На правах рукописи

Симакова Мария Николаевна

ХАРАКТЕРИСТИКИ СТРУКТУРЫ И СВОЙСТВА БЕЛКОВ СИСТЕМ ИНФИЦИРОВАНИЯ БАКТЕРИОФАГОВ T4 И PHIKZ И НЕКОТОРЫХ МЕМБРАННЫХ БЕЛКОВ

03.01.02 – биофизика

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата физико-математических наук

Санкт-Петербург – 2014

Работа выполнена в Лаборатории молекулярной биоинженерии ФГБУН “Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова” РАН (ИБХ), г. Москва.

Научный руководитель: **Мирошников Константин Анатольевич**, доктор химических наук, зав. лабораторией молекулярной биоинженерии ИБХ РАН

Официальные оппоненты: **Заседателев Александр Сергеевич**, д.ф.-м.н., профессор, зам. директора по науке ФГБУН “Институт молекулярной биологии им. В.А. Энгельгардта” РАН

Петухов Михаил Геннадьевич, д.ф.-м.н., доцент, ведущий научный сотрудник Национального исследовательского центра “Курчатовский институт” ФГБУ “Петербургский институт ядерной физики им. Б.П. Константинова”

Ведущая организация: ФГБУН “Институт микробиологии им. С.Н. Виноградского” РАН

Защита состоится «25» ноября 2014 г. в 14-00 часов на заседании Диссертационного Совета Д 212.229.25 при Федеральном государственном автономном образовательном учреждении высшего образования “Санкт-Петербургский государственный политехнический университет” (ФГАОУ ВО “СПбПУ”) по адресу: 194021, г. Санкт-Петербург, ул. Хлопина, 5, кафедра “Медицинская физика” Института физики, нанотехнологий и телекоммуникаций, ауд. 416.

С диссертацией можно ознакомиться в библиотеке ФГАОУ ВО “Санкт-Петербургский государственный политехнический университет”.

Автореферат разослан «____»_____ 2014 г.

Ученый секретарь Диссертационного Совета,
доктор физико-математических наук: Власова Ольга Леонардовна

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы. Вирусные фибриллярные белки и литические ферменты играют основную роль в процессе инфицирования вирусом клетки-хозяина. Поэтому определение структуры таких белков, изучение их функциональных доменов и механизмов взаимодействия с клеточными рецепторами являются очень важными аспектами в развитии методов борьбы с вирусными и бактериальными инфекциями.

Такие области медицинской биоинженерии, как генная терапия, основаны на использовании вируса как носителя здорового гена или как «оружия» против раковых клеток. Поэтому изучение структур и свойств вирусных белков, их классификация, а также создание искусственных белков методами белковой инженерии не только являются интересными фундаментальными направлениями, но и имеют важное прикладное значение.

Мембранные белки являются основными участниками очень большого числа клеточных процессов. С их помощью осуществляется транспорт внутрь клетки и наружу огромного количества различных ионов, питательных веществ, отходов, гормонов, лекарств и таких больших молекул, как белки и ДНК. Мембранные белки также в значительной мере ответственны за коммуникацию между клетками и окружающей их средой и за процессы энергетического обмена.

Клетки производят огромное количество мембранных белков. Так, приблизительно 25% генов в геноме человека кодируют мембранные белки. Мутации, происходящие в мембранных белках, являются причинами возникновения большинства распространенных заболеваний, таких как болезни сердца и головного мозга, а также рак. Мембранные белки составляют приблизительно 50% возможных мишеней для применения новых лекарственных препаратов, начиная от препаратов, действующих на нервную систему, и до новой антимикробной терапии. И, наконец, мембранные белки бактерий ответственны за такое опасное явление, как множественная лекарственная устойчивость. Это определяет актуальность изучения структуры и функциональных свойств мембранных белков методами молекулярной биофизики.

Таким образом, выбранные объекты исследования объединены их важностью для биомедицины, так как они могут быть молекулярной основой лекарственных препаратов, носителями активных начал или мишенями для лекарственных воздействий.

Исследования по теме данной диссертации проводились при финансовой поддержке РФФИ по проектам № 02-04-50012-а и № 05-04-49636-а, а также по плану НИР для государственного контракта № 02.740.11.0299 по теме: “Клеточная биология и физические свойства мембранных белков”.

Целью данной работы явилось исследование структуры и свойств 3-х типов белков: фибриллярных белков бактериофага T4 *Escherichia coli,* литического фермента бактериофага phiKZ *Pseudomonas aeruginosa* и интегральных мембранных белков, в частности, бактериородопсина *Halobacterium salinarum* с использованием комбинации теоретических и экспериментальных методов.

Для этого были поставлены следующие общие для 3-х типов белков **задачи**:

1. Компьютерный анализ аминокислотных последовательностей выбранных белков и прогнозирование особенностей их вторичной структуры.

2. Создание и экспрессия соответствующих векторов, синтез целевых белков и экспериментальное исследование их структуры и свойств.

Научная новизна работы заключается в следующем.

1. С использованием наиболее оптимального преобразования Фурье установлена определенная периодичность в 8-ми исследованных белках систем инфицирования бактериофагов T4 и phiKZ (для 6-ти из них *впервые)* и в 24-х мембранных белках. Для 2-х вирусных и 14-ти мембранных белков установлено, что полученные результаты о периодичности согласуются с известными данными об их структуре.
2. *Впервые* обнаружены относительно *большие* периоды расположения аминокислот, которыми последовательности всех 6-ти исследованных фибриллярных белков бактериофага T4 делятся на 4 или 6 равных частей.
3. *Впервые* получены экспериментальные данные о ферментативной активности и структуре рекомбинантных белков, выделенных при экспрессии делеционных мутантов литического фермента бактериофага phiKZ.

   Так, было установлено, что белок gp181E, полученный в результате экспрессии делеционного мутанта, кодирующего С-концевую часть литического фермента, является гидрофобным и взаимодействует с различными штаммами клеток *Pseudomonas aeruginosa*, но не лизирует их, в отличие от белка, полученного при экспрессии делеционного мутанта gp181MA, кодирующего лизоцимный домен литического фермента.

   С помощью гель-фильтрации было установлено, что белок gp181E является не мономером, а димером или тримером.

   Анализ КД-спектров полученных рекомбинантных белков привел к выводу о том, что их вторичная структура, по-видимому, представляет собой α-спиральный coiled coil.

   Теоретическое предположение о роли С-концевого домена литического фермента (белка gp181E), как сенсорной “молекулярной” иглы, участвующей в процессе инфицирования бактериофагом phiKZ клетки, получило дополнительное экспериментальное обоснование вышеприведенными *новыми* данными о свойствах делеционных мутантов gp181.
4. Плазмида pUCBM20 и модифицированная плазмидная конструкция NTS1-1233 на основе белков MBP и TrxA *впервые* были использованы для клонирования и экспрессии бактериородопсина и двух его делеционных мутантов в *E. coli*.

Теоретическая и практическая значимость работы.

1. Данные о периодичности фибриллярных белков бактериофага T4, полученные с помощью преобразования Фурье, впоследствии были использованы при конструировании методами генной инженерии биологически активного химерного белка, состоящего из фрагмента фибриллярного белка gp37 и С-концевого домена белка gp5 бактериофага T4. Технология с использованием такого химерного белка в дальнейшем позволила преодолеть ряд проблем рекомбинатной экспрессии целевого белка gp37.
2. Полученные данные о периодичности исследованных белков 3-х типов в сочетании с результатами кристаллографических экспериментов и электронной микроскопии могут быть полезными при установлении для этих белков элементов вторичной и третичной структуры.
3. Сравнение результатов, полученных в данной работе путем компьютерного анализа аминокислотных последовательностей мембранных белков с известными данными об их структуре, показало применимость обоих предложенных методов: оптимизированного метода преобразования Фурье и *нового метода*

компьютерного анализа расположения трансмембранных участков белка – как по отдельности, так и совместно, для предсказания признаков вторичной структуры 24-х мембранных белков и обусловленных этими признаками функциональных свойств этих белков. Ценность обоих методов – простота, быстрота и экономичность применения.

4. Результаты, полученные в экспериментах по клонированию и экспрессии бактериородопсина и двух его делеционных мутантов в *E. coli*, могут иметь практическое значение при решении проблемы выбора определенной гетерологичной системы для экспрессии конкретного мембранного белка. Они частично отвечают на вопрос: почему для экспрессии специфического мембранного белка стоит (или не стоит) использовать определенный вектор в сочетании с клеткой-хозяином и условиями культивирования? Так, полученные нами экспериментальные данные показали, что экспрессионная система на основе белков MBP и TrxA, впервые использованная для экспрессии бактериородопсина в *E. coli*, не является оптимальной из-за малого выхода целевого белка и трудностей, возникающих в процессе его очистки.

Автор защищает следующие положения.

1. Обоснование выбора метода преобразования Фурье для анализа периодичностей аминокислотных последовательностей в различных белках, оптимального по наибольшей достоверности частотного спектра Фурье, простоте алгоритма и экономичности вычислений.
2. Применение выбранного оптимального метода преобразования Фурье к анализу периодичности расположения аминокислот в фибриллярных белках бактериофага T4 и белке gp181 бактериофага phiKZ *Pseudomonas aeruginosa*, что в результате позволило выявить в последовательностях фибриллярных белков не только малые периоды чередования аминокислот, но и относительно большие периоды их следования.
3. Предсказание структурных особенностей белков систем инфицирования бактериофагов T4 и phiKZ на основе теоретического анализа, использованное при создании плазмидных векторов для экспрессии 3-х делеционных мутантов белка gp181 с последующим выделением и очисткой полученных рекомбинантных белков.
4. Экспериментальное исследование ферментативной активности 2-х мутантов белка gp181, анализ их КД-спектров, выводы о вторичной структуре С-концевого домена белка gp181 бактериофага phiKZ.
5. Разработку и применение, наряду с вышеупомянутым оптимизированным методом преобразования Фурье, *нового метода* компьютерного анализа расположения трансмембранных участков мембранных белков для предсказания вторичной структуры этих белков, основанного на использовании функции и шкалы гидрофобности, более предпочтительного по сравнению методом Фурье и другими известными методами.

Апробация работы. Основные результаты работы докладывались на Национальных и Международных конференциях:
“Математические методы в технике и технологиях” - XXVI Международная научная конференция (Саратов, 2013), “Математические методы в технике и технологиях” - XXV Международная научная конференция (Харьков, 2012), “Математические

методы в технике и технологиях” - XVII Международная научная конференция (Кострома, 2004), 5“Перспективные направления физико-химической биологии и биотехнологии” - XVI Зимняя молодежная научная школа ИБХ РАН (Москва, 2004), “Современные проблемы фундаментальных и прикладных наук” - XLVI Научная конференция МФТИ (Москва-Долгопрудный, 2003),
и на научных семинарах:
в Лаборатории молекулярной биоинженерии ИБХ РАН (Москва, 2003-2005), в Институте Комплексных Систем (ICS-5: Молекулярная биофизика, 2009-2012) Научно-исследовательского центра г. Юлиха (г. Юлих, Германия), Институте Структурной Биологии (г. Гренобль, Франция, 2011).

Публикации. По материалам диссертации опубликовано 11 работ.

Личный вклад автора в получение результатов, выносимых на защиту, является определяющим. Вклад автора в совместных публикациях с соавторами является основным, за исключением работы [5], где автором была выполнена лишь часть исследования.

Структура и объем диссертации. Диссертация состоит из ***Введения***, 3 глав, включая ***Литературный обзор***, ***Материалы и методы исследований***, ***Результаты исследований и их обсуждение***, и ***Заключения***, содержит 158 страниц основного текста, 48 рисунков, 7 таблиц, список цитируемой литературы из 196 наименований. Общий объем - 181 с.

СОДЕРЖАНИЕ РАБОТЫ

Во **Введении** обоснована актуальность темы диссертации, сформулированы цели и задачи, перечислены признаки научной новизны и практической значимости.

В **первой главе** проведен анализ современных проблем исследования структуры и свойств вирусных и мембранных белков. На основе проделанного анализа сделаны следующие выводы.

Исследования в структурной биологии необходимы для понимания фундаментальных природных механизмов жизни и развития медицины. Они могут базироваться на экспериментальных и теоретических методах.

Из-за объективных трудностей получения фибриллярных белков бактериофага T4 в кристаллическом состоянии решение их структуры по-прежнему остается актуальной задачей для борьбы с вирусными инфекциями. Развитие методов компьютерного моделирования и биоинформатики – альтернативные инструменты для изучения структуры этих белков.

Новым методом борьбы с бактериальными инфекциями, устойчивыми к синтетическим антибиотикам, является использование литических ферментов бактериофагов в качестве энзибиотиков. Для создания эффективных лекарств необходимо знать структуру и функции литических ферментов, а также их ферментативную активность.

Мембранные белки во всех биологических процессах играют важнейшую роль. Из-за объективных сложностей в получении мембранных белков и их кристаллов структура лишь малой их части решена. Структура бактериородопсина (bR) экспериментально установлена. Но остаются противоречия в понимании механизма протонного транспорта, осуществляемого этим белком. Поэтому развитие методов исследования, проводимое на bR, как на модельном объекте, может помочь в разрешении этих противоречий и при изучении других мембранных белков.

В дополнение к экспериментальным методам исследования структуры мембранных белков активно развиваются и применяются методы компьютерного моделирования.

В соответствии с изложенным в конце первой главы были сформулированы вышеуказанные цели и задачи настоящего исследования.

Во **второй главе** описаны ***материалы и методы, используемые в экспериментальном исследовании белков*** бактериофагов T4, phiKZ и мембранных белков, в частности, бактериородопсина *Halobacterium salinarum*, а также ***методы компьютерного анализа и биоинформатики, применяемые для теоретического исследования данных белков.***

Функции белка определяются его 3D-структурой. Главный экспериментальный метод исследования структуры белков – рентгеноструктурный анализ, применямый к кристаллам белка. Их получение очень сложно, трудоемко и включает в себя стадии: создание плазмидной конструкции, содержащей ген, кодирующий данный белок; экспрессию белка в клетках реципиента; очистку и концентрирование белка; кристаллизацию белка.

В первой части 2-й главы приведены сведения об использованных материалах, реактивах, методах генной инженерии, других биохимических и биофизических методах.

Экспериментальные трудности в получении высококачественных кристаллов фибриллярных, мембранных и других белков до конца не преодолены. Поэтому актуальной альтернативой методу рентгеновской кристаллографии при исследовании структуры белков стали методы компьютерного моделирования, основанные на использовании данных о первичной структуре белка – последовательности его аминокислот.

Во второй части 2-й главы описаны вычислительные методы предсказания вторичной структуры белков по их аминокислотным последовательностям. У фибриллярных белков эти последовательности могут иметь периодическое строение.

Для его анализа иногда используют преобразование Фурье математического образа символьной последовательности. Оно может быть выполнено не единственным образом. Известно несколько способов, различающихся отдельными деталями: 1) математическим образом аминокислотной последовательности в виде непрерывной или дискретной функции, 2) видом ее Фурье-преобразования - интегралом или рядом, 3) действительными или комплексными коэффициентами преобразования, по-разному нормированными. Проведенное нами сравнение различных способов преобразования позволило выбрать вариант дискретного преобразования последовательности а.к. рядом Фурье как наиболее оптимальный по наибольшей достоверности частотного спектра Фурье, простоте алгоритма и экономичности вычислений.

При этом математический образ символьной последовательности белка, содержащей аминокислоты числом L, можно сформировать так. Двадцать аминокислот в последовательности белка делят на группы: гидрофобные, гидрофильные и прочие. Если, например, исследуется периодичность расположения аминокислот гидрофобной группы, то в числовой последовательности, заменяющей белковую последовательность, на местах k_m, занимаемых в цепи M элементами (m=1, 2,…M) этой группы аминокислот, ставят цифру 1, а на остальных местах – цифру 0. Таким образом, получают математический образ символьной последовательности белка в виде функции

$$f(x) = f(k) = F(k_m) = \sum_{m=1}^{M} \delta(x - k_m), \quad (1)$$

заданной на дискретном множестве x=k=1, 2,…L с использованием обобщенной δ-функции Дирака. Функция (1) принимает всего 2 значения: f=1 при $k=k_m$, и f=0 при остальных k.

Выбранный нами оптимальный вариант дискретного преобразования Фурье выражается формулами

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos(\omega_n x) + b_n \sin(\omega_n x)) \cong \frac{a_0}{2} + \sum_{n=1}^{n_{max}}(a_n \cos(\omega_n x) + b_n \sin(\omega_n x)), \quad (2)$$

$$I(\omega_n)=c(\omega_n)^2=c_n^2=a_n^2+b_n^2, \quad (3)$$

$$a_n = \frac{2}{L}\sum_{m=1}^{M}\cos(\omega_n \cdot k_m), \qquad b_n = \frac{2}{L}\sum_{m=1}^{M}\sin(\omega_n \cdot k_m), \quad (4)$$

где (2) – тригонометрический ряд Фурье, которым может быть представлена функция (1), (3) – спектр Фурье, а c_n^2 – интенсивность гармоники с частотой ω_n, (4) – коэффициенты ряда Фурье.

К исследованию вторичной структуры мембранных белков применялись два частных метода компьютерного анализа: известный метод с использованием преобразования Фурье и усовершенствованный нами метод “скользящего окна”. Проводилось сравнение их между собой и с другими известными методами.

Структура мембранных белков интересна тем, что обуславливает их функциональные свойства, важные для медицины и фармакологии. Из многообразия таких белков наиболее интересны интегральные политопические трансмембранные белки, имеющие несколько гидрофобных участков, насквозь пронизывающих мембрану и функционирующих совместно в качестве шлюзов для транспортировки конкретных веществ и проведения сигналов через мембрану. Особенность этих белков в том, что повторяемость (возможно, периодическая) гидрофобных аминокислот им органически присуща.

Повторяемость расположения аминокислот гидрофобной (или гидрофильной) группы в белковых последовательностях – один из признаков вторичной и третичной структуры белка. Это верно для фибриллярных, глобулярных и мембранных белков.

Если повторяемость периодическая, то эту периодичность можно выявить, применяя известный метод преобразования Фурье к цифровому образу символьной последовательности аминокислот белка.

Число N_{TMD} трансмембранных участков-доменов (иначе говоря, число альфа-спиралей в цепи) можно приближенно определить по предлагаемой формуле

$$N_{TMD} = L/T-1 = n-1, \quad (5)$$

где L – вся длина цепи, T – главный период последовательности, определенный из ее спектра Фурье для аминокислот гидрофобной группы по пику с максимальной интенсивностью I(T) среди пиков с большими значениями периодов T. Он равен средней длине участков повторения элементов цепи.

Число N_{TMD} правильно определено по формуле (5) только для 58% исследованных нами белков. Это обусловлено отклонением от периодичности у значительной части белков. Оценить местоположение и протяженность гидрофобных участков этим методом можно лишь грубо, с точностью до T/2.

В качестве альтернативы методу Фурье в данной работе предложен ***новый метод анализа расположения трансмембранных участков путем многократного***

усреднения функции гидрофобности внутри “окна”, перемещаемого вдоль полипептидной цепи.

Трансмембранные домены (TMDs) состоят в основном из гидрофобных аминокислот – средняя по такому участку величина гидрофобности, задаваемая в последовательности белка некоторой функцией $f(k) = H_N[i(k)]$, должна быть выше, чем в примыкающих к нему гидрофильных топологических доменах (TPDs). Это локальное свойство не зависит от периодичности расположения TMDs и TPDs. Здесь $i(k) = 1, 2,\ldots 20$ – это номер той из 20-ти аминокислот, которая стоит в последовательности белка на месте с номером k. Используемая в этом методе шкала гидрофобности $H_N(i)$ может быть задана многими способами в зависимости от характеризующей это свойство и физически измеряемой величины.

Процедура усреднения функции f(k) по шкале $H_N(i)$ проводилась не один, а несколько раз по алгоритму

$$f_n(k)=\frac{1}{2n+1}\sum_{k=-n}^{n} f_{n-1}(k), \quad n=1, 2,\ldots 5, \quad f_0(k)=f(k), \tag{6}$$

где каждое новое усреднение производили над предыдущей функцией по окну большей ширины d = 2n + 1: первое усреднение по 3 элементам, второе – по пяти и т.д. Наилучший результат, на наш взгляд, достигался при n = 4, усреднении по окну шириной d = 9 а.о. (иногда при n = 5, d = 11 а.о.).

Таким образом, если повторяемость трансмембранных участков непериодическая, то для ее выявления можно и лучше использовать другой метод – метод многократного (4–5 раз) усреднения функции гидрофобности белка в пределах перемещаемого вдоль последовательности “окна” шириной 9–11 а.о.

Третья глава посвящена результатам, полученным в данной диссертационной работе, а также их обсуждению.

Исследование периодичности расположения аминокислот в фибриллярных белках и литических ферментах бактериофагов T4 и phiKZ. Вирусные белки фибриллярной природы играют первостепенную роль в процессе инфицирования. Узнавание специфических рецепторов на поверхности клетки-хозяина и адсорбция вируса на мембране происходит именно при помощи фибриллярных адгезинов. Таким образом, исследование функциональных участков этих белков, определение их структуры и механизма взаимодействия с рецепторами клетки представляют интерес для разработки способов борьбы с вирусными инфекциями.

В данной части работы вышеописанным оптимальным методом Фурье-трансформации последовательностей аминокислот (формулы (1) - (4)) сначала были исследованы пять (gp34-gp38) из шести белков, участвующих в формировании длинных хвостовых фибрилл бактериофага T4, а также два других его белка gp5 и gp12 с уже известной целиком или частично структурой. Были получены Фурье-спектры $I(\omega_n)$ как для полных последовательностей, так и для их частей, построены соответствующие графики зависимости интенсивности пиков $I(T_n)$, которая пропорциональна вероятности повторения аналогичных элементов цепи с периодом T_n, для гидрофобных (gfb), гидрофильных (gfl) групп аминокислот и объединяющей их gfb+gfl-группы.

На рисунках 1 и 2 представлены спектры, полученные для белков gp35, gp38, gp5C. Очевидно, на рисунке 1 в спектре gfb-группы присутствует спектральная компонента с *малым* периодом $T=15.3\pm0.4\cong15$, а также пик на двукратной частоте с периодом $T=7.4\pm0.1$. Такой же пик имеется в спектре для объединенной gfb+gfl-

группы. Кроме того, в обоих спектрах есть пик с *большим* периодом T=46±4. Не исключено, что он и воспроизводится на указанных малых периодах T=15 и 7.5.

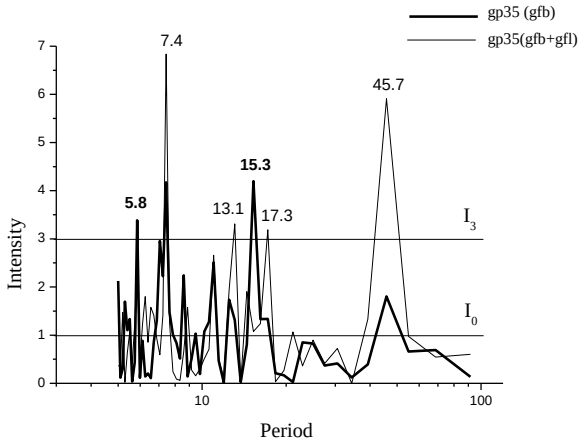


Рисунок 1 - Спектры Фурье для группы гидрофобных (gfb) аминокислот и объединенной gfb+gfl-группы в белковой последовательности gp35

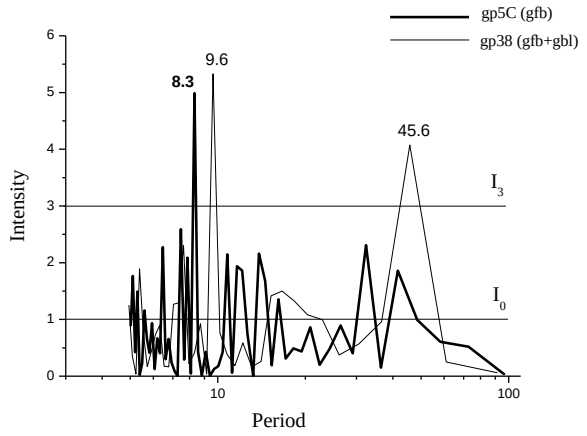


Рисунок 2 - Спектры Фурье для gfb+gfl-группы аминокислот в белковой последовательности gp38 и для аминокислот gfb-группы в С-концевой части белка gp5

Вся последовательность белка gp35 длиной L=275 делится *большим* периодом T=45.7≅275/6=L/6 на 6 равных частей.

В спектре белка gp38 на рисунке 2 среди других выделяется интенсивный пик с *малым* периодом T=9.6±0.3≅10 и еще один достаточно интенсивный пик с *большим* периодом T=45.6≅46±6, которым вся последовательность аминокислот данного белка делится на 4 равные части. С учетом погрешности пик с меньшим периодом из двух здесь указанных может рассматриваться как соответствующий пятикратной частоте другого или как самостоятельный.

Данные о периодичности, полученные методом преобразования Фурье, хорошо согласуются с экспериментальными результатами. Так, известно, что для С-концевых частей белков gp5 и gp12 характерны повторы с периодом T=8. Аналогичный период был установлен в их Фурье-спектрах, один из которых представлен на рисунке 2.

С использованием оптимального Фурье-преобразования удалось подтвердить наличие определенной периодичности в исследованных фибриллярных белках бактериофага T4: для ряда белков выявлено чередование аминокислот разных групп с относительно *малым* периодом T=15, а для некоторых – с периодами: 8, 10, 40.

Безусловно, новым результатом явилось обнаружение относительно *больших* периодов следования аминокислот, которыми вся их последовательность в белке делится на 4 или 6 равных частей.

Полученные данные о периодичности позволяют выровнять аминокислотные последовательности с учетом выявленных периодов обоих типов, что в сочетании с данными кристаллографии и электронной микроскопии может быть полезным для установления элементов вторичной и третичной структуры белков бактериофага T4.

Как указывалось выше, в полученном нами спектре Фурье для С-концевой части литического фермента gp5C бактериофага T4 были обнаружены повторы с периодом T=8, образующие уникальный паттерн. Рентгеноструктурный анализ этого белка выявил такую же периодичность. Кроме того, в структуре gp5 был обнаружен уникальный структурный мотив – тримерная β-спираль, о которой подробно говорилось в первой главе. Результаты, полученные при исследовании gp5 методом

Фурье, являются новыми и, безусловно, интересными для фундаментальной и прикладной науки.

Поэтому представлялось целесообразным исследовать также литический фермент gp181 бактериофага phiKZ, попробовать обнаружить в нем схожую с gp5 периодичность, а также изучить его структуру для понимания механизма его действия при инфицировании бактериофагом phiKZ клетки. В исследовании применялись теоретические – преобразование Фурье – и экспериментальные методы.

С помощью оптимального преобразования Фурье нам удалось подтвердить наличие определенной периодичности в белке gp181 бактериофага phiKZ, а именно, с тем же периодом T=8, который был обнаружен в белке gp5 бактериофага T4.

Экспериментальное исследование белков систем инфицирования бактериофагов T4 и phiKZ.

В ходе экспериментального исследования белков gp35 и gp35-36 бактериофага T4 было сделано следующее.

Создан плазмидный вектор для экспрессии белка gp35 в клетках *E. coli* и получена плазмидная конструкция для коэкспрессии белкового комплекса gp35-36 бактериофага T4 в клетках *E. coli*. Плазмиды были созданы с использованием вектора pET-23d(+) (*Novagen*, США) и содержали в себе полноразмерные гены g35 (1140 п.н.) и g35-36 (1800 п.н.) бактериофага T4. Ген 35 был амплифицирован с помощью ПЦР, в качестве матрицы использовалась ДНК фага T4.

Экспрессированные рекомбинатные белки gp35 и gp35-36 выделялись и очищались. Белок gp35 и комплекс белков gp35-36 экспрессировались при различных условиях в нерастворимой форме, в виде телец включения; при попытке перевода белка gp35 в растворимую форму в процессе рефолдинга он деградировал.

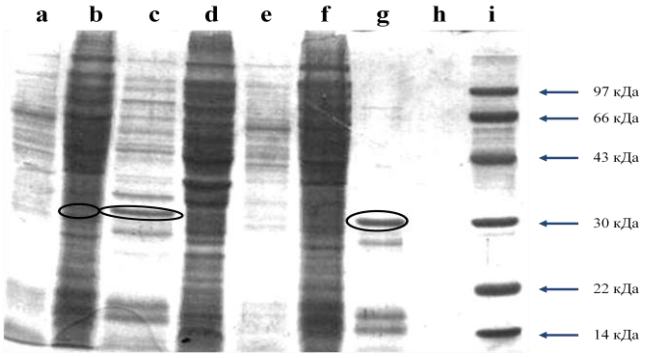
Из-за возникших объективных трудностей в экспериментальном исследовании продуктов генов 35 и 36, нам не удалось получить их в препаративных количествах для последующей кристаллизации и решения структуры X-ray-методом.

Экспериментальное исследование белка gp181 проводилось с помощью направленного делеционного мутагенеза. С учетом данных анализа аминокислотной последовательности с помощью программы BLAST, а также данных о периодичности в С-концевой части gp181, были получены три делеционных мутанта гена 181, а именно: ген g181MA, кодирующий лизоцимный домен, ген g181E, кодирующий С-концевую часть белка gp181, расположенную после лизоцимного домена и ген g181M, являющийся объединением генов g181MA и g181E.

После экспрессии и выделения был проведен анализ рекомбинантных белков на растворимость, который показал, что при температуре 37°C белок gp181M экспрессируется полностью в нерастворимой форме, белок gp181E частично находится в супернатанте, а белок gp181MA полностью растворим. При пониженной температуре 18°C белок gp181M также экспрессировался в нерастворимой форме, однако доля белка gp181E, экспрессировавшегося в растворимой форме, существенно повысилась. Три полученных рекомбинантных белка были выделены и очищены.

Также был проведен анализ взаимодействия белков с различными штаммами бактерии *Pseudomonas aeruginosa*. Белок gp181MA, являясь цитолитическим ферментом, а именно пептидогликановой гидролазой, лизирует клетки. Белок gp181E взаимодействует со всеми штаммами бактерии, но не лизирует их. Взаимодействие gp181E со штаммом PAO *Pseudomonas aeruginosa* отражено на рисунке 3. Белок gp181E находится во фракции осадка при взаимодействии со штаммом PAO *Pseudomonas aeruginosa*. В качестве контроля использовали бактерию *E. coli*. При

добавлении gp181E к *E. coli* взаимодействия не наблюдается – на рисунке 3 белок gp181E полностью во фракции супернатанта.



a, c – супернатант gp181E +PAO и *E.coli*, соответственно; e – супернатант PAO, g – супернатант gp181E; b, d – осадок gp181E +PAO и *E.coli*, соответственно; f – осадок PAO; h – осадок gp181E; i –маркер

Рисунок 3 - Картина электрофореза в 12% SDS-ПААГ, отображающая взаимодействие gp181E со штаммом PAO *Pseudomonas aeruginosa*

Перед началом экспериментальных исследований белка gp181E предполагалось, что он может играть роль сенсорной молекулярной иглы. Это предположение согласуется с наблюдавшимся взаимодействием gp181E с клетками *Pseudomonas aeruginosa*.

Таким образом, в эксперименте было установлено, что белок gp181M экспрессируется в нерастворимой форме, в виде телец включения, а белок gp181MA полностью растворим. Белок gp181E экспрессируется при пониженной температуре и становится растворимым в присутствии детергента и сахара, что косвенно свидетельствует об аффинности этого белка к полисахаридным (LPS) рецепторам клеточной стенки бактерии. Белок gp181E является гидрофобным и взаимодействует с различными штаммами клеток *Pseudomonas aeruginosa*, но не лизирует их, в отличие от gp181MA. Белок gp181E является не мономером, а димером или тримером, что было установлено с помощью гель-фильтрации.

При анализе КД-спектров белков gp181E и gp181MA (рисунки 4, 5) для оценки вкладов различных вторичных структур было установлено, что эти белки в большей степени состоят из α-спиралей: gp181E – на 71.5% и gp181MA – на 68%. Этот факт является, безусловно, интересным, т.к. большая часть структурных внутриклеточных белков, например, тропомиозин, миозин, кератин, фибриноген, трансмембранные белки оболочечных вирусов и некоторые глобулярные белки состоят из α-спиралей.

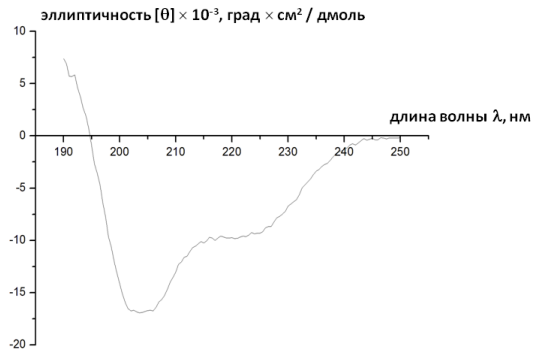


Рисунок 4 - КД-спектр белка gp181E

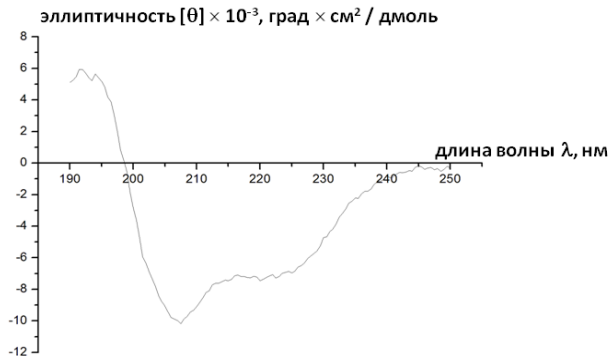


Рисунок 5 - КД-спектр белка gp181MA

Предположение о том, что С-концевой домен белка gp181 может играть роль сенсорной “молекулярной” иглы при инфицировании бактериофагом phiKZ клетки, выдвинутое с учетом установленной периодичности (T=8) этого домена и аналогии с gp5C фага T4, косвенно получило дополнительные экспериментальные обоснования. Однако, вторичная структура этого домена не является родственной β-спирали белка gp5 бактериофага T4, а представляет собой, вероятно, α-спиральный coiled coil. Указанное отличие вторичных структур этих белков может определять различие механизмов их взаимодействия с клетками. Для белка gp181 этот механизм до конца еще не исследован.

Результаты предсказания вторичной структуры мембранных белков вычислительными методами по их аминокислотной последовательности.

Исследования мембранных белков (к которым относится и bR) играют очень важную роль для развития биологической науки, фармакологии и медицины. Мембранные белки отвечают за многие функции клеток: обеспечивают избирательный обмен веществ между клеткой и окружающей ее средой, поддерживают разность электрических потенциалов внутри клетки и снаружи, обеспечивают передачу электрических сигналов в клетку и из нее, участвуют в производстве и переносе энергии в организме. Поэтому действие более половины фармацевтических препаратов направлено именно на мембранные белки.

В первой главе отмечалось, что бактериородопсин тщательно изучают уже более 40 лет. Но до сих пор не преодолены объективные экспериментальные трудности в достижении высокого кристаллографического разрешения кристаллов bR. Поэтому остается ряд противоречий, касающихся понимания механизма протонного транспорта и данных о структуре bR и его мутантов.

Альтернативой экспериментальным методам могут служить компьютерные методы. Вышеописанный оптимальный метод преобразования Фурье в данной части работы был применен к 24 трансмембранным белкам, для 14 из них – успешно.

Например, в Фурье-спектре, изображенном на рисунке 6, для последовательности bR (длиной L = 262 а.о.) проявился пик с периодом T = 33 и относительной интенсивностью I(T) = 3.5, для которого по формуле (5) находим $N_{TMD} = 7$, что совпадает с известным числом альфа-спиралей в структуре этого белка.

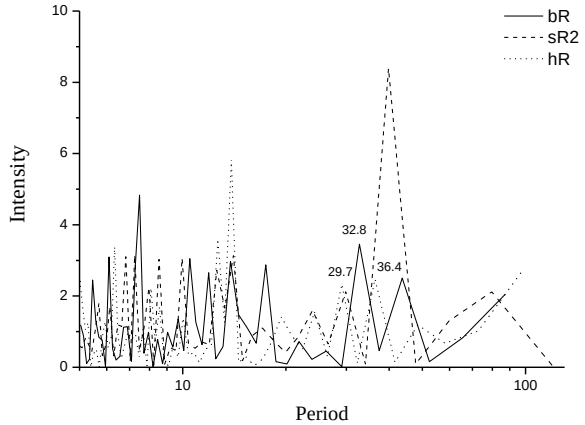


Рисунок 6 - Фурье-спектры для гидрофобной группы аминокислот в символьных последовательностях трансмембранных белков: bR, sR2, hR

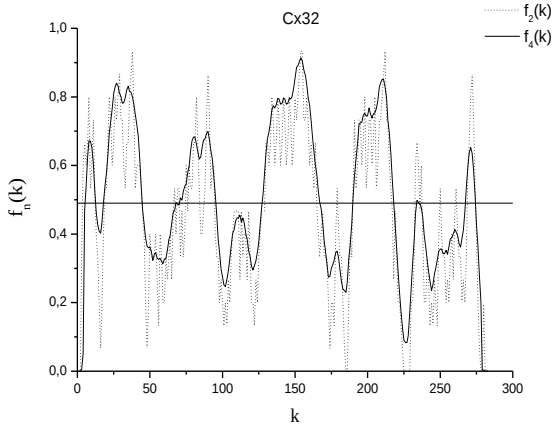


Рисунок 7 - Функции гидрофобности $f_n(k)$ для Cx32 после усреднения при n = 2 и n = 4 по шкале $H_1(i)$

Описанный во 2-ой главе *новый метод* компьютерного анализа расположения трансмембранных участков белка, основанный на использовании функции и шкалы гидрофобности, был применен к тем же 24 трансмембранным белкам (успешно для 19 из них) и показал большую, чем метод Фурье, пригодность для прогнозирования вторичной структуры таких белков и соответствующих ей функциональных свойств.

На рисунке 7 представлены результаты усреднения функции гидрофобности для последовательности белка Cx32 по шкале гидрофобности $H_1(i)$. Шкала гидрофобности $H_1(i)$ задана таким образом: гидрофобным аминокислотам (A; F; G; I; L; V; W; Y) присваивали значение 1, а всем остальным – значение 0.

Очевидно, что, если исключить из рассмотрения два узких пика на краях графика функции $f_4(k)$, то оставшиеся 4 широких пика, превышающие средний уровень $u = \langle f_n(k) \rangle = 0.49$, как раз будут соответствовать четырем TMDs в предполагаемой структуре этого белка. На графике функции $f_2(k)$ второй TMD еще не разрешен, на его месте и в других местах присутствуют несколько узких пиков.

На рисунках 8 и 9 представлены функции гидрофобности для bR после усреднения при n = 2 и n = 4 по шкалам $H_2(i)$ и $H_3(i)$ и при n = 5 по шкале $H_4(i)$. Очевидно, при n, равном 4 или 5, все 7 TMD разрешены.

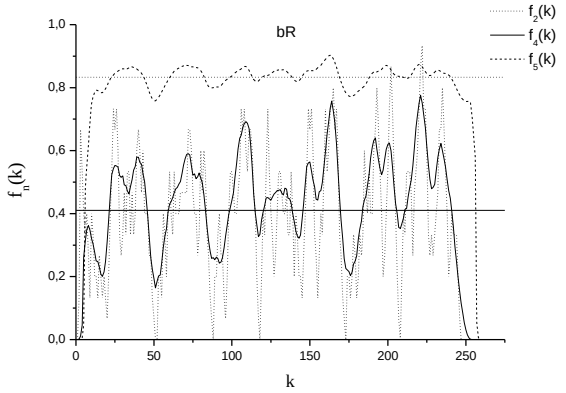


Рисунок 8 - Функции гидрофобности $f_n(k)$ для bR после усреднения при n = 2 и n = 4 по шкале $H_2(i)$ и усреднения при n = 5 по шкале $H_4(i)$ из работы [8]

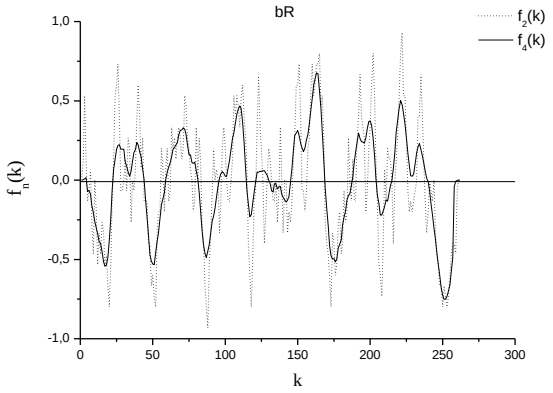


Рисунок 9 - Функции гидрофобности $f_n(k)$ для bR после усреднения при n = 2 и n = 4 по шкале $H_3(i)$ из работы [8]

Исследование применимости предложенных компьютерных методов для предсказания вторичной структуры мембранных белков по их аминокислотной последовательности позволило сделать следующие выводы.

Из двух рассмотренных в данной работе методов более предпочтительным оказался метод многократного (4–5 раз) усреднения функции гидрофобности f(k) белка внутри “окна” шириной 9–11 а.о., перемещаемого вдоль аминокислотной последовательности. Число трансмембранных участков правильно предсказано этим методом для 79% исследованных белков. Доля правильно предсказанных границ участков составляет 81%. Часть результатов представлена ниже в таблице.

Шкала и функция гидрофобности могут быть заданы многими способами – известно более 30. Сравнение различных шкал и функций гидрофобности, выполненное в данной работе, показало, что определяемое с их использованием число и расположение трансмембранных участков часто практически одинаковы, даже для очень простых (грубых) шкал. Но иногда для данного белка одна из шкал может оказаться предпочтительнее из-за того, что она позволяет лучше определить близкорасположенные трансмембранные участки.

Сравнение результатов, полученных в данной работе путем компьютерного анализа аминокислотных последовательностей 24 белков, с известными данными об их структуре показало применимость обоих предложенных методов, как по отдельности, так и совместно, для предсказания признаков неизвестной вторичной структуры мембранных белков и обусловленных этими признаками функциональных свойств этих белков. Ценность обоих методов – простота, быстрота и экономичность применения.

Таблица - Сравнение с известными данными сайта (http://www.uniprot.org) границ TMDs, вычисленных при обработке функций гидрофобности $f_n(k)$ при n = 4 и n = 5** для шести различных мембранных белков по шкалам $H_3(i)$ и $H_5(i)$ из работы [8]

Код белка	Источник данных	Номер и границы трансмембранных и интрамембранного* участков						
		1	2	3	4	5	6	7
bR	сайт	24−42	57−75	92−109	121−140	148−167	186−204	217−236
	$H_3(i)$	23−44	59−81	96−114	122−130	146−168	188−204	216−239
sR2	сайт	4−25	34−55	70−91	95−117	122−149	154−181	190−222
	$H_3(i)$	5−23	40−62	74−91	99−116	126−142	165−182	192−215
hR	сайт	31−56	63−86	121−142	146−169	173−195	208−231	241−269
	$H_3(i)$	39−55	69−77	132−140	154–164	175−197	215−240	251−265
Cx32	сайт	23−45	76−95	131−153	192−214			
	$H_3(i)$	22−41	68−96	130−169	190−216			
GRM6	сайт	586−608	623−643	655−673	698−718	749−770	784−806	820−845
	$H_5(i)$**	591−606	625−649	655−666	692−719	748−769	784−808	818−845
CNG	сайт	13−30	39−61	75−94	95−112	130−150	162−180*	186−210
	$H_3(i)$	4−28	41−59	77−90	99−107	133−149	165−172	191−206

Экспериментальное исследование бактериородопсина (bR) и его делеционных мутантов, полученных путем экспрессии в бактерии *E. coli*.

Сложная и трудоемкая технология получения мембранных белков в кристаллическом состоянии включает в себя следующие основные стадии: создание плазмидной конструкции, содержащей ген, кодирующий данный белок; экспрессию белка в клетках реципиента; разрушение клеток и отделение клеточных мембран, солюбилизацию белка в присутствии соответствующего детергента; очистку и концентрирование белка; реконституцию белка и его кристаллизацию. Наиболее критичными для достижения конечной цели являются стадия очистки из-за значительных потерь белка (до 40% на каждом из нескольких этапов) и стадия его кристаллизации из-за неопределенности необходимых условий ее проведения для данного конкретного белка. В силу указанных трудностей структура мембранных белков установлена к настоящему времени лишь для очень малой их части (< 4%).

Для получения высокого разрешения структуры белка необходимо, чтобы кристаллы этого белка имели очень высокое качество. Качество кристаллов зависит от многих факторов, одним из определяющих требований является наличие достаточного количества гомогенного и чистого продукта для кристаллизации.

Для решения проблемы негомогенности белка bR у него методами генной инженерии был удален лидерный пептид, а также Asp 249 с С-конца. То же было сделано и для двух его делеционных мутантов. Создание делеционных мутантов бактериородопсина (bRΔ10, bRΔ13) было обусловлено тем, что часть а.о. у bR дикого

типа (bRWT) не видна на карте электронной плотности при расшифровке структуры, а именно, 4 а.о. с N-конца белка и 15 а.о. с C-конца белка. Это может быть обусловлено тем, что они находятся в свободном состоянии, т.е. не встроены в мембрану, что отрицательно сказывается на качестве кристалла.

Известно, что лидерный пептид необходим для встраивания бактериородопсина в цитоплазматическую мембрану и предотвращения его деградации. Как было отмечено в первой главе, использование экзогенных N-концевых последовательностей помогло при встраивании бактериоопсина (bO) во внутреннюю мембрану бактерии и тем самым позволило повысить выход и стабильность белка. Привлекательной системой для экспрессии bR в *E.coli* представлялся гибрид bO с MBP (Maltose Binding Protein). Белок MBP кодируется геном *mal*E и имеет на своем N-конце собственный лидерный пептид для транслокации этого белка в периплазму бактерии. Поэтому можно было предположить, что конструкция MBP-bO способна обеспечить эффективное встраивание bO в цитоплазматическую мембрану. Белок MBP может быть легко отделен от bO с помощью созданного методами генной инженерии сайта узнавания TEV-протеазой, находящегося в линкерной области между двумя белками.

В этой части экспериментального исследования bR было сделано и установлено следующее. Для экспрессии бактериородопсина в *E. coli* удалось получить одну конструкцию pl_bOWT, содержащую ген bOWT, что было подтверждено с помощью секвенирования. Для двух других конструкций (pl_bOΔ10 и pl_bOΔ13) после электропорации было проанализировано порядка 300 клонов, но ни в одном из них не было обнаружено нужных вставок bOΔ10 и bOΔ13.

Экспрессию гена bO, клонированного в плазмидный вектор pl_bOWT, осуществляли в 5-ти различных штаммах *E. coli* и в 3-х различных питательных средах 2xTY, TB, LB; штамм компетентных клеток *E. coli* ER2507 и среда 2xTY оказались наиболее благоприятными для экспрессии бактериородопсина.

Проведенный скрининг 7 различных детергентов (DM, DDM, CHAPS, SDS, FOS-12, NLS, LDAO) позволил выявить наиболее подходящий из них для солюбилизации бактериородопсина, экспрессированного в системе pl_bOWT. При этом выяснилось, что детергент DM плохо подходит для солюбилизации белкового комплекса MBP-Tev-bOWT-Tev-TrxA-10×His и последующих манипуляций с ним, а детергенты DDM, NLS, FOS-12 лучше солюбилизируют белок, тем самым делая гистидиновый хвост более доступным для связывания с носителем.

Проведенный сравнительный анализ протеолитического действия TEV-протеазы в 3-х различных детергентах (NLS, DDM, FOS-12), которые лучше других подошли для солюбилизации и очистки бактериоопсина, показал, что TEV-протеаза, несмотря на свою чувствительность к действиям детергентов, функционирует во всех трех детергентах.

Как оказалось, выбранная экспрессионная система не является оптимальной для экспрессии bR в *E.coli* из-за объективных трудностей, возникающих в процессе очистки белка и последующего его протеолиза, которые приводят к практически полной потере белка.

По результатам экспериментального исследования bR можно сделать ряд общих и важных выводов.

1. Чтобы получить большое количество гомогенных функциональных мембранных белков для структурных исследований, необходимо ясно понимать процессы, которые происходят в клетках-хозяевах при рекомбинантной экспрессии

этих белков. Различные клетки-хозяева имеют индивидуальные биологические свойства, поэтому экспрессируемые в них мембранные белки могут отличаться последовательностью аминокислотных остатков, стабильностью и уровнем выхода.

2. Для разработки будущих экспериментов по экспрессии необходим переход от описательных исследований (с концентрацией внимания на сообщениях об уровнях экспрессии) к пониманию основных механизмов. Это поможет в решении проблемы, как соединить специфический мембранный белок с определенным вектором для экспрессии, клеткой-хозяином и условиями культивирования.

3. Дальнейший прогресс в совершенствовании использования шкалы биологической гидрофобности в математических моделях, например, в предложенном нами *новом методе* предсказания TMDs, может позволить рассчитывать эффективность встраивания природных полипептидов в мембрану. И, таким образом, мембранные белки для экспериментальных исследований с целью достижения наилучшей функциональной экспрессии могут быть выбраны на основе их аминокислотных последовательностей.

ЗАКЛЮЧЕНИЕ

1. Проведенное сравнение различных способов преобразования Фурье для анализа периодичности аминокислотных последовательностей в различных белках позволило выбрать дискретное преобразование рядом Фурье в качестве наиболее оптимального варианта по наибольшей достоверности частотного спектра Фурье, простоте алгоритма и экономичности вычислений.
2. При исследовании выбранным методом периодичности расположения аминокислот в фибриллярных белках бактериофага T4 были получены данные, позволившие выровнять аминокислотные последовательности с учетом установленных *малых и больших* периодов.

   Оптимальный метод преобразования Фурье был также применен к 24 трансмембранным белкам, для 14 из них – успешно.
3. Проведенный анализ аминокислотной последовательности литического фермента бактериофага phiKZ - белка gp181 - показал наличие периодичности в его C-концевой части с тем же периодом T=8, что и в C-концевой части литического фермента бактериофага T4 - белка gp5. С учетом выявленной аналогии этих белков было выдвинуто предположение о том, что C-концевой домен белка gp181 играет роль сенсорной “молекулярной” иглы в процессе инфицирования бактериофагом phiKZ клетки, которое затем получило экспериментальное обоснование.
4. В качестве альтернативы методу Фурье в данной работе предложен *новый метод* компьютерного анализа расположения трансмембранных участков белков, основанный на применении функции и шкалы гидрофобности. Его особенность заключается в многократном усреднении функции гидрофобности внутри “окна”, перемещаемого вдоль полипептидной цепи.

   Этот *новый метод* был применен к тем же 24 трансмембранным белкам, что и метод Фурье, для 19 из них − успешно. Он показал большую, чем метод Фурье, пригодность для прогнозирования вторичной структуры такого рода белков и соответствующих ей функциональных свойств.
5. Сравнение результатов, полученных путем компьютерного анализа аминокислотных последовательностей 24 трансмембранных белков, с известными данными об их структуре показало применимость обоих предложенных методов, как по отдельности, так и совместно, для предсказания признаков неизвестной

вторичной структуры мембранных белков и обусловленных этими признаками функциональных свойств этих белков.

6. Экспериментальные исследования фаговых и мембранных белков, полученных путем рекомбинантной экспрессии с использованием созданных в работе плазмид, позволило установить для них ряд структурных особенностей и свойств и сопоставить их с теоретическими предсказаниями.

   В частности, белок gp181MA фага phiKZ экспрессировался полностью в растворимой форме. Белок gp181E, экспрессированный при пониженной температуре, становился растворимым в присутствии детергента и сахара. Он оказался гидрофобным и взаимодействовал с различными штаммами клеток *Pseudomonas aeruginosa*, но не лизировал их, в отличие от gp181MA. Путем гель-фильтрации установлено, что белок gp181E – не мономер, а димер или тример. Спектры КД для белков gp181E и gp181MA показывают, что их вторичная структура примерно на 70% состоит из α-спиралей длиной около 8 а. о.

   Экспериментальные данные свидетельствуют о различиях во вторичной структуре С-концевого домена белка gp181 бактериофага phiKZ и β-спирали белка gp5 бактериофага T4, что может определять различие механизмов их взаимодействия с клетками.

Благодарности. Автор благодарен д.б.н., профессору Месянжинову В.В. за предоставленную возможность выполнить часть данной работы в Лаборатории молекулярной биоинженерии ИБХ РАН.

Автор благодарит д.х.н., зав. Лабораторией молекулярной биоинженерии ИБХ РАН Мирошникова К.А за постановку научных задач и продуктивное обсуждение результатов работы, а также к.ф-м.н., профессора Университета Аахена (г. Аахен, Германия), зав. Лабораторией клеточной и молекулярной биофизики Института Комплексных систем Научно-исследовательского Центра (г. Юлих, Германия) Горделия В.И. за предоставленную возможность выполнить часть данной работы в указанной лаборатории и в Институте Структурной Биологии (г. Гренобль, Франция) и за постановку научных задач.

Автор благодарен к.б.н. Бегуновой Е.А. и к.б.н. Эдельвейс Э.Ф. за помощь в проведении биологических экспериментов, за ценные советы и консультации, участие в обсуждении исследований.

Автор искренне благодарит своего отца д.ф-м.н., профессора Симакова Н.Н. за продуктивное обсуждение результатов работы и формирование новых идей, за помощь в разработке алгоритмов численного моделирования, за ценные замечания и полезные советы при выполнении работы, а также при ее написании и оформлении, за доброту, внимание и неизменную моральную поддержку.

ПО МАТЕРИАЛАМ ДИССЕРТАЦИИ ОПУБЛИКОВАНЫ СЛЕДУЮЩИЕ РАБОТЫ

1. Симакова М.Н., Месянжинов В.В. Исследование периодичности последовательности аминокислот в фибриллярных белках бактериофага T4 с помощью преобразования Фурье // XLVI Научная конференция МФТИ Современные проблемы фундаментальных и прикладных наук, Москва-Долгопрудный, 2003, Труды конференции. Часть IV. С. 40.
2. Симакова М.Н., Мирошников К.А., Месянжинов В.В. Исследование периодичности аминокислотной последовательности в фибриллярных белках длинных хвостовых фибрилл бактериофага T4 с помощью преобразования Фурье // XVI Зимняя молодежная научная школа Перспективные направления физико-химической биологии и

биотехнологии Института биоорганической химии им. академиков М.М.Шемякина и Ю.А.Овчинникова РАН, Москва, 2004. Тезисы докладов. С. 22-23.

3. Симакова М.Н., Месянжинов В.В. Исследование периодичности последовательности аминокислот в фибриллярных белках бактериофага Т4 // XVII Международная научная конференция Математические методы в технике и технологиях, Кострома, 2004. Сборник трудов. Т.4. С. 150-155.
4. Симакова М.Н., Симаков Н.Н. Исследование периодичности расположения аминокислот в фибриллярных белках бактериофага Т4 // Молекулярная биология. 2005. Т. 39. № 2. С. 321-329.
   Simakova M.N., Simakov N.N. Investigation of periodic distributions of amino acids in the sequences of fiber proteins of bacteriophage T4 // Molecular biology (Mosk.). 2005. Vol. 39. No. 2, pp. 284-291.
5. Чупров-Неточин Р.Н., Файзулина Н.М., Сыкилинда Н.Н., Симакова М.Н., Месянжинов В.В., Мирошников К.А. Бета-спиральный домен бактериофага Т4 управляет укладкой фрагмента длинных фибрилл в составе химерного белка // Биоорганическая химия. 2010. Т. 36. № 2. С. 193-199.
   Chuprov-Netochin R.N., Faizullina N.M., Sykilinda N.N., Simakova M.N., Mesyanzhinov V.V., Miroshnikov K.A. The β-helical domain of bacteriophage T4 controls the folding of the fragment of long tail fibers in a chimeric protein // Russian Journal of Bioorganic Chemistry, (Mosk.). 2010. Vol. 36. No. 2, pp. 172-178.
6. Симакова М.Н., Симаков Н.Н. Компьютерные методы предсказания структуры мембранных белков // XXV Международная научная конференция Математические методы в технике и технологиях, Саратов, 2012. Сборник трудов. Т.9. С. 126-129.
7. Симакова М.Н. Предсказание вторичной структуры мембранных белков компьютерными методами // Вестник СГТУ. 2012. Т.64. № 1. С. 54-61.
8. Симакова М.Н., Симаков Н.Н. Вычислительные методы предсказания вторичной структуры мембранных белков по их аминокислотной последовательности // Молекулярная биология. 2013. Т. 47. № 2. С. 347–355.
   Simakova M.N., Simakov N.N. Computational methods for predicting structure of membrane proteins using amino acids sequences // Molecular biology (Mosk.). 2013. Vol. 47. No. 2, pp. 308-316.
9. Симакова М.Н., Симаков Н.Н. Предсказание структуры мембранных белков новым компьютерным методом. // XXVI Международная научная конференция Математические методы в технике и технологиях, Саратов, 2013. Сборник трудов. Т.8. С. 184-187.
10. Simakova M.N., Simakov N.N. Topography Prediction of Helical Transmembrane Proteins by a New Modification of the Sliding Window Method // Hindawi Publishing Corporation, BioMed Research International. Vol. 2014, Article ID 921218.
    URL: http://dx.doi.org/10.1155/2014/921218 (дата обращения: 08.06.14).
11. Мирошников К.А., Симакова М.Н. Структура и свойства литического фермента gp181 бактериофага phiKZ // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2014. Т. 194. № 2. С. 75-85.