ISSN 0026-8933, Molecular Biology, 2011, Vol. 45, No. 3, pp. 522–526. © Pleiades Publishing, Inc., 2011.
Original Russian Text © K.V. Lopatovskaya, A.V. Seliverstov, V.A. Lyubetsky, 2011, published in Molekulyarnaya Biologiya, 2011, Vol. 45, No. 3, pp. 570–574.

SHORT COMMUNICATIONS

UDC 575.852

NtcA and NtcB Regulons in Cyanobacteria and Rhodophyta Chloroplasts

K. V. Lopatovskaya, A. V. Seliverstov, and V. A. Lyubetsky

Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoi Karetny per. 19, Moscow 127994 Russia;

e-mail: kristina@iitp.ru

Received July 24, 2010; in final form, January 18, 2011

DOI: 10.1134/S0026893311030083

Keywords: NtcA factor, NtcB factor, evolution of NtcA and NtcB regulons, transcription regulation, cyanobacteria, red algae.

INTRODUCTION

Regulation of some genes involving the transcription factor NtcA has been experimentally proven only for four species of cyanobacteria: *Anabaena* sp. PCC 7120 (*Nostoc* sp. PCC 7120), *Synechocystis* sp.PCC 6803, *S.* sp. WH 8102, and *S. elongatus* PCC 7942 [1–7]. The NtcA factor of *Anabaena* sp. PCC 7120 is a dimer [1]; the DNA region covered by it is somewhat asymmetric in relation to the palindromic part of the binding site [2]; this region is longer than the previously described binding site. The consensus of the sites has been determined based on the alignment of a few nucleotide sequences that are located in regulatory regions of genes, and it is 14 bp long [3]. The role of the transcription factor NtcB has been experimentally shown in the same species and in the cyanobacteria *Leptolyngbya boryana* [8–9]. In all studied cases, the binding sites of these two factors are located close to each other and are positionally linked. NtcA activates the genes encoding the nitrogen metabolism enzymes: type I glutamine synthethase (GlnA), type III glutaminesynthase (GlnN), isocitrate dehydrogenase (Isd), ferredoxin-dependent nitrite reductase (NirA), nitrate reductase (NarB), ammonium transporter (AmtB), nitrate transporter (NrtA), urea transporter (UrtA), heterocyst-specific ABC-transporter (DevBCA), as well as the gene encoding the UreE protein associated with urea amidohydrolase, and the genes encoding transcription factors, NtcA and NtcB factors, and the signal GlnB protein belonging to the pII family [10]. Several genes encoding enzymes that do not directly participate in nitroden metabolism are also activated: ribulose-5-phosphate 3-epimerase (Rpe), ribulose bis-phosphate carboxylase (RPC enzyme participates in carbon fixation and is encoded by the *rbcSL* operon), and the genes of the bicarbonate transport system (CmpABCD).

Only two genes repressed by NtcA are currently known, *gifA* and *gifB*, which encode glutamine synthetase inactivating factors. Also, only two NtcA activated genes, *nrtA* and *nirA*, which are at the same time regulated by NtcB, are known [1–9]. The plastid genomes of three Rhodophyta species (*Cyanidium caldarium*, *Porphyra purpurea*, and *P. yezoensis*) contain the *ycf28* gene that encodes NtcA; however, the gene encoding NtcB has not been found in any of the plastid genomes.

We performed a systematic search for the binding sites of the transcriptional factors NtcA and NtcB with DNA in the genomes of cyanobacteria and chloroplast genomes of Rhodophyta. The steps of the method used are given below in italics. This method is interactive and not algorithmic.

First, we *selected* genes from the published studies [1–9], for which the regulation of expression by NtcA or NtcB was demonstrated for at least one species; the cases when the factor of the change of concentration of a particular nitrogen compound was unknown were also taken into consideration. This information was taken from experimental, and, more rarely, from bioinformatic data. In the latter case, the corresponding predictions had high confidence levels, judging by the results of the publications [1–9]. These studies reported whether the regulation was an activation or a repression. Second, the clusters of the selected orthologous cyanobacterial genes (from 51 species) were *found* using the BLASTp program. Third, 5'-leader sequences 600 bp long, which are located upstream of all of the genes belonging to these clusters, were *extracted* for further studying. Sometimes these regions included relatively short genes, e.g., *cce_1215*, located upstream of the *narB* gene. And, finally, a combined *search* for the potential binding sites and promotor boxes was *performed* in each of these leader sequences. For activation cases, only the –10 box with

a TG-extension was searched for; for repression cases, both −35 box and −10 box with a TG-extension were searched for. On odd iterations, the search was performed using the pattern that was interactively adjusted as the list of conservative sites grew; and on even iterations, the program searching for a multibox regulatory signal in the set of the aligned sequences was used [11].

The second program runs significantly longer but detects promotors more precisely [11]. Both programs use the experimental data on the influence of nucleotide substitutions in the promoter region on its effectiveness [12]. Often several candidate sequences consisting of a site and a promoter were found within a single leader sequence. In this case, the best candidates in a single leader sequence of a particular species were *chosen* using the multiple alignment of the flanking sequences of the candidate sequences located upstream of the genes of a single cluster. The alignment was performed using the MEGA4 program [13] with the following parameters: Gap Opening Penalty = 15, Gap Extension Penalty = 6.66, DNA Weight Matrix = UIB, Transition Weight = 0.5, Use Negative Matrix = OFF, and Delay Divergent Cutoff = 30. This was used in alteration with the multiple sequence alignment program on the basis of the given phylogenetic tree [14]. The programs and the accompanying methods [11, 14], including the method of the search for promotors, were also described earlier [15].

Each line of the alignment was *assigned* a characteristic ("quality" of a site 20 bp long) calculated using the algorithm from [16, 17]. We must note that the smaller this characteristic, the better is the quality of the site; this characteristic defines the distance of a site to a defined set of experimentally confirmed sites that are used as standards. The lines having this characteristic greater than 29.53 (for NtcA activation) and 21.33 (for NtcA repression) are not shown in Fig. 1 (all the tables and figures for this article are located on the web-page http://www.molecbio.com/downloads/2011/3/supp_lopat_en.pdf.). The lines having this characteristic greater than 21 for NtcB regulation are not shown in Fig. 2.

The remaining lines of the alignments in Figs. 1 and 2 are retained and their characteristics are given. The characteristics for NtcA activation are in the range from 15.76 to 29.53, and for NtcA repression from 6 to 21.33; for NtcB regulation, their values range from 8 to 21. Extreme values are rare (Figs. 1, 2). The quality ("validity") of the predicted sites can be described roughly by the characteristic: e.g., changing one position of the site in comparison to the consensus on one position may sometimes slightly increase the characteristic, however, it is not substantial for the isolation of the table lines that contain the potential binding site. For every site, the positions that differ from the most conservative positions of the consensus are given (there are six such positions), and, therefore, the number of variable positions in comparison with the most conservative part of the consensus is revealed. This value may also serve as another characteristic of the quality of a site. Its value is usually 1 and rarely 2.

The frequency distribution of nucleotides was determined using the WebLogo program based on the alignment [18]. Phylogenetic trees were constructed using the neighbor-joining method by the program MEGA 4 [13]. Genome sequences were taken from the GenBank database.

Using the method described above to study cyanobacteria, we predicted 477 potential binding sites of the NtcA transcription factor, 441 of them in the species for which no experimental data is available. Also, 44 binding sites of the transcription factor NtcB were predicted, 40 of them in the species for which no experimental data is available. The sites are located upstream of the genes from 24 clusters of orthologous genes in 51 species. Nucleotide frequency distribution was calculated for each order of Cyanobacteria and separately for the genera for which many species or strains are present in databases (*Prochlorococcus*, *Synechococcus*, *Cyanothece*, as well as the genera *Nostoc* and *Anabaena* taken together) (Fig. 3).

The number of sequences that were used for the reconstruction of the consensus of the NtcA transcriptional factor binding site has increased considerably. Now the consensus became a degenerate palindrome 20 bp long, GTA-8N-TAC being its most conservative part (Fig. 4a). The respective alignment is presented in Fig. 1. The consensus of the binding sites of the NtcB factor is TGCA-5N-TGCA (Fig. 4b); it has been determined using the multiple alignment shown in Fig. 2. Our method was applied to the following 84 clusters of orthologous genes: *aarF*, *amtB (amt1)*, *apcF*, *apcE*, *apcA*, *ccmK*, *cmpA*, *cobA*, *cobB*, *cpcB*, *cynA*, *cynB*, *cynD*, *cynS*, *devB*, *icd*, *isiB*, *isiA*, *futC*, *gifA*, *gifB*, *glnA*, *glnB (glnK)*, *glnN*, *gltS*, *gor*, *hetC*, *hetR*, *hisH*, *hupS*, *hypA2*, *hypB*, *metG*, *moaA*, *moaC*, *moeA*, *mutS*, *narB*, *narK (nrtP)*, *ndhB*, *nblA*, *nifH*, *nirA*, *nirB*, *nrtA*, *nrtC*, *ntcA*, *ntcB*, *pcbD*, *pcbA*, *petH*, *petF (fdx)*, *psaI*, *psaB*, *psaL*, *psaF*, *psbA3*, *psbZ*, *psbB*, *psbO*, *psbW*, *psbE*, *psb27*, *rbcL*, *rnc*, *rnpB*, *rpe*, *rpoD* (*sigE*, *sigB*), *som*, *speB*, *tauA*, *tauB*, *tauC*, *thrC*, *trxA*, *trxM*, *ureE*, *ureG*, *urtA*, *urtB*, *urtC*, *urtD*, *urtE*, *xisA*, and three unannotated gene clusters with unknown functions (the genes SYNW0153, SYNW2097, SYNW2456 in *Synechococcus* sp. WH 8102 [19]).

Among them, the genes *glnA*, *glnN*, *glnB*, *icd*, *amtB* (*amt1*), *gifA*, *gifB*, *ntcA*, *nirA*, *nirB*, *narB*, *narK* (*nrtP*), *ntcB*, *nrtA*, *urtA*, *cynA*, *speB*, *mutS*, *rnc*, *apcF*, *som*, *psaI*, *petF* (*fdx*), and *hupS* are potentially NtcA-regulated. The information is given in detail in Fig. 1 and Table 1. The coordinates of these genes is presented in Fig. 1. Among the sequenced Rhodophyta plastid genomes, only the *glnB* gene is predicted to be regulated by the NtcA factor (see below). The *ycf28* gene is orthologous to the gene encoding the bacterial transcription factor NtcA belonging to the Crp family, which is characterized by the presence of the conservative PF00325 domain (Table 2, built using the Pfam database).

Although the homologous *ycf28* genes are present in all Rhodophyta plastid genomes (as of 2010), the PF00325 is retained only in the species whose plastid genomes contain the *glnB* gene associated with the *rps20* gene. These are the species *Cyanidium caldarium* (with the pseudogene *rps20*), *Porphyra purpurea*, and *P. yezoensis*.

The total composition of the NtcA regulon has not been determined, since the complete set of the operons that include the found NtcA-regulated genes is not always evident; the sequences of some genes belonging to such operons may even be unknown. All the NtcA regulated genes found by us have the binding site directly in the 600 bp 5'-leader region; other genes of the same operons are also regulated by this factor.

Nevertheless, we can make some conclusions on the evolution of the NtcA regulon. It evolves differently in various orders of Cyanobacteria. The regulation of the *glnA*, *nirA*, *glnB*, *ntcA*, and *amtB* is most often observed within the Chroococcales, Gloeobacterales, Nostocales, Oscillatoriales, and Prochlorales; that of the *ntcB* gene is in Chroococcales, Gloeobacterales, Nostocales, and Oscillatoriales; that of the *gifA* and *gifB* genes is in Chroococcales, Nostocales, and Oscillatoriales; that of the *glnN* gene is in Chroococcales and Gloeobacterales; that of the *nrtA* and *icd* genes is in Chroococcales and Oscillatoriales; and that of the *narB* gene is in Chroococcales only. Weakly conservative NtcA binding sites have been predicted upstream of the genes *gor*, *petH*, *rpe*, *cmpABCD*, *ureEFG*, and *urtA* in several species belonging to various taxonomical groups. The composition of the regulon sometimes differs significantly even in the closely related species. The absence of *narB* regulation in several species of *Chroococcales* is caused by its inclusion into a long operon. The genes of urea transport are also often located in a long operon. We can suggest that the sites upstream of these genes arose evolutionary recently after a chromosomal rearrangement that changed the composition of the operons. In the species of the genus *Synechococcus*, the *icd* gene is never included in the NtcA regulon, and the *amtB* gene is included in a small number of species.

The sites are rarely found upstream of the genes *cynA*, *devBCA*, *rbcL*, *xisA*, *sigE*, *nblA*, *nrtC*, and *rnpB* in a few cyanobacterial species. They are not conservative even in closely related species. Therefore, although there is experimental evidence that the genes *cynA* and *rbcL* are NtcA regulated [1–7], the sites upstream of them apparently emerged relatively recently; this was enabled by the short length of the conservative part of the site. The sites upstream of the *devBCA* operon, which were predicted for *Acaryochloris marina* and *Cylindrospermopsis raciborskii*, evidently define the regulation of heterocysts, i.e., of trichome cells that fix atmospheric nitrogen. They probably emerged evolutionary recently.

We have significantly widened the bioinformatic description of the NtcA regulon content, which includes the genes *apcF*, *som*, *psaI*, and *petF (fdx)* in many species [8]. This regulon includes the gene *apcF* (phycobilisome subunit) in the orders Chroococcales, Nostocales, and Oscillatoriales; the gene *som* (porin) in the orders Chroococcales, Nostocales, and Prochlorales; the gene *psaI* (photosystem I reaction center) in the order Prochlorales; and the gene *petF* (ferredoxin) in the orders Chroococcales, Prochlorales, Oscillatoriales, and Nostocales. On the contrary, in contrast to the hypothesis proposed in [8], regulatory sites upstream of the genes *psaI* and *petF* are rare and very divergent even in the closely related species (Fig. 1).

We have compared the previously published phylogenetic tree of the cyanobacterial species [20] with our tree for the NtcA factor (Fig. 5). The absence of a site in a certain species is correlated with the changes of the factor in these species, if other species of the given genus have this site. The species *Cyanothece* sp. PCC 7425, *Nostoc azollae* (*A. azollae*), and *Trichodesmium erythraeum* are located far from their phylogenetic relatives on the NtcA factor tree, and their regulatory sites differ significantly from the GTA-8N-TAC consensus or are absent. For example, in *Cyanothece* sp. PCC 7425, the site upstream of the *narB* gene significantly differs from the consensus, and the sites upstream of the genes *glnB*, *icd*, *nrtA*, *petF*, and *psaI* are absent. In *Nostoc azollae*, the sites upstream of the genes *glnA*, *icd*, *ntcA*, *ntcB*, *apcF* significantly differ from the consensus, and the sites upstream of the genes *glnB*, *narB*, *nrtA*, *psaI*, and *som* are absent, while in other Nostocales the sites upstream of these genes are close to the consensus. In *Trichodesmium erythraeum*, the sites upstream of the genes *nirA* and *ntcB* significantly differ from the consensus, and the sites upstream of the genes *narB*, *ntcA*, *nrtA*, *psaI*, and *apcF* are absent. These exceptional cases, as well as the generally good accordance of the NtcA factor tree with the data on the presence of the binding site, confirm our predictions on the presence of regulatory sites.

From the obtained data, we have selected the cases of the presence of the gene and the absence of the site (e.g., in the genes *icd*, *glnB*, *amtB*, *nrtA*, *narB*, *narK*, *ntcA*, *ntcB*, *nirA*, *nirB*, *apcF*, *petF*, *som*, and *psaI*), and the cases when the site is present upstream of the gene in all species (e.g., in the genes *glnN, glnA, gifA*, and *gifB*). We have also selected the genes that are usually present but are not regulated by the NtcA factor (Table 3). Such situations are observed, e.g., for the *rbcL* gene, which is not associated directly with nitrogen metabolism, and so the absence of the site upstream of it does not seem unexpected.

In the 5'-leader region of the *glnB* gene of the chloroplasts of Rhodophyta, we have found conservative regions with the consensuses GTATyATA and TTAAAnnAAAAnAA, which are adjacent to each other or are separated by three nucleotides (their complementary regions are shown in Fig. 6). The trinucleotide GTA, which is the most conservative part of the

binding site of the NtcA factor in cyanobacteria, is located there. We suggest that these regions serve as the NtcA binding sites in chloroplasts. In contrast to the bacterial consensus, these regions do not form a palindrome. Therefore, the difference between chloroplasts and cyanobacteria is in the fact that, in the chloroplasts, the sites include only one GTA triplet, while, in cyanobacteria, there are two GTA sites: on the direct strand (the first arm of the palindrome) and on the complementary strand (the second arm of the palindrome), which results in strong interaction between the NtcA dimer and DNA. Therefore, bacterial and chloroplast consensuses of the NtcA factor differ significantly.

Above, we have noted the correlation between the presence of a PF00325 domain in the NtcA protein and the presence of an NtcA binding site; in this case, the site is located near the *rps20* gene. We propose a hypotheses that can explain this correlation. Putative NtcA binding sites in the 5'-leader region of the *glnB* gene overlap the promotor upstream of the divergently located *rps20* gene (in Fig. 6, these sites are shown on the *rps20* strand). Despite the conservatism of the promotor, the position of the detected regions relative to its boxes is not constant, and so the conservatism of these regions cannot be explained only by the conservatism of the promotor itself. It is essential that no bacterial σ^{70}-promotor can be found between these regions to the beginning of the *glnB* gene, and there is no evidence of any promotors of other types. No promotor boxes for the *glnB* have been detected, but their absence could be compensated by a protein activating NtcA.

At the same time, we could assume that there is protein–DNA regulation of the *rps20* gene together with the competition of RNA polymerases, which move to each other on different DNA strands from the *rps20* promotor specified above and some other promotor in the 3'-region of the NtcA site. Regulation takes place in one of two situations: the NtcA factor is bound or not bound with its site. Our model of the competition of RNA polymerases [21] is based on the following rule: when both RNA polymerases meet, they dissociate from the DNA and cancel transcription; if an RNA polymerase and an NtcA factor meet, the RNA polymerase continues transcription and the factor dissociates from DNA. Thus, the NtcA factor acts as the potential transcription activator of the *glnB* gene, and at the same time as the transcriptional repressor of the *rps20* gene on the complementary DNA strand, by diminishing the competition of RNA polymerases transcribing the locus in the opposite directions.

Both types of regulation are in accordance with each other: in *Synechococcus* sp. PCC 7002, the NtcA factor activates the transcription of the *glnB* gene, while the *rps20* gene is located far from this gene, so that no competition occurs. On the contrary, in the red algae, the *rps20* gene is probably repressed by the NtcA, thereby activating the *glnB* gene. We should note that both these genes are absent in Streptophyta.

We have found NtcB binding sites upstream of the *nrtA* and *nirA* genes in many species (Fig. 2; the distribution of nucleotides is presented in Fig. 4b). The NtcB regulon includes the *nrtA* gene in Chroococcales and the *nirA* gene in Chroococcales, Nostocales, Oscillatoriales, and Gloeobacterales (Table 4). No NtcB binding sites have been found in Rhodophyta plastid genomes.

Thus, a great number of species has been studied in connection with the regulatory activity of the NtcA and NtcB transcription factors upstream of many genes. The evolution of NtcA and NtcB regulons has been predicted for the studied species. It has been demonstrated that the NtcA factor represses the transcription of the *gifA* and *gifB* genes and activated the transcription of other genes, specifically the *apcF* and *som* genes, in cyanobacteria. The hypothesis [8] on the essential role of the NtcA in the regulation of the genes of the nitrogen fixation pathway (*rbcL*, *rpe*, *cmpABCD*) and the photosystems seems improbable for the majority of cyanobacteria. The *petH* gene encoding the ferredoxin–NADP reductase may serve as an example: the NtcA regulatory regions upstream of this gene seem to be absent in all studied species. A hypothesis on the mechanism of the regulation of the *glnB* gene transcription in Rhodophyta has been proposed, according to which the *glnB* gene is always included in the NtcA regulon.

ACKNOWLEDGMENTS

This study was supported by the International Scientific-Technical Center (3807) and the Ministry of Education and Science of the Russian Federation (P2370 and 14.740.11.0624).

REFERENCES

1. Alfonso M., Perewoska I., Kirilovsky D. 2001. Redox control of *ntcA* gene expression in *Synechocystis* sp. PCC 6803. Nitrogen availability and electron transport regulate the levels of the NtcA protein. *Plant Physiol.* **125**, 969–981.
2. Frias J.E., Flores E., Herrero A. 2000. Activation of the *Anabaena nir* operon promoter requires both NtcA (CAP family) and NtcB (LysR family) transcription factors. *Mol. Microbiol.* **38**, 613–625.
3. Muro-Pastor M.I., Florencio F.J. 2003. Regulation of ammonium assimilation in cyanobacteria. *Plant Physiol. Biochem.* **41**, 595–603.
4. Bird C., Wyman M. 2003. Nitrate/nitrite assimilation system of the marine picoplanktonic cyanobacterium *Synechococcus* sp. strain WH 8103: Effect of nitrogen source and availability on gene expression. *Appl. Env. Microbiol.* **69**, 7009–7018.
5. Aldehni M.F., Forchhammer K. 2006. Analysis of a non-canonical NtcA-dependent promoter in *Synecho-*

coccus elongatus and its regulation by NtcA and PII. *Arch. Microbiol.* **184**, 378–386.

6. Garcia-Dominguez M., Reyes J.C., Florencio F.J. 2000. NtcA represses transcription of *gifA* and *gifB*, genes that encode inhibitors of glutamine synthetase type I from *Synechocystis* sp. PCC 6803. *Mol. Microbiol.* **35**, 1192–1201.
7. Su Z., Olman V., Mao F., Xu Y. 2005. Comparative genomics analysis of NtcA regulons in cyanobacteria: Regulation of nitrogen assimilation and its coupling to photosynthesis. *Nucleic Acids Res.* **33**, 5156–5171.
8. Maeda S.-I., Kawaguchi Y., Ohe T.-A., Omata T. 1998. *cis*-Acting sequences required for NtcB-dependent, nitrite-responsive positive regulation of the nitrate assimilation operon in the cyanobacterium *Synechococcus* sp. strain PCC 7942. *J. Bacteriol.* **180**, 4080–4088.
9. Aichi M., Takatani N., Omata T. 2001. Role of NtcB in activation of nitrate assimilation genes in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* **183**, 5840–5847.
10. Forchhammer K. 2008. PII signal transducers: Novel functional and structural insights. *Trends Microbiol.* **16**, 65–72.
11. http://lab6.iitp.ru/en/twobox/.
12. Homann A., Link G. 2003. DNA-binding and transcription characteristics of three cloned sigma factors from mustard (*Sinapis alba* L.) suggest overlapping and distinct roles in plastid gene expression. *Eur. J. Biochem.* **270**, 1288–1300.
13. Kumar S., Dudley J., Nei M., Tamura K. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief. Bioinform.* **9**, 299–306.
14. http://lab6.iitp.ru/en/treeal/.
15. Lyubetsky V.A., Rubanov L.I., Seliverstov A.V. 2010. Lack of conservation of bacterial type promoters in plastids of Streptophyta. *Biol. Direct.* **5**, 1–32.
16. Lopatovskaya K.V., Zverkov O.A., Seliverstov A.V., Lyubetsky V.A. 2009. Transcription of proline synthesis genes in bacteria of the genera *Marinobacter, Pseudomonas*, and *Shewanella* is regulated by a protein of the TetR family. *Tr. Konf. "Informatsionnye tekhnologii i sistemy"* (Proc. Conf. "Information Technologies and Systems"), Moscow, pp. 278–281.
17. Lopatovskaya K.V., Gorbunov K.Yu., Rusin L.Yu., Seliverstov A.V., Lyubetsky V.A. 2010. Evolution of transcriptional regulation of proline synthesis in gamma-proteobacteria. *Vestn. Mosk. Gos. Univ. Ser. Biol.* **4**, 92–94.
18. Crooks G.E., Hon G., Chandonia J.M, Brenner S.E. 2004. WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190.
19. Su Z., Mao F., Dam P., Wu H., Olman V., Paulsen I.T., Palenik B., Xu Y. 2006. Computational inference and experimental validation of the nitrogen assimilation regulatory network in cyanobacterium *Synechococcus* sp. WH 8102. *Nucleic Acids Res.* **34**, 1050–1065.
20. http://www.ncbi.nlm.nih.gov/genomes/MICROBES/microbial_taxtree.html.
21. Lyubetsky V.A., Zvekov O.A., Rubanov L.I., Seliverstov A.V. Modeling RNA polymerase competition: The effect of σ-subunit knockout and temperature on gene transcription level. *Biol. Direct.* (in press).