

3. Kraft A., Lutz C., Lingenhel A., Gröbner P., Piendl W. Control of ribosomal protein L1 synthesis in mesophilic and thermophilic archaea. *Genetics*, 1999, vol. 152 (4), pp. 1363-1372.
4. Fu Y., Deiorio-Haggar K., Anthony J., Meyer M.M. Most RNAs regulating ribosomal protein biosynthesis in *Escherichia coli* are narrowly distributed to Gammaproteobacteria. *Nucleic Acids Res*, 2013, vol. 41, no. 6, p. 3491-3503.
5. Tishchenko S., Gabdulkhakov A., Nevskaya N., Sarskikh A., Kostareva O., Nikonova E., Sycheva A., Moshkovskii S., Garber M., Nikonov S. High-resolution crystal structure of the isolated ribosomal L1 stalk. *Acta Crystallogr D Biol Crystallogr.*, 2012, vol. 68, pp. 1051-1057.
6. Tishchenko S., Kostareva O., Gabdulkhakov A., Mikhaylina A., Nikonova E., Nevskaya N., Sarskikh A., Piendl W., Garber M., Nikonov S. Protein-RNA affinity of ribosomal protein L1 mutants does not correlate with the number of intermolecular interactions. *Acta Crystallogr D Biol Crystallogr.*, 2015, vol. 71, pp. 376-386.
7. Korepanov A.P., Kostareva O.S., Bazhenova M.V., Bubunenko M.G., Garber M.B., Tishchenko S.V. Studying the properties of domain I of the ribosomal protein L1: incorporation into ribosome and regulation of the L1 operon expression. *Protein J.*, 2015, vol. 34 (2), pp.103-110.
8. Tishchenko S., Kljashtorny V., Kostareva O., Nevskaya N., Nikulin A., Gulak P., Piendl W., Garber M., Nikonov S. Domain II of *Thermus thermophilus* ribosomal protein L1 hinders recognition of its mRNA. *J. Mol. Biol.*, 2008, vol. 383, pp. 301-305.
9. Kostareva O., Tishchenko S., Nikonova E., Kljashtorny V., Nevskaya N., Nikulin A., Sycheva A., Moshkovskii S., Piendl W., Garber M., Nikonov S. Disruption of shape complementarity in the ribosomal protein L1-RNA contact region does not hinder specific recognition of the RNA target site. *J. Mol. Recognit.*, 2011, vol. 24, pp. 524-532.

N-КОНЦЕВОЙ МУЛЬТИМЕРИЗУЮЩИЙ ДОМЕН ИНСУЛЯТОРНОГО БЕЛКА CTCF DROSOPHILA MELANOGASTER ИМЕЕТ КОМПАКТНУЮ ПРОСТРАНСТВЕННУЮ ОРГАНИЗАЦИЮ ПРИ ОТСУТСТВИИ ВТОРИЧНОЙ СТРУКТУРЫ

Бончук А.Н.¹, Качалова Г.С.², Бойко К.М.², Максименко О.Г.¹, Георгиев П.Г.¹

¹ Институт биологии гена РАН

ул. Вавилова, 34/5, г. Москва, 119334, РФ

² Институт биохимии им. А.Н. Баха РАН

Ленинский пр., 33, г. Москва, 119071, РФ

e-mail: errinaceus@rambler.ru

Аннотация. При помощи методов кругового дихроизма и малоуглового рассеяния рентгеновских лучей (SAXS) получены данные о пространственной структуре мультимеризующего N-концевого домена инсуляторного белка CTCF *Drosophila melanogaster*. По аминокислотной последовательности данный домен не имеет гомологии ни с одним известным белковым доменом. По данным кругового дихроизма N-концевой домен белка CTCF *Drosophila* практически не имеет вторичной структуры. По данным SAXS этот домен имеет стабильную компактную структуру в растворе и является тетрамером. Получены модели низкого разрешения структуры тетрамера в растворе. Формирование тетрамеров также подтверждается при помощи методов динамического рассеяния света (DLS) и гель-фильтрации.

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

N-TERMINAL MULTIMERIZATION DAMAIN OF DROSOPHILA MELANOGASTER CTCF PROTEIN HAS COMPACT SPATIAL ORGANIZATION BUT LACKS SECONDARY STRUCTURE

Bonchuk A.N.¹, Kachalova G.S.², Boyko K.M.², Maksimenko O.G.¹, Georgiev P.G.¹

¹ Institute of gene biology RAS

Vavilova st., 34/5, Moscow, 119334, Russia

² A.N. Bach Institute of biochemistry of RAS,

Leninsky pr., 33, Moscow, 119071, Russia

e-mail: errinaceus@rambler.ru

Abstract. Spatial structure of N-terminal multimerization domain from *Drosophila melanogaster* CTCF insulator protein was studied using circular dichroism and small-angle X-ray scattering (SAXS). This domain has no significant homology in amino acid sequence with any other known protein domain. Circular dichroism spectra revealed that N-terminal domain of *Drosophila* CTCF protein almost completely lacks secondary structure. SAXS data indicated that domain possess a stable compact conformation in solution and exists as a tetramer. Low-resolution models of spatial structure of the tetramer were built. Tetramer formation is also supported by dynamic light scattering (DLS) and size-exclusion chromatography.

Keywords: multimerization, small-angle X-ray scattering, circular dichroism, chromatin.

Инсуляторы – геномные элементы, участвующие в разобщении независимых хроматиновых доменов, обладающие способностью блокировать энхансер-промоторные взаимодействия и препятствовать распространению гетерохроматина. При этом инсуляторы способны взаимодействовать между собой на больших

расстояниях в геноме (более 1000 п.н.) и таким образом принимать участие в организации пространственной структуры хроматина [1]. CTCF – консервативный среди высших многоклеточных белок, связывающийся с ДНК инсультаторов и единственный известный ДНК-связывающий инсультаторный белок у позвоночных [2, 3]. Сайты связывания белка CTCF способны обеспечивать взаимодействие между удаленными участками ДНК [1, 4]. Очевидно, в процессе установления и поддержания таких контактов в геноме принимают участие белковые домены, обеспечивающие специфичное белок-белковое взаимодействие и формирование мультимеров. Белок CTCF имеет в своем составе 11 ДНК-связывающих доменов типа «цинковые пальцы», а также N- и C- концевые домены с неизвестными функциями. Была показана способность белка CTCF мыши к димеризации, но домен, необходимый для этого не был идентифицирован [5]. Ранее нами была разработана схема экспрессии и очистки N-концевого домена белка CTCF *Drosophila melanogaster* и была показана его способность к формированию мультимеров *in vitro* и *in vivo*, мультимеризация необходима для правильного функционирования белка [6]. N-концевой домен белка CTCF *Drosophila* не имеет гомологов среди известных белковых доменов, биоинформатическими методами предсказывается наличие короткого участка coiled-coil, который, однако, не является необходимым для мультимеризации.

С целью изучения особенностей пространственной структуры N-концевого домена белка CTCF *Drosophila* (1-163 ак) мы провели его исследование при помощи метода кругового дихроизма. Полученные данные (см. рис. 1) показывают практически полное отсутствие вторичной структуры, аналогичный результат был получен ранее для N-концевого домена белка CTCF человека [7]. Однако в отличие от него N-концевой домен белка CTCF *Drosophila* способен формировать мультимеры, как следует из результатов химической сшивки при помощи глутаральдегида, и, по результатам гель-фильтрации на носителе Superdex-200, имеет больший гидродинамический радиус, который соответствует формированию тетрамеров. Для подтверждения монодисперсности полученного нами препарата был также использован метод динамического рассеяния света (DLS) при длине волны 832 нм и концентрации белка от 1 до 7 мг/мл. Гидродинамический радиус был рассчитан по уравнению Стокса-Эйнштейна по стандартным процедурам, молекулярная масса была рассчитана по уравнению Марка-Куна-Хаувинка с параметрами для жесткой сферы. Полученные значения гидродинамического радиуса 4.4–4.6 нм соответствуют молекулярной массе 110 кДа, что согласуется с данными гель-фильтрации и соответствует формированию тетрамеров или мультимеров более высокого порядка. Однако, поскольку данные методы чувствительны к форме молекулы, удлиненная форма мультимеров или нативно развернутое состояние белка потенциально могут приводить к искажению результатов. Формирование мультимеров также было подтверждено в экспериментах *in vivo* [6]. Также следует отметить, что при высоком сходстве кластера доменов «цинковые пальцы» между белками CTCF человека и *Drosophila*, гомология между N-концевыми доменами практически отсутствует [2]. Эксперимент по температурной денатурации в присутствии красителя SYPRO Orange, обладающего сродством к гидрофобным участкам белка (ThermoFluor™), также не выявил присутствия гидрофобного ядра в N-концевом домене белка CTCF *Drosophila*. Полученные данные говорят в пользу того, что изучаемый домен является нативно развернутым, но в то же время способен формировать мультимеры и, по всей видимости, имеет компактную пространственную организацию.

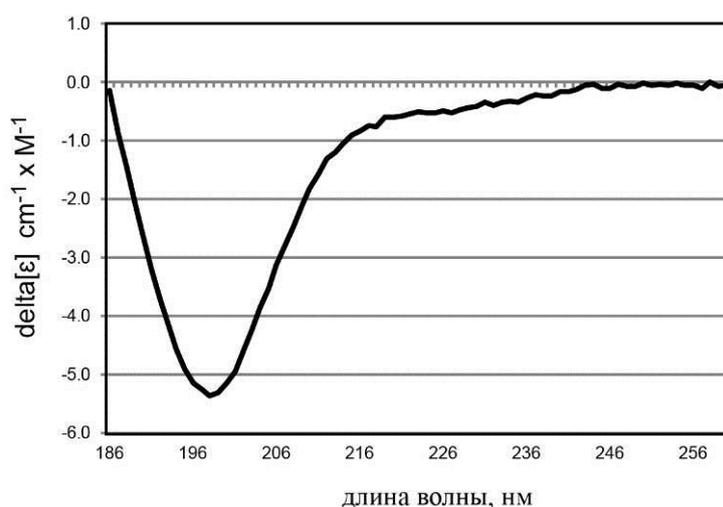


Рисунок 1 – Спектр кругового дихроизма N-концевого домена белка CTCF *Drosophila melanogaster*

Дальнейшее исследование N-концевого домена белка CTCF 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 1 до 7.5 мг/мл. Радиус гирации (R_g) белковой молекулы в растворе был рассчитан при помощи аппроксимации Гинье при малых углах ($s < 1.3/R_g$, $s = 4\pi \sin\theta/\lambda$, где 2θ – угол рассеяния), интенсивность $I(s)$ принималась равной $I_0 \exp(-sR_g)^2/3$. Для оценки максимального размера частиц D_{max} , функция парных расстояний $P(r)$ была рассчитана при помощи алгоритма GNOM из программного пакета ATSAS (см. рис. 2).

Молекулярный вес молекул был рассчитан исходя из исключенного объема гидратированных белковых молекул по закону Поро для гомогенных частиц. В результате эксперимента в зависимости от концентрации образца были получены следующие значения: $R_g = 4.65-5.0$ нм, $D_{max} = 15.2-16.0$ нм, что соответствует молекулярной массе 55.92-85.61 кДа и подтверждает формирование тетрамеров (молекулярная масса тетрамера составляет 18 кДа). *Ab initio* модели низкого разрешения для белка CTCF (1-163 ак) были получены при помощи алгоритма DAMMIN из программного пакета ATSAS (см. рис. 2). Удлиненная форма тетрамеров объясняет больший гидродинамический радиус, наблюдаемый при помощи методов гель-фильтрации и динамического рассеяния света. Двусторонняя симметрия полученных моделей позволяет предположить, что тетрамер состоит из двух прочно связанных между собой димеров, которые эффективно сшиваются между собой при химической сшивке с помощью глутаральдегида и других реактивов, что позволяет объяснить тот факт, что димер является основным продуктом реакции. Такая структура тетрамера также позволяет предположить механизм, в котором димеры белка CTCF кооперативно связываются с сайтами ДНК, которые часто находятся попарно, а их последующая тетрамеризация участвует непосредственно в установлении дальних взаимодействий в геноме. Неспособность аналогичного домена белка CTCF человека к мультимеризации *in vitro* [7], по всей видимости, объясняется проблемами, возникающими при его экспрессии в бактериальных клетках.

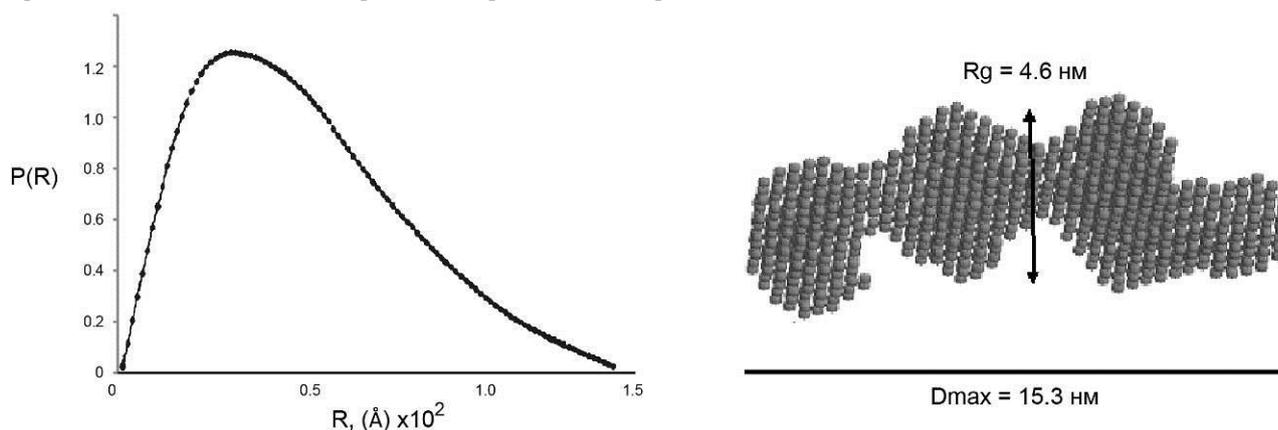


Рисунок 2 – Результаты анализа данных, полученных методом малоуглового рассеяния рентгеновских лучей (SAXS) для N-концевого домена белка CTCF *Drosophila melanogaster*. Слева – график функции парных расстояний. Справа – одна из возможных моделей низкого разрешения тетрамера

Таким образом, полученные нами данные демонстрируют, что N-концевой домен инсуляторного белка CTCF *Drosophila* не имеет вторичной структуры, но при этом имеет стабильную компактную глобулярную организацию и формирует тетрамеры, что подтверждает полученные ранее биохимические данные. К сожалению, более детальное исследование структуры домена затруднено в связи с невозможностью его кристаллизации и изучения при помощи метода ЯМР. Вероятно, явление формирования функциональных мультимеров неструктурированными белковыми доменами может быть широко распространено.

Работа выполнена при поддержке Российского научного фонда (проект 14-24-00166).

Список литературы / References:

1. Kyrchanova O., Chetverina D., Maksimenko O., Kullyev A., Georgiev P. Orientation-dependent interaction between *Drosophila* insulators is a property of this class of regulatory elements. *Nucleic Acids Res.*, 2008, vol. 36 (22), pp. 7019-7028.
2. Moon H., Filippova G., Loukinov D., Pugacheva E., Chen Q., Smith S., Munhall A., Grewe B., Bartkuhn M., Arnold R. CTCF is conserved from *Drosophila* to humans and confers enhancer blocking of the Fab-8 insulator. *EMBO Rep.*, 2005, vol. 6 (2), pp. 165-170.
3. Bell A., West A., Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, 1999, vol. 98 (3), pp. 387-396.
4. Kyrchanova O., Ivlieva T., Toshchakov S., Parshikov A., Maksimenko O., Georgiev P. Selective interactions of boundaries with upstream region of Abd-B promoter in *Drosophila* bithorax complex and role of dCTCF in this process. *Nucleic Acids Res.*, 2011, vol. 39, pp. 3042-3052.
5. Pant V., Kurukuti S., Pugacheva E., Shamsuddin S., Mariano P., Renkawitz R., Klenova E., Lobanenkov V., Ohlsson R. Mutation of a single CTCF target site within the H19 imprinting control region leads to loss of Igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Mol. Cell Biol.*, 2004, vol. 24 (8), pp. 3497-3504.
6. Bonchuk A., Maksimenko O., Kyrchanova O., Ivlieva T., Mogila V., Deshpande G., Wolle D., Schedl P., Georgiev P. Functional role of dimerization and CP190 interacting domains of CTCF protein in *Drosophila melanogaster*. *BMC Biol.*, 2015, vol. 13, p. 63.
7. Martinez S., Miranda J. CTCF terminal segments are unstructured. *Protein Sci.*, 2010, vol. 19 (5), pp. 1110-1116.