

Аннотация работы:

Данное исследование направлено на изучение пространственной структуры некоторых белков бактериофагов методами нейросетевого моделирования в сочетании с методами молекулярной динамики. В рамках проекта были исследованы предполагаемые пространственные структуры некоторых фаговых белков, гены которых обнаруживаются в DGR-касетах бактериофагов, населяющих кишечник человека. Изучение таких генов позволяет раскрыть особенности узнавания фагами бактерий, а также особенности приспособления фагов к динамичному молекулярному составу бактериальной поверхности. В рамках исследования было показано, что некоторые из генов в составе DGR-кассет кодируют новый класс белков, названных тентаклинами. Такие белки содержат N-концевой лектиновый домен С-типа, за которым следуют один или несколько иммуноглобулин-подобных доменов. Лектиновый домен, предположительно участвующий в связывании с рецепторами на поверхности бактерий, содержит характерную бета-шпильку, которая подвергается активному мутагенезу под действием DGR-системы фага.

Помимо исследования структуры тентаклинов в рамках данного проекта была исследована пространственная структура рецептор-связывающих белков бактериофага StenM_174, поражающего бактерии рода *Stenotrophomonas*. Было показано, что белок gp43 с высокой достоверностью формирует гомотримерный белок хвостового шипа, который содержит домен ветвления и рецептор-связывающий домен. Также было показано, что белок gp44 является рецептор-связывающим белком с нехарактерным типом стыковки с белком gp43. Было показано, что белки gp45 и gp46 фага StenM_174 не являются белками хвостовых шипов.

Исследование выполнено при поддержке фонда РФФ, грант № 21-14-00360 (рук. Тикунова Н.В.), а также Министерства образования и науки РФ, грант № 075-15-2021-1085 (рук. Кузнецов Н.А.).

Тема исследования: «Моделирование структуры некоторых рецептор-связывающих белков бактериофагов»

Состав коллектива:

1. Байков Иван Константинович, возраст 38 лет, Институт химической биологии и фундаментальной медицины СО РАН (ИХБФМ СО РАН), старший научный сотрудник, к.х.н., ivan_baykov@mail.ru. Является также ассистентом преподавателя кафедры ХТТ ФЕН НГУ. Координатор, основной исполнитель.
2. Морозова Вера Витальевна, Институт химической биологии и фундаментальной медицины СО РАН (ИХБФМ СО РАН), старший научный сотрудник, к.б.н., morozova@niboch.nsc.ru.
3. Бабкин Игорь Викторович, Институт химической биологии и фундаментальной медицины СО РАН (ИХБФМ СО РАН), ведущий научный сотрудник, д.б.н., i_babkin@mail.ru.
4. Тикунова Нина Викторовна, Институт химической биологии и фундаментальной медицины СО РАН (ИХБФМ СО РАН), зав. Лабораторией молекулярной микробиологии, д.б.н., доцент, tikunova@niboch.nsc.ru

Информация о финансировании: исследование финансировалось фондом РФФ, грант № 21-14-00360 (рук. Тикунова Н.В.), а также Министерством образования и науки РФ, грант № 075-15-2021-1085 (рук. Кузнецов Н.А.).

Научное содержание работы

1. Постановка задачи:

Бактериофаги представляют собой вирусы бактерий и могут быть успешно использованы для фаготерапии. За узнавание бактерий отвечают рецептор-связывающие белки, которые могут быть организованы различным образом. Детальное понимание организации и функционирования таких белков позволяет проводить рациональный дизайн этих белков. Это является одним из необходимых этапов для создания синтетических бактериофагов с измененными свойствами (в том числе, специфичностью узнавания). Синтетические бактериофаги могут иметь массу приложений, в том числе для улучшения здравоохранения населения, сохранения урожая и т.д. Именно поэтому изучение архитектуры и функционирования бактериофагов, и в частности их рецептор-связывающих белков, является актуальной и востребованной задачей.

Цель вычислительной части проекта за отчетный этап – методами молекулярной динамики провести релаксацию структур фаговых белков, генерированных в ColabFold (реализация AlphaFold2). Данный этап был выполнен на вычислительном кластере ИВЦ НГУ.

2. Подробное описание работы, включая используемые алгоритмы

В качестве объектов для данного этапа были использованы пространственные структуры фаговых белков, сгенерированные с использованием ресурсов Google Colab. Была использована версия ColabFold v. 1.5.3 (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>), которая является реализацией нейросетевого алгоритма AlphaFold2. Были использованы пространственные структуры следующих белков: 1) тентаклины nd4_tgt1 и nd12_tgt1 из бактериофагов nd4 и nd12, обнаруженных при сборке кишечного вирома здорового донора, 2) С-лектины nd4_tgt2 и nd12_tgt2 из бактериофагов nd4 и nd12, обнаруженных при сборке кишечного вирома здорового донора, 3) белки gp43, gp44, gp45 и gp46 фага StenM174.

Молекулярно-динамические эксперименты осуществляли в пакете GROMACS 2020.3 на ресурсах Информационно-вычислительного центра Новосибирского государственного университета (ИВЦ НГУ) с использованием ускорителей Nvidia Tesla V100. Молекулу белков или их части «помещали» в растворитель (100 мМ водный раствор NaCl), проводили минимизацию энергии, уравнивание при температуре 300К. Основную стадию симуляции выполняли в течение 25-100 нс с использованием силового поля Amber ff99SB. Траектории были проанализированы в программе VMD.

3. Полученные результаты

В результате проведенного исследования было показано, что ген tgt1 бактериофагов nd4 и nd12, обнаруженных при сборке кишечного вирома здорового донора, кодирует белки с необычной топологией (содержат N-концевой лектиновый домен С-типа, за которым следуют один или несколько иммуноглобулин-подобных доменов, при этом N-концевой иммуноглобулиновый домен обладает характерной аминокислотной консенсусной последовательностью). Подобные белки были обнаружены в других метагеномных бактериофагах, а также в бактериях. Белки были названы тентаклинами. Кроме того, было обнаружено, что С-лектиновые домены белков, кодируемых генами tgt1 и tgt2 фагов nd4, nd12 и им подобных, содержат сходную по

структуре бета-шпильку, аминокислотная последовательность которой изменяется под действием DGR-кассет этих бактериофагов (рис. 1). Результаты опубликованы в статье.

Также в результате проведенного исследования было показано, что белки gr43 и gr44 фага StenM_174 являются рецептор-связывающими белками. Белок gr43 является гомотримерным белком хвостового шипа с содержит домен ветвления, к которому предположительно прикрепляется белок gr44 (рис. 2). Мультимерная форма белка gr44 не установлена, но показано, что моделирование гомотримера приводит к неудовлетворительным результатам. Показано, что белок gr44 содержит нетипичный С-концевой домен, поскольку структуру этого домена не удается достоверно смоделировать с помощью AlphaFold2. Результаты опубликованы в статье.

Иллюстрации:

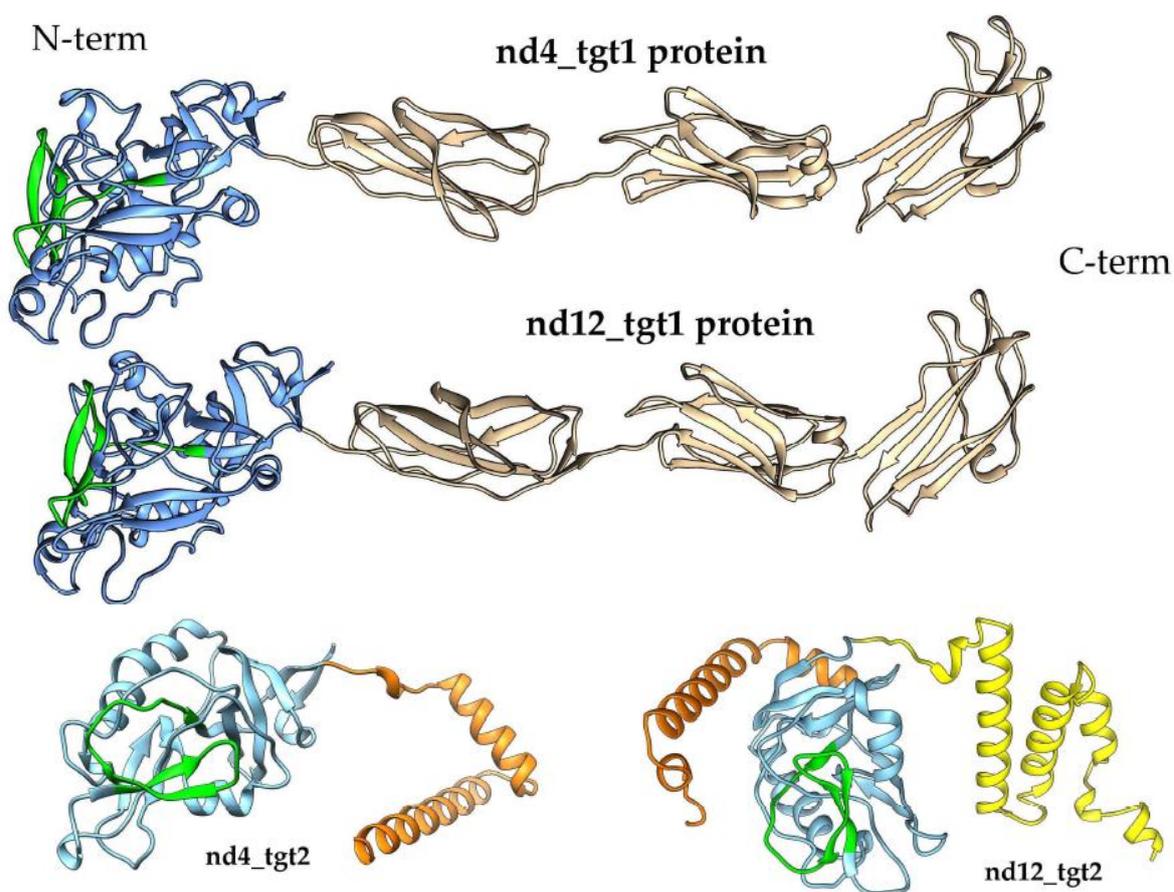
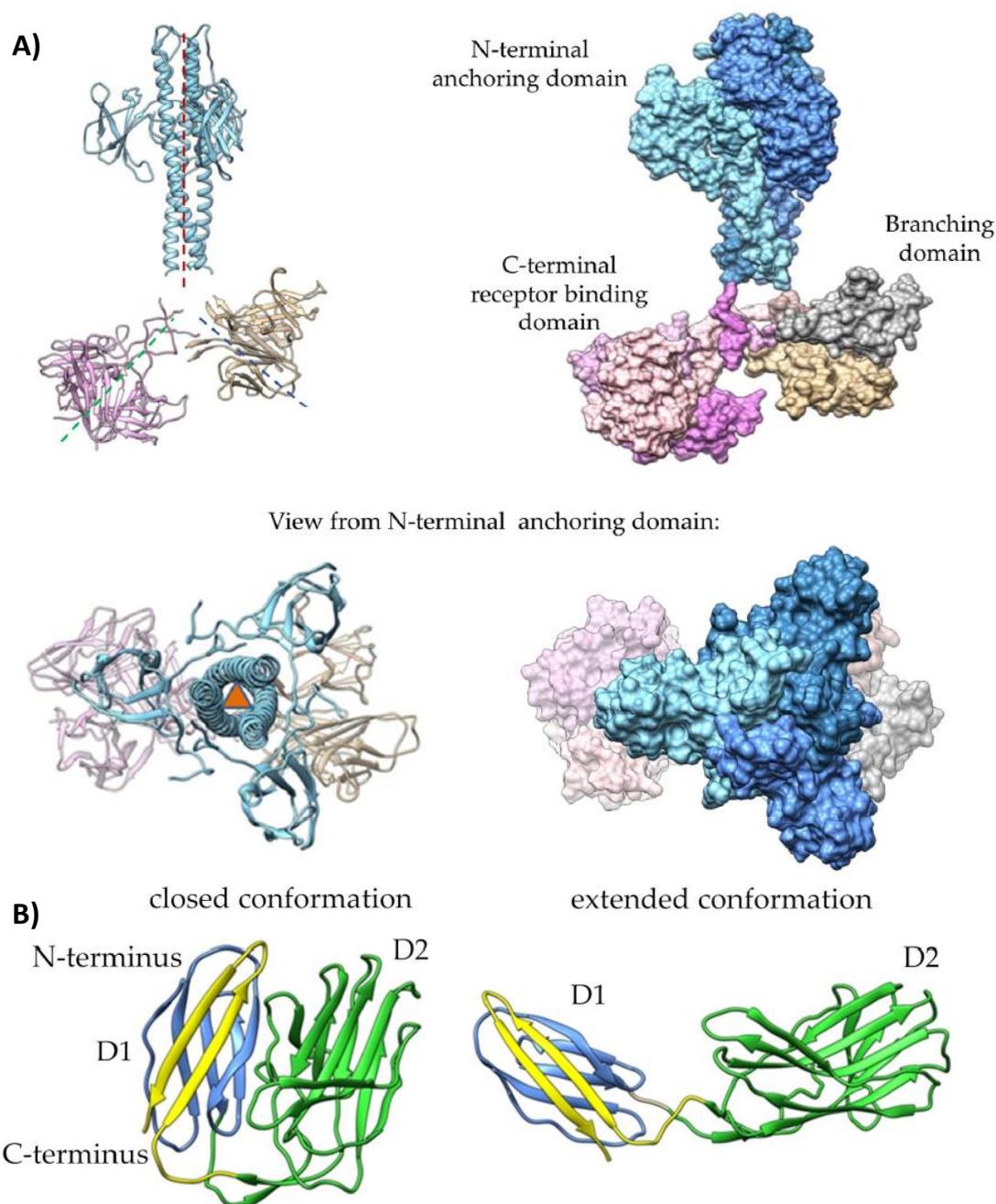


Рис. 1. Пространственная структура тентаклинов nd4_tgt1 и nd12_tgt1, а также белков, кодируемых геном tgt2 бактериофагов nd4 и nd12. Синим и голубым показаны лектиновые доменты С-типа. Зеленым отмечена бета-шпилька, последовательность которой изменяется под действием DGR-системы бактериофага. Бежевым показаны иммуноглобулин-подобные домены тентаклинов. Желтым и оранжевым показаны дополнительные альфа-спиральные домены белков tgt2.



```

WIDGIYEDYDAGDPYEGRVQIHGSVGLMKVEILGGNLPNASVFDQMTKEVVVKLKYTPPTIKVTE
VPNGSFEDNSIWETIGPGTPATAEQGWSPNGKGNLTYRDQKGEYTVFGGWAPVTNRTRQIEITGKVEH
GKSSKGNASCGVGLAWYNDKRELVRIDMGSIVDNNGGKGNWYKTSGIYSAHEPEIREVFPVIRFNRRKQ
NHPIHAGQIEWDHVYEIGYNEDETLWVEVKVTDSTHNVAIHRGIIEE

```

Рис. 2. Пространственная структура рецептор-связывающих белков gp43 (A) и gp44 (B) бактериофага StenM_174. A) Синими оттенками показан N-концевой якорный домен, за счет которого белок встраивается в хвост фага, коричнево-серыми – домен ветвления, к которому предположительно прикрепляется белок gp44. Розовыми оттенками показан C-концевой рецептор-связывающий домен. Пунктирными линиями показаны локальные оси симметрии третьего порядка. B) Желтым и синим показаны бета-листы, формирующие домен D1. Зеленым показан jelly-roll домен, предположительно обладающий рецептор-связывающей или рецептор-гидролизующей активностью.

Публикации/конференции:

1. Baykov IK, Tikunov AY, Babkin IV, Fedorets VA, Zhirakovskaia EV, Tikunova NV. **Tentaclins - A Novel Family of Phage Receptor-Binding Proteins That Can Be Hypermutated by DGR Systems.** *Int J Mol Sci.* 2023;24(24):17324. Published 2023 Dec 10. doi:10.3390/ijms242417324
2. Morozova VV, Yakubovskij VI, Baykov IK, et al. **StenM_174: A Novel Podophage That Infects a Wide Range of *Stenotrophomonas* spp. and Suggests a New Subfamily in the Family *Autographiviridae*.** *Viruses.* 2023;16(1):18. Published 2023 Dec 21. doi:10.3390/v16010018
3. Тикунова Н.В., Устный доклад «**От новых белковых структур – к новым микроорганизмам**» на школе-конференции "Современные вызовы структурной и синтетической биологии" (3-7 апреля 2024).

Эффект от использования кластера в достижении целей работы:

Использование кластера позволило выполнить продолжительную релаксацию моделей, полученных с помощью ColabFold, и получить более достоверные пространственные структуры.