

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Evaluation of the Ecological State Of Water Reservoirs Using the COI Protein of Zooplankton.

Ludmila L Frolova*, Airat R Kayumov, and Arthur M Husainov.

Kazan (Volga region) Federal University, Department of Genetics, 18 Kremlyovskaya St., 420008 Kazan, Republic of Tatarstan, Russian Federation.

ABSTRACT

Water reservoirs located in industrial areas are commonly under strong anthropogenic pressure and an efficient monitoring of their ecological state is desired. Bioindication using the zooplanktonic organisms is one of the most relevant approaches widely used for evaluation of the ecological state of waters. World widely, the water-living species are identified visually and results depend on expert's opinion and subjectivity. In this study an alternative method for organism identification using the marker protein cytochrome oxidase subunit 1 (COI) is proposed. Multiple alignments by Clustal Omega algorithm of 95 amino acid sequences of available COI proteins from zooplanktonic organisms included in the V.Sladechek's list (1973) was performed. A short unique (4-33) amino acid sequences were identified in 23 proteins of organisms normally used as bioindicators of water saprobity, a common indicator of ecological state of water body. This fact allows further developing the antibodies for direct qualification and quantification of the indicator species in the probe. We suggest that this approach represents a new technique for the ecological evaluation of ecological state of waters by using the metaproteom of water living organisms.

Keywords: COI protein, zooplankton, indicator species, saprobity, water reservoir

**Corresponding author*

INTRODUCTION

Currently the evaluation of the ecological state of the water's biocenosis is often performed using zooplankton organisms as bioindicators as has been suggested by V.Sladechek [1]. Identification of indicator species determined visually with microcopies done by zoology specialists requires extensive knowledge of morphology, sometimes as many as 25 features for each organism and depends on a researcher's subjectivity.

Recently a new method, originally named "DNA barcoding", was suggested for animal species identification [2]. DNA barcoding simplifies the recognition of small and morphologically variable species. It uses DNA sequences of variable sites of several genes, like cytochrome oxidase subunit 1 (COI) or histones. A DNA barcode is the value of a marker gene sequence that differs among species of animals. For example, for two different species it could be more than 2%, whereas in the same species it is less than 1% [3, 4]. Originally, the recognition of nucleotide sequences of COI gene was performed by its amplification in PCR and sequencing [5]. Using this approach for water organism identification is limited since it requires the isolation of each unique organism from the water as water samples generally include multiple zooplanktonic organisms.

Recent development of technologies for high-efficient sequencing as well as analysis of metagenomes and metaproteoms could simplify the qualification and quantification of the indicator species in water samples. One of the well-known approaches for the protein identification is their immunological recognition with subsequent enzyme-linked immunosorbent assay (ELISA) quantification. Because of the direct interdependence of the triplet code of the gene and encoded protein, we estimated the interspecies variability of the COI proteins. Consequently, we propose to use COI protein as a marker to identify indicator species of zooplankton. Such a method will enable a simple identification and quantification of indicator species in a water sample used to evaluate the environmental status of a water body.

MATERIALS AND METHODS

Indicator species of zooplankton used in this study were obtained from V.Sladechek's list [1]. Amino acid sequences of COI protein of indicator zooplankton species were obtained from the GenPept database (<http://www.ncbi.nlm.nih.gov/protein>). Multiple alignments of the amino acid sequences were performed using Clustal Omega online software [6]. The aligned sequences were then analyzed by Jalview software [7]. 3D-models of selected COI proteins of indicator zooplanktonic species were constructed using online SWISS-MODEL service [8], the analysis of these models in PDB format was performed using Jalview software.

RESULTS AND DISCUSSIONS

Currently, the ecological status of water bodies is widely evaluated through the saprobity coefficient which is calculated using the ratio of indicator zooplanktonic organisms in a water sample [1].

One of the recent approaches suggested for the accurate species identification is based on the analysis of nucleotide sequences of gene coding for cytochrome oxidase subunit 1 (COI) [5].

We analyzed sequences of the cytochrome oxidase subunit 1 (COI) amino acid of indicator zooplanktonic organisms which are used worldwide as bioindicators to estimate the ecological status of water bodies (Table 1). Among 614 organisms, COI protein sequences from 95 organisms were available in the GenPept database and were used in this study. These 95 species analyzed covered a saprobity weights from 1.0 (oligosaprobic clean waters) to 3.8 (polysaprobic dirty waters). Since one could expect that the described proteins belong to the most frequent and available organisms, their identification in a water sample should be enough for significant determination of the saprobity factor.

Table 1: Ecological and genetic database of indicator zooplanktonic species

No	Species	Saprobity*	Indicator weight	Accession number in the database GenPept
Phylum Nematelminthes, Classis Rotatoria, Ordo Ploimida, Familia Asplanchnidae				
1	<i>Asplanchnasieboldi</i>	o-b	1.50	AAM27169
Familia Brachionidae				
2	<i>Anuraeopsisfissa</i>	o	1.20	ADO15509
3	<i>Brachionusangularis</i>	b-a	2.50	AFQ31171
4	<i>Brachionusbidentata</i>	b	2.00	AFQ31179
5	<i>Brachionuscalyciflorus</i>	b-a	2.50	HQ336793
6	<i>Brachionusplicatilis</i>	b	2.00	AAM47413
7	<i>Brachionusquadridentatus</i>	b	2.00	ACD12386
8	<i>Brachionusrubens</i>	a	3.25	ABG67955
9	<i>Brachionusurceolaris</i>	b	2.20	ACD12394
10	<i>Keratellacochlearis</i>	o-b	1.50	AGO62846
11	<i>Keratellahiernalis</i>	o	1.50	AAP29629
12	<i>Keratellaquadrata</i>	o-b	1.55	ADO15510
13	<i>Platyiasquadricornis</i>	b	1.80	AFQ31469
Familia Cephalodella				
14	<i>Cephalodellagibba</i>	o	1.35	AAP45040
Familia Colurellidae				
15	<i>Lepadellaovalis</i>	o	1.25	AFQ31427
16	<i>Lepadellapatella</i>	-	1.25	AFQ31429
Familia Dicranophoridae				
17	<i>Dicranophorusforcipatus</i>	o-b	1.50	ABC02124
Familia Euchlanidae				
18	<i>Euchlanisdilatata</i>	o-b	1.50	AFQ31265
Familia Gastropodidae				
19	<i>Ascomorphaeacaudis</i>	o	1.30	AGO62797
Familia Lecanidae				
20	<i>Lecanebullata</i>	o-b	1.50	AFQ31310
21	<i>Lecanecornuta</i>	o-b	1.50	AFQ31341
22	<i>Lecanehamata</i>	o	1.00	AFQ31363
23	<i>Lecaneluna</i>	o-b	1.55	AFQ31386
24	<i>Lecanelunaris</i>	o-b	1.55	AFQ31394
25	<i>Lecanequadridentata</i>	o-b	1.50	AFQ31409
Familia Mytilinidae				
26	<i>Mytilinamucronata</i>	b	1.80	ABC02135
27	<i>Mytilinaventralis</i>	o	1.00	ACD12409
Familia Synchaetidae				
28	<i>Polyarthradolichoptera</i>	o	1.10	AFD23499
29	<i>Synchaetagrandis</i>	o	1.40	AFD23514
30	<i>Synchaetaoblona</i>	b	1.75	AFD23558
31	<i>Synchaetapectinata</i>	b-o	1.65	AFD23675
Familia Testudinellidae				

No	Species	Saprobity*	Indicator weight	Accession number in the database GenPept
32	<i>Testudinellapatina</i>	b	1.85	AFQ31475
Familia Trichocercidae				
33	<i>Trichocercacapucina</i>	o	1.00	AFQ31481
34	<i>Trichocercaelongata</i>	o	1.00	ABC02146
35	<i>Trichocercalongiseta</i>	o	1.20	AFQ31485
36	<i>Trichocercarattus</i>	o	1.00	ABC02147
37	<i>Trichocercatenuior</i>	o	1.10	ABC02148
Ordo Monimotrochida Familia Conochilidae				
38	<i>Conochilushippocrepis</i>	o	1.15	ABC02120
Ordo Bdelloida Familia Philodinidae				
39	<i>Dissotrochaaculeata</i>	o-b	1.60	AHA10965
40	<i>Dissotrochamacrostyla</i>	o	1.00	AHA10978
Familia Habrotrichidae				
41	<i>Habrotrichabidens</i>	o	1.00	ABV44827
Familia Philodinidae				
42	<i>Rotarianeptunia</i>	a-p	3.80	ABG46999
43	<i>Rotariarotatoria</i>	a	3.25	ACR83053
Phylum Arthropoda Classis Crustacea Ordo Daphniiformes Familia Chydoridae				
44	<i>Acroperusharpae</i>	o	1.40	ABK05909
45	<i>Alonellaexcisa</i>	o	1.20	AEL16787
46	<i>Alonellaexigua</i>	o	1.20	ABK05911
47	<i>Chydorusphaericus</i>	b	1.75	ACK76416
48	<i>Eurycercuslamellatus</i>	o	1.20	AFC87918
49	<i>Graptoleberistestudinaria</i>	o-b	1.50	ABK05922
50	<i>Pseudochydorusglobosus</i>	-	-	ABO28840
Familia Bosminidae				
51	<i>Bosminacoregoni</i>	o	0.95	AAL80020
52	<i>Bosmina longirostris</i>	o-b	1.55	ACD86050
Familia Daphniidae				
53	<i>Ceriodaphniadubia</i>	-	-	AGL10635
54	<i>Ceriodaphnialaticaudata</i>	o-b	1.60	ACD86116
55	<i>Daphnia galeata</i>	o	1.00	AED99109
56	<i>Daphnialongiremis</i>	-	-	ABS86792
57	<i>Daphnia longispina</i>	b	2.05	ABS86783
58	<i>Daphniamagna</i>	a	3.40	AEG78959
59	<i>Daphniapulex</i>	a	2.80	CAA78735
60	<i>Scapholeberismucronata</i>	b	2.00	ADQ90207
61	<i>Simocephaluserrulatus</i>	o	1.30	AHA46846
62	<i>Simocephalusvetulus</i>	o-b	1.50	AHA46828
Familia Macrothricidae				
63	<i>Ophryoxusgracilis</i>	-	-	ABK05980
Familia Moinidae				

No	Species	Saprobity*	Indicator weight	Accession number in the database GenPept
64	<i>Moinabrachiata</i>	b	2.45	AEZ36020
65	<i>Moinamacrocopa</i>	b-a	2.75	AGL10469
66	<i>Moinamicrura</i>	b	2.20	ADQ90209
Familia Sidae				
67	<i>Diaphanosomabrachyurum</i>	o-b	1.40	ADK47385
68	<i>Peniliaavirostris</i>	-	-	ACD99906
69	<i>Sida crystallina</i>	o	1.30	ABK05992
Ordo Leptodoriformes Familia Leptodoridae				
70	<i>Leptodorakindtii</i>	o-b	1.65	AEM67454
Ordo Polyphemoformes Familia Polyphemidae				
71	<i>Bythotrepheslongimanus</i>	o	1.20	AAL35829
72	<i>Polyphemuspediculus</i>	o	1.30	ADA68671
Familia Diaptomidae				
73	<i>Eudiaptomusgracilis</i>	o	1.25	ABE68707
74	<i>Eudiaptomusgraciloides</i>	o-b	1.60	ABE96892
Familia Temoridae				
75	<i>Eurytemoraaffinis</i>	-	-	ADK08466
76	<i>Eurytemoralacustris</i>	-	-	ADK08473
77	<i>Heterocopeappendiculata</i>	-	-	BAI40486
Familia Cyclopidae				
78	<i>Eucyclopsmacruroides</i>	o	1.00	AGJ74852
79	<i>Eucyclopsmacrurus</i>	o-b	1.40	AGJ74857
80	<i>Eucyclopserrulatus</i>	o	1.00	AEH50084
81	<i>Eucyclopsperatus</i>	-	-	AGJ74859
82	<i>Macrocylopsalbidus</i>	b	2.00	AGL10765
83	<i>Macrocylopsdistinctus</i>	o	1.00	AGU02292
84	<i>Tropocylopsprasinus</i>	o	1.00	AGL10838
85	<i>Acanthocylopsamericanus</i>	-	-	AGL10537
86	<i>Acanthocylopsvernalis</i>	b	1.85	AFZ64415
87	<i>Acanthocylopsrobustus</i>	-	-	AFZ64410
88	<i>Cyclopsabyssorum</i>	-	-	AGJ74811
89	<i>Cyclopsinsignis</i>	o	1.40	AGJ74816
90	<i>Cyclopskolensis</i>	-	-	ADD10348
91	<i>Cyclopsstrenuus</i>	b	2.25	AGJ74817
92	<i>Megacyclopsviridis</i>	b	1.85	AGJ74866
93	<i>Mesocylopsleuckarti</i>	o	1.25	ADQ90208
94	<i>Thermocylopscrassus</i>	-	-	AGU02291
95	<i>Thermocylopsoithonoides</i>	o	1.30	AGU02290

* x-xenosaprobity; o-oligosaprobity; b – beta-mesosaprobity; a – alpha-mesosaprobity; p – polysaprobity

Currently, species in a water sample are identified visually, and the evaluation depends on the expert's subjectivity. Since their identification by the COI gene sequence requires the manual separation of individual organisms it has a high error risk. We thought that zooplanktonic organisms could be identified by using the products of the marker genes such as COI, i.e. proteins or oligopeptides.

As an example we examined 95 available sequences of the COI proteins from indicator zooplankton species. The multiple alignments of 95 amino acid sequences of COI proteins were performed using the Clustal Omega algorithm to identify unique regions of the COI proteins (Fig 1).

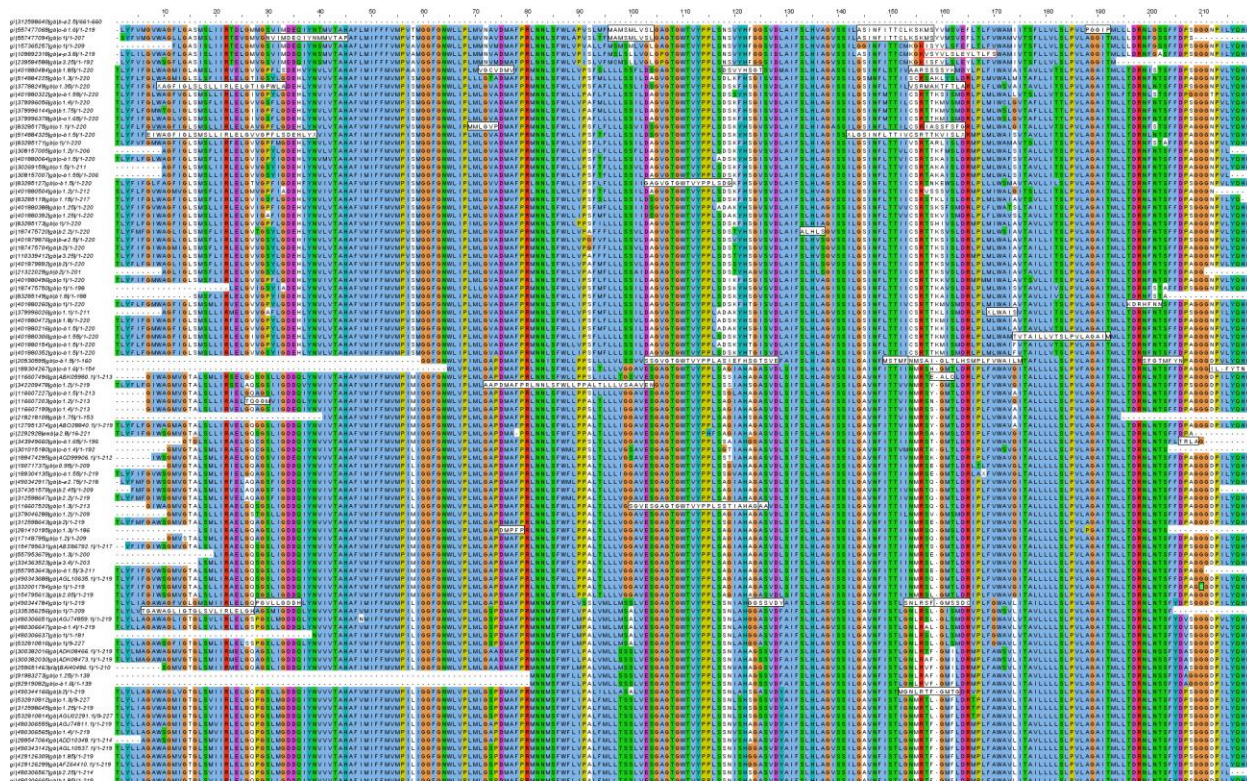


Figure 1: The multiple alignments of 95 amino acid sequences of COI proteins of indicator species

We found that 23 species have unique regions of their COI proteins consisting of 4-33 amino acids of which 6-17 are enough to be recognized by antibodies (Table 2). These indicator species cover values of all saprobities from 1.0 to 3.8. This would be a new approach for identification of these organisms by the COI protein using an immunological quantification (ELISA) of the metaproteom of water samples. We suggested new approach as a Russian patent for the evaluation of ecological state of water reservoirs using the COI protein of biosensors.

Table 2: Unique variable sequences of the COI protein of indicator zooplanktonic species

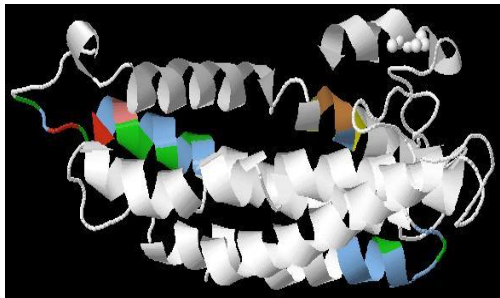
Species, Accession number in database GenPept	Saprobity*	Indicator weight	Unique fragments of amino acid sequences of the protein COI of zooplanktons with the positions in the alignment
<i>Dissotrochaaculeata</i> AHA10965	o-b	1.60	MAMSMLVSL (96-104) ASINFITTLCKSKMS (144-158) PGGIP (188-192)
<i>Dissotrochamacrostyla</i> AHA10978	o	1.00	NVIMDEQIYNMMVTAP (30-45) MAMSMLVSL (96-104) ASINFITTLCKSKMS (144-158)
<i>Rotarianeptunia</i> ABG46999	a-p	3.80	VSYVLSLEYLFLS (157-170)
<i>Testudinella patina</i> AFQ31475	b	1.85	VGCVDMV (71-77) DSVYHSGT (118-125) AARSSSYK (154-162)
<i>Cephalodellagibba</i> AAP45040	o	1.35	AGFIGLSLLIRLELGTIGPWLA (10-33) VSRMAKTFTLA (154-164)

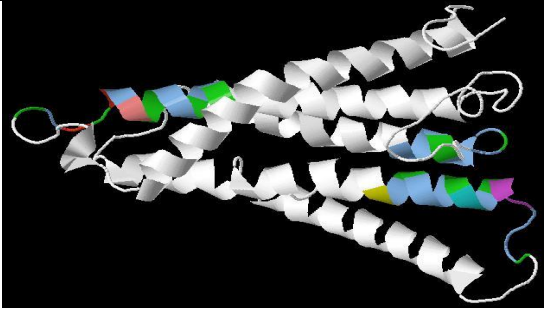
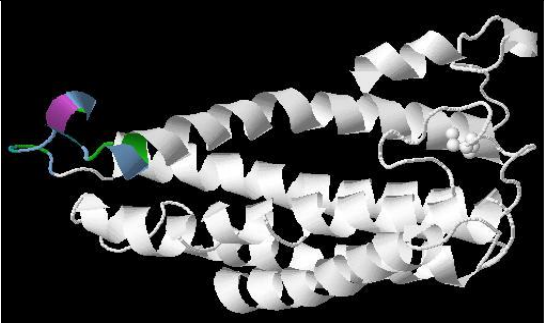
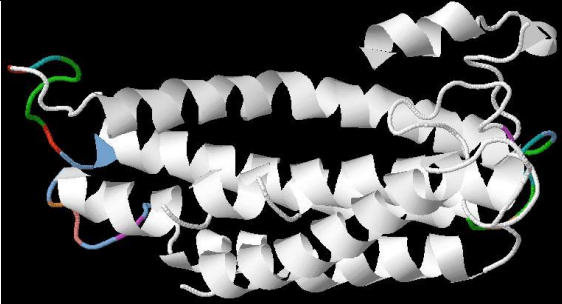
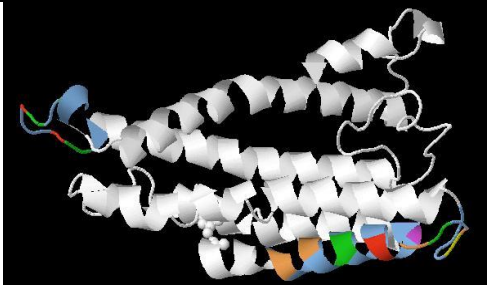
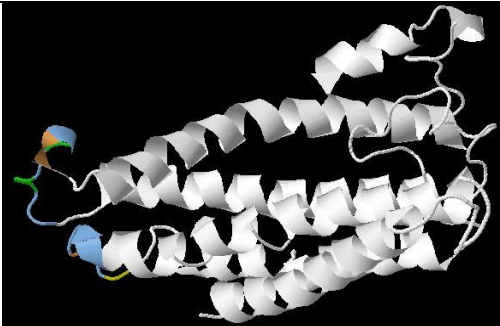
Species, Accession number in database GenPept	Saprobity*	Indicator weight	Unique fragments of amino acid sequences of the protein COI of zooplanktons with the positions in the alignment
<i>Trichocercatenuior</i> ABC02148	o	1.10	MMLGVP (69-74) IASSFSFG (157-164)
<i>Keratellacochlearis</i> AGO62846	o-b	1.50	EIWAGFIGLSMSLLIRLELGVVGPFLSDEHLYX (7-39) XLGSINFLTTIVCSRTTKVISLX (141-163)
<i>Dicranophorusforcipatus</i> ABC02124	o-b	1.50	GAGVGTGWTVYPPPLSDG (103-119)
<i>Brachionusurceolaris</i> ACD12394	b	2.20	ALHLS (133-137)
<i>Lecanehamata</i> AFQ31363	o	1.00	KDRHFNN (196-202)
<i>Polyarthradolichoptera</i> AFD23499	o	1.10	XLWAIS (160-165)
<i>Lecaneluna</i> AFQ31386	o-b	1.55	TVTAILLVTSPLVLAGAIM (174-192)
<i>Asplanchnasieboldi</i> AAM27169	o-b	1.50	SSGVGTGWTVYPPPLASIEFHSGTSVE (44-69) MSTMFNMSAIGLTLHSMPLFVWAILM (90-115) SFGTMFYN (139-146)
<i>Ceriodaphnia laticaudata</i> ACD86116	o-b	1.60	ILFYTNIY (146-153)
<i>Ophryoxus gracilis</i> ABK05980	-	-	E-ALQ (152-156)
<i>Alonella excisa</i> AEL16787	o	1.20	AAPDMAFPRLNLSFWLLPPLATLLLVSAAVEN (73-105)
<i>Alonella exigua</i> ABK05911	o	1.20	TQGGM (20-24)
<i>Leptodorakindtii</i> AEM67454	o-b	1.65	TRLA (192-195)
<i>SidacrySTALLINA</i> ABK05992	o	1.30	SGVESGAGTGWTVYPPPLSSTIAHAGAA (94-120)
<i>Polyphemus pediculus</i> ADA68671	o	1.30	DMPFP (59-63)
<i>Tropocyclops prasinus</i> AGL10838	o	1.00	PGVLLGDDH (28-36) GGSSVDY (123-129) SNLRSF-GMSSDC (153-165)
<i>Eucyclops serrulatus</i> AEH50084	o	1.00	TGAWAGLIGTGLSVLIRLELGH (12-33)
<i>Macrocylopsalbidus</i> AGL10765	b	2.00	MGNLRTFGMTG (152-162)

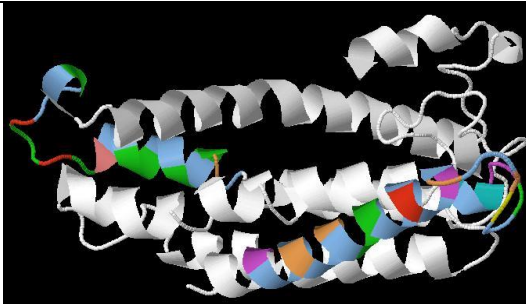
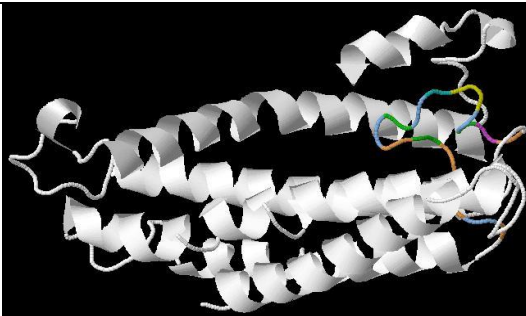
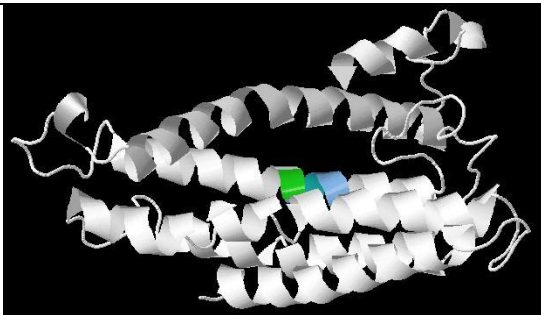

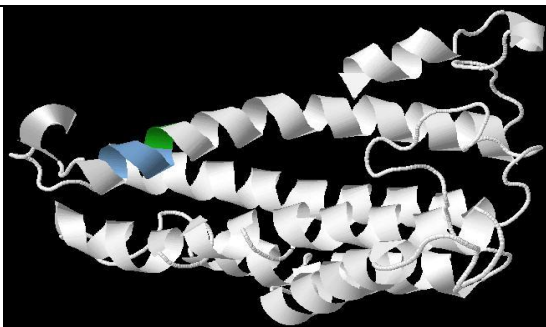
* x-xenosaprobity; o-oligosaprobity; b – beta-mesosaprobity; a – alpha-mesosaprobity; p – polysaprobity

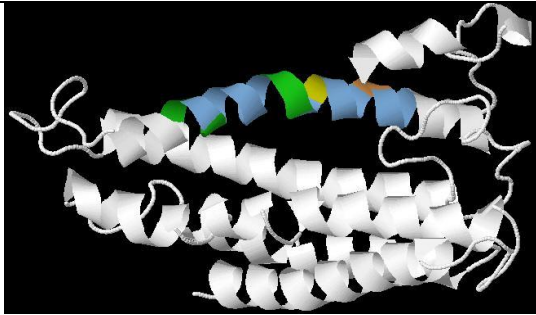
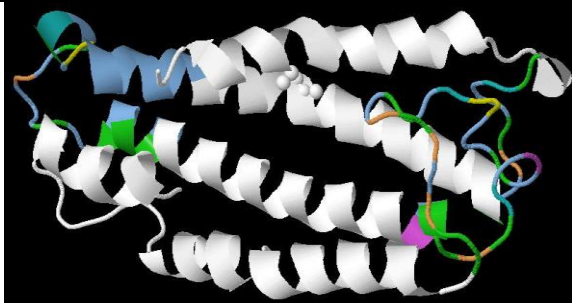
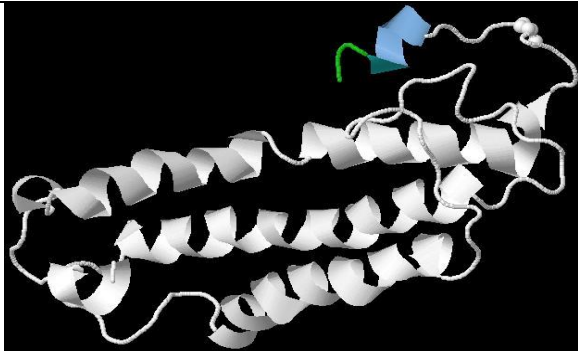
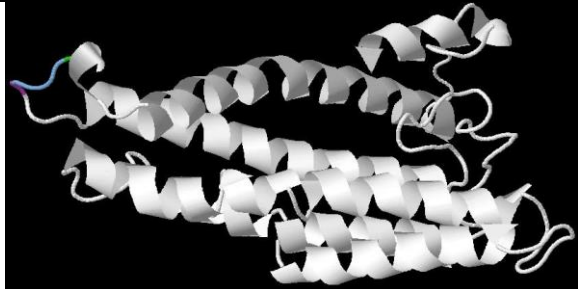
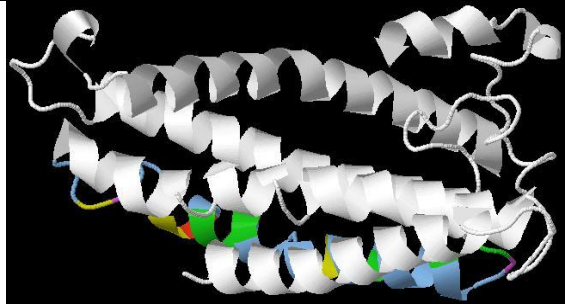
Additionally, we checked the location of these unique amino acid sequences and potential epitopes for COI proteins on the protein globule to estimate their availability for antibodies. From that, 3D-models of COI protein of indicator zooplanktonic organisms were constructed. The analysis of 3D-models shows that all sites are located on the surfaces of molecules, 26 sites of 36 are hydrophilic and located on the α -helices and therefore, more likely, could be available for antibodies (Table 3).

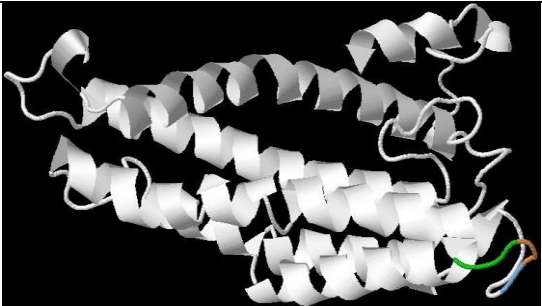
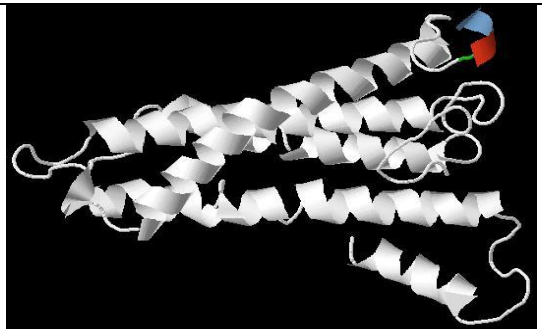
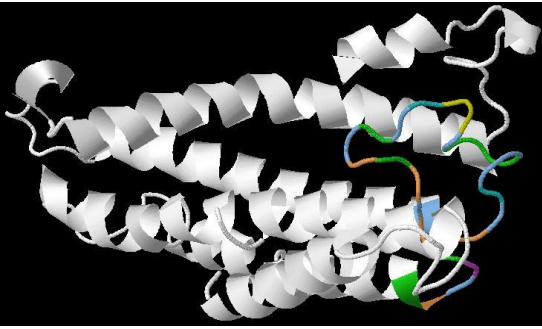
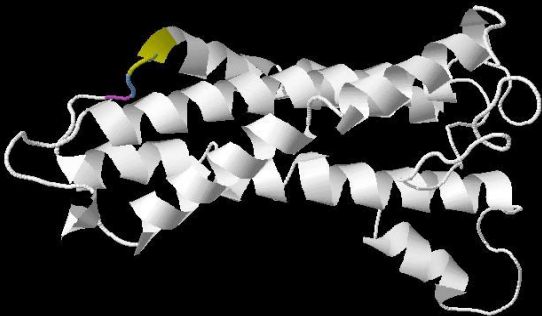
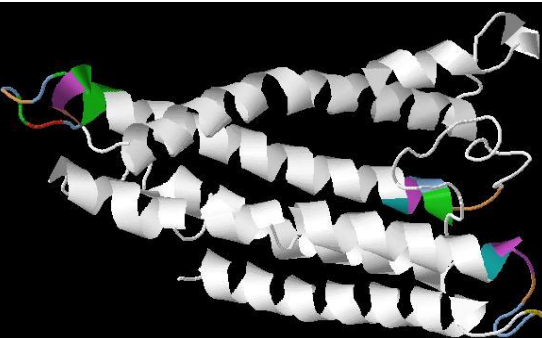
Table 3: 3D-models of amino acid sequences COI of some indicator species of zooplankton

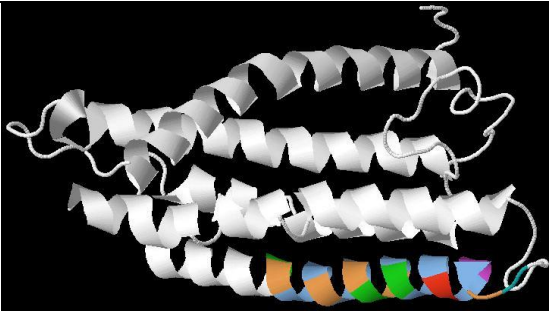

The indicator species of zooplankton / Accession number in the databank GenPept / saprobity*/ indicators weigh	3D-models of amino acid sequences COI
<i>Dissotrocha aculeate</i> AHA10965 o-b / 1.60	

The indicator species of zooplankton / Accession number in the databank GenPept / saprobity* / indicators weigh	3D-models of amino acid sequences COI
<p><i>Dissotrochamacrostyla</i> AHA10978 o / 1.00</p>	
<p><i>Rotarianeptunia</i> ABG46999 a-p / 3.80</p>	
<p><i>Testudinella patina</i> AFQ31475 b/ 1.85</p>	
<p><i>Cephalodellagibba</i> AAP45040 o/ 1.35</p>	
<p><i>Trichocercatenuior</i> ABC02148 o /1.10</p>	

The indicator species of zooplankton / Accession number in the databank GenPept / saprobity* / indicators weigh	3D-models of amino acid sequences COI
<p><i>Keratellacochlearis</i> AGO62846 o-b / 1.50</p>	
<p><i>Dicranophorusforcipatus</i> ABC02124 o-b / 1.50</p>	
<p><i>Brachionusurceolaris</i> ACD12394 b / 2.20</p>	
<p><i>Lecanehamata</i> AFQ31363 o/ 1.00</p>	
<p><i>Polyarthradolichoptera</i> AFD23499 o/ 1.10</p>	

The indicator species of zooplankton / Accession number in the databank GenPept / saprobity* / indicators weigh	3D-models of amino acid sequences COI
<p><i>Lecaneluna</i> AFQ31386 o-b / 1.55</p>	
<p><i>Asplanchnasieboldi</i> AAM27169 o-b / 1.50</p>	
<p><i>Ceriodaphnia laticaudata</i> ACD86116 o-b / 1.60</p>	
<p><i>Ophryoxus gracilis</i> ABK05980</p>	
<p><i>Alonella excisa</i> AEL16787 o / 1.20</p>	

The indicator species of zooplankton / Accession number in the databank GenPept / saprobity* / indicators weigh	3D-models of amino acid sequences COI
<p><i>Alonellaexigua</i> ABK05911 o/ 1.20</p>	
<p><i>Leptodorakindtii</i> AEM67454 o-b / 1.65</p>	
<p><i>SidacrySTALLINA</i> ABK05992 o/ 1.30</p>	
<p><i>Polyphemuspediculus</i> ADA68671 o/ 1.30</p>	
<p><i>Tropocyclopsprasinus</i> AGL10838 o / 1.00</p>	

The indicator species of zooplankton / Accession number in the databank GenPept / saprobity*/ indicators weigh	3D-models of amino acid sequences COI
<p><i>Eucyclops serrulatus</i> AEH50084 o/ 1.00</p>	
<p><i>Macrocyclus albidus</i> AGL10765 b/ 2.00</p>	

* x-xenosaprobity; o-oligosaprobity; b – beta-mesosaprobity; a – alpha-mesosaprobity; p – polysaprobity

CONCLUSION

The analysis of 95 COI proteins from zooplanktonic organisms from the V.Sladechek’s list (1973) revealed unique amino acid sequences of 23 organisms, normally used as a bioindicators of the water saprobity (the indicator of the ecological state of water biocenosis). They contain 36 unique variable amino acid sequences which are located on the protein surfaces and might be specifically recognized by antibodies. The identification of these marker proteins by ELISA should allow direct qualification and quantification of the indicator species in the probe and provide new approaches for the evaluation of ecological state of waters.

ACKNOWLEDGEMENTS

Authors gratefully acknowledge Prof. Richard W. Arnold for constructive discussion. The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University.

REFERENCES

- [1] Sladeczek V. (1973) System of water quality from the biological point of view. Arch. Hydrobiol. Ergeb. Limnol 179-191.
- [2] Hebert PD, Cywinska A, Ball SL, de Waard JR (2003) Biological identifications through DNA barcodes. Proc Roy Soc Lond B 270:313–321; doi: 10.1098/rspb.2002.2218.
- [3] Strausberger BM, Ashley MV (2001) Eggs yield nuclear DNA from egg-laying female cowbirds, their embryos and offspring. Conserv Genet 2:385–390; doi: 10.1023/A:1012526315617.
- [4] Hebert PDN, Ratnasingham S, deWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc Roy Soc Lond B 270:96– 99; doi: 10.1098/rspb.2003.0025.
- [5] Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar BiolBiotechnol 3:294–299.
- [6] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG (2011) Fast, scalable generation of high-quality protein multiple sequence



- alignments using Clustal Omega. *Molecular Systems Biology* 7:539; doi: 10.1038/msb.2011.75 (<http://www.ebi.ac.uk/Tools/msa/clustalo/>).
- [7] Waterhouse AM, Procter JB, Martin DMA, Clamp M. and Barton G J (2009) Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25 (9) 1189-1191; doi: 10.1093/bioinformatics/btp033 (<http://www.jalview.org/>).
- [8] Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, Schwede T (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research*; doi: 10.1093/nar/gku340 (<http://swissmodel.expasy.org/>).