

УДК 577.217

СЕМЕЙСТВА Zn-CysHis БЕЛКОВЫХ ФАКТОРОВ: РОЛЬ ЭЛЕКТРОСТАТИЧЕСКОГО ПОТЕНЦИАЛА Zn-ДОМЕНОВ ПРИ ФУНКЦИОНИРОВАНИИ

© 2020 г. Ю. Н. Чиргадзе^{а, *}, В. В. Иванов^{б, с}^аИнститут белка Российской академии наук, Пушкино, Московская обл., 142290 Россия^бОбъединенный институт ядерных исследований, Дубна, Московская обл., 141980 Россия^сМосковский физико-технический институт Национального исследовательского ядерного университета, Москва, 115409 Россия

*e-mail: chir@vega.protres.ru

Поступила в редакцию 04.09.2019 г.

После доработки 14.10.2019 г.

Принята к публикации 21.10.2019 г.

Большое число ДНК- и РНК-связывающих регуляторных факторов содержат Zn-домены. Все они распределяются в шесть Zn-CysHis семейств с разными функциями, которые и рассмотрены в представленном обзоре. Среди них семейство факторов транскрипции PF00096 с комплексом Zn-Cys2His2 считается наиболее важным: оно включает в свыше 340000 представителей от общего числа 365000 известных на сегодня Zn-содержащих факторов. Рассмотрена роль электростатического потенциала Zn-доменов в структуре и функциях факторов разных семейств. Показано, что положительный электростатический потенциал фактора играет существенную роль при образовании контактов с ДНК (РНК) в трех семействах: PF00096, PF12874 и PF09329, – тогда как в трех других семействах: PF03119, PF08996 и PF01258 – Zn-домены не образуют контакты с нуклеиновой кислотой и в этих случаях электростатическое взаимодействие не играет существенной роли при функционировании.

Ключевые слова: ДНК/РНК-связывающий фактор, Zn-CysHis комплекс, электростатический потенциал

DOI: 10.31857/S0026898420020032

ВВЕДЕНИЕ

Электростатическое взаимодействие в функциональных комплексах факторов транскрипции с операторной ДНК изучено достаточно хорошо [1–3]. Первый открытый в 1980 году фактор транскрипции TFIIIA – основной компонент комплекса инициации транскрипции генов 5S рРНК из *Xenopus laevis* РНК-полимеразой III [4, 5]. Этот фактор принадлежит обширному регуляторному

семейству, члены которого содержат ионы Zn в виде комплексов Zn-Cys2His2 (база данных Pfam по семействам, [6]). Это семейство включает 340000 представителей, из около 365000 факторов Zn-Cys2His2, представленных в базе GenBank [7]. К сожалению, число необходимых при рассмотрении функций пространственных структур в банке белковых данных Protein Data Bank [8] во много раз меньше (табл. 1).

Таблица 1. Статистика семейств Zn-содержащих факторов

Pfam-код семейства	Семейство Zn-содержащих регуляторных факторов	Тип Zn-комплекса	Число факторов в семействе	Число структур в банке PDB
PF00096	zf-ДНК-транскрипция	Zn-Cys2His2	340711	412
PF12874	zf-многоклеточные	Zn-Cys2His2	14429	6
PF09329	zf-праймаза	Zn-Cys3His1	71	4
PF03119	zf-ДНК лигаза	Zn-Cys4	3365	2
PF08996	zf-ДНК полимеразы	Zn-Cys4	748	13
PF01258	zf-бактериальные	Zn-Cys4	4575	15

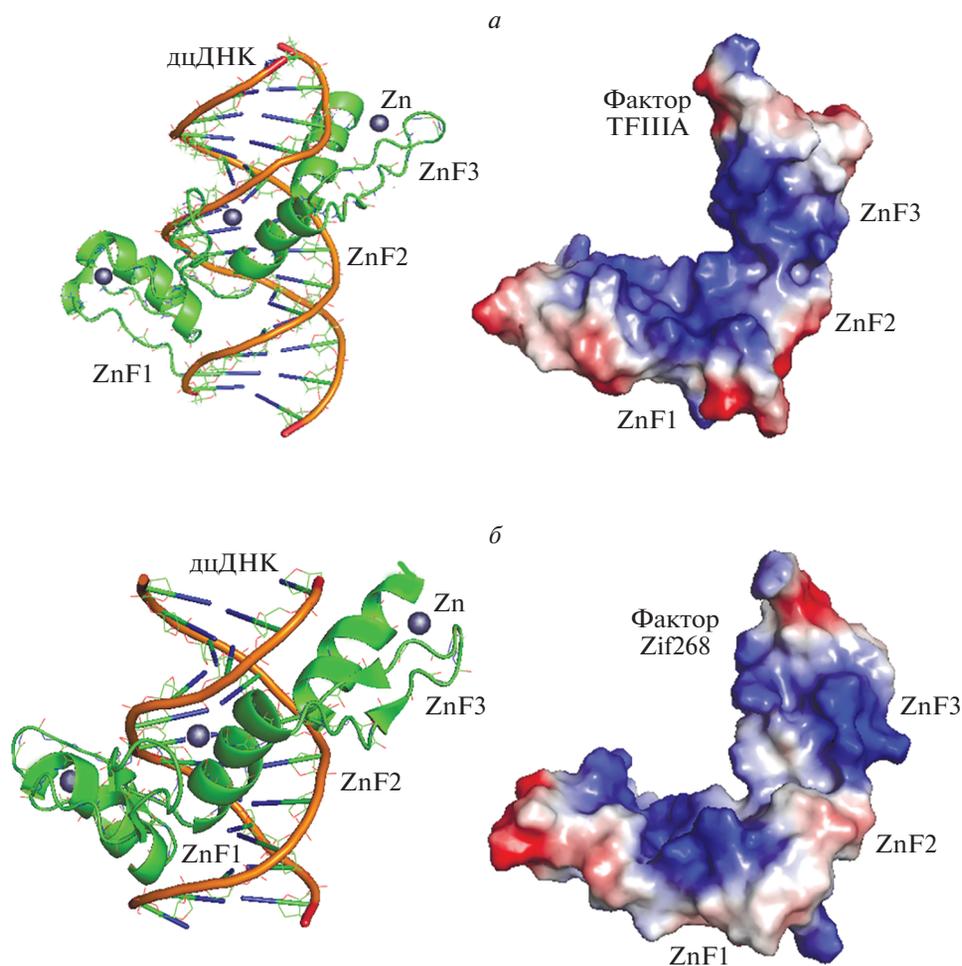


Рис. 1. Структура и электростатический потенциал двух факторов транскрипции из семейства Zn-факторов PF00096. Синий цвет обозначает положительный, красный – отрицательный электростатический потенциал поверхности молекул. *a* – Фрагмент фактора транскрипции TFIIIA лягушки в комплексе с ДНК (pdb 1tf3); *б* – фактор транскрипции Zif268 (EGR-1) мыши (pdb 1aay).

Операторная ДНК представляет собой двухцепочечную структуру, поверхность которой имеет выраженный отрицательный потенциал, обусловленный фосфатными группами нуклеотидов. Молекулы факторов транскрипции семейства Zn-Cys2His2 состоят из нескольких повторяющихся подобных по структуре белковых Zn-доменов (называемых также Zn-пальцами по форме их структуры). Структура и электростатические свойства поверхности фрагмента фактора транскрипции TFIIIA с тремя Zn-доменами (рис. 1*a*) дают нам яркий пример важности электростатических взаимодействий при образовании комплекса с операторной ДНК [9]. Каждый отдельный Zn-домен включает комплекс типа Zn-Cys2His2. Все Zn-содержащие факторы других типов сгруппированы в пять отдельных семейств. Цинковые домены этих факторов имеют похожую, но все-таки отличающуюся простран-

ственную структуру. Они включают в себя комплексы Zn-Cys3His1 и Zn-Cys4. Ниже мы рассмотрим структуры и свойства Zn-доменов разных факторов и оценим роль электростатических потенциалов при выполнении факторами их функций.

СТРУКТУРА, ФУНКЦИЯ И ЭЛЕКТРОСТАТИЧЕСКИЙ ПОТЕНЦИАЛ БЕЛКОВЫХ ФАКТОРОВ РАЗНЫХ СЕМЕЙСТВ Zn-CysHis

Первое семейство, PF00096, включает в себя факторы транскрипции с комплексом типа Zn-Cys2His2. В этом случае отдельные Zn-домены могут узнавать много разных последовательностей ДНК, и этот домен – наиболее распространенный ДНК-связывающий модуль ДНК-связывающих факторов эукариот. Мышиный фактор

транскрипции Zif268, также известный как фактор раннего эмбрионального развития EGR-1, – классическая модель, которая широко используется для изучения свойств факторов семейства Zn-Cys2His2 [10]. Этот фактор необходим при связывании РНК-полимеразы с операторной ДНК, тем самым иницируя процесс транскрипции гена. Комплекс фактора EGR-1 с ДНК показан на рис. 1б.

Корректный способ расчета электростатического потенциала этого фактора приведен в работе [3]. Общий вид потенциала может быть получен с помощью упрощенного алгоритма в программе PyMOL [11], а распределение заряженных групп на поверхности белка с помощью программы Viewer Lite (www.accelrys.com). Основная часть электростатического потенциала этого фактора образована за счет заряженных боковых групп (рис. 1б). Положительная область частично формируется боковой группой остатка His125 (His7 α), который образует контакт с фосфатной группой ДНК. В этом случае увеличение положительного потенциала фактора индуцируется, с одной стороны, за счет контакта остатка His с фосфатной группой ДНК, а с другой, его контактом с ионом Zn цинкового комплекса [3]. Другие детали связывания этого фактора с ДНК приведены в табл. 2.

Ранее мы показали на примере 12 комплексов факторов Zn-Cys2His2 с ДНК, что остаток His7 α достаточно консервативен [2]. Мы можем сделать вывод, что положительный потенциал Zn-домена играет существенную роль при связывании и образовании комплекса факторов этого семейства с операторной ДНК.

Все Zn-домены имеют вторичную структуру полипептидной цепи $\beta\beta\alpha$, и именно α -спираль узнает ДНК. Канонические позиции контактных остатков узнающей α -спирали равны –1, 3, 6, 7 в системе внутреннего отсчета остатков спирали [1, 2]. Эти остатки образуют контакты с тетра nukлеотидом ZXYZ кодирующей цепи ДНК. Анализ Zn-доменов 12 факторов транскрипции семейства Zn-Cys2His2 показал [2], что в позициях –1, 3, и 6 идентичность остатков составляет около 40%, при этом 65% образуют контакты с операторной ДНК. И только остаток His7 имеет 100%-ную идентичность и образует контакты с ДНК в 83% случаев. На основании этого можно сделать важный вывод: Zn-домены факторов транскрипции семейства Zn-Cys2His2 не содержат определенных и консервативных остатков, которые формируют положительный электростатический потенциал. Исключение составляет один остаток His7. Потенциал создают разные остатки для каждого Zn-домена. Сказанное подтверждают данные по анализу контактов на примере четырех Zn-доменов фактора транскрипции TFIIIA [12, 13]. В этом

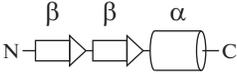
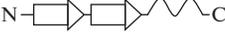
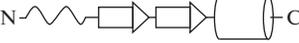
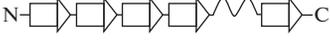
факторе фосфатные группы ДНК образуют контакты с остатками Lys36A (домен ZnF1), His 63A (домен ZnF2), Lys144A и His 155A (домен ZnF5). Однако специфичность и избирательность связывания Zn-доменов определяют контакты с нуклеотидами, их число сопоставимо с таковыми для фосфатов. Следует также отметить, что природный отбор узнающих остатков, по-видимому, сохраняет электростатический потенциал. Действительно, на рис. 1 представлен потенциал двух факторов: TFIIIA и Zif268, – структуры которых содержат в общей сложности семь разных Zn-доменов. Видно, что во всех случаях положительный потенциал практически не меняется. Все сказанное относится и к свойствам следующих РНК-связывающих факторов из семейства Zn-Cys2His2.

Факторы второго семейства, PF12874, содержат РНК-связывающие Zn-домены с комплексами Zn-Cys2His2. Эти факторы участвуют во многих биологических процессах: клеточной локализации, антивирусном ответе, РНК-интерференции (RNAi), вирусном ингибировании этой интерференции и других. В качестве примера мы рассмотрим белковый фактор JAZ человека, также известный как белок ZNF346 [14]. Он состоит из шести Zn-доменов. Структура одного из них, Zn-домена 3, имеет хорошо выраженную положительно заряженную область (рис. 2а).

Этот домен связан с двухцепочечной РНК за счет контактов трех фосфатных групп и остатков лизинов 207, 208 и 211. Интересно отметить, что домен 3 состоит из двух доменов: 3А и 3В, – каждый из которых включает в себя один и тот же цинковый комплекс. Первый домен 3А имеет структуру Zn-пальца, тогда как второй домен 3В состоит только из одной α -спирали.

Третье семейство, PF09329, объединяет Zn-праймазы и содержит комплексы Zn-Cys3His1. Цинковый домен членов этого семейства – часть мультисубъединичного комплекса Mcm10 репликации ДНК, с помощью которого происходит сборка и последующее наращивание репликационной вилки. Это существенный белковый комплекс репликации ДНК эукариот, который обеспечивает инициаторную активность праймазы. В случае *Xenopus laevis* он состоит из внутреннего домена ID и одноцепочечной ДНК [15]. Структура и распределение заряженных групп на поверхности этого комплекса приведены на рис. 2б. Он состоит из ОВ-домена (Oligonucleotide/oligosaccharide Binding fold) и Zn-домена в виде структуры Zn-пальца с комплексом Zn-Cys3His1. Цинковый палец связывает одноцепочечную ДНК за счет контактов заряженных групп остатков Lys385 и Lys386. Мы видим, что положительный потенциал Zn-пальца необходим для связывания с ДНК.

Таблица 2. Регуляторные ДНК-, РНК-связывающие факторы семейств Zn-CysHis

Рfam-код семейства	Семейство факторов и тип Zn-комплексов	Вторичная структура Zn-доменов	Функциональная роль Zn-доменов разных Zn-семейств
PF00096	zf-транскрипция Zn-Cys2His2		ДНК-связывающий фактор EGR-1, pdb 1aay (рис. 1б). Специфическое связывание между Zn-доменом и нуклеотидами ДНК. Неспецифическое связывание между остатками His и Arg и фосфатами ДНК.
PF12874	zf-РНК-связывание Zn-Cys2His2		РНК-связывающий фактор JAZ человека, pdb 2mkd (рис. 2а). Специфическое связывание в четырех Zn-доменах из шести. Неспецифическое связывание приведено для Zn-домена 3В. Три положительных остатка лизина связаны с фосфатами РНК.
PF09329	zf-праймаза Zn-Cys3His1		ДНК репликационный комплекс Mcm10, pdb 3h15 (рис. 2б). Связывание одноцепочечной ДНК с белковым фактором неспецифично. Оно обеспечено контактами фосфатных групп с двумя остатками лизина Zn-домена.
PF03119	zf-ДНК-лигаза Zn-Cys4		Репарационный комплекс ДНК-лигазы, pdb 2owo (рис. 3а). Здесь Zn-домен играет роль связки между доменами ОВ и НhН и фиксирует их расположение.
PF08996	zf-ДНК-полимераза Zn-Cys4		Фрагмент p180c из ДНК DNA полимеразы α человека, pdb 4y97 (рис. 3б). Ее каталитическая субъединица участвует в репликации ДНК. Zn-домен фиксирует положение субъединиц в полимеразе.
PF01258	zf-транскрипция dksA/traR Zn-Cys4		Фактор DksA, pdb 1tjl (рис. 3в). ppGpp-зависимая регуляция транскрипции бактерий обеспечена специфическими контактами фактора с полимеразой.

В трех рассмотренных выше семействах функция Zn-доменов зависит от их контакта с ДНК или РНК, что обеспечивает положительный электростатический потенциал Zn-доменов. В трех последующих семействах Zn-домены не образуют контактов с нуклеиновыми кислотами. Во всех этих случаях Zn-домены имеют отношение к структуре

Zn-пальцев, однако их пространственная структура в ряде случаев вырождается (табл. 2).

Ниже мы рассмотрим белковые факторы, представляющие эти семейства, более детально. Четвертое семейство PF03119 объединяет Zn-ДНК-лигазы и содержит комплексы Zn-Cys4. Цинковый домен в виде структуры пальца имеет отношение к

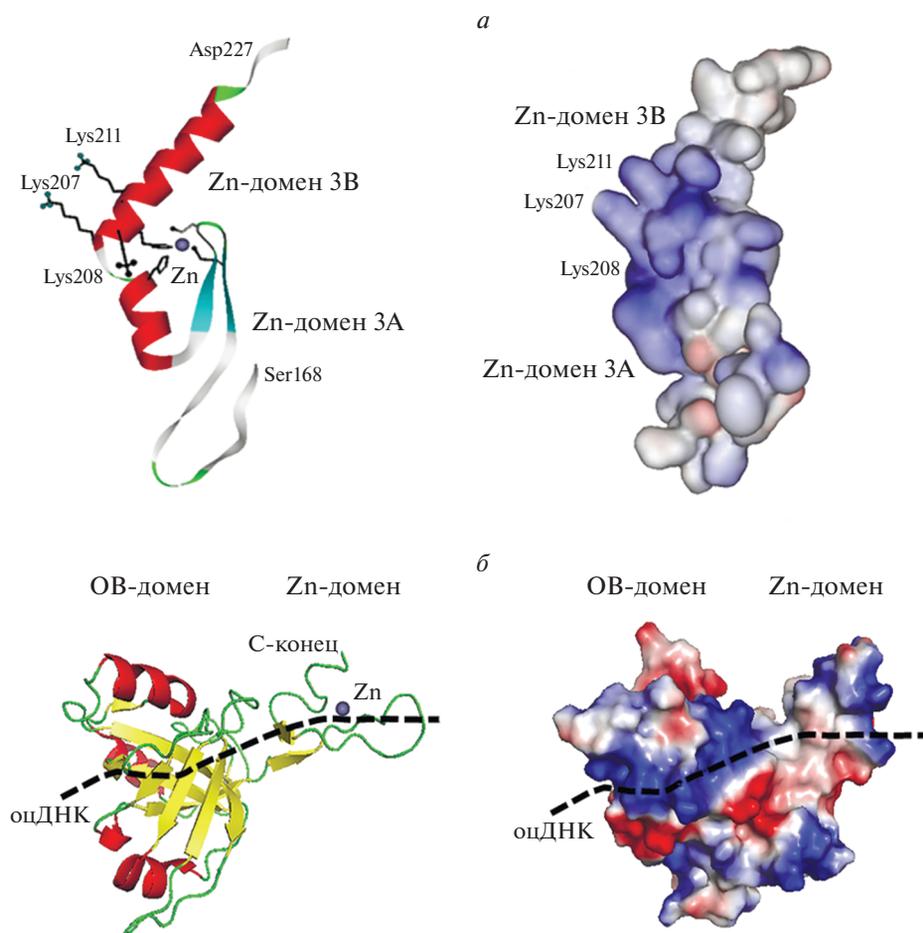


Рис. 2. Структура и распределение заряженных групп на поверхности Zn-факторов. *а* – дц-РНК-связывающий фактор, белок JAZ человека, семейство PF12874 (pdb 2mkd). *б* – Внутренний домен репликационного комплекса Msm10 лягушки, семейство PF09329 (pdb 3h15).

выполнению функции репарации ДНК как часть мультидоменного комплекса NAD^+ -зависимой лигазы. Так, комплекс из *Escherichia coli* состоит из одной длинной полипептидной цепи из 586 аминокислотных остатков [16]. Он включает в себя N-концевой инициаторный домен Ia, нуклеотидилтрансферазу (НТазу), олигонуклеотидсвязывающий домен ОВ, домен со структурой Zn-пальца и спиральный HhH-домен. В комплексе с лигазой двухцепочечная ДНК окружена доменами лигазы, как показано на рис. 3а.

Цинковый домен не имеет контактов с ДНК, и его электростатический потенциал является слабоположительным. В этом комплексе Zn-домен играет роль связующего звена между доменами ОВ и HhH.

Пятое семейство PF08996 представлено факторами, структурное тело которых занимает в функциональном комплексе пространство между ДНК и полимеразой α . Эти факторы имеют отношение к процессу репарации ДНК. В качестве примера рассмотрим комплекс полимеразы α человека (бо-

лее точно С-концевого домена p180с) с В-субъединицей p70 [17]. Структура цинкового домена из p180с показана на рис. 3б. Молекула имеет два модуля цинковых пальцев, связанных друг с другом α -спиральями. Показано, что второй домен Zn2 образует контакты с В-субъединицей полимеразы α . И это играет ключевую роль при функционировании ДНК-полимераз. Этот фактор не связан с ДНК, а распределение заряженных групп на его поверхности не имеет протяженных положительных областей.

Шестое семейство PF01258 представлено здесь фактором DksA из *E. coli* (DnaK suppressor A), критическим компонентом ppGpp -зависимой регуляции гуанозинтетрафосфата. Фактор DksA играет роль кофактора, который стабилизирует ppGpp -РНК-полимеразный комплекс [18]. Структура цинкового домена существенно отличается от всех описанных выше. Фактически, фактор DksA – полностью α -спиральным белок, который можно разделить на три разных фрагмента (рис. 3в). Пер-

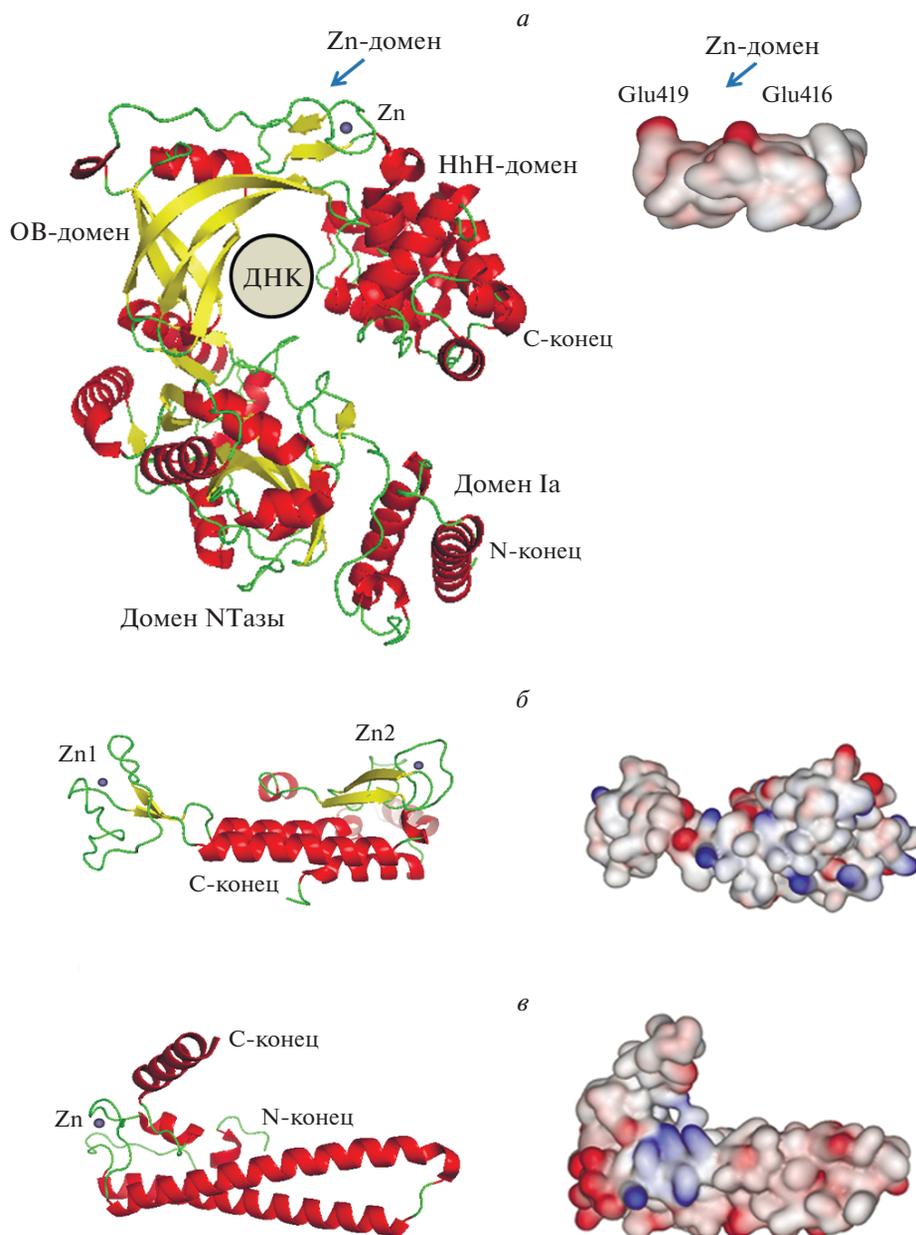


Рис. 3. Структура и распределение заряженных групп на поверхности Zn-факторов. *a* – NAD^+ -зависимая ДНК-лигаза из *E. coli*, семейство PF03119 (pdb 2owo); *б* – ДНК-полимераза человека, семейство PF08996 (pdb 4y97); *в* – бактериальный фактор DksA из *E. coli*, семейство PF01258 (pdb 1tjl).

вый – это структура из скрученных спиралей, он состоит из двух длинных спиралей. Второй фрагмент – глобула, образованная N- и C-концевыми областями фактора. Третий фрагмент представляет собой каноническую структуру цинкового домена с комплексом Zn-Cys4. Распределение заряженных групп на поверхности Zn-домена фактора DksA представлено на рисунке. Поверхность имеет две заряженные области. Отрицательно заряженная область локализована во втором глобулярном фрагменте, отрицательный заряд обусловлен остатками Glu29, Glu30, Asp108, и Glu109. По-

ложительно заряженная область расположена в первой структурной части фактора, и она сформирована кластером из остатков Lys97, Lys101 и Lys104. Как следует из предварительной модели структуры комплекса фактор-РНК-полимераза, есть два функционально важных отрицательно заряженных остатка: Asp71 и Asp74, – которые образуют координационную сферу иона Mg^{2+} , связанного с фосфатами ppGpp [18]. Эти важные остатки локализованы на той части молекулы, которая не видна на рисунке. Таким образом, электростатические свойства Zn-домена не играют заметной

роли в специфическом связывании с ДНК в полимерном комплексе.

ЗАКЛЮЧЕНИЕ

Мы обнаружили, что регуляторные Zn-факторы трех семейств: PF00096, PF12874 и PF09329 — имеют хорошо выраженный положительный потенциал поверхности Zn-доменов, который существен для выполнения их функций. Эти функции имеют непосредственное отношение к связыванию факторов с отрицательно заряженными молекулами ДНК или РНК. В факторах, трех других семейств: PF03119, PF08996 и PF01258 — Zn-домены не связываются с ДНК или РНК и, соответственно, не имеют положительных областей на своих поверхностях. Свою функциональную роль они выполняют в качестве необходимых структурных модулей мультидоменных комплексов. Во всех этих случаях выраженный положительный потенциал Zn-доменов отсутствует.

Авторы благодарят научного сотрудника Е.А. Бошкову (Институт белка РАН) за помощь при работе с базами данных и полезное обсуждение работы.

Работа поддержана грантами Российского фонда фундаментальных исследований (№№ 14-03-01091а и 17-07-01331а).

Данная статья не содержит каких-либо исследований с участием людей или животных в качестве объектов исследований.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Klug A. (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Ann. Rev. Biochem.* **79**, 213–231.
2. Polozov R.V., Sivozhelezov V.S., Chirgadze Y.N., Ivanov V.V. (2015) Recognition rules for binding of Zn-Cys2His2 transcription factors to operator DNA. *J. Biomol. Struct. Dyn.* **33**, 253–266.
3. Chirgadze Y.N., Boshkova E.A., Polozov R.V., Sivozhelezov V.S., Dzyabchenko A.V., Kuzminsky M.B., Stepanenko V.A., Ivanov V.V. (2018) The electrostatic role of the Zn-Cys2His2 complex in binding of operator DNA with transcription factors: mouse EGR-1 from the Cys2His2 family. *J. Biomol. Struct. Dyn.* **36**, 3902–3915.
4. Engelke D.R., Ng S.Y., Shastry B.S., Roeder R.G. (1980) Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell.* **19**, 717–728.
5. Miller J., McLachlan A.D., Klug A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus oocytes*. *EMBO J.* **4**, 1609–1614.
6. Punta M., Coghill P.C., Eberhardt R.Y., Mistry J., Tate J., Boursnell C., Pang N., Forslund K., Ceric G., Clements J., Heger A., Holm L., Sonnhammer E.L., Eddy S.R., Bateman A., Finn R.D. (2012) The Pfam protein families' database. *Nucleic Acids Res.* **40**, D290–301.
7. Benson D.A., Cavanaugh M., Clark K., Karsch-Mizrachi I., Lipman D.J., Ostell J., Sayers E.W. (2013) GenBank. *Nucleic Acids Res.* **41**, 36–42.
8. Berman H.M., Battistuz T., Bhat T.N., Bluhm W.F., Bourne P.E., Burkhardt K., Feng Z., Gilliland G.L., Iype L., Jain S., Fagan P., Marvin J., Padilla D., Ravichandran V., Schneider B., Thanki N., Weissig H., Westbrook J.D., Zardecki C. (2002) The protein data bank. *Acta Cryst.* **58**(Pt 6 No1), 899–907.
9. Wuttke D.S., Foster M.P., Case D.A., Gottesfeld J.M., Wright P.E. (1997) Solution structure of the first three zinc fingers of TFIIIA bound to the cognate DNA sequence: determinants of affinity and sequence specificity. *J. Mol. Biol.* **273**, 183–206.
10. Elrod-Erickson M., Rould M.A., Nekludova L., Pabo C.O. (1996) Zif268 protein-DNA complexes refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure.* **4**, 1171–1180.
11. DeLano W.L., Ultsch M.H., de Vos A.M., Wells J.A. (2000) Convergent solutions to binding at a protein-protein interface. *Science.* **287**, 1279–1283.
12. Nolte R.T., Conlin R.M., Harrison S.C., Brown R.S. (1998) Differing roles for zinc fingers in DNA recognition: structure of a six-finger transcription factor IIIA complex. *Proc. Natl. Acad. Sci. USA.* **95**, 2938–2943.
13. Chirgadze Y.N., Boshkova E.A., Yakovlev A.V., Ivanov V.V. (2019) Side projections of double-helical DNA: example of binding patterns of DNA in the complex with factor TFIIIA. *J. Biomol. Struct. Dyn.* **37**, 4433–4436.
14. Burge R.G., Martinez-Yamout M.A., Dyson H.J., Wright P.E. (2014) Structural characterization of interactions between the double-stranded RNA-binding zinc finger protein JAZ and nucleic acids. *Biochemistry.* **53**, 1495–1510.
15. Warren E.M., Huang H., Fanning E., Chazin W.J., Eichman B.F. (2009) Physical interactions between Mcm10, DNA, and DNA polymerase α . *J. Biol. Chem.* **284**, 24662–24672.
16. Nandakumar J., Nair P.A., Shuman S. (2007) Last stop on the road to repair: structure of *E. coli* DNA ligase bound to nicked DNA-adenylate. *Mol. Cell.* **26**, 257–271.
17. Suwa Y., Gu J., Baranovskiy A.G., Babayeva N.D., Pavlov Y.I., Tahirov T.H. (2015) Crystal structure of the human Pol α B subunit in complex with the C-terminal domain of the catalytic subunit domain of the catalytic subunit. *J. Biol. Chem.* **290**, 14328–14337.
18. Perederina A., Svetlov V., Vassilyeva M.N., Tahirov T.H., Yokoyama S., Artsimovitch I., Vassilyev D.G. (2004) Regulation through the secondary channel—structural framework for ppGpp-DksA synergism during transcription. *Cell.* **118**, 297–309.

Zn-CysHis PROTEIN FACTOR FAMILIES: ROLE OF ELECTROSTATIC INTERACTION OF Zn-DOMAINS IN FACTOR FUNCTIONS

Y. N. Chirgadze^{1, *} and V. V. Ivanov^{2, 3}

¹*Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, 142290 Russia*

²*Joint Institute for Nuclear Research, Dubna, Moscow Region, 141980, Russia*

³*Moscow Engineering Physics Institute, National Research Nuclear University, Moscow, 115409 Russia*

*e-mail: chir@vega.protres.ru

The large amount of the DNA and RNA binding protein regulatory factors contain Zn-domains. Total six Zn-CysHis families with different functions were considered. The transcription factors of family PF00096, complex type Zn-Cys2His2, is the most important due to their presence in over 340000 members of 365000 total members listed at present in the Zn containing factor families. We have considered the role of electrostatic potential of Zn-domains in the structure and function in the different factor families. The analysis revealed that the positive electrostatic potential of Zn-domains plays a significant role in forming contact with DNA/RNA in three protein factor families PF00096, PF12874, and PF09329. While in other three families PF03119, PF08996, and PF01258 Zn-domains do not contact with nucleic acids, and electrostatic interactions do not play a distinctive role, consequently.

Keywords: Zn-CysHis complex, Zn-finger, DNA binding factor, electrostatic potential