface of the opening is a spiral structure consisting of positively charged amino acid residues. This structure allows the protein to make close contact with the RNA–DNA heteroduplex [110]. The helices 10 and 19 interact with the major and minor grooves of the DNA–RNA duplex located in the active site [111–115]. The nucleotide binding domain is located at the “fingers”–“palm” interface; this fact is attested to by the high levels of similarity with the other polymerases. Several conserved amino acid resi-

dues forming the nucleotide-binding pocket have been determined [115]. The active site of the enzyme contains several unchangeable aspartic acid residues and a lysine residue, which activates the pyrophosphate leaving group.

As follows from an analysis of its structure, the TRBD domain of *T. thermophila* and *T. castaneum* TERT is enriched in helical structures and is divided into two parts. These two structural elements are bound via CP- and T-motifs. The CP-motif contains a positively charged pocket, whereas the T-motif is a narrow hydrophobic gap with positively charged residues near the CP-motif. Together, they form an extensive groove on the surface of TRBD, to which TER is bound [100, 103, 116–119]. The T-motif contains a β -stem-loop stretched in the direction of the C-terminal domain, which forms the “thumb,” and links it to the “fingers” of the reverse transcriptase domain. This arrangement of the TRBD domain allows the residues located on the inner side of the ring to come in direct proximity with the active site. Moreover, the gap between the TRBD and reverse transcriptase domains enables TER to get into the active site. One can assume that telomerase RNA penetrates through this gap, thus carrying the template into the enzyme’s active site [120].

TELOMERASE BIOGENESIS

As previously mentioned, telomerase consists of the two major components; however, the synthesis and processing of each component, as well as the formation of an active enzyme, require the contribution of a large number of additional factors. The regulation of TERT expression at the transcriptional phase was thoroughly discussed in the review by Skvortsov *et al.* [121]. The alternative splicing of the primary transcript of the *hTERT* gene yields 13 different mRNA variants [122–125]. Out of these variants, the so-called α - and β -forms are both the most common and well-studied. When the α -form is produced, 36 nucleotides are deleted from the sixth exon, resulting in the change of the reverse transcriptase motif A. The open reading frame is not disrupted [126, 127]. The deletion of 182 nucleotides from the exons 7 and 8 and the insertion of 38 nucleotides triggers a premature translation termination, resulting in the formation of the β -*hTERT*, which does not contain the three essential reverse transcriptase motifs [128, 129]. Splicing can independently occur at different sites; therefore, different forms of *hTERT* mRNA often co-exist in cells. The combination of different mRNA forms and their number depend on the particular cell type. Thus, one of the mRNA variants (α -/ β + form) has regulatory functions by acting as a dominant negative inhibitor of telomerase activity both in normal and tumor cells.

It remains unclear as to whether the ratio between the full-length *hTERT* mRNA and its spliced forms affects telomerase activity. The total level of *hTERT* expression was shown to correspond to the level of telomerase activity in some studies [126–129], whereas no regularities in the alteration of telomerase activity with the variation of any *hTERT* mRNA form have been revealed [130]. It is assumed that the regulatory functions executed by the products of the alternative splicing of *hTERT* are cell-type dependent. The set of *hTERT* transcripts changes during human embryonic development. During the early stages, all tissues contain full-length *hTERT* mRNA and active telomerase; subsequently, the set of mRNA forms changes depending on a tissue’s type [131]. One can assume that the variation of TERT expression is required for cell differentiation during the development of the organism.

Reversible phosphorylation of TERT plays a significant role in the regulation of the telomerase function [132]. Numerous kinases and phosphatases are already known. They affect the phosphorylation of serine, threonine, and tyrosine residues, thus changing the structure, localization, and activity of enzymes. Non-specific phosphorylation sites have been detected in the primary structure of *hTERT*; however, only a few of these can be modified, and their phosphorylation affects telomerase activity.

Telomerase RNA belongs to the family of non-coding RNAs; i.e., it does not act as a template for protein synthesis. As mentioned above, telomerase RNA contains structural elements that are characteristic both to small nucleolar RNAs and to small Cajal body specific RNAs. All human H/ACA RNAs are encoded by introns, which are synthesized in the form of pre-mRNA and are processed to yield mature RNAs without the cap structure at their 5’-terminus [133]. In contrast, human TER is transcribed by RNA polymerase II from its own promoter. The processing of the primary transcript results in the formation of the 451 nucleotide mature form containing a trimethylguanine cap at its 5’-terminus. The processing of telomerase RNA has been partially studied in yeast cells. It is known that the 3’-terminal processing of TER in *Schizosaccharomyces pombe* cells is performed by a spliceosome. Only the first splicing stage resulting in the release of the 5’-terminal exon is required for the formation of active telomerase RNA [134]. Exon ligation would yield the rapidly degradable product. It is unclear how the splicing is terminated at the intermediate stage.

Trimethylation of 5’-terminal guanine in TER in yeast cells is performed by methyltransferase Tgs1. It is assumed that in vertebrates this enzyme (hTgs1p) also participates in the hypermethylation of the 5’-cap of TER in Cajal bodies, in which it is contained [135].

The study of the processing of human telomerase RNA is complicated by its low content in cells. Human telomerase RNA is expressed and undergoes processing in yeast cells [136]. Both polyadenylated and non-polyadenylated, processed and non-processed hTER forms are produced during expression in yeast cells. Processing of hTER in this system is performed by yeast proteins Cbf5p (dyskerin homolog), Nhp2p, and Nop10p, which participate in the processing of small nucleolar RNAs containing the H/ACA-domain [137–139]. H/ACA-proteins are bound to the H/ACA-domain of telomerase RNA, which determines the 3'-boundary of the mature hTER. It is assumed that during the processing of hTER, its 3'-terminus is cleaved via exonucleases, whereas the H/ACA-proteins that have been bound determine the boundary of the mature hTER form [136]. A hypothesis that nucleases are activated as a result of H/ACA-protein binding to telomerase RNA has also been postulated [140].

It was ascertained in 2011 that the telomerase complex contains the DHX36 protein, or RHAU (known as RNA helicase). It also participates in the degradation of mRNAs containing AU-rich elements and is the resolvase both for the DNA and RNA of G-quadruplexes [141–145]. It was found that this protein interacts with the 5'-terminal region of the hTER forming the G-quadruplex and stabilizes hTER. This occurs before the 5'-terminal guanosine is trimethylated by telomerase RNA and is presumably required to protect hTER against degradation. Once hTER is capped, the formation of the G-quadruplex is no longer feasible and RHAU is no longer able to bind to hTER [146]. Furthermore, RHAU stimulates the formation of helix P1, thereby providing the correct positioning of the template domain of hTER [147].

The telomerase complex contains additional proteins, which participate in enzyme biogenesis [148]. The telomerase complex always contains the RNA-binding protein dyskerin, which is capable of recognizing the H/ACA-motif both in telomerase and in the other non-coding RNAs (small nucleolar and Cajal body specific RNAs) [149, 150]. Dyskerin is believed to participate in the biogenesis of telomerase RNP (ribonucleoprotein) and maintains the stability of telomerase RNA [151, 152]. The telomerase complex may contain the dyskerin-binding proteins NOP10, NHP2, and GAR1 [153, 154]. DNA helicases pontin and reptin, which exhibit ATPase activity, interact with hTER, hTERT, and dyskerin [155]. The content of the hTERT complex with these proteins is highest in the S-phase of the cell cycle. Telomere elongation in yeasts occurs precisely at this moment [156], whereas in human cells telomerase is associated with telomeres [157, 158]. It can be reasonably assumed that pontin and reptin can

affect the regulation of the hTERT content at different phases of the cell cycle, or affect the assembly of active telomerase in the S-phase. Dyskerin is permanently bound to hTER, whereas pontin and reptin interact with hTERT. In the S-phase, pontin and reptin interact with dyskerin by participating in a *de novo* formation of the telomerase RNP.

One of the recent studies is devoted to the identification of another protein participating in the assembly and effecting of telomerase activity [159]. It was shown that this protein is ATPase NVL2. It was demonstrated using the two-hybrid system that hTERT interacts with the NVL2 protein. The NVL gene encodes two isoforms of NVL ATPase (NVL1 and NVL2), which belong to the AAA (ATPase associated with a variety of cellular activities) family of ATPases [160, 161]. hTERT interacts with both isoforms; however, the complex with NVL2 turned out to be stronger. In cells, hTERT is co-localized with NVL2, which contains two ATPase domains. The Lys311 mutation in the first domain disrupts the binding of this protein to hTERT, whereas NVL2 knockdown reduces the telomerase activity in cells [159].

The telomerase holoenzyme contains the WDR79/TCAB1 protein (telomerase Cajal body protein 1) [58]. Cajal bodies are enriched in this protein, which is associated with TERT, TER, and dyskerin. Meanwhile, TCAB1 does not interact with the telomerase assembly factors NAF1, pontin and reptin. It is assumed that pontin and reptin at the first stage of telomerase maturation facilitate assembly of the minimally active enzyme consisting of TERT, TERC, and dyskerin. TCAB1 subsequently interacts with active telomerase and determines its localization in Cajal bodies, thus facilitating the binding to telomeres.

The data regarding the architecture of the telomerase holoenzyme lack consistency. Thus, telomerase has been demonstrated to possess catalytic activity only when in dimeric form; however, it has also been claimed that dimerization is not a prerequisite under physiological conditions [162–164]. The immunoprecipitation method was used to study the composition of the proteins isolated along with telomerase [165]. It turned out that telomerase can form several complexes differing by the proteins they consist of. It was assumed that the composition of the telomerase complex changes during maturation. At the first stage, the H/ACA-proteins are bound to the 3'-terminal stem-loop of hTER; the second complex of H/ACA-proteins with the GAR1 protein subsequently interacts with the stem of the CR4/CR5 stem-loop. At the second stage, GAR1 is replaced from the telomerase RNP by the TCAB1 protein and is bound to hTERT. Meanwhile, TERT and TCAB1 are present in the complex at a substoichiomet-

rical ratio. This fact may be evidence of the existence of either a single complex containing both components, or two different complexes, each of those containing one of the components. It was assumed that both complexes are simultaneously present in cells; however, they permanently exchange components.

The interaction between telomerase and the other proteins (La, Staufen, L22, hnRNP, C1/C2, TEP1, p23, and Hsp90) is required for the formation of the appropriate structure and its stabilization [132, 166, 167]. However, it is unclear whether these proteins affect the telomerase activity.

TELOMERE ELONGATION

The major activity of telomerase ensures the RNA-dependent telomere elongation [168]. The telomerase catalytic cycle consists of several sequential stages. One telomeric repeat is added after the substrate binding. The resulting product can either dissociate from the enzyme's active site or undergo translocation, followed by elongation. The ability of telomerase to move the DNA synthesized to the template beginning site allows one to use two processivity types to describe its function. Nucleotide addition (type I processivity) is intrinsic to all polymerases, since repeat addition (type II processivity) is unique to telomerase and determines the ability of an enzyme to repeatedly copy an RNA template region via elongation of the one substrate molecule only [169, 170].

Primer binding at the first stage of the telomerase reaction cycle is stipulated by its complementary action with the TER template region. When using primers with different sequences, the efficiency of formation of the complex with an enzyme does not correlate with the length of the resulting DNA–RNA duplex [171], since telomerase is bound to the substrate upon immediate participation of not only the template region of telomerase RNA. The structural elements of the TERT active site regulate the efficiency of duplex formation, as well as translocation of the freshly synthesized product during the processive synthesis of telomeric repeats. The anchor regions in TERT and TER also participate in the primary binding of the primer.

Nucleotides are bound to the primer at the second stage of the telomerase reaction cycle [168, 172].

The major feature of telomerase is its ability to processively add the repeats [170]. The mechanism of telomerase translocation after a repeat is synthesized remains unknown. It remains ambiguous as to whether enzyme processivity of this type is required for efficient telomere elongation or not. It has recently been ascertained that critically short telomeres elongate processively [173]. A set of products differ in the number of telomeric repeats is formed during tel-

omerase operation. After a single telomeric repeat is added, the reaction is either terminated or the rate of reaction decreases; i.e., template translocation and annealing are the rate-limiting stages. It has been demonstrated that POT1 and TPP1 proteins efficiently stimulate telomerase processivity [174]. An assumption was made that telomerase processivity is regulated by the POT1–TPP1 complex. Telomerase activity is inhibited when the complex is bound to the 3'-terminus of the primer. When it is bound to the 5'-terminus, telomerase functions processively.

It is known that telomerase is not active in all cells. Nevertheless, telomerase RNA occurs in all cells; reverse transcriptase occurs in the majority of cells. The localization of telomerase components does not necessarily coincide with the site of its "operation." Telomerase RNA frequently occurs in the cytoplasm; reverse transcriptase is found in mitochondria and other organelles. These data enable to assume that telomerase can have additional functions in the cell rather than just maintenance of the telomere length.

ALTERNATIVE FUNCTIONS OF TELOMERAISE COMPONENTS

The first batch of data on the alternative functions of telomerase were reported at the early stages of the study of this enzyme. The products of other enzyme activities were detected when studying the activity, substrate specificity, and other properties of telomerase. It turned out that telomerase is also capable of acting as a catalyst for the other reactions (*Fig. 4*).

Telomerase nuclease activity

It was demonstrated through a study of the catalytic activity and substrate specificity of telomerase derived from *Thermus thermophilus* that the length of the resulting product depends on the complementarity degree of a primer and the template region of telomerase RNA [175]. If the 3' terminus of the primer is non-complementary to the template region, a break at the boundary between the coupled and non-coupled substrate regions occurs. Moreover, the break is also possible in a entirely complementary primer. In this case, the site and possibility of a break will depend on the length and site of preferential primer annealing at the template region. Thus, telomerase derived from *Th. thermophilus* exhibits nuclease activity. It was later ascertained that human and yeast telomerases also exhibit such activity [176–181]. A thorough investigation into the mechanism of endonuclease activity has demonstrated that a substrate can be cleaved even when it is completely complementary to the template so that it could be preferentially located in the telomerase cat-

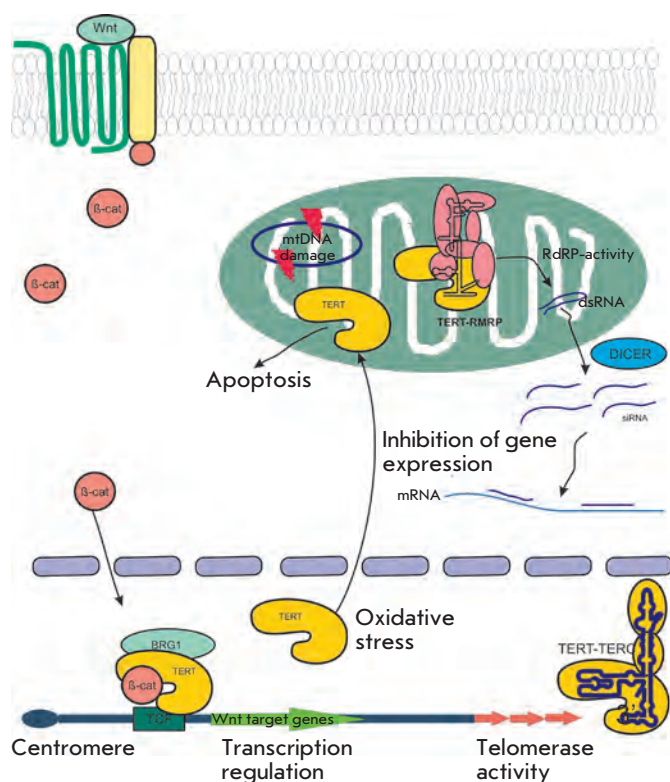


Fig. 4. Telomerase is the enzyme with versatile functions. Schematic representation of the mechanisms of telomerase functions in the cell.

alytic site. The endonuclease activity of telomerase is not sequence-specific. The primers having non-hydrolyzable internucleotide bonds at preferential cleavage sites are subjected to cleavage at other sites [179, 182].

Transferase activity

Yeast and human telomerase can also exhibit transferase activity. In the presence of Mn^{2+} -ions, telomerase can add nucleotides independently of the template. Preference is given to telomere-like primers with GT-rich 5'-end. It is unknown whether there can be *in vivo* situations when the intracellular Mn^{2+} concentration reaches values at which transferase activity is detected *in vitro*. Nevertheless, it has been assumed that certain small molecules can stimulate this ability of telomerase [183, 184].

Telomerase and mitochondria

It was revealed during the early stages of the investigations into telomerase that hTERT is also expressed in cells where telomerase activity is not detected [185, 186]. It has recently been reported that the hTERT pro-

tein occurs in each somatic cells, predominantly in the S-phase [187]; moreover, it is present not only in the nucleus, but it is also present in cytoplasm and mitochondria [188–194]. In immortalized human fibroblasts under treatment with H_2O_2 and oxidative stress conditions, hTERT is exported from the nucleus and transferred to the mitochondria [194]. Mouse embryonic fibroblasts are known to undergo earlier senescence if cultivated in conditions with increased oxygen content [195]. As shown in similar experiments using human cells, oxygen deficiency results in an increased life span [196, 197]. The oxidative stress can activate the tumor suppressor proteins p53 and Rb [198, 199]. There are several factors that cause cellular senescence under oxidative stress conditions. Firstly, they include DNA damages effected upon oxidative stress and induce the activity of the regulators of the cell cycle p21 and p16, which facilitate cellular senescence and cell cycle delay [200, 201]. As a result of decreased proliferation, telomeres critically shorten; subsequently, the cells experience a crisis. At this stage, the cells either die, or telomerase is activated, making them immortal. Secondly, oxidative stress can cause damage of telomere. The DNA within the telomeres is enriched in guanine residues that are sensitive to oxidation. Due to the oxidation of these residues telomeres become more sensitive to damage, their degree of shortening increasing [202, 203]. It has been demonstrated that the treatment of cells with mitochondria-targeted antioxidant MitoQ reduces the level of telomere damage and prolongs the life span of the fibroblasts subjected to oxidative stress [204]. Thirdly, premature senescence induced by oxidative stress may be a result of the direct inactivation of telomerase activity.

It was predicted recently that the N-terminus of TERT contains the mitochondria transport signal (MTS) consisting of 20 amino acid residues [205]. MTS is highly conserved in the TERT of higher eukaryotes, such as plants, fish, and mammals; however, yeasts and ciliate organisms contain none of this sequence [192]. The green fluorescent protein (GFP) with MTS localizes in mitochondria. The *in vitro* synthesized protein A containing MTS from hTERT at its N-terminus is transported into the purified mitochondria through the membrane potential [205]. Various methods have been used to detect the localization of hTERT in mitochondria, including immunoblotting and coimmunoprecipitation [191, 192, 194, 206–208].

Mitochondrial extracts from various human cells exhibit telomerase activity. hTERT is located in the mitochondrial matrix and is co-precipitated with the proteins TOM20, TOM40, and TIM23 [191]. It has also been demonstrated that hTERT is extracted along with the mtDNA-binding proteins TFAM, HSP60, and

TIM23, but not with TOM20 [205]. Increased hTERT content in mitochondria subjected to oxidative stress results in the stabilization of mtDNA and stimulates the mitochondrial function. This suppresses the formation of reactive oxygen species and increases the potential of the mitochondrial membrane [194]. hTERT interacts with the mtDNA regions encoding the subunits 1 and 2 of the NADH-ubiquinone oxidoreductase (ND1 and ND2). Changes in respiratory chain activity were observed in mouse cardiac (but not hepatic) cells expressing hTERT [191]. Nevertheless, non-specific interaction between hTERT and mtDNA was later detected [205]. The chromatin immunoprecipitation assay modified for mitochondria was used to demonstrate that telomerase reverse transcriptase interacts both with the regions encoding ND1, ND2 and with 12S and 16S rRNA, ND4 and ND5, COXI and COXIII, tPHK and the subunits 6 and 8 of ATP-synthase [205]. It is known that increased expression of hTERT in human fibroblasts does not prevent the stress-induced senescence, but protects them against apoptosis and necrosis [209]. The opposite effect of increased expression of hTERT under oxidative stress (i.e., increased degree of DNA damage) has also been demonstrated [192, 206]. The content of a bio-accessible iron, which can stimulate the formation of the hydroxyl radicals that damage DNA, increases in these cells [192].

The factors responsible for the intracellular transport of hTERT remain unknown. hTERT contains the mitochondria transport signal and the nuclear export signal (NES). It has been hypothesized that hTERT contains several nuclear localization signals (NLS) [210, 211]. However, the regulation mechanism of the localization of this protein inside the cell remains unclear. It is a known fact that the intracellular distribution of hTERT changes under oxidative stress conditions due to posttranslational modifications [189, 194, 208, 210, 212–216]. Src kinase was shown to regulate the export of hTERT from the nucleus to the cytoplasm under oxidative stress [189, 194], whereas dephosphorylation of hTERT by phosphatase Shp-2 results in the import of hTERT from the cytoplasm into the nucleus [212]. During oxidative stress, Src kinase phosphorylates Tyr707 in TERT. The modified hTERT interacts with the nuclear pore component, protein Ran, followed by its export from the nucleus with the participation of karyopherin CRM1. Treatment of the cells with hydrogen peroxide reduces the level of wild-type mitochondrial hTERT, whereas the level of hTERT, in which Tyr707 is substituted by Phe and thus cannot be phosphorylated by Src kinase, remains unchanged. The mutant hTERTY707P expressed in cells is accumulated in the nucleus during the oxidative stress, the apoptosis level of these cells being lower than that of the cells contain-

ing the wild-type hTERT [189]. It has also been shown that treatment of cells with H₂O₂ results in an increase in the level of wild-type mitochondrial hTERT for several hours, whereas it takes several days for this effect to develop under hyperoxic conditions. hTERT returns to the nucleus in the cells that returned to the normal state after being cultured under hyperoxic conditions [194].

Several years ago it was revealed that hTERT can form complexes with the RNA-component of mitochondrial RNA that processes endoribonuclease RMRP, along with the complexes with hTER. It turned out that the hTERT-RMRP complex exhibits the activity of RNA-dependent RNA-polymerase and synthesizes double-stranded RNAs using the RNA-component of RMRP as a template. These RNAs are processed by the enzyme complex DICER yielding small interfering RNAs (siRNA), which subsequently reduce the RMRP cell level; i.e., the RMRP function is regulated according to the principle of negative feedback [217].

It has been reliably ascertained that hTERT functioning in mitochondria is hTER-independent, and hTER is not imported into mitochondria [205]. It turned out that mitochondrial tRNAs are released from the mitochondria, along with hTERT; the mitochondrial tRNAs act as primers in the reverse transcription reaction catalyzed by hTERT. Meanwhile, this reaction can be inhibited by the addition of hTER and mutation in one of the reverse transcriptase domains of hTERT. If the mitochondria contains no hTERT, the result is mitochondrial dysfunction. It was assumed that hTERT can participate in the replication and repair of mtDNA [205].

It follows from the aforementioned facts that data regarding the functions of hTERT, and the components that interact with it in mitochondria, lack consistency. This fact can most likely be attributed to the limited amount of enzymes in a cell. All the studies devoted to the functions of telomerase in mitochondria have been carried out under conditions of protein overexpression. The additional domains that are used to extract and detect the protein can be the reason behind the unreliable results. It should be noted that at the time of writing, no clear-cut opinion has emerged within the research community regarding the functions of telomerase in mitochondria.

Telomerase and DNA damages

Non-functional telomeres are known to interact with a set of proteins involved into the DNA damage response [218–220]. These proteins participate in signal transmission in response to different impacts. ATR and ATM belong to the family of protein kinases related to phosphoinositide 3-kinases (PIKK) [221]. ATM is the

major protein activated on double-strand DNA breaks (DSB) [222], whereas ATR is activated when single-stranded DNA ends emerge upon the formation of DNA adducts, during DSB processing, or during termination of the replicative fork [223, 224]. The absence of ATM results in telomere decapping and shortening; furthermore, TRF2 is bound to ATM kinase to inhibit its activation [225]. Suppression of ATR kinase activity is observed under conditions of increased expression of hTER. On the contrary, the decrease in the amount of telomerase RNA in cells facilitates an increase in ATR activity. These processes are independent from the level of telomerase activity and telomere length. A reduced level of hTER expression in cells results in an increase in the amount of protein p53, the tumor suppressor and the major contributor to the signalling pathway upon oncogenic stress. Meanwhile, the cell content of the protein CHK1, the cell cycle regulator, increases. p53 and CHK1 are the major substrates of ATR kinase. hTER inhibits ATR kinase and disrupts the regulation of the cell cycle checkpoints upon DNA damage *in vivo* [226].

The mutations in the template region of telomerase RNA induce a decrease in the level of the protein TRF2, which stimulates the apoptosis. This effect is ATM-dependent. ATM activation results in the phosphorylation of p53, which in turn activates the transcription of the *GADD45 γ* gene. The increase in the *GADD45 γ* level results in the apoptosis. Thus, mutations in the template region of telomerase RNA trigger DNA damage; the cells systems consider these damages to be double-strand breaks [227].

Histone H₂AX in eukaryotic cells is phosphorylated by ATM kinase in response to DNA damage. The phosphorylated γ H₂AX is bound to DNA at double-strand-break sites. In cells without hTERT, which were exposed to ionized radiation, the DDR system ceases to function. In fibroblasts with hTERT expression stably suppressed by RNA interference, the ATM and γ H₂AX content is reduced. The telomere length in these cells changes to a negligible degree, whereas the chromatin structure and post-translational modifications of histones are changed [187]. It is already known that the frequency of spontaneous chromosome breaks at the G1-phase decreases by 20-fold and the ATP level increases in human foreskin fibroblasts upon overexpression of hTERT [228]. The presumable reason is that hTERT has a protective role in mitochondria. ATP is required for the functioning of chromatin remodeling factors [229] and activation of ATM kinase [230]. Whilst protecting the mitochondria, hTERT presumably has a mediated effect on ATP synthesis in a cell, as well as on all the processes for which its hydrolysis is required.

Telomerase and regulation of gene expression

The development of methods for the investigation of cell functioning and gene expression has enabled to study how the activity of some genes affects the expression of other genes. The cDNA microarray was used to ascertain that the expression level of 284 genes changes in bovine adrenal cells that overexpress TERT [231].

Nowadays, it is a known fact that telomerase can affect the cell cycle via regulation of the expression of various genes. An increase in the TERT level enhances the proliferative potential of human bone marrow stromal cells [232] and results in hyperplasia and hypertrophy of murine cardiomyocytes [233]. The increase in the level of *hTERT* expression in human breast epithelial cells with the deleted *p16* gene makes them resistant against the antiproliferative effect of the transforming factor β (TGF- β) [234]. Meanwhile, no dependence between the telomere length and cell sensitivity to TGF- β has been detected. It is also known that telomerase activation in human breast epithelial cells stimulates their processing into mitosis [235].

Evidence to the fact that telomerase affects the pRB/E2F signalling pathway has been obtained. Cyclin D upon mitotic stimulation form a complex with CDK4 and CDK6 and phosphorylate and simultaneously inactivate the retinoblastoma protein pRB. As a result, the interaction between pRB and the E2F transcription factor is disrupted. E2F is activated, triggering the expression of the genes required for cell transition from the G1- to the S-phase. Overexpression of hTERT in human crystalline lens cells induces their growth. Meanwhile, hyperphosphorylation of pRB and expression inhibition of p53, p21, and GCIP are observed [236]. p21 and GCIP are the inhibitors of cyclin complexes with cyclin-dependent kinases [237, 238], whereas p53 activates p21 transcription [239]. Thus, hTERT activates the pRB/E2F-dependent cell cycle pathway. On the other hand, hTERT stimulates the proliferation of human embryonic stem cells by shortening the G1-phase of the cell cycle [240]. This process is associated with the enhancement of the expression of cyclin D1 and hyperphosphorylation of pRB. One can assume that the transcriptional activity of E2F increases, since the level of one of its activators (CDC6) is increased. Moreover, pRB is hyperphosphorylated in hTERT-immortalized human foreskin fibroblasts and human adenoid epithelial cells, which overcome the crisis state after hTERT overexpression [241]. It is an interesting fact that in this case expression of p21 and p53 remains constant; furthermore, these cells contain no protein p16, which is an inhibitor of cyclin-dependent kinase CDK4/6 in complex with cyclin D [242].

Thus, hTERT-dependent stimulation of cell proliferation is induced by the inhibition of protein pRB and activation of the E2F transcription factor. Meanwhile, the same mechanism ensures apoptosis induction [243, 244].

It is known that hTERT overexpression in cells results in an increased content of the epidermal growth factor receptor (EGFR), the transmembrane receptor tyrosine kinase participating in the processes of growth, survival, proliferation, and differentiation of mammalian cells [235, 245]. Following ligand binding, EGFR becomes capable of activating different signaling pathways. Two of those (Ras/Raf/MEK/ERK and PI3K/Akt-kinases) participate in tumor development. These kinase cascades jointly stimulate cell entry into the S-phase of the cell cycle by affecting the expression, stability, and intracellular localization of D-type cyclins [246–250]. This fact is attested to by the results of experiments where the cells overexpressing the *hTERT* gene have the same phenotype as the one in the cells with the *EFGR* gene overexpressed or kinase cascades activated.

hTERT overexpression in epithelial cells was shown to increase the content of the fibroblast growth factor (FGF) and the fibroblast growth factor receptor (FGFR) [228, 235]. Moreover, the content of epiregulin, one of the ligands of the epidermal growth factor receptor (EGFR), which plays the key role in maintaining the proliferation status of these cells, is significantly increased in hTERT-immortalized fibroblasts [251]. Epiregulin is known to be repressed in normal human cells; however, it is activated in tumors with high proliferative potential [252]. It is plausible that telomerase stimulates its anti-apoptotic, pro-proliferative, and pro-neoplastic properties.

Expression of the two isoforms of the vascular endothelial growth factor (VEGF) is activated in human breast cancer cells, as well as in HeLa cells and in the hTERT-transfected normal embryonic lung cells [253]. VEGF is also known to stimulate hTERT expression and activate telomerase via the signal cascades of Ras- and Akt-kinases. Thus, hTERT and the growth factors interact via the positive feedback mechanism in processes of cell cycle regulation, tumor formation, and angiogenesis.

It was reported in 2003 that telomerase activation can result in the epigenetic silencing of the suppressor genes in cancer cells [254]. The DNA-methyltransferase I promoter (*DNMT1*) is activated in normal human fibroblasts upon hTERT expression. The mechanism underlying this effect has not been elucidated yet; however, one can assume that the transcription factor STAT3 is one of the major participant factors in tran-

scription activation. It is already known that STAT3 induces *DNMT1* expression in malignant T-cell tumors [255]. In this case, the signal from hTERT to STAT3 can be transduced by the previously mentioned EGFR, which phosphorylates and thereby activates STAT3 [256]. *DNMT1* participates in the regulation of gene expression by methylating the promoter regions of these genes.

It has recently been established that telomerase also interacts with the Wnt/APC/ β -catenin signalling pathway. It has been shown that TERT interacts with the chromatin-remodulating BRG1 factor. BRG1 is the β -catenin co-factor in the processes of the regulation of transcription of the genes associated with the Wnt-signalling pathway. It turned out that TERT can directly interact with the promoters regulated by Wnt and β -catenin. The signalling pathway is known to play a significant role in the cell differentiation and proliferation processes. The effects observed upon TERT expression in stem cells can presumably be explained by the effect of telomerase components on the regulatory cascade [257].

CONCLUSIONS

Data attesting to the diversity of the functions carried out by the major components of cell telomerase have recently been reported. Some of these functions (such as the nuclease and transferase activities) are associated with the major role of telomerase and its polymerase activity. The other functions (e.g., regulation of gene expression, protection against apoptosis, and contribution to the DNA response to damage) are not directly associated with polymerase activity. It should be noted that the telomerase content in higher eukaryotic cells is very low; hence, almost all the data have been obtained under conditions of artificial expression of its components. Under such conditions, conclusions can be drawn that are divorced from reality. Researchers from different laboratories obtain inconsistent data, which are difficult to interpret. The inconsistency is most likely a result of the use of different systems and models. Nevertheless, the new data reported allows one to assume that telomerase has a more versatile function, and that its impact on the cell is not confined to the regulation of the length of telomere. ●

This work was supported by Lomonosov Moscow State University Development Program (PNR 5.13), Federal Agency for Science and Innovations № 02.740.11.0706 and Russian Foundation for Basic Research (№ 11-04-01233-a and 11-04-12051-ofi-m-2011).

REFERENCES

1. Morgan T.H. // *Science*. 1911. V. 34. P. 636–638.
2. McClintock B. // *Mo. Agric. Exp. Res. Stn. Res. Bull.* 1931. V. 163. P. 4–30.
3. Muller H.J. // *Collecting Net*. 1938. V. 8. P. 182–195.
4. Blackburn E.H., Gall J.G. // *J. Mol. Biol.* 1978. V. 120. P. 33–53.
5. Blackburn E.H., Budarf M.L., Challoner P.B., Cherry J.M., Howard E.A., Katzen A.L., Pan W.C., Ryan T. // *Cold Spring Harb. Symp. Quant. Biol.* 1983. V. 47. Pt. 2. P. 1195–1207.
6. De Lange T., Lundblad V., Blackburn E.H. *Telomeres*. 2nd. Ed. Cold Spring Harbor Laboratory Press, 2006.
7. Hsu T.C., Arrighi F.E., Saunders G.F. // *Proc. Natl. Acad. Sci. USA*. 1972. V. 69. P. 1464–1466.
8. Olovnikov A.M. // *J. Theor. Biol.* 1973. V. 41. P. 181–190.
9. Watson J. // *Nat. New Biol.* 1972. V. 239. P. 197–201.
10. Greider C.W., Blackburn E.H. // *Cell*. 1987. V. 51. P. 887–898.
11. Gallardo F., Laterreur N., Cusanelli E., Ouenzar F., Querido E., Wellinger R.J., Chartrand P. // *Mol. Cell*. 2011. V. 44. P. 819–827.
12. de Lange T. // *Oncogene*. 2002. V. 21. P. 532–540.
13. Chan S.R., Blackburn E.H. // *Phil. Trans. R. Soc. Lond. B. Biol. Sci.* 2004. V. 359. P. 109–121.
14. Palm W., de Lange T. // *Annu. Rev. Genet.* 2008. V. 42. P. 301–334.
15. Riethman H., Ambrosini A., Paul S. // *Chromosome Res.* 2005. V. 13. P. 505–515.
16. Mattern K.A., Swiggers S.J., Nigg A.L., Lowenberg B., Houtsmuller A.B., Zijlmans J.M. // *Mol. Cell. Biol.* 2004. V. 24. P. 5587–5594.
17. Oganessian L., Graham M.E., Robinson P.J., Bryan T.M. // *Biochemistry*. 2007. V. 46. P. 11279–11290.
18. de Lange T. // *Genes Dev.* 2005. V. 19. P. 2100–2110.
19. Palm W., de Lange T. // *Annu. Rev. Genet.* 2008. V. 42. P. 301–334.
20. Broccoli D., Smogorzewska A., Chong L., de Lange T. // *Nat. Genet.* 1997. V. 17. P. 231–235.
21. Court R., Chapman L., Fairall L., Rhodes D. // *EMBO Rep.* 2005. V. 6. P. 39–45.
22. Bianchi A., Stansel R.M., Fairall L., Griffith J.D., Rhodes D., de Lange T. // *EMBO J.* 1999. V. 18. P. 5735–5744.
23. Lei M., Podell E.R., Cech T.R. // *Nat. Struct. Mol. Biol.* 2004. V. 11. P. 1223–1229.
24. Baumann P., Cech T.R. // *Science*. 2001. V. 11. P. 1171–1175.
25. Loayza D., de Lange T. // *Nature*. 2003. V. 423. P. 1013–1018.
26. Kim S.H., Beausejour C., Davalos A.R., Kaminker P., Heo S.J., Campisi J. // *J. Biol. Chem.* 2004. V. 15. P. 43799–43804.
27. Ye J.Z., Hockemeyer D., Krutchinsky A.N., Loayza D., Hooper S.M., Chait B.T., de Lange T. // *Genes Dev.* 2004. V. 18. P. 1649–1654.
28. Ye J.Z., Donigian J.R., van Overbeek M., Loayza D., Luo Y., Krutchinsky A.N., Chait B.T., de Lange T. // *J. Biol. Chem.* 2004. V. 279. P. 47264–47271.
29. Chen Y., Yang Y., van Overbeek M., Donigian J.R., Baciu P., de Lange T., Lei M. // *Science*. 2008. V. 319. P. 1092–1096.
30. Liu D., Safari A., O'Connor M.S., Chan D.W., Laegeler A., Qin J., Songyang Z. // *Nat. Cell Biol.* 2004. V. 6. P. 673–680.
31. Kibe T., Osawa G.A., Keegan C.E., de Lange T. // *Mol. Cell Biol.* 2010. V. 30. P. 1059–1066.
32. Chen L.Y., Liu D., Songyang Z. // *Mol. Cell Biol.* 2007. V. 27. P. 5898–5909.
33. Xin H., Liu D., Wan M., Safari A., Kim H., Sun W., O'Connor M.S., Songyang Z. // *Nature*. 2007. V. 445. P. 559–562.
34. Celli G.B., de Lange T. // *Nat. Cell Biol.* 2005. V. 7. P. 712–718.
35. Li B., de Lange T. // *Mol. Biol. Cell*. 2003. V. 12. P. 5060–5068.
36. Li B., Oestreich S., de Lange T. // *Cell*. 2000. V. 101. P. 471–483.
37. Oganessian L., Bryan T.M. // *BioEssays*. 2007. V. 29. P. 155–165.
38. Zahler A.M., Williamson J.R., Cech T.R., Prescott D.M. // *Nature*. 1991. V. 350. P. 718–720.
39. Wong H.M., Payet L., Huppert J.L. // *Curr. Opin. Mol. Ther.* 2009. V. 11. P. 146–155.
40. Neidle S., Parkinson G.N. // *Biochimie*. 2008. V. 90. P. 1184–1196.
41. Oganessian L., Moon I.K., Bryan T.M. // *EMBO J.* 2006. V. 25. P. 1148–1159.
42. Schaffitzel C., Berger I., Postberg J., Hanes J., Lipps H.J., Plückthun A. // *Proc. Natl. Acad. Sci. USA*. 2001. V. 98. P. 8572–8577.
43. Paeschke K., Simonsson T., Postberg J., Rhodes D., Lipps H.J. // *Nat. Struct. Mol. Biol.* 2005. V. 10. P. 847–854.
44. Tang J., Kan Z.Y., Yao Y., Wang Q., Hao Y.H., Tan Z. // *Nucl. Acids Res.* 2008. V. 36. P. 1200–1208.
45. Shinozuka K., Sekiguchi T., Ebara Y., Moriguchi T. // *Nucl. Acids Symp. Ser.* 2007. V. 51. P. 7–8.
46. Martinez P., Thanasoula M., Carlos A.R., Gómez-López G., Tejera A.M., Schoeftner S., Dominguez O., Pisano D.G., Tarsounas M., Blasco M.A. // *Nat. Cell Biol.* 2010. V. 12. P. 768–780.
47. Blasco M.A. // *Nat. Rev. Genet.* 2007. V. 8. P. 299–309.
48. Sfeir A., Kabir S., van Overbeek M., Celli G.B., de Lange T. // *Science*. 2010. V. 327. P. 1657–1661.
49. Martínez P., Thanasoula M., Muñoz P., Liao C., Tejera A., McNeese C., Flores J.M., Fernández-Capetillo O., Tarsounas M., Blasco M.A. // *Genes Dev.* 2009. V. 23. P. 2060–2075.
50. Sfeir A., Kosiyatrakul S.T., Hockemeyer D., MacRae S.L., Karlseder J., Schildkraut C.L., de Lange T. // *Cell*. 2009. V. 138. P. 90–103.
51. Hsu T.C., Arrighi F.E., Saunders G.F. // *Proc. Natl. Acad. Sci. USA*. 1972. V. 69. P. 1464–1466.
52. Schoeftner S., Blasco M.A. // *EMBO J.* 2009. V. 28. P. 2323–2336.
53. Benetti R., Gonzalo S., Jaco I., Schotta G., Klatt P., Jenuwein T., Blasco M.A. // *J. Cell. Biol.* 2007. V. 178. P. 925–936.
54. García-Cao M., O'Sullivan R., Peters A.H., Jenuwein T., Blasco M.A. // *Nat. Genet.* 2004. V. 36. P. 94–99.
55. Gonzalo S., Jaco I., Fraga M.F., Chen T., Li E., Esteller M., Blasco M.A. // *Nat. Cell Biol.* 2006. V. 8. P. 416–424.
56. Benetti R., Gonzalo S., Jaco I., Muñoz P., Gonzalez S., Schoeftner S., Murchison E., Andl T., Chen T., Klatt P., Li E., Serrano M., Millar S., Hannon G., Blasco M.A. // *Nat. Struct. Mol. Biol.* 2008. V. 15. P. 268–279.
57. Azzalin C.M., Reichenbach P., Khoriauli L., Giulotto E., Lingner J. // *Science*. 2007. V. 318. P. 798–801.
58. Venteicher A.S., Abrew E.B., Meng Z., McCann K.E., Terns R.M., Veenstra T.D., Terns M.P., Artandi S.E. // *Science*. 2009. V. 323. P. 644–648.
59. Venteicher A.S., Meng Z., Mason P.J., Veenstra T.D., Artandi S.E. // *Cell*. 2008. V. 132. P. 945–952.
60. Fu D., Collins K. // *Mol. Cell*. 2007. V. 28. P. 773–785.
61. Cohen S.B., Graham M.E., Lovrecz G.O., Bache N., Robinson P.J., Reddel R.R. // *Science*. 2007. V. 315. P. 1850–1853.

62. Legassie J.D., Jarstfer M.B. // *Structure*. 2006. V. 14. P. 1603–1609.
63. Chen J.L., Greider C.W. // *Trends Biochem. Sci.* 2004. V. 29. P. 183–192.
64. Theimer C.A., Feigon J. // *Curr. Opin. Struct. Biol.* 2006. V. 16. P. 307–318.
65. Sperger J.M., Cech T.R. // *Biochemistry*. 2001. V. 40. P. 7005–7016.
66. Antal M., Boros E., Solymosy F., Kiss T. // *Nucl. Acids Res.* 2002. V. 30. P. 912–920.
67. Bhattacharyya A., Blackburn E.H. // *EMBO J.* 1994. V. 13. P. 5721–5731.
68. Qiao F., Cech T.R. // *Nat. Struct. Mol. Biol.* 2008. V. 15. P. 634–640.
69. Chen J.L., Greider C.W. // *Genes Dev.* 2003. V. 17. P. 2747–2752.
70. Box J.A., Bunch J.T., Zappulla D.C., Glynn E.F., Baumann P. // *J. Biol. Chem.* 2008. V. 283. P. 24224–24233.
71. Tzfati Y., Fulton T.B., Roy J., Blackburn E.H. // *Science*. 2000. V. 288. P. 863–867.
72. Lai C.K., Miller M.C., Collins K. // *Genes Dev.* 2002. V. 16. P. 415–420.
73. O'Connor C.M., Lai C.K., Collins K. // *J. Biol. Chem.* 2005. V. 280. P. 17533–17539.
74. Miller M.C., Collins K. // *Proc. Natl. Acad. Sci. USA*. 2002. V. 99. P. 6585–6590.
75. Greider C.W., Blackburn E.H. // *Cell*. 1985. V. 43. P. 405–413.
76. Blasco M.A. // *Nat. Rev. Genet.* 2005. V. 6. P. 611–622.
77. Miller M.C., Collins K. // *Proc. Natl. Acad. Sci. USA*. 2002. V. 99. P. 6585–6590.
78. Ulyanov N.B., Shefer K., James T.L., Tzfati Y. // *Nucl. Acids Res.* 2007. V. 35. P. 6150–6160.
79. Shefer K., Brown Y., Gorkovoi V., Nussbaum T., Ulyanov N.B., Tzfati Y. // *Mol. Cell. Biol.* 2007. V. 27. P. 2130–2143.
80. Cao S., Chen S.J. // *J. Mol. Biol.* 2007. V. 367. P. 909–924.
81. Theimer C.A., Blois C.A., Feigon J. // *Mol. Cell*. 2005. V. 17. P. 671–682.
82. Kim N.K., Zhang Q., Zhou J., Theimer C.A., Peterson R.D., Feigon J. // *J. Mol. Biol.* 2008. V. 384. P. 1249–1261.
83. Theimer C.A., Finger L.D., Trantirek L., Feigon J. // *Proc. Natl. Acad. Sci. USA*. 2003. V. 100. P. 449–454.
84. Ly H., Blackburn E.H., Parslow T.G. // *Mol. Cell. Biol.* 2003. V. 23. P. 6849–6856.
85. Chen J.L., Greider C.W. // *Proc. Natl. Acad. Sci. USA*. 2005. V. 102. P. 8080–8085.
86. Kim N.W., Piatyzek M.A., Prowse K.R., Harley C.B., West M.D., Ho P.L., Coviello G.M., Wright W.E., Weinrich S.L., Shay J.W. // *Science*. 1994. V. 266. P. 2011–2015.
87. Chen Y., Fender J., Legassie J.D., Jarstfer M.B., Bryan T.M., Varani G. // *EMBO J.* 2006. V. 25. P. 3156–3166.
88. Richards R.J., Wu H., Trantirek L., O'Connor C.M., Collins K., Feigon J. // *RNA*. 2006. V. 12. P. 1475–1485.
89. Leeper T.C., Varani G. // *RNA*. 2005. V. 11. P. 394–405.
90. Chen J.L., Opperman K.K., Greider C.W. // *Nucl. Acids Res.* 2002. V. 10. P. 139–147.
91. Ueda C.T., Roberts R.W. // *RNA*. 2004. V. 10. P. 139–147.
92. Meier B., Clejan I., Liu Y., Lowden M., Gartner A., Hodgkin J., Ahmed S. // *PloS Genet.* 2006. V. 2. e. 18.
93. Gillis A.J., Schuller A.P., Skordalakes E. // *Nature*. 2008. V. 455. P. 633–636.
94. Balakin A.G., Smith L., Fournier M.J. // *Cell*. 1996. V. 86. P. 823–834.
95. Ganot P., Caizergues-Ferrer M., Kiss T. // *Genes Dev.* 1997. V. 11. P. 941–956.
96. Fu D., Collins K. // *Genes Dev.* 2006. V. 20. P. 531–536.
97. Theimer C.A., J dy B.E., Chim N., Richard P., Breece K.E., Kiss T., Feigon J. // *Mol. Cell*. 2007. V. 27. P. 869–881.
98. Gros J., Guédin A., Mergny J.L., Lacroix L. // *Chembiochem*. 2008. V. 9. P. 2075–2079.
99. Li X., Nishizuka H., Tsutsumi K., Imai Y., Kurihara Y., Uesugi S. // *J. Biochem.* 2007. V. 141. P. 755–765.
100. Xu Y., Ishizuka T., Kimura T., Komiyama M. // *J. Am. Chem. Soc.* 2010. V. 132. P. 7231–7233.
101. Jacobs S.A., Podell E.R., Cech T.R. // *Nat. Struct. Mol. Biol.* 2006. V. 13. P. 218–225.
102. Arkhipova I.R., Pyatkov K.I., Meselson M., Evgen'ev M.B. // *Nat. Genet.* 2003. V. 33. P. 123–124.
103. Rouda S., Skordalakes E. // *Structure*. 2007. V. 15. P. 1403–1412.
104. Mitchell M., Gillis A., Futahashi M., Fujiwara H., Skordalakes E. // *Nat. Struct. Mol. Biol.* 2010. V. 17. P. 513–518.
105. Sarafianos S.G., Clark Jr. A.D., Das K., Tuske S., Birktoft J.J., Ilankumaran P., Ramesha A.R., Sayer J.M., Jerina D.M., Boyer P.L., et al. // *EMBO J.* 2002. V. 21. P. 6614–6624.
106. Ollis D.L., Brick P., Hamlin R., Xuong N.G., Steitz T.A. // *Nature*. 1985. V. 313. P. 762–766.
107. Miller M.C., Liu J.K., Collins K. // *EMBO J.* 2000. V. 19. P. 4412–4422.
108. Harris D., Kaushik N., Pandey P.K., Yadav P.N., Pandey V.N. // *J. Biol. Chem.* 1998. V. 273. P. 33624–33634.
109. Legassie J.D., Jarstfer M.B. // *Biochemistry*. 2005. V. 44. P. 14191–14201.
110. Bosoy D., Lue N.F. // *Nucl. Acids Res.* 2004. V. 32. P. 93–101.
111. Lue N.F., Lin Y.C., Mian I.S. // *Mol. Cell. Biol.* 2003. V. 23. P. 8440–8449.
112. Hossain S., Singh S., Lue N.F. // *J. Biol. Chem.* 2002. V. 277. P. 36174–36180.
113. Huard S., Moriarty T.J., Autexier C. // *Nucl. Acids Res.* 2003. V. 31. P. 4059–4070.
114. Banik S.S., Guo C., Smith A.C., Margolis S.S., Richardson D.A., Tirado C.A., Counter C.M. // *Mol. Cell. Biol.* 2001. V. 22. P. 6234–6246.
115. Bosoy D., Lue N.F. // *J. Biol. Chem.* 2001. V. 276. P. 46305–46312.
116. Bryan T.M., Goodrich K.J., Cech T.R. // *Mol. Cell*. 2000. V. 6. P. 493–499.
117. Cunningham D.D., Collins K. // *Mol. Cell. Biol.* 2005. V. 25. P. 4442–4454.
118. Lai C.K., Mitchell J.R., Collins K. // *Mol. Cell. Biol.* 2001. V. 21. P. 990–1000.
119. Bosoy D., Peng Y., Mian I.S., Lue N.F. // *J. Biol. Chem.* 2003. V. 278. P. 3882–3890.
120. Sekarah V.G., Soares J., Jarstfer M.B. // *Biochim. Biophys. Acta*. 2010. V. 1804. P. 1190–1201.
121. Skvortzov D.A., Rubzova M.P., Zvereva M.E., Kiselev F.L., Donzova O.A. // *Acta Naturae*. 2009. V. 1. № 1. P. 51–67.
122. Wick M., Zubov D., Hagen G. // *Gene*. 1999. V. 232. P. 97–106.
123. Kilian A., Bowtell D.D., Abud H.E., Hime G.R., Venter D.J., Keese P.K., Duncan E.L., Reddel R.R., Jefferson R.A. // *Hum. Mol. Genet.* 1997. V. 6. P. 2011–2019.
124. Saeboe-Larssen S., Fossberg E., Gaudernack G. // *BMC Mol. Biol.* 2006. V. 7. P. 26.
125. Hisatomi H., Ohyashiki K., Ohyashiki J.H., Nagao K., Kanamaru T., Hirata H., Hibi N., Tsukada Y. // *Neoplasia*. 2003. V. 5. P. 193–197.
126. Colgin L.M., Wilkinson C., Englezou A., Kilian A., Robinson M.O., Reddel R.R. // *Neoplasia*. 2000. V. 2. P. 426–432.

REVIEWS

127. Yi X., White D.M., Aisner D.L., Baur J.A., Wright W.E., Shay J.W. // *Neoplasia*. 2000. V. 2. P. 433–440.
128. Cerezo A., Kalthoff H., Schuermann M., Schafer B., Boukamp P. // *J. Cell Sci.* 2002. V. 115. P. 1305–1312.
129. Barclay J.Y., Morris A.G., Nwokolo C.U. // *Dig. Dis. Sci.* 2005. V. 50. P. 1299–1303.
130. Song M.S., Lee S.W. // *FEBS Lett.* 2006. V. 580. P. 5033–5043.
131. Ulaner G.A., Hu J.F., Vu T.H., Giudice L.C., Hoffman A.R. // *Cancer Res.* 1998. V. 58. P. 4168–4172.
132. Aisner D.L., Wright W.E., Shay J.W. // *Curr. Opin. Genet. Dev.* 2002. V. 12. P. 80–85.
133. Jádý B.E., Bertrand E., Kiss T. // *J. Cell. Biol.* 2004. V. 1. P. 647–652.
134. Box J.A., Bunch J.T., Tang W., Baumann P. // *Nature*. 2008. V. 456. P. 910–914.
135. Girard C., Verheggen C., Neel H., Cammas A., Vagner S., Soret J., Bertrand E., Bordonné R. // *J. Biol. Chem.* 2008. V. 25. P. 2060–2069.
136. Dez C., Henras A., Faucon B., Lafontaine D., Caizergues-Ferrer M., Henry Y. // *Nucl. Acids Res.* 2001. V. 29. P. 598–603.
137. Lafontaine D.L., Tollervey D. // *Trends Biochem. Sci.* 1998. V. 23. P. 383–388.
138. Watkins N.J., Gottschalk A., Neubauer G., Kastner B., Fabrizio P., Mann M., Lührmann R. // *RNA*. 1998. V. 4. P. 1549–1568.
139. Henras A., Henry Y., Bousquet-Antonelli C., Noaillac-Depeyre J., Gélugne J.P., Caizergues-Ferrer M. // *EMBO J.* 1998. V. 17. P. 7078–7090.
140. Kufel J., Allmang C., Chanfreau G., Petfalski E., Lafontaine D.L., Tollervey D. // *Mol. Cell. Biol.* 2000. V. 20. P. 5415–5424.
141. Tran H., Schilling M., Wirbelauer C., Hess D., Nagamine Y. // *Mol. Cell.* 2004. V. 13. P. 101–111.
142. Vaughn J.P., Creacy S.D., Routh E.D., Joyner-Butt C., Jenkins G.S., Pauli S., Nagamine Y., Akman S.A. // *J. Biol. Chem.* 2005. V. 280. P. 38117–38120.
143. Creacy S.D., Routh E.D., Iwamoto F., Nagamine Y., Akman S.A., Vaughn J.P. // *J. Biol. Chem.* 2008. V. 283. P. 34626–34634.
144. Iwamoto F., Stadler M., Chalupníková K., Oakeley E., Nagamine Y. // *Exp. Cell Res.* 2008. V. 314. P. 1378–1391.
145. Lattmann S., Giri B., Vaughn J.P., Akman S.A., Nagamine Y. // *Nucl. Acids Res.* 2010. V. 38. P. 6219–6233.
146. Sexton A.N., Collins K. // *Mol. Cell. Biol.* 2011. V. 31. P. 736–743.
147. Booy E.P., Meier M., Okun N., Novakowski S.K., Xiong S., Stetefeld J., McKenna S.A. // *Nucl. Acids Res.* 2012. Jan 11.
148. McEachern M.J., Krauskopf A., Blackburn E.H. // *Annu. Rev. Genet.* 2000. V. 34. P. 331–358.
149. Meier U.T. // *Chromosoma*. 2005. V. 114. P. 1–14.
150. Matera A.G., Terns R.M., Terns M.P. // *Nat. Rev. Mol. Cell. Biol.* 2007. V. 8. P. 209–220.
151. Wang C., Meier U.T. // *EMBO J.* 2004. V. 23. P. 1857–1867.
152. Mitchell J.R., Cheng J., Collins K. // *Mol. Cell. Biol.* 1999. V. 19. P. 567–576.
153. Pogacić V., Dragon F., Filipowicz W. // *Mol. Cell. Biol.* 2000. V. 20. P. 9028–9040.
154. Dragon F., Pogacić V., Filipowicz W. // *Mol. Cell. Biol.* 2000. V. 20. P. 3037–3048.
155. Venteicher A.S., Meng Z., Mason P.J., Veenstra T.D., Artandi S.E. // *Cell*. 2008. V. 132. P. 945–957.
156. Marcand S., Brevet V., Mann C., Gilson E. // *Curr. Biol.* 2000. V. 10. P. 487–490.
157. Jádý B.E., Richard P., Bertrand E., Kiss T. // *Mol. Biol. Cell.* 2006. V. 17. P. 944–954.
158. Tomlinson R.L., Ziegler T.D., Supakorndej T., Terns R.M., Terns M.P. // *Mol. Biol. Cell.* 2006. V. 17. P. 955–965.
159. Her J., Chung I.K. // *Biochem. Biophys. Res. Commun.* 2012. V. 417. P. 1086–1092.
160. Germain-Lee E.L., Obie C., Valle D. // *Genomics*. 1997. V. 44. P. 22–34.
161. Nagahama M., Hara Y., Seki A., Yamazoe T., Kawate Y., Shinohara T., Hatsuzawa K., Tani K., Tagaya M. // *Mol. Biol. Cell.* 2004. V. 15. P. 5712–5723.
162. Errington T.M., Fu D., Wong J.M., Collins K. // *Mol. Cell. Biol.* 2008. V. 28. P. 6510–6520.
163. Sekaran V.G., Soares J., Jarstfer M.B. // *Biochim. Biophys. Acta*. 2010. V. 1804. P. 1190–1201.
164. Shcherbakova D.M., Sokolov K.A., Zvereva M.I., Dontsova O.A. // *Biochemistry*. 2009. V. 74. P. 749–755.
165. Egan E.D., Collins K. // *Mol. Cell. Biol.* 2010. V. 30. P. 2775–2786.
166. Forsythe H.L., Jarvis J.L., Turner J.W., Elmore L.W., Holt S.E. // *J. Biol. Chem.* 2001. V. 276. P. 15571–15574.
167. Hamma T., Ferré-D'Amaré A.R. // *J. Biol. Chem.* 2010. V. 285. P. 805–809.
168. Autexier C., Lue N.F. // *Annu. Rev. Biochem.* 2006. V. 75. P. 493–517.
169. Greider C.W. // *Mol. Cell. Biol.* 1991. V. 11. P. 4572–4580.
170. Lue N.F. // *BioEssays*. 2004. V. 26. P. 955–962.
171. Wallweber G., Gryaznov S., Pongracz K., Pruzan R. // *Biochemistry*. 2003. V. 42. P. 589–600.
172. Steitz T.A. // *J. Biol. Chem.* 1999. V. 274. P. 17395–17398.
173. Chang M., Arneric M., Lingner J. // *Genes Dev.* 2007. V. 21. P. 2485–2494.
174. Wang J., Yu L., Li J., Deng R., Wang X. // *Biotechnol. Appl. Biochem.* 2007. V. 48. P. 93–99.
175. Collins K., Greider C.W. // *Genes Dev.* 1993. V. 7. P. 1364–1376.
176. Cohn M., Blackburn E.H. // *Science*. 1995. V. 269. P. 396–400.
177. Lue N.F., Peng Y. // *Nucl. Acids Res.* 1997. V. 25. P. 4331–4337.
178. Lue N.F., Peng Y. // *Nucl. Acids Res.* 1998. V. 26. P. 1487–1494.
179. Oulton R., Harrington L. // *Mol. Biol. Cell.* 2004. V. 15. P. 3244–3256.
180. Huard S., Autexier C. // *Nucl. Acids Res.* 2004. V. 32. P. 2171–2180.
181. Petrov A.V., Dokudovskaya S.S., Sokolov K.A., Lavrik O.I., Favre A., Dontsova O.A., Bogdanov A.A. // *FEBS Lett.* 1998. V. 436. P. 35–40.
182. Niu H., Xia J., Lue N.F. // *Mol. Cell. Biol.* 2000. V. 20. P. 6806–6815.
183. Lue N.F., Bosoy D., Moriarty T.J., Autexier C., Altman B., Leng S. // *Proc. Natl. Acad. Sci. USA*. 2005. V. 102. P. 9778–9783.
184. Smekalova E.M., Petrova O.A., Zvereva M.I., Dontsova O.A. // *Acta Naturae*. 2012. V. 4. № 1 (12). P. 70–73.
185. Kilian A., Bowtell D.D., Abud H.E., Hime G.R., Venter D.J., Keese P.K., Duncan E.L., Reddel R.R., Jefferson R.A. // *Hum. Mol. Genet.* 1997. V. 6. P. 2011–2019.
186. Ramakrishnan S., Eppenberger U., Mueller H., Shinkai Y., Narayanan R. // *Cancer Res.* 1998. V. 58. P. 622–625.
187. Masutomi K., Possemato R., Wong J.M., Currier J.L., Tothova Z., Manola J.B., Ganesan S., Lansdorp P.M., Collins K., Hahn W.C. // *Proc. Natl. Acad. Sci. USA*. 2005. V. 102. P. 8222–8227.

188. Armbruster B.N., Banik S.S., Guo C., Smith A.C., Counter C.M. // *Mol. Cell. Biol.* 2001. V. 21. P. 7775–7786.
189. Haendeler J., Hoffmann J., Brandes R.P., Zeiher A.M., Dimmeler S. // *Mol. Cell. Biol.* 2003. V. 23. P. 4598–4610.
190. Haendeler J., Hoffmann J., Diehl J.F., Vasa M., Spyridopoulos I., Zeiher A.M., Dimmeler S. // *Circ. Res.* 2004. V. 94. P. 768–775.
191. Haendeler J., Dröse S., Büchner N., Jakob S., Altschmied J., Goy C., Spyridopoulos I., Zeiher A.M., Brandt U., Dimmeler S. // *Arterioscler. Thromb. Vasc. Biol.* 2009. V. 29. P. 929–935.
192. Santos J.H., Meyer J.N., Skorpova M., Annab L.A., van Houten B. // *Aging Cell.* 2004. V. 3. P. 399–411.
193. Santos J.H., Meyer J.N., van Houten B. // *Hum. Mol. Genet.* 2006. V. 15. P. 1757–1768.
194. Ahmed S., Passos J.F., Birket M.J., Beckmann T., Brings S., Peters H., Birch-Machin M.A., von Zglinicki T., Saretzki G. // *J. Cell Sci.* 2008. V. 121. P. 1046–1053.
195. Parrinello S., Samper E., Krtolica A., Goldstein J., Melov S., Campisi J. // *Nat. Cell. Biol.* 2003. V. 5. P. 741–747.
196. Packer L., Fuehr K. // *Nature.* 1977. V. 267. P. 423–425.
197. von Zglinicki T., Saretzki G., Döcke W., Lotze C. // *Exp. Cell Res.* 1995. V. 220. P. 186–193.
198. Shay J.W., Pereira-Smith O.M., Wright W.E. // *Exp. Cell Res.* 1991. V. 196. P. 33–39.
199. Wright W.E., Pereira-Smith O.M., Shay J.W. // *Mol. Cell. Biol.* 1989. V. 9. P. 3088–3092.
200. Chen Z., Trotman L.C., Shaffer D., Lin H.K., Dotan Z.A., Niki M., Koutcher J.A., Scher H.I., Ludwig T., Gerald W., et al. // *Nature.* 2005. V. 436. P. 725–730.
201. Kim K.S., Kang K.W., Seu Y.B., Baek S.H., Kim J.R. // *Mech. Ageing Dev.* 2009. V. 130. P. 179–188.
202. Henle E.S., Han Z., Tang N., Rai P., Luo Y., Linn S. // *J. Biol. Chem.* 1999. V. 274. P. 962–971.
203. Oikawa S., Kawanishi S. // *FEBS Lett.* 1999. V. 453. P. 365–368.
204. Saretzki G., Murphy M.P., von Zglinicki T. // *Aging Cell.* 2003. V. 2. P. 141–143.
205. Sharma N.K., Reyes A., Green P., Caron M.J., Bonini M.G., Gordon D.M., Holt I.J., Santos J.H. // *Nucl. Acids Res.* 2012. V. 40. P. 712–725.
206. Santos J.H., Meyer J.N., van Houten B. // *Hum. Mol. Genet.* 2006. V. 15. P. 1757–1768.
207. Del Bufalo D., Rizzo A., Trisciuoglio D., Cardinali G., Torrisi M.R., Zangemeister-Wittke U., Zupi G., Biroccio A. // *Cell Death Differ.* 2005. V. 12. P. 1429–1438.
208. Büchner N., Zschauer T.C., Lukosz M., Altschmied J., Haendeler J. // *Exp. Gerontol.* 2010. V. 45. P. 558–562.
209. Gorbunova V., Seluanov A., Pereira-Smith O.M. // *J. Biol. Chem.* 2002. V. 277. P. 38540–38549.
210. Seimiya H., Sawada H., Muramatsu Y., Shimizu M., Ohko K., Yamane K., Tsuruo T. // *EMBO J.* 2000. V. 19. P. 2652–2661.
211. Kovalenko O.A., Caron M.J., Ulema P., Medrano C., Thomas A.P., Kimura M., Bonini M.G., Herbig U., Santos J.H. // *Aging Cell.* 2010. V. 9. P. 203–219.
212. Jakob S., Schroeder P., Lukosz M., Büchner N., Spyridopoulos I., Altschmied J., Haendeler J. // *J. Biol. Chem.* 2008. V. 283. P. 33155–33161.
213. Akiyama M., Hideshima T., Hayashi T., Tai Y.T., Mitsiades C.S., Mitsiades N., Chauhan D., Richardson P., Munshi N.C., Anderson K.C. // *Cancer Res.* 2003. V. 63. P. 18–21.
214. Liu K., Hodes R.J., Weng N. // *J. Immunol.* 2001. V. 166. P. 4826–4830.
215. Kimura A., Ohmichi M., Kawagoe J., Kyo S., Mabuchi S., Takahashi T., Ohshima C., Arimoto-Ishida E., Nishio Y., Inoue M., et al. // *Oncogene.* 2004. V. 23. P. 4505–4515.
216. Ram R., Uziel O., Eldan O., Fenig E., Beery E., Lichtenberg S., Nordenberg Y., Lahav M. // *Clin. Cancer Res.* 2009. V. 15. P. 914–923.
217. Maida Y., Yasukawa M., Furuuchi M., Lassmann T., Possemato R., Okamoto N., Kasim V., Hayashizaki Y., Hahn W.C., Masutomi K. // *Nature.* 2009. V. 461. P. 230–235.
218. d'Adda di Fagagna F., Teo S.H., Jackson S.P. // *Genes Dev.* 2004. V. 18. P. 1781–1799.
219. Takata H., Kanoh Y., Gunge N., Shirahige K., Matsuura A. // *Mol. Cell.* 2004. V. 14. P. 515–522.
220. Verdun R.E., Crabbe L., Haggblom C., Karlseder J. // *Mol. Cell.* 2005. V. 20. P. 551–561.
221. Sancar A., Lindsey-Boltz L.A., Unsal-Kaçmaz K., Linn S. // *Annu. Rev. Biochem.* 2004. V. 73. P. 39–85.
222. Shiloh Y. // *Nat. Rev. Cancer.* 2003. V. 3. P. 155–168.
223. Flynn R.L., Zou L. // *Trends Biochem. Sci.* 2011. V. 36. P. 133–140.
224. López-Contreras A.J., Fernandez-Capetillo O. // *DNA Repair (Amst.)* 2010. V. 9. P. 1249–1255.
225. Karlseder J., Hoke K., Mirzoeva O.K., Bakkenist C., Kastan M.B., Petrini J.H., de Lange T. // *PLoS Biol.* 2004. V. 2. P. 1150–1156.
226. Kedde M., le Sage C., Duursma A., Zlotorynski E., van Leeuwen B., Nijkamp W., Beijersbergen R., Agami R. // *J. Biol. Chem.* 2006. V. 281. P. 40503–40514.
227. Mahalingam D., Tay L.L., Tan W.H., Chai J.H., Wang X. // *FEBS J.* 2011. V. 278. P. 3724–3738.
228. Sharma G.G., Gupta A., Wang H., Scherthan H., Dhar S., Gandhi V., Iliakis G., Shay J.W., Young C.S., Pandita T.K. // *Oncogene.* 2003. V. 22. P. 131–146.
229. Osley M.A., Tsukuda T., Nickoloff J.A. // *Mutat. Res.* 2007. V. 618. P. 65–80.
230. Kruhlak M.J., Celeste A., Delleire G., Fernandez-Capetillo O., Müller W.G., McNally J.G., Bazett-Jones D.P., Nussenzweig A. // *J. Cell. Biol.* 2006. V. 172. P. 823–834.
231. Perrault S.D., Hornsby P.T., Betts D.H. // *Biochem. Biophys. Res. Commun.* 2005. V. 335. P. 925–936.
232. Simonsen J.L., Rosada C., Serakinci N., Justesen J., Stenderup K., Rattan S.I., Jensen T.G., Kassem M. // *Nat. Biotechnol.* 2002. V. 20. P. 592–596.
233. Oh H., Taffet G.E., Youker K.A., Entman M.L., Overbeek P.A., Michael L.H., Schneider M.D. // *Proc. Natl. Acad. Sci. USA.* 2001. V. 98. P. 10308–10313.
234. Stampfer M.R., Garbe J., Levine G., Lichtsteiner S., Vasserot A.P., Yaswen P. // *Proc. Natl. Acad. Sci. USA.* 2001. V. 98. P. 4498–4503.
235. Smith L.L., Collier H.A., Roberts J.M. // *Nat. Cell. Biol.* 2003. V. 5. P. 474–479.
236. Xiang H., Wang J., Mao Y., Liu M., Reddy V.N., Li D.W. // *Oncogene.* 2002. V. 21. P. 3784–3791.
237. Xiong Y., Hannon G.J., Zhang H., Casso D., Kobayashi R., Beach D. // *Nature.* 1993. V. 366. P. 701–704.
238. Xia C., Bao Z., Tabassam F., Ma W., Qiu M., Hua S., Liu M. // *J. Biol. Chem.* 2000. V. 275. P. 20942–20948.
239. el-Deiry W.S., Tokino T., Velculescu V.E., Levy D.B., Parsons R., Trent J.M., Lin D., Mercer W.E., Kinzler K.W., Vogelstein B. // *Cell.* 1993. V. 75. P. 817–825.
240. Yang C., Przyborski S., Cooke M.J., Zhang X., Stewart R., Anyfantis G., Atkinson S.P., Saretzki G., Armstrong L., Lako M. // *Stem Cells.* 2008. V. 26. P. 850–863.
241. Farwell D.G., Shera K.A., Koop J.I., Bonnet G.A., Matthews C.P., Reuther G.W., Coltrera M.D., McDougall J.K., Klingelutz A.J. // *Am. J. Pathol.* 2000. V. 156. P. 1537–1547.

REVIEWS

242. Serrano M., Hannon G.J., Beach D. // *Nature*. 1993. V. 366. P. 704–707.
243. Ginsberg D. // *FEBS Lett.* 2002. V. 529. P. 122–125.
244. Lazzerini Denchi E., Helin K. // *EMBO Rep.* 2005. V. 6. P. 661–668.
245. Oda K., Matsuoka Y., Funahashi A., Kitano H. // *Mol. Syst. Biol.* 2005. V. 1. P. 1–10.
246. Liu J.J., Chao J.R., Jiang M.C., Ng S.Y., Yen J.J., Yang-Yen H.F. // *Mol. Cell. Biol.* 1995. V. 15. P. 3654–3663.
247. Lavoie J.N., L'Allemain G., Brunet A., Müller R., Pouysségur J. // *J. Biol. Chem.* 1996. V. 271. P. 20608–20616.
248. Muise-Helmericks R.C., Grimes H.L., Bellacosa A., Malstrom S.E., Tsichlis P.N., Rosen N. // *J. Biol. Chem.* 1998. V. 273. P. 29864–29872.
249. Diehl J.A., Cheng M., Roussel M.F., Sherr C.J. // *Genes Dev.* 1998. V. 12. P. 3499–3511.
250. Huang W., Chang H.Y., Fei T., Wu H., Chen Y.G. // *Oncogene*. 2007. V. 26. P. 2471–2482.
251. Lindvall C., Hou M., Komurasaki T., Zheng C., Henriksen M., Sedivy J.M., Björkholm M., Teh B.T., Nordenskjöld M., Xu D. // *Cancer Res.* 2003. V. 63. P. 1743–1747.
252. Toyoda H., Komurasaki T., Uchida D., Morimoto S. // *Biochem. J.* 1997. V. 326. P. 69–75.
253. Zhou L., Zheng D., Wang M., Cong Y.S. // *Biochem. Biophys. Res. Commun.* 2009. V. 386. P. 739–743.
254. Young J.I., Sedivy J.M., Smith J.R. // *J. Biol. Chem.* 2003. V. 278. P. 19904–19908.
255. Zhang Q., Wang H.Y., Woetmann A., Raghunath P.N., Odum N., Wasik M.A. // *Blood*. 2006. V. 108. P. 1058–1064.
256. Park O.K., Schaefer T.S., Nathans D. // *Proc. Natl. Acad. Sci. USA.* 1996. V. 93. P. 13704–13708.
257. Park J.I., Venteicher A.S., Hong J.Y., Choi J., Jun S., Shkreli M., Chang W., Meng Z., Cheung P., Ji H., et al. // *Nature*. 2009. V. 460. P. 66–72.

Coagulation Factor IX for Hemophilia B Therapy

N. A. Orlova^{1,2}, S. V. Kovnir^{1,2}, I. I. Vorobiev^{1,2*}, A.G.Gabibov¹

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

² Hematology Research Centre, Ministry of Healthcare and Social Development of the Russian Federation, Novyi Zykovskii pr., 4, Moscow, 125167, Russia

*E-mail: ptichman@gmail.com

Received 28.12.2011

Copyright © 2012 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 Factor IX is a zymogen enzyme of the blood coagulation cascade. Inherited absence or deficit of the IX functional factor causes bleeding disorder hemophilia B, which requires constant protein replacement therapy. Reviewed herein are the current state in the manufacturing of FIX, improved variants of the recombinant protein for therapy, transgenic organisms for obtaining FIX, and the advances in the gene therapy of hemophilia B.

KEYWORDS coagulation factor IX; hemophilia B; heterologous protein expression systems.

ABBREVIATIONS FIX – Factor IX; rFIX – recombinant factor IX; rhFIX – recombinant human factor IX; FIXa – activated FIX; pdFIX – plasma-derived factor IX; PTM – posttranslational modification; VKD – vitamin K dependent; IU – international unit; EGF – epidermal growth factor.

INTRODUCTION

Factor IX (FIX, Christmas factor) is a blood clotting factor, a zymogen of serine protease. Upon activation, FIX is converted into the active serine protease and, in the presence of Ca²⁺ and membrane phospholipids, it hydrolyses one arginine-isoleucine bond in factor X to form the activated factor X (Xa) [1]. The catalytic efficiency of activated FIX (FIXa) is greatly increased by the cofactor, the activated factor VIII (FVIIIa). The non-covalent complex of FIXa, FVIIIa and FX, bound to the phospholipid membrane, is called “the X-ase” or “tenase” and represents a major signal amplification loop in the blood coagulation cascade (*Fig. 1*).

Factor IX is produced in the liver, and the inactive precursor protein is processed in the endoplasmic reticulum and Golgi, where it undergoes multiple post-translational modifications and is secreted into the bloodstream upon proteolytic cleavage of the propeptide. Circulated mature FIX, 57 kDa and app. 90 nM, takes part in the blood coagulation cascade after specific proteolytic cleavage by the activated factor XI (of the contact pathway) or the activated factor VII (of the tissue factor pathway), with the formation of two polypeptide chains linked by a single disulfide bridge. Activated FIX is slowly deactivated by multiple factors – binding to antithrombin III, nexin-2, the protein Z-dependent protease inhibitor, and endocytic hepatocyte receptors or degraded by neutrophil elastase [3].

The gene of human FIX lies in the X chromosome,

has 8 exons, and spans 33.5 Kb. Various mutations in this gene can impair the functioning of the FIX protein, resulting in bleeding-disorder hemophilia B: these mutations are present in the dedicated database [4]. The rate of incidence of severe hemophilia B, requiring regular replacement therapy, is 1 in every 30,000 men, which represents approximately 20% of all hemophiliacs. Recently, it has been proved that European royalty suffered from hemophilia B: the last affected person passed away in 1940 [5]. The point mutation discovered in these kindred resulted in altered splicing and truncated form of the FIX protein.

In some cases, mutations in the promoter region of the gene result in the less severe hemophilia B Leiden [6], characterized as a nearly complete absence of FIX in childhood and steady increase in the level of endogenous FIX during puberty to the near-normal values.

Current treatment of hemophilia B is restricted to protein-replacement therapy, which is very expensive for patients and the healthcare system. Only 20% of the world’s population can afford the treatment; so hemophilia B remains lethal in childhood in poor countries [7].

REPLACEMENT THERAPY OF HEMOPHILIA B

Initial specific therapy of hemophilia B used to consist of periodic treatment by plasma transfusions, later replaced by more effective prothrombin complex concentrates (PCC), which contain a mixture of VKD pro-co-

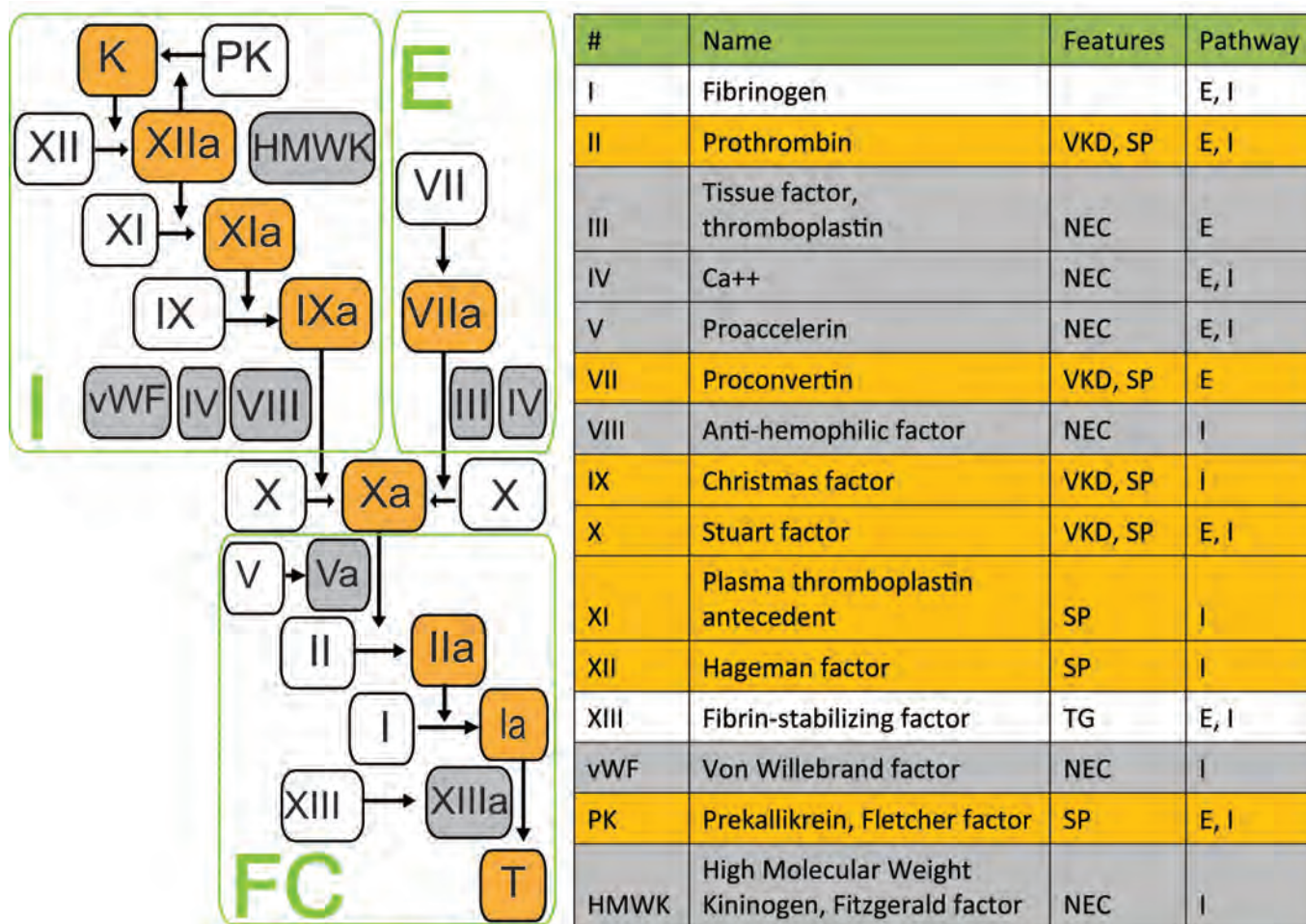


Fig. 1. Scheme of blood coagulation cascade and international nomenclature of coagulation factors [2]. VKD - vitamin K dependent; SP - serine protease, TG - transglutaminase; E – extrinsic pathway; I – intrinsic pathway; NEC - non-enzymatic cofactor, FC - final common pathway.

agulation factors, including FII, FVII and FX. The most significant drawback of PCC is the risk of thrombotic episodes. Purer preparations of FIX have been isolated through Cohn fractionation by ion-exchange chromatography. The safety of plasma-derived FIX preparations has also been improved by the introduction of various virus-inactivation steps, including heating, thiocyanate, and solvent-detergent treatment, which allow to remove enveloped viruses; and nanofiltration, to remove nonenveloped viruses [8].

Another limitation on the safety of FIX plasma concentrates is placed by the detectable amount of activated FIX (FIXa) and residual levels of other pro-coagulation factors, which result in a still significant risk of thrombotic episodes. Immunoaffinity purification of plasma-derived FIX has been sufficient to overcome these limitations [9], but as in any other plasma-derived product, the risk of viral and prion transmission remains [10].

RECOMBINANT FIX

Cloning of FIX cDNA was reported in 1982 [11], and biologically active FIX was expressed in a rat hepatoma cell line, mouse fibroblasts, and baby hamster kidney (BHK) cells in 1985 [12-14]. Expression of FIX in industrially suitable CHO cells was achieved in 1986 [15].

The first and only marketed medicinal product of recombinant FIX to date is Nonacog alpha (trade name Benefix). It was approved for clinical use in the U.S. and European Union in 1997. Nonacog alpha is expressed by CHO cells, cultivated in an animal origin components-free medium, purified through 4 chromatographic steps without the use of immunoaffinity columns, and virus-inactivated by nanofiltration with a cut-off limit of 70 kDa [16]. The final product is formulated as a lyophilized powder without human serum albumin [17]. Initial formulation allowed for 250- to 1,000-IU strength in one vial, and reformulation of recombinant FIX ex-

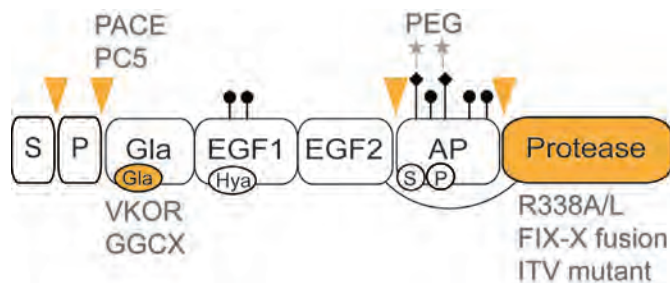


Fig. 2. Structure of FIX. S – signal peptide; P – propeptide, Gla – Gla domain, EGF1 and EGF2 – epidermal growth factor-like domains, AP – activation peptide, Protease – serine protease domain. Posttranslational modifications sites: Gla – γ -carboxylation, Hya – β -hydroxylation; S – sulfation, P – phosphorylation, \blacklozenge – N-linked and \bullet – O-linked glycosylation, cleavage points are indicated by triangles, main possible routes for factor IX production or activity improvement are typed in grey.

tended this interval up to 2000 IU/vial strength [18], additionally allowing for room-temperature storage.

Clinical studies of recombinant FIX have showed that the safety and efficacy of recombinant and plasma-derived FIX are comparable, and no evidence of viral transmission was detected after 1,514 infusions of recombinant FIX to 56 patients [19]. The immune response level to infused recombinant FIX was also comparable to plasma-derived FIX [20].

Recombinant FIX has been extensively investigated for structural deviations from natural FIX and the significance of such deviations. The germinal studies claimed that there were close similarity in post-translational modifications [21], although the adjusted recovery of FIX activity following a bolus infusion was found to be significantly lower for recombinant FIX [22], resulting in a 1.5- to 2-fold increase of the recommended dose [23]. The pattern of post-translational modifications was found to be indistinguishable between lots of recombinant FIX from two production plants [24]; thus, the differences in recombinant and plasma-derived FIX appear to be rather process-specific.

STRUCTURE AND POST-TRANSLATIONAL MODIFICATIONS OF FIX

The FIX protein is a member of vitamin K-dependent (VKD) blood clotting factors and consists of four structural domains: the Gla domain, two tandem EGF-like domains, and a C-terminal serine protease domain (Fig. 2). The N-terminal signal peptide of FIX is released after translocation to endoplasmic reticulum, and the propeptide, immediately preceding the Gla-domain, is removed at secretion. The activation peptide, which is located between the second EGF domain and the serine

protease domain, is specifically cleaved by the factor XIa at the activation of FIX.

The N-terminal Gla domain of mature FIX, common to members of the VKD group, mediates the binding of FIX and FIXa to the surface of endothelial cells; this interaction is disrupted completely if γ -carboxylation of the Asp residues in the Gla domain is blocked [25]. The first EGF-like module is known to have a high-affinity Ca^{2+} binding site and to take part in the interaction with factor VIIIa [26] and with the tissue factor [1]. It contains an unusual modification of the Asp64 residue to β -hydroxyaspartate (Hya). This modification does not affect the pro-coagulation properties of FIX [25], but exchange of the Asp64 residue to a basic or neutral amino acid impairs the activity of FIX [27]. Second EGF-like module also participates in the assembly of the FIXa-FVIIIa-FX complex [28, 29]. The EGF2 and protease domains of FIX are connected via the activation peptide and a single disulphide bond.

The activation peptide contains many sites of post-translational modifications that have various impacts on the properties of the protein (summarized in Table 1). The protease domain accounts for half of the mass of FIX. The serine protease active site in this domain is buried under the activation peptide and exposed after its cleavage. There are no known post-translational modifications in this domain; upon activation, it connects to the rest of the FIX molecule by a single disulphide bond located opposite the active site. The C-terminus of the activated FIX also lies far from the active site cavity of protease and allows the creation of functionally active fusion protein molecules.

Although the only post-translational modification with a direct and significant effect on the pro-coagulation activity of FIX is the γ -carboxylation in the Gla domain [30], other modifications also play a more or less clear role in the functioning of FIX.

The decrease in activity recovery in recombinant FIX was assigned by the research group of the Genetics Institute, Inc., to two PTM's – absent phosphorylation of Ser158 and nearly absent sulfation of Tyr 155 in the recombinant FIX [21]. It was demonstrated that infusion of recombinant FIX enriched in the sulfated variant results in increased activity recovery, and, at the same time, that isolation of recombinant FIX from infused hemophilia B dogs results in the enrichment of the material by the sulfated FIX. It should be noted that both sites of the modification are located in the activation peptide and are in very close proximity to each other, as well as to the N-linked glycan at Asn157, making the separation of sulfated and phosphorylated molecules questionable.

Natural FIX undergoes both O- and N-glycosylation. Berthing sites of O-linked oligosaccharides are present

Table 1. Post-translational modifications of FIX and its derivatives

Structural feature	Domain	Function of PTM	pdFIX	rFIX (BeneFix)	Fc fusion	PEG conjugate
Gamma-carboxylation (Glu → Gla) Total Gla and Gla share	Gla	Interaction with membrane, binding of Ca ²⁺	total 12 (12/12)	total 11,6 (60% 12/12, 35% 11/12, 5% 10/12)	Ttl 11,2	total 11,6 (33% 11/12; 64% 12/12)
beta-hydroxylation (Asp64 → Hya)	EGF	ND	37%	46% - 49%	70%	partial
N-linked glycans	AP	ND	3- and 4-antennary, sialated by Neu-5-Ac	more complex structure than pd, contain different linkages, more fucosilation and polyacetilactosamine structures		Both present, fucosylated core, 3- and 4-antennary, complex type
Asn157 heterogeneity	AP		high	low	low	sialated <i>in vitro</i>
Asn167 sialation	AP		Full	less	not full	
O-linked glycans						
Ser 53	GLA	ND	(Xy)1-2-Glc	(Xy)2-Glc	different from CHO-derived FIX, relative ratios	Present
Ser 61	GLA	ND	NeuAcGalGlcNAcFuc	NeuAcGalGlcNAcFuc		Present
Thr 159, Thr 167, Thr 172, Thr179 (?)	AP	Blocks protease active site	partial	partial		Partial
Tyr 155 sulfation	AP	Accounts for <i>in vivo</i> activity recovery	>90%	5%-15%	4%	ND
Ser 158 phosphorylation	AP	Accounts for <i>in vivo</i> activity recovery?	>90%	<10%	<10%	ND
Activated FIX		Unwanted admixture	0.21%+ _	0.11+-0.0019%	<0.013%	0.03%

in the EGF1 domain [31] and activation peptide. Two O-linked oligosaccharides in the EGF domain are present uniformly in natural and recombinant FIX, and 2 to 4 possible O-glycosylation sites in the activation peptide are modified only partially [32, 33].

There are also two sites for the mooring of N-linked oligosaccharides in the activation peptide of factor IX at the asparagine residues 157 and 167 [11]: both sites are fully occupied by predominantly sialated oligosaccharide groups in plasma-derived FIX [34]. Enzymatic cleavage of all sialic acid residues from the O- and N-linked groups does not alter the activation rate of FIX and its ability to activate factor X [35]. At the same time, the low level of sialation in N-linked glycans in the recombinant FIX may account for the differences in the binding to endothelial cells, the rate of clearance from circulation, or the susceptibility to proteolysis.

The level of the last known PTM in the FIX – β -hydroxylation of Asp64 in the EGF1 domain of FIX is slightly increased in the recombinant FIX [25], but incomplete modification of Asp64 in the natural protein clearly indicates that this PTM is not biologically important for FIX [36].

Since the currently marketed drug of recombinant FIX is not superior to plasma-derived FIX, at least in terms of the required dose strength and duration of action, further advances in studies of recombinant FIX variants and derivatives have clinical perspectives.

IMPROVEMENTS IN RECOMBINANT FIX PRODUCTION

The reported secretion level of recombinant FIX by the production cell line is relatively low: ca. 30 mg/l [37]. It may be increased by 30-50% upon addition of 1 nM of methyl testosterone to the culture medium [38] or doubled after the addition of phorbol 12-myristate, 13-acetate, and calcium ionophore [39]. The appearance of undesired activated FIX in the culture medium can be controlled by decreasing the Ca^{2+} ions concentration to 0.5 mM from 1.12 mM [40]; in another study of the same group, it was found that an increase of Ca^{2+} ions to 1.3 mM leads to a 30% increase in the production of FIX, without a significant rise in the FIXa level [41].

In the early investigation of recombinant FIX from CHO cells, it was found that propeptide processing in the secreted FIX is incomplete and that FIX with an attached propeptide is inactive [42]; complete or nearly complete propeptide processing can be achieved by co-expression of the subtilisin/kexin-like convertase PACE/furin. Homologous convertase PC5 [43] may also be employed [44].

Pro-coagulation activity of FIX requires complete gamma-carboxylation of the first 10 Glu residues in the Gla domain: last 2 residues may not be carboxylated [45]. Natural FIX is completely γ -carboxylated

in all 12 residues, and in recombinant FIX the level of γ -carboxylation is reduced at last two residues, yielding an average of 11.5 Gla residues per molecule [37] and normal clotting activity (not less than 200 IU/mg). In the case of BHK host cells, the pro-coagulant activity of the secreted FIX declined for highly producing line, and over-expression of vitamin K 2,3-epoxide reductase enzyme (VKOR), which produces the cofactor for the γ -carboxylation reaction, restored the relative pro-coagulant activity to its normal level [46].

There are no direct investigations of the rate-limiting step in the post-translational modifications cascade for recombinant FIX, and various enhancements in the CHO enzymes levels may increase the secretion of rightly processed FIX. At least one type of modification – processing of propeptide – may be carried out in the culture medium utilizing the co-expressed soluble truncated PACE variant [37]. In the case of the homologous VKD protein – human protein C (hPC) – it was found that the rate-limiting step for recombinant hPC, expressed in human 293 cells, is N-glycosylation [47]; in the case of another VKD homologue – factor VII expressed in CHO cells – both glycosylation and γ -carboxylation limit the secretion of the product [48]. It is interesting to note that the typical secretion level of factor VII is ~5 times higher than that of FIX in industrially deployed cell lines and that the only significant difference in the post-translational modifications of these two proteins is more abundant O-glycosylation of FIX.

CHO-derived cell lines, which are currently employed for the production of recombinant FIX, may be substituted for more productive ones. Natural FIX is produced by liver cells, and it has been established that the human hepatoma cell line HepG2 produces 1.5 times more recombinant FIX than the human kidney cell line 293 after transfection by the same retroviral vector [39]. Nonvertebrate cultured cells were also evaluated for the expression of FIX, and in a drosophila-derived SF2 cell line a 12-fold increase in functionally active FIX secretion, compared with CHO cells, was detected [49].

TRANSGENIC ORGANISMS

The milk of transgenic animals has been considered as a better source of recombinant therapeutic proteins for the last twenty years. FIX has been expressed in transgenic sheep as a fusion gene comprising the beta-lactoglobulin and FIX sequences, and small quantities of inactive FIX have been detected in the milk [50]. Higher levels of FIX, secreted in sheep milk, have been achieved using the nuclear transfer technique developed by PPL Therapeutics [51]. The producing species were created using the same technique as the one em-

Table 2. Main characteristics of transgenic animals as live bioreactors and calculations of expected flock size for FIX production

Animal	Gestation, months	Maturation, months	Milk output, l ^{*#}	Initiation of transgene to lactation time (month)	Estimated productivity values, g ^{*#} ,	Calculated productivity values for rFIX, g [*]	Reported FIX secretion values, corrected to actual concentration of active form
Mouse	0.75	1	0.0015	3-6	0.01- 0.02	0.000 045	30 mg/l [53]
Rabbit	1	5-6	2-5	7-8	20	-	-
Sheep	5	6-8	200-500	16-18	2500	5 – 12.5	25 mg/l, inactive [50]
Goat	5	6-8	600-800	16-18	4000	0.008 – 0.011	0,0137 mg/l [52]
Pig	4	6-8	200-400	15-16	1500	75 - 150	375 mg/l [55]
Cow	9	16	8,000	30-33	4,000-8,000	-	-

* - per year per doe

- data from [58], [59]

ployed for Dolly the Sheep and called Molly and Polly. Similar results have been obtained for transgenic goats – 13.7 µg/l with >90% of active “gamma-glycosylated” form [52], and mice – up to 60 mg/l at 50% of biologically active FIX [53].

The most successful studies were those that used pigs [54 - 56]. Despite good theoretically predicted yields, actual factor IX production levels were moderate (summarized in Table 2). The supposed rate-limiting step in the secretion of FIX by porcine mammary gland cells is γ-carboxylation. Full specific activity (i.e. complete carboxylation) of the product was noted for these animals, producing FIX at 200 mg/l [54], and only 10-20% of normal specific activity was noted for FIX from pigs, producing at the level of 2-3 g/l [55]. Nevertheless, a viable purification process was developed for this source of under-carboxylated FIX, and the purified product, highly enriched in fully carboxylated FIX, was found to be correctly glycosylated [56] and is expected to be included in clinical trials. It should be noted that usage of milk from transgenic pigs, instead of the bioreactor harvest medium, results in approximately a 10-fold higher concentration of the target protein at the expense of a much higher level of contaminating proteins, lipids, and lack of sterility. Another protein with comparable structure complexity – antithrombin III – was expressed in goat milk at a 1-2 g/l level and purified to pharmaceutical grade with a 53% yield [57]. The consumption of FIX in the U.S. can be estimated at 2 kg/year; and world consumption, at 40 kg/year. Based on the known milk output, 40 pigs will cover the U.S. market of FIX and 800 will cover the world at the present level of expression and 50% total process yield.

Other potential sources of biopharmaceutical proteins are the seeds and tissue of transgenic plants. At present, production of functionally active VKD proteins in transgenic plants is impossible, because the plants lack γ-carboxylases [60]. This limitation may be bypassed by co-expression in the plants of mammalian γ-carboxylases and their co-enzymes, but it is unlikely that such a complex task will be accomplished in the near future.

Expression of inactive FIX in transgenic plants has been reported to date in transgenic tomatoes [61] and soybean seeds [62]. The expression level of FIX in tomatoes was quite low: 15.84 µg/kg of fresh fruits. In the case of soybean seeds, a very promising 800 mg/kg expression level was achieved. Both systems yielded a mature glycosylated form of the target protein.

In some rare cases of hemophilia B (1.5–3%), large titers of neutralizing antibodies are developed by the immune system of patients in response to replacement therapy by FIX preparations [63]. These inhibitor antibodies render ordinary therapy and prophylaxis by FIX ineffective and require higher doses of FIX or immune tolerance induction (ITI) by frequent infusions of very high quantities of FIX. ITI protocols for hemophilia B last from months to years and are dangerous for patients due to the development of the nephritic syndrome and life-threatening anaphylactic reactions. At least some of the side effects in ITI treatments are caused by the excessive pro-coagulant activity of the FIX infused, and inactive variants of FIX may better serve as the immune-tolerance-inducing agent. The fusion protein of FIX and the known transmucosal carrier cholera toxin β-subunit were expressed in tobacco

chloroplasts (400 mg/kg of leaf tissue) and tested as administered orally frozen leaf powder for the prevention of inhibitor antibodies formation in FIX knockout mice treated with human FIX [64]. The control animals, fed the leaf powder from untransformed tobacco, developed 2-90 Beteshda units per ml (BU/ml) of inhibitor antibodies toward human FIX, and in mice fed the leaf powder with encapsulated fusion of the toxin subunit and FIX, the inhibitor level was indistinguishable from the baseline. These data are backed by the mortality rate of treated and control mice – 10% vs. 75% after eight weekly injections of human FIX.

VARIANTS OF RECOMBINANT FIX

Historical use of plasma-derived FIX for hemophilia B therapy dictated the utilization of intact recombinant FIX as a drug. The exact recombinant copy of the natural FIX is expected to be nonimmunogenic for patients and perform comparably to the proven plasma-derived preparations. At the same time, current replacement therapy using FIX requires very frequent infusions of large quantities of high-priced FIX drugs. Modifications to the FIX molecule making it more active or more stable in the bloodstream may prove significantly advantageous to the patients and healthcare professionals involved.

The specific activity of FIX in the blood coagulation assay may be increased threefold by a single point mutation, Arg338-> Ala338 [65], without affecting the tenase complex assembly. An exchange of 2 aminoacids and 2 short surface loops between FIX and factor X molecules was sufficient for the creation of a hybrid molecule with very high proteolytic activity and enzyme specificity, which is typical of factor X [66]. Further work limits the minimal set of point mutations required to 3 – Lys98, Tyr177, and Tyr94 [67]. No data on pro-coagulant activity were collected for these mutant forms of FIX.

Since the majority of hemophiliacs suffer from the absence of factor VIII, a protein cofactor of FIX, engineered variants of FIX capable of direct activation of FX may serve as therapy for hemophilia A. They are also insensitive to inhibitor antibodies toward factor VIII. The triple mutant of FIX Val181Ile, Lys265Thr, and Ile383Val employed as the gene therapy vector successfully bypassed factor VIII and corrected the hemophilia A phenotype in mice [68].

The pharmacokinetic properties of FIX can be improved by genetic fusions with long-lasting plasma proteins or covalent conjugation with hydrophilic polymers. Fusion of FIX and human serum albumin has been realized through the noncleavable short linker peptide or the linker peptide cleavable by factor Xia, simultaneously with the activation of FIX [69]. The

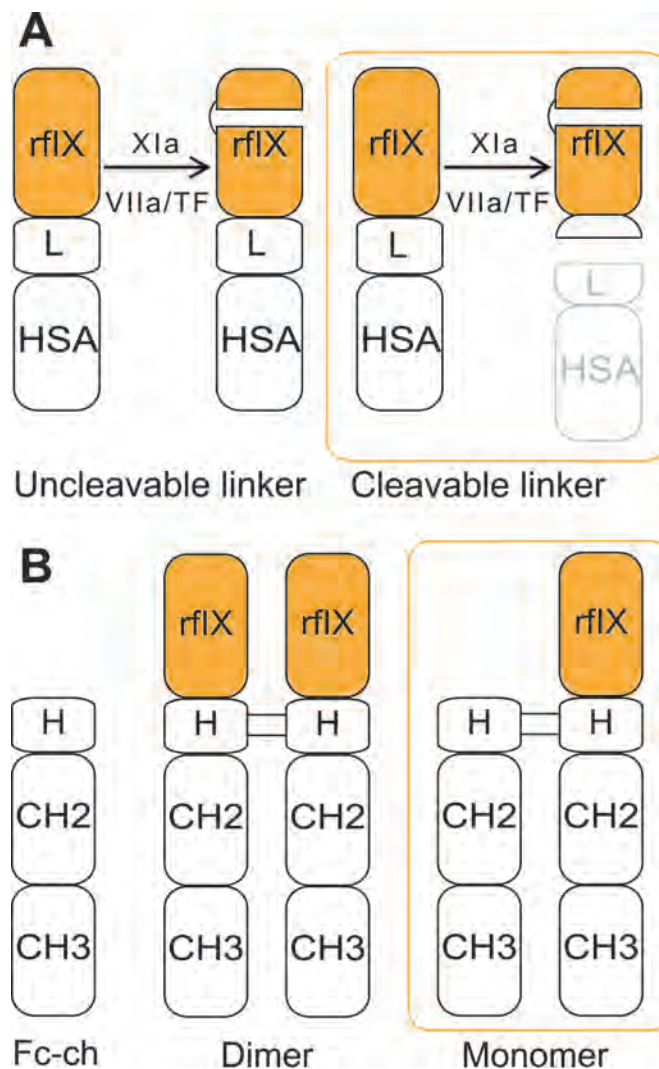


Fig. 3. Genetic fusions of FIX. Panel A fusion of FIX and HAS, panel B fusion of FIX and Fc fragment of IgG. L – linker peptide, H – hinge region of IgG, CH2 and CH3 – second and third constant domains of IgG.

noncleavable linkers were $(G)_6V$ and $SS(GGS)_6GS$, the cleavable linkers were derived from FIX and consisted of the amino-terminal activation site encompassing the amino acids 136 to 154 or 137 to 154 of mature FIX. Activation of rIX-FP with a noncleavable linker results in an activated FIXa molecule with albumin still attached to it, and activation of rIX-FP with a cleavable linker results in the release of albumin and a FIXa molecule that differs from wild-type FIXa only by a short C-terminal linker fragment derived from the FIX molecule itself (Fig. 3A). Utilization of the cleavable linker resulted in a 10- to 30-fold increase in the specific activity of fused FIX in the coagulation assay, and the variant of albumin fusion demonstrated a significant increase in

the half-life in animal models and efficacy in reducing bleeding time in FIX(-/-) mice.

The presence of the Fc domain of immunoglobulin G enables the fusion protein to bind to the neonatal Fc receptor (FcRn), a heterodimer of the MHC class I-like protein heavy chain with beta-2-microglobulin. FcRn protects Fc-containing IgG molecules from catabolism by reversible binding to the surface of endothelial cells [70].

Fusion of FIX and the Fc fragment expressed in human HEK-293H cells and isolated as the covalent heterodimer of the FIX-Fc chain and free Fc chain (Fig. 3B) has acceptable specific pro-coagulant activity (ca. 50 IU/mg) and has demonstrated a significant increase in terminal half-life in many animal models [44]. The half-life in monkey was 47.3 ± 9.1 h versus 12.7 h for intact FIX. A similar 3-fold increase in half-life was recorded in phase I/II clinical trials of the FIX-Fc fusion [71].

Group-specific attachment of PEG moiety to the N-glycans of FIX (Fig. 2) will lead to the conjugate, which will convert to natural, activated FIX, leaving the attached PEG on the released activation peptide. This kind of conjugate, with a single 40-kDa PEG group attached (N9-GP), was used in animal studies [72] and clinical trials at dose levels of 25-100 U/kg, showing a mean half-life of 93 h, 5 times higher than that of the intact FIX [73]. It is interesting to note that the incremental recovery of N9-GP was 94% higher compared with recombinant FIX, a possible indication that the PEG group may shield the FIX molecule from an undesired interaction with the cell surface or block the ability of N-glycans to mediate such an interaction. This difference in activity recovery may also be caused by the full sialation of glycans performed *in vitro* after purification of the FIX.

Prolonged action of FIX may be achieved by various encapsulation techniques, allowing a slow release of the entrapped protein into circulation. The carrier material can be a biodegradable polymer, liposome, etc. A detailed description of this field is out of scope of the present review. A very unusual method of FIX encapsulation was recently developed for human trials [74]. Red blood cells were mixed with FIX *ex vivo*, shocked for the encapsulation of the target protein inside the cells, and injected back to the patient. The entrapped FIX slowly released from the red blood cell ghosts upon their lysis in the bloodstream.

GENE THERAPY OF HEMOPHILIA B

Simultaneously with the development of recombinant FIX protein therapeutics, gene therapy strategies for hemophilia B treatment have been developed. Protein substitution therapy has obvious limitations: treatment

is not curative, and during all of the patients' lives there remains a significant risk of bleeding episodes and chronic joint damage. Other general disadvantages of constant protein infusions are the high cost of the treatment, limited availability of the medication, low half-life of the clotting factor, and risk of neutralizing antibodies (inhibitors) formation toward the administered FIX protein.

Both forms of hemophilia are a particularly good target for gene therapy since they are caused by a well-known single gene defect and have a broad therapeutic window: achievement of 1% of the normal plasma FIX level can prevent most patient's risks, and a concentration of the clotting factor as high as 150%, likewise, is not expected to cause any side-effects. For gene therapy, the therapeutic level of the expressed FIX is usually considered as 5-10% of the normal plasma level; in this case further protein injections might be avoided. A low (under-therapeutic) expression level of FIX may still be enough for immune tolerance induction in patients suffering from the inhibitory form of the disease.

Transfer of the FIX gene *in vivo* is possible even by the naked plasmid DNA, as has been shown in animal models. Hydrodynamic injection of expression plasmid containing human FIX cDNA and the hepatic control region was sufficient to achieve therapeutic levels of FIX in deficient mice for 210 days [75]. The delivery technique employed – injection of 50 µg of the plasmid in 2 ml of solution in 5-8 sec into the tail vein – is definitely unsuitable for human therapy, and a significant modification of hydrodynamic injection should be invented before clinical studies can take place.

Target cDNA may be delivered more efficiently by chromosome-integrating viral vectors of retroviral (RV) or lentiviral (LV) origin or by predominantly episomal adenoviruses (AV) or adeno-associated viruses (AAV).

Historically, RV particles had been used first for the transduction *ex vivo* of fibroblasts from model animals by FIX cDNA and subsequent re-implantation of the modified cells. A low level of human FIX was detected in the plasma of the treated animals [76]. A very small percentage of the animals with re-implanted transduced fibroblasts test positive for FIX production, but the effect remains stable for more than 600 days on the rabbit model [77]. A phase I clinical trial for hemophilia B was conducted with autologous skin fibroblasts transduced *ex vivo* with FIX-encoding γ -retroviral vectors [78] and resulted in a transient, moderate increase in the FIX plasma level in two patients.

Lentiviral vectors, in contrast to γ -retroviral vectors, are able to transduce hepatocytes of the adult liver *in vivo*. Therapeutic levels of FIX have been achieved (transiently) in adult hemophilic mice, following in-

travenous injection of LV [79]. LV are also able to effectively transduce spleen antigen-presenting cells (APC), leading to an immune response against circulating transgene proteins [80]. This unwanted ability of LV may be diminished by restricting transgene expression to certain cell types by utilizing tissue-specific promoter sequences and by co-expression of microRNA's, eliminating off-target expression. Long-term FIX expression in mice using a hepatocyte-specific promoter and hematopoietic cells-specific microRNA miR-142-3p has resulted in a more than 10% level of circulated FIX for 280 days in hemophilia B model mice [81]. No antibodies toward FIX were detected, and all animals survived after a challenge by tail-clip.

Expression of FIX may also be restricted to hematopoietic cells to ensure better availability of the target protein to the sites of its action. Integrin alpha II b promoter-bearing LV constructs, expression-targeted to megakariocytes, were used in hemophilia B mice models and showed promising results: accumulation of FIX in the alpha-granules of platelets and release after activation [82], and phenotype correction was proven by full survivability after tail-clip.

Common to all integrating viral vectors, including RV and LV, are safety concerns of insertional mutagenesis and oncogene activation after vector integration [80]; so vectors with episomic persistence, bearing FIX transgene, are attracting much more attention. High-capacity adenoviral vectors (HCAV) with episomic persistence, bearing no viral genes, are known to trigger a reduced immune response, and the use of a tissue-specific promoter (e.g. hepatic) can further diminish the response, therefore prolonging the gene expression period [83]. HCAV with a liver-specific promoter has yielded therapeutic expression levels of IX with limited toxicity in hemophilic mice [84] and hemophilic dogs [85, 86], and yet a gradual decline in transgene expression was observed. Inhibitor antibodies, as well as hematologic and hepatic toxicities, were detected in animals injected at high vector doses [86], limiting the expression period to 446-604 days in dogs.

Adeno-associated viruses are believed to be better carriers of target genes at the expense of limited packaging capacity, not exceeding 4.7 k.b.p. They are non-pathogenic, replication-deficient, and have a very low probability of chromosomal integration. [87]

AAV particles for gene therapy studies can be manufactured by the GMP-compliant process [88]. In most clinical trials with AAV vectors intramuscular injection or portal vein infusion routes have been used, directing viral particles to skeletal muscle cells or hepatocytes.

Long-term FIX expression has been achieved following muscle-directed gene transfer by the AAV2-FIX vector in hemophilia B dogs with a missense

mutation [89]. Inhibitor antibodies development depended on the nature of the FIX gene defect in the treated animals – dogs with missense mutation in the FIX gene developed virtually no inhibitors [89], and dogs with premature stop codon and unstable mRNA in the FIX gene developed a significant level of the inhibitor [90] that correlated with the AAV dose used [91]. A phase I clinical trial for intramuscular injection of AAV2-FIX vectors was performed in hemophilia B patients with missense mutations using a limited vector dose per site [92-94]. The treatment proved safe but ineffective: the achieved FIX levels were below the therapeutic level.

Utilization of a more invasive procedure aimed at delivering AAV particles to the liver resulted in FIX expression without inhibitor development in normal and hemophilic mice, hemophilia B dogs, and nonhuman primates [87, 95-99]. An 8-year study in inhibitor-prone null mutation hemophilia B dogs treated with liver-directed AAV2-FIX demonstrated long-term hemophilia correction without inhibitor development [100].

A phase I clinical trial was conducted for hemophilia B patients with intrahepatic infusion of AAV2 vectors encoding FIX and a liver-specific promoter [101]. FIX levels of up to 10% were achieved, but the transgene expression period was no longer than 6 weeks, most likely the result of an immune response toward capsid components. It has been suggested that the rapid elimination of the transduced viruses could be caused by the pre-existing immune response to wild-type AAV2 viruses, which are common in the population [101, 102]. Other serotypes of AAV, namely AAV8 and AAV9, both having a liver tropism and less common in human population [103], have been tested on model animals. AAV8 has been found to be a more effective carrier of the FIX gene than AAV2 in mice and dog models [104], [105] and has demonstrated long-term (up to 5 years) safety and efficacy in nonhuman primates [106], with a constant therapeutic level of transduced FIX in the optimal vector dose group. Phase I clinical trials are being conducted for the AAV8-FIX delivered by peripheral vein infusion. The expected immune reactions were diminished by a several-weeks course of prednisolone; the levels of FIX achieved have been in the range of 1% to 8% of normal values in the six patients treated. Two of them cut back infusions of FIX, and four have gone off infusions [107].

CONCLUSION

Despite a sustained research effort, existing therapy for hemophilia B relies mainly on the infusion of FIX, with no significant justification for the bias for plasma-derived or recombinant protein. Modification of the FIX molecule by domain fusions or conjugation with

PEG may decrease the frequency of infusions, but rather will change the cost and overall safety of the treatment. Gene therapy by FIX cDNA in viral vectors is promising for a significant proportion of patients. It has the potential to diminish the need for transfusions for the majority and completely eliminate that

need for some. Hopefully, the straight increase in the production of recombinant FIX achieved through the adoption of biosimilars produced using more productive cell lines and transgenic animals will yield benefits to the world's population of hemophilia B patients in the years to come. ●

REFERENCES

1. Zhong D, Bajaj M.S., Schmidt A.E., Bajaj S.P. // *J. Biol. Chem.* 2002. V. 277. № 5. P. 3622–3631.
2. Wright I.S. // *Can. Med. Assoc. J.* 1962. V. 86. P. 373–374.
3. Howard E.L., Becker K.C., Rusconi C.P., Becker R.C. // *Arterioscler. Thromb. Vasc. Biol.* 2007. V. 27. № 4. P. 722–727.
4. Green P.M., Giannelli F., Sommer S.S., Poon M.-C., Ludwig M., Schwaab R., Reitsma P.H., Goossens M., Yoshioka A., Figueiredo M.S., et al. The Haemophilia B Mutation Database – version 13. 2004; <http://www.kcl.ac.uk/ip/peter-green/haemBdatabase.html>.
5. Rogaev E.I., Grigorenko A.P., Faskhutdinova G., Kittler E.L., Moliaka Y.K. // *Science.* 2009. V. 326. № 5954. P. 817.
6. Reitsma P.H., Bertina R.M., Ploos van Amstel J.K., Riemens A., Briet E. // *Blood.* 1988. V. 72. № 3. P. 1074–1076.
7. Evatt B.L., Black C., Batorova A., Street A., Srivastava A. // *Haemophilia.* 2004. V. 10 Suppl. 4. P. 9–13.
8. Tabor E. // *Transfusion.* 1999. V. 39. № 11–12. P. 1160–1168.
9. Kim H.C., McMillan C.W., White G.C., Bergman G.E., Horton M.W., Saidi P. // *Blood.* 1992. V. 79. № 3. P. 568–575.
10. Ludlam C.A., Powderly W.G., Bozzette S., Diamond M., Koerper M.A., Kulkarni R., Ritchie B., Siegel J., Simmonds P., Stanley S., et al. // *Lancet.* 2006. V. 367. № 9506. P. 252–261.
11. Kurachi K., Davie E.W. // *Proc. Natl. Acad. Sci. USA.* 1982. V. 79. № 21. P. 6461–6464.
12. Anson D.S., Austen D.E., Brownlee G.G. // *Nature.* 1985. V. 315. № 6021. P. 683–685.
13. de la Salle H., Altenburger W., Elkaim R., Dott K., Dieterle A., Drillien R., Cazenave J.P., Tolstoshev P., Lecocq J.P. // *Nature.* 1985. V. 316. № 6025. P. 268–270.
14. Busby S., Kumar A., Joseph M., Halfpap L., Insley M., Berkner K., Kurachi K., Woodbury R. // *Nature.* 1985. V. 316. № 6025. P. 271–273.
15. Kaufman R.J., Wasley L.C., Furie B.C., Furie B., Shoemaker C.B. // *J. Biol. Chem.* 1986. V. 261. № 21. P. 9622–9628.
16. Harrison S., Adamson S., Bonam D., Brodeur S., Charlebois T., Clancy B., Costigan R., Drapeau D., Hamilton M., Hanley K., et al. // *Semin. Hematol.* 1998. V. 35. № 2 Suppl. 2. P. 4–10.
17. Bush L., Webb C., Bartlett L., Burnett B. // *Semin. Hematol.* 1998. V. 35. № 2 Suppl. 2. P. 18–21.
18. Lambert T., Recht M., Valentino L.A., Powell J.S., Udata C., Sullivan S.T., Roth D.A. // *Haemophilia.* 2007. V. 13. № 3. P. 233–243.
19. White G., Shapiro A., Ragni M., Garzone P., Goodfellow J., Tubridy K., Courter S. // *Semin. Hematol.* 1998. V. 35. № 2 Suppl. 2. P. 33–38.
20. Rup B. // *Dev. Biol. (Basel).* 2002. V. 109. P. 103–106.
21. Bond M., Jankowski M., Patel H., Karnik S., Strang A., Xu B., Rouse J., Koza S., Letwin B., Steckert J., et al. // *Semin. Hematol.* 1998. V. 35. № 2 Suppl. 2. P. 11–17.
22. Ewenstein B.M., Joist J.H., Shapiro A.D., Hofstra T.C., Leissinger C.A., Seremetis S.V., Broder M., Mueller-Velten G., Schwartz B.A. // *Transfusion.* 2002. V. 42. № 2. P. 190–197.
23. Bjorkman S. // *Haemophilia.* 2011. V. 17. № 2. P. 179–184.
24. Rouse J.C., McClellan J.E., Patel H.K., Jankowski M.A., Porter T.J. // *Methods Mol. Biol.* 2005. V. 308. P. 435–460.
25. Derian C.K., VanDusen W., Przysiecki C.T., Walsh P.N., Berkner K.L., Kaufman R.J., Friedman P.A. // *J. Biol. Chem.* 1989. V. 264. № 12. P. 6615–6618.
26. Spitzer S.G., Kuppaswamy M.N., Saini R., Kasper C.K., Birktoft J.J., Bajaj S.P. // *Blood.* 1990. V. 76. № 8. P. 1530–1537.
27. Rees D.J., Jones I.M., Handford P.A., Walter S.J., Esnouf M.P., Smith K.J., Brownlee G.G. // *EMBO J.* 1988. V. 7. № 7. P. 2053–2061.
28. Ahmad S.S., Rawala R., Cheung W.F., Stafford D.W., Walsh P.N. // *Biochem. J.* 1995. V. 310 (Pt. 2). P. 427–431.
29. Chang Y.J., Bajaj S.P., Wu H.L., Hamaguchi N., Hsu Y.C., Lin S.W. // *J. Biol. Chem.* 2002. V. 277. № 28. P. 25393–25399.
30. Larson P.J., Stanfield-Oakley S.A., VanDusen W.J., Kasper C.K., Smith K.J., Monroe D.M., High K.A. // *J. Biol. Chem.* 1996. V. 271. № 7. P. 3869–3876.
31. Harris R.J., van Halbeek H., Glushka J., Basa L.J., Ling V.T., Smith K.J., Spellman M.W. // *Biochemistry.* 1993. V. 32. № 26. P. 6539–6547.
32. Agarwala K.L., Kawabata S., Takao T., Murata H., Shimomishi Y., Nishimura H., Iwanaga S. // *Biochemistry.* 1994. V. 33. № 17. P. 5167–5171.
33. Kaufman R.J. // *Thromb. Haemost.* 1998. V. 79. № 6. P. 1068–1079.
34. Makino Y., Omichi K., Kuraya N., Ogawa H., Nishimura H., Iwanaga S., Hase S. // *J. Biochem.* 2000. V. 128. № 2. P. 175–180.
35. Bharadwaj D., Harris R.J., Kisiel W., Smith K.J. // *J. Biol. Chem.* 1995. V. 270. № 12. P. 6537–6542.
36. Sunnerhagen M.S., Persson E., Dahlqvist I., Drakenberg T., Stenflo J., Mayhew M., Robin M., Handford P., Tilley J.W., Campbell I.D., et al. // *J. Biol. Chem.* 1993. V. 268. № 31. P. 23339–23344.
37. McGrath B.M., Walsh G. *Directory of therapeutic enzymes.* N.Y.: Taylor & Francis, 2005. 312 pp.
38. Dadehbeigi N., Ostad S.N., Faramarzi M.A., Ghahremani M.H. // *Biotechnol. Lett.* 2008. V. 30. № 11. P. 1909–1912.
39. De Castilho Fernandes A., Fontes A., Gonsales N., Swiech K., Picanco-Castro V., Faca S., Covas D. // *Biotechnol. Appl. Biochem.* 2011. V. 58. № 4. P. 243–249.
40. Kim W.H., Kim J.S., Yoon Y., Lee G.M. // *J. Biotechnol.* 2009. V. 142. № 3–4. P. 275–278.
41. Lim I., Kim J.-S., Lee G., Choi M., Yoon Y. // *Cells and Culture / Ed. Noll T.* Amsterdam: Springer Netherlands, 2010. P. 613–618.
42. Wasley L.C., Rehemtulla A., Bristol J.A., Kaufman R.J. // *J. Biol. Chem.* 1993. V. 268. № 12. P. 8458–8465.
43. Lusson J., Vieau D., Hamelin J., Day R., Chretien M., Seidah N.G. // *Proc. Natl. Acad. Sci. USA.* 1993. V. 90. № 14. P. 6691–6695.
44. Peters R.T., Low S.C., Kamphaus G.D., Dumont J.A., Amari J.V., Lu Q., Zarbis-Papastoitis G., Reidy T.J., Merriks E.P., Nichols T.C., et al. // *Blood.* 2010. V. 115. № 10. P. 2057–2064.

REVIEWS

45. Gillis S., Furie B.C., Furie B., Patel H., Huberty M.C., Switzer M., Foster W.B., Scoble H.A., Bond M.D. // *Protein Sci.* 1997. V. 6. № 1. P. 185–196.
46. Wajih N., Hutson S.M., Owen J., Wallin R. // *J. Biol. Chem.* 2005. V. 280. № 36. P. 31603–31607.
47. McClure D.B., Walls J.D., Grinnell B.W. // *J. Biol. Chem.* 1992. V. 267. № 27. P. 19710–19717.
48. Bolt G., Steenstrup T.D., Kristensen C. // *Thromb. Haemost.* 2007. V. 98. № 5. P. 988–997.
49. Vatandoost J., Zomorodipour A., Sadeghizadeh M., Aliyari R., Bos M.H., Ataei F. // *Biotechnol. Prog.* 2012. V. 28. № 1. P. 45–51.
50. Clark A.J., Ali S., Archibald A.L., Bessos H., Brown P., Harris S., McClenaghan M., Prowse C., Simons J.P., Whitelaw C.B., et al. // *Genome.* 1989. V. 31. № 2. P. 950–955.
51. Schnieke A.E., Kind A.J., Ritchie W.A., Mycock K., Scott A.R., Ritchie M., Wilmut I., Colman A., Campbell K.H. // *Science.* 1997. V. 278. № 5346. P. 2130–2133.
52. Zhang K., Wang H., Bao Y., Lu D., Xue J., Qiu X., Huang S., Huang Y., Li B., Li H., et al. // *Chinese Sci. Bull.* 1997. V. 42. № 15. P. 1308–1313.
53. Yull F., Harold G., Wallace R., Cowper A., Percy J., Cottingham I., Clark A.J. // *Proc. Natl. Acad. Sci. USA.* 1995. V. 92. № 24. P. 10899–10903.
54. van Cott K.E., Butler S.P., Russell C.G., Subramanian A., Lubon H., Gwazdauskas F.C., Knight J., Drohan W.N., Velandier W.H. // *Genet. Anal.* 1999. V. 15. № 3–5. P. 155–160.
55. Lindsay M., Gil G.C., Cadiz A., Velandier W.H., Zhang C., van Cott K.E. // *J. Chromatogr. A.* 2004. V. 1026. № 1–2. P. 149–157.
56. Gil G.C., Velandier W.H., van Cott K.E. // *Glycobiology.* 2008. V. 18. № 7. P. 526–539.
57. Edmunds T., van Patten S.M., Pollock J., Hanson E., Bernasconi R., Higgins E., Manavalan P., Ziomek C., Meade H., McPherson J.M., et al. // *Blood.* 1998. V. 91. № 12. P. 4561–4571.
58. Dove A. // *Nat. Biotechnol.* 2000. V. 18. № 10. P. 1045–1048.
59. Panno J. *Animal Cloning: The Science of Nuclear Transfer.* N.Y.: Facts on File, 2004. 176 p.
60. Gomord V., Faye L. // *Curr. Opin. Plant Biol.* 2004. V. 7. № 2. P. 171–181.
61. Zhang H., Zhao L., Chen Y., Cui L., Ren W., Tang K. // *Biotechnol. Appl. Biochem.* 2007. V. 48. Pt 2. P. 101–107.
62. Cunha N.B., Murad A.M., Ramos G.L., Maranhao A.Q., Brigido M.M., Araujo A.C., Lacorte C., Aragao F.J., Covas D.T., Fontes A.M., et al. // *Transgenic Res.* 2011. V. 20. № 4. P. 841–855.
63. DiMichele D. // *Br. J. Haematol.* 2007. V. 138. № 3. P. 305–315.
64. Verma D., Moghimi B., LoDuca P.A., Singh H.D., Hoffman B.E., Herzog R.W., Daniell H. // *Proc. Natl. Acad. Sci. USA.* 2010. V. 107. № 15. P. 7101–7106.
65. Chang J., Jin J., Lollar P., Bode W., Brandstetter H., Hamaguchi N., Straight D.L., Stafford D.W. // *J. Biol. Chem.* 1998. V. 273. № 20. P. 12089–12094.
66. Hopfner K.P., Brandstetter H., Karcher A., Kopetzki E., Huber R., Engh R.A., Bode W. // *EMBO J.* 1997. V. 16. № 22. P. 6626–6635.
67. Sichler K., Kopetzki E., Huber R., Bode W., Hopfner K.P., Brandstetter H. // *J. Biol. Chem.* 2003. V. 278. № 6. P. 4121–4126.
68. Milanov P., Ivanciu L., Abriss D., Quade-Lysson P., Miesbach W., Alesci S., Tonn T., Grez M., Seifried E., Schuttrumpf J. // *Blood.* 2012. V. 119. № 2. P. 602–611.
69. Metzner H.J., Weimer T., Kronthaler U., Lang W., Schulte S. // *Thromb. Haemost.* 2009. V. 102. № 4. P. 634–644.
70. Dumont J.A., Low S.C., Peters R.T., Bitonti A.J. // *BioDrugs.* 2006. V. 20. № 3. P. 151–160.
71. Shapiro A.D., Ragni M., Valentino L.A., Key N.S., Josephson N., Powell J., Cheng G., Tubridy K.L., Peters R., Dumont J., et al. // *Haemophilia.* 2010. V. 16. № Suppl. s4. P. 1–158.
72. Ostergaard H., Bjelke J.R., Hansen L., Petersen L.C., Pedersen A.A., Elm T., Moller F., Hermit M.B., Holm P.K., Krogh T.N., et al. // *Blood.* 2011. V. 118. № 8. P. 2333–2341.
73. Negrier C., Knoke K., Tiede A., Giangrande P., Moss J. // *Blood.* 2011. V. 118. № 10. P. 2695–2701.
74. Sinauridze E.I., Vuimo T.A., Kulikova E.V., Shmyrev I.I., Ataullakhanov F.I. // *Med. Sci. Monit.* 2010. V. 16. № 10. P. PI19–26.
75. Kim H.S., Kim J.C., Lee Y.K., Kim J.S., Park Y.S. // *J. Gene Med.* 2011. V. 13. № 7–8. P. 365–372.
76. Palmer T.D., Thompson A.R., Miller A.D. // *Blood.* 1989. V. 73. № 2. P. 438–445.
77. Chen L., Nelson D.M., Zheng Z., Morgan R.A. // *Hum. Gene Ther.* 1998. V. 9. № 16. P. 2341–2351.
78. Qiu X., Lu D., Zhou J., Wang J., Yang J., Meng P., Hsueh J.L. // *Chin. Med. J. (Engl.)* 1996. V. 109. № 11. P. 832–839.
79. Tsui L.V., Kelly M., Zayek N., Rojas V., Ho K., Ge Y., Moskalenko M., Mondesire J., Davis J., Roey M.V., et al. // *Nat. Biotechnol.* 2002. V. 20. № 1. P. 53–57.
80. Petrus I., Chuah M., van den Driessche T. // *J. Gene Med.* 2010. V. 12. № 10. P. 797–809.
81. Brown B.D., Cantore A., Annoni A., Sergi L.S., Lombardo A., Della Valle P., D'Angelo A., Naldini L. // *Blood.* 2007. V. 110. № 13. P. 4144–4152.
82. Zhang G., Shi Q., Fahs S.A., Kuether E.L., Walsh C.E., Montgomery R.R. // *Blood.* V. 116. № 8. P. 1235–1243.
83. Pastore L., Morral N., Zhou H., Garcia R., Parks R.J., Kochanek S., Graham F.L., Lee B., Beaudet A.L. // *Hum. Gene Ther.* 1999. V. 10. № 11. P. 1773–1781.
84. Ehrhardt A., Kay M.A. // *Blood.* 2002. V. 99. № 11. P. 3923–3930.
85. Ehrhardt A., Xu H., Dillow A.M., Bellinger D.A., Nichols T.C., Kay M.A. // *Blood.* 2003. V. 102. № 7. P. 2403–2411.
86. Brunetti-Pierri N., Nichols T.C., McCorquodale S., Merriks E., Palmer D.J., Beaudet A.L., Ng P. // *Hum. Gene Ther.* 2005. V. 16. № 7. P. 811–820.
87. Snyder R.O., Miao C.H., Patijn G.A., Spratt S.K., Danos O., Nagy D., Gown A.M., Winther B., Meuse L., Cohen L.K., et al. // *Nat. Genet.* 1997. V. 16. № 3. P. 270–276.
88. Allay J.A., Sleep S., Long S., Tillman D.M., Clark R., Carney G., Fagone P., McIntosh J.H., Nienhuis A.W., Davidoff A.M., et al. // *Hum. Gene Ther.* V. 22. № 5. P. 595–604.
89. Herzog R.W., Yang E.Y., Couto L.B., Hagstrom J.N., Ellwell D., Fields P.A., Burton M., Bellinger D.A., Read M.S., Brinkhous K.M., et al. // *Nat. Med.* 1999. V. 5. № 1. P. 56–63.
90. Herzog R.W., Mount J.D., Arruda V.R., High K.A., Lothrop C.D., Jr. // *Mol. Ther.* 2001. V. 4. № 3. P. 192–200.
91. Herzog R.W., Fields P.A., Arruda V.R., Brubaker J.O., Armstrong E., McClintock D., Bellinger D.A., Couto L.B., Nichols T.C., High K.A. // *Hum. Gene Ther.* 2002. V. 13. № 11. P. 1281–1291.
92. Kay M.A., Manno C.S., Ragni M.V., Larson P.J., Couto L.B., McClelland A., Glader B., Chew A.J., Tai S.J., Herzog R.W., et al. // *Nat. Genet.* 2000. V. 24. № 3. P. 257–261.
93. Manno C.S., Chew A.J., Hutchison S., Larson P.J., Herzog R.W., Arruda V.R., Tai S.J., Ragni M.V., Thompson A., Ozelo M., et al. // *Blood.* 2003. V. 101. № 8. P. 2963–2972.

REVIEWS

94. Jiang H., Pierce G.F., Ozelo M.C., de Paula E.V., Vargas J.A., Smith P., Sommer J., Luk A., Manno C.S., High K.A., et al. // *Mol. Ther.* 2006. V. 14. № 3. P. 452–455.
95. Snyder R.O., Miao C., Meuse L., Tubb J., Donahue B.A., Lin H.F., Stafford D.W., Patel S., Thompson A.R., Nichols T., et al. // *Nat. Med.* 1999. V. 5. № 1. P. 64–70.
96. Nathwani A.C., Davidoff A.M., Hanawa H., Hu Y., Hoffer F.A., Nikanorov A., Slaughter C., Ng C.Y., Zhou J., Lozier J.N., et al. // *Blood.* 2002. V. 100. № 5. P. 1662–1669.
97. Mount J.D., Herzog R.W., Tillson D.M., Goodman S.A., Robinson N., McClelland M.L., Bellinger D., Nichols T.C., Arruda V.R., Lothrop C.D., Jr., et al. // *Blood.* 2002. V. 99. № 8. P. 2670–2676.
98. Wang L., Nichols T.C., Read M.S., Bellinger D.A., Verma I.M. // *Mol. Ther.* 2000. V. 1. № 2. P. 154–158.
99. Wang L., Takabe K., Bidlingmaier S.M., Ill C.R., Verma I.M. // *Proc. Natl. Acad. Sci. USA.* 1999. V. 96. № 7. P. 3906–3910.
100. Niemeyer G.P., Herzog R.W., Mount J., Arruda V.R., Tillson D.M., Hathcock J., van Ginkel F.W., High K.A., Lothrop C.D., Jr. // *Blood.* 2009. V. 113. № 4. P. 797–806.
101. Manno C.S., Pierce G.F., Arruda V.R., Glader B., Ragni M., Rasko J.J., Ozelo M.C., Hoots K., Blatt P., Konkle B., et al. // *Nat. Med.* 2006. V. 12. № 3. P. 342–347.
102. Mingozzi F., Maus M.V., Hui D.J., Sabatino D.E., Murphy S.L., Rasko J.E., Ragni M.V., Manno C.S., Sommer J., Jiang H., et al. // *Nat. Med.* 2007. V. 13. № 4. P. 419–422.
103. van den Driessche T., Thorrez L., Acosta-Sanchez A., Petrus I., Wang L., Ma L., Waelle D.E., Iwasaki Y., Gillijns V., Wilson J.M., et al. // *J. Thromb. Haemost.* 2007. V. 5. № 1. P. 16–24.
104. Cooper M., Nayak S., Hoffman B.E., Terhorst C., Cao O., Herzog R.W. // *Hum. Gene Ther.* 2009. V. 20. № 7. P. 767–776.
105. Wang L., Calcedo R., Nichols T.C., Bellinger D.A., Dil- low A., Verma I.M., Wilson J.M. // *Blood.* 2005. V. 105. № 8. P. 3079–3086.
106. Nathwani A.C., Rosales C., McIntosh J., Rastegarlar G., Nathwani D., Raj D., Nawathe S., Waddington S.N., Bronson R., Jackson S., et al. // *Mol. Ther.* 2011. V. 19. № 5. P. 876–885.
107. Nathwani A.C., Tuddenham E.G., Rangarajan S., Rosales C., McIntosh J., Linch D.C., Chowdary P., Riddell A., Pie A.J., Harrington C., et al. // *N. Engl. J. Med.* 2011. V. 365. № 25. P. 2357–2365.

Contribution of the *TGFB1* Gene to Myocardial Infarction Susceptibility

R.M. Barsova^{1,2}, B.V. Titov^{1,2}, N.A. Matveeva^{1,2}, A.V. Favorov^{3,4}, T.S. Sukhinina²,
R.M. Shahnovich², M.Ia. Ruda², O.O. Favorova^{1,2}

¹ Pirogov Russian National Research Medical University, Moscow, ul. Ostrovitianova, 1, Moscow, 117997, Russia

² Russian Cardiology Scientific and Production Center, Moscow, 3-ia Cherepkovskaia ul., 15a, Moscow, 121552, Russia

³ Vavilov Institute of General Genetics, Russian Academy of Science, Moscow, ul. Gubkina, 3, Moscow, 119991, Russia

⁴ Oncology Biostatistics and Bioinformatics, Johns Hopkins School of Medicine, Baltimore, MD 21205, US

*E-mail: olga_favorova@mail.ru

Received 10.02.2012

Copyright © 2012 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 Carriage frequencies of alleles and genotypes of the *TGFB1* gene polymorphous loci $-509C>T$ (rs1800469), $869T>C$ (rs1982073), $915G>C$ (rs1800471), which affect the level of cytokine TGF- β 1 production, were analyzed in the patients of Russian ethnic descent with myocardial infarction (MI) (406 cases) and in the control group of the same ethnic descent (198 controls). Significant association with MI was observed in carriage frequencies of the allele *TGFB1** $-509T$ ($p=0.046$, OR =1.45, 95% CI: 1.02-2.06), genotypes *TGFB1** $869T/T$ ($p=0.0024$, OR =1.75, 95% CI: 1.22-2.51), and *TGFB1** $915G/G$ ($p=0.048$, OR=1.76, 95% CI: 1.05-2.97). Linkage disequilibrium analysis for these SNPs has shown that the associations revealed can be considered to be independent. A complex analysis of MI association with combinations of alleles/genotypes of said SNPs indicates their cumulative effect. An analysis of susceptibility to early-onset MI (≤ 50 years old) revealed a positive association of the allele *TGFB1** $-509T$ ($p=0.002$, OR=2.24, 95% CI: 1.35-3.71) and genotype *TGFB1** $869T/T$ ($p=0.008$, OR=1.93, 95% CI: 1.18-3.15), as well as their additivity. An analysis of susceptibility to recurrent MI revealed an association of the genotype *TGFB1** $-509T/T$ ($p=0.0078$, OR=2.60, 95% CI: 1.28-5.28). The results obtained indicate the important role of the *TGFB1* gene in susceptibility to MI, including early-onset and recurrent MI, in Russians.

KEYWORDS myocardial infarction; Russians; genes; allelic polymorphism; transforming growth factor β 1; *TGFB1*; APSampler.

ABBREVIATIONS dNTP – deoxynucleoside triphosphate; LD – linkage disequilibrium; SD – standard deviation; SNP – single nucleotide polymorphism; TGF- β 1 – transforming growth factor beta 1; *TGFB1* – gene encoding transforming growth factor beta 1; CI – confidence interval; IHD – ischemic heart disease; MI – myocardial infarction; OR – odds ratio; PCR – polymerase chain reaction; PCR-SSP – polymerase chain reaction with sequence-specific primers; CVD – cardiovascular diseases.

INTRODUCTION

The absolute majority of cardiovascular diseases (CVD) are complex polygenic diseases. To date, a lot of CVD genetic markers have been revealed; among them a special place belongs to the genes, which encode proteins involved in the atherosclerotic process. This is also true for ischemic heart disease (IHD), which includes the myocardial infarction (MI) as one of the forms. Among these markers is the *TGFB1* gene, product of which (transforming growth factor beta 1) belongs to the TGF- β cytokine superfamily and functions both as a pro-atherogenic and anti-atherogenic

factor. There are three single nucleotide polymorphisms of interest: rs1800469 (SNP $-509C>T$) in the promoter, rs1982073 (SNP $869T>C$, Leu10Pro), and rs1800471 (SNP $915G>C$, Arg25Pro) in the signal sequence (exon 1) [1]. All of the polymorphisms studied affect the level of TGF- β 1 production: according to the literature [2-4], allele T of the SNP $-509C>T$, allele T of the SNP $869T>C$, and genotype G/G of the SNP $915G>C$ are associated with the higher protein plasma level.

A number of studies are dedicated to the association analysis of these SNPs with IHD or MI development;

some of them demonstrate such an association [1, 5, 6], while some do not.

Since it was finally proved that genetic susceptibility to many polygenic diseases varies in different ethnic groups, studies should be performed in ethnically homogenous populations. No studies have been performed on the association of the SNP 869T>C and 915G>C of *TGFB1* gene with MI and other CVDs occurrence in Russians. Previously, we observed a positive association of MI occurrence in a smaller sampling with the allele *TGFB1**-509T in combinations with the alleles/genotypes of the SNPs of other inflammation genes, as well as a contribution of the “alternative” allele -509*C to protective combinations [7]. In the present study, a search for the association of rs1800469 (SNP -509C>T), rs1982073 (SNP 869T>C, Leu10Pro), and rs1800471 (SNP 915G>C, Arg25Pro) of the *TGFB1* gene with MI occurrence was performed in more than 400 patients of Russian descent, as well as an analysis of haplotypes of these SNPs in healthy individuals. The distribution of alleles and genotypes of said SNPs were analyzed by a comparison of patients of different age groups and subjects of the control group, as well as groups of patients with a single and recurrent MI(s).

MATERIALS AND METHODS

For the case-control study we used blood samples from 406 patients of Russian descent with MI, with a mean age \pm standard deviation (SD) - 57.5 ± 12.8 years. 272 of the individuals were men (mean age - 53.4 ± 11.9 years), and 134 were women (65.6 ± 10.3 years). The control group consisted of 198 individuals of Russian descent without CVDs in medical history. 112 of the individuals were men (mean age - 57.1 ± 11.9 years), while 86 were women (mean age - 63.2 ± 14.2 years).

MI was diagnosed according to the 2001 AHA/ESC guidelines. The subjects of the control group underwent an examination in order to exclude IHD. All of the patients (or their relatives) and the subjects of the control group gave informed consent for this study.

The extraction of DNA with a phenol-chloroform mixture was performed according to [8].

The polymorphous loci of the *TGFB1* gene were analyzed using the PCR-SSP method. A DNA fragment of 283 bp containing SNP 869T>C was amplified using the allele-specific primers 5'-AGCAGCGGTAGCAGCAGCA-3' (SSP T), 5'-GCAGCGGTAGCAGCAGCG-3' (SSP C), and the common primer 5'-CT ACCTTTT GCC GGGAGACC -3'. In the case of SNP 915G>C, a DNA fragment of 125 bp was amplified using the allele-specific primers 5'-TGGTGCTGACGCCTGGCCG-3' (SSP G), 5'-TGGTGCTGACGCCTGGCCC-3' (SSP C),

and the common primer 5'-GGCGAGCCGCAGCTT-GGACA-3'. All primers were designed using the Vector NT I 7.1 and Primo software [9]. The amplification mixture (10 μ l) contained 70 mM Tris-HCl (pH 9.0), 20 mM (NH₄)₂SO₄, 1.0 mM MgCl₂, 0.025% Tween-20, 0.025% NP-40, 5 pmole of each primer, 0.2 mM dNTP, 0.5 U Taq-polymerase, and 100–200 ng DNA mineral oil. Amplification program: 95°C, 5 min. Then, 10 cycles: 95°C – 1 min, 64°C – 1 min, 72°C – 1 min; and 20 cycles: 95°C – 30 s, 58°C – 50 s, 72°C – 50 s. PCR was performed in MC16 amplifier (DNA-technology, LLC, Russia). The presence of amplified products was checked by electrophoresis in a 2% agarose gel with ethidium bromide. SNP -509C>T was analyzed as described in [7].

Statistical analysis. An analysis of the deviation of the observed genotype frequencies from the Hardy-Weinberg equilibrium and the linkage disequilibrium (LD) analysis were performed using the free Haploview 4.0 software [10]. The carriage frequencies of alleles and genotypes of single SNPs in different groups were compared by a two-sided Fisher's exact test using the free online GraphPad InStat software [11]. We used the APSampler software [12,13] to identify significant associations between MI and carriage of allele/genotype combinations of said *TGFB1* SNPs and to assess the level of association significance for each combination found by the general algorithm by the value of one-sided Fisher's exact test. The differences were considered significant at *p-values* ≤ 0.05 . We computed the odds ratio (OR) and its 95% confidential intervals (CI) for each allelic combination found, and we considered only those combinations, for which CI did not cross 1.

RESULTS

The genomic typing of the polymorphous loci -509C>T, 869T>C and 915G>C of the *TGFB1* gene was carried out in patients with MI and individuals without CVDs, followed by an analysis of a possible association of these SNPs with MI development. We did not observe deviations in the distribution of SNPs alleles and genotypes frequencies from the Hardy-Weinberg equilibrium in the control group. In the group of patients, SNP 869T>C and 915G>C were in the Hardy-Weinberg equilibrium, while -509C>T was not (*p* = 0.0007).

The carriage frequencies of alleles and genotypes of the *TGFB1* gene in the patients and controls are shown in *Figure 1*. The allele *TGFB1**-509T (as genotypes T/T and C/T) was more frequent in the group of patients (*p* = 0.046, OR = 1.45, 95% CI: 1.02–2.06) than in controls, while the genotype *TGFB1**-509C/C was less frequent (*p* = 0.046, OR = 0.69, 95% CI: 0.49–0.98). Furthermore, genotypes *TGFB1**869T/T (*p* = 0.0024, OR = 1.75, 95% CI: 1.22–2.51) and *TGFB1**915G/G (*p* = 0.048, OR = 1.76, 95% CI: 1.05–2.97) were also more frequent in

Table. Association of myocardial infarction with carriage of alleles/genotypes of three *TGFB1* SNPs

SNPs of <i>TGFB1</i> gene			Carriers (%)/non-carriers (%) of the combination		p-value in comparison of frequencies*	OR (95%CI) for reliable differences
-509C>T	869T>C	915G>C	Patients (N=397)	Control group (N=198)		
Predisposing combinations						
T	T/T	G	116 (29.2)/281 (70.8)	33 (16.7)/165 (83.3)	0.00048	2.06 (1.34-3.18)
T	T/T	-	116 (29.2)/281 (70.8)	33 (16.7)/165 (83.3)	0.00048	2.06 (1.34-3.18)
-	T/T	G	181 (45.6)/216 (54.4)	64 (32.3)/134 (67.7)	0.0012	1.75 (1.23-2.51)
T	-	G	273 (67.2)/ 133 (32.8)	116 (58.6)/ 82 (41.4)	0.023	1.45 (1.02-2.06)
Protective combinations						
C	C	C	17 (4.3)/380 (95.7)	23 (11.6)/175 (88.4)	0.00097	0.34 (0.18-0.65)
-	C	C	20 (5.0)/377 (95.0)	23 (11.6)/175 (88.4)	0.0036	0.40 (0.22-0.75)
C	-	C	31 (7.8)/375 (92.2)	29 (14.6)/169 (85.4)	0.0061	0.48 (0.28-0.83)
C	C	-	190 (47.9)/207 (52.1)	117 (59.1)/81 (40.9)	0.0062	0.64 (0.45-0.90)

*For each of the three SNP, the risk allele (genotype) is shown in a darker background color than the protective allele.

**Presented in decreasing order of the significance level for predisposing and protective combinations, individually.

the group of patients. Accordingly, alleles *TGFB1**869C (sum of C/C and C/T genotypes) ($p = 0.0024$, OR = 0.57, 95% CI: 0.40–0.81) and *TGFB1**915C (sum of C/C and G/C) ($p = 0.048$, OR = 0.57, 95% CI: 0.34–0.96) were more frequent in the control group. Thus, carriage of the allele *TGFB1**-509T, or genotype *TGFB1**869T/T, or genotype *TGFB1**915G/G, can be considered as risk factors for MI development. Thereby, an association of the genotype *TGFB1**869T/T with MI development is 20 times more significant than for other markers.

Due to the fact that all of the SNPs above are localized in one gene, and taking into account published data for the linkage disequilibrium of the loci in this region [14–16], we analyzed the possible haplotypes of these SNPs. We used only the control group, because only in this group Hardy–Weinberg equilibrium was observed for all SNPs. Calculation of the pairwise linkage disequilibrium between SNP -509C>T, SNP 915G>C, and

869T>C, as measured by the Haploview 4.0 software [10], revealed a weak linkage ($r^2 < 0.05$ for all pairs). Probably, the association of MI with these SNPs can be considered to be independent. It provides grounds for an analysis of the association of the alleles/genotypes-combined carriage with MI development using the results of APSampler calculations (Table).

As seen from the Table, combined carriage of the allele *TGFB1**-509T and genotype *TGFB1**869T/T, each of which is the MI risk factor (Fig. 1A, B), leads to an increase of both the significance level ($p = 0.00048$) and OR value (2.06) relative to the carriage of each of them alone. An addition of *TGFB1**915G, carriage of which wasn't associated with MI alone (Fig. 1B), to this biallelic combination maintains the p and OR values unchanged, since all (*TGFB1**-509T + *TGFB1**869T/T) combination carriers also bear the allele *TGFB1**915G. Carriage of the other two combinations (*TGFB1**869T/T

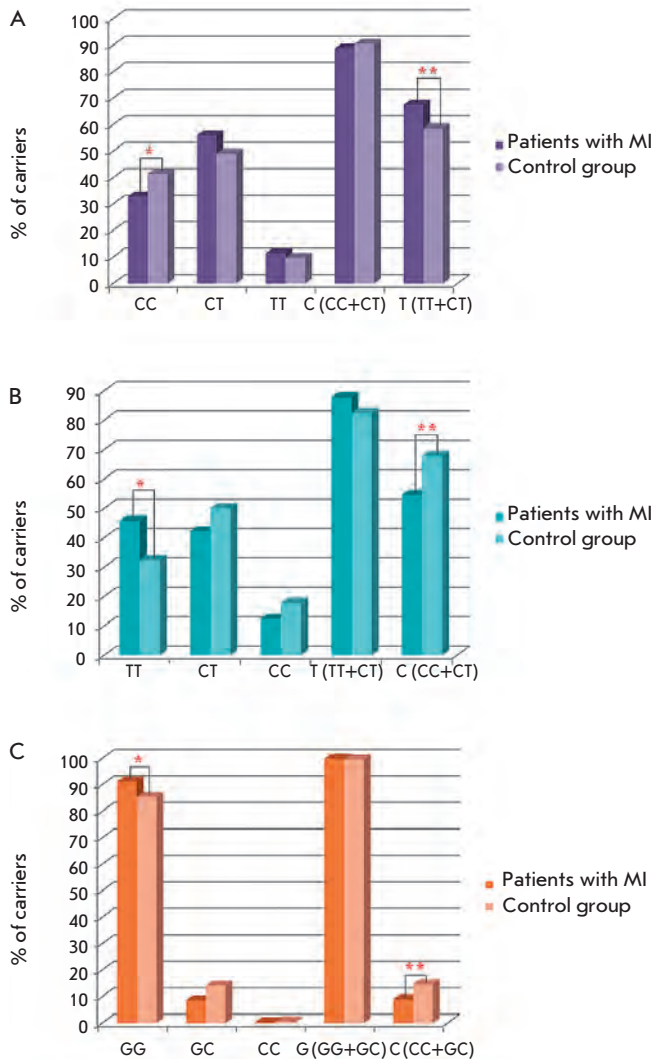


Fig. 1. Carriage frequencies of alleles and genotypes of *TGFBI* gene polymorphous loci in the patients with MI and in the control group. (A) – SNP –509C>T. * $p = 0.046$, OR = 0.69; ** $p = 0.046$, OR = 1.45; (B) – SNP 869T>C. * $p = 0.0024$, OR = 1.75, ** $p = 0.0024$, OR = 0.57; (C) – SNP 915G>C. * $p = 0.048$, OR = 1.76; ** $p = 0.048$, OR = 0.57.

+ *TGFBI**915G) and (*TGFBI**–509T + *TGFBI**915G) is characterized by the less significance level ($p = 0.0012$ and 0.023, respectively) and OR values (1.75 and 1.45, respectively) than the combination (*TGFBI**–509T + *TGFBI**869T/T).

Combined carriage of the alleles *TGFBI**–509C, *TGFBI**869C, and *TGFBI**915C is also highly significant, but it is negatively associated with MI (Table). A single negative association with MI was shown for the two latter alleles, but not for *TGFBI**–509C (Fig. 1). In this case, the significance level and OR dissimilarity

from 1 for the triallelic combination ($p = 0.00097$, OR = 0.34) were higher than for all three biallelic combinations embodied ($p = 0.0036$ to 0.0062; OR 0.40 to 0.64). Furthermore, alleles/genotypes of the three SNPs included in protective combinations are alternative to the alleles included in the predisposing combination.

It is known that genetic risk factors are often more significant for MI development at a younger age (in the case of early-onset MI). On this basis, we analyzed the alleles/genotypes distribution of polymorphous loci in patients of different age subgroups with MI. By the allocation of the a subgroup of patients who had developed the MI before 50 years of age inclusive (121 individuals), comparison of them with the total control group showed differences similar to the differences in the total sample. Thus, the allele *TGFBI**–509T ($p = 0.002$, OR = 2.24, 95% CI: 1.35–3.71) and genotype *TGFBI**869T/T ($p = 0.008$, OR = 1.93, 95% CI: 1.18–3.15) were more frequent in the group of patients younger than 50, while the genotype *TGFBI**–509C/C ($p = 0.002$, OR = 0.45, 95% CI: 0.27–0.74) and allele *TGFBI**869C ($p = 0.008$, OR = 0.52, 95% CI: 0.32–0.85) were more frequent in the control group. Furthermore, the complex analysis revealed the combination of the allele *TGFBI**–509T and genotype *TGFBI**869T/T ($p = 0.00015$, OR = 2.73, 95% CI: 1.60–4.63), which is associated with an early-onset MI and analogous to that obtained in the total group.

Then, we divided the group of patients into subgroups of patients with first MI (73 individuals) and with recurrent MI(s) (226 individuals) and compared their genotypes with each other. We found an association with only one SNP: genotype *TGFBI**–509T/T was more frequent in the group of recurrent MI(s), then in the group of patients with only one MI ($p = 0.0078$, OR = 2.60, 95% CI: 1.28–5.28), whereas allele *TGFBI**–509C was protective ($p = 0.016$, OR = 0.38, 95% CI: 0.19–0.78). Combinations found by complex analysis were less significant than single association of SNP –509C>T.

DISCUSSION

In this paper we present data on the distribution of alleles and genotypes of three functional polymorphous loci of the *TGFBI* gene in a population sample of Russians (control group). We could find published data on the frequency of alleles/genotypes only for one of the SNPs studied in Russians: namely 869T>C, notably in a group of men [17]. The genotype frequencies obtained in that study are close to the frequencies identified in the present study (Fig. 2B). Figure 2 shows data not only for Russians, but also published data on the genotype distribution of the SNP–509C>T, SNP 869T>C and SNP 915G>C of the *TGFBI* gene in different Caucasian populations. Generally, the genotype frequen-

cies we observed in Russians fall within a fairly wide range of frequencies, which are described for different populations of Europe. Besides, the picture varies for individual SNPs: if there is complete uniformity of genotype distribution for 915G>C in all the Caucasian populations studied (Fig. 2C), then in the case of SNP-509C>T and SNP 869T>C, ethnic differences reach the significance level ($p = 0.001$) when the genotype frequencies of -509T/T in Russians and Italians are compared (Fig. 2A).

These data may reflect differences in the linkage patterns of *TGFB1* polymorphous loci in different populations. Although many studies on Caucasians have revealed different haplotypes containing *TGFB1* polymorphous loci, we obtained no similar data for Russians. Probably, this is due to the population-specific character of LD pattern formation, while the differences observed reflect the significant ethno-specific variability of haploblocks [18].

We have shown the contribution of the *TGFB1* gene, specifically the contribution of carriage of the polymorphous loci -509C>T, 869T>C and 915G>C of the *TGFB1* gene, as well as their combinations to MI genetic susceptibility in Russians. The comparison of significance level and OR values for these combinations with those of individual alleles/genotypes suggests that, in case of their combined carriage, a cumulative effect occurs, which most likely reflects summation of the independent contributions of different polymorphous loci of the same gene in MI development. Since all the risk alleles/genotypes we found (*TGFB1**-509T, *TGFB1**869T/T and *TGFB1**915G/G) are associated with the higher level of gene expression [2-4], we can assume that cumulative association is determined by the unidirectionality of changes in the TGF- β 1 protein level.

In the analysis of genetic susceptibility to early-onset MI, we observed a significant association of SNP *TGFB1**-509C>T and 869T>C, but not for 915G>C. In the case of recurrent MIs, significant associations were observed only for SNP *TGFB1**-509C>T. Besides, the risk alleles were the same as for the total group of patients. Thereby, SNP -509C>T and SNP 869T>C can be considered as MI markers irrespective of age, including early-onset MI, while the SNP *TGFB1**-509C>T can serve as a predictive marker of recurrent MI development. However, we cannot exclude that the reduction in the number of associated markers in the subgroups analyzed, compared to the total group of patients, may be caused by the reduction in the size of the sample.

The associations found are in general consistent with the results obtained for other Caucasians, although it should be noted that significant variations are found in the published data. The results of study of early-onset

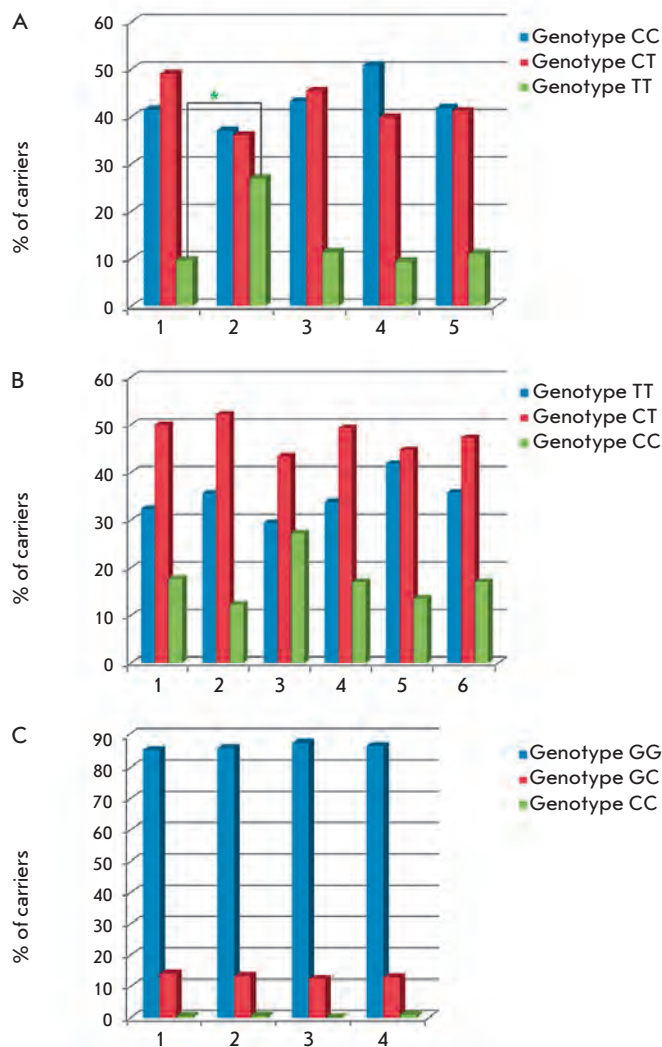


Fig. 2. Carriage frequencies of genotypes of the *TGFB1* gene polymorphous loci in different Caucasian populations. (A) – SNP -509C>T. 1 – Russia, control group in the present study, 2 – Italy [6], 3 – Germany [5], 4 – UK [14], 5 – ECTIM study (France + Northern Ireland) [1]; (B) – SNP 869T>C. 1 – Russia, control group in the present study, 2 – Russia, men [17], 3 – Italy [6], 4 – Germany [5], 5 – England [14], 6 – ECTIM study (France + Northern Ireland) [1]; (C) – SNP 915G>C. 1 – Russia. Control group in the present study, 2 – Germany [5], 3 – UK [14], 4 – ECTIM study (France + Northern Ireland) [1]. * - $p = 0.001$.

MI in Italians [6] and ECTIM study, which was performed on the French and the Northern Ireland population [1], are similar to our findings on the predisposing role of the alleles *TGFB1**-509T and *TGFB1**915G, respectively. Our results on the positive association of the allele *TGFB1**869T coincide with the findings for

Germans [5], but are in contradiction with the data for Italians [6]. A number of publications have reported no significant associations of these polymorphisms with MI in Caucasians.

Cytokine TGF- β 1, which is secreted by different cell types, including blood mononuclear cells, vascular smooth muscle cells and fibroblasts, participates in the formation and remodeling of vessels, as well as in cell differentiation and migration [19]. It plays an important role in the pathogenesis of CVDs, including atherosclerosis (including IHD and MI), essential hypertension, myocardial hypertrophy and fibrotic events, leading to heart failure and restenosis after heart surgery [20].

Some studies have found that TGF- β 1 has an anti-atherogenic effect: it inhibits inflammation and promotes atherosclerotic plaque stabilization. On the other hand, the high level of the TGF- β 1 associated with the stenosis of blood vessels and thrombosis [20] promotes fibrosis and inhibits endothelial regeneration [21]; i.e., it acts as a proatherogenic factor. In particular, it may contribute to early lipid spots by stimulating the formation of an extracellular matrix and inhibiting its

degradation [20]. On the basis of data on the role of TGF- β 1 in atherosclerosis pathogenesis, both low and high TGF- β 1 level can be unfavorable for the development of MI, depending on the combination of other factors. One such factor may be the ethnic identity of the groups studied.

CONCLUSIONS

Our data on the association of MI with alleles/genotypes of SNP *TGFB1**-509T, *TGFB1**869T/T and *TGFB1**915G/G, which are associated with the higher level of gene expression [2-4], may be indicative of the dominance of the proatherogenic functions of this cytokine in MI in Russians.

An aggregation of the results indicates the important role of the *TGFB1* gene in the formation of susceptibility to MI in Russians. It once again shows the necessity of studying the genetic factors in each ethnic group individually. ●

The study was supported by grant 8/3-280n-10 of the Moscow Government.

REFERENCES

- Cambien F., Ricard S., Troesch A., Mallet C., Genereaux L., Evans A., Arveiler D., Luc G., Ruidavets J.B., Poirier O. // *Hypertension*. 1996. V. 28. № 5. P. 881-887.
- Shah R., Hurley C.K., Posch P.E. // *Hum. Genet*. 2006. V. 120. P. 461-469.
- Nikolova P.N., Ivanova M.I., Mihailova S.M., Myhailova A.P., Baltadjieva D.N., Simeonov P.L., Paskalev E.K., Naumova E.J. // *Transpl. Immunol*. 2008. V. 18. № 4. P. 344-348.
- Khalil M.S., El Nahas A.M., Blakemore A.I. // *Nephron Exp. Nephrol*. 2005. V. 101. № 2. P. e31-41.
- Koch W., Hoppmann P., Mueller J.C., Schomig A., Kastrati A. // *Arterioscler. Thromb. Vasc. Biol*. 2006. V. 26. № 5. P. 1114-1119.
- Crobu F., Palumbo L., Franco E., Bergerone S., Carturan S., Guarrera S., Frea S., Trevi G., Piazza A., Matullo G. // *BMC Med. Genet*. 2008. V. 9. № 13.
- Sudomoina M.A., Sukhinina T.S., Barsova R.M., Favorov A.V., Shakhnovich R.M., Titov B.V., Matveeva N.A., Rybalkin I.N., Vlasik T.N., Ochs M.F. et al // *Mol Biol (Mosk)*. 2010. V. 44. № 3. P. 463-471.
- Sambrook J., Fritsch E.F., Maniatis T. // *Molecular Cloning / Ed. Nolan C. Cold Spring Harbor, N.Y.; Cold Spring Harbor Lab. Press, 1989.*
- www.changbioscience.com/primo/
- <http://www.broad.mit.edu/mpg/haploview>
- <http://www.graphpad.com/quickcalcs/index.cfm>
- <http://code.google.com/p/apsampler/>
- Favorov A.V., Andreewski T.V., Sudomoina M.A., Favorova O.O., Parmigiani G., Ochs M.F. // *Genetics*. 2005. V. 171. P. 2113-2121.
- Syrris P., Carter N.D., Metcalfe J.C., Kemp P.R., Grainger D.J., Kaski J.C., Crossman D.C., Francis S.E., Gunn J., Jeffery S., et al. // *Clin. Sci. (London)*. 1998. V. 95. № 6. P. 659-667.
- Grainger D.J., Heathcote K., Chiano M., Snieder H., Kemp P.R., Metcalfe J.C., Carter N.D., Spector T.D. // *Hum. Mol. Genet*. 1999. V. 8. № 1. P. 93-97.
- Sie M.P., Uitterlinden A.G., Bos M.J., Arp P.P., Breteler M.M., Koudstaal P.J., Pols H.A., Hofman A., van Duijn C.M., Witteman J.C. // *Stroke*. 2006. V. 37. № 11. P. 2667-2671.
- Churnosov M.I., Nekipelova E.V., Tekunova T.S., Koneva O.A., Reshetnikov E.A., Akulova L.Iu., Dobrodomova I.S., Altukhova O.B., Demin S.S. // *UNID BelGU*. 2008. V. 46. № 6.
- de Bakker P.I., Yelensky R., Pe'er I., Gabriel S.B., Daly M.J., Altshuler D. // *Nat Genet*. 2005. V. 37. № 11. P. 1217-1223.
- Khan R., Agrotis A., Bobik A. // *Cardiovasc. Res*. 2007. V. 74. № 2. P. 223-234.
- Aihara K., Ikeda Y., Yagi S., Akaike M., Matsumoto T. // *Cardiol. Res. Pract*. 2011. V. 2011. P. 175381.
- Kim I.Y., Kim M.M., Kim S.J. // *J. Biochem. Mol. Biol*. 2005. V. 38. № 1. P. 1-8.

Construction of a Full-Atomic Mechanistic Model of Human Apurinic/Apyrimidinic Endonuclease APE1 for Virtual Screening of Novel Inhibitors

I.G. Khaliullin¹, D.K. Nilov^{1,2}, I.V. Shapovalova², V.K. Švedas^{1,2*}

¹Belozersky Institute of Physicochemical Biology, Lomonosov Moscow State University, Leninskie gory, 1, str. 73, Moscow, 119991, Russia

²Lomonosov Moscow State University, Faculty of Bioengineering and Bioinformatics, Leninskie gory, 1, str. 73, Moscow, 119991, Russia

*E-mail: vytaš@belozersky.msu.ru

Received 18.04.2012

Copyright © 2012 Park-media, Ltd. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT A full-atomic molecular model of human apurinic/apyrimidinic endonuclease APE1, an important enzyme in the DNA repair system, has been constructed. The research consisted of hybrid quantum mechanics/molecular mechanics modeling of the enzyme-substrate interactions, as well as calculations of the ionization states of the amino acid residues of the active site of the enzyme. The choice of the APE1 mechanism with an Asp210 residue as a proton acceptor was validated by means of a generalization of modeling and experimental data. Interactions were revealed in the active site that are of greatest significance for binding the substrate and potential APE1 inhibitors (potential co-drugs of interest in the chemo- and radiotherapy of oncological diseases).

KEYWORDS apurinic/apyrimidinic endonuclease; QM/MM; enzymatic mechanism; molecular modeling; inhibition.

ABBREVIATIONS BER – base excision repair; AP – apurinic/apyrimidinic; APE1 – human apurinic/apyrimidinic endonuclease 1; QM/MM – quantum mechanics/molecular mechanics; MD – molecular dynamics.

INTRODUCTION

DNA damages occur frequently as a result of replication errors or on exposure to various exo- and endogenous factors, such as ultraviolet radiation and oxidative stress. In order to ensure genomic stability, mammals possess enzymatic repair systems – direct repair, base/nucleotide excision repair, and recombination mechanisms – which facilitate the elimination of most forms of DNA damage [1–3]. The pharmacological inhibition of the repair systems is a promising method for improving the efficacy of oncological therapy. The reason which accounts for this fact is that the repair systems resist the effect of the chemotherapeutic agents (e.g., temozolomide and cisplatin [4]) which damage the DNA in order to kill the tumor cell. Therefore, the selective inhibition of the enzymes participating in the DNA repair processes can be used as an accompanying therapy. The agents whose binding in the enzyme active site affects the residues directly participating in the catalytic mechanism are reasonably expected to exhibit the highest levels of efficacy. Therefore, adequate data regarding the organization of the active site of the target

enzyme, the charge distribution, and analysis of the interactions determining the strength of substrate-inhibitor binding are fundamentally necessary in the search for novel potential pharmaceuticals in the field of cancer therapy.

Apurinic/apyrimidinic endonuclease 1 (APE1) is the key enzyme in the DNA repair system, known as “base excision repair” (BER). Apurinic/apyrimidinic (AP) sites are the deoxyribose residues in the DNA molecule without a nitrogenous base; they result from the enzymatic hydrolysis of the N-glycoside bond of a damaged nucleotide and actually are the intermediates of the BER process. Furthermore, the AP sites may spontaneously emerge in cells due to apurinization [5]. According to current estimations, up to 10,000 AP sites are formed in mammalian cells per day [6]. The APE1 endonuclease recognizes the AP sites and hydrolyzes their 5'-phosphodiester bond for subsequent replacement with an undamaged nucleotide [5, 7]. The data obtained in laboratory and clinical studies attest to the significant role played by this enzyme in the development of a tumor and in the appearance of tumour resistance to antitumor agents [8].

There exists a wide variety of viewpoints concerning the catalytic mechanism underlying the action of endonuclease APE1. The first crystal structure of the enzyme was obtained in 1997 (PDB ID 1bix) [9]. In their description, the authors proposed a mechanism in which the role of the general base in the catalysis was attributed to the His309 residue. In this hypothetical mechanism, the uncharged residue His309, jointly with Asp283, forms a proton relay system similar to the one formed by serine proteases. The only difference is that a water molecule acts as an activated nucleophilic agent (*Fig. 1A*). The role of a metal ion in this mechanism consists in the binding and polarization of the negatively charged phosphate group of the substrate and in the stabilization of the transition state of the enzymatic reaction.

The fundamental significance of another residue in the active site (Asp210) for the catalysis was demonstrated in studies performed using site-directed mutagenesis: mutant forms of the enzyme with the substitutions Asp210Ala and Asp210Asn almost completely lost their catalytic properties (more than 25,000-fold reduction in activity was observed) as compared to the wild-type enzyme [10]. The determination of the crystal structures of human APE1 in complex with DNA derivatives resulted in a major revision of the assumptions regarding the mechanism of action of the enzyme [11]. One of the ascertained structures (PDB ID 1de8) is a complex of an inactive enzyme containing no metal ions with a substrate analogue, whereas the second structure (PDB ID 1de9) contains a metal (bivalent manganese) ion and the enzyme-bound DNA analogue of the substrate after catalytic cleavage. The conception was made regarding the structure of the enzyme-substrate complex, which simultaneously contains both the substrate analogue and a metal ion via the combina-

tion (spatial superposition) of the structures. Although the resulting model structure of the enzyme-substrate complex does not contain water molecules potentially capable of attacking the substrate, the arrangement of the residues in the active site before and after the catalytic process allowed making assumptions regarding the alternative mechanism of the catalytic reaction [11]. In the scheme proposed, the Asp210 residue acts as a general base activating the water molecule, whereas the His309 residue, along with the metal ion, participates in the binding and coordination to the phosphate group of the substrate (*Fig. 1B*). It is assumed that the positive charge of the His309 residue participates in the catalytic process, which is presumably facilitated by the proximate location of the Asp238 residue. The authors [11] attribute the major stabilizing function in the formation of the transition state of the enzymatic reaction to the Asn212 residue.

Hypotheses postulating that a secondary metal binding area exists in the enzyme active site have been put forward in subsequent crystallographic [12] and molecular dynamics (MD) [13] studies. The “two metal ions” mechanism of action of APE1 [12] (similar to that revealed in a related enzyme, endonuclease Endo IV) and the “moving metal” mechanism [13], involving the moving of the magnesium ion between two binding sites during the catalytic process, were proposed in these works.

It should be noted however that the NMR study utilizing the ^{25}Mg isotope [14] did not confirm the hypothesis of secondary magnesium ion binding in the active site of endonuclease APE1, thereby casting doubt on the “two metal ions” and the “moving metal” mechanisms. The authors of study [14] attribute the results of crystallographic studies [12] to the artefacts caused by

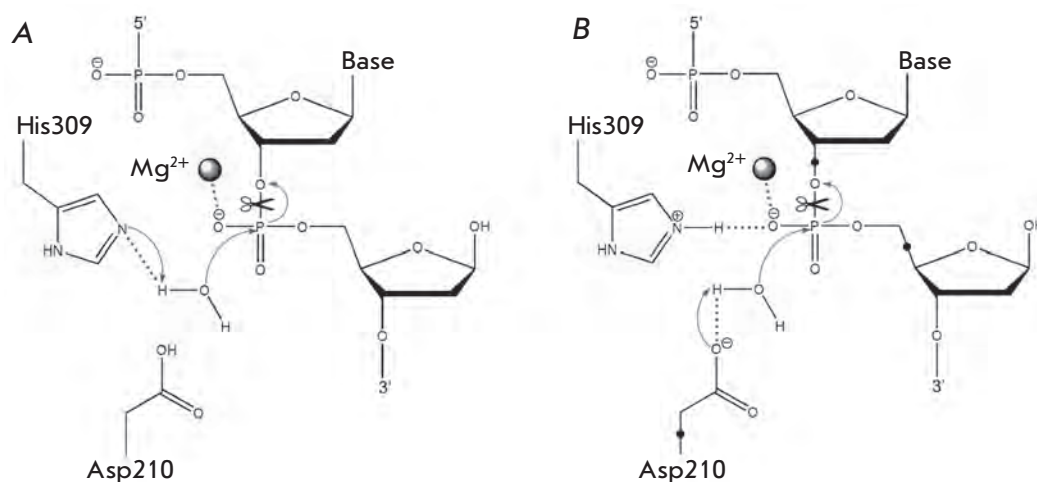


Fig. 1. Basic concepts of APE1 catalytic mechanism found in literature: (A) the His309 residue activates a water molecule acting as a base; (B) the Asp210 residue activates a water molecule acting as a base, while the His309 residue participates in substrate binding. The link atoms used in the present work on QM/MM modeling of APE1 are depicted as •.

the use of the lead ion instead of the magnesium ion. In turn, the effect of the motion of the metal ion during the MD modeling can be caused by the inaccuracy and approximations of the classical MD method.

Mundle et al. [15] proposed a two-step variant of the APE1 mechanism involving the Tyr171 residue acting in the form of the phenolate ion for a direct nucleophilic attack on the phosphate group of the substrate. This conclusion was drawn on the basis of the data obtained via site-directed mutagenesis on the 171st position. The kinetic studies of the catalytic properties of the mutated forms Tyr171Ala, Tyr171Phe, and Tyr171His demonstrated a fall in enzymatic activity by almost five orders of magnitude. It should be noted however that the authors [16] subsequently admitted the inconsistency of the previously proposed two-step scheme and supported the one-step mechanism, in which the His309 residue acts as the general base activating the water molecule, while the Tyr171 residue participates in the binding and proper orientation of the substrate.

The molecular modeling methods could assist considerably in the study of the mechanism of action of the enzyme; however, this approach has not been extensively explored in the study of APE1. Modeling of the inhibitor–enzyme interaction was performed with no allowance for the ionization state of the inhibitor, thereby making the interpretation of the obtained results more difficult [17].

Thus, there exists no unambiguous conception regarding the mechanism of action of human endonuclease APE1 or the role of the amino acids of the active site in binding and catalysis. Therefore, shedding more light on the structure of the active site, the nature of the interactions between the enzyme and substrate or inhibitors and the involvement of the active site residues into the catalytic mechanism of APE1 appears to be a topical task in molecular modeling.

MATERIALS AND METHODS

Molecular modeling software

The ionization states of the amino acid residues were calculated using the PROPKA 2.0 software [18, 19]. The preparation of the initial structure for simulations and the trajectory analysis were performed using the AmberTools 1.2 package (<http://ambermd.org>). The energy minimization and MD simulations were performed using the Amber 10 package [20, 21]; the molecular docking was performed using the Lead Finder 1.1.14 program (MolTech Ltd., Russia) [22]. Modeling of the spatial structure of 6-hydroxy-DOPA was performed using the ACD/ChemSketch 8.17 program [23]. The visualization of the structures and trajectories was carried out using the VMD 1.8.6 software [24]. Paral-

lel computations of the molecular dynamic trajectories were run on the SKIF-MSU “Chebyshev” supercomputer (MSU Research Computing Center).

Structure preparation

The initial model of the APE1 enzyme–substrate complex was built on the basis of the 1de8 crystallographic structure [11]. The coordinates of the attacking water molecule were calculated by docking. The coordinates of the manganese ion were transferred from the 1de9 structure; the metal type was subsequently substituted for magnesium. The structure of the enzyme–substrate complex was protonated and placed into a box of TIP3P type water molecules with the shortest distance of 12 Å between the box edge and protein. Sodium ions were added to neutralize the charge of the system. The *ff99SB* force field [25] was used to describe the protein and DNA molecules; parameters from the R.E.D.D.B database (<http://q4md-forcefieldtools.org>) [26] were used to describe the AP site.

Energy minimization and molecular dynamics

The model of the APE1 enzyme–substrate complex was equilibrated and subjected to calculation of the 1000 ps MD trajectory according to the following protocol. Initially, the two-stage energy minimization of the solvated system was performed. At the first stage (2,500 steps of the steepest descent algorithm followed by 2,500 steps of the conjugate gradient algorithm), the molecular mechanics description of the system was performed with the coordinates of the protein, DNA, and magnesium ion being fixed by position restraints $k(\Delta x)^2$ with a force constant of 2 kcal/(mol·Å²). At the second stage (5,000 steepest-descent steps followed by 5,000 conjugate-gradient steps), the system was divided into a quantum mechanical (QM) region and a molecular mechanical (MM) region; energy minimization was performed without any restraints. The QM region included the side chain of the Asp210 residue of the active site, the attacking water molecule, and the AP site fragment; this region was described by the RM1 semi-empirical Hamiltonian [27]. The link atom model was used to make allowance for the bonds crossing the QM–MM boundaries.

After energy minimization, using the aforementioned QM/MM division the system was heated from 0 to 300 K over 50 ps (with positional restraints of 1 kcal/(mol·Å²) on the protein, DNA, and magnesium ion), equilibrated over 500 ps at 300 K, and finally simulated for 1,000 ps. All simulations were performed using periodic boundaries and the PME (Particle Mesh Ewald) method to allow for long-range electrostatic interactions. The cut-off radius of the non-bonded interactions was 10 Å. The system was heated at a constant

volume; the equilibration and 1,000 ps trajectory simulation were performed under constant pressure. The temperature was controlled by the Langevin method. The integration time step was 0.002 ps. Interatomic distances and angles in the active site of APE1 were estimated by analyzing the 1,000 ps trajectory of equilibrium simulation.

Molecular docking

The model for performing the molecular docking procedure was obtained as follows. Water molecules, sodium ions, and the DNA substrate analogue molecule were removed from the structure of the solvated APE1 enzyme–substrate complex after energy minimization. The energy grid map surrounding the AP-site area was then calculated. Finally, a potential inhibitor molecule, 6-hydroxy-DOPA, was docked into the active site with the use of the genetic search algorithm implemented in the docking program.

RESULTS AND DISCUSSION

Ionization states of the active site residues

Based on the results of the calculation of the ionization states of the active site residues using the PROPKA 2.0 method, it was determined that the His309 residue is protonated (the calculated pK_a value of 8.6 matches the pK_2 value of the experimentally determined pH profile of enzyme activity) under optimal conditions of the hydrolysis of the phosphodiester bond (pH 7–8 [12]), whereas the Asp210 residue is deprotonated (the calculated pK_a value of 6.2 is close to the pK_1 value of the pH profile of enzyme activity). Thus, it can be said that the deprotonated and negatively charged Asp210 residue acts as a general base in catalysis, whereas the positively charged His309 residue participates in the binding of the negatively charged phosphate group of the substrate and in stabilization of the reaction intermediate product. Therefore, when building the full-atomic model of APE1, Asp210 and His309 were modelled in their charged forms.

Model of the enzyme–substrate complex and the deduced catalytic mechanism of action of APE1 endonuclease

The starting solvated model of the enzyme–substrate complex APE1 was created on the basis of the 1de8 and 1de9 crystallographic structures as described in Materials and Methods. It was then necessary to optimize the atom positions within the model (especially the coordinates of the hydrogen atoms added); therefore, a two-stage minimization of the energy of the system was performed. Molecular-mechanical minimization was performed at the first stage to remove the larg-

est strains in the system. Refinement of the active site structure was carried out at the second stage using the hybrid QM/MM method for energy minimization using the RM1 Hamiltonian. The stability of the resulting structure was confirmed by calculation of the 1,000 ps QM/MM MD trajectory. The calculated interatomic distances in the APE1 active site are listed in Table. It was demonstrated via the analysis of the resulting model that substrate binding in the active site of APE1 is accompanied by the formation of a number of bonds and interactions of a different nature. Among these, the hydrophobic interactions of deoxyribose of the AP site in the hydrophobic pocket formed by the Leu282, Phe266, and Trp280 residues should be noted. The free hydroxyl group in deoxyribose of the AP site also forms a hydrogen bond with the backbone carbonyl group of the Ala230 residue. The phosphate group at the 3' terminus of the AP site is held by the positive charge of the Arg177 residue. The phosphate group under attack electrostatically interacts with the magnesium ion and forms hydrogen bonds with the side chains of Asn174, Asn212, and His309. The hydroxyl group of the Tyr171 residue is oriented towards the oxygen atom of the leaving group.

The orientation of the attacking water molecule is provided via the interaction with the general base Asp210, carbonyl group of Asn212, and the side chain of His309; however, the interactions $H_2O:H_2 \cdots Asn212:OD1$ and $H_2O:O \cdots His309:HE2$ are not conventional hydrogen bonds, since the average $O \cdots H_2 \cdots OD1$ and $NE2 \cdots HE2 \cdots O$ angles are 137° and 122° , respectively. In the case of a hydrogen bond, these values should be at least 150° . The reactive conformation of the carboxyl group of the gen-

Table. Distance characteristics of the APE1 enzyme–substrate complex obtained via the equilibrium 1,000 ps QM/MM MD simulation. Mean values are presented together with the standard deviation

Interaction	Distance, Å
$H_2O:O \cdots AP \text{ site}:P$	1.91 ± 0.03
$H_2O:O \cdots His309:HE2$	2.52 ± 0.17
$H_2O:H_1 \cdots Asp210:OD1$	1.49 ± 0.07
$H_2O:H_2 \cdots Asn212:OD1$	2.62 ± 0.47
AP site:O1P \cdots Mg ²⁺	1.84 ± 0.04
AP site:O1P \cdots His309:HE2	1.78 ± 0.09
AP site:O2P \cdots Asn212:HD2	2.08 ± 0.32
dC5:O3' \cdots Mg ²⁺	1.95 ± 0.06
Asp210:OD2 \cdots Asn212:H	2.22 ± 0.20

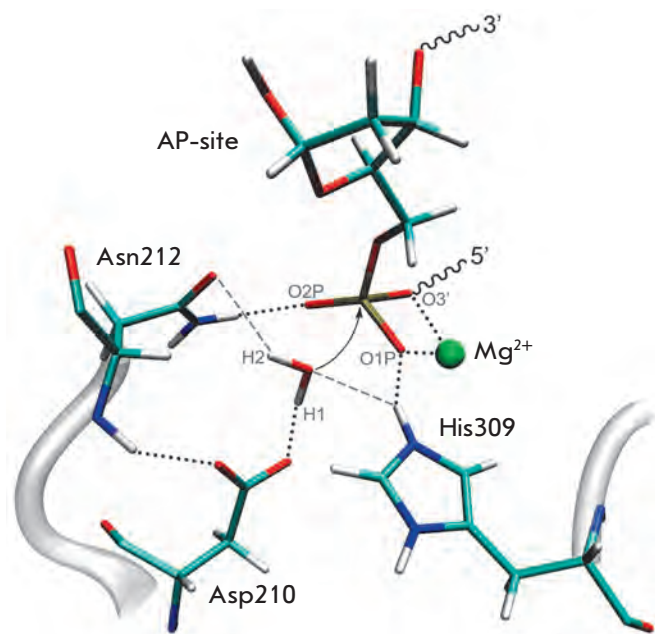


Fig. 2. The active site of the full-atomic model of APE1 enzyme–substrate complex. Hydrogen bonds are shown as dotted lines. Electrostatic interactions other than hydrogen bonds are shown as dashed lines. The arrow shows the direction of nucleophilic attack.

eral base Asp210 in the enzyme is maintained via the interaction between its side chain and the backbone amino group of Asn212 (Fig. 2).

Oriented and polarized by Asp210, His309, and the metal ion, the water molecule can attack the phosphate group of the substrate with the simultaneous transmission of the proton to the general base, the Asp210 residue. The intermediate resulting from the attack is shaped as a trigonal bipyramid and stabilized via the following interactions in the enzyme active site: oxygen atoms in the “apexes” of the bipyramid interact with the side chains of the His309 and Tyr171 residues; the trigonal base of the bipyramid is placed between Asn174, Asn212, and the magnesium ion (Fig. 3).

The data pertaining to the mutagenesis on Tyr171 [15] attest to the crucial role of this residue in the catalytic mechanism of APE1; however, in contrast to the previous suggestions [15, 16], we assume a different role for this residue in the catalysis. The proximate location of the positive charge of magnesium ion and the Arg156 residue are supposed to facilitate proton migration from the hydroxyl group of Tyr171. Therefore, we consider this residue to be a potential proton donor for the leaving group, which is a strong base. The weaker influence of the mutations at position 171 on substrate binding in comparison with the decrease in the catalytic constant [15] is attributed to the weak interaction between the

residue and the substrate at the earlier stages of the reaction preceding the catalytic process, which fully matches the assumption made. In the catalytic transformation, the leaving group apparently approaches the chain of the Tyr171 residue, enabling proton transfer.

As the reaction proceeds, the less stabilized P–O bond located at the base of the bipyramid and directed towards the Asn212 residue is subsequently transformed into a double P=O bond. The P–O bond directed towards the Tyr171 residue is simultaneously broken; as a result, the leaving group takes the proton away from the hydroxyl group of tyrosine (Fig. 3).

Restoration of the catalytically active site (deprotonation of the general base Asp210 and protonation of the acid Tyr171) occurs via the interaction with surrounding water molecules.

Analysis of the binding of substrate and mechanism-dependent inhibitors

A large number of charges and polar groups participate in the substrate binding and intermediate stabilization during the reaction. The character of these interactions and the ionization states of the amino acid residues of the active site define certain requirements to the structure of the substances capable of binding to the active site of APE1. It is necessary to implement the most significant interactions when constructing efficient inhibitors of the enzyme. The existence of the hydrophobic

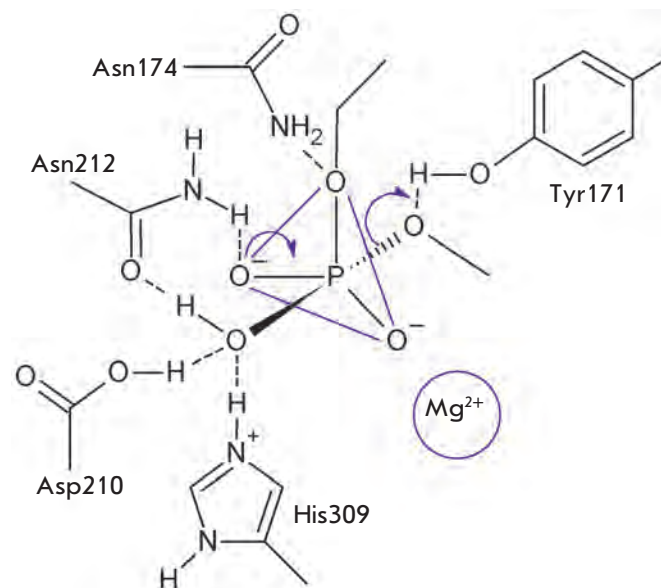


Fig. 3. The schematic picture of the hypothetical structure of the transition state (intermediate shaped as a trigonal bipyramid) and its conversion upon a hydrolysis reaction catalyzed by endonuclease APE1.

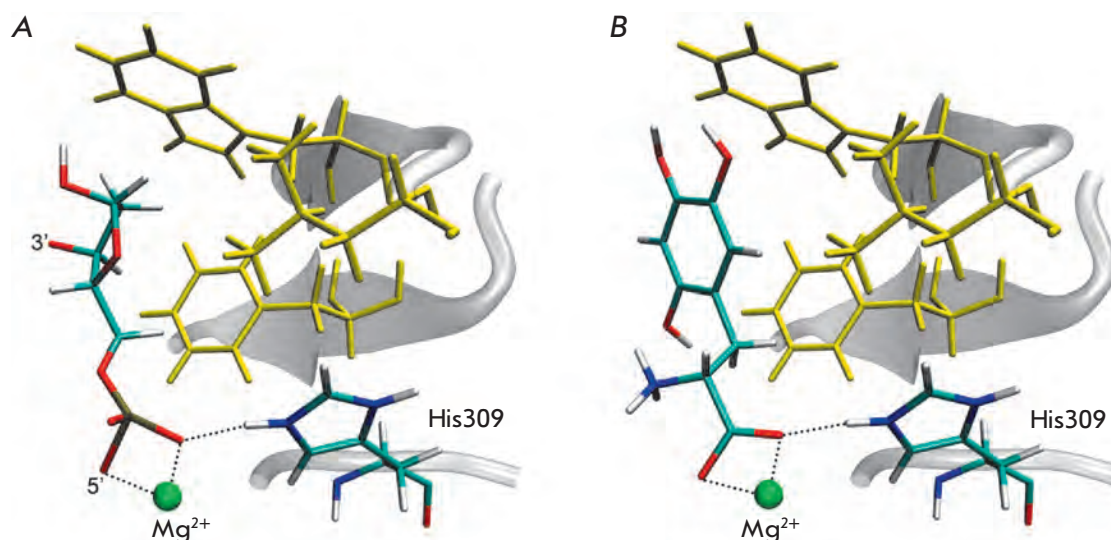


Fig. 4. Positions of the substrate (A) and the potential inhibitor (B) in the active site of the full-atomic APE1 model. Residues of the hydrophobic pocket are shown in yellow: Phe266, Trp280, Leu282.

binding site, along with the number of polar groups and groups with various charges, further complicates the search for low-molecular-weight compounds of appropriate structure. Amino acids are a class of natural compounds the structure of which simultaneously contains substituents of a different nature, capable of both performing the hydrophobic and electrostatic interactions and of acting as hydrogen bond donor or acceptor. It has been noted [17] that 6-hydroxy-DOPA is a potential inhibitor of APE1 endonuclease activity; however, an incorrect allowance was made for the ionization state of the inhibitor when modeling the enzyme–inhibitor interaction, making it more difficult to interpret the results. To clarify enzyme interactions with the potential inhibitors of this structure, molecular docking of different amino acids and their derivatives into the APE1 active site was carried out, taking into account the ionization states of the inhibitors and active site residues.

It was demonstrated via the analysis of the molecular docking results that the presence of the carboxyl group allows the selected compounds to bind to the metal ion and the His309 residue, whereas the hydrophobic substituent (e.g., phenyl radical) can occupy the hydrophobic pocket (Fig. 4). The introduction of hydroxyl substituents to the phenyl radical may lead to the formation of additional hydrogen bonds with the polar residues of the enzyme active site. Furthermore, the binding of the inhibitor to the charged residue of the general base is one of the factors presumably determining the inhibition efficiency.

CONCLUSIONS

The aim of this work was to select the most reliable mechanism of action of APE1 on the basis of a molecular modeling, analysis of the full-atomic model of APE1, and a critique of the experimental results and

assumptions previously made in the literature. For this purpose, calculation of the ionization states of the active site residues and a hybrid QM/MM modeling of the enzyme–substrate complex containing a water molecule capable of attacking the substrate were carried out. As a result of the investigation conducted, it was demonstrated that the Asp210 is likely to act as the general base in the catalytic mechanism, whereas the His309 residue, being protonated (and positively charged), participates in the binding of the phosphate group of the substrate. The analysis of the molecular dynamic trajectory of the enzyme–substrate complex attested to its high reactivity and confirmed the validity of the molecular modeling performed.

The most important interactions in the active site determining the efficiency of binding of the substrate and the potential enzyme inhibitors (which are promising co-drugs of interest in the chemo- and radiotherapy of oncological diseases) were revealed. An assumption regarding the role of the Tyr171 of the active site of APE1 as the residue capable of ceding the proton to the leaving group of the substrate was made. Thus, the investigation enabled to establish a consistent mechanism of action of the enzyme. Furthermore, it allowed to summarize MD data, as well as the experimental results of kinetic studies and the other published data. In the next step, we plan to use higher level QM/MM methods to calculate the energy barrier of the reaction catalyzed by endonuclease APE1, in compliance with the mechanism proposed, and to screen for effective inhibitors with the use of the constructed mechanistic full-atomic model of the enzyme. ●

This work was supported by the Ministry of Education and Science of the Russian Federation (Government contract № 16.512.11.2240).

REFERENCES

- 1 . Lindahl T., Wood R.D. // *Science*. 1999. V. 286. P. 1897–1905.
- 2 . Hoeijmakers J.H. // *Nature*. 2011. V. 411. P. 366–374.
- 3 . Rechkunova N.I., Krasikova Y.S., Lavrik O.I. // *Biochemistry (Moscow)*. 2011. V. 76. P. 24 - 35.
- 4 . Abbotts R., Madhusudan S. // *Cancer Treatment Rev.* 2010. V. 36. P. 425–435.
- 5 . Nevinsky G.A. // *Biochemistry (Moscow)*. 2011. V. 76. P. 94 - 117.
- 6 . Wilson D.M., Simeonov A. // *CMLS*. 2010. V. 67. P. 3621–3631.
- 7 . Dyrkheeva N.S., Khodyreva S.N., Lavrik O.I. // *Molecular biology (Moscow)*. 2007. V. 41. P. 402 - 416.
- 8 . Fishel M.L., Kelley M.R. // *Mol. Aspects Medicine*. 2007. V. 28. P. 375–395.
- 9 . Gorman M.A., Morera S., Rothwell D.G., de La Fortelle E., Mol C.D., Tainer J.A., Hickson I.D., Freemont P.S. // *EMBO J.* 1997. V. 16. P. 6548–6558.
- 10 . Erzberger J.P., Wilson D.M. // *J. Mol. Biol.* 1999. V. 290. P. 447–457.
- 11 . Mol C.D., Izumi T., Mitra S., Tainer J.A. // *Nature*. 2000. V. 403. P. 451–456.
- 12 . Beernink P.T., Segelke B.W., Hadi M.Z., Erzberger J.P., Wilson D.M., Rupp B. // *J. Mol. Biol.* 2001. V. 307. P. 1023–1034.
- 13 . Oezguen N., Schein C.H., Peddi S.R., Power T.D., Izumi T., Braun W. // *Proteins: Structure, Function and Bioinformatics*. 2007. V. 68. P. 313–323.
- 14 . Lipton A.S., Heck R.W., Primak S., McNeill D.R., Wilson D.M., Ellis P.D. // *J. Amer. Chem. Soc.* 2008. V. 130. P. 9332–9341.
- 15 . Mundle S.T., Fattal M.H., Melo L.F., Coriolan J.D., O'Regan N.E., Strauss P.R. // *DNA Repair*. 2004. V. 3. P. 1447–1455.
- 16 . Mundle S.T., Delaney J.C., Essigmann J.M., Strauss P.R. // *Biochemistry*. 2009. V. 48. P. 19–26.
- 17 . Simeonov A., Kulkarni A., Dorjsuren D., Jadhav A., Shen M., McNeill D.R., Austin C.P., Wilson D.M. III // *PloS One*. 2009. V. 4. P. e5740.
- 18 . Li H., Robertson A.D., Jensen J.H. // *Proteins*. 2005. V. 61. P. 704–721.
- 19 . Bas D.C., Rogers D.M., Jensen J.H. // *Proteins*. 2008. V. 73. P. 765–783.
- 20 . Case D.A., Darden T.A., Cheatham T.E. III, Simmerling C.L., Wang J., Duke R.E., Luo R., Crowley M., Walker R.C., Zhang W., et al. // *AMBER 10*. University of California. San Francisco. 2008.
- 21 . Walker R.C., Crowley M.F., Case D.A. // *J. Comput. Chem.* 2008. V. 29. P. 1019–1031.
- 22 . Stroganov O.V., Novikov F.N., Stroylov V.S., Kulkov V., Chilov G.G. // *J. Chem. Inf. Model.* 2008. V. 48. P. 2371–2385.
- 23 . *ACD/ChemSketch Freeware*, version 8.17. Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com. 2005.
- 24 . Humphrey W., Dalke A., Schulten K. // *J. Mol. Graphics*. 1996. V. 14. № 1. P. 33–38.
- 25 . Hornak V., Abel R., Okur A., Strockbine B., Roitberg A., Simmerling C. // *Proteins*. 2006. V. 65. P. 712–725.
- 26 . Chen J., Dupradeau F.-Y., Case D.A., Turner C.J., Stubbe J. // *Biochemistry*. 2007. V. 46. P. 3096–3107.
- 27 . Rocha G.B., Freire R.O., Simas A.M., Stewart J.J.P. // *J. Comp. Chem.* 2006. V. 27. P. 1101–1111.

Development of Chlamydial Type III Secretion System Inhibitors for Suppression of Acute and Chronic Forms of Chlamydial Infection

N.A. Zigangirova, E.S. Zayakin*, L.N. Kapotina, E.A. Kost, L.V. Didenko, D.Y. Davydova, J.P. Rummyanceva, A.L. Gintsburg

Gamaleya Research Institute of Epidemiology and Microbiology, ul. Gamalei, 18, Moscow, 123098, Russia

E-mail: e.s.zayakin@gmail.com

Received 17.01.2012

Copyright © 2012 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 Type III secretion system (T3SS) is currently considered to be one of the main pathogenicity factors in Gram-negative bacteria, which exhibit different types of parasitizing activity. The presence of this structure is essential for the development of an acute infection; the chronicity of the infection is fundamentally dependent upon its functioning. In this regard, T3SS is one of the most promising targets for the development of broad-spectrum antimicrobial drugs that do not develop resistance and are efficacious for the acute and chronic forms of infection. The mechanism of action in drug development is based on the specific inhibition of T3SS, which should interrupt the infectious process, thereby enabling the immune system to eliminate the pathogen. As a result of pilot screening using specific cellular and bacterial tests, followed by chemical optimization and detailed characterization of the biological activity, a new class of chlamydial T3SS inhibitors was obtained. The selected compounds have obvious advantages over the currently available inhibitors of T3SS pathogens thanks to the high inhibitory activity of these compounds with minimal damaging effects on eukaryotic cells. Preclinical trials of the selected inhibitors are currently under way.

KEYWORDS thiohydrazones; thiohydrazides; thiadiazines; type III secretion system; cytotoxicity; Chlamydia; inhibitors; microscopy; electron microscopy; morphology.

ABBREVIATIONS T3SS – type III secretion system; MOI – multiplicity of infection; MOMP – major outer membrane protein of Chlamydia; LPS – lipopolysaccharide; EB – elementary bodies; RB – reticular cells; IFU – inclusion forming units.

INTRODUCTION

Chlamydia are Gram-negative bacteria that parasitize intercellularly. Two species of Chlamydia are common pathogenic agents responsible for diseases in humans. The Chlamydia infection caused by *Chlamydia trachomatis* is the most prevalent among sexually transmitted diseases, causing over 100 million new cases of the disease annually [1]. According to the WHO, the number of people in the world infected with Chlamydia by the most conservative estimate has reached a billion; the number of infections is on a steady increase even in the developed world. The fraction of cases of respiratory chlamydiosis caused by *C. pneumoniae* in the total structure of pneumonias stands at 20%; epidemic outbursts of this infection occur in European countries every 4–7 years (according to WHO data). As a result, up to 80% of the world population are infected with respiratory chlamydiosis during their lives. Chronic chlamydioses pose the most serious problem; it is a proven fact that these diseases act as a

mechanism triggering severe chronic diseases, such as asthma, atherosclerosis, arthritis, female and male infertility, as well as pregnancy pathologies [2, 3].

The medical and socio-economic significance of searching for new-generation drugs using target-specific technologies is rooted in the absence of efficacious agents that can help treat chronic bacterial infections and the rapid development of pathogen resistance to the antibacterial agents used to treat acute infectious processes [4–6]. In the case of antibacterial drugs, this technology includes selecting the proteins responsible for the exhibition of pathogenic properties by the microorganism as targets; the subsequent search for specific inhibitors using computer software, organic synthesis techniques and experimental assaying; and verification of the predicted biological activity on model systems for the infectious process.

Secretion of pathogenic factors (the proteins responsible for the exhibition of pathogenic properties

by bacteria) into the macroorganism's cell is the key mechanism underlying the development of an infectious process. A total of seven secretion systems, characterized by various specificities with respect to the molecules secreted and differences in the structure of the secretory apparatus, have been described thus far. One of these systems (referred to as the type III secretion system (T3SS)) transfers protein pathogenic factors from the bacterial cell directly into the cytoplasm of the eukaryotic cell. This "molecular syringe" has been detected only in pathogenic bacteria, since it is through its functioning that the bacteria with various types of parasitizing actions, exo- and endoparasites, exhibit their pathogenic properties [7]. Because of the conservative nature of this structure, in the taxonomically distant microorganisms that are behind socially significant infections, such as *Chlamydia*, *Salmonella*, *Shigella*, *Pseudomonas*, *Escherichia*, *Yersinia*, *Brucella*, etc., it is reasonable to expect antibacterial drugs based on specific T3SS inhibitors to have a wide range of effects.

In intracellular pathogens (Chlamydia being a typical example of such organisms), the transport system makes it possible to use the regulatory pathways of a host eukaryotic cell and to subsequently suppress the cellular response. T3SS is required at each stage of Chlamydial life cycle; it provides the possibility of intracellular reproduction of the pathogen upon both acute and chronic forms of the infection. T3SS inhibition results in the suppression of the *in vitro* reproduction of Chlamydia [8].

Several T3SS inhibitors, low-molecular weight compounds of different classes, which were selected by high-efficiency screening of chemical compound databases, have been discovered thus far [9–13]. A substantial drawback of these compounds is their poor solubility in organic solvents and water. Moreover, these inhibitors exhibit significant toxicity towards mammalian cells, resulting in the death of up to 60% of the cells in the presence of a specific inhibitory concentration (50 μ M), a factor which restricts the development of these compounds for further application as antibacterial drugs.

This study was aimed at designing pharmacologically promising compounds that can suppress acute and chronic infections which inhibit the secretion of chlamydial pathogenic factors, but do not have the aforementioned drawbacks.

METHODS

Bacterial strains and cell lines

Reference strain of *C. trachomatis* BU-434 serovar L2 (ATCC VR 902B), *C. muridarum* strain Nigg(ATCC

VR-123), and *C. pneumoniae* strain K-6 kindly provided by P. Saikkii (Finland), and the McCoy B cell line (a hybrid cell line consisting of human synovial cells and mouse fibroblasts) were used in this study.

Assessment of the toxicity for eukaryotic cells

96- and 24-well plates and a one-day cell monolayer were used in the study. The cells were cultured for 24 h in the presence of various doses of inhibitors. The cytotoxic effect of the agents was assessed using three conventional procedures: methylene blue staining, the MTT assay ("Sigma"), and the calcein assay (LIVE/DEAD Viability/Cytotoxicity Kit for mammalian cells, Invitrogen, United States).

Cell infection with Chlamydia strains

The McCoy B cells were infected with Chlamydia with multiplicity of infection of 1 (MOI of 1) according to the conventional procedure [3].

Immunofluorescent detection of Chlamydia development

The intracellular chlamydial inclusions were detected via direct immunofluorescence (DIF) using monoclonal species-specific antibodies against the major outer membrane protein (MOMP) of *C. trachomatis* and fluorescein isothiocyanate (FITC) labelled genus-specific anti-chlamydial LPS antibodies (OOO Niarmedic Plus, Moscow).

Assessment of chlamydial progeny

Chlamydial progeny was assessed via a semi-quantitative analysis based on immunofluorescence. Lysates of the infected cells were seeded onto a new cell monolayer. For this purpose, the 48-hour monolayer of infected cells was removed by a sucrose phosphate glutamine buffer (SPG) and lysed by freezing. The necessary lysate dilutions were prepared and then seeded onto a new monolayer. The cells were incubated for 48 h, fixed, and stained with FITC-labelled monoclonal antibodies for subsequent assessment of the results via luminescent microscopy. The count of infected cells was determined in 10 random locations within the visual fields; the average number of inclusion-forming units (IFU) per 1 ml of the specimen was calculated (the results of three independent experiments were used).

Detection of the *C. trachomatis* effector protein IncA

A one-day monolayer of McCoy cells was infected with *C. trachomatis* with MOI of 5. The compounds under study were added at varying doses eight hours post infection (the onset of the effector protein translocation to the inclusion membrane). After 24 h of incubation the cells were stained with primary anti-IncA-antibod-

ies (Innovagen, Sweden) and secondary FITC-labelled antibodies. The cells were simultaneously stained with anti-*C. trachomatis* MOMP monoclonal antibodies.

Transmission electron microscopy (TEM)

The cells were cultured and infected in 6-well plates. The cell pellet obtained via centrifugation for 10 min at 1500 rpm (Rotanta 460R, Hettich) was fixed in aIto-Karnovsky fixative. Post-fixation with OsO₄ and contrasting in aqueous uranyl acetate were used for this purpose. The specimens were subsequently dehydrated in an ascending series of alcohols, infiltrated in a 1 : 1 mixture of a LR White resin and 100% ethanol for 1 h and in a pure resin for 12 h at +4°C. Resin polymerization was performed at +56°C for 24 h. Ultrathin sections were then prepared, then they were contrasted with a lead solution (Reynolds) and analyzed using TEM Jeol 100B.

RNA isolation and analysis of gene expression

RNA was isolated from the cell culture using the Trizol reagent (Invitrogen) 24 h post infection. The concentration of RNA pretreated with DNase I (DNA-free™, “Ambion”) was determined on a NanoDrop ND-100 spectrophotometer (ThermoFisher Scientific, United States). The reverse transcription (RT) reaction was carried out using the “Reverse Transcription System” kit (Promega, United States).

Real-time PCR with the resulting cDNA was carried out using primers to the following genes: *16S* rRNA (primer forward 5'-GGCGTATTTGGGCATCCGAGTAACG, primer reverse 5'-TCAAATC-CAGCGGGTATTAACCGCCT, Pb 5'-R6G-TGG CGG CCA ATC TCT CAA TCC GCC TAG A-BHQ2), *trpA* (primer forward 5'-CGG GAA TAA ATG GTG TGT GCG T, primer reverse 5'-TAAAGACATCCGTTCCG-GCGTT, Pb 5'-ROX-ATC TTC CAG CAC CTT TAT CAC ACG GAG A-BHQ2), *incA* (primer forward 5'-CTA CAG AAG AAA TGC GCA AAC TTT, primer reverse 5'-AAT GAT TGC TGG TTA TGC GCT AAT, Pb 5'-FAM-CGG CGA ACT TCT TCT GCT AAT GGG GTT-BHQ1), *lcrE* (primer forward 5'-GAG GCT GTG TTG AGG TAG GT, primer reverse 5'-CGA TAA ATG CGG ATA ATG AGG AT, Pb 5'-FAM-AGG TAC TGG AGC ATG AGG AGG CGT A-RTQ1). RT-PR was performed on a CFX 96 amplifier (Bio-Rad Lab., United States).

RESULTS

Analysis of the structural similarity of the known 3TSS inhibitors

Among the known 3TSS inhibitors [9–13], the compounds that belong to hydrazones based on hydrazides of aromatic carboxylic acids and various salicylic alde-

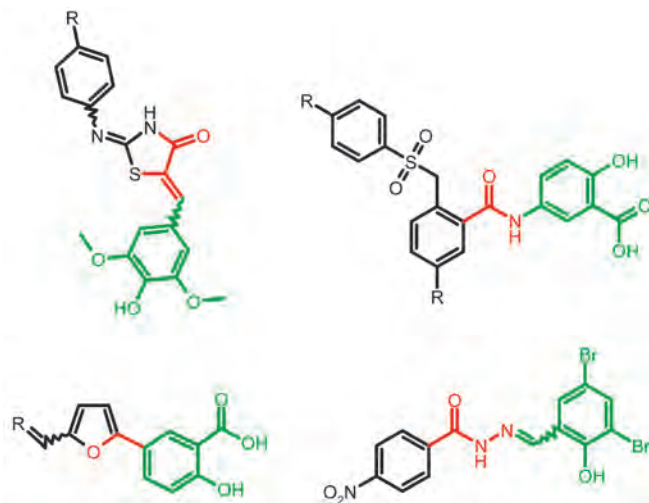


Fig 1. Structures of various classes of the known T3SS inhibitors.

hydes (IV) have been the best studied. The structures of T3SS inhibitors are shown in Fig. 1.

These molecules are similar, to a certain extent, since all of them contain residues of salicylic and 4-hydroxybenzoic acids and their derivatives (these residues are shown in green). Also worthy of note is the structural similarity of the concatenation between the derivatives of salicylic acids and the remaining molecular part (shown in red). Based on the concept of bioisosteric replacement [14, 15], according to which the carbonylic and thiocarbonylic groups in biological systems are functionally interchangeable, one can assume that the action of the thiohydrazones of oxamine acids and various salicylic aldehydes is similar to that of hydrazones. This class of compounds is relatively new; until recently, its biological properties were almost completely outside the realm of study.

Our goal was to resolve the following issues using thiohydrazones of thiohydrazides of oxamine acids as possible T3SS inhibitors:

- obtain compounds with a low toxicity with respect to eukaryotic cells;

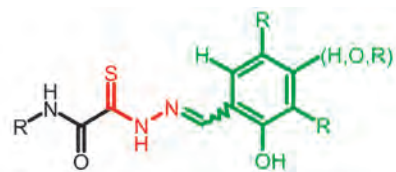


Fig. 2. The general formula of thiohydrazones of the thiohydrazides of oxamine acids.

Table 1. Toxicity indices of the selected compounds – thiohydrazones of thiohydrazides of oxamine acids

Compound №	Structure	Methylene blue staining, % of dead cells			Calcein assay, % of dead cells			MTT assay, % of metabolically inactive cells		
		12.5 μ M	25 μ M	50 μ M	12.5 μ M	25 μ M	50 μ M	12.5 μ M	25 μ M	50 μ M
1		0±2	12±2	23±2	0±2	11±2	20±2	3±1	12±3	29±3
2		2±2	13±1	24±3	3±2	12±1	21±3	3±2	15±1	31±5
3		6±4	20±2	27±2	5±4	16±2	23±2	2±2	15±5	30±4
4		5±1	15±2	28±2	6±1	12±2	23±2	3±3	13±4	28±2
5		2±1	14±2	26±3	3±1	16±2	22±3	4±3	20±4	29±6
6		2±1	15±2	26±2	5±1	12±2	24±2	2±4	17±5	27±8
7		2±2	15±4	29±5	4±2	16±4	22±5	4±5	18±3	29±9
8		2±1	14±1	24±4	6±1	12±1	24±4	3±3	15±3	28±4
9		2±1	12±1	24±4	4±1	15±1	25±4	4±5	20±4	30±4
10		2±1	12±2	24±2	5±1	14±2	22±2	3±3	19±3	31±7
11		1±1	12±1	21±4	4±1	13±1	25±4	5±1	16±2	30±5
12		0±1	11±1	22±3	3±1	10±1	24±3	5±5	15±4	29±6
13		2±1	15±2	26±3	2±1	11±2	25±3	5±4	14±3	29±7
14		4±1	15±1	25±4	3±1	12±1	22±4	3±5	17±3	28±6
15		0±1	11±2	22±3	0±1	10±2	21±3	4±1	13±3	29±5

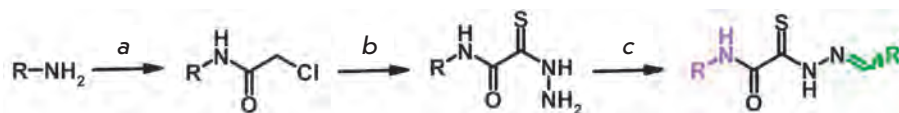
- obtain compounds with high selective activity with respect to T3SS;
- obtain compounds with good pharmacokinetic properties; and
- develop a simple scheme for the synthesis of the desired compounds.

Obtaining of thiohydrazones and testing their *in vitro* T3SS inhibition activity and toxicity

Thiohydrazones of oxamine acids (Fig. 2) were synthesized based on the appreciably simple scheme [16] shown in Fig. 3.

It can be seen that a large variety of compounds can be obtained via the use of different commercially

Fig. 3. Synthesis of thiohydrazones based on the thiohydrazides of oxamine acids: a – chloroacetyl chloride, DMF; b – 1) TEA, elemental sulphur, morpholine, DMF; 2) DMF, hydrazine hydrate; c – methanol, the corresponding aldehyde R'.



available amines and aldehydes via simple chemical conversions. A total of approximately 300 compounds have been synthesized; a number of compounds with good levels of solubility have been selected for further analysis. A total of 120 compounds were selected; despite having superior solubility as compared to that of hydrazones, thiohydrazones generally have poor levels of solubility.

The toxicity of the selected compounds for eukaryotic cells was assessed. First, the methylene blue staining technique was used. The compounds exhibiting satisfactory toxicity level were analyzed further using the calcein and MTT assays. As a result, 15 compounds were selected (Table 1). They manifested acceptable toxicity at concentration of 50 μM (death of less than 30% cells). All these compounds contained residues of various fluoroderivatives of aniline, as well as derivatives of salicylic and 4-hydroxybenzaldehyde.

Chlamydial T3SS inhibitors are known to suppress *in vitro* the intracellular reproduction of the pathogen. It was shown, via testing of the ability of the selected compounds to suppress the chlamydial infection in the cell cultures, that all the compounds exhibit good inhibitory activity (Table 2).

The ability of these compounds to suppress the translocation of *C. trachomatis* effector protein IncA was determined via immunofluorescence. It turned out that all the tested compounds inhibit the effector function of chlamydial T3SS.

Thus, new inhibitors of *C. trachomatis* T3SS that belong to the class of thiohydrazones of the thiohydrazides of oxamine acids were selected; all the compounds suppressed Chlamydia reproduction in cell cultures. The presence of a fluorine atom can be assumed to enhance the lipophilicity of the selected molecules, which ena-

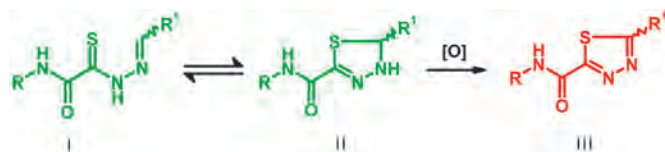


Fig. 4. Ring-chain tautomerism of thiohydrazones and the oxidation product, thiadiazole.

bles them to penetrate more easily through biological membranes, and presumably increases their resistance to various enzymes [17]. Moreover, the inclusion of the derivatives of salicylic and 4-hydroxybenzaldehydes has enabled to obtain compounds with improved solubility and activity.

Investigation of the stability and the nature of the toxicity of thiohydrazones

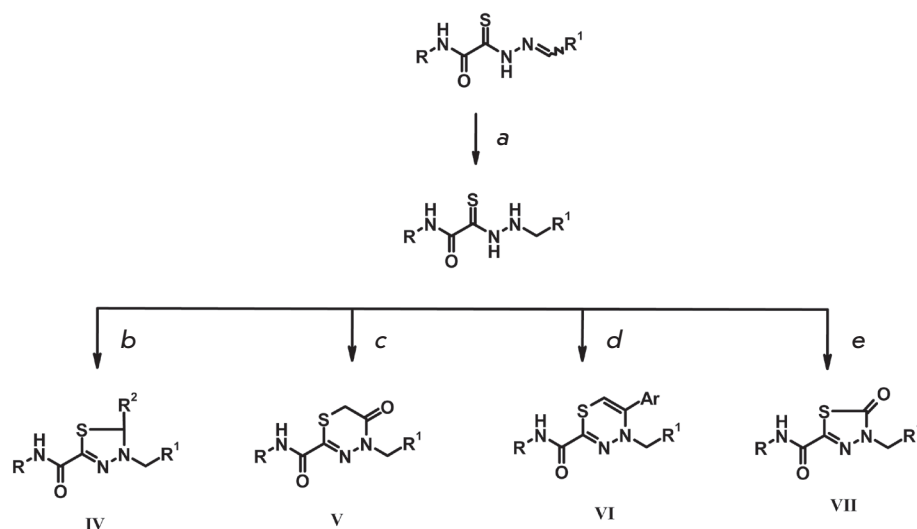
Based on the results of the analyses performed, three T3SS inhibitors were selected from a total of 15 for further study of their stability upon storage under various conditions. It turned out that these compounds in their dry form remained stable for a long time (according to the TLC data), whereas their activity in solutions decreased rapidly.

It was ascertained via the analysis of the published data [18, 19] that ring-chain tautomerism is typical of

Table 2. Inhibition of the intracellular development of *C. trachomatis* caused by the compounds under study

Compound №	Inhibition of the development of <i>C. trachomatis</i> (% of infection suppression) at different inhibitor concentrations, μM		
	12.5	25	50
1	30 \pm 4	99 \pm 7	100 \pm 1
2	70 \pm 5	98 \pm 3	100 \pm 1
3	1 \pm 1	30 \pm 6	60 \pm 5
4	5 \pm 2	10 \pm 5	60 \pm 11
5	0 \pm 1	10 \pm 3	75 \pm 6
6	20 \pm 4	40 \pm 5	90 \pm 12
7	15 \pm 3	85 \pm 8	100 \pm 2
8	5 \pm 3	10 \pm 5	90 \pm 13
9	30 \pm 8	40 \pm 7	90 \pm 10
10	0 \pm 2	10 \pm 7	98 \pm 3
11	10 \pm 2	90 \pm 7	100 \pm 1
12	0 \pm 1	70 \pm 15	100 \pm 3
13	15 \pm 5	70 \pm 14	100 \pm 15
14	15 \pm 4	45 \pm 7	100 \pm 3
15	40 \pm 3	85 \pm 8	100 \pm 2

Fig. 5. Synthesis of various heterocyclic compounds eliminating the existence of tautomerism and containing a closed thiocarbonyl group: *a* – sodium borohydride, methanol; *b* – aldehyde R^2 , isopropanol, hydrochloric acid; *c* – chloroacetic acid, isopropanol, ammonium acetate; *d* – ethanol, α -bromoketone, sodium acetate; *e* – carbonyldiimidazole, tetrahydrofuran.



thiohydrazones. Cyclic tautomers (thiadiazolines **II**) are easily oxidized by air oxygen, yielding inactive and toxic thiadiazoles **III** (Fig. 4).

These assumptions were experimentally verified via the synthesis of the hypothetical thiadiazoles and by NMR studies of the thiohydrazones of oxamine acids. It was ascertained that the both cyclic and linear forms are indeed present in thiohydrazone solutions; the oxidation products in the solution are identical to the target-synthesized thiadiazoles.

Attention was paid not only to the stability of the compounds, but also to the results of the MTT assay, which attested to an appreciably high toxicity of the compounds due to the suppression of the cellular respiration and the disruption of its oxidation-reduction potential. The analysis of the data published enables one to conclude that the open and considerably active thiocarbonyl group affects the level of toxicity; this group is bound to the reduced glutathione and induces the development of oxidative stress. This hypothesis was verified in the experiments in which glutathione was added to the cell culture medium, resulting in a 30%–40% decrease in toxicity.

Thus, the task was to enhance the stability of the resulting compounds whilst reducing their toxicity in such a manner so as to prevent the activity and specificity from being adversely affected.

Modification of thiohydrazones aimed at enhancing stability and reducing toxicity

In order to enhance the stability and reduce the toxicity of thiohydrazones, it was necessary to modify the structure so as to eliminate the formation of tautomeric forms and the open reactive thiocarbonyl group.

This was achieved by means of the synthesis of several heterocyclic compounds based on reduced thiohy-

drazones, according to the diagram shown in Fig. 5.

The synthesized heterocyclic derivatives of the thiohydrazones of oxamine acids were characterized by a higher solubility compared to the initial products; thus, their toxicity and capacity to inhibit T3SS was assessed. The compounds belonging to groups **IV** and **V** were characterized by low toxicity and exhibited specific activity with respect to T3SS, whereas the compounds belonging to groups **VI** and **VII** did not exhibit such activity.

Compounds belonging to group **IV** turned out to be instable in solutions (according to the TLC, the decomposition products were detected as early as 24 h after upon storage of the solutions at +20°C). The compounds from group **V** were characterized by acceptable indicators. A total of 12 compounds belonging to group **V** were synthesized on the basis of the fluorine-containing thiohydrazones of oxamine acids.

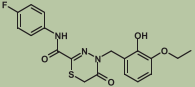
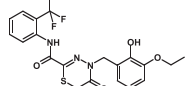
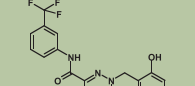
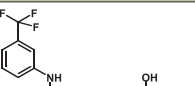
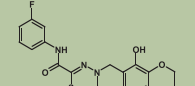
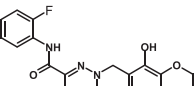
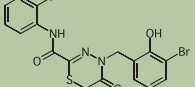
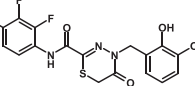
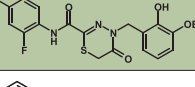
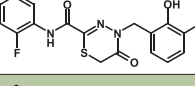
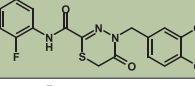
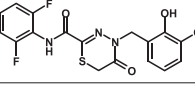
All these compounds proved considerably less toxic with respect to eukaryotic cells compared to the T3SS inhibitors that are already known and were obtained earlier (Table 3).

It was demonstrated that the synthesized thiadiazolones suppress the development of intracellular infection in dose-dependent fashion. Four compounds at a concentration of 50 μ M completely suppressed the infectious process in the cell culture (Table 4). The most efficient compound was selected via a comparison of the results of a determination of the toxicities and activities. This compound, known as CL-55, was used for a further, more advanced investigation of biological properties.

Inhibition of the effector function of T3SS by the synthesized inhibitor

The specific activity of chemical compounds (i.e., their ability to inhibit the action of the type three secretion

Table 3. Toxicity indices of thiadiazines

Compound №	Structure	Methylene blue staining, % of dead cells			Calcein assay, % of dead cells			MTT assay, % of metabolically inactive cells		
		12.5 μ M	25 μ M	50 μ M	12.5 μ M	25 μ M	50 μ M	12.5 μ M	25 μ M	50 μ M
16		0 \pm 2	5 \pm 2	10 \pm 2	0 \pm 1	6 \pm 2	12 \pm 3	1 \pm 2	8 \pm 3	13 \pm 3
17		0 \pm 2	6 \pm 1	11 \pm 3	0 \pm 3	7 \pm 2	10 \pm 3	0 \pm 2	9 \pm 3	14 \pm 4
18		0 \pm 1	4 \pm 2	9 \pm 2	0 \pm 2	8 \pm 1	12 \pm 3	0 \pm 1	9 \pm 2	15 \pm 3
19		0 \pm 1	6 \pm 2	11 \pm 2	0 \pm 1	5 \pm 1	13 \pm 4	0 \pm 1	8 \pm 3	13 \pm 3
20		0 \pm 1	5 \pm 2	12 \pm 3	0 \pm 2	6 \pm 2	14 \pm 2	0 \pm 1	8 \pm 3	14 \pm 4
21		0 \pm 1	6 \pm 2	9 \pm 2	0 \pm 1	7 \pm 2	12 \pm 3	0 \pm 1	9 \pm 3	14 \pm 3
22		0 \pm 2	5 \pm 4	10 \pm 5	0 \pm 3	6 \pm 3	12 \pm 4	1 \pm 2	8 \pm 3	15 \pm 4
23		0 \pm 2	7 \pm 1	12 \pm 4	0 \pm 1	6 \pm 2	13 \pm 3	0 \pm 2	9 \pm 3	14 \pm 5
24		0 \pm 1	4 \pm 1	10 \pm 4	0 \pm 1	6 \pm 1	14 \pm 3	0 \pm 1	10 \pm 3	17 \pm 4
25		0 \pm 3	6 \pm 2	11 \pm 2	0 \pm 2	7 \pm 2	12 \pm 3	1 \pm 3	9 \pm 2	14 \pm 3
26		0 \pm 1	7 \pm 1	12 \pm 4	0 \pm 2	6 \pm 3	13 \pm 3	0 \pm 1	8 \pm 3	14 \pm 2
27		0 \pm 2	5 \pm 1	10 \pm 3	0 \pm 3	7 \pm 2	11 \pm 3	1 \pm 2	8 \pm 3	13 \pm 3

system (T3SS) was studied using the method based on the detection of the *C. trachomatis* effector protein. The IncA protein, one of the effector proteins of this pathogen, is synthesized in the bacterial cell, followed by secretion and incorporation into the chlamydial in-

clusion membrane. This protein is known to be synthesized 6 h following the onset of infection; it emerges at the inclusion surface after 8 h. Specific antibodies can be used to detect the IncA protein within the inclusion membrane. The inhibition of the translocation of the

Table 4. Inhibition of the development of *C. trachomatis* in a cell culture

Compound №	Inhibition of the development of <i>C. trachomatis</i> (% of infection suppression) at different inhibitor concentrations, μM		
	12.5	25	50
16	15 \pm 5	40 \pm 8	87 \pm 6
17	24 \pm 3	69 \pm 6	100 \pm 5
18	7 \pm 2	26 \pm 7	63 \pm 6
19	12 \pm 3	37 \pm 6	65 \pm 8
20	9 \pm 3	26 \pm 7	59 \pm 7
21	15 \pm 7	42 \pm 7	86 \pm 10
22	2 \pm 4	23 \pm 6	52 \pm 4
23	34 \pm 6	78 \pm 8	100 \pm 2
24	30 \pm 8	69 \pm 6	98 \pm 4
25	25 \pm 5	59 \pm 6	95 \pm 5
26	24 \pm 3	82 \pm 9	100 \pm 2
27	27 \pm 4	65 \pm 10	100 \pm 2

effector protein IncA of T3SS in *C. trachomatis* with a selected inhibitor is shown in Fig. 6. Small inclusions can be detected by cell staining with anti-*C. trachomatis* major outer membrane protein (anti-MOMP) antibodies; the size of the inclusions correlates with the infection period.

An additional assay enabling the assessment of the specificity of compounds with respect to T3SS is based on the fact that the chlamydial IncA protein translocated through T3SS participates in the fusion of separate inclusions that are developed inside a cell. Staining of the cells infected with anti-MOMP antibodies has demonstrated that the introduction of a compound affecting the T3SS 8 h following infection results in the formation of several small nonfused inclusions, whereas large inclusions (one per cell) were observed in the control cells (Fig. 6).

Effect of the T3SS inhibitor on the morphology of *C. trachomatis* intracellular inclusions

The effect of the T3SS inhibitor, CL-55, on the intracellular development of the causative agent was studied by luminescence and electron microscopy. McCoy cell culture was infected with *C. trachomatis* and the inhibitor was simultaneously added to the culture medium at varying concentrations (12.5, 25 and 50 μM). Following 48 h, the cell cultures were analyzed by immunofluorescence and transmission electron microscopy.

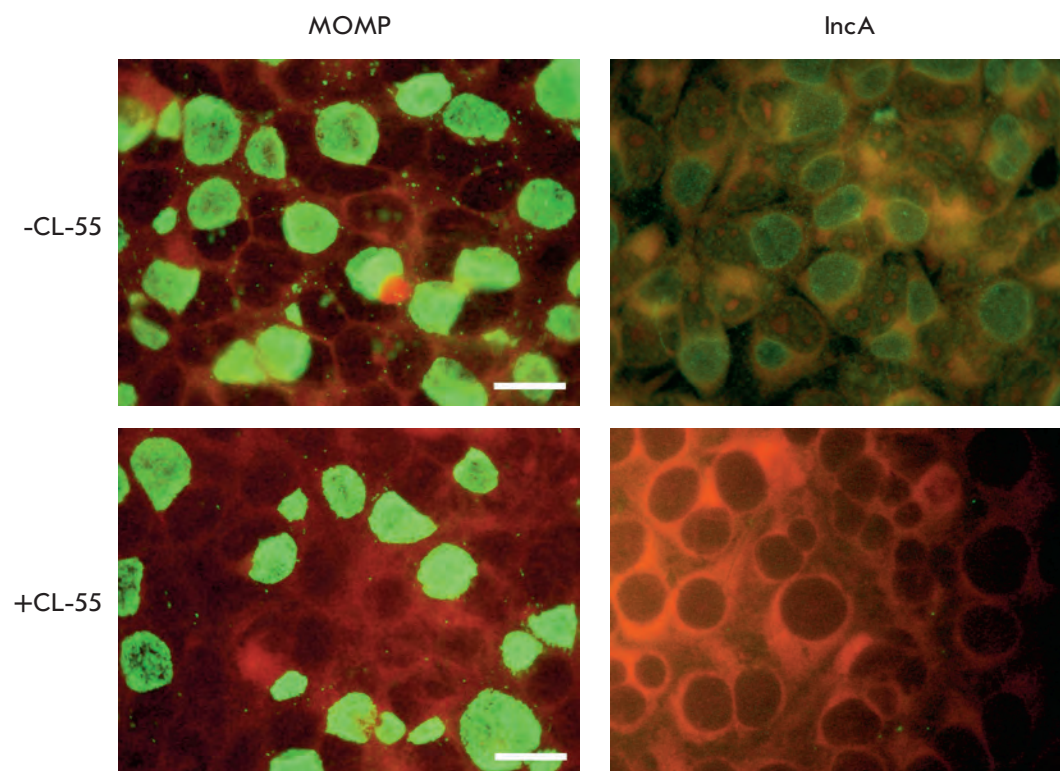


Fig. 6. Inhibition of the translocation of the *C. trachomatis* effector protein IncA into the intracellular inclusion membrane under the action of the T3SS inhibitor, CL-55. The McCoy cells were infected with *C. trachomatis*; 50 μM CL-55 was introduced into the culture medium after 8 h, and the medium was cultured for an additional 24 h. The chlamydial inclusions were stained with anti-*C. trachomatis* MOMP antibodies or anti-IncA antibodies and viewed under a luminescence microscope (green luminescence). Scale – 20 μM .

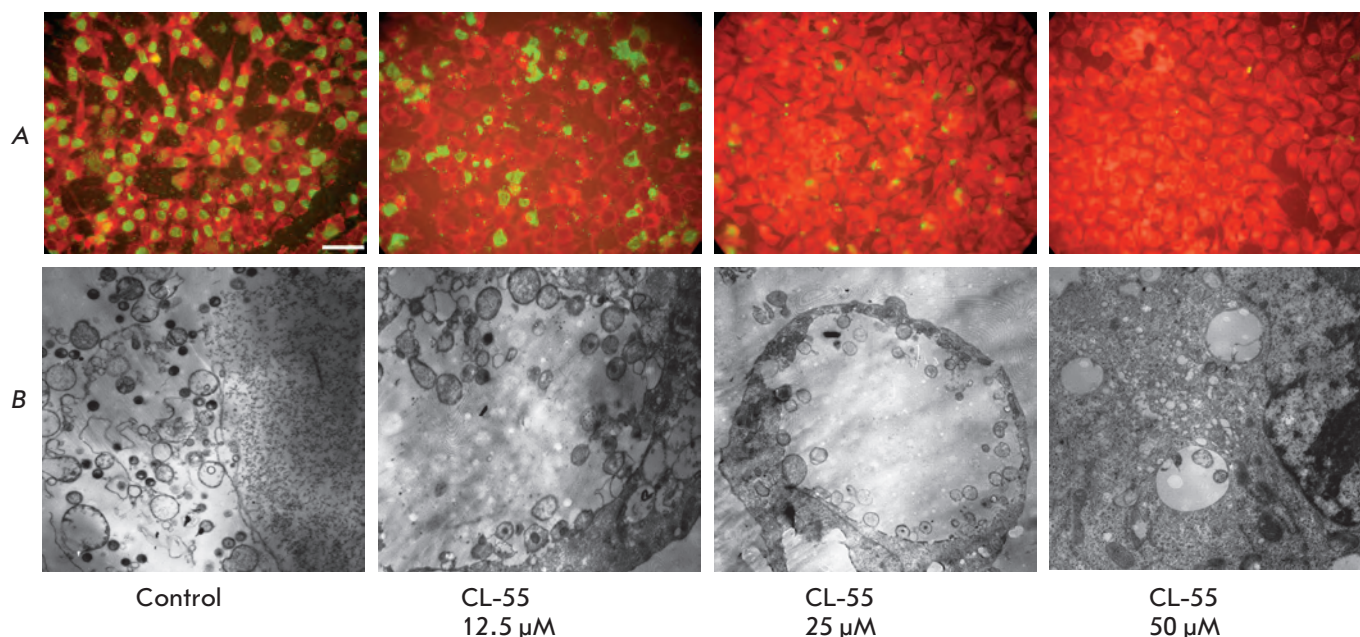


Fig. 7. Dose-dependent inhibition of the intracellular development cycle of *C. trachomatis* under the effect of CL-55. (A) – luminescent microscopy after staining with anti-*C. trachomatis* MOMP antibodies. Scale – 10 μM. (B) – Transmission electron microscopy. Control x 4000, CL-55 (12,5 μM) x 10000, CL-55 (25 μM) x 4000, CL-55 (50 μM) x 10000

The effect of CL-55 on the intracellular development of chlamydia resulted in the reduction of the count of infected cells and the inclusion size. At a dose as low as 12.5 μM, the inclusion count was 80% of the control; the inclusions were smaller in this case. At CL-55 concentration of 25 μM, the accumulation of chlamydia was inhibited by 50%. The average size of inclusions was several times lower compared to that in the control specimen. Almost no inclusions were observed when using CL-55 at a dose of 50 μM (Fig. 7).

A decrease in the inclusion size and a considerable decrease in the number of bacteria per inclusion were revealed through an electron microscopy study. Moreover, it should be noted that a typical pattern of termination of the chlamydial life cycle was observed in the control specimen: the overwhelming majority of the chlamydia inside the inclusions were of the extracellular form (the elementary bodies); some inclusions were disintegrated. Upon action of the T3SS inhibitor, the intracellular chlamydia was in the state of reticular bodies. Single, tiny inclusions containing abnormal bacterial cells were observed in the cells at a concentration of 50 μM. It has been thus demonstrated that the selected T3SS inhibitor dose-dependently suppresses the intracellular reproduction of *C. trachomatis*.

Effect of CL-55 on *C. trachomatis* progeny

The effect of the T3SS inhibitor on chlamydial devel-

opment was also assessed using the semi-quantitative method for the *in vitro* determination of the infectious properties of the pathogen. The calculation of the count of *C. trachomatis* infectious particles after the action of the inhibitor at different concentrations demonstrated a dose-dependent reduction (Fig. 8) and a complete viability inhibition in the presence of 50 μM of CL-55. Meanwhile, a significant number of the inclusions being formed possessed an atypical morphology at lower concentrations of the compound. No intracellular inclusions were formed upon further passaging.

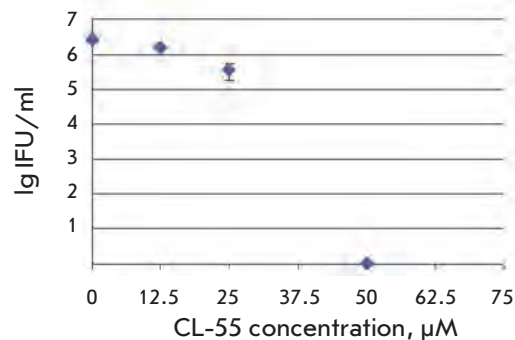


Fig. 8. Suppression of the *C. trachomatis* viability after the addition of varying concentrations of CL-55.

Effect of CL-55 on gene expression in *C. trachomatis*

The effect of the T3SS inhibitor on the activity of constitutive genes, the 16S rRNA and *trpA*, the tryptophan operon gene, and on the expression of the *incA* gene encoding the synthesis of the effector protein, and the *lcrE* gene of the T3SS regulatory protein in *C. trachomatis* were studied at this stage. The inhibitor was introduced simultaneously with cell infection; after 24 h, the RNA was isolated and the gene expression was analyzed using quantitative real-time PCR.

The use of the compound at concentrations of 25 and 50 μM reduced the activity of the 16S rRNA gene by 4 and 29 times, respectively. All specimens were subsequently normalized with respect to cDNA of the 16S rRNA DNA. The activity of the *trpA* and *incA* genes under action of 25 and 50 μM CL-55 remained unchanged compared to the control. These data attest to the following facts: firstly, the inhibitor had no pronounced effect on the pathogen metabolism; secondly, the inhibitor had no effect on the expression of the effector protein IncA gene at transcriptional level. The expression of the *lcrE* gene encoding the T3SS regulatory protein reduced by 90% in the presence of 50 μM . The study of the effect of CL-55 on the expression of T3SS-specific chlamydial genes aimed at elucidating the mechanism of action of the selected inhibitor was taken further recently.

Effect of the inhibitor CL-55 on the intracellular development of other chlamydial species

The effect of the selected T3SS inhibitor on the intracellular reproduction of representatives of two other species belonging to the *Chlamydiaceae* family, *C. pneumoniae* and *C. muridarum*, was assessed using the methods described above. The fact that CL-55 at a concentration of 50 μM completely inhibited both the intracellular accumulation and the viability of these chlamydial species attested to the universality of the effect of this inhibitor on the other Chlamydia.

DISCUSSION

The type III secretion system detected only in pathogenic bacteria was selected for use as a target for the search for new antibacterial drugs that would demonstrate efficacy with respect to both acute and chronic infections. The formation of this secretory apparatus, the so-called “molecular syringe,” starts after contact with a eukaryotic cell. It forms a pore in the target cell membrane, resulting in the direct penetration of the pathogenicity factors into the host cell’s cytoplasm [20]. T3SS can also function upon intracellular localization of the pathogen. The transport of the pathogenic factors results in cytoskeletal rearrangement, apoptosis inhibition, modification of the apparatus of transcription

and translation in the eukaryotic cells, modulation of cytokine production and other processes in the host cell which facilitates pathogen invasion, blockage of host protection, and the establishment of continuous persistence [21]. T3SS is absolutely essential for the development of an acute infectious process; the chronization of the infection fundamentally depends on how it functions. Thus, the specific inhibition of the T3SS function is expected to interrupt the infectious process both at early stages and upon a chronic course, thus allowing the immune system to eliminate the pathogen.

New efficacious T3SS inhibitors were explored according to the following scheme: similar regions were found in the molecules of organic compounds via a structural analysis of the known T3SS inhibitors, thereby enabling the construction of a new class of compounds exhibiting antibacterial activity that is specific with respect to T3SS. A significant number of such compounds have been synthesized, enabling molecular screening using cellular assays and with the selection of 15 compounds that specifically inhibit chlamydial T3SS *in vitro* in order to study the structure–activity relationship. The selected compounds were subsequently modified in order to enhance their solubility, stability, and biological activity, to reduce toxicity for eukaryotic cells, and enhance specific efficiency.

All this enabled to obtain a new T3SS inhibitor belonging to the class of heterocyclic compounds. This low-molecular-weight compound blocked the effector function of *C. trachomatis* T3SS. Thus, IncA (one of the early effector T3SS proteins) was not detected on the membrane of the chlamydial inclusion, and the process of homotypic phagosome fusion mediated by this protein was disrupted with the introduction of CL-55. An analysis of the expression of the *incA* gene has shown that the inhibitor does not reduce the transcription of this gene. A conclusion can thus be drawn that the selected compound is capable of selective inhibition of the translocation of the chlamydial effector protein.

Compound CL-55 had no pronounced inhibiting action on the level of expression of the *C. trachomatis* constitutive genes, a fact in agreement with the known mechanism of action of the T3SS inhibitors. According to this mechanism, the inhibitor affects the functioning of the secretory apparatus rather than the metabolism of a bacterial cell. Meanwhile, a considerable reduction of gene activity in one of the key regulators of chlamydial T3SS (CopN protein) was observed. Under normal conditions, this gene is expressed at all stages of the intracellular life cycle of chlamydia. Functioning as a chaperon protein, it participates in the secretion regulation of the proteins of the Inc family at the early stages of intracellular development and regulates the T3SS-mediated differentiation of reticular chlamy-

dial bodies into elementary bodies at the late stage of life cycle. It is of importance that the proliferation of reticular bodies depends on direct contact with an inclusion membrane and on interaction with the effector proteins translocated into it. Being a negative regulator, the CopN protein reduces the translocation of the major effector proteins (including IncA) on the inclusion membrane and closes the channel, thus impeding the translocation of the other T3SS effectors [22]. The resulting decrease in the expression of the gene encoding this regulatory protein can be an indicator of the specific action of the selected inhibitor on the regulation of the function of the chlamydial T3SS. The effect of the CL-55 inhibitor on the activity of a number of genes regulating the activity of *C. trachomatis* T3SS is currently under investigation.

Since the functioning of the chlamydial T3SS determines the possibility of intracellular development of the pathogen, the specific T3SS inhibitor is intended to disrupt the life cycle and block the infectious process upon both acute and persistent forms of the infection. The designed inhibitor suppressed the reproduction of three chlamydial species, *C. trachomatis*, *C. pneumoniae*, and *C. muridarum*, on cell culture models. The suppression manifested itself in the following ways: in the changes observed in the morphology of the chlamydial inclusions, by the disruption in the differentiation

of the reticular bodies into elementary bodies, and by the inhibition of the infectious properties of the pathogen. Moreover, it has been demonstrated (the results are not shown) that the selected inhibitor blocked the secretion of the T3SS effector proteins in *Salmonella*, a representative of the taxonomically non-related group of pathogenic bacteria, which may attest to the universality of the resulting T3SS inhibitor. Meanwhile, the compound exhibited no bactericidal effect on a number of Gram-negative and Gram-positive bacteria, representatives of the normal bacteria.

Thus, a new T3SS inhibitor belonging to the class of heterocyclic compounds has been obtained via targeted chemical synthesis, experimental screening, and chemical optimization. This compound is currently being studied with the use of experimental animals for its therapeutical activity and pharmacokinetic properties. These studies are intended at probing the application of the compound for the further development of an antibacterial drug that would be efficacious with respect to the acute and chronic forms of infection. ●

This work was supported by Government Contract № 16.512.11.2248 under order of the Department of Priority Directions of Science and Technology of the Ministry of Education of the Russian Federation.

REFERENCES

- Shiinsky G.E., Merzlyakov V.A., Timofeeva S.B. // *Vestn. dermatologii i venereologii*. 1999. № 1. P. 11-13.
- Dean D. // *Drugs Today (Barc)*. 2009. V. 45. P. 25-31.
- Bashmakov Y.K., Zigangirova N.A., Pashko Y.P., Kapotina L.N., Petyaev I.M. // *Comp. Hepatol*. 2010. V. 28. P. 3-9.
- Zigangirova N.A. // *Sb. nauchnikh trudov: "150 let so dnya rojdenia N.F. Gamalei"* Moscow, 2009. P. 49-61.
- Zigangirova N.A., Fedina E.D., Zorina V.V., Bortsov P.A., Tokarskaya E.A Karyagina, A.S., Alekseevsky, A.V., Krayushkin, M.M., Zayakin, E.S., Yarovenko, V.N., Gintsburg, A.L. // *Journal Microbiology and Epidemiology and Immunobiology*. 2009. № 4. P. 71-77.
- Guntsburg A.L., Zigangirova N.A., Zorina V.V. // *Vestn Ross Akad Med Nauk*. 2008. № 10. P. 34-39.
- Erhardt M., Namba K., Hughes K.T. // *Cold Spring Harb Perspect Biol*. 2010. doi: 10.1101/cshperspect.a000299
- Karyagina A.S., Alexeevsky A.V., Spirin S.A., Zigangirova N.A., Gunzburg A.L // *Molecular Biology*. 2009. V.43. № 6. P. 897-916
- Felise H.B., Nguyen H.V., Pfuertner R.A., Barry K.C., Jackson S.R., Blanc M.P., Bronstein P.A., Kline T., Miller S.I. // *Cell Host & Microbe*. 2008. V. 4. P. 325-336.
- Bailey L., Gylfe A., Sundin C., Muschiol S., Elofsson M., Nordström P., Henriques-Normark B., Lugert R., Waldenström A., Wolf-Watz H., et al. // *FEBS Lett*. 2007. V. 581. P. 587-595.
- Tautz L., Bruckner S., Sareth S., Alonso A., Becattini B., Salvesen G.S., Mustelin T. // *J. Biol. Chem*. 2005. V. 280. № 10. P. 9400-9408.
- Kauppi A.M., Andersson C.D., Norberg H.A., Sundin C., Linusson A., Elofsson M. // *Bioorg. Med. Chem*. 2007. V. 15. № 22. P. 6994-7011.
- Kauppi A.M., Nordfelth R., Uvell H., Wolf-Watz H., Elofsson M. // *Chem. Biol*. 2003. V. 10. P. 241-249.
- Patani G.A., LaVoie E.J. // *Chem Rev*. 1996. V. 96. № 8. P. 3147-3176.
- Yang H., Hendricks R.T., Arora N., Nitzan D., Yee C., Lucas M.C., Yang Y., Fung A., Rajyaguru S., Harris S.F., et al. // *Bioorg. Med. Chem. Lett*. 2010. V. 20. № 15. P. 4614-4619.
- Gunzburg A.L., Zigangirova N.A., Tokarska E.A., Zorina V.V. // *Patent №2400471 27.09.2010 r. RF. C07D 213/75, C07D 338/38, C07C 327/56, C07C/225/16*.
- Hagmann W.K. // *J. Med. Chem*. 2008. V. 51. № 15. P. 4359-4369.
- Zelenin K.N., Khrustalev V.A., Alekseev V., Sharbatyan P.A., Lebedev A.T. // *Khimia geterotsikl. soed*. 1982. № 7. P. 904-910.
- Lempert-Sreter M., Lempert K., Möller J. // *J. Chem. Soc. Perkin Trans*. 1983. V. 1. № 9. P. 2011-2020.
- Cornelis G.R. // *Nat. Rev. Microbiol*. 2006. V. 4. P. 811-825.
- Hoare A., Timms P., Bavoil P.M., Wilson D.P. // *BMC Microbiol*. 2008. V. 8. P. 5-16.
- Delphine S., Beeckman A., Daisy C., Vanrompay G. // *Curr. Issues Mol. Biol*. 2010. V. 12. P. 17-42.

Recombinant Human Cyclophilin A *in vitro* Inhibits the Formation of Fibrin Clot

Yu. A. Morozov^{1*}, L. M. Khromykh², I. I. Dementieva¹, M. A. Charnaia¹, N. L. Kulikova³, D.B. Kazansky²

¹Petrovsky Russian Research Center of Surgery, Russian Academy of Medical Sciences, Abrikosovskii per., 2, Moscow, 119991, Russia

²Carcinogenesis Institute, Blokhin Cancer Research Center, Kashirskoe sh., 24, Moscow, 115478, Russia

³Institute of Immunological Engineering, pos. Liubuchany, Chekhovskii raion, Moskovskaia obl., 142380, Russia

*E-mail: moroz_111@rambler.ru

Received 27.03.2012

Copyright © 2012 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 chemotactic properties of cyclophilin A are well-known. There exists however a poor level of understanding regarding the hemostatic effects of this protein. Herein it is shown that recombinant human cyclophilin A (rhCyA), in contrast to the granulocyte colony-stimulating factor, is capable of inhibiting *in vitro* the formation of a fibrin clot, thereby violating the spatial dynamics of clot growth; this effect is transient and dose-independent. Furthermore, the hypothesis that the conformational changes in the thrombin–rhCyA complex may mediate the anticoagulant effect of rhCyA on the autowave processes of blood clotting is postulated.

KEYWORDS recombinant human cyclophilin A; spatial dynamics of clot growth; anticoagulant activity.

ABBREVIATIONS rhCyA – recombinant human cyclophilin A; DCG – delay of clot growth, min; IGR – initial growth rate, min; SSGR – steady-state growth rate, min; CS-30 – clot size after 30-minute growth, μm ; CD – the clot density; G-CSF – granulocyte colony-stimulating factor.

INTRODUCTION

Cyclophilin A is a protein with a molecular weight of approximately 20 kDa. Cyclophilin A exhibits cystrans-isomerase activity and possesses a wide and varied range of functions. Cyclophilin A is produced by thymus cells and exhibits chemotactic activity by regulating the stem cell migration from the bone marrow to the peripheral regions [1]. Activated macrophages are also capable of secreting rhCyA, which attracts mature monocytes, neutrophils, eosinophils, and the activated T cells to the inflammatory focus [2].

However, the participation of this protein in the role played by the hemostatic system has not been sufficiently studied. The smooth muscle cells of vascular walls are known to secrete rhCyA under hypoxic and oxidative stress conditions; the activated platelets in the injured vascular wall are capable of secreting rhCyA [3], which in turn stimulates endotheliocyte proliferation [4], thereby facilitating the development of regenerative processes.

The *in vitro* effect of recombinant human cyclophilin A on the growth of a fibrin clot is studied in this work.

EXPERIMENTAL

Recombinant human rhCyA was produced using *Escherichia coli* BL21 cells transformed with plasmid derived from pCyPAwt/pGEX-2TK, which encodes the GST-CyPA fusion protein. The genetic construct was used with the kind permission from M. Bukrinsky (Albert Einstein College of Medicine of Yeshiva University, USA). The synthesis of GST-CyPA was induced by adding 0.2 mM isopropyl- β -D-thiogalactopyranoside until a finite concentration of 100 μM was achieved. The cells were then deposited via centrifugation, suspended in a Na-K phosphate buffer (pH 7.3), and subjected to ultrasonic destruction. The cell lysate was centrifugated; the superprecipitate was applied onto a GSTrap FF column (“GE Healthcare”). The GST-CyPA fusion protein was eluted with 50 mM Tris-HCl buffer (pH 8.0) containing 10 mM reduced glutathione. The fractions containing GST-CyPA were combined, transferred into a 10 mM Tris-HCl solution (pH 8.0) on a Hi-Trap Desalting column (“GE Healthcare”), and applied onto a MonoQ 5 \times 50 column (“GE Healthcare”). Elution was performed using a 1 M sodium chloride gradient (0–1 M) in 10 mM Tris-HCl, pH 8.0. The fractions con-

Effect of 30-min and 3-h incubations of platelet-rich plasma with rhCyA on the parameters of the spatial dynamics of clot growth

Probe	DCG, min	IGR, min	SSGR, min	CS-30, μm	CD, arb. units
30-min incubation					
Control	3.0 \pm 1.6	53.3 \pm 16.7	32.8 \pm 13.1	868.0 \pm 180.1	13679.0 \pm 2395.0
Test	14.8 \pm 5.7*	14.3 \pm 6.2*	16.8 \pm 7.2*	341.1 \pm 114.3*	9116.0 \pm 2086.0*
3-h incubation					
Control	1.0 \pm 0.3	59.7 \pm 4.8	37.0 \pm 9.0	1329.0 \pm 244.7	16182.0 \pm 2797.0
Test	1.4 \pm 0.6	61.8 \pm 6.7	33.5 \pm 5.2	1258.0 \pm 189.8	13971.5 \pm 3294.5

Note: * – reliability of differences $p < 0.05$ compared with the control.

taining GST-CyPA were transferred via dialysis into a Na-K phosphate buffer (pH 7.3); a thrombin solution (“Sigma-Aldrich”) on the basis of a 1 U/mg fusion protein was subsequently added. The cleavage was carried out for 10 h at 4°C. The mixture of the CyPA and GST proteins was separated on a GSTrap FF column. Thrombin was removed from the CyPA solution on a HiTrap Benzamidine FF column (“GE Healthcare”). The degree of reaction completion was controlled via polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAAG). The purity of the protein used was ~98%.

The *in vitro* effect of rhCyA on the spatial dynamics of clot growth was studied using venous blood collected by gravity flow from the peripheral vein of healthy volunteers and stabilized with a 3.7% solution of sodium citrate at a 1 : 9 citrate : blood ratio.

Blood samples were centrifugated at 1500 rpm for 7 min. The platelet-rich plasma was divided into two parts. The first probe was used as the control (C); rhCyA at a final concentration of 10 or 50 $\mu\text{g}/\text{ml}$ was added into the second probe (experimental – E), followed by incubation at 37°C for 30 min and 3 h with constant stirring to prevent platelet sedimentation. A granulocyte colony-stimulating factor (G-CSF) at a final concentration of 20 $\mu\text{g}/\text{ml}$ was added as the reference substance.

After the incubation, the platelet-rich plasma was centrifugated at 13700 rpm for 10 min. The platelet-free plasma was used in order to study the spatial dynamics of clot growth in accordance with the manual for ThromboImager-2 instrument (OOO “GemaCor”, Russia).

The values of the delay of clot growth (the so-called lag phase) (min), the initial (min) and steady-state (min) clot growth rates, clot size after the 30 minutes of growth (μm), clot density (arbitrary units), and spontaneous thrombus formation in a chamber were recorded.

The experimental results were statistically processed using the Mann–Whitney test. differences with $p < 0.05$ were considered to be statistically significant.

RESULTS

The parameters of the spatial dynamics of clot growth upon the incubation of the platelet-rich plasma with rhCyA for 30 min and 3 h are listed in *Table*.

It is clear from *Table* that the 30-min incubation of platelet-rich plasma with rhCyA at a final concentration of 10 $\mu\text{g}/\text{ml}$ resulted in a statistically significant inhibition of the spatial dynamics of clot growth; i.e., rhCyA exhibited the properties of a substance exhibiting a hypocoagulation effect. Spontaneous thrombus formation in the volume of the measuring chamber was detected for two samples in the control group. The incubation with rhCyA prevented the development of spontaneous clot formation in the system. However, this effect lasted for a short time; the differences in the parameters of the spatial dynamics of clot growth disappeared as early as after 3 h (*Table*).

No dose-dependent effect of the agent on clot growth was revealed upon incubation of the samples of platelet-rich plasma with rhCyA at the final concentration of 50 $\mu\text{g}/\text{ml}$. The parameters of the spatial dynamics of clot formation were identical to those during the incu-

bation with rhCyA at a final concentration of 10 µg/ml. No fall in the clot growth rate was observed after 3 h of incubation.

G-CSF had no effect on the values of the determined indices, which were comparable to those in the control group.

DISCUSSION

Clot formation in the organism only occurs locally; a clot is formed in a small area of a damaged blood vessel wall. Vessel wall damage results in exposure of the tissue factor (the protein immobilized on the membrane of vessel wall cells, which are covered with an endothelium layer prior to the damage) into the blood. Diffusion of the coagulation factors plays a significant role in the spatial growth of the clot and its localization. Individual pathologies have specific effects on the activity of the components of the blood coagulation system, causing defined changes at various stages of the process. The disruptions in the mechanism of haemocoagulation manifest themselves clinically either as a haemorrhage, as thrombosis, or as a combination of these factors.

It should be noted however that the clot formation process to a significant extent occurs in the organism under spatially heterogeneous conditions. Therefore, the role of individual coagulation reactions can differ for *in vitro* experiments under stirring and *in vivo* [5].

The method for recording the spatial dynamics of clot growth is based on the contact between the blood and the tissue factor in a measuring cell containing immobilized thromboplastin [6]. The activated nucleation and spatial growth of a clot is recorded on the basis of dark-field scattering microscopy. This method enables to measure such significant characteristics of the process as the clot growth rate, clot size, and spontaneous clot formation; i.e., information that is not available when using homogenous methods [7]. This allows one to perform the simultaneous and independent recording of the disruptions at all stages of the process.

It has also been ascertained (unpublished data) that rhCyA at a dose of 10 µg does not affect the internal and external pathways of blood coagulation *in vitro*. However, an eightfold increase in the level of plasma anti-Xa activity as compared to control values was observed 10 min after the onset of incubation with rhCyA. After 30- and 60-min-long incubations with the protein, the level of anti-Xa-activity decreased to some extent; however, it was still 3 times higher than that in the control sample. Identical data were obtained in animal-model experiments. The introduction of rhCyA at a dose of 100 µg to P815 intact mice was accompanied by a statistically significant threefold increase in the level of inhibition of the factor Xa; a 200 µg dose re-

sulted in the inhibition of the factor Xa, corresponding to 0.6 anti-Xa U/ml. The data presented in this paper indicate that rhCyA impedes clot growth by slowing the rate at which it forms via the inhibition of the factor Xa. Clot growth does not start immediately and is a heterogeneous process; therefore, thrombus formation is less plausible, since this clot is washed off by the bloodstream in the organism [8].

The blood coagulation has threshold and autocatalytic properties; therefore, its spatial behavior may have much in common with the active media. In 1994, Ataullakhanov and Guriya put forth a hypothesis according to which spatial (but not homogeneous) clot formation is an autowave process [9]. A deficiency in the factors of the internal coagulation pathway or their inhibition may disrupt the autowave phase of growth.

Autowave propagation of thrombin cannot be slowed by natural inhibitors, since they are incapable of ending an autowave because of their homogeneous spatial distribution. To provide deceleration, the autocatalytic generation of coagulation factors has to be suppressed. The hypothesis of an autowave propagation of the coagulation process involves the existence of a double autowave mechanism for the termination of the thrombin wave [10]. It is believed that the second wave of substance can be initiated in the environment; this wave emerges after thrombin formation and punctuates its synthesis. If this wave propagates at a higher speed than the first wave, it will be able to end the propagation of the first wave.

A thrombin inhibitor capable of propagating autocatalytically thus far remains unknown. Assumptions have been made that thrombin can exist in pro- and anticoagulant forms [11]. In its procoagulant state, thrombin is highly active with respect to fibrinogen and poorly active with respect to protein C. On the contrary, thrombin in its anticoagulant form (due to the change in the conformation of thrombin conjugated with some compounds) shows a high level of activity with respect to protein C and weak activity with respect to fibrinogen [12]. Our data on the action of rhCyA on the spatial dynamics of clot growth enables to hypothesize that rhCyA might influence the mechanisms regulating the autowave process of blood coagulation.

CONCLUSIONS

Recombinant human cyclophilin A exhibits a pronounced dose-independent anticoagulant effect.

Recombinant human cyclophilin A can prevent spontaneous thrombus formation.

The anticoagulant effect of recombinant human cyclophilin A persists for no longer than 2 h. ●

REFERENCES

1. Khromykh L.M., Kulikova N.L., Anfalova T.V., Muranova T.A., Abramov V.M., Vasiliev A.M., Khlebnikov V.S., Kazansky D.B. // *Cell Immunol.* 2007. V. 249. № 1. P. 46–53.
2. Xu Q., Leiva M.C., Fischkoff S.A., Handschumacher R.E., Lyttle C.R. // *J. Biol. Chem.* 1992. V. 267. P. 1168–1171.
3. Coppinger J.A., Cagney G., Toomey S., Kislinger T., Belton O., McRedmond J.P., Cahill D.J., Emili A., Fitzgerald D.J., Maguire P.B. // *Blood.* 2004. V. 103. № 6. P. 2094–2104.
4. Lay A., Jiang X.M., Kisker O., Flynn E., Underwood A., Condrón R., Hogg P.J. // *Nature.* 2000. V. 408. P. 869–873.
5. Tijburg P.N., Ryan J., Stern D.M., Wollitzky B., Rimón S., Rimón A., Handley D., Nawroth P., Sixma J.J., de Groot P.G. // *J. Biol. Chem.* 1991. V. 266. № 18. P. 12067–12074.
6. Ataulakhanov F.I., Guriia G.T. // *Biofizika.* 1994. V. 39. № 1. P. 89–96.
7. Kondratovich A.Y., Pokhilko A.V., Ataulakhanov F.I. // *Biochim. Biophys. Acta.* 2002. V. 1569. № 1–3. P. 86–104.
8. Ovanesov M.V. Influence of the internal path of blood clotting in the spatial dynamics of clot growth. M.: HRC RAMS, 2002.
9. Ataulakhanov F.I., Guria G.T., Safroshkina A.Y. // *Bi-fizika.* 1994. V. 39. № 1. P. 97–106.
10. Ataulakhanov F.I., Vilkova R.I., Pokhilko A.V., Sinauridze E.I. // *Biofizika.* 1994. № 39. № 4. P. 713–720.
11. Ataulakhanov F.I. // *Soros Educational Journal.* 2000. V. 6. № 7. P. 2–10.
12. Griffin J.H. // *Nature.* 1995. V. 378. № 6555. P. 337–338.

GENERAL RULES

Actae 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 *Actae 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_{pt} – temperature of phase transition). One exception is mln – million, which should be used without a period.
- Decimal numbers should always contain a period and not a comma (0.25 and not 0,25).
- The hyphen (“-”) is surrounded by two whole spaces, while the “minus,” “interval,” or “chemical bond” symbols do not require a space.
- The only symbol used for multiplication is “×”; the “×” 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) 930-80-05

ERRATUM

IN THE ARTICLE T.V. DEMIDKINA “ THE 110TH ANNIVERSARY OF ACADEMICIAN ALEXANDER EVSEEVICH BRAUNSTEIN”, PUBLISHED IN ACTA NATURAE, 2012, V. 4, № 1 (12) EDITORIAL MADE THE TRANSLATION MISTAKES:

1. Correct spelling of surname: Kritzman M. G.
2. “Many foreign scientists ...” read as H.B.F. Dixon (Great Britain), H.C. Dunathan and D.E. Metzler (USA)
3. A portrait’s of A.E. Braunstein caption supplemented with quote: “On this trail there were erroneous findings and questionable inferences, but many important results gained general acceptance as fundamental contribution to present-day knowledge of the pathways, enzymes, and integration of nitrogen and sulphur metabolism. For an active researcher there can be no happier reward. “
4. Page 7, left column - “electron” read as “electronic”.