

УДК 575.113:577.322.4:577.214

РЕКОНСТРУКЦИЯ МЕХАНИЗМОВ РЕГУЛЯЦИИ ЭКСПРЕССИИ ГЕНА *yfiA* *ESCHERICHIA COLI* В УСЛОВИЯХ СТРЕССА

© 2013 г. Т.М. Хлебодарова¹, Д.Ю. Ощепков¹, Н.В. Тикунова³,
И.В. Бабкин³, А.Д. Груздев¹, В.А. Лихошвай^{1,2}

¹ Федеральное государственное бюджетное учреждение науки Институт цитологии и генетики
Сибирского отделения Российской академии наук, Новосибирск, Россия,
e-mail: tamara@bionet.nsc.ru;

² Новосибирский национальный исследовательский государственный университет,
Новосибирск, Россия;

³ Федеральное государственное бюджетное учреждение науки Институт химической биологии
и фундаментальной медицины Сибирского отделения РАН, Новосибирск, Россия

Поступила в редакцию 3 декабря 2012 г. Принята к публикации 20 февраля 2013 г.

С использованием компьютерного ресурса SiteCon и методов математического моделирования проведены реконструкция структуры регуляторной области гена *yfiA* *E. coli* и оценка сложности механизмов его экспрессии в условиях окислительного стресса. Моделирование динамики ответа на окислительный стресс клеток *E. coli*, трансформированных плазмидой pYfi-gfp, показало, что максимальное совпадение с экспериментальными данными наблюдается в модели, предполагающей комплексное воздействие, которое осуществляется, по-видимому, через ряд транскрипционных факторов (ТФ). Поиск в регуляторной области гена *yfiA* *E. coli* потенциальных сайтов связывания ТФ показал высокий уровень достоверности распознавания для ТФ MagA, IscR, MetJ, PurR и SoxS, которые прямо или косвенно могут участвовать в ответе этого гена на окислительный стресс, а также для ТФ CRP, глобального регулятора катаболизма углеводов. Гель-шифт-анализ с очищенными рекомбинантными белками ТФ CRP, MagA и SoxS *E. coli* подтвердил их присутствие в промоторе гена *yfiA* *E. coli*, что позволяет объяснить чувствительность этого гена к митомycinу и радикал-образующим агентам.

Ключевые слова: *Escherichia coli*, *yfiA*, GFP, регуляция транскрипции.

ВВЕДЕНИЕ

Ген *yfiA* *Escherichia coli* кодирует структуру белка pY (RaiA), который, как известно, стабилизирует рибосомы и участвует в регуляции элонгации трансляции в условиях стресса (Maki *et al.*, 2000; Agafonov *et al.*, 2001; Agafonov, Spirin, 2004; Wilson, Nierhaus, 2004; Vila-Sanjurjo *et al.*, 2004; Ueta *et al.*, 2005). Данные микрочип-эксперимента свидетельствуют о чувствительности гена *yfiA* *E. coli* к различным внешним воздействиям (Pomposiello *et al.*, 2001; Zheng *et al.*, 2001; Khil, Camerini-Otero, 2002; Schembri *et al.*, 2003). К настоящему времени в литературе нет исследований структуры промотора гена *yfiA* *E. coli*. В связи с тем, что ATG кодон гена *yfiA*

расположен на расстоянии 35 нуклеотидов от конца рамки считывания впередилежащего гена *b2596*, предполагалось, что он входит в структуру оперона, включающего также и ген *ectD* (b2595), так как расстояние между последними также не превышало 30 нуклеотидов (EMBL/GenBank:U00096). Существование оперона с единой регуляцией для трех вышеперечисленных генов было подвергнуто нами сомнению на основании результатов работы К. Salmon с соавт. (Salmon *et al.*, 2003). В ней было показано, что гены *ectD* и *yfiA* по-разному реагируют на отсутствие кислорода и присутствие транскрипционного фактора FNR. В то же время экспрессии гена *b2596*, расположенного между генами *ectD* и *yfiA*, обнаружено не было (Salmon *et al.*,

2003). Кроме того, ранее при исследовании экспрессии оперона *rheLA* был обнаружен старт транскрипции, расположенный на расстоянии 30 нуклеотидов от начала старта трансляции впередилежащего гена *URF1* (синоним *yfiA*) практически в пределах гена *b2596* (Hudson *et al.*, 1984). Эти данные позволили нам поставить под сомнение существование гена *b2596* и предположить наличие независимого промотора для гена *yfiA* в пределах последовательности от конца гена *ectD* до ATG кодона гена *yfiA* (Tikunova *et al.*, 2007). Последние данные по аннотации генома *E. coli* также поставили существование гена *b2596* под сомнение (Riley *et al.*, 2006).

Ранее, исходя из вышеперечисленных данных, мы выделили потенциальную регуляторную область гена *yfiA* в пределах от -242 до +38 от старта транскрипции, обнаруженного в работе Hudson с соавт. (1984). На основе этой последовательности и репортерного гена *gfp* мы создали геносенсор, который отвечал, в частности, на оксидативный стресс (H_2O_2) и специфическое повреждение структуры ДНК (митомицин C) (Tikunova *et al.*, 2007), однако механизмы реализации этих ответов оставались неизвестными. Данные микрочип-эксперимента также свидетельствуют о чувствительности гена *yfiA E. coli* к перекиси водорода, особенно в *OxyR*-мутантах (Schembri *et al.*, 2003), митомицину C (Ueta *et al.*, 2005) и кислородному голоданию, которое зависит от присутствия транскрипционного фактора FNR (Salmon *et*

al., 2003). Мы использовали компьютерные методы анализа промоторной области гена *yfiA E. coli*, метод аппроксимации экспериментальных данных рациональными полиномами для анализа сложности механизмов экспрессии гена *yfiA E. coli* в условиях окислительного стресса и гель-шифт-анализ для экспериментального подтверждения ряда выявленных потенциальных регуляторов этого гена.

МЕТОДЫ

Компьютерные методы. Для анализа структуры потенциальной регуляторной области гена *yfiA E. coli* был использован метод SiteCon, позволяющий эффективно распознавать сайты связывания транскрипционных факторов (ССТФ). Метод заключается в выявлении набора консервативных контекстно зависимых конформационных и физико-химических свойств, определенных для позиций выравнивания ССТФ и дальнейшего сравнения выявленных консервативных свойств со свойствами анализируемой последовательности (Oshchepkov *et al.*, 2004). Для распознавания потенциальных сайтов в промоторе гена *yfiA* были созданы обучающие выборки экспериментально доказанных ССТФ PurR, MetJ, MarA, SoxS, IscR и Crp. Объем выборки для каждого ТФ и длина сайтов приведены в табл. 1. Сайты были выровнены относительно наиболее часто встречающихся нуклеотидов и использованы

Таблица 1

Потенциальные сайты связывания транскрипционных факторов, выявленные методом SiteCon, в структуре промотора гена *yfiA E. coli*

ТФ	Длина сайтов в выборке	Объем выборки	Параметры потенциальных сайтов связывания			
			позиции	P*	уровень недо-предсказания	уровень пере-предсказания
PurR	31	16	-125/-98	0,819	0,2500	1/2776
MetJ	23	24	-111/-91	0,737	0,5000	1/11106
MarA	31	16	-78/-47	0,827	0,7500	1/19987
SoxS	30	18	-74/-48	0,740	0,5556	1/1851
IscR	42	4	-73/-32	0,756	0,2500	1/2776
Crp	60	39	-51/-30	0,830	0,9231	1/100000
MetJ	23		-31/-12	0,712	0,4583	1/3570

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

для анализа. Для оценки достоверности распознавания нами были оценены величины ошибок первого и второго рода, недопредсказания и перепредсказания соответственно. При этом контроль качества распознавания для ошибки 1-го рода был произведен по методике jack-knife с последовательным исключением одной последовательности из выборки и обучением по оставшейся части выборки. Контроль ошибки второго рода производился путем распознавания сайтов связывания в последовательности длиной 100 т.п.о., которая была создана путем случайного перемешивания нуклеотидов обучающей выборки. Таким образом, нуклеотидный состав негативной и обучающей выборок был одинаковым. Для каждой выборки значения рассчитывались в виде таблицы ошибок, где в соответствие порогу распознавания были поставлены значения ошибок 1-го и 2-го родов. Для оценки достоверности каждого обнаруженного сайта брались значения ошибок 1-го и 2-го родов из таблицы, посчитанной для данного типа сайтов, и для порога, на котором был обнаружен соответствующий сайт. Исходя из этого мы считаем, что чем ниже значение ошибки 2-го рода (перепредсказания) для сайта, тем меньше вероятность обнаружить данный сайт на этом пороге в последовательности данной длины по случайным причинам и тем выше достоверность его распознавания.

Метод моделирования. Для теоретической оценки сложности механизма H_2O_2 -зависимой регуляции эффективности функционирования промоторов генов *yfiA* и *katG* *E. coli* был использован метод аппроксимации экспериментальных данных рациональными полиномами, описанный ранее (Лихошвай и др., 2009).

При оценке параметров функционирования промотора гена *yfiA* *E. coli* использовали экспериментальные значения флюоресценции клеток линии MC4100 *E. coli*, трансформированных репортерной плазмидой pYfi-gfp, в момент достижения максимального ответа на стресс через 60 мин после начала воздействия H_2O_2 , которые были получены ранее (Tikunova *et al.*, 2007). Интенсивность флюоресценции репортерного белка GFP при разных концентрациях перекиси водорода интерпретировали как величину относительной активности промотора. В силу условности шкалы измерения уровня флюо-

ресценции белка GFP экспериментальные точки, соответствующие временному срезу 60 мин, были нормированы на величину интенсивности флюоресценции при концентрации перекиси водорода, равной 0,5 mM.

В качестве контрольного был использован геносенсор на основе промотора гена *katG* *E. coli*, разработанный нами ранее на той же базовой основе, что и геносенсор *E. coli*/pYfi-gfp. При оценке параметров были использованы собственные экспериментальные данные (Khlebodarova *et al.*, 2007). Значения взяты в момент максимального ответа на стресс через 40 мин после начала воздействия перекиси. Экспериментальные точки, соответствующие временному срезу 40 мин, были нормированы на величину фоновой интенсивности флюоресценции клеток в отсутствие перекиси водорода.

Экспериментальные методы. Для получения рекомбинантных белков транскрипционных факторов Crp, SoxS, FNR и MarA *E. coli* на основе известных нуклеотидных последовательностей этих генов (GenBank: U00096) были рассчитаны праймеры для ПЦР: 5'-CTCTCCATGGTGCTTGGCAAACCGC-3' и 5'-GAGAGGGATCCACGAGTGCCGTAACGACGATG-3', 5'-AAAACCATGGATGTCCCATCAGAAAATATTCAGG-3' и 5'-CTTTGGATCC TAGGCGGTGGCGATAATCG-3', 5'-CTGGACATGTTAAAATTGACAAATATCAATTACGG-3' и 5'-CTGGAGATCTGGCAACGTTACGCGTAGACCA-3', 5'-CTCCTCATGACGATGTCCA GACGCAATACTGA-3' и 5'-TCCTAGATCTGCTGTTGTAATGATTTAATGGATGTA AAAAG-3' соответственно.

Праймеры для ТФ Crp и SoxS включали сайты рестрикции *Bam*HI и *Bsp*19I, а для FNR и MarA – *Bgl*II, *Cci*I или *Pci*I. Полученные амплифицированные фрагменты были расчетной длины и клонированы в вектор pOPE-110 (Crp, SoxS) или pQE60 (FNR, MarA) по соответствующим сайтам. Рекомбинантные белки были наработаны и очищены с помощью Ni-хелатной хроматографии согласно протоколу производителя (QIAGEN).

Гель-шифт-анализ с рекомбинантными очищенными белками проводили, как описано (Cameron, Redfield, 2006). Олигонуклеотидный дуплекс, моделирующий потенциальный сайт связывания ТФ MarA в промоторе гена *yfiA*

E. coli, был синтезирован и соответствует последовательности 5'-AAGATTCGTTGACAAAAAGTGACAAAATTAT в позиции -78/-47 от старта транскрипции.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Теоретический анализ потенциального регуляторного района гена *yfiA* *E. coli*

Как указывалось выше, проведенные ранее исследования (Tikunova *et al.*, 2007) подтвердили наше предположение о существовании собственного промотора у гена *yfiA* *E. coli* в пределах выделенной нами последовательности от конца открытой рамки считывания впередилежащего гена *ectD* (b2595) до ATG-кодона гена *yfiA*. Однако вопрос о существовании регуляции экспрессии *yfiA* (b2597) со стороны промотора гена *ectD* (b2595) оставался открытым. Дополнительно мы провели исследование нуклеотидной последовательности между генами *ectD* и *yfiA* с использованием Интернет-доступного пакета программ RNASHapes (Steffen *et al.*, 2006). На расстоянии 30 нуклеотидов от конца открытой рамки считывания гена *ectD* был обнаружен канонический Rho-независимый терминатор транскрипции, структура которого

(AAAACGGCAGCcccttgaGCTGCCGTTTT) позволяет формирование потенциальной шпильки с энергией 16,7 ккал/моль. Это значение достаточно хорошо соответствует энергии наиболее часто встречающихся экспериментально идентифицированных Rho-независимых терминаторов транскрипции у *E. coli* (de Hoop *et al.*, 2005) и в совокупности с данными микрочип-эксперимента (Salmon *et al.*, 2003) позволяет предположить отсутствие регуляции гена *yfiA* со стороны впередилежащего гена *ectD* (b2595).

Таким образом, потенциальная промоторная область гена *yfiA* *E. coli* составляет 230 пар нуклеотидов от конца сайта терминации транскрипции гена *ectD* до ATG-кодона гена *yfiA* и включает открытую рамку считывания гена b2596. Структура этой области представлена на рис. 1.

Клонированная последовательность потенциальной промоторной области гена *yfiA* *E. coli* включает часть Rho-независимого потенциального терминатора транскрипции впередилежащего гена *ectD* (Tikunova *et al.*, 2007).

Модель функционирования геносенсора *E. coli/pYfi-gfp* в условиях окислительного стресса

Данные микрочип-эксперимента свидетельствуют о том, что высокая чувствительность гена *yfiA* *E. coli* к перекиси водорода не связана с присутствием ТФ ОхуR, который является основным сенсором окислительного стресса. Более того, у ОхуR-мутантов уровень индукции гена *yfiA* перекисью водорода существенно выше такового у дикого типа (Zheng *et al.*, 2001). Для оценки сложности механизма регуляции активности гена *yfiA* *E. coli* в условиях окислительного стресса были использованы собственные экспериментальные данные по влиянию различных концентраций H₂O₂ на флюоресценцию клеток геносенсора *E. coli/pYfi-gfp* (Tikunova *et al.*, 2007), которые были нормированы, как описано в разделе «Методы». Их численные значения приведены в табл. 2.

Для описания эффективности функционирования промотора гена *yfiA* *E. coli* под действием H₂O₂ использовали рациональный полином вида:

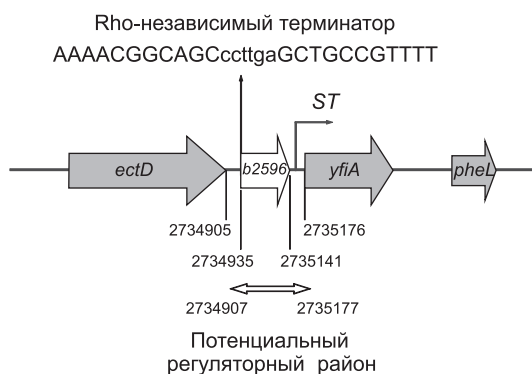


Рис. 1. Структура потенциального регуляторного района гена *yfiA* *E. coli*.

Указано расположение открытых рамок считывания генов *ectD*, *b2596*, *yfiA* и *pheL* старта транскрипции (СТ), обнаруженного в работе Hudson с соавт. (1984), потенциального Rho-независимого терминатора транскрипции гена *ectD* и границы потенциального промотора гена *yfiA*. Позиции указаны в координатах полного генома *E. coli* (GenBank: U00096).

$$V_{n,m}(s) = \left[1 + \frac{w + v \left[\frac{s}{k_1} \right]^n}{\left[1 + \left[\frac{s}{k_1} \right]^n \right] \left[1 + \left[\frac{s}{k_2} \right]^m \right]} \right] / \left[1 + \frac{w + v \left[\frac{s_0}{k_1} \right]^n}{\left[1 + \left[\frac{s_0}{k_1} \right]^n \right] \left[1 + \left[\frac{s_0}{k_2} \right]^m \right]} \right], \quad (1)$$

где n и m – параметры, которые характеризуют комплексность влияния перекиси на активацию и репрессию промотора соответственно; k_1, k_2 – константы, которые имеют размерность концентрации и определяют эффективность влияния перекиси на активность промотора; w – отношение активности промотора в отсутствие перекиси водорода к величине фонового сигнала флюоресценции; v – отношение константы эффективности инициации транскрипции активированного промотора к величине фонового сигнала флюоресценции. Величины параметров n и m характеризуют совокупную сложность механизма регуляции активности промотора перекисью водорода. Чем выше значения n и m , тем большая сложность механизма прогнозируется.

Результаты аппроксимации модели (1) к экспериментальным данным (табл. 2) приведены на рис. 2. Видно, что при $n = 1,3$ и $m = 2,8$ функция (1) хорошо описывает динамику ответа **pYfi-gfp на стресс. Отсюда следует**, что активность промотора гена *yfiA* *E. coli* регулируется через комплексный механизм и осуществляется, по-видимому, через несколько ТФ, которые могут действовать как активаторы и как ингибиторы.

Биохимический смысл значений n и m на данном этапе анализа не может быть одно-

значно распознан. Они могут отражать и количество сайтов связывания транскрипционных факторов, через которые действует перекись, и возможную каскадность механизма регуляции, и синергические взаимодействия транскрипционных факторов. Данные, использованные в настоящей работе, не позволяют разделить эти гипотезы.

Косвенно обоснованность того, что механизм ингибирования промотора гена *yfiA* в присутствии перекиси водорода связан с молекулярно-генетическими процессами, а не является следствием гибели клеток в исследуемом интервале концентраций, можно продемонстрировать при оценке предложенным методом сложности механизма регуляции экспрессии гена *katG* *E. coli* в условиях оксидативного стресса, который, как известно, осуществляется через OxyR-зависимый механизм (Tartaglia *et al.*, 1989).

Для этой оценки использованы собственные экспериментальные данные по влиянию различных концентраций H_2O_2 на флюоресценцию клеток *E. coli/pKat-gfp* геносенсора (Khlebodova *et al.*, 2007), которые были нормированы, как описано в разделе «Методы». Их численные значения приведены в табл. 2.

Таблица 2

Уровни индукции флюоресценции клеток геносенсоров *E. coli/pYfi-gfp* и *E. coli/pKat-gfp* различными концентрациями перекиси водорода

Концентрация H_2O_2 (мМ)	Значение сигнала (усл.ед)	
	<i>E. coli/pYfi-gfp</i>	<i>E. coli/pKat-gfp</i>
0	0,32	1
0,5	1	2,17
1	1,36	2,92
2	0,86	3,5
4	1,89	4,0
8	1,13	4,7

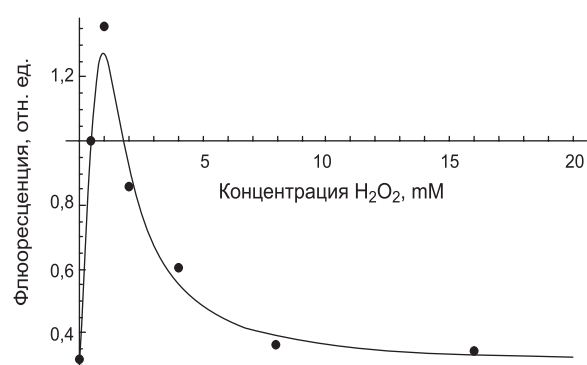


Рис. 2. Аппроксимация динамики экспрессии гена *yfiA* в условиях окислительного стресса.

Сплошная линия – теоретическая кривая, точки – экспериментальные значения, отражающие относительную активность функционирования промотора *yfiA* через 60 мин после воздействия различных концентраций H_2O_2 . Расчеты выполнены при $w = 0$, $v = 111$, $k_1 = 8,1$, $k_2 = 0,97$, $n = 1,3$, $m = 2,8$.

Результаты аппроксимации модели (1) к экспериментальным данным приведены на рис. 3. Видно, что при $n = 1$ и $m = 0$ функция (1) хорошо описывает динамику ответа $p_{Kat-gfp}$ на стресс, что согласуется с известным фактом наличия в данном промоторе сайта ТФ OxyR, через который и осуществляется активация промотора перекисью водорода (Tartaglia *et al.*, 1989).

Видно также, что концентрации перекиси водорода выше 2,0 мМ не снижают уровень ответа геносенсора *E. coli/pKat-gfp* на присутствие токсического агента. Эти данные косвенно свидетельствуют о специфичности ответа геносенсора *E. coli/pYfi-gfp* на воздействие H_2O_2 в интервале концентраций от 2,0 до 8 мМ (рис. 2), однако для выяснения конкретных механизмов его ингибирования в этом интервале концентраций H_2O_2 необходимы дополнительные исследования.

Анализ структуры потенциального промотора гена *yfiA* *E. coli* методом SiteCon

Поскольку на основании математического анализа был сделан вывод о том, что регуляция экспрессии гена *yfiA* в условиях окислительного стресса носит сложный комплексный характер, был проведен поиск в последовательности промотора гена *yfiA* потенциальных сайтов связывания ТФ, которые прямо или косвенно могут участвовать в этом ответе. Наибольший интерес для нас представляли сайты, достоверность распознавания которых достаточно высока. Мы оценили этот уровень по ошибке второго рода (уровню перепредсказания) и ограничили его снизу. 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. В дальнейшем рассматривали только те потенциальные сайты, уровень перепредсказания которых был меньше, чем 1/2400 нуклеотидов при длине анализируемой последовательности 240 нуклеотидов. Исключение было сделано для ТФ SoxS, достоверность предсказания которого была несколько ниже этой оценки. В

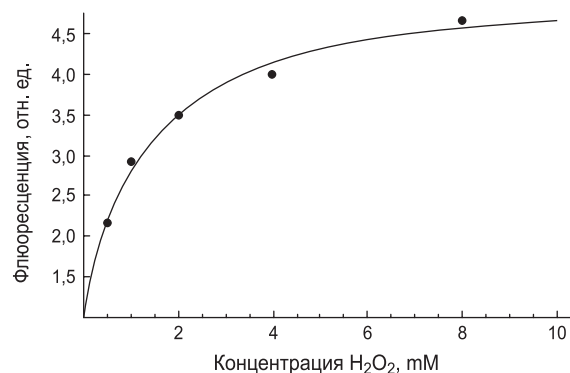


Рис. 3. Соответствие функции $V_{n,m}$ динамике экспрессии гена *katG* в условиях окислительного стресса при предположении активации его промотора через один сайт ($n = 1$, $m = 0$).

Сплошная линия — теоретическая кривая, точки — экспериментальные значения экспрессии промотора *katG* *E. coli* через 40 мин после воздействия различных концентраций H_2O_2 . Расчеты выполнены при $k_1 = 1,275$, $w = 0,6$, $v = 7,23$.

табл. 1 приведен список потенциальных сайтов связывания ТФ в промоторе гена *yfiA*, которые удовлетворяют вышеозначенным ограничениям. Таких сайтов оказалось 7, и принадлежат они 6 ТФ: MarA, IscR, MetJ, PurR, CRP и SoxS. Транскрипционные факторы MarA, IscR, MetJ, PurR и SoxS прямо или косвенно участвуют в ответе на окислительный стресс, а ТФ Crp является глобальным регулятором генов катаболизма углеводов и в целом может влиять на уровень стрессового ответа в зависимости от присутствия в среде глюкозы и уровня внутриклеточного сАМФ (Ishizuka *et al.*, 1994).

В табл. 1 сайты расположены в порядке их удаленности от старта транскрипции. Согласно данным табл. 1, выявленные потенциальные сайты связывания ТФ имеют достаточно высокий уровень конформационного сходства с известными экспериментально подтвержденными сайтами и небольшую величину вероятности их выявления по случайным причинам. Среди выявленных сайтов максимальный уровень достоверности показан для сайтов связывания ТФ Crp, MarA и MetJ, а наименьший — для IscR, PurR и SoxS.

Из данных табл. 1 следует, что центр сайта связывания Crp расположен на расстоянии -40 нуклеотидов от предполагаемого старта транскрипции. Эта позиция близка к позиции (-41,5), характерной для Crp-зависимых про-

моторов класса II (Rhodius *et al.*, 1997; Savery *et al.*, 1998). Совпадение позиций сайтов связывания для ТФ MarA и SoxS неудивительно, так как эти транскрипционные факторы имеют близкую структуру сайтов связывания, которая называется *marbox* (Martin *et al.*, 1999). Положение центра потенциального сайта MarA относительно старта транскрипции находится в позиции –63, а сайта SoxS – в позиции –61,5, что соответствует структуре *mar/sox*-зависимых промоторов класса I (Martin *et al.*, 2002). Что касается ТФ IscR, то он содержит в своей структуре [2Fe-2S] кластер, подобно семейству Mar/Sox/Rob транскрипционных факторов, а его последовательность имеет сходство с последовательностью ТФ MarA (Schwartz *et al.*, 2001). В связи с этим перекрывание его потенциального сайта связывания с последовательностью *marbox* также возможно.

В промоторной области гена *yfiA* выявлено три сайта связывания для транскрипционных репрессоров MetJ и PurR. Обнаружение в промоторе гена *yfiA* специфических регуляторов пуринового и метионинового биосинтеза кажется маловероятным, но высокая достоверность сайта MetJ в позиции –101 относительно старта транскрипции и характерное для репрессоров расположение второго сайта между боксами –35 и –10 РНК полимеразы требуют экспериментальной проверки возможности наличия этих сайтов в структуре гена *yfiA*. Анализ известных механизмов регуляции транскрипции этих двух репрессоров выявил следующую особенность. Их транскрипция репрессируется ТФ Fur, а экспрессия Fur прямо зависит от активности ТФ OxyR, SoxS и Crp (De Lorenzo *et al.*, 1988; Zheng *et al.*, 1999). Подтверждение существования этих сайтов в промоторе гена *yfiA* может свидетельствовать о наличии не прямой регуляции экспрессии гена *yfiA* в условиях оксидативного стресса.

Таким образом, комплексная чувствительность экспрессии гена *yfiA* к перекиси водорода может быть объяснена наличием как прямой активации промотора гена *yfiA* окисляющими агентами через сайты связывания ТФ SoxS и IscR, которые чувствительны к радикал-образующим агентам, так и не прямой его активацией через сайты связывания ТФ MarA, MetJ и PurR, экспрессия которых в присутствии перекиси

водорода зависит от сенсоров окислительного стресса ТФ SoxS и/или OxyR (Dempse, 1996; Zheng *et al.*, 1999).

Наличие сайта связывания Crp свидетельствует о возможности прямой зависимости уровня стрессового ответа клеток от источника питания и той роли, которую ТФ Crp играет в регуляции экспрессии *yfiA* (Ishizuka *et al.*, 1994). Более того, положение этого сайта в промоторе позволяет предположить возможность функционального взаимодействия Crp с IscR, SoxS и MarA транскрипционными факторами (табл. 1).

Экспериментальный анализ структуры промотора гена *yfiA E. coli*

На первом этапе для анализа были выбраны потенциальные сайты ТФ Crp, MarA, FNR и SoxS. Сайты связывания ТФ Crp и MarA были выявлены нами в промоторе гена *yfiA* с очень высоким уровнем достоверности: 1/100000 и 1/19987 соответственно. Предпосылкой для проверки сайта связывания ТФ FNR послужили данные микрочип-эксперимента, в которых экспрессия гена *yfiA* зависит от присутствия этого фактора, и тот факт, что в его потенциальной регуляторной области находится последовательность, близкая по структуре сайту связывания ТФ FNR (Salmon *et al.*, 2003). Уровень распознавания этой последовательности как сайта связывания ТФ FNR методом SITECON составил 1/549, что значительно выше выбранного нами уровня ограничения. Потенциальный сайт ТФ SoxS представлял интерес с двух точек зрения. Во-первых, он был выявлен в позиции –61,5, что соответствует структуре *mar/sox*-зависимых промоторов класса I (Martin *et al.*, 2002). Во-вторых, уровень его распознавания был чуть выше выбранного нами ограничения. В совокупности анализ этих сайтов позволял уточнить критический уровень распознавания метода SITECON.

На рис. 4 представлены результаты взаимодействия последовательности промотора гена *yfiA E. coli* (а, б, в) и последовательности потенциального сайта связывания MarA (AAG ATTCGTTGACAAAAAGTGACAAAATTAT) из промотора гена *yfiA* (г) в позиции –78/–47 с рекомбинантными очищенными белками ТФ SoxS (2), FNR (4), Crp (6) и MarA (7).

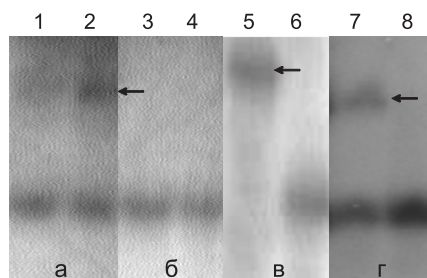


Рис. 4. Результаты анализа взаимодействия промотора гена *uflA* *E. coli* (а, б, в) и олигонуклеотидного дуплекса, моделирующего сайт связывания ТФ MagA гена *uflA* (г), с очищенными белками ТФ SoxS (2), FNR (4), Crp (5) и MagA (7) *E. coli* методом EMSA.

Стрелками показано положение комплексов ДНК/белок с ТФ SoxS (дорожка 2), Crp (дорожка 5) и MagA (дорожка 7).

Видно отсутствие сайта связывания ТФ FNR промоторе гена *uflA* (рис. 4, б, дорожка 4), что совпадает с низкой оценкой уровня его распознавания методом SITECON (1/549) и позволяет предположить не прямые механизмы его влияния на экспрессию гена *uflA*. Действие FNR на экспрессию гена *uflA* в анаэробных условиях может быть связано с участием FNR в активации транскрипции гена ТФ ArcA (Compan *et al.*, 1994), который репрессирует транскрипцию стрессовой *groS*-полимеразы (Mika, Hengge, 2005). Участие *groS*-полимеразы в транскрипции гена *uflA* не показано, однако известно, что продукт гена *uflA* появляется в клетке при изменении плотности культуры и в ответ на холодовый шок (Agafonov *et al.*, 2001; Wilson *et al.*, 2004). Эти стрессовые воздействия реализуются через *groS*-полимеразу (Takayanagi *et al.*, 1994) и позволяют предположить ее участие в транскрипции гена *uflA*.

Что касается потенциальных сайтов связывания ТФ Crp, MagA и SoxS, то результаты, представленные на рис. 4, подтверждают их наличие в структуре промотора *uflA* *E. coli* и позволяют объяснить чувствительность этого гена к митомцину и окислительным агентам (Tikunova *et al.*, 2007).

Также полученные нами результаты позволяют не рассматривать в качестве потенциальных те сайты, которые выявлены методом SiteCon с уровнем перепредсказания выше 1/1800.

Таким образом, теоретически было показано, что сложная динамика экспрессии гена

uflA *E. coli* в условиях окислительного стресса является отражением комплексного воздействия нескольких ТФ. Подтверждением этому было выявление в структуре его регуляторного района потенциальных ССТФ MagA, IscR, MetJ, PurR и SoxS, которые прямо или косвенно могут участвовать в ответе этого гена на окислительный стресс. Наличие сайтов связывания ТФ SoxS и MagA подтверждено экспериментально. Также выявлено, что экспрессия гена *uflA* *E. coli* может зависеть от присутствия в среде глюкозы и уровня внутриклеточного сАМФ, в связи с наличием в структуре его промотора экспериментально подтвержденного сайта связывания ТФ CRP.

БЛАГОДАРНОСТИ

Работа выполнена при поддержке РФФИ (проект 10-01-00717), Программ Президиума РАН «Молекулярная и клеточная биология» (проект 6.8) и СО РАН (интеграционный проект № 80) и гранта НШ-5278.2012.4.

ЛИТЕРАТУРА

- Лихошвай В.А., Степанова Т.Ю., Задорожный А.В. и др. Экспрессия гена *dps* *Escherichia coli* в присутствии токсических агентов: анализ и математическое моделирование // Информ. вестник ВОГиС. 2009. Т. 13. № 4. С. 731–740.
- Agafonov D.E., Kolb V.A., Spirin A.S. Ribosome-associated protein that inhibits translation at the aminoacyl-tRNA binding stage // EMBO Rep. 2001. V. 2. P. 399–402.
- Agafonov D.E., Spirin A.S. The ribosome-associated inhibitor A reduces translation errors // Biochem. Biophys. Res. Commun. 2004. V. 320. P. 354–358.
- 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.
- Compan I., Touati D. Anaerobic activation of *arcA* transcription in *Escherichia coli*: roles of Fnr and ArcA // Mol. Microbiol. 1994. V. 11. P. 955–964.
- de Hoon M.J., Makita Y., Nakai K., Miyano S. Prediction of transcriptional terminators in *Bacillus subtilis* and related species // PLoS Comput Biol. 2005. V. 1. No. 3. P. e25.
- De Lorenzo V., Herrero M., Giovannini F., Neilands J.B. FUR (ferric uptake regulation) protein and CAP (catabolite-activator protein) modulate transcription of *fur* gene in *Escherichia coli* // Eur. J. Biochem. 1988. V. 173. P. 537–546.
- Demple B. Redox signaling and gene control in the *Escherichia coli* *soxRS* oxidative stress regulon – a review // Gene. 1996. V. 179. No. 1. P. 53–57.
- Hudson G.S., Davidson B.E. Nucleotide sequence and transcription of the phenylalanine and tyrosine operons

- of *Escherichia coli* K12 // J. Mol. Biol. 1984. V. 180. P. 1023–1051.
- Ishizuka H., Hanamura A., Inada T., Aiba H. Mechanism of the down-regulation of cAMP receptor protein by glucose in *Escherichia coli*: role of autoregulation of the *crp* gene // The EMBO J. 1994. V. 13. P. 3077–3082.
- 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.
- Khlebodarova T.M., Tikunova N.V., Kachko A.V. *et al.* Application of bioinformatics resources for genosensor design // J. Bioinform. Comput. Biol. 2007. V. 5. P. 507–520.
- Maki Y., Yoshida H., Wada A. Two proteins, YfiA and YhbH, associated with resting ribosomes in stationary phase *Escherichia coli* // Genes Cells. 2000. V. 5. P. 965–974.
- Martin R.G., Gillette W.K., Rhee S., Rosner J.L. Structural requirements for marbox function in transcriptional activation of *mar/sox/rob* regulon promoters in *Escherichia coli*: sequence, orientation and spatial relationship to the core promoter // Mol. Microbiol. 1999. V. 34. P. 431–441.
- Martin R.G., Rosner J.L. Genomics of the *marA/soxS/rob* regulon of *Escherichia coli*: identification of directly activated promoters by application of molecular genetics and informatics to microarray data // Mol. Microbiol. 2002. V. 44. P. 1611–1624.
- Mika F., Hengge R. A two-component phosphotransfer network involving ArcB, ArcA, and RssB coordinates synthesis and proteolysis of sigmaS (RpoS) in *E. coli* // Genes Dev. 2005. V. 19. No. 22. P. 2770–2781.
- Oshchepkov D.Y., Vityaev E.E., Grigorovich D.A. *et al.* SITE-CON: 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.
- Pomposiello P.J., Bennis 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.
- Rhodus V.A., West D.M., Webster C.L. *et al.* Transcription activation at class II CRP-dependent promoters: the role of different activating regions // Nucl. Acids Res. 1997. V. 25. P. 326–332.
- Riley M., Abe T., Arnaud M.B. *et al.* *Escherichia coli* K-12: a cooperatively developed annotation snapshot—2005 // Nucl. Acids Res. 2006. V. 34. No. 1. P. 1–9.
- Salmon K., Hung S.P., Mekjian K. *et al.* Global gene expression profiling in *Escherichia coli* K12. The effects of oxygen availability and FNR // J. Biol. Chem. 2003. V. 278. P. 29837–29855.
- Savery N.J., Lloyd G.S., Kainz M. *et al.* Transcription activation at Class II CRP-dependent promoters: identification of determinants in the C-terminal domain of the RNA polymerase alpha subunit // The EMBO J. 1998. V. 17. No. 12. P. 3439–3447.
- Schembri M.A., Kjaergaard K., Klemm P. Global gene expression in *Escherichia coli* biofilms // Mol. Microbiol. 2003. V. 8. P. 253–267.
- Schwartz C.J., Giel J.L., Patschkowski T. *et al.* IscR, an Fe-S cluster-containing transcription factor, represses expression of *Escherichia coli* genes encoding Fe-S cluster assembly proteins // Proc. Natl Acad. Sci. USA. 2001. V. 98. No. 26. P. 14895–14900.
- Steffen P., Voss B., Rehmsmeier M. *et al.* RNAsHapes: an integrated RNA analysis package based on abstract shapes // Bioinformatics. 2006. V. 22. No. 4. P. 500–503.
- Takayanagi Y., Tanaka K., Takahashi H. Structure of the 5' upstream region and the regulation of the *rpoS* gene of *Escherichia coli* // Mol. Gen. Genet. 1994. V. 243. P. 525–531.
- Tartaglia L.A., Storz G., Ames B.N. Identification and molecular analysis of oxyR-regulated promoters important for the bacterial adaptation to oxidative stress // J. Mol. Biol. 1989. V. 210. P. 709–719.
- Tikunova N.V., Khlebodarova T.M., Kachko A.V. *et al.* A computational-experimental approach to designing a polyfunctional genosensor derived from the *Escherichia coli* gene *yfiA* promoter // Dokl. Biochem. Biophys. 2007. V. 417. No. 6. P. 357–361.
- Ueta M., Yoshida H., Wada C. *et al.* Ribosome binding proteins YhbH and YfiA have opposite functions during 100S formation in the stationary phase of *Escherichia coli* // Genes Cells. 2005. V. 10. No. 12. P. 1103–1112.
- Vila-Sanjurjo A., Schuwirth B.S., Hau C.W., Cate J.H. Structural basis for the control of translation initiation during stress // Nat. Struct. Mol. Biol. 2004. V. 11. No. 11. P. 1054–1059.
- Wilson D.N., Nierhaus K.H. The how and Y of cold shock // Nat. Struct. Mol. Biol. 2004. V. 11. No. 11. P. 1026–1028.
- Zheng M., Doan B., Schneider T.D., Storz G. OxyR and SoxRS regulation of *fur* // J. Bacteriol. 1999. V. 181. No. 15. P. 4639–4643.
- Zheng M., Wang X., Templeton L.J. *et al.* DNA microarray-mediated transcriptional profiling of the *Escherichia coli* response to hydrogen peroxide // J. Bacteriol. 2001. V. 183. No. 15. P. 4562–4570.

RECONSTRUCTION OF MECHANISMS REGULATING THE EXPRESSION OF THE *ESCHERICHIA COLI YFI A* GENE UNDER STRESS CONDITIONS

**T.M. Khlebodarova¹, D.Yu. Oshchepkov¹, N.V. Tikunova³,
I.V. Babkin³, A.D. Gruzdev¹, V.A. Likhoshvai^{1,2}**

¹ Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia,
e-mail: tamara@bionet.nsc.ru;

² Novosibirsk National Research State University, Novosibirsk, Russia;

³ Institute of Chemical Biology and Fundamental Medicine, SB RAS, Novosibirsk, Russia

The regulatory region of the *Escherichia coli yfiA* gene was reconstructed by using the SiteCon web resource and mathematical modeling, and its expression complexity under oxidative stress was assessed. Simulation of the response of *E. coli* cells transformed with pYfi-gfp plasmid to oxidative stress indicated that the maximum agreement with experimental data was achieved in a model implying a complex action, apparently mediated by several transcription factors (TFs). The regulatory region of the *yfiA* gene was searched for potential TF binding sites, and highly reliable recognition was predicted for TFs MarA, IscR, MetJ, PurR, and SoxS, which directly or indirectly participate in the response of the gene to oxidative stress, and for CRP, a global regulator of carbohydrate catabolism. The presence of binding sites for CRP, MarA, and SoxS in the *E. coli yfiA* promoter was confirmed by electrophoretic mobility shift assay with purified recombinant TFs. This fact explains the sensitivity of *yfiA* to mitomycin and radical-forming agents.

Key words: *Escherichia coli*, *yfiA*, GFP, transcription regulation.