

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Миронов Владимир Николаевич

**ИЗУЧЕНИЕ МЕХАНИЗМОВ ФОТОАКТИВАЦИИ  
ПОТЕНЦИАЛ-ЗАВИСИМЫХ ФЛУОРЕСЦЕНТНЫХ  
БЕЛКОВ НА ОСНОВЕ АРХЕОРОДОПСИНА-3**

Уровень образования: магистратура

Направление 28.04.04 «Наносистемы и наноматериалы»

Основная образовательная программа ВМ.5724.2019 «Фундаментальные и  
прикладные аспекты наноматериалов и нанотехнологий»

Научный руководитель:

д.х.н., доцент, кафедра лазерной  
химии и лазерного материаловедения  
Рязанцев Михаил Николаевич

Рецензент:

к.х.н., доцент, в.н.с., Вязьмин Сергей Юрьевич  
Санкт-Петербургский национальный  
исследовательский Академический  
университет имени Ж.И.Алферова РАН»,

Санкт-Петербург — 2021

## Оглавление

	Стр.
<b>Введение</b> . . . . .	4
<b>Глава 1. Обзор литературы</b> . . . . .	6
1.1 Структура бактериальных родопсинов. . . . .	6
1.2 Фотоизомеризация хромофора. . . . .	6
1.3 Фотоцикл бактериородопсина. . . . .	8
1.4 Исследование спектральных характеристик основного состояния и интермедиатов фотоцикла бактериородопсина. . . . .	9
1.4.1 Спектры комбинационного рассеяния. . . . .	9
1.4.2 Спектры инфракрасного поглощения (ИК спектры). . . . .	13
1.4.3 Флуоресценция бактериородопсина. . . . .	14
1.4.4 Сравнение фотоциклов бактериородопсина и археородопсина-3. . . . .	14
1.5 Характеристики флуоресценции археородопсина-3 и его мутантных форм. . . . .	15
1.6 Предложенные модели механизма потенциал-зависимости флуоресценции археородопсина-3. . . . .	18
1.7 Модификация стабильности О-состояния в бактериородопсине путем введения аминокислотных замен. . . . .	20
<b>Глава 2. Методы</b> . . . . .	22
2.1 Экспериментальные методы . . . . .	22
2.2 Теоретические методы. . . . .	23
2.2.1 Получение трехмерных структур белков. . . . .	23
2.2.2 Расчет спектральных свойств. . . . .	24
2.2.3 Расчет рКа аминокислот. . . . .	24
2.2.4 Расчет энергии Гиббса между состояниями белка. . . . .	24
<b>Глава 3. Результаты и обсуждение</b> . . . . .	27
3.1 Анализ экспериментальных результатов. . . . .	27
3.2 Анализ результатов компьютерного моделирования. . . . .	38

Выводы . . . . .	45
Благодарности . . . . .	47
Список литературы . . . . .	48
Приложение А. Праймеры для мутагенеза . . . . .	54
Приложение Б. Использованные буферы . . . . .	55
Приложение В. Результаты секвенирования . . . . .	56

## Введение

Археородопсин-3 – трансмембранный белок из семейства бактериальных родопсинов, применяемый в оптогенетике для визуализации активности нейронов [1–3]. Данная область применения обеспечена специфическими характеристиками флуоресценции археородопсина-3. Интенсивность флуоресценции данного белка линейно зависит от величины потенциала внешнего электрического поля. Таким образом, интенсивность флуоресценции археородопсина-3, функционирующего в мембране нейрона, изменяется при изменении величины мембранного потенциала нейрона, что позволяет проводить мониторинг активности нейронов при помощи оптических методов.

Важной задачей в оптогенетике является создание белков на основе археородопсина-3 с модифицированными характеристиками флуоресцентного сигнала. Основная часть исследований, посвященных разработке производных археородопсина-3 для оптогенетических применений, направлена на достижение следующих целей: 1) повышение интенсивности флуоресцентного сигнала; 2) повышение чувствительности интенсивности флуоресцентного сигнала к изменению величины мембранного потенциала; 3) сдвиг спектра поглощения белка в длинноволновой диапазон. Проводимые на сегодняшний день исследования основаны на переборе большого количества вариантов мутантных форм белка при помощи методов направленной эволюции и случайного мутагенеза. Более перспективным подходом является рациональный молекулярный дизайн, однако его применение ограничено отсутствием данных о механизме, определяющем потенциал-зависимость флуоресценции белка. В частности, необходимо определить, из какого состояния белка возникает флуоресцентный сигнал, а также механизм стабилизации этого состояния. Повышение концентрации состояния, из которого возникает флуоресценция, при помощи аминокислотных замен позволило бы увеличить интенсивность флуоресцентного сигнала. Помимо этого, необходимой задачей является определение механизма, определяющего зависимость интенсивности флуоресцентного сигнала от величины мембранного потенциала клетки.

Основной целью данного исследования стало определение механизма, отвечающего за потенциал-зависимость флуоресценции археородопсина-3 и его мутантных форм. Для достижения данной цели были поставлены и решены следующие задачи:

- Определение состояния белка, фотоактивация которого приводит к возникновению флуоресцентного сигнала (O-состояние), определение механизма стабилизации данного состояния;
- Определение механизма, определяющего потенциал-зависимость флуоресценции белка.

На основании полученных данных был предложен ряд аминокислотных замен, введение которых, согласно предложенному в работе механизму, приводит к стабилизации O-состояния археородопсина-3 и, соответственно, к повышению интенсивности флуоресцентного сигнала. На сегодняшний день проводятся работы по синтезу предложенных мутантных форм археородопсина-3 для проверки данной гипотезы.

## Глава 1. Обзор литературы

Основным объектом данного исследования является фоточувствительный трансмембранный белок из семейства бактериальных родопсинов, археородопсин-3 из *Halorubrum sodomense*. В данной части работы приведен обзор литературы по процессам, происходящим после фотоактивации бактериальных родопсинов на примере наиболее изученного белка, бактериородопсина из бактерии *Halobacterium salinarum*, а также приведен обзор характерных особенностей археородопсина-3 и его мутантных форм.

### 1.1 Структура бактериальных родопсинов.

Бактериальные родопсины – трансмембранные белки, которые состоят из семи альфа-спиралей, расположенных параллельно друг другу, и кофактора, связанного с лизином, расположенным на седьмой альфа-спирали. Кофактором родопсина является ретинальное основание Шиффа (Рисунок 1.1). В основном состоянии кофактор находится в протонированном состоянии. В темновом состоянии бактериородопсин представляет собой смесь двух изомерных форм: транс-формы (33%) и 13-цис-формы (67%), при фотооблучении в основном состоянии наблюдается исключительно транс-изомер.

### 1.2 Фотоизомеризация хромофора.

Первичным процессом, происходящим после фотоактивации бактериальных родопсинов, является изомеризация хромофора[4]. В результате поглощения фотона хромофор, находящийся в транс-форме, переходит в возбужденное синглетное состояние, после чего система начинает двигаться по поверхности потенциальной энергии (ППЭ) возбужденного состояния. Схематические изображения поверхности потенциальной энергии и процесса

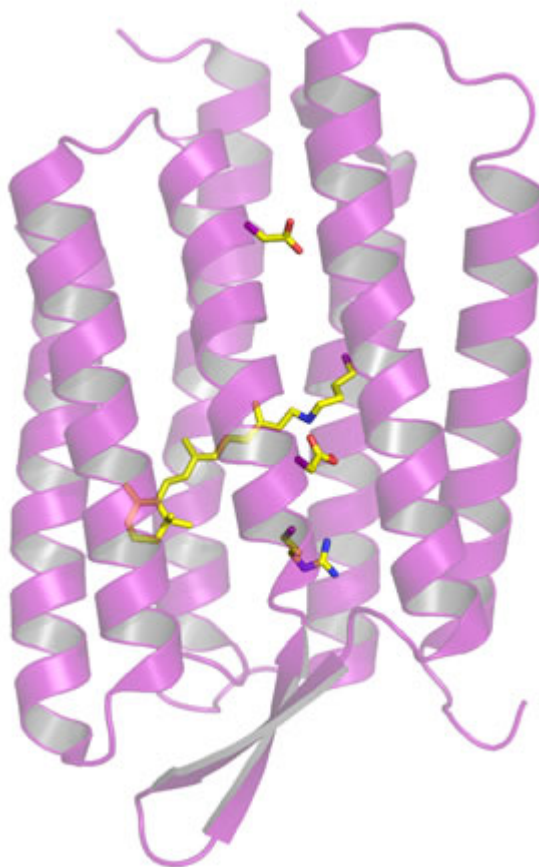


Рисунок 1.1 — Структура бактериородопсина, визуализация в программе VMD

фотоизомеризации представлены на Рисунке 1.2. В качестве основной координаты реакции выбран угол поворота вокруг двойной связи, вокруг которой происходит фотоизомеризация.

Помимо моды изомеризации, после фотовозбуждения активируется мода продольных колебаний вдоль пи-сопряженной системы хромофора. К моменту перехода системы с ППЭ возбужденного состояния на ППЭ основного состояния фаза продольных колебаний соответствует состоянию, в котором длина двойной связи  $C13=C14$ , вокруг которой происходит изомеризация, близка к длине одинарной  $C-C$  связи, что способствует изомеризации [4]. После безызлучательного перехода в основное состояние в точке конического пересечения реакция может пойти по двум путям – обратно в реагент (транс-изомер), либо в продукт (13-цис изомер). Квантовый выход фотоизомеризации для бактериородопсина равен 67%.

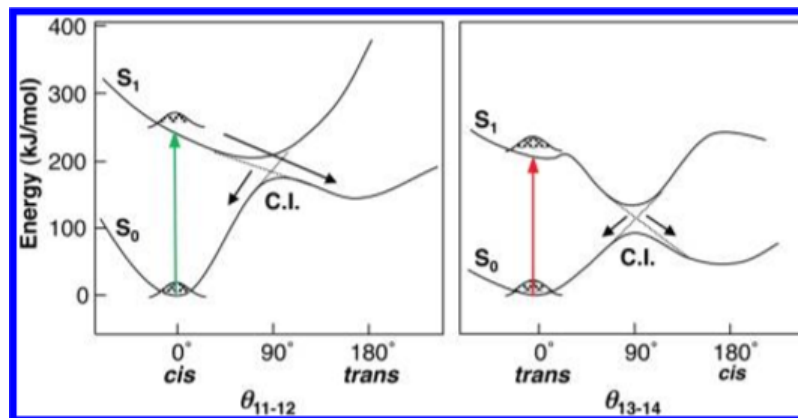


Рисунок 1.2 — Процесс фотоизомеризации и поверхности потенциальной энергии основного и возбужденного состояний. Схема взята из [4]

### 1.3 Фотоцикл бактериородопсина.

Энергия, полученная хромофором при поглощении фотона, после фотоизомеризации диссипирует в окружающую среду, что приводит к глобальной конформационной перестройке всего белка. После фотоактивации белок проходит несколько состояний – промежуточных состояний фотоцикла, которые отличаются друг от друга как конформацией белка, так и состояниями протонирования различных аминокислот. После прохождения фотоцикла родопсин возвращается в основное состояние (Рисунок 1.3). Время полного цикла для бактериородопсина составляет около 30 мс, он состоит из семи промежуточных состояний.

Первым состоянием фотоцикла бактериородопсина является J625 [5], которое переходит в короткоживущие K590 и, последовательно, в L550 интермедиаты. Наиболее существенные изменения происходят при переходе из L550 в M420, в ходе которого происходит отрыв протона от хромофора белка, а спектр поглощения сдвигается в синюю область спектра (420 нм). В дальнейшем происходит репротонирование хромофора (N550) и обратная изомеризация из 13-цис в транс-форму (O640). Затем белок возвращается в исходное состояние, результатом фотоцикла становится перенос одного протона через мембрану [6] (Рисунок 1.3).



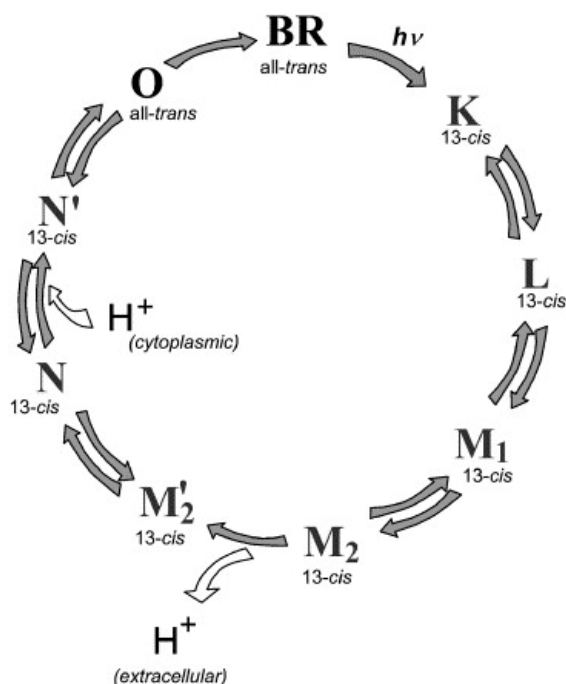


Рисунок 1.3 — Фотоцикл бактериородопсина

## 1.4 Исследование спектральных характеристик основного состояния и интермедиатов фотоцикла бактериородопсина.

### 1.4.1 Спектры комбинационного рассеяния.

Для определения состояния хромофора и его локального окружения могут использоваться данные спектроскопии комбинационного рассеяния (КР). На КР спектрах хромофора виден ряд характерных полос, соответствующих определенным колебаниям. Положение пиков меняется в зависимости от окружения хромофора – электростатического поля белка. Таким образом, интерпретация спектров комбинационного рассеяния позволяет определить изменения, происходящие с белком после фотоактивации, а также различия между различными мутантными формами белков.

1. Пик скелетного колебания соответствует продольным колебаниям вдоль  $\pi$ -сопряженной цепочки хромофора. В ходе этого колебания происходит периодическое укорачивание одиночных связей и удлинение двойных связей в сопряженной цепочке. Данный пик является самым интенсивным на КР

спектре и обычно лежит в диапазоне 1500-1550  $\text{см}^{-1}$ . Частота пика коррелирует с локальным электростатическим полем и спектром поглощения белка. Для основного состояния бактериородопсина частота его составляет 1530  $\text{см}^{-1}$ .

2. Регион характерных полос (fingerprint region) включает три основные полосы:

- на 1169-1170  $\text{см}^{-1}$  (сильно локализованное колебание C10-C11, активно у транс-изомера хромофора);
- около 1185  $\text{см}^{-1}$  – пик 13-цис изомера хромофора, локализованное колебание C14-C15;
- на 1200-1201  $\text{см}^{-1}$  – пик транс-изомера хромофора.

3. Пик, соответствующий синхронному колебанию метильных групп – сильный пик 1005-1007  $\text{см}^{-1}$ , одинаковый для цис- и транс-изомеров, может использоваться для нормировки остальных пиков.

4. Пик, соответствующий моде HOOP (hydrogen out of plane motion) – колебанию атомов водорода, связанных с атомами C13 и C14, вне плоскости углеродного скелета, расположен в области 970-958  $\text{см}^{-1}$ .

5. Пик, соответствующий C=N колебанию. Данное колебание может характеризовать наличие воды и пентагонального водного кластера рядом с хромофором.

Для бактериородопсина были получены спектры всех интермедиатов [7] (КР спектры L,N и M,O интермедиатов приведены на Рисунках 1.4, 1.5), а также спектры основного состояния в светоадаптированном (BR570) и темновом (BR540) состояниях (Рисунок 1.6) [8].

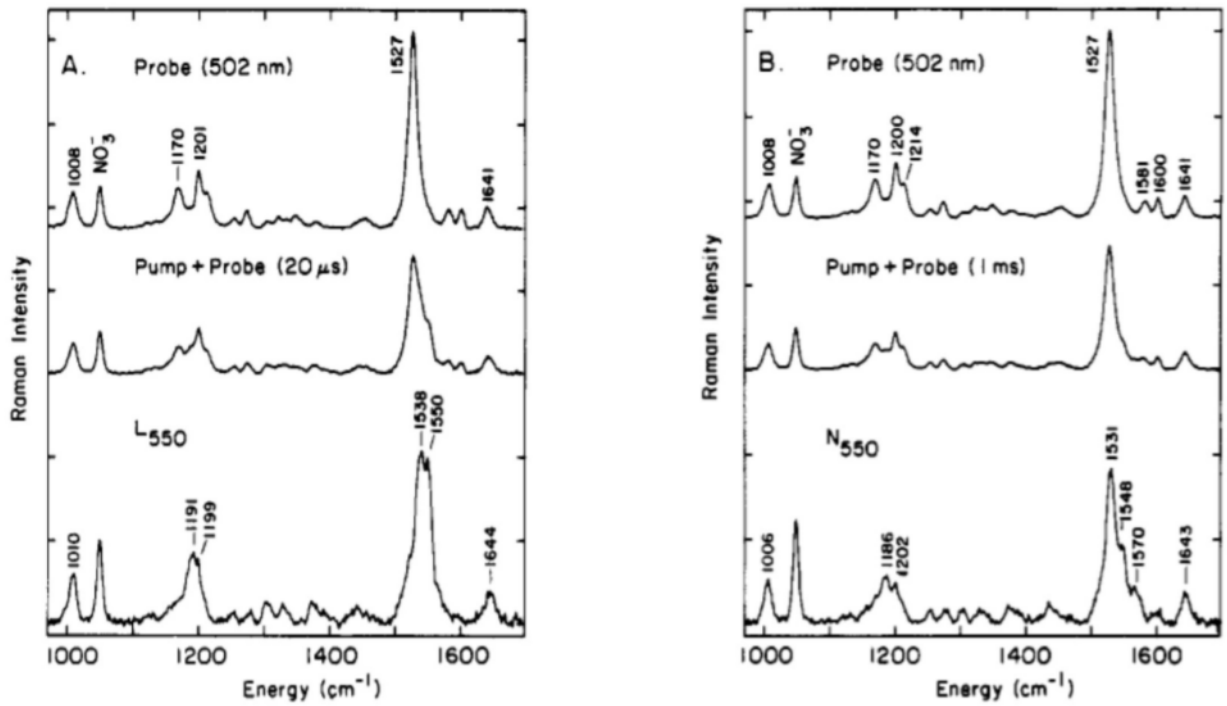


Рисунок 1.4 — КР спектры L(слева) и N(справа) интермедиатов бактериородопсина [7].

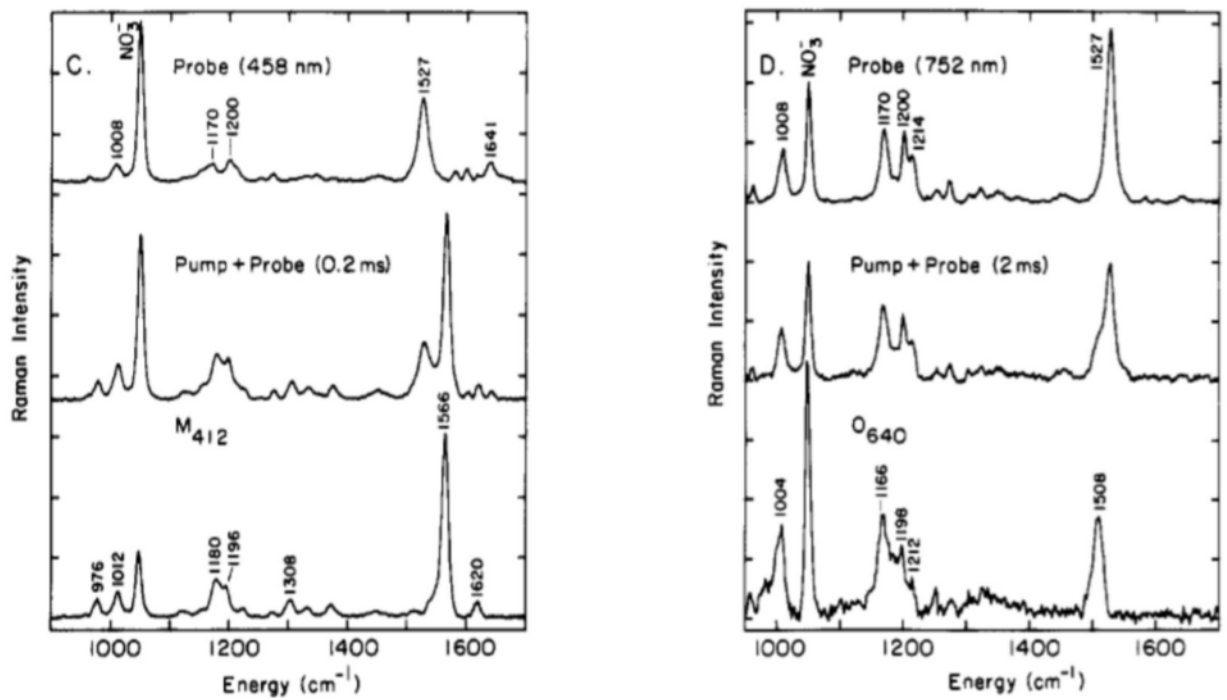


Рисунок 1.5 — КР спектры M(слева) и O(справа) интермедиатов бактериородопсина [7].

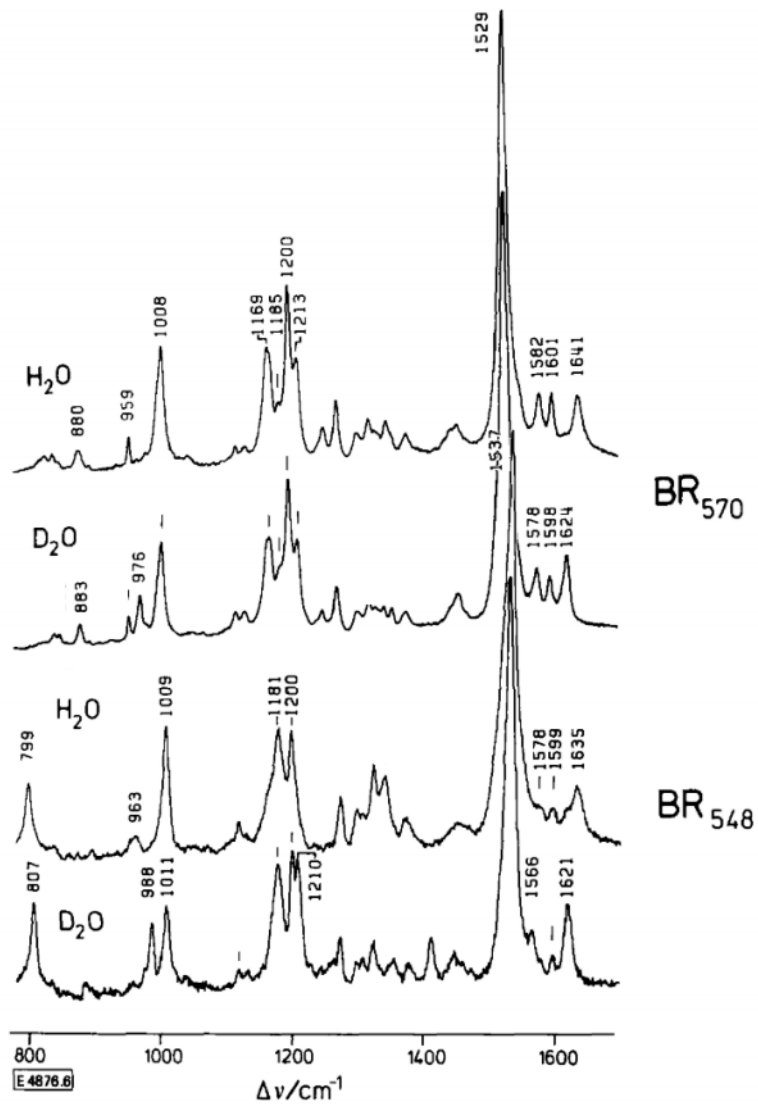


Рисунок 1.6 — КР спектры светоадаптированного (BR<sub>570</sub>) и темнового (BR<sub>548</sub>) состояний бактериородопсина из [8]

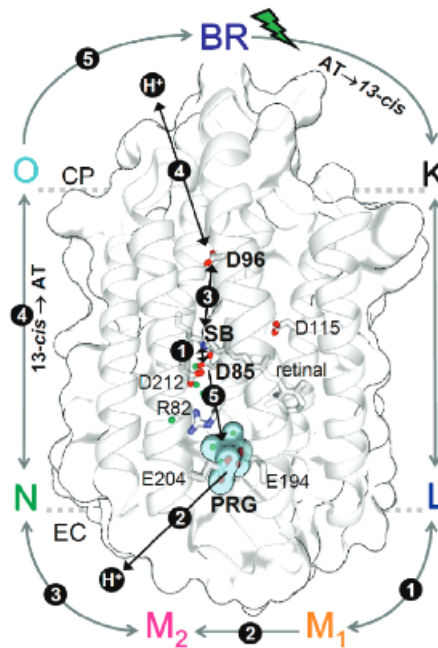


Рисунок 1.7 — Перенос протона в фотоцикле бактериородопсина, рисунок взят из [10]

#### 1.4.2 Спектры инфракрасного поглощения (ИК спектры).

На ИК спектрах бактериородопсина, помимо вышеуказанных полос, наблюдаются также колебания атомов других аминокислот белка, в частности, колебания  $\text{COO}^-$  и  $\text{COOH}$  групп. Благодаря изменению частот колебаний в зависимости от протонирования карбоксильной группы ( $\text{COO}^-$  и  $\text{COOH}$ ) могут быть определены состояния протонирования аспарагиновых и глутаминовых кислот белка.

Следует отметить, что ввиду наличия большого количества полос и перекрывания части этих полос с полосами, соответствующими колебаниям молекул воды, анализ ИК спектров трудоемок и сложен для интерпретации [9–11]. Тем не менее анализ протонирования различных аминокислот позволил определить приблизительный путь переноса протона бактериородопсином через мембрану (Рисунок 1.7). Протон переносится за счет серии последовательных переходов между  $\text{COO}^-$  группами различных аминокислот.

### 1.4.3 Флуоресценция бактериородопсина.

Ряд бактериальных родопсинов обладает слабой флуоресценцией, при этом флуоресцентным может быть как основное состояние белка, так и какой-либо из интермедиатов фотоцикла [12]. Флуоресценция бактериородопсина была изучена методами оптической спектроскопии с временным разрешением. На основании измерения времени возникновения и затухания флуоресценции было показано, что флуоресцентным является Q интермедиат [13; 14], являющийся продуктом N интермедиата. Непосредственно само возникновение флуоресценции рассмотрено в работе [15].

Предполагается, что возникновение барьера на поверхности потенциальной энергии возбужденного состояния  $S_1$  (по основной координате угла поворота) способствует замедлению изомеризации, являющейся для флуоресценции конкурирующим процессом, и ведет к увеличению квантового выхода флуоресценции. Порядок двойной связи C13=C14 коррелирует с величиной барьера, чем длина связи ближе к двойной, тем выше квантовый выход флуоресценции за счет затруднения изомеризации вокруг этой связи.

### 1.4.4 Сравнение фотоциклов бактериородопсина и археородопсина-3.

Археородопсин-3 имеет с бактериородопсином гомологию около 70 процентов. При этом все основные аминокислоты, участвующие в фотоцикле белка или находящиеся в непосредственной близости от хромофора консервативны для археородопсина-3 и бактериородопсина. Спектры поглощения в светоадаптированном состоянии похожи – 556 и 560 нм для археородопсина-3 и бактериородопсина соответственно.

На основании сопоставления ИК и КР спектров интермедиатов фотоцикла бактериородопсина (BR) и археородопсина-3 (Arch-3), а также анализов результатов спектроскопии поглощения с временным разрешением было показано, что фотоциклы Arch3 и BR практически совпадают [16; 17].

### 1.5 Характеристики флуоресценции археородопсина-3 и его мутантных форм.

У многих бактериальных родопсинов была обнаружена флуоресценция, для бактериородопсина и археородопсина-3 квантовый выход флуоресценции составляет около  $9 \cdot 10^{-5}$  и  $10^{-4}$  соответственно. Было обнаружено, что интенсивность флуоресценции археородопсина-3 и протеородопсина зависит от величины мембранного потенциала. Один из первых сенсоров был сделан на основе протеородопсина (PROPS). В связи с более простой экспрессией в эукариотах для дальнейших исследований использовался Arch-3 [3].

Основными характеристиками сенсоров мембранного потенциала являются:

- квантовый выход флуоресценции;
- чувствительность  $\frac{\Delta F}{F} * 100\%$  (или процентное изменение флуоресценции на 100мВ);
- соотношение полезного сигнала к шуму (signal to noise ratio, SNR);
- время отклика.

Также важно отсутствие влияния белка на мембранный потенциал (отключение функций белка как протонного насоса), красный спектр и достаточная величина стоксова сдвига, эффективность экспрессии.

Первыми улучшенными сенсорами мембранного потенциала стали мутанты Arch-3 по основному контриону Asp95 [18]. Эти мутации помимо усиления флуоресценции и чувствительности  $dF/F$  отключают способность белка переносить протон, что уменьшает погрешность измерений, связанных с влиянием

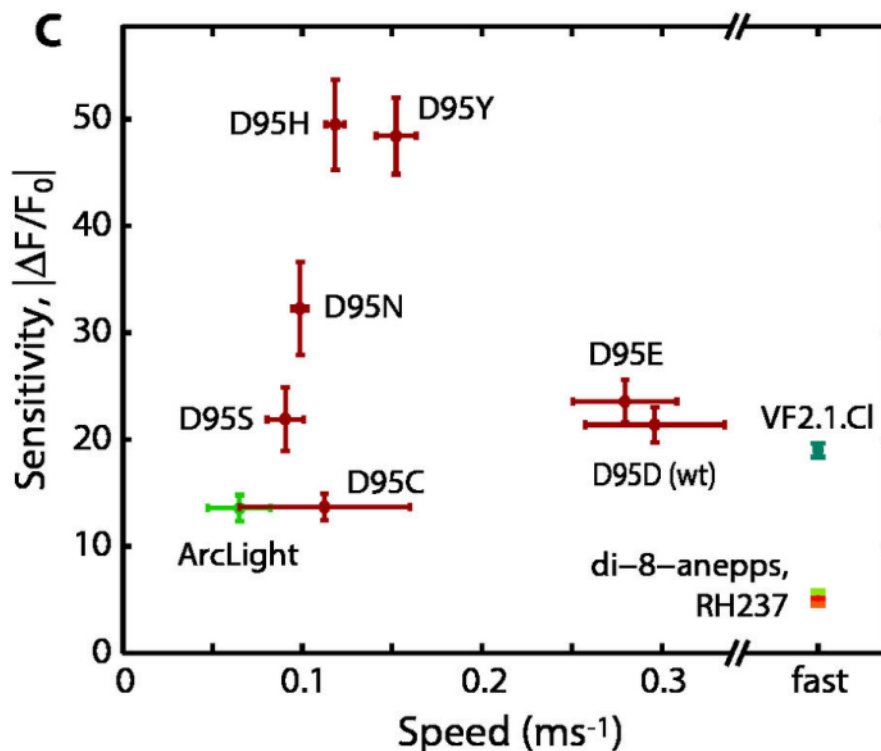


Рисунок 1.8 — Мутанты Arch-3 D95X, рисунок взят из [18]

белка на мембранный потенциал. В [18] был проведен скрининг всех мутаций вида D95X, наиболее удачными вариантами оказались D95E, D95H, D95N и D95Q (Рисунок 1.8).

В дальнейшем для поиска еще более ярких и чувствительных сенсоров был применен метод направленной эволюции, в результате чего удалось получить новую серию мутантов, характеризующихся красным спектром (626 нм) и квантовым выходом флуоресценции до  $0.5 \cdot 10^{-3}$ . При этом все белки из этой серии включают аминокислотные замены D95E и T99C [19]. Для двух белков (Arch-3 D95E/T99C, "Archer-1" и Arch-3 D95E/T99C/A225M "Archer-2") была показана зависимость флуоресценции от мембранного потенциала [20].

Другая серия сенсоров была получена на основе мутантов D95H и D95Q в сочетании с мутацией D106H [1]. Полученные белки были названы QuasAr1 и QuasAr2. Они имеют спектр поглощения около 580 нм.



Mutations	QY	$\epsilon$ , $M^{-1}\cdot cm^{-1}$	Brightness		Max emission, nm	Stokes shift, nm	Schiff base, $pK_a$
			$(QY \times \epsilon)/1,000$	$\lambda_{max}$ , nm			
Wild type	$1-9 \times 10^{-4}$ *	64,828	<0.059	556	687*	131	>10
D95E, T99C	$3.3 \times 10^{-3}$	38,436	0.12	626	731	103	7.5
D95E, T99C, P60L	$4.0 \times 10^{-3}$	52,828	0.21	624	731	105	7.7
D95E, T99C, P196S	$5.7 \times 10^{-3}$	44,116	0.22	628	731	105	6.8
D95E, T99C, V59A	$6.2 \times 10^{-3}$	68,622	0.43	622	728	105	7.9
Arch-5: D95E, T99C, V59A, P60L, P196S	$8.7 \times 10^{-3}$	65,126	0.57	622	731	107	7.6
Arch-7: D95E, T99C, V59A, P60L, P196S, D222S, A225C	$1.2 \times 10^{-2}$	104,769	1.26	616	727	113	8.4

Рисунок 1.9 — Серия мутантов Arch-3 на основе мутаций D95E/T99C, рисунок взят из [19]

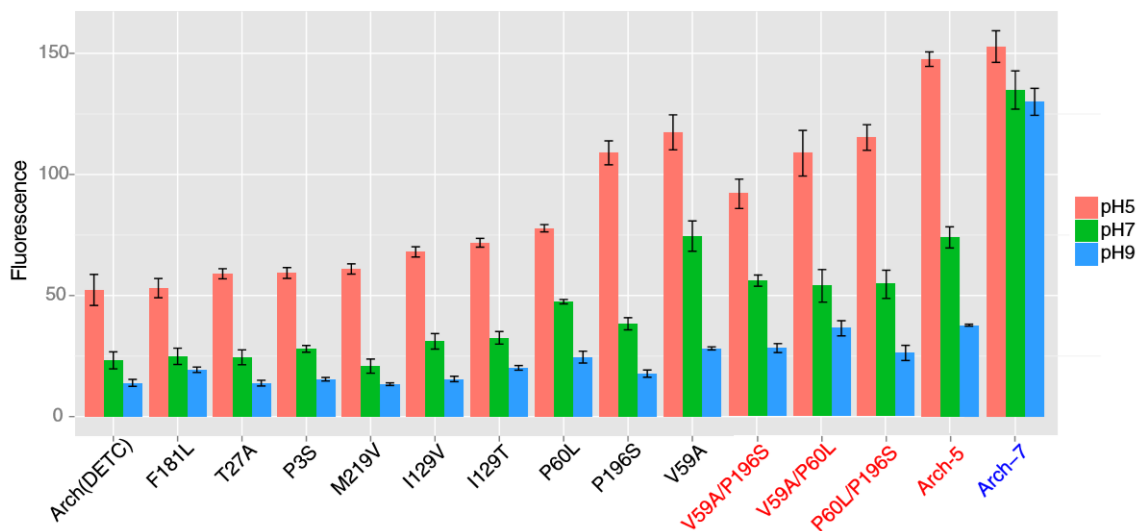


Рисунок 1.10 — Данные по флуоресценции мутантов из [19].

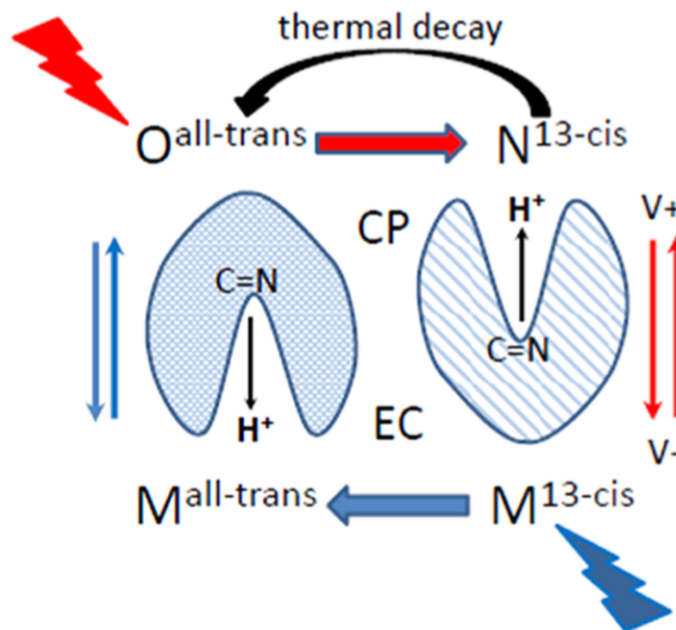


Рисунок 1.11 — Изменения равновесия между интермедиатами при изменении мембранного потенциала согласно [16].

### 1.6 Предложенные модели механизма потенциал-зависимости флуоресценции археородопсина-3.

На текущий момент механизм потенциал-зависимости исследован только для дикого типа археородопсина-3.

В работах [16; 21] предполагается, что потенциал-зависимость флуоресценции связана с соотношением N и M форм (Рисунок 1.11) [16] (протонированный и депротонированный хромофор соответственно), при этом сама флуоресценция может наблюдаться не у этих интермедиатов непосредственно, а у получающегося из после фотоактивации N-формы Q-интермедиата (Рисунок 1.12).

Возможность управляемого полем равновесия между протонированным и депротонированным хромофором также подтверждается в работе [22]. При рКа хромофора около 7, что наблюдалось в мутанте Arch-3 D95N (и BR D85N), хромофор может депротонироваться при приложении внешнего электрического поля, сопоставимого по своей величине с полем в мембране. При этом стоит отметить, что если рКа хромофора в диком типе arch-3 составляет

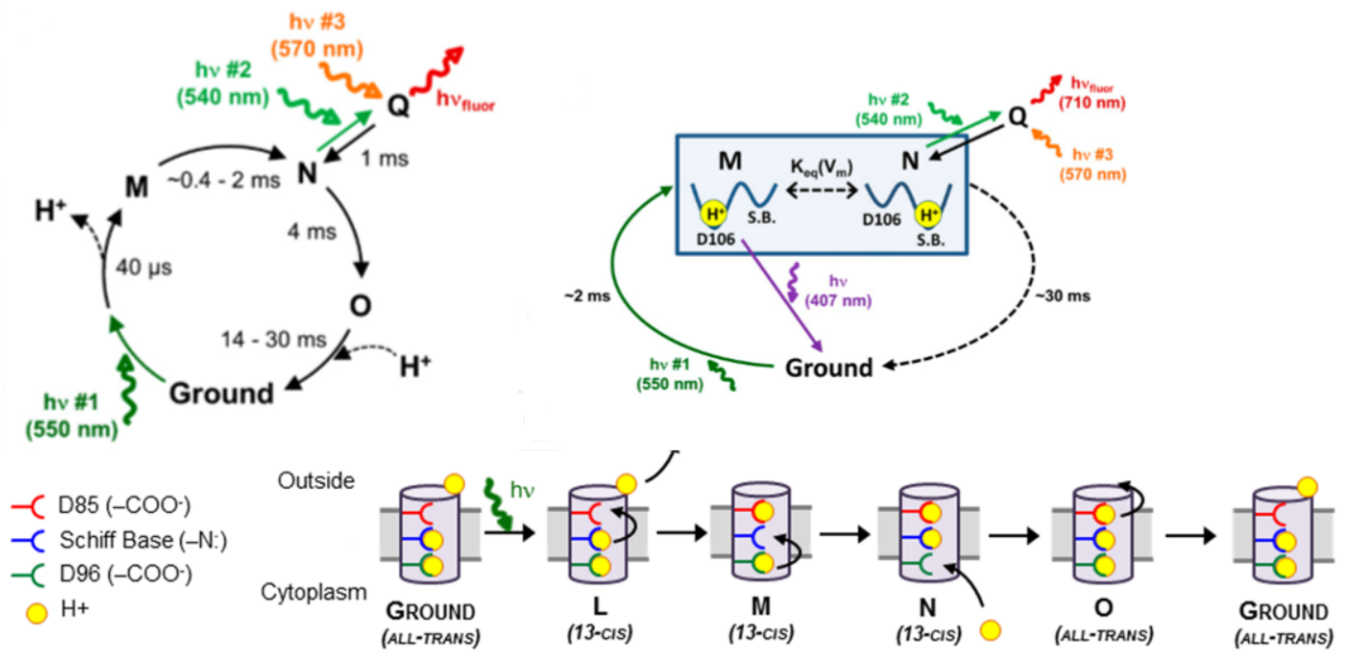


Рисунок 1.12 — Фотоцикл Arch-3, Q интермедиат и протонирования контрионов в интермедиатах согласно Maclaurin et al [21].

около 11, то во всех полученных мутантах с повышенной чувствительностью флуоресценции рКа обычно лежит в диапазоне 6-8 [18; 19].

Было выдвинута гипотеза, что флуоресценция возбуждается последовательным поглощением трех фотонов (Рисунок 1.12), первый из которых активирует фотоцикл, а второй предположительно фотоактивирует N интермедиат, об чем свидетельствует пик зависимости яркости флуоресценции от длины волны второго фотона на 550 нм и оптимальному промежутку между первым и вторым фотонами в 5 миллисекунд. Спектр возбуждения флуоресценции третьим фотоном характеризуется пиком на 580 нм, что соответствует предполагаемому пику поглощения Q интермедиата [21].

## 1.7 Модификация стабильности О-состояния в бактериородопсине путем введения аминокислотных замен.

Введение аминокислотных замен в белок может изменить соотношение интермедиатов в равновесном состоянии. Для ряда мутантных форм бактериородопсина на основе сходства красного сдвига поглощения и КР спектров было выдвинуто предположение, что они могут иметь состояние, похожее на О-интермедиат (О640) в качестве основного. В частности, были исследованы мутанты BR D85S [23–25], D85A [26], L93A [27; 28], Y185F[29].

BR L93A имеет долгоживущее О-подобное состояние, однако его спектр (565 нм) сильно отличается от интермедиата О640. D85S имеет спектр около 600 нм и возможно находится в О-подобном состоянии. Для обоих белков были получены кристаллографические структуры. Для мутанта BR Y185F было показано наличие О формы в светоадаптированном состоянии, при этом на основании данных оптической[29], КР[30] и ИК [31] спектроскопии с временным разрешением было выдвинуто предположение, что после поглощения красного фотона О640 транс-состояние, минуя короткоживущий интермедиат, переходит в N550 13-цис форму.

Кинетика О-состояния мутанта L93A имеет негауссову форму, в [27] было предположено наличие двух состояний с красным спектром, О1 и О2, имеющих хромофор в 13-цис- и транс-конформациях. В кристаллической структуре наблюдается 13-цис изомер. Однако на КР спектрах О-интермедиата BR видно преобладание транс формы. В кристаллической структуре мутанта D85S также наблюдается транс-изомер [23]. Наличие двух О-подобных состояний обсуждалось для *Acetabularia rhodopsin II* [32].

Вода не только сильно влияет на спектр белка [33], но также и играет важную роль в стабилизации интермедиатов, изомеризации хромофора [34] и переносе протона. При этом полосы О-Н колебаний различных молекул воды видны на ИК спектрах. С помощью введения различных аминокислотных замен в белок и снятия ИК спектров с полученных мутантов удалось частично понять

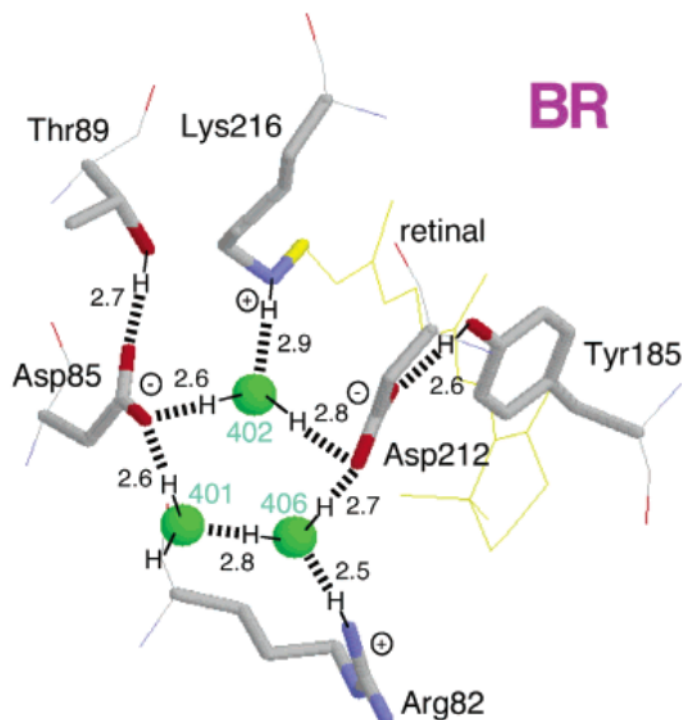


Рисунок 1.13 — Пентагональный кластер водородных связей с тремя молекулами воды в GS BR, рисунок взят из [37].

положение разных молекул воды в белке, так как в мутанте с аминокислотной заменой рядом с соответствующей молекулой воды ее полоса обычно сдвигается (или пропадает, если убирается сама молекула) [35; 36]. Три молекулы воды участвуют в образовании пентагонального кластера водородных связей с хромофором и двумя контр-ионами Asp95 и Asp212, что по-видимому приводит к стабилизации основного состояния белка с Asp95 в протонированном состоянии (Рисунок 1.13) [37]. Стоит отметить, что стабилизированное O-подобное состояние также наблюдалось для сенсорного родопсина II [38] и pharaonis Halorhodopsin [39].

## Глава 2. Методы

### 2.1 Экспериментальные методы

Мутации вносились в бактериальный вектор pET21b с геном археородопсина-3. Для сайт-направленного мутагенеза проводилась ПЦР с использованием полимеразы AccuPrime Pfx, праймеры для мутагенеза приведены в приложении А. Полученной плазмидой трансформировались клетки xl1 blue. Выделение и очистка плазмидной ДНК осуществлялось при помощи стандартного набора (Plasmid Miniprep, Евrogen) на силикатных колонках.

Каждый белок выделялся с 400 мл *E. Coli* в питательной среде LB с ампициллином, для экспрессии белка в среду добавлялся транс-ретинол (Helicon) до концентрации 5 мг/л и 40 мкл 1М IPTG (РосМедБио). Сначала проводилось центрифугирование среды в течение 10 минут на скорости 5000 об/мин, осадок лизировался ультразвуком в буфере 5 мМ MgCl<sub>2</sub>, 0.3 М NaCl, затем центрифугировался 30 минут при 14000 оборотов/минуту, полученный осадок растворялся в 1.5 процентном водном растворе детергента n-Decyl-D-Maltoside (DM), после чего белок очищался на Ni-NTA (Thermo) смоле по гистидиновому тэгу. Затем проводилась очистка белка от имидазола при помощи 0.1 процентного водного раствора DM на колонке с фильтром на 10 кДа (Amicon Ultra 0.5mL Filters, Merck).

Для оценки чистоты получаемых родопсинов от других белков использовался SDS-PAGE электрофорез на полиакриламидном геле.

Спектры комбинационного рассеяния снимались на экспресс-рамановском спектрометре SENTERRA с 785 нм лазером при мощности 100 мВт. Типичный объём капли для исследования составлял 10 мкл. Для наведения нужного рН к 9 мкл белка добавлялся 1 мкл 1М соответствующего буфера (сульфатного, фосфатного, трис-HCl и аммиачного). Концентрация белка определялась с помощью реакции Брэдфорда (набор Pierce BCA Protein Assay Kit).

Обработка полученных спектров, производилась в программе Origin 8.

## 2.2 Теоретические методы.

### 2.2.1 Получение трехмерных структур белков.

В качестве начальной трехмерной структуры транс- и 13-цис-форм археородопсина-3 в основном состоянии использовались кристаллографические структуры, полученные из базы данных RCSB (код 6GUX) [40]. Для предсказания структур мутантных форм в основном состоянии использовался метод сравнительного моделирования, в качестве шаблона использовались кристаллографические структуры археородопсина-3 дикого типа (13-цис форма для 13-цис изомеров, транс-форма для транс-изомеров). Для предсказания мутантных форм, а также археородопсина-3 дикого типа в О-состоянии также использовался метод сравнительного моделирования, в качестве шаблона использовалась кристаллографическая структура О-состояния мутантной формы бактериородопсина L93A [27]. Выравнивание аминокислотных последовательностей мутантной формы бактериородопсина и археородопсина-3 проводилось при помощи метода парного выравнивания, алгоритм AlignMe [41]. Предсказание структур проводилось при помощи программного пакета I-TASSER 5.1 [42].

Для начальных структур (предсказанных или кристаллографических) проводилась вставка хромофора, после чего проводилось предсказание наличия и положений молекул воды в полостях белка. Для предсказания положений молекул воды использовался алгоритм, основанный на оценке энергетической выгоды помещения молекулы воды в полость. Энергия молекулы воды в полости рассчитывается на основании оценки величины электростатического взаимодействия ее диполя с зарядами окружающих аминокислот, а также с дипольными моментами других молекул воды, помещенных в полость ранее (при наличии). Алгоритм реализован в программе Dowser++, программный пакет модифицировался для учета зарядов на хромофоре [43].

Для полученной молекулярной системы (родопсин + вода) проводилась оптимизация сначала на молекулярно-механическом уровне в силовом поле

Amber98, затем с использованием гибридной квантово-механической / молекулярно-механической (КМ/ММ) схемы ONIOM (B3LYP/6-31G\*//AMBER) [44]. В квантово-механическую часть включалось 50 атомов хромофора, линкерный атом располагался на связи NZ-CE лизина, связанного с ретиналем. Для оптимизации использовался программный пакет Gaussian09 [45].

### **2.2.2 Расчет спектральных свойств.**

Для расчета максимума полосы поглощения хромофора в окружении белка использовался КМ/ММ метод SORCI+Q/6-31G\*, реализованный в программном пакете ORCA 3.0.3 [46]. Электростатическое поле белка было представлено в виде соответствующего распределения точечных зарядов на атомах окружения хромофора (AMBER-98).

### **2.2.3 Расчет pKa аминокислот.**

Расчет pKa титруемых аминокислот белка проводился при помощи алгоритма, реализованного в программе PROPKA-3.1 [47]. Анализ проводился для оптимизированных моделей.

### **2.2.4 Расчет энергии Гиббса между состояниями белка.**

Для расчета энергии Гиббса между формами одного белка был построен термодинамический цикл [48].

Расчет разницы в энергии Гиббса между состояниями белка включал в себя следующие слагаемые:



1. Разница в свободной энергии между двумя изомерами хромофора, оптимизированными в газовой фазе (разница в энергии + энтропийный вклад, рассчитанный на основании расчета колебательной статсуммы,  $\Delta G(gas)$ ). Методы расчета: B3LYP/6-31G\*.
2. Разница в энергии между хромофором, оптимизированным в газовой фазе и в белке. Рассчитывался отдельно для двух изомеров ( $\Delta E_X(gc)$ ).
3. Свободная энергия помещения хромофора в белок рассчитывается отдельно для каждого изомера ( $\Delta G_{X \rightarrow 0}$ ). Для расчета данного вклада применялся стандартный подход возмущения свободной энергии, применяемый для расчета свободной энергии сольватации молекул в растворителе. Моделировался "алхимический" путь исчезновения хромофора из белка через серию промежуточных состояний, путем последовательного уменьшения энергии взаимодействия между хромофором и белковым окружением. Начальное состояние – стандартное взаимодействие между хромофором и окружением – характеризуется стандартными зарядами на атомах хромофора, а также стандартными параметрами для Ван-дер-Ваальсовых взаимодействий. Конечное состояние характеризуется нулевыми зарядами на атомах хромофора, а также отсутствием ван дер Ваальсовых взаимодействий между атомами хромофора и окружением. Путь между состояниями включал 19 промежуточных состояний, на начальном этапе проводилось уменьшение кулоновских взаимодействий (масштабирование зарядов на атомах хромофора, коэффициенты масштабирования 0.875, 0.75, 0.625, 0.5, 0.375, 0.25, 0.125, 0.0), затем проводилось масштабирование взаимодействий Ван-дер-Ваальса по аналогичной схеме. Для каждого состояния проводилась генерация больцмановского распределения конфигураций путем расчета молекулярной динамики при 298 К. Расчет молекулярной динамики проводился в программном пакете Gromacs 2019.1. Расчет проводился на молекулярно-механическом уровне в силовом поле Amber98. Алгоритм проводил расчет разницы в энергии между соседними состояниями для каждого шага молекулярной динамики. Обработка данных для расчета свободной энергии проводилась при помощи алгоритма BAR (Bennett Acceptance Ratio).

4. Итоговое значение  $\Delta G_{A \rightarrow B} = \Delta G_{A \rightarrow B}(gas) - \Delta G_{B \rightarrow 0} + \Delta E_B(gc) + \Delta G_{A \rightarrow 0} - \Delta E_A(gc)$  рассчитывается на основании вышеупомянутых вкладов.

## Глава 3. Результаты и обсуждение

### 3.1 Анализ экспериментальных результатов.

Для археородопсина-3 дикого типа и всех мутантных форм археородопсина-3 были измерены спектры комбинационного рассеяния при длине волны облучения 785 нм (Рисунок 3.1 - 3.6) и 532 нм (Рисунок 3.7, 3.8). Также были измерены электронные спектры поглощения при различных значениях рН (3.9-3.12).

У всех мутантных форм, содержащих аминокислотные замены D95E и T99C, наблюдается сдвиг пика, соответствующего скелетным колебаниям хромофора (пик ВЛА), в красную область с  $1530\text{ см}^{-1}$  до  $1509\text{-}1515\text{ см}^{-1}$ . При этом при сопоставлении КР спектров мутантов AR5, AR7, D95E/T99C/P60L со спектрами интермедиатов бактериородопсина из работы [7] обнаруживается их сходство с О интермедиатом (Рисунок 1.5): ВЛА  $1508, 1509\text{ см}^{-1}$ , регион характерных пиков состоит из более интенсивной полосы на  $1170\text{ см}^{-1}$  и менее интенсивной на  $1201\text{ см}^{-1}$ , пик 13-цис формы на  $1185\text{ см}^{-1}$  проявляется слабо.

Оптические спектры имеют несимметричный пик с максимумом около 626 нм. При разложении спектров по гауссианам выявляется вторая полоса на 540-560 нм, что близко к спектру основного состояния или N интермедиата фотоцикла (550 нм). Однако вклад второй полосы не превышает 25 процентов. Кроме того, нет доказательств, что этот пик не является вторым пиком колебательной структуры.

У мутантов D95E/T99C, D95E/T99C/P196S пик ВЛА несимметричен и может быть разложен на две гауссовы компоненты на  $1528\text{ см}^{-1}$  и  $1515\text{ см}^{-1}$ . При этом при нейтральном рН хромофор преобладает в цис состоянии, так в регионе характерных пиков наблюдается наибольший пик  $1185\text{ см}^{-1}$ . На оптическом спектре при тех же рН вклад полосы на 550 нм также порядка 20 процентов. Таким образом, можно предположить, что в этих мутантах

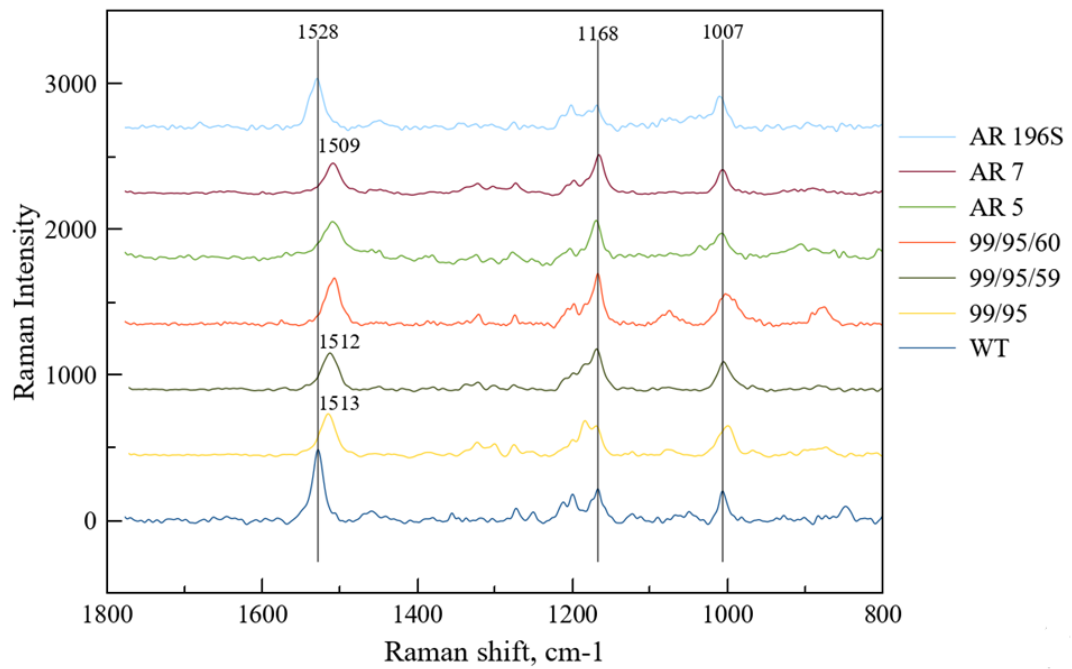


Рисунок 3.1 — Сравнение КР спектров всех мутантов при pH=6, лазер 785 нм

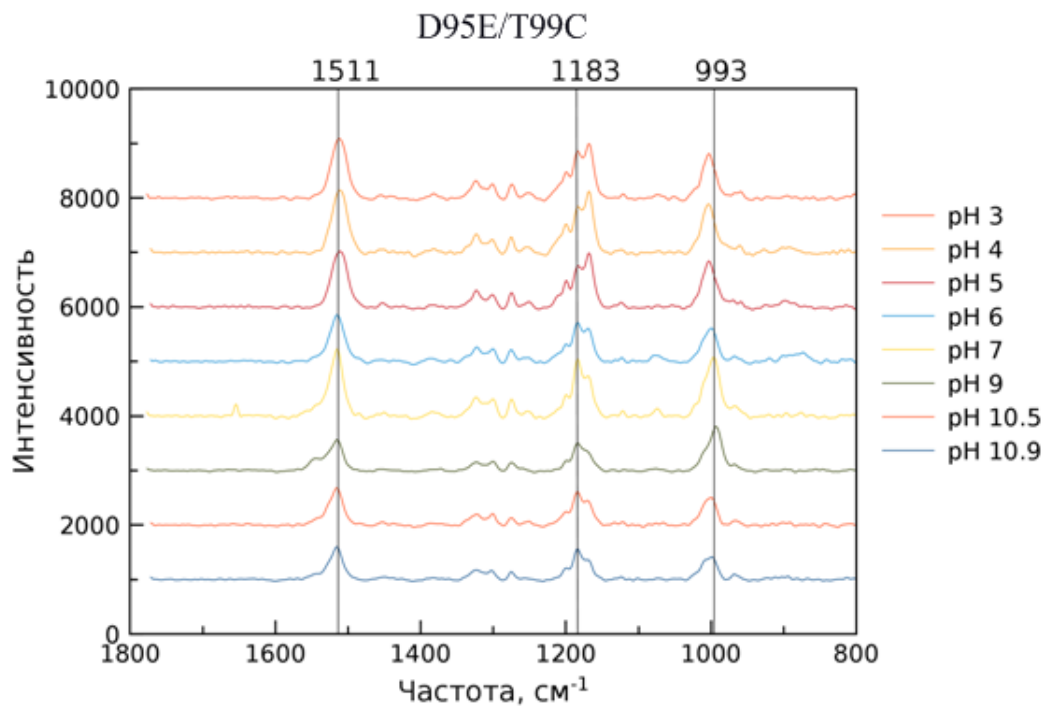


Рисунок 3.2 — КР спектры D95E/T99C, лазер 785 нм

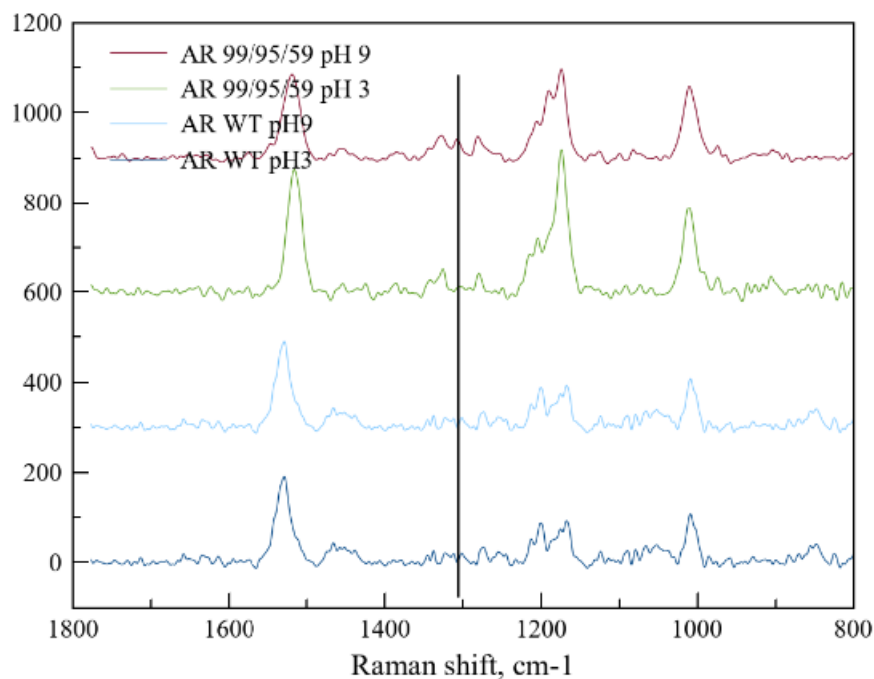


Рисунок 3.3 — КР спектры D95E/T99C/V59A в сравнении с диким типом, лазер 785 нм

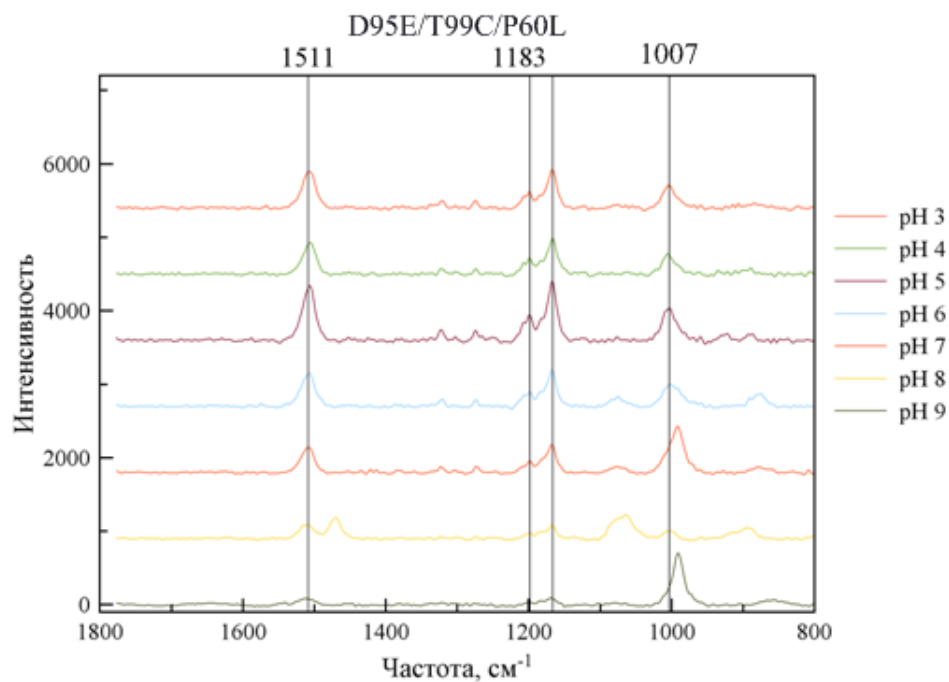


Рисунок 3.4 — КР спектры D95E/T99C/P60L, лазер 785 нм

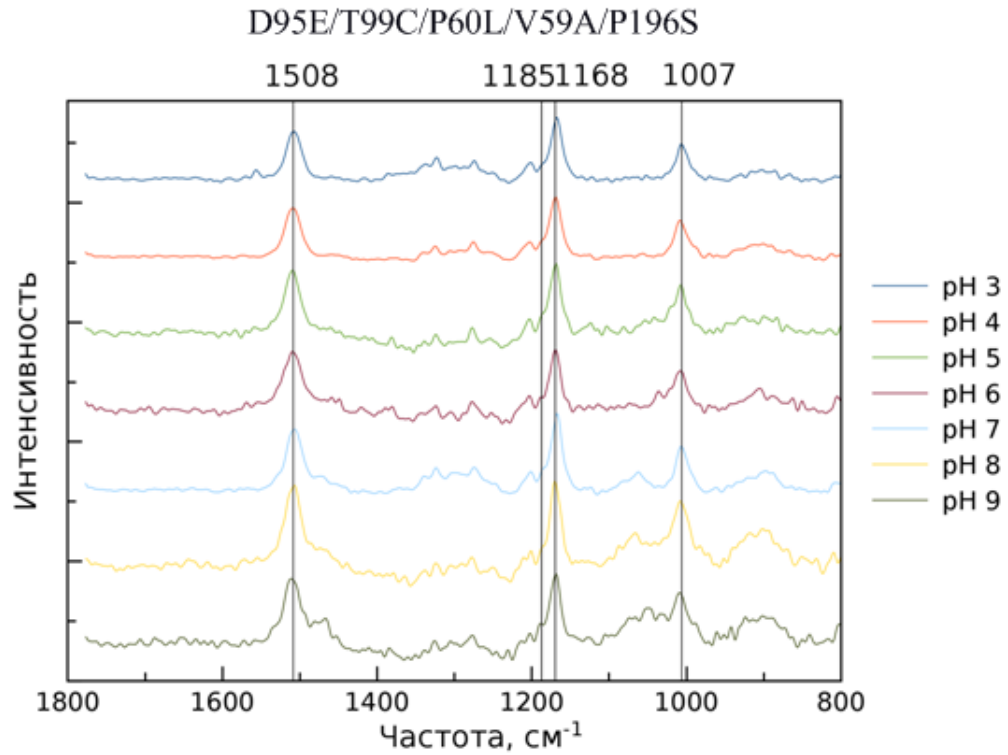


Рисунок 3.5 — КР спектры AR5, лазер 785 нм

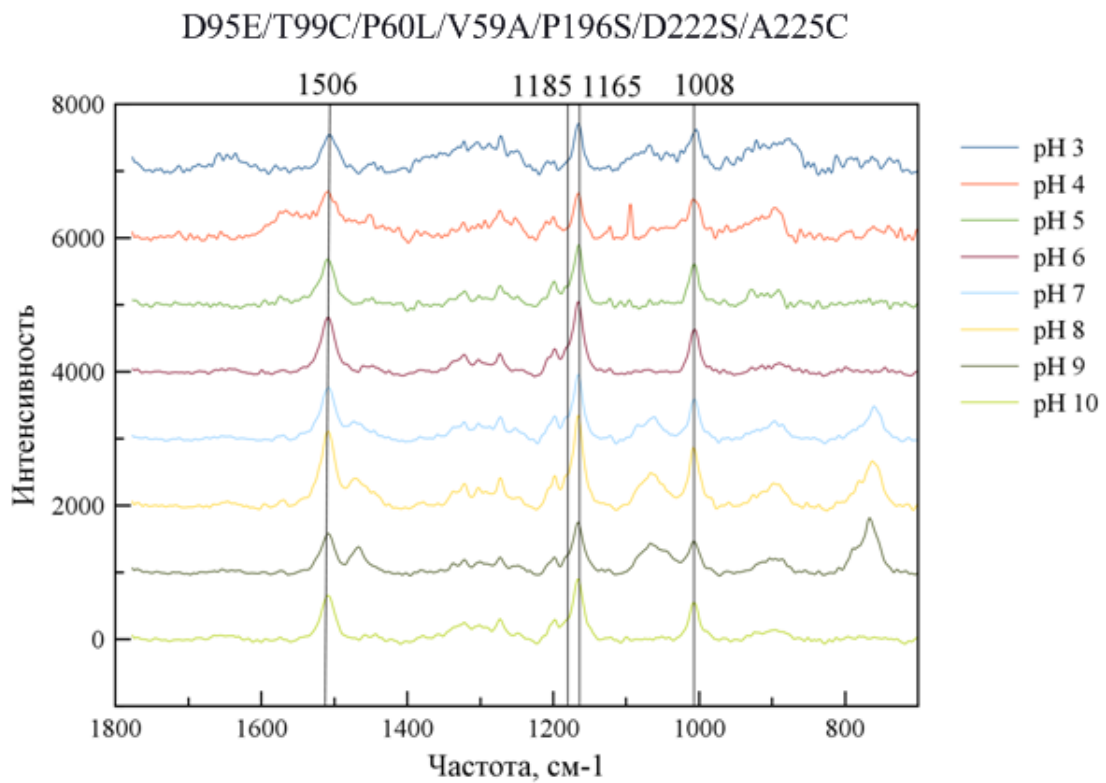


Рисунок 3.6 — КР спектры AR7, лазер 785 нм

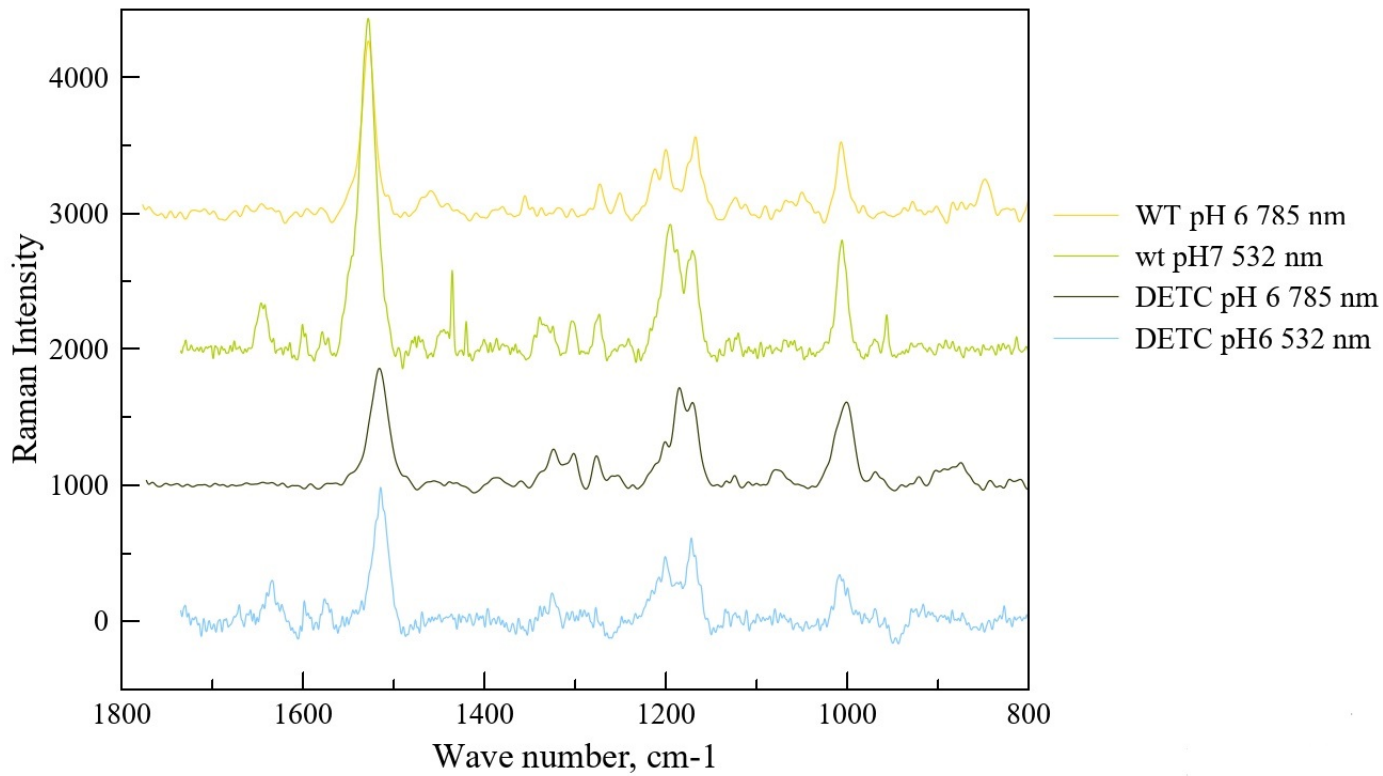


Рисунок 3.7 — Сравнение КР спектров дикого типа (WT) и Arch DETC при возбуждении лазером 532 нм и 785 нм

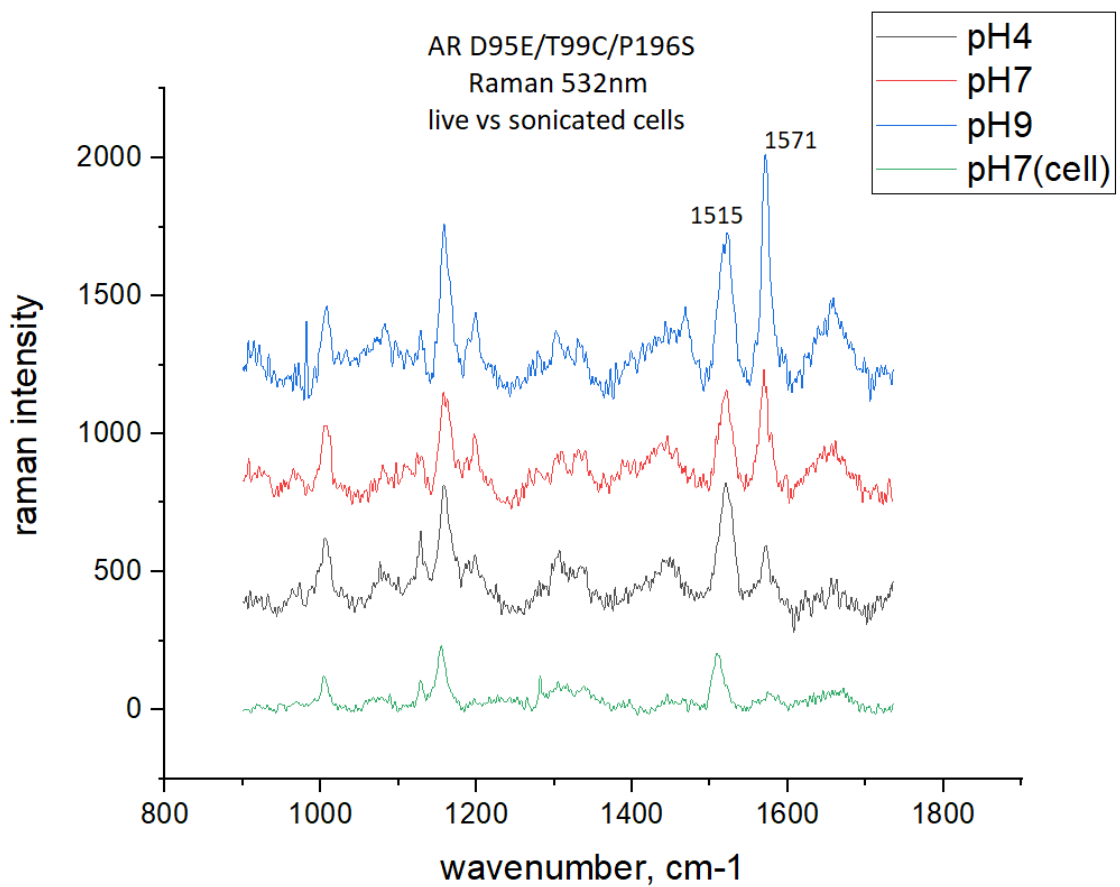


Рисунок 3.8 — КР спектр AR 99/95/196, лазер 532 нм, сравнение живых и обработанных ультразвуком клеток

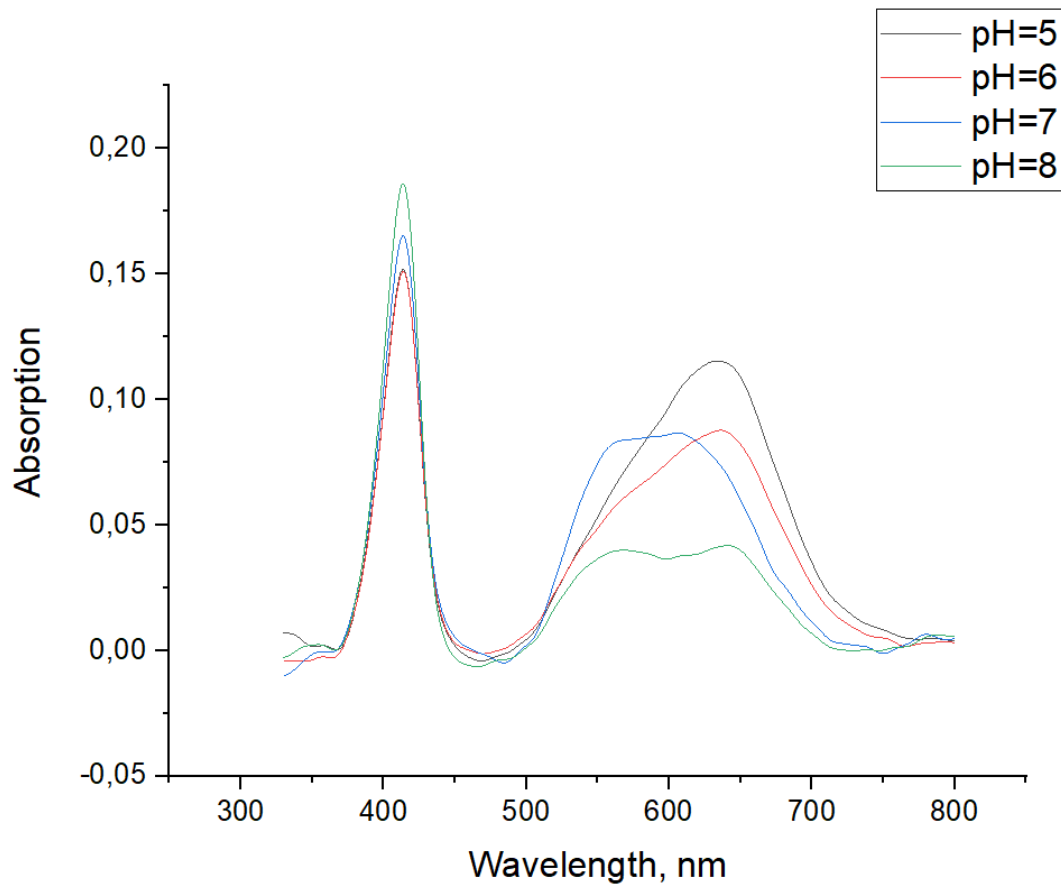


Рисунок 3.9 — Оптическое титрование AR D95E/T99C/P196S.

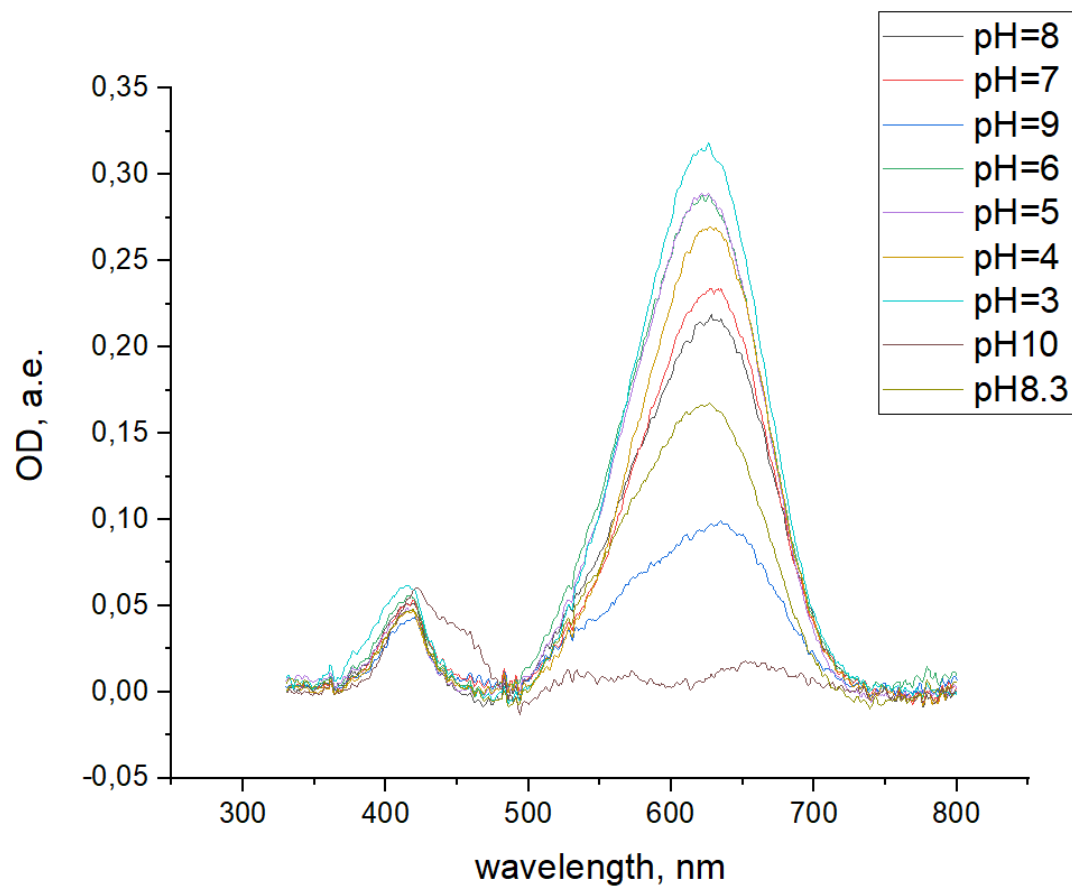


Рисунок 3.10 — Оптическое титрование AR D95E/T99C/V59A.



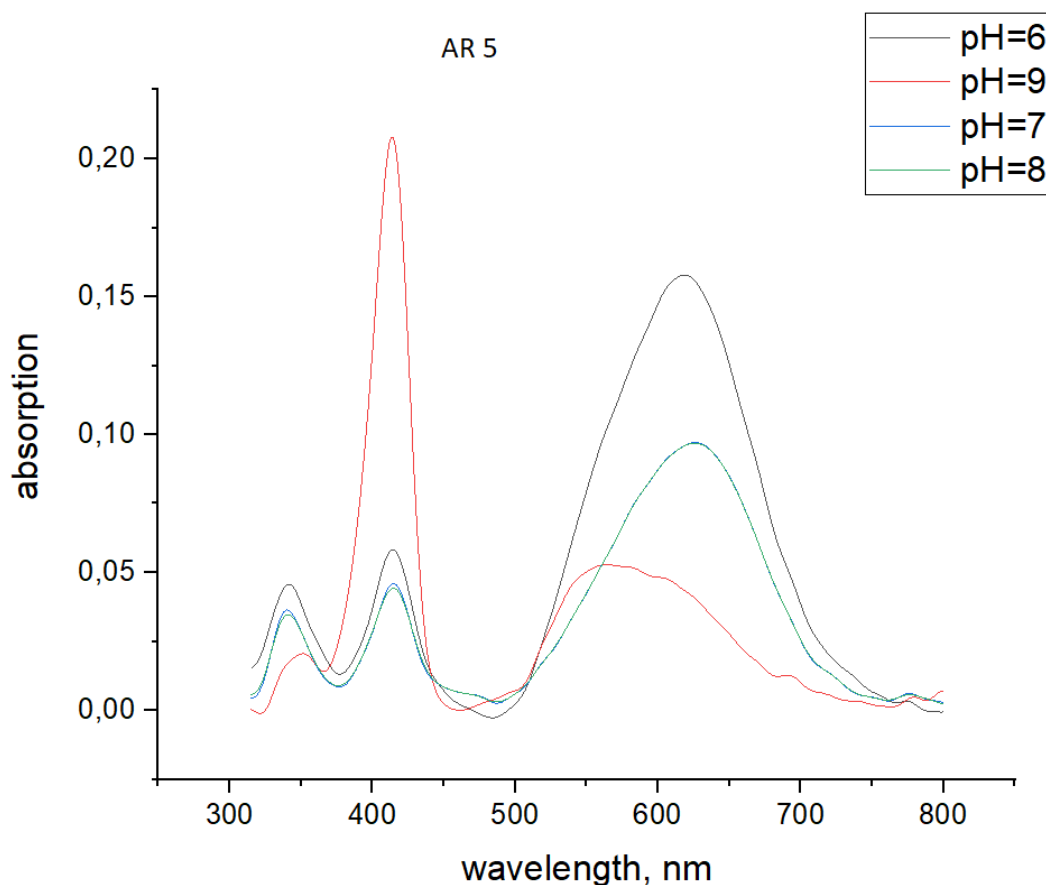


Рисунок 3.11 — Оптическое титрование AR5 (D95E/T99C/P196S/V59A/P60L).

наблюдается смесь различных состояний, причем основной вклад дают два состояния со спектром поглощения около 625 нм и находящейся в цис и транс форме, далее будем называть их О-цис и О-транс соответственно.

Так как само облучение белка лазером в ходе эксперимента может воздействовать на соотношение интермедиатов, для исследования этого фактора были измерены КР спектры не только с инфракрасным (785 нм), но и с зеленым (532 нм) лазером. Основным наблюдаемым отличием (Рисунок 3.7) стало соотношение цис- и транс-форм в этих случаях, при этом использование красного лазера приводит к накоплению состояния с хромофором в цис-форме. Таким образом, равновесие в ходе эксперимента нельзя считать исключительно термодинамическим. Данный вопрос также рассматривался в [16; 49].

Для мутанта D95N также было показано, что при уменьшении мощности облучения 785 нм лазера пик 13-цис состояния уменьшается, спектр становится более похожим на О-состояние. Это может быть объяснено появлением реи-

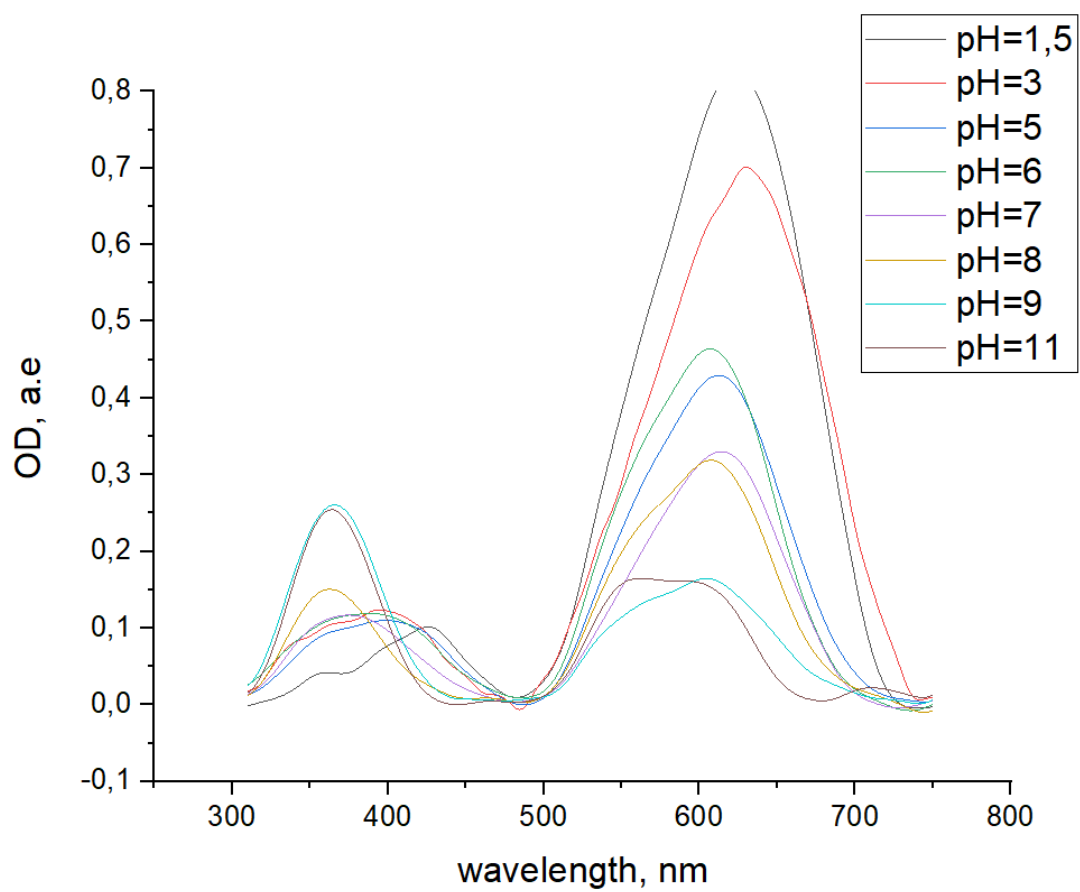


Рисунок 3.12 — Оптическое титрование AR5  
(D95E/T99C/P196S/V59A/P60L/A225M/D222S).

зомеризацией в цис интермедиата при облучении транс хромофора красным светом. В случае BR Y185F [29] полученный интермедиат интерпретируется как N (13-цис), однако для дикого типа есть сведения, что поглощение O формой красного фотона приводит к появлению 9-цис состояния [50].

У белка AR D95E/T99C/P196S при нейтральном pH наблюдается двойной пик поглощения со второй полосой 550 нм (Рисунок 3.9), рКа перехода составляет около 7. При этом на КР спектре появляется второй сильный пик ВЛА на  $1571 \text{ см}^{-1}$ , растущий с повышением pH. Также было обнаружено, что в живых не обработанных ультразвуком клетках этот пик почти исчезает (Рисунок 3.8).

На КР спектрах также наблюдается сдвиг пика C15-N колебания с  $1642(\text{WT})$  до  $1633\text{-}1635\text{см}^{-1}$  (мутантные белки). Данный сдвиг может быть объяснен исчезновением водородной связи с водой из пентагонального кластера рядом с хромофором и связанного C15 водорода. Подобные сдвиги также наблюдались в O подобном состоянии для мутанта BR D85A[26] При этом пика на  $1645\text{см}^{-1}$ , соответствующего N и L интермедиатам[9], на полученных спектрах не наблюдается.

## Оптические спектры

На рис. 3.9 - 3.12 представлены оптические спектры поглощения разных мутантов при различных pH. На спектрах видны три основных пика: состояние с депротонированным хромофором (пик 420 нм), главный пик 622-628 нм, а также пик на 550 нм.

У всех белков, кроме AR D95E/T99C/P196S полоса поглощения на 550 нм начинает появляться только начиная от pH 9, что свидетельствует о высоком рКа обоих контр-ионов (Asp95, Asp212) в этих мутантах.

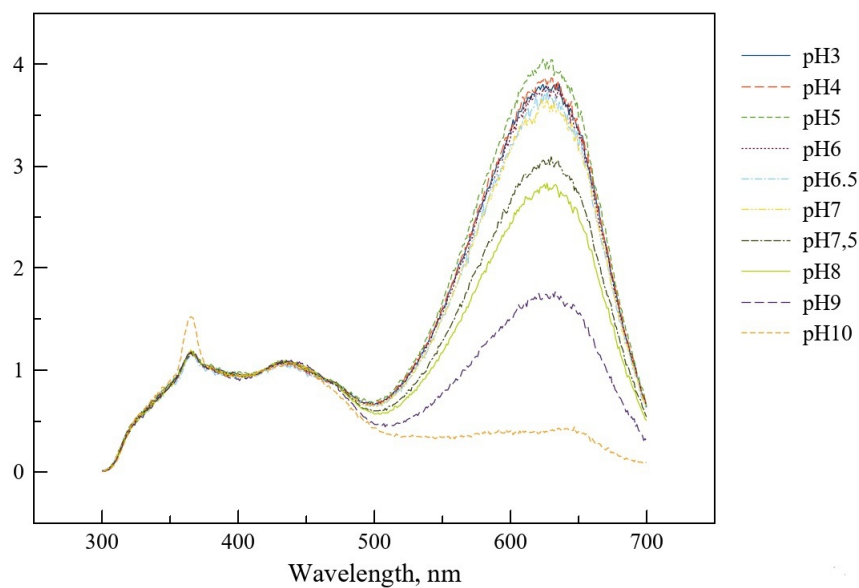


Рисунок 3.13 — Титрование спектров возбуждения флуоресценции для AR D95E/T99C. Пик на 365 нм - аппаратный, связан с порядком решетки при сканировании на 730нм

### Спектры флуоресценции и возбуждения флуоресценции

Были отсняты спектры флуоресценции мутантов при возбуждении на 420, 550 и 625 нм, а также спектры возбуждения флуоресценции при регистрации интенсивности сигнала на 730 нм при различных рН. Яркая флуоресценция наблюдалась только при возбуждении на длине волны 620-630 нм.

На основании того, что спектр возбуждения флуоресценции совпадает со спектром поглощения (пик поглощения 626 нм, разложение по гауссианам показало только слабую примесь вклада на 550 нм (не более 20% при физиологическом рН)), можно предположить, что основным флуоресцентным состоянием становится О-цис или О-транс, а не Q интермедиат (имеющий пик поглощения около 570 нм). Также модель флуоресценции из Q интермедиата плохо согласуется с тем, что в белках DETC и DETC/A225M концентрация цис-формы, интерпретируемой как N интермедиат, выше, а флуоресценция у данных белков менее яркая, чем у AR5 и AR7, где цис-форма почти не наблюдается даже при облучении красным лазером. Следует отметить, что данный вывод относится к мутантам на основе DETC, но не к мутантам типа D95X, не исследованным в данной работе.

Белок	pKa флуоресценции	pKa 550/626	pKa SB
D95E/T99C	7.8	>10	7.5
D95E/T99C/P60L	8.0	>10	7.7
D95E/T99C/P196S	7.0	7.0	6.8
AR5	8.0	8.5	7.6
AR7	8.6	10.5	8.4

Рисунок 3.14 — pKa перехода 550-620 на спектрах поглощения, pKa флуоресценции и pKa основания Шиффа(SB), последние взяты из [19]

При этом pKa флуоресценции(около 7-8 во всех мутантах) оказывается близок к pKa основания Шиффа, что свидетельствует о процессе депротонирования хромофора как причине pH-зависимости флуоресценции(прежде всего из-за того, белок перестает поглощать на длине волны возбуждения, однако и собственная флуоресценция хромофора при возбуждении 550 и 420 нм крайне слабая), а депротонирование Asp222, которое также должно приводить к падению флуоресценции, происходит при несколько большем pH, однако возможна ситуация, когда депротонирование происходит в обратном порядке. Экспериментальные данные pKa сведены в таблицу, рисунок 3.14. pKa перехода 550-620 нм были получены путем аппроксимации сигмоидальной функцией соотношения величин поглощений соответствующих пиков после разложения полосы по двум гауссианам в программе Origin 8.0

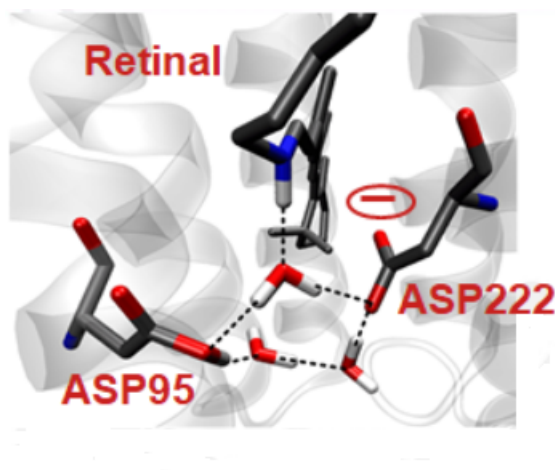


Рисунок 3.15 — Пентагональный кластер в основном состоянии Arch-3

### 3.2 Анализ результатов компьютерного моделирования.

На основании экспериментальных данных было выдвинуто предположение, что в мутантных формах археородопсина-3, характеризующихся высокой интенсивностью флуоресцентного сигнала, основной формой становится либо О-цис, либо О-транс состояние. Для каждого из исследованных белков были сконструированы компьютерные модели четырех состояний белка: основного состояния с хромофором в 13-цис и транс-форме; О-состояния с хромофором в 13-цис и транс-форме. Рассмотренные конформации белка были изучены в бактериородопсине [51] и характеризуются следующими особенностями.

Основное состояние (Рисунок 3.13, 3.14):

- кластер воды возле контрионов (Asp95 и Asp222 для археородопсина-3);
- депротонированное состояние контриона Asp222;
- наличие водородной связи между Glu204-Glu214 (ионный замок).

О-состояние (Рисунок 3.15, 3.16):

- отсутствие кластера воды возле контрионов;
- протонированное состояние контриона Asp222;
- ”открытое” состояние ионного замка Glu204-Glu214.

Согласно многим экспериментальным работам, основной контрион Asp95 считается протонированным как в N, так и в O состоянии [21], так как на

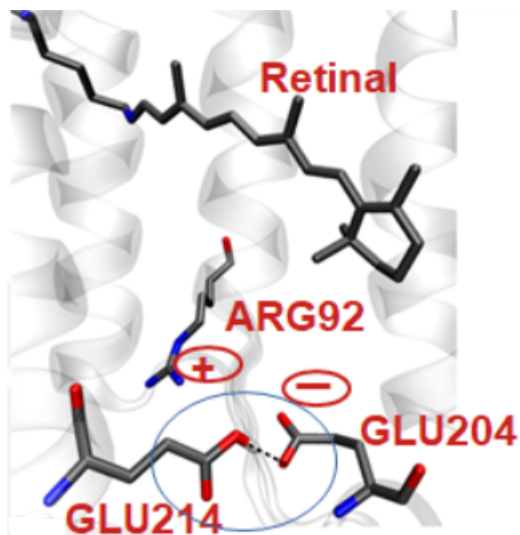


Рисунок 3.16 — Ионный замок в основном состоянии Arch-3

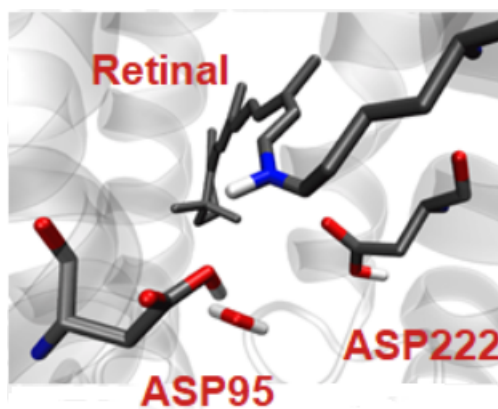


Рисунок 3.17 — Разрушение кластера воды в O форме Arch-3

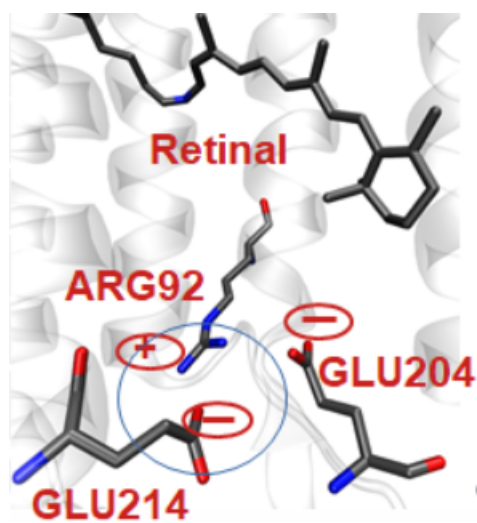


Рисунок 3.18 — Ионный замок в O форме Arch-3

ИК спектрах пик COOH группы Asp95 в этих интермедиатах одинаков [9]. Однако такие состояния протонирования не объясняют разницу в спектрах N(550) и O(640) форм. В нашем моделировании спектры удалось воспроизвести только в предположении, что в N форме основной контрион депротонирован. В основном состоянии в диком типе с тем же спектром поглощения (550 нм) рКа перехода в голубую форму(600нм), интерпретируемую как состояние с протонированным контрионом Asp95, составляет 2.5, т.е при физиологическом рН он депротонирован [49].

Относительная концентрация O состояния и, соответственно, квантовый выход флуоресценции определяется состоянием протонирования контриона Asp222. Значение рКа Asp222 и, соответственно, концентрация молекул с Asp222 в протонированном состоянии, регулируется мембранным потенциалом. Таким образом, мутантные формы, в которых рКа Asp222 находится в диапазоне 6.5-7.5 должны обладать наиболее сильной зависимостью интенсивности флуоресценции от потенциала при физиологических условиях, что наблюдается экспериментально. Для доказательства предложенного механизма были выполнены расчеты рКа контриона Asp222 и глутаминовых кислот, образующих ионный замок (Рисунок 3.18).

Показано, что рассчитанные значения максимумов поглощения для транс- и цис-состояний белка дикого типа согласуются с экспериментальными. Для мутантных форм согласие рассчитанных значений максимумов поглощения с экспериментальными получено только для моделей белков в O-состоянии с протонированным состоянием контриона Asp222(рис. 3.19). Согласно рассчитанным данным, в O-состояниях мутантных форм археородпосина-3, изученных в данной работе, рКа контриона Asp222 находятся в диапазоне 7.5 - 9.5, что согласуется с полученными экспериментальными данными по титрованию оптических спектров белков(появление полосы поглощения 550 нм при рН > 9, стоит отметить, что при депротонированном хромофоре эта рКа повышается еще на 1-2, что согласуется с результатами титрований). Для второго контриона Asp95 рКа была больше 10 во всех случаях.

При этом значения рКа контриона превышают значения рКа аминокислот, входящих в ионный замок, что также согласуется с тем, что в основном



Белок/Форма хромофор	pKa и теор. $\lambda_{\max}$	GS-trans	GS-cis	O-trans	O-cis	$\lambda_{\max}$ , эксп
Археородопсин-3 дикий тип(WT)	pKa Asp222: pKa Lock1: pKa Lock2: $\lambda_{\max}$ , nm	3.46 3.34 11.16 565	4.45 3.96 10.69 557	4.13 3.46 7.87 607	4.67 6.78 3.44 605	556
D95E/T99C	pKa Asp222: pKa Lock1: pKa Lock2: $\lambda_{\max}$ , nm	3.71 10.41 3.15 574	4,41 11.32 2.94 561	8.04 7.68 5.83 622	8.30 6.63 7.26 610	626
D95E/T99C/P60L	pKa Asp222: pKa Lock1: pKa Lock2: $\lambda_{\max}$ , nm	2.95 9.24 3.37 568	3.36 9.47 3.47 576	7.68 7.12 5.28 618	8.17 7.02 5.31 608	624
D95E/T99C/V59A	pKa Asp222: pKa Lock1: pKa Lock2: $\lambda_{\max}$ , nm	3.81 9.58 3.96 561	4.79 9.78 3.73 575	7.85 8.02 4.57 635	9.27 7.26 4.34 620	622
D95E/T99C/P196S	pKa Asp222: pKa Lock1: pKa Lock2: $\lambda_{\max}$ , nm	3.54 10.35 3.78 532	4.07 11.05 2.91 521	7.81 7.61 4.59 627	8.55 8.16 5.32 621	628
AR5: D95E/T99C/P60L/ V59A/P196S	pKa Asp222: pKa Lock1: pKa Lock2: $\lambda_{\max}$ , nm	4.22 9.99 4.13 556	4.36 9.21 3.67 575	6.17 5.90 4.41 617	8.38 6.58 4.99 613	622
AR7 D95E/T99C/P60L/ V59A/P196S/ D222S/A225C	pKa Asp222: pKa Lock1: pKa Lock2: $\lambda_{\max}$ , nm	- - - 567	- - - 567	- - - 628	- - - 610	616

Рисунок 3.19 — Результаты компьютерного моделирования основных состояний (13-цис и транс) и O-состояний (13-цис и транс) рассмотренных в работе мутантных форм археородопсина-3.



Рисунок 3.20 — Рассчитанные спектры разных состояний мутантов и экспериментальные данные

состоянии мутантных форм археородопсина-3 наблюдается O-подобное состояние, характеризуемой "открытым" состоянием ионного замка Glu204-Glu214 и протонированным состоянием контриона Asp222.

Согласно анализу полученных моделей, основными факторами, стабилизирующими O-состояние по отношению к "основному" состоянию являются:

- разрушение кластера молекул воды вблизи контрионов;
- реорганизация водородных связей белка.

Для каждого белка был проведен расчет разниц в энергии Гиббса между четырьмя изученными формами белка. Было показано, что для белка дикого типа O-состояние характеризуется более высоким значением энергии Гиббса, в то время как для мутантных форм белка "основное" состояние характеризуется более высоким значением энергии Гиббса. На Рисунке 3.20 приведена полученная в результате расчетов корреляция между рассчитанной разницей в энергии Гиббса между транс-изомерами белка (O-состояние минус основное состояние) и экспериментально измеренным квантовым выходом флуоресценции [19]. Таким образом, результаты моделирования подтверждают гипотезу о стабилизации O-состояния в флуоресцентных мутантных формах археородопсина-3, а также о возникновении флуоресцентного состояния белка из O-состояния.

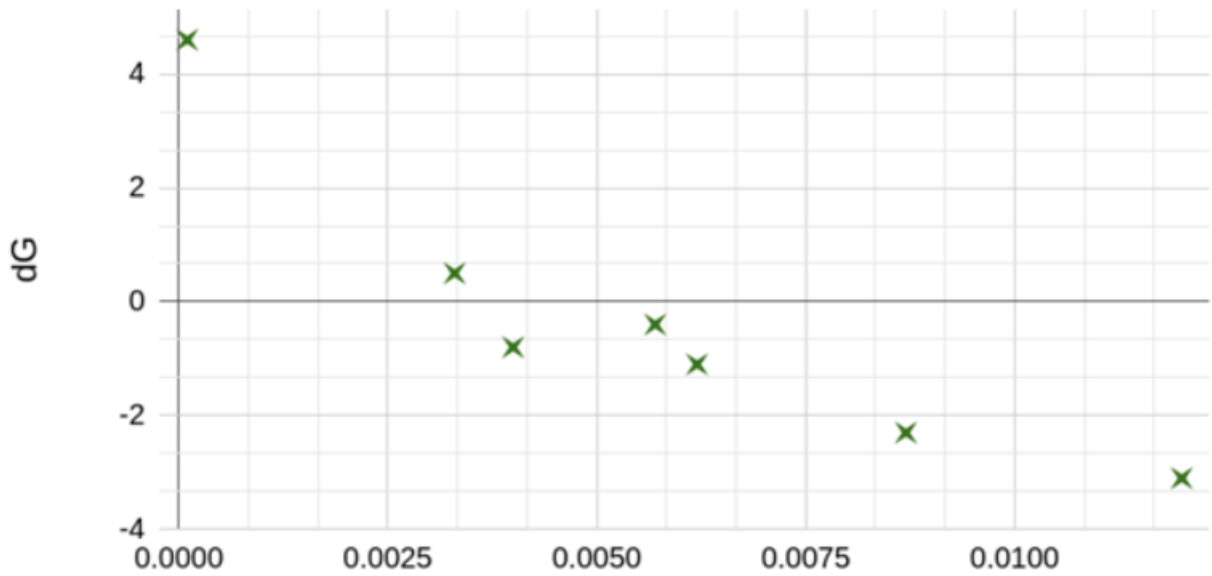


Рисунок 3.21 — Корреляция между рассчитанной разницей в энергии Гиббса между транс-изомерами белка (O-состояние минус основное состояние) и экспериментально измеренным квантовым выходом флуоресценции.

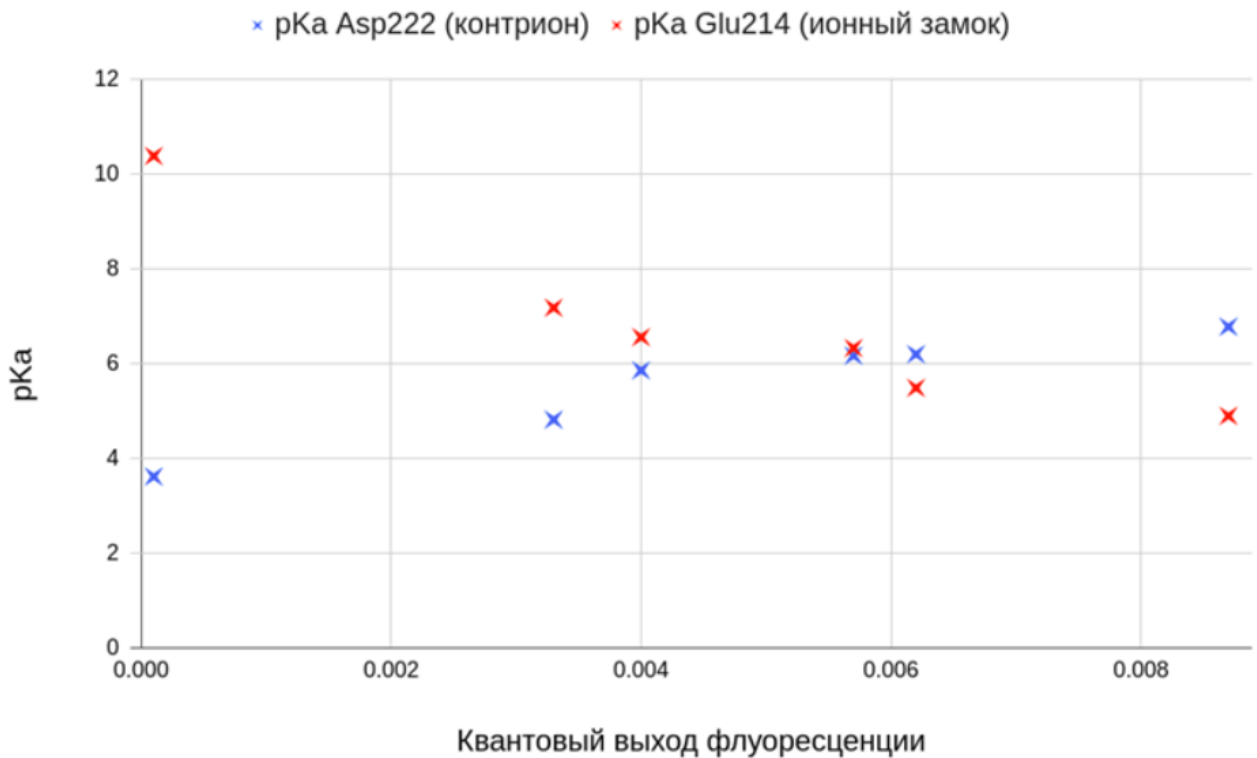
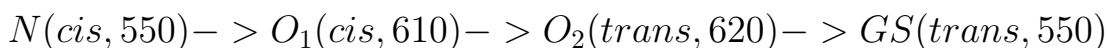


Рисунок 3.22 — Корреляция pKa контриона Asp222 и квантового выхода флуоресценции

Исходя из вышеуказанных фактов, вероятен такой вид переходов:



при этом основным состоянием является  $O_2$ (транс), под действием красного света оно переходит в  $O_1$ (цис), которое может быть флуоресцентным состоянием. Эта схема также снимает противоречие, что если флуоресцентный  $Q$  интермедиат получается из  $N$ , то при высоком рН флуоресценция должна быть больше (в эксперименте наблюдается обратная ситуация). При низком рН общее количество цис-формы падает, но это может быть за счет  $N$  интермедиата, а соотношение  $O$ -цис может расти.

На основании полученных экспериментальных и рассчитанных данных был сделан вывод, что для повышения концентрации  $O$ -формы необходимо вводит аминокислотные замены, которые либо приводят к повышению значения рКа констриона Asp222, либо приводят к понижению рКа аминокислот Glu204, Glu214, либо приводят к обоим процессам. В частности, в исследуемых мутантах основные две мутации T99C и D95E приводят к разрушению пентагонального кластера путем разрушения водородной связи и выталкиванию воды из-за большей длины глутамина соответственно.

Был проведен вычислительный скрининг ряда мутантных форм белка, в частности предложены новые мутации, приводящие к:

- разрушению кластера молекул воды (Y195F);
- изменению сети водородных связей (T215A; T203A)

Было показано, что наиболее высокое значение рКа констриона Asp222 достигается при введении четырех аминокислотных замен Tyr67Ala/Trp96Ala/Tyr195Ala/Thr204Ala.

На сегодняшний день проводится экспериментальная проверка результатов теоретического моделирования.

## Выводы

Основные результаты работы заключаются в следующем.

1. На основе анализа экспериментально измеренных спектров комбинационного рассеяния и электронных спектров поглощения археородопсина-3 дикого типа и его мутантных форм, характеризующихся высокой интенсивностью флуоресценции, а также на основании анализа результатов компьютерного моделирования различных состояний археородопсина-3 дикого типа и его мутантных форм, была выдвинута гипотеза о стабилизации О-состояния в мутантных формах археородопсина-3, характеризующихся высокой интенсивностью флуоресценции. В рассмотренных мутантных формах белка О-состояние становится основным состоянием фотоцикла.
2. На основе анализа спектров возбуждения флуоресценции археородопсина-3 и его мутантных форм была выдвинута гипотеза, что в рассмотренных мутантных формах археородопсина-3, характеризующихся высокой интенсивностью флуоресценции, флуоресценция возникает при фотовозбуждении основного(О) состояния белка.
3. Результаты компьютерного моделирования показали, что наблюдается корреляция между стабильностью О-состояния белка и экспериментально измеренным квантовым выходом флуоресценции белка.
4. На основании анализа результатов компьютерного моделирования и экспериментальных данных было выдвинуто предположение, что мембранный потенциал клетки влияет на концентрацию О-состояния белка, которое определяется протонированным состоянием хромофора, контриона Asp222 и депротонированным состоянием аминокислот Glu204 и Glu214, входящих в ионный замок. Таким образом, зависимость интенсивности флуоресценции археородопсина-3 и его производных от потенциала объясняется тем, что мембранный потенциал влияет на концентрацию О-состояния, при фотоактивации которого возникает флуоресцентный сигнал. Экспериментальные данные показывают, что наиболее сильной зависимостью интенсивности флуоресценции от мембранного потенциала обладают белки, для ко-

торых значение  $pK_a$  Asp222 и хромофора близко к физиологическому  $pH=7.2$ , что также подтверждает гипотезу.

Также на основании результатов компьютерного моделирования ряда производных археородопсина-3 были предложены аминокислотные замены, которые приводят к повышению  $pK_a$  контриона Asp222 и, согласно гипотезе, к повышению концентрации O-состояния и повышению интенсивности флуоресценции. Дальнейшим направлением исследований является более детальное изучение механизма флуоресценции археородопсина-3 и его производных при помощи методов спектроскопии с высоким временным разрешением.

## Благодарности

Исследование выполнено на базе ресурсных центров СПбГУ "Оптические и лазерные методы исследования вещества", "Вычислительный центр".

Автор выражает благодарность Николаеву Дмитрию Михайловичу и Акулову Василию Евгеньевичу за широкую помощь в работе и активное обсуждение результатов, Бондареву Станиславу Александровичу (СПбГУ, кафедра генетики и биотехнологии) и Елисееву Игорю Евгеньевичу (СПбАУ, лаборатория нанобиотехнологий) за консультации по синтезам и щедрое предоставление реактивов, сотрудникам ЛЦ ОЛМИВ (СПбГУ) Борисову Евгению Вадимовичу и Колесникову Илье Евгеньевичу за консультации по регистрации КР и флуоресцентных спектров

## Список литературы

1. All-optical electrophysiology in mammalian neurons using engineered microbial rhodopsins / D. R. Hochbaum [и др.] // *Nature methods*. — 2014. — т. 11, № 8. — с. 825—833.
2. *Zhang H., Cohen A. E.* Optogenetic approaches to drug discovery in neuroscience and beyond // *Trends in biotechnology*. — 2017. — т. 35, № 7. — с. 625—639.
3. Optical recording of action potentials in mammalian neurons using a microbial rhodopsin / J. M. Kralj [и др.] // *Nature methods*. — 2012. — т. 9, № 1. — с. 90—95.
4. Microbial and animal rhodopsins: structures, functions, and molecular mechanisms / O. P. Ernst [и др.] // *Chemical reviews*. — 2014. — т. 114, № 1. — с. 126—163.
5. *Shim S., Dasgupta J., Mathies R. A.* Femtosecond time-resolved stimulated Raman reveals the birth of bacteriorhodopsin's J and K intermediates // *Journal of the American Chemical Society*. — 2009. — т. 131, № 22. — с. 7592—7597.
6. *Lanyi J. K.* Proton transfer and energy coupling in the bacteriorhodopsin photocycle // *Journal of bioenergetics and biomembranes*. — 1992. — т. 24, № 2. — с. 169—179.
7. *Ames J. B., Mathies R. A.* The role of back-reactions and proton uptake during the N. fwdarw. O transition in bacteriorhodopsin's photocycle: a kinetic resonance Raman study // *Biochemistry*. — 1990. — т. 29, № 31. — с. 7181—7190.
8. *Alshuth T., Stockburger M.* Structural changes in the retinal chromophore of bacteriorhodopsin studied by resonance Raman spectroscopy // *Berichte der Bunsengesellschaft für physikalische Chemie*. — 1981. — т. 85, № 6. — с. 484—489.



9. Infrared and visible absolute and difference spectra of bacteriorhodopsin photocycle intermediates / R. W. Hendler [и др.] // *Applied spectroscopy*. — 2011. — т. 65, № 9. — с. 1029—1045.
10. *Lórenz-Fonfria V. A., Kandori H., Padrós E.* Probing specific molecular processes and intermediates by time-resolved Fourier transform infrared spectroscopy: Application to the bacteriorhodopsin photocycle // *The Journal of Physical Chemistry B*. — 2011. — т. 115, № 24. — с. 7972—7985.
11. Fourier transform infrared spectra of a late intermediate of the bacteriorhodopsin photocycle suggest transient protonation of Asp-212 / A. K. Dioumaev [и др.] // *Biochemistry*. — 1999. — т. 38, № 31. — с. 10070—10078.
12. Comparative studies of the fluorescence properties of microbial rhodopsins: spontaneous emission versus photointermediate fluorescence / K. Kojima [и др.] // *The Journal of Physical Chemistry B*. — 2020. — т. 124, № 34. — с. 7361—7367.
13. Fluorescence spectra of bacteriorhodopsin and the intermediates O and Q at room temperature / H. Ohtani [и др.] // *FEBS letters*. — 1995. — т. 359, № 1. — с. 65—68.
14. *Ohtani H., Itoh H., Shinmura T.* Time-resolved fluorometry of purple membrane of *Halobacterium halobium* O640 and an O-like red-shifted intermediate Q // *FEBS letters*. — 1992. — т. 305, № 1. — с. 6—8.
15. Fluorescence enhancement of a microbial rhodopsin via electronic reprogramming / M. d. C. Marin [и др.] // *Journal of the American Chemical Society*. — 2018. — т. 141, № 1. — с. 262—271.
16. Conformational changes in the archaerhodopsin-3 proton pump: detection of conserved strongly hydrogen bonded water networks / E. C. Saint Clair [и др.] // *Journal of biological physics*. — 2012. — т. 38, № 1. — с. 153—168.
17. Near-IR resonance Raman spectroscopy of archaerhodopsin 3: effects of transmembrane potential / E. C. Saint Clair [и др.] // *The Journal of Physical Chemistry B*. — 2012. — т. 116, № 50. — с. 14592—14601.
18. Screening fluorescent voltage indicators with spontaneously spiking HEK cells / J. Park [и др.] // *PloS one*. — 2013. — т. 8, № 12. — e85221.

19. Directed evolution of a far-red fluorescent rhodopsin / R. S. McIsaac [и др.] // Proceedings of the National Academy of Sciences. — 2014. — т. 111, № 36. — с. 13034—13039.
20. Archaelhodopsin variants with enhanced voltage-sensitive fluorescence in mammalian and *Caenorhabditis elegans* neurons / N. C. Flytzanis [и др.] // Nature communications. — 2014. — т. 5, № 1. — с. 1—9.
21. Mechanism of voltage-sensitive fluorescence in a microbial rhodopsin / D. Maclaurin [и др.] // Proceedings of the National Academy of Sciences. — 2013. — т. 110, № 15. — с. 5939—5944.
22. Electric-field-induced Schiff-base deprotonation in D85N mutant bacteriorhodopsin / P. Kolodner [и др.] // Proceedings of the National Academy of Sciences. — 1996. — т. 93, № 21. — с. 11618—11621.
23. Crystal structure of the D85S mutant of bacteriorhodopsin: model of an O-like photocycle intermediate / S. Rouhani [и др.] // Journal of molecular biology. — 2001. — т. 313, № 3. — с. 615—628.
24. Specificity of anion binding in the substrate pocket of bacteriorhodopsin / M. T. Facciotti [и др.] // Biochemistry. — 2004. — т. 43, № 17. — с. 4934—4943.
25. *Facciotti M. T., Rouhani S., Glaeser R. M.* Crystal structures of bR (D85S) favor a model of bacteriorhodopsin as a hydroxyl-ion pump // FEBS letters. — 2004. — т. 564, № 3. — с. 301—306.
26. Hydrogen bonding interactions with the Schiff base of bacteriorhodopsin. Resonance Raman spectroscopy of the mutants D85N and D85A. / P. Rath [и др.] // Journal of Biological Chemistry. — 1993. — т. 268, № 24. — с. 17742—17749.
27. Crystal structure of the O intermediate of the Leu93→Ala mutant of bacteriorhodopsin / J. Zhang [и др.] // Proteins: Structure, Function, and Bioinformatics. — 2012. — т. 80, № 10. — с. 2384—2396.
28. *Tóth-Boconádi R., Keszthelyi L., Stoeckenius W.* Photoexcitation of the O-intermediate in bacteriorhodopsin mutant L93A // Biophysical journal. — 2003. — т. 84, № 6. — с. 3857—3863.

29. Static and time-resolved absorption spectroscopy of the bacteriorhodopsin mutant Tyr-185. fwdarw. Phe: Evidence for an equilibrium between bR570 and an O-like species / S. Sonar [и др.] // *Biochemistry*. — 1993. — т. 32, № 9. — с. 2263—2271.
30. Fourier transform Raman spectroscopy of the bacteriorhodopsin mutant Tyr-185. fwdarw. Phe: Formation of a stable O-like species during light adaptation and detection of its transient N-like photoproduct / P. Rath [и др.] // *Biochemistry*. — 1993. — т. 32, № 9. — с. 2272—2281.
31. FTIR difference spectroscopy of the bacteriorhodopsin mutant Tyr-185. fwdarw. Phe: Detection of a stable O-like species and characterization of its photocycle at low temperature / Y. He [и др.] // *Biochemistry*. — 1993. — т. 32, № 9. — с. 2282—2290.
32. Existence of two O-like intermediates in the photocycle of *Acetabularia* rhodopsin II, a light-driven proton pump from a marine alga / J. Tamogami [и др.] // *Biophysics and physicobiology*. — 2017. — т. 14. — с. 49—55.
33. An assessment of water placement algorithms in quantum mechanics/molecular mechanics modeling: the case of rhodopsins' first spectral absorption band maxima / D. M. Nikolaev [и др.] // *Physical Chemistry Chemical Physics*. — 2020. — т. 22, № 32. — с. 18114—18123.
34. Catalysis of Ground State cis to trans Isomerization of Bacteriorhodopsin's Retinal Chromophore by a Hydrogen-Bond Network / N. Elghobashi-Meinhardt [и др.] // *The Journal of membrane biology*. — 2018. — т. 251, № 3. — с. 315—327.
35. *Shibata M., Tanimoto T., Kandori H.* Water molecules in the Schiff base region of bacteriorhodopsin // *Journal of the American Chemical Society*. — 2003. — т. 125, № 44. — с. 13312—13313.
36. *Kandori H.* Role of internal water molecules in bacteriorhodopsin // *Biochimica et Biophysica Acta (BBA)-Bioenergetics*. — 2000. — т. 1460, № 1. — с. 177—191.
37. *Shibata M., Kandori H.* FTIR studies of internal water molecules in the Schiff base region of bacteriorhodopsin // *Biochemistry*. — 2005. — т. 44, № 20. — с. 7406—7413.

38. Photocycle of Sensory Rhodopsin II from *Halobacterium salinarum* (Hs SRII): Mutation of D103 Accelerates M Decay and Changes the Decay Pathway of a 13-cis O-like Species / G. Dai [и др.] // *Photochemistry and photobiology*. — 2018. — т. 94, № 4. — с. 705—714.
39. Crystal structures of an O-like blue form and an anion-free yellow form of pharaonis halorhodopsin / S. Kanada [и др.] // *Journal of molecular biology*. — 2011. — т. 413, № 1. — с. 162—176.
40. Structures of the archaerhodopsin-3 transporter reveal that disordering of internal water networks underpins receptor sensitization / J. F. B. Juarez [и др.] // *Nature communications*. — 2021. — т. 12, № 1. — с. 1—10.
41. AlignMe—a membrane protein sequence alignment web server / M. Stamm [и др.] // *Nucleic acids research*. — 2014. — т. 42, W1. — W246—W251.
42. *Zhang Y.* I-TASSER server for protein 3D structure prediction // *BMC bioinformatics*. — 2008. — т. 9, № 1. — с. 1—8.
43. *Morozenko A., Stuchebrukhov A.* Dowser++, a new method of hydrating protein structures // *Proteins: Structure, Function, and Bioinformatics*. — 2016. — т. 84, № 10. — с. 1347—1357.
44. The ONIOM method and its applications / L. W. Chung [и др.] // *Chemical reviews*. — 2015. — т. 115, № 12. — с. 5678—5796.
45. Gaussian 09w reference / A. Frisch [и др.] // Wallingford, USA, 25p. — 2009.
46. *Neese F.* The ORCA program system // *Wiley Interdisciplinary Reviews: Computational Molecular Science*. — 2012. — т. 2, № 1. — с. 73—78.
47. PROPKA3: consistent treatment of internal and surface residues in empirical pK<sub>a</sub> predictions / M. H. Olsson [и др.] // *Journal of chemical theory and computation*. — 2011. — т. 7, № 2. — с. 525—537.
48. *Chipot C., Pohorille A.* Free energy calculations // *Springer series in chemical physics*. — 2007. — т. 86. — с. 159—184.
49. *Saint Clair E. C.* FTIR difference and resonance raman spectroscopy of rhodopsins with applications to optogenetics : дис. ... канд. / Saint Clair Erica C. — Boston University, 2013.

50. Characterization of the branched-photocycle intermediates P and Q of bacteriorhodopsin / N. B. Gillespie [и др.] // The Journal of Physical Chemistry B. — 2002. — т. 106, № 51. — с. 13352—13361.
51. *Watanabe H. C., Ishikura T., Yamato T.* Theoretical modeling of the O-intermediate structure of bacteriorhodopsin // Proteins: Structure, Function, and Bioinformatics. — 2009. — т. 75, № 1. — с. 53—61.

**Приложение А****Праймеры для мутагенеза**

AR-P196S-F GGTGCTGTGGACCGCTTACtCcATCCTGTGGATCATAGGCAC  
AR-P196S-R GTGCCTATGATCCACAGGATgGaGTAAGCGGTCCACAGCACC  
AR-P60L-F GAATATTACGCTGTGACTATCCTGGTGctCGGAATCGCATCCGCCGCATATCTG  
AR-P60L-R CAGATATGCGGCGGATGCGATTCCGgAGCACCAGGATAGTCACAGCGTAATATTC  
AR-V59A-F GGAATATTACGCTGTGACTATCCTGGcGCCCCGGAATCGCATCCGCCGCATATC  
AR-V59A-R GATATGCGGCGGATGCGATTCCGGGcCCAGGATAGTCACAGCGTAATATTC  
AR-D95ET99C-F GGATATCTATTATGCCAGGTACGCCGaaTGGCTGTTTTg  
-CACCCCACTTCTGCTGCTGGATCTGG  
AR-D95ET99C-R CCAGATCCAGCAGCAGAAGTGGGGTGcaAAACAGCCAt  
-TCGGCGTACCTGGCATAATAGATATCC

## Приложение Б

### Использованные буферы

Сульфатный буфер (рН 1,5, 2, 3): аммония сульфат 0.1 М, 1,3 %  $H_2SO_4$ . Необходимое значение рН доводилось разбавленной серной или соляной кислотой до необходимого значения потенциометрическим методом.

Фосфатный буфер (рН 3, 4, 5, 6, 7): были приготовлены растворы  $NaH_2PO_4 \cdot 12H_2O$  (0.1 М) и  $Na_2HPO_4 \cdot 2H_2O$  (0.1 М), смешением этих растворов при контроле рН-метром.

Трисовый буфер (рН 8, 9): Tris 0.1 М, далее рН потенциометрически доводился HCl до необходимого значения. Аммиачный буфер (рН 10, 11): аммоний хлорид 1 М, затем рН потенциометрически доводился до необходимого значения концентрированным 25 % раствором аммиака. Конечная концентрация хлорида аммония в образце была 0.1М.

## Приложение В

### Результаты секвенирования

#### AR DETC

```

5201 gccggccaagatcgctccgggtagaggatcgagatctcgatccccgaaattaatacagactcactataggggaattgtgagcgataacaattcccc 5300
948 CGAGCCCCAATGCTTCCGGGTAGAGGATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGAGGAATTGTGAGCGGATAACAATTCOCCTC 849
* * * * *
5301 tagaataattttgtaactttaagaaggagatacacatattggaccccatcgctctgagggctggttaagcaactgctgggtgacggcagacctgaaact 5400
848 TAGAAATAATTTTGTAACTTTAAGAAGGAGATATACATATGGACCCCATCGCTCTGAGGCTGGTTACGACCTGCTGGGTGACGGCAGACCTGAAACT 749
* * * * *
5401 ctgtggctgggcatcgccactctgctgatgctgattggaacctctactttctggtccgggatggggagtcaaccgataaggatgcccggaatattacg 5500
748 CTGTGGCTGGGCACTCGCCTCTGCTGATGCTGATTGGAACCTTCTACTTTCTGCTCCGCGGATGGGGAGTCAACGATAAGGATGCCCGGGAATATTACG 649
* * * * *
5501 ctgtgactatcctggtgccggaatcgcatccgccgcatatctgtctatgttcttgggtatcgggctactgaggtgacogtccggggcgaatgttga 5600
648 CTGTGACTATCCTGGTGCCGGAATCGCATCCGCGCATATCTGTCTATGTTCTTTGGTATCGGGCTACTGAGGTGACCGTCCGGGGCGAAATGTTGA 549
* * * * *
5601 tatctattatgccaggtacggcactggctgtttaccaccccactctctgctgctggatctggccctctctgctaaggtggatcggtgacccatcgccacc 5700
548 TATCTATTATGCCAGGTACGGCAGTGGCTGTTTTCACCCCACTCTCTGCTGCTGGATCTGGCCCTTCTGCTAAGGTGGATCGGGTGAACCATCGGCACC 449
* * * * *
5701 ctgggtgggtgaggacgacctgatgatgctcactggcctcatcggagccttgagccacagggccatagccagatacagttgggtggttctctacaatt 5800
448 CTGGTGGGTGAGGACGACCTGATGATCGTCACTGGCCTCATCGGAGCCTTGAGCCACAGGCCATAGCCAGATACAGTTGGTGGTGTCTCTACAATT 349
* * * * *
5801 gcattgatagtggtgctctatcttctggctacatccctcgatctgctgcaaaaggagcggggccccagggtggcatctaccttaacacccctgacagctct 5900
348 GCATGATAGTGGTGTCTATCTTCTGGCTACATCCCTGCGATCTGCTGCAAAAGGAGCGGGGCCCCAGGTGGCATCTACCTTTAACACCCCTGACAGCTCT 249
* * * * *
5901 ggtcttgggtgctgaggccgctaccctatcctgtggatcattgagcactgagggcgctggcggtgggctgggcatgaaactctgctgtttatgggt 6000
248 GGTCTTGGTGTGCTGAGCCGCTTACCCTATCCTGTGGATCATAGGCCACTGAGGCGCTGGCCTGGTGGGCTGGGCATCGAAACTCTGCTGTTTATGGTG 149
* * * * *
6001 ttggacgtgactgcaaggctggctttggctttatcctgttgagatccccggctattctgggacacagggcaccagaaccagtgccggtgcccgatg 6100
148 TTGGACGTGACTGCCAAGCTGGCTTTGGCTTTATCCTGTGAGATCCCGGGCTATTCTGGGCGACCCCGTGGCACCCAGAACCCAGTCCCGGTGCCGATG 49
* * * * *
6101 tcagtgccgacgactogagaccaccaccaccacc---actg-3-- 6141
48 TCAGTGCOCGACCTCGAGCACCCACCCACCCACCCATCAGCTACTGCCAG 1

```



AR 99/95/59

```

* * * * *
5397 aactctgtggctgggcatcggaactctgctgatgctgattggaaccttctacttt-ctggtcocgggatggggagtcaccgataaggatgccccggaata 5495
|||||
749 AACTCTGTGGCTGGGCATCGGCACTCTGCTGATGCTGATTGGAACTTCTACTTTTCTGGTCOCGGGATGGGGAGTCAACGATAAGGATGCCCGGGAATA 650

* * * * *
5496 ttacgctgtgactatcctgtgcccgaatcgcatccgcgcacatctgtctatgttctttggatcgggcttactgaggtgacocgctggggggaatg 5595
|||||
649 TTAOCCTGTGACTATCCTGGTGCCTCGGAATCGCATCCGCCATATCTGCTATGTTCTTTGGTATCGGGCTTACTGAGGTGACOCGCGGGGGCAAAATG 550

* * * * *
5596 ttggatctctattatgcccaggtacgccaactggctgtttaccaccccactctctgctgctggatctggccctctcctaaagtggaatgggtgaccatcg 5695
|||||
549 TTGGATATCTATTATGCCAGGTAOCGCCAATGGCTGTTTTGCACCCCACTCTGCTGCTGGATCTGGCCCTTCTGCTAAGGTGGATCGGGTGACCATCG 450

* * * * *
5696 gcacctgtgggtggtggaagcctgatgctgctcaactggcctcactggagccttgagccacacggccatagccagatacagttggtggttctctac 5795
|||||
449 GCACCCCTGGTGGGTGAGCCCTGATGATCGTCACTGGCCTCATCGGAGCCTTGAGCCACACGGCCATAGCCAGATACAGTTGGTGGTGTCTCTAC 350

* * * * *
5796 aattgcatgatagtggtgctctatcttctggctacatccctcgatctgctgcaaggagcggggcccccagggtggcatctacaccttaacacctgaca 5895
|||||
349 AATTGCAATGATAGTGGTCTATTTCTGGCTACATCCCTGCGATCTGCTGCAAAGGAGCGGGGGCCCCGAGGTGGCATCTACCTTTAACACCCCTGACA 250

* * * * *
5896 gctctgttctgtgctgctggaacgcttaacctatcctgtggatcataggcactgagggcgctggcgtggtggcctgggcatgaaactctgctgttta 5995
|||||
249 GCTCTGCTTGTGGTCTGCTGAGACCGCTTACCTATCCTGCTGATCATAGGCACTGAGGGCGCTGGCGTGGGGCTGGCATCGAAACTCTGCTGTTTAA 150

* * * * *
5996 tgggtgtggaactgactgccaaggtcggtcttggctttatcctgttgagatcccgggtattctggggacacccaggcaccagaacccagtgccggtgc 6095
|||||
149 TGGTGTGGACCTGACTGCCAAGGTGCGCTTTGGCTTTATCCTGCTGAGATCCCAGGCTATTCTGGCGACACCCAGGCAACCCAGTCCCGCGTGC 50

* * * * *
6096 cgatgtcagtcgcccgcacctcgagcaccaccaccaccactga~~~ 6141
|||||
49 CGATGTCAGTCCCGCACCTCGAGCAACCACCACCACCACCTGATCC 1

```

AR 99/95/60

```

* * * * *
5400 tctgtgctgggcatcggaactctgctgatgctgattggaaccttctactttctggtcocgggatggggagtcaccgataaggatgccccggaatattac 5495
|||||
745 TCTGTGGCTGGGCATCGGCACCTGCTGATGCTGATTGGAACTTCTACTTTCTGGTCOCGGGATGGGGAGTCAACGATAAGGATGCCCGGGAATATTAC 646

* * * * *
5500 gctgtgactatcctgtgcccgaatcgcatccgcgcacatctgtctatgttctttggatcgggcttactgaggtgacocgctggggggaatgttgg 5595
|||||
645 GCTGTGACTATCCTGGCGCCCGGAATCGCATCCGCCATATCTGCTATGTTCTTTGGTATCGGGCTTACTGAGGTGACOCGCGGGGGCAAAATGTTGG 546

* * * * *
5600 atatctattatgcccaggtacgccaactggctgtttaccaccccacttctgctgctggatctggccctctcctaaagtggaatgggtgaccatcgccac 5695
|||||
545 ATATCTATTATGCCAGGTAOCGCCAATGGCTGTTTTGCACCCCACTCTGCTGCTGGATCTGGCCCTTCTGCTAAGGTGGATCGGGTGACCATCGGCAC 446

* * * * *
5700 cctggtgggtggtggaagcctgatgctgctcaactggcctcactggagccttgagccacacggccatagccagatacagttggtggttctctacaatt 5795
|||||
445 CCTGTGGGTGTTGAGCCCTGATGATCGTCACTGGCCTCATCGGAGCCTTGAGCCACACGGCCATAGCCAGATACAGTTGGTGGTGTCTCTACAAAT 346

* * * * *
5800 tgcattgatagtggtgctctatcttctggctacatccctcgatctgctgcaaggagcggggcccccagggtggcatctacaccttaacacctgacagctc 5895
|||||
345 TGCAATGATAGTGGTCTATTTCTGGCTACATCCCTGCGATCTGCTGCAAAGGAGCGGGGGCCCCGAGGTGGCATCTACCTTTAACACCCCTGACAGCTC 246

* * * * *
5900 tggctctggtgctggaagccttaacctatcctgtggatcataggcactgagggcgctggcgtggtggcctgggcatgaaactctgctgtttatggt 5995
|||||
245 TGGCTTGGTCTGCTGGACCGCTTACCCTATCCTGTGGATCATAGGCACTGAGGGCGCTGGCCTGGTGGGCTGGGCATCGAAACTCTGCTGTTTATGTT 146

* * * * *
6000 gttggaactgactgccaaggtcggtcttggctttatcctgttgagatcccggctattctggggacacccaggca-ccagaacccagtgccggtgcca 6095
|||||
145 GTTGGACGTGACTGCCAAGGTGCGCTTTGGCTTTATCCTGTTGAGATCCCAGGCTATTCTGGCGACACCCAGGCAAGCCAGAACCCAGTCCCGGTGCCGA 46

* * * * *
6099 tgtcagtcgcccgcacctcgagcaccaccaccaccaccactga~~~ 6141
|||||
45 TGTCAAGTCCCGCACCTCGAGCAACCACCACCACCAC-CAGCC 1

```

AR5

```

5501 ctgtgactatcctgggtgcccgaatcgcatccgcccgcatactgctctatgttctttgggtatcggg-cttactgaggtgacgctggggggcgaatgttg 5599
626 -----AATCGCATCCGCGCCATATCTGTCTATGTTCTTTGGTATCGGGGCTTACTGAGGTGACCGTCGGGGGCGAAATGTTG 549

5600 atatctatatgcccaggtacgcgcactggctgtttaccaccccactctgctgctggatctggccctctctgctaaggtggatcgggtgaccatcggcac 5699
548 ATATCTATTATGCCAGGTACGCCGAATGGCTGTTTGCACCCCACTTCTGCTGCTGGATCTGGCCCTTCTCGCTAAGGTGGATCGGGTGACCATCGGCAC 449

5700 cctggtgggtgtggacgcccctgatgatctgactggcctcatcggagccttgagccacacggccatagccagatacagttggtggttctctacaatt 5799
448 CCTGGTGGGTGTGGACGCCCTGATGATCGTCACTGGCCTCATCGGAGCCTTGAGCCACACGGCCATAGCCAGATACAGTTGGTGGTGTCTCTACAATT 349

5800 tgcattgatagtggtgctctattttctggctacatccctcgatctgctgcaaggagcggggccccagggtggcatctaccttaacacctgacagctc 5899
348 TGCATGATAGTGGTGTCTCTATTTCTGGCTACATCCCTGCGATCTGCTGCAAAGGAGCGGGGCCCCAGGTGGCATCTACCTTAAACACCTGACAGCTC 249

5900 tggctcttggctgctgggacgcttaacctatccctggtgcatagggcactgaggcgctggcgtggtgggctgggcatcgaaactctgctgtttatggt 5999
248 TGGTCTTGGTGTGGACCGCTTACTCCATCCCTGGATCATAGGCCTGAGGGCGCTGGCGTGGTGGGCTGGGCATCGAAACTCTGCTGTTATGGT 149

6000 gttggaactgactgccaaggtcggctttggctttatcctggtgagatccgggctattctggcgacaccaggcaccagaaccagtgccggtgcccgat 6099
148 GTTGGACGTGACTGCCAAGTCCGGCTTTGGCTTTATCCCTGTTGAGATCCCGGGCTATTCTGGCGACACCAGGCACCAGAACCCAGTGCCTGCGCAT 49

6100 gtcagtgcgcgcgactcgagcaccacccaccacca-ctga~~~~~ 6141
48 GTCAGTGCOCGCGACTTCGAGCACCAACCCACCACTTATGGG 1

```

AR7

```

5501 ctgtgactatcctgggtgcccgaatcgcatccgcccgcatactgctctatgttctttgggtatcgggcttactgaggtgacgctggggggcgaatgttga 5600
639 -----CGGAATCGCATCCGCGCCATATCTGTTATGTTCTTTGGTATCGGGGCTTACTGAGGTGACCGTCGGGGGCGAAATGTTGA 559

5601 tatcta-ttatgcccaggtacgcgcactggctgtttaccaccccactctgctgctggatctggccctctctgctaaggtggatcgggtgaccatcggcac 5699
558 TATCTATTATGCCAGGTACGCCGAATGGCTGTTTGCACCCCACTTCTGCTGCTGGATCTGGCCCTTCTCGCTAAGGTGGATCGGGTGACCATCGGCAC 459

5700 cctggtgggtgtggacgcccctgatgatctgactggcctcatcggagccttgagccacacggccatagccagatacagttggtggttctctacaatt 5799
458 CCTGGTGGGTGTGGACGCCCTGATGATCGTCACTGGCCTCATCGGAGCCTTGAGCCACACGGCCATAGCCAGATACAGTTGGTGGTGTCTCTACAATT 359

5800 tgcattgatagtggtgctctattttctggctacatccctcgatctgctgcaaggagcggggccccagggtggcatctaccttaacacctgacagctc 5899
358 TGCATGATAGTGGTGTCTCTATTTCTGGCTACATCCCTGCGATCTGCTGCAAAGGAGCGGGGCCCCAGGTGGCATCTACCTTAAACACCTGACAGCTC 259

5900 tggctcttggctgctgggacgcttaacctatccctggtgcatagggcactgaggcgctggcgtggtgggctgggcatcgaaactctgctgtttatggt 5999
258 TGGTCTTGGTGTGGACCGCTTACTCCATCCCTGGATCATAGGCCTGAGGGCGCTGGCGTGGTGGGCTGGGCATCGAAACTATGCTGTTATGGT 159

6000 gttggaactgactgccaaggtcggctttggctttatcctggtgagatccgggctattctggcgacaccaggcaccagaaccagtgccggtgcccgat 6099
158 GTTGGACGTGACTGCCAAGTCCGGCTTTGGCTTTATCCCTGTTGAGATCCCGGGCTATTCTGGCGACCCCTGAGGCACCAGAACCCAGTGCCTGCGCAT 59

6100 gtcagtgcgcgcgactcgagcacca-----caccacacca-----ctg--a 6141
58 GTCAGTGCOCGCGACTTCGAGCACCAACCCACCACTTATGGGCTTATCGGCGACCCCTGAGGCACCAGAACCCAGTGCCTGCGCAT 1

```