

УДК: 57.056:575.117.2:546.48

Механизмы регуляции экспрессии гена *dps Escherichia coli* в условиях стресса: реконструкция по кинетическим данным

Хлебодарова Т.М.^{*1}, Степанова Т.Ю.¹, Ощепков Д.Ю.¹,
Тикунова Н.В.³, Бабкин И.В.³, Лихошвай В.А.^{**1,2}

¹Институт цитологии и генетики, Сибирское отделение Российской академии наук,
Новосибирск, 630090, Россия

²Новосибирский национальный исследовательский государственный университет,
Новосибирск, 630090, Россия

³Институт химической биологии и фундаментальной медицины, Сибирское отделение
Российской академии наук, Новосибирск, 630090, Россия

Аннотация. Механизмы экспрессии гена *dps* в присутствии солей кадмия реконструированы по динамике флюоресценции геносенсорной конструкции на основе промотора гена *dps Escherichia coli* и репортерного гена *gfp* с использованием метода математического моделирования и биоинформационного анализа. Оценка непротиворечивости различных гипотез о генетических механизмах регуляции экспрессии гена *dps* показала, что наиболее вероятным механизмом регуляции его активности в ответ на кадмий может быть взаимодействие трех различных транскрипционных факторов, два из которых являются активаторами гена *dps*, а один репрессором, обладающие разной чувствительностью к кадмию. Модель позволяет выделить неконкурентные отношения между факторами как наиболее предпочтительные. Анализ регуляторной области гена *dps* показал, что данным условиям удовлетворяют, если учитывать также способность кадмия индуцировать окислительный стресс, сайты связывания транскрипционных факторов OxyR, H-NS и Rob (или IscR). OxyR и Rob (или IscR) могут быть активаторами экспрессии гена, а H-NS – репрессором. Перекрывающиеся потенциальные сайты связывания факторов Rob и IscR выявлены в промоторе гена *dps* методом SITECON и экспериментально подтверждены гель-шифт анализом с рекомбинантными белками Rob и IscR *E. coli*.

Ключевые слова: моделирование, распознавание сайтов связывания транскрипционных факторов, регуляция экспрессии генов, *Escherichia coli*, *dps*, кадмий.

ВВЕДЕНИЕ

Ген *dps Escherichia coli* кодирует структуру ДНК-связывающего белка, который защищает клетки от множественных стрессов, включая окислительный и термальный стресс, UV и гамма облучение, кислотный и щелочной шок, а также повышает устойчивость клеток к токсическому воздействию меди и железа [1–3]. Данные микрочип-эксперимента также свидетельствуют о чувствительности гена *dps E. coli* к различным внешним воздействиям [4–7].

* tamara@bionet.nsc.ru

** likho@bionet.nsc.ru

В настоящее время известно, что в экспоненциальной фазе роста экспрессия гена *dps E. coli* в условиях окислительного стресса регулируется через транскрипционный фактор (ТФ) ОхуR [8], а индукция экспрессии в условиях голодания, холода и осмотического стресса зависит от присутствия *groS* полимеразы [7, 9]. Однако спектр агентов, активирующих экспрессию гена *dps* значительно шире [4, 5], и механизмы реализации этих ответов неизвестны.

Негативная роль ТФ Fis, H-NS и MntR в регуляции экспрессии гена *dps E. coli* не связывается авторами с реализацией какого-либо стрессового ответа [10, 11].

Ранее нами на основе промотора гена *dps E. coli* и репортерного гена *gfp* был создан геносенсор *E. coli/pDps-gfp* [12], который оказался высокочувствителен не только к перекиси водорода, но и к солям кадмия [13].

Как указывалось, чувствительность гена *dps* к перекиси водорода объясняется наличием в его промоторе сайта связывания ТФ ОхуR [8], который является сенсором окислительного стресса и реализует свое действие через *groD* полимеразу. У мутантов по этому фактору индукция экспрессии гена *dps* в присутствии перекиси водорода резко падает [6]. Что касается солей кадмия, то в настоящее время, в промоторе гена *dps* не описано сайтов связывания для ТФ, через которые мог бы реализоваться ответ на тяжелые металлы.

В тоже время известно, что кадмий индуцирует окислительный стресс [14], вероятно за счет его способности образовывать формы активного кислорода [15–17]. Исходя из этого, можно было предположить возможность реализации ответа на кадмий через сайт связывания ТФ ОхуR. Однако простейший анализ оценки динамики флюоресценции геносенсора *E. coli/pDps-gfp* в ответ на Cd^{2+} показал, что, в отличие от перекиси водорода, зависимость его уровня индукции от концентрации Cd^{2+} носит сложный нелинейный характер [13], которая, не может быть реализована через один сайт. В условиях токсического воздействия на клетку ионов Cd^{2+} динамика флюоресценции геносенсора *E. coli/pDps-gfp*, являясь не прямым показателем экспрессии промотора гена *dps*, может не совпадать с истиной картиной регуляции гена *dps* в ответ на кадмий. Предположив, что динамика активности геносенсора *E. coli/pDps-gfp* в стрессовых условиях является отражением, как генетических механизмов регуляции экспрессии промотора гена *dps*, так и метаболического статуса рекомбинантных клеток, который может изменяться в процессе развития стрессорной реакции, мы методами математического моделирования оценили генетическую составляющую ответа клеток геносенсора на Cd^{2+} и провели реконструкцию механизмов регуляции экспрессии гена *dps E. coli* в присутствии Cd^{2+} . Теоретические предсказания были верифицированы с помощью биоинформационных методов, в результате чего был определен паттерн потенциальных регуляторов гена *dps* в присутствии ионов кадмия. Было предсказано, что сложная нелинейная динамика флюоресценции клеток геносенсора *E. coli/pDps-gfp* в условиях токсического действия Cd^{2+} может быть отражением взаимодействия трех транскрипционных факторов – ОхуR, H-NS и Rob (или IscR), два из которых (ОхуR и Rob/IscR) могут быть активаторами экспрессии гена, а H-NS – репрессором. Присутствие перекрывающихся потенциальных сайтов связывания ТФ Rob и IscR в промоторе гена *dps* экспериментально подтверждено гель-шифт анализом с рекомбинантными белками Rob и IscR *E. coli*.

1. МЕТОДЫ

1.1. Метод оценки непротиворечивости механизмов регуляции активности промотора гена *dps E. coli*

Непротиворечивость предположительных механизмов регуляции активности промотора гена *dps E. coli* проверяли с использованием обобщенных функций Хилла

[18], которыми описывали удельную активность эффективности инициации транскрипции в зависимости от концентрации ионов кадмия, добавленных в среду. Примеры конкретных функций приведены в разделе «Результаты».

Качество модели оценивали по минимальному значению функционала

$$\sum_i \left(\frac{V(s_i)}{U(s_i)} + \frac{U(s_i)}{V(s_i)} - 2 \right), \quad (1)$$

которое вычисляли методом градиентного спуска. В (1) функция V соответствует рассматриваемому механизму, а суммирование ведется по точкам, в которых проводились измерения: i – номер измерения, s_i – концентрация добавленных ионов кадмия, $U(s_i)$ – оцененная из экспериментов относительная генетическая активность промотора гена *dps* в зависимости от концентрации добавленных ионов кадмия s_i , $V(s_i)$ – относительная генетическая активность промотора гена *dps*, вычисленная по обобщенной функции Хилла, вид которой соответствует тестируемой модели регуляции активности промотора *pDps* ионами кадмия. Полученный оптимальный набор параметров и значение достигнутого минимума анализировались с точки зрения соответствия биологической природе рассматриваемых процессов и на основании этого делался вывод о непротиворечивости тестируемого регуляторного механизма.

Значения $U(s_i)$ оценивали из экспериментальных данных по динамике изменения интенсивности флюоресценции клеток геносенсора в присутствии различных концентраций Cd^{2+} при температуре 26°C [13]. Методика оценки значений $U(s_i)$ описана в разделе 2.1 главы «Результаты и обсуждение».

1.2. Компьютерные методы

Для анализа структуры регуляторной области гена *dps* *E. coli* был использован метод SITECON, позволяющий эффективно распознавать сайты связывания ТФ. Метод заключается в выявлении набора консервативных контекстно-зависимых конформационных и физико-химических свойств, определенных для позиций выравнивания сайтов связывания ТФ и дальнейшего сравнения выявленных консервативных свойств со свойствами анализируемой последовательности [19]. Для распознавания в промоторе гена *dps* потенциальных сайтов ТФ, были созданы обучающие выборки экспериментально доказанных сайтов связывания ТФ Rob, PurR, MetJ, MarA, SoxS, и IscR, которые прямо или косвенно могут участвовать в ответе гена на тяжелые металлы и окислительный стресс [5, 20, 21]. В качестве контроля для распознавания использовали также выборки сайтов связывания ТФ OxyR и IHF, локализация которых в промоторе гена *dps* была определена ранее [8]. Объем выборки для каждого ТФ и длина сайтов приведены в табл. 1.

Таблица 1. Параметры обучающих выборок экспериментально доказанных сайтов связывания ТФ, использованных для выявления потенциальных сайтов методом SITECON в промоторе гена *dps* *E. coli*

Параметры выборки	Транскрипционные факторы							
	Rob	IscR	SoxS	MarA	OxyR	MetJ	PurR	IHF
Длина сайтов в выборке	24	42	30	31	31	23	31	33
Объем выборки	6	4 (тип 1) 4 (тип 2)	18	16	17	24	16	42

Сайты были выровнены относительно наиболее часто встречающихся нуклеотидов и использованы для анализа. Методика оценки достоверности распознавания сайтов по величине ошибок первого и второго рода, описана ранее [22]. В качестве

потенциальных мы рассматривали только те сайты связывания ТФ, уровень перепредсказания (ошибка II рода) которых был меньше, чем 1/2000 нуклеотидов. Уровень перепредсказания отражает количество потенциальных сайтов, которые могут быть найдены в случайной последовательности по случайным причинам на данном пороге распознавания, и поэтому данный параметр может быть использован как простая оценка уровня достоверности предсказания.

1.3. Экспериментальные методы

Для получения рекомбинантных белков транскрипционных факторов IscR, Rob и MarA *E. coli*, на основе известных нуклеотидных последовательностей этих генов (GenBank: U00096), были рассчитаны праймеры для ПЦР: 5'-CTGGTCATGAGACTG-ACATCTAAAGGGCGCTATG-3' и 5'-CTGGAGATCTAGCGCGTAACTTAACGTCG-ATCGC-3', 5'-CCACCATGGATCAGGC-CGGCATTATTC-3' и 5'-CCAAGATCTACGACGGATCGGAATCAGCAGT-3', 5'-CTCCTCATGACGATGTCCAGACGCAATACTGA-3' и 5'-TCCTAGATCTGCTGTTGT-AATGATTTAATGGATGAAAAAG-3', соответственно. Праймеры включали сайты рестрикции *Bgl*III, *Cci*I или *Nco*I. Полученные амплифицированные фрагменты были расчетной длины и клонированы в вектор pQE60 по соответствующим сайтам. Рекомбинантные белки были наработаны и очищены с помощью Ni-хелатной хроматографии согласно протоколу (QIAGEN).

Гель-шифт анализ с рекомбинантными очищенными белками проводили, как описано [23]. Олигонуклеотидные дуплексы, моделирующие потенциальные сайты связывания ТФ IscR, Rob и MarA в промоторе гена *dps E. coli*, были синтезированы и соответствовали последовательностям сайта ТФ IscR 5'-ATCTTTATTAGTTTTGTT-TTTCACGCTTGTTACCACTA-3' в позиции -97/-60, сайта ТФ Rob 5'-ACAAGCGTGAAAAACAATAAAGAT-3' в позиции -97/-67 и сайта ТФ MarA 5'-TATGCAAAAGCATATTTATCCGAAAATTCC-3' в позиции -154/-124 от старта транскрипции. В качестве позитивного контроля использовали экспериментально доказанные сайты связывания ТФ IscR, Rob и MarA из промоторов соответствующих генов *iscR*, *rob* и *marA E. coli*.

Для определения функциональности потенциального сайта RpoE полимеразы в -35-область его последовательности были внесены точковые мутации, которые не затрагивали сайт связывания RpoD полимеразы. Для синтеза измененной последовательности промотора были сконструированы праймеры для ПЦР, в последовательность которых были внесены необходимые мутации: 5'-TATGTGACCGCTATTCTGGCTGTTCCT-3' и 5' AGGAACAGCCAGAATAGCGG-TCCGACATAG-3'. В качестве матрицы использовали плазмиду pDps-gfp [12]. Процесс синтеза был двухэтапным. Полученные ампликоны были обработаны эндонуклеазами рестрикции *Apa*I и *Sph*I, выделены с помощью электрофореза в 1% агарозном геле, очищены согласно протоколу производителя (QIAGEN) и лигированы с вектором. Соответствие последовательности сайта внесенным изменениям проверяли секвенированием.

2. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

2.1. Оценка генетической составляющей механизма регуляции активности геносенсора *E. coli/pDps-gfp* под воздействием ионов кадмия

Для оценки активности промотора гена *dps*, как указывалось выше, были использованы полученные ранее экспериментальные данные [13, рис. 2]. В экспериментах использованы 8 различных концентраций s_k , $k = 1, \dots, 8$, а измерения проведены в 18 временных точках культивирования t_j , $j = 0, 1, \dots, 17$. Поэтому

экспериментальные данные представляют прямоугольную числовую матрицу $Flu(s_k, t_j)$, $k = 1, \dots, K, j = 1, \dots, J, K = 7, J = 17$.

Было показано, что рекомбинантные клетки *E. coli/pDps-gfp* при фиксированной концентрации s_k добавленных в среду ионов кадмия реагировали на присутствие Cd^{2+} в среде монотонным повышением уровня флюоресценции $Flu(s_k, t_j)$ с течением времени t_j (рис. 1,а, точки). В фиксированные времена t_j характер изменения уровня флюоресценции $Flu(s_k, t_j)$ демонстрирует нетривиальную зависимость: в диапазоне концентраций s от 0.0 мМ до 0.125 мМ интенсивность флюоресценции растет; в интервале [0.125 мМ, 0.5 мМ] интенсивность флюоресценции падает и при концентрации $s = 0.5$ мМ достигает минимальных значений; при $s \geq 0.5$ мМ значение $Flu(s_k, t_j)$ растет с насыщением. Таким образом, данные, приведенные на рис. 1,а, показывают сложную зависимость изменения интенсивности флюоресценции Flu от значений концентраций добавленных ионов кадмия, что свидетельствует о зависимости активности промотора гена *dps* от концентраций ионов кадмия в среде.

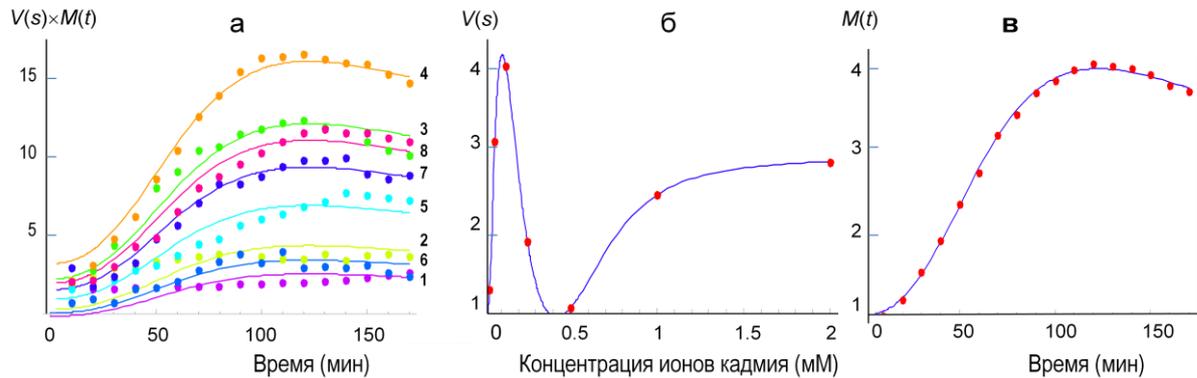


Рис. 1. Динамика экспрессии гена *dps* *E. coli* в зависимости от времени с момента добавления в среду ионов кадмия. (а) точки – уровень флюоресценции $Flu(s_k, t_j)$, измеренный в эксперименте при различных концентрациях Cd^{2+} (кривая 1 – $s_0 = 0$ мМ, кривая 2 – $s_1 = 0.031$ мМ, кривая 3 – $s_2 = 0.062$ мМ, кривая 4 – $s_3 = 0.125$ мМ, кривая 5 – $s_4 = 0.25$ мМ, кривая 6 – $s_5 = 0.5$ мМ, кривая 7 – $s_6 = 1$ мМ, кривая 8 – $s_7 = 2$ мМ), кривые – уровень флюоресценции $V(s) \times M(t)$, рассчитанные как произведение функций (2) и (3); (б) точки относительной активности $V(s)$ промотора *dps*, представленные в табл. 2, кривая рассчитана на основе аппроксимирующей функции (2); (в) точки – значения коэффициентов $M(t)$ для пересчета относительной активности промотора *dps* в фиксированные моменты времени в интенсивность флюоресценции Flu , представленные в табл. 2, кривая рассчитана на основе аппроксимирующей функции (3). Для расчета функций (2) и (3) использованы следующие значения параметров: $v = 0.85, v_1 = 5.42, v_2 = 2.92, k_1 = 0.061, k_2 = 0.106, k_3 = 0.208, n_1 = 3.32, n_2 = 6.22, n_3 = 9.96, m_0 = 0.43, t_{m,0} = 39.21, t_{m,1} = 86.21, n_{m,1} = 3.32, n_{m,2} = 3.32$.

Чтобы выделить из наблюдаемых данных, представленных на рис 1,а, вклад активности экспрессии гена в интенсивность флюоресценции, мы предложили следующую модель структуры данных: значение $Flu(s_k, t_j)$ представляет произведение величины собственной удельной активности промотора *pDps* на коэффициент пересчета данной активности в интенсивность флюоресценции белка Gfp: $Flu(s_k, t_j) = U(s_k)M(t_j)$. Здесь U – удельная активность инициации транскрипции с промотора гена *dps* (генетическая составляющая), M – коэффициент метаболической активности клетки, не связанный с активностью функционирования промотора *dps* (плотность культуры, копияность плазмид, эффективность трансляции, стабильность РНК и белков и т.д.).

Предложенная модель приводит к следующей оптимизационной задаче. Имеется прямоугольная положительная матрица чисел $Flu(s_k, t_j)$. Необходимо найти два положительных вектора $\mathbf{U} = (x_k)_{k=1, \dots, K}$ и $\mathbf{M} = (y_j)_{j=1, \dots, J}$, для которых функционал

$$D = \sqrt{\sum_{j=1}^J \sum_{k=1}^K (x_k y_j - Flu(s_k, t_j))^2} \quad (2)$$

имеет минимальное значение. Так как функционал (2) имеет одну степень свободы, то необходимо задать одно ограничение. В работе приняли, что значение коэффициента метаболической активности $M(0)$ в нулевой момент времени равно единице: $y_1 = 1$.

По условию задачи поиск решения проводили в положительной области изменения переменных. Так как матрица экспериментальных данных является положительной, а оптимизируемая функция является билинейной формой, то в данной области гарантированно существует единственный минимум, который был установлен методом градиентного спуска.

В табл. 2 представлены оцененные значения векторов \mathbf{X} и \mathbf{Y} . На рис. 1,б и 1,в полученные результаты представлены точками. Для повышения наглядности представления качества решенной задачи мы построили функции $V_3(s)$ и $M_h(t)$, аппроксимирующие соответствующие данные из табл. 2. Активность промотора $V(s)$ была аппроксимирована функцией

$$V_3(s) = \frac{v_{3,0} + v_{3,1} \left(\frac{s}{k_{3,1}}\right)^{n_{3,1}} + v_{3,2} \left(\frac{s}{k_{3,3}}\right)^{n_{3,2}}}{1 + \left(\frac{s}{k_{3,1}}\right)^{n_{3,1}} + \left(\frac{s}{k_{3,2}}\right)^{n_{3,2}} + \left(\frac{s}{k_{3,3}}\right)^{n_{3,3}}}, \quad (3)$$

а коэффициент пересчета $M(t)$ – функцией

$$M_h(t) = m_0 + \frac{1 - m_0 + \left(\frac{t}{t_m}\right)^{n_{m,1}}}{1 + \left(\frac{t}{t_m}\right)^{n_{m,2}}}. \quad (4)$$

Значения параметров были оценены на основе функционала (1). Кривая на рис. 1,б рассчитана по формуле (3), а кривая на рис. 1,в – по формуле (4). На рис. 1,а приведены кривые, которые рассчитаны как произведение функций (3) и (4). Значения параметров, при которых проводились расчеты, приведены в легенде к рис. 1. Достигнутое качество соответствия теоретических кривых экспериментальным данным (рис. 1,а), позволяет нам сделать вывод, что предложенная модель структуры данных является непротиворечивой и построенная в рамках этой модели кривая (рис. 1,б) действительно отражает удельную активность промотора гена *dps* в клетке в зависимости от концентрации ионов кадмия в среде и может быть использована для анализа механизмов его регуляции.

Таблица 2. Оцененные значения относительной активности $V(s)$ промотора гена *dps* *E. coli* и коэффициента пересчета $M(t)$

s (mM)	$V(s)$	t (мин)	$M(t)$	t (мин)	$M(t)$	t (мин)	$M(t)$
0	0.85	0	1	60	2.76	120	4.10
0.031	1.28	10	1	70	3.22	130	4.08
0.062	3.13	20	1.19	80	3.48	140	4.05
0.125	4.1	30	1.53	90	3.75	150	3.97
0.25	1.89	40	1.92	100	3.90	160	3.83
0.5	1.06	50	2.37	110	4.03	170	3.76
1.0	2.47						
2.0	2.88						

Таким образом, генетическая составляющая ответа клеток геносенсора *E. coli/pDps-gfp* на ионы кадмия при температуре 26°C имеет два пика активности (рис. 1,б), которые можно условно описать тремя состояниями промотора гена *dps*, зависящими от концентрации Cd^{2+} . Первое и второе состояние промотора гена *dps* определяется его активацией при низких и высоких концентрациях кадмия, а третье – репрессией при средних концентрациях Cd^{2+} . В связи с этим, возникает вопрос о том, какой молекулярно-генетический механизм регуляции активности промотора гена *dps* способен воспроизводить подобную кинетику.

2.2. Оценка непротиворечивости различных гипотез о генетических механизмах регуляции активности промотора гена *dps E. coli*

Для ответа на поставленный выше вопрос, мы рассмотрели разные варианты регуляторных механизмов и для каждого построили специальную обобщенную функцию Хилла. Затем для каждой функции, методом, описанным выше, определили оптимальные значения параметров, при которых она наилучшим образом приближает оцененные значения относительной активности промотора гена *dps*, приведенные в табл. 1. Величина отклонения служила критерием непротиворечивости. Чем меньше была величина отклонения, тем более вероятным считался механизм. Было рассмотрено девять различных комбинаций регуляторных механизмов. В первых 2-х вариантах исследованы возможности реализации механизма регуляции транскрипции гена *dps* с двух промоторов, регуляторные области которых не конкурируют между собой. При этом первый промотор активируется малыми дозами Cd^{2+} и ингибируется средними и высокими, а второй – активируется высокими дозами Cd^{2+} . В первом варианте активирующая и ингибирующая активности первого промотора конкурируют между собой, а во втором – нет.

Данные предположения приводят к формулам (5) и (6)

$$V_1(s) = \left(\frac{v_{1,0} + v_{1,1} \left(\frac{s}{k_{1,1}} \right)^{n_{1,1}}}{1 + \left(\frac{s}{k_{1,1}} \right)^{n_{1,1}} + \left(\frac{s}{k_{1,2}} \right)^{n_{1,2}}} + \frac{v_{1,3} \left(\frac{s}{k_{1,3}} \right)^{n_{1,3}}}{1 + \left(\frac{s}{k_{1,3}} \right)^{n_{1,3}}} \right), \quad (5)$$

$$V_2(s) = \left(\frac{v_{2,0} + v_{2,1} \left(\frac{s}{k_{2,1}} \right)^{n_{2,1}}}{\left(1 + \left(\frac{s}{k_{2,1}} \right)^{n_{2,1}} \right) \left(1 + \left(\frac{s}{k_{2,2}} \right)^{n_{2,2}} \right)} + \frac{v_{2,2} \left(\frac{s}{k_{2,3}} \right)^{n_{2,3}}}{1 + \left(\frac{s}{k_{2,3}} \right)^{n_{2,3}}} \right). \quad (6)$$

Решение задачи оптимизации параметров приводит к следующим значениям параметров для модели (5): $v_{1,0} = 0.85$; $v_{1,1} = 5.51$; $v_{1,3} = 2.91$; $k_{1,1} = 0.062$; $k_{1,2} = 0.105$; $k_{1,3} = 0.67$; $n_{1,1} = 3.31$; $n_{1,2} = 6.18$; $n_{1,3} = 3.93$, а для модели (10): $v_{2,0} = 0.85$; $v_{2,1} = 5.61$; $v_{2,2} = 2.91$; $k_{2,1} = 0.062$; $k_{2,2} = 0.195$; $k_{2,3} = 0.67$; $n_{2,1} = 3.32$; $n_{2,2} = 2.89$; $n_{2,3} = 3.93$. В обоих случаях, при данных значениях функционал (1) имеет нулевое отклонение и кривые аппроксимации данных табл. 2, практически не различаются (рис. 2,а).

Далее были рассмотрены пять вариантов регуляции гена *dps* с единственного промотора, что также аналогично ситуации 2-х конкурирующих между собой промоторов. Во всех случаях предполагается, что промотор активируется малыми и высокими дозами Cd^{2+} , причем активация малыми дозами ингибируется средними и

высокими дозами, а активация высокими дозами Cd^{2+} не подвержена негативной регуляции.

В данной ситуации, если регуляторные активности конкурируют между собой, то предположения приводят к формуле (3), если не конкурируют – к формуле (6). Если активность малыми и репрессия средними дозами Cd^{2+} конкурируют между собой и не конкурируют с активностью большими дозами, то это приводит к формуле (9). Для этих моделей решение задачи оптимизации параметров приведено выше (рис. 1,б и 2,а). Предположение, что активация малыми и репрессия средними дозами Cd^{2+} не конкурируют между собой и конкурируют с активацией большими дозами Cd^{2+} , приводит к формуле (7), если активация малыми и большими дозами Cd^{2+} конкурируют между собой, а репрессия не конкурирует с процессами активации, то это приводит к формуле (8):

$$V_6(s) = \frac{v_{6,0} + v_{6,1} \left(\frac{s}{k_{6,1}} \right)^{n_{6,1}} + v_{6,2} \left(\frac{s}{k_{6,3}} \right)^{n_{6,3}}}{\left(1 + \left(\frac{s}{k_{6,1}} \right)^{n_{5,1}} \right) \left(1 + \left(\frac{s}{k_{6,2}} \right)^{n_{6,2}} \right) + \left(\frac{s}{k_{6,3}} \right)^{n_{6,3}}}, \quad (7)$$

$$V_7(s) = \frac{v_{7,0} + v_{7,1} \left(\frac{s}{k_{7,1}} \right)^{n_{7,1}} + v_{7,2} \left(\frac{s}{k_{7,3}} \right)^{n_{7,3}} \left(1 + \left(\frac{s}{k_{7,2}} \right)^{n_{7,2}} \right)}{\left(1 + \left(\frac{s}{k_{7,1}} \right)^{n_{7,1}} + \left(\frac{s}{k_{7,3}} \right)^{n_{7,3}} \right) \left(1 + \left(\frac{s}{k_{7,2}} \right)^{n_{7,2}} \right)}. \quad (8)$$

Решение задачи оптимизации параметров приводит к следующим значениям параметров для модели (7): $v_{6,0} = 0.85$; $v_{6,1} = 5.59$; $v_{6,2} = 2.92$; $k_{6,1} = 0.062$; $k_{6,2} = 0.2$; $k_{6,3} = 0.21$; $n_{6,1} = 3.3$; $n_{6,2} = 2.87$; $n_{6,3} = 3.96$, и для модели (12): $v_{7,0} = 0.85$; $v_{7,1} = 5.58$; $v_{7,2} = 2.92$; $k_{7,1} = 0.062$; $k_{7,2} = 0.2$; $k_{7,3} = 0.22$; $n_{7,1} = 3.33$; $n_{7,2} = 2.89$; $n_{7,3} = 7.17$. В обоих случаях, при данных значениях функционал (1) имеет нулевое отклонение и кривые аппроксимации данных, приведенных в табл. 2, аналогичны друг другу (рис. 2,б).

В заключение рассмотрим еще два варианта инициации транскрипции гена *dps* с единственного промотора, когда активация осуществляется только малыми концентрациями Cd^{2+} , а репрессия – средними. Если активация малыми дозами Cd^{2+} и репрессия средними конкурируют между собой, то эти предположения приводят к формуле (9), если активация малыми дозами и репрессия средними дозами Cd^{2+} не конкурируют между собой, то это приводит к формуле (10):

$$V_8(s) = \frac{v_{8,0} + v_{8,1} \left(\frac{s}{k_{8,1}} \right)^{n_{8,1}}}{\left(1 + \left(\frac{s}{k_{8,1}} \right)^{n_{8,1}} + \frac{\left(\frac{s}{k_{8,2}} \right)^{n_{8,2}}}{1 + \left(\frac{s}{k_{8,3}} \right)^{n_{8,3}}} \right)}, \quad (9)$$

$$V_9(s) = \frac{v_{9,0} + v_{9,1} \left(\frac{s}{k_{9,1}} \right)^{n_{9,1}}}{\left(1 + \left(\frac{s}{k_{9,1}} \right)^{n_{9,1}} \right) \left(1 + \frac{\left(\frac{s}{k_{9,2}} \right)^{n_{9,2}}}{1 + \left(\frac{s}{k_{9,3}} \right)^{n_{9,3}}} \right)}. \quad (10)$$

Для модели (9) решение задачи оптимизации параметров приводит к следующим значениям параметров: $v_{8,0} = 0.85$; $v_{8,1} = 3.92$; $k_{8,1} = 0.048$; $k_{8,2} = 0.13$; $k_{8,3} = 0.31$; $n_{8,1} = 4.22$; $n_{8,2} = 11.88$; $n_{8,3} = 9.47$, а для модели (10) – к значениям: $v_{9,0} = 0.85$; $v_{9,1} = 4.03$; $k_{9,1} = 0.049$; $k_{9,2} = 0.235$; $k_{9,3} = 0.335$; $n_{9,1} = 4.35$; $n_{9,2} = 5.0$; $n_{9,3} = 6.81$. В обоих случаях, при данных значениях функционал (1) имеет отклонение 0.0269838, что визуально отражается в существенном отклонении теоретической кривой от трех экспериментальных точек (рис. 2,в).

На основании полученных данных, мы предполагаем, что наиболее вероятным вариантом молекулярно-генетического механизма регуляции активности промотора гена *dps* в ответ на кадмий может быть взаимодействие трех различных ТФ, два из которых являются активаторами гена *dps*, а один – репрессором, обладающие разной чувствительностью к кадмию. Следует отметить, что модель не позволяет различить ситуации реализации ответа через один или через два промотора, однако позволяет выделить неконкурентные отношения между ТФ как более предпочтительные. Об этом свидетельствуют менее напряженные параметры модели (6) в сравнении с остальными. Хотя некоторые варианты конкурентных взаимоотношений между ТФ активаторами (модель (7)) по параметрам близки к модели (6).

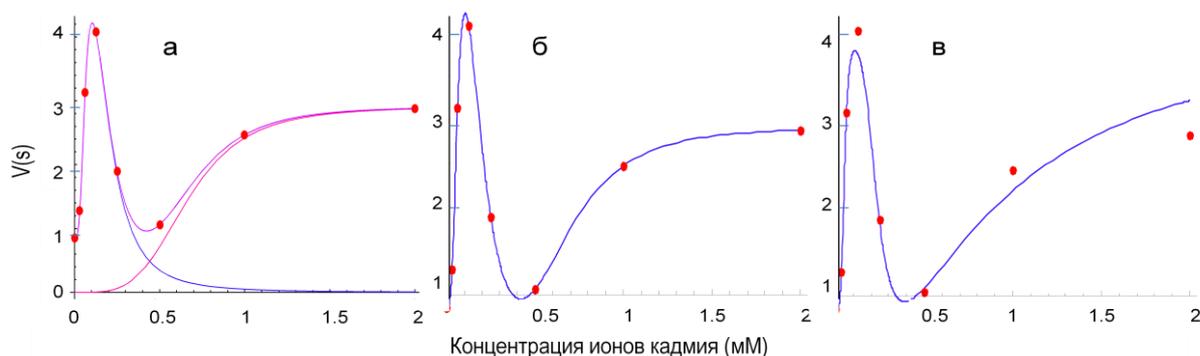


Рис. 2. Кривые, аппроксимирующие кинетику экспрессии гена *dps* *E. coli* в зависимости от концентрации Cd^{2+} , при различных механизмах его регуляции: а) – кривая рассчитана на основе аппроксимирующей функции (5), б) – функции (8), в) – функции (9). Точки – относительная активность $V(s)$ промотора гена *dps*, согласно данным приведенным в табл. 2.

Какие есть основания для реализации такого варианта регуляции экспрессии гена *dps* в присутствии кадмия? К настоящему времени в промоторе гена *dps* не описано сайтов связывания для ТФ, участвующих в ответе клеток *E. coli* на тяжелые металлы. Однако, учитывая способность кадмия образовывать формы активного кислорода [15–17], можно предположить участие в реализации этого ответа ОхуR и других факторов, чувствительных к радикал-образующим агентам.

2.3. Анализ структуры промотора гена *dps* *Escherichia coli* методом SITECON

Исходя из выше высказанных предположений, мы провели компьютерный анализ структуры промотора гена *dps* *E. coli* для выявления потенциальных сайтов связывания ТФ, которые прямо или косвенно могут участвовать в ответе гена на тяжелые металлы и окислительный стресс (см. раздел «Методы»).

В табл. 3 приведены результаты анализа последовательности промотора гена *dps* методом SITECON [19]. Дополнительно к ранее локализованным сайтам ТФ OxyR и IHF [8], в промоторе гена *dps* выявлены потенциальные сайты связывания трех ТФ, которые удовлетворяют описанным в разделе «Методы» ограничениям и могут участвовать в ответе гена на кадмий. Среди них ТФ Rob, который участвует в защите клеток от тяжелых металлов [20], и ТФ IscR (тип 2) и MarA, которые, как и OxyR, являются сенсорами окислительного стресса [5, 21, 24]. Присутствие сайтов связывания ТФ SoxS, MetJ и PurR в промоторе гена *dps* маловероятно (табл. 3). Позиции сайтов связывания ТФ Rob и IscR перекрываются (табл. 3), однако это не удивительно, поскольку они содержат в своей структуре [2Fe-2S] кластер и относятся к одному семейству Mar/Sox/Rob транскрипционных факторов. Обращает внимание нехарактерное для ТФ MarA удаленное расположение его потенциального сайта связывания от старта транскрипции.

Следует отметить, что в области $-35/-10$ промотора гена *dps* *E. coli* присутствует также хорошо выраженный консенсус сайта связывания стрессовой полимеразы RpoE – GGAACACatagccgggtctatacTtAAT (рис. 4) [25], экспрессия которой индуцируется, в том числе, и металлами [26]. Мы полагаем, что через эту полимеразу также может реализоваться ответ на кадмий.

Таблица 3. Сайты связывания транскрипционных факторов, выявленные методом SiteCon, в промоторе гена *dps* *E. coli*

ТФ	Параметры сайтов связывания			
	Позиции	P*	Уровень перепредсказания	Уровень недопредсказания
IHF	-109/-91	0.866	1/2082	0.8571
IscR (тип 2)	-97/-60	0.771	1/3843	0.5000
Rob	-97/-67	0.839	1/5260	0.8333
OxyR	-70/-38	0.835	1/50000	0.8235
MarA	-154/-124	0.766	1/2701	0.5625
SoxS	-7/-32	0.651	1/270	0.3333
MetJ	+4/+37	0.668	1/819	0.1875
PurR	-112/-90	0.804	1/1408	0.3333

Примечание: P* – уровень конформационного сходства с известными сайтами связывания.

2.4. Экспериментальный анализ структуры промотора гена *dps* *E. coli*

Гель-шифт анализ результатов взаимодействия нуклеотидных последовательностей, моделирующих потенциальные сайты связывания ТФ Rob, IscR и MarA в промоторе гена *dps* (табл. 3) с соответствующими рекомбинантными, очищенными белками ТФ представлен на рис. 3. Видно, что в промоторе гена *dps* *E. coli* присутствуют сайты связывания ТФ Rob (дорожка 6) и IscR (дорожка 4) и отсутствует сайт связывания ТФ MarA (дорожка 2).

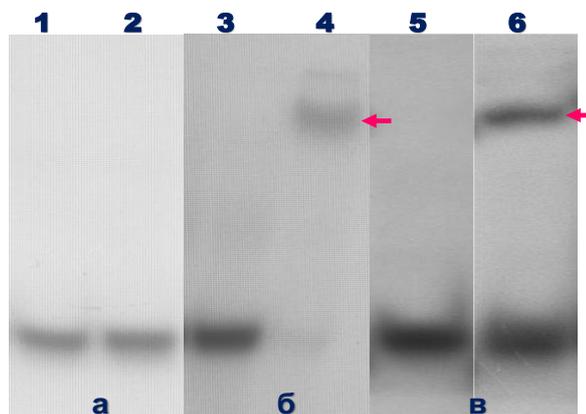


Рис. 3. Гель-шифт анализ взаимодействия последовательностей сайтов связывания транскрипционных факторов MarA (а), IscR (б) и Rob (в) в промоторе гена *dps* с 50 pmol белка ТФ MarA (2), IscR (4) и Rob (6) *E. coli*. Стрелками показано положение комплексов ДНК/белок с ТФ IscR (дорожка 4) и Rob (дорожка 6).

Для проверки функциональности потенциального сайта связывания RpoE полимеразы в промоторе гена *dps E. coli*, мы ввели мутации в -35 -область ее сайта (рис. 4) и сравнили динамику активности геносенсора с мутантным и нативным промотором в присутствии различных концентраций Cd^{2+} . Различий обнаружено не было (данные не показаны).

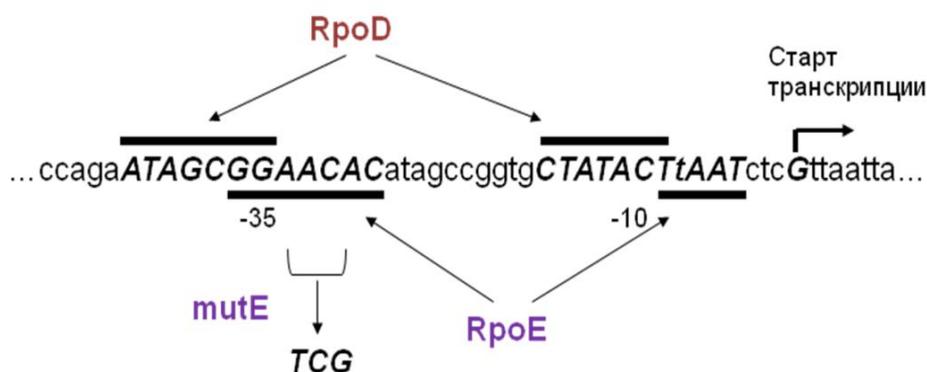


Рис. 4. Структура сайтов связывания RpoD и RpoE полимераз в промоторе гена *dps E. coli*. Показано расположение -35 - и -10 -областей сайтов связывания RpoD и RpoE полимераз относительно старта транскрипции и измененные нуклеотиды в -35 -области RpoE сайта.

Таким образом, в структуре промотора гена *dps E. coli* экспериментально подтверждено присутствие сайтов связывания ТФ Rob и IscR и не подтверждена функциональность потенциального сайта связывания RpoE полимеразы.

2.5. Реконструкция механизмов регуляции экспрессии гена *dps Escherichia coli* в присутствии ионов кадмия

В совокупности, согласно полученным и литературным данным, в промоторе гена *dps* присутствуют сайты связывания трех ТФ – Rob, IscR и OxyR, которые могут участвовать в реализации ответа гена *dps* на Cd^{2+} через взаимодействие с RpoD полимеразой. Все факторы могут быть как активаторами, так и репрессорами по отношению к своим генам-мишеням. Но если для ТФ OxyR известно, что он активирует транскрипцию гена *dps* [6], то в отношении ТФ IscR и Rob таких данных нет. Известно, что через сайты второго типа ТФ IscR может, как активировать, так и репрессировать транскрипцию своих генов-мишеней [27], поэтому решить вопрос о

действии ТФ IscR на экспрессию гена *dps* без дополнительных исследований нельзя. Ситуация с ТФ Rob несколько проще, так как он в основном является активатором своих генов-мишеней, а репрессорное действие проявляет только в случае расположения его сайта вблизи старта транскрипции. Достаточно удаленное расположение его сайта связывания в промоторе гена *dps* от старта транскрипции (табл. 2) позволяет предположить, что он может быть активатором этого гена.

В то же время, однозначно определить вклад ТФ Rob в реализацию ответа гена *dps* на кадмий без дополнительных экспериментов сложно, поскольку его сайт перекрывается с сайтом связывания ТФ IscR. Косвенно, исходя из результатов гель-шифт анализа (рис. 2,б и 2,в), согласно которым сродство ТФ Rob к своему сайту связывания выше, чем у ТФ IscR, можно предположить, что его реакция на кадмий может быть реализована при меньших концентрациях индуктора, чем у ТФ IscR.

Таким образом, мы полагаем, что активное состояние промотора гена *dps* в области высоких и низких концентраций Cd^{2+} может быть реализовано в первую очередь через ТФ OxyR и Rob, однако исключить действие ТФ IscR мы не можем.

Что касается репрессорного действия описанных выше ТФ на экспрессию гена *dps*, то такое предположение может быть сделано только для ТФ IscR, при этом у нас нет никаких данных о том, что он может также обладать селективной чувствительностью к различным концентрациям кадмия.

В то же время, из литературных данных известно, что в экспоненциальной фазе роста экспрессия гена *dps* *E. coli* негативно регулируется ТФ Fis и H-NS [10]. Имеют ли эти ТФ отношение к защите клеток от токсического воздействия кадмия? В работе Feriánc с соавторами [28] показано, что в период адаптации клеток *E. coli* к токсическому действию кадмия происходит остановка роста и деления клеток. Именно в этот период в клетках накапливается значительное количество белка ТФ H-NS, который исчезает после восстановления способности клеток к росту и делению [28]. Авторы исследовали эту ситуацию при концентрации кадмия в среде равной 273 μM и плотности культуры $OD_{420} = 0.5$. При более низкой плотности культуры эта концентрация кадмия была для клеток летальной [14]. В нашем исследовании ингибиторный эффект кадмия на экспрессию промотора гена *dps* *E. coli* наблюдался в среднелогарифмической фазе роста культуры и концентрации кадмия 0.25–0.5 мМ [13].

Сопоставление этих фактов позволили нам предположить, что именно транскрипционный фактор H-NS обеспечивает состояние репрессии промотора гена *dps* *E. coli*, однако только при тех концентрациях кадмия, при которых в период адаптации к токсическому эффекту кадмия происходит временная остановка роста и деления клеток.

В целом, учитывая результаты моделирования, согласно которым наиболее вероятным вариантом молекулярно-генетического механизма регуляции экспрессии гена *dps* *E. coli* в присутствии Cd^{2+} может быть неконкурентное взаимодействие трех ТФ, два из которых являются активаторами и один репрессором, мы считаем, что данный механизм ответа гена *dps* на кадмий вероятнее всего обеспечен взаимодействием транскрипционных факторов OxyR, Rob и H-NS с RpoD полимеразой. Их сайты связывания практически не пересекаются. Однако, исключить участие ТФ IscR, сайт связывания которого перекрывается с сайтом ТФ Rob, мы не можем.

Работа Хлебодаровой Т.М., Лихошвая В.А. и Ощепкова Д.Ю. поддержана Российским научным фондом (грант № 14-24-00123).

СПИСОК ЛИТЕРАТУРЫ

1. Martinez A., Kolter R. Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J. Bacteriol.* 1997. V. 179. P. 5188–5194.

2. Nair S., Finkel S.E. Dps protects cells against multiple stresses during stationary phase. *J. Bacteriol.* 2004. V. 186. P. 4192–4198.
3. Calhoun L.N., Kwon Y.M. Structure, function and regulation of the DNA-binding protein Dps and its role in acid and oxidative stress resistance in *Escherichia coli*: a review. *J. Appl. Microbiol.* 2011. V. 110. P. 375–386.
4. Khil P.P., Camerini-Otero R.D. Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. *Mol. Microbiol.* 2002. V. 44. P. 89–105.
5. Pomposiello P.J., Bennik M.H., Demple B. Genome-wide transcriptional profiling of the *Escherichia coli* responses to superoxide stress and sodium salicylate. *J. Bacteriol.* 2001. V. 183. P. 3890–3902.
6. Zheng M., Wang X., Templeton L.J., Smulski D.R., LaRossa R.A., Storz G. DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide. *J. Bacteriol.* 2001. V. 83. P. 4562–4570.
7. Phadtare S., Kato I., Inouye M. DNA microarray analysis of the expression profile of *Escherichia coli* in response to treatment with 4,5-dihydroxy-2-cyclopenten-1-one. *J. Bacteriol.* 2002. V. 184. P. 6725–6729.
8. Altuvia S., Almirón M., Huisman G., Kolter R., Storz G. The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol. Microbiol.* 1994. V. 13. P. 265–272.
9. Lomovskaya O.L., Kidwell J.P., Matin A. Characterization of the sigma 38-dependent expression of a core *Escherichia coli* starvation gene, *pexB*. *J. Bacteriol.* 1994. V. 176. P. 3928–3935.
10. Grainger D.C., Goldberg M.D., Lee D.J., Busby S.J. Selective repression by Fis and H-NS at the *Escherichia coli* *dps* promoter. *Mol. Microbiol.* 2008. V. 68. P. 1366–1377.
11. Yamamoto K., Ishihama A., Busby S.J., Grainger D.C. The *Escherichia coli* K-12 MntR miniregulon includes *dps*, which encodes the major stationary-phase DNA-binding protein. *J. Bacteriol.* 2011. V. 193. P. 1477–1480.
12. Khlebodarova T.M., Tikunova N.V., Kachko A.V., Stepanenko I.L., Podkolodny N.L., Kolchanov N.A. Application of bioinformatics resources for genosensor design. *J. Bioinform. Comput. Biol.* 2007. V. 5. P. 507–520.
13. Лихошвай В.А., Степанова Т.Ю., Задорожный А.В., Тикунова Н.В., Хлебодарова Т.М. Экспрессия гена *dps* *Escherichia coli* в присутствии токсических агентов: анализ и математическое моделирование. *Информационный вестник ВОГИС.* 2009. Т. 13. С. 731–740.
14. VanBogelen R.A., Kelley P.M., Neidhardt F.C. Differential induction of heat shock, SOS, and oxidation stress regulons and accumulation of nucleotides in *Escherichia coli*. *J. Bacteriol.* 1987. V. 169. P. 26–32.
15. Stohs S.J., Bagchi D. Oxidative mechanisms in the toxicity of metal ions. *Free Radic. Biol. Med.* 1995. V. 18. P. 321–336.
16. Stohs S.J., Bagchi D., Hassoun E., Bagchi M. Oxidative mechanisms in the toxicity of chromium and cadmium ions. *J. Environ. Pathol. Toxicol. Oncol.* 2001. V. 20. P. 77–88.
17. Valko M., Morris H., Cronin M.T. Metals, toxicity and oxidative stress. *Curr. Med. Chem.* 2005. V. 12. P. 1161–1208.
18. Likhoshvai V., Ratushny A. Generalized hill function method for modeling molecular processes. *J. Bioinform. Comput. Biol.* 2007. V. 5. № 2B. P. 521–531.
19. Oshchepkov D.Y., Vityaev E.E., Grigorovich D.A., Ignatieva E.V., Khlebodarova T.M. SITECON: a tool for detecting conservative conformational and physicochemical properties in transcription factor binding site alignments and for site recognition. *Nucl. Acids Res.* 2004. V. 32. P. W208–W212.

20. Nakajima H., Kobayashi K., Kobayashi M., Asako H., Aono R. Overexpression of the *robA* gene increases organic solvent tolerance and multiple antibiotic and heavy metal ion resistance in *Escherichia coli*. *Appl. Environ. Microbiol.* 1995. V. 61. P. 2302–2307.
21. Zheng M., Doan B., Schneider T.D., Storz G. OxyR and SoxRS regulation of *fur*. *J. Bacteriol.* 1999. V. 181. № 15. P. 4639–4643.
22. Khlebodarova T.M., Oshchepkov D.Yu., Tikunova N.V., Babkin I.V., Gruzdev A.D., Likhoshvai V.A. Reconstruction of the mechanisms that regulate the expression of the *Escherichia coli yfiA* gene under stress conditions. *Russ. J. Genet.: Applied Research.* 2013. V. 3. № 4. P. 271–278.
23. Cameron A.D., Redfield R.J. Non-canonical CRP sites control competence regulons in *Escherichia coli* and many other gamma-proteobacteria. *Nucl. Acids Res.* 2006. V. 34. P. 6001–6014.
24. Demple B. Redox signaling and gene control in the *Escherichia coli soxRS* oxidative stress regulon – a review. *Gene.* 1996. V. 179. P. 53–57.
25. Rezuchova B., Miticka H., Homerova D., Roberts M., Kormanec J. New members of the *Escherichia coli* sigma E regulon identified by a two-plasmid system. *FEMS Microbiol. Lett.* 2003. V. 225. P. 1–7.
26. Egler M., Grosse C., Grass G., Nies D.H. Role of the extracytoplasmic function protein family sigma factor RpoE in metal resistance of *Escherichia coli*. *J. Bacteriol.* 2005. V. 187. P. 2297–2307.
27. Giel J.L., Rodionov D., Liu M., Blattner F.R., Kiley P.J. IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O₂-regulated genes in *Escherichia coli*. *Mol. Microbiol.* 2006. V. 60. P. 1058–1075.
28. Ferianc P., Farewell A., Nyström T. The cadmium-stress stimulon of *Escherichia coli* K-12. *Microbiology.* 1998. V. 144. P. 1045–1050.

Материал поступил в редакцию 12.12.2014, опубликован 13.01.2015.