

Московский Государственный Университет  
имени М.В.Ломоносова  
Биологический факультет  
кафедра биоинженерии

## **Дипломная работа**

на тему

**«Расчет спектральных свойств бактериородопсина гибридными методами  
квантовой химии и молекулярной механики»**

Выполнил студент 5-ого курса

Ф.С. Орехов

Научные руководители:

к.б.н. О.С. Соколова

к.ф.-м.н. А.К. Шайтан

Москва

2010

## Принятые сокращения

В работе, главным образом для облегчения восприятия текста, используются следующие общепринятые в специальной литературе сокращения:

bR (bacteriorhodopsin) – бактериородопсин;  
CPMD (Car-Parinello molecular dynamics) – молекулярная динамика Кар-Паринелло;  
DFT, TD-DFT (density functional theory, time-dependent density functional theory) – теория функционала электронной плотности, временное расширение теории функционала электронной плотности;  
HF, ХФ (Hartree-Fock) – метод Хартри-Фока;  
HBN (hydrogen bond network) – сеть водородных связей;  
(H)RSB ((hydrogenated) retinal Schiff base) – (протонированное) Шиффово основание ретиналя;  
НООР – внеплоскостные колебания атомов водорода;  
MD, МД – молекулярная динамика;  
QC, КХ – квантовая химия;  
MM – молекулярная механика;  
ПМ – пурпурная мембрана;  
КХ/ММ, КМ/ММ, QM/ММ – гибридная методика (метод) квантовой химии и молекулярной механики.

## ОГЛАВЛЕНИЕ

1.1. ВВЕДЕНИЕ.....	5
1.2. Задачи .....	8
2. ОБЗОР ЛИТЕРАТУРЫ .....	10
2.1. Структура бактериородопсина и его комплексов с липидами ПМ.....	10
2.1.1. Сеть водородных связей .....	13
2.1.2. Межспиральные контакты и спиральная подвижность.....	15
2.1.3. Обзор существующих структур бактериородопсина.....	16
2.2. Фотоцикл. Механизм переноса протона .....	18
2.2.1. Общие замечания.....	18
2.2.2. Фотоизомеризация ретиналя. Интермедиаты К и J .....	20
2.2.3. Депротонирование Шиффова основания. Интермедиаты L и M.....	24
2.2.4. Выброс протона во внеклеточное пространство. Интермедиат M.....	27
2.2.5. «Переключение» Шиффова основания (Protonation Switch).....	29
2.2.6. Репротонирование Шиффова основания. Интермедиаты M и N.....	30
2.2.7. Захват протона на цитоплазматической поверхности и реизомеризация ретиналя. Интермедиат N, N' и O .....	33
2.2.8. Восстановление исходного состояния. Интермедиат O .....	34
2.2.9. Предотвращение обратной изомеризации .....	35
2.3. Использование в биоэлектронике.....	36
2.4. Компьютерная фотохимия бактериородопсина (обзор работ) .....	38
2.4.1. Компьютерный расчет спектров ретиналя и его производных.....	38
2.4.2. Изучение влияния белкового окружения на спектральные характеристики ретиналя .....	39
3. МЕТОДЫ ИССЛЕДОВАНИЯ .....	41
3.1. Введение.....	41
3.2. <i>Ab initio</i> методы квантовой химии.....	46
3.2.1. Метод Хартри-Фока .....	46

3.3. Метод функционала плотности (DFT) .....	48
3.3.1. Основы метода .....	48
3.3.2. Временное расширение метода DFT .....	52
3.3.3. Эволюция метода DFT: SCC-DFTB, DFTBA.....	53
3.5. Метод молекулярной динамики.....	53
3.5.1. Общие замечания.....	53
3.5.2. Силовые поля и их структура.....	54
3.5.3. Температура. Термостаты.....	56
3.5.4. Численное интегрирование уравнений движения.....	57
3.5.5. Методы учета электростатических взаимодействий.....	58
3.6. Ab initio молекулярная динамика .....	59
3.6.1. Молекулярная динамика Борна-Оппенгеймера.....	59
3.6.2. Молекулярная динамика Кара-Парринелло.....	60
3.7. Гибридные методы КМ/ММ .....	61
4. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ .....	65
4.1. Подготовка моделей бактериородопсина и его мутантов.....	65
4.2. Расчет максимумов спектров поглощения для белка дикого типа и мутантов .....	70
4.3. Методика учета взаимной поляризации ретиналя и белкового окружения .....	77
4.4. Учет микросостояний бактериородопсина при расчете максимума поглощения.....	78
5. ЗАКЛЮЧЕНИЕ .....	80
ПРИЛОЖЕНИЕ 1 .....	81
ПРИЛОЖЕНИЕ 2 .....	83
СПИСОК ЛИТЕРАТУРЫ.....	84

## 1.1. ВВЕДЕНИЕ

С самого момента своего открытия [7] в начале семидесятых годов прошлого века бактериородопсин продолжает оставаться объектом пристального изучения. Лаборатории по всему миру проводят многочисленные исследования, ставящие своей целью создания детальной и полной картины строения и функционирования этого белка.

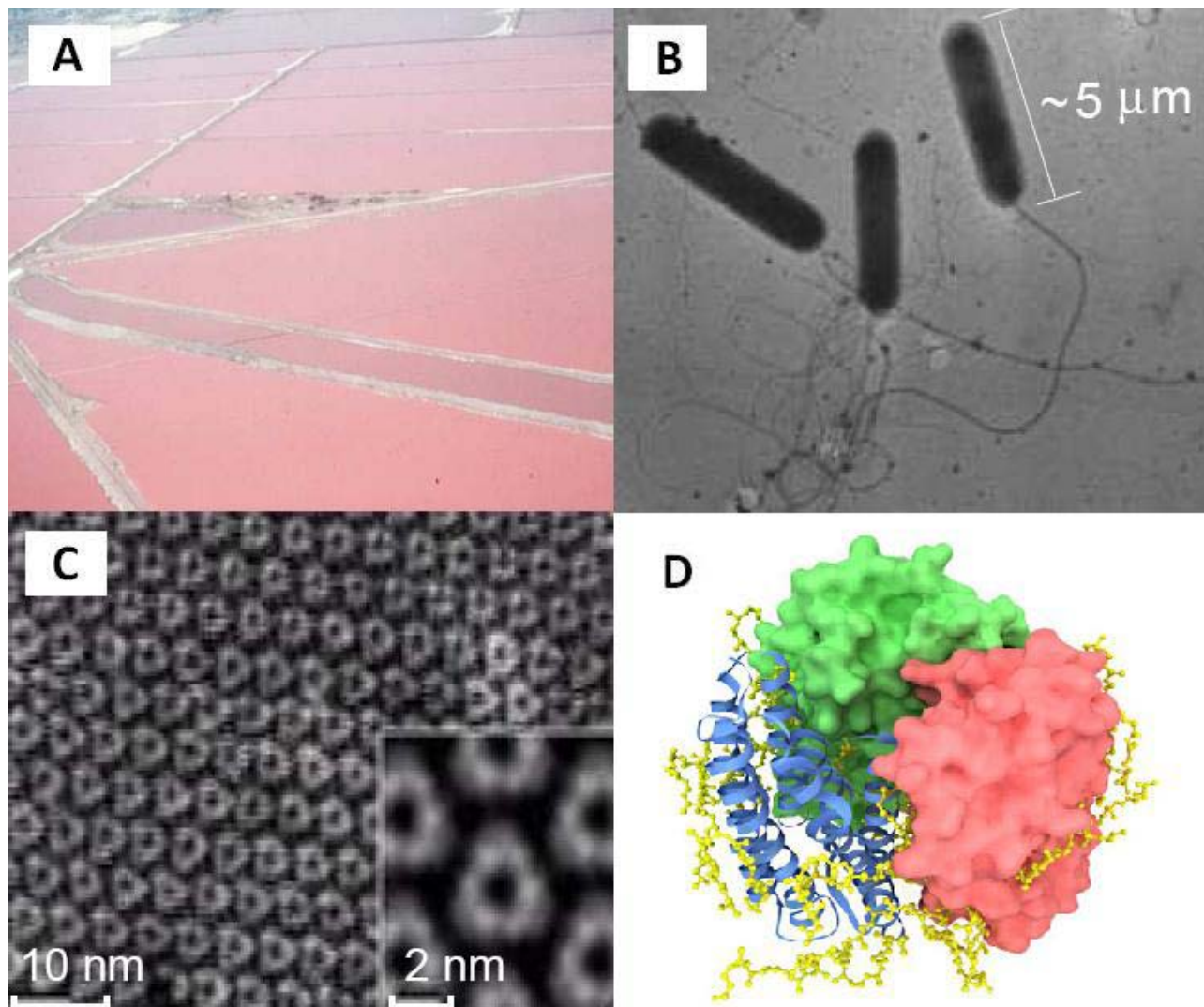


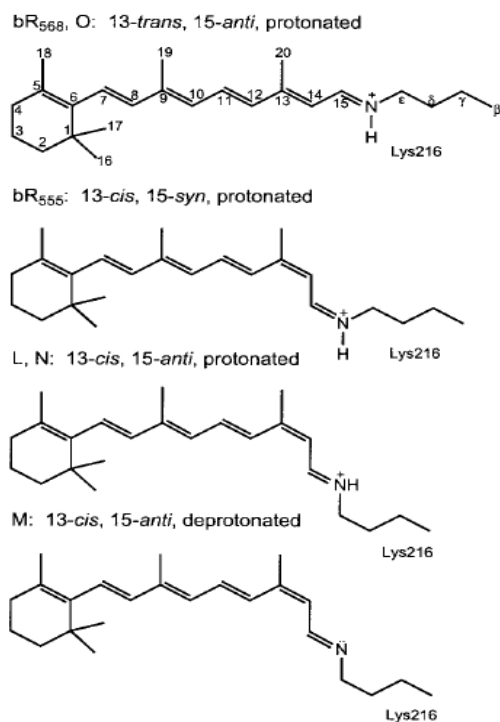
Рисунок 1. А – искусственные водоемы, предназначенные для промышленного получения соли, заселенные *Halobacterium salinarum*; В – изображение отдельных бактерий *Halobacterium salinarum*, полученное на электронном микроскопе; С – изображение цитоплазматической стороны пурпурной мембраны *Halobacterium salinarum* с отдельными тримерами бактериородопсина, полученное на атомным силовым микроскопом [8]; D – трехмерная визуализация тримера бактериородопсина, построенная на основе данных кристаллографического анализа [9].

Бактериородопсин представляет собой небольшой ( $\approx 24$  кДа) интегральный мембранный белок, встречающийся у галобактерий вида *Halobacterium salinarum* (раньше *Halobacterium halobium*). Археи этого вида – единственные организмы, способные существовать в экстремальных условиях солевых озер, концентрация соли в которых достигает очень высоких значений (5 М, или 25%), создавая тем самым т.н. «автостерильные» физиологические условия. Для эффективного существования в подобных условиях

галобактерии обладают уникальным в своем роде фотосинтетическим аппаратом, главную роль в котором играет бактериородопсину [10], [11]. Этот белок на свету создает трансмембранный градиент протонов, выкачивая протоны из цитоплазмы во внеклеточное пространство, переводя световую энергию в химическую [12]. Эффективность такого механизма заметно ниже хлорофилльного фотосинтеза (35%), но, тем не менее, достигает 15% [13]. Образованный в результате работы бактериородопсина протонный градиент в дальнейшем используется для различных нужд бактериальной клетки, в наибольшей степени, будучи переведенным в форму АТФ посредством АТФ-аз.

Следует отметить ряд других bR-подобных белков, имеющих у галобактерий. Это два рецепторных белка, известных как сенсорные родопсины I и II (SR I и SR II). Они позволяют бактерии детектировать наличие в спектре света диапазона длин волн, пригодного для осуществления фотоцикла. Это связано с тем фактом, что бактериородопсин работает как протонная помпа лишь при его освещении зеленым или желтым светом. Избыточное синее освещение ингибирует фотоцикл. SR I и SR II, реагируя на различные длины волн, запускают сигнальный каскад, приводящий в конечном счете к активации бактериальной локомоторной системы, осуществляющей перемещение клетки в область пространства с подходящим для фотосинтеза спектральным составом света. Еще один белок этой группы, которым обладают галобактерии – галородопсин (hR). Это светозависимая помпа, поддерживающая в клетке физиологическую концентрацию ионов хлора [14].

У галобактерий бактериородопсин *in vivo* организован в двухмерные гексагональные комплексы, в которые, помимо собственно трех молекул бактериородопсина, повернутых на 120° друг относительно друга, входят специфические липиды; такие структуры, представляющие собой двумерные кристаллы, получили название пурпурных мембран. До 80% поверхности бактериальных клеток могут быть покрыты ПМ. ПМ отличаются очень высокой стабильностью к различным физико-химическим воздействиям. В частности, эксперименты (из обзора [14]) показывают, что ПМ сохраняет свои свойства после длительного (несколько лет) пребывания на солнечном свете, в присутствии кислорода воздуха, а также при высушивании. Также ПМ выдерживает в воде температуры до 80°C, а в сухом состоянии – до 140°C [15], однако подобная стабильность утрачивается при разрушении кристаллической структуры ПМ; устойчивость к значениям pH наблюдается в диапазоне от 0 до 12 едениц; ПМ устойчива в растворах с высокой ионной силой, стабильна в растворах NaCl с концентрацией до 3 М. ПМ частично чувствительна к воздействию полярных органических растворителей, однако устойчива в неполярных растворителях, таких как гексан [16], [17], [18].



**Рисунок 2. Структурные формулы остатка лизина, с ковалентно пришитым к нему через Шиффово основание ретиналем, в различных конформациях, с указанием интермедиата фотоцикла, которому соответствует данная конформация.**

Трансмембранный домен бактериородопсина состоит из семи гидрофобных  $\alpha$ -спиралей, расположенных перпендикулярно плоскости мембраны. К остатку Lys 216, расположенному в средней части спирали G, ковалентно, через Шиффово основание, присоединен ретиналь. В основном, темновом, состоянии ретиналь Шиффова основания находится в состоянии термодинамического равновесия между *all-trans* формой и 13-*cis*, 15-*syn* [19]. При поглощении кванта света определенной энергии ретиналь претерпевает процесс фотоизомеризации и переходит, в конечном счете, в состоянии 13-*cis*, 15-*anti*. Фотоизомеризация запускает процесс переноса протона с цитоплазматической стороны мембраны во внешнюю среду, детальный механизм которого будет рассмотрен ниже.

В определенной степени процесс фотоизомеризации ретиналя в бактериородопсине можно рассматривать [14] как энзиматический процесс, в котором осуществляется светозависимая каталитическая реакция преобразования ретиналя из одной формы в другую под действием апопротеина.

Разносторонние исследования бактериородопсина включают исследования его структуры (с привлечением методов дифракции электронов, рентгеноструктурного анализа, метода ядерного магнитного резонанса) и динамических свойств (методами импульсной спектроскопии, инфракрасной спектроскопии, молекулярной динамики и квантовой химии). Эти работы интересны как с чисто теоретической стороны, так и в плане использования их результатов в прикладных нанобиотехнологических задачах, как то использование модифицированных форм бактериородопсина в качестве протонных помп, или создание на основе модифицированного бактериородопсина оптических носителей информации.

## 1.2. Задачи

Замечательная особенность ретинальсодержащих белков, остающаяся на протяжении десятилетий в фокусе самых современных исследований – спектральная подстройка (англ. spectral tuning). Хорошо известно, что свободный ретиналь в вакууме обладает максимумом поглощения при 610 нм [20], в неполярных растворителях типа метанола – 440 нм, а в бактериородопсине – 568 нм. То есть, белковое окружение существенно влияет на максимум поглощения (вызывает т.н. опсиновый сдвиг), причем для разных белков, отличающихся по аминокислотному составу при сохранении общего плана строения ретинальсодержащих белков, максимумы поглощения отличаются на значительные величины (см. таблицу 1).

Таблица 1.

*Ретинальсодержащие белки и соответствующие им максимумы спектров поглощения.*

<i>Белок</i>	<i>Конформер ретиналя</i>	<i><math>\lambda_{max}</math> нм</i>		<i>Источник</i>
		<i>Протонированное Шиффово основание (PSB)</i>	<i>Нейтральное Шиффово основание (SB)</i>	
Родопсин (Rh)	11- <i>cis</i>	500		[21]
Синий сенсорный белок	11- <i>cis</i>	425		[22]
Зеленый сенсорный белок	11- <i>cis</i>	530		[22]
Красный сенсорный белок	11- <i>cis</i>	560		[22]
Иодородопсин (красные колбочки кур)	11- <i>cis</i>	570		[23]
Сенсорный родопсин I (SR I)	all- <i>trans</i>	580		[24]
Сенсорный родопсин II (SR II)	all- <i>trans</i>	500		[24]
Галородопсин (phR, NpHR)	all- <i>trans</i>	570		[25]
Родопсин осьминога (octR)	11- <i>cis</i>	472		[26]
Каналородопсин II (ChR2)	all- <i>trans</i>	480		[27]
<b>Бактериородопсин (bR)</b>	all- <i>trans</i>	568		[28]
Метародопсин II (форма Rh)	all- <i>trans</i>		380	[21]
М интермедиат bR	11- <i>cis</i>		412	[28]

Кроме того, в литературе имеются многочисленные данные, указывающие и на эффекты точечных аминокислотных замен на положение спектрального максимума



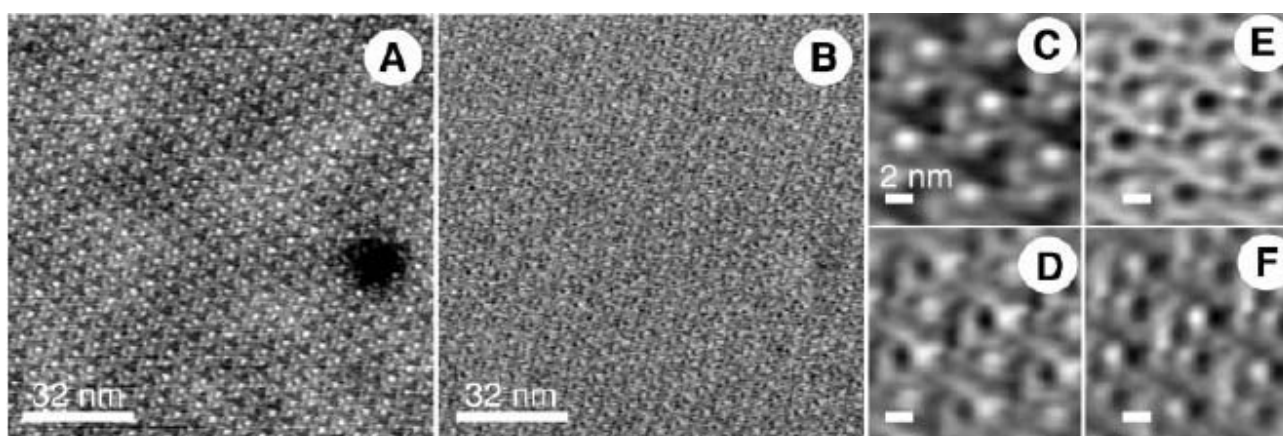
бактериородопсина. Компьютерное предсказание спектральных сдвигов является главной целью настоящей работы. Для достижения выбранной цели, учитывая существующие тенденции развития этой области компьютерной биологии, были сформулированы следующие задачи:

1. Создание адекватной модели бактериородопсина, выбор набора мутантов и получение их атомных структур, в том числе выбор оптимальной методики оптимизации структур.
2. Расчет спектров для белка дикого типа и набора мутантов различными методами квантовой химии по схеме КМ/ММ.
3. Расчет спектров бактериородопсина по методике, позволяющей оценить вклад взаимной поляризации в спектральный сдвиг.
4. Расчет спектров для набора микросостояний белка, полученных при проведении CPMD. Оценка вклада сэмплирования микросостояний (т.е. учета ансамбля микросостояний) белка в спектральный сдвиг.

## 2. ОБЗОР ЛИТЕРАТУРЫ

### 2.1. Структура бактериородопсина и его комплексов с липидами ПМ

Первые экспериментальные данные о структуре бактериородопсина были получены с помощью метода криоэлектронной микроскопии (cryo-electron microscopy) двухмерных кристаллов с разрешением вплоть до 3,0 Å [29], [30]. Эти эксперименты позволили судить о примерной геометрии белка: расположении трансмембранных  $\alpha$ -спиралей, аминокислотных остатков, ретиналя. Однако разрешение полученных моделей давало возможности с достаточной точностью определить положение связанных молекул воды, создать атомную модель белка. Появившиеся в дальнейшем модели, полученные в экспериментах по дифракции электронов (разрешение 3,0 Å [29] и в ранних экспериментах по рентгеноструктурному анализу (разрешение 2,95 Å [31], также не давали четкого представления о строении бактериородопсина на атомном уровне. Ситуация изменилась с разработкой методики кристаллизации в липидном кубике (cubic lipid phase crystallization method) [32], которая позволила повысить разрешение моделей, получаемых в экспериментах по рентгеноструктурному анализу, последовательно до 2,5 Å [4], 2,30 Å [33], 1,90 Å [34], 1,55 Å [2], а затем и до 1,43 Å [35].



**Рисунок 3. Атомно-силовая фотография ПМ, на фото C-F видна организация комплексов молекул бактериородопсина. Из [36].**

In vivo бактериородопсин, организованный в тримеры, входит в состав т.н. пурпурной мембраны (ПМ) археобактерий (см. рис. 3). Отдельные молекулы бактериородопсина в тримере контактируют очень плотно, межспиральные контакты, занимающие площадь в  $15,942 \text{ Å}^2$  (что соответствует 26,8% поверхности каждого из мономеров), представлены, в том числе, солевым мостиком между Lys 40 и Asp 104 [31]. Массовая доля бактериородопсина в ПМ может достигать 75%, также известно, что с каждым тримером контактируют 30 молекул липидов: 6 внутри, между тремя молекулами бактериородопсина и 24 вокруг тримера. Эти контакты, по большей части, имеют гидрофобную природу



**Рисунок 4. Псевдотрехмерное представление бактериородопсина (1QNJ). Апопротеин изображен в виде  $\alpha$ -спиралей, ретиналь визуализирован в виде ball-stick. Изображение получено в пакете Chimera [6].**

(гидрофобные алкильные цепи взаимодействуют с гидрофобными остатками аминокислот), также играют существенную роль в формировании белок-липидного комплекса водородные связи, образующиеся между гидроксильными группами некоторых остатков тирозина и кислородами эфирных связей липидов (к примеру, между Туг 64 и липидом 500 [34]).

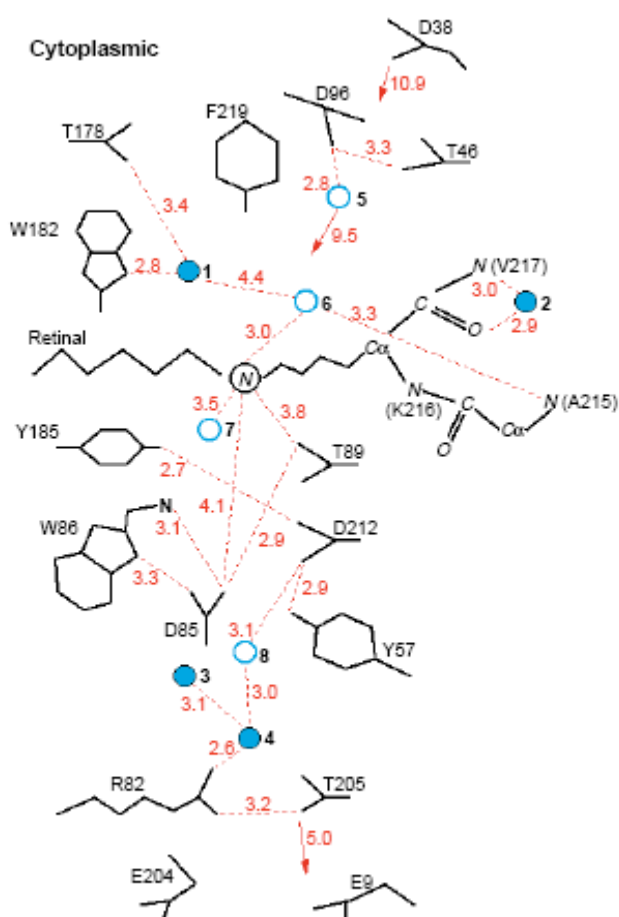
Липидный бислой ПМ состоит на 90% из полярных липидов и, соответственно, на 10% – из неполярных. Полярные липиды представлены производными 2,3- 2,3-ди-О-фитанил-sn-глицерол и включают в себя фосфатидилглицерол, фосфатидилглицеролфосфат, метиловый эфир фосфатидилглицеролфосфата, фосфатидилглицеролсульфат,

триглицозилдиэфир, триглицозит сульфат, дифитанилглицерол тетраглицозил [37].

Пространственная структура бактериородопсина, близка к белкам семейства GPCR (G-protein coupling receptors, англ. G-белок сопряженные рецепторы; представители этого семейства,

среди которых – зрительный родопсин, адренорецепторы, широко распространены в организмах млекопитающих, где выполняют главным образом рецепторные функции), что отражено и в самом названии бактериородопсина. Из 248 аминокислот, входящих в состав бактериородопсина, большая часть образует 7 трансмембранных гидрофобных  $\alpha$ -спиралей, обычно обозначаемых литерами от А до G. Небольшие внеклеточные сегменты, включая внеклеточный N-конец и цитоплазматический C-конец, согласно исследованиям [38], не играют существенной роли в функционировании белка, как и организация бактериородопсина в тримеры в мембранах *in vivo*.

Спираль E, F и G расположены практически перпендикулярно к плоскости мембраны, в то время как спирали A, B, C и D пересекают мембрану под небольшим углом к нормали (см. рис. 4). Вместе семь спиралей образуют трансмембранный канал, разделенный ретиналем, присоединенным к  $\epsilon$ -аминогруппе Lys 216, на две полости: цитоплазматический и внеклеточный полу-каналы. Термин «канал» в данном контексте не совсем уместен (во всяком случае, здесь он не обозначает пору): на самом деле, как будет видно из дальнейшего



**Рисунок 5.** Схематическое изображение основных остатков, экспонированных в белковую полость. Изображение взято из [5].

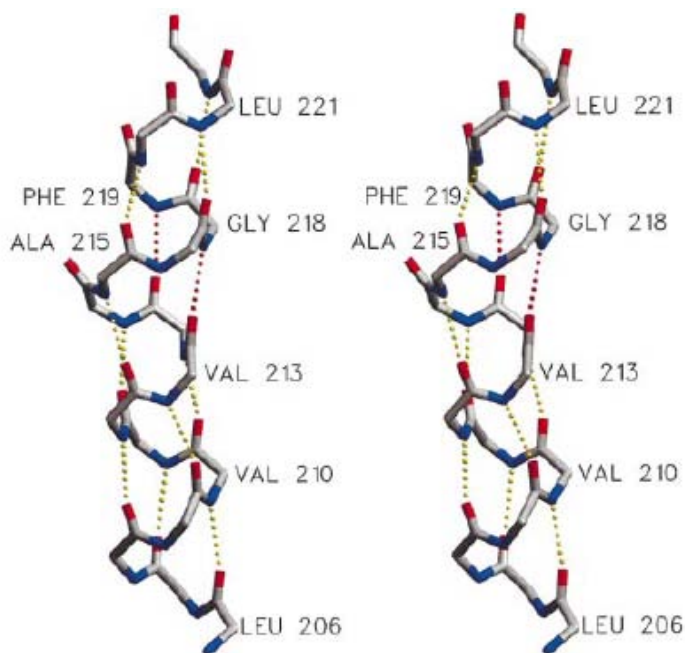
донором протона для Шиффова основания в процессе протонного транспорта. Эти аминокислоты участвуют в организации сложной сети водородных связей, по которой протон доставляется к Asp 96 из цитоплазмы.

Ряд гидрофобных ароматических остатков, как в цитоплазматическом, так и во внеклеточном полуканале закрепляют ретиналь, ограничивая его подвижность в ходе фотоизомеризации. С дистальной (по отношению к Шиффову основанию) стороны это осуществляется благодаря гидрофобным взаимодействиям с Trp 138 и Trp 189, а с проксимальной – за счет Trp 86 и Trp 182. Стерическое взаимодействие между 13-метильной группой и Trp 182, а также между 9-метильной группой и Leu 93 показано в опытах с аминокислотными заменами. Таким образом, гидрофобный карман, в котором расположен ретиналь, является достаточно жестким.

рассмотрения, реального переноса протона из цитоплазматического пространства во внешнюю среду в фотоцикле не происходит.

Во внеклеточном сегменте трансмембранного канала расположено большое число полярных аминокислотных остатков, принимающих участие в акцепции протона с Шиффова основания и его дальнейшем транспорте. Самый важный из них – остаток Asp 85, основной компонент сложного противоиона к протонированному Шиффову основанию и акцептор протона Шиффова основания. Другие важные остатки – это Asp 212 (входит в состав противоиона вместе с Asp 85), Arg 82 (стабилизирует заряженные Asp 85 и Asp 212), Glu 204, Glu 194.

Цитоплазматический полуканал составляют, главным образом, гидрофобные аминокислоты, за исключением Asp 96, который является



**Рисунок 6. Стереонизображение  $\pi$ -изгиба.**  
Изображение взято из [2].

Как показывают эксперименты с точечными мутациями, замена любой из этих аминокислот резко тормозит протонный транспорт, хотя полная его остановка происходит лишь при замене Lys 216 и Asp 85.

Также стоит заметить, что ряд  $\alpha$ -спиралей, формирующих бактериородопсин, имеют нерегулярные участки и петли. В спиральях В, С и F нерегулярные участки возникают из-за остатков пролина, а в спирали G – за счет образования  $\pi$ -спирального участка (один оборот спирали) в районе Lys 216.

### 2.1.1. Сеть водородных связей

Хорошо известно из литературы (к примеру, [2]), что связанная (англ. bound water) вода организована в бактериородопсине в сеть, будучи связанной водородными связями между собой и с полярными аминокислотами трансмембранного канала. При этом ряд молекул воды ограничен в своей подвижности и имеет фиксированное положение в структуре белка (что характеризуется низким значением В-фактора). Эти молекулы особенно важны для функционирования белка, т.к. образуемая ими сеть водородных связей непосредственным образом участвует в процессе трансмембранного переноса протона. Последнее подтверждается экспериментами по обезвоживанию (как за счет высушивания, так и за счет осмотических агентов) бактериородопсина. Показано, что при умеренной дегидратации главным образом ингибируется переход протона с Asp 96 на Шиффово основание [39].

Положение этих молекул воды, каждой из которых приписан определенный номер (в работе использованы обозначения для молекул воды как в структуре [2]) известно из данных рентгено-структурного анализа.

Во внеклеточном полуканале Asp 85 связан с Thr 89, водой 402 и водой 401; Asp 212 связан с водой 406, Tyr 57 и Tyr 185.

Несколько молекул воды (вода 406, 407, 408 и 409) расположены ближе к наружной поверхности, они образуют сеть водородных связей, включающую аминокислотные остатки Arg 82, Thr 205 и Glu 204. Эта сеть представляет собой путь, по которому протон с первичного акцептора переносится во внеклеточное пространство [4].

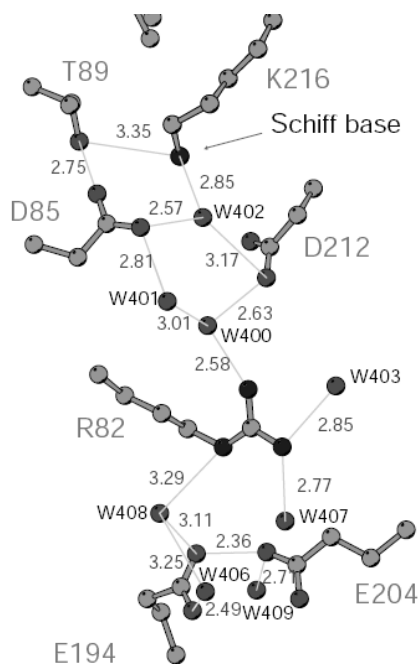


Рисунок 7. Сеть водородных связей во внеклеточном полуканале. Изображение из [4].

В цитоплазматическом полуканале неподалеку от середины спирали G расположены две молекулы: вода 501 и вода 502, связанные водородными связями с C=O группами Ala 215 и Lys 216 соответственно. Вода 501 также образует водородную связь с индоловым кольцом Trp 182, расположенным в спирали F, связывая таким образом спирали F и G. Водородные связи в спирали G нарушают ее жесткую  $\alpha$ -спиральную геометрию, приводя к образованию короткого  $\pi$ -спирального сегмента, т.н. « $\pi$ -bulge» [4] (см. рис. 6). Этот участок делает возможной относительную пространственную подвижность спиралей F и G, связанных, как было сказано выше, водородными связями, которая, как показывают исследования, и происходит при фотоизомеризации ретиналя.

Для цитоплазматического полуканала, точнее для его части, находящейся в непосредственной близости от первичного донора протона, Asp 96, характерно отсутствие полярных аминокислотных остатков и воды. Это, а также водородная связь, образованная COOH-группой аспартата и Thr 46, обуславливает необычайно высокое значение  $pK_a$  Asp 96.

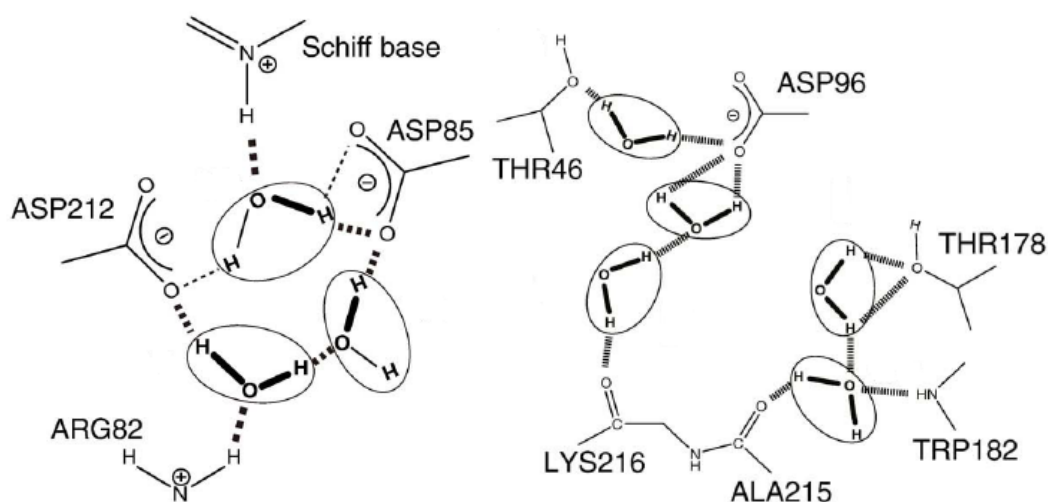


Рисунок 8. Слева - основные молекулы воды во внеклеточном полуканале и сеть водородных связей. Справа - основные молекулы воды в цитоплазматическом полуканале и сеть водородных связей. Изображение из [40].



### 2.1.2. Межспиральные контакты и спиральная подвижность

Межспиральные контакты играют существенную роль как в организации и поддержании третичной структуры бактериородопсина, так и в обеспечении его функционирования. Данные, приведенные в нижеследующей таблице, взяты из статьи [2] и получены из кристаллической структуры 1C3W (X-ray, 1.55 Å).

Таблица 2.

*Перечень остатков, формирующих межспиральные водородные связи (по [2]).*

<b>Спирали</b>	<b>Остатки, формирующие водородную связь</b>
A - B	Arg7 (NH2) and Leu61 (O)
	Arg7 (NH2) and Met60 (O)
B - C	Thr46 (OG1) and Asp96 (OD2)
	Tyr57 (OH) and Arg82 (NH1) via water 407
C - D	Leu87 (O) and Asp115 (OD1) via water 511
	Thr90 (OG1) and Asp115 (OD1)
D - E	Met118 (O) and Ser141 (OG)
	Ala126 (O) and Arg134 (NH1)
	Thr128 (O) and Arg134 (NH1)
E - F	Trp138 (NE) and Pro186 (O)
	Arg134 (NE) and Glu194 (O)
	Arg134 (NH2) and Glu194 (O)
	Leu152 (O) and Arg175 (NH1)
F - G	Trp182 (NE) and Ala215 (O) via water 501
	Tyr185 (OH) and Asp212 (OD1)
	Ser193 (OG) and Glu204 (OE1)
	Ser193 (N) and Glu204 (OE1)
	Glu194 (OE1) and Glu204 (OE2)
G - A	Leu224 (O) and Lys30 (NZ)
A - C	Arg7 (NH1) and Tyr79 (OH)
	Glu9 (OE1) and Tyr79 (OH)
B - F	Tyr57 (OH) and Asp212 (OD2)
B - G	Tyr57 (OH) and Thr205 (O) via water 407
C - F	Tyr79 (N) and Glu194 (OE1) via water 404
	Tyr83 (OH) and Trp189 (NE1)

C - G	Thr46 (O) and Lys216 (O) via water 502
	Tyr79 (O) and Glu204 (OE2) via water 405
	Arg82 (NH1) and Thr205 (O) via water 407
	Arg82 (NH2) and Thr205 (O) via water 407
	Asp85 (OD2) and Asp212 (OD1) via water 402
	Asp85 (OD2) and Lys216 (NZ) via water 402

### 2.1.3. Обзор существующих структур бактериородопсина

В таблице приведены большинство имеющихся в белковом банке данных (Protein Data Bank, <http://www.rcsb.org/pdb/>) структур бактериородопсина и его мутантов в основном состоянии или на той или иной стадии фотоцикла.

Таблица 3.

*Структуры бактериородопсина и его мутантов с атомным разрешением, представленные в базе данных PDB.*

<b>PDB код</b>	<b>Метод</b>	<b>Разре- шение</b>	<b>Остатки, представленные в модели</b>	<b>Ссылка</b>	<b>Комментарии</b>
<b>Модели бактериородопсина в основном состоянии</b>					
1BRD	EM	3.5	8-32, 38-62, 74-100, 106-127, 137-157, 166-191, 202-225	[41-42]	WT, all- <i>trans</i>
2BRD	EM	3.5	7-227	[29]	WT, all- <i>trans</i>
1AT9	EM	3.0	2-231	[30]	WT, all- <i>trans</i>
2AT9	EM	3.0	6-227	[43]	WT, all- <i>trans</i>
1FBB	EM	3.2	4-227	[44]	WT, all- <i>trans</i>
1AP9	X-ray	2.35	7-225	[4]	WT, all- <i>trans</i>
1BRR	X-ray	2.9	3-232	[31]	WT, all- <i>trans</i>
1BRX	X-ray	2.3	6-152, 167-228	[33]	WT, all- <i>trans</i>
1C3W	X-ray	1.55	5-156-162-231	[2]	WT, all- <i>trans</i>
1BM1	X-ray	3.5	7-227	[45]	WT, all- <i>trans</i>
1QHJ	X-ray	1.9	5-232	[34]	WT, all- <i>trans</i>
1E0P(A)	X-ray	2.1	5-232	[46]	WT, all- <i>trans</i>
1CWQ(A)	X-ray	2.25	2-239	[47]	WT, all- <i>trans</i>
1QM8	X-ray	2.5	2-230	в печати	WT, all- <i>trans</i>



1KGB	X-ray	1.81	5-156, 162-231	[48]	WT, all- <i>trans</i>
1IW6	X-ray	2.3	5-231	[49]	WT, all- <i>trans</i>
1KME	X-ray	2.0	5-231	[50]	WT, all- <i>trans</i>
1XOS	X-ray	2.5	5-231	[51]	WT, 13- <i>cis</i> , 15- <i>syn</i>
1MOL	X-ray	1.47	5-156, 162-231	[35]	WT, all- <i>trans</i>
<b>Интермедиаты фотоцикла</b>					
1FBK	EM	3.2	4-228		M, D96G/F171C/F219L
1QKO	X-ray	2.1	5-232		K, WT
1QKP	X-ray	2.1	5-232		K, WT
1C8R	X-ray	1.8	5-156, 162-231		основное состояние, D96N
1C8S	X-ray	2.0	5-153, 176-222		поздний M, D96N
1F50	X-ray	1.7	5-156, 162-231		основное состояние, E204Q
1F4Z	X-ray	1.8	5-156, 162-231		ранний M, E204Q
1E0P	X-ray	2.1	5-232		L, WT
1CWQ	X-ray	2.25	2-239		M, WT
1DZE	X-ray	2.5	6-230		M, WT
1JV7	X-ray	2.25	9-63, 78-232		O, D85S
1JV6	X-ray	2.0	9-63, 78-230		D85S/F219L
1KG8	X-ray	2.0	5-155, 167-231		ранний M, WT
1KG9	X-ray	1.81	5-156, 163-231		WT
1IXF	X-ray	2.6	5-231		K, WT
1M0K	X-ray	1.43	5-156, 162-231		K, WT
1M0M	X-ray	1.43	5-156, 162-231		M1, WT

## 2.2. Фотоцикл. Механизм переноса протона

### 2.2.1. Общие замечания

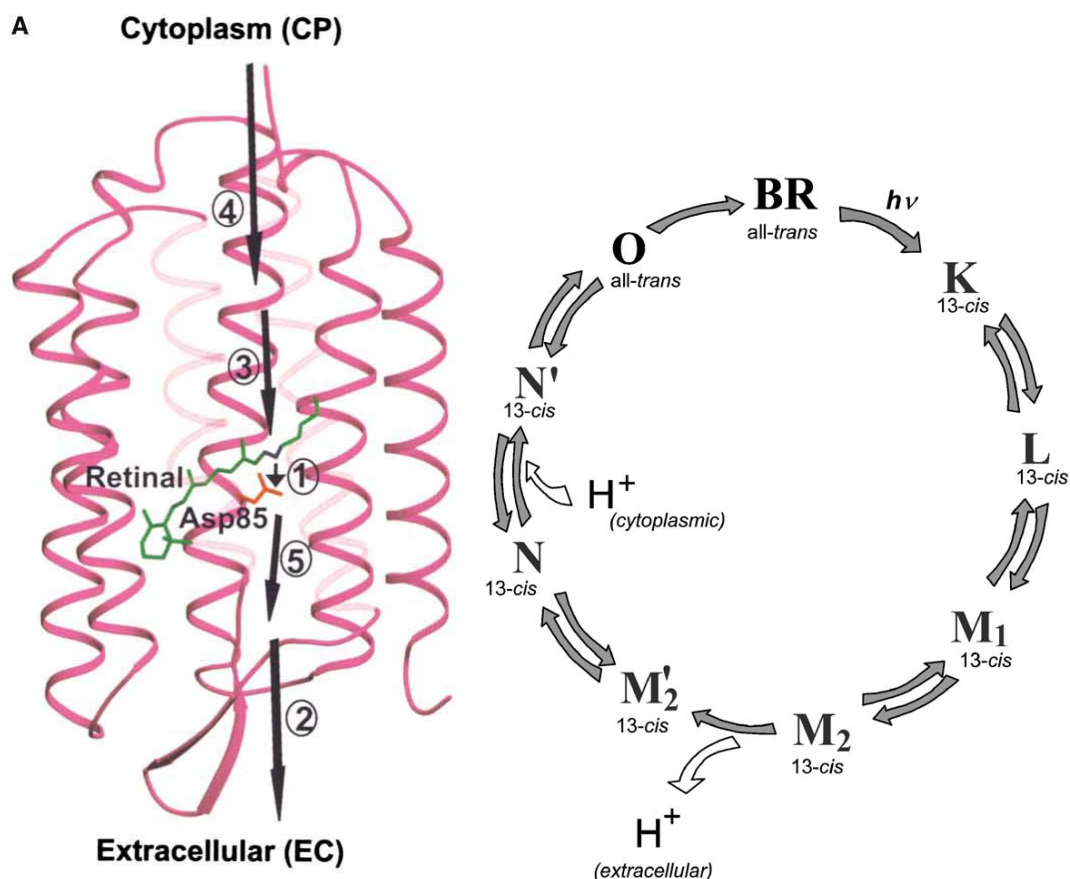


Рисунок 9. Слева - схема фотоцикла бактериородопсина. Справа - схематическое изображение бактериородопсина, на котором отмечены основные этапы фотоцикла. Изображения взяты из [52].

Еще в самых ранних работах, посвященных изучению динамики протонного транспорта бактериородопсином, были описаны квази-стабильные состояния белка, различимые спектроскопически. Эти стадии, последовательно сменяющие друг друга в т.н. *фотоцикле*, были обозначены латинскими буквами от J до O. Лишь первая из реакций фотоцикла зависит от света, остальные реакции представляют собой реакции термического типа, сходные с такими же реакциями в других биологических транспортных системах. Каждая реакция соответствует определенному физическому процессу, поэтому ряд спектроскопически неразличимых (батохромных) состояний был разделен впоследствии на под-состояния: M1, M2, M2' и N, N'.

На рисунке представлены основные этапы фотоцикла с точки зрения процесса трансмембранного переноса протона. На первом этапе, следующем за изомеризацией ретиналя, осуществляется первичный перенос протона с Шиффова основания на Asp 85; за ним следует выброс протона во внеклеточное пространство; следующий этап –

репротонирование Шиффова основания, заключающееся в переносе на него протона с Asp 96, который затем протонируется за счет цитоплазматического протона; заключительный этап цикла – депротонирование Asp 85 и реизомеризация ретиналя. Весь фотоцикл занимает порядка 10 мс, причем бактериородопсин не обладает рефрактерным периодом [53], т.е. по завершении одного фотоцикла, он готов к следующему, максимальная достижимая скорость протонного транспорта составляет 100 протонов в секунду на одну молекулу белка, однако даже на ярком свете, бактериородопсин не достигает светового насыщения.

Таблица 4.

*Характеристика интермедиатов фотоцикла бактериородопсина.*

<b>Назв.</b>	<b>Максимум поглощения (нм)</b>	<b>Конф. ретиналя (C13=C14)</b>	<b>Конф. Шиффова основания (C15=N)</b>	<b>Состояние протонирования Шиффова основания</b>	<b>Температура «замерзания» стадии, °C</b>	<b>Квантовый выход, %</b>
bR	570	All-trans	anti	+	-	64 <sub>B→K</sub>
K	586	13-cis	?	+	-190	
L	544	13-cis	?	+	-100	
M1	409	13-cis	anti	-	-30	
M2	409	13-cis	anti	-	-10	
N	562	13-cis	anti	+	-60	
O	629	All-trans	anti	+		
D	550	13-cis	syn	+		

## 2.2.2. Фотоизомеризация ретиналя. Интермедиаты К и J

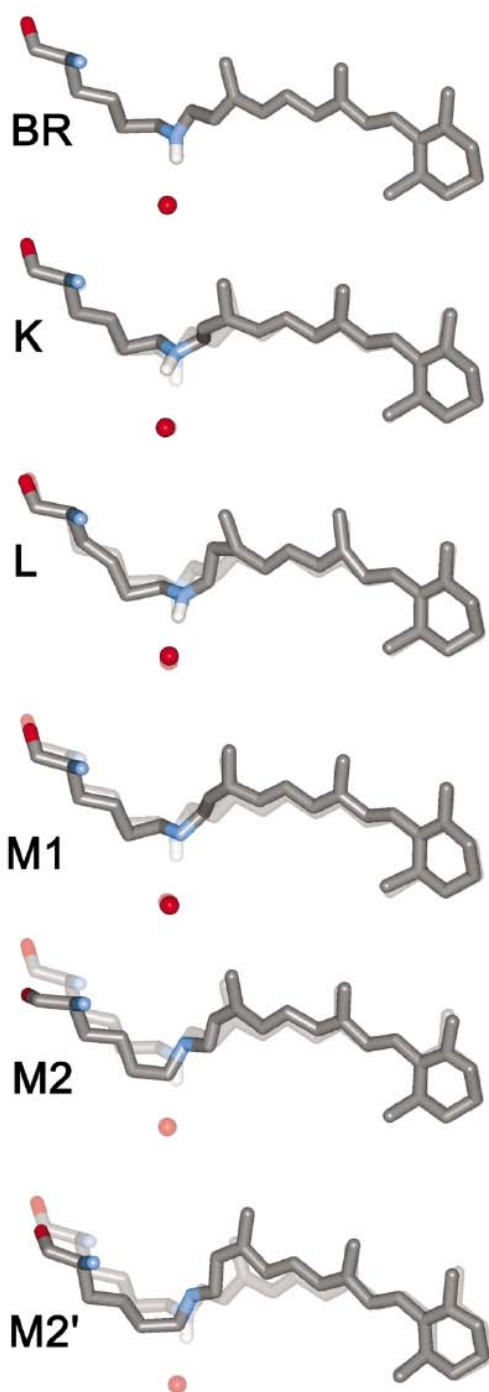


Рисунок 10. Атомные структуры ретиналя, взятые из кристаллических структур интермедиатов фотоцикла, наложенные на структуру бактериородопсина в основном состоянии. Изображение взято из [1].

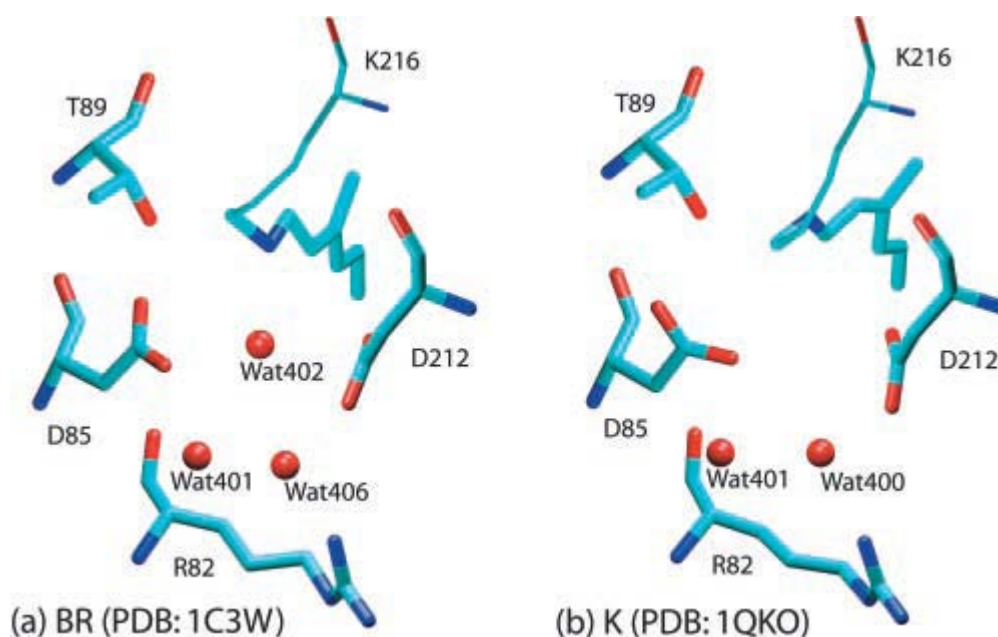
При освещении ретиналя в *all-trans* форме происходит его фотоизомеризация в форму *13-cis*, *15-anti*. Т.к. ретиналь жестко закреплен в связывающем его кармане, то фотоизомеризация по связи  $C_{12}-C_{13}$  вызывает напряженность, реализующуюся в виде изгиба участка ретиналя между  $C_{14}-N_Z$  [54]. Сам фотоцикл можно рассматривать, как релаксацию этой напряженной конформации, движимую избыточной энтальпией (порядка 50 кДж/моль, [55]), запасенной после распада возбужденного состояния.

После поглощения кванта за несколько пикосекунд образуется форма K610, которая является стабильной при температурах ниже 100 К, и распадается за времена порядка микросекунд при физиологической температуре. Максимум спектра поглощения этой формы сдвинут в красную область (т.н. *red-shift*), что обусловлено, однако, не самим фактом фотоизомеризации ретиналя, но изменениями в его белковом окружении. Как видно из кристаллической структуры этого интермедиата [56], основные изменения, происходящие при переходе bR – K, затрагивают, главным образом, молекулу воды 402, остаток Asp 85 и ретиналь (в особенности, N-H связь). В исходной структуре вода 402 жестко фиксирована за счет образования водородных связей с Lys 216, Asp 85 и Asp 212. В образовавшемся интермедиате K вода 402 делокализована, водородная связь между Thr 89 и Asp 85, как ранее считалось, разрывается, и

карбоксильная группа последнего смещается, однако, по новым данным, в том числе полученным из молекулярной динамики [54] и из экспериментов по LT-FTIR [57], [58-60] эта водородная связь сохраняется. Шиффово основание на этой стадии фотоцикла остается протонированным. Как показывают эксперименты по молекулярной динамике [54], [61]

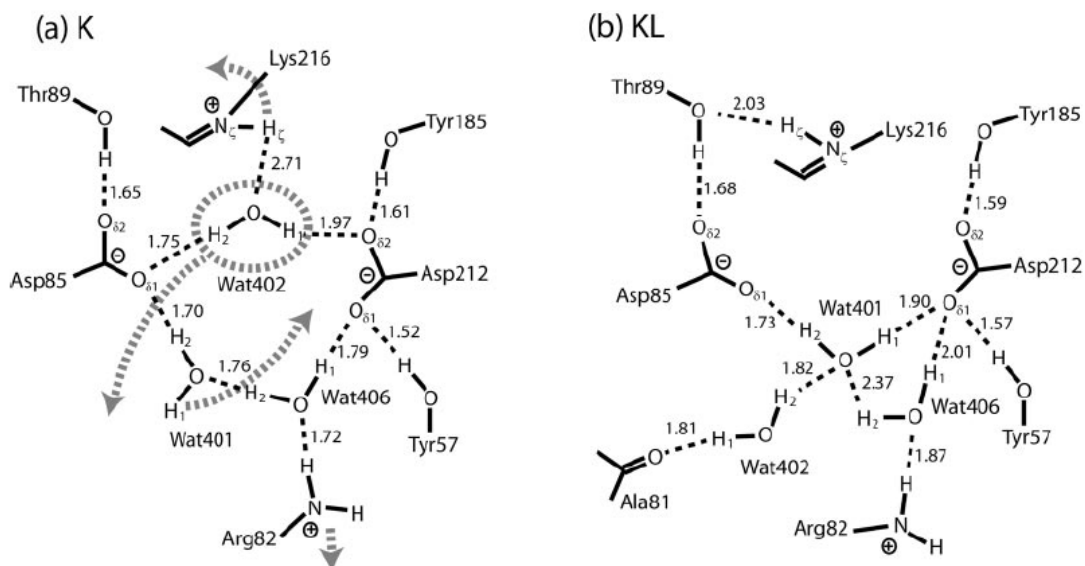
изомеризация сопровождается на начальном этапе поворотом N-H связи Шиффова основания по направлению к Asp 212, а не к Asp 85.

В релаксированной структуре ретиналя в форме *13-cis*, *15-anti* протон Шиффова основания повернут относительно своего исходного состояния на  $180^\circ$  сторону гидрофобного цитоплазматического полуканала. Однако на стадии интермедиата К эта конфигурация еще не реализована не до конца, из-за стерических конфликтов ретиналя с белковым окружением, не успевшим еще на этой стадии существенно поменяться. Для ее достижения необходимо, во-первых, изменение структуры белкового окружения, во-вторых, перестройка сети водородных связей, организованной вокруг молекул воды 401 и 402 и иммобилизирующей Шиффово основание, и, в-третьих, удаление заряда на Шиффовом основании, электростатически взаимодействующего с отрицательно заряженными остатками Asp 85, Asp 212 и другими составляющими контриона. Эти изменения растянуты во времени и полностью реализуются к стадии M1 фотоцикла.



**Рисунок 11.** Структура внеклеточного полуканала, прилегающего к Шиффову основанию в основном состоянии и в интермедиате К, полученная из кристаллических структур. Изображение взято из [5].

В ряде исследований интермедиат К разделяют на две батохромные формы: раннюю К форму и позднюю KL форму (в том числе по данным спектроскопии в видимом диапазоне волн [62], FTIR [63], [64-66], [67], [68], [69] резонансной Раман-спектроскопии [70], молекулярной динамики [54]. KL форма образуется спустя 10 нс (по данным спектроскопии в видимом диапазоне волн) или 50-100 нс (по данным TR-FTIR). В [54] приводятся следующие структурные изменения сайта связывания ретиналя, связанные с переходом К-KL, полученные из экспериментов по молекулярной динамике.



**Рисунок 12. Схема организации внеклеточного полуканала на стадии интермедиата К и KL. По [54].**

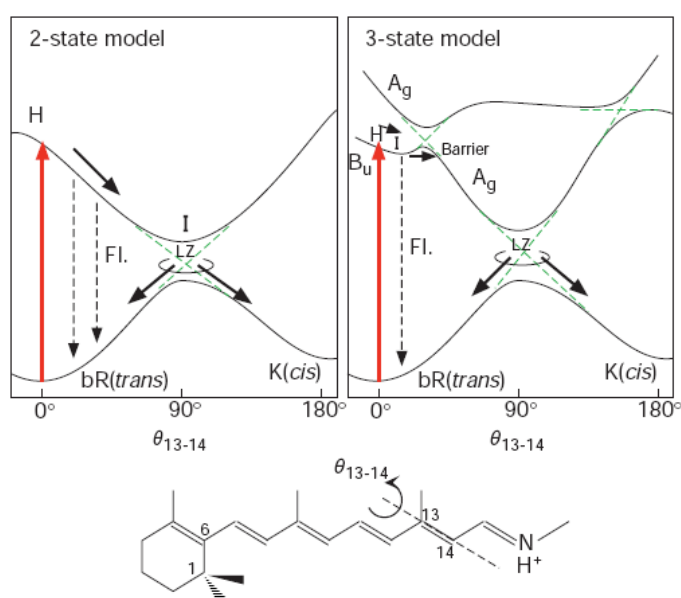
Вода 402, связанная водородной связью с протоном Шиффова основания уходит, при этом разрываются и ее связи с Asp 85 и Asp 212, а N-H связь Шиффова основания поворачивается в сторону цитоплазматического канала, где образует слабую водородную связь с Thr 89, из-за чего участок ретиналя с Шиффовым основанием снова деформируется, но теперь N-H связь поворачивается в сторону, противоположную положению этой связи в К интермедиате. Также на стадии KL происходит смещение положительно заряженного аминокислотного остатка Arg 82 (на 1,5 Å в сторону внеклеточного пространства), которое не обнаруживается в спектроскопических экспериментах (в которых, однако, наблюдается еще большее смещение этого аминокислотного остатка на более поздних этапах фотоцикла, L [46] и M [71], [47], [48], но выявляется в экспериментах по молекулярной динамике [54], что, вероятно, связано с отличной от нативной кинетикой этих процессов при низкотемпературных условиях проведения спектроскопических экспериментов.

Форме K610 предшествует J625. Эта форма образуется через 500 фемтосекунд после поглощения кванта и переходит в интермедиат К спустя несколько пикосекунд. Последний переход сопровождается диссипацией избытка энергии ретиналя и глубокой перестройкой его молекулярного скелета. Сильное красное смещение спектра этой формы обусловлено поляризацией белковой матрицы под действием измененного дипольного момента возбужденного ретиналя. Считается, что начальный этап фотоизомеризации происходит за времена порядка 100-200 фемтосекунд (т.е. в белковом окружении значительно быстрее, чем в растворе, где характерное время фотоизомеризации составляет  $10^{-11}$  с).

Рассмотрим подробнее ранние события, происходящие при изомеризации ретиналя. При поглощении одного кванта света ретиналь совершает адиабатический переход на первый

возбужденный электронный уровень, область потенциальной поверхности на которой он при этом находится сразу после перехода получила название области Франка-Кондона (т.к. переход происходит по принципу Франка-Кондона) и обозначается литерой Н. На этом этапе молекулярный скелет ретиналя остается неизменным, изменяется лишь его электронное состояние. Затем Н форма распадается, частично возвращаясь за счет флуоресценции в основное состояние, а частично релаксируя к следующему интермедиату изомеризационного процесса, который различается спектроскопически,  $I_{460}$ . Напряженность структуры постепенно растет, появляются внеполоскостные колебания протона при  $C_{14}$ , изменяется частота колебаний  $C_{13}-C_{14}$  связи. Из состояния  $I_{460}$  ретиналь переходит за несколько сотен фемтосекунд в т.н. область Ландау-Зенера (или англ. CI: conical intersection), где происходит адиабатический переход на основной электронный уровень. На этом этапе структура ретиналя соответствует некоему среднему между *cis*- и *trans*- формами состоянию, угол связи  $C_{13}-C_{14}$  составляет  $\approx 90^\circ$ . Такая структура, после перехода на основной электронный уровень, может либо возвратиться в начальное *trans*- состояние, или релаксировать в направлении *cis*- формы (за  $\approx 500$  фс), образуя последовательно интермедиаты J и K.

Существует две модели, описывающие структуру потенциальных поверхностей электронных уровней ретиналя. Это двухуровневая и трехуровневая схемы. В настоящее время предпочтение отдается второй модели, которая находит подтверждения, как в экспериментах [72], [73], [74], [75], [76], так и в расчетах *ab initio* [77], [78]. Существуют, тем не менее, исследования свидетельствующие в пользу двухуровневой модели [79], [80].



**Рисунок 13.** Двухстадийная и трехстадийная схемы фотоизомеризации бактериородопсина. Объяснения в тексте. Изображение взято из [76].



Основное отличие двух обозначенных схем заключается в наличии в случае трехуровневой модели локального минимума энергии на первом возбужденном уровне, который возникает за счет взаимодействия первого и второго возбужденных уровней. Этот минимум ограничен потенциальным барьером, за счет чего происходит накопление возбужденного интермедиата не в области Ландау-Зенера, а в области локального минимума, где он находится в *trans*-форме в течение нескольких сотен фемтосекунд. Это позволяет объяснить наблюдающееся в эксперименте отсутствие корреляции между временем жизни возбужденного состояния и квантовым выходом реакции фотоизомеризации, которая была бы логична, в случае справедливости двухуровневой модели.

Как видно, интермедиат К влияет на развитие всех последующих событий фотоцикла: все они, в том числе и термическая реизомеризация ретиналя обратно в *all-trans* форму [73], осуществляются за счет запасенной свободной энергии этой формы.

### 2.2.3. Депротонирование Шиффова основания. Интермедиаты L и M

При смещении Шиффова основания ретиналя ( $pK_a$  изначально  $>11$ ) в более гидрофобное окружение цитоплазматического полуканала в интермедиате KL, его  $pK_a$  падает. Одновременно с этим претерпевает глобальные изменения HBN, в том числе разрываются водородные связи между водой 402, Thr 89 и Asp 85, в результате чего  $pK_a$  последнего возрастает (изначально 2,5 [81]). Это, в целом, объясняет смещение равновесия реакции  $L \leftrightarrow M$  в сторону перехода протона на Asp 85 (характерное время реакции составляет 10  $\mu s$ ).

Столь важный для течения всего фотоцикла процесс перехода протона с Шиффова основания на Asp 85, обусловленный вышеуказанными изменениями окружения Шиффова основания, возможен благодаря особенностям структуры ретиналя. Как отмечалось ранее, при фотоизомеризации ретиналя образуется форма *13-cis 15-anti*. В отличие от *13-cis 15-syn* изомера, устроенного таким образом, что направление и положение N-H связи Шиффова основания практически не меняется при фотоизомеризации связи C<sub>13</sub>-C<sub>14</sub>, в *15-anti* форме эта связь обращена в цитоплазматический канал.



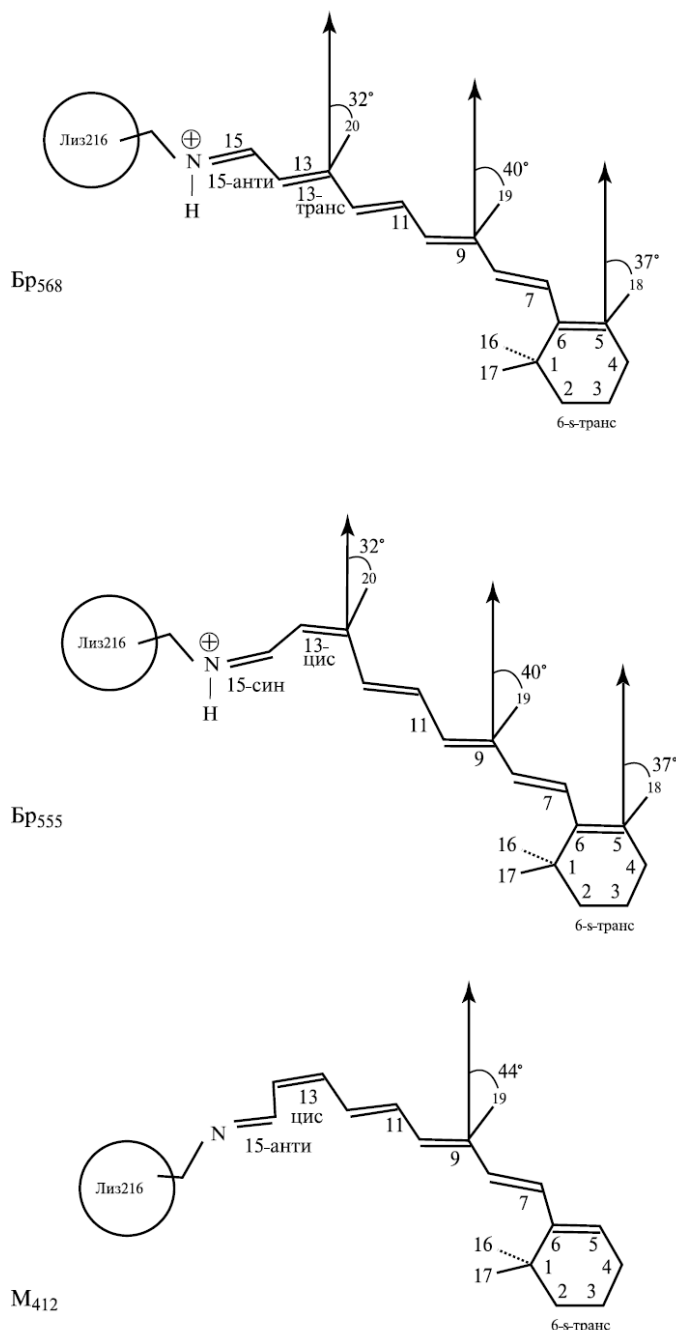


Рисунок 14. Геометрия ретиналя для различных этапов фотоцикла. Из [82].

В интермедиате L перенос протона с Шиффова основания на Asp 85 еще происходит, однако смещение спектра этой формы в коротковолновую область (610 нм – 550 нм) говорит об усилении акцепторного влияния Asp 85 и о существенных изменениях в белковом окружении Шиффова основания [82]. Переход протона на Asp 85 в M стадии сопровождается еще большим смещением спектра в красную область (550 нм – 412 нм).

Интересный факт, заключающийся в том, что первичным акцептором протона является Asp 85, а не симметрично расположенный по отношению к нему Asp 212, объясняется, исходя из кристаллической структуры интермедиата K. Низкие значения  $pK_a$  анионных форм обоих остатков обусловлены образующимися водородными связями, однако, в случае Asp 212

водородные связи образуются с остающимися неподвижными в течение фотоцикла аминокислотными остатками Tyr 57 и Tyr 185, а в случае Asp 85 водородные связи образуются с молекулой воды 401 и Thr 89, которые оба подвижны и изменяют в ходе фотоцикла свое положение, в результате чего увеличивается  $pK_a$  Asp 85 [83].

Хотя, как указано выше, основная причина перехода протона на Asp 85 ясна, детали этого механизма нельзя считать совершенно ясными. Существует ряд гипотез, объясняющих этот процесс, и, по всей видимости, некоторые из них соответствуют реально происходящим в тех или иных условиях процессам, что находит подтверждение в вычислительных экспериментах [84], [40].

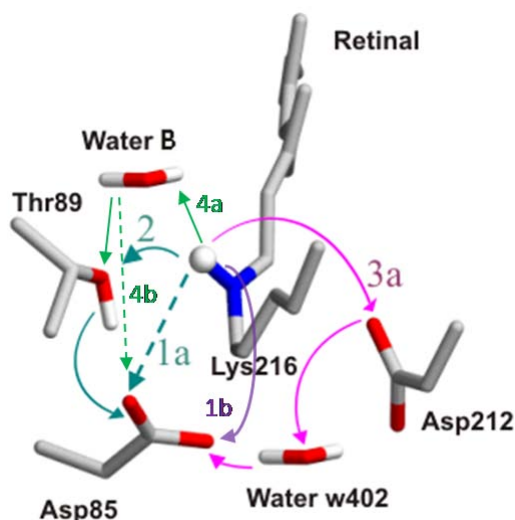


Рисунок 15. Возможные пути перехода протона с Шиффова основания на Asp 85. Из [84].

Рассмотрим возможные сценарии перехода протона с Шиффова основания на Asp 85.

- Переход протона непосредственно с HSB на Asp 85 (1a). Активизационный барьер этого процесса составляет 12,4 ккал/моль, что удивительно мало, учитывая, что дистанция между Asp 85 и N-H связью Шиффова основания составляет  $\text{\AA}4\text{\AA}$  (в переходном состоянии она сокращается до  $2,5\text{\AA}$ ). Отмечено также, что при ограничении белковой подвижности сайта связывания ретиналя активизационный барьер вырастает до 23,4 ккал/моль. Существует гипотетическая схема передачи протона напрямую на Asp 85, но с другой стороны от ретиналя (1b); энергетический барьер такой реакции – 27,8 ккал/моль, из чего следует малая вероятность ее протекания *in vivo*.
- Переход протона через Thr 89 (2) [44], [85]. Энергетический барьер реакции 13,6 ккал/моль. Энергетический барьер этого процесса, как и предыдущего, сильно снижается при удалении воды 402. Это обусловлено разрывом водородной связи между этой молекулой воды и Asp 85, в результате которого подвижность остатка аспартата

увеличивается (к примеру, в случае прямого переноса протона (путь 1a) активизационный барьер уменьшается с 12,4 ккал/моль до 6,3 ккал/моль).

- Переход протона через Asp 212 и воду 402 (3a). Процесс начинается с поворота N-H связи с цитоплазматической стороны в сторону внеклеточного канала, что сопряжено с преодолением энергетического барьера в 8,2 ккал/моль. Затем происходит переход протона с Шиффова основания на Asp 212 (барьер – 11,5 ккал/моль), а затем – на Asp 85 через воду 402 (барьер – 4,6 ккал/моль).
- Переход через воду В и Thr 89 (4a) [84]. Перенос протона в этом случае осуществляется через молекулу воды В, расположенную неподалеку от Шиффова основания в цитоплазматическом полуканале и аминокислотный остаток Thr 89; энергетический барьер процесса равен 9,7 ккал/моль.
- Переход непосредственно через воду В (4b) [84]. В этом случае перенос протона осуществляется за счет одной лишь воды В (обозначение из модели 1QKO, соответствует воде 401), мигрирующей в ходе процесса по направлению к Asp 85 и обратно. Энергетический барьер процесса – 13,8 ккал/моль.

Экспериментально определенная энтальпия активации [86] равняется 13 ккал/моль, что соответствует механизмам 1a, 2, 3a, 4b. Энергия активации для механизма 4a несколько ниже. Что касается частных условий возможности тех или иных механизмов, то все они соблюдаются *in vivo*: подвижность белкового окружения ретиналя позволяет осуществить прямой перенос протона по механизму 1a, сеть водородных связей обуславливает принципиальную возможность первичного переноса протона по механизмам 2, 3a, 4a, 4b. Вода 402, важная для передачи протона по механизму 3a, занимает, согласно последним исследованиям [87], [35], [5] (ранее этот факт ставился под сомнение), определенную фиксированную позицию во внеклеточном полуканале, мало отличающуюся от ее позиции на начальной стадии фотоцикла (образует водородные связи с Asp 85 и Asp 212).

По всей видимости, *in vivo* все принципиально возможные (энергетически и стерически не запрещенные) схемы первичного переноса протона сосуществуют, реализуясь в зависимости от условий (наличие или отсутствие воды 402 и воды В, температура) с различными вероятностями [40].

#### **2.2.4. Выброс протона во внеклеточное пространство. Интермедиат М**

По ходу депротонирования Шиффова основания ретиналя, т.е. при переходе от L интермедиата к М, параллельно протекает процесс выброса протона во внеклеточное пространство. Это протон не Шиффова основания и не Asp 85, он принадлежит сети

водородных связей внеклеточного канала, основной вклад в организацию которой вносят остатки Arg 82, Glu 194 и Glu 204.

Совершенно четко можно говорить о решающей роли протонирования Asp 85 для осуществления процесса выброса протона во внеклеточное пространство. Как было показано в [88], кривая титрования Asp 85 при значениях pH от 5 до 9 имеет комплексный характер, что означает, что и Asp 85, и некоторый аминокислотный остаток во внеклеточном полуканале могут находиться в протонированном состоянии, но не оба одновременно: протонирование одного исключает подобное состояние для другого.

Согласно рентгеноструктурным данным, уже на стадии L интермедиата наблюдается пространственное смещение аминокислотного остатка Arg 82 в сторону внеклеточной поверхности. На последующей стадии M, характеризующейся полным депротонированием Шиффова основания и, таким образом, исчезновением отрицательного заряда на Asp 85, оно становится еще более заметным и достигает 1,6 Å (доль нормали к мембране) [2]. Это движение, наблюдаемое в кристаллических структурах всех интермедиатов фотоцикла после M<sub>1</sub> [89], объясняется исчезновением отрицательного заряда на Asp 85 и перестройками в HBN.

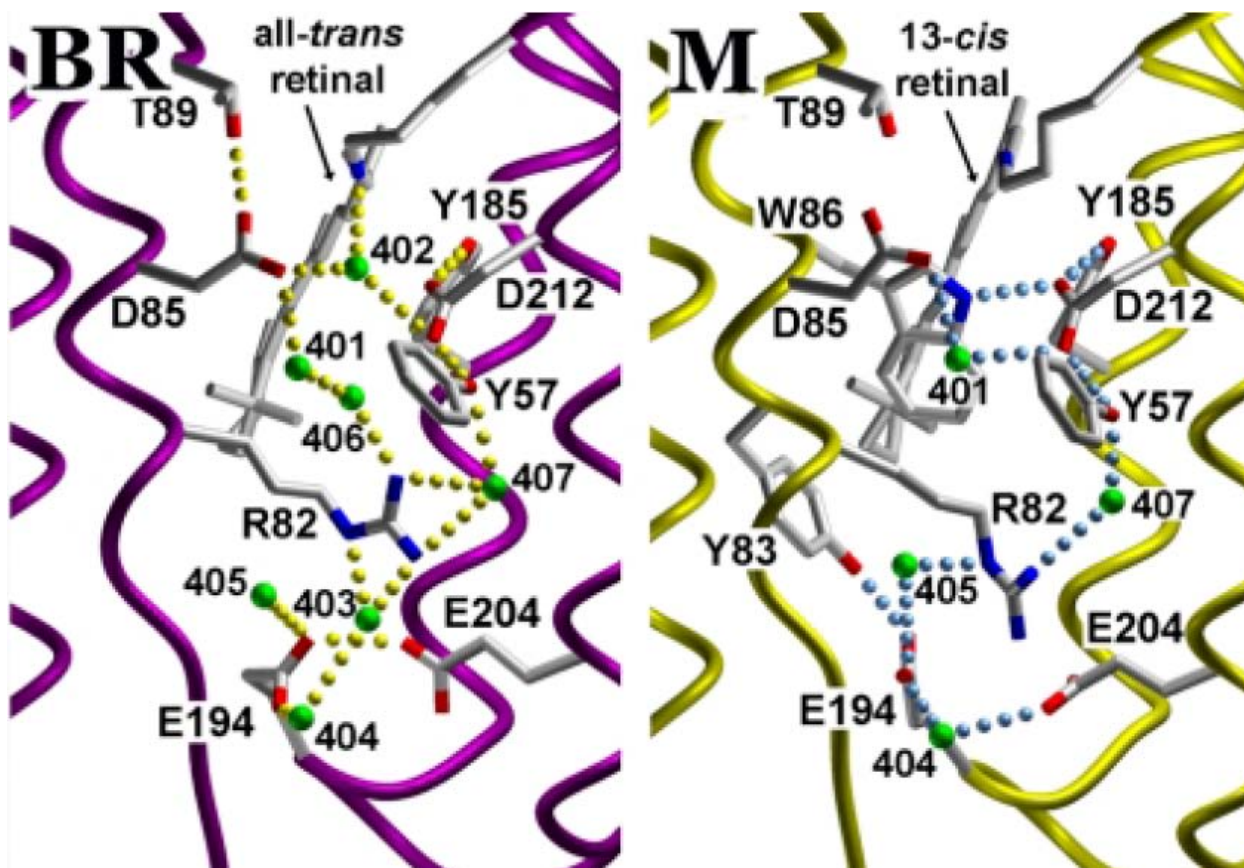


Рисунок 16. Организация сети водородных связей в основном состоянии и в интермедиате M.

В темноадаптированном состоянии Arg 82 связан с Asp 85 водородными связями через воду 406 и 401. На стадии интермедиата М водородная связь Arg 82 с молекулой воды 406 заменяется связью с водой 405, которая теперь, помимо этого, образует водородные связи с Glu 194 и Tyr 83.

Непосредственный источник протона, выбрасываемого во внеклеточное пространство, не установлен. В различных работах выдвигались предположения, что им может быть протонированный Glu 204 [90], комплекс Glu 204 и Glu 194 [91], общий протон Glu 204 и Glu 194 [31]. Против этих версий, однако, говорит факт отсутствия в инфракрасном спектре полосы, характерной для отрицательных C=O колебаний [92], возникающих при депротонировании COOH-группы глутамата. Более вероятным источником протона представляется сеть водородных связей (HBN). Собственно донором может выступать водный кластер  $H_5O_2^+$  [93].

### 2.2.5. «Переключение» Шиффова основания (Protonation Switch)

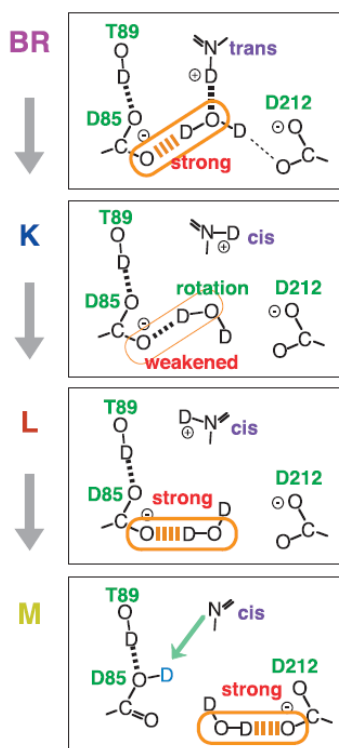


Рисунок 17. Схема начальных этапов фотоцикла.

Фотоцикл бактериородопсина должен содержать этап, на котором ориентация Шиффова основания меняется на противоположную, обращаясь в сторону цитоплазматического полуканала и Asp 96, предотвращая тем самым обратную миграцию протона на Шиффова основание с Asp 85.

Различными авторами (по [14]) предлагались всевозможные механизмы такого переключения:

- изменение кривизны полиеновой цепи, приводящее к переориентации азота Шиффова основания из внеклеточного во внутриклеточный полуканал, после депротонирования Шиффова основания;
- переориентация азота Шиффова основания во внутриклеточный полуканал в результате утраты электростатического взаимодействия с Asp 85;
- движение сегмента спирали С с Asp 85 от Шиффова основания;
- изменение  $pK_a$  Asp 85;
- переход  $OH^-$  из внеклеточного полуканала в цитоплазматический.

Таких этапов может быть несколько (и, по всей вероятности, именно так дело и обстоит), но совершенно необходимо для полноценного осуществления бактериородопсином протонного транспорта, чтобы Шиффово основание вначале было обращено к Asp 85, а после протонирования последнего поворачивалось в сторону Asp 96 [38].

#### 2.2.6. Репротонирование Шиффова основания. Интермедиаты M и N

В реакции  $M_2' \leftrightarrow N$  происходит (за несколько мс) перенос протона с первичного донора, Asp 96, на депротонированное на предыдущих этапах Шиффово основание, протекающий через стадию неустойчивого равновесия, когда и Asp 96, и Шиффово основание частично протонированы. Для осуществления такого переноса необходимо, чтобы значение  $pK_a$  Asp 96 уменьшалось по ходу процесса с 11 [94], до, приблизительно, 7,5 (т.е. не выше, чем  $pK_a$  депротонированного Шиффова основания) [95], [96], [97], [98], а также, необходима организация пути для переноса протона на 10-12 Å в достаточно гидрофобных условиях цитоплазматического полуканала.

Как было в многочисленных работах как экспериментального [39], [99], так и теоретического [100], [101] толка, функцию посредника при переносе протона в этом процессе выполняют молекулы воды, что вполне логично хотя бы из тех соображений, что альтернативных путей для этого процесса, учитывая гидрофобность цитоплазматического канала, нет [102]. Непрямым доказательством этого служит то, что равновесие  $M_2' \leftrightarrow N$  очень чувствительно к осмотическому [39] и гидростатическому [99] давлению.

В структуре  $M_2$  интермедиата, полученной рентгеноструктурным анализом кристалла мутанта E204Q бактериородопсина с разрешением 1,9 Å, поблизости от Asp 96 появляется водный кластер, структурированный водородными связями с Asp 96 и другими аминокислотными остатками [103]. Кластер содержит воду 502, связанную с карбонильной группой Lys 216 и присутствующую еще в основном состоянии бактериородопсина, воду 504, расположенную между Asp 96 и Thr 46, и еще одну молекулу воды. Видимо образование водородных связей снижает  $pK_a$  Asp 96, как то и требуется для отдачи протона на Шиффово основание. Этот водный кластер расположен так, что расстояние от ближайшей к Шиффову основанию воды до Шиффова основания сокращается и составляет на этой стадии 7,4 Å [38].



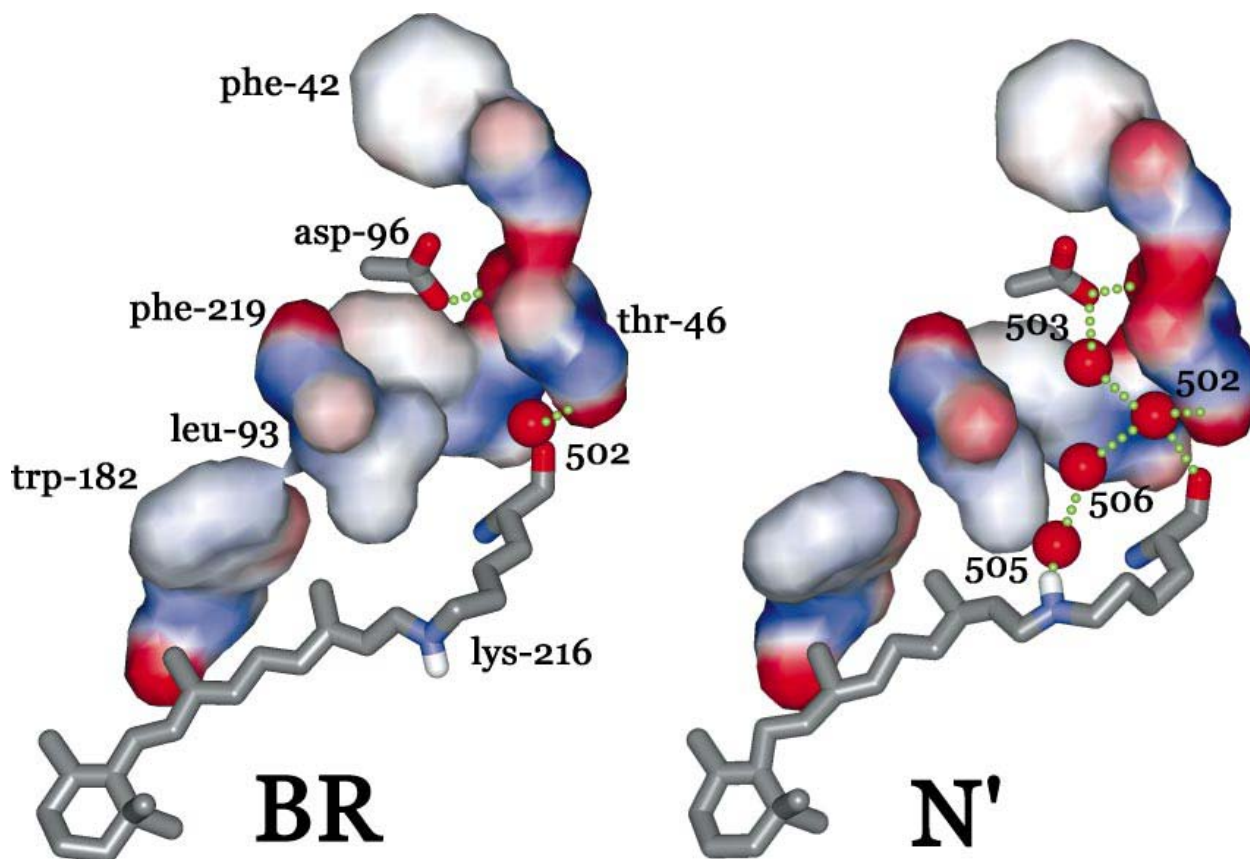


Рисунок 18. Сравнительная структура основного состояния и интермедиата N. Из [38].

Из другой структуры интермедиата М, полученной рентгеноструктурным анализом кристалла бактериородопсина дикого типа с разрешением 2,25 Å, HBN на этом этапе развита в цитоплазматическом полуканале сильнее, чем в предыдущем случае [47]. Молекулы воды 501 и 502 смешаются относительно своих позиций в основном состоянии, к ним прибавляется еще четыре новые молекулы воды. Как и в структуре мутанта E204Q одна из этих молекул расположена между Asp 96 и Thr 46. HBN распространяется от Asp 96 до Шиффова основания так, что до последнего остается расстояние 2,8 Å, причем это расстояние молекулы воды могут покрывать за счет тепловых флуктуаций.

Еще одна структура, полученная также рентгеноструктурным анализом кристаллов бактериородопсина дикого типа, но с лучшим разрешением (1,52Å) [5]. В этой структуре отсутствуют четыре дополнительные молекулы воды, как в предыдущей структуре, но возле Asp 96 имеется водный кластер, сходный с таковым структуры мутанта E204Q. Asp 96 и Thr 46 не связаны водой. Видимо эта структура соответствует более ранней форме интермедиата М.

Еще в одной известной структуре, соответствующей интермедиату N' фотоцикла (на этой стадии Шиффова основание уже репротонировано, но ретиналь еще в конформации *13-cis*), присутствует полная цепь связанных водородными связями молекул воды, протянувшаяся от

Asp 96 до Шиффова основания. В состав этой цепи входят следующие молекулы: вода 503–вода 502–вода 506–вода 505 [5].

Причиной формирования водного кластера и его дальнейшего развития считают образование внутрибелковых полостей, увеличивающихся в размере за счет структурной перестройки бактериородопсина, в которую вовлечены цитоплазматические концы  $\alpha$ -спиралей G и F [103]. F-спираль двигается в ходе этого процесса наружу, что подтверждается исследованиями рентгеноструктурного анализа [104-105], [106], [107], [108] и дифракции нейтронов [109]. В одной работе указывается также на вращательную компоненту движения F-спирали [110]. Имеются экспериментальные данные о вовлечении в эту конформационную перестройку E-F петли [111]. Движение цитоплазматического конца G-спирали в ходе этой перестройки направлено внутрь белка [112]. Помимо этого в опытах по иммобилизации F-спирали показано заметное уменьшение скорости протонного транспорта [113].

Переход протона по цепи молекул воды занимает времена порядка нескольких нс, однако реакция  $M_2' \rightarrow N$  протекает за несколько мс. Очевидно, существуют обстоятельства, затрудняющие ход репротонирования. Можно назвать три таких обстоятельства, не исключая, тем не менее, друг друга. Во-первых, такая разница в скоростях этих процессов может быть вызвана конформационными изменениями в белке, которые формируют условия для организации цепи молекул воды в цитоплазматическом полуканале. Во-вторых, цепь молекул воды может частично разрываться и вновь объединяться в результате конформационных флуктуаций белка, из-за чего в непрерывном виде она существует редко. И, в-третьих, образование и существование цепи молекул воды может не быть самым важным фактором для репротонирования. В этом случае, решающая роль отводится преодолению протоном энергетического барьера кулоновского взаимодействия с анионом Asp 96.

Как показывают структурные исследования [5] и более ранние работы [39] по термодинамике фотоцикла, основным барьером является депротонирование Asp 96, в то время как конформационные изменения белка происходят постепенно и проходят последовательно ряд этапов [83]. Замена Asp 96 на Asn существенно замедляет протонный транспорт и делает его зависимым от значения pH [39], [114], [115], что говорит о том, что в этом случае источником протона служит цитоплазматическая поверхность мембраны. В этом случае не требуется разделять заряды протона и аниона Asp 96, что снижает энтальпию реакции на шесть порядков. Однако при этом сильно возрастает энтропийный барьер, что можно объяснить отсутствием Asp 96, выполняющего роль центра организации водного



кластера, а с другой стороны способствующего захвату цитоплазматического протона (за счет отрицательного заряда).

Как уже было указано, важные структурные изменения бактериородопсина, происходящие на этом этапе фотоцикла, заключаются в удалении  $\alpha$ -спирали F от  $\alpha$ -спирали G. Причиной этого служит перестройка Lys 216 и  $\pi$ -bulge G-спирали, а также боковых аминокислотных цепей в окрестности 13-метильной группы ретиналя. 13-метильная группа постепенно, на протяжении стадий K, L и M фотоцикла, перемещается вверх, вступая в стерический конфликт с индольным кольцом Trp 182, из-за чего последний приходит в движение. Изгиб цепи Lys 216 передается на  $\pi$ -bulge G-спирали, в результате чего перемещается кислород пептидной связи Ala 215, что, в свою очередь, приводит к разрыву водородной связи между Ala 215 и Thr 178 (через молекулу воды 501), связывающей  $\alpha$ -спирали G и F. Движение Ala 215 также вызывает смещение остатков Leu 181 и Phe 219, образующих межспиральных контакт (F и G  $\alpha$ -спиралями). В итоге Asp 96 и Thr 46 располагаются таким образом, что становится возможным проникновение в пространство между этими остатками молекулы воды 504 [103].

#### 2.2.7. Захват протона на цитоплазматической поверхности и реизомеризация ретиналя. Интермедиат N, N' и O

Кажется весьма вероятным наличие сопряжения между изомерной формой ретиналя и состоянием (протонированный/депротонированный) Asp 96, из-за которого процесс депротонирования Asp 96 из внешнего цитоплазматического источника и процесс реизомеризации ретиналя в состояние *all-trans* протекают параллельно. В частности, это подтверждается экспериментами с направленным мутагенезом. Показано, что время жизни N фотоинтермедиата увеличивается в мутантах с аминокислотными заменами не только поблизости от ретиналя, но и в окрестности Asp 96. Более убедительным доказательством гипотезы сопряжения реизомеризации ретиналя и депротонирования аспартата являются свойства двойного мутанта Asp 85 Asn/Phe 42 Cys. Этот мутант достаточно точно имитирует интермедиат N: на Asp 85 отсутствует отрицательный заряд, а  $pK_a$  Asp 96 снижен (с более чем 13 у дикого типа, до 8,2) за счет замены фенилаланина на цистеин. Депротонирование Asp 96 можно вызвать повышением pH до 8, что влечет за собой в темноте изомеризацию ретиналя в 13-*cis*, 15-*anti* [116]. С другой стороны, у некоторых мутантов (в частности, у V49A и F171C), для которых характерно долгое время жизни интермедиата N, захват протона с цитоплазматической поверхности (т.е. депротонирование аспартата) происходит в 10-100 раз быстрее изомеризации ретиналя. Таким образом, у этих мутантных форм при pH < 7 четко различимы две формы интермедиата N: с анионной (N) и протонированной (N')

формами Asp 96 [116]. Вероятно, разобщение реизомеризации и репротонирования в этих экспериментах связано с нарушением устройства специфической цепи водных молекул, о структуре которой речь пойдет ниже.

Механизм сопряжения связан с цепью молекул воды, связанных водородными связями. Молекула воды 505 образует водородную связь с N-H группой Шиффова основания, но только в том случае, если эта группа ориентирована в полость цитоплазматического канала. При повороте Шиффова основания, эта водородная связь рвется, и цепь молекул воды коллапсирует, в результате чего окружение Asp 96 становится более гидрофобным и его  $pK_a$  растет, что приводит к его протонированию. Важно, что такой механизм позволяет исключить возможность обратного переноса протона на Asp 96 с Шиффова основания (т.е. передачи протона по футильному циклу).

Ряд структурных изменений на этом этапе фотоцикла способствуют проникновению протона с цитоплазматической стороны мембраны внутрь белка, к Asp 96. Здесь также образуется короткая цепь молекул воды, включающая воду 504 (в структуре интермедиата М она занимала положение между Asp 96 и Thr 46, образуя с этими остатками водородные связи, а на стадии N интермедиата молекула перемещается ближе к цитоплазматической поверхности мембраны) и воду 520 (связанную водородными связями с Asp 38 и кислородом пептидной связи Leu 99, оба остатка расположены на поверхности белка).

К структурным особенностям бактериородопсина, играющим важную роль на стадии репротонирования Asp 96, стоит отнести шесть кислых аминокислотных остатка на обращенной в цитоплазму поверхности белка: Asp 36, Asp 38, Asp 102, Asp 104, Glu 161 и Asp 166. Эти группы формируют своего рода «протонную антенну» [117], увеличивая концентрацию протонов в окрестности устья белкового канала и направляя их к молекулам воды цитоплазматического полуканала и, далее, Asp 96. Важность этих аминокислот (особенно Asp 38) на стадии репротонирования Asp 96 подтверждается исследованиями мутантов (Asp38Arg [118]), у которых М интермедиат фотоцикла распадается в десять раз медленнее, по сравнению с бактериородопсином дикого типа, что, впрочем, может быть связано с невозможностью, в этом случае, образования водородной связи между Asp 38 и водой 520.

#### 2.2.8. Восстановление исходного состояния. Интермедиат О

Как уже было замечено, происходящая в ходе реакции  $N' \leftrightarrow O$  реизомеризация ретиналя связана с процессом репротонирования Asp 96. Данные ИК спектроскопии О интермедиата (наличие в спектре НООР) говорят, что вновь образовавшийся *all-trans* изомер находится в

напряженном повернутом состоянии, которое релаксирует одновременно с депротонированием Asp 85, в результате чего восстанавливается исходное состояние бактериородопсина и исходный же заряд ретиналя. Приблизительно о структуре интермедиата О можно судить на основании рентгеноструктурной модели мутанта D85S [119]. В этой структуре сдвиги  $\alpha$ -спиралей F и G на цитоплазматической стороне белка (см. раздел *Репротонирование Шиффова основания. Интермедиаты М и N*) сменяются структурными изменениями апопротеина на внеклеточной половине белка: сдвигами  $\alpha$ -спиралей A, B, D и E. Как эти изменения связаны с депротонированием Asp 85 пока не вполне ясно. Показано [120], [96] также, что в ИК спектре появляются полосы характерные для протонированного состояния карбоксильной группы C=O, что говорит о нахождении, хотя бы и временном протона, ушедшего с Asp 85, на Asp 212 [38].

### 2.2.9. Предотвращение обратной изомеризации

В работе [121], посвященной изучению энергетического профиля бактериородопсина в ходе фотоцикла и обратной термической реизомеризации методом QM/MM, освещается немаловажный аспект работы бактериородопсина, а именно, особенности строения и кинетики протонного транспорта, позволяющие избежать обратной изомеризации ретиналя, повышая тем самым эффективность протонной помпы.

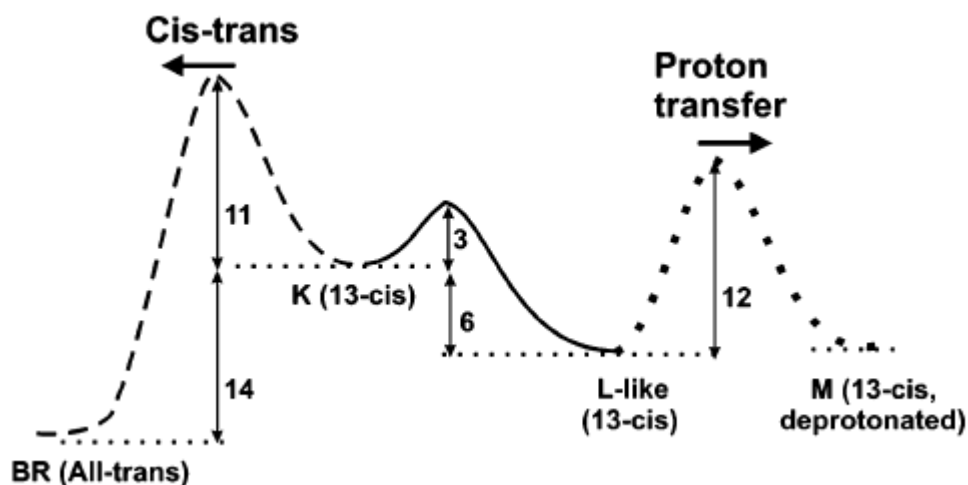


Рисунок 19. Энергетический профиль процесса фотоизомеризации бактериородопсина и начальных этапов фотоцикла, числами обозначены высоты энергетических барьеров (в ккал/моль).

Исследование показало, что, хотя наличие в фотоцикле напряженных (повернутых, англ. twisted) состояний ретиналя понижает активизационный барьер для обратной термической реизомеризации ретиналя в all-trans форму, он все равно остается на 5-6 ккал/моль выше энергетического барьера процесса протонного транспорта. Как указывается в статье, структура повернутого состояния ретиналя очень точно настроена так, чтобы запасть

максимум энергии для осуществления последующих этапов фотоцикла, не допуская при этом обратной реизомеризации ретиналя.

### 2.3. Использование в биоэлектронике

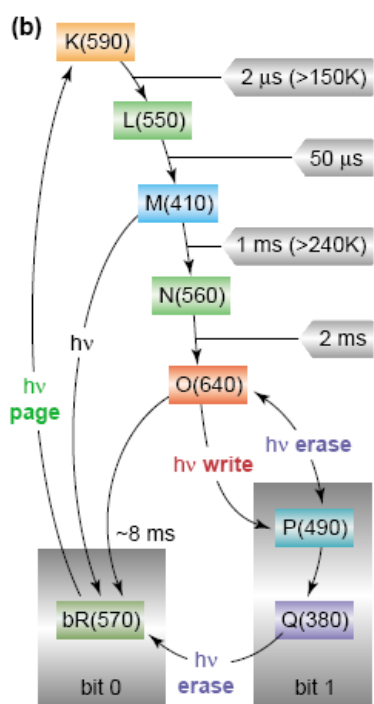
Предположения использовать бактериородопсин (в диком виде, или его мутанты) для различных технических нужд начали возникать практически сразу после его открытия. Этому, в частности, способствовали детальные исследования белка учеными самого разного профиля.

Технические приложения бактериородопсина собираются в ряд масштабных направлений [14]:

- протонный транспорт:
  - генерация АТФ в реакторах;
  - опреснение морской воды;
  - генерация электрической энергии из света;
- фотоэлектрические применения:
  - ультрабыстрая световая детекция;
  - искусственная сетчатка;
  - детекция подвижности;
- фотохромные применения:
  - хранение информации:
    - 2D-носители;
    - 3D-носители;
    - голографические носители;
  - обработка информации:
    - световые переключатели;
    - оптическая фильтрация;
    - нейросети;
    - распознавание паттернов;
    - интерферометрия;
- различные применения:
  - детекция радиации;
  - биосенсорные приложения;
  - генерация вторых гармоник.

Так как конечной целью проекта, в рамках которого выполняется настоящая работа, является разработка оптических носителей информации нового поколения, построенных на базе мутантных форм бактериородопсина, далее речь в этом разделе пойдет о результатах, достигнутых в этой конкретной области.

Для использования бактериородопсина в качестве запоминающего элемента оптических устройств хранения информации необходимо выполнения ряда условий. Во-первых, стадии (интермедиаты) фотоцикла (нативного или каким-либо образом модифицированного) бактериородопсина, используемые в качестве логических «0» и «1», должны быть максимально различимы спектроскопически, т.е. максимумы их спектров поглощения должны быть максимально раздвинуты. Во-вторых, интермедиаты должны быть термо- и фотостабильными: не должны распадаться и переходить друг в друга под действием света или высоких температур. В-третьих, процессы взаимопревращения интермедиатов должны обладать высоким квантовым выходом.



**Рисунок 20.** Модифицированный фотоцикл бактериородопсина, который предлагается использовать для записи информации в работе [3].

Ни одна из опубликованных на данный момент схем не удовлетворяет полностью всем трем условиям. В частности, самая разработанная на данный момент схема предполагает использование в качестве механизма записи/удаления информации переход между интермедиатами Q и bR мутанта D85N [3], [122]. В нативном белке интермедиаты P и Q не образуются, однако при условии протонирования Asp 85 интермедиат O может при дополнительном освещении переходить в короткоживущий интермедиат P (максимум спектра поглощения соответствует 490 нм), характеризующийся 9-cis конформацией ретиналя, который раз распадается с образованием стабильного интермедиата Q (максимум спектра поглощения – 380 нм). Последний может существовать, не изменяясь, на протяжении многих лет, пока не будет освещен светом той длины волны, который переводит его в основное состояние. Минус данной схемы заключается в низкой эффективности фотоконверсии как интермедиата O в интермедиат P (0.02%), так и обратно, P в bR (1%). Также существенным недостатком подобного подхода является необходимость дополнительного освещения бактериородопсина в основном состоянии для перевода его в стадию O фотоцикла, из которой только возможен дальнейший переход в интермедиат P.

Другое перспективное направление исследований в этой области предполагает построение 3D устройств хранения данных, использующих двухфотонные переходы ретиналя. В схемах такого рода используется два перпендикулярно направленных лазера, причем фотоизомеризация ретиналя возможна только в той точке, в которой два световых пучка пересекаются. Такая схема обладает тем достоинством, что позволяет увеличить объем хранящихся данных на порядки, однако ее реализация сталкивается со значительными затруднениями, в частности в отношении позиционирования лазерных лучей, стабильности интермедиатов, используемых для записи информации и низкого квантового выхода реакций фотоконверсии интермедиатов.

## 2.4. Компьютерная фотохимия бактериородопсина (обзор работ)

### 2.4.1. Компьютерный расчет спектров ретиналя и его производных

Компьютерное предсказание спектральных свойств ретиналя особенно важно в свете поиска и апробации методик, позволяющих корректно описывать на уровне КХ молекулярные системы такого рода, в частности для дальнейшего применения этих методов при реализации гибридных расчетов КХ/ММ спектров целых белков. Появление экспериментально полученных спектров ретиналя в вакууме [20] дало мощный толчок данному направлению исследований. Приведем краткий обзор достигнутых результатов.

На данный момент определенно существует ряд КХ методов, позволяющих с высокой степенью точности воспроизвести спектры ретиналя и его производных в вакууме. Хотя, в вычислительном плане, многие из этих методов остаются весьма затратными, в тоже время, теперь стало возможным говорить и о широкой их доступности (расчеты на одном современном компьютере методами группы CASSCF занимают время порядка нескольких суток).

Наилучший результат для  $S_1$  перехода *all-trans* формы ретиналя достигнут [123] на уровне *ab initio* методики высокого уровня CASPT2 с большим базисным набором (ANO) на базе геометрии ретиналя, полученной на уровне теории MP2, – 606 нм, что практически точно повторяет экспериментальное значение в 610 нм. Использование других методов приводит к ошибкам различной величины: в случае TD-DFT – -100/-50 нм на геометрии DFT(B3LYP)/CASSCF [124-125], в случае полуэмпирической методики OM2-MRCI – -10/-40 нм на геометрии DFT(B3LYP)/CASSCF [124], в случае SORCI – 25/-20 нм на геометрии DFT(B3LYP)/CASSCF [124], более 200 нм для расчета методом SAC-CI на геометрии DFT (B3LYP) [126], и -65 нм при расчете по методу CASPT2 на геометрии, оптимизированной методом CASSCF [127]. В работе [128] проведен расширенный анализ существующих

методов и достигнут близкий к экспериментальному результат для *all-trans* ретиналя в вакууме (604 нм) на уровне MRCISD+Q теории (геометрия оптимизирована DFT (B3LYP)). Также показано, что использование MR-DDCI2+Q также дает хорошие результаты, при учете систематической погрешности метода (возникающей из-за пренебрежения двойными возбуждениями) величиной -65 нм.

В упоминавшейся уже работе [123] получены значения максимумов поглощения, посчитанные методом CASPT2, для ретиналя в состоянии *all-trans* и *13-cis* с протонированным и нейтральным Шиффовым основанием. Полученные максимумы находятся в очень хорошем согласии с экспериментом.

#### 2.4.2. Изучение влияния белкового окружения на спектральные характеристики ретиналя

Существующий массив результатов по гибридным расчетам ретинальсодержащих белков впечатляет своим объемом и исчисляется сотнями статей и других печатных материалов. Ретинальсодержащие белки являются типичными объектами для апробации и тестирования новых методов КХ и КХ/ММ.

В разделе в очень сжатой форме перечисляются те из результатов, которые имеют преимущественно отношение к расчетам бактериородопсина, являющегося объектом настоящей работы.

Хотя изложенные различными авторами оригинальные методики расчетов в той или иной степени уникальны, справедливо утверждать, что большая их часть в основе своей сходна. Так, для реализации гибридного вычисления преимущественно используются схемы вроде ONIOM с учетом электростатического взаимодействия между внутренней и внешней подсистемами на уровне квантовой химии (electronical embedding, см. в разделе, посвященном теории КХ/ММ расчетов) и неполяризуемыми силовыми полями CHARMM и Amber, т.е. без учета поляризации ретиналом белкового окружения. В основном расчеты проводят на предварительно минимизированных на уровне КХ/ММ структурах белка; практически нет работ, посвященных учету многих микросостояний системы хромофор-белок (англ. sampling) в динамических исследованиях [129].

Особое внимание стоит уделить работам групп Thiel (Max-Planck Institut für Kohlenforschung, Германия), Elstner (Universität Paderborn, Германия) и Morokuma (Gaussian Inc., США), являющих своего рода «первопроходцами» в сфере применения гибридных методов к исследованию ретинальсодержащих белков и их спектральной «настройки».

Thiel с коллегами последовательно применяют оригинальные полуэмпирические методики OM1 и OM2 (англ. orthogonalized model) к различным ретинальсодержащим и иным фотоактивным [130-131] белкам. Группой достигнут примечательный уровень предсказательности в рамках широкораспространенной модели КХ/ММ расчетов ONIOM [124, 132-133]. Группой был проведен масштабный анализ влияния аминокислотных замен на смещение спектрального максимума ретинальсодержащих белков, путем последовательной замены всех аминокислот белка на глицин и дальнейшего расчета спектра быстрым полуэмпирическим методом OM2/MRCI [124].

Morokuma из корпорации Gaussian является автором КХ/ММ методов IMMOM [134] и ONIOM [135], последний, включенный изначально в пакет квантово-химических расчетов Gaussian 98, получил в последствии широкое распространение. Группой, работающей под его руководством, проведен ряд исследований, ставящих целью выяснение молекулярных основ спектральной «настройки» ретинальсодержащих белков [125, 128, 136-140]. В частности, в работах были предложены три основных источника спектрального сдвига [129]:

- наличие стерических взаимодействий ретиналя с аминокислотами, формирующими ближайшее окружение хромофора, т.н. гидрофобный карман;
- электростатическое взаимодействие заряженного Шиффова основания с комплексным контрионом;
- воздействие заряженных и поляризованных аминокислот на электронную плотность хромофора.



### 3. МЕТОДЫ ИССЛЕДОВАНИЯ

#### 3.1. Введение

Зарождение численных методов исследования молекулярных процессов относится к началу 70-х годов прошлого века, т.е. первые работы в этой области появились, приблизительно, в тоже время, когда был открыт бактериородопсин (1971 год). В дальнейшем, в виду постоянно наблюдавшегося повышенного интереса исследователей к этому белку, практически все новые вычислительные методики находили применение в теоретических исследованиях бактериородопсина, его структуры и механизма работы. Такая же тенденция сохраняется и в настоящее время – бактериородопсин можно назвать традиционным объектом современных гибридных QM/MM (квантовомеханических/молекулярномеханических) методов, о которых, помимо прочего, пойдет речь далее.

Теоретические (в данном контексте имеются в виду численные) методы исследования достаточно разнообразны, и различаются как собственно общими подходами к описанию изучаемой системы, так и различной степенью точности, аккуратности расчетов, а также скоростью практической реализации на вычислительных машинах.

В общем, наибольшая точность характеризует *ab initio* квантовые методы, базирующиеся на приближенном решении уравнения Шредингера и построении многоэлектронной волновой функции, которые, в тоже время, отличаются и наибольшими затратами вычислительных ресурсов. Простейший из *ab initio* методов, метод Хартри-Фока, не учитывает никаких электронных корреляций, а потому считается наименее точным. Для учета электрон-корреляционных эффектов был создан ряд расширенных методов, основанных на HF. Простейший из таких методов – MP2 (теория Мюллера-Плессета 2-ого порядка). Даже это приближение в значительной мере улучшает результаты расчетов энергий и геометрии молекул, однако во многих случаях, в частности для многих органических молекул, поправка на электронную корреляцию оказывается несущественной, и ею можно безболезненно пренебречь. Другие методы, развившиеся из HF: теория Мюллера-Плессета n-ого порядка (англ. MPn), GVB (англ. generalized valence bond), MCSCF (англ. multi-configurational self-consistent field), CI(S) (англ. configurational interaction (single)), CC (англ. coupled cluster theory). Если говорить об относительной точности *ab initio* методов, то, к примеру, Young, в своем руководстве по численным методам (Computational Chemistry, Young D.C., 2001), приводит следующую зависимость: HF << MP2 < CISD  $\approx$  MP4  $\approx$  CCSD < Full CI. При этом автор указывает на четыре потенциальных источника ошибки в *ab initio* расчетах:

приближение Борна-Оппенгеймера, использование неполных базисных наборов, неполный учет электронных корреляций и релятивистские эффекты.

Совершенно другой подход к описанию молекулярной системы реализуется в методах, основанных на функционале плотности, объединенных под общим названием DFT (англ. density functional theory). В методах группы DFT энергия системы и другие ее молекулярные свойства выводятся из электронной плотности, а не из волновой функции, как в предыдущих методах. Теоретическое обоснование для такого подхода дает теорема Хоэнберга-Кона (Hohenberg & Kohn), впервые сформулированная в 1964 году. По своей структуре DFT в каком-то смысле близка к методу HF, в частности, электронная плотность выражается в этом случае в виде линейной комбинации базисных функций, близких по своему математическому представлению к базисным орбиталам метода HF, и получивших название орбиталей Кона-Шэма (Kohn & Sham). Электронные корреляции в DFT учитываются через обменно-корреляционные термы энергии. Стоит отметить, что в вычислительном плане DFT расчеты много выгоднее соответствующих методов, основанных на HF, однако их повсеместному распространению мешает сложность подбора (осуществляемого зачастую чисто эмпирическим путем) наиболее соответствующих конкретной системе параметров DFT (базиса и функционалов), от которых в значительной степени зависит точность получаемых результатов.

Для расчета электронных спектров молекул используются соответствующие расширения HF и DFT – временные методы (англ. TD – time dependent) HF и DFT.

Для ускорения и, одновременно, упрощения квантовохимических расчетов был разработан ряд полуэмпирических методов (англ. semiempirical methods). В этих методиках поступаются рядом малосущественных (в определенных ситуациях) положений *ab initio* подхода, вводя различные эмпирические параметры, значительно упрощающие расчеты. Полуэмпирические методы широко распространены в практике, в частности для предсказания геометрии молекулярных систем и приблизительных значений их энергий, однако их использование всегда сопряжено с определенным риском, т.к. обычно такие методы разрабатываются для применения к достаточно узкому набору систем, и оправданность их использования для другого класса соединений всегда требует особого анализа. К наиболее распространенным полуэмпирическим методам относят метод Хюкеля, CNDO, MINDO, INDO, ZINDO, AM1, PM3, TNDО, SAM1 и G1 (G2, G3).

Все перечисленные выше методы рассматривают молекулярную систему как квантовую, в методах же молекулярной механики она описывается как классическая система, динамика

ядер которой подчиняется уравнениям движения Ньютона, а электроны не учитываются вовсе. Методы молекулярной механики много быстрее квантовохимических, однако, в угоду скорости страдает точность результатов и, в целом, реалистичность такого подхода, который в принципе не способен описывать процессы образования и разрыва химических связей, электронного возбуждения, изменения атомных зарядов моделируемых молекул. Хотя перечисленные эффекты часто играют первостепенную роль в функционировании молекулярно-биологических объектов (в том числе и бактериородопсина), в некоторых случаях (на практике, довольно часто) ими можно пренебречь и принять такое классическое описание системы. Как уже было сказано, динамика ММ системы рассчитывается, исходя из решения системы уравнений движения всех атомов системы, взаимодействия между которыми описываются соответствующими потенциалами, общий математический вид которых и силовые постоянные выводятся из общих теоретических соображений и квантовохимических расчетов. В зависимости от способа интегрирования уравнений движения системы выделяют метод молекулярной динамики и метод Монте-Карло.

Объединить, во всяком случае, потенциально, два подхода, квантовый и классический, к описанию молекулярных систем, призвана группа гибридных методов, QM/MM. Принципиальным моментом гибридного подхода является разделение исходной системы на две подсистемы, поведение одной из которых описывается каким-либо из квантовых методов, а другой – методом ММ. При этом за счет использования квантовохимических методик становится возможным моделировать ряд эффектов, ускользающих при классическом рассмотрении, но при этом на исследуемую систему не накладываются жесткие ограничения по количеству атомов, что позволяет моделировать на достаточно больших временах сложные молекулярно-биологические системы, вплоть до белковых комплексов, насчитывающих многие тысячи атомов. Основной проблемой, с которой сталкиваются разработчики подобных методик, является сопряжение квантовой и классической подсистем, которое реализуется в нескольких различных подходах.

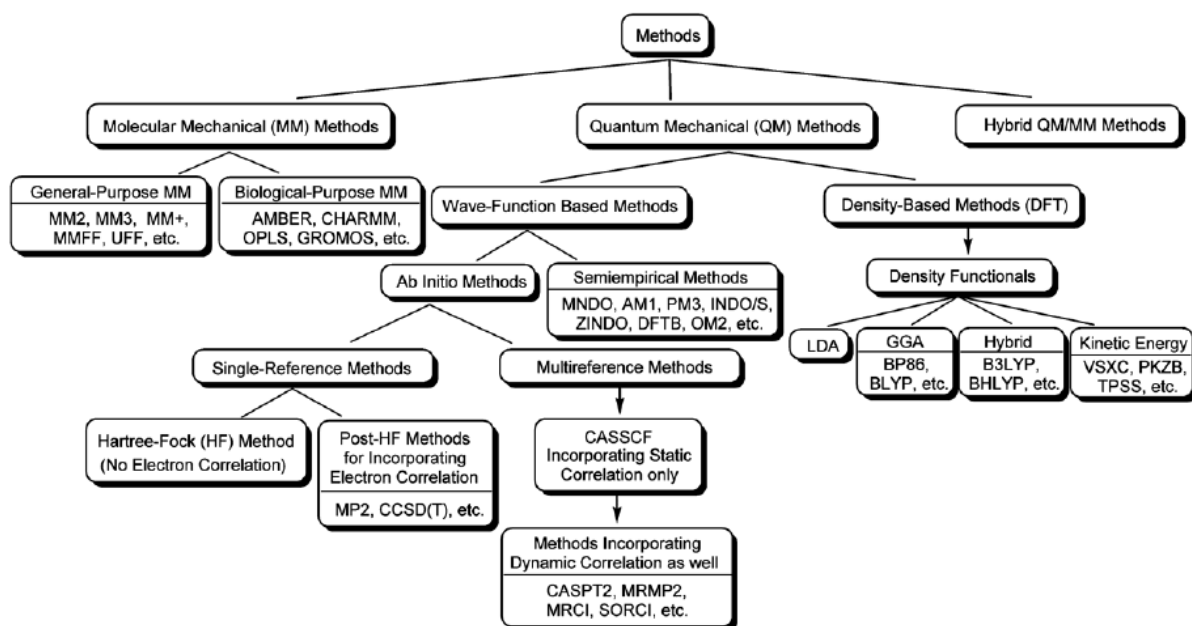


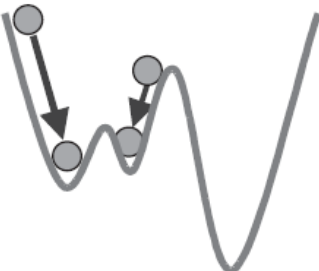


Рисунок 21. Схема, отображающая многообразие существующих на данный момент вычислительных методов (из [137]).

Другой подход к объединению преимуществ методов квантовой химии и молекулярной динамики реализуется в методе *ab initio* молекулярной динамики («из первых принципов»). Суть метода заключается в определении сил, действующих на атомы системы не из эмпирических потенциалов, а из рассчитанной из первых принципов методами квантовой химии поверхности потенциальной энергии. К примеру в наиболее разработанном методе *ab initio* МД, методе молекулярной динамики Кара-Парринелло, расчет потенциальных поверхностей производится в рамках DFT, хотя принципиально возможно использовать для этой цели и методы, базирующиеся на HF.

Таблица 5.

Сравнение основных подходов к моделированию биологических систем (по книге *Computational Chemistry, Joung, с дополнениями*).

Квантовая химия	Классическая МД	Ab initio МД
+ Высокая точность результатов	- Точность ограничена адекватностью подобранного эмпирического силового поля	+ Точность как в методах квантовой химии
+ Исходит из первых принципов, результат в	- Точность зависит от эмпирически подобранных	+ Свободная от эмпирических параметров

первом приближении прозрачен	параметров	точность, исходит из первых принципов
- Требуется больших вычислительных мощностей, реально возможен расчет систем, состоящих из десятков атомов.	+ Относительно нетребовательны к вычислительным ресурсам, возможно исследовать системы, состоящие из тысяч атомов.	+ Относительно нетребовательны к вычислительным ресурсам, возможно исследовать системы, состоящие из тысяч атомов.
+ Описывает протекание химических реакций	- Нельзя описывать протекание химических реакций	+ Описывает протекание химических реакций
+ Возможно моделирование ионов переходных металлов	- Невозможно моделирование ионов переходных металлов	+ Возможно моделирование ионов переходных металлов
- При 0 К	+ Конечные температуры	+ Конечные температуры
- В основном, моделирование в газовой фазе	+ Возможно моделирование конденсированных фаз	+ Возможно моделирование конденсированных фаз
- При минимизации исходной геометрии молекула переходит в ближайший минимум, даже если он не самый глубокий (см. рис.)	+ За счет тепловых флуктуаций молекула способна преодолевать энергетические барьеры локальных минимумов и достигать состояния, отвечающего абсолютному энергетическому минимуму системы	+ За счет тепловых флуктуаций молекула способна преодолевать энергетические барьеры локальных минимумов и достигать состояния, отвечающего абсолютному энергетическому минимуму системы
		

## 3.2. *Ab initio* методы квантовой химии

### 3.2.1. Метод Хартри-Фока

Данный квантово-химический метод является исторически одним из первых. В изначальном виде он был сформулирован Д. Хартри в 1927 году как метод самосогласованного поля, а в 1930 году В. Фок доработал метод, введя в него учет антисимметричности электронных волновых функций. Несмотря на некоторую простоту, расчеты с применением данного метода оказываются весьма точными, что обеспечивает данному подходу известную популярность и в настоящее время.

В основе метода лежит приближение центрального поля, в рамках которого движение каждого электрона рассматривается, независимо явно от других электронов, в некотором усредненном поле, создаваемом остальными электронами системы. Энергии каждого из электронов системы считаются вариационным методом и, таким образом, всегда превосходят реальные значения энергии, сходясь к т.н. пределу Хартри-Фока, зависящему от применяемого базисного набора.

Основное преимущество метода заключается в том, что он позволяет разбить многоэлектронное уравнение Шрёдингера на несколько простых одноэлектронных уравнений. Решением каждого такого уравнения служит одноэлектронная волновая функция, называемая электронной орбиталью, а энергия такой орбитали – орбитальной энергией. Электронная орбиталь отражает поведение электрона в среднем поле остальных электронов.

Одноэлектронное уравнение Шрёдингера выглядит следующим образом:

$$\begin{array}{c} \hline \\ \hline \end{array}$$

Здесь оператор представляет собой кулоновский потенциал, создаваемый в точке нахождения первого электрона, распределенным в пространстве зарядом второго электрона, причем плотность этого распределения задается квадратом модуля орбитали  $\psi_2$ :

. Этот оператор называют орбитальным кулоновским. Полный кулоновский потенциал,  $V = \sum_{j=2}^N \frac{e^2}{r_{1j}}$ , определяет кулоновское поле, которое действует на первый электрон в каждой точке пространства со стороны всех остальных электронов системы.

Второй оператор,  $\hat{P}$ , имеет более сложный физический смысл. Этот оператор отражает антисимметричность волновой функции  $\Psi$  системы к перестановкам индексов электронов, «обменам». Поэтому этот оператор называют обменным оператором (также, полный и орбитальный).

Оператор в левой части одноэлектронного уравнения Шрёдингера по аналогии с гамильтонианом называют фокианом, или оператором Фока.

Другое приближение, вводимое в методе Хартри-Фока, утверждает, что волновую функцию можно представить как комбинацию простых волновых функций, известных как точные решения уравнения Шрёдингера для случая одноэлектронных систем. Обычно в качестве таких функций выбирают т.н. орбитали гауссова типа (сокр. GTO), радиальная часть которых пропорциональна  $e^{-\alpha r}$ , значительно реже – орбитали слетеровского типа (сокр. STO), с радиальной частью, которая выглядит как  $r^n e^{-\alpha r}$ . Совокупность таких базисных функций, используемых для аппроксимации волновых функций системы в виде их линейной комбинации, называют базисным набором. Именно от полноты базисного набора зависит в наибольшей степени достоверность получаемого в методе Хартри-Фока результата.

Для того чтобы получаемая в таком приближении волновая функция обладала симметрией, гауссовы орбитали домножают на т.н. угловую функцию. В случае сферической симметрии  $s$  орбитали это константа; для  $p$  симметрии используются  $x$ ,  $y$  и  $z$  угловые термы; угловые термы  $xy$ ,  $xz$ ,  $yz$ ,  $x^2-y^2$ ,  $4z^2-2x^2-2y^2$  дают симметрию  $d$  орбитали и т.д.

Полученные в таком приближении орбитали затем комбинируют в детерминант (т.н. детерминант Слетера), что позволяет удовлетворить двум условиям, накладываемым на системы электронов квантовой механикой. Во-первых, электроны должны быть неразличимы, т.е. можно только говорить о том, занята та или иная орбиталь электроном или нет, но никак нельзя определить какой конкретно электрон расположен на конкретной орбитали. Второе ограничение связано с антисимметричностью волновой функции относительно перестановки двух частиц, т.е. при такой перестановке знак волновой функции меняется на противоположный.



Орбитали, из которых составляют детерминант, могут представлять собой отдельные ГТО, но чаще в их качестве выбирают некоторую суперпозицию базисных функций одного типа атомов или разных.

Алгоритм расчета по методу Хартри-Фока начинается с выбора орбитальных коэффициентов, обычно с помощью полуэмпирического метода. Полученная волновая функция используется для расчета энергии системы и следующего набора орбитальных коэффициентов, которые, в свою очередь, могут быть использованы для очередной итерации. Процедура продолжается до тех пор, пока не будет достигнута нулевая разница по энергии между двумя итерационными шагами. Из-за этого метод Хартри-Фока также называют методом само-согласованного поля (в ряду итераций электростатическое поле, создаваемое электронами, приходит в равновесие само с собой).

### 3.3. Метод функционала плотности (DFT)

#### 3.3.1. Основы метода

Методы функционала электронной плотности представляют собой группу методик принципиально отличных от формализма Хартри-Фока и методов, основанных на последнем – рассмотренных выше. Все свойства системы (в том числе и энергия системы) вычисляются исходя из электронной плотности, а не волновой функции, как в методе Хартри-Фока, теоретическим обоснованием чему служит опубликованная в 1964 году теорема Хоэнберга-Кона, устанавливающая однозначную связь между внешним потенциалом  $V_{ext}$ , приложенным к системе электронов, и электронной плотностью  $\rho$ .

В случае молекулы, внешний потенциал, действующий на электронную подсистему, – это электростатическое поле ядер атомов, и, таким образом, согласно теореме Хоэнберга-Кона, электронная плотность определяет гамильтониан системы, а через него и все свойства системы, производимые из гамильтониана.

По аналогии с построением исходной волновой функции в методе Хартри-Фока из элементарных взаимноортогональных волновых функций базисного набора, в методах DFT электронная плотность строится из базисных функций (т.н. орбиталей Кона-Шэма), сходных по математическому представлению с элементарными базисными волновыми функциями, но аппроксимирующих электронную плотность молекулярной системы, а не волновую функцию. Полученное таким образом начальное приближение электронной плотности затем подвергается оптимизации в соответствии с вариационным принципом, утверждающим, что

истинная энергия моделируемой системы не больше минимальной энергии модели. Существенное преимущество методов DFT – гораздо меньшие требования к вычислительным ресурсам, так как, в отличие от волновой функции, зависящей от  $3N$  координат, где  $N$  - число электронов в системе, электронная плотность зависит только от трех координат, независимо от числа электронов в системе, что значительно сокращает вычисления и позволяет исследовать методами DFT системы с большим числом атомов, вплоть до тысяч, что особенно важно при изучении крупных биомакромолекулярных комплексов.

Полная энергия системы в теории функционала плотности может быть записана в виде функционала от электронной плотности.

Из вариационного принципа Ритца следует, что энергии основного состояния при заданном внешнем потенциале соответствует минимум функционала энергии.

Первый член в правой части выражения для функционала энергии (1) определяет взаимодействие электронной плотности с внешним потенциалом, а второй член,  $E_{\text{Hxc}}$ , определяется следующим образом:

$$E_{\text{Hxc}} = E_{\text{H}} + E_{\text{xc}}$$

и включает в себя электронный кинетический терм  $T_s$  невзаимодействующих электронов, формирующих электронную плотность  $n(\mathbf{r})$ , и терм, соответствующий электрон-электронным взаимодействиям. Последний состоит из классической энергии электрон-электронных взаимодействий, т.н. энергии Хартри, и электрон-корреляционной составляющей, обозначенной через функционал  $E_{\text{xc}}$ . Нахождение минимума энергии, соответствующей этому функционалу, энергии электронных корреляций, составляет основную вычислительную задачу метода DFT, а сам вид электрон-корреляционного функционала - принципиальную проблему, от успешности решения которой в первую очередь зависит качество полученных результатов. В настоящее время разработано немало количество различных функционалов, заточенных под какие-то специфические задачи, или же универсальных.

Также, Кон и Шэм предложили удобный прием, существенно облегчающий применение методов функционала плотности в квантовохимических расчетах – концепцию т.н.

фиктивных невзаимодействующих электронов, формирующих электронную плотность, идентичную электронной плотности реальных, взаимодействующих электронов. Фиктивные электроны описываются ортонормальными одночастичными волновыми функциями (они, как правило, аппроксимируются согласно методу ЛКАО (линейной комбинации атомных орбиталей, англ. LCAO) как линейная комбинация атомных орбиталей, формируемых из базисных функций гауссова или слетеровского типа, хотя часто с этой целью используются и базисы плоских волн, особенно когда моделируются периодические или обширные системы), и электронная плотность фиктивных невзаимодействующих электронов вычисляется как:

Фактор величиной в два возникает здесь в силу вырожденности функций по спину, а суммирование ведется по всем занятым одночастичным состояниям. В обозначениях уравнение для функционала энергии записывается следующим образом:

$$- \frac{1}{2} \sum_i \int d\mathbf{r} \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r})$$

В этом выражении третий терм в правой части уравнения представляет собой кинетическую энергию невзаимодействующих электронов, в то время как все электронные корреляции оказываются включенными в электронно-корреляционный терм  $E_{\text{corr}}$ . И минимум функционала энергии вычисляется уже не для электронной плотности, а для набора одноэлектронных волновых функций, решением системы уравнений вида, называемых уравнениями Кона-Шэма, согласно вариационному принципу, чем существенно облегчается компьютерная задача, формулируемая в рамках метода:

$$\begin{aligned} & \sum_i \int d\mathbf{r} \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) + \int d\mathbf{r} \psi^*(\mathbf{r}) \nabla^2 \psi(\mathbf{r}) \\ & + \int d\mathbf{r} \psi^*(\mathbf{r}) \nabla^2 \psi(\mathbf{r}) = 0 \end{aligned}$$

Уравнения Кона-Шэма имеют вид, сходный с уравнениями Хартри метода Хартри-Фока, но, в отличие от них, они носят не приближенный характер, а, в принципе, точный. После решения системы уравнений вида (10), электронная плотность вычисляется по уравнению (8), а соответствующая ей энергия основного состояния – по уравнению (6).

Хотя, как уже говорилось выше, метод функционала плотности в принципиальном отношении позволяет получить точное значение энергии системы в основном состоянии, для этого необходимо знать точное представление обменно-корреляционного функционала, что, увы, невозможно. Ниже мы рассмотрим основные его аппроксимации, имеющие применение в квантовой химии.

Простейшее приближение обменно-корреляционного функционала, предложенное в оригинальной работе Кона и Шэма в 1965 году – т.н. приближение локальной плотности (англ. LDA, local density approximation):

здесь обменно-корреляционная энергия в расчете на электрон, параметризованная с высокой точностью в расчетах однородного электронного газа с плотностью квантовым методом Монте-Карло.

Лучшую аппроксимацию обменно-корреляционного функционала, по сравнению с LDA, дает приближение обобщенного градиентного разложения (англ. GGA, generalized gradient approximation):

в котором значение функционала зависит не только от локального значения плотности, но и от ее градиента. Разумный выбор подынтегральной функции позволяет уменьшить ошибку расчета в сравнении с LDA на значительную величину, вплоть до порядка. К функционалам группы GGA, в частности, относится широко распространенный функционал BLYP.

Также необходимо упомянуть обменно-корреляционный функционал B3LYP, включающий определенный вклад точной обменной энергии Хартри-Фока (из-за чего подобные функционалы называют гибридными), и гибридный функционал B2PLYP, в котором обменная энергия аппроксимируется на уровне MP2 теории. Гибридные функционалы удачно применяются при расчетах многих органических молекулярных систем, в том числе – биомакромолекул. Кроме того, в сравнении с функционалами GGA, функционалы этой группы успешно описывают слабые взаимодействия, в том числе количественно.

На современном этапе развития вычислительных приложений теории функционала плотности стало возможным использование результатов, полученных на этом уровне без дополнительного сравнения с результатами расчетов *ab initio* методами.

### 3.3.2. Временное расширение метода DFT

Для изучения возбужденных состояний квантовой системы (энергий возбужденных состояний, заселенностей) в 1984 году Рунге и Гроссом (Runge & Gross) было разработано расширение метода DFT, т.н. временной метод DFT. Теоретическое основание методу дает теорема Рунге-Гросса, являющаяся временным аналогом теоремы Хоэнберга-Кона, и утверждающая, что под действием меняющегося во времени внешнего потенциала система электронов эволюционирует таким образом, что в каждый конкретный момент времени существует однозначное соответствие между внешним потенциалом и электронной плотностью.

В теории Кона-Шэма для поиска энергии квантовой системы используется вариационный принцип Ритца, что невозможно в случае временного варианта теории функционала плотности, где используется динамический аналог принципа Ритца – принцип стационарного действия.

Решение временного уравнения Шредингера в терминах теоретической механики является поиском точки стационарности интеграла действия :

—

А, с учетом взаимосвязи между волновой функцией системы и электронной плотностью, действие можно рассматривать как функционал электронной плотности, стационарное состояние которого соответствует точной электронной плотности системы:

—————

Решение уравнения (), являющегося уравнением Эйлера-Лагранжа, проводится в рамках формализма, близкого формализму стационарной теории DFT, однако точное решение возникающей при этом системы временных уравнений Кона-Шэма (англ. TDKS – time-dependent Kohn-Sham) требует значительных вычислительных ресурсов, поэтому, в настоящее время в квантово-химических приложениях широкое распространение получило решение TDKS в рамках теории возмущения, позволяющей вычислить значения ряда параметров системы, таких как поляризованность и энергии возбуждения, исходя из

Подробное рассмотрение имплементации теории возмущения во временную теорию DFT выходит за пределы обзора литературы, приводимого в настоящей работе и может быть получено например здесь [141].

### 3.3.3. Эволюция метода DFT: SCC-DFTB, DFTBA

Метод SCC-DFTB (self-consistent-charge density functional tight-binding) представляет собой приближенный метод квантовой химии [142-143], получивший широкое распространение в последние годы. Этот подход базируется на пренебрежении либо параметризации интегралов взаимодействия, возникающих при решении KS уравнений стандартного метода DFT, а также т.н. приближении сильной связи, и формально близок к расширенному методу Хюккеля и методу CNDO. Однако, в отличие от обычных полуэмпирических методик, основным источником для параметризации метода служат расчеты в рамках DFT, а не эмпирические данные. Тем самым избегается проблема привязки метода к конкретной группе соединений или химических систем, что, наряду с применением неортогональных базисных наборов, позволяет добиться универсальности. Метод позволяет достичь замечательного ускорения квантово-химического расчета вплоть до 10-100 раз.

В математическом смысле, SCC-DFTB выводится путем разложения функционала энергии метода DFT по флуктуациям электронной плотности в окрестности исходной электронной плотности.

Детальному рассмотрению методики посвящен обзор [144].

Иная реализация идеи метода SCC-DFTB – DFTBA, была разработана в Gaussian Inc. [145]. При этом, в отличие от SCC-DFTB, для аппроксимации интегралов взаимодействия используются аналитические выражения, а не набор табличных параметров.

## 3.5. Метод молекулярной динамики

### 3.5.1. Общие замечания

Метод молекулярной динамики был разработан Б. Алдером в 1950-х годах прошлого века. В основе метода лежит классическое рассмотрение моделируемого объекта как системы взаимодействующих материальных точек, движение которых описывается классическими уравнениями Ньютона:

$$\frac{d^2 \mathbf{r}_i}{dt^2} = -\frac{1}{m_i} \nabla_i V(\{\mathbf{r}_j\})$$

Здесь  $i$  – номер атома,  $N$  – количество атомов в системе,  $m_i$  – масса атома,  $\mathbf{r}_i$  – радиус-вектор атома,  $\mathbf{F}_i$  – равнодействующая сил, действующих на атом,  $E$  – потенциальная энергия системы.

Если взять достаточно маленький интервал времени  $\hbar$ , такой, что изменением сил, действующих на частицы системы, можно пренебречь, то можно записать следующие равенства для координат и скоростей частиц в момент времени :

—

Отсюда видно, что если мы можем вычислять действующие на систему силы, то мы можем предсказать динамику исследуемой системы в будущем.

### 3.5.2. Силовые поля и их структура

Равнодействующая сил , действующих на -ый атом, находится здесь как градиент потенциала, который представляется в виде суммы ряда термов, отвечающих различным межатомным взаимодействиям. В таблице представлены основные типы потенциальных термов, используемые в большинстве современных программных пакетов, предназначенных для расчета молекулярной динамики.

Здесь - потенциал всех валентных взаимодействий, - потенциал невалентных взаимодействий, - потенциал валентной связи, - потенциал валентного угла, - потенциал торсионного угла, – потенциал Ван дер Ваальсовых взаимодействий, – потенциал электростатического взаимодействия.

Таблица 6.

*Математическое представление потенциалов, применяемых для моделирования молекулярной динамики.*



Название	Тип взаимодействия	Функциональный вид
Гармонический	Валентная связь	
Гармонический	Валентный угол	
Морзе	Валентная связь	
Кубический	Валентная связь	
Urey-Bradley	Валентный угол	
Полином	Валентный угол	
Гармонический	Ложный торсионный угол	
Периодический	Торсионный угол	
Ryckaert-Bellemans	Торсионный угол	
Фурье-ряд	Торсионный угол	
Ленард-Джонс 6-12	Ван дер Ваальс	— —
Ленард-Джонс 10-12	Ван дер Ваальс	— —
Кулон	Электростатическое	—

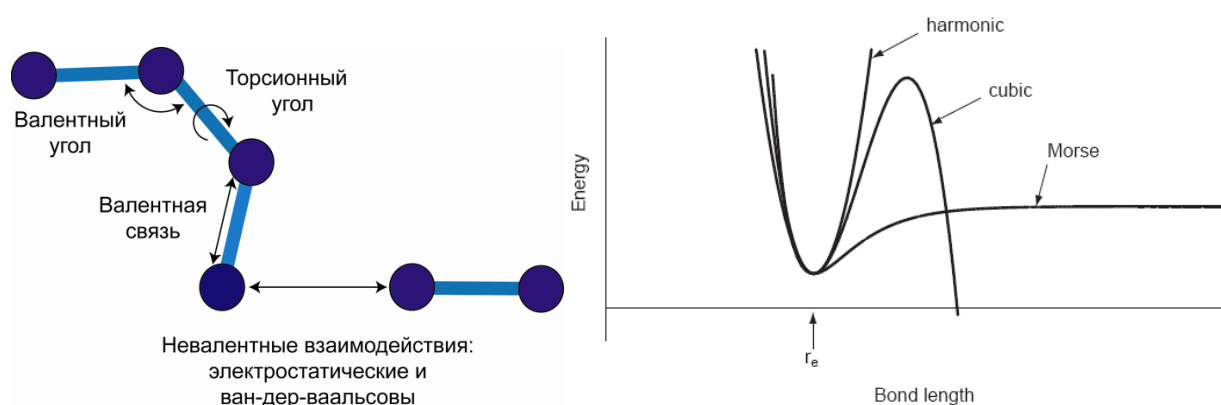


Рисунок 22. Слева - схема, отображающая типы межатомных взаимодействий, справа - типы функций, используемые для моделирования химической связи ( $r_e$  – равновесная длина связи).

В молекулярной динамике используются различные варианты стандартизации термов межатомных взаимодействий, а также наборы силовых постоянных, входящих в функции

потенциалов этих взаимодействий. Совокупность математических представлений потенциалов межатомных взаимодействий и силовых констант называют силовым полем. В практике молекулярной динамики используют различные силовые поля, часто заточенные под применение с конкретной группой химических соединений. В частности, для моделирования молекулярных биологических систем, главным образом, ДНК и белков, используют силовые поля OPLS, Amber, GROMOS, GMX (GROMACS), CHARMM. Силовые поля типов MM2, MM3, UFF, MMFF имеют более общую направленность и применимы к широкому спектру органических и неорганических соединений.

### 3.5.3. Температура. Термостаты

Системы с большим числом частиц удобно характеризовать таким макропараметром как температура. Температура связана со средней (по ансамблю состояний) кинетической энергией системы через соотношение:

$$\overline{E_{kin}} = \frac{1}{2} k_B T$$

где  $k_B$  - константа Больцмана.

В молекулярной динамике усреднение кинетической энергии осуществляется не по ансамблю, а по траектории и указанное равенство выполняется лишь приблизительно, однако с увеличением длины траектории точность совпадения средних по ансамблю и по траектории растет. Исходя из этого, можно записать формулу для молекулярно-динамического аналога температуры :

$$T = \frac{2}{3N} \overline{E_{kin}}$$

где  $\overline{E_{kin}}$  - средняя по траектории кинетическая энергия, а  $N$  – полное число степеней свободы системы. Помимо этого вводится понятие мгновенной температуры системы, равной:

$$T_{inst} = \frac{2}{3N} E_{kin}$$

Таким образом, задав в начальный момент времени координаты и импульсы частиц системы, мы задаем полную энергию системы, а температура вычисляется через кинетическую энергию системы, усредненную по траектории. Это весьма неудобно, т.к. сложно определить какие начальные параметры надо задать системе, чтобы в ней поддерживалась необходимая нам температура. Для решения этой проблемы был предложен ряд методов

термостатирования молекулярной системы, т.е. поддержания ее температуры на определенном заданном значении.

Широко распространенным и весьма простым вариантом термостата является термостат Берендсена [146]. В уравнения движения добавляется термостатирующий член:

$$\ddot{r}_i = -\frac{1}{m_i} \nabla_i U(r) - \frac{\gamma}{m_i} (\dot{r}_i - \sqrt{\langle \dot{r}_i^2 \rangle})$$

здесь коэффициент  $\gamma$  имеет смысл величины, обратной времени релаксации флуктуаций температуры системы к ее равновесному значению  $\langle T \rangle$ . Этот термостат обладает рядом недостатков. Он не отвечает каноническому распределению Гиббса и нарушает закон равнораспределения кинетической энергии по степеням свободы. В вычислительных экспериментах с небольшими молекулярными системами наблюдалось перекачивание энергии из высокочастотных в низкочастотные моды колебаний.

Другой вариант термостата – термостат Нозе-Гувера [147]. В этом случае в уравнение движения также дописывается дополнительный член, масштабным образом меняющий скорости частиц системы при отклонении температуры от равновесного значения. Однако в этом случае коэффициент этого масштабирующего члена вычисляется при решении дополнительного дифференциального уравнения.

$$\ddot{r}_i = -\frac{1}{m_i} \nabla_i U(r) - \frac{\gamma}{m_i} (\dot{r}_i - \sqrt{\langle \dot{r}_i^2 \rangle})$$

где  $\gamma$  – параметр, обычно выбираемый близким к 1, а  $f$  – число степеней свободы.

### 3.5.4. Численное интегрирование уравнений движения

Уравнения Ньютона, описывающие динамику моделируемой молекулярной системы,

$$m_i \ddot{r}_i = -\nabla_i U(r)$$

интегрируются численно, с применением различных методов, позволяющих экономить вычислительные ресурсы, таких как алгоритм «перескоков» (англ. «leap-frog»).

Пусть  $X(t)$ ,  $V(t)$  и  $a(t)=F(t)/m$  – координата, скорость и ускорение, для любой из компонент решения уравнений движения. Разложим функцию  $X(t)$  в ряд Тейлора в окрестности произвольной точки  $t_0$ , учитывая, что  $t-t_0=h$ , получаем:

$$X(t) = X(t_0) + V(t_0)h + \frac{1}{2}a(t_0)h^2 + \dots$$

Складывая первое и второе уравнения, получим выражение для координаты частицы в следующий момент времени через ее координаты и ускорение в предыдущие моменты времени:

а, вычитая из второго первое, получим выражение для скорости частицы через координаты:

Простейший метод численного интегрирования уравнений движения – алгоритм Верле. В начальный момент времени,  $t=0$ , полагаются известными координаты и скорости всех частиц системы: . На следующем этапе вычисляются новые координаты системы:

. Затем вычисления следуют по циклу:

- 1) по координатам находят силы, а затем ускорения: — —;
- 2) вычисляются новые координаты системы с следующий момент времени:

Скорости напрямую в этом алгоритме не вычисляются, если возникает такая необходимость, они могут быть посчитаны из известных координат.

Алгоритм с «перескоками», или leap-frog, является модификацией предыдущего подхода. На регулярном шаге вычисляются как скорости, так и координаты атомов:

Следует правда обратить внимание, что скорости и координаты находятся со сдвигом на полшага. По координатам в момент времени  $t$  находятся силы и соответствующие им ускорения —. Затем находятся скорости в момент времени  $t+0.5h$ . После чего – координаты в момент времени  $t+h$ . И снова вся процедура повторяется.

### 3.5.5. Методы учета электростатических взаимодействий

В самом простом случае, при расчете невалентных взаимодействий потенциальная энергия рассчитывается не для каждой пары атомов системы, а для атомов, находящихся на расстоянии меньше определенного значения – радиуса обреза. Обычно при учете

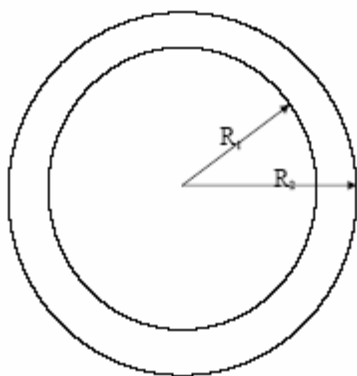


Рисунок 23. Радиус обреза. См. текст.

невалентных взаимодействий с радиусом обреза используется т.н. метод Верле, который позволяет значительно экономить вычислительные ресурсы за счет составления списков пар невалентно взаимодействующих атомов. Эти списки обновляются не на каждом шаге, а лишь когда в системе происходят достаточные изменения в положении атомов друг относительно друга. Трудоемкость расчета невалентных

взаимодействий по этому методу пропорциональна  $(C1 \cdot S \cdot N + (C2/K) \cdot N^2)$  (по сравнению с  $N^2$  при учете всех электростатических взаимодействий в системе  $N$  частиц), где  $C1$  и  $C2$  некоторые небольшие константы,  $S$  – число частиц, в среднем попадающих в сферу радиуса  $R_2$ , а  $K = \Delta R / (2v_{\max} \cdot h)$ ,  $\Delta R = R_2 - R_1$ . Обычно коэффициент при  $N^2$  достаточно мал, но при больших  $N$  этот член в выражении для трудоемкости становится доминирующим, снижая тем самым эффективность методики.

### 3.6. *Ab initio* молекулярная динамика

#### 3.6.1. Молекулярная динамика Борна-Оппенгеймера

Молекулярную динамику, классическому представлению которой был посвящен предыдущий раздел, можно в самом фундаментальном и общем смысле рассматривать как движение системы в соответствии с законами ньютоновской механики по поверхности потенциальной энергии, определяемой, в случае классической МД, топологией системы и набором эмпирических параметров, называемых силовыми постоянными или силовым полем. Впрочем, как уже указывалось неоднократно выше, такой подход оказывается весьма ограниченным, а в некоторых случаях совершенно неудовлетворительным для описания динамики сложных систем. Это наводит на мысль о необходимости выработки метода, позволяющего с высокой степенью реалистичности и достоверности описывать динамику молекулярных систем, за счет более точного расчета поверхности потенциальной энергии, что оказывается возможным на уровне квантовой химии.

Математическая формулировка *ab initio* МД весьма проста и сводится, по сути, к паре формул:

---

Здесь  $H_{el}$  – полный электронный гамильтониан системы, параметрически зависящий от координат ядер  $\mathbf{R}$ . Таким образом, в случае такой динамики атомы движутся по классическим законам, однако силы, действующие на них, вычисляются «из первых принципов».

Т.к. этот подход базируется на приближении Борна-Оппенгеймера, заключающемся в разделении электронных и ядерных степеней свободы, он называется МД Борна-Оппенгеймера (англ. BOMD).

### 3.6.2. Молекулярная динамика Кара-Парринелло

Схема BOMD требует КХ-расчета моделируемой системы на каждом шагу МД, поэтому в вычислительном плане BOMD – очень требовательный метод. С целью упрощения расчета без существенного ущерба для реалистичности модели, в 1985 году Кар и Парринелло предложили свой оригинальный метод *ab initio* МД, базирующийся на DFT [148].

CPMD формулируется в рамках альтернативного формализма классической механики, формализма Лагранжа. Центральной величиной, определяющей динамику системы в этом случае, является лагранжиан  $L$ , с учетом разбиения системы на ядерную и электронную подсистемы:

---

В выражении для лагранжиана первый терм соответствует классической кинетической энергии ядер; второй – квазиклассической кинетической энергии электронных орбиталей, рассматриваемых в CPMD в классическом приближении, параметр  $m_e$  – фиктивная масса электрона; третий терм – энергия взаимодействия электронной и ядерной подсистем, в формализме DFT этот член соответствует просто энергии Кона-Шэма (см. обсуждение теории

DFT в разделе 4.3.1); четвертый терм вводится для сохранения ортонормированности орбиталей, теряющейся при введении в лагранжиан второго термина.

Уравнения движения для ядер и для электронов выводятся следующим стандартным образом:

$$\begin{aligned} & \frac{d^2 \mathbf{R}}{dt^2} = -\frac{1}{M} \nabla_{\mathbf{R}} E \\ & \frac{d^2 \mathbf{r}}{dt^2} = -\frac{1}{m} \nabla_{\mathbf{r}} E \end{aligned}$$

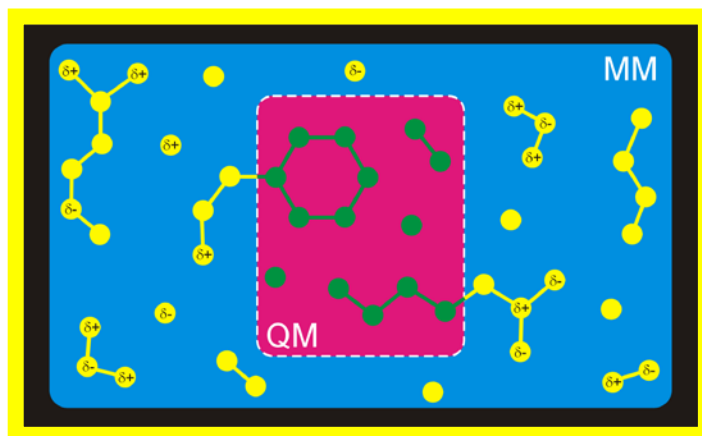
Для интегрирования уравнений движения в подходе CPMD используется алгоритм Верле, рассмотренный в разделе 4.5.4.

### 3.7. Гибридные методы КМ/ММ

Квантово-механическое описание поведения молекулярных систем отличается высокой точностью полученных "из первых принципов" результатов, в которых учитываются существенные квантовые эффекты – процессы образования и разрыва химических связей, электронного возбуждения, изменения атомных зарядов моделируемых молекул. Однако такие вычисления требуют значительных вычислительных ресурсов, нелинейно растущих (за исключением, быть может, ряда современных полуэмпирических методик) при увеличении размеров моделируемой системы, из-за чего в настоящее время в зависимости от сложности метода, доступны для изучения системы, содержащие от нескольких десятков атомов (full CI, CASSCF) до сотен атомов (DFT). Однако, даже небольшие по размеру белки часто состоят из нескольких тысяч атомов (в частности, в бактериородопсине порядка 3000 атомов).

В свою очередь, широко распространенные для моделирования биологических макромолекул, методы молекулярной механики позволяют осуществлять расчеты огромных надмолекулярных комплексов, вплоть до вирусных капсидов, содержащих миллионы атомов. Впрочем, в угоду скорости расчетов страдает точность и реалистичность таких расчетов, анализируя результаты которых всегда нужно помнить о принципиальных ограничениях классического рассмотрения.





**Рисунок 24.** Схема, демонстрирующая разделение моделируемой системы на две подсистемы: внутреннюю (квантовую) и внешнюю (классическую).

Объединить, во всяком случае, потенциально, два подхода, квантовый и классический, к описанию молекулярных систем, призвана группа гибридных методов, КХ/ММ (QM/MM), первоначально предложенный в [149]. Принципиальным моментом гибридного подхода является разделение исходной системы на две подсистемы, поведение одной из которых (назовем ее внутренней) описывается каким-либо из квантовых методов, а другой (внешней) – методом ММ. При этом за счет использования квантовохимических методик становится возможным моделировать ряд эффектов, ускользающих при классическом рассмотрении, а на исследуемую систему не накладываются жесткие ограничения по количеству атомов, что позволяет моделировать на достаточно больших временах сложные молекулярно-биологические системы, вплоть до белковых комплексов, насчитывающих многие тысячи атомов.

Рассмотрим подробно проблемы, с которыми приходится сталкиваться при реализации гибридного подхода. Основной трудностью, с которой сталкиваются разработчики подобных методик, является сопряжение квантовой и классической подсистем, которое по-разному реализуется в различных альтернативных схемах.

Общая энергия системы может быть записана двумя принципиально разными путями, в зависимости от того для каких подсистем и систем производится расчет методами квантовой химии и молекулярной механики. Эти схемы носят названия *субтрактивной* и *аддитивной*.

В случае субтрактивной схемы, на каждом шаге расчета гибридным методом, производится три вспомогательных расчета:

- расчет всей системы на уровне ММ, дающий энергию ;
- расчет внутренней подсистемы на уровне КХ, дающий энергию ;
- расчет внутренней подсистемы на уровне ММ, дающий энергию .

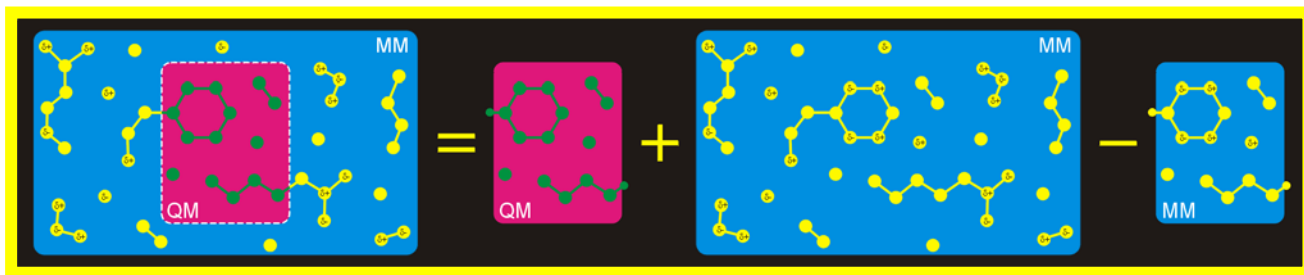


Рисунок 25. Схема, иллюстрирующая субтрактивную схему КХ/ММ (в том числе схему ONIOM).

Общая энергия системы для гибридного расчета по этой схеме находится через «разностное» (откуда и название схемы) выражение:

Т.к. для внутренней подсистемы производится расчет в том числе и с применением метода ММ, этот подход требует наличия силовых параметров для описания взаимодействий между атомами внутренней подсистемы, а также атомных парциальных зарядов для атомов, входящих в эту подсистему.

В другой схеме, именуемой аддитивной, минуеться расчет внутренней подсистемы на уровне ММ, за счет чего этот подход кажется в какой-то степени более удобным и привлекательным: параметры силового поля для внутренней подсистемы в этом случае не нужны. ММ расчет (энергетический терм) производится только для внешней подсистемы, при этом к выражению для ММ энергии подсистемы добавляется терм, описывающий Ван-дер-Ваальсовы взаимодействия внутренней и внешней подсистем. Взаимодействие же между подсистемами описывается за счет введения в выражение для итоговой энергии специального члена, на том или ином уровне теории описывающим электростатическое взаимодействие подсистем:

В обеих схемах, учет электростатического взаимодействия подсистем может быть реализован одним из двух способов: учет на уровне молекулярной механики (англ. mechanical embedding) и учет на уровне квантовой химии (англ. electronic embedding). В первом случае, рассчитывается энергия кулоновского взаимодействия точечных парциальных зарядов двух подсистем, причем эти заряды не меняются в течение расчета, т.е. в данном методе оказывается невозможным учитывать поляризацию внутренней подсистемы внешней подсистемой. В случае же electronic embedding, электростатическое взаимодействие учитывается за счет включения в квантовый гамильтониан внутренней подсистемы члена, отражающего взаимодействие матрицы точечных парциальных зарядов

внешней подсистемы с электронной плотностью внутренней подсистемы. Этот метод учитывает поляризацию и особо важен при расчетах свойств (в том числе спектральных характеристик хромофоров), зависящих от процессов поляризации и переноса заряда.

Еще один немаловажный аспект гибридных расчетов – реализация интерфейса между внутренней и внешней подсистемами при наличии между ними ковалентных связей. Существует два общепринятых подхода, направленных на решение этой задачи. В первом, появляющиеся при разрыве связи (связей) валентности насыщаются виртуальными *линкерными атомами*. В таком качестве обычно используют атомы водорода или галогены. Линкерными атомы не несут заряда при расчетах ММ. Второй подход заключается в специальном описании образующихся валентных орбиталей и называется методом «замороженных» орбиталей (англ. frozen orbitals method). Подробно этот подход, не использующийся в настоящей работе, описан в [150].

В работе гибридные расчеты преимущественно проводились в программном пакете Gaussian (Gaussian Inc., [151]) в рамках модели ONIOM, предложенной Morokuma с коллегами в [134-135, 152]. Этот подход, в рамках изложенной выше терминологии, является субтрактивной схемой, с возможностью использовать обе схемы учета электростатического взаимодействия (ЕЕ, *electronical embedding* и ММ, *mechanical embedding*). Разрыв ковалентных связей на границе расчетных подсистем компенсируется добавлением линкерных атомов водорода.

Для эффективной минимизации энергии системы с целью получения оптимальной геометрии изучаемых структур в схеме ONIOM применяется методика микро-итераций, в ходе которой после каждого шага минимизации внутренней подсистемы, производится оптимизация всей системы на уровне ММ с «замороженной» внутренней подсистемой. Это позволяет заметно быстрее приводить энергию гибридной системы к минимуму.

Кроме того, для расчетов спектров в программе Orca (<http://www.thch.uni-bonn.de/tc/orca/>) использовалась более примитивная реализация гибридного метода, представляющая собой *electronical embedding* расчет внутренней подсистемы в матрице точечных атомных зарядов внешней подсистемы. Расчет молекулярной механикой в этом случае не производится (проводится точечный расчет, то есть без учета динамического поведения белка).

## 4. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

### 4.1. Подготовка моделей бактериородопсина и его мутантов

Для выполнения поставленных в работе целей были созданы модели бактериородопсина и его мутантов атомного разрешения. Структура бактериородопсина дикого типа в основном состоянии с ретиналом в *all-trans* состоянии была взята из базы данных PDB (код структуры 1QNH, получена методом рентгеноструктурного анализа с разрешением 1,9 Å [34]). Структура бактериородопсина с ретиналом в форме 13-*cis*, 15-*syn* также была получена из базы данных PDB (код структуры 1X0S, получена методом рентгеноструктурного анализа с разрешением 2,5 Å [51]). Предпочтение было отдано этим структурам как самым полным (1QNH включает аминокислотные остатки с 5 по 232; структура 1X0S – аминокислотные остатки с 5 по 231) и полученным с самым высоким разрешением.

Также был произведен отбор тестового набора мутантных форм бактериородопсина. Выбирались мутанты, для которых известен из литературных данных экспериментально измеренный максимум оптического поглощения. Так как структурные исследования с определенностью указывают на наличие в экспериментальных образцах, как белка дикого типа, так и мутантов различных конформаций ретиналя (*all-trans*; 13-*cis*, 15-*syn*; 13-*cis*, 15-*anti* и иногда иных конформеров), были выбраны мутанты, для которых экспериментально установлено процентное соотношение конформаций ретиналя, и в которых одна из конформаций (а именно *all-trans*) преобладает. Таким образом было выбрано 9 мутантов, данные по которым, а также по структурам дикого типа, представлены в таблице ниже и на рисунке.

Таблица 7.

*Перечень структур, использовавшихся в работе.*

<b>Структура</b>	<b>Экспериментальный максимум поглощения, нм</b>	<b>Сдвиг максимума поглощения, нм</b>	<b>Литературный источник</b>
Дикий тип	568	0	[28]
13- <i>cis</i> , 15- <i>syn</i>	556	-12	[32]
Arg 82 Gln	598	30	[32]

Asp 85 Asn	617	49	[122]
Thr 89 Val	455	-113	[131]
Asp 96 Asn	569	1	[131]
Ser 141 Cys	475	-93	[75]
Trp 182 Phe	489	-79	[75]
Pro 186 Leu	479	-89	[39]
Asp 212 Ala	548	-20	[39]
Asp 212 Asn	579	11	[39]

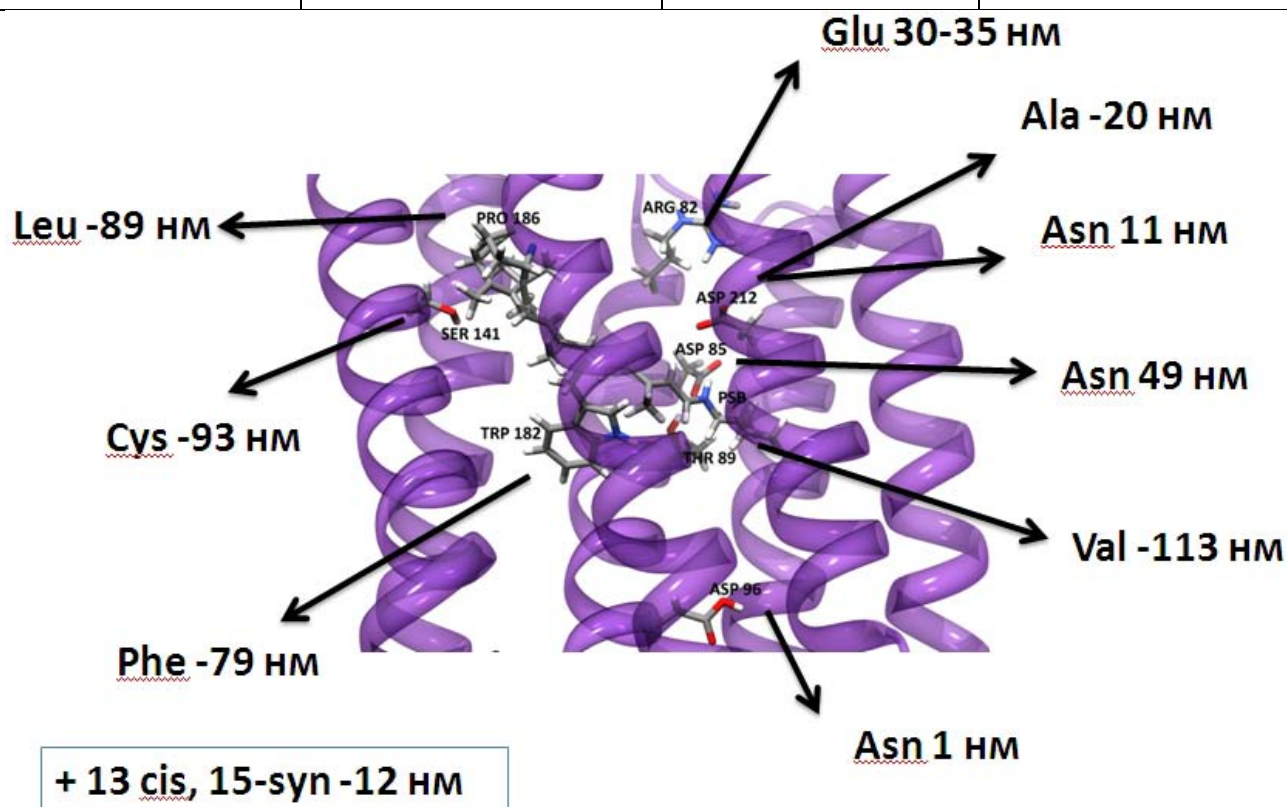


Рисунок 26. Схема, изображающая набор отобранных структур: 9 мутантов (для каждого указан спектральный сдвиг, с минусом, если сдвиг в синюю область, без знака – если в красную) и форма бактериородопсина с конформером ретиналя 13-cis, 15-syn. PSB – Шиффово основание ретиналя, аминокислоты обозначены трехбуквенным кодом.

Из кристаллические структуры, взятых из базы данных PDB, были удалены молекулы липидов и молекулы воды, расположенные вне внутрибелковой полости.

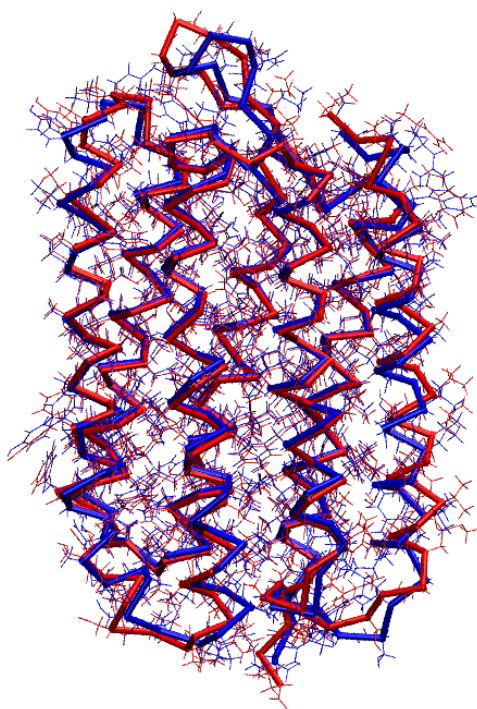
Состояния протонирования всех аминокислотных остатков были выбраны в соответствии с результатами эмпирического расчета программой PROPKA [153-154], за исключением Asp 96, Asp 115, Glu 204, состояния протонирования которых были выбраны протонированными,

что соответствует их состоянию в нативном белке и приводится во многих исследованиях бактериородопсина молекулярной динамикой (к примеру, в [155], [90, 113, 156], [54]).

Все модели были дополнены атомами водорода (таким образом использовалось полноатомное представление) в пакете Maestro (Schrodinger, LLC), в том же программном пакете были получены структуры мутантов путем внедрения необходимых мутаций в структуру бактериородопсина дикого типа 1QNJ с разрешением возникающих стерических конфликтов.

Так как для моделирования бактериородопсина и его мутантов методами КМ/ММ было использовано поле Amber 99 [157], то все структуры предварительно были минимизированы на уровне ММ. Сравнение полученных после минимизации структур, с исходными показало, что существенных изменений в структуре при минимизации не происходит, о чем говорит значение среднеквадратичного смещения координат атомов структур и визуальный анализ (см. рис. и табл. ниже).

От релаксации структур мутантов было решено отказаться вслед за [158], чтобы не допустить двусмысленности в трактовке результатов расчетов (в том смысле, что при проведении оптимизации сложно судить об истинной причине спектрального сдвига на атомном уровне).



**Рисунок 27.** Выровненные друг относительно друга и наложенные структуры бактериородопсина дикого типа до и после минимизации на уровне ММ в поле Amber. Синим обозначена структура до минимизации, красным – после.

Таблица 8.

*Значения среднеквадратичного смещения координат атомов структур после минимизации моделей на уровне ММ.*

<i>Структура</i>	<i>СКС, Å</i>	<i>Структура</i>	<i>СКС, Å</i>
Дикий тип	1,13 Å	Ser 141 Cys	1,14 Å
13- <i>cis</i> , 15- <i>syn</i>	2,20 Å	Trp 182 Phe	1,13 Å
Arg 82 Gln	1,14 Å	Pro 186 Leu	1,14 Å
Asp 85 Asn	1,13 Å	Asp 212 Ala	1,13 Å
Thr 89 Val	1,12 Å	Asp 212 Asn	1,13 Å
Asp 96 Asn	1,13 Å		

Отсутствующие силовые постоянные для имеющихся в остатке ретиналя валентных взаимодействий были взяты из поля ММ2. Эти новые параметры представлены в таблицах ниже. Заряды для ретиналя были рассчитаны по методу RESP на уровне DFT (B3LYP), аналогичным образом получены заряды в использовавшейся реализации силового поля Amber 99. Заряды приведены в Приложении.

Таблица 9.

*Параметры валентных углов для остатка ретиналя, адаптированные из параметров для сходных взаимодействий в поле ММ2.*

<i>Типы атомов</i>	<i>Тип функции</i>	<i><math>r_{eq}</math>, Å</i>	<i><math>K</math>, ккал/(моль·Å<sup>2</sup>)</i>
N2 CM	гармонический	1.365	448
CM HC	гармонический	1.08	367
CM H1	гармонический	1.08	367
CA HC	гармонический	1.08	367



Таблица 10.

*Параметры валентных углов для остатка ретиналя, адаптированные из параметров для сходных взаимодействий в поле ММ2.*

<b>Типы атомов</b>	<b>Тип функции</b>	<b><math>\Theta_0</math>, °</b>	<b><math>K_\theta</math>, ккал/моль/рад<sup>2</sup></b>
N2 CA CA	гармонический	122	20
N2 CA HA	гармонический	118	30
CT CM CA	гармонический	120	25
CM CA CA	гармонический	132	25
CM CA HA	гармонический	114	20
CT CM CT	гармонический	110	25
CT CT CM	гармонический	114	20
N2 CM CM	гармонический	121.2	70
N2 CM H1	гармонический	119.1	35
CM CM CM	гармонический	117	63
CM N2 H	гармонический	121.2	30
CT N2 CM	гармонический	121.2	70
CM CM HC	гармонический	119.7	35
CM CM H1	гармонический	119.7	35

#### 4.2. Расчет максимумов спектров поглощения для белка дикого типа и мутантов

С целью выработки оптимальной методики расчета спектров, были апробированы различные квантово-химические методы, относящиеся к различным группам: *ab initio* методы, DFT, полуэмпирические методы.

Минимизация структур проводилась по схеме ONIOM в программном пакете Gaussian 09 (Gaussian Inc.). В качестве MM метода было выбрано поле Amber, в качестве квантовых методов использовались следующие: полуэмпирический метод PM6, DFT с функционалом B3LYP, SCC-DFTB (набор параметров mio, оптимизированный для расчетов биомакромолекул [159]), MP2.

В качестве т.н. внутреннего региона (квантовой подсистемы, т.е. той части системы, которая рассчитывается высокоточным квантовым методом) был выбран только ретиналь вплоть до атома C<sub>ε</sub>. Asp 85 не был включен в квантовую область вслед за рядом работ [126, 137], в которых было показано, что включение этого аминокислотного остатка в область квантового рассмотрения не приводит к существенному улучшению предсказания. Разрыв между двумя подсистемами пришелся на одинарную связь C<sub>ε</sub>-C<sub>δ</sub>, возникшие валентности были насыщены линкерными атомами водорода. Внутренняя «квантовая» подсистема таким образом состояла из 53 атомов.

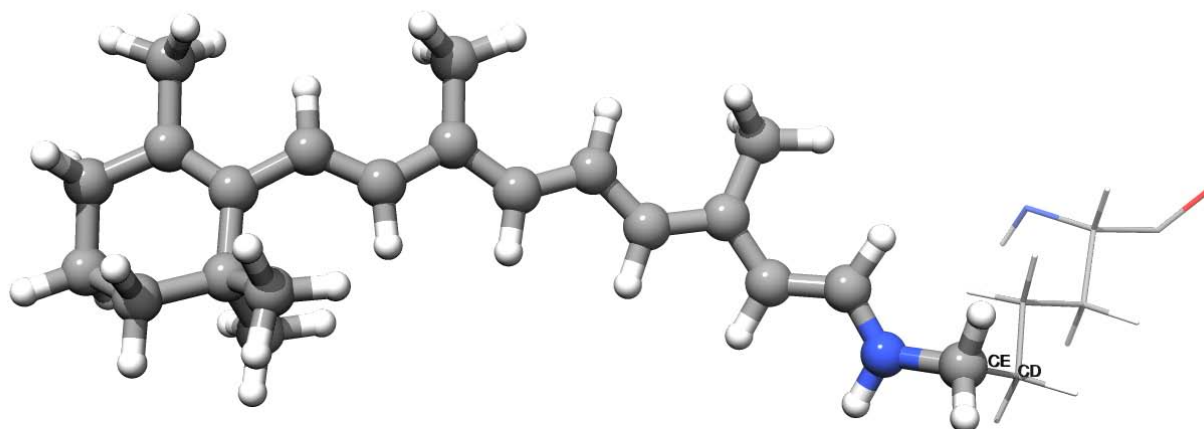


Рисунок 28. Схема деления системы на квантовую и классическую подсистемы. Квантовая подсистема визуализирована шариками и цилиндрами, а классическая – тонкими линиями.

Для расчета спектров использовались следующие КХ методы: TD-DFT (функционал B3LYP), полуэмпирический ZINDO/S, оптимизированный для расчетов спектров органических

хромофоров, высокоточные и требовательные к ресурсам мультikonфигурационные методы SORCI и MRDDCI2+Q (программный пакет Orca).

Результаты апробации перечисленных методов на бактериородопсине дикого типа приведены в таблице.

Таблица 11.

*Результаты апробации различных методов минимизации структуры бактериородопсина дикого типа по схеме ONIOM и расчета максимума спектра поглощения. Величины максимумов поглощения приведены в нм. Во всех вычислениях использовался базисный набор группы Поупла 6-31G\*\* (см. дополнительное сравнение результатов расчета спектра с применением разных базисных наборов). (Экспериментальное значение максимума поглощения bR дикого типа – 568 нм).*

	Метод минимизации структуры				
Метод расчета спектра		PM6	DFT(B3LYP)	DFTB	MP2
	TD-DFT (B3LYP)	514	542	542	545
	ZINDO/S	516	557	558	556
	SORCI	518	540		541
	MRDDCI2+Q	453	475		

Для определения оптимального для расчетов базисного набора была проведена серия расчетов спектров по методу TD-DFT (B3LYP) на геометрии, оптимизированной DFT, с разными базисными наборами. Результаты приведены в таблице ниже.

Таблица 12.

*Результаты расчетов спектра с применением разных базисных наборов, проведенных по схеме ONIOM TD-DFT (B3LYP) на геометрии, оптимизированной DFT. (Экспериментальное значение максимума поглощения bR дикого типа – 568 нм).*

	3-21G*	6-31G**	TZVP	ANO
--	--------	---------	------	-----

Число базисных функций	132	354	579	753
$\lambda_{\text{max}}$ для белка дикого типа (exp. 558-560 нм)	513	542	546	546

Из таблицы видно, что существенного улучшения результата при переходе от базиса 6-31G\*\* к более крупным базисам TZVP и ANO не происходит, поэтому, в целях экономии вычислительных ресурсов, во всех расчетах в данной работе использовался базис группы Поупла 6-31G\*\*.

Результаты апробации КХ методик показали, что оптимальной комбинацией методик оптимизации/расчета спектра по соотношению расчетного времени и точности результата оказались метод DFT для расчета оптимальной геометрии и методы TD-DFT и ZINDO/S для расчетов спектров. Было показано, что результаты расчетов оптимальных геометрий по полуэмпирическому методу SCC-DFTB практически совпадают с геометриями, оптимизированными полным методом DFT.

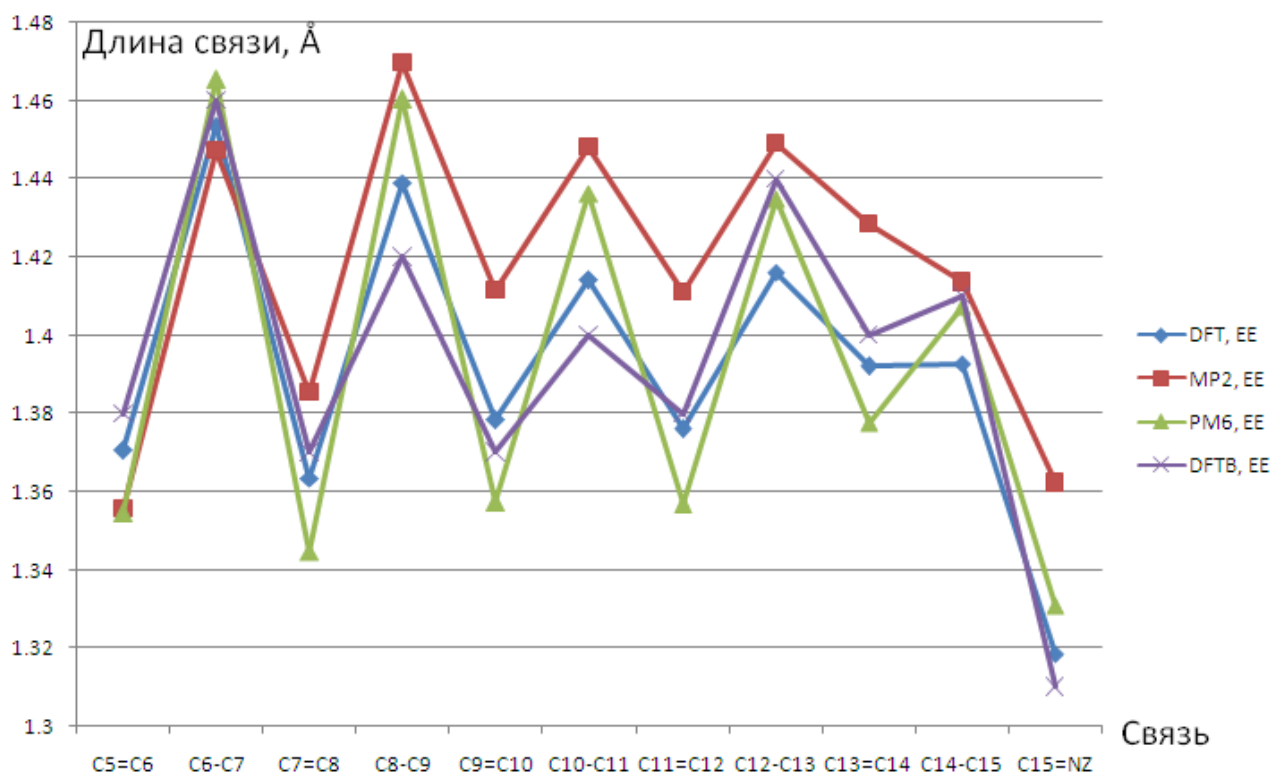


Рисунок 29. Диаграмма отражающая изменение длин двойных и одинарных связей вдоль полиеновой цепи. EE – КХ/ММ с *electronic embedding*.

Как видно на диаграмме, демонстрирующей изменение длин связей полиеновой цепочки, геометрия которой была рассчитана четырьмя разными КХ методами, наилучшее описание сопряженной  $\pi$ -системы предоставляет метод MP2 – разница между длинами двойных и

одинарных связей в геометрии, оптимизированной на этом уровне, минимальна ( $BLA = 0.053181 \text{ \AA}$ ), ( $BLA$ , англ. bond length alteration, «изменение длин связей», равен разнице между длиной средней двойной и средней одинарной связей и является одним из основных параметров геометрии сопряженных полиенов), а в случае полуэмпирического метода PM6 – параметр  $BLA$  максимален ( $BLA = 0.087193 \text{ \AA}$ ).

Использование для оптимизации геометрии метода MP2 кажется нецелесообразным, т.к. при значительном увеличении времени расчета при этом не наблюдается ощутимого улучшения результата.

Максимумы поглощения, вычисленные в рамках мультikonфигурационного подхода по методам SORCI и MRDDCI2+Q не оказались ближе к экспериментальным, чем вычисленные методами теории функционала электронной плотности. Метод SORCI дал величину максимума близкую к рассчитанной TD-DFT. Максимум поглощения, рассчитанный по методу MRDDCI2+Q, оказался сильно смещенным в красную область, однако, если учесть систематическую погрешность метода в аналогичных расчетах [128] в са. 65 нм, он дает величину максимума аналогичную величине максимума поглощения, полученной методом SORCI.

На следующем этапе был проведен расчет спектров для ряда мутантов и бактериородопсина дикого типа с ретиналем в конформации 13-*cis*, 15-*syn* методами ZINDO/S и TD-DFT (B3LYP), предварительно минимизированных методом DFT (B3LYP, по схеме ONIOM-EE). Для определения эффекта на спектральный сдвиг, оказываемого поляризацией электронной плотности, для расчета использовалось две схемы учета электростатики: *electronical embedding* и *mechanical embedding*, первая из которых подразумевает поляризацию электронной плотности матрицей зарядов белка, а в случае *mechanical embedding* такое влияние не учитывается.

Результаты приведены на рис. 30.

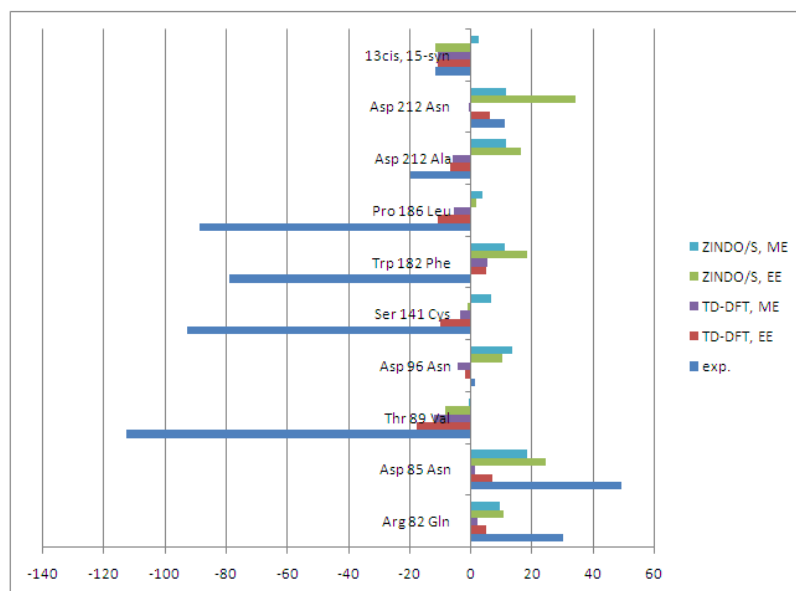


Рисунок 30. Диаграмма, отражающая вычисленные спектральные сдвиги. EE – KX/MM с *electronical embedding*; ME – KX/MM с *mechanical embedding*.

Во всех случаях, кроме одного (мутант Trp 182 Phe) метод DFT предсказал корректно знак спектрального сдвига, в случае полуэмпирической методики ZINDO/S результаты предсказания оказались хуже. Для бактериородопсина с 13-*cis*, 15-*syn* ретином величина сдвига была предсказана количественно, что может быть объяснено с б реалистичностью модели этой формы бактериородопсина – за основу структуры была взята структура из базы данных кристаллографических структур, в отличие от методики подготовки структур мутантов (см. выше). Средние абсолютные погрешности расчетов приведены в таблице ниже, из которой видно, что наименьшая погрешность характеризует метод TD-DFT с учетом электростатического взаимодействия подсистем KX и MM.

Таблица 13.

*Средние абсолютные погрешности (в нм) предсказанных по ряду методов спектральных сдвигов мутантов бактериородопсина.*

<i>TD-DFT, EE</i>	<i>TD-DFT, ME</i>	<i>ZINDO/S, EE</i>	<i>ZINDO/S, ME</i>
42	47	50	53

В рамках анализа результатов изучавшиеся формы бактериородопсина можно разделить на три группы. К первой относится белок дикого типа с изомерной формой ретиналя 13-*cis*, 15-*syn*; ко второй – мутанты с заменами заряженных аминокислотных остатков; к третьей – замены аминокислот с гидрофобным радикалом. Для единственного представителя первой

группы структур был количественно предсказан сдвиг спектра в синюю область. Сдвиг может быть связан с напряжением полиеновой цепочки изомера ретиналя, обуславливающим рост BLA (за счет нарушения планарности сопряженной  $\pi$ -системы ретиналя) и, как следствие, сдвиг спектра в синюю область (последний факт теоретически обоснован в [78] и находит подтверждение и в настоящей работе, см. графики зависимости рассчитанного спектрального максимума от BLA в Приложении 2).

Похожим образом можно объяснить сдвиги в синюю область, наблюдающиеся для третьей группы мутантов (с гидрофобными радикалами). Стерические конфликты с боковыми группами этих аминокислот приводят к росту напряженности структуры ретиналя и сдвигу в синюю область по описанному выше механизму. Значительное занижение величины сдвига в этом случае, а также ошибка в определении знака сдвига для одного из мутантов, вероятно, связаны с неудовлетворительным описанием структурных перестроек, включая перераспределение водородных связей и изменения в положении связанных молекул воды и заряженных аминокислотных остатков поблизости от ретиналя, в этом случае.

В целом, для второй группы мутантов, расчеты спектральных максимумов дали относительно лучшие результаты, в двух случаях сдвиги были предсказаны на количественном уровне.

Отсутствие спектрального сдвига в случае мутанта Asp96Asn обуславливается выбранным состоянием протонирования этого аминокислотного остатка в белке дикого типа (такой выбор оправдан экспериментальными данными, см. в разделе выше). В протонированном состоянии Asp 96 не несет отрицательного заряда и близок по своему электростатическому влиянию и характеру образующихся водородных связей к Asn 96.

Разные по знаку спектральные сдвиги в случае двух мутантов по Asp 212: Asn и Ala, объясняется характером образующихся водородных связей (см. рисунок 31).

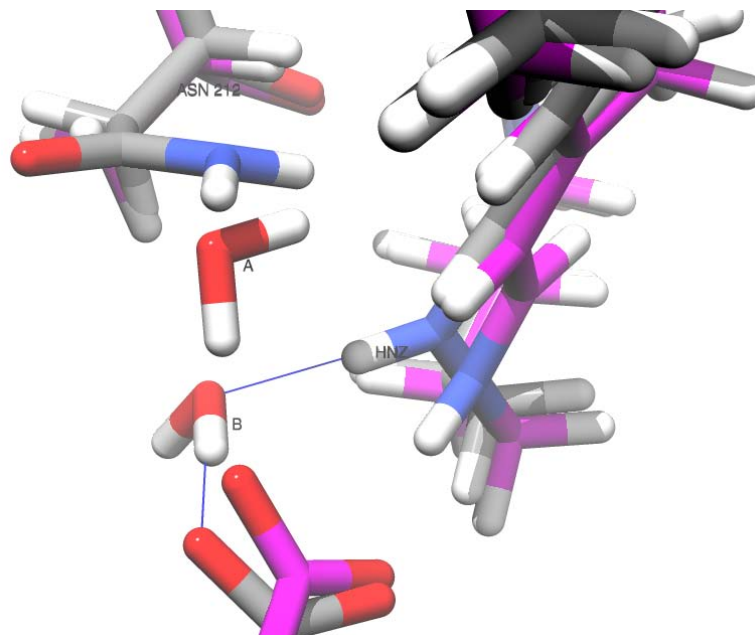


Рисунок 31. Сравнение структур мутантов Asp212Asn углероды раскрашены серым, вода 402 помечена литерой B) и Asp212Ala (малинового цвета, вода 402 – обозначена литерой A). Пояснения в тексте.

К сдвигу в синюю область спектрального максимума мутанта Asp212Asn приводит напряжение и изгиб полиеновой структуры ретиналя за счет электростатического взаимодействия с Asn 212, стабилизирующееся образованием водородной связи с молекулой воды 402. Это электростатическое влияние отсутствует при замене Asp212 на Ala, положительный заряд сопряженной  $\pi$ -системы уменьшается и спектральный максимум сдвигается в красную область.

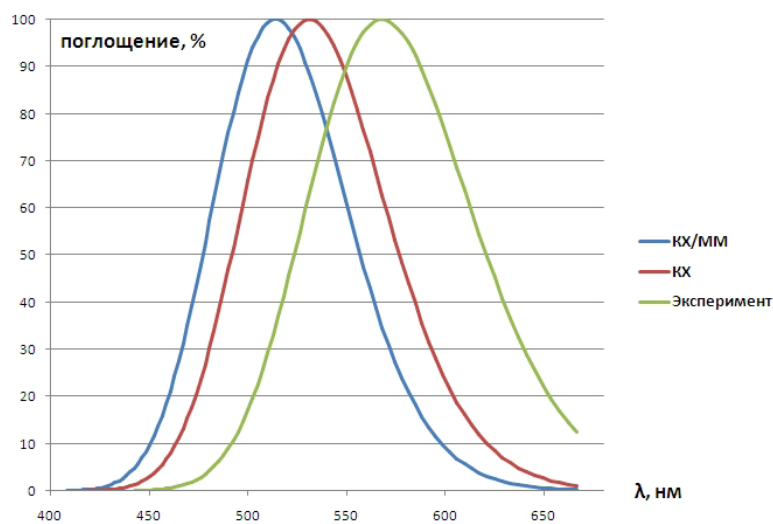
Указанный механизм электростатического модулирования положения спектральных максимумов, за счет электростатического влияния электроотрицательных группировок, вызывающих увеличение положительного заряда  $\pi$ -системы, особенно важен при внедрении отрицательно заряженных аминокислотных радикалов поблизости от азота Шиффова основания и метильных групп полиеновой цепи (атомы  $N_Z$ ,  $C_5$ ,  $C_9$ ,  $C_{13}$ ) [137]. В настоящем исследовании этот механизм превалирует при формировании спектрального сдвига мутантов заряженных аминокислот (Asp 85, Asp 212, Arg 82).



#### 4.3. Методика учета взаимной поляризации ретиналя и белкового окружения

Широко распространенные в настоящее время методы КХ/ММ обладают существенным недостатком. Подавляющее большинство реализаций КХ/ММ расчетов включает алгоритмы учета электростатического воздействия окружения на квантовую подсистему, но не наоборот. В случае бактериородопсина, хромофор которого представляет собой в случае белка дикого типа сильно положительно заряженную сопряженную  $\pi$ -систему, кажется разумным и необходимым для реалистичного описания такой системы учитывать не только поляризацию ретиналя белковым окружением, но и поляризацию белка – ретиналем. Нами была осуществленная попытка учета такого воздействия в вычислительном эксперименте по современной полуэмпирической методике MOZYME (англ. LO-PM6, метод PM6 с локализованными орбиталями), позволяющей осуществлять КХ расчет целого белка за разумное вычислительное время. Результаты расчета были сравнены с результатами расчета по аналогичной методике PM6, но осуществленной по схеме классического КХ/ММ, без учета взаимной поляризации.

Предсказанный в таком чисто квантовом эксперименте максимум поглощения составил 531 нм, что на 17 нм ближе к экспериментальному значению, чем величина максимума поглощения, предсказанная в расчете КХ/ММ без учета взаимной поляризации (514 нм) (см. рис. ниже).



**Рисунок 32.** Относительное расположение максимумов поглощения, рассчитанных методами КХ/ММ и КХ, и экспериментального спектра поглощения бактериородопсина.

#### 4.4. Учет микросостояний бактериородопсина при расчете максимума поглощения

Другим немаловажным аспектом компьютерных расчетов спектральных характеристик, остающийся за пределами приближения стандартных методов КХ/ММ, является учет микросостояний белка (называемый в англоязычной научной литературе «sampling»). В работе предпринята попытка учесть этот фактор, предположительно значительно влияющий на реалистичность описания фотоактивных белков.

В качестве исходной структуры была взята предварительная оптимизированная на уровне КХ/ММ (DFT, B3LYP, electronical embedding) модель бактериородопсина дикого типа. Используя эту структуру как стартовую, была рассчитана молекулярная динамика CPMD по гибридной схеме, когда динамика хромофора рассчитывалась по методу CPMD, а белок – в приближении ММ поля Amber. Расчет проводился в программных пакетах Gromacs 4 [160] и CPMD<sup>1</sup> 3.12 (<http://www.cpmc.org/>).

Протокол гибридной молекулярной динамики приводится ниже.

Таблица 14.

*Протокол гибридной CPMD/Amber молекулярной динамики бактериородопсина дикого типа.*

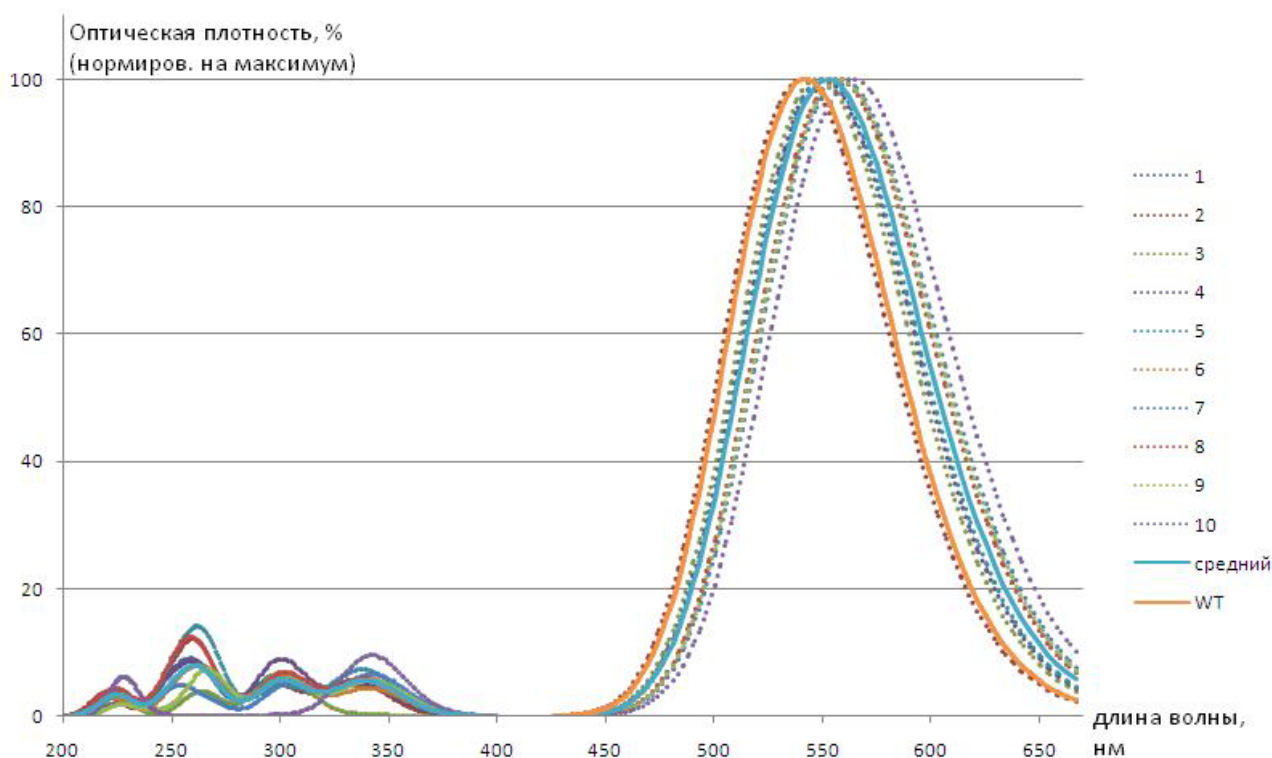
Тип интегратора, термостат	стохастическая динамика
Шаг интегрирования	1 фс
Длина траектории	500 фс
Тип учета электростатики, радиус обрезания	PME, 12 Å
Тип учета взаимодействий Ван дер Ваальса, радиус обрезания	cut-off, 12 Å
Температура	300 К
Баростат	Парриннело-Рамана
Функционал DFT	LDA
Набор псевдопотенциалов	VDB LDA
Cut-off	70.0 Å
Начальные скорости взяты из распределения Максвелла при 300 К	

<sup>1</sup> Copyright IBM Corp 1990-2008, Copyright MPI für Festkörperforschung Stuttgart 1997-2001.

Из траектории были вырезаны кадры через каждые 50 фс (всего 10 кадров), и для каждого кадра был рассчитан спектр поглощения на уровне KX/MM (TD-DFT, B3LYP, EE), с предшествующей оптимизацией структур на уровне KX/MM (SCC-DFTB, EE).

Полученные в результате спектры изображены на рисунке, где спектры приближены гауссовыми кривыми на основании полученных в расчете максимумов поглощения, а высоты (соответствующие оптической плотности) нормированы на максимальное значение оптической плотности.

Средняя величина максимума по ансамблю микросостояний оказалась равной 553 нм, показав тем самым улучшение в 11 нм в сравнении с расчетом методом DFT на минимизированной структуре.



**Рисунок 33.** График, изображающий рассчитанные для кадров KX/MM МД спектры, средний по набору микросостояний спектр и спектр, рассчитанный для минимизированной на уровне DFT структуры.

## 5. ЗАКЛЮЧЕНИЕ

Результаты, достигнутые в работе, можно обобщить в следующих выводах:

1. Отобраны экспериментальные пространственные структуры бактериородопсина, на основе которых созданы и оптимизированы методами КХ/ММ атомные модели 10 мутантных и двух диких форм бактериородопсина.
2. Исследованы различные методы оптимизации структур бактериородопсина и расчета спектральных максимумов. Наилучшие результаты показаны методами DFT (функционал B3LYP) и SCC-DFTB при расчете оптимальных геометрий, методами TD-DFT (B3LYP) и ZINDO/S при расчете спектральных максимумов.
3. Рассчитаны спектральные максимумы для набора мутантов бактериородопсина. Предложены гипотезы молекулярных механизмов спектральной настройки бактериородопсина. Полученные результаты проанализированы с точки зрения предложенных гипотез.
4. Предложена методика учета взаимной поляризации ретиналя и его белкового окружения. Показано, что учет взаимной поляризации существенно улучшает результат расчета: при прочих равных абсолютная погрешность уменьшается с -54 нм до -37 нм (на 17 нм).
5. Предложена методика оценки спектров по набору микросостояний белка. Расчет спектра для ансамбля микросостояний бактериородопсина позволяет достигнуть улучшения предсказания метода на 11 нм (абсолютная ошибка уменьшается с -26 нм до -15 нм).
6. Показана линейная зависимость максимума поглощения от параметра BLA, что согласуется и с уже имеющимися в литературе результатами.

## ПРИЛОЖЕНИЕ 1

Атомные заряды для остатка лизина, с ковалентно присоединенным к нему через Шиффово основание ретиналом.

Атом	Заряды по Малликену, DFT (B3LYP)	RESP заряды, DFT (B3LYP)	Заряды по Малликену	RESP заряды
CG	-	-	-	-0.1115
HG (2)	-	-	-	0.0279
CD	-	-	-	0.0178
HD (2)	-	-	-	0.0256
CE	-0.316	-0.29	-0.102	0.0407
HE (2)	0.349 0.349	0.16 0.16	0.146	0.0665
NZ	-0.677	0.19	-0.260	-0.0601
HZ	0.444	0.34	0.346	0.4037
C15	0.240	0.20	0.054	-0.0881
H15	0.281	0.16	0.205	0.2646
C14	-0.371	-0.48	-0.487	-0.178
H14	0.232	0.18	0.213	0.1401
C13	0.190	0.41	0.305	0.0043
C20	-0.561	-0.42	-0.358	-0.0004
H20 (3)	0.226 0.180 0.228	0.15	0.136	0.044
C12	-0.287	-0.37	-0.413	-0.1661
H12	0.211	0.17	0.227	0.1869
C11	-0.101	0.12	0.032	-0.0439
H11	0.224	0.10	0.157	0.1394
C10	-0.271	-0.34	-0.345	-0.1435
H10	0.209	0.15	0.204	0.1252
C9	0.090	0.33	0.173	-0.1179
C19	-0.531	-0.38	-0.263	-0.0006
H19 (3)	0.207 0.202 0.170	0.12	0.116	0.0625

C8	-0.213	-0.30	-0.300	-0.0951
H8	0.212	0.17	0.205	0.0823
C7	-0.170	0.04	-0.076	-0.0665
H7	0.205	0.10	0.172	-0.0058
C6	-0.029	-0.30	-0.243	-0.0337
C5	0.031	0.27	0.081	0.026
C18	-0.511	-0.44	-0.329	0.0046
H18 (3)	0.174 0.170 0.177	0.12	0.128	0.0138
C4	-0.334	-0.27	-0.108	0.0031
H4 (2)	0.174 0.176	0.10	0.092	0.0206
C3	-0.323	0.03	-0.089	0.0167
H3 (2)	0.174 0.165	0.02	0.066	-0.0089
C2	-0.308	-0.22	-0.192	0.0033
H2 (2)	0.168 0.163	0.06	0.057	0.0089
C1	-0.043	0.48	0.604	0.0094
C16	-0.482	-0.43	-0.454	-0.0261
H16 (3)	0.147 0.165 0.174	0.11	0.101	0.002
C17	-0.467	-0.45	-0.427	-0.0139
H17 (3)	0.165 0.167 0.157	0.11	0.101	-0.0005

## ПРИЛОЖЕНИЕ 2

Графики зависимостей величин рассчитанных спектральных максимумов от среднего BLA. Пояснения в разделе 4.2.

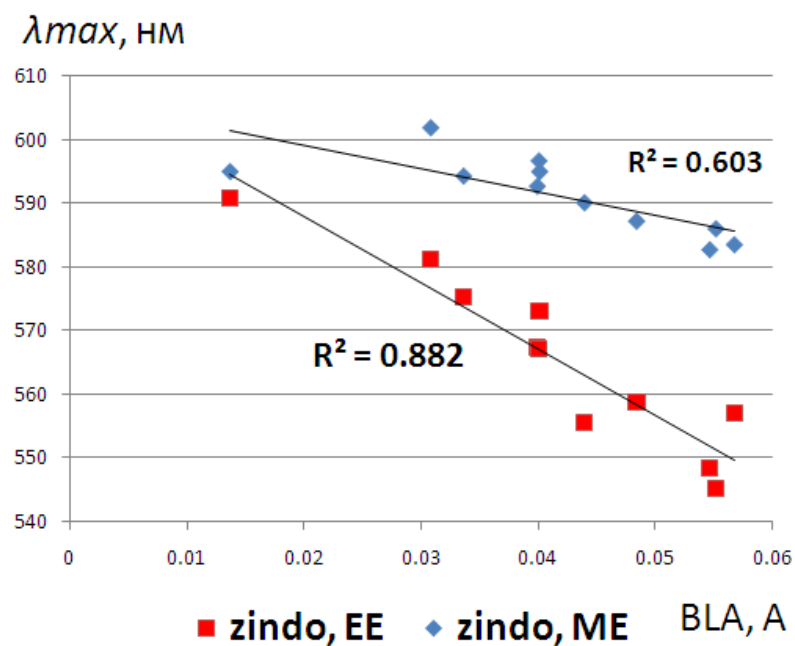


Рисунок 34. Зависимость величин рассчитанных методом ZINDO/S (по схемам ONIOM расчета EE и ME) спектральных максимумов от среднего BLA.

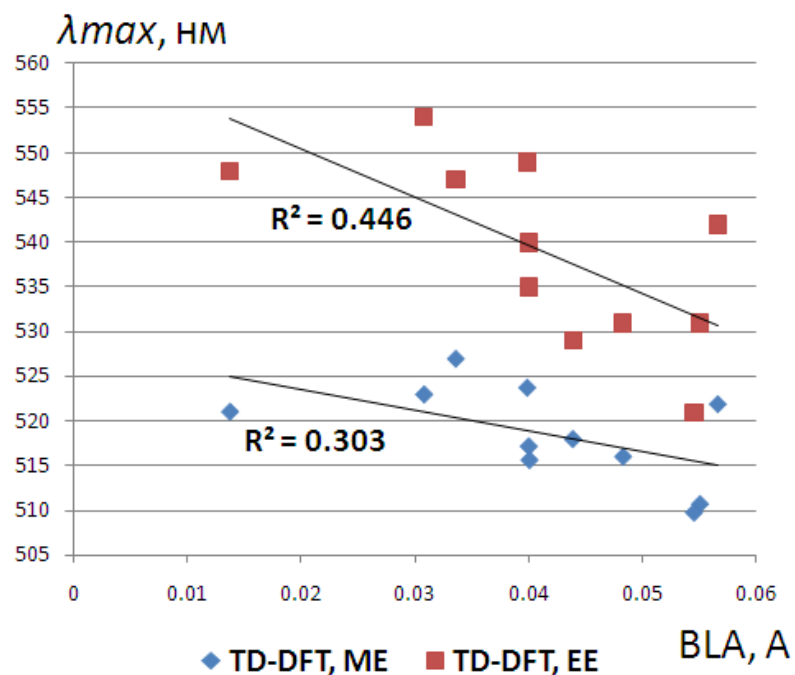


Рисунок 35. Зависимость величин рассчитанных методом TD-DFT (B3LYP) (по схемам ONIOM расчета EE и ME) спектральных максимумов от среднего BLA.

## СПИСОК ЛИТЕРАТУРЫ

1. Lanyi, J. and H. Luecke, *Bacteriorhodopsin. Review*. Current Opinion in Structural Biology, 2001. **11**(4): p. 415-419.
2. Luecke, H., et al., *Structure of bacteriorhodopsin at 1.55 Å resolution*. Journal of Molecular Biology, 1999. **291**(4): p. 899-911.
3. Wise, K., *Optimization of bacteriorhodopsin for bioelectronic devices*. Trends in Biotechnology, 2002. **20**(9): p. 387-394.
4. Pebay-Peyroula, E., et al., *X-ray Structure of Bacteriorhodopsin at 2.5 Å from Microcrystals Grown in Lipidic Cubic Phases*. Science, 1997. **277**(5332): p. 1676-1681.
5. Lanyi, J. and B. Schobert, *Mechanism of Proton Transport in Bacteriorhodopsin from Crystallographic Structures of the K, L, M1, M2, and M22 Intermediates of the Photocycle*. Journal of Molecular Biology, 2003. **328**(2): p. 439-450.
6. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. J Comput Chem, 2004. **25**(13): p. 1605-12.
7. Oesterhelt, D. and W. Stoeckenius, *Rhodopsin-like protein from the purple membrane of Halobacterium halobium*. Nat New Biol, 1971. **233**(39): p. 149-52.
8. Muller, D.J., et al., *Imaging purple membranes in aqueous solutions at sub-nanometer resolution by atomic force microscopy*. Biophys J, 1995. **68**(5): p. 1681-6.
9. Kolbe, M., et al., *Structure of the light-driven chloride pump halorhodopsin at 1.8 Å resolution*. Science, 2000. **288**(5470): p. 1390-6.
10. Haupts, U., J. Tittor, and D. Oesterhelt, *CLOSING IN ON BACTERIORHODOPSIN: Progress in Understanding the Molecule*. Annual Review of Biophysics and Biomolecular Structure, 1999. **28**(1): p. 367-399.
11. Findlay, J.B. and D.J. Pappin, *The opsin family of proteins*. Biochem J, 1986. **238**(3): p. 625-42.
12. Dannon, A., Stoeckenius, W., *Photophosphorylation in Halobacterium halobium*. Proc. Natl. Acad. Sci. U.S.A., 1974. **71**: p. 1234-1238.
13. Oesterhelt, D., *Light-energy transformation in halobacteria - a second path to the nature of photosynthesis*. Nova Acta Leopoldina, 1982. **55**(246).
14. Hampp, N., *Bacteriorhodopsin as a Photochromic Retinal Protein for Optical Memories*. Chemical Reviews, 2000. **100**(5): p. 1755-1776.
15. Shen, Y., et al., *Stabilization of the Membrane Protein Bacteriorhodopsin to 140 °C in two Dimensional Films*. Nature 1993. **366**: p. 48-50.
16. Bamberg, E., et al., *Transmembranous incorporation of photoelectrically active bacteriorhodopsin in planar lipid bilayers*. Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7502-6.
17. Eisenbach, M., S.R. Caplan, and G. Tanny, Biochim. Biophys. Acta, 1979. **554**: p. 269-280.
18. Eisenbach, M. and S.R. Caplan, Biochim. Biophys. Acta, 1979. **554**: p. 281-292.
19. Harbison, G.S., et al., *Dark-adapted bacteriorhodopsin contains 13-cis, 15-syn and all-trans, 15-anti retinal Schiff bases*. Proc Natl Acad Sci U S A, 1984. **81**(6): p. 1706-9.
20. Andersen, L.H., et al., *Absorption of schiff-base retinal chromophores in vacuo*. Journal of the American Chemical Society, 2005. **127**(35): p. 12347-12350.
21. Wald, G., *Molecular basis of visual excitation*. Science, 1968. **162**(850): p. 230-9.
22. Kochendoerfer, G.G., et al., *How color visual pigments are tuned*. Trends Biochem Sci, 1999. **24**(8): p. 300-5.
23. Tachibanaki, S., et al., *Effect of chloride on the thermal reverse reaction of intermediates of iodopsin*. Biochemistry, 1995. **34**(40): p. 13170-5.
24. Kloppmann, E., T. Becker, and G.M. Ullmann, *Electrostatic potential at the retinal of three archaeal rhodopsins: implications for their different absorption spectra*. Proteins, 2005. **61**(4): p. 953-65.



25. Kouyama, T., et al., *Crystal structure of the light-driven chloride pump halorhodopsin from Natronomonas pharaonis*. J Mol Biol, 2010. **396**(3): p. 564-79.
26. Cooper, A., S.F. Dixon, and M. Tsuda, *Photoenergetics of octopus rhodopsin. Convergent evolution of biological photon counters?* Eur Biophys J, 1986. **13**(4): p. 195-201.
27. Bamann, C., et al., *Spectral characteristics of the photocycle of channelrhodopsin-2 and its implication for channel function*. J Mol Biol, 2008. **375**(3): p. 686-94.
28. Bogomolni, R.A., et al., *Action spectrum and quantum efficiency for proton pumping in Halobacterium halobium*. Biochemistry, 1980. **19**(10): p. 2152-9.
29. Grigorieff, N., et al., *Electron-crystallographic refinement of the structure of bacteriorhodopsin*. J Mol Biol, 1996. **259**(3): p. 393-421.
30. Kimura, Y., et al., *Surface of bacteriorhodopsin revealed by high-resolution electron crystallography*. Nature, 1997. **389**(6647): p. 206-11.
31. Essen, L., et al., *Lipid patches in membrane protein oligomers: crystal structure of the bacteriorhodopsin-lipid complex*. Proc Natl Acad Sci U S A, 1998. **95**(20): p. 11673-8.
32. Landau, E.M. and J.P. Rosenbusch, *Lipidic cubic phases: a novel concept for the crystallization of membrane proteins*. Proc Natl Acad Sci U S A, 1996. **93**(25): p. 14532-5.
33. Luecke, H., H.T. Richter, and J.K. Lanyi, *Proton transfer pathways in bacteriorhodopsin at 2.3 angstrom resolution*. Science (New York, N.Y.), 1998. **280**(5371): p. 1934-1937.
34. Belrhali, H., et al., *Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at 1.9 Å resolution*. Structure, 1999. **7**(8): p. 909-917.
35. Schobert, B., et al., *Crystallographic structure of the K intermediate of bacteriorhodopsin: conservation of free energy after photoisomerization of the retinal*. J Mol Biol, 2002. **321**(4): p. 715-26.
36. Stark, M., et al., *From images to interactions: high-resolution phase imaging in tapping-mode atomic force microscopy*. Biophys J, 2001. **80**(6): p. 3009-18.
37. Kates, M., *Biology of halophilic bacteria, Part II. Membrane lipids of extreme halophiles: biosynthesis, function and evolutionary significance*. Experientia, 1993. **49**(12): p. 1027-36.
38. Lanyi, J.K., *Bacteriorhodopsin*. Annu Rev Physiol, 2004. **66**: p. 665-88.
39. Cao, Y., et al., *Water is required for proton transfer from aspartate-96 to the bacteriorhodopsin Schiff base*. Biochemistry, 1991. **30**(45): p. 10972-9.
40. Bondar, A., et al., *Mechanism of Primary Proton Transfer in Bacteriorhodopsin*. Structure, 2004. **12**(7): p. 1281-1288.
41. Henderson, R., et al., *An atomic model for the structure of bacteriorhodopsin*. Biochem Soc Trans, 1990. **18**(5): p. 844.
42. Henderson, R., et al., *Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy*. J Mol Biol, 1990. **213**(4): p. 899-929.
43. Mitsuoka, K., et al., *The structure of bacteriorhodopsin at 3.0 Å resolution based on electron crystallography: implication of the charge distribution*. J Mol Biol, 1999. **286**(3): p. 861-82.
44. Subramaniam, S., *Crystallographic analysis of protein conformational changes in the bacteriorhodopsin photocycle*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2000. **1460**(1): p. 157-165.
45. Sato, H., et al., *Specific lipid-protein interactions in a novel honeycomb lattice structure of bacteriorhodopsin*. Acta Crystallogr D Biol Crystallogr, 1999. **55**(Pt 7): p. 1251-6.
46. Royant, A., et al., *Helix deformation is coupled to vectorial proton transport in the photocycle of bacteriorhodopsin*. Nature, 2000. **406**(6796): p. 645-8.
47. Sass, H.J., et al., *Structural alterations for proton translocation in the M state of wild-type bacteriorhodopsin*. Nature, 2000. **406**(6796): p. 649-53.
48. Facciotti, M.T., et al., *Structure of an early intermediate in the M-state phase of the bacteriorhodopsin photocycle*. Biophys J, 2001. **81**(6): p. 3442-55.
49. Matsui, Y., et al., *Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin*. J Mol Biol, 2002. **324**(3): p. 469-81.

50. Faham, S. and J.U. Bowie, *Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure*. J Mol Biol, 2002. **316**(1): p. 1-6.
51. Nishikawa, T., M. Murakami, and T. Kouyama, *Crystal structure of the 13-cis isomer of bacteriorhodopsin in the dark-adapted state*. J Mol Biol, 2005. **352**(2): p. 319-28.
52. Lanyi, J., *Bacteriorhodopsin*. Current Opinion in Structural Biology, 2001. **11**(4): p. 415-419.
53. Groma, G.I. and Z. Dancshazy, *How Many M Forms are there in the Bacteriorhodopsin Photocycle?* Biophys J, 1986. **50**(2): p. 357-366.
54. Hayashi, S., E. Tajkhorshid, and K. Schulten, *Structural Changes during the Formation of Early Intermediates in the Bacteriorhodopsin Photocycle*. Biophys. J., 2002. **83**(3): p. 1281-1297.
55. Birge, R.R. and T.M. Cooper, *Energy storage in the primary step of the photocycle of bacteriorhodopsin*. Biophys J, 1983. **42**(1): p. 61-9.
56. Edman, K., et al., *High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle*. Nature, 1999. **401**(6755): p. 822-6.
57. Kandori, H., et al., *Structural change of threonine 89 upon photoisomerization in bacteriorhodopsin as revealed by polarized FTIR spectroscopy*. Biochemistry, 1999. **38**(30): p. 9676-83.
58. Nagata, T., et al., *The Hydrogen-Bonding Network of Water Molecules and the Peptide Backbone in the Region Connecting Asp83, Gly120, and Glu113 in Bovine Rhodopsin*. Biochemistry, 1998. **37**(49): p. 17216-17222.
59. Kandori, H., et al., *Structural Changes of pharaonis Phoborhodopsin upon Photoisomerization of the Retinal Chromophore: Infrared Spectral Comparison with Bacteriorhodopsin*. Biochemistry, 2001. **40**(31): p. 9238-9246.
60. Furutan, Y., et al., *FTIR Spectroscopy of the M Photointermediate in pharaonis Phoborhodopsin*. Biophysical Journal, 2002. **83**(6): p. 3482-3489.
61. Warshel, A. and Z.T. Chu, *Nature of the Surface Crossing Process in Bacteriorhodopsin: Computer Simulations of the Quantum Dynamics of the Primary Photochemical Event*. The Journal of Physical Chemistry B, 2001. **105**(40): p. 9857-9871.
62. Shichida, Y., et al., *Absorption spectra of intermediates of bacteriorhodopsin measured by laser photolysis at room temperatures*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 1983. **723**(2): p. 240-246.
63. Rothschild, K.J., P. Roepe, and J. Gillespie, *Fourier transform infrared spectroscopic evidence for the existence of two conformations of the bacteriorhodopsin primary photoproduct at low temperature*. Biochim Biophys Acta, 1985. **808**(1): p. 140-8.
64. Maeda, A., et al., *Tryptophan perturbation in the L intermediate of bacteriorhodopsin: fourier transform infrared analysis with indole-15N shift*. Biochemistry, 1992. **31**(50): p. 12543-5.
65. Maeda, A., et al., *Structures of aspartic acid-96 in the L and N intermediates of bacteriorhodopsin: analysis by Fourier transform infrared spectroscopy*. Biochemistry, 1992. **31**(19): p. 4684-90.
66. Sasaki, J., et al., *Protein changes associated with reprotonation of the Schiff base in the photocycle of Asp96-->Asn bacteriorhodopsin. The MN intermediate with unprotonated Schiff base but N-like protein structure*. J Biol Chem, 1992. **267**(29): p. 20782-6.
67. Fahmy, K., et al., *Aspartic acid-212 of bacteriorhodopsin is ionized in the M and N photocycle intermediates: an FTIR study on specifically 13C-labeled reconstituted purple membranes*. Biochemistry, 1993. **32**(22): p. 5862-9.
68. Hage, W., et al., *Protein Dynamics in the Bacteriorhodopsin Photocycle: A Nanosecond Step-Scan FTIR Investigation of the KL to L Transition*. J. Phys. Chem., 1996. **100**(39): p. 16026-16033.
69. Dioumaev, A.K. and M.S. Braiman, *Nano- and microsecond time-resolved FTIR spectroscopy of the halorhodopsin photocycle*. Photochem Photobiol, 1997. **66**(6): p. 755-63.

70. Einfeld, W., T. Althaus, and M. Stockburger, *Evidence for parallel photocycles and implications for the proton pump in bacteriorhodopsin*. Biophys Chem, 1995. **56**(1-2): p. 105-12.
71. Luecke, H., et al., *Structural changes in bacteriorhodopsin during ion transport at 2 angstrom resolution*. Science, 1999. **286**(5438): p. 255-61.
72. Atkinson, G.H., L. Ujj, and Y. Zhou, *Vibrational Spectrum of the J-625 Intermediate in the Room Temperature Bacteriorhodopsin Photocycle*. J. Phys. Chem. A, 2000. **104**(18): p. 4130-4139.
73. Hasson, K.C., G.A.I. Feng, and P. Anfinrud, *The photoisomerization of retinal in bacteriorhodopsin: Experimental evidence for a three-state model* PNAS. PNAS, 1996.
74. Haran, G., *Excited state dynamics of bacteriorhodopsin revealed by transient stimulated emission spectra*. Chemical Physics Letters, 1996. **261**(4-5): p. 389-395.
75. Gai, F., et al., *Chemical Dynamics in Proteins: The Photoisomerization of Retinal in Bacteriorhodopsin*. Science, 1998. **279**(5358): p. 1886-1891.
76. Kobayashi, T., T. Saito, and H. Ohtani, *Real-time spectroscopy of transition states in bacteriorhodopsin during retinal isomerization*. Nature, 2001. **414**(6863): p. 531-534.
77. Vreven, T.B., F., et al., *Ab Initio Photoisomerization Dynamics of a Simple Retinal Chromophore Model*. J. Am. Chem. Soc., 1997. **119**(51): p. 12687-12688.
78. Garavelli, M., F. Negri, and M. Olivucci, *Initial excited-state relaxation of the isolated 11-cis protonated Schiff base of retinal: Evidence for in-plane motion from ab initio quantum chemical simulation of the resonance Raman spectrum*. Journal of the American Chemical Society, 1999. **121**(5): p. 1023-1029.
79. Peteanu, L.A., et al., *The first step in vision occurs in femtoseconds: complete blue and red spectral studies*. Proc Natl Acad Sci U S A, 1993. **90**(24): p. 11762-6.
80. Mathies, R.A., et al., *Direct observation of the femtosecond excited-state cis-trans isomerization in bacteriorhodopsin* Science, 1988. **240**(4853): p. 777-779.
81. Zimanyi, L., et al., *Pathways of proton release in the bacteriorhodopsin photocycle*. Biochemistry, 1992. **31**(36): p. 8535-43.
82. Рубин, А.Б., *Биофизика, учебник в 2 тт.* 1999.
83. Lanyi, J., *Proton transfers in the bacteriorhodopsin photocycle*. Biochimica et biophysica acta, 2006. **1757**(8): p. 1012-1018.
84. Bondar, A.N., et al., *Key role of active-site water molecules in bacteriorhodopsin proton-transfer reactions*. J Phys Chem B, 2008. **112**(47): p. 14729-41.
85. Edman, K., et al., *Deformation of helix C in the low temperature L-intermediate of bacteriorhodopsin*. J Biol Chem, 2004. **279**(3): p. 2147-58.
86. Ludman, K., C. Gergely, and G. Varo, *Kinetic and thermodynamic study of the bacteriorhodopsin photocycle over a wide pH range*. Biophys. J., 1998 **75**: p. 3110-3119.
87. Tanimoto, T., Y. Furutani, and H. Kandori, *Structural changes of water in the Schiff base region of bacteriorhodopsin: proposal of a hydration switch model*. Biochemistry, 2003. **42**(8): p. 2300-6.
88. Balashov, S.P., et al., *Titration of aspartate-85 in bacteriorhodopsin: what it says about chromophore isomerization and proton release*. Biophys J, 1996. **70**(1): p. 473-81.
89. Lanyi, J., *Crystallographic Structure of the Retinal and the Protein after Deprotonation of the Schiff Base: The Switch in the Bacteriorhodopsin Photocycle*. Journal of Molecular Biology, 2002. **321**(4): p. 727-737.
90. Brown, L.S., et al., *Glutamic acid 204 is the terminal proton release group at the extracellular surface of bacteriorhodopsin*. J Biol Chem, 1995. **270**(45): p. 27122-6.
91. Balashov, S.P., et al., *Glutamate-194 to cysteine mutation inhibits fast light-induced proton release in bacteriorhodopsin*. Biochemistry, 1997. **36**(29): p. 8671-6.
92. Rammelsberg, R., et al., *Bacteriorhodopsin's intramolecular proton-release pathway consists of a hydrogen-bonded network*. Biochemistry, 1998. **37**(14): p. 5001-9.

93. Spassov, V., *pKa calculations suggest storage of an excess proton in a hydrogen-bonded water network in bacteriorhodopsin*. Journal of Molecular Biology, 2001. **312**(1): p. 203-219.
94. Szaraz, S., D. Oesterhelt, and P. Ormos, *pH-induced structural changes in bacteriorhodopsin studied by Fourier transform infrared spectroscopy*. Biophys J, 1994. **67**(4): p. 1706-12.
95. Brown, L.S. and J.K. Lanyi, *Determination of the transiently lowered pKa of the retinal Schiff base during the photocycle of bacteriorhodopsin*. Proc Natl Acad Sci U S A, 1996. **93**(4): p. 1731-4.
96. Zscherp, C., et al., *In situ determination of transient pKa changes of internal amino acids of bacteriorhodopsin by using time-resolved attenuated total reflection Fourier-transform infrared spectroscopy*. Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5498-503.
97. Li, Q., et al., *On the protein residues that control the yield and kinetics of O(630) in the photocycle of bacteriorhodopsin*. Biophys J, 2000. **78**(1): p. 354-62.
98. Dioumaev, A.K., et al., *Coupling of the reisomerization of the retinal, proton uptake, and reprotonation of Asp-96 in the N photointermediate of bacteriorhodopsin*. Biochemistry, 2001. **40**(38): p. 11308-17.
99. Varo, G., et al., *Light-driven chloride ion transport by halorhodopsin from Natronobacterium pharaonis. 1. The photochemical cycle*. Biochemistry, 1995. **34**(44): p. 14490-9.
100. Humphrey, W., et al., *Molecular dynamics study of bacteriorhodopsin and artificial pigments*. Biochemistry, 1994. **33**(12): p. 3668-78.
101. Roux, B., et al., *Thermodynamic stability of water molecules in the bacteriorhodopsin proton channel: a molecular dynamics free energy perturbation study*. Biophys J, 1996. **71**(2): p. 670-81.
102. Wikstrom, M., *Proton translocation by bacteriorhodopsin and heme-copper oxidases*. Curr Opin Struct Biol, 1998. **8**(4): p. 480-8.
103. Luecke, H., *Atomic resolution structures of bacteriorhodopsin photocycle intermediates: the role of discrete water molecules in the function of this light-driven ion pump*. Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2000. **1460**(1): p. 133-156.
104. Kamikubo, H., et al., *Structure of the N intermediate of bacteriorhodopsin revealed by x-ray diffraction*. Proc Natl Acad Sci U S A, 1996. **93**(4): p. 1386-90.
105. Kamikubo, H., et al., *The last phase of the reprotonation switch in bacteriorhodopsin: the transition between the M-type and the N-type protein conformation depends on hydration*. Biochemistry, 1997. **36**(40): p. 12282-7.
106. Hendrickson, F.M., F. Burkard, and R.M. Glaeser, *Structural characterization of the L-to-M transition of the bacteriorhodopsin photocycle*. Biophys J, 1998. **75**(3): p. 1446-54.
107. Oka, T., et al., *Time-resolved x-ray diffraction reveals multiple conformations in the M-N transition of the bacteriorhodopsin photocycle*. Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14278-82.
108. Oka, T., et al., *Time-resolved X-ray diffraction reveals movement of F helix of D96N bacteriorhodopsin during M-MN transition at neutral pH*. Biophys J, 2002. **82**(5): p. 2610-6.
109. Dencher, N.A., et al., *Structural changes in bacteriorhodopsin during proton translocation revealed by neutron diffraction*. Proc Natl Acad Sci U S A, 1989. **86**(20): p. 7876-9.
110. Xiao, W., et al., *Light-induced rotation of a transmembrane alpha-helix in bacteriorhodopsin*. J Mol Biol, 2000. **304**(5): p. 715-21.
111. Brown, L.S., R. Needleman, and J.K. Lanyi, *Conformational change of the E-F interhelical loop in the M photointermediate of bacteriorhodopsin*. J Mol Biol, 2002. **317**(3): p. 471-8.
112. Oka, T., et al., *Conformational change of helix G in the bacteriorhodopsin photocycle: investigation with heavy atom labeling and x-ray diffraction*. Biophys J, 1999. **76**(2): p. 1018-23.

113. Brown, L.S., et al., *Functional significance of a protein conformation change at the cytoplasmic end of helix F during the bacteriorhodopsin photocycle*. Biophys J, 1995. **69**(5): p. 2103-11.
114. Otto, H., et al., *Aspartic acid-96 is the internal proton donor in the reprotonation of the Schiff base of bacteriorhodopsin*. Proc Natl Acad Sci U S A, 1989. **86**(23): p. 9228-32.
115. Gerwert, K., et al., *Role of aspartate-96 in proton translocation by bacteriorhodopsin*. Proc Natl Acad Sci U S A, 1989. **86**(13): p. 4943-7.
116. Dioumaev, A.K., et al., *Partitioning of free energy gain between the photoisomerized retinal and the protein in bacteriorhodopsin*. Biochemistry, 1998. **37**(28): p. 9889-9893.
117. Checover, S., et al., *Mechanism of proton entry into the cytoplasmic section of the proton-conducting channel of bacteriorhodopsin*. Biochemistry, 1997. **36**(45): p. 13919-28.
118. Riesle, J., et al., *D38 is an essential part of the proton translocation pathway in bacteriorhodopsin*. Biochemistry, 1996. **35**(21): p. 6635-43.
119. Rouhani, S., et al., *Crystal structure of the D85S mutant of bacteriorhodopsin: model of an O-like photocycle intermediate*. J Mol Biol, 2001. **313**(3): p. 615-28.
120. Dioumaev, A.K., et al., *Fourier transform infrared spectra of a late intermediate of the bacteriorhodopsin photocycle suggest transient protonation of Asp-212*. Biochemistry, 1999. **38**(31): p. 10070-8.
121. Bondar, A.N., et al., *Tuning of retinal twisting in bacteriorhodopsin controls the directionality of the early photocycle steps*. The journal of physical chemistry. B, 2005. **109**(31): p. 14786-14788.
122. Tan, E.H. and R.R. Birge, *Correlation between surfactant/micelle structure and the stability of bacteriorhodopsin in solution*. Biophys J, 1996. **70**(5): p. 2385-95.
123. Sekharan, S., O. Weingart, and V. Buss, *Ground and excited states of retinal schiff base chromophores by multiconfigurational perturbation theory*. Biophysical journal, 2006. **91**(1).
124. Hoffmann, M., et al., *Color Tuning in Rhodopsins: The Mechanism for the Spectral Shift between Bacteriorhodopsin and Sensory Rhodopsin II*. J. Am. Chem. Soc., 2006. **128**(33): p. 10808-10818.
125. Vreven, T. and K. Morokuma, *Investigation of the S0 S1 excitation in bacteriorhodopsin with the ONIOM(MO:MM) hybrid method*. Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta), 2003. **109**(3): p. 125-132.
126. Fujimoto, K., et al., *Mechanism of color tuning in retinal protein: SAC-CI and QM/MM study*. Chemical Physics Letters, 2005. **414**(1-3): p. 239-242.
127. Cembran, A., et al., *The retinal chromophore/chloride ion pair: Structure of the photoisomerization path and interplay of charge transfer and covalent states*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(18): p. 6255-6260.
128. Altun, A., S. Yokoyama, and K. Morokuma, *Mechanism of spectral tuning going from retinal in vacuo to bovine rhodopsin and its mutants: multireference ab initio quantum mechanics/molecular mechanics studies*. J Phys Chem B, 2008. **112**(51): p. 16883-90.
129. Wanko, M., et al., *Computational photochemistry of retinal proteins*. Journal of computer-aided molecular design, 2006. **20**(7-8): p. 511-518.
130. Lan, Z., E. Fabiano, and W. Thiel, *Photoinduced nonadiabatic dynamics of 9H-guanine*. Chemphyschem, 2009. **10**(8): p. 1225-9.
131. Sanchez-Garcia, E., M. Doerr, and W. Thiel, *QM/MM study of the absorption spectra of DsRed.M1 chromophores*. J Comput Chem, 2010. **31**(8): p. 1603-12.
132. Wanko, M., et al., *Calculating absorption shifts for retinal proteins: computational challenges*. The journal of physical chemistry. B, 2005. **109**(8): p. 3606-3615.
133. Hoffmann, M., et al., *Color tuning in rhodopsins: the mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II*. J Am Chem Soc, 2006. **128**(33): p. 10808-18.

134. Maseras, F. and K. Morokuma, *IMOMM : A New Integrated Ab Initio + Molecular Mechanics Optimization Scheme of Equilibrium Structure*. Journal of Computational Chemistry, 1995. **16**.
135. Dapprich, S., et al., *A New ONIOM Implementation in Gaussian 98. Part I. The Calculation of Energies, Gradients, Vibrational Frequencies and Electric Field Derivatives*. J. Mol. Str. (Theochem), 1999: p. 461-462.
136. Vreven, T., et al., *Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints*. Journal of Computational Chemistry, 2003. **24**(6): p. 760-769.
137. Altun, A., S. Yokoyama, and K. Morokuma, *Spectral tuning in visual pigments: an ONIOM(QM:MM) study on bovine rhodopsin and its mutants*. J Phys Chem B, 2008. **112**(22): p. 6814-27.
138. Altun, A., S. Yokoyama, and K. Morokuma, *Quantum mechanical/molecular mechanical studies on spectral tuning mechanisms of visual pigments and other photoactive proteins*. Photochem Photobiol, 2008. **84**(4): p. 845-54.
139. Altun, A., S. Yokoyama, and K. Morokuma, *Color tuning in short wavelength-sensitive human and mouse visual pigments: ab initio quantum mechanics/molecular mechanics studies*. J Phys Chem A, 2009. **113**(43): p. 11685-92.
140. Sekharan, S. and K. Morokuma, *Drawing the Retinal Out of Its Comfort Zone: An ONIOM(QM/MM) Study of Mutant Squid Rhodopsin*. J Phys Chem Lett, 2010. **1**(3): p. 668-672.
141. van Gisbergen, S.J.A., J.G. Snijders, and E.J. Baerends, *A density functional theory study of frequency-dependent polarizabilities and Van der Waals dispersion coefficients for polyatomic molecules*. J. Chem. Phys., 1995. **103**(21): p. 9347-9354.
142. Elstner, M., et al., *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties*. Phys. Rev. B, 1998. **58** p. 7260-68.
143. Porezag, D., et al., *Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon*. Phys Rev B Condens Matter, 1995. **51**(19): p. 12947-12957.
144. Frauenheim, T., et al., *A self-consistent charge density-functional based tight-binding method for predictive materials simulations in physics, chemistry and biology*. Physica Stat. Sol. B, 2000. **217**: p. 41-62.
145. Zheng, G., et al., *Parameter calibration of transition-metal elements for the spin-polarized self-consistent-charge density-functional tight-binding (DFTB) method: Sc, Ti, Fe, Co and Ni*. J. Chem. Theory and Comput., 2007. **3**: p. 1349-67.
146. van Gunsteren, W.F. and H.J. Berendsen, *Computer simulation as a tool for tracing the conformational differences between proteins in solution and in the crystalline state*. J Mol Biol, 1984. **176**(4): p. 559-64.
147. Hoover, W.G., *Canonical dynamics: Equilibrium phase-space distributions*. Phys Rev A, 1985. **31**(3): p. 1695-1697.
148. Car, R. and M. Parrinello, *Unified approach for molecular dynamics and density-functional theory*. Phys Rev Lett, 1985. **55**(22): p. 2471-2474.
149. Warshel, A. and M. Levitt, *Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme*. J Mol Biol, 1976. **103**(2): p. 227-49.
150. Kairys, V. and J.H. Jensen, *QM/MM Boundaries Across Covalent Bonds: A Frozen Localized Molecular Orbital-Based Approach for the Effective Fragment Potential Method*. The Journal of Physical Chemistry A, 2000. **104**(28): p. 6656-6665.
151. Frisch, M.J.T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.;

- Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., *Gaussian 09, Revision A.1*. Gaussian, Inc., 2009.
152. Morokuma, K., *New challenges in quantum chemistry: Quests for accurate calculations for large molecular systems*. Philos Transact A Math Phys Eng Sci, 2002. **360**(1795): p. 1149-64.
  153. Bas, D.C., D.M. Rogers, and J.H. Jensen, *Very fast prediction and rationalization of pKa values for protein-ligand complexes*. Proteins, 2008. **73**(3): p. 765-83.
  154. Li, H., A.D. Robertson, and J.H. Jensen, *Very fast empirical prediction and rationalization of protein pKa values*. Proteins, 2005. **61**(4): p. 704-21.
  155. Sasaki, J., et al., *Complete identification of C = O stretching vibrational bands of protonated aspartic acid residues in the difference infrared spectra of M and N intermediates versus bacteriorhodopsin*. Biochemistry, 1994. **33**(11): p. 3178-84.
  156. Brown, L.S., et al., *The complex extracellular domain regulates the deprotonation and reprotonation of the retinal Schiff base during the bacteriorhodopsin photocycle*. Biochemistry, 1995. **34**(39): p. 12903-11.
  157. Wang, J., P. Cieplak, and P.A. Kollman, *How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules*. J. Comp. Chem., 2000. **21**: p. 1049-1074.
  158. Altun, A., S. Yokoyama, and K. Morokuma, *Spectral tuning in visual pigments: an ONIOM(QM:MM) study on bovine rhodopsin and its mutants*. The journal of physical chemistry. B, 2008. **112**(22): p. 6814-6827.
  159. Niehaus, T.A., M. Elstner, and T. Frauenheim, *Application of an approximate density-functional method to sulfur containing compounds*. J. Mol. Struc. (THEOCHEM), 2001. **541**(1-3): p. 185.
  160. Van Der Spoel, D., et al., *GROMACS: fast, flexible, and free*. J Comput Chem, 2005. **26**(16): p. 1701-18.