Principles of Creation of Protein Carriers of DNA New Derivatives of Human Epidermal Growth Factor for Gene Therapy

G. E. Pozmogova*, and A. N. Chuvilin***

Translated from *Byulleten' Eksperimental'noi Biologii i Meditsiny*, Supplement 2, pp. 87-93, April, 2007
Original article submitted February 18, 2006

We formulated a new approach to the creation of transport proteins for the delivery of foreign DNA to target cells and used it for obtaining a polypeptide PGEk. Structural and functional analysis of PGEk—DNA complexes demonstrated good prospects for the creation of a wide spectrum of targeted preparations for gene therapy. These approaches and regularities are necessary for construction of new DNA carriers selective for various cell types.

Key Words: *gene therapy; protein complexes; epidermal growth factor; nucleoprotein complexes; DNA conformation*

Gene therapy implies the use of oligo- or polynucleotides acting via expression of a transgene or interaction with intracellular proteins. Various strategies of gene therapy of diseases of different etiologies were proposed and are developing now. However, the chasm between splendid laboratory results and clinical practice can be explained by peculiarities of genetic material (GM) as a therapeutic agent. The development of systems for targeted delivery of foreign DNA can help to overcome these difficulties.

Directed transport of GM implies delivery of a polynucleotide to certain tissues and organs, its transport across the plasma membrane of target cells, and traffic to a specific intracellular target. It is very important to prevent transport of exogenous GM into other cells of the organism not requiring transformation. Creation of an effective and selective DNA transporter is a complex task determined by the mechanism of therapeutic effect, nature of GM, and other parameters. None known methods of directed DNA transport can solve all the above listed problems. In light of this, creation of an effective delivery system adequate to a certain clinical situation is a key problem of gene therapy.

Viral vectors, transporters of expressed GM obtained by fusion of exogenous DNA with viral genome are most effective. These transporters usually retain their initial tissue specificity. In recent years, affinity of the viral system (*e.g.* based on adenovirus) to the target cells was attained via exposure of additional specific peptides of the cell surface. Achievements and problems of obtaining and use of viral vectors are discussed in many recent papers.

Let us note some most important aspects in this field. First of all, we should note a well-known problem of safety of clinical application of viral vectors, which stems from their possible recombination with homologous natural viruses, development of immune and inflammatory reactions, selectivity to the target tissue and regulation of transgene expression. Another open question is related to discussion of necessary and safe load of foreign DNA. For instance, missing function in hemophilia or

*Laboratory of Artificial Antibodygenesis, Research Institute of Physicochemical Medicine, Federal Agency for Health Care and Social Development; **Bioengineering Center, Russian Academy of Sciences, Moscow. ***Address for correspondence:*** pozm@mail.ru. G. E. Pozmogova

0007-4888/07/1443-0457

mucoviscidosis was successfully corrected after transfection of just 5-10% cells [16]. Some authors observed changes not only in transfected, but also in neighbor cells, in which no foreign DNA was delivered [8,16,18,20]. Moreover, some genotherapeutic strategies are based on the use of oligonucleotide derivatives that cannot be delivered by virus-based vectors. On the other hand, the size of virus-delivered DNA is limited. In light of this, creation of targeted nonviral GM delivery systems for gene therapy is an actual problem [1,8-14,15-20]. The biochemical approach to the targeted DNA delivery is based on natural mechanisms of cell metabolism. The method consists in construction of complexes or covalent conjugates of the drug with receptor-binding domain (usually the whole peptide or protein) responsible for selective internalization of the whole construct by the target cells.

Principles of Construction of Protein Vectors

Targeting of biochemical transport systems is based on differences in tissue composition of surface cell receptors and antigens or its changes during pathological processes. GM is internalized via receptor-mediated endocytosis. Therefore, the minimum targeted structure for DNA delivery should contain a transmembrane receptor-specific domain and a DNA-binding domain.

The search for specific ligands is a specific and very complex problem; the ways for this search is just outlined. The problem is that the receptors of pathogenesis and their polymorphism are described for just few pathologies.

Targeting domain. The most numerous group of nonviral vectors for selective delivery of foreign GM to target cells includes protein and peptide constructs, where targeting ligand is presented by a synthetic peptide, natural or recombinant protein, or protein fragment. Asialoglycoprotein, transferring, insulin, melanocyte-stimulating hormone, viral proteins and capsids, α-fetoprotein (AFP), and epidermal growth factor (EGF) receptor ligands were used for this purpose [1,2,8-20,24].

The use of targeting peptides is promising, but still is only a trend. Some studies showed that the use of oligonucleotide—peptide conjugates or noncovalent complexes with oligonucleotides can increase bioavailability of oligonucleotides, improve stability of their complexes with polynucleotide targets and intracellular stability, and changes in pharmacokinetic properties [9,13].

Of special interest is high membrane tropism of oligonucleotide complexes with amphiphilic peptide MPG (Ac-GALFLGFLGAAGSTMGAWSQP KSKRKV-NHCH$_2$CH$_2$SH) consisting of a hydrophobic fragment of HIV-1 fusion protein gp41 and nuclear localization signal (NLS) sequence of SV-40 virus T antigen [9,13,19,20].

However, delayed biological consequences of therapeutic use of viral protein fragments cannot be predicted. Therefore, construction of vectors on the basis of natural intrinsic proteins or homologous proteins maximally close to them is a more actual and promising approach.

Not all mentioned ligands ensure high specificity for the target cells and detailed study of their selectivity and transport properties is required in each case. We hope that further investigations will reveal high affinity compounds capable of distinguishing any cell line.

In our experiments we used human AFP and recombinant human EGF, the most typical ligands for tumor cells, as the targeting domains. Chemical conjugates of phosphodiester and thiophosphoryl antisense oligonucleotides with AFP and EGF residues obtained by us not only 5-20 fold increased the efficiency of internalization of oligonucleotides, but also ensured their selectivity to the target cells [9,12]. The following two approaches were chosen for condensation. First, the synthesized antisense oligonucleotides with terminal aminoalkyl group were treated with 5-maleimidohexanoic acid N-hydroxysuccinimide ester and AFP was treated with a bifunctional agent SPDP (Sigma) and then with dithiothreitol. After purification, the modified components were mixed and the obtained conjugate was isolated by gel chromatography [12].

Another method of condensation implied the synthesis of oligonucleotides with 3'-terminal uridine, where the cis-diol group was directionally cleaved with sodium periodate. A small excess (10%) of obtained and purified dialdehyde derivatives interacted with EGF (obtained according to out patented method) without catalysis and condensing reagents [7]. Evidently, condensation occurs at one N-terminal amino group of the protein due to the formation of amide group, which is more labile under physiological conditions. The reaction was controlled by HPLC (Fig. 1), the yields of conjugates after purification were 78-87% (per initial EGF). The composition of conjugates (protein:oligonucleotide 1:1) was determined by the ratio of absorption at 260 and 280 nm and by MALDY-TOF mass spectrometry.

DNA-binding domain. Introduction of a special additional sequence capable of effectively binding DNA into proteins made it possible to construct vectors with universal length and GM content.

Now, synthetic or natural polycations (polylysine, aminodextran, chitosan, protamine, spermin, fragments of histone protein, oligomers of ornithine, arginine, and lysine, and synthetic polycationic polymers) or fragments of DNA-binding proteins (topoisomerase I, DNA-binding domain of GAL4 yeast protein) are used as DNA-binding domain [1,2,8-14,16-20,24]. GM is retained in the carrier structure due to covalent, ionic, or affinity interactions. Intercalators, agents strongly but reversibly binding with DNA duplex and triplex and increasing their stability, were also used for this purpose [9,25]. In some studies, the formation of the complex was based on biotin—streptavidin interaction [9,14,18,24], for instance, a three-domain protein carrying streptavidin residue and effectively binding biotinylated DNA was obtained [12].

Other studies showed that DNA can be carried also by short peptide sequences like nuclear localization signal (NLS) sequence of SV40 virus T antigen [1,9-11,13,15,19,20]. It is evident that introduction of DNA-binding domain should not change the receptor-binding properties of the targeting component and function of the carried GM.

Despite encouraging results of selective *in vitro* delivery of foreign GM with various protein vectors obtained in different laboratories (*e.g.* our experiments with AFP polylysine derivatives, EGF conjugates, *etc*. [8-12]) their clinical use faces a number of problems.

Low solubility and unacceptable geometrical size of DNA—vector complexes (diameter >1000 nm) and total cytotoxicity of polylysine are the main obstacles for practical use of polyplex vectors [9,16-19].

Additional domains Additional domains were introduced into protein carriers of DMA for determination of the intracellular traffic of internalized DNA fragments. The use of signal or endosomolytic sequences in some cases considerably increased the efficiency of transfection [8-11,14,16-18, 20,24]. Unfortunately, these multidomain constructs are difficult to prepare because of the problem of adequate polypeptide folding, while their use is limited due to immunogenicity of foreign proteins.

Addressed supramolecular nucleoprotein complexes. We believe that the problem of delivery of exogenous GM with protein carriers can be more fruitfully discussed from the viewpoint of not molecular, but supramolecular chemistry (chemistry of molecular assemblies). During the development of the structure of transport proteins, the properties of their nucleoprotein complexes should be predicted and the interaction of these nucleoprotein assemblies with natural biopolymers (which, in turn, are also associated with intracellular molecular structures) should be taken into account. The present approach is based on the concept that transport protein should spontaneously bind GM with the formation of a stable (under physiological conditions) low-immunogenic complex exhibiting selectivity towards the target cells. Moreover, the vector should be technologically available and stable.

Chemical and enzymatic synthesis of the gene, growing a producer strain, and biotechnological accumulation of the recombinant polypeptide are now the most promising, adaptable, and available method for creating new proteins. The authors took part in the creation of producers of both natural and artificial multidomain proteins [4-8,14,23].

Another important aspect in the formation of a transport nucleoprotein complex is the characteristics (*e.g.* conformation) of the transported GM. Since biological activity of oligonucleotide aptamers often depends on their tertiary structure, it is interesting to study of the effect of complex formation with the protein on oligonucleotide conformation. The methods of investigation of kinetic and thermodynamic characteristics of various DNA structures [2,3,21,22] provided the basis for physicochemical analysis of the formation and properties of protein—nucleic acid associates [1,5].

In our search for new vector protein we hoped to find a carrier combining targeted action and technological availability of recombinant proteins with universal nature and high efficiency of cationic structures.

Structure of PGEk, a protein recombinant DNA carrier

A new recombinant protein vector PGEk (**P**rotein **G**en-carrier based on **E**pidermal Growth Factor, 64 àmino acid residues) was obtained for the delivery of oligonucleotides into cells superproducing EGF receptor. Human EGF was chosen as the receptor-binding domain. The DNA-binding domain was presented by a sequence containing classic NLS motif [13,18,20]. The NLS domain is known to retain karyophilic properties after complexation with DNA [9-11,24] and is a oligocation, which ensures at least electrostatic interaction with negatively changed phosphate groups in DNA molecule. Comparison of proliferative properties of commercial EGF preparation and chemical conjugates of oligonucleotides with recombinant EGF confirmed our assumption that N-terminal modification of this protein does not prevent ligand—receptor inter-

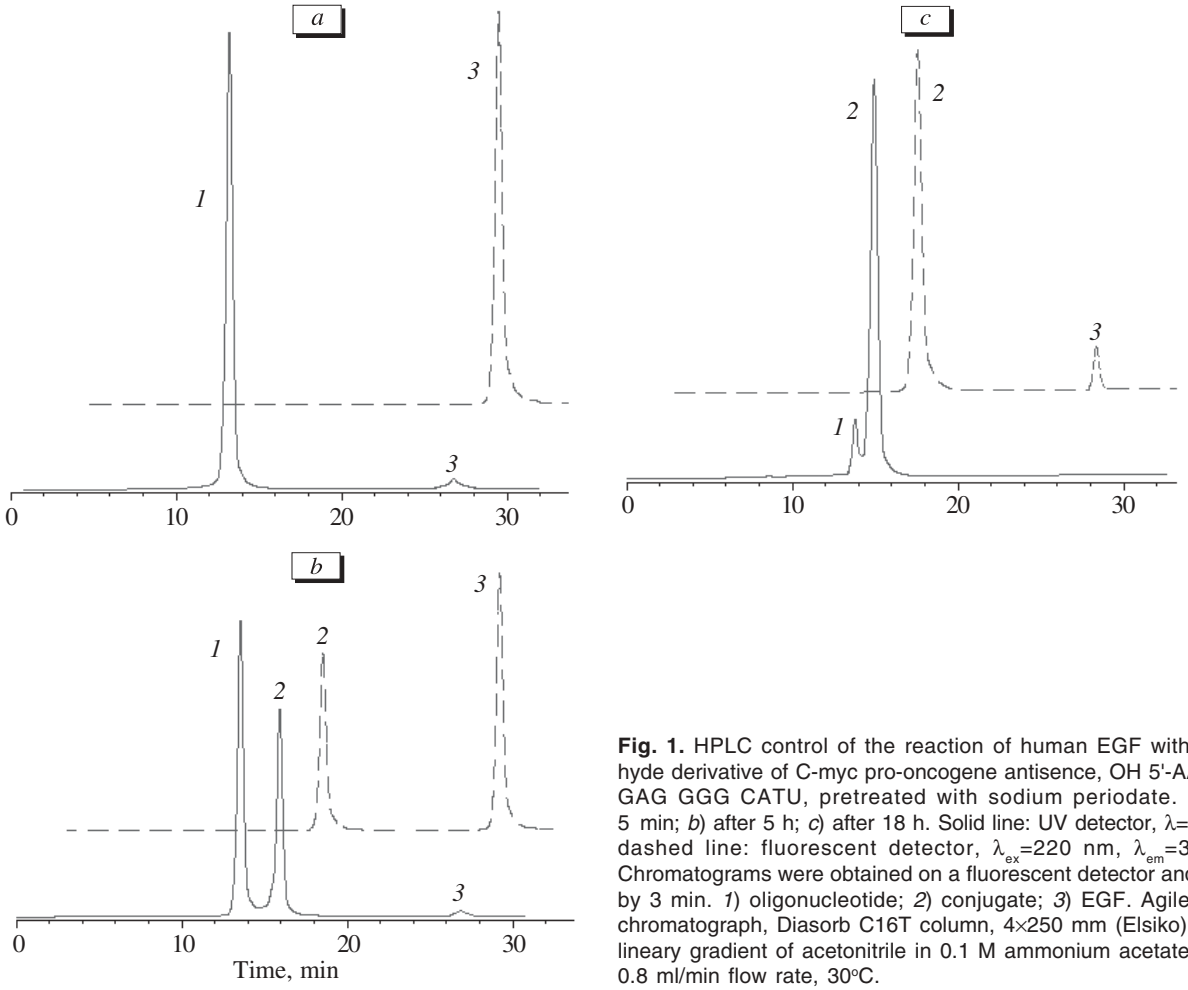

Fig. 1. HPLC control of the reaction of human EGF with dialdehyde derivative of C-myc pro-oncogene antisence, OH 5'-AAC GTT GAG GGG CATU, pretreated with sodium periodate. *a*) after 5 min; *b*) after 5 h; *c*) after 18 h. Solid line: UV detector, λ=260 nm; dashed line: fluorescent detector, λ_{ex}=220 nm, λ_{em}=380 nm. Chromatograms were obtained on a fluorescent detector and shifted by 3 min. *1*) oligonucleotide; *2*) conjugate; *3*) EGF. Agilent 1100 chromatograph, Diasorb C16T column, 4×250 mm (Elsiko); eluent: lineary gradient of acetonitrile in 0.1 M ammonium acetate (pH 7); 0.8 ml/min flow rate, 30°C.

action, which determined the choice of PGEk structure.

Stability of the transport complex is of principal importance; the bond should be stable enough for attaining the target tissue and cell, but somewhat weaker that the bond between the nucleotide and the intracellular target.

The following requirements were fulfilled during designing the protein PGFk vector:

1) to simplify maximally the protein construct due to multifunctionality of its two domains;

2) to use the largest and best characterized fragments of the hydrophobic and hydrophilic domains;

3) to minimize changes in natural sequences;

4) to provide for the possibility of intracellular cleavage of the complex with the formation of nontoxic metabolites;

5) to guarantee universal properties of the protein with respect to the nature and length of GM;

6) to determine technological parameters for obtaining and use of vector protein.

On the basis of published reports, our data, and some preliminary experiments we constructed and synthesized a vector protein PGEk consisting of 64 amino acid residues. Of these, 11 resudies corresponded to NLS motif (underlined) fused with EGF sequence: <u>KKKKRKVEDPY</u>NSDSECPLSHDGYC LHDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELR.

For creation of PGEk structure and methods of its biological synthesis and purification we carried out the following experiments. Strains of *E. coli* and *Saccharomyces cerevisiae* producing human EGF were obtained using methods of gene engineering; methods of isolation and purification of this factor were developed [7], and the identity of recombinant and natural EGF was proven. For evaluation of the position of domain junction and evaluation of the addressing properties of human EGF fragment we synthesized and studied a number of covalent conjugates of EGF with antisense oligonucleotides. For obtaining modified EGF we developed a genetic construct and expression system, obtained a *E. coli* strain (V-8389 VKPM) producing PGEk fussed with thioredoxin, and developed methods for PGEk isolation and purification [8,10,11].

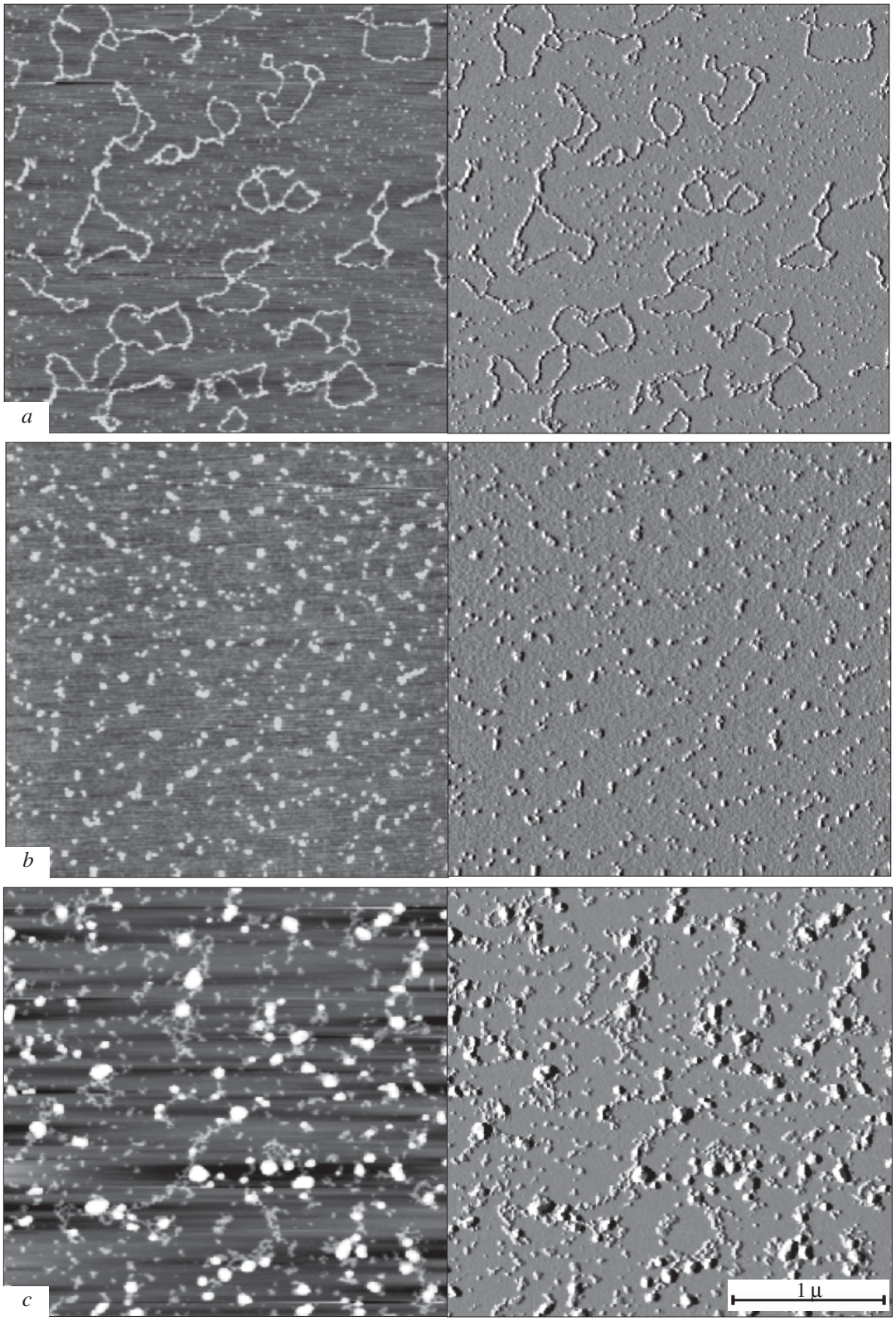

Fig. 2. Formation of complexes between PGEk protein and PEGFP-C1 plasmid. *a*) PEGFP-C1 plasmid (Clontech), 2 μg/ml; *b*) PGEk protein, 2 nM; *c*) PGEk:PEGFP-C1 mixture, 2:1. Photos were prepared using Nanoscope IIIa scanning force microscope (Digital Instruments) by I. A. Besscetnova (trapping mode, air medium). Samples dissolved in buffer (5 mM HEPES, 5 mM Mg acetate, 10 mM KCl, pH 7.5) were applied onto a mica matrix, after 2 min the matrix was washed with water and dried.

Using known antisense oligonucleotides and plasmids carrying marker gene encoding green fluorescent protein we showed that PGEk is a universal GM carrier rapidly and spontaneously forming soluble associates with DNA fragments of varying length and different nature. Complexes of oligo- and polynucleotides with PGEk are effectively internalized by target cells via receptor-mediated endocytosis, which determines selective effect on cells carrying surface EGF receptors. The presence of PGEk decreases the probability of penetration of modified membranotropic antisense oligonucleotides into non-target cells, thus protecting them from unsafe transfection [8,10,11].

Recent atomic force microscopy data directly confirmed our assumption on compactization of plasmid DNA in the presence of even small molar excess of PGEk (Fig. 2).

For more precise evaluation of the formation and structure of supramolecular complexes of recombinant PGEk protein with oligonucleotides with different conformation we used a telomeric oligomer $d(TTAGGG)_4$ (TMO) and its thiophosphoryl analogue thio-$d(TTAGGG)_4$ (TMS) characterized by high antitumor potential.

A strict correlation between the structure of the complexes and their biological properties was found. Analysis of polarized UV fluorescence and circular dichroism spectra of complexes showed that, similarly to their free state, TMO molecules form a G-quadruplex structure, while TMS molecules are presented by unfolded threads. The first two PGEk molecules bind with TMO and TMS via a non-cooperative mechanism: $K_{1TMO}=(7\pm1)\times10^7$ M^{-1} and $K_{1TMS}=(3.0\pm0.5)\times10^7$ M^{-1}. The following binding up to 6 PGEk molecules per oligonucleotide proceeded cooperatively. $K_{2TMO}=(4.0\pm1.5)\times10^6$ M^{-1} and $K_{1TMS}=(0.8\pm0.2)\times10^6$ M^{-1}.

We analyzed the effect of PGEk on the process of selective internalization of oligonucleotides by tumor cells and their subsequent translocation into the nucleoplasm. It was found that the antitumor effect of TMS considerably increased in the presence of PGEk in the found optimal proportions and was selective for the target tumor cell. IC_{50} for different cell lines varied from 470 to 90 nM. Biological properties of nontoxic TMO step-wise changed after addition of PGEk. The PGEk—TMO complexes (5:1) effectively and selectively suppressed the growth of target cells similarly to PGEk—TMS complexes.

Thus, using TMO oligonucleotide we demonstrated the possibility of protecting natural DNA structure from biodegradation and hence, the possibility of using nontoxic phosphodiester oligonucleotides for gene therapy [1,8-11,15].

The above listed characteristics of PGEk, a simple, universal, and technologically available DNA carrier protein, open prospects for the development of new effective preparations for gene therapy on this basis. Moreover, the found approaches to constructing and obtaining protein vectors and regularities of self-assembly of the transport complexes and their interaction with intracellular biopolymers provide the basis for the creation of new DNA-carrying proteins selective to other types of target cells.

The study was carried out in collaboration with laboratories of M. M. Shemyakin and Yu. A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Novosibirsk Institute of Bioorganic Chemistry, Siberian Division of Russian Academy of Sciences, All-Russian Research Center of Molecular Diagnostics and Therapy, Russian Ministry of Health, Max Plank Institute of Biophysical Chemistry (Germany), and V. A. Engelgardt Institute of Molecular Biology, Russian Academy of Sciences.

REFERENCES

1. I. A. Besschetnova, G. E. Pozmogova, A. N. Chuvilin, *et al., Mol. Biol*., **40**, No. 3, 489-545 (2006).
2. O. F. Borisova, A. K. Shchelkina, I. A. Il'icheva, and G. E. Pozmogova, *Ibid*., **35**, 860-867 (2001).
3. O. F. Borisova, A. K. Shchelkina, V. A. Kharitonov, *et al., Ibid*., **33**, No. 3, 503-511 (1999).
4. I. V. Gulina, O. A. Shul'ga, V. N. Mironov, *et al., Ibid*., **28**, No. 5, 1166-1175 (1994).
5. G. S. Zhvirblis, V. G. Gorbulev, P. M. Rubtsov, *et al., Ibid*., **22**, No. 1, 145-150 (1988).
6. M. N. Nikolaidis, A. Sh. Pasadanyan, G.E. Pozmogova, *et al., Ibid*., **28**, No. 5, 1098-1105 (1994).
7. *Patent of Russian Federation No. 21505001 (reg. 17.02.99). S. cerevisiae Yeast Strain VKM CR-349D, a Producer of Human Epidermal Growth Factor*. M. A. El'darov, G. E. Pozmogova, and S. M. Kagiyants.
8. *Patent of Russian Federation No. 2248983 (reg. 18.08.03). Peptide Vector, Method of Obtaining, Nucleotide Sequence, REcombinant Plasmid DNA, and Escherichia coli B-8389 VKPM Strain for Its Preparation, Method of Genetic Modification of Animals and Human Cells*, G. E. Pozmogova, A. N. Chuvilin, G. A. Posypanova, *et al*.
9. G. E. Pozmogova and D. G. Knorre, *Vopr. Med. Khimii*, **44**, No. 4, 331-347 (1998).
10. G. E. Pozmogova and G. A. Posypanova, *Novye Lek. Preparaty*, No. 11, 72-80 (2005).
11. G. E. Pozmogova and A. N. Chuvilin, *Ibid*., pp. 66-71.
12. G. A. Posypanova, N. N. Kireeva, V. A. Makarov, *et al., Vopr. Biol. Med. Farm. Khimii*, No. 3, 15-20 (2005).
13. D. A. Stetsenko, A. A. Arzumanov, V. A. Korshun, and M. J. Gait, *Mol. Biol*., **34**, No. 6, 998-1006 (2000).
14. I. G. Shemyakin, V. A. Anisimova, P. Kh. Kopylov, *et al., Ibid*., **31**, No. 5, 790-794 (1997).
15. I. A. Besschetnova, G. E. Pozmogova, A. K. Shchyolkina, and O. F. Borisova, *J. Biomol. Struct. Dyn*., **22**, 859-860 (2005).

16. C. Grignet-Debrus, V. Cool, N. Badson, *et al., Cancer Gene Ther*., **7**, 1456-1468 (2000).
17. M. A. Kay, C. S. Manno, M. V. Ragni, *et al., Nat. Genet*., No. 24, 257-261 (2000).
18. R. I. Mahat, O. D. Monera, L. C. Smith, and A. Rolland, *Curr. Opin. Mol. Ther*., **1**, No. 2, 226-243 (1999).
19. M. Manoharan, *Antisense Nucleic Acid Drug Dev*., **12**, No. 2, 103-128 (2002).
20. P. D. Richardson, B. T. Kren, and C. J. Steer, *Hepatology*, **35**, No. 3, 512-518 (2002).
21. A. K. Shchyolkina, O. F. Borisova, M. A. Livshits, *et al. Biochemistry*, **39**, No. 33, 10,034-10,044 (2000).
22. *A. K. Shchyolkina, O. F. Borisova, M. A. Livshits, et al. J. Biomol. Struct. Dyn*., **16**, No. 6, 1264-1265 (1999).
23. A. A. Shulga, I. V. Levichkin, F. T. Kurbanov, *et al., Nucl. Acids Res*., **22**, No. 18, 3808-3810 (1994).
24. C. Uherek and W. Wels, *Adv. Drug Delivery Rev*., **44**, Nos. 2-3, 153-166 (2000).
25. E. Wagner, M. Cotten, K. Mechtler, *et al., Bioconjug. Chem*., No. 2, 226-226 (1991).