

# Регулоны факторов NtcA и NtcB у цианобактерий и багряннок

К.В. Лопатовская, А.В. Селиверстов, В.А. Любецкий

ИППИ РАН

lyubetsk@iitp.ru

## Аннотация

Рассмотрено большое количество видов цианобактерий и хлоропластов *Rhodophyta* в связи с регуляторной активностью факторов *NtcA* и *NtcB* перед различными генами, большей частью связанными с метаболизмом азота. Это позволило предсказать много новых сайтов связывания этих факторов, существенно уточнить консенсус сайтов связывания, предсказать эволюцию *NtcA*- и *NtcB*-регулонов и опровергнуть гипотезу о существенной роли у подавляющего большинства видов *NtcA*-регуляции генов, вовлеченных в фиксацию углерода, и генов фотосистем.

## 1. Введение

Регуляция с участием транскрипционного фактора *NtcA* экспериментально показана только у четырех видов: *Nostoc* sp. PCC 7120, *Synechocystis* sp. PCC 6803 и *Synechococcus* sp. WH 8102 и *Synechococcus elongatus* PCC 7942, соответствующие ссылки приведены в [1-8]. У *Nostoc* sp. PCC 7120 фактор *NtcA* является димером [3], и закрываемая им область на ДНК несколько асимметрична [2] и шире, чем наиболее консервативная часть соответствующего сайта связывания, полученного в [4] с помощью выравнивания небольшого числа регуляторных областей. Сайты связывания фактора *NtcB* описаны перед генами *nrtA* и *nirA* у *Synechococcus* sp. PCC 7942, у *Synechocystis* sp. PCC 6803, у *Nostoc* sp. PCC 7120 и *Leptolyngbya boryana* [2]. В этих случаях сайты связывания *NtcB* позиционно связаны с сайтами связывания *NtcA*. Нами проведен систематический поиск сайтов связывания с ДНК транскрипционных факторов *NtcA* (*Ycf28*) и *NtcB* вблизи промоторов в геномах цианобактерий и хлоропластов *Rhodophyta* перед всеми генами, которые имеют хотя бы в одном виде экспериментально подтвержденную *NtcA*- или *NtcB*-регуляцию. Это – гены *glnA*, *glnN*, *glnB* (*glnK*), *icd*, *amtB* (*amt1*), *rbcL*, *gifA*, *gifB*, *gor*, *petH*, *ntcA*, *devB*, *xisA*, *hetC*, *nirA*, *nirB*, *narB*, *narK* (*nrtP*), *ntcB* и *nrtA* [1-8]. А также – перед многими генами, упомянутыми в [8], у которых эта регуляция экспериментально не подтверждена. У четырех перечисленных выше

видов гены, активируемые *NtcA*, кодируют ферменты метаболизма азота: глутаминсинтетаза первого типа *GlnA*; глутаминсинтетаза третьего типа *GlnN*; НАД-зависимая изоцитратдегидрогеназа *Isd*; ферредоксин-зависимая нитритредуктаза *NirA*; нитратредуктаза *NarB*; транспортер аммония *AmtB*; транспортер нитрата *NrtA*; транспортер карбамида *UrtA*; белок, связанный с карбамид-амидогидролазой *UreE*; специфичный для гетероцист ABC-транспортер *DevBCA*; транскрипционные факторы *NtcA* и *NtcB*; сигнальный белок из семейства pII *GlnB*; а также несколько ферментов, непосредственно не связанных с метаболизмом азота, включая рибулозо-5-фосфат-3-эпимеразу *Rpe*, рибулозо-дифосфат карбоксилазу, фермент, вовлеченный в фиксацию углерода *RbcLS*, система транспорта бикарбоната *CmpABCD*. Гены *gifA* и *gifB*, репрессируемые *NtcA*, кодируют белки, которые осуществляют регуляторные белок-белковые взаимодействия.

## 2. Материалы и методы

Было исследовано более 50 геномов бактерий, относящихся к типу цианобактерии. Геномы брались из базы данных ГенБанк [9]. Поиск гомологов в базах данных проводился с помощью программы BLAST [10]. Сайты связывания перед указанными выше и некоторыми другими генами искались с помощью множественного выравнивания, которое строилось программой MEGA4 и оригинальными программами [11-12]. Полученные для случаев *NtcA* и *NtcB* выравнивания приведены на рис. 1 в форме частотного распределения нуклеотидов по позициям, которое получено программой WebLogo [13].

## 3. Результаты и обсуждение

### 3.1. Эволюция *NtcA*- и *NtcB*-регулонов у цианобактерий

Нами впервые предсказана *NtcA*-регуляция у большинства видов цианобактерий и *Rhodophyta*. А также – существенно расширен консенсус сайта

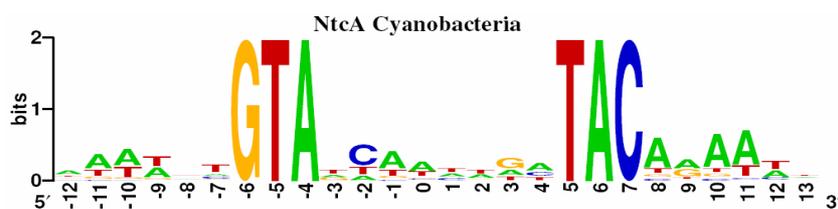
связывания с ДНК фактора NtcA, это – вырожденный палиндром с длиной 24, наиболее консервативная часть которого имеет вид GTA-8-TAC, рис. 1а. На этой основе нами изучена эволюция NtcA-регулона: у цианобактерий он значительно меняется между порядками. Наиболее широко распространена регуляция генов *glnA*, *nirA*, *glnB*, *ntcA*, *amtB* (*amt1*) у представителей Chroococcales, Gloeobacterales, Nostocales, Oscillatoriales, Prochlorales; гена *ntcB* – у Chroococcales, Gloeobacterales, Nostocales, Oscillatoriales; генов *gifA*, *gifB* – у Chroococcales, Nostocales, Oscillatoriales; *glnN* – у Chroococcales, Gloeobacterales; *nrtA*, *icd* – у Chroococcales, Oscillatoriales; *narB* – только у Chroococcales; *hetC* – только у Nostocales. Слабо консервативный сайт связывания NtcA предсказан у нескольких видов из разных таксономических групп перед генами *gor*, *petH*, *rpe*, *cmpABCD*, *ureEFG*, *urtA*. Состав регулона иногда значительно отличается даже у близких видов. Например, исчезновение регуляции гена *narB* у некоторых Chroococcales связано с включением его в длинный оперон, в состав которого часто входят и гены транспорта мочевины. Вероятно, перед этими генами сайты возникли эволюционно недавно, после перестройки хромосом, изменившей состав оперонов. У видов рода *Synechococcus* ген *icd* никогда не входит в NtcA-регулон, а ген *amtB* (*amt1*) входит в его состав у небольшого числа видов.

Перед генами *synA*, *devBCA*, *rbcl*, *xisA* сайты не консервативны даже у близких видов и изредка встречаются у единичных видов цианобактерий. Таким образом, хотя имеются экспериментальные указания на NtcA-регуляцию генов *synA* и *rbcl* [3, 5, 8], возникновение сайтов перед ними

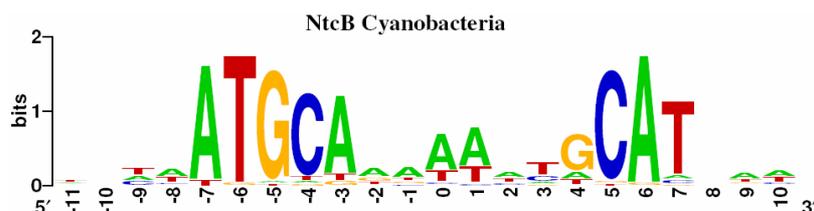
представляются случайными событиями, которые оказались возможными вследствие небольшой длины консервативной части сайта. Сайты перед опероном *devBAC* предсказаны у *Acaryochloris marina* и *Cylindrospermopsis raciborskii* и связаны с регуляцией гетероцист, связывающих атмосферный азот клеток таллома. Вероятно, они возникли эволюционно недавно.

Существенно расширено биоинформатическое предсказание в [8] регуляции генов *apcF* и *som* у цианобактерий. Для гена *apcF* (субъединица фикобилисомы) оно расширено на Chroococcales, Nostocales и Oscillatoriales, а для гена *som* (порин) – на Chroococcales, Nostocales и Prochlorales. Напротив, предсказание в [8] регуляции генов *psal* (первой фотосистемы) и *petF* не нашло подтверждения: регуляторные сайты встречаются редко и мозаично, заметно меняются даже у близких видов.

У Rhodophyta пластомы *Cyanidium caldarium*, *Porphyra purpurea* и *P. yezoensis* содержат гены *glnB* и *ntcA* (*ycf28*). В 5'-лидерной области гена *glnB* найдены консервативные участки с консенсусами GTATyATA или TTAAnnAAAAnAA, которые примыкают друг к другу или разделены тремя нуклеотидами. Предположительно они служат сайтами связывания транскрипционного фактора NtcA. Найденная пара консервативных участков близко соответствует более короткому консенсусу GTA-8-TAC для сайтов связывания фактора NtcA у *Synechococcus* sp. PCC 7002. Этому короткому консенсусу соответствуют сайты GTA-9-TAA у *Porphyra* и GTA-6-TAA у *Cyanidium*, которые предположительно служат для связывания фактора NtcA.



**Рисунок 1а. Распределение частот нуклеотидов в предсказанном нами регуляторном мотиве связывания NtcA с ДНК у цианобактерий**



**Рисунок 1б. Распределение частот нуклеотидов в предсказанном нами регуляторном мотиве связывания NtcB с ДНК у цианобактерий**

Сайты связывания фактора NtcB обнаружены нами у многих видов перед генами *nrtA* и *nirA*, частотное распределение нуклеотидов показано на рис. 1b. NtcB-регулон включает ген *nrtA* у Chroococcales и ген *nirA* – у Chroococcales, Nostocales, Oscillatoriales, Gloeobacterales. В пластомах Rhodophyta сайты связывания NtcB не найдены.

### 3.2. Совместная регуляция генов *rps20* и *glnB*

Ген *rps20* кодирует рибосомный белок S20. Найдены потенциальные промоторы для этого гена у всех семи секвенированных красных и криптофитовых водорослей. Они близки к консенсусу бактериальных  $\sigma^{70}$ -зависимых промоторов и включают 5'-расширение (-10)-бокса, рис. 2.

Нами предположена совместная регуляция генов *rps20* и *glnB* с участием транскрипционного фактора Ycf28. Ген *glnB* позиционно связан с геном *rps20* и кодирует белок из семейства PII, который регулирует активность ферментов, связанных с

метаболизмом азота [14], посредством белок-белкового взаимодействия. Ген *ycf28* ортологичен гену бактериального транскрипционного фактора NtcA из семейства *crp*, которое определяется наличием консервативного домена PF00325 (по базе Pfam). Несмотря на гомологию между собой белков Ycf28 из хлоропластов всех пяти доступных красных водорослей, сам домен присутствует в белках только трех видов: *Cyanidium caldarium* (здесь *rps20* – псевдоген), *Porphyra purpurea* и *Porphyra yezoensis*. И отсутствует в белках из *Cyanidioschyzon merolae* и *Gracilaria tenuistipitata*, табл. 1.

Таким образом, имеется корреляция между присутствием этого домена и наличием найденного нами сайта связывания репрессора, см. рис. 2. Выше упомянуты найденные нами консервативные участки с консенсусами GTATyATA или TTAAnnAAAAAnAA 5'-лидерной области гена *glnB*. Они перекрывают промотор перед дивергентно расположенным геном *rps20* – на рис. 2 эти участки показаны на цепи гена *rps20*.

**Таблица 1. Домены PF00325, характеризующие белковое семейство Crp, вблизи C-концов белков Ycf28 из хлоропластов красных водорослей.** Показаны парные выравнивания каждого из найденных доменов с консенсусом таких доменов у всех белков из семейства Crp (по базе Pfam). Числа указывают координаты начала и конца каждого домена. Значение E указывает качество E-value парного выравнивания; участок с качеством большим 1 не приведен.

Консенсус белкового семейства Crp			LpmsLRqeIAdylG1TrETVsR1LtrLrekGLI
<i>Cyanidioschyzon merolae</i>	E=0.9	114	WRLS=QASLARILGTSRAAIGQVLGDWKKQAWL 145
<i>Cyanidium caldarium</i>	E=0.00019	157	IYIS=QHDIASILSTTRSTITRLINQLRKDNII 188
<i>Porphyra purpurea</i>	E=0.00052	184	LTIT=HKVLAQIIIGSNRVSITRIISKLIHTKFI 215
<i>Porphyra yezoensis</i>	E=0.0021	184	FTIT=HKILAQIIIGSNRVSVTRILANLLKTKLI 215
<i>Gracilaria tenuistipitata</i>	E>1		

Species	Promoter upstream <i>rps20</i>	
<i>Cyanidioschyzon merolae</i>	ACTCTTGCTTTTGGCATCTGCT=ATTTTATCTTTATGTAGACT	-33
<i>Cyanidium caldarium</i>	AAATTTGTTTATTTTACGTTAAAT=ATGATACAGTAATTTATAAC	-32
<i>Porphyra purpurea</i>	GCTATTGCCTATTTCTTTTTCGTTAAATGTTATAATACGGCGCATA	-78
<i>Porphyra yezoensis</i>	ACTATTGCCTATTTCTTTTTCGTTAAATGTTATAATACGCCGCATA	-78
<i>Gracilaria tenuistipitata</i>	GTTCTTGTCTATTTTAAATGATTAATGATATAATCCCAATTAGAT	-63
<i>Guillardia theta</i>	TAAATTTATTCCATTAATTTCTTATATGTTATAATCTTTTATTAC	-59
<i>Rhodomonas salina</i>	TCTTCTTATTTC=ATAATTTGTTCTATGTTATAATCACTAATCGT	-55
	-35box      17-18      EX -10box	

**Рисунок 2. Потенциальные промоторы перед геном *rps20*.** Подчёркнуты (-35)- и (-10)-боксы промотора и 5'-расширение (-10)-бокса. Цветом выделены консервативные участки, которые предположительно служат сайтами связывания репрессора.

Несмотря на консервативность промотора положение найденных участков не постоянно относительно его боксов, поэтому консервативность участков не может быть объяснена просто консервативностью самого промотора. Существенно, что от этих участков и вплоть до начала гена *glnB* нет бактериального  $\sigma^{70}$ -промотора и нет оснований предполагать наличие там промоторов других типов.

Наличие этих участков коррелирует с присутствием в ортологе фактора NtcA домена PF00325 для связывания с участками ДНК. Поэтому можно предположить, что имеет место белок-ДНКовая регуляция на сайте, состоящем из этих двух участков. А также – предположить, что она сочетается с конкуренцией в этом локусе полимераз на двух цепях ДНК.

#### 4. Выводы

Изучение большого количества видов в связи с регуляторной активностью факторов NtcA и NtcB перед различными генами позволило: 1) предсказать много новых сайтов связывания этих факторов и существенно уточнить консенсусы сайтов; 2) предсказать эволюцию NtcA- и NtcB-регулонов; 3) показать, что у цианобактерий фактор NtcA репрессирует транскрипцию генов *gifA*, *gifB* у многих цианобактерий и активирует транскрипцию других генов у многих цианобактерий и Rhodophyta; в частности, генов *apcF* и *som*, для которых, по-видимому, отсутствует экспериментальное подтверждение регуляции; 4) опровергнуть гипотезу из [9] о существенной роли у подавляющего большинства видов NtcA-регуляции генов, вовлеченных в фиксацию углерода (*rbcL*, *rpe*, *cmpABCD*), или генов фотосистем: например, *petH*, кодирующего ферредоксин-НАДФ-редуктазу, хотя его NtcA-регуляция экспериментально подтверждена у *Nostoc* sp. PCC 7120; это также относится к гену *psaI*, кодирующему белок I первой фотосистемы.

#### 5. Список литературы

[1] M. Garcia-Dominguez, J.C. Reyes, F.J. Florencio, "NtcA represses transcription of *gifA* and *gifB*, genes that encode inhibitors of glutamine synthetase type I from *Synechocystis* sp. PCC 6803", *Molecular Microbiology*, 2000, 35 (5), pp. 1192-1201.

[2] J.E. Frias, E. Flores, A. Herrero, "Activation of the *Anabaena nir* operon promoter requires both NtcA (CAP family) and NtcB (LysR family) transcription factors", *Molecular Microbiology*, 2000, 38(3), pp. 613-625.

[3] M. Alfonso, I. Perewoska, D. Kirilovsky, "Redox control of *ntca* gene expression in *Synechocystis* sp. PCC 6803. Nitrogen availability and electron transport regulate the levels of the NtcA protein", *Plant Physiology*, 2001, 125. pp. 969-981.

[4] M.I. Muro-Pastor, F.J. Florencio. "Regulation of ammonium assimilation in cyanobacteria", *Plant Physiology and Biochemistry*, 2003, 41, pp. 595-603.

[5] C. Bird, M. Wyman, "Nitrate/nitrite assimilation system of the marine picoplanktonic cyanobacterium *Synechococcus* sp. strain WH 8103: effect of nitrogen source and availability on gene expression", *Applied and environmental microbiology*, 2003, 69 (12), pp. 7009-7018.

[6] M.F. Aldehni, K. Forchhammer, "Analysis of a non-canonical NtcA-dependent promoter in *Synechococcus elongatus* and its regulation by NtcA and PII", *Arch Microbiol*, 2006, 184, pp. 378-386.

[7] S.-I. Maeda, Y. Kawaguchi, T.-A. Ohe, T. Omata, "Cis-Acting sequences required for NtcB-dependent, nitrite-responsive positive regulation of the nitrate assimilation operon in the cyanobacterium *Synechococcus* sp. strain PCC 7942", *Journal of bacteriology*, 1998, 180 (16), pp. 4080-4088.

[8] Z. Su, V. Olman, F. Mao, Y. Xu, "Comparative genomics analysis of NtcA regulons in cyanobacteria: regulation of nitrogen assimilation and its coupling to photosynthesis", *Nucleic Acids Research*, 2005, 33(16), pp. 5156-5171.

[9] D.A. Benson, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell and D.L. Wheeler, "GenBank", *Nucleic Acids Research, Oxford J.*, 2008, 36 (Database issue), pp. D25-30.

[10] L. Cummings, L. Riley, L. Black, A. Souvorov, S. Resenchuk, I. Dondoshansky, T. Tatusova, "Genomic BLAST: custom-defined virtual databases for complete and unfinished genomes", *FEMS Microbiol. Lett.*, 2002, 216, pp. 133-138.

[11] Finding of multi-box regulatory signal in the set of unaligned sequences [<http://lab6.iitp.ru/en/twobox/30.03.2010>].

[12] A tree-based method of sequence alignment [<http://lab6.iitp.ru/en/treeal/30.03.2010>].

[13] G.E. Crooks, G. Hon, J.M. Chandonia and S.E. Brenner, "WebLogo: A sequence logo generator", *Genome Research, Cold Spring Harbor Laboratory Press*, 2004, 14, pp. 1188-1190.

[14] K. Forchhammer, "PII signal transducers: novel functional and structural insights", *Trends in Microbiology*, 2008, 16(2), pp. 65-72.