

СПЕЦИФИКА ПРОСТРАНСТВЕННОЙ СТРУКТУРЫ БЕЛКОВ ОБОЛОЧКИ ПОТЕКСВИРУСОВ

Долгов А.А.¹, Богачева Е.Н.¹, Ксенофонов А.Л.²

¹ ФИЦ химической физики им. Н.Н. Семенова РАН

ул. Косыгина, 4, г. Москва, 119991, РФ, e-mail: bogen_chph@mail.ru

² НИИ физико-химической биологии им. А.Н. Белозерского

Московского государственного университета им. М.В. Ломоносова

Ленинские горы, дом 1, стр. 40, г. Москва, 119992, РФ, e-mail: ksenofon@belozersky.msu.ru

Поступила в редакцию: 14.07.2020

Аннотация. Изучение структуры вирусных белков оболочки, а также возможностей их модификации и структурной перестройки крайне важно для разработки новых подходов для создания биотехнологической продукции, в том числе медицинского назначения. Вирусы растений обладают безусловным преимуществом при получении новых функциональных и биологически активных материалов, в частности перспективных вакцин, так как они не патогенны для млекопитающих, в том числе для человека. На примере типичного представителя группы потексвирусов - X-вируса картофеля (ХВК) исследована специфика пространственной структуры белков оболочки. Проведен системный анализ причин отклонений от глобулярной формы. Одна из причин - это биологическая функция защиты РНК. Проведено моделирование полноразмерного белка с помощью сервера I-Tasser и впервые рассмотрена конформация его N-конца. Предложен критерий компактности укладки полипептидной цепи в пространстве. В отличие от используемых в литературе эмпирических формул представлен способ расчета радиуса гирации белка по его первичной структуре, приведены значения радиусов гирации для различных типов аминокислотных остатков.

Ключевые слова: потексвирусы, X-вирус картофеля, моделирование пространственной структуры, радиус гирации.

Вирусы растений обладают безусловным преимуществом при получении новых функциональных и биологически активных материалов, в частности перспективных вакцин, так как они не патогенны для млекопитающих, в том числе для человека [1-4].

Потексвирусы представляют собой особую группу среди всех вирусов растений. Несмотря на низкую стабильность вирионов, многие вирусы этой группы накапливаются в растениях в очень больших количествах, уступая в этом отношении, наверное, только вирусу табачной мозаики. Кроме того, потексвирусы (как и другие гибкие вирусы растений) с трудом поддаются исследованию многими методами структурного анализа (ядерный магнитный резонанс, дифракция рентгеновских лучей на нитях и другие). Несмотря на то, что репликация потексвирусов в зараженных клетках изучена лучше многих других групп вирусов растений, о тонкой структуре частиц потексвирусов и их свободных белков оболочки известно немного.

Все это в полной мере относится к типичному представителю данной группы - X-вирусу картофеля (ХВК). Кроме того, белок оболочки (БО) ХВК обладает немалым числом собственных аномалий. Среди них можно назвать аномальную электрофоретическую подвижность и аномальные гидродинамические свойства БО, неспособность осаждаться в изоэлектрической точке и формировать вирусоподобные спиральные агрегаты и ряд других особенностей [5-7].

Вирионы ХВК представляют собой гибкие нитевидные частицы длиной 515 нм. 1270 субъединиц оболочки образуют левую спираль с диаметром $D = 13,0$ нм, шагом $H = 3,45$ нм и $n = 8,89$ субъединиц на виток [8]. Из опытных данных криоэлектронной микроскопии [8, 9] и тритиевой планиграфии [10-12] следует, что структура субъединицы ХВК существенно отличается от глобулярной, в частности, не существует гидрофобного ядра.

Высказано предположение, что изолированный белок оболочки ХВК, несмотря на высокое содержание альфа-спиральных участков, принадлежит к особой группе класса так называемых природно-неструктурированных белков. Они характеризуются повышенной гибкостью и, следовательно, аномальными гидродинамическими свойствами [13]. Главное отличие этих белков от структурированных (глобулярных) белков состоит в том, что они не имеют уникальной третичной структуры в изолированном виде, а приобретают ее после взаимодействия с партнерами. Их конформация в комплексе определяется партнером по взаимодействию, а не только собственной аминокислотной последовательностью, как это характерно для структурированных (глобулярных) белков.

Ключевым вопросом для сравнения структур является выбор критериев.

При анализе структуры использовали данные PDB-файла 6r7g [8], содержащего информацию о координатах центров атомов фрагмента вирусной оболочки ХВК. Нами предложен безразмерный нормированный критерий компактности укладки полипептидной цепи в пространстве

$$\sigma = \frac{R_g 2\sqrt{3}}{N * d_{CA}}$$

где d_{CA} – расстояние между последовательными α -атомами, N – число остатков, R_g – радиус гирации или инерции – среднее расстояние между центрами атомов $|r_k - r_j|$ объекта. По определению

$$R_g^2 = \frac{\sum_k (r_k - \langle r \rangle)^2}{N} = \frac{\sum_{j,k} (r_k - r_j)^2}{2N^2}, \quad (\langle r \rangle = \frac{\sum_{k=1}^N r_k}{N}).$$

Для полностью развернутой цепи $\sigma = 1$. В различных моделях цепей $R_g \sim N^\gamma$, где $\gamma = 0,5$ для свободно сочлененной цепи, $\gamma = 0,6$ в модели Флори, $\gamma = 0,588 \pm 0,001$ по расчетам теории возмущений [14]. С ростом N $\sigma \sim N^{\gamma-1} \rightarrow 0$. Выбор такого критерия обусловлен следующими доводами.

Выберем в качестве меры компактности системы из N точек среднее расстояние между точками, т.е. R_g . Для нормировки используем максимальное значение R_{gmax} , достигаемое при полностью развернутой цепи. Если длина звена d , то $R_{gmax}^2 = N^2 d^2 / 12$. Таким образом величина σ характеризует степень уменьшения среднего расстояния между точками по сравнению с максимальным для конфигурации цепочки, характеризующейся R_g .

Величину R_g можно непосредственно определить в экспериментах по малоугловому рассеянию нейтронов или рентгеновских лучей или ее можно рассчитать по данным о структуре, полученным другими методами (например, pdb-файлов). На возможность использования радиуса гирации для структурных оценок указано в работе [15].

В отличие от используемых в литературе эмпирических формул покажем, что радиус гирации системы, содержащей N точек и состоящей из k частей, так что в j -той части находится n_j точек ($j = 1..k$) можно представить в виде

$$R_g^2 = \sum_{j=1}^k n_j \cdot (d_j^2 + R_{gj}^2) / N$$

где d_j – расстояние от центра j -го компонента до общего центра, R_{gj} – радиус гирации j -того фрагмента. Радиус-вектор r_i i -той точки j -того фрагмента можно представить в виде $r_i = d_j + \rho_{ji}$, где ρ_{ji} – вектор, соединяющий центр j -го компонента с i -той точкой. Тогда

$$\sum_{j=1}^k \sum_{i=1}^{n_k} (d_j + \rho_{ji})^2 = \sum_{j=1}^k n_j d_j^2 + \sum_{i=1}^{n_k} \rho_{ji}^2 + 2 \cdot d_j \cdot \sum_{i=1}^{n_k} \rho_{ji} = \sum_{j=1}^k n_j (d_j^2 + R_{gj}^2)$$

или

$$R_g^2 = \sum_{j=1}^k v_j (d_j^2 + R_{gj}^2)$$

, где $v_j = n_j / N$ – доля j -го компонента.

Это дает возможность оценить вклад радиусов гирации аминокислотных остатков по первичной структуре белка.

Для элемента вторичной структуры – α -спирали из k точек с шагом H и n витков на оборот $r_j = [a \cdot \cos(j\gamma), a \cdot \sin(j\gamma), j \cdot h]$, $j = 0..k-1$, $h = H/n$, $\gamma = 2\pi/n$ можно получить формулу

$$R_{gh}^2 = a^2 \left(1 - \frac{\sin^2(\frac{k\gamma}{2})}{\sin^2(\frac{\gamma}{2})} \right) + \frac{h^2(k^2 - 1)}{12}$$

Для идеальной α -спирали $a = 0.75$ нм, $n = 3.6$, $H = 5.4$ нм, $h = 1.5$ нм, $k \geq 5$

$$R_{gh}^2 = 0.5625 \left(1 - \frac{\sin^2(\frac{k\pi}{3.6})}{\sin^2(\frac{\pi}{3.6})} \right) + 0.1875(k^2 - 1)$$

Для нахождения радиуса гирации и вычисления σ нами разработана программа, детали которой можно узнать по адресу alaldolgov48@gmail.com. С ее помощью проведены расчеты радиуса гирации субъединиц оболочки ХВК и σ на базе PDB-файла 6r7g [8], содержащего информацию о координатах центров атомов фрагмента вирусной оболочки ХВК, содержащего 13 последовательных субъединиц спиральной оболочки, а также используемого в качестве модельного глобулярного белка лизоцима. Вычисленные значения σ для каждой из этих субъединиц вируса практически совпадают. Найденные значения R_g сопоставлялись с расчетами по программе CRY SOL [16]. Отличие не превышало 0,2%. Радиусы гирации аминокислотных остатков рассчитывали, как среднее расстояние центра атома субъединицы до ее центра. Величина d_{CA} составила

Таблица 1. Значения R_g для различных типов аминокислотных остатков (АО)

АО	ALA	ARG	ASN	ASP	CYS	GLN	GLU	GLY	HIS	ILE
R_g , нм	0,147± 0,003	0,291± 0,022	1,952± 0,008	0,1955± 0,009	0,1737± 0,008	0,234± 0,009	0,233± 0,010	0,133± 0,006	0,226± 0,015	1,994± 0,005
АО	LEU	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	VAL
R_g , нм	0,208± 0,004	0,2590± 0,014	0,228± 0,012	0,240± 0,010	0,166± 0,006	0,161± 0,006	0,175± 0,003	0,274± 0,016	0,271± 0,002	0,178± 0,002

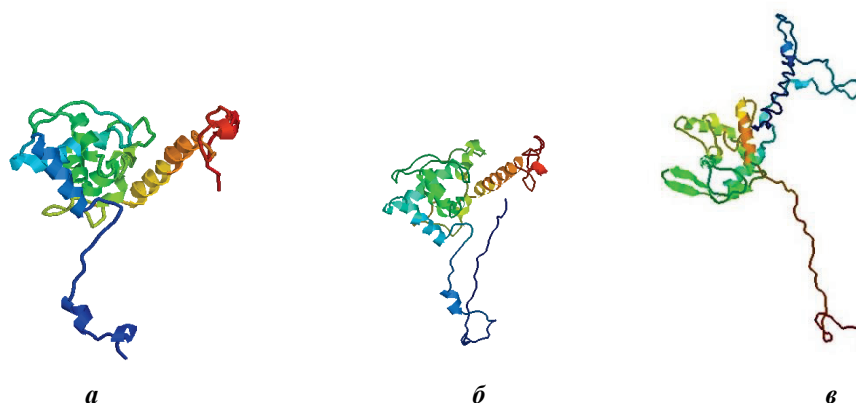


Рисунок 1. Пространственные модели X- и A-вирусов картофеля, построенные: а- по банку структур белков PDB (6r7g), б, в- по наиболее вероятным моделям белка X- и A- вируса, полученным с помощью сервера I-Tasser соответственно

$0,3793 \pm 0,0025$ нм. Для глобулярного белка лизоцима $\sigma = 0,09$. С увеличением σ силы дальнего взаимодействия между атомами ослабевают из-за увеличения расстояний между атомами, лабильность белка возрастает. Рассчитанные по данным 6r7g значения R_g для различных типов аминокислотных остатков приведены в таблице 1. Они в пределах ошибки совпадают с аналогичными значениями для лизоцима.

Основой для моделирования третичной структуры служила первичная структура X-вируса картофеля. Предсказание пространственной структуры выполняли с помощью сервера I-Tasser [17, 18]. При моделировании программа использовала, в основном шаблоны, полученные на основании данных криоэлектронной микроскопии (PDB-файлы 5fn1 и 6r7g). Впервые предсказан находящийся преимущественно в неупорядоченном виде N-конец макромолекулы. Отметим изменение вторичной структуры при моделировании полноразмерного белка. I-tasser практически не учитывает структурированность спиралей на уровне одного витка и часто уменьшает размеры спиралей, что, однако, не является критичным. Процент спиральности при этом уменьшился с 0,475 до 0,397 для модели 6r7g и построенной с помощью I-Tasser соответственно. Параметры соответствия для лучшей модели составили: C-score = -0,16; Расчетная ТМ-оценка = $0,69 \pm 0,12$ и RMSD = $6,0 \pm 3,7$ Å. С помощью сервера I-Tasser мы также выполнили предсказание и для субъединицы A-вируса картофеля (ABK) (269 АО). C-score=-1,34. Расчетная ТМ-оценка = $0,55 \pm 0,15$ и RMSD = $9,0 \pm 4,6$ Å. При этом рассчитанные значения σ для субъединицы ХВК из 6r7g (остатки 29-237), субъединиц ХВК и АВК (расчеты по I-Tasser) составили соответственно 0,09, 0,861 и 0,105. что указывает на большую неупорядоченность структуры АВК в сравнении с ХВК, что ранее было показано нами рядом физико-химических методов [19]. Построение и визуализацию пространственных структур модельных белков проводили с помощью программы RasMol 2.7.1. (<http://www.sourceforge.net/projects/openrasmol>).

Выводы. На примере типичного представителя группы потексвирусов - X-вируса картофеля (ХВК) исследована специфика пространственной структуры белков оболочки. Проведен системный анализ причин отклонений от глобулярной формы. Одна из причин - это биологическая функция защиты РНК. Предложен критерий компактности укладки полипептидной цепи в пространстве. В отличие от используемых в литературе эмпирических формул представлен способ расчета радиуса гирации белка по его первичной (или пространственной?) структуре, приведены значения радиусов гирации для различных типов аминокислотных остатков. Проведено моделирование полноразмерного белка с помощью сервера I-Tasser и впервые рассмотрена конформация его N-конца, отсутствующего в PDB- структуре.

Работа выполнена при поддержке гранта РФФИ 18-04-00525а и в рамках государственного задания Минобрнауки России (Тема 0082-2019-0015, № АААА-А20-120031490003-7).

Список литературы / References:

1. Урываев Л.В., Альховский С.В., Самохвалов Е.И., Беляев А.М и др. Вирусы как объекты и инструменты нанобиотехнологий. *Вопросы вирусологии*, 2012, т. S1, с. 52-65. [Uryvaev L.V., Alchovskii S.V., Samochvalov E.I., Belyaev A.M. et al. Viruses as objects and instruments in nanobiotechnology. *Problems of virology*, 2012, vol. S1, pp. 52-65. (In Russ.)]
2. Архипенко М.В., Петрова Е.К., Никитин Н.А., Дубровин Е.В. и др. Искусственные вирусоподобные частицы, полученные in vitro из белка оболочки X-вируса картофеля и чужеродных вирусных РНК. *Acta Naturae*, 2011, т. 3(10), с. 42-48. DOI: 10.32607/20758251-2011-3-3-40-46. [Arkhipenko M.V., Petrova E.K., Nikitin N.A., Dubrovin E.V. et al. Artificial virus-like particles obtained in vitro from potato X virus envelope protein and foreign viral RNA. *Acta Naturae*, 2011, vol. 3 (10), pp 42-48. (In Russ.)]
3. Lico C., Schoubben A., Baschieri S., Blasi P., Sant, L. Nanoparticles in biomedicine: new insights from plant viruses. *Curr. Med. Chem.*, vol. 20, 3471-3487 (2013). DOI: 10.2174/09298673113209990035.

4. Lico C., Benvenuto E., Baschieri S. Te two-faced potato virus X: from plant pathogen to smart nanoparticle. *Front. Plant Sci.*, 2015, vol. 6, p. 1009. DOI: 10.3389/fpls.2015.0109.
5. Немых М.А., Новиков В.К., Арутюнян А.М., Калмыков П.В., Драчев В.А., Добров Е.Н. Сравнительное исследование стабильности структуры субъединиц белка оболочки X-вируса картофеля в растворе и в составе вирусных частиц. *Молекулярная биология*, 2007, т. 41, № 4, с. 697-705. DOI: 10.1134/S0026893307040164. [Nemuch M.A., Novikov V.K., Arutyunyan A.M. et al. Comparative structural stability of subunits of the potato virus X coat protein in solution and virus particles. *Molecular biology*, vol. 41, no. 4, pp. 697-705. (In Russ.)]
6. Добров Е.Н., Немых М.А., Лукашина Е.В., Баратова Л.А., Драчев В.А., Ефимов А.В. Модифицированная модель структуры субъединиц белка оболочки X-вируса картофеля. *Молекулярная биология*, 2007, т. 41, № 4, с. 706-710. [Dobrov E.N., Novikov V.K., Lukashina E.V. et al. Modified model of the structure of the potato virus X coat protein. *Molecular Biology*, 2007, vol. 41, no. 4, pp. 638-641. (In Russ.)]
7. Архипенко М.В., Козловский С.В., Карпова О.В. Влияние N-концевой области белка оболочки X-вируса картофеля на структуру вирусных частиц. *Докл. Акад. Наук*, 2003, т. 391, № 1, с. 117-119. [Arkhipenko M.V., Kozlovsky S.V., Karpova O.V., Arkhipenko M.V., Zayakina O.V., Rodionova N.P., Atabekov I.G. *Effect of the N-terminal domain of the coat protein of the potato virus X on the structure of viral particles. Doklady Biochemistry and Biophysics*, 2003, vol. 391, no. 1, pp. 189-191. (In Russ.)]
8. Grinzato A., Kandiah E., Lico C., Betti C., Baschieri B., Zanotti G. Atomic structure of potato virus X, the prototype of the Alphaflexiviridae family. *Nat. Chem. Biol.*, 2020, vol. 16, pp. 564-569. DOI: 10.1038/s41589-020-0502-4.
9. Kendall A., McDonald M., Bian W. et al. Structure of Flexible Filamentous Plant Viruses. *Virology*, 2008, pp. 9546-9554. DOI: 10.1128 / JVI.00895-08.
10. Lukashina E., Ksenofontov A., Fedorova N. et al. Analysis of the role of the coat proteins n-terminal segment in potato virus x stability and functional activity. *Mol. Plant Pathology*, 2012, vol. 12, pp. 38-45. DOI: 10.1111/J.1364-3703.2011.00725.
11. Baratova L.A., Grebenshchikov N.I., Dobrov E.N., Gedrovich A.V., Kashirin I.A., Shishkov A.V., Efimov A.V., Järvekülg L., Radavsky Y.A., Saarma M. The organization of potato virus X coat proteins in virus particles studied by tritium planigraphy and model building. *Virology*, 1992, vol. 188, pp. 175-180. DOI: 10.1016 / 0042-6822 (92) 90747-d.
12. Lukashina E. et al. Tritium planigraphy study of structural alterations in the coat protein of Potato virus X induced by binding of its triple gene block 1 protein to virions. *FEBS J.*, 2009, vol. 276, pp. 7006-7015. DOI: 10.1111/j.1742-4658.2009.07408.x.
13. Сердюк В.Н. Структурированные белки и белки с внутренней неупорядоченностью. *Молекулярная биология*, 2007, т. 41, № 2, с. 297-313. [Serdyuk I.N. Structured proteins and proteins with intrinsic disorder. *Molecular Biology*, 2007, vol. 41, no. 2, pp. 262-27. (In Russ.)]
14. Дой М., Эдвардс С. *Динамическая теория полимеров*. Москва, «Мир», 1998. [Doi M., Edwards S.F. *The Theory of Polymer Dynamics*. Moscow, "Mir", 1998. (In Russ.)]
15. Лобанов М.Ю., Богатырева Н.С., Галзитская О.В. Радиус инерции – индикатор компактности белковых структур. *Молекуляр. биология*, 2008, т. 42, № 4, с. 701-706. [Lobanov M.Yu., Bogatyreva N.S., Galzitskaya O.V. Radius of gyration as an indicator of protein structure compactness. *Molecular Biology*, 2008, vol. 42, no. 4, pp. 623-628. DOI: 10.1134/S0026893308040195. (In Russ.)]
16. Svergun D., Barberato C., and Koch M. H. J. CRYSOLO – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography*, 1995, vol. 28, pp. 768-773. DOI: 10.1107/s0021889895007047.
17. Yang Zhang. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 2008, vol. 9, p. 40. DOI: 10.1186/1471-2105-9-40.
18. Jianyi Yang, Renxiang Yan, Ambrish Roy, Dong Xu, Jonathan Poisson, Yang Zhang. The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 2015, vol. 12, pp. 7-8. DOI: 10.138/nmeth.3213.
19. Ksenofontov A.L., Paalme V., Arutyunyan A.M., Semenyuk P.I., Fedorova N.V., Rumvold R., Baratova L.A., Jarvekulg L., Dobrov E. N. Partially disordered structure in intravirus coat protein of potyvirus potato virus A. *PLoS one*, 2013, vol. 8, e67830. DOI: 10.1371/journal.pone.0067830.

PECULARITY OF THE POTEXVIRUS COAT SPATIAL STRUCTURE

Dolgov A.A.¹, Bogacheva E.N.¹, Ksenofontov A.L.²¹N.N. Semenov Federal Research Center for chemical physics Russian Academics
Kosygin St. 4, Moscow, 117977, Russia; e-mail: bogen_chph@mail.ru²A.N. Belozersky Research Institute of Physico-Chemical Biology MSU
Leninskiye gory, 1, build. 40, Moscow, 119992, Russia; e-mail: ksenofon@belozersky.msu.ru

Abstract. The study of the structure of viral envelope proteins, as well as the possibilities of their modification and structural rearrangement is extremely important for the development of completely new approaches for creating biotechnological products, including medical ones. Plant viruses have an absolute advantage in the production of new functional and biologically active materials, in particular promising vaccines, since they are not pathogenic for mammals, including humans. On the example of a typical representative of the potexvirus group - potato X-virus, the specificity of the spatial structure of envelope proteins is studied. A systematic analysis of the causes of deviations from the globular form is carried out. One reason is the biological function of protecting RNA. A full-sized protein was simulated using the I-Tasser server and the conformation of its N-terminus was first considered. A compactness criterion for laying a polypeptide chain in space is proposed. In contrast to the empirical formulas used in the literature, a method for calculating the radius of gyration of a protein by its primary structure is presented, the values of the gyration radii for various types of amino acid residues are presented.

Key words: *potexvirus X-virus potato, simulation of the spatial structure, radius of gyration.*