МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА
ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ,
НИИ ФИЗИКО-ХИМИЧЕСКОЙ БИОЛОГИИ имени А. Н. БЕЛОЗЕРСКОГО

На правах рукописи

Шарапова Яна Александровна

Поиск новых путей регуляции функциональных свойств нейраминидазы NanA как ключевого фермента патогенеза *Streptococcus pneumoniae* с использованием методов компьютерной биологии

Специальность 03.01.08 – Биоинженерия

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата биологических наук

Научный руководитель:
д.х.н., профессор Швядас Витаутас Юозапас-Каятоно

Москва – 2021

ОГЛАВЛЕНИЕ

1. СПИСОК СОКРАЩЕНИЙ 5

2. ВВЕДЕНИЕ 6

3. ОБЗОР ЛИТЕРАТУРЫ 14

3.1. Суперсемейство сиалидаз 14

3.1.1. Общие сведения 14

3.1.2. Структурно-функциональная классификация по версии CATH и CAZy 15

3.1.3. Сиалидазы вирусов и микроорганизмов – ключевые факторы патогенности 17

3.2. Нейраминидаза вируса гриппа 18

3.2.1. Общие сведения о вирусе гриппа и вирусной нейраминидазе 18

3.2.2. Структура и каталитический механизм 23

3.3. Бактериальные нейраминидазы из Streptococcus pneumoniae 31

3.3.1. Патогенез *Streptococcus pneumoniae*: колонизация и инвазия 31

3.3.2. Нейраминидазы NanA, NanB и NanC из *Streptococcus pneumoniae*: структура и каталитический механизм 34

3.3.3. Роль NanA, NanB и NanC в патогенезе пневмококка и формировании биопленок, синергия с нейраминидазой вируса гриппа 41

3.4. Нейраминидазы - перспективные мишени для создания лекарств от заболеваний дыхательных путей человека 47

3.4.1. Современное состояние исследований и доступные лекарственные препараты: нейраминидаза вируса гриппа 47

3.4.2. Современное состояние исследований и доступные лекарственные препараты: бактериальные нейраминидазы из *Streptococcus pneumoniae* 51

4. МЕТОДЫ 54

4.1. Сбор первичной информации: структуры и последовательности 54

4.2. Биоинформатический анализ 55

4.2.1. Поиск гомологов и построение множественных выравниваний нейраминидаз из различных источников 55
4.2.1.1. Сравнительный анализ аминокислотных последовательностей 55
4.2.1.2. Сравнительный анализ 3D-структур 56
4.2.1.3. Построение структурно-опосредованного множественного выравнивания 56
4.2.2. Анализ множественных выравниваний последовательностей и 3D-структур 59
4.3. Молекулярное моделирование 60
4.3.1. Подготовка полноатомных молекулярных моделей 60
4.3.1.1. Подготовка моделей лигандов 60
4.3.1.2. Подготовка модели фермента/фермент-субстратного комплекса 61
4.3.2. Молекулярная динамика 62
4.3.2.1. Использование силовых полей FF14SB и FF15IPQ для параметризации моделей 62
4.3.2.2. Протокол оптимизации моделей (минимизация и уравновешивание) 64
4.3.2.3. Протокол свободной динамики и эффективное использование HPC-ресурсов для его исполнения 65
4.3.3. Моделирование полноразмерной структуры NanA 66
4.3.3.1. Подготовка исходных данных с использованием биоинформатического анализа 66
4.3.3.2. Подготовка исходных данных с использованием белок-белкового докинга 67
4.3.3.3. Гомологичное моделирование полноразмерной структуры NanA 68
4.3.3.4. Контрольные эксперименты для оценки полученных моделей NanA 72
4.3.4. Моделирование межмолекулярной агрегации NanA 76
4.3.5. Моделирование взаимодействия низкомолекулярных лигандов с NanA 79
4.3.5.1. Докинг с использованием программы LeadFinder 79
4.3.5.2. Моделирование связывания артокарпина в структуре NanA 80
4.4. Визуализация 84
5. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ 85

5.1. Подавление функциональных свойств нейраминидазы NanA ингибиторами межмолекулярной агрегации 85

5.1.1. Пространственная организация каталитического и лектинового доменов в структурах бактериальных нейраминидаз 85

5.1.2. Молекулярное моделирование уникальной структуры NanA, в которой каталитический и лектиновый домены соединенны гибким линкером из 16 аминокислот 90

5.1.3. Биоинформатический анализ пространственной обособленности лектинового и каталитического доменов в гомологичных сиалидазах 101

5.1.4. Моделирование межмолекулярной агрегации NanA 101

5.1.5. Моделирование взаимодействия ингибитора артокарпина с новым аллостерическим центром межмолекулярной агрегации 108

5.1.6. Роль конформационной подвижности и межмолекулярной агрегации NanA в формировании биопленок пневмококка 111

5.2. Подавление функциональных свойств нейраминидазы при связывании ингибиторов во втором аллостерическом центре NanA 118

5.2.1. Аннотация нового аллостерического центра в структуре NanA с использованием биоинформатического анализа 119

5.2.2. Молекулярное моделирование взаимодействия ингибитора оптактина с аллостерическими центрами в структурах NanA и NanB 120

5.2.3. Молекулярное моделирование взаимодействия олигосахаридного рецептора с аллостерическим центром в структуре NanA 125

5.3. Подавление функциональных свойств нейраминидазы NanA ингибиторами двойного действия, взаимодействующими с консервативными остатками активного центра нейраминидаз патогенов 129

5.3.1. Сравнительный анализ активных центров бактериальных и вирусных нейраминидаз. 129

5.3.2. Молекулярное моделирование комплексов бактериальных и вирусных нейраминидаз с ингибиторами двойного действия 131

6. ЗАКЛЮЧЕНИЕ 136

7. ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ 138

8. СПИСОК ЛИТЕРАТУРЫ 140

1. СПИСОК СОКРАЩЕНИЙ

S. pneumoniae – *Streptococcus pneumoniae*

NanA – нейраминидаза A из *S. pneumoniae*

NanB – нейраминидаза B из *S. pneumoniae*

NanC – нейраминидаза C из *S. pneumoniae*

PDB – Protein Data Bank

EC – Enzyme Commission

CAZy – Carbohydrate-Active enZYmes

GH – Glycoside Hydrolase – гликозилгидролазы

CBM – углевод-связывающий модуль

BgaA – β-галактозидаза A из *S. pneumoniae*

DANA – 2-дезокси-2,3-дегидро-N-ацетилнейраминовая кислота

МД – молекулярная динамика / молекулярно-динамический

а.о. – аминокислотный остаток

CPU – central processing unit, или процессор

GPU – graphical processing unit, или видеокарта

APBS – Adaptive Poisson-Boltzmann Solver

2. ВВЕДЕНИЕ

1. Актуальность темы исследования

Среди трех нейраминидаз (NanA, NanB и NanC), продуцируемых пневмококком (*Streptococcus pneumoniae),* наибольшее внимание привлекает нейраминидаза А (NanA) – один из ключевых белков в патогенезе. Попав в носоглотку человека, эта болезнетворная бактерия прикрепляется к эпителию и образует биопленку, становясь постоянным, трудноискоренимым очагом инфекции и воспаления. Из носоглотки бактерия может распространяться в другие части организма, вызывая пневмонию, отит и менингит. Многочисленные исследования показали, что нейраминидаза А необходима пневмококку для колонизации: отрезая сиаловую кислоту от олигосахаридных рецепторов на поверхности эпителия, NanA способствует бактериальной адгезии, питанию и росту бактерий, а также образованию биопленки. Нейраминидаза А пневмококка является гомологом активно изучаемой нейраминидазы вируса гриппа. Известно, что совместное инфицирование человека вирусом гриппа и пневмококком может привести к опасным осложнениям. По этим причинам нейраминидазы обоих патогенов считаются перспективными мишенями для действия лекарств.

2. Степень разработанности темы исследования

Несмотря на важность NanA в жизнедеятельности бактерии, этот фермент еще мало изучен: не расшифрована полная кристаллографическая структура, не известно ни одного функционального/регуляторного участка связывания помимо каталитического центра (хотя недавние исследования предполагают наличие аллостерических центров в практически любых белках), неясны

механизмы участия NanA в формировании биопленки. Недостаточное понимание структурной организации и функциональных особенностей действия NanA затрудняет возможность использования этого фермента в качестве мишени для разработки антибактериальных препаратов.

3. Цель и задачи исследования

Целью настоящего исследования являлся поиск новых путей регуляции функциональных свойств нейраминидазы NanA как ключевого фермента в патогенезе *Streptococcus pneumoniae*. Для достижения поставленной цели были сформулированы следующие задачи:

1. Провести сравнительный биоинформатический анализ информации о последовательностях и структурах родственных нейраминидаз/сиалидаз для выявления общих черт структурной организации, закономерностей функционирования и регуляции эволюционно удаленных гомологов из различных бактерий и вирусов;
2. С использованием биоинформатического анализа и молекулярного моделирования изучить особенности организации участков связывания лигандов в структуре NanA и охарактеризовать неизвестные ранее функциональные сайты;
3. С использованием молекулярного моделирования изучить механизмы ингибирования NanA при связывании низкомолекулярных лигандов в активном центре и иных топологически независимых от него участках в структуре фермента.

4. Научная новизна

Впервые проведено детальное исследование особенностей структурной организации NanA из *Streptococcus pneumoniae* на основании сравнения с известными структурами гомологичных сиалидаз и комплексного моделирования. Установлено, что, в отличие от других сиалидаз, домены в молекуле NanA не вступают в образование цельной структуры, что вкупе с наличием неупорядоченных участков структуры придает этому модульному ферменту гибкость и подвижность. Впервые показано, что, благодаря гибкости структуры NanA, две молекулы фермента способны образовать стабильный межмолекулярный комплекс. Поскольку каждая молекула NanA прикреплена к поверхности бактерии длинным гибким участком полипептидной цепи, агрегация молекул NanA, находящихся на соседних бактериальных клетках, способна связывать их в сеть, что может являться одним из механизмов участия NanA в формировании биопленки. В структуре NanA описаны два новых, неизвестных ранее аллостерических центра, охарактеризовано их строение и возможность использования для функциональной регуляции NanA. Один из них расположен в интерфейсе взаимодействия двух молекул NanA; связывание в этом аллостерическом центре артокарпина – экспериментально подтвержденного ингибитора каталитической функции и образования биопленок, препятствует образованию устойчивого межмолекулярного агрегата. Другой аллостерический центр расположен с противоположной стороны от активного центра, способен связывать сахара, входящие в состав олигосахаридных рецепторов и, возможно, способствует бактериальной адгезии. Выявлены различия в структурной организации участков связывания ингибиторов бактериальных и вирусных нейраминидаз, которые необходимо учитывать при создании ингибиторов двойного действия.

5. Теоретическая и практическая значимость

Изучение структурно-функциональных взаимосвязей в NanA и поиск новых путей ее регуляции представляет фундаментальный и практический интерес. В данной работе с помощью методов компьютерной биологии показано, что в отличие от большинства гомологичных сиалидаз, NanA из *S. pneumoniae* обладает уникальной гибкой, подвижной организацией, что, помимо лучшего понимания ее структуры и роли в патогенезе, открывает новые практические возможности для дизайна препаратов, нацеленных на ингибирование адгезии бактерий к эпителию и образования ими биопленки.

6. Методология и методы исследования

В работе использован комплекс современных методов биоинформатики и молекулярного моделирования, в том числе молекулярная динамика, докинг белок-белок и белок-лиганд, моделирование полноразмерной структуры белка по гомологии, а также сравнительный анализ структур и аминокислотных последовательностей. Важным методологическим аспектом в работе являлось использование ресурсов суперкомпьютера «Ломоносов-2». Для разных типов задач подбиралась конкретная конфигурация вычислительного оборудования, доступного на разных узлах и разделах суперкомпьютера с целью оптимизации эффективности расчета.

Все последовательности и структуры нейраминидаз брали из публичных баз данных, таких как PDB, UniProt, SwissProt, TrEMBL. Модифицированную версию PDB2PQR v. 1.9.0, совместимую с PROPKA v. 3.1, использовали для протонирования моделей ферментов. Пакет AmberTools использовали для параметризации систем в силовом поле FF15IPQ (дополнительную серию вычислений проводили также в силовом поле FF14SB и другом растворителе

для исключения влияния выбора параметров моделирования на результат). Модели ферментов помещали в ячейку с растворителем так, чтобы расстояние от любого атома белка до края периодической ячейки было не менее 12 Å (а в некоторых случаях 30 Å). Для молекулярной динамики использовали многоступенчатый протокол, описанный в полном тексте диссертации.

Докинг лигандов в белок проводили в программе LeadFinder. При поиске сайтов их связывания использовали большие ансамбли конформеров белка, полученные из длительных молекулярно-динамических симуляций. Белок-белковый докинг проводили с использованием веб-серверов ClusPro и ZDOCK (для исключения влияния выбора программы на результат); для него также использовали ансамбли конформеров.

Построение множественных структурно-опосредованных выравниваний гомологов проводили при помощи веб-сервера Mustguseal, который объединяет и автоматизирует классические методы биоинформатики, такие как BLAST, MAFFT, и др. Анализ полученных выравниваний осуществляли с использованием веб-сервера Zebra2 для поиска консервативных и специфических позиций.

7. Положения, выносимые на защиту

1) NanA из *S. pneumoniae* характеризуется гибкой структурной организацией благодаря соединению каталитического и лектинового доменов подвижным линкером из 16 аминокислотных остатков;
2) Молекулы NanA, прикрепленные к поверхности разных клеток пневмококка, способны образовать агрегаты за счет межмолекулярного взаимодействия лектиновых и каталитических доменов, что может являться одним из механизмов образования биопленок пневмококком;

3) В структуре NanA охарактеризованы два новых, неизвестных ранее аллостерических центра, расположенные в интерфейсе взаимодействия двух молекул NanA при межмолекулярной агрегации и в воронке шестилопастного β-пропеллера на противоположной стороне от активного центра;
4) Определены различия в структурной организации активных центров в NanA из патогенной бактерии *S. pneumoniae* и вирусных нейраминидазах, идентифицированы консервативные позиции, а также специфические остатки, отличающиеся у нейраминидаз из различных источников и обуславливающие селективность каждого центра. Полученные результаты создают основу для дизайна ингибиторов двойного действия, способных подавлять активность как бактериальных, так и вирусных нейраминидаз.

8. Степень достоверности

Результаты диссертационного исследования получены с помощью современных методов молекулярного моделирования и биоинформатического анализа. Вычислительные эксперименты проводились в несколько повторений с соответствующими контрольными экспериментами и хорошо воспроизводились. Дополнительные серии вычислений проводили в другом силовом поле и растворителе, а также с использованием другой программы для исключения влияния выбора параметров моделирования на результат. При поиске сайтов связывания лигандов использовали большие ансамбли конформеров белка, полученные из длительных молекулярно-динамических симуляций. Результаты, полученные в данной работе, согласуются с экспериментальными данными из публикаций других исследовательских групп. Литературный обзор и обсуждение результатов подготовлены с использованием

всей доступной литературы по теме. Результаты диссертационного исследования опубликованы в рецензируемых научных журналах.

9. Личный вклад автора в проведение исследования

Личный вклад соискателя заключался в анализе литературных источников, планировании и проведении вычислительных экспериментов, анализе и интерпретации полученных результатов, подготовке научных публикаций, представлении результатов на конференциях. Основные результаты работы получены самим автором.

10. Публикации по теме диссертации

По материалам диссертации опубликованы 4 статьи в рецензируемых отечественных и зарубежных журналах, индексируемых в наукометрических базах данных Web of Science и/или Scopus.

11. Апробация работы

Часть материалов диссертации защищена в 2019 году в рамках НКР (научно-квалификационной работы), подготовленной во время обучения в аспирантуре факультета биоинженерии и биоинформатики МГУ имени М. В. Ломоносова. Результаты диссертационного исследования были представлены на следующих конференциях: International Symposium «Systems Biology and Bioinformatics» (Санкт-Петербург, Россия, 2016), Moscow Conference on Computational Molecular Biology (MCCMB’17 и MCCMB’19, Москва, Россия, 2017 и 2019 гг.), Международных научных конференциях студентов, аспирантов и молодых учёных «Ломоносов-2018», «Ломоносов-2019» и «Ломоносов-2020» (Москва, Россия, 2018, 2019 и 2020 гг.), Russian Supercomputing Days (Москва, Россия, 2019), The 44th FEBS Congress (Краков, Польша, 2019).

12. Структура и объем диссертации

Диссертационная работа состоит из следующих разделов: оглавление, список сокращений, введение, обзор литературы, методы, результаты и их обсуждение, заключение, основные результаты и выводы, список литературы. Работа изложена на 160 страницах, содержит 44 иллюстрации и цитирует 261 литературный источник.

3. ОБЗОР ЛИТЕРАТУРЫ

3.1. Суперсемейство сиалидаз

3.1.1. Общие сведения

Сиалидазы (или, как их часто называют, нейраминидазы) широко распространены в природе и встречаются чаще всего у бактерий и вирусов, а также у вторичноротых животных и некоторых протистов. Считается, что впервые сиалидазы появились у древних вторичноротых животных, а от них попали к микроорганизмам, среди которых широко распространились с помощью горизонтального переноса генов [1]. Сиалидазы катализируют гидролитическое отщепление α-присоединенной сиаловой кислоты от таких субстратов, как гликопротеины, гликолипиды, полисахариды. Сиаловая кислота – это общее наименование для ~50 производных нейраминовой кислоты, встречающихся в биологических системах (Рисунок 1) [2]. Несмотря на низкую идентичность аминокислотных последовательностей, большое разнообразие в субстратной специфичности и других биохимических свойствах различных сиалидаз, гомология прослеживается на молекулярном уровне, что позволяет говорить о сиалидазах как о далеких гомологах, объединенных в суперсемейство общим происхождением и основными структурно-функциональными характеристиками [3, 4]. Для каталитического домена большинства сиалидаз характерна укладка 6-лопастного β-пропеллера. Считается, что для правильного сворачивания полипептидной цепи важны т.н. «Asp-boxes» – мотивы SxDxGxxW/F, встречающиеся в β-поворотах лопастей пропеллера в количестве 1-5 штук на последовательность в большинстве сиалидаз. Также характерной чертой сиалидаз является наличие RIP-мотива, который содержит один из ключевых каталитических аргининов [1, 4]. Структура каталитического центра сохраняется неизменной у всех экзо-α-

сиалидаз (кроме семейства GH156, см. главу 3.1.2) и включает триаргининовый кластер, необходимый для связывания карбоксилатной группы, общей для всех сиаловых кислот, а также Tyr/Glu нуклеофильную пару и кислоту/основание Asp [5]. Каталитические домены экзо- и эндо-сиалидаз структурно схожи, но в активном центре последних отсутствует один из трех аргининов, тирозин и аспартат или глутамат [6]. Дополнительные карбоксильные группы, присутствующие в полисиаловом субстрате эндо-сиалидаз, вероятно, компенсируют отсутствие ключевых каталитических остатков [7].

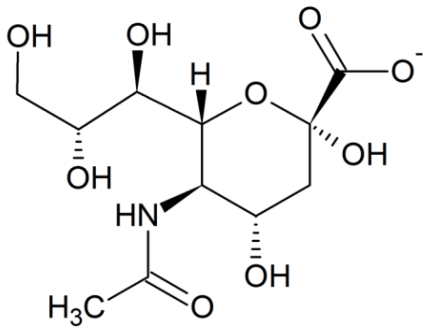

Рисунок 1. N-ацетилнейраминовая кислота (Neu5Ac) – наиболее распространенная сиаловая кислота в организме человека.

3.1.2. Структурно-функциональная классификация по версии CATH и CAZy

В соответствии с субстратной специфичностью и каталитическим механизмом, сиалидазы делятся на четыре класса:

- Гидролитические экзо-α-сиалидазы, или нейраминидазы (EC 3.2.1.18), которые отщепляют концевой остаток сиаловой кислоты и высвобождают свободную сиаловую кислоту. Обычно имеют широкую субстратную специфичность к α2-3-, α2-6-, α2-8-присоединенным концевым сиаловым кислотам.

- Транс-сиалидазы (или сиалилтрансферазы, EC 3.2.1.18) отщепляют концевой остаток сиаловой кислоты и переносят его на другой гликоконъюгат. Отдают предпочтение α2-3-присоединенным сиаловым кислотам.

- Интрамолекулярные транс-сиалидазы, или ангидросиалидазы (EC 4.2.2.15), специфичны к α2-3-присоединенным сиаловым кислотам и производят 2,7-ангидрид сиаловой кислоты.

- Эндо-сиалидазы (EC 3.2.1.129) катализируют эндо-гидролиз α2-8-гликозидных связей в олиго- и поли-сиаловых кислотах.

Согласно классификации CAZy, основанной на сходстве аминокислотных последовательностей, суперсемейство сиалидаз представлено семействами гликозилгидролаз GH33, GH34, GH83 и GH156 (все экзо-α-сиалидазы – нейраминидазы, транс-сиалидазы, ангидросиалидазы) и GH58 (эндо-α-сиалидазы). Семейства GH33, GH34 и GH83 относятся к подгруппе «retaining» (*с англ.* «сохраняющие»), поскольку конфигурация аномерного центра образующегося продукта реакции остается неизменной (α-конфигурация и субстрата, и продукта), и объединяются в клан GH-E. Сиалидазы семейств GH58 и GH156 – «инвертирующие» (*англ.* «inverting»): в результате катализа происходит, т.о., мутаротация α-субстрата в β-продукт. Сиалидазы из семейства GH156 кардинальным образом отличаются от всех остальных – вместо 6-лопастного β-пропеллера их последовательность уложена в $(\beta/\alpha)_8$-баррель, каталитические аминокислоты и механизм также другие [8].

Согласно аннотации CAZy, большинство бактериальных экзо-α-сиалидаз, в т.ч. нейраминидазы NanA, NanB, NanC из *S. pneumoniae*, а также сиалидазы эукариот относятся к семейству GH33. К семейству GH34 относятся

нейраминидазы вируса гриппа A и B, а нейраминидазы других вирусов – к семейству GH83. Вспомогательные домены для связывания субстрата, присутствующие в структурах многих гликозилгидролаз, также разделены на семейства в базе данных CAZy. Лектиновые домены, ассоциированные с сиалидазами семейства GH33, относятся к семейству CBM40 (в т.ч. и лектиновые домены пневмококковых нейраминидаз)[9, 10].

По структурной классификации CATH каталитические домены сиалидаз объединены в суперсемейство 2.120.10.10, что означает, что вторичная структура сиалидаз представлена в основном β-листами, представители имеют общую архитектуру 6-лопастного β-пропеллера и топологию и считаются произошедшими от общего предка, т.е. гомологичными. Лектиновые домены входят в суперсемейство 2.60.120.200 (β-сэндвичи, топология «jelly roll») [11].

3.1.3. Сиалидазы вирусов и микроорганизмов – ключевые факторы патогенности

Представители суперсемейства сиалидаз играют важную роль в жизнедеятельности многих организмов. Вирусы, вызывающие различные ОРВИ (парамиксовирусы), грипп, паротит (*Mumps rubulavirus*), болезнь Ньюкасла(*Avian avulavirus 1*), имеют в составе своего капсида сиалидазы или гемагглютинин-нейраминидазы – белки, которым присуща и сиалидазная, и гемагглютинирующая активность. Сиалидазы являются ключевыми факторами патогенности бактерий-возбудителей перитонита (*Bacterioides fragilis*, *Escherichia sp.*, *Enterococcus sp.*), холеры (*Vibrio cholerae*), гангрены (*Clostridium sp.*), пневмонии (*Streptococcus pneumoniae*) и других заболеваний [3, 16]. Некоторые бактерии мимикрируют под богатую сиаловыми кислотами среду в организме-хозяине путем сиалилирования собственных поверхностных белков. Например, транс-сиалидазы из *Neisseria meningitidis* и *Haemophilus influenzae*

отрезают остатки сиаловой кислоты с рецепторов человека и переносят их на поверхность бактерии, что помогает патогенам «спрятаться» от иммунной системы [12]. Было также показано, что нейраминидаза А (NanA) из *S. pneumoniae* отщепляет сиаловую кислоту с поверхности *N. meningitidis* и *H. influenzae* – сиалидазная активность, таким образом, может служить конкурентным преимуществом в борьбе с другими бактериями, занимающими ту же нишу [13]. Некоторые микроорганизмы, в т.ч. и непатогенные, используют отщепленную сиалидазой сиаловую кислоту организма-хозяина в качестве источника питания [14, 15, 16, 17, 18].

У млекопитающих было обнаружено четыре сиалидазы, различающиеся по внутриклеточной локализации, субстратной специфичности, кинетическим и физико-химическим свойствам – лизосомная NEU1, цитозольная NEU2, NEU3 в плазматической мембране и NEU4 из митохондриальной мембраны и эндоплазматического ретикулума. Сиалидазная активность в организме человека влияет на такие процессы, как метаболизм ганглиозидов и гликопротеинов, рост клеток, их пролиферацию и дифференцировку, на межклеточные взаимодействия, функционирование мембран и др. [19, 20].

3.2. Нейраминидаза вируса гриппа

3.2.1. Общие сведения о вирусе гриппа и вирусной нейраминидазе

Грипп – это острое инфекционное заболевание дыхательных путей, вызываемое вирусами родов А и В семейства *Orthomyxoviridae*. Генетический материал вируса гриппа представлен восемью сегментами одноцепочечной антисмысловой РНК. Сегменты вирусной РНК связаны с нуклеопротеином NP и гетеротримерным комплексом РНК-зависимой РНК-полимеразы (PB1, PB2 и PA), образуя рибонуклеопротеидный комплекс (RNP). RNP упакован в оболочку из матриксного белка M1 и, поверх M1, в билипидную мембрану, полученную

вирусной частицей при выходе из зараженной клетки. Внутри капсида, наряду с RNP, содержится также белок ядерного экспорта (NEP, или NS2), который необходим для транспорта вирусного генетического материала в клеточное ядро. Мембрану вириона пронизывают три вида интегральных белков – ионный канал M2, гемагглютинин (HA) и нейраминидаза (NA) (Рисунок 2). Организация вируса типа B сходна с вирусом типа A, но вместо ионного канала M2 в мембране присутствуют белки BM2 и NB. Также известны роды *Influenza C* и *D*, которые содержат лишь 7 сегментов РНК, функционально отличны от A и B и, по-видимому, не вызывают серьезных заболеваний у человека [21, 22]. Родовую принадлежность вируса гриппа определяют антигенные свойства главных внутренних белков вириона – M1 и NP [23]. Вирус гриппа типа A далее классифицируется по серотипам поверхностных белков гемагглютинина (HA) и нейраминидазы (NA) – ключевых факторов вирулентности. В соответствии с их антигенной специфичностью выделяют 18 подтипов HA и 11 подтипов NA. Два новых изолята из летучих мышей H17N10 и H18N11 сильно отличаются от известных ранее подтипов вируса: их гемагглютинин и нейраминидаза не проявляют своей классической активности [24].

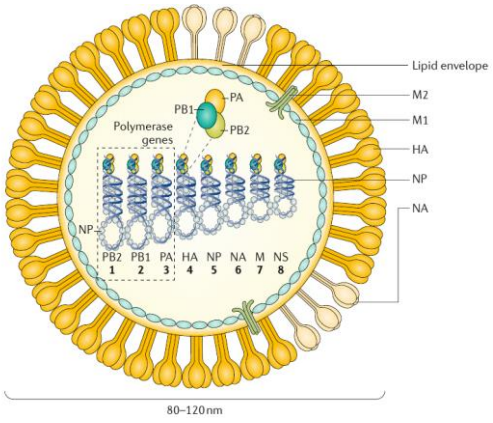

Рисунок 2. Строение вируса гриппа A. См. пояснения в тексте. Рисунок взят из [21].

Сегментированный геном вируса гриппа создает возможность для рекомбинации поверхностных антигенов – гемагглютинина и нейраминидазы. Если в клетку организма попали несколько разных подтипов вируса, то вирусное потомство может содержать сегменты РНК, кодирующие гемагглютинин и нейраминидазу из разных подтипов, в результате чего может образоваться совершенно новый серотип вируса. Этот процесс рекомбинации поверхностных антигенов называется *антигенным шифтом*. Также для вируса гриппа известен процесс *антигенного дрейфа*, отвечающий за вариацию в пределах подтипа. Антигенный дрейф обеспечивается отсутствием корректирующей (экзонуклеазной) активности у РНК-зависимой РНК-полимеразы. Вирус гриппа, таким образом, может быстро мутировать. Если мутации возникают в участках, отвечающих за связывание гемагглютинина или нейраминидазы с антителом, и препятствуют формированию комплекса антиген-антитело, то организм оказывается незащищенным перед натиском инфекции. Мутации вблизи участка связывания субстрата в гемагглютинине могут привести к изменению его субстратной специфичности. В таком случае может оказаться, что вирус, изначально инфицирующий лишь, например, птиц, приобретает способность заражать людей. Мутации в нейраминидазе способны делать вирус резистентным к существующим ингибиторам. Тем не менее, несмотря на потенциально огромное количество комбинаций HA и NA, только три подтипа вируса вызывали пандемии: H1N1 (1918 и 2009), H2N2 (1957), H3N2 (1968). На данный момент лишь H1N1 и H3N2, а также две линии вируса типа B (Victoria и Yamagata) присутствуют в человеческой популяции, вызывая сезонные эпидемии [25, 26, 27, 28, 54]. Вирус гриппа B не способен вызывать пандемии, вероятно, из-за крайней ограниченности его природного резервуара: к настоящему моменту он был обнаружен только у людей и тюленей [29]. Природным резервуаром вируса гриппа A являются дикие птицы и, возможно,

летучие мыши. От них вирус через зараженную воду передается домашним животным (свиньям, уткам, курам, лошадям, кошкам и собакам), а от них воздушно-капельным путем заражается человек [21].

Жизненный цикл вируса гриппа происходит, главным образом, в эпителиальных клетках дыхательных путей человека и других млекопитающих, и в эпителиальных клетках желудочно-кишечного тракта птиц. Инфекция начинается с прикрепления вируса к клетке-мишени. Гемагглютинин распознает концевой остаток сиаловой кислоты олигосахаридных рецепторов на поверхности эпителиальной клетки. У человека сиаловая кислота присоединена к предыдущему сахару, в основном, α2,6-кетозидной связью, а у птиц – α2,3-связью; α2,3-присоединенная сиаловая кислота в небольшом количестве также присутствует на дыхательном эпителии человека, что объясняет возможность заражения человека вирусом птичьего гриппа [30, 31]. И хотя эффективность такого заражения в целом снижена, оно представляет серьезную угрозу здоровью и даже жизни. α2,3-присоединенная сиаловая кислота чаще встречается в нижних дыхательных путях (в бронхиолах и альвеолах) – если вирус попадает в легкие, это приводит к стремительному развитию пневмонии, часто заканчивающейся летальным исходом [32].

Связавшийся с рецептором вирус попадает в клетку посредством эндоцитоза (Рисунок 3). Низкий pH в эндосоме вызывает конформационное изменение в молекуле гемагглютинина, что приводит к слиянию вирусной мембраны с эндосомальной и образованию поры, через которую рибонуклеопротеидный комплекс выходит в цитоплазму клетки хозяина [33]. Параллельно ионный канал M2 накачивает ионы водорода из полости эндосомы внутрь капсида – закисление среды облегчает разворачивание генетического материала вируса и его выход в цитоплазму. Далее RNP транспортируется в ядро с помощью NLS (сигнал ядерной локализации) его белков [34, 35]. РНК-

зависимая РНК-полимераза использует вирусную антисмысловую РНК для синтеза матричных РНК и комплементарной РНК, с которых транскрибируется множество копий антисмысловой РНК для новых вирионов. Матричные РНК экспортируются из ядра с помощью вирусных белков M1 и NEP/NS2 для синтеза белков в цитоплазме. Белки вирусной мембраны (гемагглютинин, нейраминидаза, ионный канал M2) синтезируются в шероховатом ЭПР и претерпевают посттрансляционную модификацию (гликозилирование) в комплексе Гольджи, что служит сигналом для транспорта в апикальную клеточную мембрану. Далее происходит образование новых вирусных частиц и упаковка генетического материала. «Отпочковывание» вирионов с поверхности клетки инициируется накоплением матриксного белка M1 под мембраной.

Нейраминидаза играет ключевую роль при высвобождении новообразованных вирионов с поверхности клетки. Гемагглютинин, цепляясь за сиалосодержащие рецепторы, удерживает вирусные частицы в слое поверхностных углеводов до тех пор, пока нейраминидаза не отщепит сиаловую кислоту. Сиалидазная активность нейраминидазы способствует распространению вируса по дыхательным путям и заражению соседних клеток [27, 33, 35, 36]. Также нейраминидаза отрезает сиаловую кислоту в составе посттрансляционных модификаций поверхностных белков самого вируса гриппа, что препятствует взаимной агрегации вирусных частиц. Хотя большинство исследований и обзоров литературы сосредоточены, главным образом, на роли нейраминидазы при выходе вируса из клетки, с годами стало ясно, что она необходима также и на начальной стадии инфекции [37, 54]. Для успешного размножения вируса гриппа чрезвычайно важен баланс функциональной активности гемагглютинина и нейраминидазы: если баланс смещен в сторону НА, вирус застрянет в поверхностном слое слизи, так и не добравшись до олигосахаридных рецепторов на эпителиальных клетках; если

же в сторону NA – вирус отщепит слишком много остатков сиаловой кислоты и не сможет проникнуть в клетку. HA и NA обладают, таким образом, комплементарными активностями и должны эволюционировать вместе [38, 39, 40, 41].

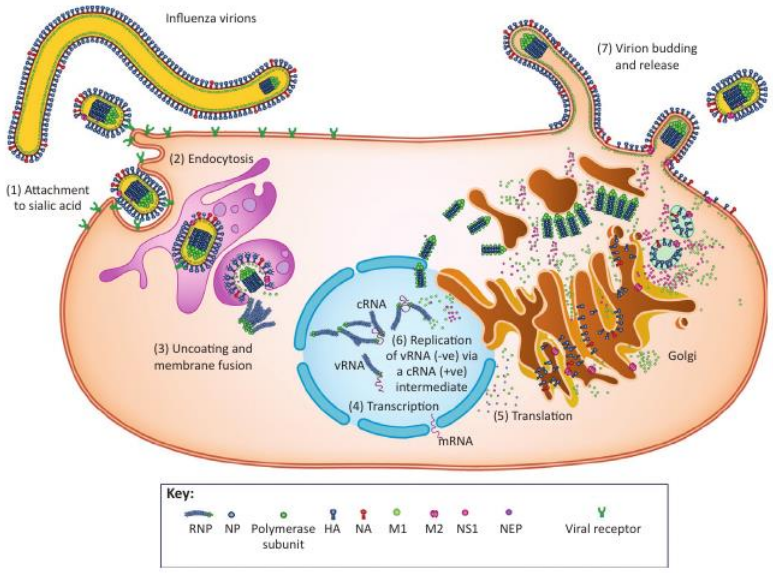

Рисунок 3. Цикл репликации вируса гриппа. См. пояснения в тексте. Рисунок взят из [42].

3.2.2. Структура и каталитический механизм

Нейраминидаза – поверхностный гликопротеин вируса гриппа, который выглядит как «гриб на ножке». Нейраминидаза состоит из четырех одинаковых субъединиц; фермент функционален именно в тетрамерной форме [43]. Каждая субъединица имеет глобулярную часть (6-лопастной β-пропеллер) на длинной α-спиральной ножке, трансмембранный участок, заякоривающий белок в липидной мембране вируса, а также маленький цитоплазматический участок (Рисунок 4). Цитоплазматический участок почти на 100% консервативен у всех подтипов вируса гриппа А, что говорит о его функциональной значимости. Было показано, что мутации в этом участке / его делеция приводят к изменению морфологии вирионов, уменьшению их количества и задержке формирования

[44, 45, 46]. Предполагается, что такой эффект связан с потерей взаимодействия нейраминидазы образующейся вирусной частицы с матриксным белком M1 [47, 48]. «Грибная шляпка» образована четырьмя доменами с укладкой 6-лопастного β-пропеллера и нейраминидазной активностью (Рисунок 5А). «Шляпка» сидит на «ножке», образованной четырьмя длинными α-спиралями. Длина и последовательность «ножки» сильно варьируют у разных подтипов вируса, однако в каждой из четырех спиралей содержится, по крайней мере, один остаток цистеина. Цистеины на соседних спиралях образуют дисульфидные связи, которые, как принято считать, стабилизируют тетрамер [49]. Было показано, что делеции в «ножке» приводят к увеличению вирулентности вируса за счет уменьшения нейраминидазной активности. Такой эффект обусловлен влиянием «ножки» на подвижность петель активного центра нейраминидазы [50].

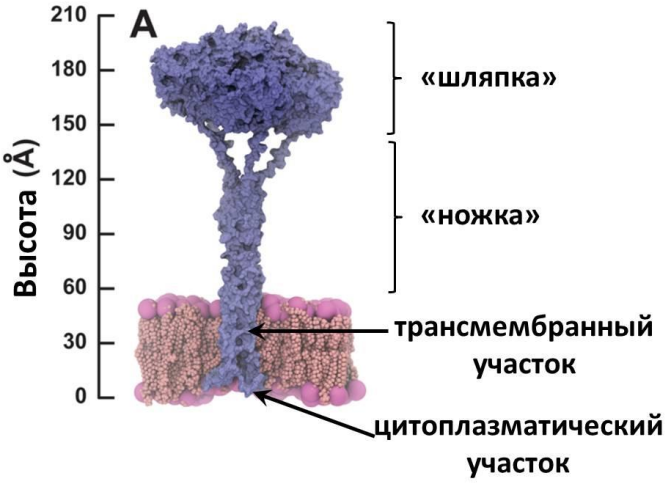

Рисунок 4. Доменная организация нейраминидазы вируса гриппа А. Рисунок взят из [50].

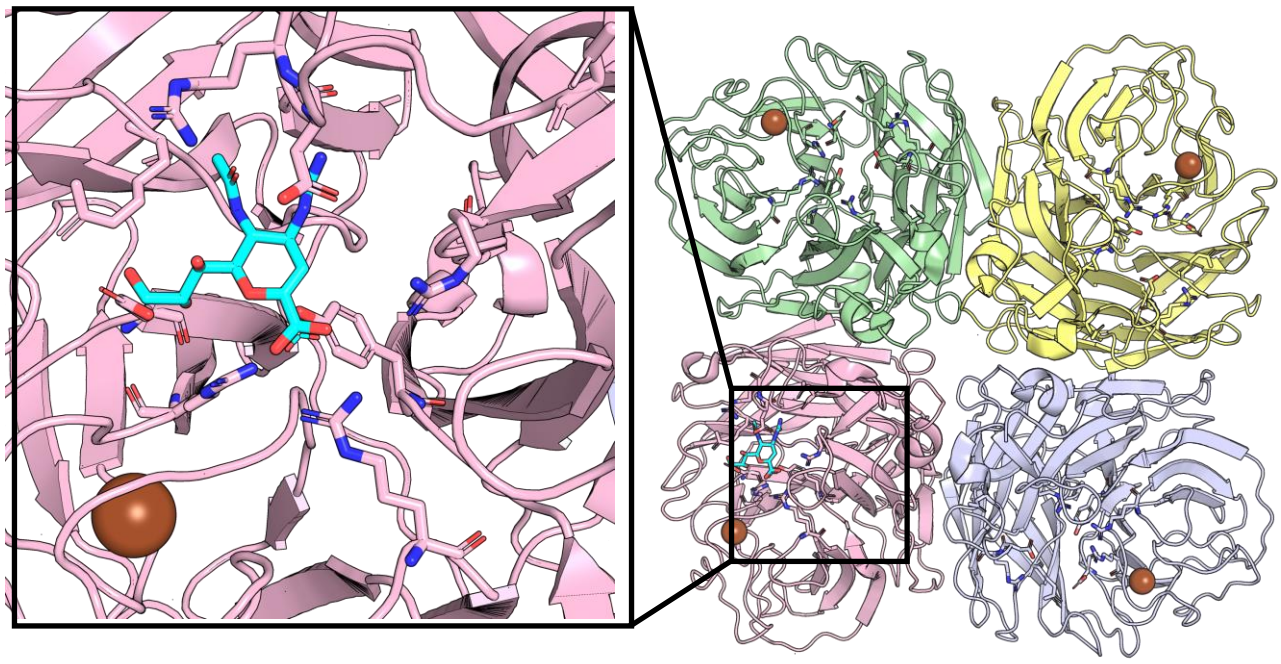

Рисунок 5. (А) Структура активного центра каталитического домена нейраминидазы вируса гриппа (на примере PDB ID 2hty; координаты ингибитора занамивира (ярко-голубой), имитирующего переходное состояние природного субстрата, взяты из кристаллографической структуры 3b7e и наложены на 2hty для наглядности); (Б) Кристаллографическая структура тетрамера каталитических доменов нейраминидазы вируса гриппа (на примере PDB ID 2hty). Четыре домена с β-пропеллерной укладкой покрашены разными цветами. Палочками изображены ключевые остатки активных центров. Ионы кальция показаны коричневыми сферами.

Нейраминидаза содержит несколько сайтов гликозилирования [51], а также сайт связывания иона кальция, который необходим для стабильности и активности фермента. Активный центр нейраминидазы вируса гриппа расположен в одной из воронок в центре пропеллера и содержит много заряженных аминокислот (Рисунок 5). Остатки Arg118, Asp151, Arg152, Arg224, Glu276, Arg292, Arg371 и Tyr406 непосредственно контактируют с субстратом, образуя сеть водородных связей и солевых мостиков, и консервативны у большинства вирусов типов A и B. Остатки Arg118, Arg292 и Arg371 формируют т.н. «триаргининовый кластер», характерный для большинства сиалидаз и играющий ключевую роль в связывании сиаловой кислоты. Три аргинина образуют сильные ионные взаимодействия (солевые мостики) с

отрицательно заряженной карбоксильной группой субстрата. Центральный Arg371 кластера взаимодействует с карбоксильной группой сразу двумя атомами азота, образуя «вилочковый солевой мостик». Остатки Trp178 и Ile222 образуют гидрофобные контакты с метильной группой атома C-5 сиаловой кислоты, а алифатическая часть Arg224 – с глицериновым фрагментом субстрата [52]. Также высоко консервативны остатки Glu119, Arg156, Ser179, Asp198, Glu227, Glu277, Asn294 и Glu425, которые не взаимодействуют с субстратом напрямую, но важны для поддержания структуры активного центра [53, 54].

В предполагаемом каталитическом механизме нейраминидазы концевой остаток сиаловой кислоты олигосахаридного рецептора связывается в активном центре, при этом кольцо сиаловой кислоты изменяет конформацию с кресла на α-лодку из-за взаимодействия ее карбоксильной группы с триаргининовым кластером. Далее снова происходит изменение конформации сиалового кольца, которое становится плоским у атома углерода C2. Это переходное состояние (сиалозил-катион) координируется взаимодействиями с Tyr406 и Asp151. Затем происходит нуклеофильная атака по C2-атому каталитическим остатком Tyr406 (здесь Glu277 выступает в роли общего основания), что приводит к разрыву α-кетозидной связи, соединявшей сиаловую кислоту с предыдущим по цепи сахаром и образованию ковалентного интермедиата (Рисунок 6). Десиалилированный рецептор уходит из активного центра. На этой стадии реакции сиаловая кислота в конформации кресла ковалентно связана с белком через остаток тирозина. Гидролиз этого интермедиата (нуклеофильная атака молекулой воды) происходит через стадию образования сиалозил-катиона с последующей диффузией α-сиаловой кислоты в раствор [52, 53, 55].

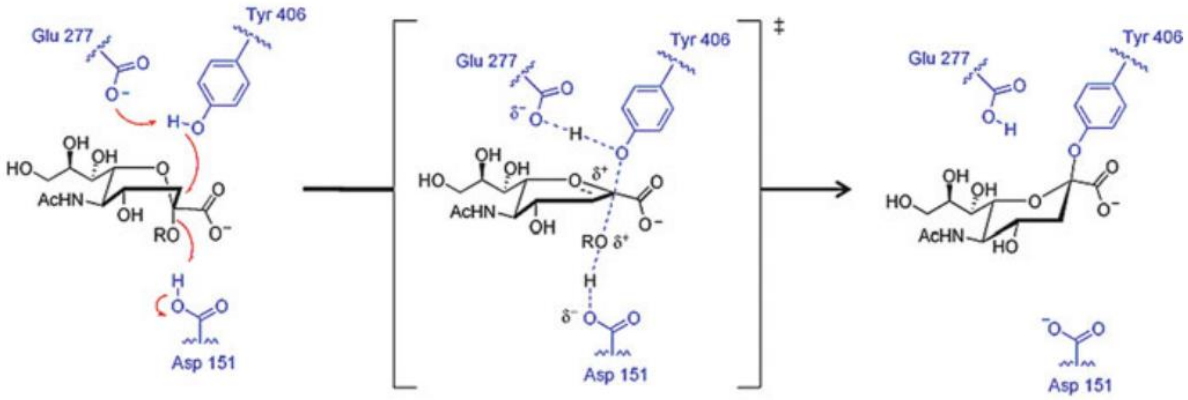

Рисунок 6. Предполагаемый каталитический механизм нейраминидазы. В квадратных скобках – переходное состояние реакции – сиалозил-катион. Справа – ковалентный интермедиат реакции. Рисунок взят из [55].

Помимо участка в активном центре, где связывается сиаловая кислота и происходит ферментативный катализ, в структуре нейраминидазы вируса гриппа А принято выделять еще три топологически независимых участка связывания. Это «полость-150» и «полость-430», обозначаемые так в соответствии с нумерацией петель, образующих эти участки, а также так называемый «вторичный участок» связывания сиаловой кислоты, получивший свое имя из-за способности взаимодействовать с концевой частью природного субстрата нейраминидазы (Рисунок 7).

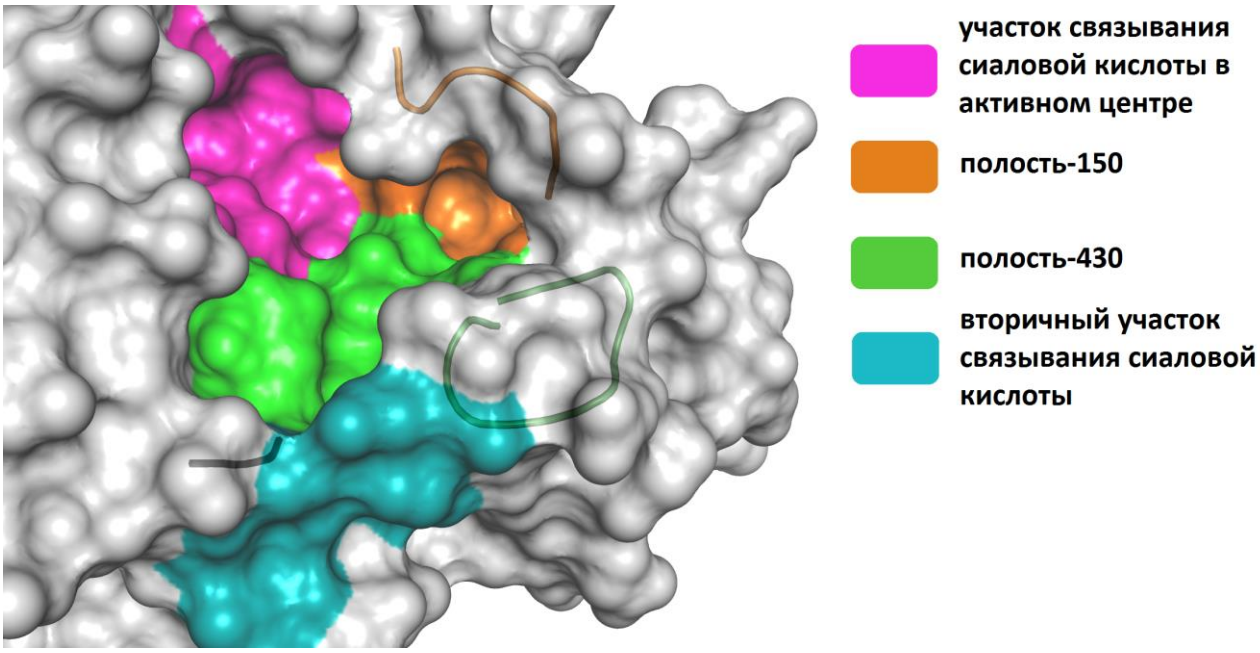

Рисунок 7. Схема расположения различных участков связывания в структуре нейраминидазы вируса гриппа N1 (на примере PDB ID 3beq). Оранжевой «проволкой» показана петля-150, зеленой – петля-430, черной – петля-370.

Полость-150

Нейраминидазы N1 и N2, циркулирующие в человеческой популяции в настоящий момент, принадлежат к двум разным филогенетическим группам. Первая группа включает подтипы N1, N4, N5 и N8, вторая – N2, N3, N6, N7 и N9. Лекарства, применяющиеся сегодня для лечения гриппа, создавались исходя из доступных на тот момент кристаллографических структур нейраминидаз N2 и N9 [56, 57]. Кристаллизация нейраминидаз из первой филогенетической группы показала, что структуры активных центров ферментов двух групп идентичны, за исключением области петли, образованной остатками 147-152. В кристаллографических структурах апо-форм нейраминидаз первой группы эта петля-150 ориентирована наружу от активного центра и образует дополнительную полость (полость-150). В присутствии лиганда в активном центре петля-150 взаимодействует с ним и принимает конформацию,

аналогичную той, что наблюдается у ферментов второй группы, с лигандом или без. Отсюда наиболее вероятно следует, что лиганд (субстрат / ингибитор) связывается с нейраминидазой в открытой конформации петли-150, после чего происходит медленная конформационная перестройка – петля переходит в закрытую конформацию, образуя взаимодействия с молекулой лиганда. Было предложено, таким образом, использовать полость-150 для дизайна новых ингибиторов, селективных к нейраминидазам первой группы [58]. С помощью молекулярно-динамических исследований, однако, была показана подвижность петли-150 у представителей второй филогенетической группы, которые, следовательно, тоже могут иметь полость-150, хотя популяционное распределение и смещено таким образом, что открытая петля-150 не наблюдается в кристаллах [59, 60, 61].

Полость-430

Молекулярно-динамические симуляции апофермента нейраминидазы птичьего гриппа H5N1 показали, что движение петли-150 часто сопровождается движением находящейся поблизости петли-430, что приводит к значительному расширению воронки активного центра. Авторы высказали предположение, что наблюдаемые изменения в структуре фермента можно использовать для создания новых ингибиторов нейраминидазы [62]. Существование нового участка связывания под петлей-430 было предсказано и другими компьютерными и экспериментальными методами [63, 64, 65]. На сегодняшний день полость-430 изучена достаточно поверхностно, ее функциональная роль в структуре нейраминидазы неизвестна, и иногда этот участок путают с т.н. «вторичным сайтом» связывания сиаловой кислоты [65, 66]. Предполагается, что полость-430 может являться одним из промежуточных мест связывания субстрата / ингибитора на пути в каталитический центр [66, 67].

«Вторичный участок» связывания сиаловой кислоты

Так называемый «вторичный участок» связывания сиаловой кислоты был впервые обнаружен в 1984 году, когда в молекулах нейраминидазы N9 обнаружили гемагглютинирующую активность [68]. Позднее была получена кристаллографическая структура нейраминидазы N9 в комплексе с сиаловой кислотой, при этом температурные условия кристаллизации отличались от использовавшихся ранее [69]. Анализ полученной структуры показал, что субстрат связался не только в участке связывания сиаловой кислоты активного центра, но еще и во «вторичном участке связывания». Этот участок находится по другую сторону от петель -371 и -430 и полости-430, которую эти петли образуют. Сиаловая кислота во «вторичном участке» находится в конформации «кресло» и образует взаимодействия с остатками Ser367, Ser370, Ser372, Asn400, Trp403 и Lys432. Сравнив известные на тот момент аминокислотные последовательности нейраминидаз вируса гриппа А, авторы пришли к выводу, что остатки, образующие «вторичный участок» связывания сиаловой кислоты, консервативны в основном у изолятов, полученных из птиц. Существование этого участка долгое время ставилось под сомнение, пока не было доказано с помощью компьютерных и экспериментальных методов, которые также показали его наличие и у некоторых изолятов из свиньи и человека [70, 71]. Существует гипотеза, согласно которой «вторичный участок» повышает каталитическую эффективность фермента за счет связывания мультивалентных субстратов и их физического удерживания вблизи активного центра. Предполагается также, что аминокислотные замены в этом участке одновременно с изменением субстратной специфичности гемагглютинина являются причиной возникновения пандемических штаммов вируса гриппа, что объясняется тем, что «вторичный участок» связывания сиаловой кислоты перекрывается с одним из антигенных сайтов нейраминидазы [72, 73].

3.3. Бактериальные нейраминидазы из *Streptococcus pneumoniae*

3.3.1. Патогенез *Streptococcus pneumoniae*: колонизация и инвазия

Streptococcus pneumoniae, или пневмококк, – грам-положительная бактерия, распространенный представитель микрофлоры слизистой оболочки верхних дыхательных путей человека. Приблизительно 20-50% детей и 5-20% взрослых в странах с высоким уровнем жизни являются носителями *S. pneumoniae*, в менее развитых странах эти показатели достигают 90% и 50% среди детей и взрослых, соответственно [74]. Колонизация дыхательных путей пневмококком протекает асимптоматически у большинства здоровых людей вплоть до постепенной элиминации бактерий иммунной системой.

Колонизация подразумевает прикрепление пневмококка к невоспаленному эпителию верхних дыхательных путей посредством связывания с олигосахаридными рецепторами на поверхности эпителиальных клеток и образование биопленки, которая может сохраняться недели и даже месяцы. Первое, с чем сталкивается пневмококк при попадании в носовую полость – это слизистый слой, покрывающий эпителий – гликокаликс. Внешняя полисахаридная капсула пневмококка – один из ключевых факторов его вирулентности, позволяющий избежать фагоцитоза и пробраться к эпителию сквозь слой слизи. Почти все пневмококковые капсулы отрицательно заряжены, что способствует отталкиванию от сиалозидов, в большом количестве присутствующих в слизи [167]. Как только бактерия преодолела слой слизи, толстая капсула становится не нужной, поскольку препятствует адгезии к эпителиальной поверхности [75].

Адгезия осуществляется посредством взаимодействия множества различных белков на поверхности бактерии с соответствующими рецепторами на апикальной поверхности эпителиальных клеток. В литературных источниках

представлено много информации о механизмах адгезии пневмококка, что говорит об интересе научного сообщества к этому процессу. Эти механизмы, однако, изучены плохо, и, в большинстве случаев, информация неполная и противоречивая. Существует путаница относительно того, какие именно белки пневмококка участвуют в адгезии в разных тканях и органах, какие ответственны за адгезию на здоровом и воспаленном эпителии, с какими рецепторами взаимодействуют эти белки в каждом конкретном случае. На стадии начального прикрепления бактерии к здоровому, невоспаленному эпителию верхних дыхательных путей, вероятно, основную роль играют поверхностный адгезин А (PsaA), белок RrgA на концах пилей, а также поверхностные гликозилгидролазы, в том числе и нейраминидаза А (NanA) [76, 77]. Первые прикрепившиеся бактериальные клетки делятся и формируют клональные однослойные микроколонии на поверхности эпителия. Клетки в микроколониях выделяют особый пептид-феромон (CSP, *competence-stimulating peptide*), который привлекает свободных (планктонных) пневмококков – так микроколонии становятся многослойными. Это явление коммуникации между клетками называется «чувством кворума» (*англ. quorum sensing*) [78, 79, 80]. Бактерии в многослойной колонии начинают выделять в среду ДНК, нерастворимые полисахариды, белки, липиды, которые образуют слой внеклеточного матрикса [81, 82]. Пневмококки, окруженные матриксом, растут и развиваются, обмениваются сигналами и генетическим материалом [83, 84]. Для улучшения диффузии ДНК и нутриентов в слое матрикса образуются каналы. Поверхностные гликозил-гидролазы *S. pneumoniae* обеспечивают его моно-сахарами, использующимися в качестве источника питания. Внеклеточный матрикс в зрелой биопленке довольно плотный и структурированный [85]. Антитела, фагоциты не могут сквозь него проникнуть. Фагоциты, реагируя на постоянное присутствие патогена (планктонного,

рассеивающегося из биопленки) в организме, выделяют активные формы кислорода, которые инактивируются еще в наружных слоях биопленки, так и не проникнув глубже. Постоянное выделение активных форм кислорода, таким образом, никак не влияет на биопленку, но приводит к повреждениям тканей самого организма. Также внеклеточный матрикс может замедлять диффузию некоторых антибиотиков [86]. По крайней мере часть бактерий в биопленке находится в состоянии замедленного метаболизма, когда клетке не хватает питания и она не растет. Такие нерастущие клетки невосприимчивы к антимикробным средствам, и потому могут пережить самые суровые условия [87, 88].

Переход *S. pneumoniae* из «пассивного сожителя» в «активного агрессора» происходит при наличии местного воспаления, вызванного, например, вирусной инфекцией, снижением иммунитета, повреждением эпителиального слоя [89, 90]. Каскад воспалительных сигналов приводит к изменению типа и количества рецепторов на эпителиальной клетке. Усиливается экспрессия рецептора фактора активации тромбоцитов (PAFr), с которым должны связываться лейкоциты для того, чтобы проникнуть через эпителий; также экспрессируется больше полимерного иммуноглобулинового рецептора (PIGr). Связывание одного из основных компонентов клеточной стенки пневмококка фосфорилхолина (ChoP) с PAFr и холин-связывающего белка А (CbpA) с PIGr индуцирует эндоцитоз бактерии эпителиальной клеткой [88, 91, 92]. Содержимое эндосомы транспортируется на противоположную сторону, прилегающую к базальной мембране. Гликолитические ферменты енолаза и глицеральдегид-3-фосфат дегидрогеназа, а также холин-связывающий белок Е (CbpE), расположенные на поверхности *S. pneumoniae*, связывают и активируют плазминоген человека. Плазмин, разрушая белки внеклеточного матрикса, помогает патогену добраться до эндотелия кровеносных сосудов [93,

94]. Миграция сквозь эндотелий осуществляется так же, как и через эпителий. Живые бактерии, таким образом, попадают в кровоток и разносятся в другие ткани. Наряду с описанным механизмом проникновения в кровоток, пневмококк использует также такие белки как, например, пневмолизин (Ply), который полимеризуется в клеточной мембране в большую пору, пируват оксидазу (SpxB) и α-глицерофосфат оксидазу (GlpO), которые продуцируют пероксид, для непосредственного повреждения тканей человека [89, 167, 95].

И хотя такой переход в инвазивное состояние случается нечасто, очень широкая распространенность пневмококка делает его основным возбудителем многих инфекционных заболеваний человека. Попадание микроорганизма в нижние части дыхательных путей, которые в норме стерильны, приводит к стремительному развитию пневмонии (в т.ч. вторичной пневмонии, сопровождающей гриппову инфекцию); бактерия может также проникать в полость среднего уха, вызывая острый отит, в мозговые оболочки, вызывая гнойный пневмококковый менингит, а также в кровь, что приводит к общему сепсису. В более редких случаях может также вызывать артрит, плеврит, синусит, остеомиелит [89].

3.3.2. Нейраминидазы NanA, NanB и NanC из *Streptococcus pneumoniae*: структура и каталитический механизм

S. pneumoniae экспрессирует три сиалидазы (NanA, NanB и NanC), которые различаются структурно и функционально и присутствуют в 100%, 96% и 51% штаммов, соответственно [96]. Сиалидазы NanA, NanB и NanC имеют молекулярные массы 115, 78 и 82 кДа. Анализ их последовательностей показал, что все три сиалидазы имеют N-концевую сигнальную последовательность, которая служит сигналом экспорта белка из клетки и отрезается сигнальной пептидазой [97]. За сигнальной последовательностью

идет лектиновый домен (CBM – *carbohydrate binding module*), принадлежащий к семейству CBM40 по классификации CAZy (см. раздел 3.1.2.). Лектиновый домен связывает концевую часть олигосахаридного рецептора – сиаловую кислоту. За лектиновым доменом следует каталитический (GH33 по CAZy), образующий структуру 6-лопастного β-пропеллера с дополнительной небольшой вставкой – инсерционным (вставочным) доменом. Сиалидаза NanA, кроме того, имеет дополнительный С-концевой домен, содержащий аминокислотный мотив LPETG, заякоривающий белок в клеточной стенке (Рисунок 8) [98].

Все три сиалидазы секретируются клеткой наружу, но только NanA остается «заякоренной» в клеточной стенке. Ассоциированный с мембраной фермент сортаза (SrtA) разрезает полипептидную цепь NanA между треонином и глицином в мотиве LPETG и ковалентно сшивает треонин с пептидогликаном клеточной стенки [83].

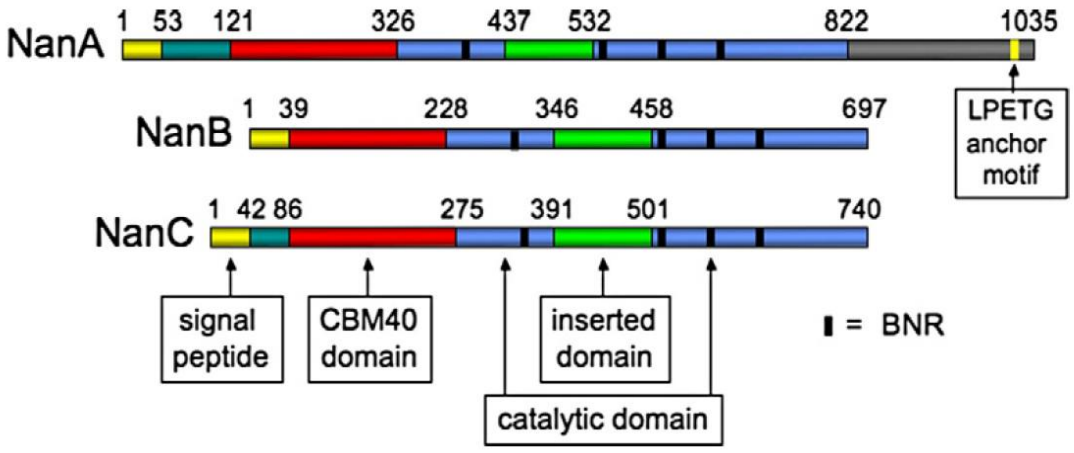

Рисунок 8. Диаграмма, отображающая разбивку последовательностей трех сиалидаз *S. pneumoniae* на домены. Желтым выделен сигнальный пептид, красным – лектиновый домен, голубым – каталитический со вставочным доменом (зеленый). Серым цветом обозначен С-концевой домен сиалидазы NanA. Рисунок взят из [98].

Попытки получить полную структуру NanA не увенчались успехом: белок не образовывал кристалл. Структуры отдельных доменов – каталитического (включая вставочный) [99, 100, 101] и лектинового [102, 103] были получены после применения ограниченного протеолиза либо путем экспрессии отдельных доменов с генетических конструктов. Взаимное расположение лектинового и каталитического доменов в NanA неизвестно, их взаимодействие не изучалось. Для NanB и NanC известны полные кристаллические структуры, включающие одновременно лектиновый и каталитический домен. NanB и NanC идентичны по последовательности на ~50%, их идентичность с NanA составляет ~20%. При этом структуры доменов очень схожи у трех сиалидаз пневмококка (см. Таблица 1 в секции Методы).

NanA – это типичная экзо-α-сиалидаза, катализирующая гидролитическое отщепление α-2,3-, α-2,6- и, по некоторым данным, α-2,8-присоединенной сиаловой (N-ацетилнейраминовой) кислоты с конца олигосахаридных субстратов с высвобождением α-сиаловой кислоты, которая в растворе превращается в более выгодный β-аномер. NanB – интрамолекулярная транс-сиалидаза, отщепляющая концевой остаток N-ацетилнейраминовой кислоты от α-2,3-сиалозидных субстратов с образованием 2,7-ангидрида этой кислоты. NanB может превращать и α-2,6- и α-2,8-субстраты, но в 5-10 раз хуже, чем α-2,3-сиалозиды [98]. NanC – это транс-сиалидаза, которая сперва отщепляет 2-дезокси-2,3-дидегидро-N-ацетилнейраминовую кислоту от α-2,3-сиалозидов, а затем медленно гидратирует полученное производное до N-ацетилнейраминовой кислоты (Рисунок 9) [104, 105].

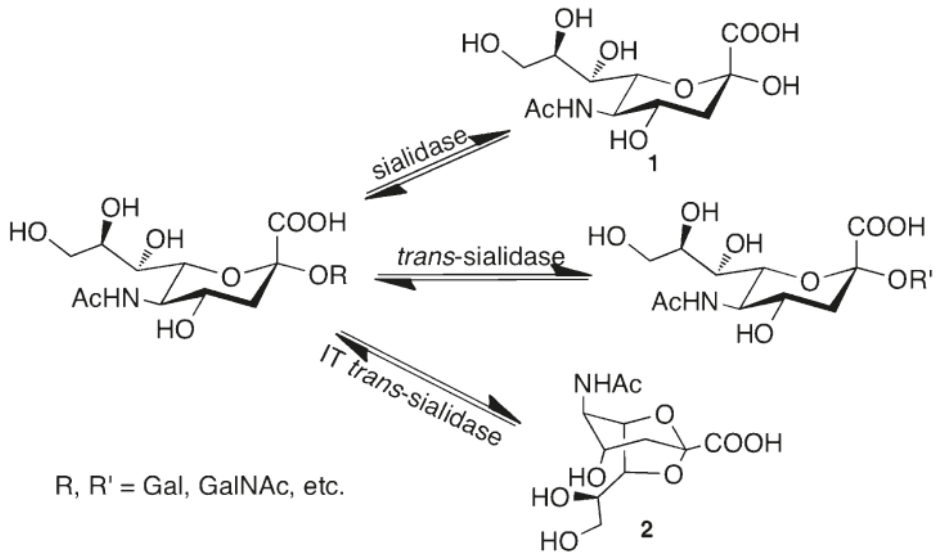

Рисунок 9. Реакции, катализируемые тремя сиалидазами *S. pneumoniae* (взято из [104]).

Полость активного центра в NanA просторная и широкая, что делает ее легко доступной для разных субстратов и молекул воды. В отличие от NanA, активный центр NanB и NanC представляет собой узкую щель из-за присутствия остатков триптофана и тирозина (Trp589 и Tyr674 в NanB; Trp 632 и Tyr 716 в NanC), образующих стэкинг-взаимодействие и ограничивающих вход в активный центр, что определяет разницу в субстратной специфичности пневмококковых сиалидаз (Рисунок 11). Каталитические остатки в активных центрах трех пневмококковых сиалидаз не отличаются от таковых в нейраминидазе вируса гриппа. Это триаргининовый кластер (Arg332, Arg706, Arg648), взаимодействующий с карбоксильной группой сиаловой кислоты, Tyr737/Glu632 нуклеофильная пара и кислота/основание Asp357 (нумерация по NanA, PDB 2YA8) (Рисунок 11). C5-ацетамидная группа субстрата также окружена гидрофобными остатками. Механизм катализа сиалидаз пневмококка и вируса гриппа одинаков вплоть до стадии второй нуклеофильной атаки (см. главу 3.2.2). В зависимости от того, что выступает в роли нуклеофила – молекула воды в гидролитической сиалидазе, O7-атом сиалозил-катиона в

интрамолекулярной транс-сиалидазе или другой углевод в транс-сиалидазе – продуктом реакции может быть N-ацетилнейраминовая кислота, ее 2,7-ангидрид или сиалил-β-галактоза в случаях NanA, NanB и NanC, соответственно [104, 106].

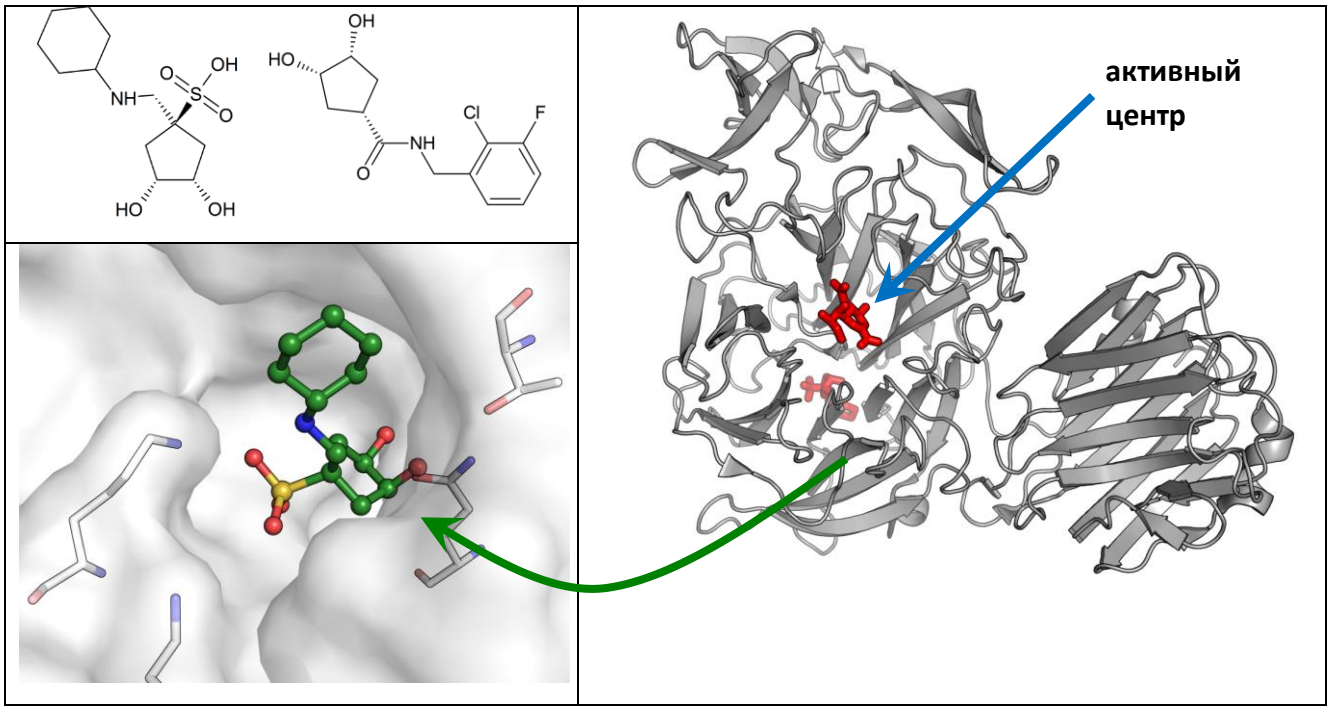

Рисунок 10. Аллостерический центр в структуре каталитического домена бактериальной нейраминидазы NanB (показан зеленой стрелкой). Приведены химические формулы предполагаемых ингибиторов – оптактина и оптактамида, а также показано связывание оптактина в аллостерическом центре (левый нижний угол). Активный центр находится на противоположной стороне каталитического домена NanB и показан синей стрелкой.

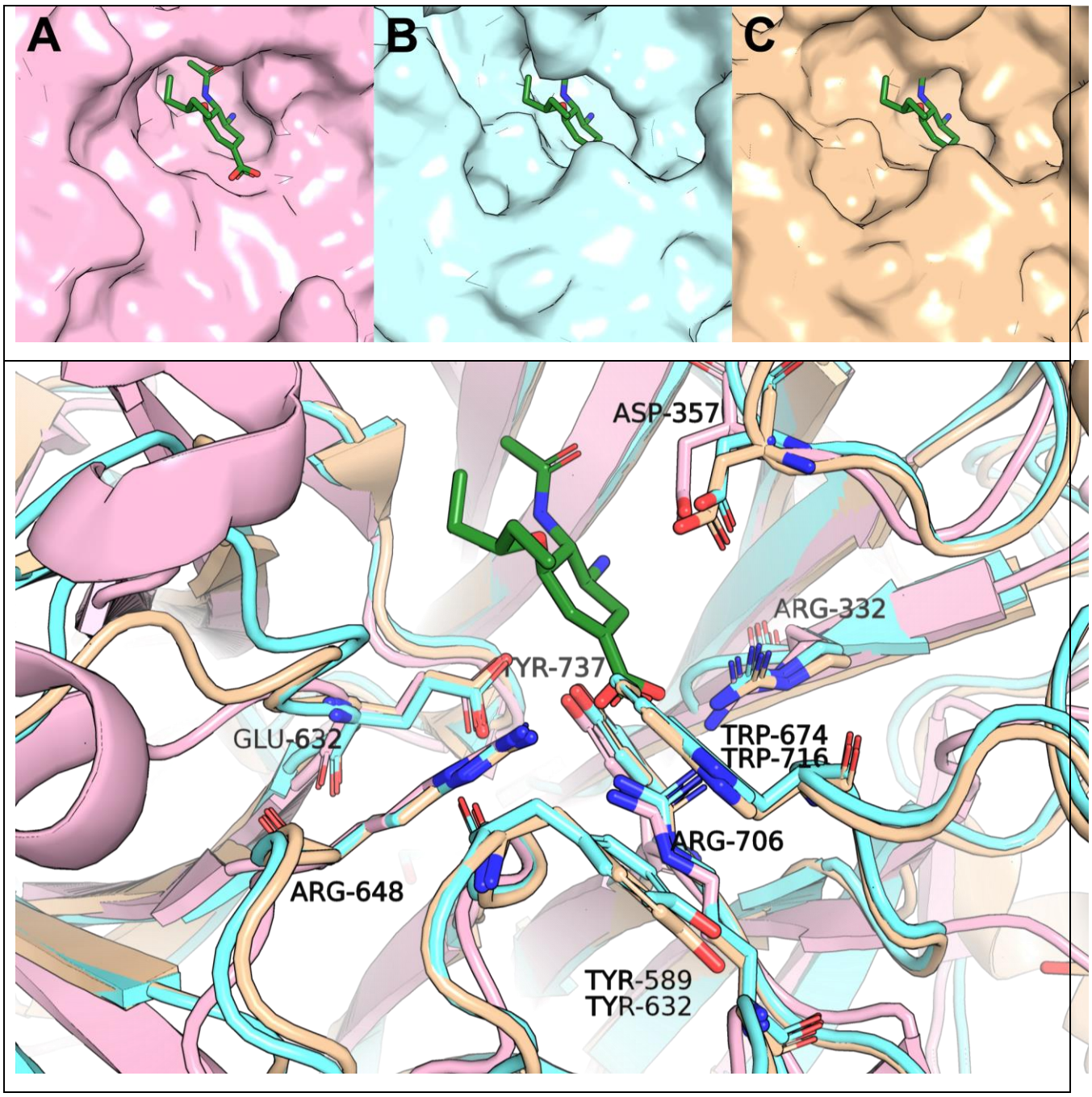

Рисунок 11. Сравнение структур активных центров сиалидаз NanA (розовая), NanB (голубая) и NanC (бежевая) пневмококка. Верхнее изображение: полости активного центра нейраминидаз A, B и C. Нижнее изображение: подписаны и выделены палочками каталитические остатки, а также триптофан и тирозин, ограничивающие вход в активный центр и определяющие различие в субстратной специфичности трех ферментов. Зеленым показана молекула озельтамивира.

В 2016 году в PDB появилась структура нейраминидазы NanB пневмококка в комплексе с низкомолекулярным соединением оптактином/оптактамидом, связанным с противоположной стороны от активного центра, в другой воронке, образованной β-пропеллером (Рисунок 10). Однако исследование о предполагаемом аллостерическом ингибиторе до сих пор не опубликовано (Rogers et al., 2016). В структурах остальных нейраминидаз пневмококка на сегодняшний день не обнаружено никаких других потенциальных участков связывания, кроме каталитических центров.

Все три сиалидазы пневмококка содержат лектиновый домен с N-конца. Лектиновый домен не имеет ферментативной активности. Он связывает концевой остаток олигосахаридного субстрата – сиаловую кислоту. Участок связывания субстрата в лектиновом домене включает остатки Phe167, Asp180, Tyr182, Thr193, Glu195, Arg197, Gln203, Asn209, Arg274 и Trp280 (нумерация по структуре PDB 4ZXK). Было показано, что присутствие лектинового домена в структурах бактериальных сиалидаз значительно повышает их каталитическую эффективность по отношению к большим поливалентным субстратам. Такой эффект объясняется, вероятно, одновременным взаимодействием центров связывания в лектиновом и каталитическом доменах с разными остатками сиаловой кислоты в составе поливалентного субстрата [107]. Это помогает ферменту найти богатые сиаловой кислотой участки поверхности дыхательного тракта и привести каталитический домен в непосредственную близость к его субстрату [108].

3.3.3. Роль NanA, NanB и NanC в патогенезе пневмококка и формировании биопленок, синергия с нейраминидазой вируса гриппа

То, что сиалидазная активность необходима для жизнедеятельности *S. pneumoniae*, является общепризнанным фактом. Конкретные механизмы участия сиалидаз NanA, NanB и NanC в патогенезе пневмококка, однако, изучены плохо. Считается, что одна из них – NanA – играет наиболее важную роль; NanB, возможно, нужна для питания бактерии, а также, наряду с NanA, помогает бактерии не застрять в слое слизи [109]; единственное спекулятивное предположение относительно функции NanC – это регуляция активности NanA [104]. В эксперименте на мышах было показано, что NanA необходима для колонизации носоглотки, распространения бактерии из носоглотки в нижние дыхательные пути и начальной колонизации легких, однако не было выявлено роли NanA в распространении/развитии инфекции в других тканях организма [110]. При изучении колонизации дефектным по NanA пневмококком полостей носоглотки и среднего уха шиншиллы было показано, что уже в течение 24 часов после заражения наблюдается значительное уменьшение количества дефектных по NanA бактерий в сравнении с диким типом. При этом полная элиминация мутанта в среднем ухе происходит к десятому дню, а в носоглотке – к концу второй недели; дикий тип сохраняется в два раза дольше. Тот факт, что мутант по NanA все же способен выживать какое-то время, может объясняться некоторой компенсацией нейраминидазной активности сиалидазой NanB [111]. В экспериментах на мышах с использованием пневмококка с нефункциональными нейраминидазами было показано, что и NanA, и NanB важны для выживания микроорганизма, причем не только в дыхательном тракте, но и в кровотоке. *S. pneumoniae*, дефектный по NanA, в течение 12 часов элиминировался из дыхательных путей иммунной системой мышей, в то время как мутант по NanB сохранялся двое суток в легких и неделю в верхних

дыхательных путях, хотя и не мог размножаться. Оба мутанта были неспособны вызвать сепсис вслед за интраназальным введением патогена; при введении внутривенно оба мутанта элиминировались в течение двух суток [112]. Известно, что NanA способна отрезать сиаловую кислоту с некоторых белков иммунной системы, таких как лактоферрин, секреторный компонет и иммуноглобулины, что помогает бактерии избежать фагоцитоза, следующего за опсонизацией [113]. Это, возможно, объясняет неспособность дефектных по нейраминидазам А и В бактерий выжить в крови в исследовании Manco et al, 2006 [112].

NanA отрезает сиаловую кислоту с концов олигосахаридов на эпителиальных клетках, открывая доступ к следующим по цепи сахарам, которые могут использоваться бактерией как рецепторы для адгезии. Пневмококк, дефектный по NanA, обладает значительно худшей способностью к адгезии, чем дикий тип. При этом предварительная обработка клеток очищенной нейраминидазой может восстановить адгезию до исходного уровня [114, 115, 116]. Последовательное дегликозилирование рецепторов организма-хозяина поверхностными гликозил-гидролазами пневмококка (NanA, BgaA – β-галактозидаза, StrH – N-ацетилглюкозаминидаза) не только помогает «обмануть» иммунную систему и открывает рецепторы для адгезии, но и снабжает бактерию источником питания в виде отщепленных моно-сахаров, что приводит к усилению роста патогена [117, 118, 119]. Отщепление нейраминидазой А сиаловой кислоты уменьшает электростатическое отталкивание внешней капсулы бактерии от слизистого слоя; вместе с общим уменьшением вязкости слизи это облегчает пневмококку задачу добраться до эпителиальной поверхности [120].

Особый интерес научного сообщества вызывает механизм формирования пневмококком биопленки. Дефектный по NanA *S. pneumoniae* оказался неспособным образовать зрелую биопленку, хотя наблюдался начальный слой прикрепившихся бактерий. Это позволило предположить, что для дальнейшего прикрепления планктонных бактерий и эффективного образования толстых, многослойных микроколоний внутри биопленки требуется NanA [121]. Хотя сущность этого явления все еще плохо изучена, были предложены две общие стратегии участия NanA в формировании биопленки – каталитическая и некаталитическая. Многочисленные исследования показали, что ферментативная активность NanA важна для начальной адгезии бактерий к эпителиальным клеткам [114, 115, 116, 118]. Свободная сиаловая кислота – продукт каталитического превращения олигосахаридных рецепторов в активном центре NanA, служит сигналом к образованию биопленки. Добавление сиаловой кислоты значительно увеличивало количество пневмококков, в том числе и прикрепившихся к эпителию. Увеличение общего количества бактериальных клеток сопровождалось также их распространением в легкие мышей. Сиаловая кислота приводит к усилению экспрессии нейраминидазы NanA и, в меньшей степени, NanB, но никак не влияет на уровень экспрессии NanC [122]. В общем и целом, экспрессия NanA в пневмококках, растущих в условиях биопленки, увеличена по сравнению с планктонным пневмококком [123]. Кроме того, совместная ферментативная активность NanA и β-галактозидазы A высвобождает галактозу, которая через реакции метаболизма пирувата и ацетил-КоА также способствует образованию биопленки [124].

Последние исследования выявили альтернативную роль NanA в формировании биопленок, которая не зависит от ее каталитической функции. Экспрессия пневмококковой NanA в *Lactococcus lactis* – непатогенной бактерии,

которая используется в производстве молочных продуктов и не образует биопленку, способствовала образованию плотного слоя клеток, организованных в микроколонии. Добавление ингибитора каталитической активности NanA в среду не повлияло на наблюдаемую сложную многоклеточную структуру [116]. В аналогичном эксперименте образование биопленок было вызвано экспрессией NanA в *Streptococcus gordonii*, который лишен нейраминидазной активности. Было высказано предположение, что в условиях биопленки нейраминидаза А может действовать как адгезин [116]. В другом исследовании было продемонстрировано, что пневмококк, дефектный по NanA, не способен колонизировать носовую полость мышей, а также вызвать инфекцию среднего уха; коинфекции мышей вирусом гриппа А, который также обладает нейраминидазной активностью, оказалось недостаточно для того, чтобы полностью компенсировать отсутствие NanA у бактерий, что позволило предположить, что NanA участвует в формировании пневмококковой биопленки независимо от своей ферментативной активности [125]. В том же исследовании был проведен эксперимент по выращиванию пневмококковой биопленки на искусственной полистиреновой поверхности, в среде, в которой сиаловая кислота отсутствовала вовсе. Даже в таких условиях, когда ферментативная активность нейраминидазы никак не могла проявиться, было зафиксировано ухудшение жизнеспособности биопленки (т.е. количества живых клеток в ней) в 3,25 раза для дефектного по NanA штамма по сравнению с диким типом [125]. Как именно происходит участие NanA в образовании биопленки, независимое от ее ферментативной активности, пока не известно. Предполагается, что это явление может быть связано со способностью к адгезии лектинового домена. Другая поверхностная гликозил-гидролаза пневмококка – BgaA (β-галактозидаза А) – может действовать как адгезин, независимо от ее

ферментативной активности. Рецептор, с которым связывается BgaA, пока не был выявлен [126].

NanA играет важную роль не только при асимптоматической колонизации верхних дыхательных путей, но и в инвазивных пневмококковых инфекциях. Известно, например, что, проникнув через эпителиальный слой в кровоток, микроорганизм может попасть в мозг человека. В исследовании на мышах было продемонстрировано, что нейраминидаза NanA является необходимым и достаточным фактором в преодолении пневмококком гемато-энцефалического барьера. При этом ключевую роль в адгезии к эндотелиальным клеткам (*англ. hBMECs, human brain microvascular endothelial cells*), их активации и инвазии играет лектиновый домен [127, 128]. Впрочем, процессу миграции *S. pneumoniae* из крови в нервную ткань способствует также и каталитическая активность: отрезая сиаловую кислоту с пептида LAP (*latency-associated peptide*), NanA активирует TGF-β-сигнальный путь. Это приводит к повреждению контактов между клетками hBMECs и способствует проникновению патогена в мозг [129, 130]. Было также обнаружено, что *S. pneumoniae* может проникать в мозг через обонятельные нервы и луковицы, минуя кровоток, а индуктором этого процесса служит свободная сиаловая кислота [131]. В недавнем исследовании Tseng et al. было показано, что каталитическая активность NanA нарушает взаимодействие Siglec-5 с TLR-2-рецепторами, что приводит к активации инфламмасом и каспаз и вызывает чрезмерно сильную воспалительную реакцию в организме-хозяине и гибель зараженных клеток [132].

Поскольку пневмония, отит, менингит зачастую являются осложнениями после перенесенного гриппа, и, поскольку нейраминидаза является одним из основных факторов вирулентности обоих патогенов, активно изучается взаимосвязь бактериальной и вирусной инфекции и, в частности,

взаимодействие бактериальной и вирусной сиалидаз [125, 133, 134, 135, 136, 137, 138, 139, 140]. В экспериментах *in vivo* было показано, что животные, предварительно зараженные вирусом гриппа, в большей степени страдают от пневмококковой инфекции: патоген активнее колонизирует носоглотку, образуя биопленку, а затем распространяется и в легкие. Такой эффект обеспечивается общим повреждением эпителия, нарушением секреции слизи и биения эпителиальных ресничек и воспалением, вызванным вирусом; кроме того, вирусная нейраминидаза, отрезая сиаловую кислоту, экспонирует рецепторы бактериальной адгезии и стимулирует бактериальную пролиферацию, а также снабжает пневмококка источником питания [141, 142, 143, 144]. Организм человека реагирует на нейраминидазу вируса гриппа активацией TGF-β-сигнального пути и повышенной секрецией таких клеточных белков, как фибронектин и α5-интегрин, что приводит к усилению адгезии пневмококка к эпителию легких [145].

Несмотря на довольно интенсивное и разностороннее исследование влияния вирусной нейраминидазы на последующую бактериальную инфекцию, роль бактериального фермента в этом синергизме практически не изучалась. Wren et al. показали, что активность вирусной нейраминидазы неспособна полностью заменить бактериальную NanA в NanA-дефектной бактерии, что, по предположению авторов, связано с пока неизвестной неэнзиматической функцией NanA [125]. Был также проведен *in vitro* эксперимент, в котором новообразованные вирусные частицы агрегировали на поверхности клеток в присутствии сильного ингибитора вирусной нейраминидазы – занамивира. Однако при добавлении в систему экзогенной нейраминидазы из *S. pneumoniae* агрегаты вирионов практически полностью расцеплялись, и возобновлялся перенос вирусной инфекции к соседним клеткам. Бактериальная нейраминидаза, таким образом, «спасала» вирус от действия ингибитора [146].

3.4. Нейраминидазы - перспективные мишени для создания лекарств от заболеваний дыхательных путей человека

3.4.1. Современное состояние исследований и доступные лекарственные препараты: нейраминидаза вируса гриппа

Анализ структуры активного центра и каталитического механизма нейраминидазы вируса гриппа А позволил создать несколько ингибиторов (Рисунок 12), два из которых коммерчески доступны в России и используются для лечения гриппа – занамивир и озельтамивир (Рисунок 13) [147, 148]. Их константы ингибирования равны 0,6 и 3,9 нМ, соответственно [149]. Озельтамивир синтезируется в виде предшественника – пролекарства, у которого к карбоксильной группе присоединена этильная или фосфатная, которые отщепляются печеночными эстеразами при приеме внутрь. В случае с занамивиром в защите карбоксильной группы нет необходимости, так как препарат Реленза выпускается в виде ингалятора, и активное вещество поступает непосредственно в полость верхних дыхательных путей [150]. Препарат Rapiacta (он же Rapivab, Peramiflu) (действующее вещество – перамивир) применялся лишь в случаях крайней необходимости, когда остальные лекарства не действовали [151]. С 2014 года коммерчески доступен в США, Южной Корее и Японии. Еще один ингибитор нейраминидазы – ланинамивир – выпускается в форме препарата Inavir, но используется пока только в Японии [152].

Все вышеперечисленные лекарственные препараты структурно похожи на переходное состояние природного субстрата, они связываются в активном центре и ингибируют сиалидазную активность. Однако вирус постоянно мутирует, и все чаще появляются штаммы, устойчивые к действию озельтамивира; также были описаны случаи появления резистентности к занамивиру и перамивиру [153, 154, 155, 156].

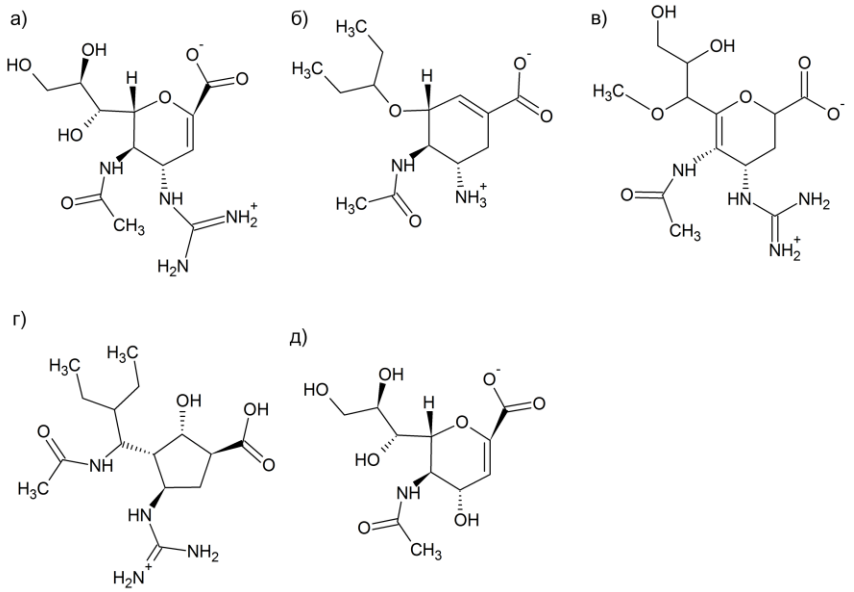

Рисунок 12. Известные ингибиторы нейраминидазы вируса гриппа, имитирующие переходное состояние субстрата в активном центре: а) занамивир; б) озельтамивира карбоксилат; в) ланинамивир; г) перамивир; д) DANA – ненасыщенное производное сиаловой кислоты.

Предпринимаются попытки использовать другие центры связывания в структуре вирусной нейраминидазы для создания новых ингибиторов, которые либо полностью связывались бы в полости-150/430, либо связывались в активном центре, захватывая некоторыми функциональными группами и полость-150/430 [157, 158]. Например, антоцианин, выделенный из ягод черной бузины и связывающийся в полости-430, показал при измерении ингибирующей активности IC_{50} 72μM, что на три порядка ниже IC_{50} занамивира, и это означает, что сам по себе антоциан – не очень сильный ингибитор нейраминидазы, однако его формулу можно пытаться улучшить [65]. Ацилгидразоны также связываются в полости-430 и ингибируют нейраминидазу в микромолярном диапазоне [159].

Рисунок 13. Противовирусные препараты, нацеленные на ингибирование нейраминидазы и успешно использующиеся в борьбе с вирусом гриппа: а – занамивир (Releza™), б – озельтамивир (Tamiflu™), в – ланинамивир (Inavir™),г – перамивир (Rapiacta™).

Также ведутся активные исследования по созданию ингибиторов других белков вируса гриппа, в числе которых протонный насос М2, вирусная полимераза, гемагглютинин, нуклеопротеин [160, 161]. Старые лекарства-ингибиторы протонного насоса М2 – Амантадин и Римантадин – потеряли эффективность в связи с широким развитием резистентности к ним.

Широкое распространение получили два типа вакцин против гриппа: инактивированные (IIV, *inactivated influenza vaccine*) и живые аттенуированные (LAIV, *live attenuated influenza vaccine*). Традиционно используются трехвалентные формы этих вакцин, т.е. содержащие вирусы A(H1N1), A(H3N2) и одну из двух линий вируса типа В. В некоторых странах стали доступны четырехвалентные вакцины, включающие также и вторую линию вируса типа В. Вакцина IIV одобрена для вакцинации людей в возрасте от 6 месяцев и вводится внутримышечно; LAIV вводится интраназально и одобрена только для

людей в возрасте от 2 до 49 лет, не имеющих хронических заболеваний. LAIV также не предназначена для беременных в связи с высоким риском развития осложнений. IIV и LAIV сделаны таким образом, чтобы вызвать иммунный ответ организма на молекулы гемагглютинина, который, однако, быстро мутирует [162]. Эффективность IIV по отношению к здоровым людям моложе 65 лет в случае, когда штамм в вакцине оказывается похож на циркулирующий штамм, составляет 60%. Эффективность в предотвращении заболевания среди людей старше 65 лет не очень велика, однако на 50-60% сокращает количество связанных с осложнениями госпитализаций и на 80% количество летальных исходов. Согласно некоторым исследованиям, среди детей эффективность вакцины LAIV выше, чем IIV, и достигает 87%. Для здоровых взрослых эффективность обеих вакцин приблизительно одинакова. Недавно были выделены антитела, блокирующие активный центр широкого спектра подтипов вирусной нейраминидазы; нейраминидаза мутирует медленнее гемагглютинина, поэтому эти антитела могут способствовать разработке новой вакцины в будущем [163, 164]. Стоит заметить, что существующие вакцины могут оказаться абсолютно не эффективными в отношении новых пандемических штаммов, возникших в результате антигенного шифта. На производство моновалентной пандемической вакцины может уйти несколько месяцев, поэтому необходимо иметь «запас» антивирусных препаратов на случай появления новых штаммов – пандемических, либо устойчивых к существующим лекарствам [165, http://www.euro.who.int/ru/health-topics/communicable-diseases/influenza].

3.4.2. Современное состояние исследований и доступные лекарственные препараты: бактериальные нейраминидазы из *Streptococcus pneumoniae*

По данным официальной статистики Министерства здравоохранения Российской Федерации в 2016 году общая заболеваемость пневмонией составила 678 425 случаев, в 2017 – 671 907 случаев; острым средним отитом – 1 260 826 случаев в 2016 году и 1 227 076 в 2017. В группе риска находятся дети и пожилые люди, люди, проживающие или работающие в больших коллективах, люди с ослабленным иммунитетом, в т.ч. ВИЧ-инфицированные, люди с хроническими заболеваниями сердца, почек, легких, а также курящие [166]. Повсеместная распространенность пневмококка и тяжесть вызываемых им заболеваний побуждает ученых и врачей по всему миру работать над созданием вакцин и антимикробных препаратов. На сегодняшний день для профилактики пневмококковых инфекций применяются две вакцины – поливалентная полисахаридная пневмококковая вакцина (PPSV23) и пневмококковая конъюгированная вакцина (PCV13). Эти вакцины представляют собой смеси капсульных полисахаридов из 23 и 13 серотипов бактерии, соответственно; PCV13 содержит также нетоксичный вариант дифтерийного токсина CRM197. Покрывающая бактерию полисахаридная капсула – важный фактор ее вирулентности, помогающий избежать фагоцитоза и пробраться к эпителию сквозь слой слизи [167, 168]. Подавляющее большинство клинических изолятов *S. pneumoniae* имеют капсулу. Более 80% здоровых взрослых приобретают антитела к пневмококку в течение 2-3 недель после введения PPSV23. Однако у детей, пожилых и людей с хроническими заболеваниями уровень иммуногенности PPSV23 очень низок. PCV13 была разработана с целью преодоления указанных недостатков PPSV23. Эффективность PCV13 против содержащихся в ней серотипов в рандомизированном двойном слепом плацебо-контролируемом исследовании

CAPiTA (~85 000 взрослых старше 65 лет) составила от 45 до 75% по разным показателям. Стоит заметить, что всего насчитывается более 90 серотипов пневмококка, и ни PPSV23, ни PCV13 не эффективны против остальных из 90 серотипов, что не входят в состав вакцин. Также, поскольку состав антигенов, входящих в вакцины, разрабатывается исходя из серотипов, доминирующих в развитых странах, это делает их менее эффективными в странах третьего мира, а высокая стоимость производства PCV13 снижает ее доступность. Более того, во время клинических испытаний наблюдалось замещение серотипов, включенных в вакцину, другими инвазивными серотипами. В долгосрочной перспективе это может привести к полной потере эффективности существующих вакцин из-за перераспределения серотипов в популяции [166, 169]. В связи с вышеозначенными недостатками полисахаридных вакцин прилагается много усилий к созданию вакцин на основе белков *S. pneumoniae*, как более универсальных мишеней среди разных серотипов. Наиболее привлекательными кандидатами пока выбраны пневмококковый поверхностный белок А (PspA), пневмолизин (Ply), PhtD, PhtE и холин-связывающий белок А (CbpA, также известный как PspC). Вакцины на основе этих белков находятся в стадии клинических испытаний [169, 170, 171].

Для лечения инфекций, вызванных *S. pneumoniae*, применяют различные антибиотики β-лактамного ряда, цефалоспорины, макролиды, фторхинолоны, тетрациклин, клиндамицин и др. Однако количество случаев заболеваний штаммами, устойчивыми к одному или нескольким антибиотикам, возрастает с устрашающей скоростью, а в связи с быстрым развитием болезни, времени на определение чувствительности к антибиотикам, как правило, нет, и пациент может погибнуть несмотря на прием стандартных лекарств [172, 173, 174, 175]. Поэтому активно ведется поиск новых антибиотиков, а также различных низкомолекулярных соединений, ингибирующих, например, такие процессы,

как образование биопленок пневмококком, прикрепление поверхностных белков к клеточной стенке, чувство кворума и т.д. [176, 177, 178, 179, 180]. Также, поскольку вирусная и бактериальная сиалидазы родственны и обладают некоторой степенью структурного сходства, и, поскольку вторичная бактериальная инфекция является главной причиной смертности при заболевании гриппом, создание ингибиторов двойного действия, т.е. подавляющих активность обеих сиалидаз, стало в последние годы популярной стратегией, чтобы попытаться разрушить «смертельный синергизм» вируса гриппа и пневмококка [181, 182, 183]. При этом в литературе не удалось найти исследований по поиску соединений, селективно ингибирующих именно бактериальную нейраминидазу NanA, что, вероятно, связано с отсутствием ее полной кристаллографической структуры, а также тем обстоятельством, что для NanA пока не было охарактеризовано иных участков связывания кроме того, где связывается субстрат.

4. МЕТОДЫ

4.1. Сбор первичной информации: структуры и последовательности [i]

Подструктуры NanA из *S. pneumoniae* представлены в базе данных PDB отдельными записями. Для дальнейшей работы были выбраны структуры 4ZXK (лектиновый домен) и 2YA8 (каталитический домен с инсерционным включительно). Два домена представлены в PDB тремя и одиннадцатью структурами и закристаллизованы с разными низкомолекулярными лигандами, но в остальном практически не отличаются. Наибольшее RMSD основных цепей выбранных структур в попарном наложении с остальными составляет 0,506 Å и 0,488 Å, для 4ZXK и 2YA8, соответственно (согласно алгоритму SSM [184]). Последовательность междоменного линкера (16 аминокислотных остатков), который соединяет две отдельные подструктуры в NanA и отсутствует в соответствующих PDB-записях, была взята из базы данных UniProtKB (P62575). PDB-записи 2JKB и 4YW2 были выбраны в качестве наиболее важных в этом исследовании гомологов NanA – NanB и NanC из *S. pneumoniae*. NanB и NanC представлены в PDB тридцатью семью и одиннадцатью структурами и закристаллизованы с разными низкомолекулярными лигандами, но в остальном практически не отличаются. Наибольшее RMSD основных цепей выбранных структур в попарном наложении с остальными составляет 0,611 Å и 0,618 Å, для 2JKB и 4YW2, соответственно. Структура нейраминидазы вируса гриппа H1N1 представлена в PDB двадцатью шестью записями. Для дальнейшей работы была выбрана запись 3B7E с лучшим на момент проведения исследования

[i] При подготовке данного раздела диссертации использована следующая публикация, выполненная автором в соавторстве, в которой, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: **Sharapova Y.**, Suplatov D., Švedas V. Neuraminidase A from Streptococcus pneumoniae has a modular organization of catalytic and lectin domains separated by a flexible linker //The FEBS journal. – 2018. – Т. 285. – №. 13. – С. 2428-2445; объем вклада автора диссертации составляет 75% и заключался в задумке, планировании и выполнении экспериментов, а также написании статьи.

разрешением 1,45 Å. Структуры остальных нейраминидаз/сиалидаз были выбраны с помощью биоинформатического анализа (см. раздел 4.3.3.1).

4.2. Биоинформатический анализ

4.2.1. Поиск гомологов и построение множественных выравниваний нейраминидаз из различных источников

Построение качественного выравнивания белков суперсемейства является сложной задачей из-за плохого сходства последовательностей удаленных гомологов. Для решения этой проблемы может быть использовано сопоставление 3D-структур, потому что трехмерная организация белков более консервативна в эволюции, чем аминокислотная последовательность. Несмотря на то, что объем базы данных белковых структур PDB все сильнее растет с каждым годом, в ней по-прежнему содержится не более 180 тысяч записей, в то время как общее число известных на сегодняшний день белков составляет около 168 миллионов. Необходимо совместное использование и последовательностей, и структур родственных белков для построения качественного выравнивания, содержащего всю известную информацию о вариабельности функционально разнообразных гомологов в рамках общей структурной организации. Для реализации этой технически сложной задачи использовался метод Mustguseal [185, 186]. Подробное описание применения подхода описано ниже.

4.2.1.1. Сравнительный анализ аминокислотных последовательностей

Выравнивание аминокислотных последовательностей использовали для сравнительного анализа эволюционно близких гомологов, которые сохранили достаточную степень сходства – в среднем, не менее 40% [187]. Для поиска эволюционно близких гомологов в базах данных Swiss-Prot и TrEMBL использовали GPU-реализацию программы BLAST [188]. Для фильтрации

собранной выборки с целью устранения избыточных записей при пороге 95-100% идентичности по последовательности, а также артефактов поиска (очень далеких белков со сходством <0,5-бит/колонка с репрезентативным белком) использовали пакет программ HHSearch [189]. Для выравнивания полученной выборки аминокислотных последовательностей использовали программу MAFFT [190].

4.2.1.2. Сравнительный анализ 3D-структур

Выравнивание трехмерных структур использовали для сравнительного анализа далеких гомологов, которые претерпели значительные функциональные изменения в процессе эволюции от общего предка и, как результат, потеряли сходство по аминокислотной последовательности, но сохранили достаточное сходство структурной организации. Для поиска эволюционно далеких гомологов в базе данных PDB использовали алгоритм superpose (PDBeFOLD) [184]. Выбор совпадений основывался на проценте эквивалентностей вторичной структуры (по умолчанию не менее 70-90% запроса («query») должно составлять не менее 70-90% цели («target»)) с последующей фильтрацией обнаруженных белков по сходству аминокислотных последовательностей на уровне 40-95%. Выравнивание собранной неизбыточной выборки структур выполняли алгоритмом MATT [218].

4.2.1.3. Построение структурно-опосредованного множественного выравнивания

Построение структурно-опосредованного множественного выравнивания нейраминидаз из различных источников выполняли с использованием веб-сервера Mustguseal, который позволяет конструировать выравнивания с использованием всей доступной информации о последовательностях и

структурах родственных белков, а также программ и алгоритмов, описанных выше (https://biokinet.belozersky.msu.ru/mustguseal) [185, 186]. Уникальные возможности достигаются за счет последовательного сочетания четырех типов алгоритмов биоинформатики – поиск по структурному сходству используется для поиска эволюционно далеких родственных белков, которые приобрели функциональное/регуляторное разнообразие за счет значительного изменения аминокислотных последовательностей и структур в процессе естественного отбора (Рисунок 14, этап №1); структурное выравнивание – для наложения найденных эволюционно далеких белков, которые представляют различные семейства в рамках большого суперсемейства (Рисунок 14, этап №2); поиск по сходству аминокислотных последовательностей – для поиска эволюционно близких белков, представителей выбранных семейств, и последующее выравнивание аминокислотных последовательностей – для наложения последовательностей эволюционно близких белков (представителей одного семейства; Рисунок 14, этап №3). На последнем этапе выравнивание доступных структур эволюционно удаленных гомологов используется в качестве «направляющей» для совмещения аминокислотных последовательностей всех собранных представителей суперсемейства (Рисунок 14, этап №4).

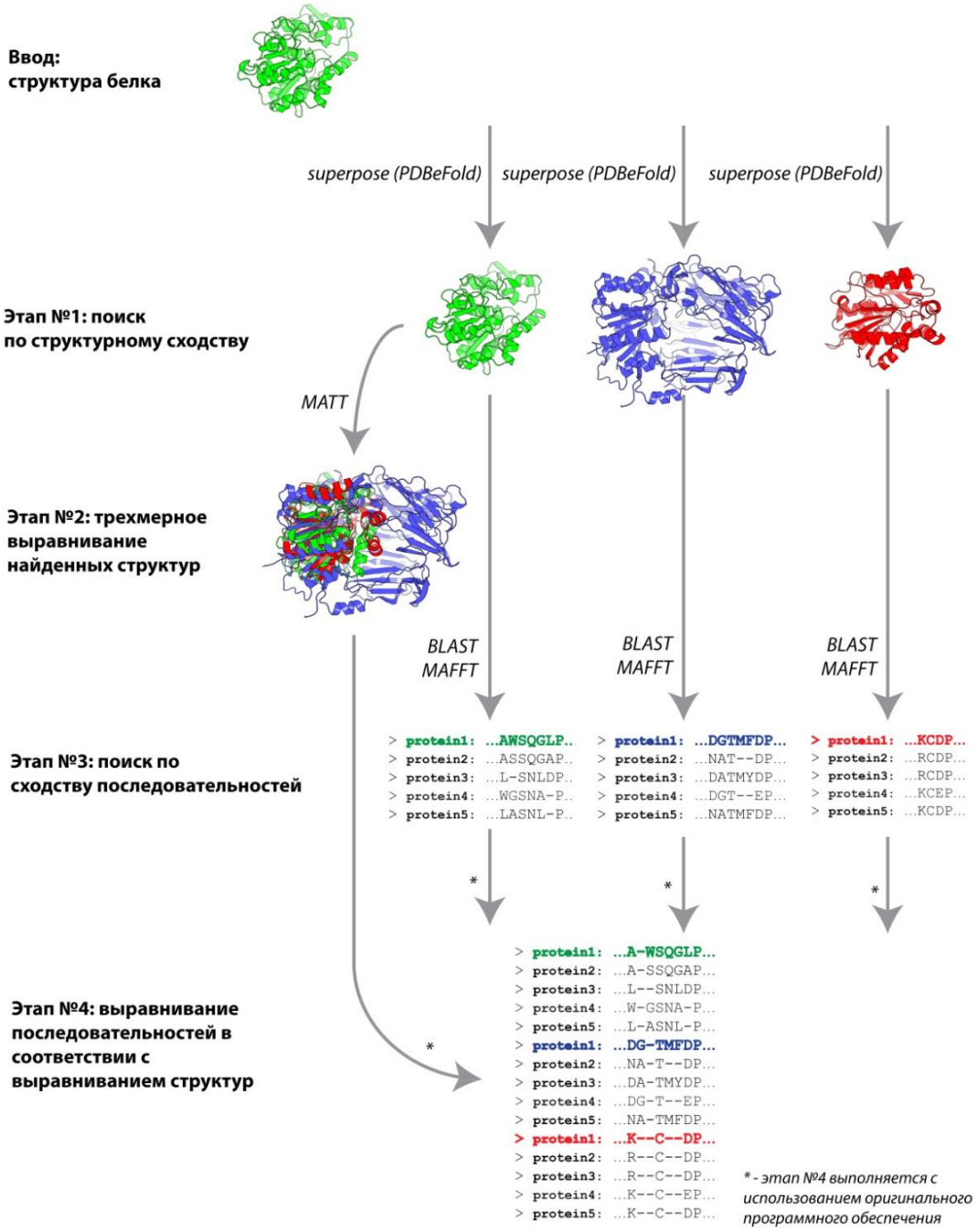

Рисунок 14. Схема протокола Mustguseal для построения структурно-опосредованного множественного выравнивания суперсемейств белков [185].

4.2.2. Анализ множественных выравниваний последовательностей и 3D-структур

Сравнительный анализ гомологичных белков в функционально разнообразном суперсемействе, т.е. совокупности эволюционно родственных белков, которые реализуют различные функции в рамках общей структурной организации, может выявить сложное взаимодействие между структурой, функцией и регуляцией белка. Идентификация позиций в структуре белков, которые демонстрируют консенсус или изменчивость среди гомологов, может помочь в изучении молекулярных механизмов действия и регуляции белков, разработке улучшенных вариантов ферментов/белков и селективных лигандов для модуляции их функциональных свойств [191]. В этой работе веб-сервер Mustguseal использовали для построения множественных выравниваний белков, после чего использовали дополнительные веб-серверы для анализа полученного выравнивания.

Множественное выравнивание для анализа строили с использованием рекомендованных параметров: в качестве затравок (queries) использовали трехмерные структуры вирусной нейраминидазы и NanA; пороги эквивалентности вторичной структуры не менее 90% для искомого белка и не менее 90% для записи из базы данных PDB; для фильтрации собранной выборки последовательностей и структур с целью устранения избыточных записей использовали порог 95% идентичности. Остальные параметры для построения выравнивания использовали по умолчанию, как описано выше. На основании собранных выборок последовательностей и структур бактериальных и вирусных нейраминидаз строили единое (общее) выравнивание. Полученное выравнивание анализировали с использованием веб-сервера Zebra2 [192, 193]. Количество перетасовок (shuffles) каждой колонки множественного выравнивания, которое регулирует точность случайной модели для оценки

статистической значимости найденных консервативных и специфических позиций, соответствовало 10000. Для коррекции оценки специфичности позиций с использованием окружения в трехмерном пространстве использовался 3D-mode, учитывались соседние остатки в радиусе 4Å. Остальные параметры использовались со значениями по умолчанию. Обнаруженные консервативные и специфические позиции были проранжированы в соответствии с полученными оценками статистической значимости. Для последующего изучения были отобраны только те позиции, статистическая значимость которых была выше глобального порога, автоматически рассчитанного по формуле знако-ранговой статистики Бернулли, т.н. «B-cutoff» [194].

4.3. Молекулярное моделирование [ii]

4.3.1. Подготовка полноатомных молекулярных моделей

4.3.1.1. Подготовка моделей лигандов

Структуру оптактина ((1s,3R,4S)-1-[(циклогексиламино)метил]-3,4-дигидроксициклопентансульфокислоты) построили в программе ChemSketch v.12.01 [ACD/ChemSketch Freeware, version 12.01, Advanced Chemistry Development, Inc., Toronto, ON, Canada, www.acdlabs.com, 2012] и

[ii] При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: **Sharapova Y.**, Suplatov D., Švedas V. Neuraminidase A from Streptococcus pneumoniae has a modular organization of catalytic and lectin domains separated by a flexible linker //The FEBS journal. – 2018. – Т. 285. – №. 13. – С. 2428-2445; **Sharapova Y.,** Švedas V., Suplatov D. Catalytic and lectin domains in neuraminidase A from Streptococcus pneumoniae are capable of an intermolecular assembly: Implications for biofilm formation //The FEBS Journal. – 2021. – Т. 288. – №. 10. – С. 3217-3230; **Шарапова Я. А.**, Швядас В. К. Молекулярное моделирование связывания аллостерического ингибитора оптактина в новом сайте в структуре нейраминидазы a из Streptococcus pneumoniae //Вестник Московского университета. Серия 2. Химия. – 2018. – Т. 59. – №. 5. – С. 323-331; **Sharapova Y. A.**, Suplatov D. A., Švedas V. K. Simulating the long-timescale structural behavior of bacterial and influenza neuraminidases with different HPC resources //Supercomputing Frontiers and Innovations. – 2018. – Т. 5. – №. 3. – С. 30-33. Объем вклада автора диссертации в трех первых работах составляет 75%, в последней – 50%, и заключался в задумке, планировании и выполнении экспериментов, а также написании статей.

оптимизировали с использованием программы квантовой химии Firefly (http://classic.chem.msu.su). Для расчета зарядов использовали пакет R.E.D. tools [195], остальные параметры были взяты из силового поля GAFF.

Трехмерная модель артокарпина была взята из базы данных PubChem (ID 5458461), заряды для нее были рассчитаны с помощью веб-сервера R.E.D. (https://upjv.q4md-forcefieldtools.org/REDServer-Development/) [196] с использованием Gaussian16. Остальные параметры также были взяты из силового поля GAFF.

Структуры сахаров, использованных в работе, строились на основе кристаллографической структуры 3HTT. Параметры для молекул сахаров брались из силового поля GLYCAM_06j-1.

4.3.1.2. Подготовка модели фермента/фермент-субстратного комплекса

Модифицированная версия PDB2PQR v. 1.9.0 [197, 198], совместимая с PROPKA v. 3.1 [199], использовалась для оценки значений pKa ионизируемых групп и их протонирования при физиологическом pH, устранения стерических перекрываний и оптимизации сети водородных связей. Пакет AmberTools использовался для параметризации системы в выбранном силовом поле Amber (выбор силовых полей обсуждается ниже). Начальные модели белков помещались в кубическую ячейку с растворителем так, чтобы расстояние от любого атома белка до края периодической ячейки было не менее 12Å. Для гомологичных моделей сиалидаз (т.е. моделей NanA и моделей NanB и NanC, полученных методом обратного гомологичного моделирования), предсказанных ZDOCK моделей NanC, сиалидазы *T. rangeli* и сиалидазы *V. cholerae*, а также моделей межмолекулярных комплексов NanA, размер системы был увеличен так, чтобы расстояние от любого атома белка до края ячейка было не менее 30 Å [200, 201]. Это было сделано для того, чтобы предотвратить возникновение

артефактов в периодической системе, т.е., чтобы избежать взаимодействия расширяющейся молекулы NanA с ее изображением в соседней ячейке в МД-симуляции в случае диссоциации комплекса двух доменов.

Для подготовки моделей пневмококковых нейраминидаз в комплексе с оптактином использовали кристаллографическую структуру NanB из *S. pneumoniae* в комплексе с оптактином (PDB 4XHX) и структуру NanA из *S. pneumoniae* без оптактина (PDB 2YA8). Для подготовки начальной модели комплекса NanA с оптактином производили структурное выравнивание аллостерического центра в NanB (4XHX) и нового центра в NanA (2YA8), после чего координаты лиганда копировали из одной структуры в другую [202].

4.3.2. Молекулярная динамика

4.3.2.1. Использование силовых полей FF14SB и FF15IPQ для параметризации моделей

При изучении доменной организации молекулы NanA все ее модели, построенные по гомологии, а также структуры самих гомологов были параметризованы и изучены независимо в двух силовых полях Amber (FF14SB и FF15IPQ) и с двумя моделями воды (TIP4P-EW и SPC/*E*b), чтобы оценить влияние различных параметров симуляции на результаты моделирования.

Обычно используемое силовое поле FF14SB [203] использовалось как описано в работе Pierce *et al.* [210]. С этим полем использовалась четырехцентровая модель воды TIP4P-Ew, т.к. ранее было продемонстрировано, что она лучше описывает вращательное движение белков по сравнению с обычно используемой трехцентровой моделью TIP3P [204]; кроме того, TIP4P-Ew хорошо воспроизводит свойства растворителя как жидкости при температурах и давлениях, релевантных в биологии [205]. Размер молекулярных систем, построенных на основе кристаллографических структур

нейраминидаз/сиалидаз в водной ячейке составлял 110 000 – 140 000 атомов, а размер систем гомологичных моделей в растворителе – 305 000 – 350 000 атомов (из-за увеличенного размера периодической ячейки). Средняя скорость МД-расчета на четырех Tesla K40 суперкомпьютера составила 23–30 нс/день и 9–11 нс/день, соответственно.

Недавно было разработано силовое поле FF15IPQ с использованием Q-схемы неявной поляризации с целью повысить точность современных силовых полей (в т.ч., FF14SB). Было отмечено, что поле FF15IPQ более корректно моделирует солевые мостики, вторичную и третичную структуру белка во время длительных симуляций, и может быть использовано для моделирования упорядоченности/иррегулярности в белковых структурах [206]. Атомные заряды в силовом поле FF15IPQ были рассчитаны с учетом недавно разработанной трехточечной модели воды SPC/*E*b, которая аккуратно описывает вращательную диффузию белков в растворе [207]. Размер молекулярных систем, построенных на основе кристаллографических структур нейраминидаз/сиалидаз в водной ячейке составлял 85 000 – 110 000 атомов, а размер систем гомологичных моделей и полученных белок-белковым докингом – 230 000 – 275 000 атомов (из-за увеличенного размера периодической ячейки). Средняя скорость МД-расчета на четырех Tesla K40 суперкомпьютера составила 34–45 нс/день и 13–16 нс/день, соответственно.

Основной вывод, который можно сделать из всех МД-симуляций сиалидаз – в контексте данного исследования результат не зависел от выбора силового поля, т.к. все модели NanA, построенные по гомологии, в конечном итоге разделились на лектиновый и каталитический домены, связанные междоменным линкером, в то время как все модели NanB, NanC, транс-сиалидазы из *M. decora*, сиалидаз из *T. rangeli*, *T. cruzi* и *V. cholerae* сохраняли компактную глобулярную структуру на протяжении всей МД-траектории. Поэтому одна иллюстрация МД-

траектории при произвольно выбранных параметрах симуляции будет показана далее в тексте для каждой модели. Однако с учетом того, что были получены качественно эквивалентные результаты, стоит отметить значительную разницу в скорости компьютерных расчетов [208, 209]. Использование трехцентровой модели воды SPC/*E*b с полем FF15IPQ вместо четырехцентровой модели воды TIP4P-Ew с полем FF14SB оставляет пустой сайт в каждом молекуле растворителя, уменьшая общее количество атомов в периодической ячейке, что уменьшает сложность расчета электростатических взаимодействий, и, таким образом, дает улучшение скорости расчета молекулярной динамики на четырех графических процессорах Tesla K40 в 1,45–1,50 раза. Поэтому в дальнейшем, при изучении межмолекулярной агрегации NanA, расчеты велись в поле FF15IPQ с растворителем SPC/*E*b с целью ускорения расчета траекторий.

4.3.2.2. Протокол оптимизации моделей (минимизация и уравновешивание)

Протокол классической молекулярной динамики описан в работе Pierce *et al.* [210]. Сначала исходные системы проходят минимизацию энергии в вакууме, затем помещаются в ячейку с растворителем, как описано выше, и затем энергия снова минимизируется, но уже в воде, для того, чтобы удалить стерические перекрывания и способствовать образованию комплементарности двух подструктур в гомологичных моделях. Далее на все тяжелые атомы белка накладываются гармонические позиционные ограничения величиной 3 ккал/моль*A^2, и система нагревается от 0 до 300 К в течение 500 пс в ансамбле NVT. На следующем этапе система уравновешивается (эквилибрируется) в ансамбле NPT при температуре 300 К и постоянном изотропном давлении 1 атм, а позиционные ограничения постепенно уменьшаются с 3 до 0 ккал/моль*A^2 в 13 шагов (по 200 пс каждый). Графики плотности, температуры и общей

энергии сошлись к концу периода эквилибрации. Полученные системы использовались для свободных (без позиционных ограничений) молекулярно-динамических симуляций.

4.3.2.3. Протокол свободной динамики и эффективное использование HPC-ресурсов для его исполнения

Свободные (без позиционных ограничений) молекулярно-динамические симуляции проводились в ансамбле NVT при температуре 300 К. Продолжительность симуляций составила по 150 нс для гомологичных моделей (т.е. моделей NanA и моделей NanB и NanC, полученных методом обратного гомологичного моделирования), а также предсказанных ZDOCK моделей NanC, сиалидазы *T. rangeli* и сиалидазы *V. cholerae*, и по 500 нс для кристаллографических структур сиалидаз. Для расчета электростатических взаимодействий в периодических системах использовался метод PME (Particle Mesh Ewald), радиус отсечения электростатических взаимодействий составил 10 Å [211]. Шаг интегрирования был установлен на 2 фс. Термостат Ланжевена использовался во время МД-симуляций для поддержания температуры на уровне 300 К. Расчет молекулярной динамики проводился в AMBER с использованием GPU-версии PME с явно заданным растворителем [212, 213, 214]. Этапы минимизации, нагревания и уравновешивания проводилось на локально установленном GeForce GTX 980 Ti GPU-ускорителе, свободная молекулярная динамика – на четырех GPU-ускорителях Tesla K40 суперкомпьютера Ломоносов-2 в режиме MPI.

4.3.3. Моделирование полноразмерной структуры NanA

4.3.3.1. Подготовка исходных данных с использованием биоинформатического анализа

Комплекс биоинформатических методов сравнения белковых структур и последовательностей был применен для отбора репрезентативной выборки всех сиалидаз, в которых лектиновый и каталитический домены, схожие с NanA, были закристаллизованы в виде единой глобулы [185]. Структуры лектинового и каталитического (включая инсерционный здесь и далее) доменов NanA из *S. pneumoniae* были использованы в качестве «queries» для запуска двух независимых поисков по структурному сходству по всей базе данных PDB с использованием алгоритма SSM [184]. В процессе поиска по структурному сходству структуры отбирались на основании процента идентичных элементов вторичной структуры – «target»-структура должна была содержать по меньшей мере 50% идентичных элементов вторичной структуры с «query». Для дальнейшего изучения были отобраны лишь те белки, которые нашлись в обоих независимых поисках, а значит, содержат в своей структуре лектиновый и каталитический домены, схожие с NanA и закристаллизованные в виде единой глобулы. Отобранные белки были проранжированы по значению Q-score, которое выдается алгоритмом SSM и отображает качество попарного пространственного наложения каждого белка с каталитическим доменом NanA. Из похожих белков, идентичных по аминокислотной последовательности на 95% и более, были удалены те, что имели более низкий Q-score. Таким образом, была получена репрезентативная, неизбыточная выборка белков, в которых закристаллизованы в виде единой глобулы домены, похожие на лектиновый и каталитический домены NanA из *S. pneumoniae*. Аннотация доменов в структурах белков производилась согласно классификации CATH и

литературным источникам [98, 108, 215, 216]. Лектиновый и каталитический домены NanA из *S. pneumoniae* принадлежат к семействам CBM40 и GH33 классификации CAZy, соответственно [10]. Экспертный анализ базы данных CAZy показал, что биоинформатический анализ корректно идентифицировал все известные сиалидазы, в которых CBM40-лектиновый домен закристаллизован вместе с GH33-каталитическим, а также отобрал те белки, в которых лектиновые домены похожи на лектиновый домен NanA, но еще не были отнесены к соответствующему семейству CAZy.

4.3.3.2. Подготовка исходных данных с использованием белок-белкового докинга

Гомологичное моделирование, основанное на биоинформатическом анализе родственных нейраминидаз/сиалидаз и белок-белковом докинге, было использовано независимо друг от друга для предсказания взаимной ориентации лектинового и каталитического доменов в структуре NanA пневмококка.

Предсказание взаимного расположения доменов в NanA осуществлялось путем белок-белкового докинга доменов программой ZDOCK [217]. Кристаллографические структуры лектинового и каталитического доменов, не содержащие междоменный линкер (т.е. PDB ID 4ZXK и 2YA8, Рисунок 15 А), независимо друг от друга прошли минимизацию энергии, эквилибрацию, а затем подверглись молекулярно-динамической симуляции длительностью 250 нс в силовом поле FF14SB, как описано выше. Шесть структурных вариантов каждого домена были отобраны из молекулярно-динамических траекторий с 50-нс-шагом (а именно, конформации доменов в 0 нс, 50 нс, 100 нс, 150 нс, 200 нс, 250 нс), после чего были использованы для тридцати шести запусков белок-белкового докинга (т.е. каждый из шести структурных вариантов лектинового домена докировался независимо в каждый из шести вариантов каталитического

домена). Затем вся процедура была повторена, но молекулярная динамика проводилась уже в силовом поле FF15IPQ, как описано выше. Всего, таким образом, было проведено 72 запуска белок-белкового докинга, и в каждом запуске программа ZDOCK предложила 10 разных ориентаций лектинового домена на поверхности каталитического (включая инсерционный) для дальнейшего анализа – в общей сложности 720 комплексов. Программа Modeller использовалась для того, чтобы оценить, могут ли два домена в ориентации, предсказанной ZDOCK, быть соединенными междоменным линкером в 16 остатков; структурный анализ программой PyMol использовался для проверки того, смогут ли участок связывания субстрата в лектиновом домене (который включает остатки Phe167, Asp180, Tyr182, Thr193, Glu195, Arg197, Gln203, Asn209, Arg274 и Trp280, нумерация по структуре PDB 4ZXK) и активный центр в каталитическом домене (остатки Arg332, Asp357, Glu632, Arg648, Arg706 и Tyr737, нумерация по структуре PDB 2YA8) связать субстрат. В результате проверки предсказаний ZDOCK с помощью Modeller и PyMol выяснилось, что все 720 предсказаний некорректны с точки зрения структуры или функции – в подавляющем большинстве предложенных белок-белковых комплексов учасок связывания субстрата в одном из доменов был заблокирован вторым доменом, а в остальных случаях два домена невозможно было соединить 16-а.о.-линкером. Поэтому результаты белок-белкового докинга далее не рассматривались.

4.3.3.3. Гомологичное моделирование полноразмерной структуры NanA

Предсказание структуры по гомологии использовалось для построения нескольких моделей NanA путем состыковывания ее доменов по аналогии с тем, как они состыкованы в структурах эволюционно родственных сиалидаз (Рисунок 15 и Таблица 1). На первом этапе все возможные ориентации

лектинового домена на поверхности каталитического были смоделированы в соответствии с результатами биоинформатического анализа. Структурное выравнивание отдельных доменов NanA (т.е. PDB ID 4ZXK и 2YA8, Рисунок 15 A) с полными структурами каждого гомолога выполнялось с использованием программы MATT [218]. MATT ищет совместимые в пространстве пары фрагментов структуры с учетом поворотов и трансляций. Эта программа хорошо справляется с выравниванием удаленных гомологов разного размера [219], и потому была использована в этом исследовании для сравнения NanA с ее эволюционно далекими родственниками. На втором этапе каждая предсказанная ориентация доменов NanA была подана на вход программе Modeller [220], чтобы соединить две подструктуры междоменным линкером и избавиться от стерического перекрывания аминокислотных остатков в месте контакта двух доменов. В каждом случае применялось два альтернативных способа смоделировать междоменный линкер: как свободную петлю или с позиционными ограничениями, наложенными на остатки 306-310 (нумерация в соответствии с PDB ID 4ZXK), чтобы они сформировали α-спираль согласно предсказанию сервера-предиктора вторичной структуры Jpred [221]. В случаях, когда ориентация доменов NanA моделировалась по гомологии с сиалидазами трипаносом, применялся также третий способ моделирования линкера – остатки 308-322 укладывались в α-спираль – так же, как они уложены в линкерах гомологов (Рисунок 15 E и F). Для каждой ориентации доменов и каждого варианта организации вторичной структуры междоменного линкера программа Modeller предсказывала 20 альтернативных структурных вариантов. Отбирался вариант с лучшим значением DOPE-score [222], проходил минимизацию энергии, эквилибрацию и изучался с помощью APBS-анализа (Adaptive Poisson-Boltzmann Solver) и молекулярной динамики (см. раздел 4.3.2.3). Все предсказанные молекулярные модели NanA из *S. pneumoniae* включали N-

концевой лектиновый домен, междоменный линкер и каталитический домен (включая инсерционный) в пределах единой полипептидной цепи. С-концевой домен NanA не был включен в модели, поскольку его структура неизвестна ни для NanA, ни для любого другого гомологичного белка и поскольку считается, что этот домен играет роль «якоря» и никак не участвует в каталитической функции фермента [223]. Результаты APBS-анализа и МД-симуляций для моделей NanA, в которых отличалась исходная вторичная структура линкера, оказались, по существу, одинаковыми. Поэтому каждая ориентация доменов NanA, предсказанная биоинформатическим анализом, проиллюстрирована в дальнейшем одним произвольно выбранным вариантом линкера (Рисунок 16).

Таблица 1. Гомологичные нейраминидазы/сиалидазы, которые имеют в своей структуре лектиновый и каталитический домены, схожие с доменами NanA и закристаллизованные одной глобулой.

PDB	Название	Лектиновый домен			Каталитический домен	
		Топология	NanA SSE (%)	SeqID (%)	NanA SSE (%)	SeqID (%)
2JKB	NanB из *Streptococcus pneumoniae*	N-концевой	82	21	85	23
4YW2	NanC из *Streptococcus pneumoniae*	N- концевой	87	22	83	26
2SLI	Транс-сиалидаза из *Macrobdella decora*	N- концевой	93	23	84	29
1MZ6	Сиалидаза из *Trypanosoma rangeli*	C- концевой	83	13	71	14
3PJQ	Сиалидаза из *Trypanosoma cruzi*	C- концевой	82	12	69	15
1W0O	Сиалидаза из *Vibrio cholerae*	N- концевой	65	7	67	17
		Инсерционный	68	9		

Сходство по структуре и аминокислотной последовательности между лектиновым доменом и каталитическим доменом (включая инсерционный домен) NanA и соответствующими доменами в структурах гомологичных нейраминидаз/сиалидаз. «NanA SSE» соответствует проценту совпадений элементов вторичной структуры между NanA и гомологом, т. е. доле аминокислот, участвующих в элементах вторичной структуры (SSE) NanA, которые могут быть выровнены с эквивалентными остатками в структуре соответствующего гомолога. «SeqID» соответствует проценту идентичности по последовательности NanA и гомолога.

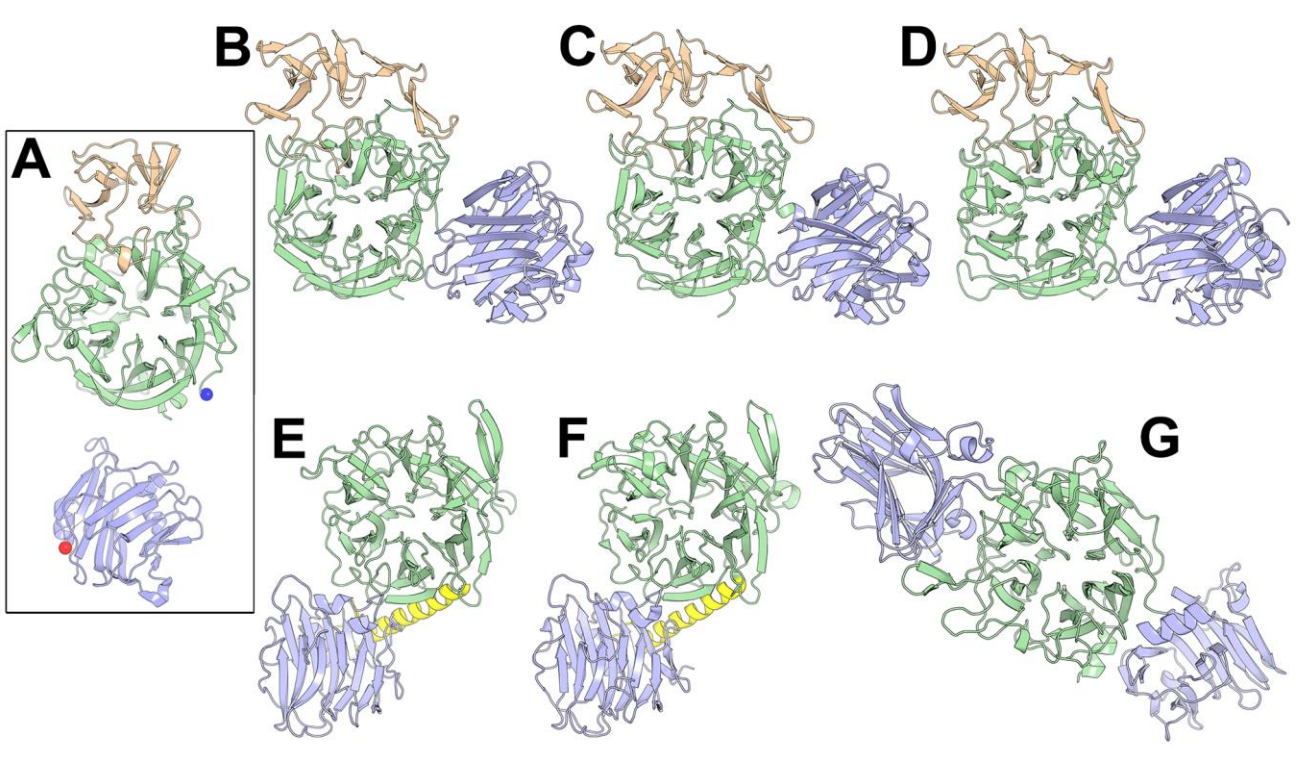

Рисунок 15. Структуры PDB отдельных доменов NanA из *S. pneumoniae* и структуры PDB гомологичных нейраминидаз/сиалидаз, которые имеют каталитические и лектиновые домены, гомологичные доменам NanA и закристаллизованные в одной глобуле. (A) NanA из *S. pneumoniae*, PDB: 2YA8 (каталитический домен, включая инсерционный домен) и PDB: 4ZXK (лектиновый домен); (B) NanB из *S. pneumoniae*, PDB: 2JKB; (C) NanC из *S. pneumoniae*, PDB: 4YW2; (D) транс-сиалидаза из *M. decora*, PDB: 2SLI; (E) сиалидаза из *T. rangeli*, PDB: 1MZ6; (F) сиалидаза из *T. cruzi*, PDB: 3PJQ; (G) сиалидаза из *V. cholerae*, PDB: 1W0O. Лектиновый домен выделен голубым цветом, каталитический домен – зеленым, с инсерционным доменом (там, где он есть), выделенным бежевым. Междоменный линкер выделен желтым (там, где он есть). На изображении (A) Cα-атомы последнего кристаллизованного аминокислотного остатка в лектиновом домене (красный) и первого кристаллизованного аминокислотного остатка в каталитическом домене (синий) NanA показаны в виде сфер. Изображения были созданы с использованием PyMol.

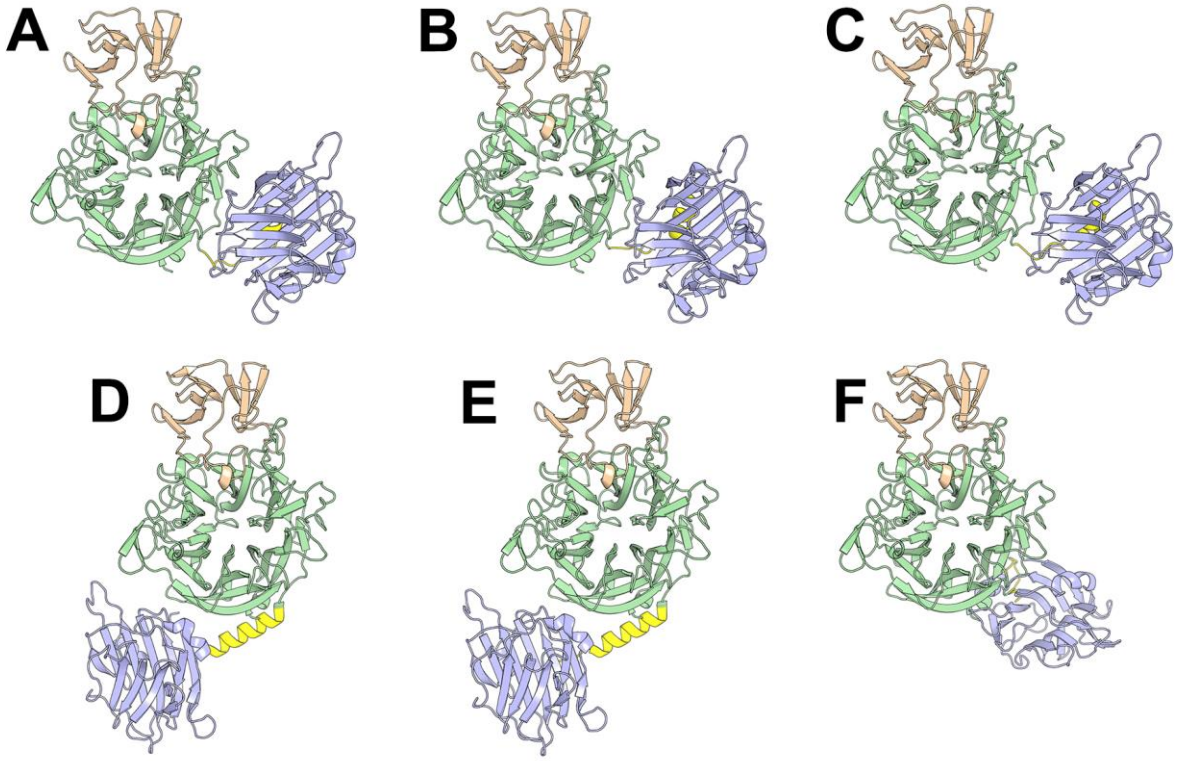

Рисунок 16. Ориентации лектинового домена на поверхности каталитического домена (включая инсерционный домен) NanA из *S. pneumoniae*, смоделированные в соответствии с биоинформатическим анализом гомологичных нейраминидаз/сиалидаз (т.е. гомологичные модели NanA до молекулярно-динамических симуляций): (A) NanB из *S. pneumoniae*; (B) NanC из *S. pneumoniae*; (C) транс-сиалидаза из *M. decora*; (D) сиалидаза из *T. rangeli*; (E) сиалидаза из *T. cruzi*; (F) сиалидаза из *V. cholerae* (на основе ее N-концевого лектинового домена).

4.3.3.4. Контрольные эксперименты для оценки полученных моделей NanA

Контрольный эксперимент по обратному гомологичному моделированию структур NanB и NanC из *S. pneumoniae* был проведен для того, чтобы оценить описанную выше процедуру предсказания структуры NanA по гомологии. Полученные молекулярные модели NanA, содержащие лектиновый и каталитический домены, состыкованные так же, как в кристаллографических структурах NanB и NanC (Рисунок 16 A и B), были использованы как образцы для обратного гомологичного моделирования соответствующих сиалидаз.

Полные структуры NanB и NanC были разделены на две подструктуры, соответствующие лектиновому и каталитическому (включая инсерционный) доменам. Была проведена минимизация энергии, эквилибрация и МД-симуляция в течение 50 нс для отдельных подструктур, независимо друг от друга. Полученные после симуляции подструктуры NanB и NanC были наложены на соответствующие исходные модели NanA (Рисунок 16 A и B) алгоритмом структурного выравнивания MATT, а затем поданы на вход программе Modeller, чтобы соединить подструктуры в единую полипептидную цепь и избавиться от стерического перекрывания аминокислотных остатков в месте контакта двух доменов. Для полученных моделей была проведена минимизация энергии и эквилибрация. Молекулярная динамика показала, что модели NanB и NanC, полученные обратным гомологичным моделированием, были стабильны в тех же условиях, что использовались для оценки гомологичных моделей NanA (Рисунок 17, протокол молекулярной динамики см. выше). Из этого следует вывод, что описанная процедура для предсказания структуры по гомологии и последующая оптимизация молекулярной динамикой корректна и может использоваться для моделирования полных белковых структур из независимых подструктур отдельных доменов, имеющих комплементарные поверхности контакта.

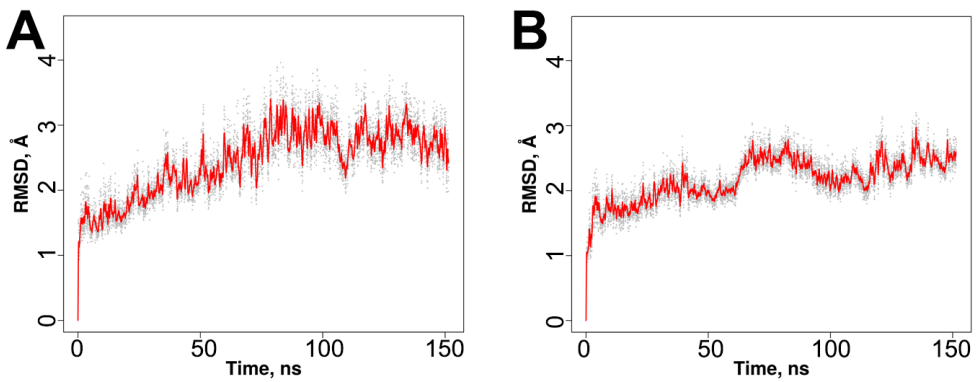

Рисунок 17. Графики RMSD моделей (A) NanB и (B) NanC из *S. pneumoniae*, полученные с помощью обратного моделирования по гомологии, в течение 150 нс МД-симуляции при 300K. RMSD рассчитывали по атомам основной цепи всех доменов; кадры траекторий были выровнены по каталитическому домену. Серые точки соответствуют фактическим значениям RMSD, красные линии - средним значениям, рассчитанным в пределах временного окна в 400 пс. Изображения были подготовлены с использованием языка R.

Также был проведен контрольный эксперимент по реконструкции известного междоменного интерфейса программой ZDOCK для оценки способности алгоритма белок-белкового докинга правильно предсказывать интерфейс между каталитическим и лектиновым доменами. PDB-структуры NanC из *S. pneumoniae* (Рисунок 15 C), сиалидазы из *T. rangeli* (Рисунок 15 E) и сиалидазы из *V. cholerae* (Рисунок 15 G) были выбраны для эксперимента и разделены на две подструктуры: каталитический домен (включая вставочный) и N-концевой лектиновый в случае NanC; каталитический домен и C-концевой лектиновый в случае сиалидазы из *T. rangeli* (то есть, междоменный линкер был удален); каталитический домен (с инсерционным лектиновым доменом) и N-концевой лектиновый домен в случае сиалидазы из *V. cholerae*. Эти подструктуры были независимо друг от друга проминимизированы, эквилибрированы, а затем просимулированы в течение 50 нс в силовом поле FF15IPQ (см. протокол молекулярной динамики выше) для того, чтобы сгенерировать структурное разнообразие. Для каждой нейраминидазы/сиалидазы из МД-траекторий было отобрано по три

конформационных варианта каждой подструктуры (0 нс, 25 нс и 50 нс), которые были далее использованы для девяти запусков белок-белкового докинга (т.е. каждый из трех вариантов первой подструктуры был независимо докирован в каждый из трех вариантов второй подструктуры). Таким образом, для каждого белка было предсказано по 90 комплексов двух соответствующих подструктур. Для оценки предсказаний ZDOCK были использованы те же самые критерии, что использовались при попытке состыковать домены NanA (см. раздел 4.3.3.2), т.е. Modeller был использован, чтобы проверить, могут ли два домена быть соединены в предложенных ориентациях (в случаях NanC и сиалидазы из *V. cholerae* программа пыталась напрямую соединить два домена, а в случае сиалидазы из *T. rangeli* учитывалось наличие междоменного линкера длиной 25 а.о.), и структурный анализ с помощью PyMol использовался для того, чтобы оценить, смогут ли участок связывания субстрата в лектиновом домене и активный центр каталитического домена связывать субстрат. Эти фильтры были успешно пройдены 5, 11 и 5 комплексами из 90 для NanC, сиалидазы из *T. rangeli* и сиалидазы из *V.cholerae,* соответственно. Для сравнения – ни один комплекс из 720, предсказанных ZDOCK для доменов NanA не прошел ту же самую процедуру оценки. Дальнейшее сравнение полученных комплексов с соответствующими полноразмерными кристаллографическими структурами NanC, сиалидазы из *T. rangeli* и сиалидаза из *V. cholerae* показало, что две конформации с RMSD 6 Å, две конформации с RMSD 6 Å и одна конформация с RMSD 1,25 Å были корректно предсказаны ZDOCK в каждом случае (значения RMSD были рассчитывались для атомов основной цепи лектинового домена после выравнивания предсказанного комплекса с кристаллографической структурой PDB по атомам остова каталитического домена). Эти полноразмерные модели сиалидаз, которые включали оба домена (а также междоменный линкер - в случае сиалидазы из *T. rangeli*) были затем изучены

методом молекулярной динамики. МД-симуляции показали, что полученные модели были стабильны в тех же условиях, которые использовались для оценки гомологичных моделей NanA (Рисунок 18, см. протокол МД в разделе 4.3.2.3). Из этого следует вывод, что алгоритм ZDOCK способен корректно воспроизводить известные междоменные интерфейсы при затрате гораздо меньшего количества усилий, чем было предпринято при попытке докировать домены NanA, что, в свою очередь, говорит о том, что неспособность ZDOCK предсказать стабильное расположение доменов для NanA является значимой.

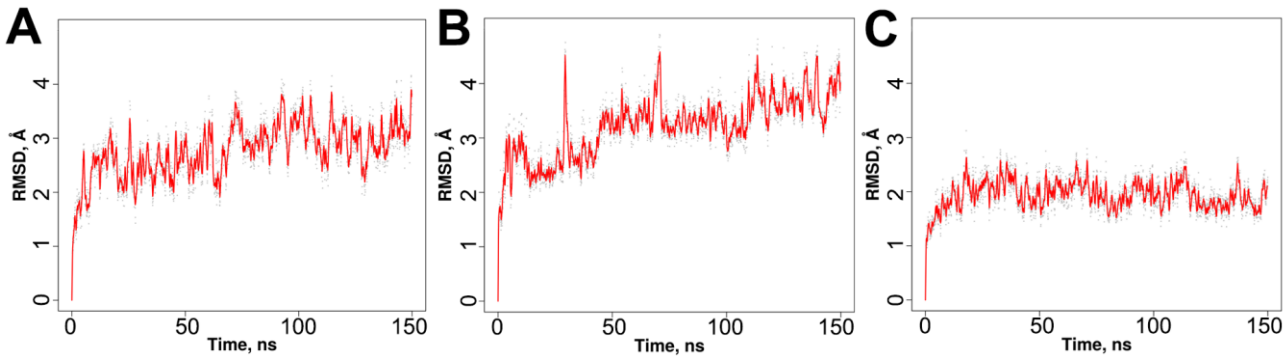

Рисунок 18. Графики RMSD моделей (A) NanC, (B) сиалидазы из *T. rangeli*, (C) сиалидазы из *V. cholerae*, реконструированных методом белок-белкового докинга, в течение 150 нс МД-симуляции при 300K. RMSD рассчитывали по атомам основной цепи всех доменов; кадры траекторий были выровнены по каталитическому домену. Серые точки соответствуют фактическим значениям RMSD, красные линии - средним значениям, рассчитанным в пределах временного окна в 400 пс. Изображения были подготовлены с использованием языка R.

4.3.4. Моделирование межмолекулярной агрегации NanA

Кристаллографические структуры лектинового и каталитического доменов (включая инсерционный домен), не содержащие междоменного полипептидного линкера, независимо друг от друга претерпели 1000 нс МД-симуляции при 310 K (протокол МД обсуждается в разделе 4.3.2) для того, чтобы сгенерировать ансамбли конформеров соответствующих подструктур NanA. Для каждого домена было отобрано по сотне структурных вариантов с

временным шагом 10 нс (т. е. 10 нс, 20 нс, 30 нс и т.д.), которые затем использовались для проведения ста запусков белок-белкового докинга с помощью сервера ClusPro [224], т. е. структурные варианты лектинового домена докировались в конформеры каталитического домена, отобранные в тот же временной интервал в независимой симуляции. Проводился т. н. «слепой» докинг полноразмерных доменов, без указания каких-либо притягивающихся / отталкивающихся аминокислотных остатков. Для дальнейшей работы брали по десять лучших комплексов каталитического и лектинового доменов согласно ранжированию результатов каждого запуска докинга ClusPro. Лучший комплекс доменов (т. е. с рангом #1) из каждого докинга был подвергнут процедуре МД-симуляции в течение 150 нс при 310 К, за исключением тех случаев, когда субстрат-связывающий центр одного домена был блокирован другим доменом в предложенном докингом комплексе (как обсуждалось в разделе 4.3.3.2). Далее рассматривались только те комплексы, которые были стабильны в течение всей МД-траектории; все они были качественно эквивалентны и соответствовали комплексу лектинового и каталитического доменов, который далее обсуждается в секции результатов 5.1.4 (т.е. Рисунок 27, А). Структурный анализ показал, что этот комплекс был успешно предсказан ClusPro в 81% случаев, то есть в большинстве запусков докинга соответствующий комплекс структурных вариантов лектинового и каталитического доменов входил в первую десятку предсказаний алгоритма. Структуры этого комплекса, полученные в результате докинга разных конформеров доменов, характеризовались максимальным значением RMSD 9,73 Å при среднем 6,01 Å (расчет по атомам основной цепи лектинового домена). Наконец, десять случайно выбранных структурных вариантов лектинового и каталитического доменов были поданы на вход программе белок-белкового докинга ZDOCK [217], после чего прошли молекулярно-динамическую симуляцию для того, чтобы оценить, влияет ли

выбор программного обеспечения на результат докинга. Результаты ZDOCK были эквивалентны результатам ClusPro, однако степень воспроизводимости оказалась ниже – комплекс каталитического и лектинового доменов (т.е. Рисунок 27, А) наблюдался в числе десяти лучших предсказаний лишь в одном запуске из десяти.

Аналогичным образом был проведен докинг полученных из МД структурных вариантов доменов с целью изучения возможности взаимодействия одинаковых доменов из двух молекул NanA (т. е. двух каталитических или двух лектиновых). Наиболее высоко ранжированные алгоритмом ClusPro димеры каталитического домена (т.е. с рангом #1) оказались нестабильны в течение МД. Было замечено, что все доступные кристаллографические структуры каталитического домена NanA структурно эквивалентны и представляют собой димер (т. е. цепи A и B в записях PDB 2VVZ, 2W20, 2YA4, 2YA5, 2YA6, 2YA7, 2YA8, 3H71, 3H72, 3H73 и 5KKY) (Рисунок 27, В). Соответствующий димер наблюдался в числе двадцати лучших предсказаний ClusPro (т. е. с рангом #3–#18) в 72% запусков докинга. Модели димера, построенные докингом, характеризовались максимальным значением RMSD 12,16 Å при среднем 7,44 Å, в сравнении с доступными записями PDB (расчет по атомам основной цепи второй каталитической субъединицы после выравнивания моделей по первой субъединице), и оставались стабильными в течение 150 нс МД-симуляции при 310 К. Количественную разницу в результатах докинга лектинового домена в каталитический и каталитического домена в каталитический можно объяснить значительно большим размером каталитической субъединицы по сравнению с лектиновой (470 против 185 аминокислотных остатков), что приводит к гораздо большему числу степеней свободы во втором случае. Наконец, белок-белковый докинг не предложил стабильных структур димера лектиновых доменов. PDB-запись 4ZXK

лектинового домена NanA содержит димер; он, однако, не содержит междоменных взаимодействий, а также не был стабилен в течение МД-симуляций. Две другие доступные PDB-записи 4C1X и 4C1W содержали лишь мономер лектинового домена. Таким образом, димер лектинового домена в PDB-записи 4ZXK был определен как артефакт кристаллизации.

4.3.5. Моделирование взаимодействия низкомолекулярных лигандов с NanA

4.3.5.1. Докинг с использованием программы LeadFinder

Метод молекулярного докинга используют, чтобы предсказывать взаимодействие лигандов с макромолекулярными мишенями, например, при дизайне лекарств *in silico* [225]. Для докинга низкомолекулярных соединений в структуру белка (кристаллографическую, или же структурный вариант, полученный методом молекулярной динамики) использовалась программа LeadFinder [226]. Метод докинга заключается в поиске конформации лиганда с наилучшей энергией связывания. Свободная энергия связывания может задаваться следующим образом: $\Delta G = \Delta GVdW + \Delta Gelec + \Delta GH\text{-}bonds + \Delta Gtors + \Delta Gsol$. Первые три члена в этом выражении представляют вклад Ван-дер-Ваальсовых взаимодействий, электростатических взаимодействий и водородных связей в энергию связывания. $\Delta Gtors$ описывает потери в энтропии лиганда, $\Delta Gsol$ описывает изменения в энтропии растворителя на границе раздела фаз. Вычисление энергии проводится по всем парам атомов лиганд – макромолекула, однако, поскольку конформация макромолекулы остается неизменной, количество вычислений можно существенно сократить. Для этой цели непосредственно перед докингом рассчитывается карта потенциала вокруг макромолекулы. Эта карта представляет собой трехмерную решетку, окружающую интересующий участок поверхности макромолекулы (например,

активный центр фермента). В каждом узле решетки рассчитывается потенциал взаимодействия пробного атома с макромолекулой. В качестве стартовой точки для докинга берется случайное положение лиганда. На каждом шаге симуляции применяется небольшое случайное смещение по каждой степени свободы субстрата: сдвиг центра тяжести, изменение ориентации и вращение торсионных углов. Поскольку макромолекула неподвижна, энергия связывания зависит только от положения лиганда. Энергия новой конформации сравнивается с энергией предыдущего шага; если новая энергия ниже, конформация принимается. Для расчета докинга используют генетические алгоритмы, обеспечивающие достаточно полный анализ возможных конформаций лиганда, в сочетании с локальной минимизацией для более точной подстройки конформации.

4.3.5.2. Моделирование связывания артокарпина в структуре NanA

Комплексный подход, основанный на сочетании методов биоинформатики и молекулярного моделирования, был применен для поиска ранее неизвестного аллостерического центра распознавания ингибитора артокарпина в структуре нейраминидазы А (NanA) из патогенной бактерии *Streptococcus pneumoniae*. Новизной использованного подхода является учет пластичности структуры белка за счет анализа большого ансамбля конформеров. Конформационная пластичность является исключительно важным свойством белка – как для реализации его функции, так и при его эволюции. Тем не менее, в большинстве проводимых исследований при изучении связывания/распознавания лиганда в структуре сайта-мишени, а также при скрининге больших библиотек низкомолекулярных соединений, применяются стратегии, которые можно назвать «статическими» – изучаемые белки упрощенно представляются в виде единственной структуры, полученной,

как правило, из банка данных PDB. Последние исследования позволяют говорить о низкой эффективности таких «статических» подходов. Активные и аллостерические центры белков/ферментов следует рассматривать не как статические (жесткие) структуры, а в виде ансамбля конформеров [227].

По сотне структурных вариантов лектинового и каталитического доменов (включая инсерционный), полученных из молекулярно-динамических траекторий как описано в главе 4.3.4, были использованы для идентификации сайта связывания артокарпина в структуре NanA. Каждый структурный вариант был независимо исследован с помощью программы FPOCKET [228] для того, чтобы найти все потенциальные сайты связывания на его поверхности. Значение параметра «минимальное количество α-сфер на карман» было установлено на 27 [229]. Для каждого структурного варианта каталитического домена карман, соответствующий каталитическому центру, был исключен из дальнейшего рассмотрения, поскольку, исходя из кинетических исследований, было предположено, что связывание артокарпина происходит в отличном от каталитического центре, т.е. в аллостерическом [249]. Каждый детектированный программой FPOCKET карман в каждом структурном варианте двух доменов независимым образом использовался в качестве рецептора для докинга белок-лиганд с помощью программы LeadFinder [226]. Размеры решетки для докинга (т.е. карты поверхности электростатического потенциала) для каждого кармана определялись таким образом, чтобы она могла вместить артокарпин в самой вытянутой его конформации. Расчет велся в высокоточном режиме «extra precision». Из конфигураций лиганда, предложенных программой докинга, рассматривалась только одна лучшая поза артокарпина в каждом кармане (то есть с рангом #1 в выходном файле LeadFinder). Для каждого структурного варианта рецептора избыточные позы лиганда (т. е. возникшие в результате перекрытия соседних решеток) удалялись, пороговым значением RMSD было

выбрано 2 Å. Структурные варианты каждого из доменов с докированными в их поверхность лигандами были наложены в пространстве друг на друга с помощью программы структурного выравнивания parMATT [230]. Позы артокарпина, предложенные докингом, были скластеризованы [231] и проранжированы по степени воспроизводимости. В каталитическом домене кластер #1 соответствовал координатам лиганда, связанного в ранее неизвестном сайте связывания (см. главу 5.1.5), и характеризовался воспроизводимостью 84% (т.е. артокарпин наблюдался в соответствующей полости в 84 структурных вариантах каталитического домена NanA из 100, а в оставшихся 16 запусках докинга лиганд оказался значительно смещенным в координатном пространстве, хотя и находился в пределах решетки). Докинг не определил достоверного связывания артокарпина в лектиновом домене, поэтому оно в дальнейшем не рассматривалось [201].

Конформация артокарпина, связанного в новом сайте каталитического домена NanA, с наилучшей энергией ΔG_0, была взята из кластера #1 и подвергнута длительной молекулярно-динамической симуляции. Артокарпин оставался связанным в новом сайте в течение 1000 нс МД-симуляции при 310 К (Рисунок 32). Из последних 500 нс МД-траектории были отобраны с временным шагом 5 нс структуры комплекса белок-лиганд, которые использовались для оценки свободной энергии связывания с помощью LeadFinder; при этом в явном виде учитывались молекулы воды, участвующие в образовании мостиковых водородных связей (Рисунок 31). ΔG_0 связывания артокарпина с новым сайтом в структуре NanA рассчитывали как среднее из значений ΔG_0, полученных при докинге в ансамбль конформеров, что составило -7,08 ккал/моль. Для сравнения, ΔG_0 связывания артокарпина, рассчитанная из экспериментально определенной константы ингибирования K_I, равной 9,68 ± 1,49 μM при 37°C

[249], соответствует диапазону ΔG_0 связывания [-7,22; -7,02] ккал/моль (рассчитано по формуле из [232]).

С использованием трехмерного структурного наложения в PyMOL, артокарпин был введен в комплементарный сайт в структурах междоменных комплексов NanA для изучения его влияния на формирование межмолекулярных ансамблей. Во всех моделях комплексов лиганд, связанный с каталитическим доменом, приводил к стерическим конфликтам с аминокислотными остатками второго домена (лектинового или каталитического), из чего можно сделать предположение, что присутствие артокарпина в новом сайте может затруднять формирование межмолекулярного комплекса [201]. Для того чтобы проверить, будет ли такой комплекс стабильным в случае, если все же сформируется, даже несмотря на препятствия, созданные артокарпином, было использовано молекулярное моделирование. Десять случайно выбранных комплексов двух доменов с ингибитором, связанным с каталитическим доменом, прошли минимизацию энергии в NAMD [233] (где реализован более мощный алгоритм минимизации энергии), чтобы поспособствовать формированию локальных структурных перестроек, вызванных связыванием лиганда, и устранить стерические конфликты в межмолекулярном интерфейсе; затем комплексы прошли эквилибрацию и молекулярно-динамическую симуляцию длительностью 1000 нс с использованием AMBER (см. раздел 4.3.2). Присутствие артокарпина вызывало индуцированные лигандом конформационные изменения в межмолекулярном интерфейсе, что в конечном итоге инициировало диссоциацию 7 из 10 комплексов каталитического и лектинового доменов и 5 из 10 димеров каталитического домена. В трех и пяти случаях, соответственно, комплексы оставались стабильными несмотря на присутствие артокарпина, что, вероятно, можно объяснить чрезмерной оптимизацией NAMD, способствовавшей

образованию неестественной, избыточной комплементарности между доменами. В контрольном эксперименте, в котором использовались те же самые комплексы доменов, подвергнутые тем же процедурам, но уже без ингибитора, лишь два комплекса каталитического и лектинового доменов диссоциировали за тот же период времени, что, вероятно, можно объяснить изначально неоптимальной стыковкой доменов; все димеры каталитического домена оставались стабильными.

4.4. Визуализация

Для визуализации результатов биоинформатического анализа использовали интерактивные HTML5 инструменты, встроенные в соответствующие веб-серверы платформы Mustguseal [185]. Для просмотра множественных выравниваний использовали программное обеспечение JalView [234]. Для визуального анализа трехмерных структур макромолекул, в том числе, молекулярно-динамических траекторий и 3D-аннотаций результатов биоинформатического анализа, использовали программы PyMol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC; pymol.org) и VMD (Visual Molecular Dynamics; www.ks.uiuc.edu/Research/vmd/).

5. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

5.1. Подавление функциональных свойств нейраминидазы NanA ингибиторами межмолекулярной агрегации [iii]

5.1.1. Пространственная организация каталитического и лектинового доменов в структурах бактериальных нейраминидаз

Взаимодействие между доменами в NanB и NanC из *S. pneumoniae* было изучено с использованием молекулярного моделирования [200]. Длительные молекулярно-динамические симуляции продолжительностью 500 нс при 300 К не выявили каких-либо существенных структурных изменений – на протяжении всей траектории каталитический и лектиновый домены оставались плотно упакованными в жесткие, стабильные глобулы (Рисунок 19). Междоменные контакты в обоих ферментах представлены плотной сетью стабилизирующих взаимодействий (Рисунок 20 и Рисунок 21). Шесть водородных связей и три солевых мостика формируются между шестнадцатью аминокислотными остатки в лектиновом и каталитическом доменах NanB: Asp108, Arg129, Trp134, Gly135, Lys136, His137, Lys138, Glu169, Arg310, Asp643, Ile644, Asp645, Ser648, Ser682, Gln686, Asp694. Аналогично, семь водородных связей и три солевых мостика образуются между восемнадцатью аминокислотными остатками лектинового и каталитического доменов в NanC: Asp152, Arg173, Gly179, Lys180, Tyr181, Lys182, Lys213, Val272, Thr273, His275, Leu276, Glu356, Asp685, Asp687, Asn690, Lys724, Thr731, Asp736. Большинство этих взаимодействий были либо консервативны (например, солевой мостик между консервативными остатками

[iii] При подготовке данного раздела диссертации использованы следующие публикации, выполненная автором в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: **Sharapova Y.**, Suplatov D., Švedas V. Neuraminidase A from Streptococcus pneumoniae has a modular organization of catalytic and lectin domains separated by a flexible linker //The FEBS journal. – 2018. – Т. 285. – №. 13. – С. 2428-2445; **Sharapova Y.,** Švedas V., Suplatov D. Catalytic and lectin domains in neuraminidase A from Streptococcus pneumoniae are capable of an intermolecular assembly: Implications for biofilm formation //The FEBS Journal. – 2021. – Т. 288. – №. 10. – С. 3217-3230. Объем вклада автора диссертации в обоих случаях составляет 75% и заключался в задумке, планировании и выполнении экспериментов, а также написании статей.

Asp694 / Asp736 и Lys138 / Lys182) или эквивалентны (например, солевой мостик между противоположно заряженными группами Arg310 / Glu356 и Glu169 / Lys213, или водородная связь между гуанидиновой группой Arg129 / Arg173 и атомами главной цепи Ser682 / Lys724) в обоих ферментах. Карты поверхностей электростатического потенциала, рассчитанные с использованием подхода APBS, показали, что комплементарные области поверхностей контакта лектинового и каталитического доменов в NanB и NanC имеют тенденцию быть противоположно заряженными на периферии, а остатки ближе к центру интерфейса, как правило, гидрофобные (Рисунок 22 A, B). В дополнение к описанной выше роли полярных и заряженных остатков, гидрофобные остатки, расположенные в центре междоменного интерфейса, также способствуют стабильности глобул NanB и NanC.

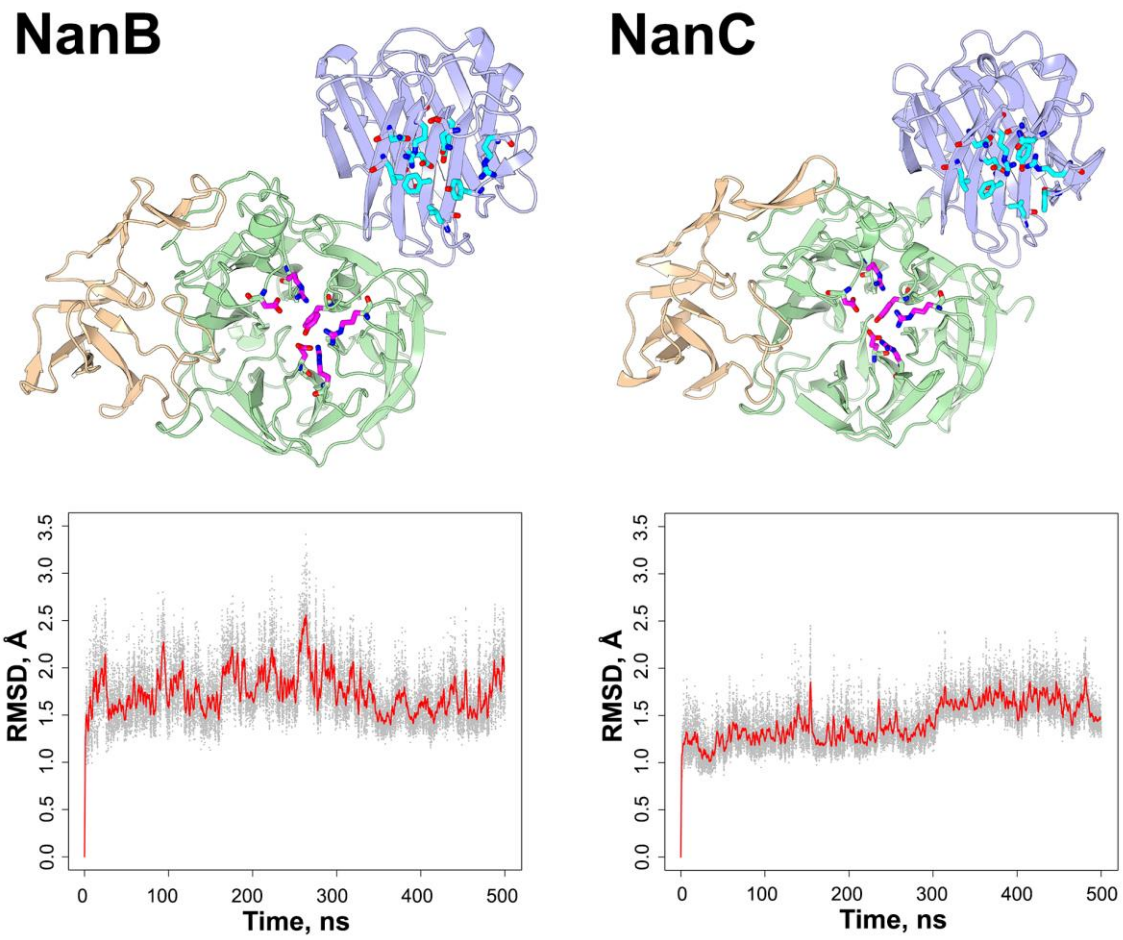

Рисунок 19. Структуры NanB и NanC из *S. pneumoniae* в течение 500 нс МД-симуляций при 300K. Лектиновый домен выделен голубым цветом, каталитический домен – зеленым, с инсерционным доменом, выделенным бежевым. Аминокислотные остатки в участке связывания субстрата лектинового домена окрашены в голубой цвет, а ключевые остатки активного центра каталитического домена окрашены в розовый цвет. Графики RMSD рассчитывались по атомам основной цепи обоих доменов; кадры из траектории белка были выровнены по каталитическому домену (включая инсерционный). Серые точки соответствуют фактическим значениям RMSD, красные линии - средним значениям, рассчитанным в пределах временного окна 2 нс. Изображения были подготовлены с использованием PyMol и языка R.

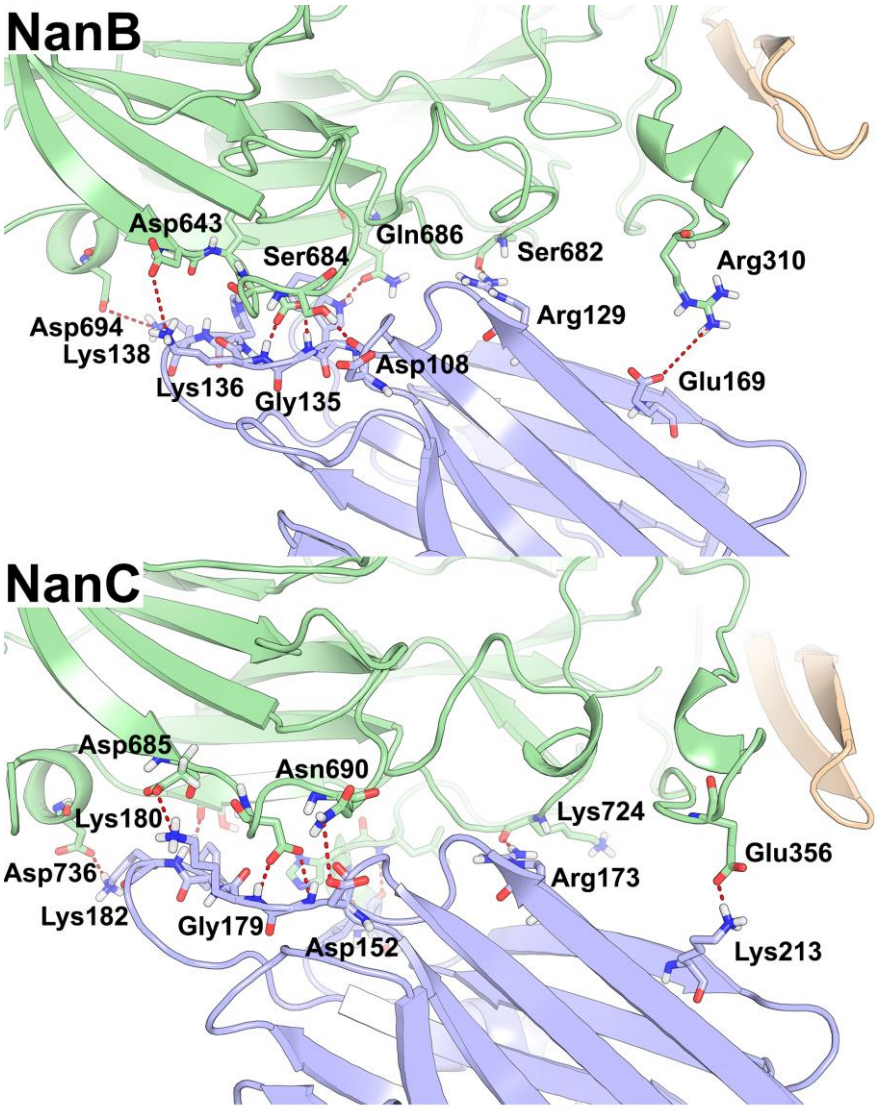

Рисунок 20. Водородные связи и солевые мостики между лектиновым и каталитическим доменами в структурах NanB и NanC из *S. pneumoniae*. Нумерация остатков дается в соответствии с файлами PDB 2JKB и 4YW2 соответственно. Лектиновый домен выделен голубым цветом, каталитический домен – зеленым, с инсерционным доменом, выделенным бежевым. Изображения были созданы с использованием PyMol.

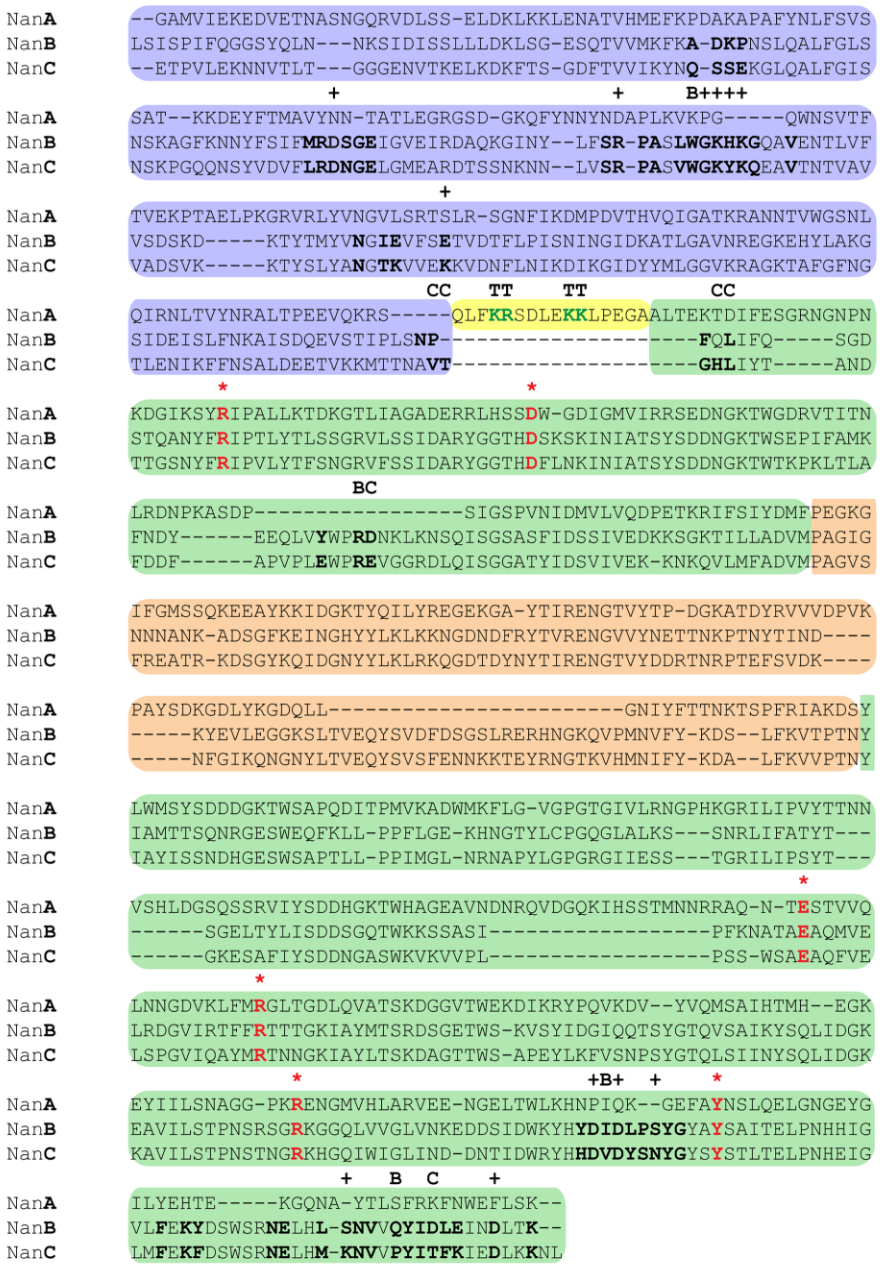

Рисунок 21. Структурно-опосредованное выравнивание последовательностей гомологичных NanA, NanB и NanC из *S. pneumoniae*. Лектиновый домен выделен голубым цветом, каталитический домен – зеленым, с инсерционным доменом, выделенным бежевым. Междоменный линкер в NanA выделен желтым. Знак «*» обозначает колонки с консервативными каталитически остатками. Аминокислоты, образующие поверхность контакта между лектиновым и каталитическим доменами в NanB и NanC, показаны жирными черными буквами. Знаки «+», «B» и «C» обозначают колонки с остатками, которые участвуют в образовании водородных связей или солевых мостиков между лектиновым и каталитическим доменами в NanB и NanC – в обоих белках, только в NanB, только в NanC соответственно. Знак «T» обозначает четыре потенциальных сайта трипсинолиза в последовательности линкера NanA, согласно предсказанию веб-сервера PeptideCutter с вероятностью расщепления в этих сайтах в пределах 82% -100%.

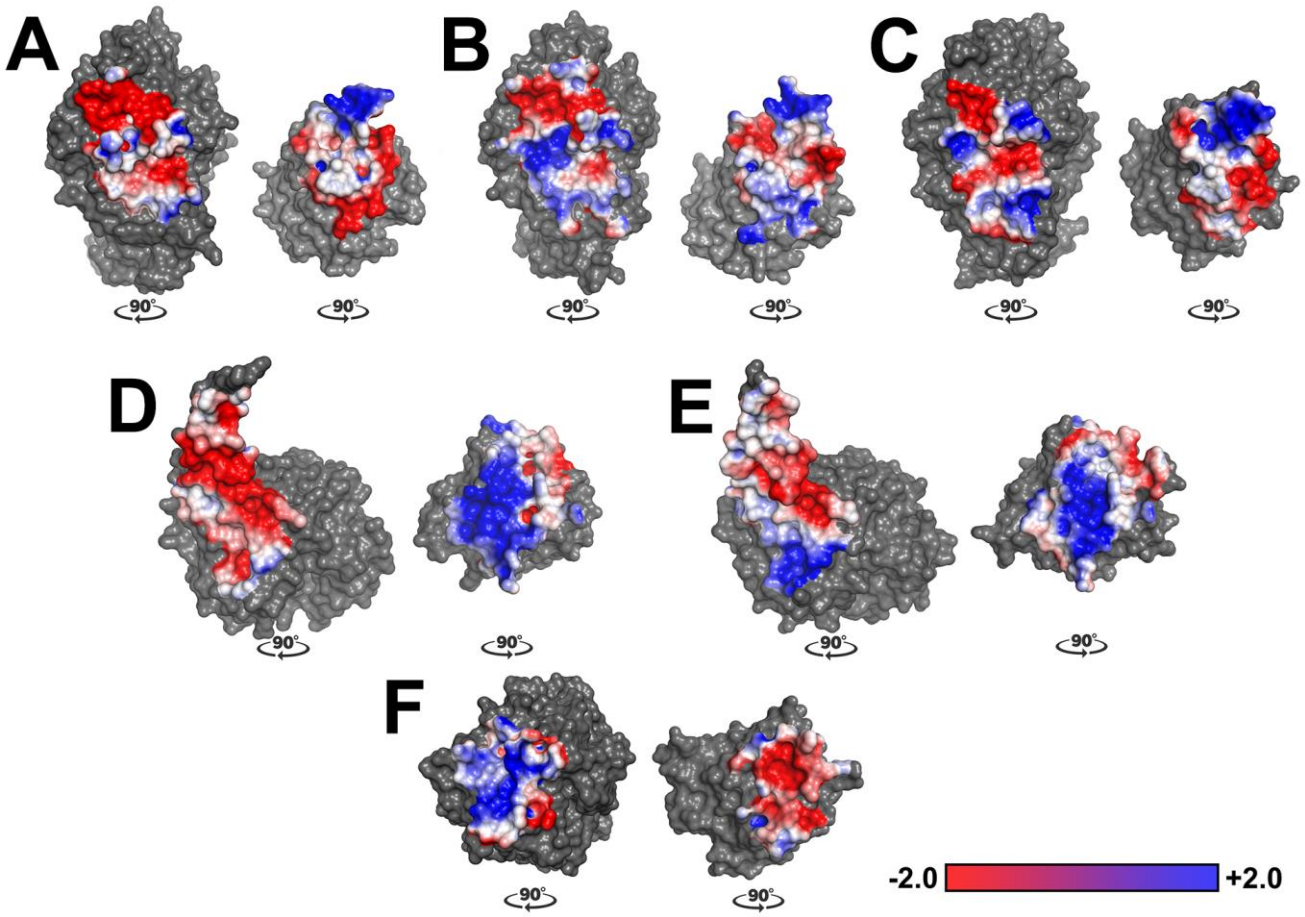

Рисунок 22. APBS карта электростатического потенциала поверхностей контакта между каталитическим (слева) и лектиновым (справа) доменами в (A) NanB из *S. pneumoniae*; (B) NanC из *S. pneumoniae*; (C) транс-сиалидазе из *M. decora*; (D) сиалидазе из *T. rangeli*; (E) сиалидазе из *T. cruzi*; (F) сиалидаза из *V. cholera* (показан интерфейс с N-концевым лектиновым доменом). Два домена были повернуты на 90º в противоположных направлениях, так, чтобы поверхности контакта можно было наблюдать в одной плоскости. Электростатический потенциал окрашен от красного (-2 kT/e) до синего (+2 kT/e). Изображения были созданы с использованием PyMol.

5.1.2. Молекулярное моделирование уникальной структуры NanA, в которой каталитический и лектиновый домены соединенны гибким линкером из 16 аминокислот

Все известные структуры сиалидаз, имеющие CBM40-лектиновый домен, кристаллизованный вместе с каталитическим доменом GH33, были отобраны с

помощью биоинформатического анализа (Рисунок 15, Таблица 1). Эти гомологичные белки сохраняют стабильную ориентацию доменов, упакованных в жесткие глобулы, о чем свидетельствуют длительные молекулярно-динамические симуляции (Рисунок 15, Рисунок 23). Поэтому эти гомологи использовались в качестве образцов при молекулярном моделировании, чтобы учесть все возможные ориентации лектинового домена на поверхности каталитического домена (вместе с инсерционным доменом здесь и далее) в NanA из *Streptococcus pneumoniae* [200]. Структурная организация в целом и ориентация доменов в транс-сиалидазе из *M. decora* была очень схожа с таковой у сиалидаз NanB и NanC из *S. pneumoniae*, с N-концевым лектиновым доменом, напрямую соединенным с каталитическим. Следовательно, гомологичные модели NanA, основанные на структурах NanB и NanC из *S. pneumoniae*, а также на структуре транс-сиалидазы из *M. decora*, были эквивалентны (Рисунок 16, A, B и C). Сиалидазы из *T. rangeli, T. cruzi* и *V. cholerae* оказались более далекими гомологами бактериальных NanA, NanB и NanC и обладают существенно отличающейся структурной организацией. Каталитический домен GH33 трипаносомных сиалидаз не содержит инсерционного домена; за каталитическим следует C-концевой лектиновый домен, ковалентно связанный с ним линкером (Рисунок 15, E и F). Линкер состоит из 25 аминокислотных остатков и частично свернут в α-спираль в кристаллографических структурах [215]. Классификация CAZy недоступна для лектиновых доменов трипаносомных сиалидаз, однако биоинформатический анализ выявил значительное структурное сходство с доменом CBM40 в NanA (Таблица 1). Положения C-концевого лектинового домена в сиалидазах из *T. rangeli* и *T. cruzi* были взяты в качестве основы для моделирования возможных положений N-концевого лектинового домена на поверхности каталитического домена в NanA,

так как длины линкера NanA было достаточно для ковалентного соединения двух подструктур (Рисунок 16, D и E).

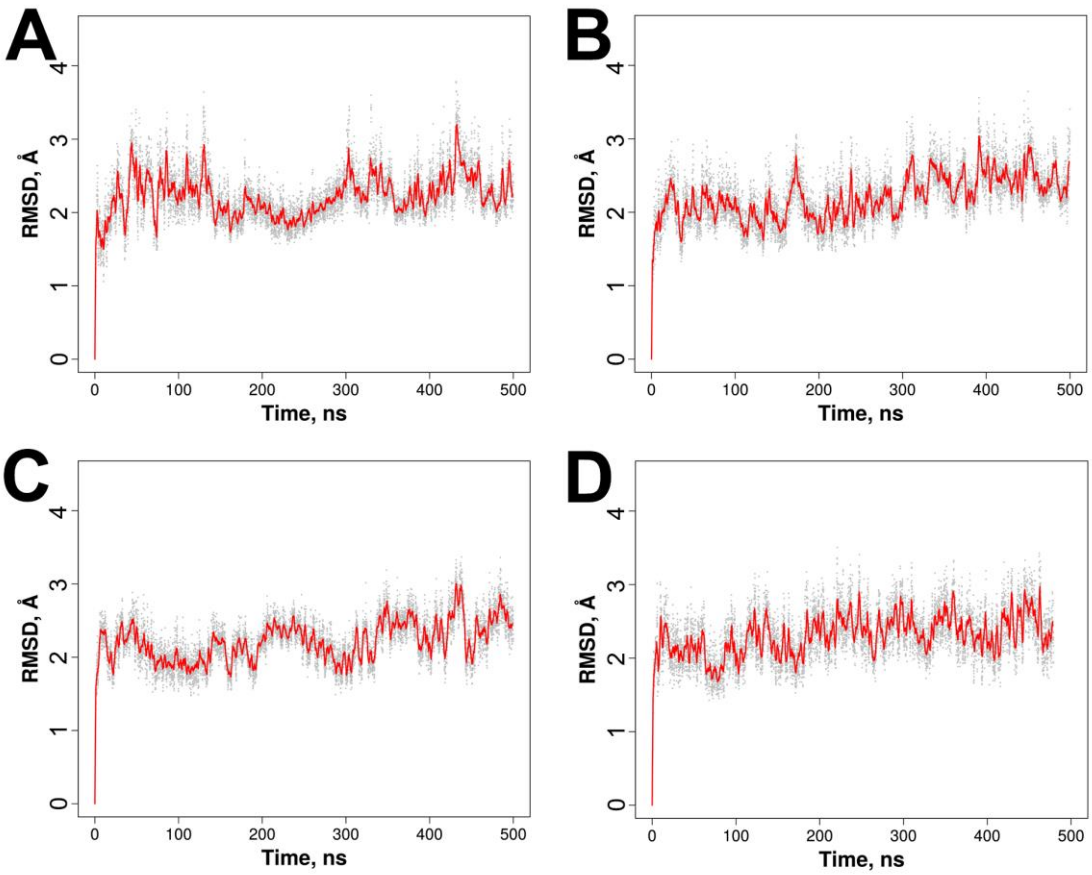

Рисунок 23. Графики RMSD (A) транс-сиалидазы из *M. decora*, (B) сиалидазы из *T. rangeli*, (C) сиалидазы из *T. cruzi* и (D) сиалидазы из *V. cholerae* в течение 500 нс МД-симуляции при 300K. RMSD рассчитывали по атомам основной цепи всех доменов; кадры траекторий были выровнены по каталитическому домену. Серые точки соответствуют фактическим значениям RMSD, красные линии - средним значениям, рассчитанным в пределах временного окна в 2 нс. Изображения были подготовлены с использованием языка R.

Мы также рассмотрели особый случай с сиалидазой из *V. cholerae*, которая содержит центральный каталитический домен GH33, фланкированный двумя CBM40-подобными лектиновыми доменами – один N-концевой, а другой вставлен в β-пропеллер каталитического домена (Рисунок 15, G). Предыдущие исследования этого белка показали, что N-концевой лектиновый домен может

связывать сиаловую кислоту, но не выявили роли инсерционного/вставочного лектинового домена [108]. Хотя два лектиновых домена в сиалидазе из *V. cholerae* структурно схожи, они идентичны по последовательности лишь на 23%, а ключевые остатки, взаимодействующие с лигандом в N-концевом лектиновом домене, во вставочном лектиновом домене либо замещены неэквивалентными остатками, либо смещены в пространстве по сравнению с правильными позициями для связывания лиганда. Поэтому маловероятно, что инсерционный лектиновый домен в этом белке может связывать сиаловую кислоту [108]. Тем не менее, мы попытались учесть позиции обоих лектиновых доменов в сиалидазе из *V. cholerae* при моделировании структуры NanA, но только вариант, основанный на положении N-концевого лектинового домена сиалидазы из *V. cholerae*, оказался возможным (Рисунок 16, F). При моделировании на основании ориентации инсерционного лектинового домена две субструктуры NanA было невозможно связать линкером, поэтому эта модель далее не рассматривалась. Также нами был проведен отдельный эксперимент белок-белкового докинга структурных конформеров лектинового и каталитического доменов NanA (всего было построено 720 комплексов доменов), однако мы не нашли никакого нового интерфейса взаимодействия доменов в пределах досягаемости линкера (см. раздел 4.3.3.2). Следовательно, далее изучались только гомологичные модели NanA.

Взаимодействие между доменами в гомологичных моделях NanA было изучено с использованием молекулярной динамики [200]. В течение молекулярно- динамических симуляций в структурах отдельных доменов не наблюдалось значительных изменений, о чем свидетельствуют плоские графики RMSD (Рисунок 24, A и B). Наиболее значительные изменения наблюдались в расположении доменов относительно друг друга: уже в течение первых 20 нс лектиновый домен начал удаляться от каталитического домена, и к концу

траекторий во всех моделях NanA две подструктуры оказались полностью разделенными в пространстве, оставаясь при этом ковалентно связанными междоменным линкером (Рисунок 24, C). Расчеты APBS также показали недостаточную электростатическую комплементарность потенциальных поверхностей контакта между лектиновым и каталитическим доменами во всех моделях NanA (Рисунок 25), в отличие от эффективных междоменных взаимодействий, наблюдаемых в гомологичных сиалидазах (Рисунок 22). Биоинформатическое сравнение сиалидаз *S. pneumoniae* подтвердило, что остатки, которые образуют стабилизирующие междоменные взаимодействия в NanB и NanC, в NanA либо замещены аминокислотами с другими физико-химическими свойствами, либо соответствующие участки структуры имеют другую организацию (Рисунок 21). В результате, структура NanA оказалась модульной, с лектиновым доменом, отделенным в пространстве от каталитического (включая инсерционный) гибким междоменным линкером длиной в 16 остатков. Модель междоменного пептидного линкера NanA была построена в PyMol, чтобы оценить его линейную длину. Расстояние между крайними атомами основной цепи достигает 52 Å. На протяжении МД-симуляций две разошедшиеся подструктуры NanA свободно двигались в растворителе и принимали различные ориентации друг относительно друга благодаря гибкости линкера. Следует отметить, что в течение 150 нс можно наблюдать, как лектиновый домен сначала удаляется от каталитического на всю длину линкера (рис. 10, A), а затем меняет направление движения, начиная снова приближаться к каталитическому домену. При этом обе субстрат-связывающие поверхности, находятся в одной плоскости (рис. 10, B), как и у других многодоменных сиалидаз (Рисунок 19). Биологическая роль этой структурной адаптации NanA как ключевого фактора вирулентности дополнительно обсуждается ниже.

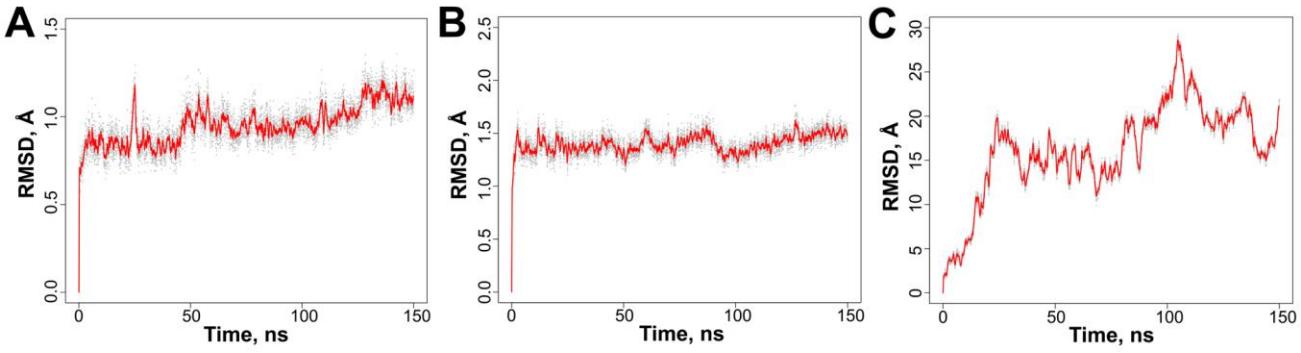

Рисунок 24. RMSD атомов основной цепи (A) лектинового и (B) каталитического доменов (включая инсерционный домен) в структуре NanA в течение 150 нс МД-симуляции при 300K. (C) RMSD всех атомов основной цепи обоих доменов в ковалентно связанной модели NanA; кадры траектории были выровнены по каталитическому домену (включая инсерционный). Серые точки соответствуют фактическим значениям RMSD, красные линии - средним значениям, рассчитанным в пределах временного окна 400 пс. Графики были подготовлены с использованием языка R. Графики относятся к гомологичной модели NanA, основанной на структуре NanB из *S. pneumoniae* (Рисунок 16, A). Результаты для других гомологичных моделей были качественно одинаковыми.

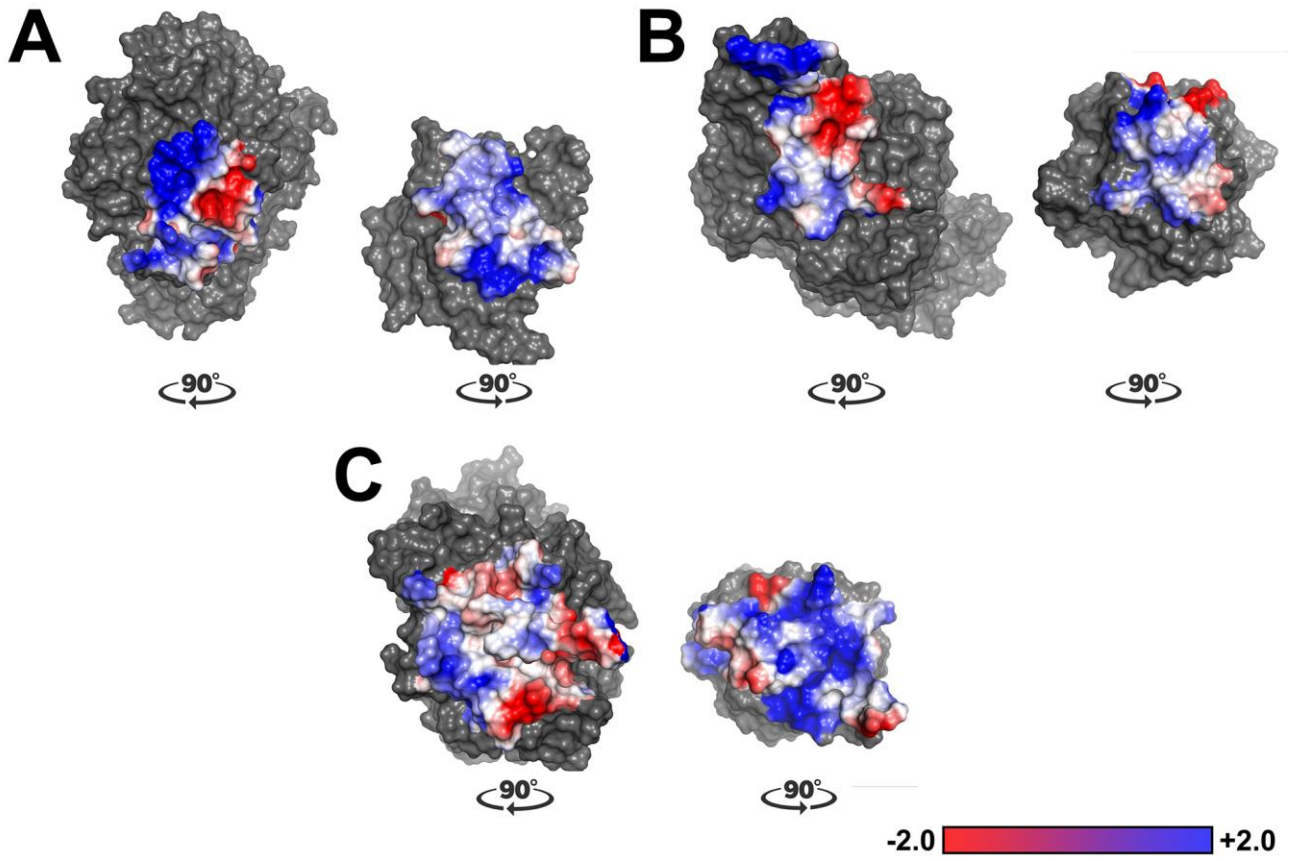

Рисунок 25. APBS карты электростатического потенциала поверхностей контакта между каталитическим (слева) и лектиновым (справа) доменами в исходных гомологичных моделях NanA из *S. pneumoniae* после минимизации энергии и эквилибрации, полученных на основе (A) NanB из *S. pneumoniae*; (B) сиалидазы из *T. rangeli*; (C) сиалидазы из *V. cholera* (N-концевого лектинового домена). Изображения APBS для моделей NanA, построенных по гомологии с NanC из *S. pneumoniae* и транс-сиалидазой из *M. decora*, были эквивалентны изображению (A); по гомологии с сиалидазой из *T. cruzi* – эквивалентны изображению (B). Два домена повернуты на 90º в противоположных направлениях, так, чтобы поверхности контакта можно было наблюдать в одной плоскости. Электростатический потенциал окрашен от красного (-2 kT/e) до синего (+2 kT/e). Изображения были созданы с использованием PyMol.

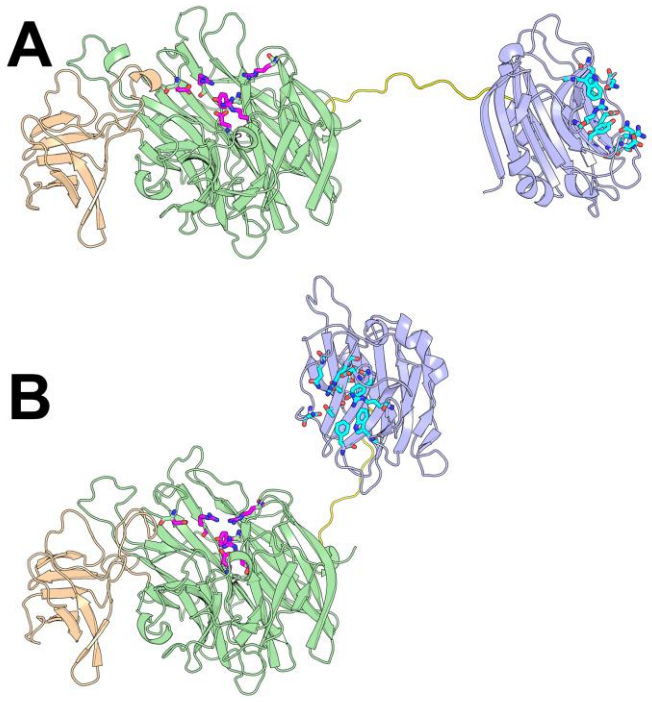

Рисунок 26. Ориентация доменов в структуре NanA, изученная методом молекулярной динамики. Изображения (A) и (B) соответствуют структурам, наблюдавшимся в течение 150 нс МД-траекторий при 300K. Лектиновый домен (синий) и каталитический домен (зеленый), включая инсерционный домен (бежевый), не образуют жесткую глобулу и пространственно обособлены гибким линкером (желтый). Аминокислотные остатки в участке связывания субстрата лектинового домена окрашены в голубой цвет (Phe167, Asp180, Tyr182, Thr193, Glu195, Arg197, Gln203, Asn209, Arg274 и Trp280, нумерация по PDB 4ZXK); ключевые остатки активного центра каталитического домена окрашены в розовый цвет (Arg332, Asp357, Glu632, Arg648, Arg706 и Tyr737, нумерация согласно PDB 2YA8). Изображения были созданы с использованием PyMol.

Гомологичные нейраминидазы NanA, NanB и NanC из *Streptococcus pneumoniae* принадлежат к суперсемейству сиалидаз. Представители этого суперсемейства из вирусов и патогенных бактерий являются ключевыми факторами вирулентности и основными мишенями при разработке лекарств [52, 53, 112, 167, 169, 181, 183, 235]. Бактериальная NanA содержит лектиновый и каталитический домены в составе единой полипептидной цепи, и оба считаются важными антимикробными мишенями. Кристаллографическая структура полноразмерного фермента могла бы способствовать поиску эффективных

ингибиторов NanA, однако попытки получить ее не увенчались успехом. В то же время для NanB и NanC доступны записи PDB, в которых каталитические и лектиновые домены являются частью глобулярных структур. Молекулярное моделирование и биоинформатика были использованы в этой работе для сравнительного исследования гомологичных нейраминидаз / сиалидаз. Было обнаружено существенное различие в пространственной организации лектинового и каталитического доменов в структурах гомологичных NanA, NanB и NanC. Междоменные контакты в NanB и NanC стабилизируются водородными связями, солевыми мостиками и гидрофобными взаимодействиями, приводящими к образованию жестких глобул. Напротив, моделирование показало, что в NanA нет эффективного взаимодействия между доменами, и лектиновый домен пространственно обособлен от каталитического гибким линкером длиной в 16 аминокислот. Вероятно, именно поэтому не удалось кристаллизовать полноразмерную NanA, а структуры отдельных доменов поддаются кристаллизации после трипсинолиза [100]. Действительно, междоменный линкер NanA содержит сайты расщепления трипсином согласно предсказанию веб-сервера PeptideCutter [236] (Рисунок 21).

Модульные ферменты, состоящие из нескольких компактных доменов, соединенных гибкими линкерами, иногда встречаются в природе. Такая структурная организация не очень распространена, но может быть обнаружена в тех белках, которым конформационная гибкость необходима для функции. Например, многие РНК-связывающие белки (например, РНК-редактирующий фермент ADAR2 и прокариотический фактор инициации трансляции IF3) обладают модульными структурами, обеспечивающими гибкость для функциональных доменов и позволяющими им взаимодействовать с конкретными сайтами в молекулах РНК различной длины и структуры [237, 238]. Большинство ферментов, участвующих в деградации клеточной стенки

растений, представляют собой модульные ферменты, состоящие из каталитического домена, присоединенного гибким линкером к связывающему целлюлозу домену (углевод-связывающий модуль). Наличие углевод-связывающего модуля (CBM) позволяет повысить эффективность биокаталитического гидролиза целлюлозы за счет увеличения эффективной концентрации фермента на поверхности субстрата [239, 240, 241]. Например, каталитический и углевод-связывающий модули целлюлазы семейства 7 из *Trichoderma reesei* связаны гибким длинным линкером из 27 аминокислотных остатков. CBM прикрепляет весь фермент к большой поверхности клеточной стенки, и линкер позволяет каталитическому домену искать доступные восстанавливающие концы целлюлозы в радиусе 80 Å от CBM [242]. Другой белок, домены которого подвижны друг относительно друга – нейраминидаза из непатогенной почвенной бактерии *Micromonospora viridifaciens*. В состав этого фермента входит связывающий галактозу домен, присоединенный к каталитическому домену через линкерный домен [243]. CBM этой нейраминидазы помогает захватить субстрат и доставить его в активный центр, расположенный в каталитическом домене, что требует движения доменов вокруг точки вращения, образованной парой соседних остатков глицина между каталитическим и линкерным доменами.

Известно, что лектиновый домен играет важную роль в функционировании бактериальных нейраминидаз. Подобно CBM в целлюлазах, лектиновый домен в нейраминидазах может притянуть фермент к поверхностям дыхательных путей, богатым сиаловой кислотой, приблизив, таким образом, каталитический домен к его субстрату [108]. Присутствие лектинового домена в структурах бактериальных нейраминидаз значительно повышает их каталитическую эффективность по отношению к поливалентным сиалозидам, что позволяет предположить, что каталитический и лектиновый домены

взаимодействуют одновременно с разными остатками сиаловой кислоты в составе разветвленных полисахаридных субстратов [107]. Структуры всех трех нейраминидаз *S. pneumoniae* содержат N-концевой лектиновый домен, за которым следует каталитический домен, и все три белка секретируются во внеклеточное пространство; но только в NanA каталитический домен прикрепляется к клеточной стенке бактерии дополнительным С-концевым доменом [244]. Данные о структурной подвижности, полученные методом молекулярной динамики, позволяют предположить, что лектиновый домен NanA может простираться в окружающую среду на всю длину междоменного линкера (Рисунок 26), чтобы захватить удаленный субстрат и, таким образом, зафиксировать каталитический домен у поверхностей, богатых сиаловой кислотой. Гибкий линкер, соединяющий N-концевой лектиновый домен с каталитическим, может являть собой пример эволюционной адаптации NanA с целью преодолеть ограниченную подвижность прикрепленного к поверхности бактериальной клетки каталитического домена и расширить субстратную специфичность фермента по отношению к большим разветвленным полисахаридным субстратам.

Кристаллизация многих мультидоменных белков представляет трудность, вероятно, потому, что линкеры, соединяющие домены, могут быть гибкими. Наиболее очевидным решением в таких случаях было кристаллизовать изолированные подструктуры. Последние достижения в области компьютерных технологий и программного обеспечения увеличили потенциал вычислительной биологии для изучения модульных белков и структурной подвижности [245, 246, 247]. В этой работе молекулярное моделирование на GPU-ускорителях было применено для симуляции большой молекулярной системы бактериальной NanA и изучения модульной организации ее доменов. В молекуле NanA, в отличие от родственных NanB и NanC, домены не формируют жесткой

глобулярной структуры и пространственно обособлены гибким междоменным линкером из 16 аминокислот [200]. Эта структурная адаптация может оказывать влияние на функцию и роль белка в патогенезе (см. раздел 5.1.4).

5.1.3. Биоинформатический анализ пространственной обособленности лектинового и каталитического доменов в гомологичных сиалидазах

С использованием метода Mustguseal была изучена распространенность такого явления как гибкий линкер среди гомологичных сиалидаз. Анализ сходства структур и последовательностей из публичных баз данных показал, что присутствие участка последовательности между каталитическим и лектиновым доменами, который может соответствовать гибкому междоменному линкеру, характерно, прежде всего, для NanA из различных штаммов *S. pneumoniae*. Похожая организация была обнаружена также в NanI и NanJ из *C. perfringens*, в которых данный участок состоит из 14 и 16 аминокислот, соответственно.

5.1.4. Моделирование межмолекулярной агрегации NanA

Суперкомпьютерное моделирование показало, что один каталитический и один лектиновый домены, а также два каталитических домена, принадлежащие двум разным молекулам NanA, могут агрегировать, взаимодействуя через ранее неохарактеризованный междоменный интерфейс (Рисунок 27), электростатически комплементарный (Рисунок 28) и стабилизированный сетью межмолекулярных водородных связей и солевых мостиков (Рисунок 29). В комплексе каталитического домена с лектиновым межмолекулярная поверхность контакта располагается по «левую» сторону от каталитической субъединицы (Рисунок 27), в отличие от гомологичных NanB и NanC из *S. pneumoniae*, у которых взаимодействующие друг с другом домены относятся к

одной и той же молекуле нейраминидазы, а поверхность контакта доменов находится по «правую» сторону каталитической субъединицы (Рисунок 19). При этом участок связывания лиганда в лектиновом домене и активный центр каталитической субъединицы ориентированы в одном направлении в межмолекулярном комплексе NanA (Рисунок 27, A). Наблюдаемая сонаправленная ориентация центров связывания в разных доменах также характерна для NanB и NanC [98, 106], а также других модульных сиалидаз [200, 248, 216], состоящих из доменов CBM40 и GH33 по классификации CAZy. Предполагается, что лектиновый домен в таком комплексе приводит каталитическую субъединицу в непосредственную близость к ее субстратам на участках поверхности дыхательных путей, богатых сиаловой кислотой [108], а также способствует связыванию поливалентных субстратов за счет одновременного взаимодействия с обоими доменами [107, 200].

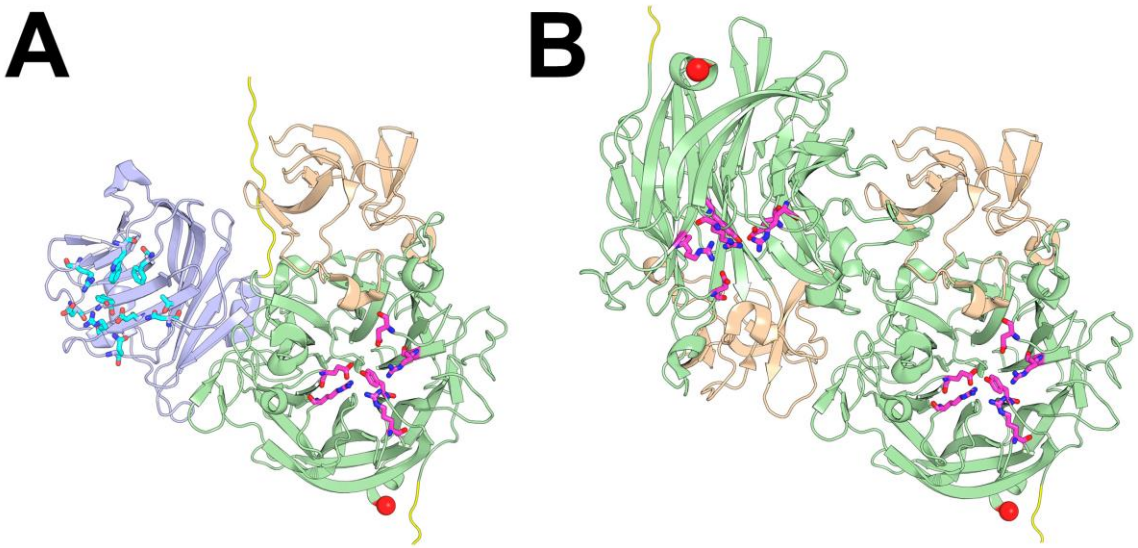

Рисунок 27. Пространственная организация димеров из (A) каталитического и лектинового и (B) двух каталитических доменов NanA из *S. pneumoniae* согласно результатам суперкомпьютерного моделирования (т. е. два домена принадлежат разным молекулам NanA). Лектиновый домен показан голубым цветом, каталитический домен – зеленым, с инсерционным доменом (бежевый), междоменный линкер в NanA – желтым. Ключевые аминокислотные остатки центров связывания субстрата показаны как на Рисунок 26. Изображения были созданы с использованием PyMOL.

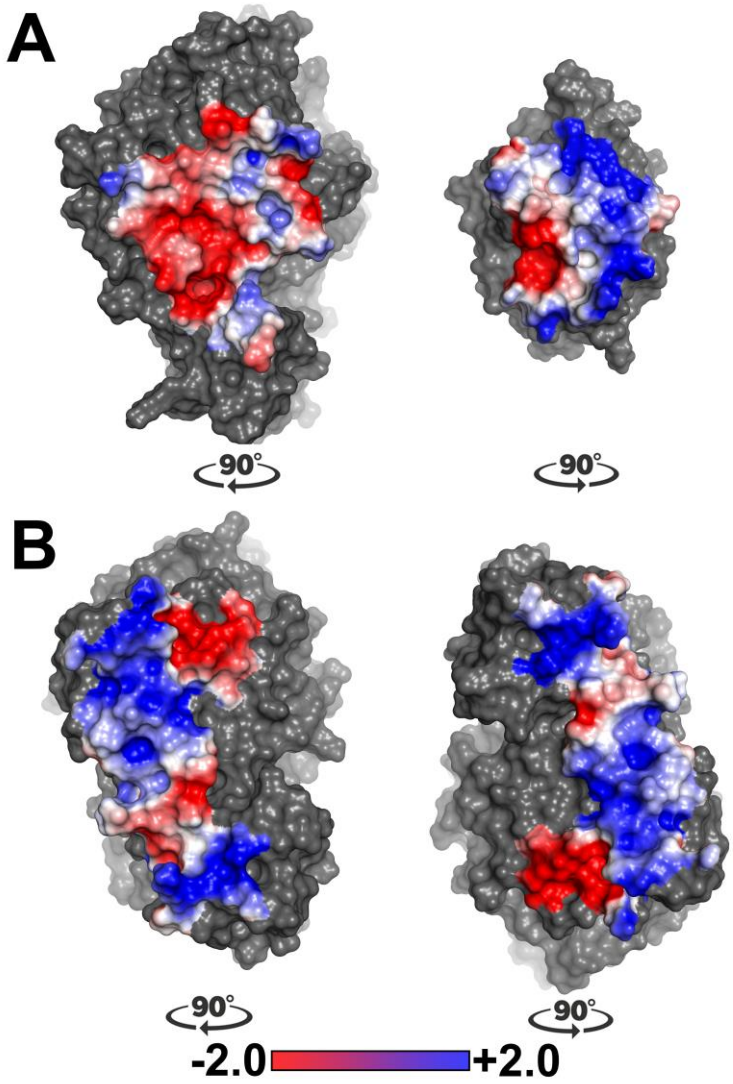

Рисунок 28. Электростатический потенциал поверхностей контакта между (A) каталитическим (слева) и лектиновым (справа) и (B) двумя каталитическими доменами в межмолекулярных комплексах NanA из *S. pneumoniae*. Два домена были повернуты на 90º в противоположных направлениях так, чтобы поверхности контакта можно было наблюдать в одной плоскости. Электростатический потенциал показан от красного (-2 kT/e) к синему (+2 kT/e). Изображения были созданы с использованием метода APBS в PyMol.

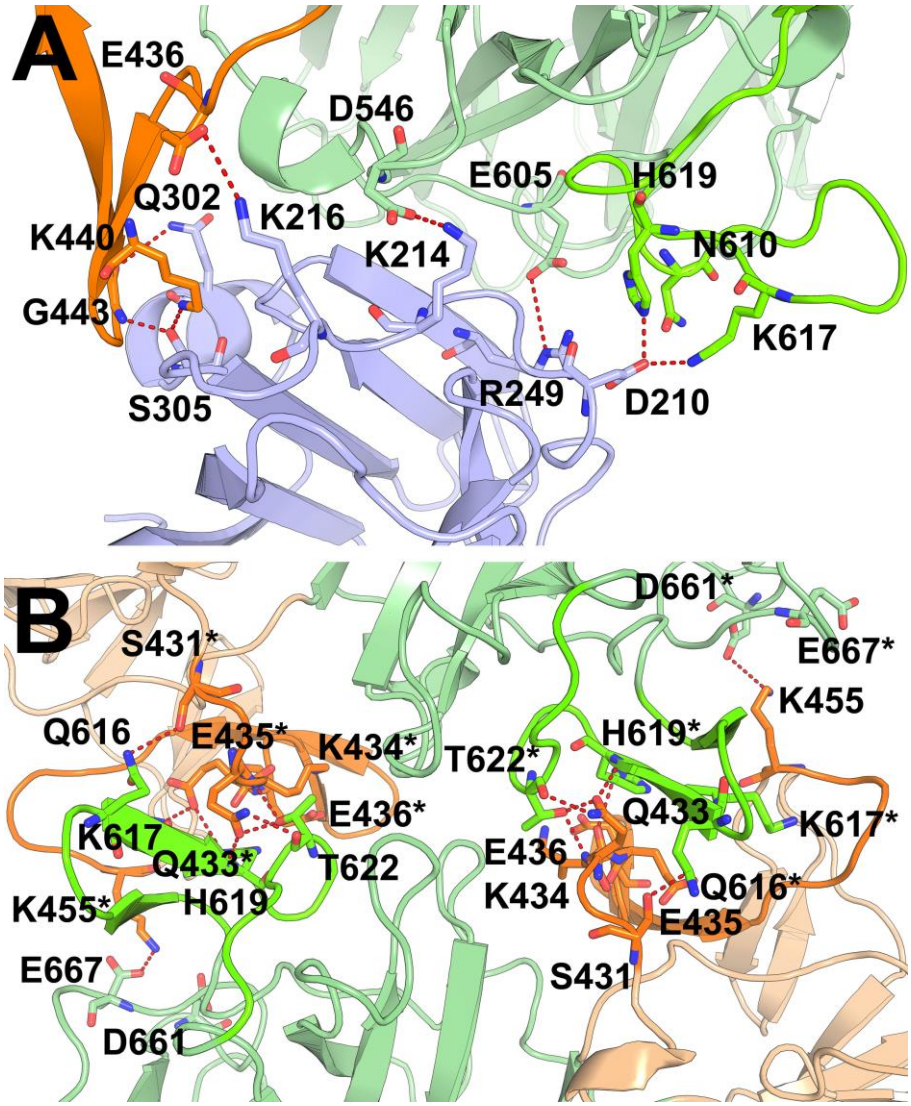

Рисунок 29. Водородные связи и солевые мостики в междоменном интерфейсе (A) каталитического и лектинового доменов и (B) димера каталитических доменов NanA из *S. pneumoniae*. Нумерация остатков приведена в соотвествии с PDB-файлами доменов (т. е. PDB 4ZXK и 2YA8). На нижнем рисунке (B) символом «*» обозначены аминокислотные остатки второй каталитической субъединицы. Лектиновый домен показан голубым цветом, каталитический домен – зеленым и неоновым зеленым, с инсерционным доменом (бежевый и оранжевый). Участок полипептидной цепи 431–455 и петля 609–626 выделены оранжевым и неоновым зеленым цветом, соответственно. Изображения были созданы с использованием PyMOL.

Карта поверхности электростатического потенциала, рассчитанная с использованием подхода APBS, показала, что комплементарные участки поверхностей контакта лектинового и каталитического доменов NanA противоположно заряжены (Рисунок 28, А). Молекулярно-динамическая симуляция показала, что структура межмолекулярного комплекса поддерживается сетью стабилизирующих взаимодействий (Рисунок 29, А): до шести солевых мостиков и пяти водородных связей образовалось между атомами боковых и основных цепей Asp210, Lys214, Lys216, Arg249, Gln302 и Ser305 лектинового домена (нумерация как в PDB 4ZXK) и Glu436, Lys440, Gly443, Asp546, Glu605, Asn610, Lys617 и His619 каталитического домена (нумерация как в PDB 2YA8). Второй наблюдавшийся тип взаимодействия молекул нейраминидазы А представлял собой комплекс двух каталитических доменов из двух молекул NanA (Рисунок 27, В). Такой димер был симметричным и обладал электростатической комплементарностью (Рисунок 28, В), поддерживаемой плотной сетью стабилизирующих взаимодействий (Рисунок 29, В): до пяти солевых мостиков и десяти водородных связей образовалось между атомами боковых и основных цепей Ser431, Gln433, Lys434, Glu435, Glu436, Lys455, Gln616, Lys617, His619, Thr622, Asp661 и Glu667 обеих субъединиц. Хотя образование комплексов доменов NanA ранее не обсуждалась в литературе, описанный в этой работе димер каталитического домена явно присутствует во всех доступных кристаллографических структурах этого белка, опубликованных независимыми лабораториями в период с 2008 по 2017 годы (а именно цепи А и В в PDB-записях 2VVZ, 2W20, 2YA4, 2YA5, 2YA6, 2YA7, 2YA8, 3H71, 3H72, 3H73 и 5KKY). Охарактеризованные димеры каталитического и лектинового и двух каталитических доменов имеют определенное сходство с точки зрения топологии доменов и участвующих в димеризации элементов структуры: в обоих типах межмолекулярных комплексов NanA интерфейс взаимодействия доменов

расположен по «левую» сторону каталитической субъединицы (Рисунок 27), а в димеризации в обоих случаях принимают участие регион 431–455 и петля 609–626, и, в частности, остатки Glu436, Lys617 и His619, непосредственно участвующие в образовании сети стабилизирующих взаимодействий в обоих комплексах (нумерация как в PDB 2YA8; Рисунок 29). Описанные выше интерфейсы димеризации являются специфической характеристикой, присущей NanA, но не NanB и NanC, поскольку соответствующие аминокислотные остатки в обоих комплексах либо расположены в петле 609–626, наблюдаемой только в структуре NanA, либо заменены с заряженных типов аминокислот на полярные или гидрофобные в NanB и NanC (Рисунок 30). Также стоит отметить, что наблюдаемое взаимодействие между каталитическим и лектиновым доменами невозможно в пределах одной молекулы NanA, поскольку этого не позволяет длина междоменного линкера [200].

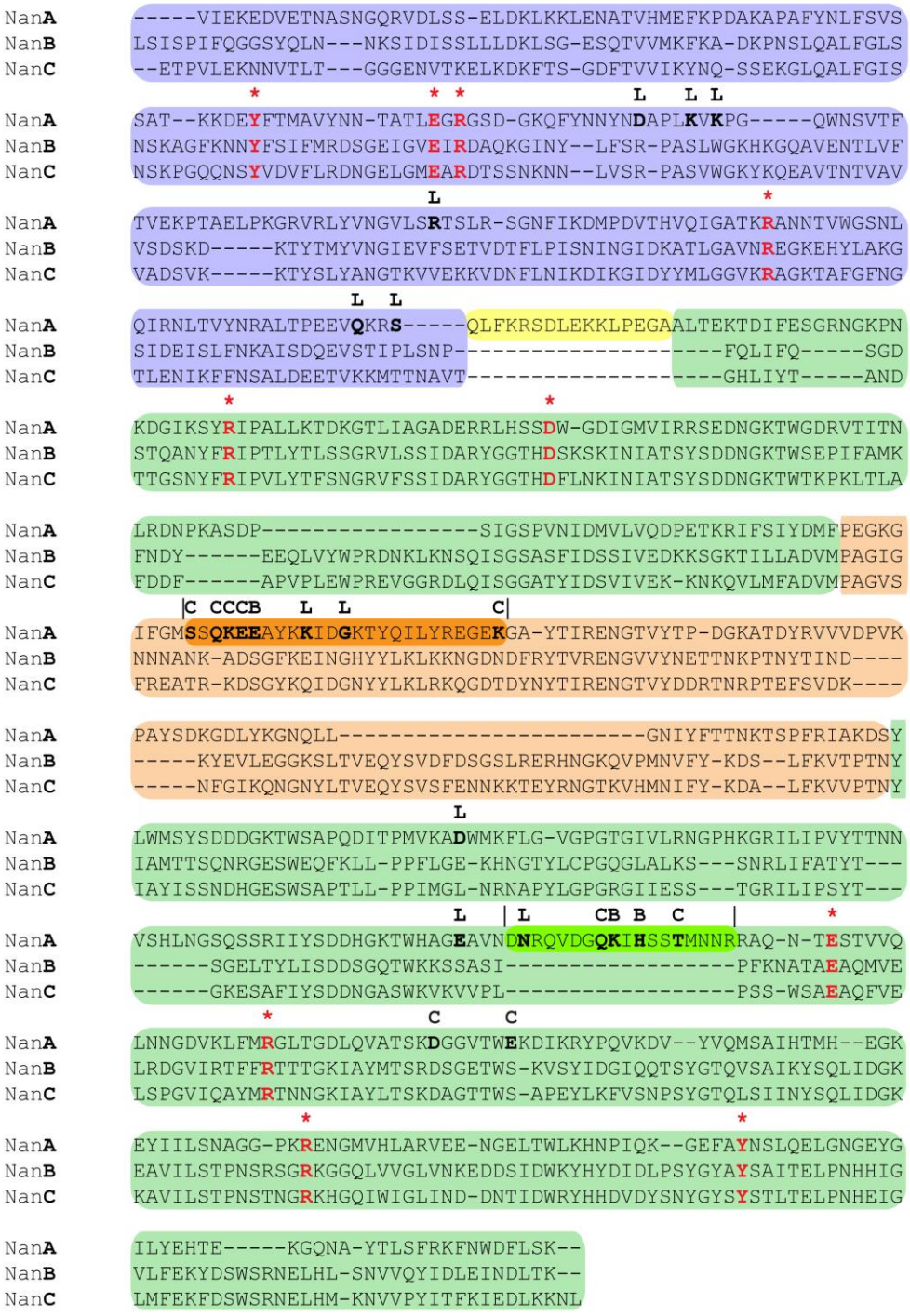

Рисунок 30. Структурно-опосредованное выравнивание последовательностей гомологичных NanA, NanB и NanC из *S. pneumoniae*. Было выполнено пространственное наложение PDB-структур 4ZXK, 2YA8, 2JKB и 4YW2, выравнивание аминокислотных последовательностей было получено из структурного выравнивания программой parMATT [230]. Последовательность гибкого междоменного линкера в NanA (16 а.о. между лектиновым и каталитическим доменами) добавлена вручную из UniProt P62575 и выделена на рисунке желтым цветом. Лектиновый

домен выделен голубым цветом, каталитический домен – зеленым, с инсерционным доменом, выделенным бежевым. С-концевая часть NanA, прикрепляющая белок к клеточной стенке, отсутствует в PDB-файлах и не показана на рисунке. Аминокислоты, образующие поверхность контакта между двумя доменами в межмолекулярных комплексах NanA, показаны жирными черными буквами и отмечены знаками: «L» – только в комплексе лектинового домена с каталитическим; «C» – только в димере двух каталитических доменов; «B» – в обоих комплексах. Участок 431-455 и петля 609-626 (см. объяснения в тексте) выделены оранжевым и неоновым зеленым, соответственно, и дополнительно отмаркированы знаками «|». Знак «*» обозначает колонки с консервативными каталитически остатками каталитического домена и консервативными остатками субстрат-связывающего центра лектинового домена.

5.1.5. Моделирование взаимодействия ингибитора артокарпина с новым аллостерическим центром межмолекулярной агрегации

Для установления возможной биологической роли межмолекулярной агрегации NanA было изучено влияние изопренилированного флавона артокарпина – экспериментально обнаруженного ингибитора каталитической активности NanA, также оказывающего ингибирующее действие на образование пневмококковой биопленки [249]. Артокарпин, согласно кинетическим исследованиям, ингибитор смешанного типа. Исходя из этого, ранее было высказано предположение о наличии альтернативного (т.е. аллостерического) сайта связывания для артокарпина, отличного от каталитического центра; однако точное место связывания лиганда в структуре NanA и механизм его действия оставались неизвестными [249]. Проведенное нами молекулярное моделирование показало, что артокарпин может связываться в сайте, расположенном между каталитическим и инсерционным доменами (Рисунок 31): дигидроксифенильный фрагмент лиганда образует две водородные связи с атомами кислорода основной цепи Gln433 и Glu435; метокси-группа и гетероциклический атом кислорода в бензопирановой части координируются сетью мостиковых водородных связей, образованных молекулами воды, атомами основной цепи остатков Ala437, Lys544 и Ala545 и атомами боковых цепей Asp518 и Lys544; изопреновые и метокси-

группы артокарпина образуют гидрофобные взаимодействия с гидрофобными частями Tyr438, Thr445, Lys440, Pro541, Lys434 и Lys544; бензопирановое и дигидроксифенильное кольца взаимодействуют с алифатическими фрагментами Glu436, Asp546 и Lys549. Оценка свободной энергии Гиббса связывания артокарпина с новым сайтом в структуре каталитического домена NanA составила -7,08 ккал/моль, что эквивалентно экспериментально определенному значению -7,12 ккал/моль. Артокарпин, связанный с каталитическим доменом в вышеобозначенном сайте, физически пересекается либо с лектиновым, либо с каталитическим доменом в межмолекулярных комплексах NanA (Рисунок 32). Связывание лиганда вызывает локальную структурную перестройку и изменение ориентации аминокислотных остатков двух доменов, что приводит к возникновению стерических конфликтов в межмолекулярном интерфейсе и увеличивает гидрофобность электростатически комплементарных поверхностей. Эти изменения препятствуют образованию стабильного межмолекулярного комплекса. Длительные МД-симуляции показали, что даже если два домена агрегируют в присутствии артокарпина, такой комплекс становится менее стабильным из-за нарушения межмолекулярных связей и снижения электростатической комплементарности интерфейса, что в конечном итоге инициирует разделение межмолекулярного комплекса на субструктуры отдельных доменов в течение 300-900 нс симуляции (Рисунок 33).

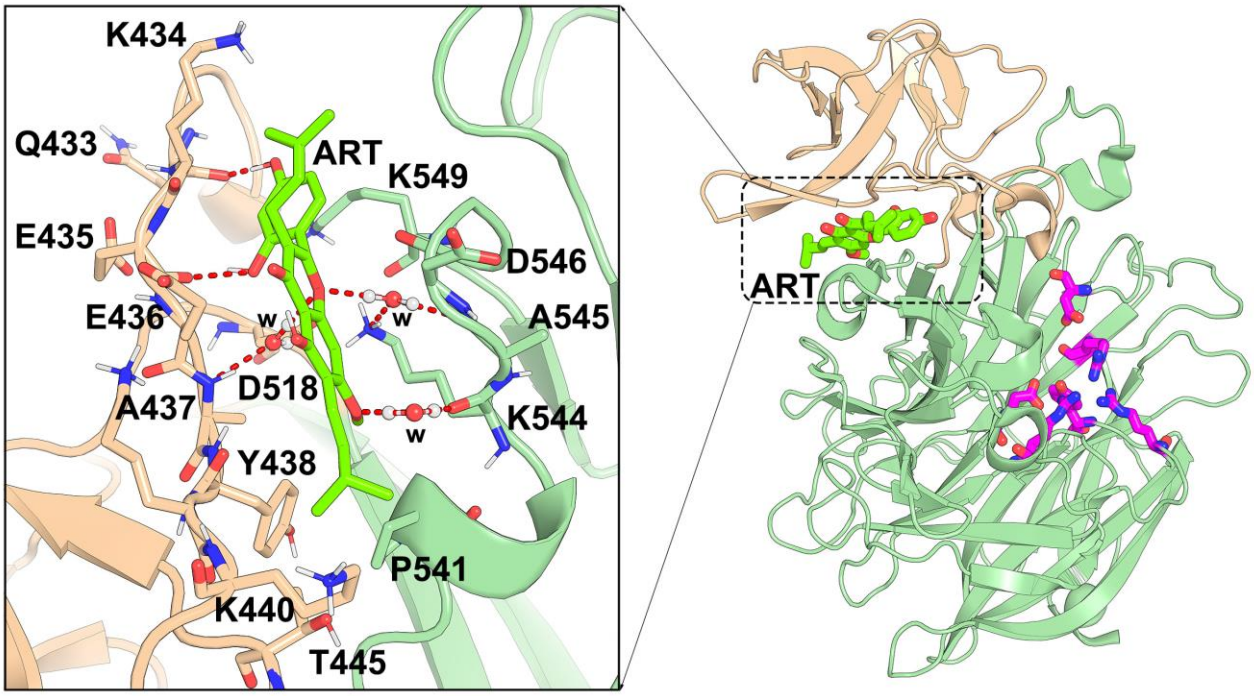

Рисунок 31. Сайт связывания артокарпина (ART) в структуре каталитического домена (включая инсерционный домен) NanA. Нумерация остатков дана в соответствии с PDB 2YA8. Знаки «w» отмечают расположение молекул воды, участвующих в мостиковых водородных связях. Для наглядности, ключевые остатки активного центра каталитического домена показаны в виде палочек пурпурного цвета (как на Рисунок 27). Каталитический домен показан зеленым цветом, инсерционный домен – бежевым. Изображения были созданы с использованием PyMOL.

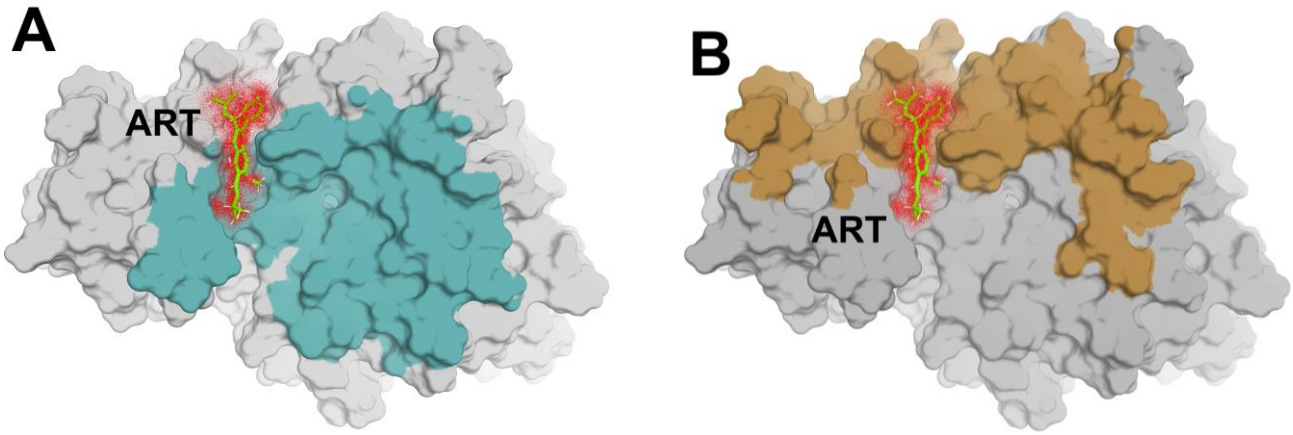

Рисунок 32. Центр связывания артокарпина в структуре каталитического домена NanA (серый). Участки поверхности каталитического домена, комплементарные (A) лектиновому домену и (B) каталитическому домену другой молекулы NanA в межмолекулярных комплексах, окрашены в бирюзовый и песочный оттенки, соответственно. Референсное положение артокарпина (ART) показано зелеными палочками; красные точки соответствуют фактическому положению атомов лиганда в течение 1000 нс МД-траектории. Изображения были созданы в PyMOL.

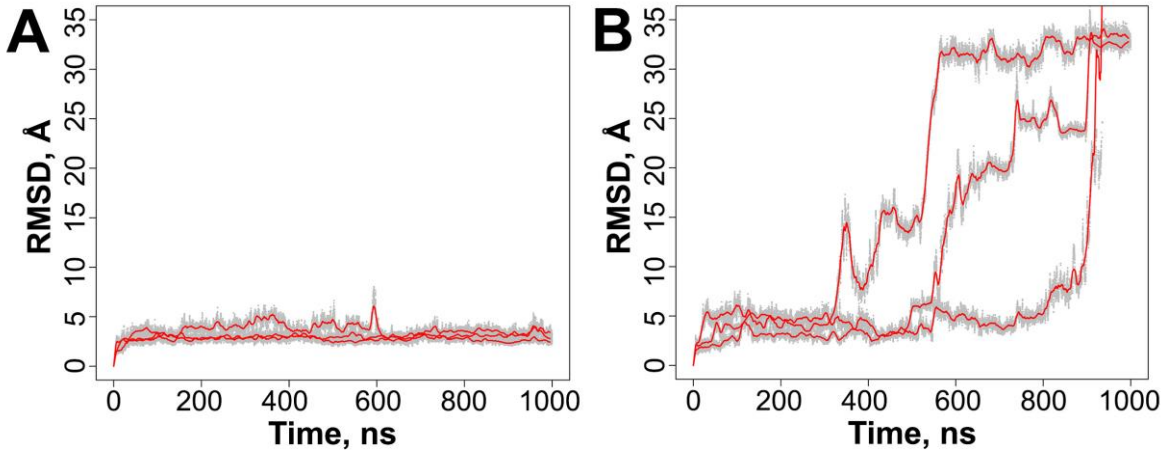

Рисунок 33. RMSD атомов основной цепи межмолекулярных комплексов NanA из *S. pneumoniae*, содержащих один лектиновый и один каталитический домен, в течение 1000 нс МД-симуляции при 310 К: (A) комплекс доменов и (B) комплекс доменов с артокарпином, связанным в комплементарном сайте в структуре каталитического домена. Показаны три МД-траектории комплексов с ингибитором и без него. RMSD рассчитывали по атомам основной цепи всех остатков; кадры белковой структуры, взятые из траектории, выравнивали по каталитическому домену. Серые точки соответствуют фактическим значениям RMSD; красные линии соответствуют средним значениям, рассчитанным во временном окне 10 нс. Графики подготовлены с использованием языка R.

5.1.6. Роль конформационной подвижности и межмолекулярной агрегации NanA в формировании биопленок пневмококка

Это исследование представляет собой первую попытку изучить на молекулярном уровне агрегацию каталитических и лектиновых доменов из разных молекул нейраминидазы A из *S. pneumoniae*. Систематическое использование суперкомпьютерного моделирования позволило описать ранее неохарактеризованный междоменный интерфейс. Он отличается высокой степенью электростатической комплементарности, которая способствует образованию множества солевых мостиков и водородных связей между двумя каталитическими или одним каталитическим и одним лектиновым доменами, принадлежащим разным молекулам NanA. Нейраминидаза A является поверхностным белком: за ее каталитическим доменом следует С-концевая часть

длиной 244 а.о. с LPXTG-мотивом, который прикрепляет фермент к клеточной стенке бактерии. Ранее проведенный биоинформатический анализ показал, что эта область молекулы наименее консервативна, и потому, вероятно, существует в виде множества структур, включая очень гибкие и/или неупорядоченные структуры, обеспечивающие гибкое прикрепление модульного фермента к поверхности бактерии [97, 103, 250, 251]. Взаимодействие нескольких молекул NanA, прикрепленных длинным гибким участком их полипептидных цепей к разным клеткам пневмококка, в принципе, может приводить к образованию супрамолекулярных комплексов (Рисунок 34). Поскольку участие NanA в формировании пневмококковой биопленки было описано в литературных источниках [116, 121, 125], мы попытались дополнительно изучить возможную роль межмолекулярной агрегации NanA в патогенезе.

Колонизация верхних дыхательных путей человека пневмококком подразумевает начальное прикрепление бактерий к рецепторам респираторного эпителия с последующим образованием биопленки. Учитывая важность этого процесса, прилагаются значительные усилия для его изучения. Первые прикрепившиеся к эпителию бактериальные клетки начинают делиться и инициируют систему «чувства кворума» для привлечения планктонных (т.е. еще не прикрепленных к рецепторам) пневмококков. Такие микроколонии начинают секрецию внеклеточного матрикса, состоящего из полисахаридов, нуклеиновых кислот, белков и липидов, что приводит к образованию зрелой биопленки – сложной толстой структуры, содержащей пустоты и каналы, чтобы пневмококки могли обмениваться сигналами, питательными веществами и генетическим материалом. Образование биопленки имеет решающее значение как при бессимптомной колонизации, так и при инвазивном заболевании. Она защищает бактерию от противомикробных агентов, иммунного ответа и способствует ее выживанию и распространению в другие части дыхательного тракта, что создает

значительную проблему для лечения инфекции (подробнее см. в главе « 3.3.1. Патогенез Streptococcus pneumoniae: колонизация *и* инвазия »).

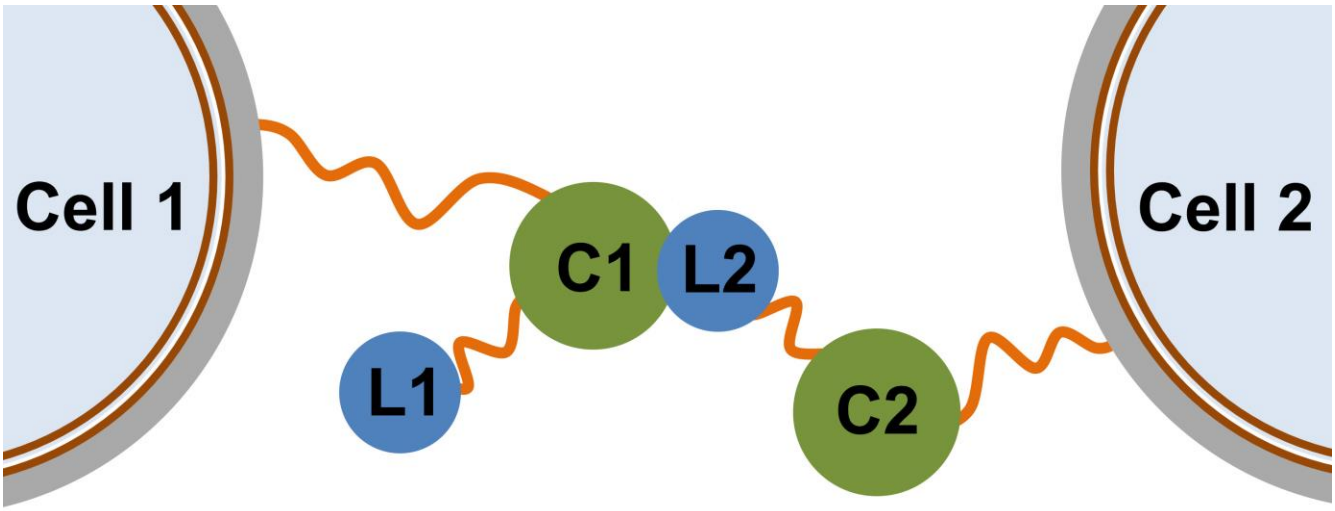

Рисунок 34. Схема межмолекулярного комплекса двух молекул NanA, взаимодействующих через интерфейс между каталитическим и, как пример, лектиновым доменами и прикрепленных к разным клеткам *S. pneumoniae*. Каталитический и лектиновый домены обозначены буквами «C» и «L» и пронумерованы в соответствии с тем, к какой клетке они принадлежат. Оранжевыми кривыми изображены С-концевой домен и гибкий междоменный линкер.

Недавние исследования свидетельствуют об участии нейраминидазы А в колонизации организма-хозяина и формировании биопленок пневмококком. Показано, что экспрессия фермента повышается в клетках *S. pneumoniae*, растущих в условиях биопленки, по сравнению с планктонным пневмококком [116, 123, 125]. Дефектный по NanA *S. pneumoniae* оказался неспособным образовать зрелую биопленку, хотя наблюдался начальный слой прикрепившихся бактерий, что позволило предположить, что для дальнейшего прикрепления планктонных бактерий и эффективного образования толстых, многослойных микроколоний внутри биопленки требуется NanA [121]. Хотя сущность этого явления все еще плохо изучена, были предложены две общие стратегии участия NanA в формировании биопленки – каталитическая и некаталитическая. Многочисленные исследования показали, что ферментативная активность NanA

важна для начальной адгезии бактерий к эпителиальным клеткам [115, 116, 118]. Сиаловая кислота – продукт каталитического превращения олигосахаридных рецепторов в активном центре NanA, потребляется бактерией в качестве источника питания, а также служит химическим сигналом для усиления адгезии и роста пневмококка и способствует образованию биопленки [119, 122]. Сиаловая кислота также индуцирует экспрессию гена *nanA* [122]. Последние исследования выявили альтернативную роль NanA в формировании биопленок, которая не зависит от ее каталитической функции. Экспрессия пневмококковой NanA в *Lactococcus lactis* – непатогенной бактерии, которая используется в производстве молочных продуктов и не образует биопленку, способствовала образованию плотного слоя клеток, организованных в микроколонии. Добавление ингибитора каталитической активности NanA в среду не повлияло на наблюдаемую сложную многоклеточную структуру [116]. В аналогичном эксперименте образование биопленок было вызвано экспрессией NanA в *Streptococcus gordonii*, который лишен нейраминидазной активности. Было высказано предположение, что в условиях биопленки нейраминидаза А может действовать как адгезин [116]. В другом исследовании было продемонстрировано, что пневмококк, дефектный по NanA, не способен колонизировать носовую полость мышей, а также вызвать инфекцию среднего уха; коинфекции мышей вирусом гриппа А, который также обладает нейраминидазной активностью, оказалось недостаточно для того, чтобы полностью компенсировать отсутствие NanA у бактерий, что позволило предположить, что NanA участвует в формировании пневмококковой биопленки независимо от своей ферментативной активности [125]. В том же исследовании наблюдалось значительное (в 3,25 раза) снижение жизнеспособности биопленок, образованных NanA-дефектным *S. pneumoniae*, растущим на абиотической полистирольной поверхности в среде YSK, где сиаловая кислота отсутствует вовсе, по сравнению со штаммом дикого типа. Наконец, ингибитор кверцетин

препятствует образованию пневмококковой биопленки, напрямую блокируя закрепление NanA в клеточной стенке бактерии [179].

Низкомолекулярное природное соединение артокарпин, ингибирующее каталитическую активность NanA, продемонстрировало способность предотвращать образование пневмококковой биопленки и препятствовать адгезии бактерий к легочным эпителиальным клеткам, а также снижало жизнеспособность пневмококков более чем в 1000 раз, не причиняя очевидного вреда эпителию, в связи с чем артокарпин представляется перспективным прототипом новых лекарств [249]. Оставались неясными, однако, молекулярные механизмы воздействия артокарпина на бактериальные клетки и точное влияние ингибирования NanA на образование биопленки. Кинетические исследования предположили наличие альтернативного (т.е. аллостерического) сайта связывания артокарпина в структуре NanA, отличного от каталитического центра [249]. В рамках этой работы современные методы вычислительной биологии были использованы для изучения связывания артокарпина с нейраминидазой А. Комплементарный сайт был обнаружен в той части структуры каталитического домена NanA, которая принадлежит интерфейсу связывания лектинового или каталитического домена другой молекулы NanA – артокарпин, таким образом, напрямую препятствует межмолекулярной агрегации доменов, что объясняет, как это низкомолекулярное соединение подавляет образование биопленок *S. pneumoniae*. Артокарпин ингибирует также и каталитическую активность NanA [249]. Мы полагаем, что причиной наблюдаемого влияния на каталитическую функцию фермента является аллостерическая коммуникация между предложенным сайтом связывания артокарпина и каталитическим центром через гибкую петлю 425–437, которая соединяет две полости (Рисунок 35). Эта петля способна принимать различные ориентации, и, как предполагается на основании анализа кристаллографических структур и молекулярного моделирования,

участвует в процессе доставки субстратов в активный центр NanA [101, 223]. Локальные изменения в подвижности/динамике структуры при связывании регуляторного лиганда являются одними из наиболее распространенных механистических паттернов, лежащих в основе явления аллостерии в белках [227, 252]. Связывание артокарпина с комплементарным участком, возможно, приводит к смещению петли 425–437, что в конечном итоге препятствует продуктивному связыванию субстратов в активном центре. Точная последовательность структурных событий, которые представляют собой аллостерическую коммуникацию между двумя сайтами в структуре NanA, остается неясной и требует дальнейшего изучения.

Формирование биопленок – это сложный многоступенчатый процесс, в котором микроорганизм задействует избыточные механизмы, чтобы совершенно точно обеспечить успешную колонизацию [253, 254]. Эта избыточность является основным препятствием при выявлении факторов патогенности, изучении их роли в заболевании и разработке новых терапевтических агентов для борьбы с инфекцией [83]. Экспериментальные методы все еще ограничены в их возможности выявить конкретные механизмы образования биопленки. Развитие современной вычислительной структурной биологии, подкрепляемое постоянно растущими суперкомпьютерными мощностями, позволяет достичь такого уровня детализации в понимании механизмов действия белков, которого трудно или даже невозможно достичь в экспериментах [255], и мы использовали ее для изучения уникальной структурной организации нейраминидазы А из *S. pneumoniae*. Суперкомпьютерное моделирование показало, что два каталитических домена, а также один каталитический и один лектиновый домен, принадлежащие разным молекулам NanA, связанным с поверхностью соседних бактериальных клеток, способны образовать супрамолекулярный комплекс. Это, вероятно, представляет собой механизм некаталитического участия этого потенциально ключевого в

патогенезе белка в образовании биопленки. Эти результаты расширяют наше понимание структурного разнообразия модульных сиалидаз/нейраминидаз и могут поспособствовать более глубоким экспериментальным исследованиям этих ферментов. Дальнейшие исследования сайтов связывания в структуре NanA, их особенностей и влияния на функцию белка необходимы для облегчения поиска селективных модуляторов его патогенной роли.

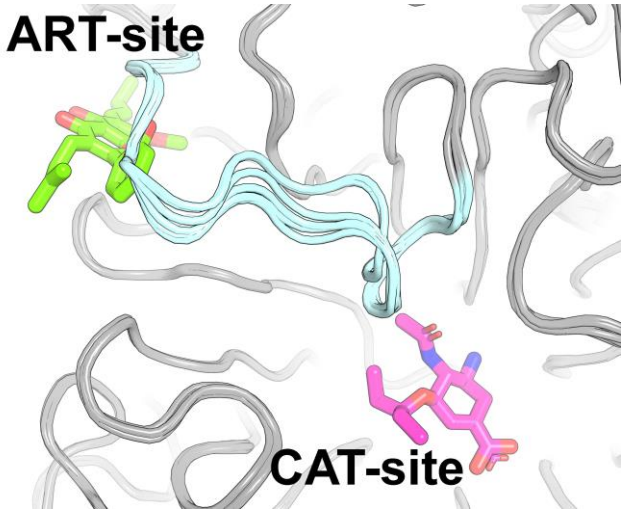

Рисунок 35. 3D-наложение PDB-структур каталитического домена NanA (PDB-структуры 2VVZ, 2W20, 2YA5, 2YA7, 2YA8, 3H71 наложены с помощью команды «alignto» в PyMOL). Гибкая петля 425–437 (голубая) соединяет предложенный в этой работе сайт связывания артокарпина («ART-site») с известным каталитическим сайтом («CAT-site») в структуре NanA. «ART-site» содержит артокарпин (показан в виде зеленых палочек), связанный в соответствии с суперкомпьютерным моделированием. Каталитический сайт занят известным ингибитором каталитической функции фермента озельтамивиром (обозначен пурпурными палочками), координаты которого взяты из PDB 2YA8.

5.2. Подавление функциональных свойств нейраминидазы при связывании ингибиторов во втором аллостерическом центре NanA [iv]

Биоинформатический анализ и молекулярное моделирование использовали для выявления ранее неизвестных участков связывания в структурах гомологичных нейраминидаз пневмококка. Сравнительный анализ показал, что нейраминидазы NanA и NanB из *Streptococcus pneumoniae* характеризуются высокой степенью структурного сходства каталитических доменов (85% совпадений элементов вторичной структуры в соответствии с алгоритмом SSM [184]) и низкой степенью идентичности последовательностей (23% совпадений аминокислотных остатков при парном выравнивании). При помощи метода pocketZebra обнаружено, что в структуре NanA есть функционально важный участок связывания с высоким содержанием специфических позиций, который эквивалентен участку связывания ингибитора оптактина в кристаллографической структуре NanB (PDB 4XHX) и топологически независим от активного центра. Молекулярное моделирование показало, что в стабилизации оптактина в структуре NanB принимают участие остатки Tyr250, Thr251, Lys334, Gln494, Lys499, Lys597, Thr657 и Glu658, при этом важную роль в координации лиганда также играют молекулы воды. Однако в NanA соответствующие остатки в структуре аллостерического сайта заменены на другие аминокислоты, что препятствует связыванию оптактина.

[iv] При подготовке данного раздела диссертации использована следующая публикация, выполненная автором в соавторстве, в которой, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: **Шарапова Я. А.**, Швядас В. К. Молекулярное моделирование связывания аллостерического ингибитора оптактина в новом сайте в структуре нейраминидазы a из Streptococcus pneumoniae //Вестник Московского университета. Серия 2. Химия. – 2018. – Т. 59. – №. 5. – С. 323-331. объем вклада автора диссертации составляет 75% и заключался в задумке, планировании и выполнении экспериментов, а также написании статьи.

5.2.1. Аннотация нового аллостерического центра в структуре NanA с использованием биоинформатического анализа

Для поиска и аннотации участков связывания в структуре NanA из *S. pneumoniae* проводили биоинформатический анализ суперсемейства сиалидаз с использованием программы pocketZebra [260]. Метод основан на определении специфических позиций, консервативных внутри функциональных подсемейств, но различающихся между ними, в большом множественном выравнивании гомологов, и ранжировании полостей/карманов на поверхности структур белков в соответствии с предсказанной специфичностью. Ранее проведенный структурный анализ известных каталитических и аллостерических центров показал, что они обогащены специфическими позициями, которые играют ключевую роль в формировании функционального разнообразия в суперсемействах (например, способности связывать разные субстраты) [260]. Проведенный в этой работе биоинформатический анализ суперсемейства сиалидаз показал высокую специфичность каталитического центра (Рисунок 36, Pocket1), а также ранее неохарактеризованного сайта в структуре NanA, который расположен на противоположной стороне β-пропеллера каталитического домена фермента (Рисунок 36, Pocket2). Сравнительный анализ показал, что Pocket2 в структуре NanA эквивалентен участку связывания ингибитора оптактина в кристаллографическом комплексе гомологичного фермента NanB (PDB 4XHX). Таким образом, наличие двух центров связывания в противоположных воронках каталитического домена является свойством не только NanB, но и других нейраминидаз, в том числе NanA, а высокое содержание специфических позиций в новом сайте позволяет предположить его участие в функционировании этих ферментов.

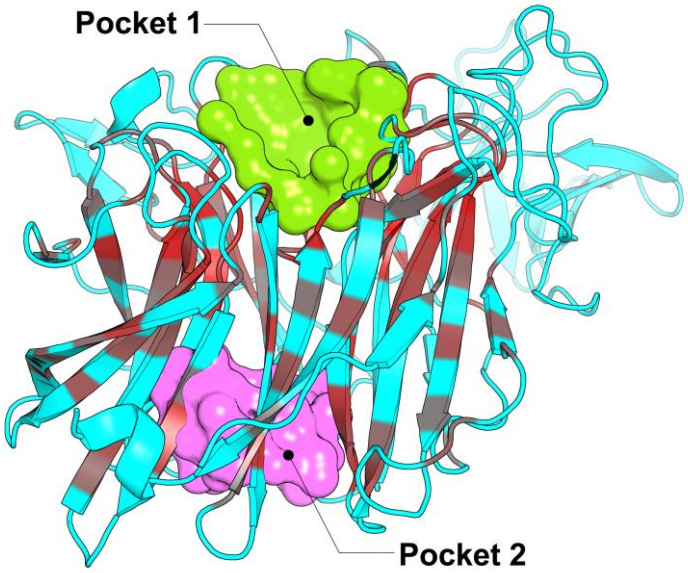

Рисунок 36. Сайты связывания лигандов в структуре NanA из *S. pneumoniae* (PDB 2YA8), предсказанные методом pocketZebra на основании биоинформатического анализа суперсемейства сиалидаз. Pocket1 соответствует ранее описанному каталитическому центру, Pocket2 – ранее неохарактеризованному новому сайту в структуре NanA. Нумерация центров связывания соответствует ранжированию pocketZebra. Градиентная окраска основной цепи белка соответствует статистической оценке специфичности позиций в колонках множественного выравнивания суперсемейства сиалидаз (интенсивность красного соответствует степени специфичности). Иллюстрация подготовлена с использованием программы PyMol.

5.2.2. Молекулярное моделирование взаимодействия ингибитора оптактина с аллостерическими центрами в структурах NanA и NanB

Молекулярное моделирование было использовано для изучения связывания оптактина в аллостерическом сайте NanB [202]. Было показано, что ориентация лиганда сохраняется неизменной на протяжении 10 нс молекулярно-динамической траектории и в целом соответствует его положению в кристалле 4XHX (Рисунок 37 A). Среднеквадратичное отклонение (RMSD) оптактина от его первоначального положения в течение всего периода симуляции в среднем составило 1,15 Å, что идентично среднему значению RMSD 1,14 Å для всех тяжелых атомов белка за аналогичный период симуляции. Стабилизация лиганда в структуре фермента обеспечивается за счет водородных и ионных взаимодействий между его

функциональными группами и аминокислотами центра связывания, в которых важную роль играют молекулы растворителя (Рисунок 37 Б). Дигидроксициклопентановая группа ингибитора образует водородные связи с остатками Thr251, Gln494 и Thr657 и формирует гидрофобный контакт с ароматической группой Tyr250. Вторичная аминогруппа лиганда взаимодействует с Lys597 и Glu658 опосредованно через молекулу воды. Сульфогруппа координируется тремя водородными связями с участием молекул воды и остатков Tyr250, Thr251, Ala497 и Lys597. Кроме того, моделирование показало, что сульфогруппа оптактина дополнительно стабилизируется одним ионным мостиком, в образовании которого по очереди участвуют остатки Lys334 и Lys499 (Рисунок 38). Циклогексиловая группа оптактина формирует контакт с гидрофобной поверхностью аминокислотных остатков, расположенных у основания β-поворота 658-664. Оценка свободной энергии связывания оптактина в описанной ориентации в структуре NanB составила -11,17 ккал/моль [202].

Молекулярное моделирование было использовано для изучения связывания оптактина в новом сайте в структуре NanA (Рисунок 36, Pocket2). Ориентация ингибитора в исходной модели комплекса с NanA была аналогична его ориентации в аллостерическом центре NanB (PDB 4XHX). Молекулярное моделирование показало, что связывание лиганда в таком положении в NanA нестабильно. В течение первых 10 нс молекулярно-динамической симуляции среднеквадратичное отклонение (RMSD) оптактина от его первоначального положения составило до 5 Å, а в течение последующих 20 нс наблюдается либо полная диссоциация лиганда и его переход в раствор, либо его дальнейшее смещение вдоль поверхности сайта связывания (Рисунок 39 А) [202].

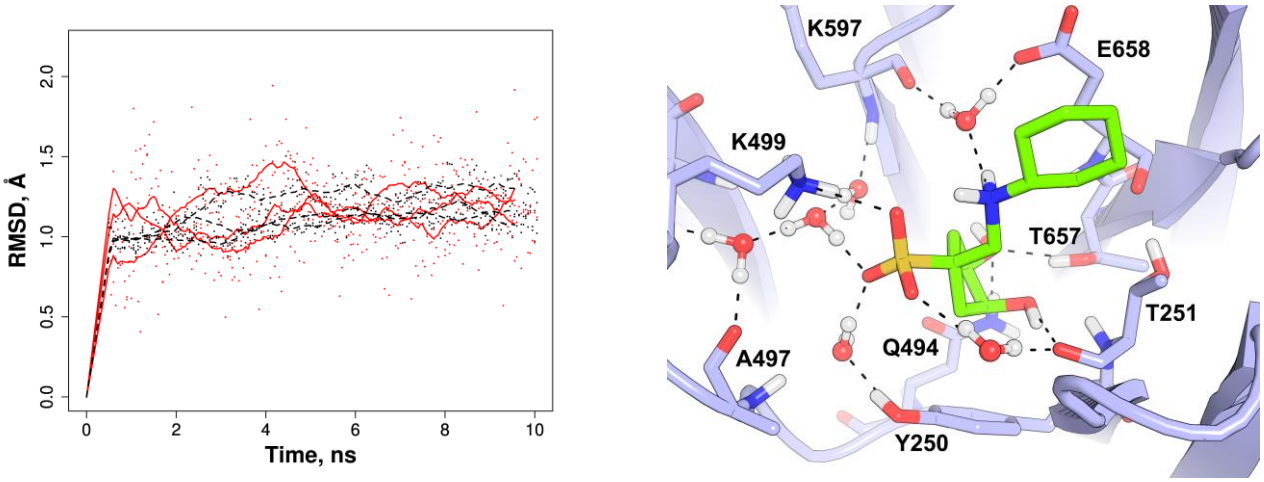

Рисунок 37. (А) Среднеквадратичное отклонение (RMSD) ингибитора оптактина в аллостерическом сайте NanB из *S. pneumoniae* в процессе молекулярно-динамической симуляции. При расчете RMSD учитывались только тяжелые атомы оптактина (красная сплошная линия, красные точки) и тяжелые атомы основной цепи белка (черная пунктирная линия, черные точки). Приведены результаты расчета трех траекторий. Точки соответствуют замеренным значениям RMSD, линии соответствуют усредненным значениям с шагом 1 нс. Иллюстрация подготовлена с использованием пакета R. (Б) Связывание ингибитора оптактина (зеленый) в аллостерическом центре NanB из *S. pneumoniae*. Пунктирными линиями изображены водородные и ионные связи между аминокислотными остатками белка и оптактином, в том числе с участием молекул воды. Нумерация аминокислотных остатков приведена в соответствии с PDB 4XHX. Иллюстрация подготовлена с использованием программы PyMol.

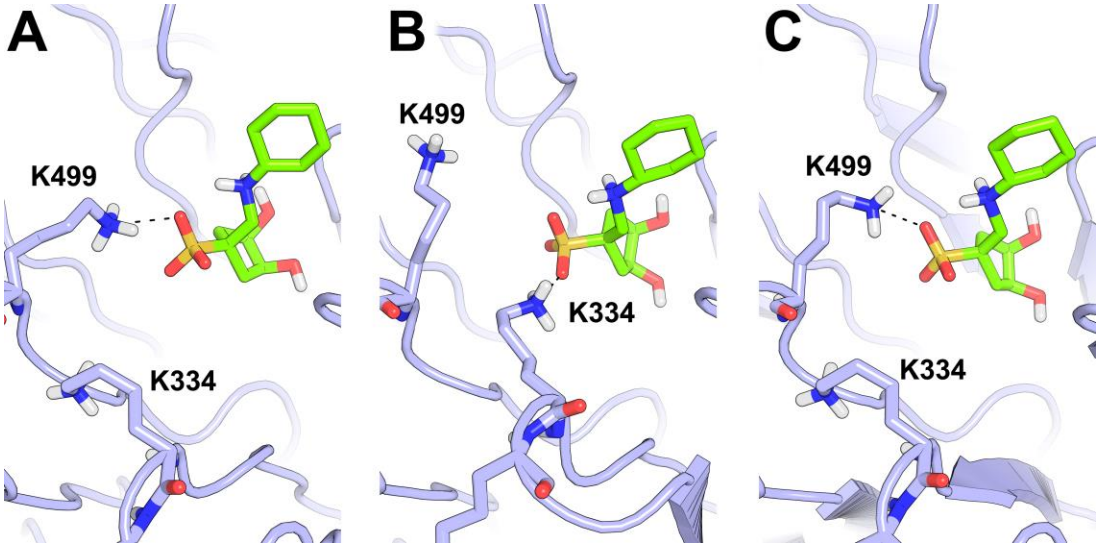

Рисунок 38. Координация сульфогруппы оптактина (зеленый) в аллостерическом центре NanB из *S. pneumoniae* с участием Lys334 и Lys499 – на 1 нс (А), 5 нс (В) и 8.5 нс (С) молекулярно-динамической симуляции. Нумерация аминокислотных остатков приведена в соответствии с PDB 4XHX. Иллюстрация подготовлена с использованием программы PyMol.

Сравнительный анализ аллостерического сайта в NanB и нового сайта в NanA показал, что, при схожей структурной организации их воронок, аминокислоты, принимающие участие в координации оптактина в NanB, в структуре NanA заменены на остатки с другими физико-химическими свойствами (Рисунок 39 Б). Остатки Lys334 и Lys499 в NanB, принимающие участие в связывании сульфогруппы, заменены в NanA на Pro409 и Arg562, при этом гуанидиновая группа аргинина образует водородные связи с атомом кислорода основной цепи Gln637 и боковым радикалом Asn639 и, таким образом, ориентирована в противоположную сторону от участка связывания лиганда. Остаток Tyr250 в NanB заменен на Leu337 в NanA, что увеличивает объем центра связывания. Другие ключевые остатки Thr657, Thr251 и Gln494 заменены соответственно на Gln741, Lys338 и Thr557. Кроме того, β-поворот 658-664, остатки которого образуют гидрофобный контакт с атомами циклогексиловой группы оптактина в NanB, в NanA имеет другой аминокислотный состав и ориентацию в структуре [202].

Поиск новых участков связывания в структурах белков и комплементарных к ним лигандов, взаимодействие с которыми может селективно регулировать функциональную активность биополимера, представляется новым, активно развивающимся направлением наук о белках. Это связано как с фундаментальным интересом к поиску новых путей регуляции белков/ферментов, так и обусловлено преимуществами аллостерических модуляторов – возможностью создания новых лекарственных средств, обладающих меньшей токсичностью за счет большей селективности связывания и специфичности механизма действия [256, 257, 258, 259]. В этой работе биоинформатический анализ был использован для аннотации нового, ранее неохарактеризованного сайта в структуре фермента нейраминидазы А – ключевого фактора вирулентности патогенной бактерии *S. pneumoniae* [202].

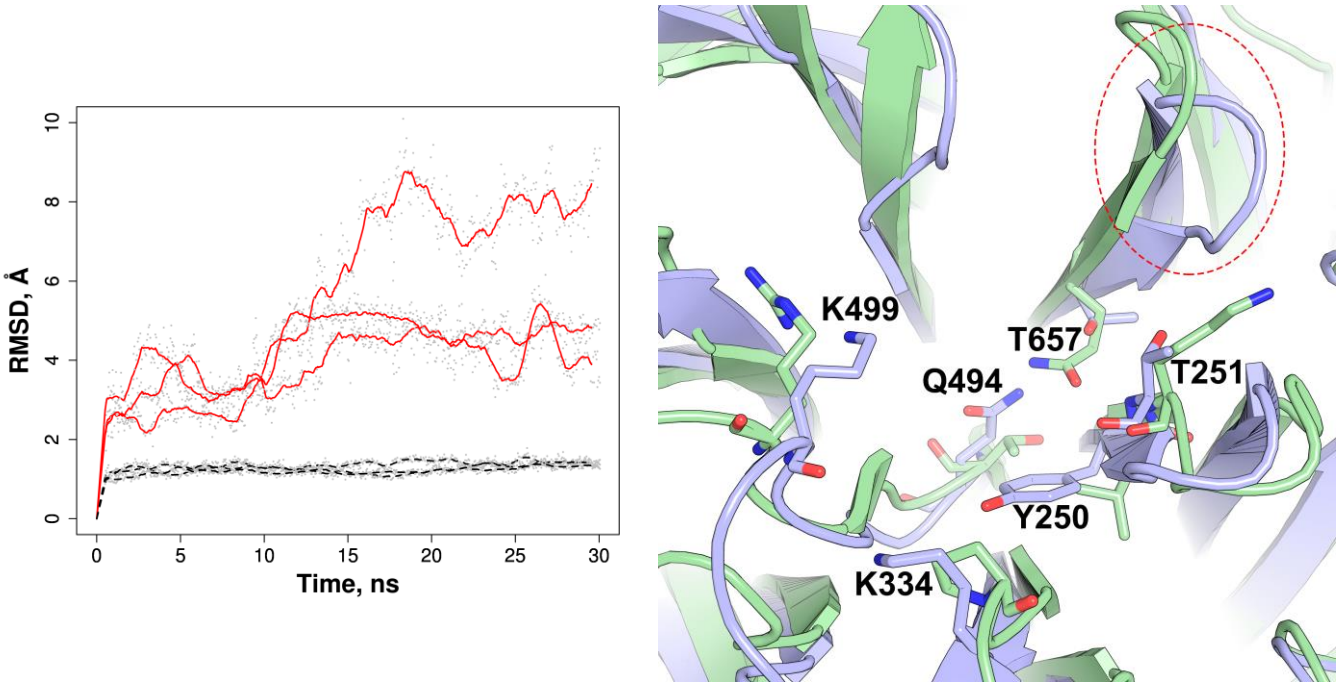

Рисунок 39. (А) Среднеквадратичное отклонение (RMSD) тяжелых атомов оптактина в новом сайте NanA из *S. pneumoniae* в процессе молекулярно-динамической симуляции. При расчете RMSD учитывались только тяжелые атомы оптактина (красная сплошная линия) и тяжелые атомы основной цепи белка (черная пунктирная линия). Приведены результаты расчета трех траекторий. Точки соответствуют замеренным значениям RMSD, линии соответствуют усредненным значениям с шагом 1 нс. Иллюстрация подготовлена с использованием пакета R. (Б) Структурное выравнивание нового центра в NanA (зеленый) и аллостерического центра в NanB (голубой) из *S. pneumoniae*. Палочками показаны аминокислотные остатки в структуре NanB, которые участвуют в связывании оптактина, и соответствующие им остатки в структуре NanA. Красным пунктиром обведен β-поворот, ориентация которого в структурах NanA и NanB отличается (см. объяснение в тексте). Нумерация аминокислотных остатков приведена в соответствии с PDB 4XHX. Иллюстрация подготовлена с использованием программы PyMol.

При этом было показано, что обнаруженный сайт образован позициями, специфическими в суперсемействе сиалидаз, что говорит о его участии в функционировании фермента. Важным фактором, способствовавшим проведению исследования, было появление в базе данных PDB структуры родственного фермента нейраминидазы B в комплексе с ингибитором оптактином, который связан в сайте, эквивалентном новому сайту в NanA. Молекулярное моделирование показало, что связывание оптактина в NanA затруднено другим

аминокислотным составом нового участка связывания. Однако сам факт наличия в NanA такого сайта говорит о необходимости продолжить исследование его взаимосвязи с функцией фермента и поиск комплементарных регуляторных лигандов с использованием методов биоинформатики и молекулярного моделирования [202].

5.2.3. Молекулярное моделирование взаимодействия олигосахаридного рецептора с аллостерическим центром в структуре NanA

Изучение связывания сахаров в новом аллостерическом центре NanA проводилось с помощью методов докинга и молекулярной динамики. Для докинга были выбраны сахара, в изобилии присутствующие на поверхности дыхательного эпителия и в покрывающей его слизи – это α- и β- манноза, β-N-ацетилглюкозамин, β-галактоза, α-N-ацетилнейраминовая кислота. Эти моно-сахара составляют цепочки олигосахаридных рецепторов. Последовательная деятельность гликозил-гидролаз пневмококка и других обитателей микрофлоры может открывать доступ к этим рецепторам. Молекулярная динамика комлексов NanA, полученных молекулярным докингом, на протяжении 40 нс при температуре 310 К выявило высокую аффиность различных сахаров к аллостерическому сайту в NanA. Во всех случаях рассмотренные лиганды оставались в воронке аллостерического центра на протяжении всего времени наблюдения. Наиболее стабильное связывание в новом центре продемонстрировала β-галактоза (Рисунок 40 и Рисунок 41 А). Ориентации α-маннозы и α-N-ацетилнейраминовой кислоты сохранялись неизменными по меньшей мере в одной траектории из трех повторений, а в остальных случаях претерпевали незначительные изменения за счет переключения некоторых функциональных групп субстратов на взаимодействие с соседними аминокислотными остатками аллостерического центра. Наименее стабильное

распознавание наблюдалось для β-N-ацетилглюкозамина и β-маннозы, которые во всех повторениях молекулярно-динамических траекторий существенно смещались относительно первоначального положения, описанного с использованием докинга (Рисунок 41).

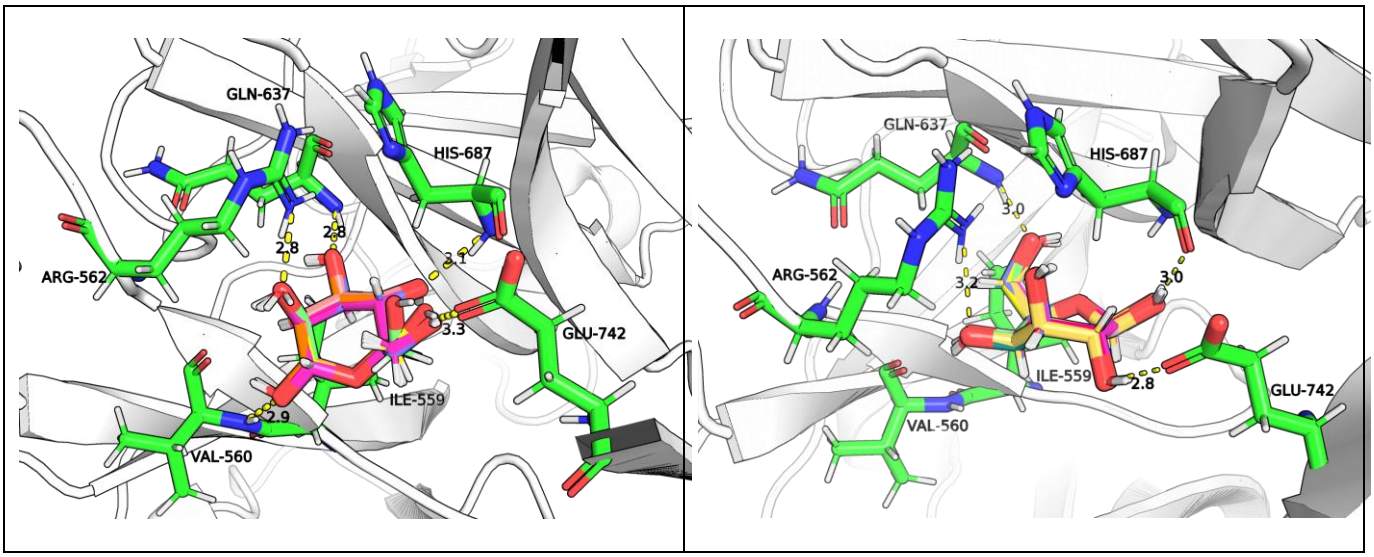

Рисунок 40. Связывание α-маннозы (слева) и β-галактозы (справа) в аллостерическом центре NanA из *S. pneumoniae*. Зелеными палочками показаны аминокислотные остатки в структуре NanA, которые участвуют в связывании сахаров. Желтыми пунктирными линиями показаны водородные связи, образуемые молекулой сахара с остатками аллостерического центра. Иллюстрация подготовлена с использованием программы PyMol.

Недавно стало известно, что поверхностные гликозил-гидролазы пневмококка, в т.ч. BgaA и NanA, могут действовать как адгезины на стадии асимптоматической колонизации пневмококком верхних дыхательных путей, связываясь с олигосахаридными рецепторами. Конкретные механизмы этого явления практически не изучены. Предполагается, что адгезивная способность NanA может быть связана с ее лектиновым доменом (и непосредственно участком связывания сиаловой кислоты в нем) [125, 126]. Наше исследование показывает, что новый аллостерический центр в NanA имеет высокую афинность к разным сахарам, что свидетельствует о том, что этот участок связывания может играть важную роль в адгезии и колонизации *S. pneumoniae* (Рисунок 42).

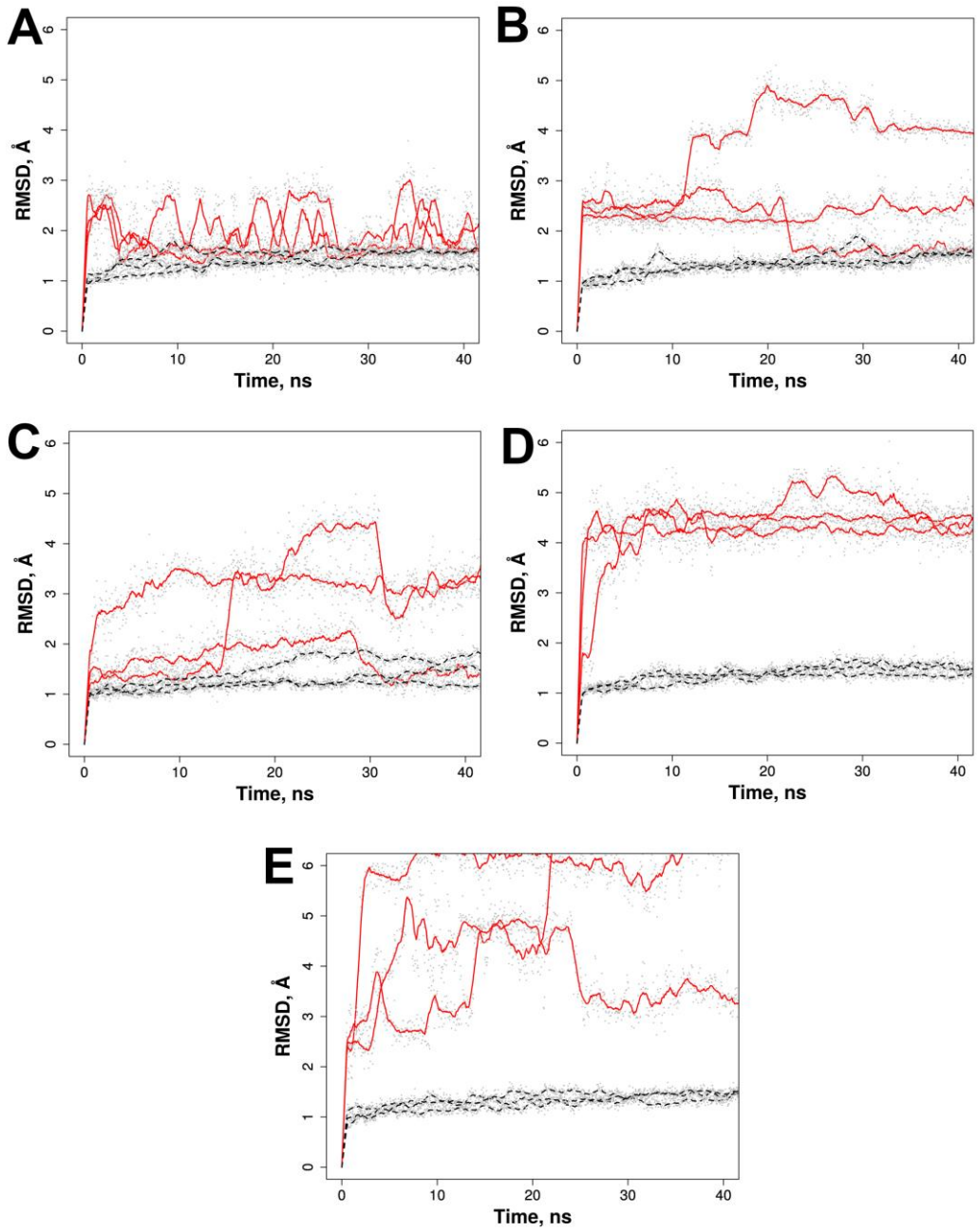

Рисунок 41. Структуры комплексов NanA из *S. pneumoniae* с (A) β-галактозой, (B) α-маннозой, (C) α-N-ацетилнейраминовой кислотой, (D) β-N-ацетилглюкозамином и (E) β-маннозой в течение 40 нс МД-симуляций при 310K.

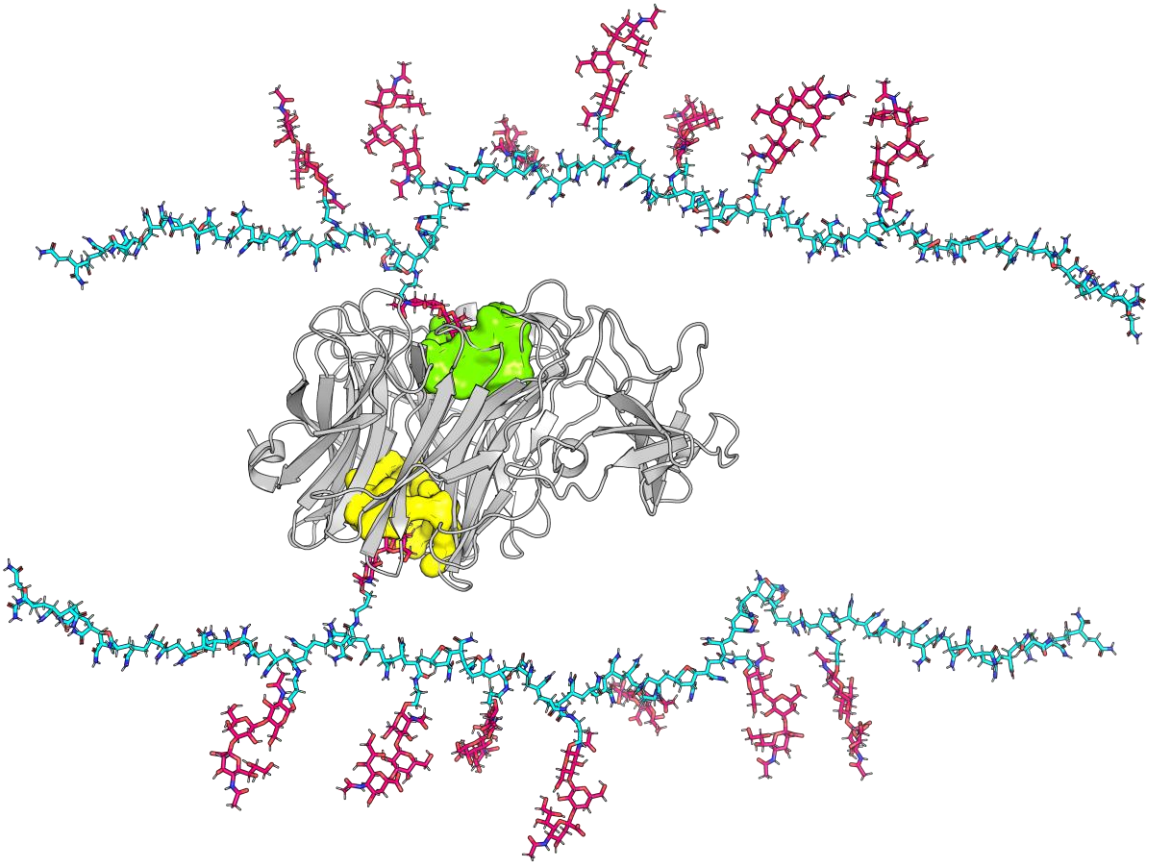

Рисунок 42. Схема взаимодействия двух олигосахаридных рецепторов (ярко-розовые палочки) на поверхности больших поливалентных субстратов с каталитическим и аллостерическим центрами связывания в структуре NanA из *S. pneumoniae* (зеленый, вверху – активный центр, желтый, внизу – аллостерический центр). Полимерная цепь, связывающая много олигосахаридных группировок в единый большой субстрат, выделена голубыми палочками. Иллюстрация подготовлена с использованием программы PyMol.

5.3. Подавление функциональных свойств нейраминидазы NanA ингибиторами двойного действия, взаимодействующими с консервативными остатками активного центра нейраминидаз патогенов

5.3.1. Сравнительный анализ активных центров бактериальных и вирусных нейраминидаз

Для анализа активных центров в структурах бактериальных и вирусных нейраминидаз было построено структурно-опосредованное множественное выравнивание соответствующих ферментов с их гомологами с использованием метода Mustguseal [185]. Полученное выравнивание содержало 429 последовательностей и структур нейраминидаз/сиалидаз из различных бактериальных источников и вирусных штаммов. Для характеристики участков связывания и определения консервативных и специфических аминокислотных остатков был использован метод pocketZebra [260]. Было показано, что центральная часть структур каталитических доменов бактериальных и вирусных нейраминидаз соответствует 6-лопастному-β-пропеллеру и характеризуется высокой степенью структурного сходства у гомологов из разных источников, однако периферийные элементы – петли, соединяющие β-тяжи, а также участки структуры, удаленные от центральной оси симметрии – характеризуются значительными отличиями. Следующие аминокислотные остатки консервативны в изученных бактериальных и вирусных нейраминидазах (нумерация приведена по структуре каталитического домена бактериальной нейраминидазы, PDB: 2YA8) – Asp408, Arg648, Tyr525, Glu632, Leu551, Arg706, Asp678, Asp357, Pro334, Tyr737, Trp547, Glu753, Arg332, Val553, Gly568, Val681. Консервативные позиции преимущественно локализованы в области активного центра рассмотренных ферментов и представляют интерес в контексте разработки ингибиторов двойного действия, способных подавлять активность обоих ферментов. Специфические

позиции, т.е. эквивалентные участки структур нейраминидаз, которые заняты разными типами аминокислотных остатков в ферментах из вирусов и патогенных бактерий, напротив, интересны как механизм распознавания ингибиторов, селективных к каждому ферменту по-отдельности (Рисунок 43).

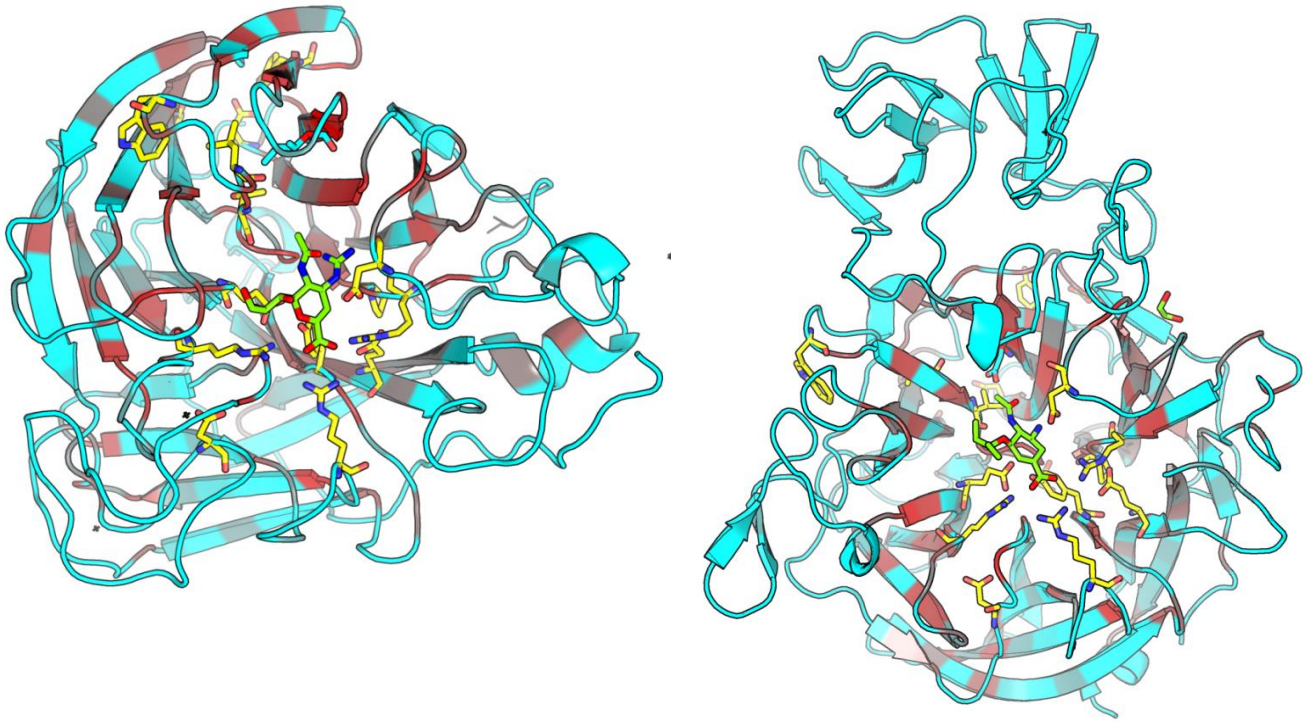

Рисунок 43. Аннотация консервативных (желтый) и специфических (красный) остатков в структурах каталитических доменов нейраминидазы вируса гриппа H1N1 (PDB: 3B7E, слева) и *S. pneumoniae* (PDB: 2YA8, справа), аннотированные в соответствии с результатами биоинформатического анализа. Лиганд, связанный в активном центре в кристаллографической структуре фермента, показан зеленым цветом. Градиентная окраска красный-бирюзовый соответствует степени специфичности позиций: чем выше интенсивность красного, тем выше специфичность.

5.3.2. Молекулярное моделирование комплексов бактериальных и вирусных нейраминидаз с ингибиторами двойного действия

Поскольку сиалидазы являются важными факторами вирулентности для вируса гриппа и многих патогенных бактерий, в т.ч. *S. pneumoniae*, а вторичная бактериальная инфекция является главной причиной смертности при заболевании гриппом, создание ингибиторов двойного действия, способных подавлять активность обеих сиалидаз, стало в последние годы популярной стратегией, чтобы разрушить «смертельный синергизм» вируса гриппа и пневмококка [181, 182, 183]. Ведется активный поиск ингибиторов двойного действия путем скрининга библиотек соединений, однако пока не было обнаружено ни одного вещества, ингибирующего оба фермента достаточно эффективно. Известно, что некоторые ингибиторы нейраминидазы вируса гриппа способны подавлять активность и нейраминидазы А, однако при значительно более высоких концентрациях, чем вирусный фермент. Так, занамивир и озельтамивир являются сильными ингибиторами вирусной нейраминидазы, действуя в нижнем наномолярном диапазоне; в то же время озельтамивир ингибирует NanA на три порядка хуже, чем вирусную, а занамивир практически не действует. Еще одно соединение, которое часто используется для изучения и сравнения кинетических свойств разных сиалидаз – 2-дезокси-2,3-дидегидро-N-ацетилнейраминовая кислота (также известная как DANA или Neu5Ac2en) – ингибирует оба фермента в микромолярном диапазоне (Таблица 2).

Таблица 2. Оценка связывания ингибиторов нейраминидаз, полученная докингом, в сравнении с экспериментальными значениями.

Ингибитор	Нейраминидаза вируса гриппа H1N1 (PDB 3B7E)		NanA из *S. pneumoniae* (PDB 2YA8)	
	Воспроизводимость корректной позы при докинге, %	IC_{50}, CL-assay, μm	Воспроизводимость корректной позы при докинге, %	IC_{50}, CL-assay, μm
Занамивир	90	0,00032 ± 0,00011	0	Не активен
DANA	30	0,59 ± 0,26	20	2,59 ± 0,81
Озельтамивир	100	0,00022 ± 0,00005	80	0,65 ± 0,13

Средние значения IC_{50} указаны для хемилюминесцентного теста (CL-assay) и взяты из [261].

Чтобы понять, каким образом ингибиторы двойного действия способны подавлять активность и бактериальной, и вирусной нейраминидаз, мы сравнили связывание занамивира, озельтамивира и DANA в активных центрах обоих ферментов с использованием методов компьютерной биологии. Молекулярное моделирование показало, что консервативные остатки в активных центрах NanA из *S. pneumoniae* и нейраминидазы вируса гриппа H1N1 участвуют в связывании озельтамивира, DANA и занамивира. Однако разница в действии этих ингибиторов на вирусный и бактериальный гомологи в значительной степени обусловлена наличием специфических позиций в активных центрах (Рисунок 44). Наиболее простое в структурном смысле соединение из трех вышеперечисленных – DANA, примерно одинаково взаимодействует с одними и теми же консервативными остатками, и в обоих случаях корректная поза связывания

молекулы (такая поза, которая наблюдается в структурах кристаллографических комплексов нейраминидаз с соответствующим лигандом) воспроизводится при докинге не очень часто, в 20-30% случаев. Занамивир – это производное DANA, в котором O4-гидроксильная группа замещена гуанидиновой для взаимодействия с Glu119 и Glu227 в активном центре вирусной нейраминидазы. Эти остатки являются специфическими для вирусных нейраминидаз, в бактериальной NanA они замещены Ile333 и Pro555, а аналогичные отрицательно заряженные остатки расположены в пространстве по-другому. Гуанидиновая группа занамивира, таким образом, теряет свои взаимодействия в структуре NanA. Кроме того, в активный центр бактериального фермента характеризуется наличием еще одной специфической позиции, которая мешает связыванию занамивира: боковая цепь Arg351 ориентирована в полость центра связывания и вытесняет оттуда занамивир. Молекулярный докинг хорошо воспроизводит продуктивное положение занамивира в активном центре вирусной нейраминидазы, в то время как связывание в бактериальной нейраминидазе не обнаруживается. В активном центре нейраминидазы вируса гриппа присутствуют также две специфические позиции, которые при связывании озельтамивира образуют солевой мостик – Glu276 и Arg224. Неполярная поверхность, образованная их боковыми цепями, образует гидрофобное взаимодействие с 3-пентильной группой озельтамивира. В NanA эти остатки заменены на Thr631 и Gly552, что обуславливает ухудшение ингибирования бактериального фермента озельтамивиром на три порядка. Результаты молекулярного моделирования согласуются с экспериментальными данными об ингибировании бактериальной и вирусной нейраминидаз (Таблица 2).

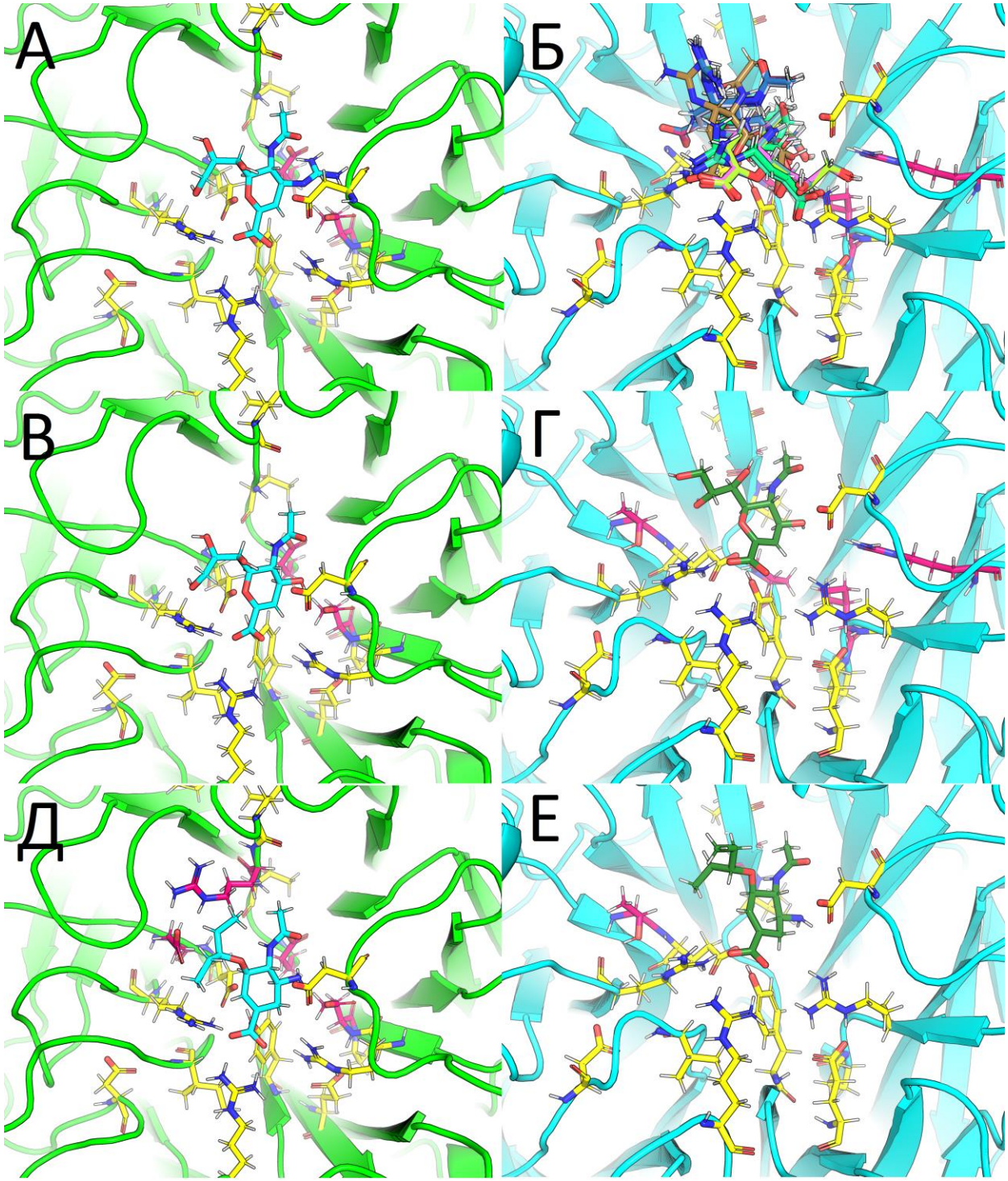

Рисунок 44. Моделирование связывания ингибиторов двойного действия занамивира (А и Б), DANA (В и Г) и озельтамивира (Д и Е) в активных центрах нейраминидазы вируса гриппа H1N1 (изображения слева) и NanA из *S. pneumoniae* (изображения справа). Позиции ингибиторов получены молекулярным докингом. Консервативные остатки показаны желтыми палочками, специфические позиции, обуславливающие разницу в связывании ингибиторов в вирусном и бактериальном гомологах, показаны ярко-розовыми палочками. Пояснения см. в тексте.

Таким образом, в активных центрах NanA из *S. pneumoniae* и вирусных нейраминидазах идентифицированы консервативные позиции (триаргининовый кластер, ключевые каталитические остатки Tyr, Glu и Asp), а также специфические остатки, отличающиеся у нейраминидаз из различных источников и обуславливающие селективность каждого центра. Полученные результаты создают основу для дизайна ингибиторов двойного действия, способных подавлять активность как бактериальных, так и вирусных нейраминидаз.

6. ЗАКЛЮЧЕНИЕ

Нейраминидаза А (NanA) играет важную роль в патогенезе *Streptococcus pneumoniae* – отщепляя остатки сиаловой кислоты от олигосахаридных рецепторов на поверхности клеток дыхательного эпителия человека, она участвует в процессе колонизации организма-хозяина и формировании бактериальной биопленки. Кроме того, NanA может кооперировать с вирусной нейраминидазой, приводя к развитию тяжелых вторичных бактериальных инфекций (пневмонии, отиту, менингиту). Изучение структурно-функциональных взаимосвязей в NanA и поиск новых путей ее регуляции представляет фундаментальный и практический интерес. Для решения поставленной задачи был применен комплекс методов биоинформатического анализа и молекулярного моделирования. Использование высокопроизводительных вычислений, в том числе ресурсов суперкомпьютерного комплекса МГУ имени М.В.Ломоносова, для анализа всей доступной информации о структурах и последовательностях гомологов в биоинформатических базах данных, расчета длинных траекторий молекулярной динамики и обработки больших наборов конформационных вариантов белка позволило выйти на качественно новый уровень понимания роли динамической структуры NanA в проявлении ее функциональных свойств. Проведенные исследования позволяют сделать вывод о том, что в дополнение к «классическому» способу регуляции активности, который подразумевает поиск ингибиторов, селективно взаимодействующих с каталитическим центром конкретного фермента из патогенного организма, для регуляции функции NanA применимы, по меньшей мере, три новых подхода:

(1) создание ингибиторов, связывающихся в новом, ранее неизвестном аллостерическом центре распознавания артокарпина и подавляющих участие NanA в формировании биопленок пневмококка, что становится возможным

благодаря впервые описанной нами уникальной структурной организации фермента, а именно пространственной обособленности лектинового и каталитического доменов, соединенных гибким линкером из 16 аминокислот, что позволяет им образовывать межмолекулярные агрегаты;

(2) создание селективных ингибиторов ко второму аллостерическому центру NanA, расположенному в воронке на противоположной стороне шестилопастного β-пропеллера каталитического домена фермента;

(3) создание ингибиторов множественного действия, взаимодействующих с консервативными остатками активного центра широкого ряда гомологичных нейраминидаз/сиалидаз из патогенных организмов, в том числе вирусной нейраминидазы, и подавляющих индивидуальную активность этих белков и их синергию.

7. ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

1) Лектиновый и каталитический (включая инсерционный) домены в структуре NanA соединены подвижным, гибким линкером из 16 аминокислотных остатков в отличие от близких гомологов NanB и NanC, где соответствующие домены не разделены в пространстве и образуют компактную глобулу;

2) Молекулы NanA, прикрепленные к клеточной стенке разных клеток пневмококка, способны образовать межмолекулярные агрегаты за счет взаимодействия их лектиновых и каталитических доменов, что может являться одним из механизмов формирования биопленок пневмококком;

3) Механизм действия артокарпина – ингибитора формирования биопленок пневмококком, основан на его взаимодействии с аллостерическим центром на поверхности каталитического домена, расположенным в области контакта с лектиновым или каталитическим доменом другой молекулы NanA при межмолекулярной агрегации;

4) В структуре NanA есть второй аллостерический центр, топологически независимый от активного центра и расположенный в воронке на противоположной стороне шестилопастного β-пропеллера каталитического домена фермента;

5) Второй аллостерический центр в NanA эквивалентен ранее описанному аллостерическому центру в гомологичной сиалидазе NanB пневмококка, однако не обладает достаточным сродством к низкомолекулярному лиганду оптактину, аллостерическому ингибитору NanB, в связи с заменами

аминокислотных остатков в ключевых позициях данного участка связывания;

6) Второй аллостерический центр в NanA способен эффективно связывать сахара, входящие в состав олигосахаридных рецепторов эпителия верхних дыхательных путей человека, в том числе β-галактозу, α-маннозу, α-N-ацетилнейраминовую кислоту, что свидетельствует о роли этого участка связывания в адгезии и колонизации *S. pneumoniae*;

7) В активных центрах NanA из *S. pneumoniae* и вирусных нейраминидазах идентифицированы консервативные позиции (триаргининовый кластер, ключевые каталитические остатки Tyr, Glu и Asp), а также специфические остатки, отличающиеся у нейраминидаз из различных источников и обуславливающие селективность каждого центра. Полученные результаты создают основу для дизайна ингибиторов двойного действия, способных подавлять активность как бактериальных, так и вирусных нейраминидаз.

8. СПИСОК ЛИТЕРАТУРЫ

1. Roggentin, P., Schauer, R., Hoyer, L. L., & Vimr, E. R. (1993). The sialidase superfamily and its spread by horizontal gene transfer. *Molecular microbiology*, **9(5)**, 915-921.

2. Varki, A. (2007). Glycan-based interactions involving vertebrate sialic-acid-recognizing proteins. *Nature*, **446(7139)**, 1023.

3. Schwerdtfeger, S. M., & Melzig, M. F. (2010). Sialidases in biological systems. *Die Pharmazie-An International Journal of Pharmaceutical Sciences*, **65(8)**, 551-561.

4. Giacopuzzi, E., Bresciani, R., Schauer, R., Monti, E., & Borsani, G. (2012). New insights on the sialidase protein family revealed by a phylogenetic analysis in metazoa. *PloS one*, **7(8)**, e44193.

5. Buschiazzo, A., & Alzari, P. M. (2008). Structural insights into sialic acid enzymology. *Current opinion in chemical biology*, **12(5)**, 565-572.

6. Morley, T. J., Willis, L. M., Whitfield, C., Wakarchuk, W. W., & Withers, S. G. (2009). A new sialidase mechanism bacteriophage K1F endo-sialidase is an inverting glycosidase. *Journal of Biological Chemistry*, **284(26)**, 17404-17410.

7. Stummeyer, K., Dickmanns, A., Mühlenhoff, M., Gerardy-Schahn, R., & Ficner, R. (2005). Crystal structure of the polysialic acid–degrading endosialidase of bacteriophage K1F. *Nature structural & molecular biology*, **12(1)**, 90.

8. Bule, P., Chuzel, L., Blagova, E., Wu, L., Gray, M. A., Henrissat, B., Rapp, E., Bertozzi, C. R., Taron, C. H. & Davies, G. J. (2019). Inverting family GH156 sialidases define an unusual catalytic motif for glycosidase action. *Nature communications*, **10(1)**, 1-11.

9. Tailford, L.E., Owen, C. D., Walshaw, J., Crost, E. H., Hardy-Goddard, J., Le Gall, G., de Vos, W.M., Taylor, G. L., Juge, N. (2015). Discovery of intramolecular trans-sialidases in human gut microbiota suggests novel mechanisms of mucosal adaptation. *Nature communications*, **6**, 7624.

10. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., & Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic acids research*, **42(D1)**, D490-D495.

11. Knudsen, M., & Wiuf, C. (2010). The CATH database. *Human genomics*, **4(3)**, 207.

12. Harvey, H. A., Swords, W. E., & Apicella, M. A. (2001). The Mimicry of Human Glycolipids and Glycosphingolipids by the Lipooligosaccharides of Pathogenic Neisseria andHaemophilus. *Journal of autoimmunity*, **16(3)**, 257-262.

13. Shakhnovich, E. A., King, S. J., & Weiser, J. N. (2002). Neuraminidase expressed by Streptococcus pneumoniae desialylates the lipopolysaccharide of Neisseria meningitidis and Haemophilus influenzae: a paradigm for interbacterial competition among pathogens of the human respiratory tract. *Infection and immunity*, **70(12)**, 7161-7164.

14. Corfield, T. (1992). Bacterial sialidases—roles in pathogenicity and nutrition. *Glycobiology*, **2(6)**, 509-521.

15. Wiggins, R., Hicks, S. J., Soothill, P. W., Millar, M. R., & Corfield, A. P. (2001). Mucinases and sialidases: their role in the pathogenesis of sexually transmitted infections in the female genital tract. *Sexually transmitted infections*, **77(6)**, 402-408.

16. Taylor, G. (1996). Sialidases: structures, biological significance and therapeutic potential. *Current opinion in structural biology*, **6(6)**, 830-837.

17. Lewis, A. L., & Lewis, W. G. (2012). Host sialoglycans and bacterial sialidases: a mucosal perspective. *Cellular microbiology*, **14(8)**, 1174-1182.

18. Sudhakara, P., Sellamuthu, I., & Aruni, A. W. (2019). Bacterial sialoglycosidases in Virulence and Pathogenesis. *Pathogens*, **8(1)**, 39.

19. Miyagi, T., & Yamaguchi, K. (2012). Mammalian sialidases: physiological and pathological roles in cellular functions. *Glycobiology*, **22(7)**, 880-896.

20. Monti, E., Bonten, E., D'Azzo, A., Bresciani, R., Venerando, B., Borsani, G., Shauer, R. & Tettamanti, G. (2010). Sialidases in vertebrates: a family of enzymes tailored for several cell functions. *In Advances in carbohydrate chemistry and biochemistry*, **64**, 403-479. Academic Press.

21. Krammer, F., Smith, G. J., Fouchier, R. A., Peiris, M., Kedzierska, K., Doherty, P. C., Palese, P., Shaw, M.L., Treanor, J., Webster, R.G. & García-Sastre, A. (2018). Influenza. *Nature Reviews: Disease Primers*, **4**, 3.

22. Ducatez, M. F., Pelletier, C., & Meyer, G. (2015). Influenza D virus in cattle, France, 2011–2014. *Emerging infectious diseases*, **21(2)**, 368.

23. Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M., & Kawaoka, Y. (1992). Evolution and ecology of influenza A viruses. *Microbiological reviews*, **56(1)**, 152-179.

24. Wu, Y., Wu, Y., Tefsen, B., Shi, Y., & Gao, G. F. (2014). Bat-derived influenza-like viruses H17N10 and H18N11. *Trends in microbiology*, **22(4)**, 183-191.

25. Xu, X., Zhu, X., Dwek, R. A., Stevens, J., & Wilson, I. A. (2008). Structural characterization of the 1918 influenza virus H1N1 neuraminidase. *Journal of virology*, **82(21)**, 10493-10501.

26. Yoon, S. W., Webby, R. J., & Webster, R. G. (2014). Evolution and ecology of influenza A viruses. *In Influenza Pathogenesis and Control*, **I**, 359-375. Springer, Cham.

27. Dou, D., Revol, R., Östbye, H., Wang, H., & Daniels, R. (2018). Influenza A virus cell entry, replication, virion assembly and movement. *Frontiers in immunology*, **9**, 1581.

28. Taubenberger, J. K., & Kash, J. C. (2010). Influenza virus evolution, host adaptation, and pandemic formation. *Cell host & microbe*, **7(6)**, 440-451.

29. Bodewes, R., Morick, D., de Mutsert, G., Osinga, N., Bestebroer, T., van der Vliet, S., Smits, S.L., Kuiken, T., Rimmelzwaan, G.F., Fouchier, R.A.M. & Osterhaus, A. D. (2013). Recurring influenza B virus infections in seals. *Emerging infectious diseases*, **19(3)**, 511.

30. Beare, A. S., & Webster, R. G. (1991). Replication of avian influenza viruses in humans. *Archives of virology*, **119(1-2)**, 37-42.

31. Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A., & Klenk, H. D. (2004). Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proceedings of the National Academy of Sciences*, **101(13)**, 4620-4624.

32. Gambotto, A., Barratt-Boyes, S. M., de Jong, M. D., Neumann, G., & Kawaoka, Y. (2008). Human infection with highly pathogenic H5N1 influenza virus. *The Lancet*, **371(9622)**, 1464-1475.

33. Bouvier, N. M., & Palese, P. (2008). The biology of influenza viruses. *Vaccine*, **26**, D49-D53.

34. Cros, J. F., & Palese, P. (2003). Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. *Virus research*, **95(1)**, 3-12.

35. Samji, T. (2009). Influenza A: understanding the viral life cycle. *The Yale journal of biology and medicine*, **82(4)**, 153.

36. Portela, A., & Digard, P. (2002). The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. *Journal of general virology*, **83(4)**, 723-734.

37. Matrosovich, M. N., Matrosovich, T. Y., Gray, T., Roberts, N. A., & Klenk, H. D. (2004). Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. *Journal of virology*, **78(22)**, 12665-12667.

38. Mitnaul, L. J., Matrosovich, M. N., Castrucci, M. R., Tuzikov, A. B., Bovin, N. V., Kobasa, D., & Kawaoka, Y. (2000). Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. *Journal of virology*, **74(13)**, 6015-6020.

39. Wagner, R., Matrosovich, M., & Klenk, H. D. (2002). Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Reviews in medical virology*, **12(3)**, 159-166.

40. Xu, R., Zhu, X., McBride, R., Nycholat, C. M., Yu, W., Paulson, J. C., & Wilson, I. A. (2012). Functional balance of the hemagglutinin and neuraminidase activities accompanies the emergence of the 2009 H1N1 influenza pandemic. *Journal of virology*, **86(17)**, 9221-9232.

41. Gaymard, A., Le Briand, N., Frobert, E., Lina, B., & Escuret, V. (2016). Functional balance between neuraminidase and haemagglutinin in influenza viruses. *Clinical Microbiology and Infection*, **22(12)**, 975-983.

42. Hutchinson, E. C. (2018). Influenza virus. *Trends in microbiology*, **26(9)**, 809-810.

43. Fujisaki, S., Takashita, E., Yokoyama, M., Taniwaki, T., Xu, H., Kishida, N., Sato, H., Tashiro, M., Imai, M., & Odagiri, T. (2012). A single E105K mutation far from the active site of influenza B virus neuraminidase contributes to reduced susceptibility to multiple neuraminidase-inhibitor drugs. *Biochemical and biophysical research communications*, **429(1-2)**, 51-56.

44. Mitnaul, L. J., Castrucci, M. R., Murti, K. G., & Kawaoka, Y. (1996). The cytoplasmic tail of influenza A virus neuraminidase (NA) affects NA incorporation into virions, virion morphology, and virulence in mice but is not essential for virus replication. *Journal of virology*, **70(2)**, 873-879.

45. Jin, H., Leser, G. P., Zhang, J., & Lamb, R. A. (1997). Influenza virus hemagglutinin and neuraminidase cytoplasmic tails control particle shape. *The EMBO journal*, **16(6)**, 1236-1247.

46. Barman, S., Adhikary, L., Chakrabarti, A. K., Bernas, C., Kawaoka, Y., & Nayak, D. P. (2004). Role of transmembrane domain and cytoplasmic tail amino acid sequences of influenza a virus neuraminidase in raft association and virus budding. *Journal of virology*, **78(10)**, 5258-5269.

47. Enami, M., & Enami, K. (1996). Influenza virus hemagglutinin and neuraminidase glycoproteins stimulate the membrane association of the matrix protein. *Journal of virology*, **70(10)**, 6653-6657.

48. Barman, S., Ali, A., Hui, E. K. W., Adhikary, L., & Nayak, D. P. (2001). Transport of viral proteins to the apical membranes and interaction of matrix protein with glycoproteins in the assembly of influenza viruses. *Virus research*, **77(1)**, 61-69.

49. Blok, J., & Air, G. M. (1982). Variation in the membrane-insertion and" stalk" sequences in eight subtypes of influenza type A virus neuraminidase. *Biochemistry*, **21(17)**, 4001-4007.

50. Durrant, J. D., Bush, R. M., & Amaro, R. E. (2016). Microsecond molecular dynamics simulations of influenza neuraminidase suggest a mechanism for the increased virulence of stalk-deletion mutants. *The Journal of Physical Chemistry B*, **120(33)**, 8590-8599.

51. Kim, P., Jang, Y., Kwon, S., Lee, C., Han, G., & Seong, B. (2018). Glycosylation of hemagglutinin and neuraminidase of influenza a virus as signature for ecological spillover and adaptation among influenza reservoirs. *Viruses*, **10(4)**, 183.

52. Von Itzstein, M. (2007). The war against influenza: discovery and development of sialidase inhibitors. *Nature reviews Drug discovery*, **6(12)**, 967.

53. Штыря, Ю. А., Мочалова, Л. В., & Бовин, Н. В. (2009). Нейраминидаза вируса гриппа: структура и функция. *Acta Naturae*, **1(2)**, 28-34.

54. McAuley, J., Gilbertson, B., Trifkovic, S., Brown, L. E., & McKimm-Breschkin, J. (2019). Influenza virus neuraminidase structure and functions. *Frontiers in microbiology*, **10**, 39.

55. Chan, J., & Bennet, A. J. (2012). Enzymology of influenza virus sialidase. In *Influenza Virus Sialidase-A Drug Discovery Target* (pp. 47-66). Springer, Basel.

56. Varghese, J. N., Laver, W. G., & Colman, P. M. (1983). Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature*, **303(5912)**, 35.

57. Baker, A. T., Varghese, J. N., Laver, W. G., Air, G. M., & Colman, P. M. (1987). Three-dimensional structure of neuraminidase of subtype N9 from an avian influenza virus. *Proteins: Structure, Function, and Bioinformatics*, **2(2)**, 111-117.

58. Russell, R. J., Haire, L. F., Stevens, D. J., Collins, P. J., Lin, Y. P., Blackburn, G. M., Hay, A.J., Gamblin, S.J. & Skehel, J. J. (2006). The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature*, **443(7107)**, 45-49.

59. Amaro, R. E., Cheng, X., Ivanov, I., Xu, D., & McCammon, J. A. (2009). Characterizing loop dynamics and ligand recognition in human-and avian-type influenza neuraminidases via generalized born molecular dynamics and end-point free energy calculations. *Journal of the American Chemical Society*, **131(13)**, 4702-4709.

60. Amaro, R. E., Swift, R. V., Votapka, L., Li, W. W., Walker, R. C., & Bush, R. M. (2011). Mechanism of 150-cavity formation in influenza neuraminidase. *Nature communications*, **2**, 388.

61. Wu, Y., Qin, G., Gao, F., Liu, Y., Vavricka, C. J., Qi, J., Jiang, H., Yu, K. & Gao, G. F. (2013). Induced opening of influenza virus neuraminidase N2 150-loop suggests an important role in inhibitor binding. *Scientific reports*, **3**, 1551.

62. Amaro, R. E., Minh, D. D., Cheng, L. S., Lindstrom, W. M., Olson, A. J., Lin, J. H., Li, W.W. & McCammon, J. A. (2007). Remarkable loop flexibility in avian influenza N1 and its implications for antiviral drug design. *Journal of the American Chemical Society*, **129(25)**, 7764-7765.

63. Landon, M. R., Amaro, R. E., Baron, R., Ngan, C. H., Ozonoff, D., Andrew McCammon, J., & Vajda, S. (2008). Novel druggable hot spots in avian influenza neuraminidase H5N1 revealed by computational solvent mapping of a reduced and representative receptor ensemble. *Chemical biology & drug design*, **71(2)**, 106-116.

64. Craig, I. R., Pfleger, C., Gohlke, H., Essex, J. W., & Spiegel, K. (2011). Pocket-space maps to identify novel binding-site conformations in proteins. *Journal of chemical information and modeling*, **51(10)**, 2666-2679.

65. Swaminathan, K., Dyason, J. C., Maggioni, A., von Itzstein, M., & Downard, K. M. (2013). Binding of a natural anthocyanin inhibitor to influenza neuraminidase by mass spectrometry. *Analytical and bioanalytical chemistry*, **405(20)**, 6563-6572.

66. Le, L., Lee, E. H., Hardy, D. J., Truong, T. N., & Schulten, K. (2010). Molecular dynamics simulations suggest that electrostatic funnel directs binding of Tamiflu to influenza N1 neuraminidases. *PLoS computational biology*, **6(9)**, e1000939.

67. Tran, D., Le, L. T., & Truong, T. N. (2013). Discover binding pathways using the sliding binding-box docking approach: application to binding pathways of oseltamivir to avian influenza H5N1 neuraminidase. *Journal of computer-aided molecular design*, **27(8)**, 689-695.

68. Laver, W. G., Colman, P. M., Webster, R. G., Hinshaw, V. S., & Air, G. M. (1984). Influenza virus neuraminidase with hemagglutinin activity. *Virology*, **137(2)**, 314-323.

69. Varghese, J. N., Colman, P. M., Van Donkelaar, A., Blick, T. J., Sahasrabudhe, A., & McKimm-Breschkin, J. L. (1997). Structural evidence for a second sialic acid binding site in avian influenza virus neuraminidases. *Proceedings of the National Academy of Sciences*, **94(22)**, 11808-11812.

70. Sung, J. C., Wynsberghe, A. W., Amaro, R. E., Li, W. W., & McCammon, J. A. (2010). Role of secondary sialic acid binding sites in influenza N1 neuraminidase. *Journal of the American Chemical Society*, **132(9)**, 2883-2885.

71. Lai, J. C., Garcia, J. M., Dyason, J. C., Böhm, R., Madge, P. D., Rose, F. J., Nicholls, J.M., Peiris, J.S.M., Haselhorst, T. & von Itzstein, M. (2012). A secondary sialic acid binding site on influenza virus neuraminidase: fact or fiction? *Angewandte Chemie International Edition*, **51(9)**, 2221-2224.

72. Dai, M., McBride, R., Dortmans, J. C., Peng, W., Bakkers, M. J., de Groot, R. J., van Kuppeveld, F.J.M., Paulson, J.C., de Vries, E. & de Haan, C. A. (2017). Mutation of the second sialic acid-binding site, resulting in reduced neuraminidase activity, preceded the emergence of H7N9 influenza A virus. *Journal of virology*, **91(9)**, e00049-17.

73. Uhlendorff, J., Matrosovich, T., Klenk, H. D., & Matrosovich, M. (2009). Functional significance of the hemadsorption activity of influenza virus neuraminidase and its alteration in pandemic viruses. *Archives of virology*, **154(6)**, 945-957.

74. Chao, Y., Marks, L. R., Pettigrew, M. M., & Hakansson, A. P. (2015). Streptococcus pneumoniae biofilm formation and dispersion during colonization and disease. *Frontiers in cellular and infection microbiology*, **4**, 194.

75. Hammerschmidt, S., Wolff, S., Hocke, A., Rosseau, S., Müller, E., & Rohde, M. (2005). Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infection and immunity*, **73(8)**, 4653-4667.

76. Hammerschmidt, S. (2006). Adherence molecules of pathogenic pneumococci. *Current opinion in microbiology*, **9(1)**, 12-20.

77. Rajam, G., Anderton, J. M., Carlone, G. M., Sampson, J. S., & Ades, E. W. (2008). Pneumococcal surface adhesin A (PsaA): a review. *Critical reviews in microbiology*, **34(3-4)**, 131-142.

78. Azimi, S., Klementiev, A. D., Whiteley, M., & Diggle, S. P. (2020). Bacterial quorum sensing during infection. *Annual Review of Microbiology*, **74**, 201-219.

79. Shanker, E., & Federle, M. J. (2017). Quorum sensing regulation of competence and bacteriocins in Streptococcus pneumoniae and mutans. *Genes*, **8(1)**, 15.

80. Vidal, J. E., Howery, K. E., Ludewick, H. P., Nava, P., & Klugman, K. P. (2013). Quorum-sensing systems LuxS/autoinducer 2 and Com regulate Streptococcus pneumoniae biofilms in a bioreactor with living cultures of human respiratory cells. *Infection and immunity*, **81(4)**, 1341-1353.

81. Тутельян, А. В., Романова, Ю. М., Маневич, Б. В., Юшина, Ю. К., Федорова, Л. С., Синицына, О. А., Синицын, А.П. & Емшанов, О. В. (2020). Методы борьбы с биологическими пленками на пищевых производствах. *Молочная промышленность*, **10**, 4-9.

82. Романова, Ю. М., Тутельян, А. В., Синицын, А. П., Писарев, В. М., Алексеева, Н. В., Филипова, Н. И., Толордава, Э. Р., Синицына, О. А. & Емшанов, О. В. (2020). Ферменты из группы карбогидраз разрушают структуру матрикса биопленок грамположительных и грамотрицательных бактерий. *Медицинский алфавит*, **4(34)**, 40-45.

83. Nobbs, A. H., Lamont, R. J., & Jenkinson, H. F. (2009). Streptococcus adherence and colonization. *Microbiology and Molecular Biology Reviews*, **73(3)**, 407-450.

84. Cowley, L. A., Petersen, F. C., Junges, R., Jimenez, M. J. D., Morrison, D. A., & Hanage, W. P. (2018). Evolution via recombination: cell-to-cell contact facilitates larger recombination events in Streptococcus pneumoniae. *PLoS genetics*, **14(6)**, e1007410.

85. Domenech, M., García, E., & Moscoso, M. (2012). Biofilm formation in Streptococcus pneumoniae. *Microbial biotechnology*, **5(4)**, 455-465.

86. Vyas, H. K., Indraratna, A. D., Everest-Dass, A., Packer, N. H., De Oliveira, D. M., Ranson, M., McArthur, J. D. & Sanderson-Smith, M. L. (2020). Assessing the role of pharyngeal cell surface glycans in Group A Streptococcus biofilm formation. *Antibiotics*, **9(11)**, 775.

87. Costerton, J. W., Stewart, P. S., & Greenberg, E. P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science*, **284(5418)**, 1318-1322.

88. Loughran, A. J., Orihuela, C. J., & Tuomanen, E. I. (2019). Streptococcus pneumoniae: invasion and inflammation. *Microbiology spectrum*, **7(2)**, 7-15.

89. Bogaert, D., de Groot, R., & Hermans, P. W. M. (2004). Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *The Lancet infectious diseases*, **4(3)**, 144-154.

90. Marks, L. R., Davidson, B. A., Knight, P. R., & Hakansson, A. P. (2013). Interkingdom signaling induces Streptococcus pneumoniae biofilm dispersion and transition from asymptomatic colonization to disease. *MBio*, **4(4)**, e00438-13.

91. Cundell, D. R., Gerard, N. P., Gerard, C., Idanpaan-Heikkila, I., & Tuomanen, E. I. (1995). Streptococcus pneumoniae anchor to activated human cells by the receptor for platelet-activating factor. *Nature*, **377(6548)**, 435.

92. Rosenow, C., Ryan, P., Weiser, J. N., Johnson, S., Fontan, P., Ortqvist, A., & Masure, H. R. (1997). Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of Streptococcus pneumoniae. *Molecular microbiology*, **25(5)**, 819-829.

93. Bergmann, S., Rohde, M., Preissner, K. T., & Hammerschmidt, S. (2005). The nine residue plasminogen-binding motif of the pneumococcal enolase is the major cofactor of plasmin-mediated

degradation of extracellular matrix, dissolution of fibrin and transmigration. *Thrombosis and haemostasis*, **94(08)**, 304-311.

94. Attali, C., Durmort, C., Vernet, T., & Di Guilmi, A. M. (2008). The interaction of Streptococcus pneumoniae with plasmin mediates transmigration across endothelial and epithelial monolayers by intercellular junction cleavage. *Infection and immunity*, **76(11)**, 5350-5356.

95. Weiser, J. N., Ferreira, D. M., & Paton, J. C. (2018). Streptococcus pneumoniae: transmission, colonization and invasion. *Nature Reviews Microbiology*, **16(6)**, 355.

96. Pettigrew, M. M., Fennie, K. P., York, M. P., Daniels, J., & Ghaffar, F. (2006). Variation in the presence of neuraminidase genes among Streptococcus pneumoniae isolates with identical sequence types. *Infection and immunity*, **74(6)**, 3360-3365.

97. Camara, M., Boulnois, G. J., Andrew, P. W., & Mitchell, T. J. (1994). A neuraminidase from Streptococcus pneumoniae has the features of a surface protein. *Infection and immunity*, **62(9)**, 3688-3695.

98. Xu, G., Potter, J. A., Russell, R. J., Oggioni, M. R., Andrew, P. W., & Taylor, G. L. (2008). Crystal structure of the NanB sialidase from Streptococcus pneumoniae. *Journal of molecular biology*, **384(2)**, 436-449.

99. Hsiao, Y. S., Parker, D., Ratner, A. J., Prince, A., & Tong, L. (2009). Crystal structures of respiratory pathogen neuraminidases. *Biochemical and biophysical research communications*, **380(3)**, 467-471.

100. Xu, G., Li, X., Andrew, P. W., & Taylor, G. L. (2008). Structure of the catalytic domain of Streptococcus pneumoniae sialidase NanA. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, **64(9)**, 772-775.

101. Gut, H., Xu, G., Taylor, G. L., & Walsh, M. A. (2011). Structural basis for Streptococcus pneumoniae NanA inhibition by influenza antivirals zanamivir and oseltamivir carboxylate. *Journal of molecular biology*, **409(4)**, 496-503.

102. Connaris, H., Govorkova, E. A., Ligertwood, Y., Dutia, B. M., Yang, L., Tauber, S., Taylor, M.A., Alias, N., Hagan, R., Nash, A.A., Webster, R. G. & Taylor, G.L. (2014). Prevention of influenza by targeting host receptors using engineered proteins. *Proceedings of the National Academy of Sciences*, **111(17)**, 6401-6406.

103. Yang, L., Connaris, H., Potter, J. A., & Taylor, G. L. (2015). Structural characterization of the carbohydrate-binding module of NanA sialidase, a pneumococcal virulence factor. *BMC structural biology*, **15(1)**, 15.

104. Xu, G., Kiefel, M. J., Wilson, J. C., Andrew, P. W., Oggioni, M. R., & Taylor, G. L. (2011). Three Streptococcus pneumoniae sialidases: three different products. *Journal of the American Chemical Society*, **133(6)**, 1718-1721.

105. Xiao, K., Wang, X., & Yu, H. (2019). Comparative studies of catalytic pathways for Streptococcus pneumoniae sialidases NanA, NanB and NanC. *Scientific reports*, **9(1)**, 1-13.

106. Owen, C. D., Lukacik, P., Potter, J. A., Sleator, O., Taylor, G. L., & Walsh, M. A. (2015). Streptococcus pneumoniae NanC: structural insights into the specificity and mechanism of a sialidase that produces a sialidase inhibitor. *Journal of Biological Chemistry*, **290(46)**, 27736-27748.

107. Thobhani, S., Ember, B., Siriwardena, A., & Boons, G. J. (2003). Multivalency and the mode of action of bacterial sialidases. *Journal of the American Chemical Society*, **125(24)**, 7154-7155.

108. Moustafa, I., Connaris, H., Taylor, M., Zaitsev, V., Wilson, J. C., Kiefel, M. J., von Itzstein, M. & Taylor, G. (2004). Sialic acid recognition by Vibrio cholerae neuraminidase. *Journal of Biological Chemistry*, **279(39)**, 40819-40826.

109. Hammond, A. J., Binsker, U., Aggarwal, S. D., Ortigoza, M. B., Loomis, C., & Weiser, J. N. (2021). Neuraminidase B controls neuraminidase A-dependent mucus production and evasion. *PLoS pathogens*, **17(4)**, e1009158.

110. Orihuela, C. J., Gao, G., Francis, K. P., Yu, J., & Tuomanen, E. I. (2004). Tissue-specific contributions of pneumococcal virulence factors to pathogenesis. *Journal of Infectious Diseases*, **190(9)**, 1661-1669.

111. Tong, H. H., Blue, L. E., James, M. A., & DeMaria, T. F. (2000). Evaluation of the virulence of a Streptococcus pneumoniae neuraminidase-deficient mutant in nasopharyngeal colonization and development of otitis media in the chinchilla model. *Infection and immunity*, **68(2)**, 921-924.

112. Manco, S., Hernon, F., Yesilkaya, H., Paton, J. C., Andrew, P. W., & Kadioglu, A. (2006). Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. *Infection and immunity*, **74(7)**, 4014-4020.

113. King, S. J., Hippe, K. R., Gould, J. M., Bae, D., Peterson, S., Cline, R. T., Fasching, C., Janoff, E.N. & Weiser, J. N. (2004). Phase variable desialylation of host proteins that bind to Streptococcus pneumoniae in vivo and protect the airway. *Molecular microbiology*, **54(1)**, 159-171.

114. Tong, H. H., McIver, M. A., Fisher, L. M., & DeMaria, T. F. (1999). Effect of lacto-N-neotetraose, asialoganglioside-GM1 and neuraminidase on adherence of otitis media-associated serotypes of Streptococcus pneumoniae to chinchilla tracheal epithelium. *Microbial pathogenesis*, **26(2)**, 111-119.

115. Tong, H. H., Liu, X., Chen, Y., James, M., & Demaria, T. (2002). Effect of neuraminidase on receptor-mediated adherence of Streptococcus pneumoniae to chinchilla tracheal epithelium. *Acta oto-laryngologica*, **122(4)**, 413-419.

116. Brittan, J. L., Buckeridge, T. J., Finn, A., Kadioglu, A., & Jenkinson, H. F. (2012). Pneumococcal neuraminidase A: an essential upper airway colonization factor for Streptococcus pneumoniae. *Molecular oral microbiology*, **27(4)**, 270-283.

117. King, S. J., Hippe, K. R., & Weiser, J. N. (2006). Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by Streptococcus pneumoniae. *Molecular microbiology*, **59(3)**, 961-974.

118. Burnaugh, A. M., Frantz, L. J., & King, S. J. (2008). Growth of Streptococcus pneumoniae on human glycoconjugates is dependent upon the sequential activity of bacterial exoglycosidases. *Journal of bacteriology*, **190(1)**, 221-230.

119. Marion, C., Burnaugh, A. M., Woodiga, S. A., & King, S. J. (2011). Sialic acid transport contributes to pneumococcal colonization. *Infection and immunity*, **79(3)**, 1262-1269.

120. Hilleringmann, M., Kohler, S., Gámez, G. & Hammerschmidt, S. (2015). Pneumococcal Pili and Adhesins. In *Streptococcus pneumoniae: Molecular Mechanisms of Host-Pathogen Interactions.* Brown, J., Hammerschmidt, S. & Orihuela, C. (eds). Academic Press.

121. Parker, D., Soong, G., Planet, P., Brower, J., Ratner, A. J., & Prince, A. (2009). The NanA neuraminidase of Streptococcus pneumoniae is involved in biofilm formation. *Infection and immunity*, **77(9)**, 3722-3730.

122. Trappetti, C., Kadioglu, A., Carter, M., Hayre, J., Iannelli, F., Pozzi, G., Andrew, P.W. & Oggioni, M. R. (2009). Sialic acid: a preventable signal for pneumococcal biofilm formation, colonization, and invasion of the host. *The Journal of infectious diseases*, **199(10)**, 1497-1505.

123. Oggioni, M. R., Trappetti, C., Kadioglu, A., Cassone, M., Iannelli, F., Ricci, S., Andrew, P.W. & Pozzi, G. (2006). Switch from planktonic to sessile life: a major event in pneumococcal pathogenesis. *Molecular microbiology*, **61(5)**, 1196-1210.

124. Blanchette, K. A., Shenoy, A. T., Milner, J., Gilley, R. P., McClure, E., Hinojosa, C. A., Kumar, N, Daugherty, S.C., Tallon, L.J., Ott, S., King, S.J., Ferreira, D.M., Gordon, S.B., Tettelin, H. & Orihuela, C.J. (2016). Neuraminidase A-Exposed Galactose Promotes Streptococcus pneumoniae Biofilm Formation during Colonization. *Infection and immunity*, **84(10)**, 2922-2932.

125. Wren, J. T., Blevins, L. K., Pang, B., Roy, A. B., Oliver, M. B., Reimche, J. L., Wozniak, J.E., Alexander-Miller, M.A. & Swords, W. E. (2017). Pneumococcal neuraminidase A (NanA) promotes biofilm formation and synergizes with influenza A virus in nasal colonization and middle ear infection. *Infection and immunity*, **85(4)**, e01044-16.

126. Limoli, D. H., Sladek, J. A., Fuller, L. A., Singh, A. K., & King, S. J. (2011). BgaA acts as an adhesin to mediate attachment of some pneumococcal strains to human epithelial cells. *Microbiology*, **157(8)**, 2369.

127. Uchiyama S, Carlin AF, Khosravi A, Weiman S, Banerjee A, Quach D, Hightower G, Mitchell TJ, Doran KS & Nizet V (2009) The surface-anchored NanA protein promotes pneumococcal brain endothelial cell invasion. *Journal of Experimental Medicine*, **206(9)**, 1845-1852.

128. Banerjee A, Van Sorge NM, Sheen TR, Uchiyama S, Mitchell TJ & Doran KS (2010) Activation of brain endothelium by pneumococcal neuraminidase NanA promotes bacterial internalization. *Cell Microbiology*, **12(11)**, 1576-1588.

129. Gratz, N., Loh, L. N., Mann, B., Gao, G., Carter, R., Rosch, J., & Tuomanen, E. I. (2017). Pneumococcal neuraminidase activates TGF-β signalling. *Microbiology*, **163(8)**, 1198-1207.

130. Anil, A., & Banerjee, A. (2020). Pneumococcal Encounter With the Blood–Brain Barrier Endothelium. *Frontiers in cellular and infection microbiology*, **10**, 614.

131. Hatcher, B. L., Hale, J. Y., & Briles, D. E. (2016). Free Sialic Acid Acts as a Signal That Promotes Streptococcus pneumoniae Invasion of Nasal Tissue and Nonhematogenous Invasion of the Central Nervous System. *Infection and immunity*, **84(9)**, 2607-2615

132. Tseng, Y. W., Chang, C. C., & Chang, Y. C. (2021). Novel Virulence Role of Pneumococcal NanA in Host Inflammation and Cell Death Through the Activation of Inflammasome and the Caspase Pathway. *Frontiers in Cellular and Infection Microbiology*, **11**, 125.

133. Cauley, L. S., & Vella, A. T. (2015). Why is co-infection with influenza virus and bacteria so difficult to control? *Discovery medicine*, **19(102)**, 33.

134. Rudd, J. M., Ashar, H. K., Chow, V. T., & Teluguakula, N. (2016). Lethal Synergism between Influenza and Streptococcus pneumoniae. *Journal of infectious pulmonary diseases*, **2(2)**.

135. Lee, K. H., Gordon, A., & Foxman, B. (2016). The role of respiratory viruses in the etiology of bacterial pneumonia: an ecological perspective. *Evolution, medicine, and public health*, **2016(1)**, 95-109.

136. Martin-Loeches, I., van Someren Gréve, F., & Schultz, M. J. (2017). Bacterial pneumonia as an influenza complication. *Current opinion in infectious diseases*, **30(2)**, 201-207.

137. Dutta, A., Chen, Y. Y., Chen, T. C., Chang, C. S., Huang, Y. L., Chen, T. A., Lin, Y. C., Lin, C. Y., Hsieh, Y. C. & Huang, C. T. (2021). Role of Pneumococcal NanC in the Severe Disease of Streptococcus pneumoniae Superinfection with Influenza. *ImmunoHorizons*, **5(4)**, 210-218.

138. Short, K. R., Habets, M. N., Hermans, P. W., & Diavatopoulos, D. A. (2012). Interactions between Streptococcus pneumoniae and influenza virus: a mutually beneficial relationship? *Future microbiology*, **7(5)**, 609-624.

139. LeMessurier, K. S., Tiwary, M., Morin, N. P., & Samarasinghe, A. E. (2020). Respiratory barrier as a safeguard and regulator of defense against influenza A virus and Streptococcus pneumoniae. *Frontiers in immunology*, **11**, 3.

140. Rowe, H. M., Meliopoulos, V. A., Iverson, A., Bomme, P., Schultz-Cherry, S., & Rosch, J. W. (2019). Direct interactions with influenza promote bacterial adherence during respiratory infections. *Nature microbiology*, **4(8)**, 1328-1336.

141. McCullers, J. A. (2006). Insights into the interaction between influenza virus and pneumococcus. *Clinical microbiology reviews*, **19(3)**, 571-582.

142. Siegel, S. J., Roche, A. M., & Weiser, J. N. (2014). Influenza promotes pneumococcal growth during coinfection by providing host sialylated substrates as a nutrient source. *Cell host & microbe*, **16(1)**, 55-67.

143. Sender, V., Hentrich, K., & Henriques-Normark, B. (2021). Virus-Induced Changes of the Respiratory Tract Environment Promote Secondary Infections With Streptococcus pneumoniae. *Frontiers in Cellular and Infection Microbiology*, **11**, 199.

144. Rossi, G. A., Fanous, H., & Colin, A. A. (2020). Viral strategies predisposing to respiratory bacterial superinfections. *Pediatric pulmonology*, **55(4)**, 1061-1073.

145. Li, N., Ren, A., Wang, X., Fan, X., Zhao, Y., Gao, G. F., Cleary, P. & Wang, B. (2015). Influenza viral neuraminidase primes bacterial coinfection through TGF-β–mediated expression of host cell receptors. *Proceedings of the National Academy of Sciences*, **112(1)**, 238-243.

146. Nishikawa, T., Shimizu, K., Tanaka, T., Kuroda, K., Takayama, T., Yamamoto, T., Hanada, N. & Hamada, Y. (2012). Bacterial neuraminidase rescues influenza virus replication from inhibition by a neuraminidase inhibitor. *PloS one*, **7(9)**, e45371.

147. von Itzstein, M., Wu, W. Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., Phan, T.V., Smythe, M.L., White, H.F., Oliver, S.W., Colman, P.M., Varghese, J.N., Ryan, D.M., Woods, J.M., Bethell, R.C., Hotham, V.J., Cameron, J.M. & Penn, C. R. (1993). Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature*, **363**, 418-423.

148. Kim, C. U., Lew, W., Williams, M. A., Liu, H., Zhang, L., Swaminathan, S., Bischofberger, N., Chen, M.S., Mendel, D.B., Tai, C.Y., Laver, W.G. & Stevens, R. C. (1997). Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *Journal of the American Chemical Society*, **119(4)**, 681-690.

149. Ilyushina, N. A., Seiler, J. P., Rehg, J. E., Webster, R. G., & Govorkova, E. A. (2010). Effect of neuraminidase inhibitor–resistant mutations on pathogenicity of clade 2.2 A/Turkey/15/06 (H5N1) influenza virus in ferrets. *PLoS pathogens*, **6(5)**, e1000933.

150. Eisenberg, E. J. (2007). in *Prodrugs* (Stella V. J. et al., ed.). Springer New York, pp. 1323-1334.

151. Birnkrant, D. & Cox, E. (2009). The emergency use authorization of peramivir for treatment of 2009 H1N1 influenza. *New England Journal of Medicine*, **361(23)**, 2204-2207.

152. Ikematsu, H. & Kawai, N. (2011). Laninamivir octanoate: a new long-acting neuraminidase inhibitor for the treatment of influenza. *Expert Review of Anti-Infective Therapy*, **9(10)**, 851-857.

153. Thorlund, K., Awad, T., Boivin, G., & Thabane, L. (2011). Systematic review of influenza resistance to the neuraminidase inhibitors. *BMC infectious diseases*, **11(1)**, 134.

154. McKimm-Breschkin, J. L. (2013). Influenza neuraminidase inhibitors: antiviral action and mechanisms of resistance. *Influenza and other respiratory viruses*, **7**, 25-36.

155. Leang, S. K., Kwok, S., Sullivan, S. G., Maurer-Stroh, S., Kelso, A., Barr, I. G., & Hurt, A. C. (2014). Peramivir and laninamivir susceptibility of circulating influenza A and B viruses. *Influenza and other respiratory viruses*, **8(2)**, 135-139.

156. McKimm-Breschkin, J. L., & Barrett, S. (2015). Neuraminidase mutations conferring resistance to laninamivir lead to faster drug binding and dissociation. *Antiviral research*, **114**, 62-66.

157. Mahal, A., Duan, M., Zinad, D. S., Mohapatra, R. K., Obaidullah, A. J., Wei, X., Pradhan, M. K., Das, D., Kandi, V., Zinad, H. S. & Zhu, Q. (2021). Recent progress in chemical approaches for the development of novel neuraminidase inhibitors. *RSC Advances*, **11(3)**, 1804-1840.

158. Cheng, L. S., Amaro, R. E., Xu, D., Li, W. W., Arzberger, P. W., & McCammon, J. A. (2008). Ensemble-based virtual screening reveals potential novel antiviral compounds for avian influenza neuraminidase. *Journal of medicinal chemistry*, **51(13)**, 3878-3894.

159. Zhao, Z. X., Cheng, L. P., Li, M., Pang, W., & Wu, F. H. (2019). Discovery of novel acylhydrazone neuraminidase inhibitors. *European journal of medicinal chemistry*, **173**, 305-313.

160. La Frazia, S., Piacentini, S., Riccio, A., Rossignol, J. F., & Santoro, M. G. (2018). The second-generation thiazolide haloxanide is a potent inhibitor of avian influenza virus replication. *Antiviral research*, **157**, 159-168.

161. Shin, W. J., & Seong, B. L. (2019). Novel antiviral drug discovery strategies to tackle drug-resistant mutants of influenza virus strains. *Expert opinion on drug discovery*, **14(2)**, 153-168.

162. Heaton, N. S., Sachs, D., Chen, C. J., Hai, R., & Palese, P. (2013). Genome-wide mutagenesis of influenza virus reveals unique plasticity of the hemagglutinin and NS1 proteins. *Proceedings of the National Academy of Sciences*, **110(50)**, 20248-20253.

163. Stadlbauer, D., Zhu, X., McMahon, M., Turner, J. S., Wohlbold, T. J., Schmitz, A. J., Strohmeier, S., Yu, W., Nachbagauer, R., Mudd, P. A., Wilson, I. A., Ellebedy, A. H. & Krammer, F. (2019). Broadly protective human antibodies that target the active site of influenza virus neuraminidase. *Science*, **366(6464)**, 499-504.

164. Giurgea, L. T., Morens, D. M., Taubenberger, J. K., & Memoli, M. J. (2020). Influenza Neuraminidase: A Neglected Protein and Its Potential for a Better Influenza Vaccine. *Vaccines*, **8(3)**, 409.

165. Hamborsky, J., Kroger, A., Wolfe, S., eds. 13th ed. (2015). Influenza. In *Centers for Disease Control and Prevention. Epidemiology and Prevention of Vaccine-Preventable Diseases*. Washington, DC: Public Health Foundation

166. Hamborsky, J., Kroger, A., Wolfe, S., eds. 13th ed. (2015). Pneumococcal Disease. In *Centers for Disease Control and Prevention. Epidemiology and Prevention of Vaccine-Preventable Diseases*. Washington, DC: Public Health Foundation

167. Kadioglu, A., Weiser, J. N., Paton, J. C., & Andrew, P. W. (2008). The role of Streptococcus pneumoniae virulence factors in host respiratory colonization and disease. *Nature Reviews Microbiology*, **6(4)**, 288-301.

168. Hyams, C., Camberlein, E., Cohen, J. M., Bax, K., & Brown, J. S. (2010). The Streptococcus pneumoniae capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. *Infection and immunity*, **78(2)**, 704-715.

169. Bittaye, M., & Cash, P. (2015). Streptococcus pneumoniae proteomics: determinants of pathogenesis and vaccine development. *Expert review of proteomics*, **12(6)**, 607-621.

170. Pichichero, M. E., Khan, M. N., & Xu, Q. (2016). Next generation protein based Streptococcus pneumoniae vaccines. *Human vaccines & immunotherapeutics*, **12(1)**, 194-205.

171. Chan, W. Y., Entwisle, C., Ercoli, G., Ramos-Sevillano, E., McIlgorm, A., Cecchini, P., Bailey, C., Lam, O., Whiting, G., Green, N., Goldblatt, D., Wheeler, J. X. & Brown, J. S. (2019). A novel, multiple-antigen pneumococcal vaccine protects against lethal Streptococcus pneumoniae challenge. *Infection and immunity*, **87(3)**, e00846-18.

172. Guitor, A. K., & Wright, G. D. (2018). Antimicrobial resistance and respiratory infections. *Chest*, **154(5)**, 1202-1212.

173. Young, N., & Thomas, M. (2018). Meningitis in adults: diagnosis and management. *Internal medicine journal*, **48(11)**, 1294-1307.

174. Peyrani, P., Mandell, L., Torres, A., & Tillotson, G. S. (2019). The burden of community-acquired bacterial pneumonia in the era of antibiotic resistance. *Expert review of respiratory medicine*, **13(2)**, 139-152.

175. Cho, J. C., Childs-Kean, L. M., Zmarlicka, M. T., & Crotty, M. P. (2018). Return of the tetracyclines: omadacycline, a novel aminomethylcycline antimicrobial. *Drugs of today (Barcelona, Spain: 1998)*, **54(3)**, 209-217.

176. Cools, F., Delputte, P., & Cos, P. (2021). The search for novel treatment strategies for Streptococcus pneumoniae infections. *FEMS Microbiology Reviews*, **fuaa072**, 1-23.

177. Ishii, S., Fukui, K., Yokoshima, S., Kumagai, K., Beniyama, Y., Kodama, T., Fukuyama, T., Okabe, T., Nagano, T., Kojima, H. & Yano, T. (2017). High-throughput screening of small molecule inhibitors of the Streptococcus quorum-sensing signal pathway. *Scientific reports*, **7(1)**, 4029.

178. Zimmerman, T., & Ibrahim, S. (2017). Choline kinase, a novel drug target for the inhibition of Streptococcus pneumoniae. *Antibiotics*, **6(4)**, 20.

179. Wang, J., Song, M., Pan, J., Shen, X., Liu, W., Zhang, X., Li, H. & Deng, X. (2018). Quercetin impairs Streptococcus pneumoniae biofilm formation by inhibiting sortase A activity. *Journal of cellular and molecular medicine*, **22(12)**, 6228-6237.

180. Bhattacharyya, P., Agarwal, B., Goswami, M., Maiti, D., Baruah, S., & Tribedi, P. (2018). Zinc oxide nanoparticle inhibits the biofilm formation of Streptococcus pneumoniae. *Antonie Van Leeuwenhoek*, **111(1)**, 89-99.

181. Walther, E., Xu, Z., Richter, M., Kirchmair, J., Grienke, U., Rollinger, J. M., Krumbholz, A., Saluz, H.P., Pfister, W., Sauerbrei, A. & Schmidtke, M. (2016). Dual acting neuraminidase inhibitors open new opportunities to disrupt the lethal synergism between Streptococcus pneumoniae and influenza virus. *Frontiers in microbiology*, **7**, 357.

182. Grienke, U., Richter, M., Walther, E., Hoffmann, A., Kirchmair, J., Makarov, V., Nietzsche, S., Schmidtke, M. & Rollinger, J. M. (2016). Discovery of prenylated flavonoids with dual activity against influenza virus and Streptococcus pneumoniae. *Scientific reports*, **6**, 27156.

183. Hoffmann, A., Richter, M., von Grafenstein, S., Walther, E., Xu, Z., Schumann, L., Grienke, U., Mair, C.E., Kramer, C., Rollinger, J.M., Liedl, K. R., Schmidtke, M. & Kirchmair, J. (2017). Discovery and Characterization of Diazenylaryl Sulfonic Acids as Inhibitors of Viral and Bacterial Neuraminidases. *Frontiers in microbiology*, **8**, 205.

184. Krissinel, E., & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D: Biological Crystallography*, **60(12)**, 2256-2268.

185. Suplatov, D. A., Kopylov, K. E., Popova, N. N., Voevodin, V. V., & Švedas, V. K. (2018). Mustguseal: a server for multiple structure-guided sequence alignment of protein families. *Bioinformatics*, **34(9)**, 1583-1585.

186. Suplatov, D., Sharapova, Y., & Švedas, V. (2021). Mustguseal and Sister Web-Methods: A Practical Guide to Bioinformatic Analysis of Protein Superfamilies. In *Multiple Sequence Alignment* (pp. 179-200). Humana, New York, NY.

187. Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein engineering*, **12(2)**, 85-94.

188. Vouzis, P.D., and Sahinidis, N.V. (2011). GPU-BLAST: using graphics processors to accelerate protein sequence alignment. *Bioinformatics*, **27(2)**, 182-188

189. Fischer, J.D., Mayer, C.E., and Söding, J. (2008). Prediction of protein functional residues from sequence by probability density estimation. *Bioinformatics*, **24(5)**, 613-620.

190. Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*, **30(4)**, 772-780.

191. Suplatov, D., Kirilin, E. & Švedas, V. "Bioinformatic analysis of protein families to select function-related variable positions." In *Understanding Enzymes* (Ed. Allan Svendsen), pp. 375-410. Jenny Stanford Publishing, 2016.

192. Suplatov, D., Sharapova, Y., Geraseva, E., & Švedas, V. (2020). Zebra2: advanced and easy-to-use web-server for bioinformatic analysis of subfamily-specific and conserved positions in diverse protein superfamilies. *Nucleic acids research*, **48(W1)**, W65-W71.

193. Suplatov, D., Kirilin, E., Takhaveev, V., & Švedas, V. (2014). Zebra: a web server for bioinformatic analysis of diverse protein families. *Journal of Biomolecular Structure and Dynamics*, **32(11)**, 1752-1758.

194. Suplatov, D., Shalaeva, D., Kirilin, E., Arzhanik, V., & Švedas, V. (2014). Bioinformatic analysis of protein families for identification of variable amino acid residues responsible for functional diversity. *Journal of Biomolecular Structure and Dynamics*, **32(1)**, 75-87.

195. Dupradeau, F. Y., Pigache, A., Zaffran, T., Savineau, C., Lelong, R., Grivel, N., Lelong, D., Rosanski, W. & Cieplak, P. (2010). The R.E.D. tools: Advances in RESP and ESP charge derivation and force field library building. *Physical Chemistry Chemical Physics*, **12(28)**, 7821-7839.

196. Vanquelef, E., Simon, S., Marquant, G., Garcia, E., Klimerak, G., Delepine, J. C., Cieplak, P. & Dupradeau, F. Y. (2011). RED Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments. *Nucleic acids research*, **39(suppl_2)**, W511-W517.

197. Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., & Baker, N. A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic acids research*, **35(suppl_2)**, W522-W525.

198. Krossa, S., Faust, A., Ober, D., & Scheidig, A. J. (2016). Comprehensive structural characterization of the bacterial homospermidine synthase–an essential enzyme of the polyamine metabolism. *Scientific reports*, *6*, 19501.

199. Søndergaard, C. R., Olsson, M. H., Rostkowski, M., & Jensen, J. H. (2011). Improved treatment of ligands and coupling effects in empirical calculation and rationalization of p K a values. *Journal of chemical theory and computation*, **7(7)**, 2284-2295.

200. Sharapova, Y., Suplatov, D., & Švedas, V. (2018). Neuraminidase A from Streptococcus pneumoniae has a modular organization of catalytic and lectin domains separated by a flexible linker. *The FEBS journal*, **285(13)**, 2428-2445.

201. Sharapova, Y., Švedas, V., & Suplatov, D. (2021). Catalytic and lectin domains in neuraminidase A from Streptococcus pneumoniae are capable of an intermolecular assembly: Implications for biofilm formation. *The FEBS Journal*, **288(10)**, 3217-3230.

202. Шарапова, Я. А., & Швядас В. К. (2018). Молекулярное моделирование связывания аллостерического ингибитора оптактина в новом сайте в структуре нейраминидазы a из Streptococcus pneumoniae. *Вестник Московского университета. Серия 2. Химия*, **59(5)**, 323-331.

203. Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of chemical theory and computation*, **11(8)**, 3696-3713.

204. Wong, V., & Case, D. A. (2008). Evaluating rotational diffusion from protein MD simulations. *The Journal of Physical Chemistry B*, **112(19)**, 6013-6024.

205. Horn, H. W., Swope, W. C., Pitera, J. W., Madura, J. D., Dick, T. J., Hura, G. L., & Head-Gordon, T. (2004). Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of chemical physics*, **120(20)**, 9665-9678.

206. Debiec, K. T., Cerutti, D. S., Baker, L. R., Gronenborn, A. M., Case, D. A., & Chong, L. T. (2016). Further along the road less traveled: AMBER ff15ipq, an original protein force field built on a self-consistent physical model. *Journal of chemical theory and computation*, **12(8)**, 3926-3947.

207. Takemura, K., & Kitao, A. (2012). Water model tuning for improved reproduction of rotational diffusion and NMR spectral density. *The Journal of Physical Chemistry B*, **116(22)**, 6279-6287.

208. Суплатов, Д. А., Шарапова, Я. А., Попова, Н. Н., Копылов, К. Е., Воеводин, В. В., & Каятоно, Ш. В. (2019). Молекулярная динамика в силовом поле FF14SB в воде TIP4P-Ew, и в силовом поле FF15IPQ в воде SPC/Eb: сравнительный анализ на GPU и CPU. *Вестник Южно-Уральского государственного университета. Серия: Вычислительная математика и информатика*, **8(1)**, 705-716.

209. Sharapova, Y. A., Suplatov, D. A., & Švedas, V. K. (2018). Simulating the long-timescale structural behavior of bacterial and influenza neuraminidases with different HPC resources. *Supercomputing Frontiers and Innovations*, **5(3)**, 30-33.

210. Pierce, L. C., Salomon-Ferrer, R., Augusto F. de Oliveira, C., McCammon, J. A., & Walker, R. C. (2012). Routine access to millisecond time scale events with accelerated molecular dynamics. *Journal of chemical theory and computation*, **8(9)**, 2997-3002.

211. Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N· log (N) method for Ewald sums in large systems. *The Journal of chemical physics*, **98(12)**, 10089-10092.

212. Götz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., & Walker, R. C. (2012). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized born. *Journal of chemical theory and computation*, **8(5)**, 1542-1555.

213. Le Grand, S., Götz, A. W., & Walker, R. C. (2013). SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations. *Computer Physics Communications*, **184(2)**, 374-380.

214. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *Journal of chemical theory and computation*, **9(9)**, 3878-3888.

215. Buschiazzo, A., Tavares, G. A., Campetella, O., Spinelli, S., Cremona, M. L., París, G., Amaya, M.F., Frasch, A.C. & Alzari, P. M. (2000). Structural basis of sialyltransferase activity in trypanosomal sialidases. *The EMBO Journal*, **19(1)**, 16-24.

216. Luo, Y., Li, S. C., Li, Y. T., & Luo, M. (1999). The 1.8 Å structures of leech intramolecular trans-sialidase complexes: evidence of its enzymatic mechanism. *Journal of molecular biology*, **285(1)**, 323-332.

217. Pierce, B. G., Wiehe, K., Hwang, H., Kim, B. H., Vreven, T., & Weng, Z. (2014). ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics*, **30(12)**, 1771-1773.

218. Menke, M., Berger, B., & Cowen, L. (2008). Matt: local flexibility aids protein multiple structure alignment. *PLoS computational biology*, **4(1)**, e10.

219. Kalaimathy, S., Sowdhamini, R., & Kanagarajadurai, K. (2011). Critical assessment of structure-based sequence alignment methods at distant relationships. *Briefings in bioinformatics*, **12(2)**, 163-175.

220. Šali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology*, **234(3)**, 779-815.

221. Drozdetskiy, A., Cole, C., Procter, J., & Barton, G. J. (2015). JPred4: a protein secondary structure prediction server. *Nucleic acids research*, **43(W1)**, W389-W394.

222. Shen, M. Y., & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein science*, **15(11)**, 2507-2524.

223. Xu, Z., Von Grafenstein, S., Walther, E., Fuchs, J. E., Liedl, K. R., Sauerbrei, A., & Schmidtke, M. (2016). Sequence diversity of NanA manifests in distinct enzyme kinetics and inhibitor susceptibility. *Scientific reports*, **6**, 25169.

224. Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D. & Vajda, S. (2017). The ClusPro web server for protein–protein docking. *Nature protocols*, **12(2)**, 255.

225. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry*, **19(14)**, 1639-1662.

226. Stroganov, O. V., Novikov, F. N., Stroylov, V. S., Kulkov, V., & Chilov, G. G. (2008). Leadfinder: an approach to improve accuracy of protein− ligand docking, binding energy estimation, and virtual screening. *Journal of chemical information and modeling*, **48(12)**, 2371-2385.

227. Суплатов, Д. А., & Швядас, В. К. (2015). Изучение функциональных и аллостерических сайтов в суперсемействах белков. *Acta Naturae*, **7(4 (27))**, 39-52.

228. Le Guilloux, V., Schmidtke, P., & Tuffery, P. (2009). Fpocket: an open source platform for ligand pocket detection. *BMC bioinformatics*, **10(1)**, 1-11.

229. Suplatov, D., Sharapova, Y., Timonina, D., Kopylov, K., & Švedas, V. (2018). The visualCMAT: A web-server to select and interpret correlated mutations/co-evolving residues in protein families. *Journal of bioinformatics and computational biology*, **16(02)**, 1840005.

230. Shegay, M. V., Suplatov, D. A., Popova, N. N., Švedas, V. K., & Voevodin, V. V. (2019). parMATT: parallel multiple alignment of protein 3D-structures with translations and twists for distributed-memory systems. *Bioinformatics*, **35(21)**, 4456-4458.

231. Suplatov, D., Timonina, D., Sharapova, Y., & Švedas, V. (2019). Yosshi: a web-server for disulfide engineering by bioinformatic analysis of diverse protein families. *Nucleic acids research*, **47(W1)**, W308-W314.

232. Suplatov, D., Kopylov, K., Sharapova, Y., & Švedas, V. (2019). Human p38α Mitogen-Activated Protein Kinase in the Asp168-Phe169-Gly170-in (DFG-in) state can bind allosteric inhibitor Doramapimod. *Journal of Biomolecular Structure and Dynamics*, **37(8)**, 2049-2060.

233. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot C., Skeel, R. D., Kale, L. & Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of computational chemistry*, **26(16)**, 1781-1802.

234. Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25(9)**, 1189-1191.

235. Ilyushina, N. A., Bovin, N. V., Webster, R. G., & Govorkova, E. A. (2006). Combination chemotherapy, a potential strategy for reducing the emergence of drug-resistant influenza A variants. *Antiviral research*, **70(3)**, 121-131.

236. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S.E., Wilkins, M.R., Appel, R.D. & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook* (Walker, J.M., ed), pp. 571-607. Humana Press.

237. Kycia, J.H., Biou, V., Shu, F., Gerchman, S.E., Graziano, V. & Ramakrishnan, V. (1995). Prokaryotic translation initiation factor IF3 is an elongated protein consisting of two crystallizable domains. *Biochemistry*, **34(18)**, 6183-6187.

238. Lunde, B. M., Moore, C., & Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nature reviews Molecular cell biology, **8(6)**, 479.

239. Булахов, А. Г., Гусаков, А. В., Рожкова, А. М., Волков, П. В., Матыс, В. Ю., Зоров, И. Н., & Синицын, А. П. (2017). Свойства химерной полисахаридмонооксигеназы с присоединенным целлюлозосвязывающим модулем и ее применение для гидролиза целлюлозосодержащего сырья в составе целлюлазного комплекса. *Катализ в промышленности*, **17(6)**, 554-560.

240. Gill, J., Rixon, J.E., Bolam, D.N., McQueen-Mason, S., Simpson, P.J., Williamson, M.P., Hazlewood, G.P. & Gilbert, H.J. (1999). The type II and X cellulose-binding domains of Pseudomonas

xylanase A potentiate catalytic activity against complex substrates by a common mechanism. *Biochemical journal*, **342(2)**, 473-480.

241. Rabinovich, M.L., Melnick, M.S. & Bolobova, A.V. (2002). The structure and mechanism of action of cellulolytic enzymes. *Biochemistry (Moscow)*, **67(8)**, 850-871.

242. Beckham, G.T., Bomble, Y.J., Matthews, J.F., Taylor, C.B., Resch, M.G., Yarbrough, J.M., Decker, S.R., Bu, L., Zhao, X., McCabe, C., Wohlert, J., Bergenstråhle, M., Brady, J.W., Adney, W.S., Himmel, M.E. & Crowley, M.F. (2010). The O-glycosylated linker from the Trichoderma reesei Family 7 cellulase is a flexible, disordered protein. *Biophysical journal*, **99(11)**, 3773-3781.

243. Gaskell, A., Crennell, S. & Taylor, G. (1995). The three domains of a bacterial sialidase: a β-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure*, **3(11)**, 1197-1205.

244. Camara, M., Boulnois, G. J., Andrew, P. W., & Mitchell, T. J. (1994). A neuraminidase from Streptococcus pneumoniae has the features of a surface protein. *Infection and immunity*, **62(9)**, 3688-3695.

245. Shaw, D.E., Maragakis, P., Lindorff-Larsen, K., Piana, S., Dror, R.O., Eastwood, M.P., Bank, J.A., Jumper, J.M., Salmon, J.K., Shan, Y. & Wriggers, W. (2010). Atomic-level characterization of the structural dynamics of proteins. *Science*, **330(6002)**, 341-346.

246. Bernardi, R. C., Melo, M. C., & Schulten, K. (2015). Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta (BBA)-General Subjects*, **1850(5)**, 872-877.

247. Doshi, U., & Hamelberg, D. (2015). Towards fast, rigorous and efficient conformational sampling of biomolecules: Advances in accelerated molecular dynamics. *Biochimica et Biophysica Acta (BBA)-General Subjects*, **1850(5)**, 878-888.

248. Luo, Y., Li, S. C., Chou, M. Y., Li, Y. T., & Luo, M. (1998). The crystal structure of an intramolecular trans-sialidase with a NeuAcα2→ 3Gal specificity. *Structure*, **6(4)**, 521-530.

249. Walther, E., Richter, M., Xu, Z., Kramer, C., Von Grafenstein, S., Kirchmair, J., Grienke, U., Rollinger, J.M., Liedl, K.R., Slevogt, H., Sauerbrei, A., Saluz, H.P., Pfister, W. & Schmidtke, M. (2015). Antipneumococcal activity of neuraminidase inhibiting artocarpin. *International Journal of Medical Microbiology*, **305(3)**, 289-297.

250. Jedrzejas, M. J. (2001). Pneumococcal virulence factors: structure and function. *Microbiology and Molecular Biology Reviews*, **65(2)**, 187-207.

251. King, S. J., Whatmore, A. M., & Dowson, C. G. (2005). NanA, a neuraminidase from Streptococcus pneumoniae, shows high levels of sequence diversity, at least in part through recombination with Streptococcus oralis. *Journal of bacteriology*, **187(15)**, 5376-5386.

252. Greener, J. G., & Sternberg, M. J. (2018). Structure-based prediction of protein allostery. *Current opinion in structural biology*, **50**, 1-8.

253. Ghosh, S., & O'Connor, T. J. (2017). Beyond paralogs: the multiple layers of redundancy in bacterial pathogenesis. *Frontiers in cellular and infection microbiology*, **7**, 467.

254. Domenech, M., & García, E. (2020). The N-acetylglucosaminidase LytB of Streptococcus pneumoniae is involved in the structure and formation of biofilms. *Applied and environmental microbiology*, **86(10)**, e00280-20.

255. Nussinov, R., Tsai, C. J., Shehu, A., & Jang, H. (2019). Computational structural biology: Successes, future directions, and challenges. *Molecules*, **24(3)**, 637.

256. Eglen, R., & Reisine, T. (2011). Drug discovery and the human kinome: recent trends. *Pharmacology & therapeutics*, **130(2)**, 144-156.

257. Nussinov, R., Tsai, C. J., & Csermely, P. (2011). Allo-network drugs: harnessing allostery in cellular networks. *Trends in pharmacological sciences*, **32(12)**, 686-693.

258. Nussinov, R., & Tsai, C. J. (2013). Allostery in disease and in drug discovery. *Cell*, **153(2)**, 293-305.

259. Wenthur, C. J., Gentry, P. R., Mathews, T. P., & Lindsley, C. W. (2014). Drugs for allosteric sites on receptors. *Annual review of pharmacology and toxicology*, **54**, 165-184.

260. Suplatov, D., Kirilin, E., Arbatsky, M., Takhaveev, V., & Švedas, V. (2014). pocketZebra: a web-server for automated selection and classification of subfamily-specific binding sites by bioinformatic analysis of diverse protein families. *Nucleic acids research*, **42(W1)**, W344-W349.

261. Richter, M., Schumann, L., Walther, E., Hoffmann, A., Braun, H., Grienke, U., Rollinger, J.M., von Grafenstein, S., Liedl, K.R., Kirchmair, J., Wutzler, P., Sauerbrei, A. & Schmidtke, M. (2015). Complementary assays helping to overcome challenges for identifying neuraminidase inhibitors. *Future Virology*, **10(2)**, 77-88.