ANALYSIS OF CONSERVED HYDROPHOBIC CORES IN PROTEINS AND SUPRAMOLECULAR COMPLEXES

ANNA KARYAGINA
Gamaleya Institute of Epidemiology and Microbiology
Gamaleya Str. 18, Moscow 123098, Russia
Institute of Agricultural Biotechnology
Timiryazevskaya Str. 42, Moscow 127550, Russia
akaryagina@rambler.ru

ANNA ERSHOVA
Department of Bioengineering and Bioinformatics
Moscow State University, Moscow, 119992, Russia
aershova@rambler.ru

MIKHAIL TITOV
Institute of Agricultural Biotechnology
Timiryazevskaya Str. 42, Moscow 127550, Russia
mlt@iab.ac.ru

IVAN OLOVNİKOV, EVGENIY AKSIANOV, ALEXANDRA RYAZANOVA, ELENA KUBAREVA, SERGEI SPIRIN and ANDREI ALEXEEVSKI
Belozersky Institute and Department of Bioengineering and Bioinformatics
Moscow State University, Moscow, 119992, Russia
aba@belozersky.msu.ru

Received 7 October 2005
Revised 24 November 2005
Accepted 9 December 2005

The conserved hydrophobic core is an important feature of a family of protein domains. We suggest a procedure for finding and the analysis of conserved hydrophobic cores. The procedure is based on using an original program called CluD (http://monkey.belozersky.msu.ru/CluD/cgi-bin/hftri.pl). Conserved hydrophobic cores of several families including homeodomains and interlock-containing domains are described. Hydrophobic clusters on some protein-DNA and protein-protein interfaces were also analyzed.

Keywords: Hydrophobic core; protein family; hydrophobic interaction.
1. Introduction

Non-polar molecules, as well as non-polar atomic groups (like –CH₃) of proteins, tend to aggregate in water to minimize the surface accessible to the polar solvent. This phenomenon is known as hydrophobic effect.

Hydrophobic effect is significant for folding and stabilization of a protein domain structure as well as for protein-protein, protein-nucleic acid and protein-ligand interactions. A promising approach in the investigation of hydrophobic effect is the detection of hydrophobic clusters in resolved 3D structures of proteins and macromolecular complexes. By definition, hydrophobic clusters are spatial areas filled mainly by non-polar atomic groups.

The detection of hydrophobic clusters in proteins is useful for comparison of proteins with similar topology, for identification of structural domains and for theoretical prediction of early folding intermediates. Analysis of hydrophobic clusters at interfaces is useful for understanding the mechanisms of intermolecular interaction. A number of algorithms for detecting hydrophobic clusters in protein structures was suggested. The approach of Swindells is implemented in a program for structural domain detection, which is included in CATH structural classification software (http://www.biochem.ucl.ac.uk/bsm/cath/cath.html). Unfortunately, no implementations of algorithms for hydrophobic clusters detection are publicly available, which prevents from their systematic testing and routine usage. Also, the mentioned algorithms are not applicable for the detection of interface clusters.

Earlier we have developed CluD (Cluster Detector), a program for the detection of spatial areas occupied by non-polar atomic groups in 3D structures of proteins and multimolecular complexes (http://monkey.belozersky.msu.ru/CluD/cgi-bin/hftri.pl).

In a structure of a protein domain, the biggest hydrophobic cluster detected by CluD is the hydrophobic core of the domain. In a family of related structures, hydrophobic cores vary in details but usually have an extended common part, the conserved hydrophobic core of the family. In a family of multimolecular complexes with a similar structure, conserved hydrophobic clusters at the interface(s) can be defined in a similar manner. We believe that conserved hydrophobic clusters are important features of structural families.

In the present work, we determined the conserved hydrophobic cores for a number of structural families (homeodomains and their complexes with DNA, lambda repressor-like HTH domains and their complexes with DNA, V-set domains). The elaborated procedure was also used for the description of hydrophobic clusters at the DNA-protein interface in complexes of HhaI C5-cytosine DNA-methyltransferase with DNA, and in capsids of icosahedral viruses.

In order to detect conserved hydrophobic clusters, a structural correspondence between related structures should be established. In the case of closely related structures this can be done by standard backbone superimposition tools. An example of establishing a structural correspondence between weakly related structures that
cannot be superimposed by existing tools, is explored in our work. It is the superimposition of sandwich-like domains by fitting so-called interlock structural motifs, which are typical for the sandwich-like domains.\(^6\)

The determination of conserved hydrophobic cores or clusters allows predicting the functional role of a number of amino acid residues for homologous proteins. We have done such predictions for several protein groups.

We suggest our procedure as a routine tool for describing structural families and annotating amino acid residues. The procedure was included into a practical course on structural bioinformatics in Moscow State University.

2. Materials and Methods

We used an original program CluD\(^5\) and its modification for multi-model files describing biological units (such as viral capsids). This program detects spatial regions occupied by non-polar atomic groups in a 3D structure. As a non-polar atomic group we consider a group of a protein or a nucleic acid that includes one carbon or sulfur atom and covalently bound hydrogen atom(s). In proteins and nucleic acids those are \(-\text{CH}_3\), \(-\text{CH}_2\text{-}\), \(-\text{CH} =\), \(-\text{SH}\), and \(-\text{S-}\) groups.

CluD is an implementation of an original algorithm based on dividing the total set of non-polar atomic groups into separate or weakly connected subsets called hydrophobic clusters (for an overview of the algorithm, see http://monkey.belozersky.msu.ru/CluD/help.html). Input data of CluD includes the analyzed structure, the maximal distance of hydrophobic interaction (5.4 \(\text{Å}\) by default, recommended values are 4.5 to 5.4 \(\text{Å}\)), and the list of the types of non-polar atomic groups. The latter has two variants: (1) all carbon and sulfur groups (the extended list), (2) only those ones that do not covalently bind to oxygen or nitrogen atoms (the strict list).

For the distance of hydrophobic interaction we consider the distance between the centers of carbon (sulfur) atoms of the interacting groups. The maximal value of it is 5.4 \(\text{Å}\), which is the doubled minimal distance between a carbon atom of a \(-\text{CH}_3\) group and a water oxygen atom in protein structures: the hydrophobic interaction occurs if a water molecule cannot be placed between two non-polar groups.

The output of CluD is a list of clusters found. There exists a web-interface to CluD (http://monkey.belozersky.msu.ru/CluD/cgi-bin/hftri.pl); the output is presented in three forms: a table, a RasMol script and a Chime (see http://www.mdl.com/products/framework/chime/) page visualizing the clusters.

Several additional modules of CluD were developed:

(1) a module identifying the clusters on the interface of two molecules (two domains); two interacting chains with residue numbers should be chosen; and
(2) a module working with so-called “biological units”, available at PDB site and arranged as a number of models. This module was used for dealing with complete virus capsids.
In the present work we used CluD with the default parameters: the strict atom list and the threshold equal to 5.4 Å.

The structures used in the work were obtained from PDB and VIPERdb (Virus Particle Explorer, http://viperdb.scripps.edu/) as files in PDB format. The multiple sequence alignments of the analyzed proteins are obtained (except for the case of the representatives of Ig-like fold) by the following procedure. A structural superimposition of the structures was obtained using SwissPDBViewer. Then the sequence alignment based on the structural superimposition was generated by SwissPDBViewer. The obtained sequence alignment was imported into GeneDoc alignment editor and manually corrected if needed.

The procedure of the conserved clusters detection is as follows:

1. aligning the sequences of the analyzed proteins;
2. detecting hydrophobic clusters in each structure, selecting clusters of interest (the hydrophobic core of a protein domain or the clusters involved in an intermolecular interaction); and
3. highlighting those alignment positions that contain residues involved in the selected clusters in all structures. The highlighted residues are considered as forming the conserved hydrophobic core (conserved hydrophobic clusters).

3. Results and Discussion

3.1. Conserved hydrophobic cores of several protein domain families

3.1.1. Homeodomains

Homeodomains are a family of ∼60 amino acid residue three-helical DNA-binding protein domains from eukaryotic transcriptional factors. Hydrophobic cores of individual homeodomains vary in details (Fig. 1, highlighted in gray). The conserved hydrophobic core of homeodomains includes atomic groups from 15 residues (Fig. 1, columns in frames).

The detected conserved hydrophobic core occupies the central part of the protein globule (Fig. 2), which is in agreement with the theoretical expectation.

Interestingly, in a number of cases the conserved hydrophobic core includes non-polar groups of bulky polar residues like Lys and Arg. Moreover, one position of the conserved hydrophobic core can contain amino acid residues with quite different properties: aliphatic and aromatic residues (Ile, Phe, etc.), polar (Thr, Gln) and even charged residues with non-polar atomic groups (Arg, Lys), see the positions marked by asterisks in Fig. 1. The same situation is observed for other analyzed families, see below (Figs. 3 and 4). Thus, for example, in some cases a substitution of Leu to Lys conserves the residue function of hydrophobic core formation.
Fig. 1. Multiple alignment of amino acid sequences for 16 homeodomains with known structure. Amino acid residues forming the hydrophobic cores are highlighted in gray. Fifteen positions that include residues contributing in the hydrophobic core in all structures (the conserved hydrophobic core) are in frames; among them are four positions, marked by asterisks below the alignment, that include residues with quite different properties. The secondary structure of homeodomains is shown above the alignment; H is for alpha-helix. The analyzed structures are: Antennapedia (9ANT, chain A), HoxB1 (1B72, chain A), Ultrabithorax (1BS1, chain A), HoxA9 (1PUF, chain A), Engrailed (1HDD, chain C), Even-skipped (1JGG, chain A), Msx-1 (1IG7, chain A), NK-2 (1NK2, model 1, chain P), Paired (1FJL, chain B), Mata1 (1LE8, chain A), Extradenticle (1BS1, chain B), Pbx1 (1PUF, chain B), Pit-1 (1AU7, chain A), Oct-1 (1GT0, chain C), and Mata2 (1YRN, chain B).

Fig. 2. Antennapedia homeodomain in complex with DNA (9ANT, chain A). DNA is shown as a wireframe model. The protein backbone is shown in the “cartoon” style. The atoms forming the hydrophobic core of the protein are shown as light gray balls and those forming the conserved hydrophobic core of homeodomains are shown as dark gray balls (a spacefill model).

It can be predicted that in all other homeodomains, the residues located in those 15 positions are involved in the hydrophobic core formation, too. An analysis of the sequence alignment of 825 homeodomains shows that in 12 of those positions more than 95% of amino acid residues can be involved in the hydrophobic core formation.
3.1.2. Lambda repressor-like HTH domains

The SCOP\textsuperscript{10,11} superfamily “lambda repressor-like HTH domains” includes HTH-containing DNA binding domains of several transcriptional regulators from bacteriophages (Lambda, 434, P22 and other) and bacteria (HTH\textsubscript{3} family according to Pfam databank), and POUs domains. The POU-specific (POUs) domains are DNA-binding domains of so-called POU proteins (these eukaryotic transcriptional factors contain also homeodomains of POU\textsubscript{h} family). HTH\textsubscript{3} domains and POUs domains have similar 3D structure consisting of four alpha-helices and could be analyzed as a single (although rather divergent) family. The conserved hydrophobic core of this family includes atom groups from 20 residues (Fig. 3).

Thus, the conserved hydrophobic core reveals a similarity between domains of a divergent structural family. This is in contrast with low (undetectable by the standard tools) sequence similarity and even with weak structural coincidence (e.g. for superimposition of 53 corresponding C\textsubscript{\alpha} atoms of Pit-1 and 434cro the value of r.m.s.d. is 2.4 Å).

3.2. Conserved hydrophobic cores in families of the sandwich-like architecture

Sandwich-like domains consist of two beta-sheets packed against each other. Protein domains of the sandwich-like architecture are highly divergent. According to SCOP

<table>
<thead>
<tr>
<th>Family</th>
<th>Protein</th>
<th>HTH-3</th>
<th>POUs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P22 repr.</td>
<td>HHHHHHHHHHHHH</td>
<td>HHHHHH</td>
</tr>
<tr>
<td></td>
<td>SinI/SinR</td>
<td>tournaments</td>
<td>tournament</td>
</tr>
<tr>
<td></td>
<td>lambda repr.</td>
<td>-EQLEDARKGTTGMGTS-</td>
<td>-GQTLSKIGRALKAG-</td>
</tr>
<tr>
<td></td>
<td>434 repr.</td>
<td>-DSSYHAKTTTALGA-</td>
<td>-MQLSL RCFAKAAG-</td>
</tr>
<tr>
<td></td>
<td>434 cro</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Pit-1</td>
<td>ALEQFANREPRYRTGGA-</td>
<td>ST QNGMA LAAV</td>
</tr>
<tr>
<td></td>
<td>Oct-1</td>
<td>ELEQFAKXGQSKRKLGA-</td>
<td>ST QNGMA LAAGL</td>
</tr>
</tbody>
</table>

Fig. 3. Multiple alignment of amino acid sequences for five HTH\textsubscript{3} domains and two POUs domains with known structure. The designations are the same as in Fig. 1. The analyzed structures are: P22 C2 repressor (1ADR, model 1), SinI/SinR (1B0N, chain A), lambda repressor (1LMB, chain 3), 434 repressor (1RPE, chain L), 434 cro (2CRO), Pit-1 (1AU7, chain A), and Oct-1 (1GT0, chain C).
classification, domains of 96 different superfamilies joined into 52 folds are described as sandwich-like. The hydrophobic core seems to be an important commonality of sandwich-like domains. Below, conserved hydrophobic cores are determined for two sets of domains: (1) representatives of the V-set domains family, which consists of domains related to the antibody variable domains; (2) representatives of the immunoglobulin-like (Ig-like) fold. At the fold level, the domains lack detectable similarity in sequence and generally cannot be structurally superimposed by standard tools. The Ig-like fold includes the V-set family.

3.2.1. V-set domains

The SCOP family “V-set domains (antibody variable domains-like)” includes 1687 structures. They have a similar topology composed of seven to nine beta-strands forming a beta-sandwich structure. We have analyzed structures of five arbitrary selected, distantly related domains of this family. The conserved hydrophobic core of these representatives includes atom groups from 23 residues (Fig. 4).

![Fig. 4. Multiple alignment of amino acid sequences for five V-set domains. The designations are the same as in Fig. 1. The analyzed structures are: murine sialoadhesin (1QFO, chain A), human virus receptor (1EAJ, chain B), murine catalytic antibody (1CT8, chain D, residues 1–113), rat myelin membrane adhesion molecule (1NEU), and murine immunoglobulin (1AFV, chain H, residues 1–120). Eight key positions of interlock motif are shown by arrows below the alignment.](image-url)
Fig. 5. Conserved hydrophobic core in murine sialoadhesin (1QFO, chain A). The protein backbone is shown in the “cartoon” style. The atoms of the conserved hydrophobic core are shown as dark gray balls.

In agreement with expectations, the conserved hydrophobic core is located between two merely parallel beta-sheets forming a sandwich architecture (Fig. 5).

Earlier in domains of the Ig-like fold eight positions were distinguished as key positions of the interlock structural motif. The conserved hydrophobic core of the analyzed V-set domains includes all these positions. This observation is explored in the next section.

3.2.2. Ig-like fold
According to SCOP, domains of the Ig-like fold are classified into 23 structural superfamilies presenting various proteins from all kingdoms of life. Structures from different superfamilies cannot be superimposed by existing superimposition tools. Despite this, all Ig-like domains have common structural features. First, they are of sandwich architecture. Second, the majority of Ig-like domains contain the interlock structural motif consisting of four beta-strands (Fig. 6). It was hypothesized that interlocks are the most conserved part of the Ig-like fold and also other folds of sandwich architecture. In all interlocks eight key positions with distinguished functional role were defined.

Hydrophobic cores were detected for 94 representatives of 15 families from the Ig-like fold. The correspondence between amino acid residues in all structures cannot be established by sequence or structural alignment. For this purpose, we used interlock and its eight key positions detected in each structure. This allows to align, at least, four strands forming the interlock, and therefore, to find the conserved hydrophobic core within the interlocks (Fig. 7, the complete alignment is
Analysis of Conserved Hydrophobic Cores in Proteins and Supramolecular Complexes

Fig. 6. The interlock structural motif. (a) The scheme of the motif structure. (b) The structure of the interlock motif from N-terminal domain of T-cell surface glycoprotein CD4 (PDB code 3CD4, residues 1–100). The protein backbone is shown as a cartoon model, the strands 1 and 2 are in dark gray, the strands 3 and 4 are in light gray; the C\(^\alpha\) atoms of the key interlock positions are shown as balls.

Fig. 7. Multiple alignment of amino acid sequences for eight representatives of the Ig-like fold. Only four strands of the interlock motif are shown. The designations are the same as in Fig. 1. The key interlock positions coincide with the conserved hydrophobic core positions. The analyzed structures are: CD4 V-set domain (3CD4, 1-102), Class II MHC alpha chain, C-terminal domain (1AQD, chain A:82-181), CD2-binding domain of CD58, second domain (1CCZ, chain A:94-171), Mucosal addressin cell adhesion molecule-1 (MADCAM-1) (1BQS, chain A:1-90), p50 subunit of NF-kappa B (1SVC, chain P:253-351), Growth hormone receptor (1AXI, chain B:35-124), beta-Glucuronidase (1BHG, chain A:232-324), N-cadherin (neural) (1NCI, chain A:1-99), Neocarzinostatin (1NOA, 1-108), and Invasin (1CWV, 503-596).

available at http://monkey.belozersky.msu.ru/gs/). The result is that the conserved hydrophobic core within the interlock is composed exclusively of amino acid residues from eight key positions. Thus, the distinguished role of amino acid residues in eight key interlock positions of Ig-like domains\(^6\) is confirmed.
3.3. Hydrophobic cores on the DNA-protein interfaces

3.3.1. Homeodomain-DNA complexes

A detailed analysis of hydrophobic clusters at DNA-protein interface in homeodomain-DNA complex structures was done previously. Five conserved interface hydrophobic clusters were detected. Each of them was observed in more than 10 from 41 analyzed structures of homeodomain-DNA complexes. In Fig. 8(a) two clusters formed by atomic groups of (i) residue 54 of homeodomain and nucleotide 200 of DNA and (ii) residue 47 of homeodomain and nucleotide 104 of DNA are shown (the numbering of sequences of homeodomain and its recognition site are as in Refs. 5, 12).

By the analysis of those clusters some “rules” of DNA recognition by homeodomains can be derived, e.g., “if Val or Ile occupy the position 47 of a homeodomain, then T is in the position 104 of the recognition site and the hydrophobic cluster 47-104 is formed in the homeodomain-DNA complex” or “if Ile or Met occupy the position 54 of homeodomain, then the position 200 of recognition DNA site is occupied by A or C and the hydrophobic cluster 54–200 is formed”. Thus the analysis of hydrophobic clusters can be useful for prediction of recognition sites of DNA-binding proteins.

3.3.2. Methyltransferase HhaI — DNA complexes

DNA-methyltransferase HhaI recognizes the sequence 5'-GCGC-3' and methylates the C5 atom of the inner cytosine of this sequence. During this reaction the target cytosine residue flips from the DNA double helix (Fig. 8(b)). Hydrophobic clusters were identified on DNA-protein interfaces in complexes with unmethylated DNA molecules are shown as wireframe models; backbone of protein is in spline style.

![Fig. 8. Hydrophobic clusters at the DNA-protein interface in complexes of homeodomain HoxB1 with DNA (a) and methyltransferase HhaI with DNA (b). The PDB codes of structures are 1B72 (a) and 5MHT (b). The atoms forming hydrophobic clusters are shown as balls of different shades of gray.](image-url)
(3MHT), hemimethylated (5MHT) and fully methylated (4MHT) DNA duplexes. According to the experimental data an affinity of M.HhaI to DNA substrate changes in the order: hemimethylated > unmethylated > fully methylated. This correlates with the number of hydrophobic clusters on DNA-protein interface: 4 (hemimethylated) > 3 (unmethylated) > 2 (fully methylated). Thus, an analysis of hydrophobic clusters at intermolecular interfaces can be useful for prediction of protein-ligand affinity in certain complexes.

3.4. Hydrophobic core structures in virus capsids

Hydrophobic interaction is known to play an important role in capsid assembly. We investigated hydrophobic clusters in the satellite virus structural superfamily. A satellite virus capsid consists of 60 equal subunits with the icosahedral symmetry (triangulation number $T=1$). The structures of capsid proteins were solved for three satellite viruses. Models of complete capsids of these viruses were downloaded from VIPERdb.$^7$

Adjusted for biological units evaluation, CluD program was used for the detection of hydrophobic clusters in complete capsids. The scheme of found clusters are presented in Figs. 9(b–d).

Single subunits of these three virus capsids are structurally related one-domain proteins of sandwich architecture. Like other sandwich-like domains, all of them have main hydrophobic core located between two beta-sheets and a few hydrophobic clusters on both external sides of “sandwich”. Despite of the definite similarity of single subunits, intersubunit hydrophobic interactions vary in many details (Fig. 9). Nevertheless, limited similarity can be found. In all three capsids main hydrophobic cores are involved in the formation of intersubunit hydrophobic clusters of capsids. Similar small clusters surround vertices of order 5 (Figs. 9(b–d); the center). A net of hydrophobic clusters connects all capsid subunits.

We believe that the developed approach can be useful for the investigation of capsid assembly.

3.5. Discussion

Comparative analysis of protein structures allows to determine conserved in the evolution features. The conservation of polypeptide chain folding is a well-known phenomenon and is extensively explored for many purposes. However, other important conserved features of protein structure can be distinguished and investigated. Among them are the protein domain hydrophobic cores and other hydrophobic clusters. Thus, special tools for detecting hydrophobic clusters in a given protein structures are needed.

Several algorithms solving this problem were suggested. The majority of them detect clusters of hydrophobic amino acid residues or compact side-chain clusters, which often consist of hydrophobic residues.$^{1-3}$ We believe that a more adequate
approach is the usage of atomic groups as “indivisible units” instead of amino acid residues. The approach was realized in the program CluD5 developed earlier in our group.

One of the advantages of the approach is that non-polar atomic groups of such polar residues as Lys or Arg are considered equally to side chain atomic groups of aliphatic or aromatic residues. Another advantage is that non-protein molecules, nucleic acids and ligands can be involved into consideration; this possibility is not covered by “aa residue” based algorithms.

In the present work we propose a method, based on the usage of CluD program, for the detection of conserved hydrophobic clusters both in protein structures (the “conserved hydrophobic core” of a domain) and on intermolecular interfaces. Of course, conserved hydrophobic clusters can be determined only if a set of related or weakly related complexes is given. We have demonstrated some possibilities of the method in a series of examples.

First, we have detected hydrophobic cores of homeodomains, lambda-repressor-like HTH domains and sandwich-like domains. Interestingly, in all three cases...
non-polar atomic groups of polar amino acids are involved into hydrophobic core formation together with non-polar amino acid residues (Figs. 1, 3, 4). Moreover, in several cases (the positions marked by asterisks in Figs. 1, 3, 4) substitutions of non-polar residues by polar ones could be considered as keeping the function of hydrophobic core formation. This can explain the fact that the conserved hydrophobic core often includes more residues than the number of conserved positions in a sequence alignment. It is in the case for homeodomains, the family with low but detectable sequence homology; even more for lambda repressor-like HTH domains, which can be aligned only by structural superimposition. Moreover, the analysis of Ig-fold representatives revealed the conserved hydrophobic core for sandwich-like structures that have different topology and cannot be aligned structurally.

We believe that detecting conserved hydrophobic cores can help in prediction of the function of several residues in proteins with unknown 3D structure.

Second, we analyzed the conservation of hydrophobic clusters on the DNA-protein interface in homeodomain-DNA and HhaI methyltransferase-DNA complexes. In the first case, a few partially conserved clusters were detected, which putatively reflects different site-specificity of certain homeodomains. Some rules for prediction of homeodomain recognition sites have been derived from this analysis. In the case of HhaI methyltransferase we have demonstrated the dependence of DNA-protein affinity on the number of interface hydrophobic clusters.

Third, we have analyzed hydrophobic clusters in such large macromolecular complexes as complete virus capsids. The role of intermolecular hydrophobic clusters in capsid assembly was confirmed. Thus, the critical for capsid assembly amino acid residues can be determined.

In summary, we believe that the method of conserved hydrophobic core detection provides a potential for a structural families characterization, functional annotation of amino acid residues, and localization and measurement of intermolecular interactions. Further development of the method, including a module for detection surface hydrophobic clusters in protein structures, are planned in our group.

Acknowledgments

This work was supported by RFBR (grants 06–04–49558, 06–07–89143 and 04–04–48714) and Ludwig Institute for Cancer Research (CRDF GAP grant RB0–1277).

References


**Anna Karyagina** received her Diploma degree in virology from Moscow State University (MSU) in 1983, Ph.D. and D.Sci. degrees in molecular biology from Institute of Agricultural Biotechnology (IAB), Moscow, Russia, in 1989 and 1998, respectively. Since 1983 till 2005, she was in IAB. Now she is a leading researcher in Gamaleya Institute of Epidemiology and Microbiology, Moscow, Russia, and a leading researcher and a head of research group in IAB.

**Anna Ershova** received her Diploma degree in genetics and selection from Timiryazev Academy of Agricultural Sciences, Moscow, Russia, in 2005. Now she is a graduate student of the Department of Bioengineering and Bioinformatics of the Moscow State University.
Mikhail Titov received his diploma degree in Computer Science from Moscow Institute of Physics and Techniques, Moscow, Russia, in 2004. Now he is a graduate student of Institute of Agricultural Biotechnology of Russian Academy of Agricultural Sciences, Moscow.

Ivan Olovnikov is a fourth-level student of the Department of Bioengineering and Bioinformatics of the Moscow State University.

Eugeny Aksianov received his Diploma degree in virology from Moscow State University (MSU), Moscow, Russia, in 2003. Now he is a graduate student of Biology Department of the MSU.

Alexandra Ryazanova is a fourth-level student of the Department of Bioengineering and Bioinformatics of the Moscow State University.
Elena Kubareva received the Diploma and Doctoral degrees, both in Chemistry of Natural Compounds, from Moscow State University (MSU), Moscow, Russia, in 1984 and 1989, respectively. From 1984 up to now she was at the laboratory of chemistry of nucleic acids in Belozersky Institute of Physico-Chemical Biology of the MSU, and currently as a leading researcher and a head of a scientific group.

Sergei Spirin received his Diploma and Doctoral degrees, both in Mathematics, from the Moscow State University (MSU), Moscow, Russia, in 1985 and 2000, respectively. From 1989 up to now he was at the laboratory of mathematical methods in biology in Belozersky Institute of Physico-Chemical Biology of the MSU, and currently as a senior researcher.

Andrei Alexeevski received his Diploma and Doctoral degrees, both in Mathematics, from the Moscow State University (MSU), Moscow, Russia, in 1970 and 1983, respectively. From 1989 up to now he was at the laboratory of mathematical methods in biology in Belozersky Institute of Physico-Chemical Biology of the MSU, and currently as a leading researcher.