

ИССЛЕДОВАНИЕ СТРУКТУРЫ N-КОНЦЕВОГО ДИМЕРИЗУЮЩЕГО ДОМЕНА БЕЛКА CLAMP МЕТОДОМ МАЛОУГЛОВОГО РАССЕЯНИЯ РЕНТГЕНОВСКИХ ЛУЧЕЙ (SAXS)

Тихонова Е.А.¹, Бончук А.Н.¹, Качалова Г.С.², Георгиев П.Г.¹, Максименко О.Г.¹

¹Институт биологии гена РАН

ул. Вавилова, д. 34/5, г. Москва, 119334, РФ; e-mail: errinaceus@rambler.ru

²ФИЦ Биотехнологии РАН

Ленинский проспект, д. 33, г. Москва, 119071, РФ

Поступила в редакцию: 20.06.2018

Аннотация. CLAMP – транскрипционный фактор *Drosophila*, участвующий в привлечении комплекса дозовой компенсации на специфические сайты ДНК. При помощи малоуглового рассеяния рентгеновских лучей (SAXS) получены данные о пространственной структуре N-концевого домена данного белка, содержащего домен типа «цинковый палец» и прилегающую аминокислотную последовательность с неизвестной структурой. Для этого домена ранее была показана способность к гомо-димеризации и взаимодействию с комплексом дозовой компенсации. Домен типа «цинковый палец» важен для функциональной активности N-концевого домена. Данные SAXS подтверждают, что N-концевой домен CLAMP является димером, свернут и имеет компактную симметричную структуру.

Ключевые слова: мультимеризация, малоугловое рассеяние рентгеновских лучей, хроматин.

Дозовая компенсация у дрозофилы осуществляется посредством увеличения уровня экспрессии генов X-хромосомы самца (X/Y) в два раза по сравнению с самками (X/X). За увеличение экспрессии генов X-хромосомы самцов отвечает мультисубъединичный комплекс дозовой компенсации (КДК), состоящий из пяти белков, MSL1, MSL2, MSL3, MOF и MLE, и двух некодирующих РНК - гоX1 и гоX2 [1]. Белки MSL1 и MSL2 создают структурную основу КДК [2, 3]. Полногеномный анализ сайтов связывания белков КДК показал, что комплекс рекрутируется в кодирующие области активно транскрибируемых генов [4, 5]. Одним из неразрешенных вопросов является, каким образом КДК специфично связывается с X-хромосомой самцов. При инактивации белков MSL3 или MLE наблюдается связывание неполного комплекса MSL1-MSL2 с определенным и воспроизводимым набором, состоящим из примерно 200 сайтов на X-хромосоме, которые были названы сайтами первичного связывания (СПС) [6]. Наиболее вероятным механизмом привлечения КДК является присутствие в составе СПС-сайтов для ТФ, которые участвуют в специфичном связывании с КДК. В скрининге транскрипционных факторов, определяющих способность СПС-элемента стимулировать транскрипцию в культуре клеток S2, был найден белок CLAMP, который содержит семь доменов «цинковые пальцы» C2H2-типа [7]. *In vitro* было показано, что часть белка, содержащая с 4 по 7 C2H2-домены, специфично связывается с GA-повторами [8]. Полногеномный ChIP-seq анализ показал, что белок CLAMP связывается с районами СПС, но одновременно имеет более 5 000 сайтов по всему геному, большая часть из которых распределена на аутосомах [8]. В настоящее время предполагается, что CLAMP совместно с белком MSL2 участвует в привлечении КДК на СПС.

На N-конце белка Clamp находится одиночный домен типа «цинковый палец» C2H2-типа, окруженный последовательностями, не имеющими выраженной гомологии с известными белковыми доменами. Цинковые пальцы C2H2-типа обычно участвуют в связывании ДНК, однако имеются данные и о том, что они могут принимать участие в белок-белковых взаимодействиях, в том числе и в димеризации. Однако структурные данные для таких взаимодействий отсутствуют. Предварительные исследования в дрожжевой двугибридной системе показали, что N-концевой домен Clamp обладает способностью к мультимеризации. Анализ делеционных мутантов выявил, что C2H2-домен важен для правильного функционирования данного домена, вместе с 30-аминокислотной последовательностью, примыкающей к этому домену. Эта же последовательность участвует во взаимодействии с коровым компонентом MSL-комплекса – белком MSL2. Учитывая очень широкую распространенность доменов C2H2 у высших многоклеточных, и слабую изученность их способности к белок-белковым взаимодействиям, структурные данные о таком взаимодействии представляет большой интерес. Формирование димеров было подтверждено при помощи гель-фильтрации и химической сшивки с помощью глutarальдегида.

Дальнейшее исследование N-концевого домена белка Clamp проводилось при помощи метода малоуглового рассеяния рентгеновского излучения (SAXS). Эксперимент проводился на источнике рентгеновского излучения BM29 BioSAXS в ESRF (Гренобль, Франция). Измерения проводились при длине волны 0.099 нм, концентрация белка варьировалась от 2.3 до 11.8 мг/мл. Радиус гирации (Rg) белковой молекулы в растворе был рассчитан при помощи аппроксимации Гинье при малых углах ($s < 1.3/Rg$, $s = 4\pi\sin\theta/\lambda$, где 2θ – угол рассеяния), интенсивность $I(s)$ принималась равной $I_0 \exp(-sRg)^2/3$. Для оценки максимального размера частиц D_{max} , функция парных расстояний $P(r)$ была рассчитана при помощи алгоритма GNOM из программного пакета ATSAS (рис. 1А, Б). Молекулярный вес молекул был рассчитан исходя из исключенного объема гидратированных белковых молекул по закону Поро для гомогенных частиц.

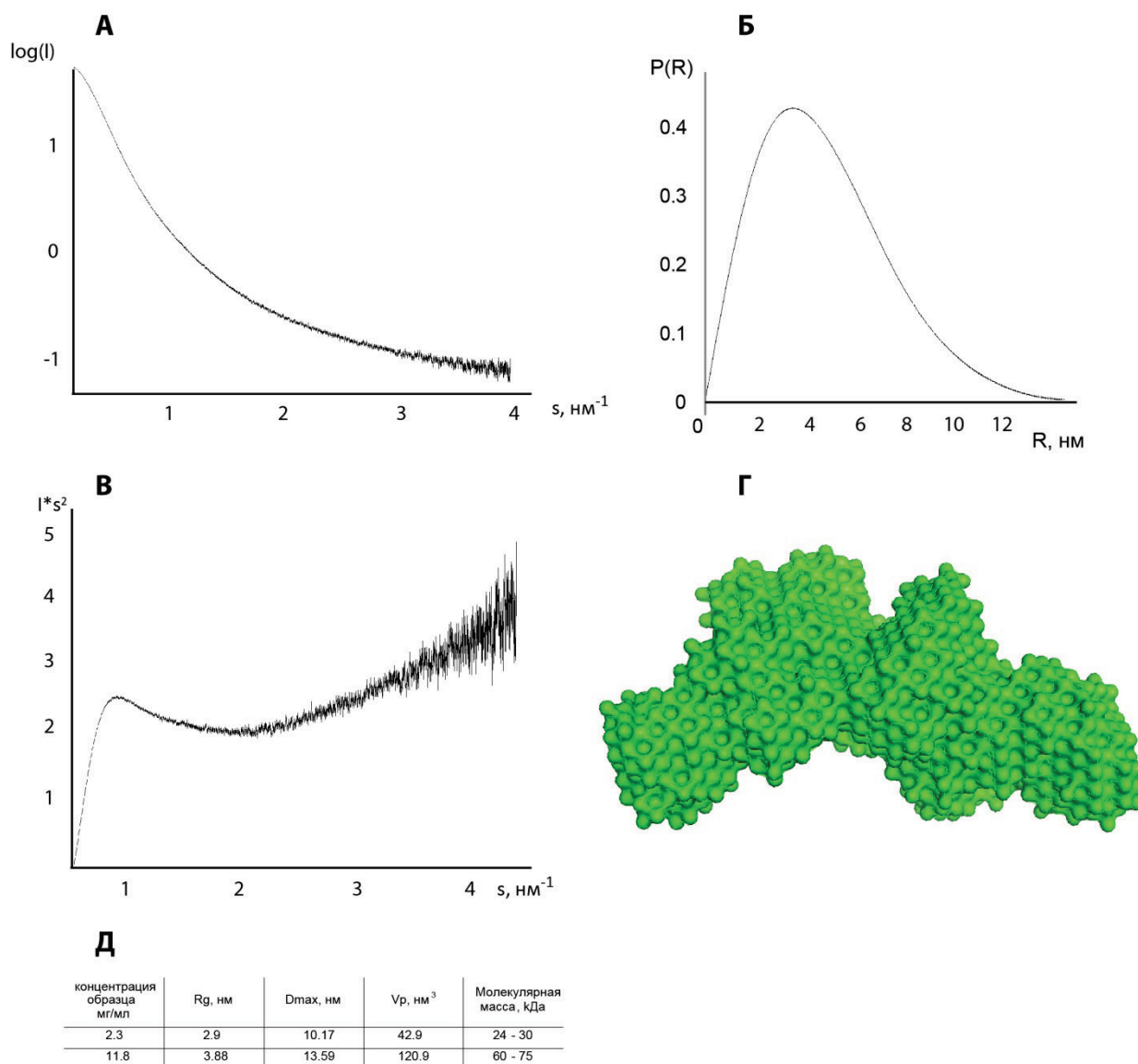


Рисунок 1. Результаты эксперимента по малоугловому рассеянию рентгеновских лучей N-концевого домена белка Clamr в растворе (1-153 ак, мол. масса мономера 17 кДа). А) Кривая распределения интенсивности рассеяния (2,3 мг/мл); Б) График функции парных расстояний (2,3 мг/мл); В) График Кратки (2,3 мг/мл); Г) Одна из возможных *ab initio* моделей низкого разрешения при концентрации 2,3 мг/мл; Д) Рассчитанные параметры олигомеров N-концевого домена белка Clamr в растворе (мол. масса мономера 17 кДа) при разных концентрациях образца

В результате эксперимента в зависимости от концентрации образца были получены следующие значения: $R_g = 2,9-3,9$ нм, $D_{max} = 10,2-13,7$ нм, что соответствует молекулярной массе 25-35 кДа при концентрации 2,3 мг/мл и 60-75 при 11,8 мг/мл (рис. 1Д) и подтверждает формирование димеров (молекулярная масса мономера составляет 17 кДа) и указывает на формирование олигомеров более высокого порядка при высокой концентрации, однако учитывая высокую концентрацию, при которой наблюдается их сборка, можно заключить, что данные тетрамеры имеют достаточно невысокую константу диссоциации и не являются физиологическими их появление по всей видимости связано с высокой концентрацией и заморозкой-разморозкой образца.

Дальнейший анализ проводился для данных, полученных при концентрации 2,3 мг/мл. Для анализа структурированности исследуемого домена проводился анализ графика Кратки ($I \times s^2$ относительно s), форма которого сильно зависит от степени свернутости исследуемого белка. Для структурированных полипептидов характерна куполообразная форма кривой при малых значениях s , в то же время для развернутых полипептидов кривая имеет логарифмическую форму. Данные представлены на рисунке 1В. Как видно, для N-концевого домена Clamr характерно наличие небольшого куполообразного изгиба, что свидетельствует о том, что часть аминокислотной цепи не структурирована. *Ab initio* модели низкого разрешения для белка Clamr (1-153 ак) были получены при помощи алгоритма DAMMIN из программного пакета ATSAS (рис. 1Г). Большинство моделей являются симметричными, что подтверждает формирование димеров. Каждый из мономеров состоит в

свою очередь из двух частей – одна из них участвует в формировании димера, а вторая, по всей видимости, не имеет структуры. Сопоставление данных моделей с известными кристаллическими структурами C2H2-доменов также позволяет предположить, что C2H2-домен находится вблизи димерного интерфейса, поскольку его размеры слишком велики и не позволяют предположить присутствия цинкового пальца в части домена, не участвующей в формировании димера. Полученные данные хорошо соотносятся с результатами анализа делеционных мутантов и подтверждают участие C2H2-домена в белок-белковых взаимодействиях.

Работа выполнена при поддержке Российского научного фонда (проект 17-74-20155).

Список литературы / References:

1. Kuroda M.I., Hilfiker A., Lucchesi J.C. Dosage Compensation in Drosophila—a Model for the Coordinate Regulation of Transcription. *Genetics*, 2016, vol. 204, pp. 435-450.
2. Hallaçli E., Lipp M., Georgiev P., Spielman C., Cusack S., Akhtar A., Kadlec J. Msl1-mediated dimerization of the dosage compensation complex is essential for male X-chromosome regulation in Drosophila. *Mol Cell*, 2012, vol. 48, pp.587-600.
3. Kadlec J., Hallaçli E., Lipp M., Holz H., Sanchez-Weatherby J., Cusack S., Akhtar A. Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. *Nat. Struct. Mol. Biol.*, 2011, vol. 18, pp. 142-149.
4. Alekseyenko A.A., Ho J.W., Peng S., Gelbart M., Tolstorukov M.Y., Plachetka A., Kharchenko P.V., Jung Y.L., Gorchakov A.A., Larschan E. Sequence-specific targeting of dosage compensation in Drosophila favors an active chromatin context. *PLoS Genet.*, 2012, vol. 8, e1002646.
5. Straub T., Zabel A., Gilfillan G.D., Feller C., Becker P.B. Different chromatin interfaces of the Drosophila dosage compensation complex revealed by high-shear ChIP-seq. *Genome Res.*, 2013, vol. 23, pp. 473-485.
6. Straub T., Grimaud C., Gilfillan G.D., Mitterweger A., Becker P.B. The chromosomal high-affinity binding sites for the Drosophila dosage compensation complex. *PLoS Genet*, 2008, e1000302.
7. Larschan E., Soruco M.M., Lee O.K., Peng S., Bishop E., Chery J., Goebel K., Feng J., Park P.J., Kuroda M.I. Identification of chromatin-associated regulators of MSL complex targeting in Drosophila dosage compensation. *PLoS Genet.*, 2012, vol. 8, e1002830.
8. Marcela M.I., Chery J., Bishop E.P., Siggers T., Tolstorukov M.Y., Leydon A.R., Sugden A.U., Goebel K., Feng J., Xia P., Vedenko A., Bulyk M.I., Park P.J. and Larschan E. The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. *Genes & development*, 2013, vol. 27, pp. 1551-1556.

STRUCTURAL STUDIES OF N-TERMINAL DIMERIZATION DOMAIN OF CLAMP PROTEIN USING SMALL-ANGLE X-RAY SCATTERING (SAXS)

Tikhonova E.A.¹, Bonchuk A.N.¹, Kachalova G.S.², Georgiev P.G.¹, Maksimenko O.G.¹

¹ Institute of gene biology RAS

Vavilova str., 34/5, Moscow, 119334, Russia; e-mail: errinaceus@rambler.ru

² Research Center of Biotechnology RAS

Leninsky prospekt, 33, Moscow, 119071, Russia

Abstract. CLAMP is Drosophila transcription factor involved in the recruitment of Dosage Compensation Complex. Spatial structure of N-terminal domain of CLAMP protein from Drosophila was studied using small-angle X-ray scattering (SAXS). This domain contains zinc-finger and adjacent amino-acid sequence with unknown structure. Earlier dimerization activity and binding of dosage compensation complex were shown for this domain. SAXS data confirm dimer formation and provide insight into the structure of the domain, which is folded, and have compact symmetrical structure.

Key words: *multimerization, small-angle X-ray scattering, chromatin.*