

Федеральное государственное бюджетное учреждение науки
Институт биоорганической химии
им. академиков М.М. Шемякина и Ю.А. Овчинникова
Российской Академии Наук

На правах рукописи

Божанова Нина Георгиевна

**Разработка и изучение флуоресцентных меток методами моделирования и
молекулярной эволюции белков**

специальность 03.01.03 (молекулярная биология)

Диссертация на соискание ученой степени

кандидата биологических наук

Научный руководитель:
член-корреспондент РАН
доктор биологических наук
Лукьянов Константин Анатольевич

Содержание

1. ВВЕДЕНИЕ.....	5
2. ОБЗОР ЛИТЕРАТУРЫ.....	6
2.1. ФЛУОРЕСЦЕНТНОЕ МЕЧЕНИЕ БЕЛКОВ.....	6
2.1.1. Генетически кодируемые метки.....	6
2.1.1.1. Пептиды	6
Тетрацистеиновая последовательность	6
Олигогистидиновая последовательность.....	7
Flag tag.....	7
Ферментативная модификация пептидов	8
2.1.1.2. Белки.....	10
Флуоресцентные белки.....	10
Самомодифицирующиеся белки.....	11
Флуорофор-связывающие белки	13
Флуороген-активирующие белки	14
2.1.1.3. Неприродные аминокислоты	17
2.1.2. Химические метки	19
2.1.2.1. Неспецифическое мечение	19
2.1.2.2. Специфическое мечение	19
2.2. МЕТОДЫ МОЛЕКУЛЯРНОЙ ЭВОЛЮЦИИ БЕЛКОВ.....	22
2.2.1. Случайный мутагенез.....	22
2.2.2. Сайт-специфический мутагенез.....	24
2.2.3. In vitro рекомбинация	27
2.3. МОДЕЛИРОВАНИЕ В БИОЛОГИИ	30
2.3.1. Модельные соединения.....	30
2.3.2. Компьютерное моделирование.....	33
2.3.2.1. Определение позиций для мутагенеза.....	34
2.3.2.2. Выбор перспективных замен.....	37
2.3.2.3. In vitro рекомбинация.....	38
2.3.2.4. Дизайн белков.....	38
Редизайн	38
De novo дизайн	39
3. МАТЕРИАЛЫ И МЕТОДЫ.....	42
3.1. АМПЛИФИКАЦИЯ ФРАГМЕНТОВ ДНК	42

3.2.	ЭЛЕКТРОФОРЕЗ В АГАРОЗНОМ ГЕЛЕ	42
3.3.	ОЧИСТКА ДНК.....	43
3.4.	САЙТ-СПЕЦИФИЧЕСКИЙ МУТАГЕНЕЗ.....	43
3.4.1.	<i>Гибридизация N- и C- концевых фрагментов</i>	<i>43</i>
3.4.2.	<i>Самособирающееся клонирование</i>	<i>44</i>
3.4.3.	<i>“Аква” клонирование.....</i>	<i>45</i>
3.5.	СОЗДАНИЕ БИБЛИОТЕК СЛУЧАЙНЫХ МУТАНТОВ	45
3.6.	ТРАНСФОРМАЦИЯ БАКТЕРИЙ.....	46
3.6.1.	<i>Химическая трансформация</i>	<i>46</i>
3.6.2.	<i>Электрическая трансформация</i>	<i>46</i>
3.7.	ЭКСПРЕССИЯ И ОЧИСТКА РЕКОМБИНАНТНЫХ БЕЛКОВ.....	46
3.8.	ДЕНАТУРИРУЮЩИЙ ЭЛЕКТРОФОРЕЗ В ПОЛИАКРИЛАМИДНОМ ГЕЛЕ	47
3.9.	ОПРЕДЕЛЕНИЕ СПЕКТРАЛЬНЫХ СВОЙСТВ.....	48
3.10.	РН-ТИТРОВАНИЕ	48
3.11.	ОПРЕДЕЛЕНИЕ КОНСТАНТ ДИССОЦИАЦИИ	48
3.12.	ПОЛУЧЕНИЕ ГЕНЕТИЧЕСКИХ КОНСТРУКЦИЙ	50
3.12.1.	<i>Клонирование гена blc</i>	<i>50</i>
3.12.2.	<i>Создание конструкций для временной трансфекции культур эукариотических клеток.....</i>	<i>50</i>
3.13.	КУЛЬТИВИРОВАНИЕ И ВРЕМЕННАЯ ТРАНСФЕКЦИЯ КУЛЬТУР ЭУКАРИОТИЧЕСКИХ КЛЕТОК..	51
3.14.	ФЛУОРЕСЦЕНТНАЯ МИКРОСКОПИЯ	51
3.15.	КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ.....	52
3.15.1.	<i>MODELLER</i>	<i>53</i>
3.15.2.	<i>AutoDoc Vina.....</i>	<i>53</i>
3.15.3.	<i>Rosetta</i>	<i>53</i>
4.	РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ	56
4.1.	РАЗРАБОТКА И ИЗУЧЕНИЕ НОВЫХ СПЕКТРАЛЬНЫХ ВАРИАНТОВ ФЛУОРЕСЦЕНТНЫХ БЕЛКОВ, СОДЕРЖАЩИХ ТРИПТОФАН В СОСТАВЕ ХРОМОФОРА.....	56
4.1.1.	<i>Получение оранжевых триптофановых флуоресцентных белков.....</i>	<i>56</i>
4.1.2.	<i>Изучение роли замен</i>	<i>58</i>
4.1.3.	<i>Возможные структуры хромофора.....</i>	<i>61</i>
4.1.4.	<i>Поведение при денатурации</i>	<i>64</i>
4.1.5.	<i>pH титрование</i>	<i>66</i>
4.1.6.	<i>Фотоконверсия</i>	<i>69</i>
4.1.7.	<i>Модельное соединение.....</i>	<i>70</i>

4.1.8.	Компьютерное моделирование.....	73
4.2.	РАЗРАБОТКА И ИЗУЧЕНИЕ ФЛУОРЕСЦЕНТНЫХ МЕТОК НА ОСНОВЕ ФЛУОРОГЕН-АКТИВИРУЮЩИХ БЕЛКОВ.....	76
4.2.1.	Выбор флуорогенов.....	76
4.2.2.	Поиск флуороген-активирующих белков.....	81
4.2.2.1.	<i>In silico</i> скрининг	81
4.2.2.2.	Клонирование генов белков 3HO2, 1DOS, 1PVS и 2QRY	83
4.2.2.3.	Характеристика найденных белков	83
4.2.2.4.	Анализ причин различий в флуорогенности	87
4.2.3.	Создание флуороген-активирующих белков	90
4.2.3.1.	Выбор белка для моделирования.....	90
4.2.3.2.	<i>In silico</i> мутагенез липокалина.....	91
4.2.3.3.	<i>In vitro</i> анализ библиотеки мутантов.....	92
4.2.3.4.	Характеристика выбранных пар белок-флуороген	92
4.2.3.5.	Изучение роли замен.....	96
4.2.3.6.	Использование мутантов липокалина в качестве метки для флуоресцентной микроскопии	98
4.2.4.	Улучшение созданных белков.....	105
4.2.4.1.	Случайный мутагенез	105
4.2.4.2.	Характеристика белков, полученных случайным мутагенезом	105
4.2.4.3.	<i>In silico</i> мутагенез.....	107
4.2.4.4.	Характеристика белков, полученных в результате <i>in silico</i> мутагенеза	108
4.2.4.5.	Использование улучшенных вариантов в качестве метки для флуоресцентной микроскопии	111
4.2.5.	Дальнекрасные флуорогены.....	113
5.	ЗАКЛЮЧЕНИЕ	116
6.	ВЫВОДЫ.....	117
7.	СПИСОК СОКРАЩЕНИЙ	118
8.	СПИСОК ЛИТЕРАТУРЫ	120

1. Введение

В настоящее время комплексные процессы, происходящие в живых организмах, часто изучают путём детального рассмотрения распределения, локализации, движения и взаимодействия молекул на всех уровнях организации живых систем, от клеточного до организменного.

Большая часть соединений в клетке бесцветна, поэтому наблюдение за биомолекулами практически всегда требует от исследователя введения какой-либо метки. Сегодня для этой цели доступен целый ряд разнообразных инструментов: изотопные маркеры, радиоактивные трассеры, электрохимические сенсоры, колориметрические биосенсоры, флуоресцентные метки и т.д. Применение флуоресцентных меток имеет ряд существенных преимуществ. Среди них — безопасность, низкая токсичность, разнообразие и высокая чувствительность современных флуоресцентных детекторов, позволяющая регистрировать сигнал с большой скоростью, в малых объемах (индивидуальных клетках) и при низкой концентрации.

Идеальная флуоресцентная метка для биоимиджинга должна быть яркой, небольшой, нетоксичной, простой в применении и подходить для решения широкого спектра задач на всех уровнях организации (от *in vitro* до *in vivo*). Она должна флуоресцировать в красной области спектра (облучение таким светом минимизирует проблемы фототоксического эффекта и автофлуоресценции биологических образцов), обладать максимальными эффективностью и специфичностью мечения.

Хотя за последние десятилетия были достигнуты значительные успехи в области флуоресцентного мечения живых систем, позволившие ответить на многие биологические вопросы, задача создания идеальной флуоресцентной метки пока что остается нерешенной и актуальной.

Понимание основ работы существующих флуоресцентных меток позволит рационально совершенствовать имеющиеся и разрабатывать новые инструменты. Активно развивающиеся в последнее время мультидисциплинарные исследования демонстрируют свою силу как в фундаментальных областях, находя решения старых вопросов, порой недоступных методологическому аппарату одной науки, так и в прикладных, позволяя создавать новые технологии.

2. Обзор литературы

2.1. Флуоресцентное мечение белков

Сложно точно определить время изобретения флуоресцентного мечения. Пожалуй, первой открытой флуоресцентной молекулой можно считать сульфат хинина, описанный Джоном Гершем в 1845 году [1]. Термин “флуоресценция”, однако, был введен несколько позже — Джорджем Стоксом в 1852 году [2], а “флуорофор” — лишь в 1897 году Ричардом Майером для описания группы, связанной с флуоресценцией химического вещества [3].

Первые флуоресцентные микроскопы появились в 1911 году [4]. Первоначально с их помощью наблюдали бактерий, ткани и клетки растений и животных, которые обладают собственной флуоресценцией. Между тем уже в 1914 году чешский естествоиспытатель Станислав Провачек изучал связывание красителей с клетками для того, чтобы сделать их флуоресцентными [5]. Эти ранние работы стали толчком для разработки широкого спектра агентов для флуоресцентного мечения клеток, применяемых сегодня в биологических и медицинских исследованиях.

Основную массу органических веществ клетки и большой интерес для исследователей представляют разнообразные белки. Все используемые для их изучения флуоресцентные метки можно разделить на две большие группы: генетически кодируемые и химические.

2.1.1. Генетически кодируемые метки

К генетически кодируемым меткам относят те системы, в которых сама метка или специфически распознаваемая меткой часть встраивается в исследуемый белок во время его синтеза. В зависимости от длины встраиваемого участка различают пептидные и белковые генетически кодируемые метки, а также неприродные аминокислоты.

2.1.1.1. Пептиды

Генетически кодируемые метки на основе пептидов имеют ряд существенных преимуществ. Благодаря небольшому размеру значительно снижается вероятность нарушить фолдинг или функцию исследуемого белка. Кроме того, короткая последовательность может быть внесена не только на N- или C-конец полипептидной цепи, но и встроена внутрь белковой молекулы (например, в петлю), что увеличивает шансы на успешное мечение. С другой стороны, с уменьшением длины распознаваемого участка возрастает вероятность присутствия подобного же сайта в нецелевом белке, что может приводить к появлению неспецифического сигнала.

Тетрацистеиновая последовательность

Первым пептидом, разработанным для специфического мечения белков, была

тетрацистеиновая последовательность -Cys-Cys-X-X-Cys-Cys- (X — любая аминокислота, кроме цистеина, обычно -Cys-Cys-Pro-Gly-Cys-Cys-), имеющая высокую аффинность к молекулам с двумя атомами мышьяка в составе. Содержащие такую последовательность белки способны флуоресцировать при связывании некоторых мышьяк-содержащих красок (например, легко проникающих сквозь клеточную мембрану мышьяковых производных флуоресцеина и родамина FAsH и ReAsH) (рисунок 1) [6].

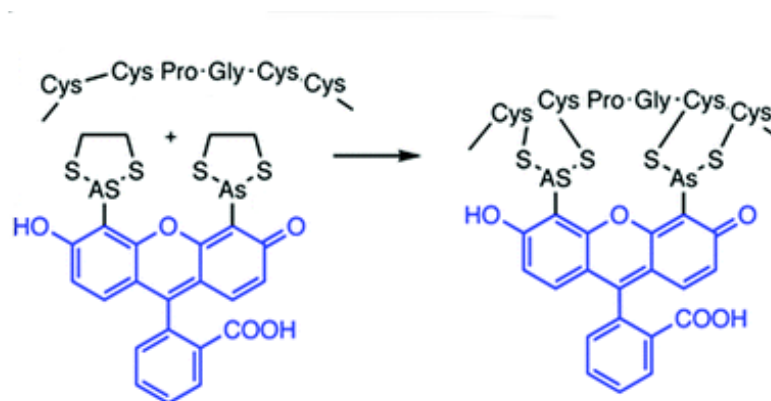


Рисунок 1. Схема взаимодействия тетрацистеинового мотива с красителем FAsH [7].

К сожалению, специфичность красителей, содержащих атомы трехвалентного мышьяка, не высока: белки с несколькими или даже одним цистеином также могут быть помечены, что приводит к появлению фоновой флуоресцентной сигналы. Кроме того, такие красители очень токсичны, что затрудняет их широкое использование *in vivo* [8].

Олигогистидиновая последовательность

Последовательность из 6 и более гистидинов также может быть использована для мечения за счёт способности к нековалентному взаимодействию с нитрилотриацетоникелем(II), ковалентно связанным с каким-либо флуорофором [9]. Данный метод был успешно применен как для *in vitro*, так и для *in vivo* мечения белков интереса [10]. В отличие от вышеупомянутой тетрацистеиновой последовательности, олигогистидиновая или схожие с ней последовательности встречаются в природных белках гораздо реже, что уменьшает вероятность неспецифического мечения.

Flag tag

Ещё одной последовательностью, способной к специфическому нековалентному связыванию флуоресцентной метки, содержащей ионы металла, является последовательность, включающая четыре остатка аспартата [11], например, Flag tag (-Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys-) [12].

Ферментативная модификация пептидов

Селективность реакций мечения коротких аминокислотных последовательностей может быть значительно улучшена, если использовать для их распознавания высокоспецифичные биологические молекулы. Для создания подобных систем чаще всего используются ферменты, проводящие посттрансляционные модификации.

Одной из первых реализаций данного подхода было создание системы флуоресцентного мечения на основе фермента биотинлигазы. Этот фермент ковалентно присоединяет биотин к боковой цепи определенного остатка лизина в акцепторном пептиде. Так как биотин имеет очень высокое сродство к стрептавидину, то последний в комплексе с флуоресцентной меткой можно использовать для визуализации белков, несущих распознаваемый биотинлигазой пептид [13]. Кроме этого были обнаружены другие биотинлигазы, принимающие в качестве субстрата не только биотин, но и его алкильные и азид-содержащие производные, которые могут быть затем специфически помечены флуоресцентной меткой [14]. К сожалению, применение этой системы *in vivo* существенно затруднено в связи с наличием природных мишеней биотинлигазы в клетках.

Вскоре был разработан ряд методов на основе лигазы липоевой кислоты. Сначала было обнаружено, что помимо природного субстрата — липоевой кислоты — фермент способен также осуществлять перенос производного алифатической карбоновой кислоты, содержащей азидную группу, на остаток лизина 22-членного пептида. В дальнейшем азид может вступать в реакцию азид-алкинового циклоприсоединения с алкиновой группой, соединенной с флуорофором (рисунок 2).

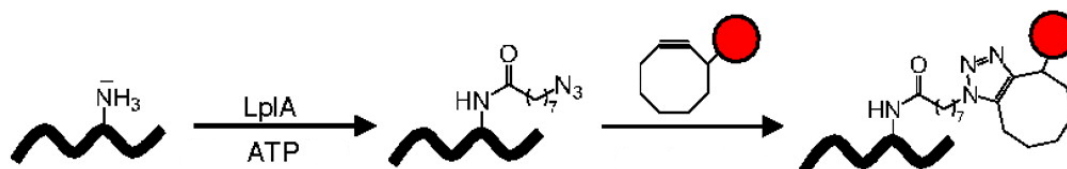


Рисунок 2. Схема двухстадийного присоединения флуоресцентной метки при помощи лигазы липоевой кислоты (LplA) [14].

Позже на основе той же лигазы липоевой кислоты был создан ряд мутантных белков, которые, пожалуй, являются самой успешной на сегодняшний день реализацией системы фермент-пептид. Сначала был получен вариант, узнающий небольшую флуоресцентную молекулу — 7-гидроксикумарин — и катализирующий её ковалентную сшивку с остатком лизина в 13-аминокислотном распознаваемом пептиде [15]. Позже, при помощи программы Rosetta, был создан второй мутант, катализирующий специфическое присоединение резорурфина к остатку лизина той же последовательности. Лигаза резорурфина может быть использована для специфического мечения белков, расположенных как в цитоплазме и ядре, так и на внешней стороне цитоплазматической мембраны [16].

Описанный подход к флуоресцентному мечению помогает совместить достоинства пептидных и белковых генетически кодируемых меток. При использовании ферментов достигается высокая специфичность, что уменьшает фоновый сигнал, а небольшой размер метки снижает влияние на функции исследуемых белков. Главным недостатком этого метода является ограниченный выбор подходящих флуорофоров, отсутствие возможности для многоцветного мечения, а также достаточно трудоемкий (если вообще возможный) процесс расширения репертуара используемых меток.

2.1.1.2. Белки

Флуоресцентные белки

Флуоресцентные белки составляют необычное семейство белков, характеризующееся высококонсервативной пространственной структурой (бачонок из 11 β -цепей, внутри которого находится α -спираль) и отличающееся от других способностью к формированию протяжённой ароматической структуры — хромофора. Хромофор формируется автокаталитически из собственных аминокислотных остатков и обеспечивает способность флуоресцентных белков к поглощению и испусканию видимого света (рисунок 3).

Реакции, приводящие к формированию хромофора, требуют лишь наличия в окружающей среде кислорода, и потому практически независимы от того, в какой системе экспрессируется ген: белок способен к сворачиванию и флуоресценции как в клетках прокариот, так и в клетках эукариот.

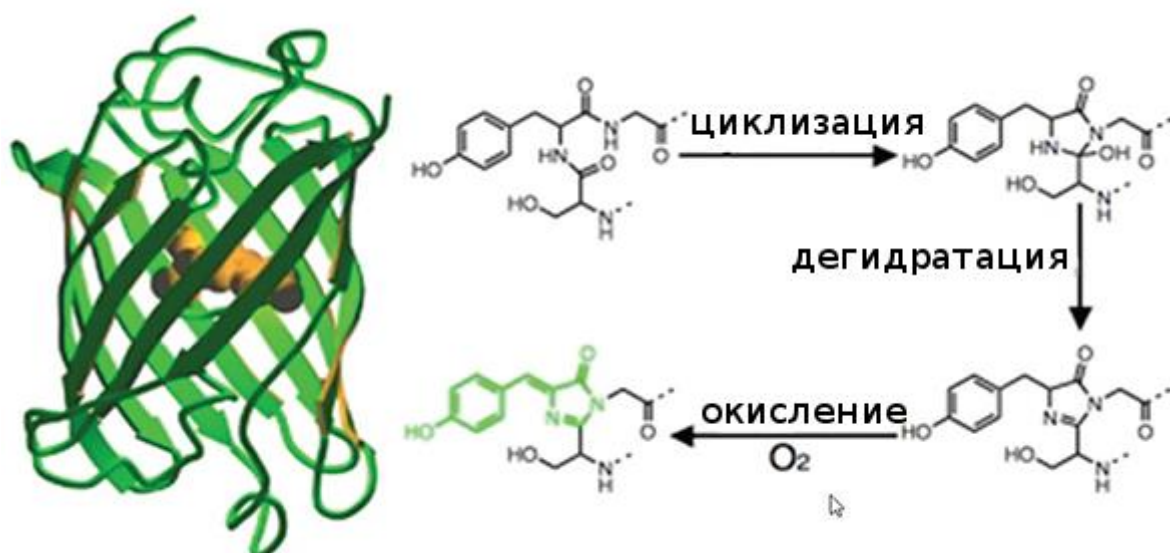


Рисунок 3. Структура флуоресцентного белка GFP и схема образования его хромофора из трех аминокислотных остатков Ser65, Tyr66 и Gly67 [17].

Первый флуоресцентный белок (зелёный флуоресцентный белок, GFP) был выделен в 1962 году Шимомурой из медузы *Aequorea victoria* при изучении ее биолюминесценции [18]. Но лишь спустя 30 лет потенциал флуоресцентных белков был осознан, когда был клонирован ген GFP [19], а затем Чалфи и коллегами было показано, что GFP может быть использован в качестве генетически кодируемой метки в гетерологических системах [20,21].

В последующие несколько лет было обнаружено, что мутагенезом GFP можно получить спектрально отличающиеся белки с цветом флуоресценции от синего до жёлтого. Например, замена 66 тирозина на другую ароматическую аминокислоту приводит к гипсохромному сдвигу (сдвигу в коротковолновую область спектра) спектров возбуждения и эмиссии флуоресценции

белков. Таким образом был получен ряд циановых (Cerulean [22], mTurquoise [23], mTurquoise2 [24] и т.д.), синих (Azurite [25], EBFP2 [26] , mTagBFP2 [27] и т.д.) и фиолетовых (Sirius [28]) вариантов. Мутация T203Y может приводить к расширению сопряжённой системы хромофора за счёт стекинг-взаимодействия фенольного кольца 203 тирозина с остатком тирозина хромофора и, как следствие, к значительному батохромному сдвигу (сдвигу в длинноволновую область спектра) спектров возбуждения и эмиссии флуоресценции. Эффект подобной замены был предсказан на основе анализа кристалла GFP и позже был подтверждён структурой жёлтого флуоресцентного белка YFP [29]. Чуть позже из кораллов были клонированы первые красные флуоресцентные белки, обладающие более протяжённой сопряжённой ароматической системой. Это открытие существенно расширило палитру доступных флуоресцентных меток [30].

На сегодняшний день для биотехнологических применений разработаны сотни мутантных белков с различными свойствами, позволяющие исследовать живые системы на самых разных уровнях. Флуоресцентные белки помогают изучать локализацию, перемещение, скорость деградации и даже возраст исследуемых белков. Спектральное разнообразие позволяет использовать сразу несколько меток. Фотоактивируемые флуоресцентные белки используют для получения изображений сверхвысокого разрешения методами флуоресцентной микроскопии. С использованием флуоресцентных белков могут проводиться и более сложные функциональные исследования: визуализация белок-белковых взаимодействий, определение активности ферментов (например, киназ или протеаз), концентрации внутриклеточных ионов, различных метаболитов и медиаторов (H^+ , Ca^{2+} , Cl^- , H_2O_2 , cAMP, и др.), а также других параметров в живых клетках и тканях

Автономность формирования хромофора во флуоресцентных белках и их невероятное разнообразие стали причиной того, что сегодня они являются одним из самых используемых инструментов для визуализации клеточных структур и процессов в биологических исследованиях. Однако они всё же не являются универсальной флуоресцентной меткой, так как имеют ряд ограничений в использовании. Так, необходимость кислорода для реакции созревания хромофора исключает использование в анаэробных условиях, значительное время созревания затрудняет изучение быстрых процессов, а сравнительно большой размер (220-240 аминокислотных остатков) не исключает влияния метки на исследуемый объект.

Самомодифицирующиеся белки

Самомодифицирующиеся белки представляют собой ферменты, способные взаимодействовать с меченым субстратом с образованием ковалентной связи.

Первая подобная метка была разработана на основе человеческого белка репарации O^6 -алкилгуанин-ДНК-алкилтрансферазы (hAGT), который катализирует необратимый перенос алкильной группы с субстрата (O^6 -алкилгуанин-ДНК) на один из своих цистеиновых остатков (рисунок 4А). Так как субстратная специфичность этого фермента довольно низкая, он хорошо реагирует с похожими соединениями, например O^6 -бензилгуанином, а также с соединениями с замещенным бензильным кольцом. Например, O^6 -бензилгуанин, несущий флуорофор в пара-положении бензильной группы, реагирует с hAGT, в результате чего флуорофор ковалентно связывается с ферментом (рисунок 4Б).

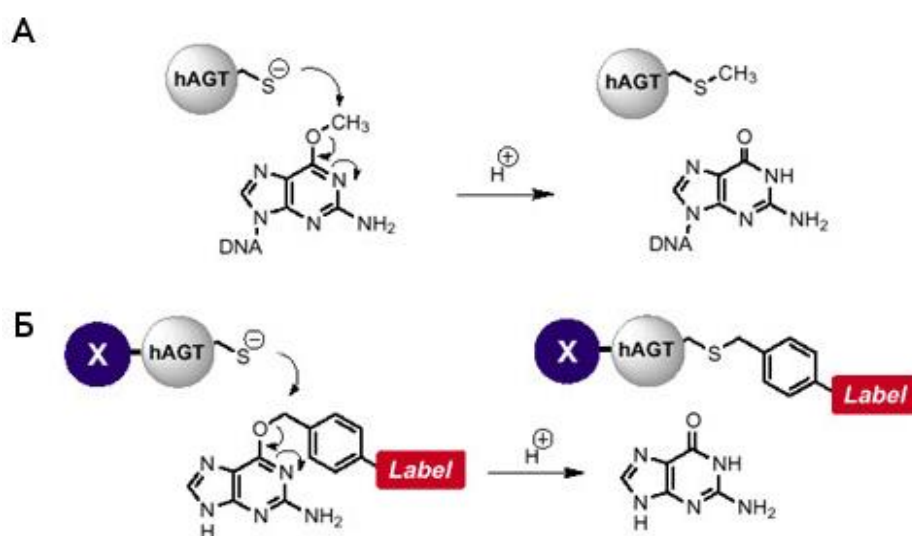


Рисунок 4. А. Механизм репарации ДНК O^6 -алкилгуанин-ДНК алкилтрансферазой (hAGT); Б. Ковалентное мечение белка слияния (X-hAGT) при помощи производного O^6 -бензилгуанина [31].

Широкому использованию данного подхода мешало взаимодействие производных O^6 -бензилгуанина с эндогенными белками репарации, вызывавшее значительное фоновое мечение. В начале для решения этой проблемы использовали линии клеток млекопитающих, дефицитные по O^6 -алкилгуанин-ДНК-алкилтрансферазе, однако это существенно сужало область возможного применения данной метки [31]. Выходом из сложившейся ситуации стало обнаружение ряда мутаций, повышающих сродство фермента к O^6 -бензилгуанину и нарушающих его взаимодействие с ДНК [32]. Последующие модификации привели к уменьшению размера метки, получившей название SNAP-tag, и дальнейшему повышению специфичности к модифицированному субстрату [33,34].

Позднее той же группой исследователей был сделан ещё один мутант O^6 -алкилгуанин-ДНК-алкилтрансферазы, получивший название CLIP-tag. Новая метка была сделана методом направленной эволюции. Она имеет высокое сродство к O^2 -бензилцитозину и в 100 раз менее эффективно взаимодействует с производными O^6 -бензилгуанина, субстратом SNAP-tag [35].

Другой сходный способ мечения основан на взаимодействии модифицированной бактериальной галогеналкандегалогеназы с синтетическими лигандами, связанными с флуоресцентными красителями. Методика получила название HaloTag. Реакция между белком и его лигандом также необратима и приводит к образованию ковалентной связи [36].

Поскольку SNAP-tag, CLIP-tag и HaloTag имеют различные субстраты, но одинаковые pK_a и схожую стабильность, они могут быть эффективно использованы для одновременного мечения различных мишеней в любом клеточном компартменте. За счёт последовательной окраски разными производными субстрата, данные метки дают возможность отличить старые белки от новосинтезированных.

Неоспоримым преимуществом использования самомодифицирующихся белков является сочетание специфичности мечения с доступностью разнообразия имеющихся химических флуорогенов и простотой использования. Однако, как и в случае с флуоресцентными белками, самомодифицирующиеся белки из-за своего большого размера способны изменить или нарушить работу исследуемого белка.

Флуорофор-связывающие белки

К данной группе генетически кодируемых меток относят белки, способные специфически связывать молекулы, способные к флуоресценции — флуорофоры (рисунок 5А). Часто флуорофор-связывающие белки относят к группе флуоресцентных белков, однако совершенно иной принцип возникновения флуоресценции побуждает некоторых авторов выделять их в отдельную группу.

Например, существует целая группа флуоресцирующих репортерных белков, связывающих флавины (FbFPs — flavin mononucleotide (FMN)-based fluorescent proteins — флуоресцентные белки на основе флавинмононуклеотида). К ним относятся белки: PpFbFP, сделанный на основе сенсорного белка SB2 из *Pseudomonas putida*; BsFbFP — N-концевой LOV-домен (light, oxygen or voltage-sensing) рецептора синего цвета YtvA из *Bacillus subtilis* — и оптимизированный для экспрессии в *E.coli* EcFbFP [37]; iLOV (improved LOV), полученный генно-инженерными методами также из LOV-домена рецептора синего света, но фототропина II из *Arabidopsis thaliana* [38]; сконструированный на основе того же белка фотосенсибилизатор miniSOG (mini Singlet Oxygen Generator) [39] и т.д. Все они имеют меньший размер, чем флуоресцентные белки, хорошую pH-стабильность, высокую “скорость созревания”, независимую от наличия кислорода, и эмиссию в зелёной области спектра [40]. Однако отсутствие цветового разнообразия и низкая яркость [41] пока препятствует полному вытеснению ими флуоресцентных белков.

Флуороген-активирующие белки

В отличие от флуорофор-связывающих белков, флуороген-активирующие белки взаимодействуют с низкомолекулярными веществами — флуорогенами, не обладающими значимой флуоресценцией в свободном виде, но приобретающие такую способность в комплексе с белком за счёт стабилизации структуры или модификации (рисунок 5Б). Использование флуорогенов позволяет избавиться от одного из недостатков флуорофор-связывающих белков — фонового сигнала несвязанного лиганда — не прибегая к отмыванию.

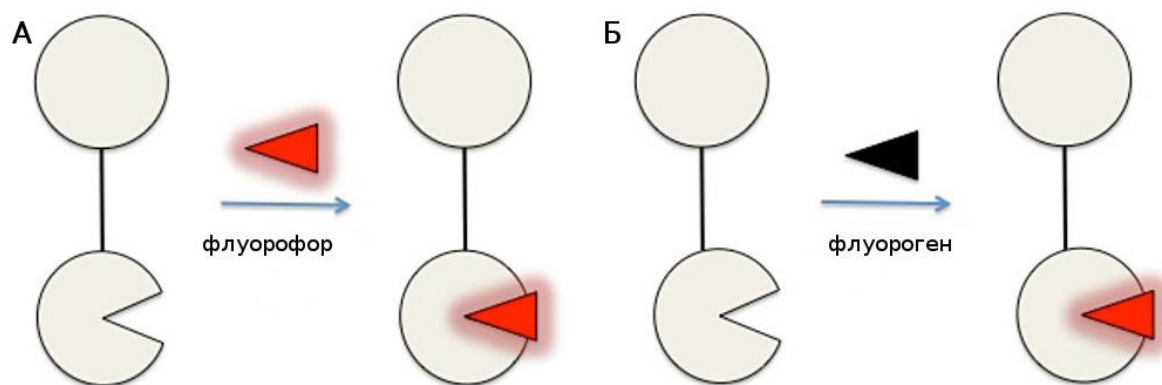


Рисунок 5. А. Принцип работы флуорофор-связывающих белков; Б. Принцип работы флуороген-активирующих белков.

В качестве примера природного флуороген-активирующего белка можно привести UnaG — белок, выделенный из мышечных клеток японского речного угря и принадлежащий к семейству белков, связывающих жирные кислоты. Флуоресценция в зелёной области спектра обеспечивается высокоаффинным нековалентным связыванием эндогенного лиганда — билирубина (рисунок 6), а следовательно, не зависит от наличия кислорода в окружающей среде [42]. Билирубин постоянно присутствует в клетках млекопитающих, поэтому UnaG всё время флуоресцентен. Однако изучение структуры этого белка может быть полезно для создания искусственных флуороген-активирующих белков, флуоресценция которых может быть индуцирована в необходимый исследователю момент добавлением не природного субстрата с заданными свойствами.



Рисунок 6. Возникновение флуоресценции при смешивании апопротеина UnaG (apoUnaG) с билирубином (BR) [42].

Ряд других флуороген-активирующих белков был сделан на основе бактериальных фитохромов. Первый был получен в результате сайт-специфического случайного мутагенеза хромофор-связывающего домена фитохрома DrBphP из *Deinococcus radiodurans*. Мономерный белок IFP1.4 связывает биливердин IX α и имеет максимум флуоресценции при 708 нм [43]. Через два года был опубликован улучшенный вариант дальне-красного белка iRFP (infrared fluorescent protein), сделанный на основе фитохрома RpBphP2 из *Rhodospseudomonas palustris*, отличающийся повышенной фотостабильностью и яркостью, но являющийся димером [44], а позже — ещё ряд вариантов: Wi-Phy [45], IFP1.4_{rev} [46] и т.д. Флуоресценция в дальне-красной области спектра позволяет эффективно использовать эти белки для мечения клеток, тканей и неинвазивного *in vivo* имиджинга животных. Однако сравнительно большой размер (более 300 аминокислотных остатков) и довольно низкий квантовый выход флуоресценции накладывают ряд ограничений на их применение.

Первые флуороген-активирующие белки, исходно не имеющие лиганда в живых системах, были сконструированы на основе человеческих одноцепочечных антител, отобранных на связывание двух известных флуорогенов — тиазолового оранжевого и малахитового зелёного. Антитело, отобранное для связывания производного тиазолового оранжевого, характеризовалось наномолярной константой связывания и увеличением флуоресценции комплекса относительно несвязанного флуорогена в 2 600 раз. Коэффициент экстинкции и квантовый выход флуоресценции комплекса были сравнимы с таковыми для широко используемого флуоресцентного белка eGFP. Комплекс антитела с производным малахитового зелёного демонстрировал 18 000-кратное возрастание флуоресценции, яркость, превышающую яркость флуоресцентного белка mCherry, и батохромный сдвиг спектров эмиссии и возбуждения флуоресценции относительно последнего на 50 нм [47]. Эти белки были успешно применены для мечения поверхностных и внутриклеточных структур, многоцветного мечения и получения изображений сверхвысокого разрешения [48].

Позднее той же группой ученых были представлены флуороген-активирующие белки, связывающие новый флуороген, имеющий максимумы возбуждения и эмиссии флуоресценции в комплексе равные, соответственно, 701 и 731 нм. Данные значения соответствуют так называемому окну прозрачности биологических объектов — области спектра, в которой поглощение света биологическими молекулами (гемоглобином, меланином) и молекулами воды минимально. Благодаря этому представленная пара флуороген-белок является перспективной для исследований на интактных животных [49].

Самой новой и, пожалуй, на данный момент самой перспективной разработкой является белок Y-FAST [50] (Yellow Fluorescence-Activating and absorption-Shifting Tag — жёлтый

активирующий флуоресценцию и сдвигающий поглощение таг), созданный на основе фотоактивного жёлтого белка (PYP) — рецептора синего цвета из *Halorhodospira halophila* [51]. Этот белок массой всего 14 кДа быстро и обратимо связывает свободно проходящий через мембрану живых клеток 4-гидрокси-3-метилбензилиденроданин, что приводит к батохромному сдвигу максимума абсорбции на 80 нм и увеличению квантового выхода флуоресценции примерно в 800 раз (рисунок 7).

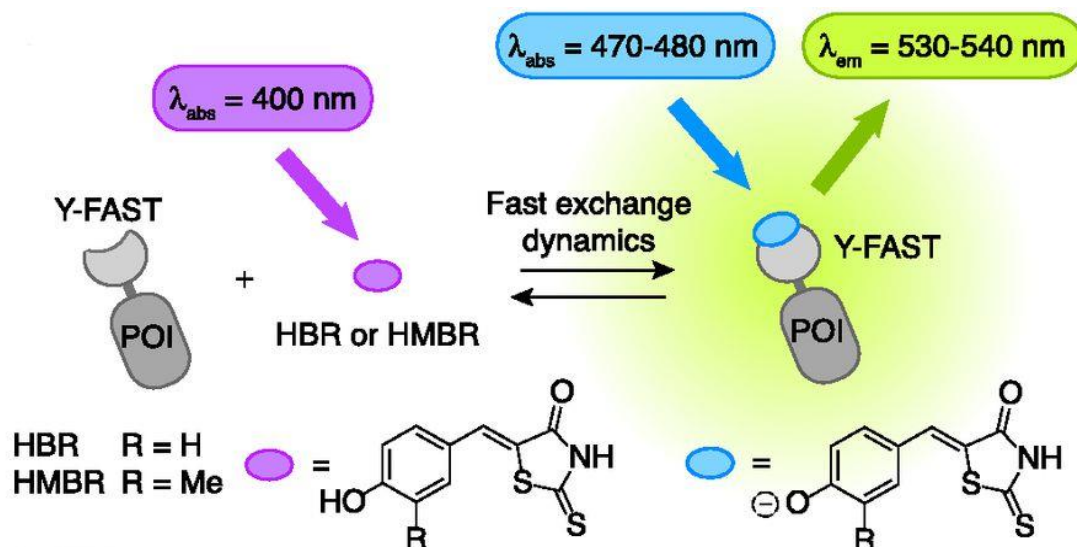


Рисунок 7. Принцип работы белка Y-FAST [50].

В связи со сравнительно недавним созданием первых флуороген-активирующих белков и пока что небольшим количеством работ, сделанных при помощи флуоресцентных меток такого типа, в настоящий момент довольно сложно говорить о преимуществах и недостатках данного метода. Представленное на сегодняшний день цветовое разнообразие невелико, однако существующие варианты уже демонстрируют эмиссию в красной и дальне-красной областях спектра, а также не имеют явных серьезных ограничений на расширение имеющейся палитры.

2.1.1.3. Неприродные аминокислоты

Неприродные аминокислоты — это аминокислоты, которые не кодируются генетически и в норме не включаются в белки в процессе трансляции. Существует три основных способа введения неприродных аминокислот в полипептидную цепь:

1. Обычная аминоацил-тРНК синтетаза может катализировать присоединение неприродной аминокислоты, аналогичной канонической, к тРНК. При культивировании бактерий, ауксотрофных по данной канонической аминокислоте, в присутствии её аналога, аналог будет замещать эту каноническую аминокислоту во всех белках. Данный метод был использован, например, для получения меченного флуоресцентным аналогом триптофана человеческого аннексина А5 [52].

2. *In vitro* тРНК может быть соединена с любой неприродной аминокислотой. Если синтез белка будет происходить в присутствии таких “неправильно заряженных” тРНК, то неприродные аминокислоты будут встраиваться в белок при трансляции кодона использованной тРНК. Этот метод обычно применяется при синтезе белков в бесклеточных системах, однако также показана возможность инъекции “неправильно заряженных” тРНК в ооциты лягушки [53].

3. Самым элегантным способом является создание специальных пар тРНК — аминоацил-тРНК синтетаза, распознающих неприродные аминокислоты. Если новая аминокислота кодируется “свободным” кодоном (кодоном, не используемым для кодирования одной из канонических аминокислот), то белки с неприродными аминокислотами могут синтезироваться обычными клетками, снабженными дополнительной парой тРНК — аминоацил-тРНК синтетаза в присутствии неприродных аминокислот. Этот подход был успешно реализован как в бактериях, так и в клетках млекопитающих [54]. Тем не менее, высокая концентрация свободных неприродных аминокислот, необходимая для данного метода, является существенной проблемой в случае введения флуоресцентной метки, поэтому в таком качестве в настоящее время практически не используется.

Использование неприродных аминокислот позволяет осуществлять как прямое, так и не прямое введение флуоресцентных меток. Прямое введение метки предполагает использование флуоресцентных неприродных аминокислот (рисунок 8), не прямое — неприродных аминокислот с непротеиногенными химически активными группами (например, кетонными), которые позволяют осуществлять селективное, сайт-специфическое мечение уже после синтеза белка.

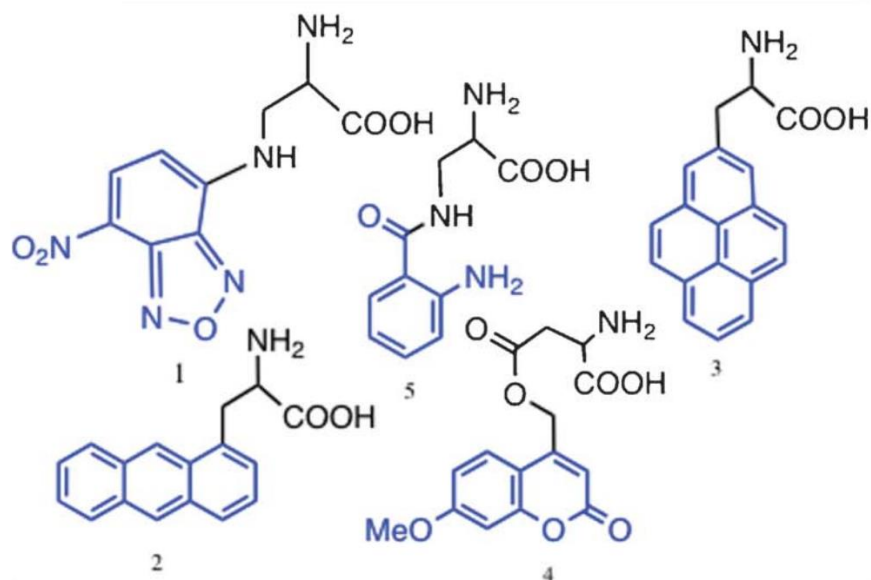


Рисунок 8. Примеры флуоресцентных неприродных аминокислот: 3-(4-нитробензooksадиазол-7-иламино)-2-аминопропионовая кислота (1), 2-амино-3-(антрацен-5-ил)пропионовая кислота (2), 2-амино-3-(пирен-2-ил)пропионовая кислота (3), аминокислоты, несущие 7-метоксикумарин (4) и б-антранилоил (5). Синим цветом обозначена флуоресцентная группа [7].

К сожалению, флуоресцентное мечение при помощи неприродных аминокислот в настоящее время имеет ряд существенных недостатков, таких как низкая специфичность мечения, трудоемкость при использовании в клетках млекопитающих, сложность перехода к новой метке, высокий фоновый сигнал и т.д. В связи с этим данный метод пока что не нашел широкого применения.

2.1.2. Химические метки

Химические красители имеют ряд преимуществ перед генетически кодируемыми метками. Среди них — меньший размер (редко больше 1 кДа), большее разнообразие цветов, отсутствие периода созревания, зачастую большая фотостабильность. Однако бывают и существенные недостатки: токсичность, фототоксичность, плохое проникновение сквозь клеточную мембрану. Главной же проблемой химического мечения остается низкая специфичность взаимодействия и высокий фоновый сигнал.

2.1.2.1. Неспецифическое мечение

Для неспецифического химического флуоресцентного мечения используют реакции модификации функциональных групп канонических аминокислотных остатков, поэтому оно применимо лишь для *in vitro* исследований.

Среди широко используемых методов можно отметить мечение тиоловых групп биомолекул при помощи флуорофоров, содержащих малеимидную группу, способную образовывать стабильную тиоэфирную связь [55]; мечение аминокислотных остатков (ε-аминогрупп остатков лизина и N-концевых аминокислот) с использованием флуоресцентных конъюгатов с N-гидроксисукцинимидным эфиром [56] и реакции с азидной группой, введенной в биомолекулу постсинтетически, используя ковалентные ингибиторы или модифицированные предшественники [57].

2.1.2.2. Специфическое мечение

Лишь в 2014 году была показана практическая возможность создания высокоспецифичной в живых системах, нетоксичной, флуорогенной химической метки, имеющей максимумы возбуждения и эмиссии флуоресценции в дальне-красной области спектра. Для этой цели в качестве флуорогена были использованы производные кремний-родамина (SiR), которые обычно находятся в равновесии между двумя формами: флуоресцентной (открытой) и нефлуоресцентной (закрытой) (рисунок 9А). Агрегация или связывание с гидрофобными поверхностями стабилизирует нефлуоресцентную форму соединения, в то время как взаимодействие с полярной поверхностью белковых молекул благоприятствует открытой конформации. Флуороген был соединен с лигандами, которые связываются с микротрубочками и F-актином. После перебора различных линкеров, были созданы варианты SiR-tubulin и SiR-actin (рисунок 9Б), флуоресцентный сигнал которых при взаимодействии с белком-мишенью возрастает более чем в 10 и 100 раз соответственно.

SiR-tubulin и SiR-actin пригодны для мечения интактных живых клеток. Это особо ценно для изучения тех клеток, которые практически не поддаются трансфекции и, как следствие, не могут быть изучены при помощи генетически кодируемых меток.

Фотостабильность SiR-tubulin и SiR-actin делает их пригодными для STED-микроскопии (Stimulated Emission Depletion Microscopy — микроскопия истощения флуоресценции вынужденным излучением). Используя SiR-tubulin, ученым удалось получить изображения микротрубочек и centrosom с наивысшим на сегодняшний день разрешением для живых клеток, что ещё раз подчеркивает перспективность использования низкомолекулярных соединений, непосредственно взаимодействующих с изучаемой биомолекулой [58].

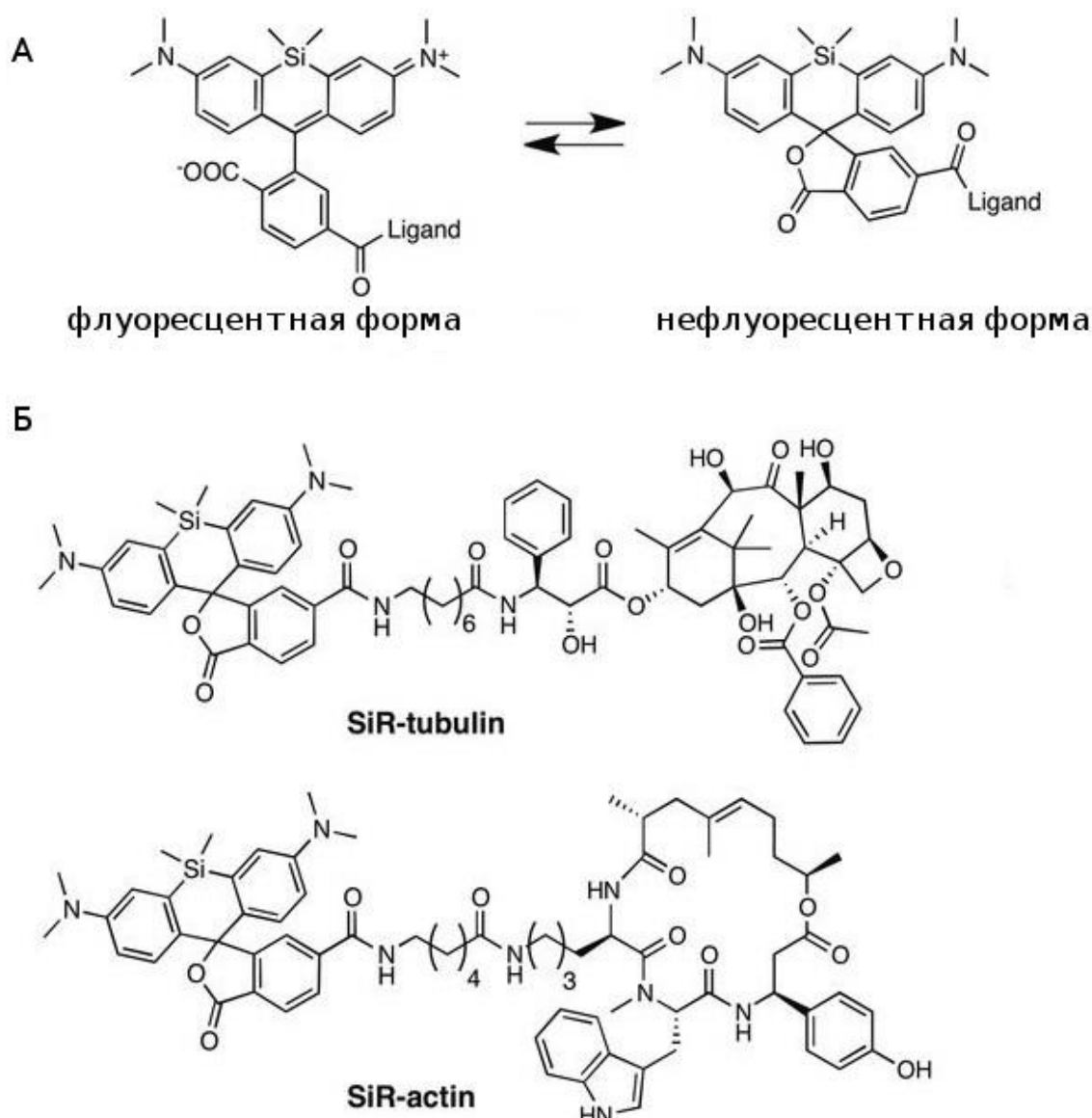


Рисунок 9. Производные кремний-родамина (SiR) **А.** Равновесие между флуоресцентной и нефлуоресцентной формами в производных кремний-родамина; **Б.** Структуры красителей SiR-tubulin и SiR-actin (показана только закрытая форма).

Несмотря на наличие приятных исключений, таких как SiR-actin и SiR-tubulin, особых предпосылок для появления системы простого и быстрого подбора подобных молекул для произвольного белка пока нет. Поэтому сегодня основным направлением в области улучшения методов флуоресцентного мечения живых систем всё же остается область генетически кодируемых меток.

2.2. Методы молекулярной эволюции белков

Мутации, накапливавшиеся в течение многих тысячелетий, обеспечили огромное разнообразие существующих живых организмов и, соответственно, наличие природных белков с самыми различными функциями. Однако скорость спонтанного мутагенеза в живых системах недостаточно велика, чтобы удовлетворить всё возрастающие потребности по созданию новых белков для прикладного использования в биологии. В связи с этим за последние десятилетия было разработано множество подходов, облегчающих задачу получения белков с заданными свойствами.

Основной стратегией методов молекулярной эволюции белков является получение библиотеки, в которой в дальнейшем ищут варианты с требуемыми свойствами. Так как перебор всех аминокислотных сочетаний даже в сравнительно небольшом фрагменте белка физически невозможен, все эти методы ставят своей задачей получение библиотеки с максимально равномерным и полным покрытием возможного разнообразия.

Если известны положения аминокислот, которые с большой вероятностью определяют функцию белка (например, такая информация может быть получена из анализа кристаллической структуры модифицируемого белка, сравнения с гомологичными белками или компьютерными методами), то прибегают к различным методам сайт-специфического мутагенеза. При отсутствии таких данных обычно обращаются к методам случайного мутагенеза. Кроме того, в любом из перечисленных случаев можно воспользоваться тем или иным методом *in vitro* рекомбинации.

2.2.1. Случайный мутагенез

Первые протоколы случайного мутагенеза основывались на использовании различных химических и физических агентов, индуцирующих повреждения в молекулах ДНК. Так, было показано применение азотистой кислоты (вызывает дезаминирование цитозина, аденина и гуанина), муравьиной кислоты (вызывает депуринизацию ДНК), гидразина (вызывает депиримидинизацию ДНК) [59], этилметансульфоната (способствует алкилированию гуанина и тимина, что приводит к транзициям АТ→GC и GC→АТ) [60], ультрафиолетового облучения [61] и т.д.

Другим подходом стало создание специальных штаммов для осуществления *in vivo* мутагенеза за счёт увеличения количества ошибок при репликации ДНК. Например, широко используются и коммерчески доступны *E. coli* штамма XL1-Red, содержащие мутации в ϵ -субъединице ДНК-полимеразы III, обладающей 3'-5' экзонуклеазной активностью (mutD [62]), а также в ферментах, ответственных за исправление неправильно спаренных оснований (mutS [63]) и специфическую деградацию 8-оксо-7,8-дигидрогуанина (mutT [64]), которые демонстрируют

примерно 5 000-кратное увеличение частоты спонтанных мутаций в сравнении со штаммом дикого типа [65]. К сожалению, в таких бактериях с одинаковой эффективностью изменяются как плазмидная, так и геномная ДНК, что плохо подходит для создания библиотек мутантов какого-то конкретного гена. Это ограничение было преодолено при помощи использования *Saccharomyces cerevisiae* с ортогональной цитоплазматической репликационной системой из *Kluyveromyces lactis* [66], состоящей из линейной двухцепочечной высококопийной плазмиды и специфической полимеразы TP-DNAPI (TP(terminal protein)-primed DNA polymerase), использующей белок в качестве затравки для репликации. Специфическое праймирование вместе с пространственным разделением приводит к тому, что изменение точности TP-DNAPI полимеразы не влияет на качество репликации геномной ДНК, а значит позволяет мутировать лишь исследуемую область, помещенную в плазмиду. Позже был предложен похожий подход, использующий модифицированный ретротранспозон *Ty1* в качестве независимо копируемого элемента [67].

Перечисленные выше методы имеют ряд существенных недостатков. Так, химические и физические агенты небезопасны для человека, а получаемый с их помощью спектр мутаций зачастую смещен в сторону каких-то определенных замен; *in vivo* мутагенез демонстрирует довольно низкую частоту мутаций и его сложно контролировать. Несмотря на это, многие из описанных подходов до сих пор развиваются и успешно используются для направленной эволюции целых организмов [68,69], однако они значительно вытеснены предложенными позднее *in vitro* методами и редко применяются для работы с отдельными белками.

Протоколы случайного *in vitro* мутагенеза, известные под названием “ПЦР с ошибками”, основаны на создании специальных условий (например, за счёт увеличения концентрации магния [70], добавления в реакционную среду солей марганца [71] или аналогов нуклеотидов (например, 8-оксогуанина [72], бромуридина [73], 2-гидроксиаденина [74] и т.д.)), при которых ДНК-полимераза совершает множество ошибок при ПЦР-амплификации гена, кодирующего исследуемый белок, или использовании специальных ДНК-полимераз с пониженной точностью (например, Mutazyme). При этом, варьируя концентрации дополнительных компонентов в ПЦР-смеси, количество матрицы и число циклов амплификации, можно легко регулировать среднее число вносимых мутаций и изменять его в зависимости от конкретной задачи. Несмотря на то, что проблема смещения спектра мутаций для ПЦР с ошибками также существует, использование несбалансированных смесей дезоксирибонуклеозидтрифосфатов [75] или сочетаний различных полимераз [76] существенно уменьшает дисбаланс в частоте вносимых замен.

Полностью избавиться от различий в частоте вносимых мутаций позволяет предложенный в 2004 году метод SeSaM (sequence saturation mutagenesis — насыщающий мутагенез) [77].

Протокол состоит из четырёх стадий (рисунок 10). Сначала проводят ПЦР-амплификацию гена, используя биотинилированный прямой и обычный обратный праймеры и смесь обычных и α -фосфотиоат-содержащих дезоксирибонуклеозидтрифосфатов. В отличие от фосфоэфирной связи, фосфотиоатная чувствительна к действию иодидов [78], поэтому обработка ПЦР-продукта спиртовым раствором йода с последующими выделением биотинилированных цепей на стрептавидиновой смоле и денатурацией ДНК приводит к получению библиотеки одноцепочечных фрагментов мутируемого гена разной длины. Последующая обработка терминальной трансферазой позволяет присоединить на 3'-конец каждого из фрагментов универсальное основание (универсальными основаниями называют соединения, которые могут заменить любое из 4 оснований ДНК, значительно не изменяя свойства и геометрию получаемого олигонуклеотида. В качестве универсального основания могут быть использованы, например, инозин, 3-нитропиррол, 5-нитроиндол и т.д. [79]). На следующей стадии каждый из фрагментов достраивают до полноразмерного гена, используя одноцепочечную ДНК в качестве матрицы, после чего синтезируют комплементарную цепь, которая содержит случайные замены напротив универсального основания. Полученную двухцепочечную ДНК амплифицируют стандартным образом для избавления от аналогов нуклеотидов и используют для последующего клонирования.

Хотя метод SeSaM, несомненно, приводит к получению максимально приближенных к идеалу мутантных библиотек, простота, доступность, разнообразие и дешевизна протоколов “ПЦР с ошибками” делает их, пожалуй, самым используемым на сегодняшний день способом случайного *in vitro* мутагенеза.

2.2.2. Сайт-специфический мутагенез

Для проведения сайт-специфического случайного мутагенеза используют синтетические ДНК-олигонуклеотиды, которые внедряют в необходимый участок гена либо при помощи ПЦР [80], либо используя уникальные сайты рестрикции [81].

Самый простой способ — использование олигонуклеотидов с вырожденными кодонами NNN или NNB (где N — любое из 4 оснований, B — G, C или T) — дает возможность получить любую из 20 основных аминокислот, но также может привести к включению стоп-кодона. Правда, из-за вырожденности генетического кода, получаемые библиотеки обогащены одними аминокислотами (лейцином, аргинином, серином и т.д.) и практически не содержат других (метионина, триптофана). Кроме того, дисбаланс в частоте вносимых замен экспоненциально возрастает с увеличением количества сайтов, подвергаемых мутагенезу, а требуемый размер библиотеки очень быстро начинает превышать максимально достижимый любым из известных на данный момент методов.

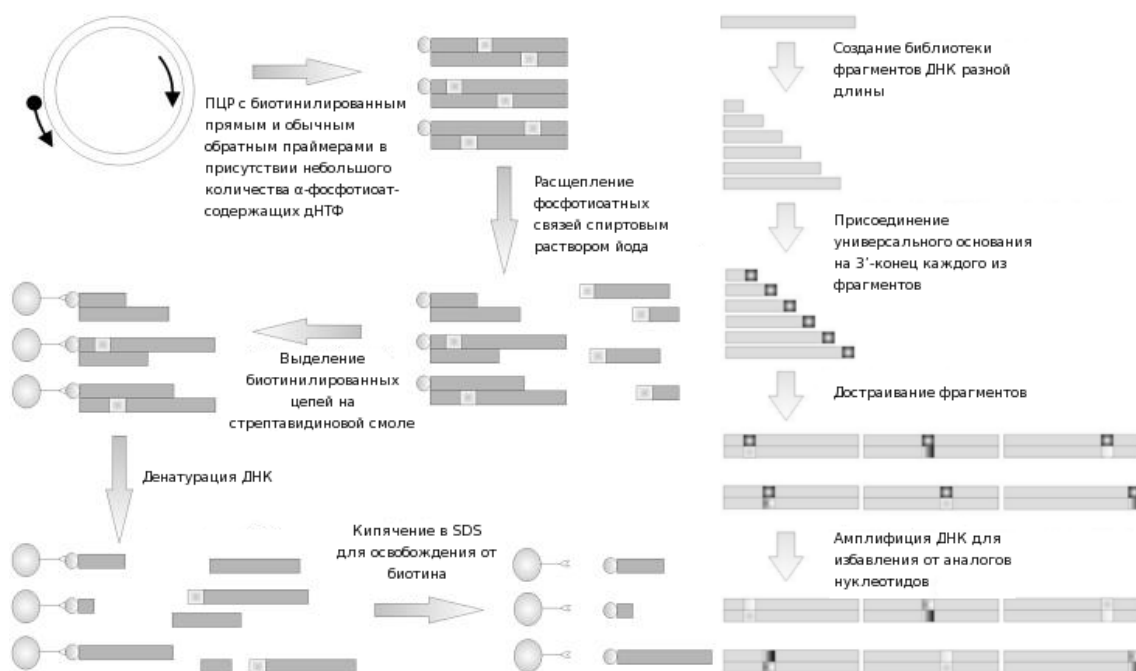


Рисунок 10. Протокол случайного *in vitro* мутагенеза SeSaM [77].

Различий в частоте вносимых мутаций и появления стоп-кодонов можно избежать несколькими способами. Один из подходов основан на использовании множества отдельных колонок для синтеза олигонуклеотидов. На каждом этапе на каждой из колонок к олигонуклеотиду последовательно присоединяют три основания, которые кодируют определенную аминокислоту. После синтеза каждого кодона, носитель из всех колонок смешивают и снова распределяют по колонкам для синтеза следующего кодона [82]. Альтернативой является использование для синтеза олигомеров динуклеотид [83] или тринуклеотид [84] фосфорамидитов в качестве мономеров. Однако из-за низких выходов при синтезе мономеров, а также из-за их различной реакционной способности эти методы являются дорогостоящими и трудоемкими, что служит причиной ограниченности их применения.

Возможность одновременного внесения случайных мутаций в несколько удаленных друг от друга участков гена была показана в методе OmniChange. Для этого использовали специальные олигонуклеотиды состава 5'-(nt*)₁₂-NNK-(nt)₁₂-3', где 12 5'-концевых нуклеотидов (nt*) соединены между собой фосфотиоатной связью, а все остальные (nt) — обычной фосфоэфирной. Амплифицированные при помощи этих праймеров фрагменты ДНК обрабатывали спиртовым раствором йода, что приводило к расщеплению фосфотиоатных связей и образованию выступающих 5'-концов длиной 12 нуклеотидов. Полученные фрагменты гибридизовали и использовали для трансформации (рисунок 11) [85].

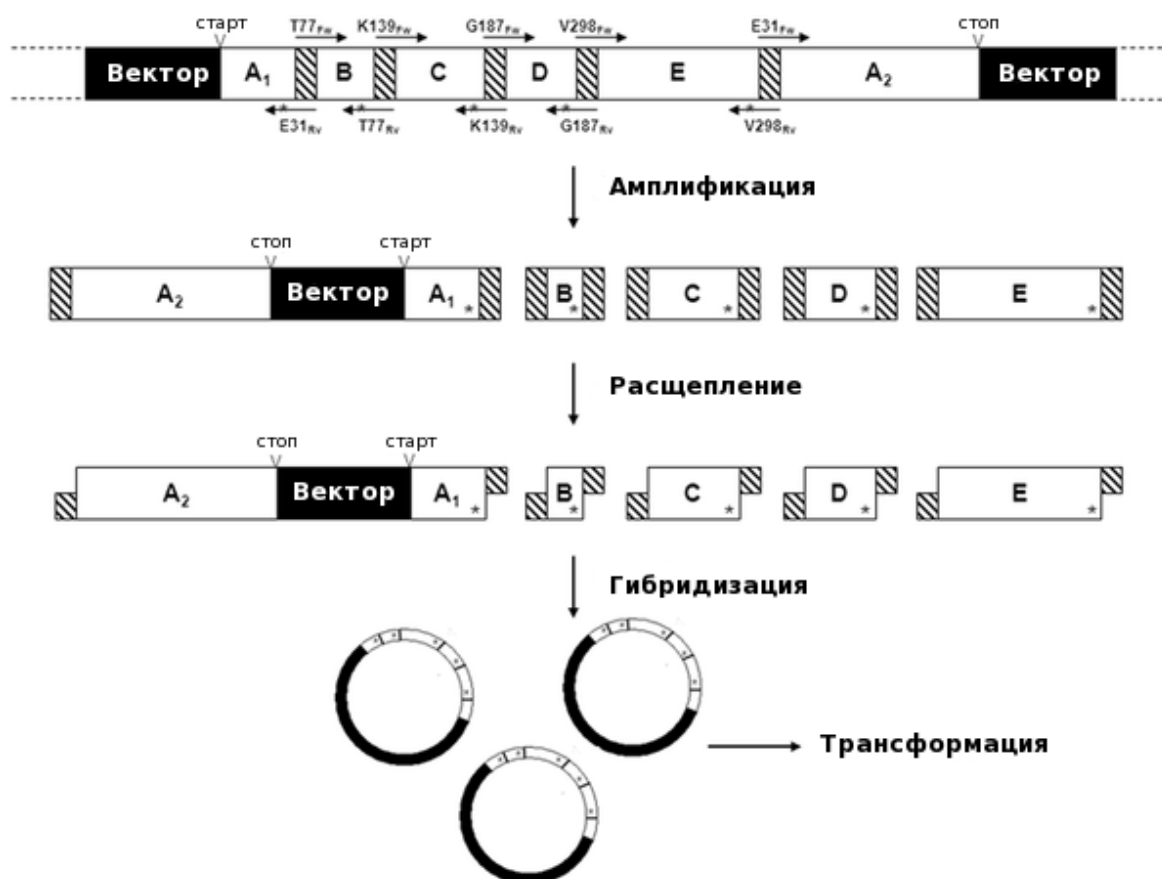


Рисунок 11. Одновременное внесение случайных замен в пять позиций методом OmniChange [85]

Позже был описан ряд способов, позволяющих добиться равномерного покрытия всего аминокислотного разнообразия за счёт использования смесей обычных олигонуклеотидов. Так, для перебора всех возможных вариантов в одной позиции достаточно четырёх праймеров, содержащих, соответственно, два вырожденных кодона — NDT и VMA (где N — любое из 4 оснований; D — A, G или T; V — A, C или G; M — A или C) и два специфических — ATG и TGG (для метионина и триптофана соответственно), взятых в соотношении 12:6:1:1 [86]. Однако для мутагенеза двух положений понадобится уже 16 праймеров, трех — 64 и т.д, поэтому рандомизацию в большом количестве сайтов легче, удобней и дешевле проводить, используя один из множества существующих в данный момент методов клонирования на основе гомологичной сборки ДНК, используя праймеры, содержащие вырожденные кодоны. Например, такая возможность была показана при тестировании метода MUPAC (multiple patch cloning — множественное лоскутное клонирование) [87]. Для того же, чтобы эффективно вносить замены в близко или подряд расположенные сайты, было предложено объединить ранее опубликованные подходы: использовать 4 колонки, на каждой из которых синтезируется одна из 4 вышеупомянутых последовательностей (NDT, VMA, ATG или TGG), после чего носитель перемешивать и повторять процедуру необходимое количество раз [88].

2.2.3. *In vitro* рекомбинация

Рекомбинация генов, несущих разнообразные мутации, является одним из основных механизмов появления белков с новыми свойствами в процессе эволюции. Этот естественный процесс можно имитировать и *in vitro*, ускоряя и модифицируя его в зависимости от поставленной задачи.

Первый опубликованный протокол *in vitro* рекомбинации включал в себя неспецифическое расщепление гомологичных генов при помощи ДНКазы I, после чего выделенные короткие фрагменты денатурировали и амплифицировали, используя их одновременно и как матрицу, и как праймеры. Получившиеся гибридные полноразмерные гены клонировали и исследовали на наличие искомым функций (рисунок 12А) [89]. Похожий протокол, отличающийся лишь способом обратной сборки фрагментов, был опубликован позже и известен под названием RACHITT (random chimeragenesis on transient template — случайный химерогенез на временной матрице). В этом случае для получения полноразмерных гибридных молекул в качестве матрицы для отжига фрагментов используют одноцепочечную урацил-содержащую ДНК. После удаления свободных 5'- и 3'-концов фрагментов, достраивания промежутков и лигирования, получившийся гетеродуплекс обрабатывают урацил-ДНК-гликозилазой, что делает исходную матрицу непригодной для амплификации. При помощи ПЦР синтезируют вторую цепь, и химерные гены используют для клонирования (рисунок 12Б) [90]. Хотя варьирование концентрации ДНКазы I и времени обработки ферментом дает возможность регулировать размер получаемых ДНК-фрагментов и, соответственно, частоту рекомбинации, позднее предложенный метод, основанный на добавлении уридина в смесь дезоксирибонуклеозидтрифосфатов с последующей обработкой ПЦР-продуктов урацил-ДНК-гликозилазой и расщеплением пиперидином, позволяет делать это ещё эффективней [91]. Использование синтетических олигонуклеотидов для тех же целей также возможно [92,93].

In vitro рекомбинацию также можно проводить без предварительной фрагментации ДНК. Так, в методе StEP (staggered extension process — метод колеблющегося удлинения) было предложено использовать модифицированную ПЦР. Используя в качестве матрицы смесь гомологичных генов и сокращая время элонгации, можно добиться того, что за один цикл будет синтезироваться лишь небольшая часть гена, а последующие денатурация и новый отжиг будут способствовать смене матрицы (рисунок 12В) [94].

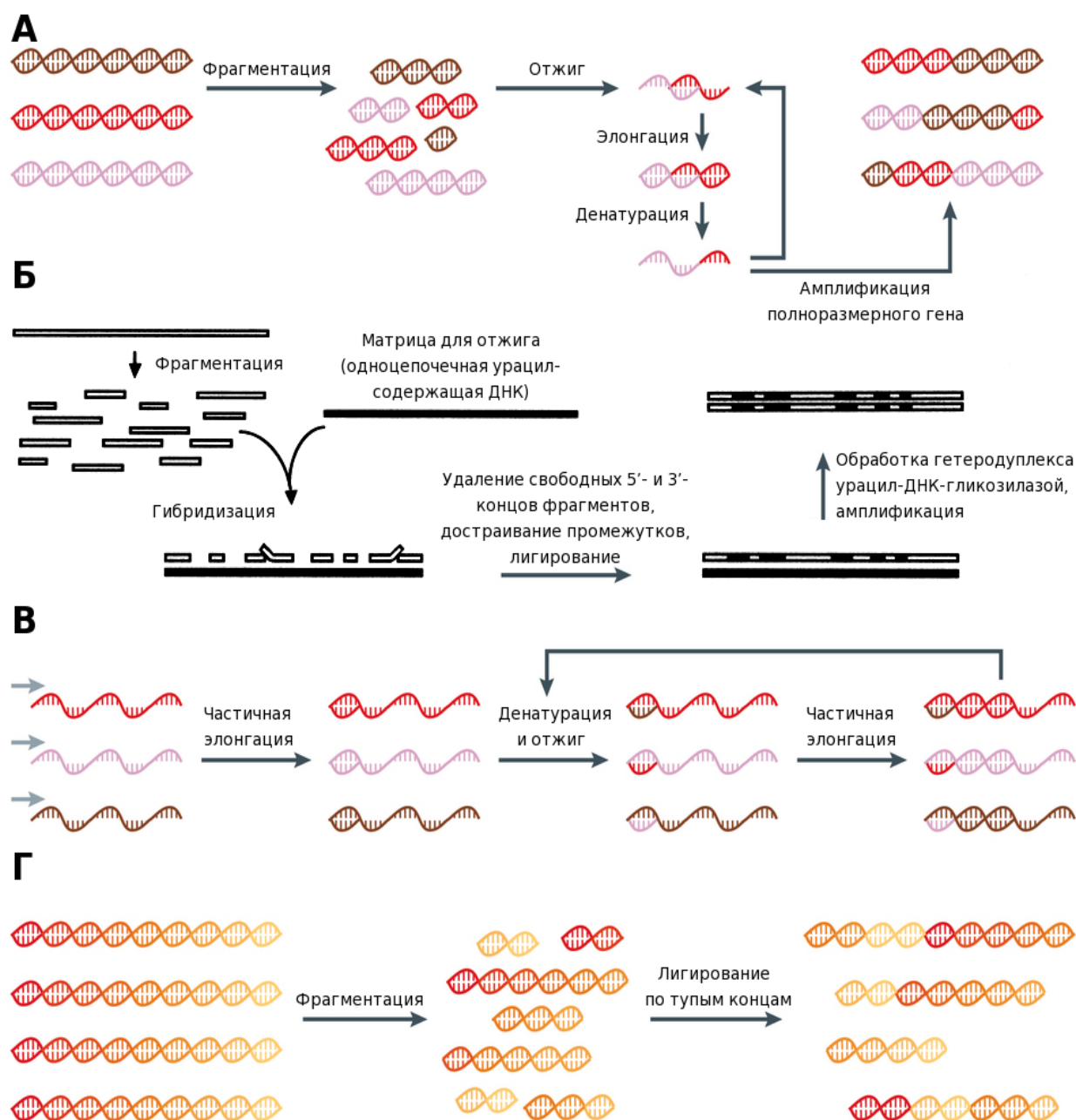


Рисунок 12. Методы *in vitro* рекомбинации. **А.** Неспецифическое расщепление гомологичных генов; **Б.** RACHITT (random chimeragenesis on transient template — случайный химерогенез на временной матрице); **В.** StEP (staggered extension process — метод колеблющегося удлинения); **Г.** NRR (nonhomologous random recombination — негомологичная случайная рекомбинация) [90,95].

Возможности *in vitro* рекомбинации не ограничиваются только гомологичными белками: в качестве матрицы также можно использовать один белок, белки со схожими функциями, но разными аминокислотными последовательностями, или несколько совершенно разных белков. В целом все эти методы основаны на фрагментации исходных генов с их последующим сиквенс-независимым лигированием. Отличия же состоят в количестве генов, участвующих в рекомбинации, частоте рекомбинации, возможности регулировать длину конечного продукта и т.д. Например, первые варианты методов позволяли делать химеры заданной [96] или случайной [97] длины комбинацией всего двух белков в строго заданной последовательности. Позднее

опубликованный подход NRR (nonhomologous random recombination — негомологичная случайная рекомбинация) позволяет соединять фрагменты из неограниченного количества исходных генов в случайном порядке и контролировать длину продукта рекомбинации [98].

Кроме *in vitro* методов для создания новых белков путём рекомбинации можно использовать и *in vivo* системы, позволяющие проводить селекцию полученных конструкций без дополнительных процедур клонирования и трансформации, приводящих к значительным потерям на каждом цикле отбора. Например, такая возможность была показана на дрожжах *S. cerevisiae*. Для этого использовали целевую плазмиду, библиотеку донорных кассет и индуцируемую эндонуклеаз-зависимую систему рекомбинации. Сначала дрожжи, несущие целевую плазмиду, спаривали с дрожжами, содержащими плазмиду с одной из донорных кассет, после чего индуцировали рекомбинацию и выбирали варианты с необходимыми свойствами. При необходимости отобранные особи можно размножить и провести новый цикл [99].

2.3. Моделирование в биологии

Моделирование в биологии применяется для изучения структур, функций и процессов на всех уровнях организации живого: от молекулярного до популяционного. Все используемые модели можно отнести к одной из трёх больших групп: биологические, физико-химические и математические.

Если биологические модели (экспериментальные системы, воссоздающие ключевые аспекты исследуемого более сложного объекта) широко распространены и используются в биологии практически повсеместно, то применение остальных видов моделирования зачастую связано с различными трудностями: поиском коллабораторов или освоением методов другой области. Тем не менее, именно использование “гетерологичных” подходов (например, небιологических методов в биологии) зачастую приводит к нахождению ответов на старые вопросы или созданию совершенно новых отраслей и технологий.

Ниже рассмотрены более подробно лишь два частных примера небιологического моделирования, использованных в рамках данной работы: изучение модельных соединений (физико-химическое моделирование) и использование компьютерных подходов для создания белков с новыми свойствами (математическое моделирование).

2.3.1. Модельные соединения

Химики с их врожденной тягой к пониманию процессов на молекулярном уровне часто изучают биологические системы при помощи модельных соединений.

В прошлом потребность в модельных соединениях была очевидной: детальное изучение сложных биологических систем доступными в то время методами было просто невозможно. Сегодня это не так: новые технологии изучения биомолекул (такие как, например, ЯМР и кристаллография) позволили сделать огромный шаг вперед. Тем не менее, модели по-прежнему могут обеспечивать исследователей ценной информацией, которую трудно получить, изучая объект напрямую. Действительно, большие биологические системы характеризуются множеством взаимосвязанных параметров, которые сложно отделить друг от друга. Модели же позволяют сосредоточиться на одном конкретном аспекте биологической структуры, комплекса или пути. Кроме того, всё ещё существует множество объектов, для которых прямое изучение до сих пор не доступно. Таким образом, модельные системы остаются важными для понимания биологии, просто типы изучаемых молекул и задаваемые вопросы становятся сложнее.

В качестве примера успешного использования модельных соединений для изучения биологических вопросов можно привести исследование структуры ДНК. Например, для изучения

влияния электростатических эффектов (в частности, водородных связей между нуклеотидами) на структуру и свойства нуклеиновых кислот был синтезирован ряд неполярных аналогов нуклеозидов [100]. Некоторые из этих аналогов (например, 4-метилбензимидазол и 2,4-дифтортолуол (рисунок 13)) были достаточно близки по размеру и форме к природным азотистым основаниям, что позволяло эффективно отделять электростатические эффекты от стерических [101–103].

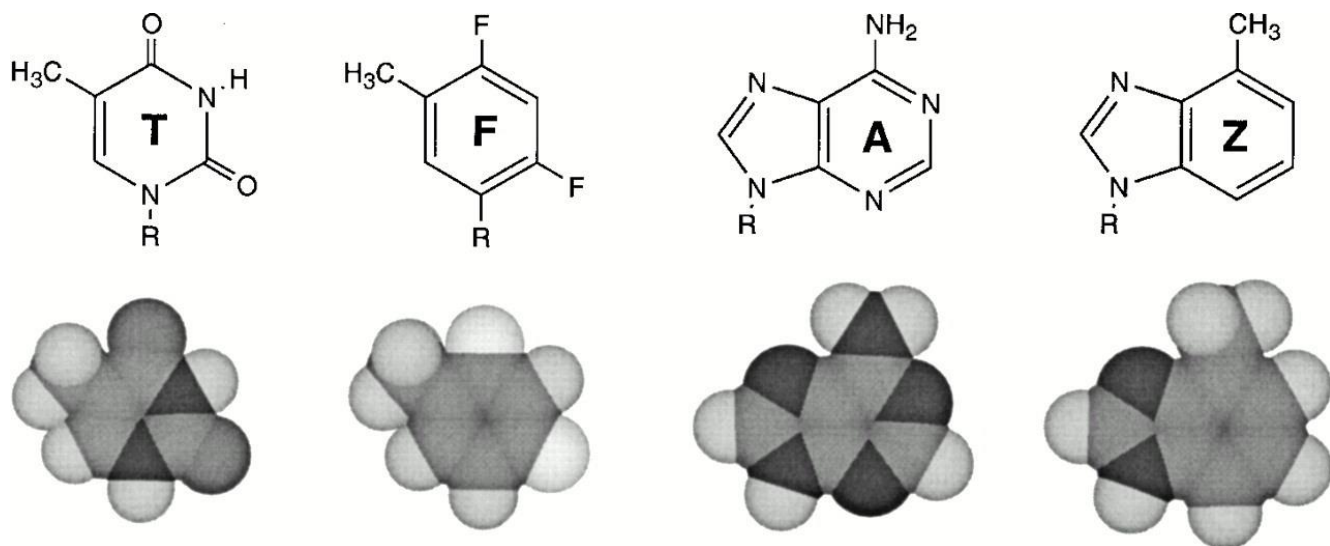


Рисунок 13. Структуры двух природных азотистых оснований: аденина (A) и тимина (T), и их неполярных аналогов: 4-метилбензимидазола (Z) и 2,4-дифтортолуола (F) соответственно [104].

Благодаря активному развитию и совершенствованию методов синтетического и полусинтетического получения полипептидов, сегодня стало возможным изучение не только небольших модельных соединений, но и целых модельных белков. Это позволило, например, начать активные исследования роли различных посттрансляционных модификаций белков (например, гликозилирования [105] или убиквитинилирования [106]), которые раньше были затруднены из-за невозможности выделения достаточного количества гомогенно модифицированного субстрата из биологических образцов. Также на примере синтеза аденин-содержащего аналога РНКазы А был продемонстрирован новый подход к исследованию ферментативного катализа, не доступный в рамках традиционной биохимии или молекулярной биологии [107].

Метод синтеза модельных соединений был применен и для изучения флуоресцентных меток, в частности, флуоресцентных белков. Как уже было упомянуто, хромофор таких белков формируется автокаталитически из трёх аминокислотных остатков, расположенных в центре β -бочонка. При этом он остаётся ковалентно связанным (с одной или с двух сторон в зависимости от типа хромофора) с соседними аминокислотными остатками, а также образует ряд водородных связей с отдалённо расположенными аминокислотами и молекулами воды (рисунок 14). Эти взаимодействия могут влиять как на спектральные свойства белка, так и в целом на способность

белка формировать хромофор, поэтому исследования методами мутагенеза не всегда оказываются информативными.

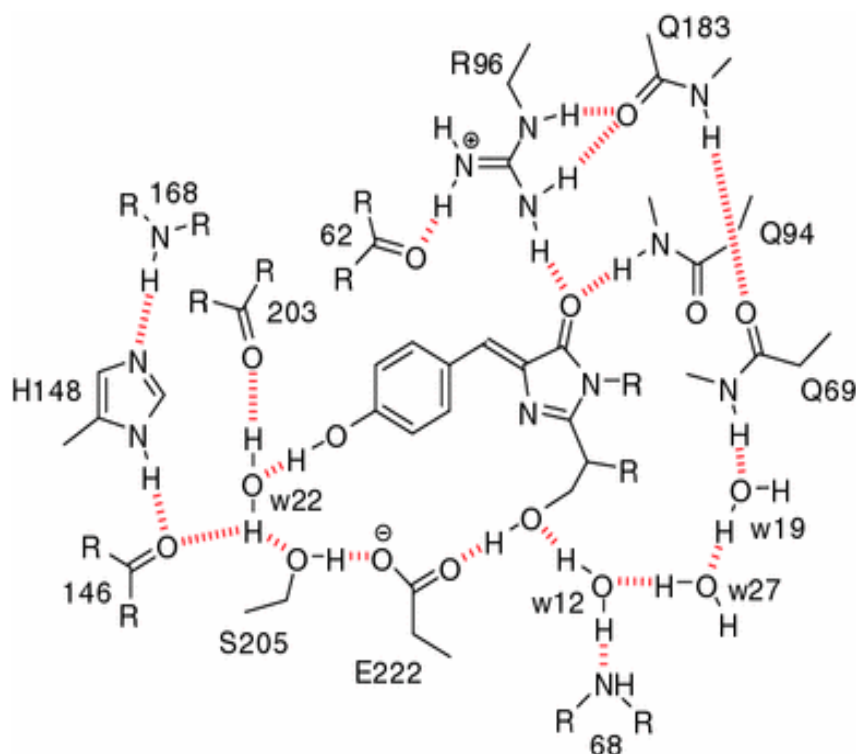


Рисунок 14. Схема взаимодействий хромофора флуоресцентного белка GFP с окружающими аминокислотными остатками и молекулами воды [108].

Для того чтобы изучить вклад отдельных контактов, образуемых хромофором, в изменение спектров белка, был синтезирован ряд аналогов хромофора GFP, моделирующих разные природные взаимодействия (рисунок 15). Спектры поглощения этих модельных соединений были затем исследованы методом газовой спектроскопии [109].

Несмотря на внушительные успехи и открытия, сделанные с использованием модельных соединений, их применение, как и любое другое моделирование, должно проводиться с полным пониманием существующих ограничений, основным из которых является то, что они упрощают задачу, которая должна быть изучена, но используемое упрощение не всегда корректно для исследуемой системы.

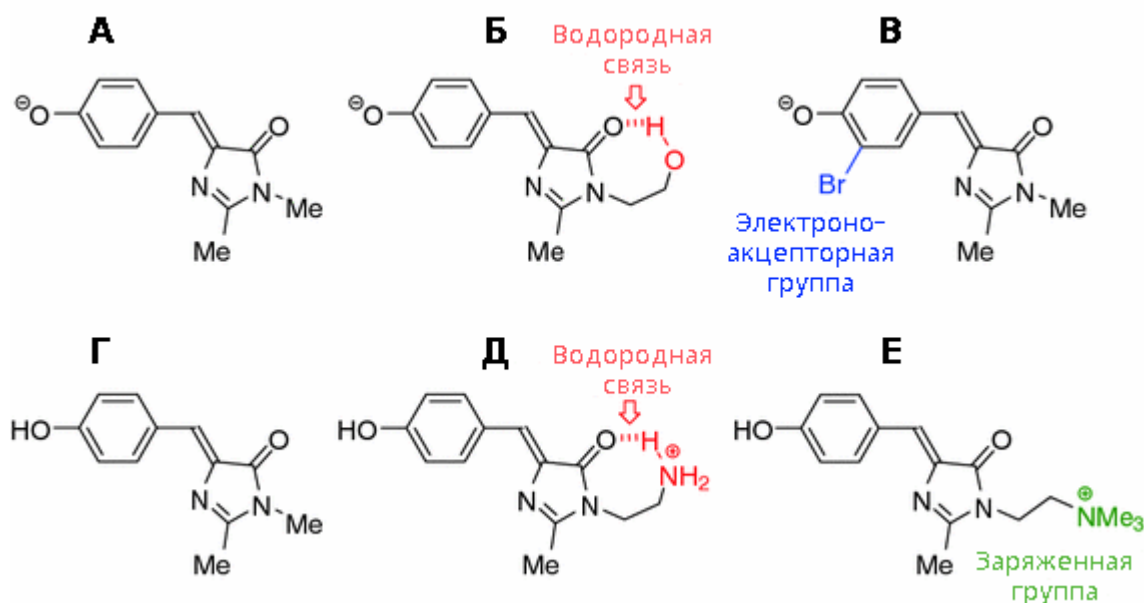


Рисунок 15. Структуры исследованных аналогов хромофора GFP (в анионной (А-В) и нейтральной (Г-Е) формах), моделирующие различные природные взаимодействия хромофора с окружающими аминокислотными остатками [108]

2.3.2. Компьютерное моделирование

Так как количество доступных программ и приложений крайне велико и с каждым днём увеличивается, данный раздел призван лишь обозначить основные задачи, возникающие при создании белков с новыми свойствами, которые можно сегодня решить с помощью компьютерных методов, и привести отдельные примеры их реализации.

Компьютерные методы способны выполнять задачи различной степени сложности, справляться с ними с разной эффективностью, требуют разного уровня подготовки пользователей, а также существенно отличаются по время- и трудозатратности. Ряд подходов помогает получить некоторую полезную информацию: определить позиции в белке, которые в первую очередь стоит подвергать случайному мутагенезу, или выбрать наиболее перспективные замены для постановки в данное положение. Часто они представлены в виде онлайн приложений и просты в использовании, что, соответственно, делает их доступными для широкого круга исследователей, в том числе для не имеющих значительного опыта в области структурной биологии и биоинформатики. Другие методы способны гораздо точнее предсказать необходимые мутации или даже смоделировать белок с заданной функцией с нуля. Однако для получения хороших результатов такие программы требуют тонкой настройки параметров для каждого конкретного случая, длительных расчётов и проведения сложного анализа результатов, что значительно повышает требования к навыкам пользователя и необходимым вычислительным мощностям.

2.3.2.1. Определение позиций для мутагенеза

Самые частые задачи при создании белков с новыми свойствами — изменение активности, субстратной специфичности фермента или типа катализируемой им реакции. Для этого в первую очередь анализируют и заменяют аминокислотные остатки в лиганд-связывающем кармане и/или каталитическом центре.

Проще всего определить аминокислотные остатки, участвующие в координации лиганда, если уже известна кристаллическая структура комплекса. В таком случае можно использовать PyMol — самую распространенную в настоящее время программу для 3D визуализации молекул (например, можно посмотреть на образуемые водородные связи, используя встроенные функции “Action → find → polar contacts → to other atom in object” или “Action → present → ligands sites → cartoon”, или выделить все атомы на заданном расстоянии от лиганда с помощью функции командной строки “around”). Несмотря на бóльшую информативность 3D моделей (рисунок 16), они требуют установки специальных программ для просмотра и навыков работы с ними. Для упрощения процедуры анализа (а также для подготовки иллюстраций для печатных изданий) были разработаны алгоритмы, представляющие рассчитанные водородные связи и другие лиганд-белковые взаимодействия в виде двухмерных диаграмм. Например, широкое распространение получило приложение LigPlot+ (<http://www.ebi.ac.uk/thornton-srv/software/LigPlus/>) [110], являющееся новой версией ранее опубликованной программы LIGPLOT [111]. В дополнение к прежним функциям, LigPlot+ способно анализировать несколько комплексов одновременно, предоставляя возможность увидеть сходства и различия в связывании одного лиганда несколькими белками или разных лигандов одним белком. Другое приложение — PLIP (projects.biotec.tu-dresden.de/plip-web) [112] — позволяет провести полное исследование взаимодействия белка с лигандом онлайн, не устанавливая никакого программного обеспечения на компьютер и даже не требуя предварительной регистрации.

Если кристаллическая структура комплекса отсутствует, анализ во всех вышеупомянутых программах можно проводить на основе моделей, полученных в результате докинга — метода молекулярного моделирования, позволяющего предсказать предпочтительную ориентацию двух (или более) взаимодействующих молекул. Процесс докинга низкомолекулярного вещества состоит из двух стадий: получения набора потенциальных положений лиганда в лиганд-связывающем кармане (систематическим или стохастическим поиском) и последующей оценки энергии взаимодействия предсказанных комплексов. Так как докинг часто является одной из стадий процесса создания новых лекарственных средств, многие зарекомендовавшие себя программы (например, FlexX (<https://www.biosolveit.de/FlexX/>) [113], Glide (<https://www.schrodinger.com/glide>) [114,115]) и т.д.) доступны лишь на платной основе. Наиболее

часто используемой программой на протяжении многих лет [116] остается свободное программное обеспечение с открытым исходным кодом AutoDock [117] (в настоящее время доступны и поддерживаются две версии: AutoDock 4 (<http://autodock.scripps.edu/>) [118] и AutoDock Vina (<http://vina.scripps.edu/>) [119]). Также существует возможность воспользоваться не требующим установки приложением, например, онлайн-сервером программы Rosetta ROSIE (http://rosie.rosettacommons.org/ligand_docking) [120–122].

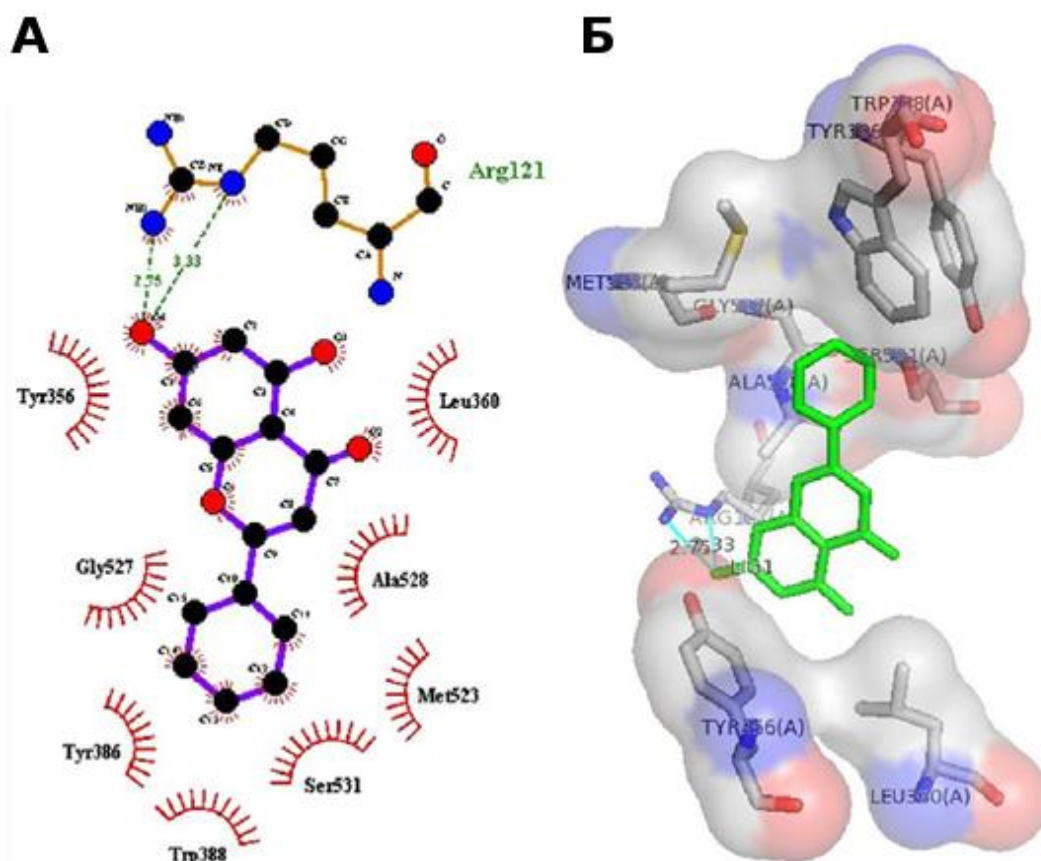


Рисунок 16. Двумерная диаграмма (А) и 3D модель (Б) взаимодействия белка с лигандом [123].

Кроме того, так как активные центры ферментов и аминокислотные остатки, участвующие в связывании лигандов, чаще всего находятся в углублениях белковой молекулы, потенциальные мишени для мутагенеза можно довольно точно предсказать, идентифицировав лишь положение лиганд-связывающего (каталитического) кармана. Поиск таких карманов можно осуществлять, используя чисто геометрические подходы (например, при помощи онлайн сервера CASTp (<http://cast.engr.uic.edu>) [124,125]), проводя сравнение третичной структуры и/или аминокислотной последовательности с белками с известными функциями (по такому принципу работает FINDSITE (<http://cssb.biology.gatech.edu/findsite>) [126,127] или вычисляя энергию взаимодействия между белком и набором химических зондов, так как сайт связывания часто характеризуется энергетическими свойствами, которые отличаются от остальной части поверхности белка и могут быть достоверно идентифицированы (одной из реализаций данного

метода является приложение SiteHound (<http://scbx.mssm.edu/sitehound/sitehound-web/Input.html>) [128]). Каждый из перечисленных методов имеет свои сильные и слабые стороны. Для увеличения надежности предсказаний рекомендуется проводить анализ каждого нового случая в нескольких программах, использующих разные алгоритмы поиска. Благодаря наличию онлайн приложений и высокой скорости подобных расчетов, такая проверка не занимает много времени.

Если структура исследуемого белка неизвестна, её можно попытаться предсказать при помощи гомологичного (сравнительного) моделирования или даже *de novo* фолдинга. Процесс гомологичного моделирования состоит из нескольких стадий: выбор матриц(ы), множественное выравнивание аминокислотных последовательностей, построение 3D модели исследуемого белка и последующая её оптимизация. Его эффективность зависит от степени схожести аминокислотной последовательности моделируемого белка с последовательностями доступных кристаллических структур и от количества таких структур. Хорошие результаты удается получить, если гомология составляет более 30-40%. Каждую из описанных стадий можно проводить независимо, используя разные программы и приложения. Кроме того, существует множество программ, позволяющих провести весь процесс от начала до конца. Например, широкое применение нашла программа MODELLER (<https://salilab.org/modeller/>) [129] и онлайн сервер Swiss-Model (<https://swissmodel.expasy.org/>) [130]. При низкой степени гомологии с известными белками рекомендуется использовать программы, разработанные для *de novo* фолдинга (актуальный рейтинг программ можно узнать на сайте <http://www.predictioncenter.org/>, содержащем результаты CASP (The Critical Assessment of protein Structure Prediction — критическая оценка предсказания белковых структур) — проходящих раз в два года “соревнований” программ, предсказывающих третичную структуру белков по их аминокислотной последовательности) или по возможности вообще отказаться от данного белка в качестве матрицы для мутагенеза.

Не менее частой потребностью биотехнологии является задача повышения термостабильности существующего фермента. Один из способов её решения — внесение замен в положения с большой подвижностью. Такие позиции можно предсказать, анализируя тепловые факторы (b-factors) аминокислотных остатков в кристаллических структурах белков. Например, для этой цели можно использовать программу B-FITTER (доступна для скачивания по адресу <http://www.kofo.mpg.de/en/research/biocatalysis>) [131], анализирующую все атомы за исключением водорода. Аминокислоты с самыми высокими значениями теплового фактора являются первыми кандидатами для мутагенеза. Если белок не был кристаллизован, выбор может быть сделан на основе предсказания влияния замен на стабильность белка. Для этого каждую аминокислоту в белке последовательно заменяют на каждую из 20 возможных, после чего

рассчитывают изменение энергии структуры ($\Delta\Delta G$) относительно исходной. Учитывая рекомендации авторов программы (энергия в большинстве подобных программ приводится в условных единицах, не имеющих физического аналога; границы значимости изменений энергии устанавливаются при тестировании программы на известном наборе данных, для которых $\Delta\Delta G$ измерены экспериментально, и обычно указываются в описании), все мутации делят на стабилизирующие, нейтральные и дестабилизирующие. Позиции, замены в которых в основном классифицируются как стабилизирующие, являются перспективными кандидатами для проведения сайт-специфического мутагенеза. Одной из хорошо зарекомендовавших себя для подобных расчётов программ является FoldX (<http://foldxsuite.crg.eu/>) [132].

2.3.2.2. Выбор перспективных замен

Большинство вносимых при сайт-специфическом случайном мутагенезе мутаций либо нейтральны, либо вредны. Поэтому для сокращения размера анализируемых библиотек и увеличения эффективности такого мутагенеза (а значит, и уменьшения стоимости) был разработан ряд подходов, помогающих оценить рациональность проведения мутагенеза по выбранной позиции и отобрать потенциально полезные замены.

Один из простых и в то же время эффективных подходов — сравнение с гомологичными белками, обогащенными в ходе длительного процесса естественной эволюции функциональными в данном контексте заменами. Аминокислотные остатки, которые необходимы для поддержания структурных или функциональных свойств белка, как правило, высококонсервативны [133]. Рациональность изменения консервативных остатков зависит от цели работы. Если задачей является изменение функции белка, консервативные остатки стоит подвергать мутагенезу. Если же требуется лишь модифицировать какое-либо физико-химическое свойство (термостабильность, pH-стабильность, чувствительность к химическим агентам среды, склонность к димеризации и т.д.), менять такие позиции не рекомендуется, во всяком случае, не в первую очередь и с большой осторожностью. Оценить степень консервативности аминокислотных остатков в белке можно, например, с помощью сервера ConSurf (<http://consurf.tau.ac.il/>) [134]. В отличие от многих других подобных программ, ConSurf может как самостоятельно осуществлять поиск гомологов и множественное выравнивание на основе предоставленной структуры или аминокислотной последовательности белка, так и использовать вариант, предложенный пользователем. Кроме информации о степени консервативности положения, выходные данные содержат список всех встреченных в этом месте аминокислотных остатков. Эти данные могут быть использованы для ограничения разнообразия вносимых мутаций [135], а, например, приложения ANT [136] и TopLib (<http://stat.haifa.ac.il/~yuval/toplib/>) [137] — для подбора подходящих вырожденных кодонов для мутагенеза и оценки требуемого размера библиотеки, соответственно.

Другим способом сокращения количества анализируемых вариантов является выявление и внесение консенсусных замен. Такой подход особенно зарекомендовал себя в создании термостабильных белков [138,139]. Схожим методом, известным под названием REAP (reconstructed evolutionary adaptive paths — реконструированный путь адаптивной эволюции), является выбор замен для изменения существующих белков на основе моделирования процесса их эволюции и выявления аминокислотных остатков, ответственных за изменение функций [140].

Любым способом полученный сокращенный список замен может быть дополнительно проверен и, при необходимости, скорректирован при помощи уже упомянутых программ, рассчитывающих изменение энергии ($\Delta\Delta G$) структуры белка в ответ на внесенную мутацию.

2.3.2.3. *In vitro* рекомбинация

Для более успешного осуществления *in vitro* рекомбинации был разработан алгоритм SCHEMA, позволяющий разделять белки на фрагменты, которые с большой вероятностью будут нормально сворачиваться и функционировать в составе гибридных белков. Алгоритм учитывает взаимодействия между аминокислотными остатками и рассчитывает, сколько из них будет нарушено при исключении данного фрагмента. Позиции, разрыв полипептидной цепи по которым приводит к минимальному количеству разрушенных связей, являются благоприятными для рекомбинации [141].

2.3.2.4. Дизайн белков

Под дизайном подразумевается предсказание либо первичной последовательности (*de novo* дизайн), либо полного набора аминокислотных замен в модифицируемом белке (редизайн), которое приведет к получению белка с заданными свойствами.

Редизайн

Многочисленные методы редизайна используют биоинформатический анализ известных белков, молекулярную динамику, *in silico* мутагенез, молекулярный докинг, квантово-механические расчёты, а также различные их сочетания.

Чисто биоинформатические методы в основном служат не столько цели создания новых или улучшения существующих белков, сколько лучшему пониманию их устройства. Они обычно базируются на информации о гомологичных белках. Правда, в отличие от описанных выше методов, использующих эти данные лишь для предсказания потенциально полезных мутаций, биоинформатические подходы к редизайну учитывают дополнительные факторы. Так, метод Janus, предсказывающий мутации, необходимые для взаимопревращения структурно схожих, но функционально различных белков, анализирует консервативность и ковариантность физико-химических свойств (а не просто аминокислот) последовательностей, а также учитывает их удалённость от активного центра фермента [142].

Комбинированные же подходы могут быть применены и для получения белков с новыми свойствами. Например, объединение биоинформатического метода SABER (Selection of Active/Binding sites for Enzyme Redesign — выбор каталитических сайтов/сайтов связывания для редизайна ферментов), анализирующего данные из базы PDB (<http://www.rcsb.org/>) и находящего структуры со специфическим пространственным расположением каталитических групп, с *in silico* мутагенезом при помощи программы Rosetta позволяет как изменить каталитическую активность белка на присущую другому, так и успешно расположить теоретически рассчитанные аминокислотные остатки для получения новой ферментативной активности [143]. Редизайн с использованием молекулярного докинга, молекулярной динамики и квантово-механических расчётов привёл к получению фермента с инвертированной энантиоселективностью [144]. Также с задачей редизайна успешно справляются программы, созданные для *de novo* дизайна белков.

***De novo* дизайн**

Современные подходы *de novo* дизайна белков основываются на гипотезе, высказанной Нобелевским лауреатом Кристианом Анфинсеном [145], что белок сворачивается в конформацию с наименьшей доступной для данной аминокислотной последовательности энергией. Тогда, при наличии аккуратного метода для расчёта энергии белковой цепи и эффективного алгоритма перебора вариантов из пространства возможных аминокислотных последовательностей, *de novo* дизайн белков должен быть осуществимой задачей. Проблема состоит в том, что энергию такой огромной системы как белок крайне трудно рассчитать с достаточной точностью, а количество возможных комбинаций аминокислот слишком велико. Тем не менее эта задача уже не кажется абсолютно неосуществимой.

De novo дизайн можно разделить на два направления: теоретический (“истинный”) и прикладной. Целью “истинного” *de novo* дизайна является изучение белковых структур, установление принципов их строения с последующей проверкой их правильности путём создания не встречающихся в природе молекул [146]. В случае прикладного *de novo* дизайна (моделирование новых ферментов или лиганд-связывающих белков) в качестве основы обычно используют известные топологии. При этом весь процесс состоит из четырёх стадий (рисунок 17). Сначала определяют, каким образом (с помощью каких функциональных групп) лиганд или переходное состояние может быть стабилизировано. Затем создают библиотеку, содержащую все возможные сочетания аминокислотных остатков, имеющих необходимые функциональные группы, во всех возможных взаимных ориентациях. На третьей стадии элементы этой библиотеки пытаются поместить в каждую из выбранных белковых топологий. Успешные варианты дополнительно оптимизируют, и лучшие модели проверяют экспериментально [147].

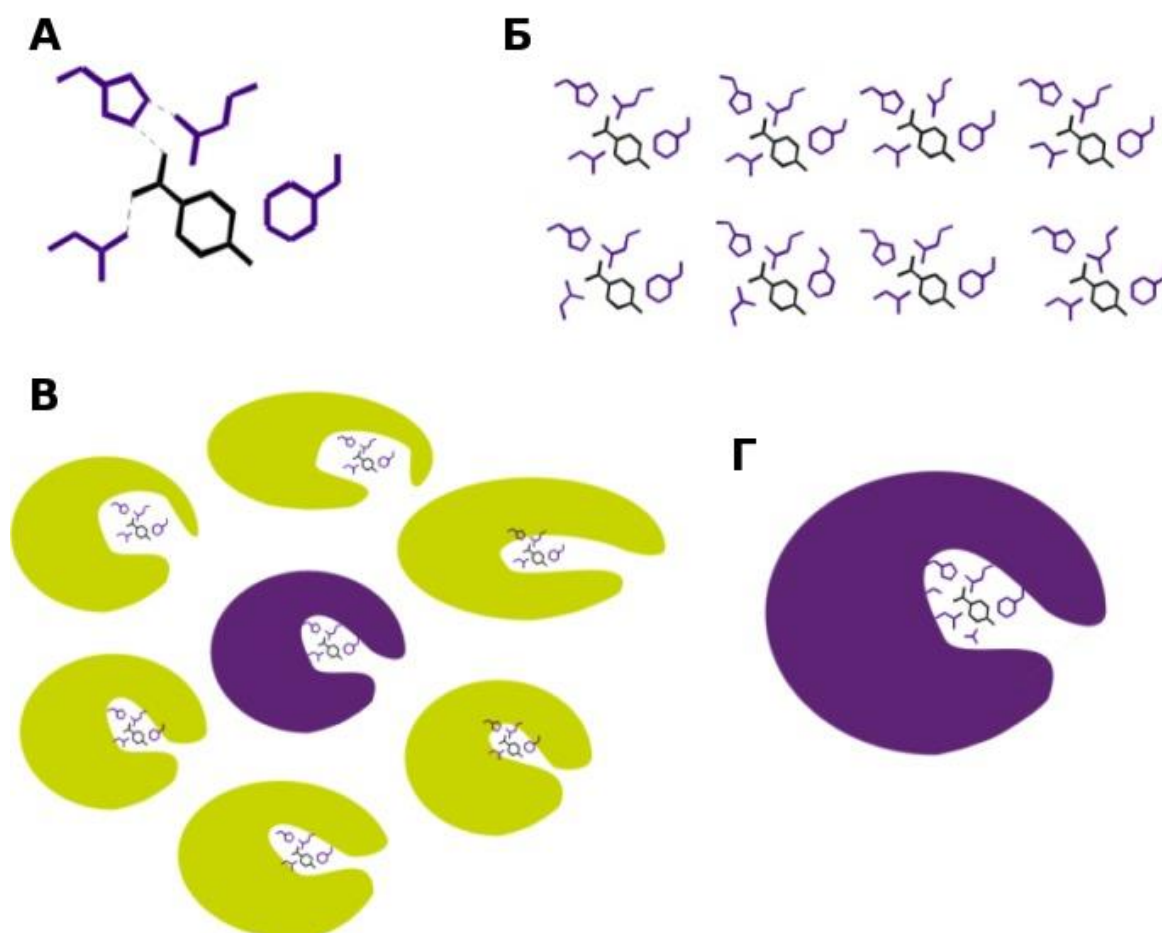


Рисунок 17. Стадии *de novo* дизайна белков на основе известных топологий. **А.** Моделирование способов стабилизации переходного состояния; **Б.** Создание библиотеки ротамеров, способных стабилизировать смоделированное переходное состояние; **В.** Помещение элементов библиотеки в выбранные белковые топологии, оценка полученных структур, выбор лучшей; **Г.** Оптимизация лучшей структуры [148].

Пожалуй, самой успешной на сегодняшний день программой, справляющейся с *de novo* дизайном белковых молекул, является программа Rosetta (<https://www.rosettacommons.org/>). С её помощью была показана возможность получения белков с топологией, не встречающейся среди изученных природных белков [149], связывающих низкомолекулярные лиганды (например, дигоксигенин — агликон сердечного гликозида дигоксина — и прогестерон [150]), а также был создан ряд ферментов, катализирующих реакции, не найденные до сих пор в живых организмах (например, реакцию Дильса-Альдера (рисунок 18А) (реакция [4+2]-циклоприсоединения диенофилов и сопряжённых диенов с образованием шестичленного цикла) [151] и элиминацию Кемпа (рисунок 18Б) (отщепление атома водорода от неионизируемого углеродного атома) [152]). Также широкое применение находит программа ORBIT (<http://www.cs.duke.edu/donaldlab/osprey.php>) [153,154].

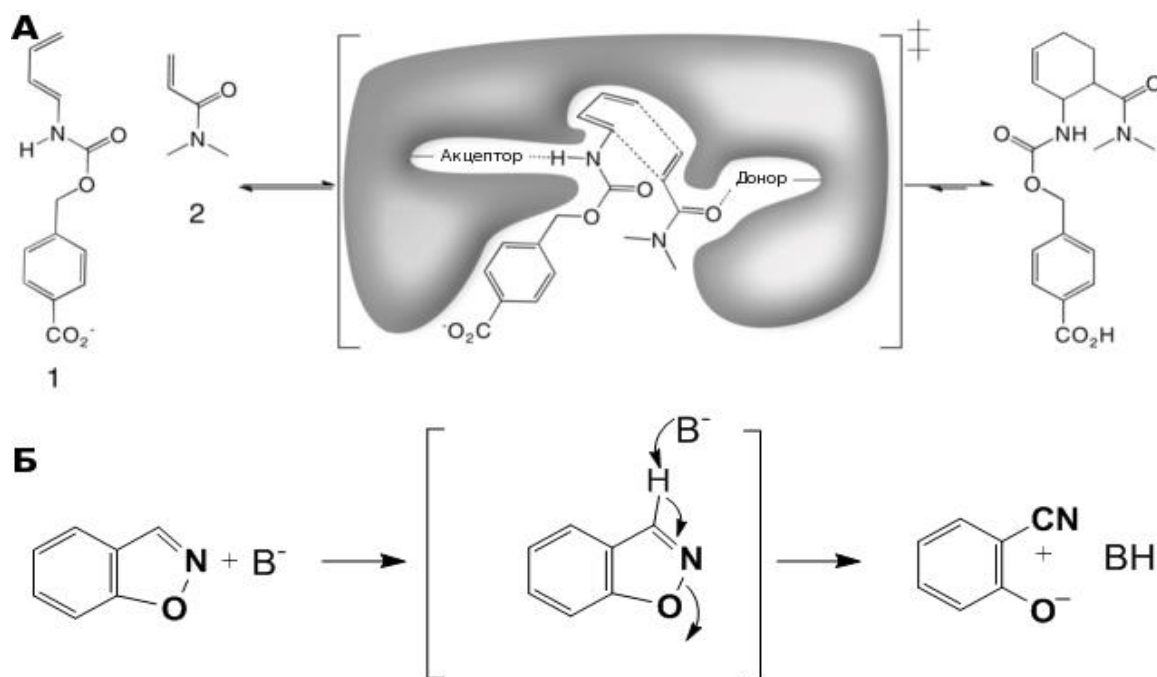


Рисунок 18. Схемы химических реакций, катализируемых ферментами, созданными при помощи программы Rosetta. **А.** Реакция Дильса-Альдера — реакция [4+2]-циклоприсоединения диенофилов (2) и сопряжённых диенов (1) с образованием шестичленного цикла (3) [151]; **Б.** Элиминация Кемпа — отщепление атома водорода от неионизируемого углеродного атома.

Несмотря на существующие успехи компьютерного моделирования, созданные с их помощью белки пока что, как правило, уступают природным (по эффективности катализа, константе связывания с лигандом и т.д.), поэтому для получения эффективного биотехнологического инструмента их дорабатывают при помощи тех или иных методов направленной эволюции, чаще всего, используя описанные выше методы случайного мутагенеза. Однако это не умаляет возможностей и пользы расчётных методов и лишь стимулирует их дальнейшее развитие и совершенствование.

3. Материалы и методы

3.1. Амплификация фрагментов ДНК

Для амплификации фрагментов ДНК применялась полимеразная цепная реакция (ПЦР). Амплификация проводилась в приборах Thermal Cycler 2720 (Applied Biosystems) или PTC-200 Peltier Thermal Cycler (MJ Research) с использованием наборов реактивов и буферов для ПЦР компании Евроген.

Состав реакционной смеси (50 мкл):

- 5 мкл 10х буфера;
- 1 мкл 50х смеси полимераз (Encyclo или Tersus);
- 1 мкл 5'-концевого праймера (конечная концентрация 0.2 мкМ);
- 1 мкл 3'-концевого праймера (конечная концентрация 0.2 мкМ);
- 0.5 мкл 50х смеси дезоксирибонуклеозидтрифосфатов (конечная концентрация каждого 0.1 мМ);
- 1 мкл матричной ДНК (0.1-1 нг);
- 40.5 мкл mQ H₂O.

Программа полимеразной цепной реакции:

- Денатурация — 95°C, 1 минута.
- Денатурация — 94°C, 10-20 секунд.
- Отжиг праймеров — 60-70°C (расчет температуры отжига проводился с использованием бесплатного программного обеспечения NetPrimer (Premier Biosoft)), 10-20 секунд.
- Элонгация — 72°C (время элонгации устанавливалось из расчета скорости работы полимеразы 1000 пар оснований/мин).
- Стадии 2-4 повторяли 12-25 раз.

3.2. Электрофорез в агарозном геле

Электрофорез в агарозном геле проводили для визуализации результатов ПЦР, а также для разделения и последующей очистки фрагментов ДНК (после амплификации или рестрикции). 1-2% агарозный гель готовили на трис-ацетатном буфере (ТАЕ) с добавлением бромистого этидия. Электрофорез проводили в камере для горизонтального электрофореза (BioRad) с использованием источника постоянного тока EPS 250 (C.B.S. Scientific Company) при напряжении 140 — 180 В (в зависимости от длины анализируемых или разделяемых фрагментов). В качестве маркера наносили 1 Kb DNA Ladder (СибЭнзим). Для визуализации результатов электрофореза использовали систему гель-документации G:BOX Chemi (Syngene); в случае последующего выделения ДНК из геля использовали УФ-трансиллюминатор, снабженный УФ-лампой ($\lambda=365$ нм).

3.3. Очистка ДНК

Для выделения ДНК из реакционных смесей (после ПЦР, рестрикции или лигирования) или агарозного геля (после разделения фрагментов) использовали наборы для очистки ДНК Cleanup Standard (Евроген) и Cleanup Mini (Евроген) согласно прилагаемой к наборам инструкции.

Для выделения плазмидной ДНК использовали набор для выделения плазмидной ДНК Plasmid Midi Kit (Qiagen) (плазмиды, предназначенные для временной трансфекции клеток млекопитающих) или Plasmid Miniprep (Евроген) согласно прилагаемой к наборам инструкции.

3.4. Сайт-специфический мутагенез

В зависимости от конкретной ситуации (расстояния между уникальными сайтами рестрикции, количества замен, которые необходимо единовременно ввести в ген, нуклеотидного окружения места внесения замены, настроения автора и т.д.) использовался один из следующих трех методов внесения сайт-специфических мутаций:

- Гибридизация N- и C- концевых фрагментов (рисунок 19А) [155].
- Самособирающееся клонирование (рисунок 19Б) [156].
- “Аква” клонирование (рисунок 19В) (от англ. AQUA — advanced quick assembly — улучшенная быстрая сборка) [157].

3.4.1. Гибридизация N- и C- концевых фрагментов

Для данного метода использовали пару комплементарных праймеров, несущих желаемую замену, и пару концевых праймеров. На первой стадии проводили независимую амплификацию двух фрагментов гена (N-концевого и C-концевого) по стандартной методике, используя для каждой реакции один концевой праймер и один праймер, несущий замену. Полученные ПЦР-продукты чистили через гель.

На второй стадии продукты, полученные на первой стадии, объединяли и проводили амплификацию ДНК, при которой смешанные фрагменты гена одновременно служили и матрицей, и праймерами. После 5 циклов в реакционную смесь добавляли концевые праймеры и проводили амплификацию полноразмерного гена, содержащего желаемую замену.

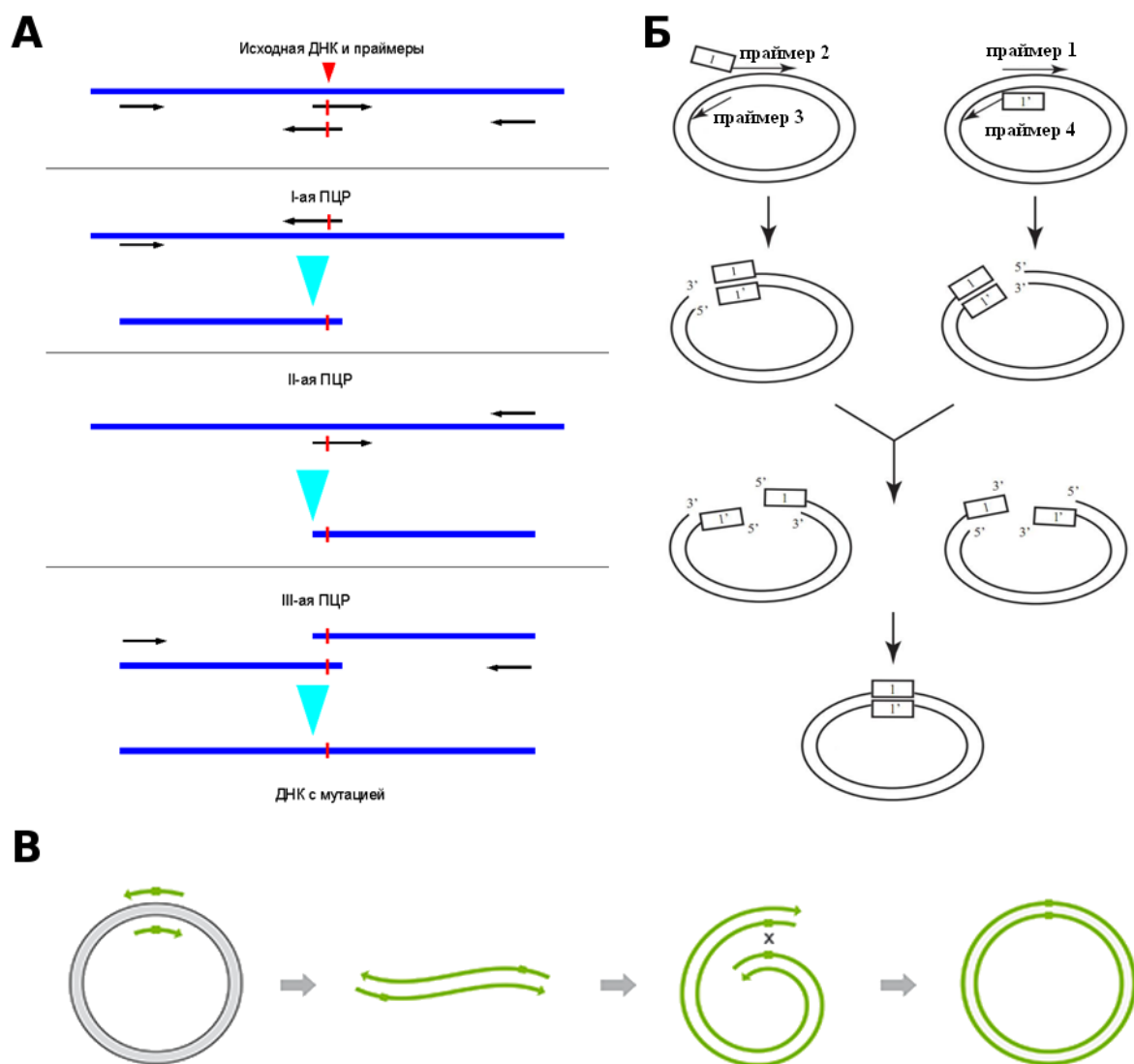


Рисунок 19. Принципы внесения сайт-специфических замен. **А.** Гибридизация N- и C- концевых фрагментов; **Б.** Самособирающееся клонирование; **В.** “Аква” клонирование

Продукт амплификации чистили, обрабатывали рестриктазами и лигировали в линейризованный по тем же сайтам вектор.

3.4.2. Самособирающееся клонирование

Для внесения точечных мутаций методом самособирающегося клонирования также необходимо использовать четыре праймера: два “длинных” (2 и 4 на рисунке 19Б) и два “коротких” (1 и 3 на рисунке 19Б). Прямой и обратный “длинных” праймеры должны иметь комплементарную часть длиной 11-13 нуклеотидов с соотношением АТ/ГС близким к 50%. Оба “коротких” праймера, наоборот, не должны содержать эту последовательность. Замена может находиться как в комплементарной части “длинных”, так и в общей части прямых или обратных праймеров.

Используя “длинный” прямой и “короткий” обратный праймеры (или наоборот), проводили две независимые амплификации всего вектора. Для уничтожения исходного вектора

ПЦР-продукт обрабатывали рестриктазой DpnI (Fermentas) или MspI (СибЭнзим), специфичной к метилированной ДНК. Очищенные ПЦР-продукты смешивали друг с другом, добавляли буфер (состав 4х буфера: 100 мМ Tris-HCl, 500 мМ KCl, 15 мМ MgCl₂, pH 8.3), после чего денатурировали (95°C — 3 мин) и остужали до 25°C. В результате, в реакционной смеси образовывалось некоторое количество гибридных молекул ДНК, имеющих выступающие комплементарные концы.

Температура отжига последовательности длиной 11-13 нуклеотидов с соотношением АТ/ГС близким к 50% составляет примерно 25°C, поэтому полученные гибридные молекулы ДНК замыкаются в кольцо при комнатной температуре и могут быть непосредственно использованы для химической трансформации бактерий (возможность использования электрической трансформации сильно ограничена из-за высокого содержания солей в гибридизационном буфере).

3.4.3. “Аква” клонирование

Для внесения точечных мутаций методом аква клонирования требуются два праймера (прямой и обратный), содержащие требуемую замену и имеющие комплементарный участок длиной 16-32 нуклеотидов. Используя эти праймеры проводили амплификацию всего вектора. Для уничтожения исходного вектора ПЦР-продукт обрабатывали рестриктазой DpnI (Fermentas) или MspI (СибЭнзим), специфичной к метилированной ДНК. После очистки ПЦР-продукт растворяли в воде и использовали для химической или электрической трансформации бактерий, полагаясь на *in vivo* рекомбинацию коротких гомологичных участков.

3.5. Создание библиотек случайных мутантов

Случайный мутагенез кодирующей последовательности генов проводили с помощью ПЦР с прогнозируемой повышенной частотой ошибок, которой добивались изменением соотношения и увеличением концентрации дезоксирибонуклеозидтрифосфатов в реакционной смеси, добавлением марганца и использованием и так не очень точной Taq ДНК-полимеразы.

Состав реакционной смеси (50 мкл):

- 5 мкл 10х буфера для Taq ДНК-полимеразы;
- 0.5 мкл 100х Taq ДНК-полимеразы;
- 4 мкл MnSO₄ (8 мМ);
- 3 мкл dGTP (2 мМ);
- 2 мкл 25х смеси дезоксирибонуклеозидтрифосфатов (концентрация каждого нуклеотида в конечной смеси: 0.2 мМ dGTP, 0.2 мМ dATP, 1 мМ dCTP, 1 мМ dTTP);

- 1 мкл 5'-праймера (конечная концентрация 0.2 мкМ);
- 1 мкл 3'-праймера (конечная концентрация 0.2 мкМ);
- 1 мкл матричной ДНК с концентрацией 1-100 пг;
- 32.5 мкл H₂O mQ.

Стандартную программу ПЦР повторяли 25-35 раз (в зависимости от концентрации матричной ДНК и желаемого среднего количества замен в гене). Продукт амплификации чистили, обрабатывали рестриктазами и лигировали в линейаризованный по тем же сайтам вектор.

3.6. Трансформация бактерий

В зависимости от цели фасовку компетентных клеток *E.coli* штамма XL1-Blue (Evrogen) или XJb(DE3) Autolysis (Zymo Research) размораживали на льду. После этого к суспензии бактериальных клеток добавляли лигат (пересажженный или очищенный на колонке в случае электропорации), плазмидную ДНК или ПЦР-продукт(ы) (в случае самособирающегося или “аква” клонирования).

3.6.1. Химическая трансформация

Смесь клеток с ДНК инкубировали на льду в течение 20 мин, а затем 60-90 секунд в водяной бане или термостате при 42 °C (тепловой шок). После теплового шока клетки снова помещали в лед на 1-3 минуты, после чего к ним добавляли 200 мкл среды SOB.

3.6.2. Электрическая трансформация

Смесь клеток с ДНК переносили в предварительно охлажденную кювету для электропорации (Bio-Rad). Электропорацию проводили при помощи прибора MicroPulser (Bio-Rad). Сразу после электропорации в кювету добавляли 1 мл среды SOB.

Независимо от метода трансформации суспензию клеток после добавления SOB переносили в пробирку, инкубировали при 37°C в течение 40 минут при постоянном перемешивании, а затем использовали для посева на чашки Петри с агаризованной средой LB с добавлением селективного антибиотика: ампициллина (100 мкг/мл) или канамицина (20 мкг/мл). Чашки инкубировали при 37°C в течение ночи (примерно 16 часов).

3.7. Экспрессия и очистка рекомбинантных белков

Для экспрессии и выделения белков использовали вектор pQE30 (индукция при помощи IPTG) или pBAD (активируемым арабинозой промотор) и *E.coli* штаммов XLBlue, XJb(DE3) Autolysis (под арабинозным промотором у данного штамма находится ген лизоцима) и BW25113 (DaraBAD штамм — не способен потреблять арабинозу).

После трансформации бактерии высевали на чашки Петри с твердой средой, содержащей

селективный антибиотик, и растили в течение 16 часов при 37°C. Выросшие колонии использовали для инокуляции жидкой среды (необходимый объем определялся из расчета 1 мл на 200 мл конечной бактериальной культуры), содержащей селективный антибиотик, которую выращивали при 37°C и перемешивании в течении 16 часов. Суспензию использовали для инокуляции свежей среды (200 мл, 500 мл или 1 л).

Бактериальную культуру растили при 37°C и перемешивании до оптической плотности 0.6, после чего индуцировали экспрессию гена интереса и инкубировали культуру при 37°C и перемешивании до оптической плотности 3 (от 4 до 16 часов в зависимости от объема среды и скорости перемешивания).

Клетки осаждали центрифугированием, супернатант сливали, а осадок ресуспендировали в буферном растворе (PBS, pH 7.4). Бактериальные клетки разрушали трехкратным замораживанием и размораживанием суспензии, после чего с помощью ДНКазы или ультразвука разрушали ДНК (при использовании штамма XJb(DE3) Autolysis), либо разрушали при помощи ультразвука (при использовании штамма BW25113). Полученный лизат центрифугировали (12 000 g, 25 мин, 4°C), супернатант отбирали и наносили на колонку, содержащую металло-аффинную смолу (TALON, Clontech), предварительно уравновешенную буферным раствором. Смолу последовательно промывали 10 объемами следующих четырех растворов: PBS (pH 7.4); PBS (pH 7.4), 0.25M NaCl; PBS (pH 7.4), 0.5M NaCl; PBS (pH 7.4), 10мМ имидазола. Белок, иммобилизованный на смоле, использовали для дальнейших экспериментов. Для получения раствора белок элюировали буфером (PBS, pH 7.4), содержащим 200мМ имидазола, после чего удаляли имидазол диализом, гель-фильтрацией (PD-10 Desalting Column, GE Healthcare) или при помощи концентрационных колонок (Merck Millipore).

3.8. Денатурирующий электрофорез в полиакриламидном геле

Экспрессию а также чистоту выделенных рекомбинантных белков оценивали при помощи вертикального денатурирующего электрофореза в полиакриламидном геле по Лэммли [158].

Концентрирующий гель (4%) готовили на 0.5M Трис-HCl буфере (pH 6.8) с добавлением 0.1% SDS, разделяющий (12%) — на 1.5M Трис-HCl (pH 8.8). Буфер для нанесения содержал 1% SDS, 10% глицерина и 2.5% β-меркаптоэтанола. Электрофорез проводили в вертикальной камере Mini-Protean Tetra (Bio-Rad) с использованием источника постоянного тока EPS 250 (C.B.S. Scientific Company) при напряжении 100 В (при прохождении проб концентрирующего геля) и 200 В (при прохождении проб разделяющего геля). В качестве маркера наносили Protein Ladder 10-250 кДа (СибЭнзим). Для избежания размывания полос белки в геле фиксировали 3.5% раствором хлорной кислоты в 50% растворе этанола, после чего окрашивали, используя

Coomassie Blue R250.

При анализе экспрессии белков 70 мкл суспензии бактерий ($OD = 3$) центрифугировали и осадок ресуспендировали в 50 мкл 1х буфера для нанесения. При анализе выделенных рекомбинантных белков раствор, содержащий 5-15 мкг белка, смешивали с 12.5 мкл 4х буфера для нанесения и доводили до объема 50 мкл. В зависимости от цели эксперимента, пробы либо прогревали в течении 0.5-10 минут при температуре 95°C и центрифугировали, либо наносили немедленно (7.5-10 мкл на дорожку).

3.9. Определение спектральных свойств

Для снятия спектров поглощения использовали спектрофотометр Cary 100 Bio UV/VIS (Varian), для детекции спектров флуоресценции использовали спектрофлуориметр Cary Eclipse (Agilent Technologies).

3.10. pH-титрование

pK_a хромофоров определяли прямым титрованием водного раствора соединения. Для изучения pH-зависимости спектров флуоресцентных белков использовали набор 50мМ буферов с pH от 2.5 до 11.0 с шагом 0.5 (цитратные буферы для pH 2.5-4.5, фосфатные буферы для pH 5.0-7.5 и боратные буферы для pH 8.0-11.0). Аликвоту выделенного белка растворяли в соответствующем буфере. Спектры детектировали после 15 минут инкубации.

3.11. Определение констант диссоциации

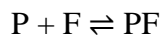
В опытах по титрованию использовали свежевыделенные белки и свежеприготовленные растворы флуорогенов. Концентрацию белков определяли по поглощению на 280 нм и коэффициентам экстинкции, рассчитанным в приложении ProtParam (<http://web.expasy.org/protparam/>).

Для приготовления растворов флуорогенов небольшое количество сухого вещества растворяли в DMSO. Эти растворы (предварительно разбавленные) использовали для определения концентрации по поглощению в максимуме абсорбции и ранее рассчитанным коэффициентам экстинкции, после чего готовили рабочие растворы в 96% этаноле.

Непосредственно перед измерением смешивали необходимое количество растворов белка и флуорогена, добавляли буфер (PBS, pH 7.4) до общего объема 1 мл. Раствор аккуратно перемешивали и немедленно детектировали спектры флуоресценции и возбуждения флуоресценции. Для каждой пары белок-флуороген было измерено не менее 30 точек.

Для оценки констант диссоциации комплексов белок-флуороген было предположено, что белки являются мономерами, обладают единственным сайтом связывания с флуорогеном и

взаимодействуют с ним обратимо. В этом случае реакция взаимодействия белка с флуорогеном может быть выражена уравнением:



где P — свободный белок, F — свободный флуороген, PF — комплекс белок-флуороген. Тогда константа диссоциации (K_d) равна:

$$K_d = \frac{[P]_{eq}[F]_{eq}}{[PF]_{eq}}$$

где $[P]_{eq}$ — концентрация свободного белка в состоянии равновесия, $[F]_{eq}$ — концентрация свободного флуорогена в состоянии равновесия, $[PF]_{eq}$ — концентрация комплекса белок-флуороген в состоянии равновесия.

Так как нам известны лишь начальные концентрации белка и флуорогена, и мы не обладаем каким-либо методом для прямого измерения концентрации свободного белка, свободного флуорогена или комплекса флуороген-белок в состоянии равновесия, то для выражения константы диссоциации можно использовать уравнения массового баланса:

$$[P]_{init} = [P]_{eq} + [PF]_{eq}$$

$$[F]_{init} = [F]_{eq} + [PF]_{eq}$$

где $[P]_{init}$ и $[F]_{init}$ — начальные концентрации белка и флуорогена соответственно. Откуда можно выразить равновесные концентрации белка и флуорогена:

$$[P]_{eq} = [P]_{init} - [PF]_{eq}$$

$$[F]_{eq} = [F]_{init} - [PF]_{eq}$$

Сделав эту замену в исходном выражении для K_d , получим:

$$K_d = \frac{([P]_{init} - [PF]_{eq})([F]_{init} - [PF]_{eq})}{[PF]_{eq}}$$

Откуда:

$$([P]_{init} - [PF]_{eq})([F]_{init} - [PF]_{eq}) = K_d [PF]_{eq}$$

$$[P]_{init}[F]_{init} - [PF]_{eq}[F]_{init} - [P]_{init}[PF]_{eq} + [PF]_{eq}^2 - K_d[PF]_{eq} = 0$$

$$[PF]_{eq}^2 - [PF]_{eq} ([F]_{init} + [P]_{init} + K_d) + [P]_{init}[F]_{init} = 0$$

Решая данное квадратное уравнение, можно выразить концентрацию комплекса белок-флуороген:

$$[PF]_{eq} = \frac{([F]_{init} + [P]_{init} + K_d) - \sqrt{([F]_{init} + [P]_{init} + K_d)^2 - 4[P]_{init}[F]_{init}}}{2}$$

Интенсивность флуоресценции (F) смеси равна:

$$F = [PF]_{eq}\Phi_{[PF]} + ([F]_{init} - [PF]_{eq})\Phi_{[F]}$$

где $\Phi_{[PF]}$ и $\Phi_{[F]}$ — относительные квантовые выходы флуоресценции комплекса белок-флуороген и свободного флуорогена соответственно.

На основе экспериментальных данных по титрованию флуорогена различными концентрациями исследуемого белка с помощью метода наименьших квадратов *leastsq* из пакета Scipy (Python) подбирались K_d , $\Phi_{[F]}$ и $\Phi_{[PF]}$.

3.12. Получение генетических конструкций

3.12.1. Клонирование гена *blc*

Полная копия закристаллизованного фрагмента белка Blc [159], не содержащего сигнальной последовательности, была клонирована в две стадии. Сначала укороченный фрагмент гена *blc* был амплифицирован с использованием *E.coli* штамма K12 Rosetta-gami 2(DE3)pLysS в качестве матрицы и двух пар праймеров (сайты распознавания рестриктаз подчеркнуты):

Blc-F: ATGCCGTGTGGACGGGTATT

Blc-R: CGCCGCCTGGCAGATTCTTA

Blc-Nhe-F: GTACCGCTAGCAGCTCCACGCCGCCGCGTGGCGTGAC

Blc-Hind-R: GTACAAAGCTTCTAACTACCAGGCTGCTGTACCCAAATA

ПЦР-продукт был обработан рестриктазами NheI и HindIII (Fermentas) и лигирован в линеаризованный по тем же сайтам экспрессионный вектор pBAD вместо находившейся в нем кодирующей последовательности белка iRFP. Недостающий линкерный участок между полигистидиновой последовательностью и кодирующей последовательностью белка Blc был добавлен при помощи самособирающегося клонирования:

Blc-link-F: AAGAAATCCAGCTCCACGCCGCCGC

Blc-link-F-long: ACCAGCTTGTACAAGAAATCCAGCTCCACGCCGCCGC

Blc-link-R: TGATTCGAGATGATGATGATGATGATGATGAGAACCC

Blc-link-R-long: GTACAAGCTGGTTGATTCGAGATGATGATGATGATGATGATGAGAACCC

3.12.2. Создание конструкций для временной трансфекции культур эукариотических клеток

Конструкции FR66W-C165A-N, FR66W-Yellow-N и FR66W-Orange-mito были созданы при помощи рестрикции: последовательность, кодирующую соответствующий флуоресцентный белок, получали рестрикцией бактериального вектора pQE-30 по сайтам AgeI/NotI и лигировали в соответствующий вектор (TagRFP-N (Evrogen) или TagRFP-mito (Evrogen)) вместо TagRFP.

Конструкции H2B-TagBFP-Blc, H2B-Blc-TagBFP, H2B-Blc, keratin-TagBFP-Blc и keratin-Blc (где Blc — это белок дикого типа или любой его мутант) были получены при помощи самособирающегося клонирования, используя векторы Blc-pBAD, H2B-TagBFP-Blc, H2B-TagBFP (Evrogen) и keratin-TagRFP (Evrogen) в качестве матрицы.

Конструкции actinin-Blc (где Blc — это любой мутант белка Blc) были получены лигированием амплифицированного и обработанного рестриктазами гена мутанта белка Blc в вектор actinin-TagRFP (Evrogen) вместо флуоресцентного белка.

Конструкции actinin-TagBFP-Blc, vimentin-TagBFP-Blc и Lifeact-TagBFP-Blc (где Blc — это любой мутант белка Blc) были получены с помощью рестрикции: последовательность TagBFP-Blc или Blc, полученную рестрикцией вектора H2B-TagBFP-Blc или H2B-Blc, лигировали в соответствующий вектор (actinin-TagRFP (Evrogen), vimentin-Citrine или Lifeact-TagBFP (Evrogen)) вместо флуоресцентного белка.

Правильность всех конструкций, при создании которых проводилась амплификация ДНК, была подтверждена секвенированием.

3.13. Культивирование и временная трансфекция культур эукариотических клеток

В работе были использованы культуры эукариотических клеток линий HEK293T, NIH-3T3 и HeLa Kyoto. Для культивации клеток использовали среду DMEM (ПанЭко), дополненную 2мМ глутамина (ПанЭко), смесью антибиотиков (пенициллин и стрептомицин, ПанЭко) и содержащую 10% фетальной телячьей сыворотки (HyClone, Thermo Scientific). Культуры растили в инкубаторе с 5% CO₂ при 37°C и пересевали каждые 2 дня.

Перед трансфекцией и последующей микроскопией клетки переносили на 35 мм чашки Петри со стеклянным дном Fluorodish (World Precision Instruments) или в μ -Slide VI^{0.4} (IBIDI) и культивировали до достижения 70% плотности. Для трансфекции использовали реагент FuGene 6 или FuGene HD (Promega) согласно инструкции производителя.

Анализ клеток проводили через 24-48 часов после трансфекции. Перед микроскопией клетки дважды промывали раствором Версена (ПанЭко) и заменяли среду для культивации на раствор Хэнкса (ПанЭко), дополненный 20мМ HEPES (ПанЭко).

3.14. Флуоресцентная микроскопия

Для анализа библиотек случайных мутантов использовали флуоресцентный биноккуляр Olympus SZX12.

Для анализа взаимодействия иммобилизованного на смоле белка с флуорогенами применяли флуоресцентный микроскоп BZ-9000 (Keyence), оснащенный фильтрами SB GFP-BP (OP-79301), SB TexasRed (OP-79302) и SB Cy5 (OP-79305). Съемку проводили, используя объектив CFI Plan Apo 2x/NA0.1 (971826).

Для визуализации трансфицированных клеток использовали:

- Флуоресцентный широкопольный микроскоп BZ-9000 (Keyence, Япония), оснащенный фильтрами SB DAPI-BP (OP-79304), SB GFP-BP (OP-79301), SB TexasRed (OP-79302) и SB Cy5 (OP-79305). Съемку проводили, используя следующие объективы (Nikon, Япония) CFI Plan Apo 20x/NA0.75 (971804), CFI Plan Apo 40x/NA0.95 (971844) и CFI Plan Apo VC60xH/NA1.40 (971805).

- Флуоресцентный широкопольный микроскоп AF6000B (Leica, Германия) оборудованный камерой ZYLA-5.5-CL10 (Andor, Англия), оснащенный иммерсионным объективом HCX PL APO lbd.BL 63x/NA1.4 oil (Leica, Германия) и стандартным набором фильтров: BFP (возбуждение BP390/18, эмиссия BP460/60), GFP (возбуждение BP470/40, эмиссия BP525/50) и TX2 (возбуждение BP560/40, эмиссия BP645/75).

- Конфокальный инвертированный микроскоп DMIRE TCS SP2 (Leica, Германия), оснащенный иммерсионным объективом HCX PL APO lbd.BL 63x/NA1.4 oil (Leica, Германия), аргоновым (488 нм) и гелий-неоновым (543 нм) лазерами.

Обработка изображений производилась с использованием программ LAS AF Lite, BZ-9000 Analyzer и Fiji [160].

Микроскопия сверхвысокого разрешения была выполнена в лаборатории флуоресцентного биоимиджинга (НижГМА, Нижний Новгород, Россия) с использованием микроскопа Eclipse Ti N-STORM (Nikon, Япония), оборудованного EM-CCD камерой iXon3 DU-897 (Andor, Англия) и оснащенного иммерсионным объективом Apo TIRF 100x/NA1.49. В работе использовались лазеры с длиной волны излучения 488 и 561 нм. Для обработки изображений был использован плагин ThunderStorm [161] для Fiji [160].

3.15. Компьютерное моделирование

Расчет геометрических параметров лигандов был выполнен Михаилом Барановым и Романом Гриценко (группа синтеза природных соединений ИБХ РАН) на основе теории функционала плотности (Density Functional Theory, DFT) с применением гибридного обменно-корреляционного функционала B3LYP (Becke, three-parameter, Lee-Yang-Parr) [162], который был ранее успешно использован для сходных систем [163–165].

Вычисления проводились в рамках базисных наборов def2-SVP или 6-311+G** без учета влияния растворителя, а стационарный характер полученных структур был подтвержден отсутствием мнимых частот колебаний в расчетных спектрах.

Расчет оптических свойств лигандов был выполнен Романом Гриценко с использованием программы ORCA 3.0.3[166].

Для различных задач компьютерного моделирования белков в работе использовались программы MODELLER, AutoDoc Vina и Rosetta. Подробные описания областей и способов применения каждой из программ приведены ниже.

3.15.1. MODELLER

Программу MODELLER [129] применяли для гомологичного моделирования 3D структур мутантов белка B1c с одной или двумя заменами при первичном *in silico* скрининге. Для каждой уникальной аминокислотной последовательности генерировали одну 3D структуру, которую затем использовали для докинга.

3.15.2. AutoDoc Vina

При первичном *in silico* скрининге проводили докинг с жесткой структурой белка, который осуществляли в программе AutoDoc Vina согласно ранее опубликованному протоколу [119]. Для подготовки лиганда и рецептора использовали программу AutoDockTools [118] версии 1.5.6. Область связывания рассчитывали в программе PyMol. Максимально сохраняли 20 различных сайтов связывания.

3.15.3. Rosetta

Программу Rosetta использовали для более детального исследования (моделирования и докинга) структур, представляющий особый интерес, а также для *in silico* дизайна.

Перед работой энергии всех кристаллических структур, полученных из базы PDB, были минимизированы с помощью протокола Relax [167] со следующими свойствами:

```
-flip_HNQ  
-no_optH false  
-relax:constrain_relax_to_start_coords  
-nstruct50  
-ex1  
-ex2  
-use_input_sc
```

Структуру с минимальной общей энергией использовали для всех дальнейших расчетов.

После внесения аминокислотных замен в кристаллические структуры известных белков процедуру минимизации энергии повторяли. Если замен было внесено немного (одна-две), то приведенный выше протокол не изменяли. При внесении большего количества замен (три и больше) минимизацию проводили, не используя свойство `-relax:constrain_relax_to_start_coords`.

Докинг низкомолекулярных лигандов (флуорогенов) осуществляли согласно опубликованному протоколу [121], заменив функции “Translate”, “Rotate” и “SlideTogether” на

позднее добавленную функцию “Transform” [122]. Для каждой пары белок-флуороген проводили три раунда моделирования. Сначала получали 5 000 структур, разрешив алгоритму двигать лиганд на расстояние до 5 Å и свободно вращать. Для следующего раунда отбирали 50 лучших структур (50 структур с самыми низким значением `interface_delta` (параметр, показывающий энергию взаимодействия лиганда с белком) среди 2 000 структур с наименьшей общей энергией). Во втором раунде снова получали 5 000 структур (по 100 из каждой выбранной по результатам первого раунда), ограничив максимальное расстояние передвижения лиганда до 1 Å, а вращение — до 45°. Аналогично выбранные 50 лучших структур второго раунда использовались для третьего, последнего, призванного скорректировать и улучшить расположение флуорогена в найденном сайте связывания. При этом разрешённое расстояние перемещения составляло 0.2 Å, а угол вращения — 5°. 50 лучших (из 5 000) структур третьего раунда использовали для анализа положения лиганда в белке, энергии их взаимодействия, вклада отдельных аминокислот во взаимодействие между белком и лигандом и т.д.

Для улучшения имеющихся белков методом *in silico* мутагенеза использовали алгоритм, разработанный с помощью сотрудников лаборатории Йенса Майлера (Meiler Lab, Vanderbilt University, Nashville (TN), США). Алгоритм вносил минимальное количество изменений: точечные замены в аминокислотные остатки, непосредственно контактирующие с лигандом (Cα атомы которых расположены на расстоянии не более 6 Å от флуорогена) или координирующие положение таких аминокислотных остатков (Cα атомы расположены на расстоянии не более 8 Å от флуорогена и боковые цепи повернуты в сторону лиганда). Для минимизации количества вносимых замен был введен дополнительный энергетический бонус за присутствие исходных аминокислот. В качестве стартовых структур использовали структуры с самыми низкими показателями общей энергии и энергии взаимодействия белка с лигандом, полученные в результате докинга. Пример используемого для работы алгоритма XML-скрипта приведен на рисунке 20.

```

<ROSETTASCRIPTS>
<SCOREFXNS>
  <ligand_soft_rep weights=ligand_soft_rep>
  </ligand_soft_rep>
  <hard_rep weights=ligandprime>
  </hard_rep>
</SCOREFXNS>
<SCORINGGRIDS ligand_chain="X" width="16">
  <vdw grid_type="ClassicGrid" weight="1.0"/>
</SCORINGGRIDS>
<TASKOPERATIONS>
  <DetectProteinLigandInterface name=design_interface cut1=6.0 cut2=8.0 cut3=10.0 cut4=12.0
  design=1/>
</TASKOPERATIONS>
<LIGAND_AREAS>
  <docking_sidechain chain=X cutoff=6.0 add_nbr_radius=true all_atom_mode=false
  minimize_ligand=10/>
  <final_sidechain chain=X cutoff=6.0 add_nbr_radius=true all_atom_mode=false/>
  <final_backbone chain=X cutoff=7.0 add_nbr_radius=false all_atom_mode=true
  Calpha_restraints=0.3/>
</LIGAND_AREAS>
<INTERFACE_BUILDERS>
  <side_chain_for_docking ligand_areas=docking_sidechain/>
  <side_chain_for_final ligand_areas=final_sidechain/>
  <backbone ligand_areas=final_backbone extension_window=3/>
</INTERFACE_BUILDERS>
<MOVEMAP_BUILDERS>
  <docking_sc_interface=side_chain_for_docking minimize_water=true/>
  <final_sc_interface=side_chain_for_final bb_interface=backbone minimize_water=true/>
</MOVEMAP_BUILDERS>
<MOVERS>
  <Transform name="transform" chain="X" box_size="8.0" move_distance="5.0" angle="360"
  cycles="5000" temperature="5" initial_perturb="1.0"/>
  <HighResDocker name=high_res_docker cycles=1 repack_every_Nth=1 scorefxn=ligand_soft_rep
  movemap_builder=docking/>
  <FinalMinimizer name=final scorefxn=hard_rep movemap_builder=final/>
  <InterfaceScoreCalculator name=add_scores chains=X scorefxn=hard_rep/>
  <FavorNativeResidue name=favor_native bonus=1.0/>
  <ddG name=calculateDDG jump=1 per_residue_ddg=1 scorefxn=hard_rep/>
  <PackRotamersMover name=designinterface scorefxn=hard_rep task_operations=design_interface/>
</MOVERS>
<PROTOCOLS>
  <Add mover_name=transform/>
  <Add mover_name=favor_native/>
  <Add mover_name=designinterface/>
  <Add mover_name=high_res_docker/>
  <Add mover_name=final/>
  <Add mover_name=calculateDDG/>
  <Add mover_name=add_scores/>
</PROTOCOLS>
</ROSETTASCRIPTS>

```

Рисунок 20. Пример XML-скрипта для дизайна белков, связывающих флуорогены.

4. Результаты и обсуждение

4.1. Разработка и изучение новых спектральных вариантов флуоресцентных белков, содержащих триптофан в составе хромофора

Замена тирозина хромофора на триптофан привела к получению ряда флуоресцентных белков с выдающимися свойствами: цианового белка с 93%-ным квантовым выходом флуоресценции [24], зелёного белка с самым длинным временем жизни флуоресценции [168,169] и оранжевого генетически кодируемого фотосенсибилизатора [170]. Кроме того, белки с хромофорами на основе триптофана оказались прекрасными FRET-донорами: в настоящее время наиболее часто используемыми FRET-парами являются пары на основе цианового и жёлтого белков [171,172]. Однако существует лишь несколько таких белков, характеризующихся флуоресценцией в жёлтой и/или оранжевой областях спектра [88,170,173]. Рациональные подходы к разработке улучшенных вариантов затруднены из-за недостаточной информации о структуре хромофора таких белков.

4.1.1. Получение оранжевых триптофановых флуоресцентных белков

Методом направленного мутагенеза в кодирующую последовательность белка FusionRed была внесена мутация, приводящая к замене 66 тирозина (здесь и далее нумерация в соответствии с выравниванием относительно GFP из *Aequorea victoria* (avGFP)), входящего в состав хромофора, на триптофан. При экспрессии полученного белка (FR-Y66W) в *E. coli* слабая флуоресценция колоний при облучении синим светом наблюдалась лишь на пятый день инкубации при 4°C. Чтобы найти компенсаторные мутации, восстанавливающие флуоресценцию, был проведен случайный мутагенез, в результате которого были обнаружены две группы белков, обладающие жёлтой и оранжевой флуоресценцией. Самые яркие представители этих двух популяций, FR66W-Yellow и FR66W-Orange, соответственно, были выделены и отсекарованы.

Белок FR66W-Yellow содержал две (M14T / N192K), а FR66W-Orange три (M14T / K15R / C165G) аминокислотные замены относительно FR-Y66W (рисунок 21).

	10	20	30	40	50		
avGFP	MSKGEELFTGVV	PILVELDGDVNGHKFS	VS	SGEGEGDATY	GKLT	LKFI	CTTG-KLPVPWPT
FusionRed	MVSELIKENM	PMKLYMEGTVNNHHFKCT	SE	GEGKPYEGT	QTMRIK	VVEGGPL	PFAFDI
FR66W-Yellow	MVSELIKENM	PTKLYMEGTVNNHHFKCT	SE	GEGKPYEGT	QTMRIK	VVEGGPL	PFAFDI
FR66W-N192K	MVSELIKENM	PMKLYMEGTVNNHHFKCT	SE	GEGKPYEGT	QTMRIK	VVEGGPL	PFAFDI
FR66W-C165S	MVSELIKENM	PMKLYMEGTVNNHHFKCT	SE	GEGKPYEGT	QTMRIK	VVEGGPL	PFAFDI
FR66W-C165A	MVSELIKENM	PMKLYMEGTVNNHHFKCT	SE	GEGKPYEGT	QTMRIK	VVEGGPL	PFAFDI
FR66W-C165G	MVSELIKENM	PMKLYMEGTVNNHHFKCT	SE	GEGKPYEGT	QTMRIK	VVEGGPL	PFAFDI
FR66W-Orange	MVSELIKENM	PTRLYMEGTVNNHHFKCT	SE	GEGKPYEGT	QTMRIK	VVEGGPL	PFAFDI
	60	70	80	90	100	110	
avGFP	LVTTFSYGVQCFS	RYPDHMKQH	DFFKSAMPEGYVQERT	IFFKDDGNYKTRAEV	KFE	GDT	
FusionRed	LATSFMYGSRTFI	KHPPGIP--	DFFKQSFPEGFTWERV	TTYEDGGVLTATQD	TS	LQDGC	
FR66W-Yellow	LATSFMYGSRTFI	KHPPGIP--	DFFKQSFPEGFTWERV	TTYEDGGVLTATQD	TS	LQDGC	
FR66W-N192K	LATSFMYGSRTFI	KHPPGIP--	DFFKQSFPEGFTWERV	TTYEDGGVLTATQD	TS	LQDGC	
FR66W-C165S	LATSFMYGSRTFI	KHPPGIP--	DFFKQSFPEGFTWERV	TTYEDGGVLTATQD	TS	LQDGC	
FR66W-C165A	LATSFMYGSRTFI	KHPPGIP--	DFFKQSFPEGFTWERV	TTYEDGGVLTATQD	TS	LQDGC	
FR66W-C165G	LATSFMYGSRTFI	KHPPGIP--	DFFKQSFPEGFTWERV	TTYEDGGVLTATQD	TS	LQDGC	
FR66W-Orange	LATSFMYGSRTFI	KHPPGIP--	DFFKQSFPEGFTWERV	TTYEDGGVLTATQD	TS	LQDGC	
	120	130	140	150	160	170	
avGFP	LVNRIELKGIDF	KEDGNILGHKLE	YNNYNSHN	VYIMADKQKNGIK	VNF	KIRHNIEDGSVQL	
FusionRed	LIYNVKVRGVNF	PANGPVMQKKTL	-GWEASTET	MYPA--DGGLEGAC	DMAL	KLVGGGHLI	
FR66W-Yellow	LIYNVKVRGVNF	PANGPVMQKKTL	-GWEASTET	MYPA--DGGLEGAC	DMAL	KLVGGGHLI	
FR66W-N192K	LIYNVKVRGVNF	PANGPVMQKKTL	-GWEASTET	MYPA--DGGLEGAC	DMAL	KLVGGGHLI	
FR66W-C165S	LIYNVKVRGVNF	PANGPVMQKKTL	-GWEASTET	MYPA--DGGLEGAC	DMAL	KLVGGGHLI	
FR66W-C165A	LIYNVKVRGVNF	PANGPVMQKKTL	-GWEASTET	MYPA--DGGLEGAC	DMAL	KLVGGGHLI	
FR66W-C165G	LIYNVKVRGVNF	PANGPVMQKKTL	-GWEASTET	MYPA--DGGLEGAC	DMAL	KLVGGGHLI	
FR66W-Orange	LIYNVKVRGVNF	PANGPVMQKKTL	-GWEASTET	MYPA--DGGLEGAC	DMAL	KLVGGGHLI	
	180	190	200	210	220	230	
avGFP	ADHYQQNTPIGD	-GPVLLPDNHYL	STQSALSKDPNE	KRDH	MVLL	EFVTAAGITHGMDELY	
FusionRed	CNLETTYRSKK	PATNLKMPGVYN	VDRRLERIKE	-ADDETYVEQ	HEVAVARYST	GAGDGG	
FR66W-Yellow	CNLETTYRSKK	PATNLKMPGVYN	VDRRLERIKE	-ADDETYVEQ	HEVAVARYST	GAGDGG	
FR66W-N192K	CNLETTYRSKK	PATNLKMPGVYN	VDRRLERIKE	-ADDETYVEQ	HEVAVARYST	GAGDGG	
FR66W-C165S	CNLETTYRSKK	PATNLKMPGVYN	VDRRLERIKE	-ADDETYVEQ	HEVAVARYST	GAGDGG	
FR66W-C165A	CNLETTYRSKK	PATNLKMPGVYN	VDRRLERIKE	-ADDETYVEQ	HEVAVARYST	GAGDGG	
FR66W-C165G	CNLETTYRSKK	PATNLKMPGVYN	VDRRLERIKE	-ADDETYVEQ	HEVAVARYST	GAGDGG	
FR66W-Orange	CNLETTYRSKK	PATNLKMPGVYN	VDRRLERIKE	-ADDETYVEQ	HEVAVARYST	GAGDGG	

Рисунок 21. Выравнивание аминокислотных последовательностей белков. Цветом выделены внесенные замены. Серым выделены аминокислоты, боковые цепи которых направлены внутрь β -бочонка флуоресцентного белка.

4.1.2. Изучение роли замен

Для определения роли обнаруженных в ходе случайного мутагенеза замен, мы протестировали каждую из них индивидуально. Только два из пяти белков, FR66W-C165G и FR66W-N192K, обладали заметной флуоресценцией в первый день после экспрессии. FR66W-C165G демонстрировал оранжевую флуоресценцию и был практически таким же ярким, как и FR66W-Orange. FR66W-N192K обладал жёлтой флуоресценцией и был заметно тусклее, чем FR66W-Yellow.

Кристаллическая структура белка FusionRed не решена, однако существует кристаллическая структура его предшественника, белка mKate [174]. Используя эту структуру, а также литературные данные, мы попытались проанализировать роль внесенных замен (рисунок 22). Боковая цепь аминокислотного остатка в положении 165 находится напротив остатка второй (ароматической) аминокислоты хромофорной триады и известна своей способностью влиять на свойства флуоресцентных белков, в том числе за счёт влияния на степень протонированности, цис-транс изомеризацию и поляризацию хромофора [175]. Способ влияния аминокислотного остатка в положении 192 на спектральные свойства флуоресцентного белка менее очевиден: он располагается в петле и не имеет возможности непосредственно взаимодействовать с хромофором. Возможно, он влияет на микроокружение хромофора за счёт репозиции 9 и/или 10 β -тяжа, соединенных этой петлей.

Чтобы подробнее изучить влияние замен в положении 165, мы провели насыщающий мутагенез белка FR-Y66W по этому положению. Большая часть колоний была практически нефлуоресцентной после 20 часов на 37°C. Лишь мутанты с серином и аланином в указанном положении (белки FR66W-C165S и FR66W-C165A, соответственно) обладали ярко выраженной жёлтой флуоресценцией. Они, а также белки FR66W-C165G, FR66W-Yellow и FR66W-Orange были выделены и охарактеризованы более подробно (таблица 1, рисунок 23).

Таблица 1. Оптические свойства изученных мутантов белка FusionRed и их сравнение со свойствами ранее опубликованных триптофан-содержащих флуоресцентных белков с эмиссией в жёлто-оранжевой области спектра.

Мутации (помимо Y66W)	Возбуждение, нм	Эмиссия, нм	Стоксов сдвиг, нм	FQY, %	ϵ , $M^{-1} \cdot cm^{-1}$
M14T, N192K	490	535	45	28	22000
C165S	499	543	44	n.d.	n.d.
C165A	490	553	63	25	29000
C165G	529	565	36	n.d.	n.d.
M14T, K15R, C165G	530	565	35	26	19000
emKate_Y64W [176]	479	558	79	n.d.	n.d.
mHoneydew [173]	487/504	537/562	50/58	12	17000

FQY (fluorescence quantum yield) — квантовый выход флуоресценции

ϵ — коэффициент экстинкции, n.d. — не определяли

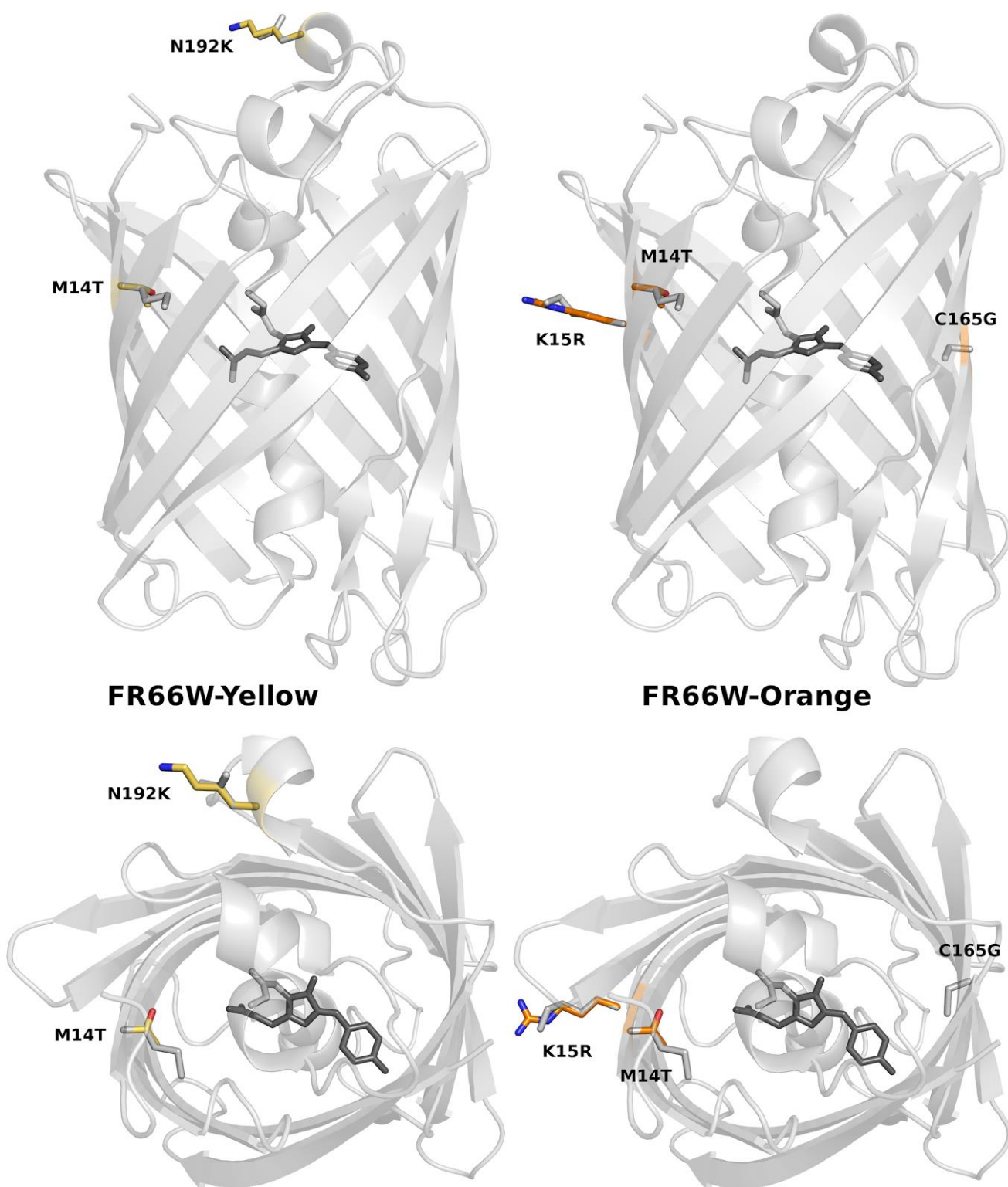


Рисунок 22. Предполагаемые позиции аминокислотных остатков, замененных в белках FR66W-Yellow и FR66W-Orange. Моделирование на основе предшественника белка FusionRed, белка mKate (PDB ID 3BXB).

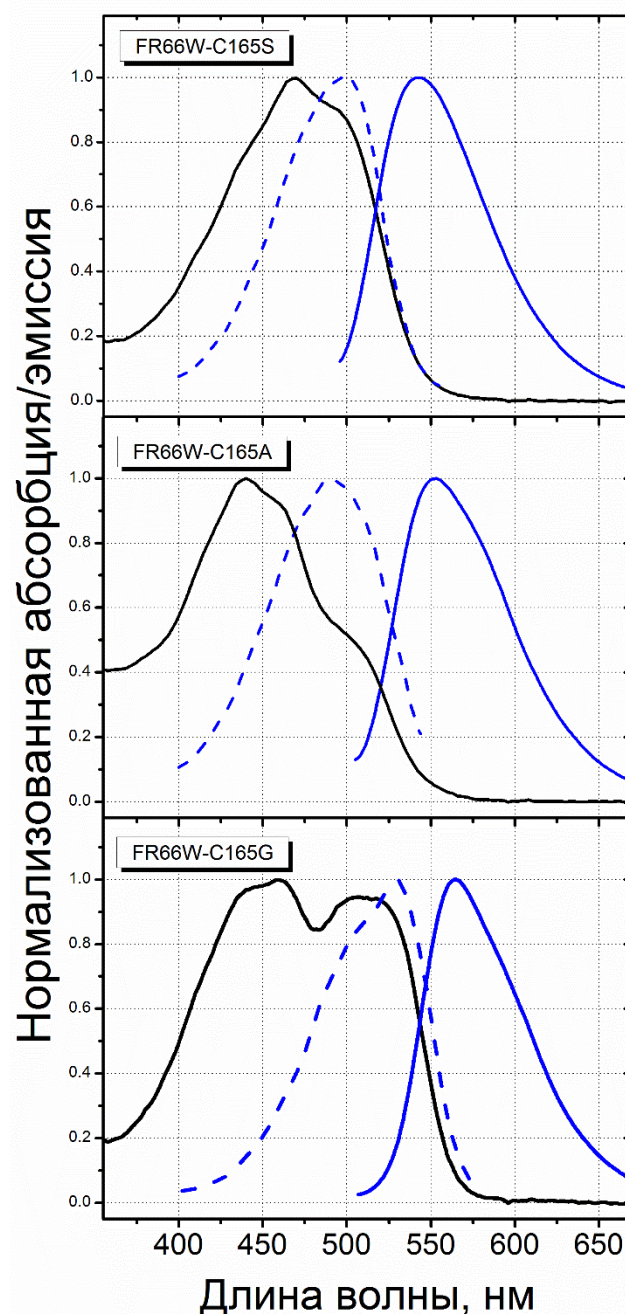
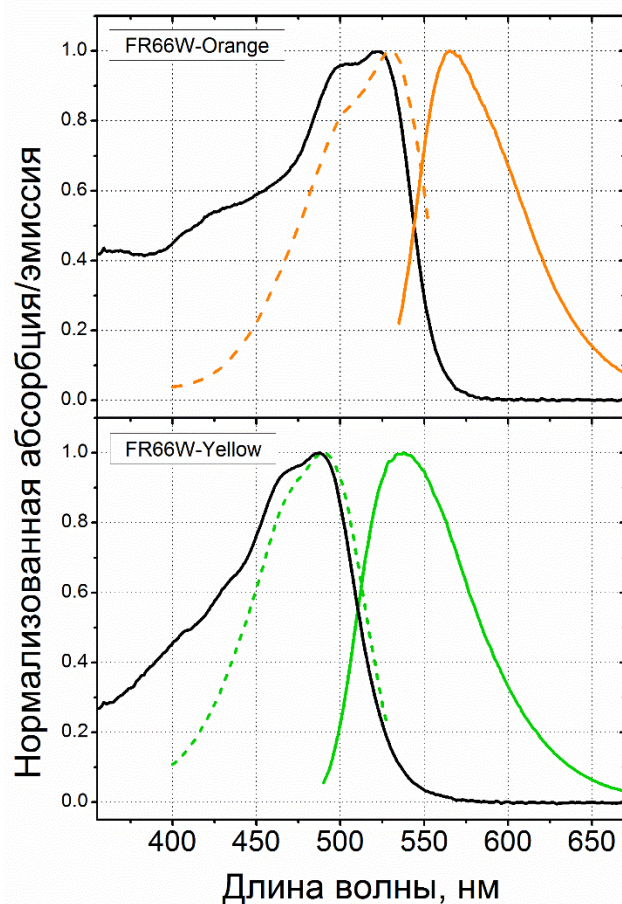


Рисунок 23. Спектральные свойства белков FR66W-C165S, FR66W-C165A, FR66W-C165G, FR66W-Yellow и FR66W-Orange. Чёрная сплошная линия — спектр поглощения, прерывистая цветная линия — спектр возбуждения флуоресценции, сплошная цветная линия — спектр эмиссии флуоресценции.

В отличие от первого опубликованного триптофан-содержащего флуоресцентного белка с эмиссией в жёлто-оранжевой области спектра mHoneydew [173], а также всех ранее охарактеризованных триптофан-содержащих циановых белков, спектры триптофан-содержащих мутантов белка FusionRed имели лишь один пик возбуждения и эмиссии флуоресценции, а не «двугорбые» пики (рисунок 24), которые, как было принято ранее считать, представляют собой фотофизическую особенность хромофора, содержащего индольное кольцо [177]. Мы решили

исследовать природу различия спектральных свойств полученных нами жёлтых и оранжевых флуоресцентных белков, подробнее изучив два их представителя — белки FR66W-Yellow и FR66W-Orange.

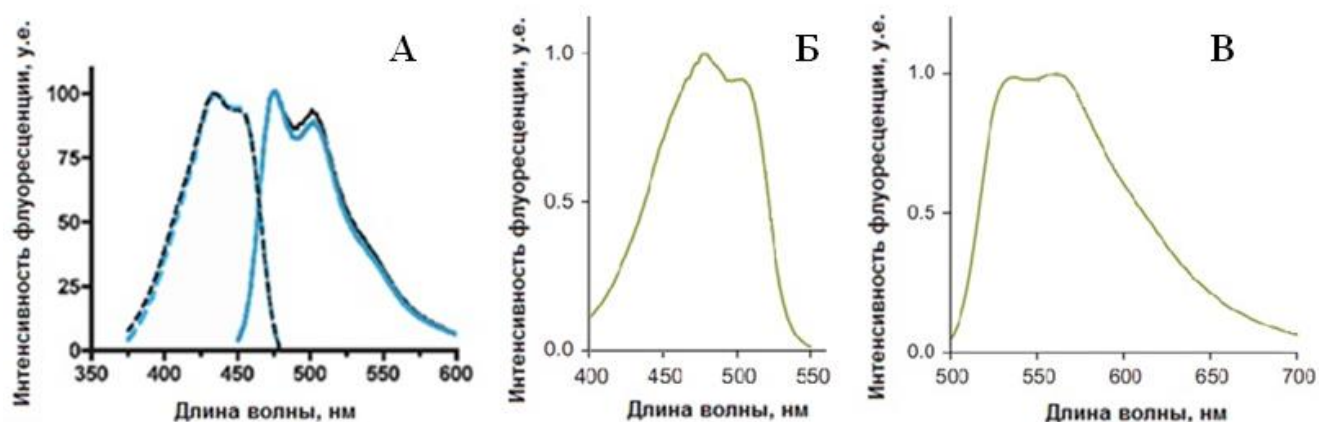


Рисунок 24. “Двугорбые” спектры возбуждения и эмиссии флуоресценции ранее опубликованных белков, содержащих триптофан в составе хромофора. **А.** Спектры возбуждения (пунктирная линия) и эмиссии (сплошная линия) флуоресценции белков ECFP (черная линия) и Cerulean (синяя линия) [178]; **Б.** Спектр возбуждения флуоресценции белка mHoneydew; **В.** Спектр эмиссии флуоресценции белка mHoneydew [173].

4.1.3. Возможные структуры хромофора

На основе данных о структурах и свойствах хромофоров других, в основном тирозин-содержащих, флуоресцентных белков [179–184], можно сделать вывод, что подобные спектральные различия могут объясняться как формированием разных хромофоров в данных группах белков, так и нахождением одного и того же хромофора в разных состояниях.

Созревание хромофора в исследуемых белках может идти разными путями (рисунок 25). Замыкание имидазолонового кольца и окисление C_{α} -N связи 65-ой аминокислоты будет приводить к образованию ацилиминового (DsRed-подобного) хромофора. Образовавшаяся связь $C_{\alpha}=N$ может быть затем гидролизована, что приведет к разрыву белковой цепи и формированию кетонного (asFP595-подобного [184]) хромофора (рисунок 25, путь А). Окисление C_{α} -N связи может и не произойти, в результате чего образуется GFP-подобный хромофор (рисунок 25, путь В).

При прочих равных спектры триптофан-содержащих хромофоров всегда сдвинуты в коротковолновую область спектра относительно тирозин-содержащих белков (GFP-подобный триптофан-содержащий хромофор характерен для циановых белков). Наиболее длинноволновый GFP-подобный тирозин-содержащий хромофор содержится в жёлтых флуоресцентных белках и формируется за счёт стэкинг-взаимодействия тирозина хромофора с другим белковым

тирозином. Максимумы спектров эмиссии флуоресценции GFP-подобных тирозин-содержащих жёлтых белков находятся в районе 530 нм. Оба исследуемых триптофан-содержащих белка не содержат каких-либо ароматических аминокислот, способных образовать стэкинг с хромофором. Однако даже гипотетический стэкинг в этих белках проявлялся бы в сдвиге их спектров эмиссии в синюю область спектра, и мы бы наблюдали зелёную, а не оранжевую, флуоресценцию. Таким образом, образование GFP-подобного хромофора в исследуемых белках представляется маловероятным, хотя наличие дополнительного максимума поглощения в спектре FR66W-S165G может быть связано как раз с образованием такого хромофора с очень низким квантовым выходом флуоресценции.

При одинаковом белковом окружении кетонный хромофор будет краснее своего ацилиминового предшественника [179,180,184,185]. Таким образом, если разница между жёлтыми и оранжевыми белками обеспечивается формированием разных хромофоров, FR66W-Yellow должен содержать DsRed-подобный, а FR66W-Orange — asFP595-подобный хромофор.

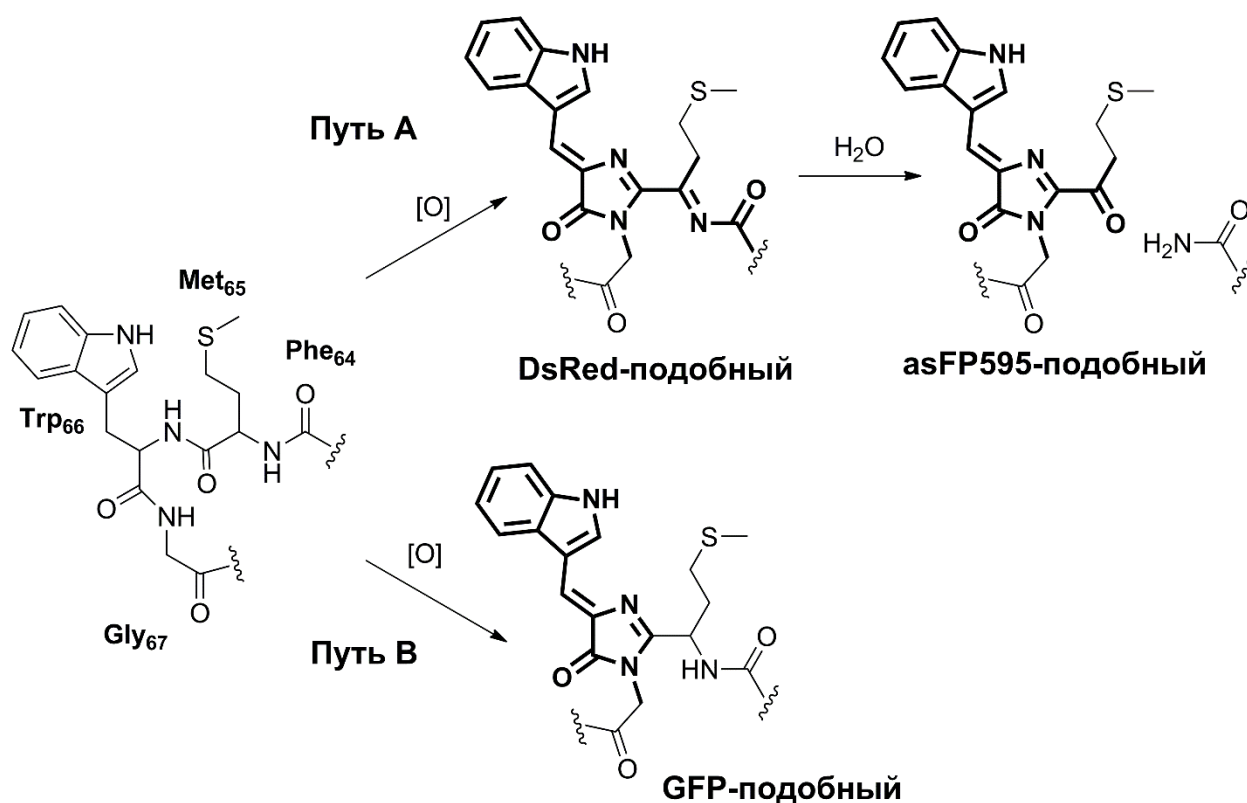


Рисунок 25. Возможные пути созревания хромофора.

4.1.4. Поведение при денатурации

Денатурация белка делает хромофор, находящийся в нативном состоянии в центре белковой глобулы, доступным для растворителя. Электрофорез в полиакриламидном геле дает информацию о размере и целостности белковой цепи. Таким образом, денатурирующий электрофорез в полиакриламидном геле может дать некоторую информацию о состоянии хромофора в нативном белке.

Мы проанализировали предварительно прогретые и не прогретые образцы белков FR66W-Yellow и FR66W-Orange при помощи денатурирующего электрофореза в полиакриламидном геле (рисунок 27). Во всех случаях на геле хорошо различимы три полосы с кажущимися молекулярными массами примерно 27, 20 и 8 кДа. Полоса 27 кДа соответствует стандартному размеру полноразмерного флуоресцентного белка. Две короткие полосы (20 и 8 кДа) похожи на результат разрыва белковой цепи в районе хромофора. Слабая полоса с кажущейся молекулярной массой примерно 22 кДа в непрогретых образцах скорее всего является неденатурированным или частично денатурированным белком.

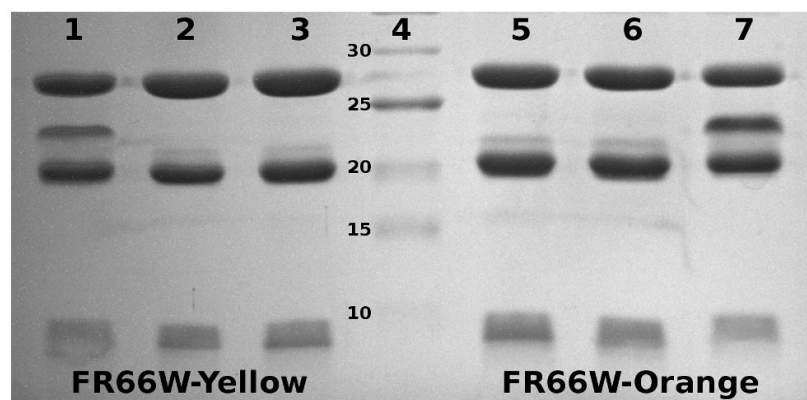


Рисунок 27. Анализ белков FR66W-Yellow и FR66W-Orange при помощи денатурирующего электрофореза в полиакриламидном (15%) геле. 1,2,3 — белок FR66W-Yellow (непрогретый образец, образец, прогретый в течение 0.5 и 10 минут, соответственно). 4 — Маркер. 5,6,7 — белок FR66W-Orange (образец, прогретый в течение 10 и 0.5 минут, соответственно и непрогретый образец).

Так как мы не увидели значительной разницы в интенсивности коротких полос (20 и 8 кДа) между тремя образцами, можно предположить, что разрыв белковой цепи либо существует в нативном белке, либо происходит очень легко.

Мы также исследовали изменения спектральных свойств белков при денатурации (рисунок 28). В присутствии 1% SDS оба белка потеряли способность к флуоресценции. При этом спектр поглощения белка FR66W-Yellow не претерпел видимых изменений даже после 10 минут прогревания при температуре 95°C, в то время как спектр поглощения белка FR66W-Orange сместился влево и стал совпадать со спектром FR66W-Yellow.

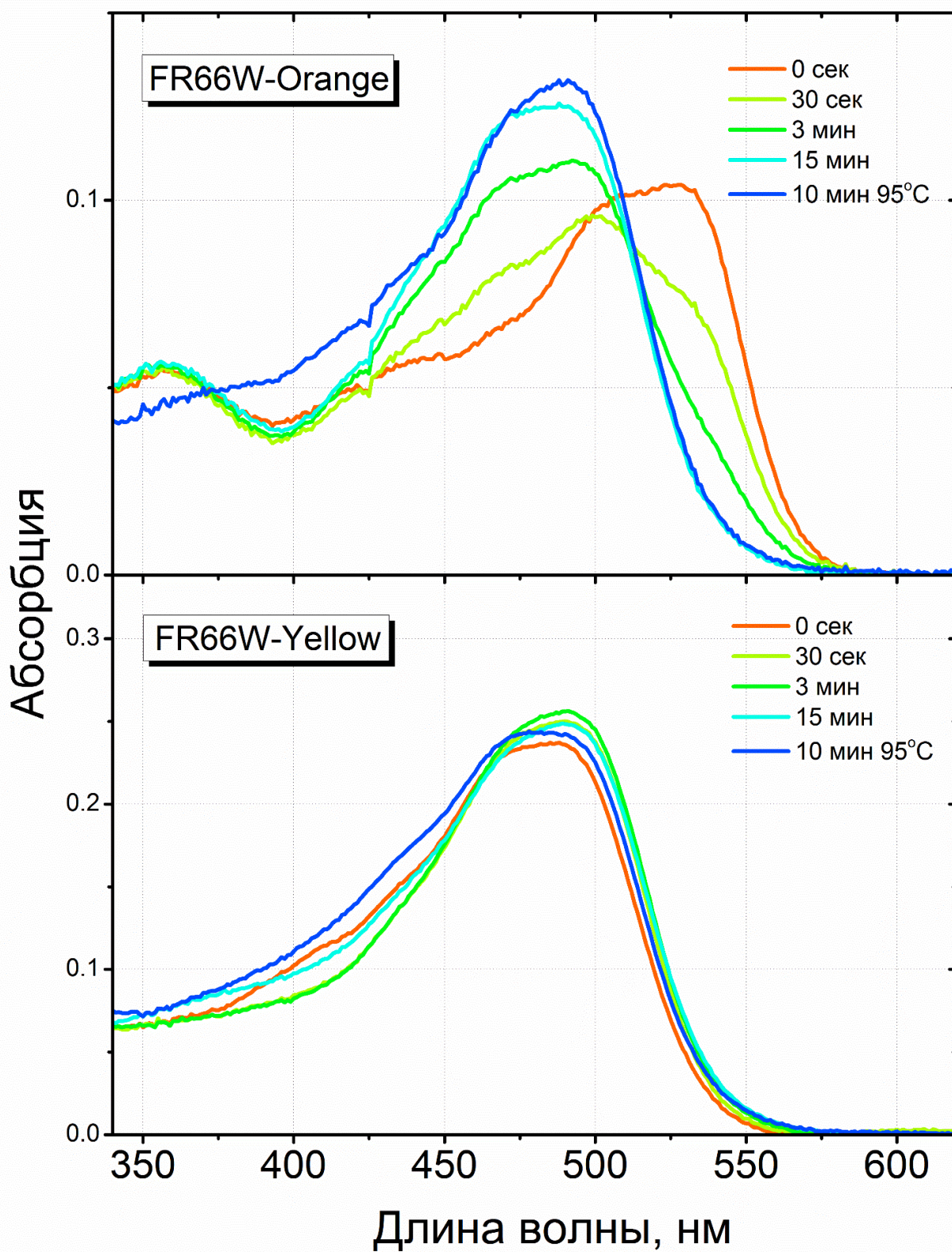


Рисунок 28. Спектры поглощения белков FR66W-Orange и FR66W-Yellow при до и после добавления 1% SDS.

Так как ацилиминовый (DsRed-подобный) хромофор нестабилен в данных условиях, а спектры FR66W-Yellow не изменились, мы пришли к выводу, что FR66W-Yellow содержит кетонный хромофор.

Хотя изменение спектров поглощения белка FR66W-Orange при денатурации теоретически может быть объяснено гидролизом ацилиминового хромофора, для тирозин-содержащих хромофоров ранее было показано, что кетонные хромофоры имеют спектры, сдвинутые в красную область относительно ацилиминовых. Следовательно, формирование ацилиминового хромофора в белке FR66W-Orange также маловероятно.

Таким образом, мы пришли к выводу, что оба белка формируют кетонный хромофор. При этом разница в спектрах может быть вызвана либо его ионизацией в одном из белков, либо присутствием разных изомеров.

4.1.5. pH титрование

Флуоресцентные белки с ионизированными хромофорами обычно крайне чувствительны к изменению pH среды либо из-за прямого титрования хромофора, либо из-за титрования аминокислотных остатков белка, вовлеченных в стабилизацию иона. Особенно это должно быть заметно на триптофан-содержащих хромофорах из-за крайне высокого pK_a индольного фрагмента.

Мы провели титрование обоих белков и не обнаружили значительной pH-чувствительности (рисунки 29, 30). FR66W-Orange характеризуется стабильной флуоресценцией в широком диапазоне pH (4.9-11.0). Изменение интенсивности флуоресценции FR66W-Yellow в диапазоне pH от 6.5 до 11.0 не превышает 20%. Дальнейшее закисление или защелачивание обоих белков приводило к значительному снижению интенсивности флуоресценции, однако мы так и не обнаружили никаких признаков сдвига максимумов спектров.

Таким образом, полученные результаты не подтвердили присутствие ионизованного хромофора ни в одном из исследованных белков.

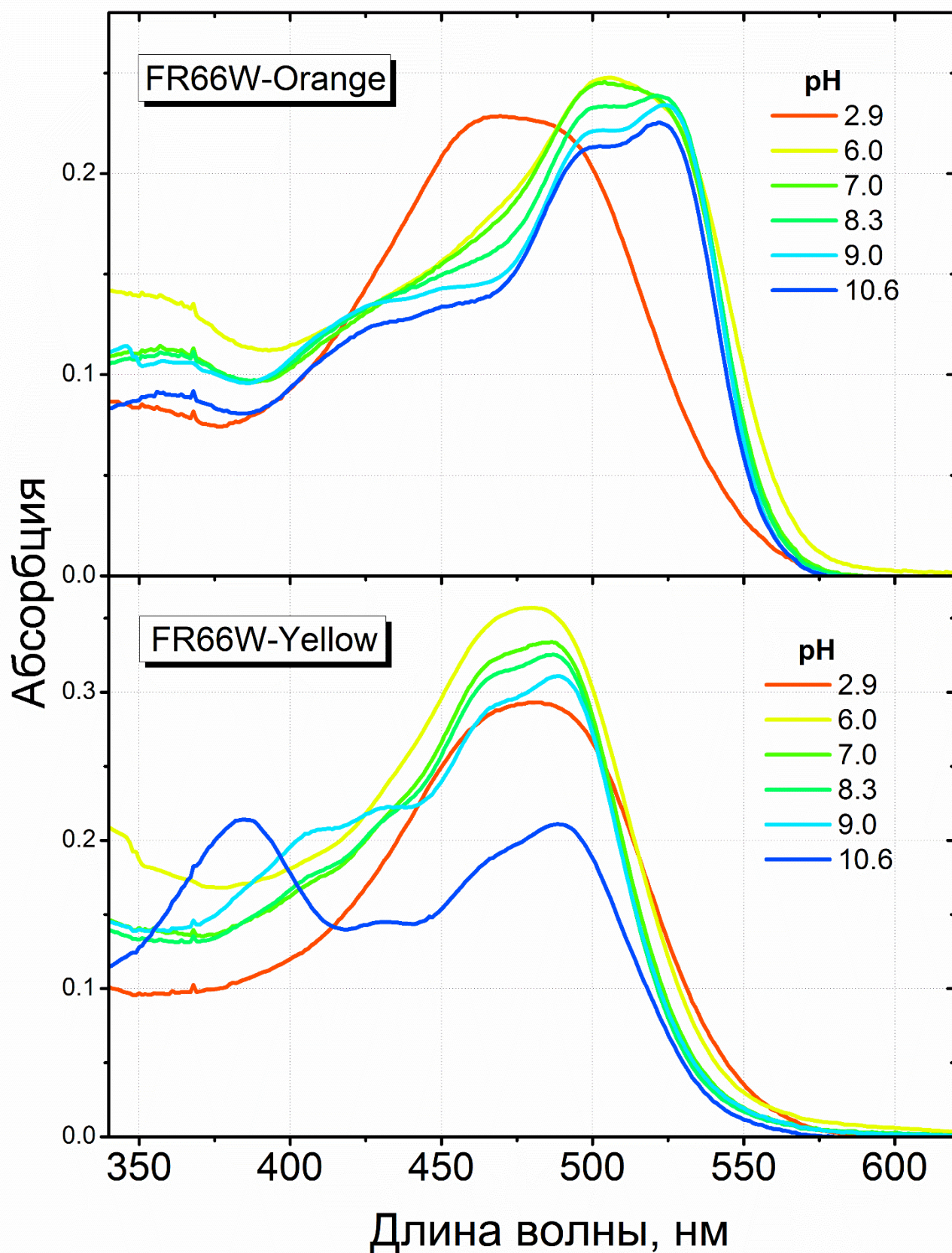


Рисунок 29. Нормализованные спектры поглощения белков FR66W-Yellow и FR66W-Orange при разных значениях pH.

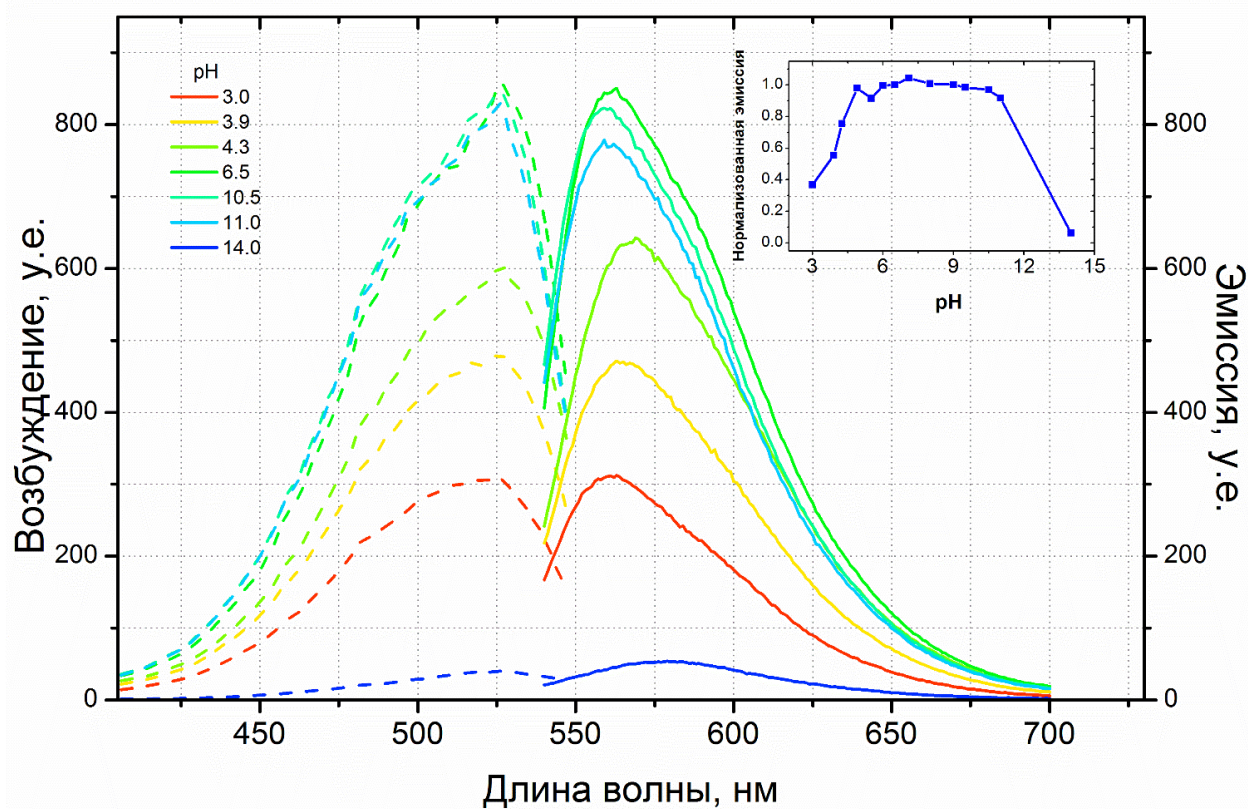
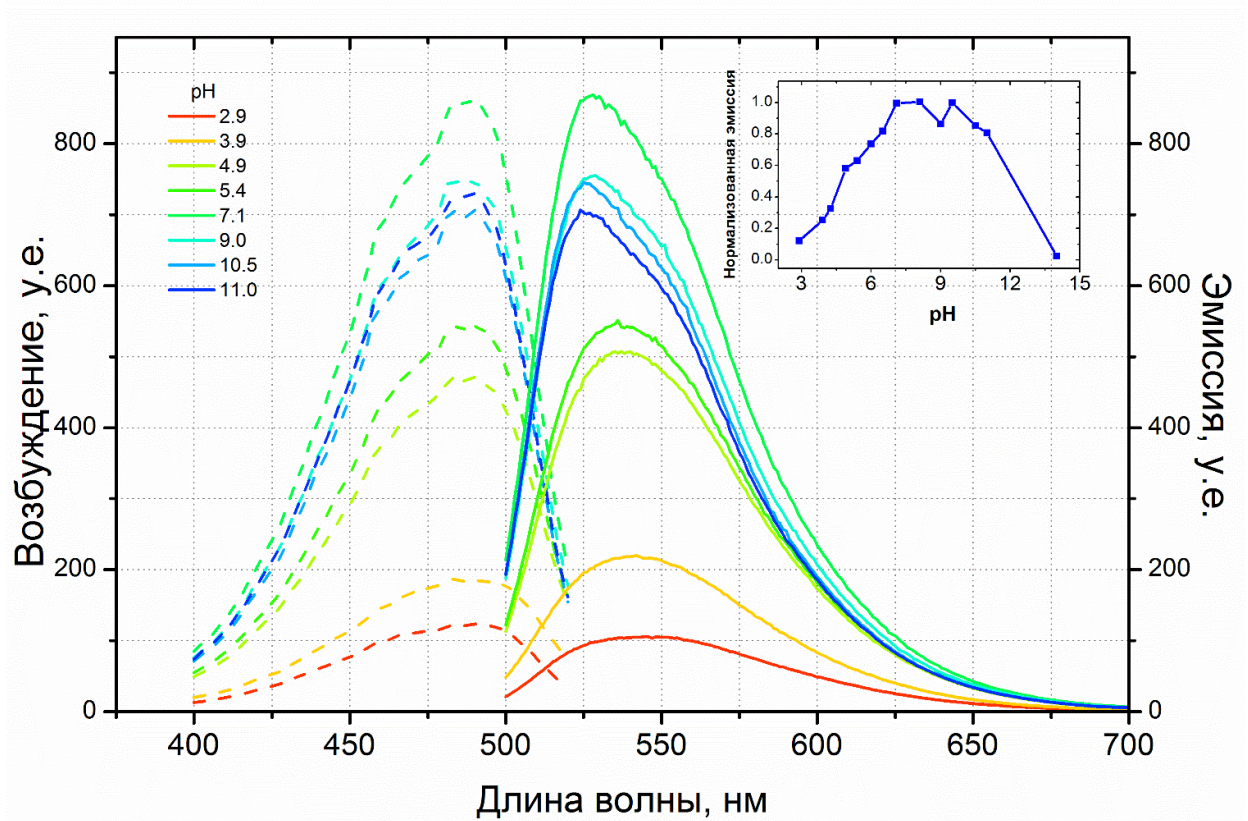


Рисунок 30. Спектры возбуждения и эмиссии флуоресценции белков FR66W-Yellow (верхняя панель) и FR66W-Orange (нижняя панель) при разных значениях pH. Пунктирная линия — возбуждение флуоресценции, сплошная линия — эмиссия флуоресценции.

4.1.6. Фотоконверсия

В тирозин-содержащих хромофорах обычно только один из двух возможных изомеров стабилизирован белковым окружением [188]. Однако изомеризация хромофора в нативном белке иногда возможна и может быть индуцирована изменением pH [189] или облучением светом определенной длины волны [190].

Мы протестировали свет разной длины волны и обнаружили, что облучение бактерий, экспрессирующих FR66W-Yellow, интенсивным синим светом приводит к обратимому сдвигу спектра флуоресценции белка в красную область спектра (рисунок 31)

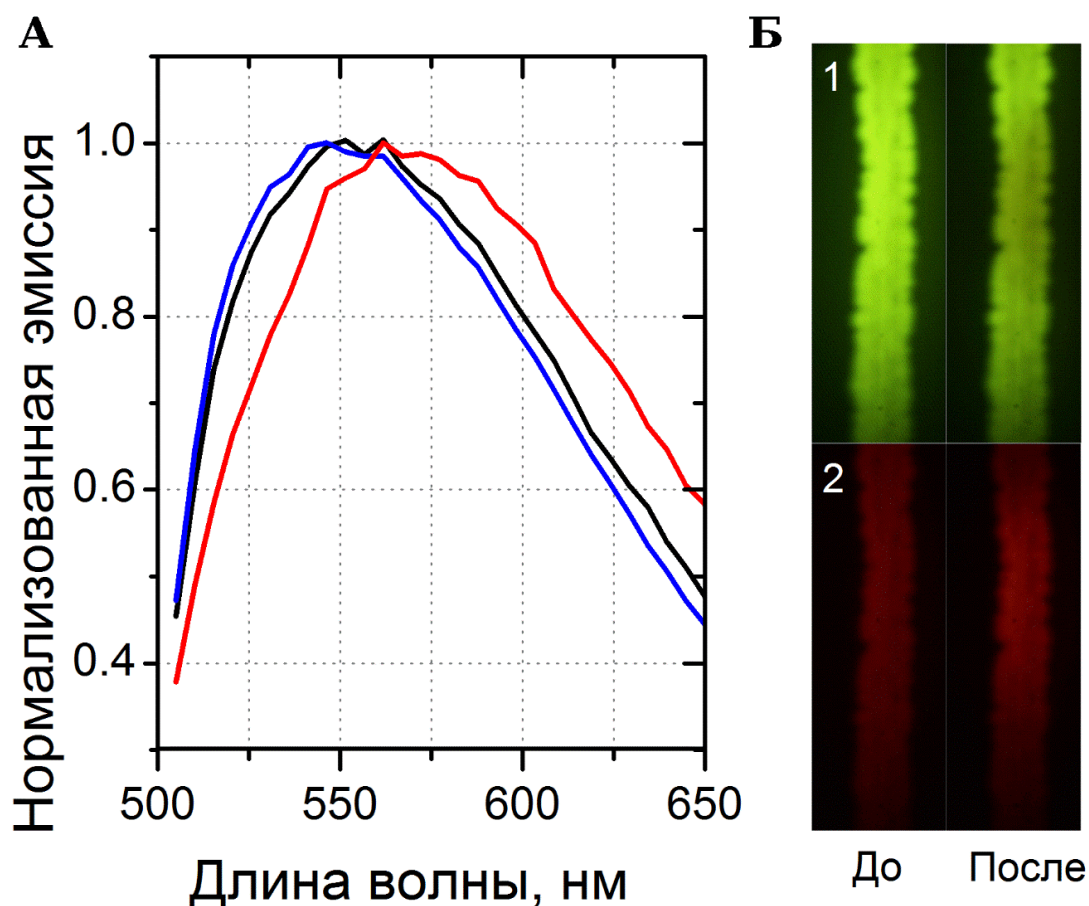


Рисунок 31. Фотоконверсия белка FR66W-Yellow. **А.** Спектры эмиссии флуоресценции бактерий, экспрессирующих белок FR66W-Yellow, измеренные при помощи конфокального микроскопа. Синяя линия — до облучения, красная линия — после 7 секунд облучения синим светом (488 нм, 100 мкВт), черная линия — через 10 минут после облучения. **Б.** Штрих бактерий, экспрессирующих белок FR66W-Yellow, до (слева) и после (справа) 5 минут облучения синим светом (460-490 нм, 14 мВт/см²) в зелёном (1) и красном (2) каналах флуоресцентного бинокуляра.

Обратимость наблюдаемого процесса фотоконверсии ещё раз подтверждает наше предположение о том, что оба белка содержат хромофор одного типа, так как гидролиз ацилиминового хромофора необратим.

4.1.7. Модельное соединение

Синтез аналогов хромофоров флуоресцентных белков является методом изучения их свойств, успешно примененным для ряда тирозин-содержащих [184,191–193], а также одного триптофан-содержащего [168] цианового белка. Поэтому для дальнейшей проверки нашей гипотезе о структуре хромофора FR66W-Yellow и FR66W-Orange, сотрудниками группы синтеза природных соединений ИБХ РАН было синтезировано модельное соединение, имитирующее предполагаемый триптофан-содержащий кетонный хромофор белков (рисунок 32Б). Строение данного соединения было подтверждено с помощью ЯМР-анализа, который был проведен Константином Минеевым из лаборатории биомолекулярной ЯМР-спектроскопии. Двойная связь в структуре хромофора определяет существование Z и E изомеров. В соответствии с данными ЯМР, в растворе синтезированного вещества доминировала Z-форма (93+%). Однако в присутствии сильной кислоты удалось получить смесь, обогащенную (~75%) E-формой. В растворе она довольно быстро превращалась обратно в Z-форму, что хорошо видно на протонных спектрах ЯМР (рисунок 32А), однако могла быть зафиксирована в твердом состоянии.

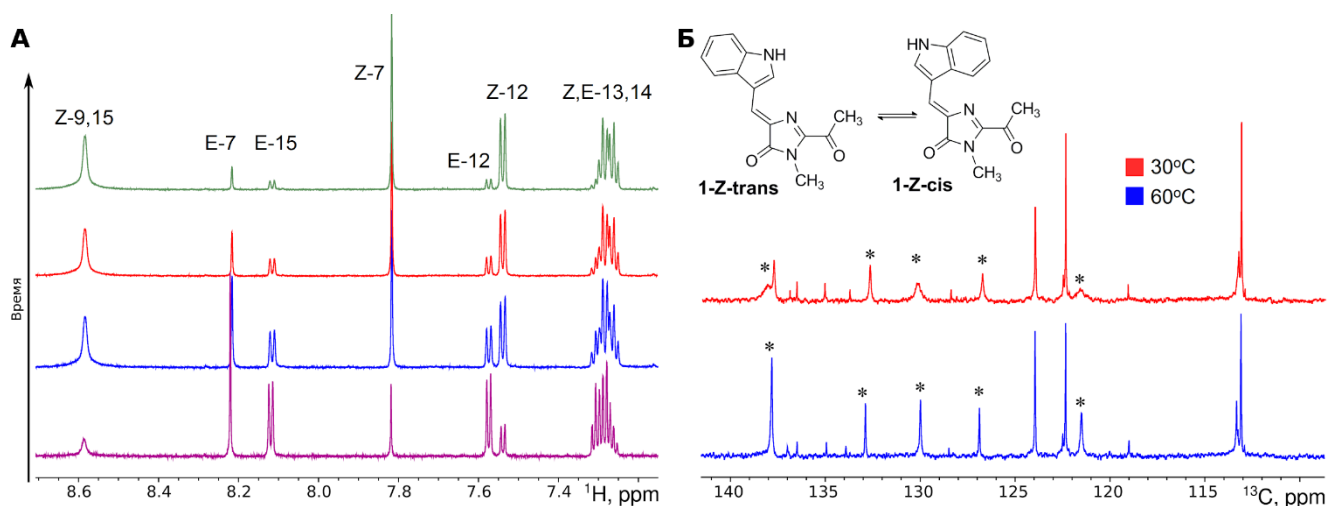


Рисунок 32. Фрагменты ЯМР спектров модельного соединения. **А.** Фрагменты протонных спектров смеси, обогащенной E-формой, снятые в течение 24 часов при 30°C. **Б.** Структуры двух конформеров (Z-цис и Z-транс) и фрагменты ¹³C ЯМР спектров Z-формы, снятые при 30°C (красный) и 60°C (синий). Пики, уширенные из-за вращения индольного кольца, отмечены звездочками.

Аналогично другим синтезированным аналогам хромофоров, синтезированное модельное соединение было практически не флуоресцентным, поэтому нами были проанализированы лишь спектры поглощения. рН титрование соединения подтвердило, что молекула может быть депротонирована, и этот процесс в воде характеризуется $pK_a = 10.5$ (рисунок 33).

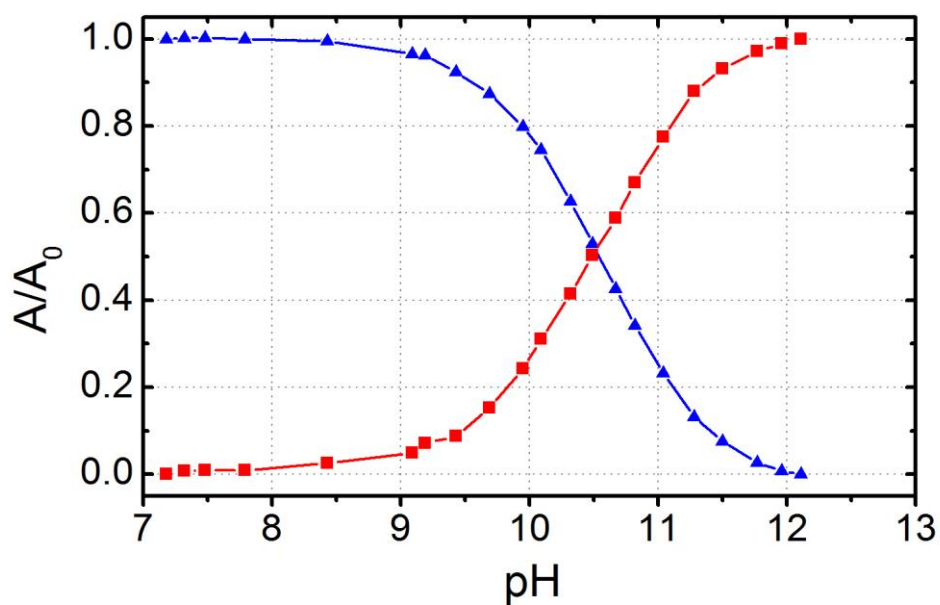
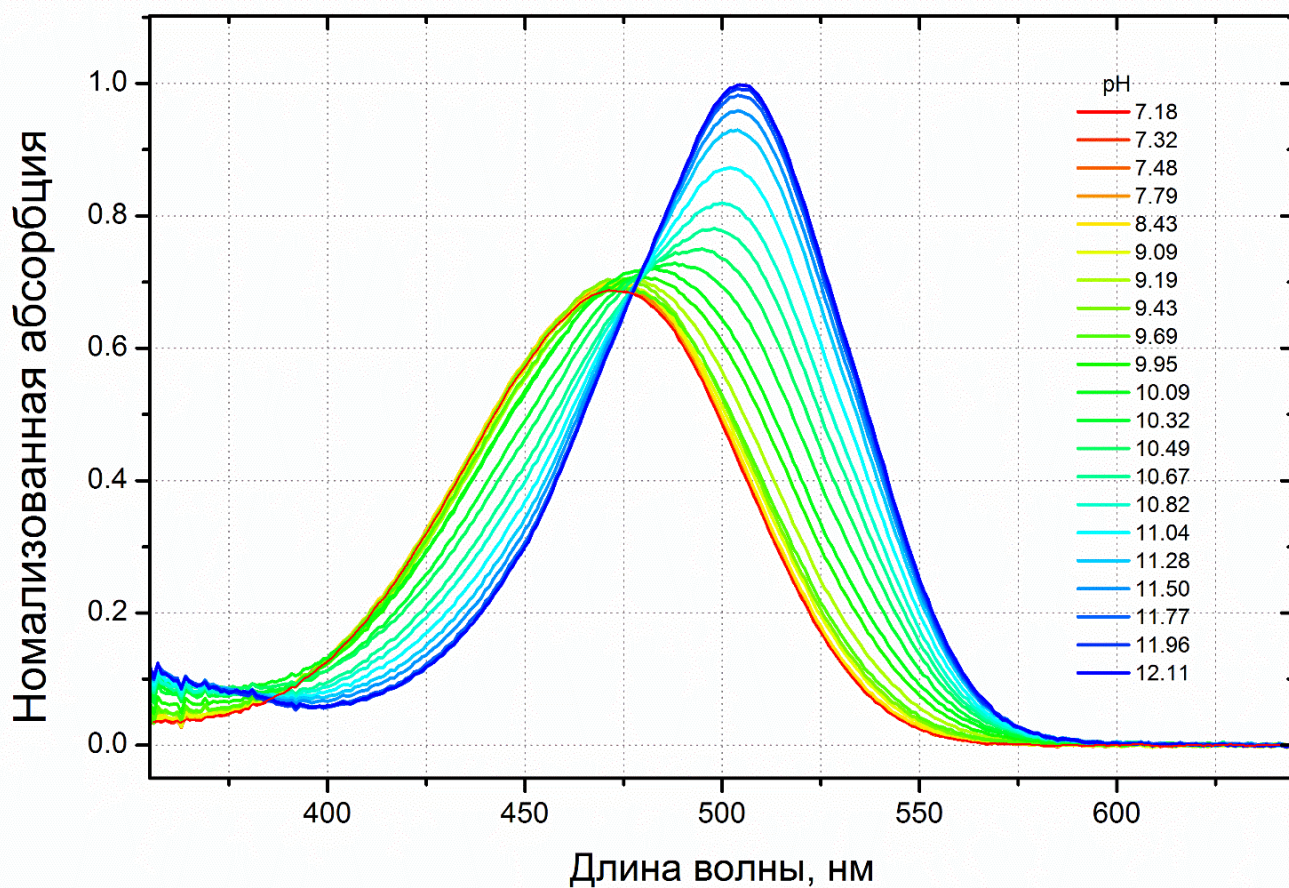


Рисунок 33. Спектры поглощения модельного соединения при разных значениях pH (верхняя панель) и кривые фотометрического титрования раствора Z-формы модельного соединения при 450 (синяя линия) и 505 (красная линия) нм (нижняя панель).

Положения максимумов абсорбции депротонированной формы модельного хромофора в различных органических растворителях, хорошо моделирующих белковое окружение (диметилформамид, ацетонитрил), оказались заметно сдвинутыми в длинноволновую область спектра относительно FR66W-Orange. Это дополнительно подтвердило наш вывод о том, что ни в одной из групп белков хромофор не находится в ионизированном состоянии.

К нашему большому удивлению мы также не обнаружили никакой разницы между спектрами поглощения Z- и E-форм модельного соединения: в полярных растворителях, которые лучше всего моделируют условия внутри белковой глобулы, спектры обоих изомеров были схожи с таковыми для FR66W-Yellow (рисунок 34).

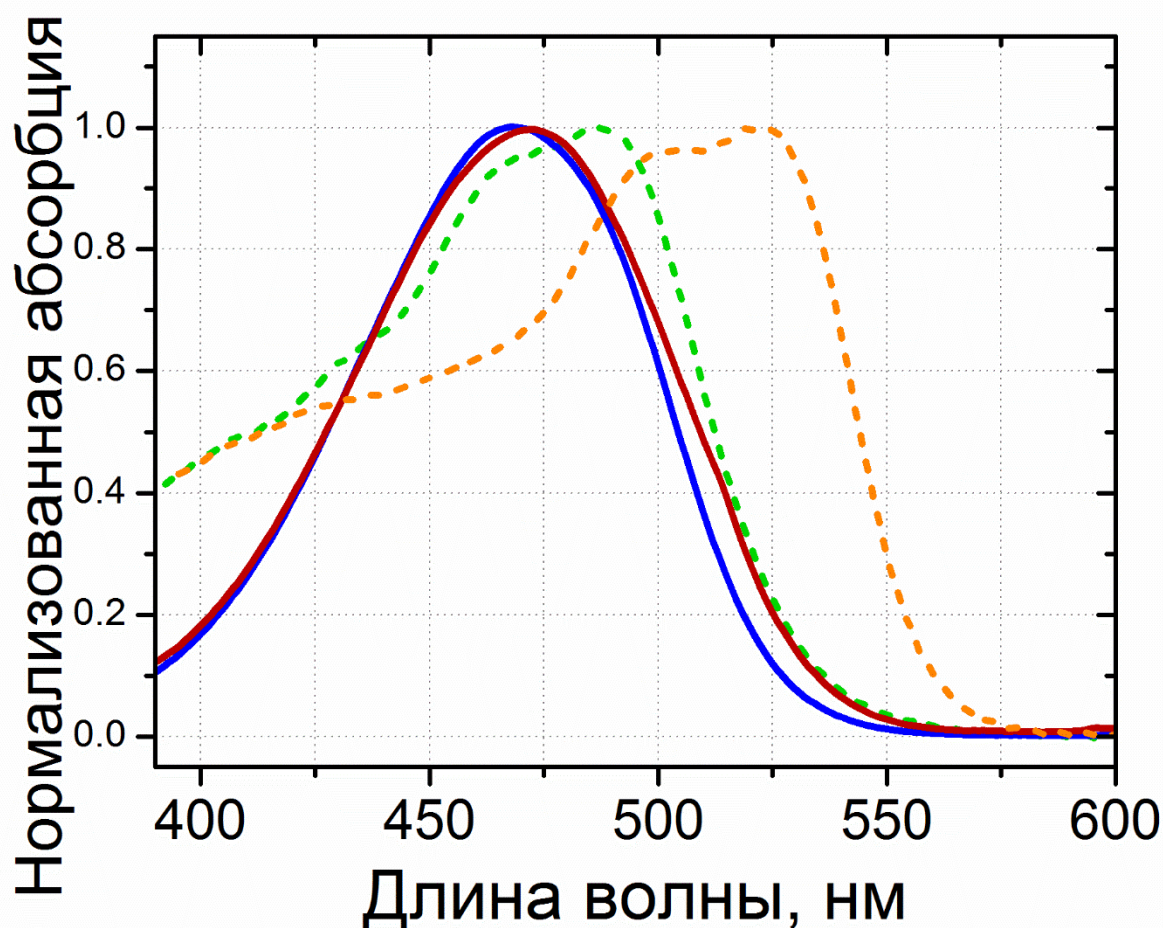


Рисунок 34. Спектры поглощения Z-формы (сплошная красная линия) и обогащенной E-формой смеси (сплошная синяя линия) модельного соединения в сравнении со спектрами FR66W-Yellow (зелёная пунктирная линия) и FR66W-Orange (оранжевая пунктирная линия).

Некоторые сигналы на ЯМР спектрах модельного соединения в Z-форме были уширены. Это было особенно заметно на ^{13}C ЯМР спектрах, где для разрешения всех сигналов пришлось нагревать образец (рисунок 32Б). Такое уширение может свидетельствовать о наличии некоторого динамического процесса в системе, например, вращения индолильной части вокруг одинарной С-С связи (рисунок 26, Z-цис и Z-транс).

Для того, чтобы проверить это предположение, нашим коллегой из университета г. Лунд (Швеция) Романом Гриценко были проведены теоретические расчёты термодинамических свойств исследуемой системы. Согласно расчётам в водном растворе в основном присутствуют три изомера: Z-транс (доминантная форма, 76.8%), Z-цис (менее стабильный, 19.6%) и E-транс (3.6%) (рисунок 35). При этом предсказанная энергия перехода между Z-транс и Z-цис формами довольно низкая, и этот процесс может спонтанно протекать при комнатной температуре.

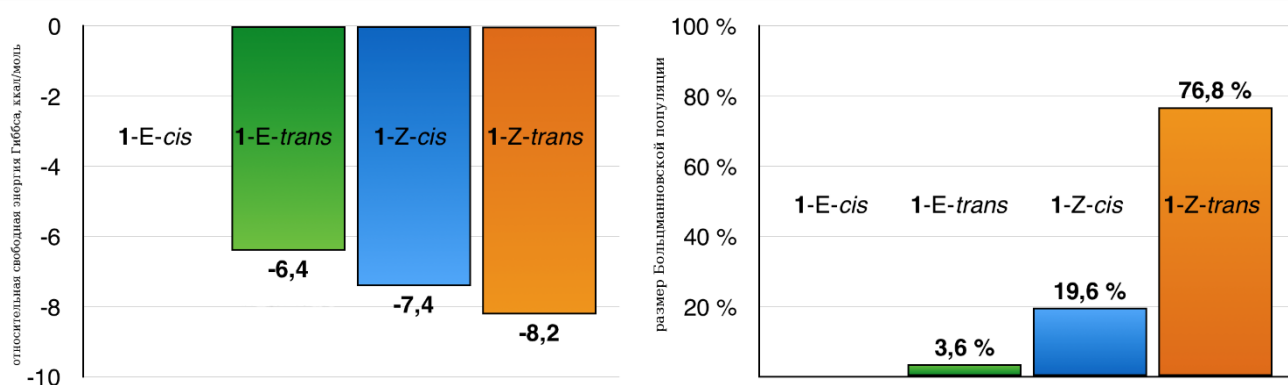


Рисунок 35. Рассчитанные значения свободной энергии Гиббса (относительно E-цис формы, левая панель) и соответствующие размеры Больцмановских популяций (правая панель) для изомеров в исследуемой системе.

Полученные результаты свидетельствуют о том, что ранее измеренные для Z-формы спектры представляют собой свойства смеси различных конформеров и не могут быть использованы для сравнения со спектрами белков, где хромофор зафиксирован строго в одном положении.

4.1.8. Компьютерное моделирование

Так как невозможно получить раствор, содержащий только один из изомеров модельного соединения, чтобы проверить нашу гипотезу о наличии разных изомеров хромофора в белках с жёлтой и оранжевой эмиссией флуоресценции, мы решили прибегнуть к компьютерному моделированию.

Существует несколько подходов к изучению данного вопроса с помощью компьютерных расчётов. Можно моделировать структуру белка и проверять, какой из вариантов хромофора

энергетически и геометрически совместим с данной аминокислотной последовательностью. К сожалению, использование традиционных протоколов моделирования белков в нашей ситуации невозможно: химические модификации, происходящие с тремя аминокислотами во время созревания хромофора, делают их недоступными для корректной обработки с помощью стандартных библиотек ротамеров аминокислотных остатков, используемых большинством программ. И хотя возможность моделирования структур флуоресцентных белков при помощи программы Rosetta за счёт включения хромофора в виде неканонической аминокислоты нами была ранее показана на примере расчёта стабильности мутантов GFP (рисунок 36), в случае кетонного хромофора, образование которого сопровождается разрывом белковой цепи, алгоритм требует ещё больше изменений (включение модифицированных концевых неканонических аминокислот). Подобная адаптация является отдельной проблемой, которая находится за пределами данной работы. Поэтому мы обратились к другому методу — расчётам спектральных свойств химических веществ.

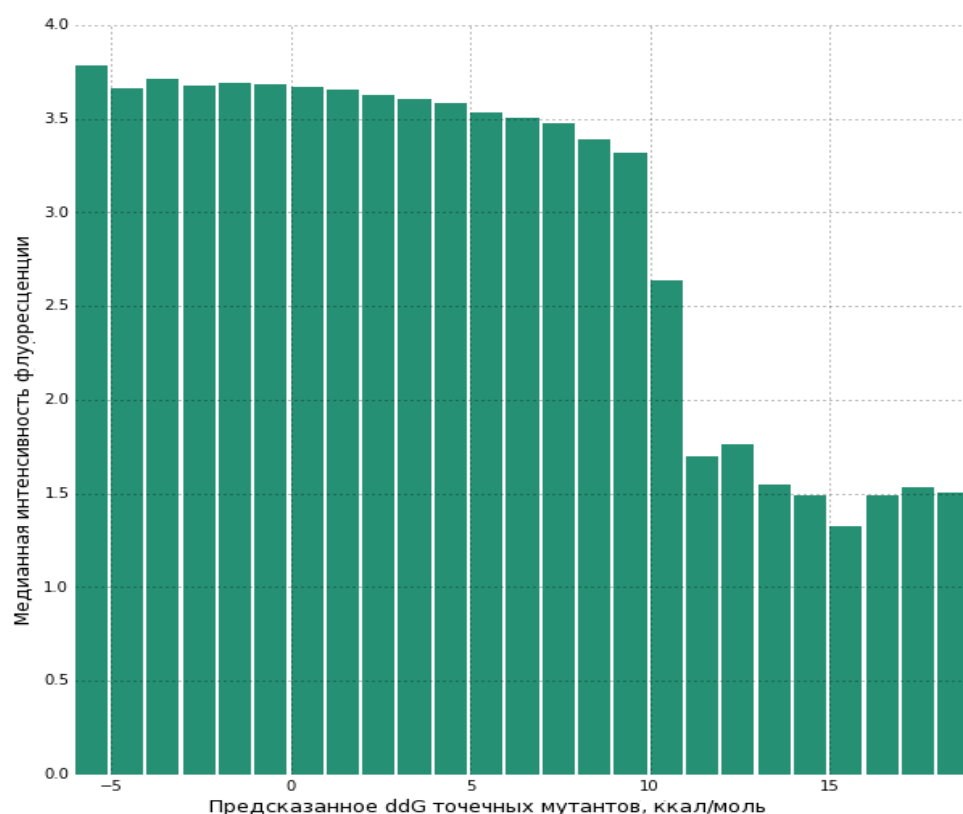


Рисунок 36. Моделирование связи фенотипов и генотипов флуоресцентного белка при помощи программы Rosetta. Зависимость медианной интенсивности флуоресценции точечных мутантов GFP от предсказанного изменения энергии этих мутантов относительно белка дикого типа

Романом Гриценко были проведены расчёты зависимости максимумов абсорбции хромофора от его геометрии. Они показали, что максимум спектра поглощения Z-транс изомера значительно сдвинут в длинноволновую область по сравнению с Z-цис и E-транс формами

(рисунок 37А). При этом теоретически рассчитанный спектр смеси всех изомеров (сумма спектров с учётом представленности изомера в смеси) хорошо согласовывался с экспериментально полученным спектром поглощения модельного соединения (рисунок 37Б).

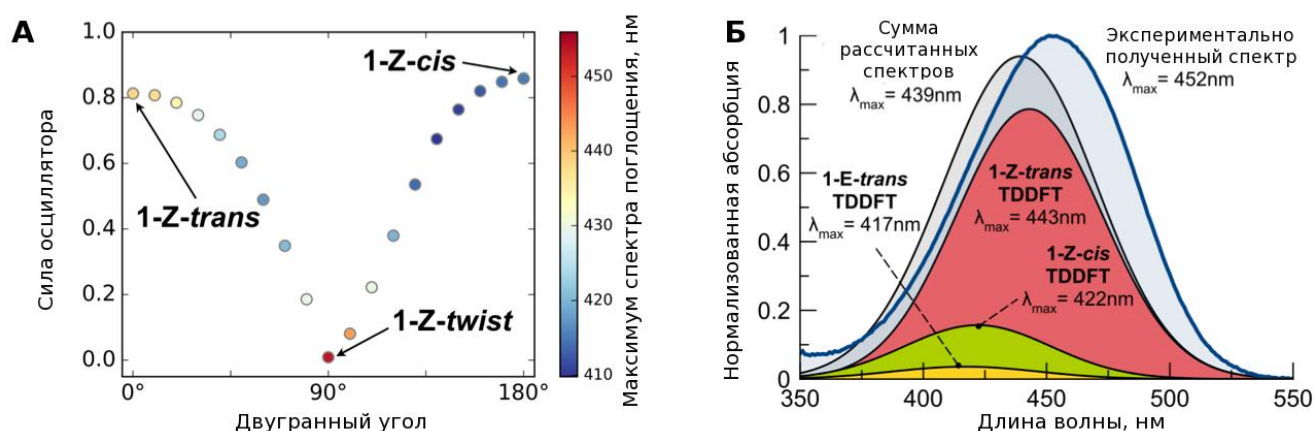


Рисунок 37. Влияние геометрии модельного соединения на его спектральные свойства. **А.** зависимость длины волны максимума спектра поглощения и силы осциллятора от величины двугранного угла между двумя частями сопряжённой π -системы. **Б.** Сравнение теоретически рассчитанного спектра смеси всех изомеров с экспериментально полученным спектром поглощения модельного соединения в ацетонитриле.

Таким образом, на основе анализа свойств белков FR66W-Yellow и FR66W-Orange, исследования модельного соединения хромофора и компьютерных расчётов мы установили, что оба белка содержат нейтральный кетонный триптофан-содержащий хромофор. В FR66W-Orange он находится в Z-транс конформации, в то время как FR66W-Yellow может содержать либо Z-цис, либо E-транс форму, что ещё предстоит установить.

4.2. Разработка и изучение флуоресцентных меток на основе флуороген-активирующих белков

Системы мечения на основе флуороген-активирующих белков, способные объединить достоинства генетически-кодируемых меток и синтетических красителей, представляются на сегодняшний день одной из самых перспективных групп флуоресцентных меток. Они состоят из двух равнозначных компонентов — белка и флуорогена, совместно определяющих основные параметры, характеризующие спектральные свойства системы флуоресцентного мечения (положения максимумов абсорбции и эмиссии, коэффициенты экстинкции, квантовый выход флуоресценции, фотостабильность и т.д.), а также ее совместимость с живыми объектами (отсутствие токсичности, возможность изучать процессы *in vivo*, степень воздействия на исследуемый белок и т.д.).

4.2.1. Выбор флуорогенов

При выборе флуорогена были выдвинуты следующие требования: используемое вещество должно быть нетоксичным, легко и быстро проникать через неповрежденную мембрану клетки, обладать небольшим размером, флуоресценцией в видимой или более длинноволновой областях спектра и высокой флуорогенностью.

Перечисленным свойствам обладают хромофоры флуоресцентных белков. Квантовый выход флуоресценции этих веществ в свободном состоянии крайне низкий ($<0.1\%$) [194], в то время как при стабилизации, например в составе белка или РНК-аптамера, он может превышать 90% [195,196]. Такая флуорогенность объясняется наличием разного рода безызлучательных процессов, гасящих флуоресценцию свободного хромофора: подвижностью бензилиденовой части молекулы [197] и возможностью формирования нефлуоресцентного возбужденного состояния с переносом заряда (ICT state) [198], которые могут быть ограничены за счёт окружения.

Разнообразие спектральных свойств флуоресцентных белков достаточно велико [175], однако оно основывается на различных модификациях (ионизации, изомеризации, ковалентных и нековалентных взаимодействий с окружающими аминокислотными остатками и т.д.) довольно ограниченного количества структур хромофоров (рисунок 38), не все из которых при этом стабильны в свободном состоянии. Впрочем, разнообразие синтетических соединений не ограничено двадцатью каноническими аминокислотами и доступными с их помощью автокаталитическими реакциями.

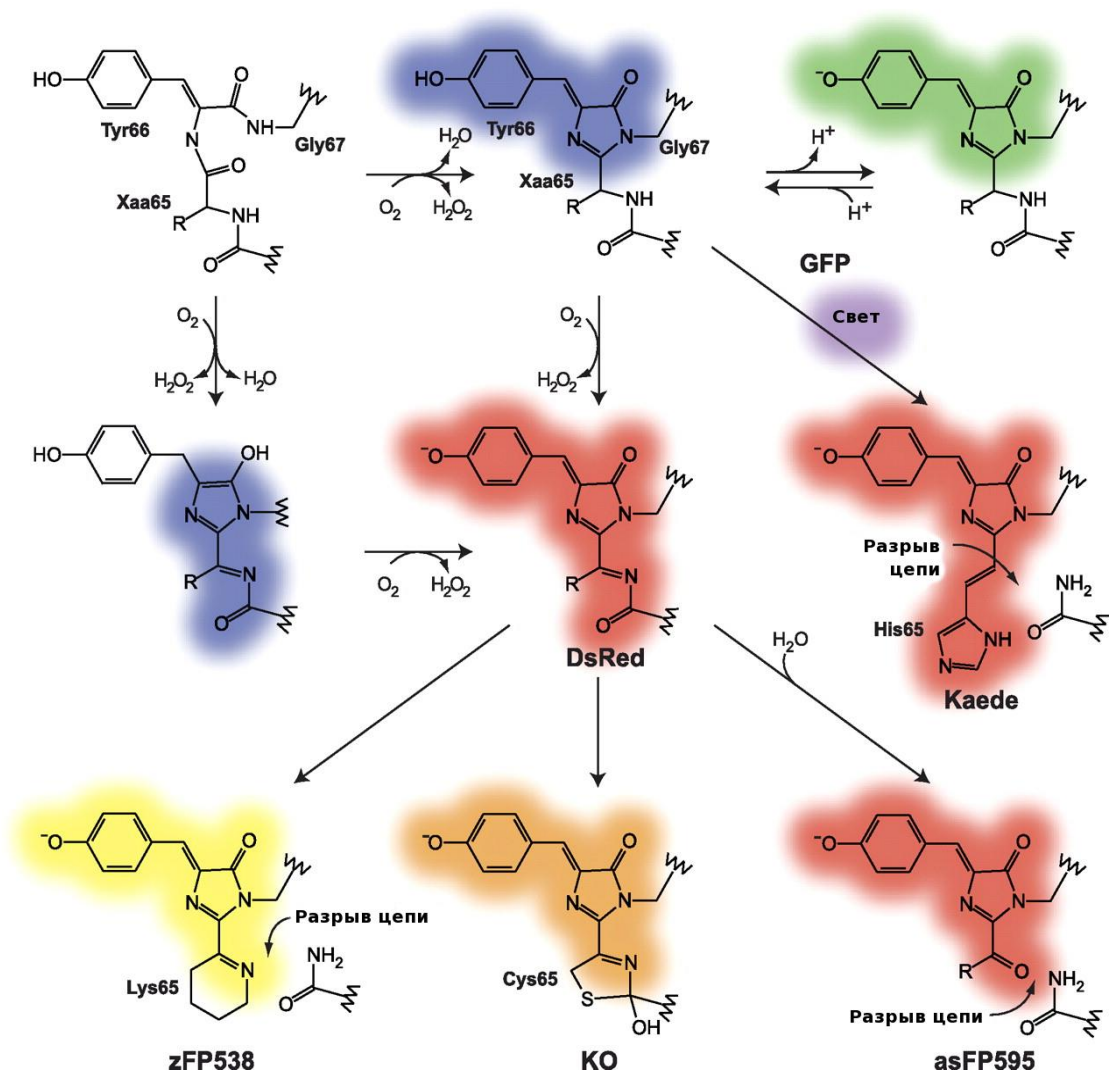


Рисунок 38. Структуры хромофоров природных флуоресцентных белков [175].

Сотрудниками группы синтеза природных соединений ИБХ РАН были синтезированы библиотеки производных стабильных в водных растворах хромофоров флуоресцентных белков GFP (таблица 2) и Kaede (таблица 3) [199], а также их конформационно фиксированных аналогов (таблица 4) [197,200], не способных к изомеризации бензилиденового фрагмента, которые и были использованы в данной работе.

Чтобы не тестировать заведомо непригодные для использования в намеченных целях вещества, были проанализированы различные физико-химические свойства данных соединений.

Сначала были исключены из рассмотрения флуорогены со слишком малой растворимостью в воде (например, M1035, GA08 и т.д.) из-за невозможности их использования *in vivo*.

Таблица 2. Библиотека GFP-подобных флуорогенов.

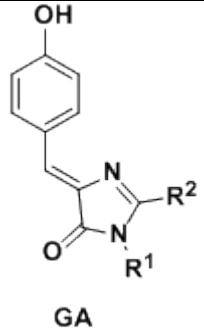
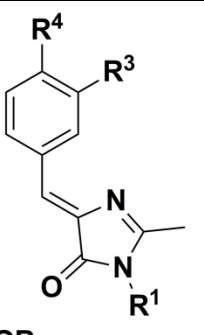
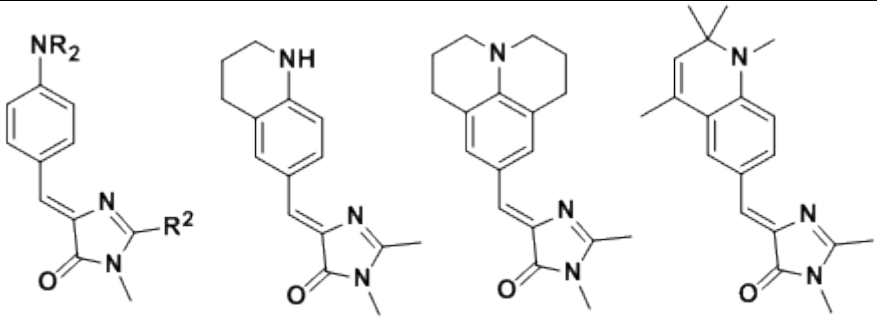
 <p>GA</p>			 <p>GB</p>			 <p>GC GC08 GC09 GC10</p>			
Название	R ²	R ¹	Название	R ¹	R ³	R ⁴	Название	R ²	R
GA01	H	CH ₃	GB01	CH ₃	H	Br	GC01	CH ₃	Et
GA02	CH ₃	H	GB02	CH ₃	H	I	GC02	Et	Et
GA03	CH ₃	CH ₃	GB03	NO ₂ h	H	Br	GC04	CH ₃	CH ₃
GA04	CH ₃	Et	GB04	CH ₃	NO ₂	Br	GC05	CH ₃	H
GA05	CH ₃	Pr	GB05	CH ₃	OH	H	GC06	CH ₃	n-Bu
GA06	CH ₃	iPr	GB06	CH ₃	CH ₂ CH=CH ₂	OH	GC08	CH ₃	См. выше
GA07	CH ₃	tBu	GB07	CH ₃	H	OCH ₃	GC09	CH ₃	См. выше
GA08	CH ₃	Ph	GB08	CH ₃	H	NO ₂	GC10	CH ₃	См. выше
GA10	CH ₃	CF ₃	GB09	CH ₂ COOEt	H	Br			
GA17	OC ₃	Et	GB10	CH ₃	H	OAc			
GA18	CH ₃	iBu	GB11	CH ₃	NO ₂	H			

Таблица 3. Библиотека Каеде-подобных флуорогенов.

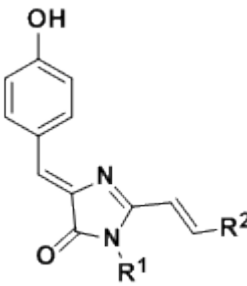
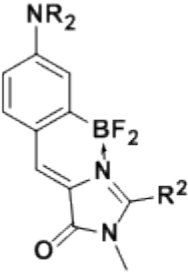
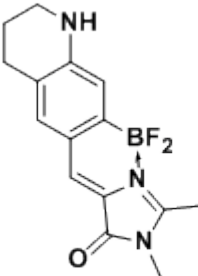
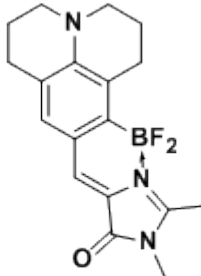
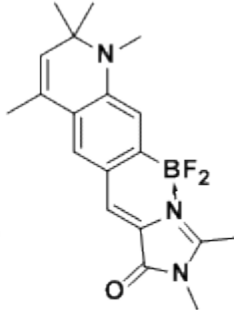
	Название	R ¹	R ²
	A001	CH ₃	Ph
	A002	CH ₃	4-OH-C ₆ H ₄
	A003	CH ₃	4-NO ₂ -C ₆ H ₄
	A004	CH ₃	3-индолил
	A005	CH ₃	C ₆ H ₄ -CH=CH
	A006	CH ₃	CO(NH ₂)
	A007	CH ₃	4-Br-C ₆ H ₄
	A008	CH ₃	CH ₃
	A009	CH ₃	2-пиридил
	A010	CH ₃	3-Br-4-OH-C ₆ H ₃
	A011	CH ₃	3-NO ₂ -4-OH-C ₆ H ₃
	A012	CH ₃	3,5-Br-4-OH-C ₆ H ₂
	A012H	H	3,5-Br-4-OH-C ₆ H ₂
	A013	CH ₃	4-OCH ₃ -C ₆ H ₄
	A014	CH ₃	3-NO ₂ -C ₆ H ₄
	A015	CH ₃	2-OH-C ₆ H ₄
	A016	CH ₃	4-COOH-C ₆ H ₄
	A017	CH ₃	9-антраценил
	A018	CH ₃	2-F-C ₆ H ₄
	A019	CH ₃	2-тиофенил
	A020	CH ₃	2-OCH ₃ -C ₆ H ₄
	A022	CH ₃	4-N(CH ₃) ₂ -C ₆ H ₄
	A023	CH ₃	3,4-OCH ₃ -C ₆ H ₃
	A024	CH ₃	4-CH ₃ -C ₆ H ₄
	A025	CH ₃	3-CH ₃ -C ₆ H ₄
	A026	CH ₃	2-CH ₃ -C ₆ H ₄
	A027	CH ₃	3,4-CH ₃ -C ₆ H ₃
	A028	CH ₃	4-Cl-C ₆ H ₄
	A029	CH ₃	3-F-C ₆ H ₄
	A030	CH ₃	4-NEt ₂ -C ₆ H ₄
	A031	CH ₃	4-NBu ₂ -C ₆ H ₄

Таблица 4. Библиотека борированных GFP-подобных флуорогенов.

			
ABDI-BF ₂	M1167	M1035	M1169

Название	R ²	R
M739	CH ₃	Et
M774	Et	Et
M766	CH ₃	CH ₃
M816	CH ₃	H
M1078	CH ₃	n-Bu
M1035	CH ₃	См. выше
M1169	CH ₃	См. выше
M1167	CH ₃	См. выше

Анализ спектральных свойств оставшихся элементов библиотеки привел к отказу от флуорогенов, содержащих электроноакцепторные заместители в бензилиденовой части (например GB01, GB02, GB04, GB11 и т.д.), так как они способствуют смещению спектров абсорбции и, соответственно, максимумов возбуждения и эмиссии флуоресценции в синюю область спектра, что мешает эффективному применению данных соединений для мечения живых систем.

Флуорогенный потенциал молекул оценивали, измеряя квантовый выход флуоресценции в различных растворителях и при сильном охлаждении. Соединения с изначально высоким квантовым выходом флуоресценции в водных растворах или продемонстрировавшие незначительные изменения в протестированных условиях также были изъяты из библиотеки.

Всё это позволило существенно сократить количество тестируемых *in silico* и *in vitro* флуорогенов.

4.2.2. Поиск флуороген-активирующих белков

В первую очередь было решено проверить, нет ли среди белков с известными структурами тех, что способны связывать выбранные синтетические флуорогены.

4.2.2.1. *In silico* скрининг

Так как нашей целью являлось создание новой метки для флуоресцентного мечения белков в клетках млекопитающих, из всех доступных в базе данных PDB структур были выбраны лишь белки *E.coli* (чтобы снизить вероятность функциональности и нахождения белковых партнеров и/или лигандов при экспрессии в гетерологичной системе) длиной менее 500 аминокислотных остатков. Для того чтобы результаты докинга были более достоверными, рассматривались лишь структуры, снятые с разрешением не хуже 3.0 Å. Всего этим критериям удовлетворяло чуть более 3 000 белков.

Докинг проводился в автоматизированном режиме с помощью программы AutoDoc Vina [119], при этом всегда использовалась лишь первая цепь структуры, освобожденная от всех лигандов.

Сначала все структуры были проверены на возможность связывания хромофора GFP (GA03). Лучшие 500 белков были затем использованы для анализа взаимодействия с различными Kaede-подобными флуорогенами (Kaede-подобные флуорогены выглядят более перспективными в качестве потенциальной *in vivo* метки так как их поглощение, а также спектры возбуждения и эмиссии флуоресценции смещены в красную область).

После дополнительного анализа (из рассмотрения были исключены, например, белки, имеющие ковалентно-связанный лиганд, олигомерную структуру, известные белки-партнеры, связывающиеся с высокой аффинностью, нежелательную каталитическую функцию и т.д.) лучших моделей (таблица 5), полученных в результате двух раундов докинга, было принято решение клонировать и экспрессировать гены четырёх белков (здесь и далее названия соответствуют четырёхбуквенным кодам в базе данных PDB): 3HO2 (β -кетоацил-АЦФ-синтаза II) [201], 1DOS (фруктозобисфосфат-альдолаза) [202], 2QRY (периплазматический тиамин-связывающий белок) [203] и 1PVS (3-метиладенин-ДНК-гликозилаза II) [204] (таблица 6).

Таблица 5. Лучшие 50 белков по результатам докинга хромофоров GFP и Kaede и их оценки (чем меньше число, тем лучше).

Докинг хромофора Kaede			Докинг хромофора GFP		
№	PDB ID	Оценка	№	PDB ID	Оценка
1	1DOS	-12.5	1	2GFX	-9.75
2	1GVH	-11	2	2QRY	-9.6
3	1PVS	-10.8	3	1DOS	-9.6
4	3FBR	-10.8	4	3HNZ	-9.55
5	1VB6	-10.7	5	1XDQ	-9.45
6	3ASV	-10.4	6	2GFY	-9.45
7	1TJ1	-10.2	7	3DNT	-9.2
8	1V9Z	-10.2	8	3FBR	-9.15
9	3NR0	-10.2	9	3HO2	-9.15
10	1TLZ	-10.2	10	3G11	-9.1
11	1TLY	-10.2	11	1GVH	-9.05
12	1TLW	-10.2	12	1B3N	-9.05
13	3IP0	-10.2	13	3I8P	-9.05
14	2R46	-10.2	14	2GFV	-9
15	1TJ2	-10.2	15	3EPS	-9
16	3OW7	-10.1	16	1HO5	-9
17	1RP7	-10.1	17	1KFY	-9
18	2FQ1	-10.1	18	2CGL	-9
19	3DNT	-10.1	19	3HZI	-8.95
20	1I8T	-10.1	20	1XDY	-8.95
21	2FZM	-10.1	21	2FDK	-8.95
22	1FFT	-10.1	22	2PUA	-8.9
23	2ANB	-10.1	23	3HO9	-8.9
24	2UDP	-10	24	3MR8	-8.9
25	2QCU	-10	25	3ASV	-8.85
26	1TIW	-10	26	2PUE	-8.85
27	2FZN	-10	27	2CGJ	-8.85
28	1TJ0	-10	28	1IL2	-8.8
29	1MPG	-10	29	1TJ2	-8.75
30	2CGL	-9.9	30	2WDR	-8.75
31	2R45	-9.9	31	1MWJ	-8.75
32	1L8A	-9.9	32	1FDI	-8.7
33	1MWJ	-9.9	33	3O7Q	-8.7
34	1TLC	-9.9	34	1PNS	-8.7
35	1W7K	-9.9	35	1K0G	-8.7
36	3DMQ	-9.9	36	3NQ8	-8.7
37	3Q2D	-9.8	37	2WS3	-8.7
38	3D4V	-9.8	38	2R4E	-8.65
39	3SEX	-9.8	39	2ANB	-8.65
40	3ITG	-9.8	40	1K87	-8.65
41	2OWO	-9.8	41	1NEN	-8.65
42	1NAI	-9.8	42	2WDV	-8.6
43	3CW7	-9.8	43	2FZN	-8.6
44	2R4J	-9.8	44	2HGP	-8.6
45	1LQA	-9.8	45	1HPU	-8.6
46	1W78	-9.8	46	2Q29	-8.6
47	2G28	-9.8	47	2QOW	-8.6
48	2VET	-9.8	48	1TLC	-8.6
49	1K87	-9.8	49	3OAS	-8.6
50	3CVS	-9.8	50	1SPA	-8.6

Таблица 6. Характеристики выбранных для клонирования белков.

PDB ID	место в докинге хромофора GFP	оценка докинга хромофора GFP	место в докинге хромофора Kaede	оценка докинга хромофора Kaede
1PVS	87	-8.35	3	-10.8
1DOS	3	-9.6	1	-12.5
2QRY	2	-9.6	419	-8.8
3HO2	9	-9.15	247	-9.0

4.2.2.2. Клонирование генов белков 3HO2, 1DOS, 1PVS и 2QRY

Гены, кодирующие белки 3HO2, 1DOS, 1PVS и 2QRY, были амплифицированы с использованием *E. coli* штамма XL1 Blue в качестве матрицы и клонированы в экспрессионный вектор pBAD при помощи самособирающегося клонирования. Праймеры, использованные для клонирования, приведены в таблице 7.

4.2.2.3. Характеристика найденных белков

Предварительную полуколичественную оценку взаимодействия выбранных белков с флуорогенами проводили следующим образом: небольшое количество металл-аффинной смолы TALON с иммобилизованным на ней выделенным исследуемым белком помещали в лунку 96-луночного планшета с одним из исследуемых флуорогенов в концентрации 1 мкМ. Детекцию осуществляли в полуавтоматическом режиме при помощи флуоресцентного микроскопа Keyence BZ-9000 используя фильтры DAPI-BP (377/50, 447/60), GFP-BP (472.5/30, 520/35), TexasRed (562/40, 624/40) и Cy5 (628/40, 692/40). Съемку каждого из исследуемых флуорогенов проводили в двух каналах, наиболее близко расположенных к максимуму спектра флуоресценции (рисунок 39).

Флуоресцентный сигнал смолы в каждой из лунок сравнивали с сигналом контроля (чистая смола в растворе флуорогена) для исключения ложноположительных результатов за счёт связывания флуорогена с носителем и заносили в итоговую таблицу, которая использовалась для выбора потенциальных пар белок-флуороген для дальнейших исследований. Пары белок-флуороген, демонстрировавшие более яркий флуоресцентный сигнал в сравнении с контролем, были затем изучены более детально.

Таблица 7 Список праймеров, использованных для клонирования генов белков 3HO2, 1DOS, 1PVS и 2QRY

1DOS внешний прямой	GGCCCGACGATACAGGACAAGAGACATGTCTAAGATTTTTGATTTTCGTAAAACCTGGCG
1DOS внешний обратный	AGACCCGCAGAGCGGGCCTTGAGATAAGCAGAAAGGAATATCTTACAGAACG
1DOS внутренний прямой “длинный”	AGCTCGAGATCTATGTCTAAGATTTTTGATTTTCGTAAAACCTGGCGTAATCACTGGTGATGACGTACAG
1DOS внутренний прямой “короткий”	ATGTCTAAGATTTTTGATTTTCGTAAAACCTGGCGTAATCACTGGTGATGACGTACAG
1DOS внутренний обратный “длинный”	CAGCCAAGCTTTTACAGAACGTCGATCGCGTTCAGTTCCTGGAATGCTTTCTCCAGACGAGCG
1DOS внутренний обратный “короткий”	TTACAGAACGTCGATCGCGTTCAGTTCCTGGAATGCTTTCTCCAGACGAGCG
2QRY прямой “длинный”	AGCTCGAGATCTATGTCTGCCCCTGCTGTTGCTGTGACAGCGCCCCG
2QRY прямой “короткий”	ATGTCTGCCCCTGCTGTTGCTGTGACAGCGCCCCG
2QRY обратный “длинный”	CAGCCAAGCTTTTAACGGCTGACGGCGCGTTGCCATTCGC
2QRY обратный “короткий”	TTAACGGCTGACGGCGCGTTGCCATTCGC
1PVS обратный “короткий”	TCATGCTTCGTCTGGTTGCCAGCCTTCCGTATACCAG
1PVS обратный “длинный”	CAGCCAAGCTTTCATGCTTCGTCTGGTTGCCAGCCTTCCGTATACCAG
1PVS прямой “короткий”	ATGTATACCCTGAACTGGCAGCCGCGTATGACTGGTCGTGG
1PVS прямой “длинный”	AGCTCGAGATCTATGTATACCCTGAACTGGCAGCCGCGTATGACTGGTCGTGG
3HO2 внешний прямой	GTCCCACTAGAATCATTTTTTCCCTCCCTGGAGGACAAACGTGTCTAAGCGTCGTG
3HO2 внешний обратный	GGCCCGCAAGCGGACCTTTTATAAGGGTGGAAAATGACAACCTAGATCTTTTAAAG
3HO2 внутренний прямой “длинный”	AGCTCGAGATCTGTGTCTAAGCGTCGTGTAGTTGTGACCGGACTGGGCATGTTGTC
3HO2 внутренний прямой “короткий”	GTGTCTAAGCGTCGTGTAGTTGTGACCGGACTGGGCATGTTGTC
3HO2 внутренний обратный “длинный”	CAGCCAAGCTTTTAGATCTTTTAAAGATCAAAGAACCATTAGTGCCACCGAAGCCG
3HO2 внутренний обратный “короткий”	TTAGATCTTTTAAAGATCAAAGAACCATTAGTGCCACCGAAGCCG
pBAD-His прямой “длинный”	AAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCC
pBAD-His прямой “короткий”	TTTTGGCGGATGAGAGAAGATTTTCAGCC
pBAD-His обратный “длинный”	AGATCTCGAGCTATCCTTATCGTCATCGTCGTACAGATCC
pBAD-His обратный “короткий”	ATCCTTATCGTCATCGTCGTACAGATCC

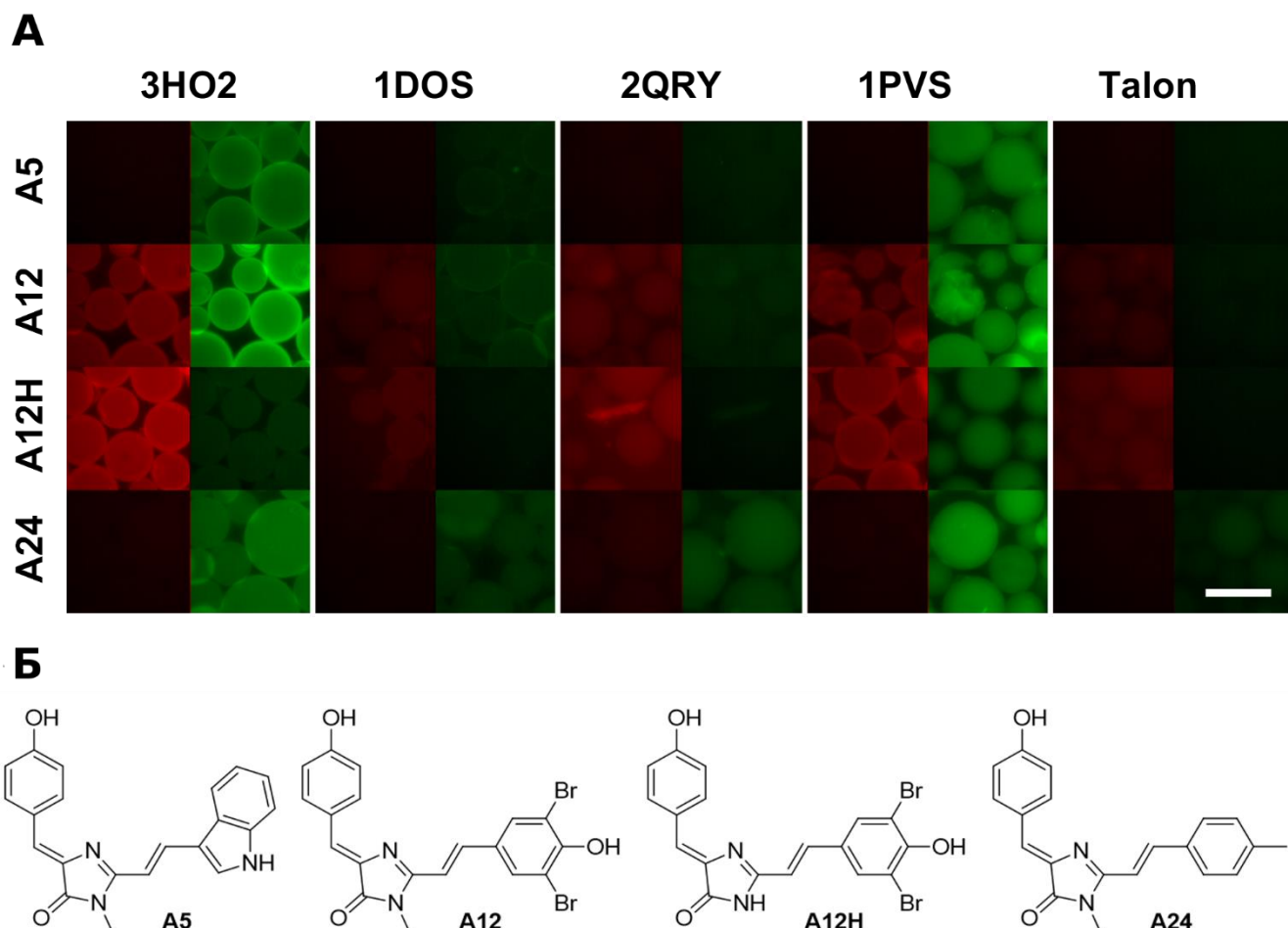


Рисунок 39. Пример *in vitro* анализа взаимодействия белков, иммобилизованных на металл-аффинной смоле, с флуорогенами при помощи флуоресцентной микроскопии. **А.** Микрофотографии смеси металл-аффинной смолы TALON с иммобилизованным на ней выделенным исследуемым белком (название столбца) с одним из исследуемых флуорогенов (название строки). Правая колонка — съёмка с использованием набора фильтров GFP-BP (472.5/30, 520/35), Левая колонка — съёмка с использованием набора фильтров TexasRed (562/40, 624/40). Последний столбец — негативный контроль — чистая металл-аффинная смола. Масштабная линейка 100 мкм. **Б.** Структуры флуорогенов, представленных на панели А.

Несмотря на то, что все выбранные белки имели высокий рейтинг в скрининге на связывание хромофора GFP, ни один из них не продемонстрировал возрастания флуоресценции при добавлении GFP-подобных флуорогенов. Этот факт не отрицает того, что эти белки, возможно, связывают данные соединения, так как не любое даже высокоаффинное взаимодействие обязано приводить к возрастанию флуоресценции (например, флуороген может быть зафиксирован в непланарной конформации, не способной к флуоресценции). Однако такой вид взаимодействия не соответствовал целям нашей работы и дальнейшее детальное исследование подобных пар не осуществлялось.

Из семи выбранных для анализа Kaede-подобных флуорогенов добавление четырёх (A5, A12, A12H и A24) вызвало заметное увеличение флуоресцентного сигнала. Их свойства, а также их взаимодействие с белками было изучено более подробно (таблица 8).

Таблица 8. Свойства свободных флуорогенов и комплексов белок-флуороген в водных растворах (PBS, pH 7.4).

Вещество	λ_{exc} (нм) free/ +1PVS/ +3HO2	λ_{em} (нм) free/ +1PVS/ +3HO2	E ($\text{M}^{-1}\cdot\text{cm}^{-1}$)	Растворимость, мкМ			FQY (%)	λ_{Abs} (нм)	λ_{Abs} аниона (нм)
				PBS	PBS +1% EtOH	PBS +5% EtOH			
A05	478 475 480	568 563 565	31000	29.6	n.d.	n.d.	0.40	468	525
A12	520 503 520	608 601 599	39500	1.1	10.5	13.8	0.05	481	n.d.*
A12H	480 481 485	572 577 576	35000	5.9	29.9	n.d.	0.03	465	n.d.*
A24	440 438 433	535 512 506	34500	8.4	13.5	31.0	0.15	436	511

*- возможны несколько анионных форм

λ_{ex} — максимум спектра возбуждения флуоресценции

λ_{em} — максимум спектра эмиссии флуоресценции

λ_{Abs} — максимум спектра поглощения

FQY (fluorescence quantum yield) — квантовый выход флуоресценции

E — коэффициент экстинкции, n.d. — не определяли

Для каждой пары белок-флуороген было проведено титрование для определения константы связывания, а в насыщающей концентрации были сравнены спектры эмиссии флуоресценции свободного флуорогена и его комплекса с белком (рисунок 40).

Многие из исследованных пар белок-флуороген демонстрировали микромолярные и даже субмикромолярные константы диссоциации. При этом мы с удивлением обнаружили, что в большинстве случаев возрастание флуоресценции при связывании флуорогена белком было незначительным и не превышало двух раз. Наблюдаемое несоответствие между результатами микроскопического скрининга и измеренными *in vitro* значениями можно объяснить тем, что флуорогены сами по себе всё же обладают ненулевым квантовым выходом флуоресценции. Поэтому их накопление на поверхности носителя за счёт высокоаффинного связывания с иммобилизованным белком может приводить к видимому значительному возрастанию интенсивности флуоресценции, не связанному с изменениями спектральных свойств флуорогенов.

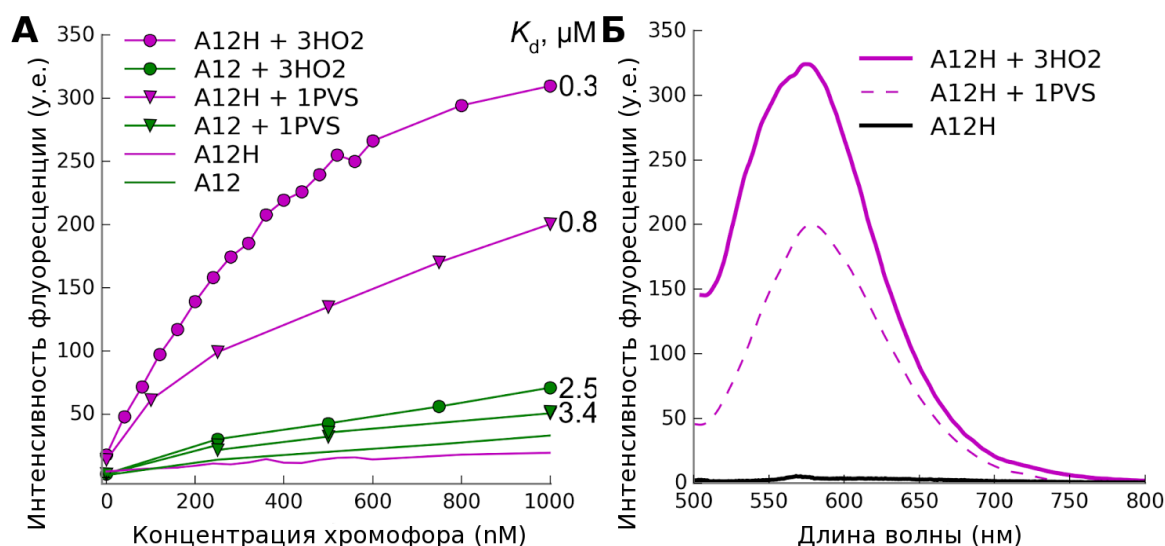


Рисунок 40. Пример *in vitro* анализа взаимодействия белков с флуорогенами. **А.** Кривые титрования белков 3HO2 и 1PVS флуорогенами A12 (зелёные линии) и A12H (малиновые линии). Маркеры, показывающие измеренные значения, соединены для наглядности линиями. Справа от кривой указана рассчитанная константа диссоциации комплекса. **Б.** Спектры эмиссии флуоресценции раствора чистого флуорогена (черная линия) и комплексов флуорогена с белками.

Тем не менее, истинно флуорогенное поведение (приводящее к возрастанию квантового выхода флуоресценции и/или коэффициента экстинкции комплекса в сравнении с соответствующими показателями свободного флуорогена) было также обнаружено. Так, взаимодействие флуорогена A12H с белками 3HO2 и 1PVS характеризовалось увеличением яркости флуоресценции более чем на два порядка.

4.2.2.4. Анализ причин различий в флуорогенности

Чтобы попытаться понять причину различий в флуорогенности комплексов, мы провели анализ взаимодействия некоторых изученных *in vitro* пар белок-флуороген с помощью протокола докинга низкомолекулярных лигандов программы Rosetta. Хотя в целом как общая энергия взаимодействия одного белка со всеми лигандами, так и природа, количество и энергия лиганд-стабилизирующих взаимодействий (таблица 9) для всех флуорогенов были схожими, в случае комплексов, проявивших высокую флуорогенность, докинг сходился к единственному положению флуорогена в лиганд-связывающем кармане (рисунок 41Б), в то время как для других даже среди лучших 50 моделей (что составляет всего 1% от всех посчитанных и проанализированных для каждого взаимодействия моделей) наблюдались различные альтернативные варианты (рисунок 41А,В,Г). Такое разнообразие может означать недостаточную фиксацию флуорогена в белковой глобуле, приводящую к возможности безызлучательной диссипации поглощенной энергии за счёт подвижности лиганда.

Таблица 9. Вклад аминокислотных остатков в стабилизацию флуорогена в белковой молекуле. Учитывались все взаимодействия с энергией взаимодействия менее -1 REU.

	3HO2			1PVS		
	остаток	тип взаимодействия	$\Delta\Delta G$, REU	остаток	тип взаимодействия	$\Delta\Delta G$, REU
A5	F399	π - π	-2.30	Y239	π - π	-1.48
	F397	π - π	-1.42	T219	...H	-1.05
	P271	...H	-1.08	W218	π - π	-2.93
				L125	HI	-1.64
A12	F399	π - π	-2.15	Y239	π - π	-2.10
	D264	...H	-1.12	W218	π - π	-2.40
				L125	HI	-1.14
A12H	F399	π - π	-2.19	Y239		-1.57
	D264	...H	-1.12	W218	π - π	-2.61
				L125	HI	-1.10
A24	F399	π - π	-2.33	Y239		-1.88
	T304	...H	-1.55	T219	...H	-1.11
	P271	...H	-1.18	W218	π - π	-2.46
				L125	HI	-1.41

HI - гидрофобное взаимодействие

π - π — стэкинг-взаимодействие

...H — водородная связь

REU — единица измерения энергии программы Rosetta

Описанный докинг в программе AutoDoc Vina несомненно выигрывает по скорости и поэтому, в отличие от программы Rosetta, пригоден для скрининга больших библиотек. Выигрыш в скорости чаще всего достигается за счёт упрощения протокола. Таким упрощением в нашем случае является полное исключение из рассмотрения подвижности белка, которая, однако, существует и в некоторых случаях может быть существенной и играть критическую роль во взаимодействии белка с лигандом. Это даёт некоторое количество ложноположительных результатов, хотя, на наш взгляд, и не умаляет достоинства метода. Последующий анализ лучших результатов при помощи более медленных, но более точных алгоритмов докинга может быть использован в качестве дополнительного критерия отбора для специфических задач.

Несмотря на то, что в некоторых парах белок-флуороген наблюдалось значительное возрастание яркости флуоресценции (более 170 раз для A12H-3HO2), итоговые квантовые выходы флуоресценции комплексов (несколько процентов) в сочетании с не очень высокими коэффициентами экстинкции ($30\,000$ – $40\,000\text{ M}^{-1}\text{cm}^{-1}$) и неустойчивостью флуорогена в клеточной культуре препятствовали успешному применению найденных белков в качестве флуоресцентной метки *in vivo*.

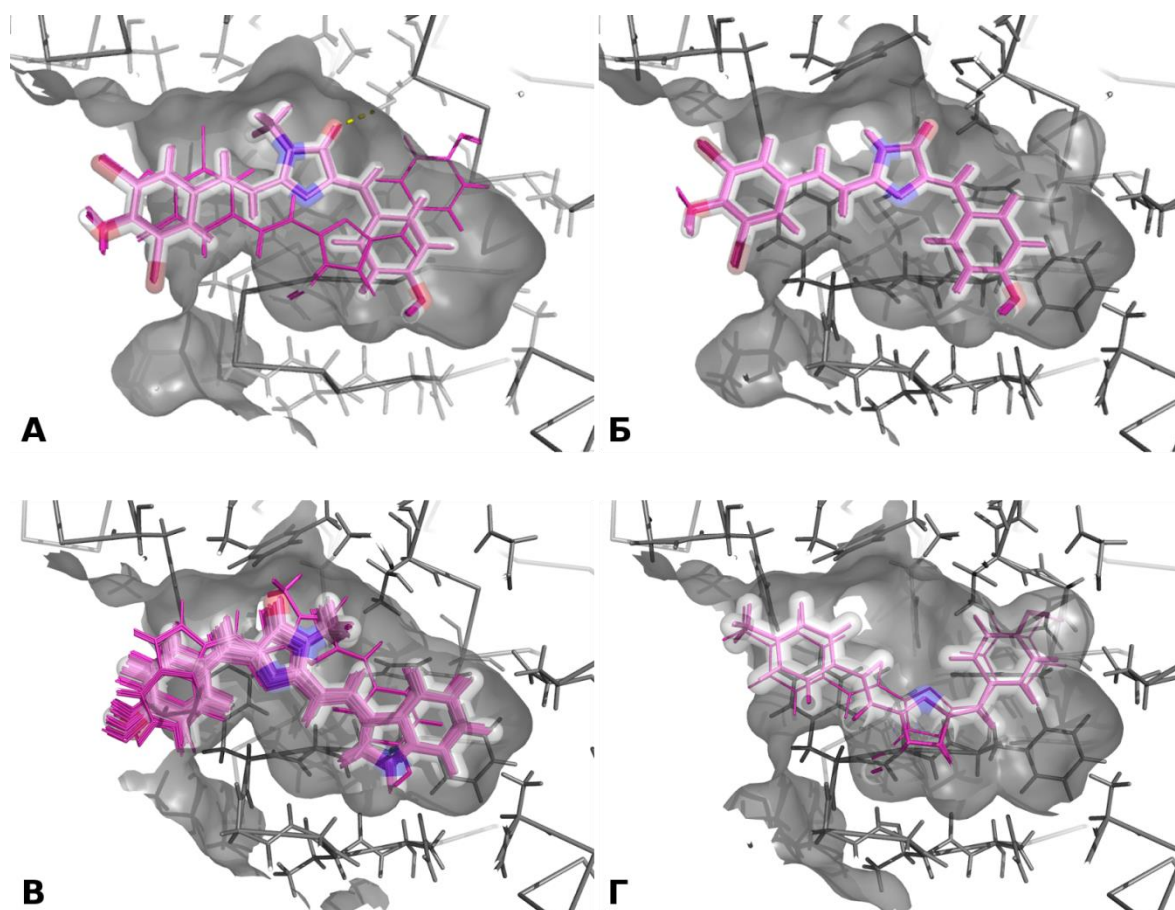


Рисунок 41. Результаты докинга флуорогенов А12 (А), А12Н (Б), А5 (В) и А24 (Г) с помощью программы Rosetta в белок ЗНО2. Малиновым цветом показано положение хромофора в 50 лучших моделях. Серым цветом обозначен лиганд-связывающий карман белка.

4.2.3. Создание флуороген-активирующих белков

Поиск готовых флуороген-активирующих белков среди известных структур в базе данных PDB показал высокий предсказательный потенциал выбранного метода предварительного компьютерного скрининга. На следующем этапе работы мы решили объединить разработанный протокол с дополнительным этапом *in silico* мутагенеза.

4.2.3.1. Выбор белка для моделирования

В качестве основы для создания флуороген-активирующего белка был выбран белок B1c из *E. coli*. B1c является представителем семейства липокалинов — обширной группы небольших транспортных белков с разнообразными аминокислотными последовательностями и характерной вторичной структурой: они содержат восемь антипараллельных β -сегментов, которые формируют цилиндр, содержащий внутри лиганд-связывающий карман (рисунок 42) [205].

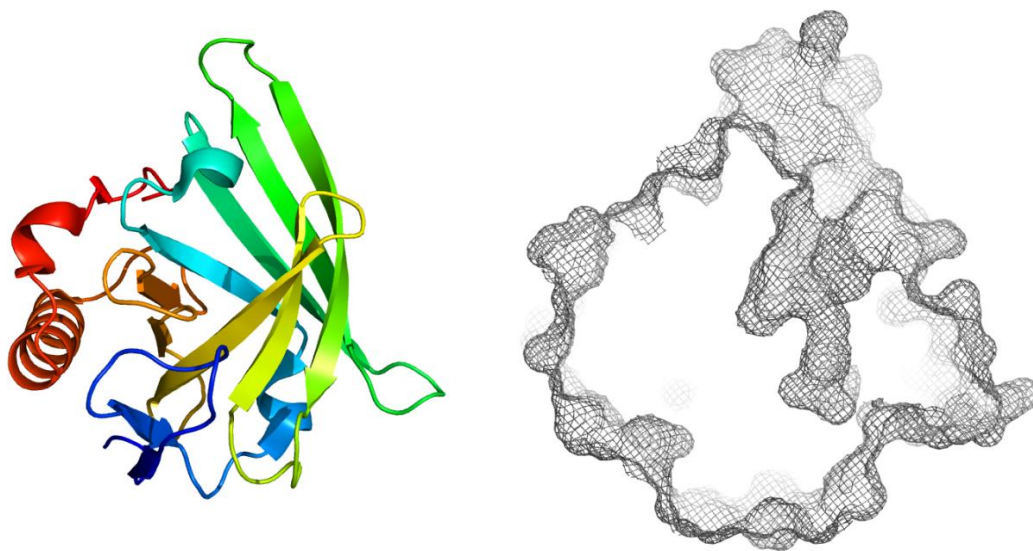


Рисунок 42. Пространственная структура липокалина B1c (слева) и его лиганд-связывающего кармана (справа).

Мы остановили свой выбор на этом белке сразу по нескольким причинам. Во-первых, это один из немногих липокалинов, который не имеет в своем составе цистеинов, образующих дисульфидные мостики. Во-вторых, в базе данных PDB для белка B1c существуют три кристаллические структуры (1QWD [159], 2ACO [206] и 3MBT [207]), что позволяет производить поиск аминокислотных остатков, участвующих в связывании лиганда, гомологичное моделирование структуры его мутантов и прочие компьютерные вычисления. В-третьих, как уже ранее обсуждалось, использование гетерологичного белка уменьшает шансы нежелательного влияния на физиологию исследуемого организма. И наконец он имеет сравнительно небольшой размер — всего 160 аминокислотных остатка (без сигнального пептида, направляющего белок в периплазматическое пространство).

4.2.3.2. *In silico* мутагенез липокалина

Анализ имеющихся кристаллических структур позволил выбрать 23 позиции в исходном белке, боковые цепи аминокислотных остатков в которых вносят существенный вклад в формирование лиганд-связывающего кармана. С помощью программы MODELLER [129] было произведено моделирование структур всех возможных мутантов белка B1c по выбранным позициям с не более чем двумя заменами относительно исходного сиквенса (437 точечных мутантов и 91 333 варианта с заменами по двум положениям, то есть, с учетом белка дикого типа, была проанализирована 91 771 структура).

Для каждой из полученных моделей как и в случае с белками из базы данных PDB с помощью программы AutoDoc Vina [119] было проведено моделирование взаимодействия с аналогом хромофора флуоресцентного белка GFP (в нейтральной и ионизованной форме), по результатам которого для последующей проверки *in vitro* было отобрано 20 мутантов, каждый из которых содержал по две точечные мутации (рисунок 43).

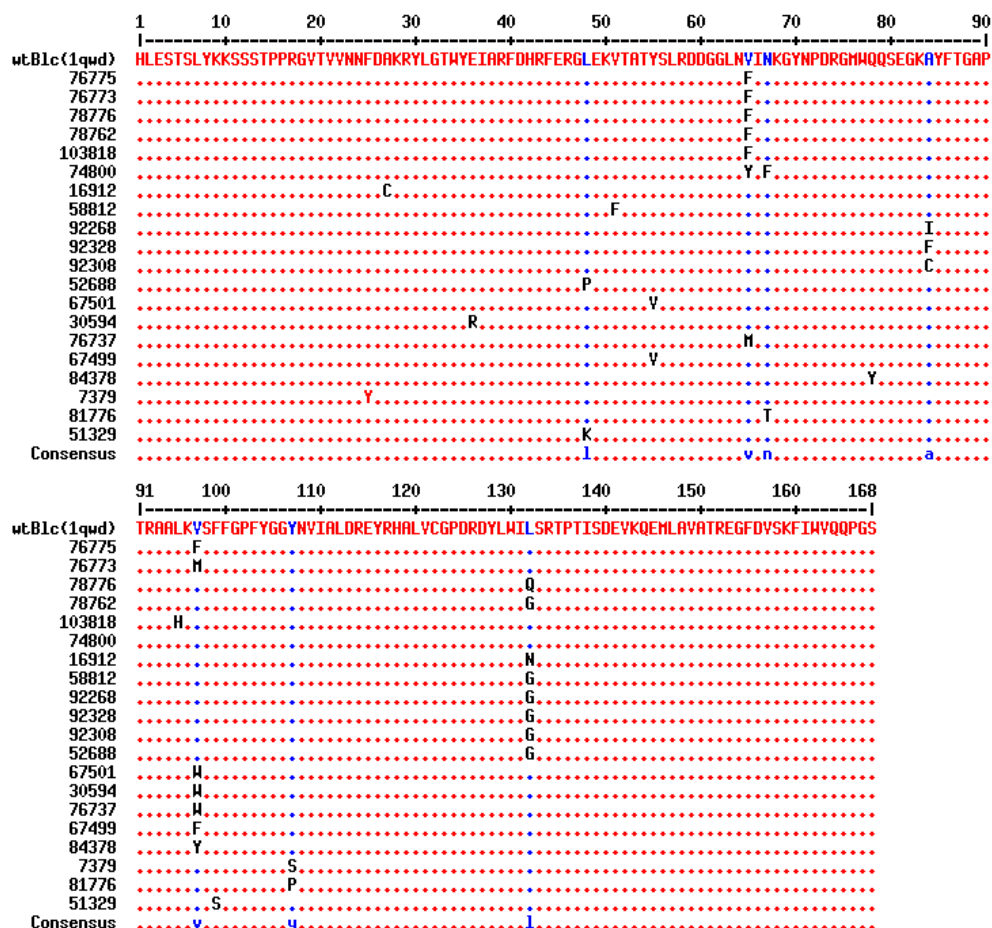


Рисунок 43. Аминокислотная последовательность белка B1c и 20 выбранных для дальнейшего анализа мутантов.

4.2.3.3. *In vitro* анализ библиотеки мутантов

In vitro анализ библиотеки мутантов проводился по описанному ранее протоколу при помощи флуоресцентного микроскопа (рисунок 44). Фрагмент итоговой таблицы представлен на рисунке 45.

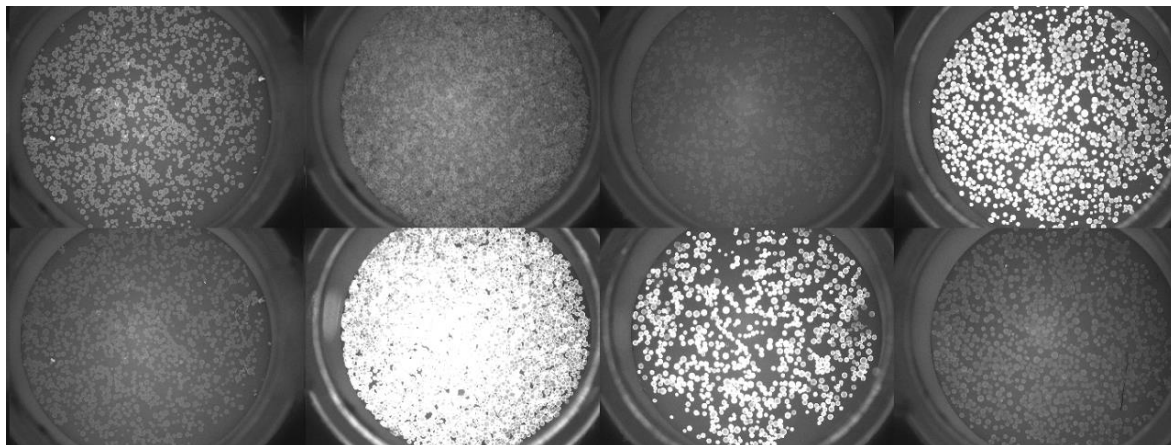


Рисунок 44. Пример *in vitro* анализа взаимодействия мутантов белка B1c, иммобилизованных на металл-аффинной смоле, с флуорогенами при помощи флуоресцентной микроскопии. Верхний ряд — набор фильтров GFP-BP (472.5/30, 520/35), нижний ряд — набор фильтров TexasRed (562/40, 624/40).

Лишь 5 из 20 выбранных в ходе *in silico* моделирования мутантов не продемонстрировали значительного увеличения интенсивности флуоресценции ни с одним из исследованных в скрининге флуорогенов, ещё раз подтвердив высокую эффективность использованного метода предсказания взаимодействий. Среди оставшихся 15 белков из рассмотрения были исключены мутанты, имеющие признаки высокоаффинного связывания неизвестного лиганда (обладающие интенсивной окраской после диализа и/или гель-фильтрационной очистки). Из флуорогенов же были выбраны только те, которые давали яркий сигнал, но лишь с небольшим количеством мутантов (эффективное взаимодействие с множеством мутантов указывает на неспецифичность связывания и, соответственно, увеличивает вероятность нахождения флуорогеном других белков-партнеров в клетке). Таким образом, дальнейшая работа была продолжена с двумя белками (16912 и 78776) и тремя борированными флуорогенами (M739, M766 и M774).

4.2.3.4. Характеристика выбранных пар белок-флуороген

Наблюдаемое возрастание флуоресцентного сигнала в используемой системе скрининга, как мы уже выяснили, может быть вызвано двумя процессами: простым связыванием флуорогена с белком (в результате чего флуороген, обладающий низкой, но детектируемой флуоресценцией в свободном виде, накапливается на поверхности смолы, что приводит к более яркой флуоресценции в сравнении с чистым носителем) или связыванием, сопровождающимся

увеличением квантового выхода флуоресценции и/или изменением спектральных свойств флуорогена. Поэтому для каждой выбранной пары белок-флуороген были подробно изучены спектральные свойства комплекса, константа связывания лиганда и изменение интенсивности флуоресценции (таблица 10).

Таблица 10. Свойства выбранных на основе *in vitro* анализа пар белок-флуороген.

	Белок	16912	78776	свободный флуороген
	Замены	A27C / L132N	V65F / L132Q	n.d.
M739	λ_{ex} , нм	513	546	520
	λ_{em} , нм	542	565	563
	FQY, %	32	15	3.5
	FI @ λ_{max}	86	10.5	n.d.
	K _d , мкМ	0.1	9	n.d.
M766	λ_{ex} , нм	n.d.	n.d.	511
	λ_{em} , нм	542	559	564
	FQY, %	8	15	5.0
	FI @ λ_{max}	9.4	5.5	n.d.
	K _d , мкМ	0.9	35	n.d.
M774	λ_{ex} , нм	515	552	522
	λ_{em} , нм	545	568	568
	FQY, %	20	12	2.2
	FI @ λ_{max}	65	17	n.d.
	K _d , мкМ	0.5	5.0	n.d.

λ_{ex} — максимум спектра возбуждения флуоресценции

λ_{em} — максимум спектра эмиссии флуоресценции

FQY (fluorescence quantum yield) — квантовый выход флуоресценции

FI @ λ_{max} — максимальное возрастание флуоресцентного сигнала

K_d — константа диссоциации

n.d. — не определяли

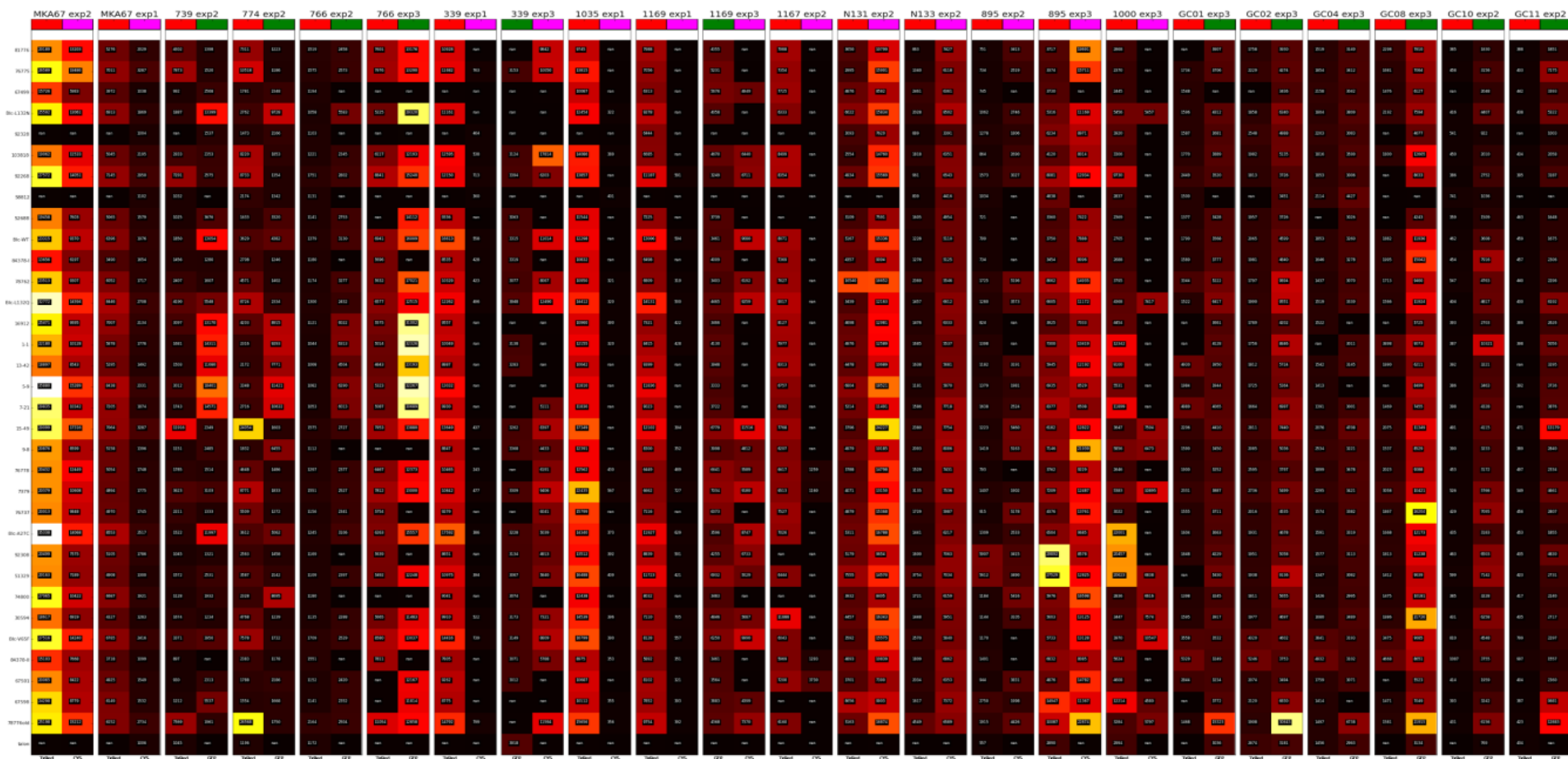


Рисунок 45. Пример таблицы с результатами полуколичественной оценки взаимодействия мутантов белка B1c с флуорогенами. По вертикали отмечены белки, по горизонтали — флуорогены. Цветом на пересечении строки с белком и столбца с флуорогеном отмечено возрастание интенсивности флуоресценции данного комплекса относительно флуоресценции раствора флуорогена (черный — возрастания интенсивности флуоресценции нет, белый — возрастание интенсивности флуоресценции максимально).

Спектры возбуждения флуоресценции комплексов белка 16912 с различными флуорогенами практически совпадали со спектрами растворов свободных флуорогенов, а максимумы спектров эмиссии флуоресценции были сдвинуты в синюю область примерно на 20 нм (рисунок 46 и таблица 10). Наблюдалось обратимое связывание белка с флуорогенами, характеризующееся константами диссоциации от 100 нМ до практически 1 мкМ. С учетом увеличения квантового выхода флуоресценции и сдвига максимума эмиссии, наибольшее возрастание флуоресцентного сигнала относительно раствора флуорогена достигало двух порядков.

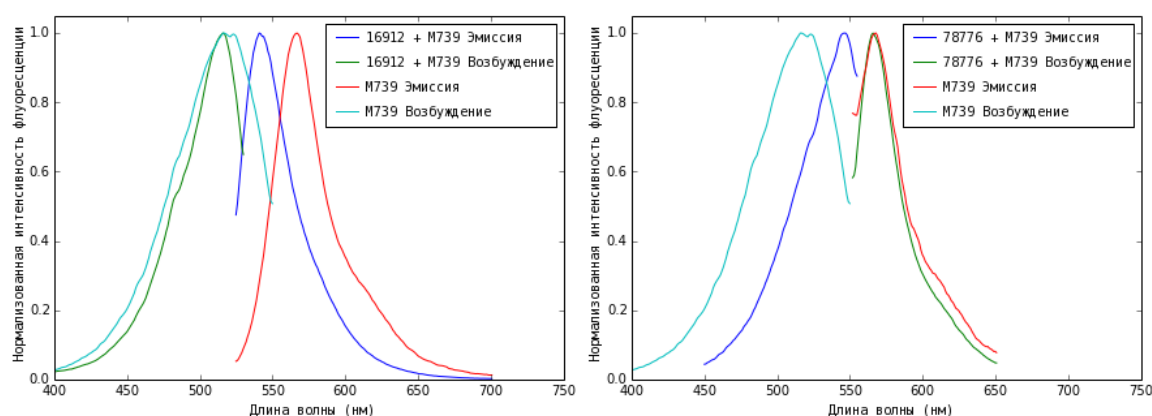


Рисунок 46. Нормализованные спектры возбуждения и эмиссии флуоресценции комплексов белков 16912 и 78776 с флуорогеном M739 и чистого флуорогена.

Спектры эмиссии флуоресценции комплексов белка 78776, наоборот, либо совпадали со спектрами раствора флуорогена, либо были незначительно сдвинуты в ту или иную сторону на несколько нанометров (несовпадение также может быть вызвано неточным определением максимумов эмиссии флуоресценции свободного флуорогена из-за его низкого квантового выхода флуоресценции), при этом наблюдались существенные сдвиги спектров возбуждения флуоресценции в красную область (рисунок 46 и таблица 10). Связывание с флуорогенами происходило с меньшей (микромолярной) аффинностью. Всё это суммарно приводило к меньшему максимальному возрастанию флуоресцентного сигнала по сравнению с комплексами с белком 16912.

В случае флуорогенов M739 и M774 наблюдалось как специфическое связывание лигандов белками, так и значительное изменение спектральных свойств при образовании комплексов. Возрастание флуоресценции при взаимодействии с флуорогеном M766 было выражено намного слабее, поэтому он был исключен из дальнейшего рассмотрения.

4.2.3.5. Изучение роли замен

Каждый из двух выбранных белков (16912 и 78776) имеет по две аминокислотные замены относительно белка дикого типа (рисунок 43). Для того чтобы попытаться понять роль внесенных мутаций в изменении свойств белков, были сделаны четыре мутанта, каждый из которых содержал лишь одну аминокислотную замену (A27C, V65F, L132N или L132Q, соответственно). Эти белки, а также белок дикого типа, были выделены, после чего было исследовано их взаимодействие с флуорогенами M739 и M774. Полученные результаты приведены в таблице 11.

Таблица 11. Свойства белка дикого типа и точечных мутантов.

Белок	Окраска	M739					M774				
		λ_{ex} , нм	λ_{em} , нм	FQY, %	FI @ λ_{max}	K_d , мкМ	λ_{ex} , нм	λ_{em} , нм	FQY, %	FI @ λ_{max}	K_d , мкМ
A27C	++	510	539	32	168	4	512	548	15	58	8
V65F	+++++	507	536	27	63	4	546	571	14	8	7.5
L132N	+	514	542	30	67	0.18	520	546	20	55	0.6
L132Q	+	513	546	34	50	3.5	544	566	30	20	7
wtBlc	++++	510	540	60	180	3	514	552	19	50	10
свободный флуороген	n.d.	520	563	3.5	n.d.	n.d.	522	568	2.2	n.d.	n.d.

λ_{ex} — максимум спектра возбуждения флуоресценции

λ_{em} — максимум спектра эмиссии флуоресценции

FQY (fluorescence quantum yield) — квантовый выход флуоресценции

FI @ λ_{max} — максимальное возрастание флуоресцентного сигнала

K_d — константа диссоциации

n.d. — не определяли

Наиболее очевидной из всех исследованных является роль замены L132N. В обоих случаях (белок с одной заменой L132N и 16912, содержащий эту замену) мутанты по данному положению характеризовались более чем на порядок меньшей константой связывания с флуорогенами в сравнении с белком дикого типа. Это может объясняться, например, образованием аспарагином водородной связи с флуорогеном, стабилизирующей его в лиганд-связывающем кармане.

Так как ни один из точечных мутантов не демонстрировал спектральных свойств, схожих со свойствами 78776 (сдвинутый в длинноволновую область спектр возбуждения флуоресценции и совпадение спектра эмиссии флуоресценции со спектром свободного хромофора), можно предположить их кумулятивный эффект. Из предыдущих исследований [198] известно, что многие борированные GFP-подобные флуорогены обладают ярко выраженным сольватохромизмом — их спектральные свойства сильно зависят от растворителя (с ростом

полярности растворителя спектры смещаются в красную область) (рисунок 47). Спектральные свойства, демонстрируемые комплексами флуорогенов с белком дикого типа и всеми охарактеризованными точечными и двойными мутантами, кроме 78776, схожи со спектральными свойствами флуорогенов в неполярных растворителях. Это согласуется с высказанным предположением о связывании белком B1c лизофосфолипидов [206], что должно быть обусловлено сильной гидрофобностью лиганд-связывающего кармана. Введение двух аминокислот с большими боковыми цепями (V65F и L132Q) могло уменьшить размер кармана, сделав его менее доступным для лигандов, и привести к увеличению экспонированности связанных флуорогенов в воду (а значит смещению спектров в красную область) и повышению констант диссоциации.

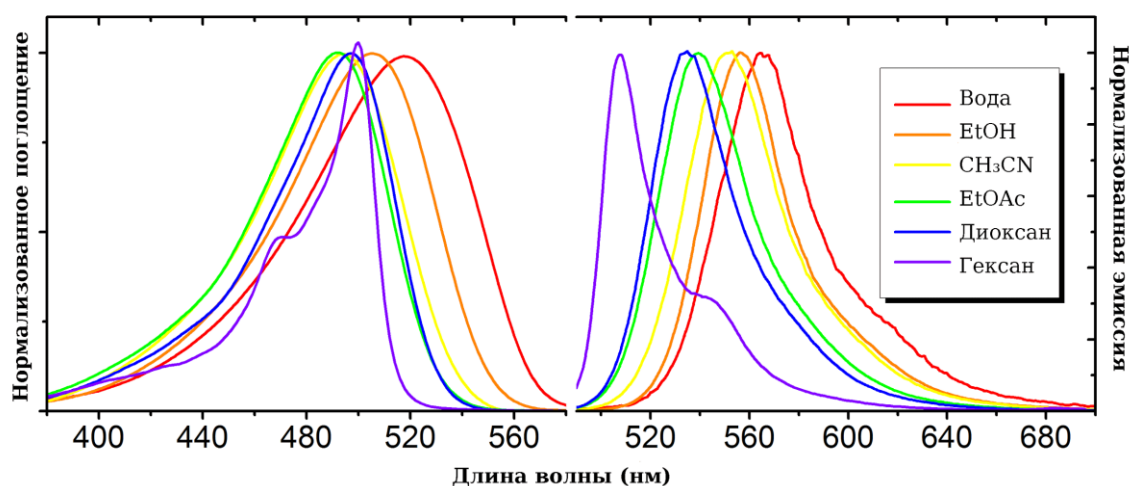


Рисунок 47. Нормализованные спектры поглощения и эмиссии флуоресценции флуорогена M739 в различных растворителях. [198]

Роль цистеина, заменившего аланин в 27 положении в мутанте 16912, вызывает больше всего вопросов, так как он располагается вдали от предсказанного при моделировании положения лиганда в белке. Существенным отличием мутанта A27C от белка дикого типа является его значительно менее выраженная окрашенность. Возможно данная замена неполярной аминокислоты на полярную затрудняет взаимодействие белка с природным гидрофобным лигандом или влияет на фолдинг.

Два из проанализированных белков (мутант A27C и белок дикого типа) существенно превосходили по яркости ранее отобранные для работы 16912 и 78776. Их также было решено проверить *in vivo*.

4.2.3.6. Использование мутантов липокалина в качестве метки для флуоресцентной микроскопии

Для проверки возможности использования пар белок-флуороген *in vivo* самые, на наш взгляд, перспективные варианты (wtBlc, 16912, 78776 и A27C) были заклонированы в вектор H2B-TagBFP (Evrogen). При этом нуклеотидные последовательности белка Blc или его мутантов помещали после последовательности, кодирующей флуоресцентный белок TagBFP, отделяя от него нуклеотидной последовательностью, кодирующей гибкий линкер, состоящий из 6 аминокислот (SGLRSR). В результате было получено и проверено посредством секвенирования четыре конструкции для экспрессии в культуре клеток млекопитающих:

H2B-TagBFP-wtBlc
H2B-TagBFP-16912
H2B-TagBFP-78776
H2B-TagBFP-A27C

Полученные конструкции были использованы для временной трансфекции клеток линии HEK293T. Анализ результатов проводился через 48 часов после трансфекции при помощи флуоресцентного микроскопа Keyence BZ-9000, используя фильтры DAPI-BP (377/50, 447/60), GFP-BP (472.5/30, 520/35) и TexasRed (562/40, 624/40). Перед съемкой среду для культивации (DMEM) заменяли на раствор Хэнкса, а в процессе съемки добавляли исследуемый флуороген (M739 или M774) до конечной концентрации 1 мкМ.

Все конструкции, экспрессированные в культуре клеток млекопитающих, при добавлении флуорогена демонстрировали моментальное появление флуоресцентного сигнала в ядре, колокализованного с сигналом TagBFP (рисунок 48), тем самым подтверждая результаты, полученные *in vitro* на выделенных белках. Однако при экспрессии конструкции H2B-TagBFP-wtBlc мы также могли видеть слабую (<10% интенсивности от сигнала, получаемого при добавлении лиганда) красную флуоресценцию (детекция в канале TexasRed (562/40, 624/40)) в ядре до добавления флуорогена. Вероятно, какие-то вещества (или их аналоги), обуславливающие интенсивную окраску выделенного из бактериальной культуры белка, также присутствуют и в эукариотических клетках и способны к флуоресценции. Мы проверили выделенный белок wtBlc (рисунок 49) и обнаружили флуоресценцию с максимумом эмиссии при 650 нм и двумя основными максимумами в спектре возбуждения при 340 и 600 нм соответственно. К сожалению, все попытки выделить и охарактеризовать связываемое белком дикого типа вещество не увенчались успехом. Поэтому, несмотря на данные о высоком квантовом выходе флуоресценции комплекса wtBlc + флуороген M739, полученные *in vitro* (60%, таблица 11), нами было принято решение не использовать белок дикого типа в дальнейшей работе

из-за возможных проблем, связанных, по-видимому, с высокоаффинным связыванием данным белком неустановленного вещества (например, из-за вероятности влияния на физиологию исследуемой клетки).

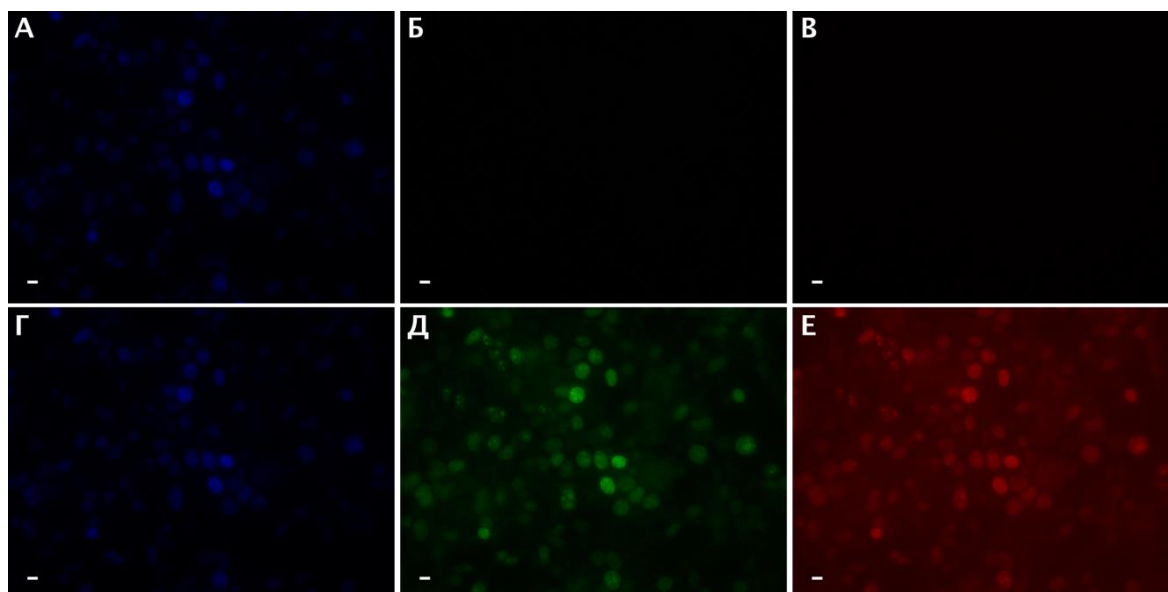


Рисунок 48. Флуоресцентная широкопольная микроскопия живых клеток линии HEK293T, временно трансфицированных конструкцией H2B-TagBFP-16912; масштабная линейка 10 мкм. **А-В.** До после добавления флуорогена M739; **Г-Е.** После добавления флуорогена M739. **АГ.** Набор фильтров DAPI (360/40, 460/50); **БД.** Набор фильтров GFP (470/40, 525/50); **ВЕ.** Набор фильтров TexasRed (562/40, 624/40).

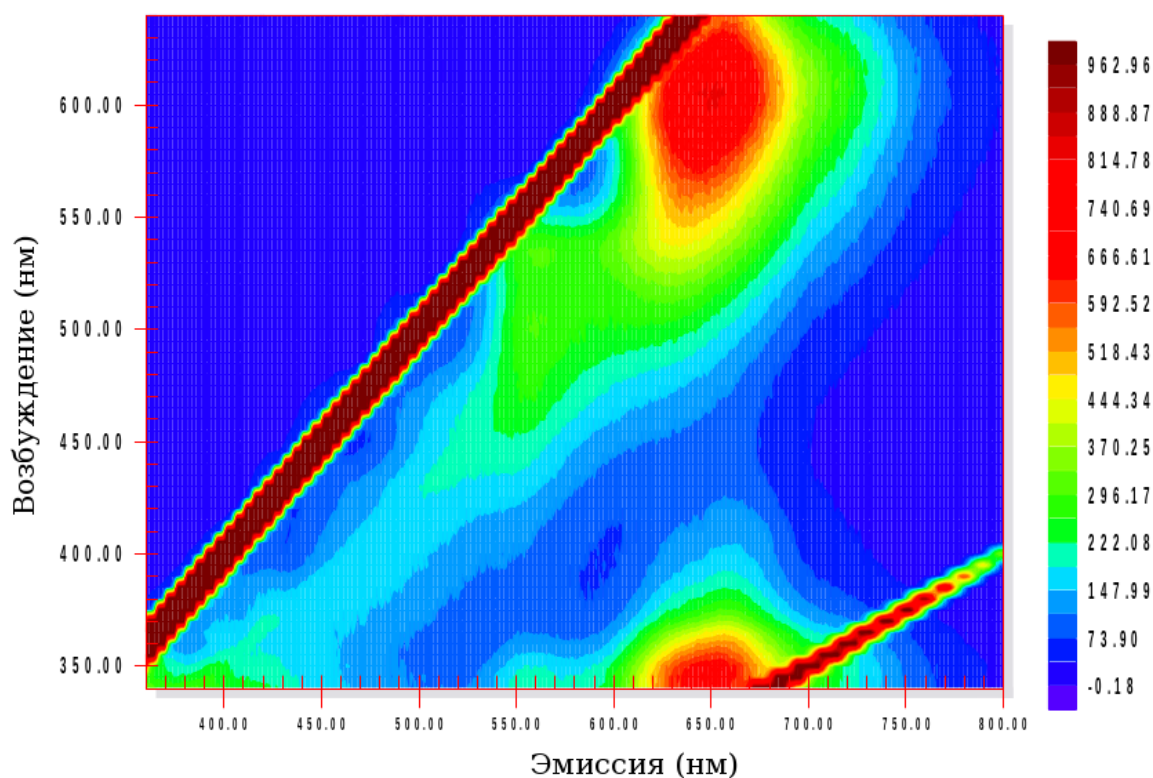


Рисунок 49. 2D спектр флуоресценции белка wtBls.

Во всех оставшихся конструкциях более яркий сигнал был получен при использовании флуорогена M739. При этом, в полном соответствии с *in vitro* данными, в парах H2B-TagBFP-16912 + флуороген M739 и H2B-TagBFP-A27C + флуороген M739 флуоресцентный сигнал был ярче в зелёной области спектра (детекция в канале GFP-BP (472.5/30, 520/35)), в то время как флуоресценция в паре H2B-TagBFP-78776 + флуороген M739 была смещена в красную область и лучше детектировалась при помощи набора фильтров TexasRed (562/40, 624/40). Так как химически, спектрально, а также по взаимодействию с исследуемыми мутантами липокалина B1c флуорогены M739 и M774 похожи, дальнейшая работа *in vivo* проводилась лишь с флуорогеном M739, комплексы которого отличались более высоким квантовым выходом флуоресценции и, в большинстве случаев, более низкими константами связывания.

Некоторые белки в силу своего строения или выполняемой функции могут быть помечены только с одной стороны [208], следовательно очень важно, чтобы разрабатываемая метка была способна эффективно образовывать как N-, так и C-концевые белки слияния. Поэтому следующим этапом работы стала проверка способности выбранных мутантов белка B1c присоединять другой белок со стороны их C-конца. Для этого на основе вектора H2B-TagBFP (Evrogen) были созданы конструкции H2B-*<мутант белка B1c>*-TagBFP. Добавление флуорогена M739 к клеткам линии HEK293T, временно трансфицированным данными векторами, вызывало появление локализованной в ядре флуоресценции, сравнимой по интенсивности с флуоресценцией соответствующих конструкций H2B-TagBFP-*<мутант белка B1c>*. Таким образом, мы показали, что разрабатываемая на основе липокалина B1c флуоресцентная метка сохраняет свои свойства при присоединении дополнительного белка как с N-, так и с C-конца. Это позволяет предположить, что она может быть успешно использована для мечения белков интереса с любой стороны.

Фотостабильность флуоресцентных меток также является одним из их ключевых характеристических параметров [209]. Теоретически нековалентное обратимое взаимодействие флуороген-активирующего белка с флуорогеном может привести к созданию практически не фотообесцвечивающейся флуоресцентной метки [210]. Во-первых, индуцируемое светом разрушение лиганда (если только не происходит фотохимического взаимодействия с белком, блокирующего дальнейший обмен) может быть компенсировано его избыточной начальной концентрацией или добавлениями в процессе съемки. Во-вторых, небольшое время жизни комплекса само по себе уменьшает вероятность фотообесцвечивания. Таким образом, от нашей системы мечения мы ожидали увеличения фотостабильности в сравнении с флуоресцентными белками, содержащими схожий по своей структуре хромофор, а также различия в фотостабильности в парах с разными константами связывания.

Фотостабильность комплексов белок-флуороген проверяли на модели культуры клеток человека (НЕК293Т), используя лазерный сканирующий конфокальный микроскоп. Клетки временно трансфицировали векторами, кодирующими белок слияния гистона H2В либо с контрольным флуоресцентным белком (мы использовали хорошо зарекомендовавшие себя, широко используемые и спектрально близкие флуоресцентные белки: EGFP в качестве контроля для комплексов с максимумом эмиссии флуоресценции до 540 нм и mKate в качестве контроля для комплекса с максимумом эмиссии флуоресценции выше 540 нм), либо с флуоресцентным белком TagBFP и одним из мутантов белка B1c.

Как мы и предполагали, фотостабильность комплексов оказалась выше фотостабильности флуоресцентных белков. При этом в случае пары белок-флуороген с низкой константой диссоциации (16912 + M739, $K_d = 100$ нМ) наблюдаемое увеличение составляло 2-3 раза. Фотостабильность менее аффинных комплексов (A27C + M739, $K_d = 4$ мкМ и 78776 + M739, $K_d = 9$ мкМ) отличалась более чем на порядок (рисунок 50А). Вместе с тем скорости возрастания и убывания флуоресцентного сигнала при, соответственно, добавлении и отмывании флуорогена также согласовывались с определенными *in vitro* константами диссоциации: быстрее всего появлялся и исчезал сигнал от взаимодействия лиганда с белком 78776, 16912 демонстрировал самую медленную динамику (рисунок 50Б). Это ещё раз подтверждает высказанное предположение о влиянии времени жизни комплекса на его фотостабильность.

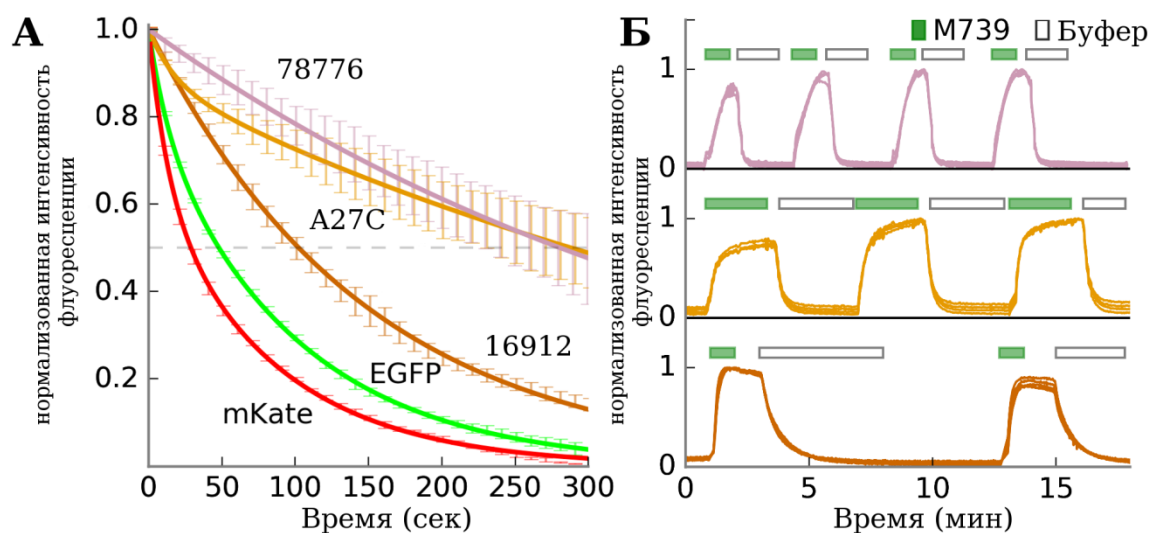


Рисунок 50. Характеристика комплексов *in vivo*. **А.** Кривые фотообесцвечивания флуоресцентного сигнала в ядрах живых клеток линии НЕК293Т, полученные при помощи лазерной сканирующей конфокальной микроскопии; **Б.** Изменение интенсивности флуоресцентного сигнала в клетках линии HeLa Kyoto, временно трансфицированных векторами, кодирующими белок слияния гистона H2В с флуоресцентным белком TagBFP и одним из мутантов белка B1c (сверху вниз: 78776, A27C, 16912), в последовательных циклах добавления и отмывания флуорогена M739, широкопольный флуоресцентный микроскоп.

Для того чтобы показать, что созданная метка пригодна, как и флуоресцентные белки, для мечения разнообразных структур клетки *in vivo*, нами был создан ряд конструкций, представляющих собой белки слияния белков цитоскелета с мутантами липокалина B1c (A27C, 16912 и 78776), как содержащих, так и не содержащих в своем составе флуоресцентный белок TagBFP (см. материалы и методы).

Временная трансфекция клеток линии HeLa Kyoto созданными векторами подтвердила, что созданные белки могут быть успешно использованы в качестве флуоресцентной метки для широкопольной и конфокальной микроскопии. При этом добавление флуорогена никак не влияет на локализацию исследуемых белков в клетке, а наличие сигнала в конструкциях без флуоресцентного белка TagBFP служит ещё одним доказательством того, что флуоресценция действительно связана с взаимодействием мутантов белка B1c с флуорогеном и не зависит от присутствия TagBFP (рисунок 51). Наиболее перспективным выглядит использование белков 16912 и A27C, обладающих более высоким квантовым выходом флуоресценции.

Так как максимально достижимое разрешение флуоресцентной микроскопии составляет 200-350 нм, а многие клеточные структуры имеют меньший размер, в последнее время активно развиваются различные подходы, позволяющие повысить разрешающую способность флуоресцентного микроскопа (так называемые методы микроскопии сверхвысокого разрешения). Среди них — микроскопия структурированного освещения (SIM, structured illumination microscopy), микроскопия истощения флуоресценции вынужденным излучением (STED, stimulated emission depletion), локализационная микроскопия одиночных молекул (SMLM, single-molecule localization microscopy) и др. [211].

В методах локализационной микроскопии увеличения разрешения достигают за счёт получения серии изображений с очень низкой плотностью флуоресцентных сигналов, вычисления положения каждой флуоресцирующей молекулы и последующего наложения всех обработанных кадров [212]. Получения необходимой плотности флуоресцентных сигналов в течении длительного времени можно добиться разными способами, в том числе за счёт использования обратимого взаимодействия с флуорогенным красителем, что было показано на примере малахитового зелёного, связываемого одноцепочечным антителом [48].

Наши коллеги из лаборатории флуоресцентного биоимиджинга (НижГМА, Нижний Новгород, Россия) протестировали и подтвердили пригодность разработанной системы флуоресцентного мечения для локализационной микроскопии сверхвысокого разрешения (рисунок 52). При этом лучшие результаты были получены для белков A27C и 78776,

обладающих более низкими константами диссоциации и, соответственно, быстрее обменивающих флуороген.

Таким образом, имеющиеся данные подтверждают, что созданная на основе мутантов липокалина B1c и GFP-подобных флуорогенов система флуоресцентного мечения является перспективной альтернативой флуоресцентным белкам и химическим красителям для визуализации структур и процессов в живых системах и совместима с основными используемыми в настоящее время техниками флуоресцентной микроскопии.

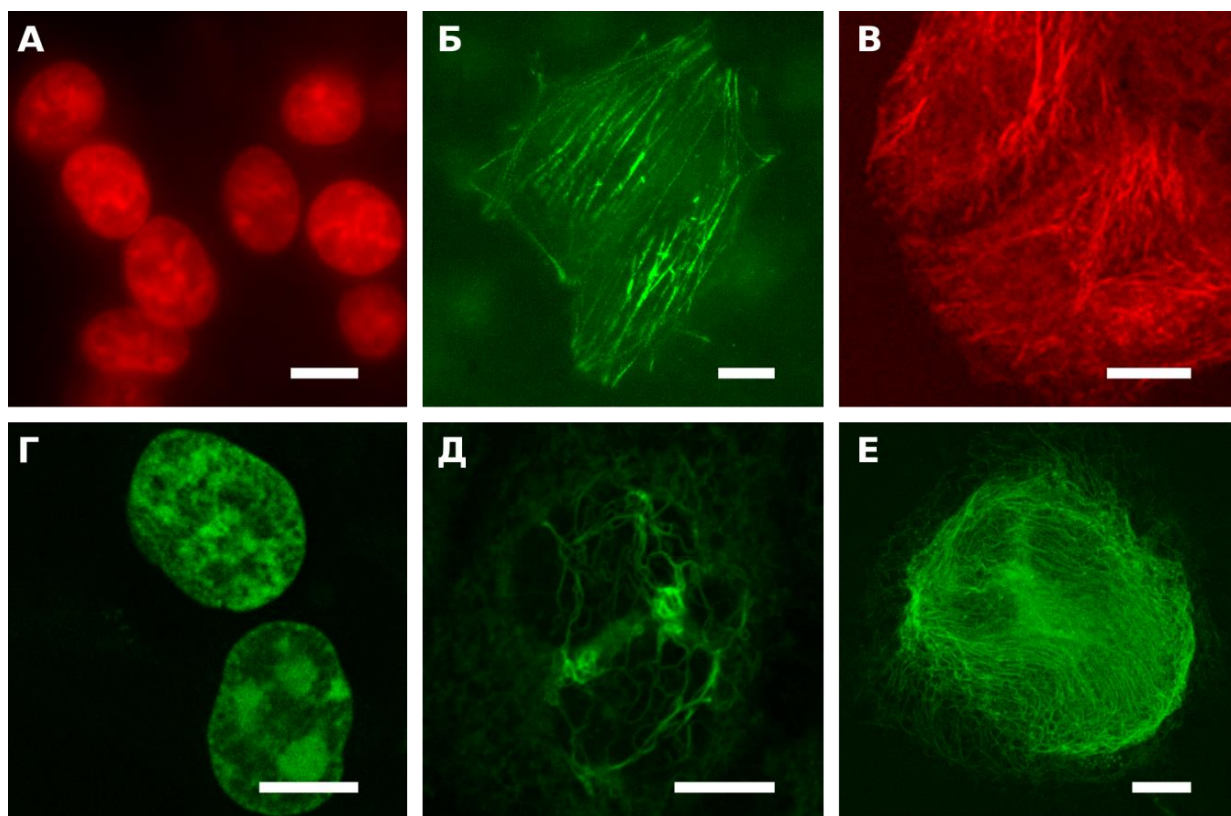


Рисунок 51. Флуоресцентная микроскопия живых клеток, экспрессирующих конструкции с мутантами белка B1c, с флуорогеном M739; масштабная линейка 10 мкм. **А.** Клетки линии HEK293T, временно трансфицированные вектором H2B-TagBFP-78776 (широкопольная микроскопия, набор фильтров TexasRed (562/40, 624/40)); **Б.** Клетки линии NIH-3T3, временно трансфицированные вектором α -actinin-16912 (широкопольная микроскопия, набор фильтров GFP (470/40, 525/50)); **В.** Клетки линии HeLa Kyoto, временно трансфицированные вектором α -actinin-78776 (конфокальная микроскопия, возбуждение: 543 нм, эмиссия: 560-600 нм); **Г.** Клетки линии HEK293T, временно трансфицированные вектором H2B-TagBFP-16912 (конфокальная микроскопия, возбуждение: 488 нм, эмиссия: 520-560 нм); **Д.** Клетки линии HeLa Kyoto, временно трансфицированные вектором vimentin-TagBFP-A27C (конфокальная микроскопия, возбуждение: 488 нм, эмиссия: 520-560 нм); **Е.** Клетки линии HeLa Kyoto, временно трансфицированные вектором keratin-16912 (конфокальная микроскопия, возбуждение: 488 нм, эмиссия: 520-560 нм).

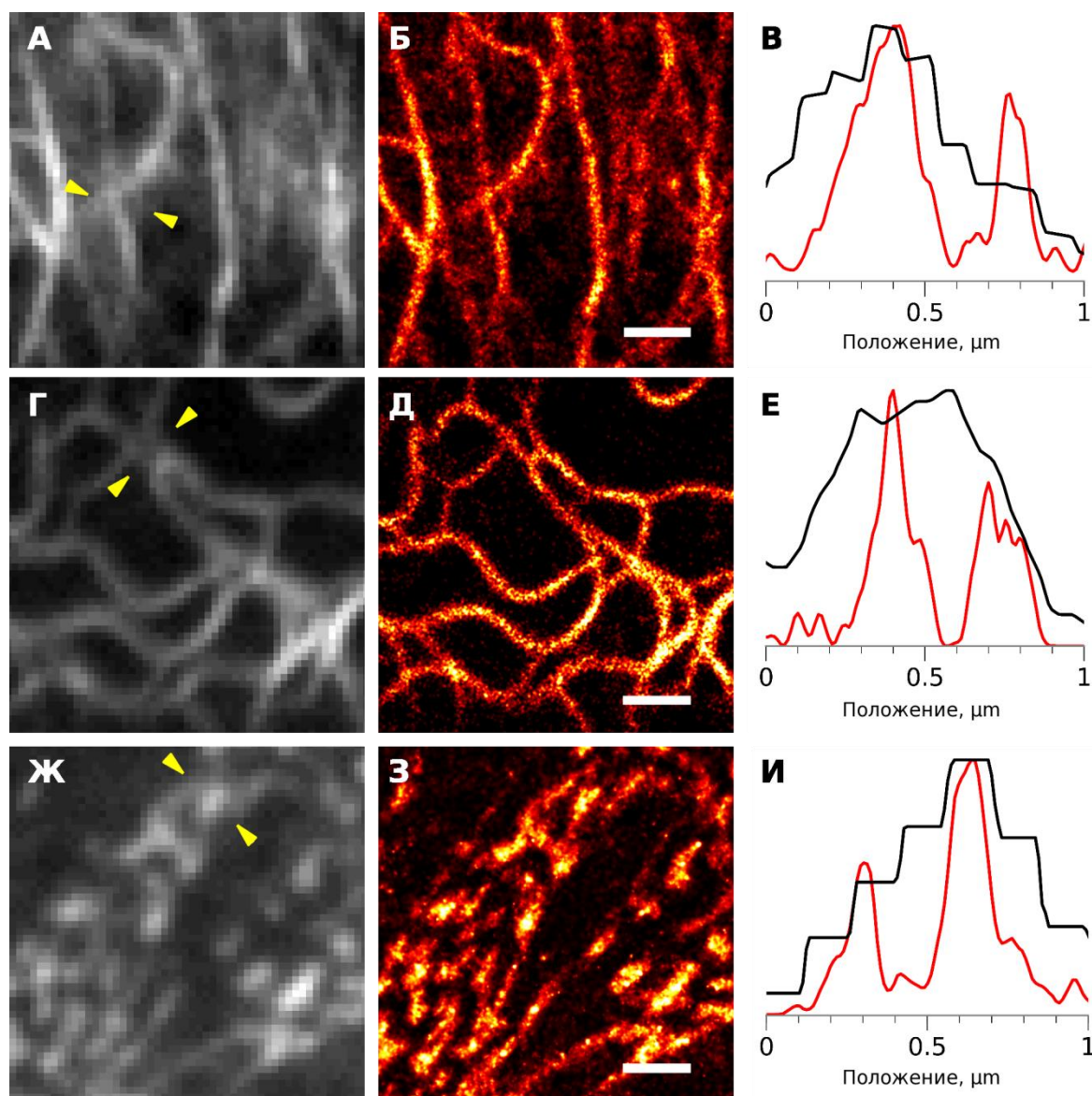


Рисунок 52. Микроскопия сверхвысокого разрешения. Живые клетки линии HeLa Kyoto, временно трансфицированные векторами, кодирующими белки слияния с мутантами белка B1c; масштабная линейка 1 мкм. **АБВ.** Keratin-16912; **ГДЕ.** Vimentin-TagBFP-A27C; **ЖЗИ.** α -actinin-78776. **АГЖ.** Широкопольная микроскопия; **БДЗ.** Реконструкция по 5000 кадрам изображения сверхвысокого разрешения; **ВЕИ.** Нормализованные профили интенсивности флуоресцентного сигнала участка изображения между стрелками, показанными на кадрах широкопольной микроскопии; черная кривая — широкопольная микроскопия, красная кривая — реконструкции изображений сверхвысокого разрешения.

4.2.4. Улучшение созданных белков

Компьютерное моделирование белков в последнее время демонстрирует огромные успехи, однако до сих пор точность таких вычислений ограничена. Часто удается получить лишь минимально детектируемый эффект, который затем усиливают до необходимого уровня другими способами, например путём внесения случайных замен в ген с последующим отбором мутантов с требуемыми свойствами [150]. В нашем случае моделирование сразу привело к получению вариантов белка, пригодных для использования, однако было решено сделать попытку их улучшения.

4.2.4.1. Случайный мутагенез

Библиотека случайных мутантов была получена на основе белка 16912, обладающего высокими аффинностью и квантовым выходом флуоресценции. Бактерии высевали на чашки Петри с агаризованной средой LB, дополненной индуктором синтеза белка (арабинозой) и 2ММ флуорогена M739. Анализ проводили через 16 часов после трансформации и выбирали самые яркие колонии.

Первый раунд случайного мутагенеза выявил два новых белка, значительно превосходящих 16912 по яркости, названных 16912rnd1 и 16912rnd20. Данные мутанты содержали, соответственно, три и две аминокислотные замены в сравнении с исходным белком (рисунок 53). При этом большая часть замен располагалась в удаленных от предполагаемого места связывания флуорогена частях белка, поэтому не могла быть нами предсказана во время компьютерного моделирования в связи с собственноручно наложенными ограничениями. Тем не менее, нахождение замен на значительном расстоянии от лиганд-связывающего кармана вовсе не отрицает возможности их влияния на связывание белка с флуорогеном.

Матрицей для второго раунда случайного мутагенеза были взяты оба новых варианта липокалина (16912rnd1 и 16912rnd20). Анализ библиотеки на основе 16912rnd1 выявил ещё один перспективный вариант — 16912rnd1rnd1, имеющий одну обратную замену (N52T) и одну новую замену (K50R) относительно 16912rnd1 (рисунок 53). В библиотеке на основе 16912rnd20 более ярких мутантов найдено не было.

4.2.4.2. Характеристика белков, полученных случайным мутагенезом

Увеличение яркости флуоресценции мутантных белков, выявленное при скринировании библиотек, может быть обусловлено как возрастанием их аффинности к флуорогену, так и повышением квантового выхода флуоресценции комплекса. Для того, чтобы выяснить природу наблюдаемых эффектов, все три новых белка (16912rnd1, 16912rnd20 и 16912rnd1rnd1) были

выделены, и их взаимодействие с флуорогенами M739 и M774 было проанализировано более детально (таблица 12).

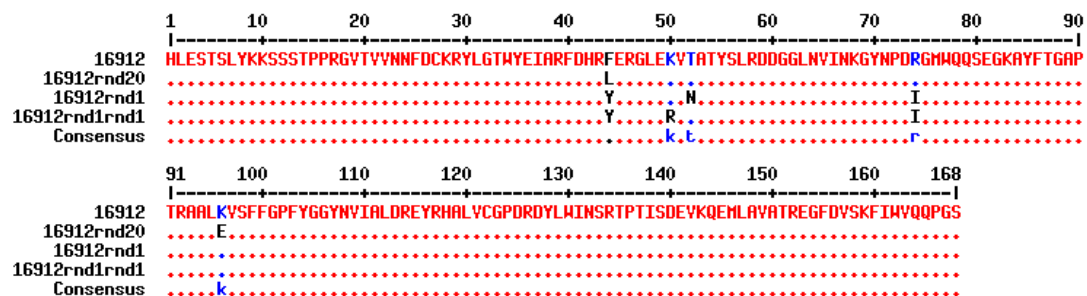


Рисунок 53. Аминокислотная последовательность белка 16912 и трёх выбранных для дальнейшего анализа мутантов, полученных в ходе случайного мутагенеза.

Спектры эмиссии и возбуждения флуоресценции комплексов белков с флуорогенами совпадали по форме и положению максимумов со спектрами комплексов исходного белка 16912. При этом для флуорогена M739, который использовался в качестве селективного лиганда в процессе скрининга библиотек, как константы диссоциации, так и квантовый выход флуоресценции во всех трех случаях были улучшены: константы диссоциации в этих парах снизились до 30-50 нМ (что примерно на два порядка ниже, чем у белка дикого типа), а квантовый выход при этом достиг значений, сравнимых или превышающих таковые для wtB1c. Несмотря на сильную схожесть флуорогенов M739 и M774, взаимодействие полученных путём случайного мутагенеза мутантов с флуорогеном M774 отличалось от взаимодействия исходного белка 16912 гораздо меньше: в двух из трех случаях константа диссоциации не изменилась, улучшения квантового выхода флуоресценции хоть и наблюдались, но были выражены гораздо менее значительно.

Результаты, полученные в эксперименте по случайному мутагенезу белка 16912, показали, что яркость и аффинность полученных ранее мутантов могут быть улучшены. Однако все новые варианты, имеющие большую аффинность к лиганду, скорее всего, будут менее фотостабильны и не применимы для методов микроскопии сверхвысокого разрешения. Способа селективного отбора ведущего к улучшению квантового выхода флуоресценции без понижения константы связывания нами не было обнаружено.

Таблица 12. Свойства белков, полученных случайным мутагенезом.

Белок	Замены	M739					M774				
		λ_{ex} , нм	λ_{em} , нм	FQY, %	FI @ λ_{max}	K_d , мкМ	λ_{ex} , нм	λ_{em} , нм	FQY, %	FI @ λ_{max}	K_d , мкМ
16912	A27C / L132N	513	542	32	86	0.1	515	545	20	65	0.5
16912rnd1	A27C / L132N / F44Y / N52T / R74I	513	541	55	141	0.04	516	545	26	63	0.5
16912rnd20	A27C / L132N / F44L / K96E	511	541	66	126	0.03	519	545	29	75	0.15
16912rnd1rnd1	A27C / L132N / F44Y / K50R / R74I	515	540	44	130	0.05	515	544	29	135	0.6
свободный флуороген	n.d.	520	563	3.5	n.d.	n.d.	522	568	2.2	n.d.	n.d.

λ_{ex} — максимум спектра возбуждения флуоресценции

λ_{em} — максимум спектра эмиссии флуоресценции

FQY (fluorescence quantum yield) — квантовый выход флуоресценции

FI @ λ_{max} — максимальное возрастание флуоресцентного сигнала

K_d — константа диссоциации

n.d. — не определяли

4.2.4.3. *In silico* мутагенез

Мы также попробовали улучшить имеющиеся белки при помощи компьютерного моделирования. Наш выбор остановился на хорошо зарекомендовавшей себя программе Rosetta, которая имеет два подходящих для этой цели алгоритма — «Match» и «Design».

Алгоритм «Match» принимает описание желаемого взаимодействия лиганда с белковым микроокружением и библиотеку структур, в которые программа пытается поместить и зафиксировать требуемые для координации лиганда аминокислотные остатки. В нашем случае вместо библиотеки структур был использован лишь липокалин дикого типа Blc.

Алгоритм «Design» в качестве исходной структуры использует комплекс белка с лигандом. В белок случайным образом вносятся мутации, после чего оценивается энергия взаимодействия и общая энергия комплекса. За основу для моделирования при помощи алгоритма «Design» были взяты результаты докинга флуорогена M739 в липокалин дикого типа и мутанты 16912, 78776, A27C, V65F, L132Q и L132N, также осуществленного с помощью программы Rosetta.

Моделирование с использованием алгоритма «Match» привело к внесению значительного количества аминокислотных замен в кодирующую последовательность липокалина дикого типа

при незначительном уменьшении общей энергии структуры и энергии взаимодействия белка с лигандом, в то время как использование алгоритма «Design» в некоторых случаях предсказало существенные улучшения в показателях взаимодействия белков с лигандом при внесении всего 5-7 аминокислотных замен. Дальнейшая работа проводилась только с результатами, полученными с помощью алгоритма «Design».

Самыми низкими показателями энергии взаимодействия белка с лигандом обладали дизайны на основе липокалина дикого типа, а также вариантов 16912, A27C, L132N и L132Q. Анализ мутаций в 50 лучших дизайнах на основе каждого из этих белков (рисунок 54А) позволил выявить наиболее перспективные замены: F44A/Q, N67F, S80Y, E81V и S98A. Более того, один из лучших предсказанных мутантов белка 16912 содержал только эти пять замен (рисунок 54Б).

На основе повторного моделирования взаимодействия структур с перспективными мутациями (F44A, F44Q, N67F, S80Y, E81V и S98A и их различными сочетаниями) с флуорогеном M739 используя стандартный алгоритм докинга, было принято решение создать шесть новых мутантов со следующими заменами:

Qwd_4_F44A (wtBlc F44A/N67F/S80Y/E81V/S98A)
Qwd_4_F44Q (wtBlc F44Q/N67F/S80Y/E81V/S98A)
Qwd_4 (wtBlc N67F/S80Y/E81V/S98A)
16912_4_F44A (16912 F44A/N67F/S80Y/E81V/S98A)
16912_4_F44Q (16912 F44Q/N67F/S80Y/E81V/S98A)
16912_4 (16912 N67F/S80Y/E81V/S98A)

Все замены были внесены в кодирующую последовательность соответствующих исходных белков с помощью методики самособирающегося клонирования.

4.2.4.4. Характеристика белков, полученных в результате *in silico* мутагенеза

Белки Qwd_4_F44A, Qwd_4_F44Q, Qwd_4, 16912_4_F44A, 16912_4_F44Q и 16912_4 были наработаны и выделены по стандартному протоколу. В отличие от белка дикого типа (wtBlc), ни один из мутантов не характеризовался ярко выраженной окраской.

Как и в случае вариантов, полученных случайным мутагенезом, спектры эмиссии и возбуждения флуоресценции комплексов, образованных мутантами белка 16912, демонстрировали лишь незначительные (несколько нм) сдвиги относительно исходных. Свойства мутантов белка дикого типа отличались лишь во взаимодействии с флуорогеном M774: их спектры эмиссии флуоресценции были смещены более чем на 10 нм влево и совпадали со спектрами белка 16912 и его производных.

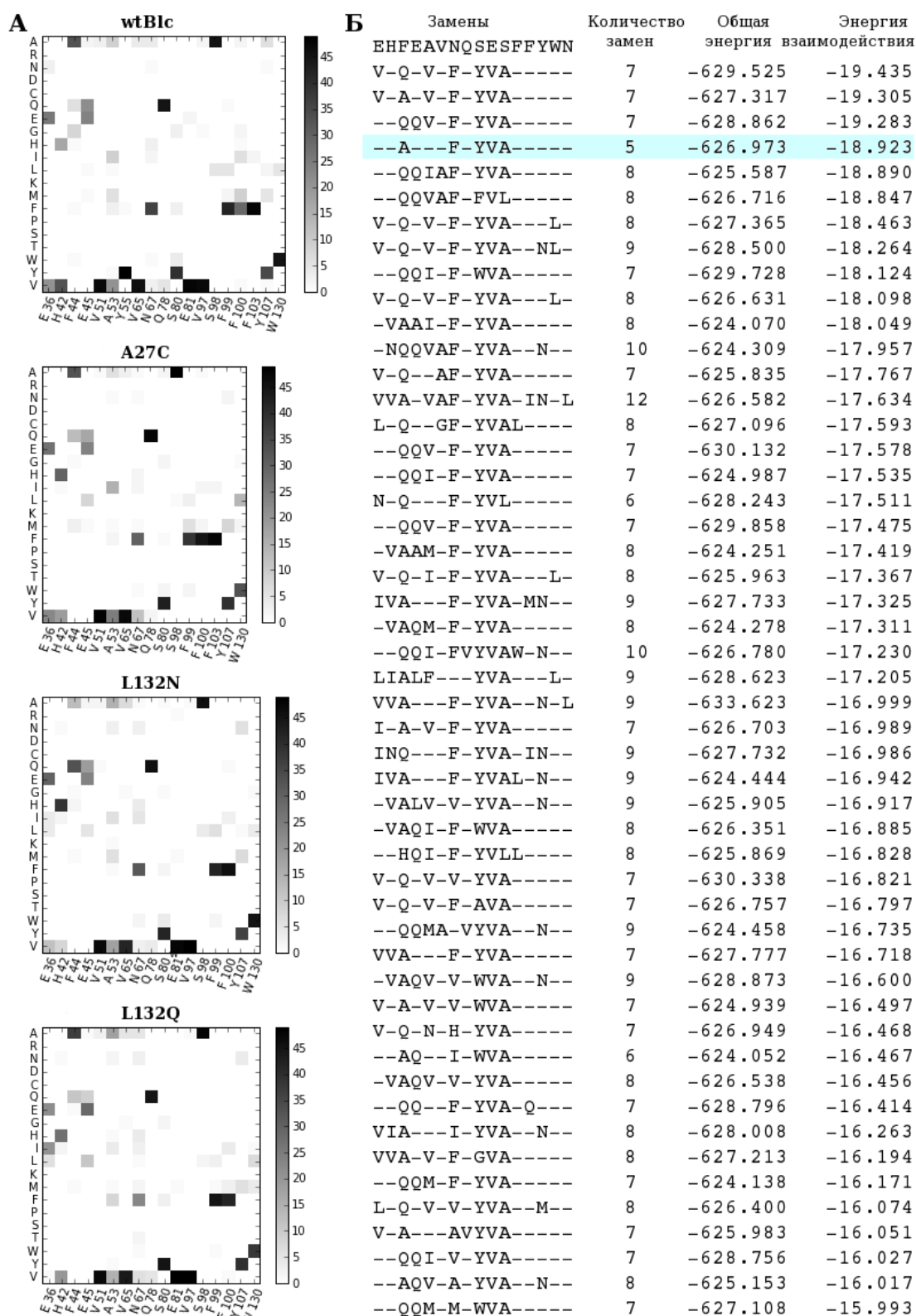


Рисунок 54. Результаты моделирования при помощи алгоритма «Design» программы Rosetta. **А.** Анализ мутаций в 50 лучших дизайнах на основе липокалина дикого типа, а также вариантов A27C, L132N и L132Q. По горизонтали показаны положения в которых были обнаружены замены, по вертикали — 20 аминокислот. Интенсивность окраски квадрата на пересечении отражает частоту встречаемости данной замены в данном положении; **Б.** Характеристики 50 лучших дизайнов на основе белка 16912. Голубым цветом выделен мутант, содержащий только пять наиболее перспективных замен.

В большинстве случаев (за исключением некоторых вариантов, содержащих замену F44Q) в соответствии с ожиданиями смоделированные белки характеризовались увеличенными квантовым выходом флуоресценции и максимальной разницей между флуоресценцией свободного и связанного флуорогенов (таблица 13). Константы диссоциации мутантных белков в основном были либо незначительно, либо существенно выше таковых для исходных белков. Так, для Qwd_F44A было показано взаимодействие с флуорогеном M739 с константой диссоциации 3.5 мкМ, тогда как для белка дикого типа K_d была определена равной 3.0 мкМ. В то же время взаимодействие 16912_F44Q с M739 хромофором было на порядок менее стабильным.

Таблица 13. Свойства белков, полученных результате *in silico* мутагенеза.

Белок	Замены	M739					M774				
		λ_{ex} , нм	λ_{em} , нм	FQY, %	FI @ λ_{max}	K_d , мкМ	λ_{ex} , нм	λ_{em} , нм	FQY, %	FI @ λ_{max}	K_d , мкМ
16912	A27C / L132N	513	542	32	86	0.1	515	545	20	65	0.5
16912_4	A27C / L132N / N67F / S80Y / E81V / S98A	511	541	62	148	0.24	512	543	35	98	0.8
16912_F44Q	A27C / L132N / N67F / S80Y / E81V / S98A / F44Q	521	550	57	70	1.9	521	548	29	65	0.75
16912_F44A	A27C / L132N / N67F / S80Y / E81V / S98A / F44A	513	541	87	230	0.3	520	546	38	115	0.85
wtBlc	-	510	540	60	180	3	514	552	19	50	10
Qwd_4	N67F / S80Y / E81V / S98A	506	536	70	340	3.6	513	540	28	142	21
Qwd_F44Q	N67F / S80Y / E81V / S98A / F44Q	514	543	45	112	19	518	543	28	104	11
Qwd_F44A	N67F / S80Y / E81V / S98A / F44A	509	539	86	330	3.5	515	542	37	156	6.5
Свободный флуороген	n.d.	520	563	3.5	n.d.	n.d.	522	568	2.2	n.d.	n.d.

λ_{ex} — максимум спектра возбуждения флуоресценции

λ_{em} — максимум спектра эмиссии флуоресценции

FQY (fluorescence quantum yield) — квантовый выход флуоресценции

FI @ λ_{max} — максимальное возрастание флуоресцентного сигнала

K_d — константа диссоциации

n.d. — не определяли

В целом умеренное снижение константы связывания с лигандом (особенно для исходно высокоафинного белка 16912), следует рассматривать как положительный результат, так как, как было показано ранее, более высокие константы диссоциации обеспечивают более быстрый обмен флуорогена, что увеличивает флотабильность комплекса и, следовательно, улучшает свойства флуоресцентной метки.

Таким образом, нами была показана применимость современных методов компьютерного моделирования не только для создания прототипа, пригодного для усовершенствования традиционными экспериментальными методами, но и для улучшения свойств существующих белков: при помощи алгоритма «Design» программы Rosetta удалось создать ряд новых мутантных вариантов липокалина B1c с увеличенными квантовым выходом флуоресценции и максимальной разницей между флуоресценцией свободного и связанного флуорогенов.

4.2.4.5. Использование улучшенных вариантов в качестве метки для флуоресцентной микроскопии

Для того чтобы проверить, что новые мутанты липокалина B1c, продемонстрировавшие улучшенные свойства *in vitro*, по-прежнему пригодны для использования в качестве флуоресцентной метки *in vivo* и при этом сохраняют свои преимущества, нами были созданы вектора, содержащие белки слияния белков цитоскелета (Actinin, Vimentin, LifeAct) с наиболее на наш взгляд перспективными мутантами из каждой из двух групп (из вариантов, полученных методом случайного мутагенеза, был взят 16912rnd1rnd1, из продуктов *in silico* мутагенеза — 16912_F44A и Qwd_F44A) (см. материалы и методы).

При временной трансфекции клеток линии HeLa Kyoto все проверенные белки продемонстрировали правильную локализацию и увеличенную яркость в сравнении со спектрально схожими исходными мутантами 16912 и A27C (рисунок 55). При этом, как и предполагалось, конструкции с белком 16912rnd1rnd1, имеющим низкую константу диссоциации, не обладали высокой флотабильностью.

Несмотря на то, что за счёт небольшого гипсохромного сдвига максимумов спектров возбуждения и эмиссии флуоресценции рассчитанное наибольшее возрастание флуоресцентного сигнала относительно раствора флуорогена комплекса Qwd_F44A+M739 было выше, чем у 16912_F44A+M739 (таблица 13), последний продемонстрировал гораздо большую яркость и контрастность в экспериментах *in vivo*. Скорее всего данный эффект связан с разницей в константах диссоциации комплексов этих двух мутантов (отличаются более чем на порядок).

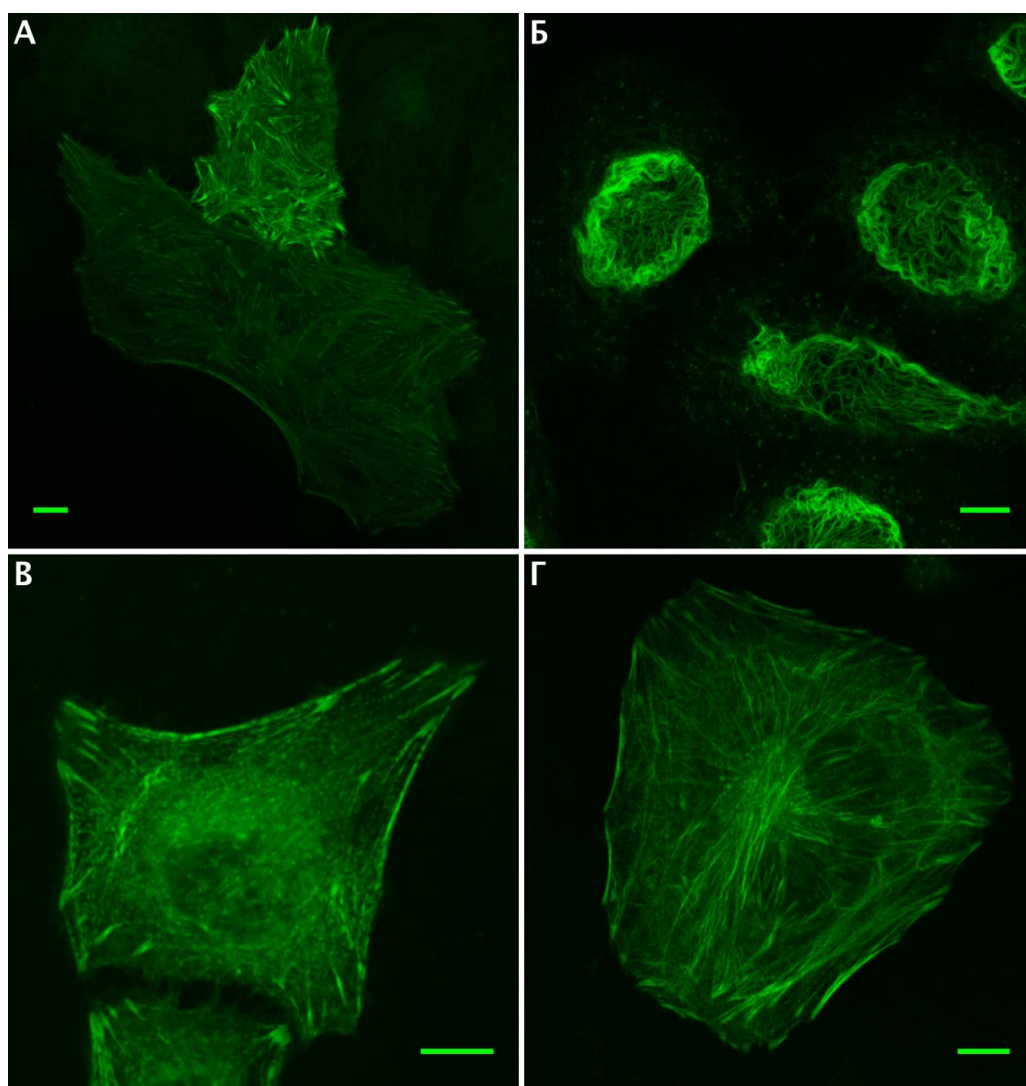


Рисунок 55. Флуоресцентная конфокальная микроскопия конструкций с мутантами белка B1c, полученными результате *in silico* и случайного мутагенеза, и флуорогеном M739 в живых клетках линии HeLa Kyoto; возбуждение 488 нм, эмиссия 520-560 нм; масштабная линейка 10 мкм. **А.** LifeAct-16912_F44A; **Б.** Vimentin-16912_F44A; **В.** Actinin-Qwd_F44A; **Г.** LifeAct-16912rnd1rnd1.

Проявивший себя лучшим образом в опытах *in vivo* белок 16912_F44A был также сравнён нашими коллегами из Нижнего Новгорода с предыдущими вариантами (A27C и 78776) в экспериментах по локализационной микроскопии сверхвысокого разрешения. Их данные тоже подтвердили превосходство улучшенного при помощи *in silico* моделирования мутанта.

Таким образом, применив два различных подхода, нам удалось создать ряд новых мутантов липокалина B1c, связывающих GFP-подобные флуорогены с образованием ещё более ярких комплексов и характеризующихся широким (более двух порядков) разнообразием констант связывания. Один из них — вариант 16912_F44A — продемонстрировал значительное превосходство над всеми ранее полученными мутантами во всех доступных нам видах флуоресцентной микроскопии.

4.2.5. Дальнекрасные флуорогены

Очень многие вещества (такие как вода, гемоглобин, оксигемоглобин, меланин, флавины и т.д.), являющиеся жизненно важными для живых организмов, поглощают свет в видимой или близкой к видимой областях спектра. Это свойство живого является одной из проблем флуоресцентной микроскопии, так как существенно ограничивает глубину, на которую может проникнуть свет определенной длины волны и, соответственно, затрудняет применение ряда флуоресцентных меток для исследования крупных интактных организмов. Однако существует небольшой промежуток — примерно между 700 и 900 нм (так называемое окно прозрачности биологических объектов), в котором поглощение света биологическими объектами сравнительно мало (рисунок 56). Флуоресцентные метки, возбуждающиеся светом и испускающие свет в пределах или близко к окну прозрачности, являются очень востребованными.

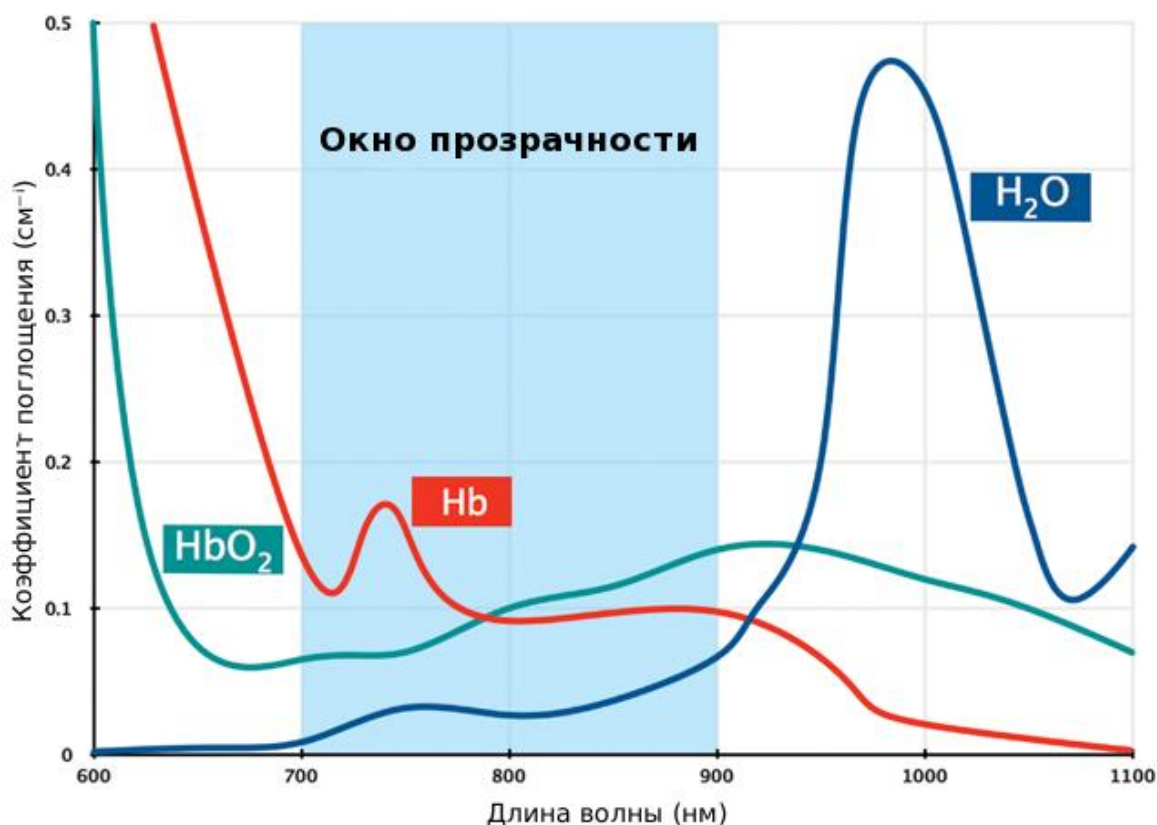


Рисунок 56. Спектры поглощения основных “поглотителей” биологических тканей: гемоглобина (Hb), оксигемоглобина (HbO₂) и воды (H₂O), и окно прозрачности (700-900 нм) [213].

Возбуждение и эмиссия флуоресценции используемых нами в работе флуорогенов могут быть смещены в красную область спектра за счёт добавления заместителей, увеличивающих систему сопряжённых связей соединения. По нашей просьбе сотрудники группы синтеза природных соединений ИБХ РАН синтезировали ряд подобных веществ (рисунок 57).

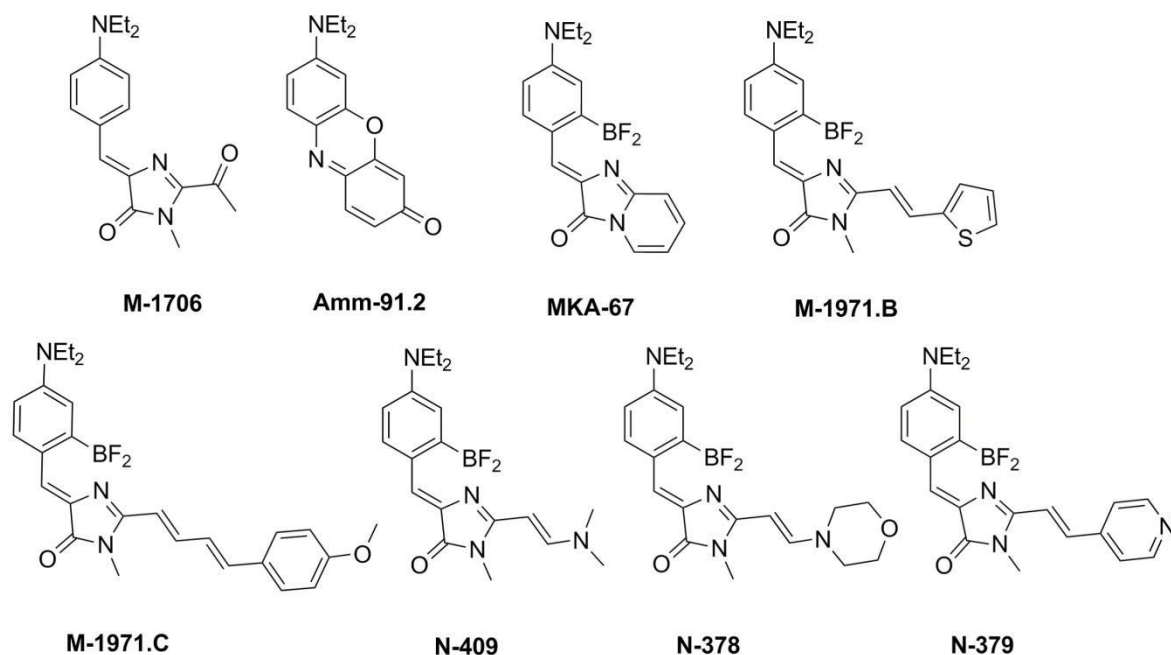


Рисунок 57. Структуры соединений-производных флуорогена M739, содержащих заместители, увеличивающие систему сопряжённых связей.

Так как в наших руках уже было достаточно большое количество разнообразных конструкций H2B-TagBFP-*<мутант белка Blc>*, а основной целью является получение метки, пригодной для флуоресцентного мечения в живых системах, скрининг новой библиотеки флуорогенов было решено провести сразу *in vivo*. Клетки линии НЕК293Т трансфицировали вышеупомянутыми векторами и анализировали через 48 часов при помощи флуоресцентного микроскопа Keyence BZ-9000, используя фильтры DAPI-BP (377/50, 447/60), TexasRed (562/40, 624/40) и Cy5 (628/40 и 692/40), добавляя перед съёмкой в среду исследуемые флуорогены. Для каждой пары белок-флуороген оценивали наличие и контрастность флуоресцентного сигнала (средняя интенсивность флуоресценции в области ядра трансфицированных клеток (определялись по синему сигналу от TagBFP), делённая на среднюю интенсивность сигнала нетрансфицированных клеток) в красном и дальнекрасном каналах сразу после добавления лиганда и через полчаса.

Возрастание флуоресценции было отмечено для двух флуорогенов — N379 и MKA-67. При этом в первом случае флуоресценция детектировалась лишь в красном (TexasRed) канале, тогда как сигнал от добавления MKA-67 можно было видеть в обоих (TexasRed и Cy5) (рисунок 58АБ), что больше соответствовало нашим целям.

Значимое увеличение флуоресценции флуорогена MKA-67 демонстрировали клетки с белками 16912, A27C, L132Q и 16912_F44A, при этом наибольшей контрастностью (около 10 раз) обладал 16912. Однако последующие *in vitro* эксперименты показали, что увеличение квантового

выхода флуоресценции при взаимодействии 16912 с МКА-67 не максимально даже среди имеющихся и проверенных мутантов (а значит есть потенциал для улучшения). Но, видимо, как и в случае пар Qwd_F44A+M739 и 16912_F44A+M739, разница в константах диссоциации комплексов снова сыграла решающую роль (константа диссоциации комплекса 16912+МКА-67 составляет 20 нМ, что является самым высокоаффинным взаимодействием, полученным в ходе данной работы).

Мы проверили новую пару 16912 + МКА-67 на более сложной, чем гистоны, модели. Временная трансфекция клеток линии HeLa Kyoto вектором Keratin-16912 с последующим добавлением флуорогена показала, что, несмотря на не очень высокую контрастность, структура цитоскелета хорошо видна, и данный комплекс может быть успешно использован для микроскопии (рисунок 58В).

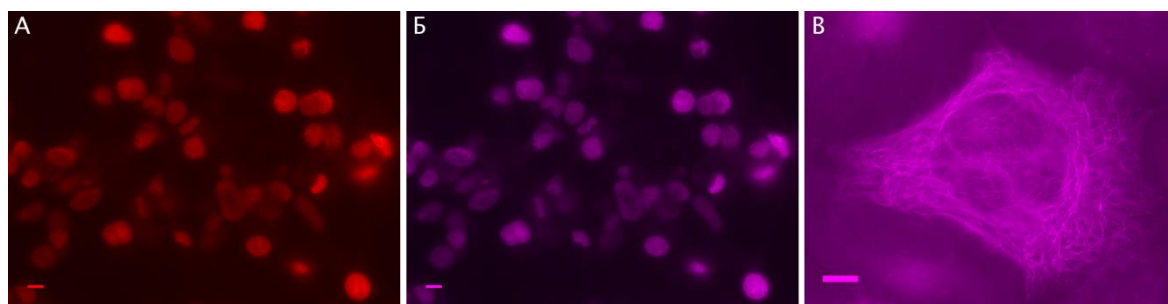


Рисунок 58. Флуоресцентная широкопольная микроскопия конструкций с флуорогеном МКА-67, масштабная линейка 10 мкм. **А.** Живые клетки линии HEK293Т, временно трансфицированные плазмидой H2В-TagBFP-16912, набор фильтров TexasRed (562/40, 624/40); **Б.** Живые клетки линии HEK293Т, временно трансфицированные плазмидой H2В-TagBFP-16912, набор фильтров Cy5 (628/40, 692/40); **В.** Живые клетки линии HeLa Kyoto, временно трансфицированные плазмидой Keratin-16912, набор фильтров Cy5 (628/40, 692/40).

В отличие от флуорогенов М739 и М774, которые дают яркий флуоресцентный сигнал немедленно после их добавления в среду, флуоресцентный сигнал от МКА-67, детектируемый в широкопольном флуоресцентном микроскопе, нарастал со временем, становясь хорошо видимым через 10-20 минут и достигая максимума примерно через час. Такая динамика может быть обусловлена как скоростью прохождения соединения через мембрану клеток, так и особенностью его взаимодействия с белком и требует дальнейшего изучения. Возможно, повторение всех стадий, проделанных для получения эффективного активирующего белка для GFP-подобного флуорогена М739, в сочетании с синтезом и исследованием библиотеки производных флуорогена МКА-67 приведет к созданию улучшенной дальнекрасной метки.

5. Заключение

Границы между науками начинают стираться, методы из одной области дополняют информацию, полученную в другой, позволяя делать ранее недоступные выводы и открытия. В данной работе сила междисциплинарных исследований проиллюстрирована на двух примерах: фундаментальном и прикладном.

В первой части работы, объединив классические биохимические подходы с анализом модельных химических соединений и компьютерным моделированием, нам удалось предложить структуру хромофора и объяснение различия спектральных свойств двух полученных родственных мономерных флуоресцентных белков FR66W-Yellow и FR66W-Orange, содержащих триптофан в составе хромофора и обладающих жёлтой (535 нм) и оранжевой (565 нм) эмиссией флуоресценции соответственно.

Наши данные могут послужить основой для рационального дизайна и получения улучшенных вариантов пока что мало изученных, но потенциально перспективных триптофан-содержащих флуоресцентных белков с эмиссией в жёлто-оранжевой области спектра.

Во второй части работы, сочетая современные компьютерные подходы с традиционными экспериментальными методами, мы разработали новый протокол поиска потенциальных флуороген-активирующих белков — перспективной и активно разрабатываемой группы флуоресцентных меток, объединяющих в себе достоинства генетически кодируемых белков и химических красителей.

Работа протокола была подтверждена созданием с его помощью новой системы флуоресцентного мечения белков, пригодной как для традиционной широкопольной или конфокальной флуоресцентной микроскопии, так и для микроскопии сверхвысокого разрешения, в том числе локализационной микроскопии сверхвысокого разрешения.

Разработанный протокол *in silico* скрининга и наличие двух имеющих разную природу компонентов должны облегчить как дальнейший процесс оптимизации и изменения параметров описанной системы за счёт комбинации независимых модификаций каждой из составных частей, так и создание новых пар на основе других белков и флуорогенов.

6. Выводы

1. На основе красного мономерного флуоресцентного белка FusionRed методами направленного и последующего случайного мутагенеза получены белки FR66W-Yellow и FR66W-Orange, содержащие триптофан в составе хромофора и обладающие соответственно жёлтой (535 нм) и оранжевой (565 нм) эмиссией.
2. На основе свойств полученных белков и модельного соединения, а также компьютерного моделирования предложены структуры хромофоров белков FR66W-Yellow и FR66W-Orange.
3. Разработаны протоколы *in silico* скрининга библиотек белков с известными структурами, а также библиотек близких мутантов белков с известными структурами для поиска потенциальных флуороген-активирующих белков.
4. При помощи разработанного протокола среди белков *E. coli* с известной структурой найдены и охарактеризованы *in vitro* белки ЗНО2 (β -кетоацил-АЦФ-синтаза II) и 1PVS (3-метиладенин-ДНК-гликозилаза II), способные связывать Kaede-подобный флуороген A12H с субмикромольной аффинностью и вызывать увеличение яркости его флуоресценции более чем на два порядка.
5. При помощи разработанного протокола на основе липокалина Blc из *E. coli* созданы и охарактеризованы три флуороген-активирующих белка (16912, 78776 и A27C), связывающие конформационно фиксированное производное хромофора GFP (M739). Показана возможность их применения для флуоресцентного мечения белков в клетках млекопитающих используя методы широкопольной и конфокальной микроскопии, а также микроскопии сверхвысокого разрешения.
6. Методами случайного мутагенеза и *in silico* мутагенеза при помощи программы Rosetta получена и протестирована *in vivo* панель улучшенных флуороген-активирующих белков, характеризующихся увеличенной аффинностью и/или квантовым выходом флуоресценции.
7. Произведен скрининг созданной на основе флуорогена M739 библиотеки соединений с увеличенной системой сопряжённых связей, и обнаружены пары белок-флуороген, характеризующиеся спектрами возбуждения и эмиссии флуоресценции, сдвинутыми в красную область.

7. Список сокращений

Ac — ацетил

Bu — бутил

CASP (The Critical Assessment of protein Structure Prediction) — критическая оценка предсказания белковых структур

CH₃CN — ацетонитрил

DMSO — диметилсульфоксид

EtOAc — этилацетат

EtOH — этанол

Et — этил

FMN (flavin mononucleotide) — флавиномононуклеотид

FQY (fluorescence quantum yield) — квантовый выход флуоресценции

FRET (Förster resonance energy transfer) — фёрстеровский резонансный перенос энергии

GFP (green fluorescent protein) — зелёный флуоресцентный белок

hAGT — O⁶-алкилгуанин-ДНК-алкилтрансфераза

iBu — изобутил

iPr — изопропил

K_d — константа диссоциации

MUPAC (multiple patch cloning) — множественное лоскутное клонирование

NRR (nonhomologous random recombination) — негомологичная случайная рекомбинация

PBS (phosphate buffered saline) — натрий-фосфатный буфер

PDB (Protein Data Bank) — банк данных 3D структур белков и нуклеиновых кислот (www.rcsb.org)

Ph — фенил

Pr — пропи́л

RACHITT (random chimeragenesis on transient template) — случайный химерогенез на временной матрице

REAP (reconstructed evolutionary adaptive paths) — реконструированный путь адаптивной эволюции

SDS (sodium dodecyl sulfate) — додецилсульфат натрия

SeSaM (sequence saturation mutagenesis) — насыщающий мутагенез

SIM (structured illumination microscopy) — микроскопия структурированного освещения

SiR — кремний-родамин

SMLM (single-molecule localization microscopy) — локализационная микроскопия одиночных молекул

STED (Stimulated Emission Depletion Microscopy) — микроскопия истощения флуоресценции вынужденным излучением

StEP (staggered extension process) — метод колеблющегося удлинения

tBu — третбутил

Y-FAST (Yellow Fluorescence-Activating and absorption-Shifting Tag) — жёлтый активирующий флуоресценцию и сдвигающий абсорбцию таг

АЦФ — ацилтрансфераза

ДНК — дезоксирибонуклеиновая кислота

ПЦР — полимеразная цепная реакция

РНК — рибонуклеиновая кислота

ТАЕ — трис-ацетатный буфер

УФ — ультрафиолет

тРНК — транспортная РНК

8. Список литературы

- [1] Herschel J.F.W. On a case of superficial colour presented by a homogeneous liquid internally colourless // *Phil. Trans. R. Soc. Lond.*– 1845.– Т. 135.– С. 143–145.
- [2] Stokes G.G. On the Change of Refrangibility of Light // *Phil. Trans. R. Soc. Lond.*– 1852.– Т. 142.– С. 463–562.
- [3] Meyer R. Über Einige Beziehungen Zwischen Fluoreszenz und Chemischer Constitution // *Festschrift der Herzoglichen Technischen Hochschule Carolo-Wilhelmina.*– 1897.– С. 167–205.
- [4] Kricka L.J., Fortina P. Analytical ancestry: “firsts” in fluorescent labeling of nucleosides, nucleotides, and nucleic acids // *Clin. Chem.*– 2009.– Т. 55, № 4.– С. 670–683.
- [5] Masters B.R. The development of fluorescence microscopy // *Elsevier Oceanogr. Ser.*– 2010.
- [6] Griffin B.A., Adams S.R., Tsien R.Y. Specific covalent labeling of recombinant protein molecules inside live cells // *Science.*– 1998.– Т. 281, № 5374.– С. 269–272.
- [7] Sahoo H. Fluorescent labeling techniques in biomolecules: a flashback // *RSC Adv.*– 2012.– Т. 2, № 18.– С. 7017.
- [8] Griffin B.A., Adams S.R., Jones J., Tsien R.Y. Fluorescent labeling of recombinant proteins in living cells with FAsH // *Methods Enzymol.*– 2000.– Т. 327.– С. 565–578.
- [9] Kapanidis A.N., Ebright Y.W., Ebright R.H. Site-specific incorporation of fluorescent probes into protein: hexahistidine-tag-mediated fluorescent labeling with (Ni^{2+} : nitrilotriacetic acid) n-fluorochrome conjugates // *J. Am. Chem. Soc.*– 2001.– Т. 123, № 48.– С. 12123–12125.
- [10] Guignet E.G., Hovius R., Vogel H. Reversible site-selective labeling of membrane proteins in live cells // *Nat. Biotechnol.*– 2004.– Т. 22, № 4.– С. 440–444.
- [11] Ojida A., Honda K., Shinmi D., Kiyonaka S., Mori Y., Hamachi I. Oligo-asp tag/Zn(II) complex probe as a new pair for labeling and fluorescence imaging of proteins // *J. Am. Chem. Soc.*– 2006.– Т. 128, № 32.– С. 10452–10459.
- [12] Hopp T.P., Prickett K.S., Price V.L., Libby R.T., March C.J., Cerretti D.P., Urdal D.L., Conlon P.J. A short polypeptide marker sequence useful for recombinant protein identification and purification // *Biotechnology.*– 1988.– Т. 6, № 10.– С. 1204–1210.
- [13] Reche P., Perham R.N. Structure and selectivity in post-translational modification: attaching the biotinyl-lysine and lipoyl-lysine swinging arms in multifunctional enzymes // *EMBO J.*– 1999.– Т. 18,

№ 10.– C. 2673–2682.

[14] Slavoff S.A., Chen I., Choi Y.-A., Ting A.Y. Expanding the substrate tolerance of biotin ligase through exploration of enzymes from diverse species // *J. Am. Chem. Soc.*– 2008.– T. 130, № 4.– C. 1160–1162.

[15] Uttamapinant C., White K.A., Baruah H., Thompson S., Fernández-Suárez M., Puthenveetil S., Ting A.Y. A fluorophore ligase for site-specific protein labeling inside living cells // *Proc. Natl. Acad. Sci.* – 2010.– T. 107, № 24.– C. 10914–10919.

[16] Liu D.S., Nivón L.G., Richter F., Goldman P.J., Deerinck T.J., Yao J.Z., Richardson D., Phipps W.S., Ye A.Z., Ellisman M.H., Drennan C.L., Baker D., Ting A.Y. Computational design of a red fluorophore ligase for site-specific protein labeling in living cells // *Proc. Natl. Acad. Sci.*– 2014.– T. 111, № 43.– C. E4551–E4559.

[17] Chen S., Chen Z.-J., Ren W., Ai H.-W. Reaction-based genetically encoded fluorescent hydrogen sulfide sensors // *J. Am. Chem. Soc.*– 2012.– T. 134, № 23.– C. 9589–9592.

[18] Shimomura O., Johnson F.H., Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea* // *J. Cell. Comp. Physiol.*– 1962.– T. 59.– C. 223–239.

[19] Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G., Cormier M.J. Primary structure of the *Aequorea victoria* green-fluorescent protein // *Gene*.– 1992.– T. 111, № 2.– C. 229–233.

[20] Chalfie M., Tu Y., Euskirchen G., Ward W.W., Prasher D.C. Green fluorescent protein as a marker for gene expression // *Science*.– 1994.– T. 263, № 5148.– C. 802–805.

[21] Inouye S., Tsuji F.I. *Aequorea* green fluorescent protein // *FEBS Lett.*– 1994.– T. 341, № 2-3.– C. 277–280.

[22] Rizzo M.A., Springer G.H., Granada B., Piston D.W. An improved cyan fluorescent protein variant useful for FRET // *Nat. Biotechnol.*– 2004.– T. 22, № 4.– C. 445–449.

[23] Goedhart J., van Weeren L., Hink M.A., Vischer N.O.E., Jalink K., Gadella T.W.J. Jr. Bright cyan fluorescent protein variants identified by fluorescence lifetime screening // *Nat. Methods*.– 2010.– T. 7, № 2.– C. 137–139.

[24] Goedhart J., von Stetten D., Noirclerc-Savoye M., Lelimousin M., Joosen L., Hink M.A., van Weeren L., Gadella T.W.J. Jr, Royant A. Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93% // *Nat. Commun.*– 2012.– T. 3.– C. 751.

- [25] Mena M.A., Treynor T.P., Mayo S.L., Daugherty P.S. Blue fluorescent proteins with enhanced brightness and photostability from a structurally targeted library // *Nat. Biotechnol.*— 2006.— T. 24, № 12.— C. 1569–1571.
- [26] Ai H.-W., Shaner N.C., Cheng Z., Tsien R.Y., Campbell R.E. Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins // *Biochemistry.*— 2007.— T. 46, № 20.— C. 5904–5910.
- [27] Subach O.M., Cranfill P.J., Davidson M.W., Verkhusha V.V. An enhanced monomeric blue fluorescent protein with the high chemical stability of the chromophore // *PLoS One.*— 2011.— T. 6, № 12.— C. e28674.
- [28] Tomosugi W., Matsuda T., Tani T., Nemoto T., Kotera I., Saito K., Horikawa K., Nagai T. An ultramarine fluorescent protein with increased photostability and pH insensitivity // *Nat. Methods.*— 2009.— T. 6, № 5.— C. 351–353.
- [29] Wachter R.M., Elsliger M.A., Kallio K., Hanson G.T., Remington S.J. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein // *Structure.*— 1998.— T. 6, № 10.— C. 1267–1277.
- [30] Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G., Markelov M.L., Lukyanov S.A. Fluorescent proteins from nonbioluminescent *Anthozoa* species // *Nat. Biotechnol.*— 1999.— T. 17, № 10.— C. 969–973.
- [31] Keppler A., Gendreizig S., Gronemeyer T., Pick H., Vogel H., Johnsson K. A general method for the covalent labeling of fusion proteins with small molecules in vivo // *Nat. Biotechnol.*— 2003.— T. 21, № 1.— C. 86–89.
- [32] Keppler A., Pick H., Arrivoli C., Vogel H., Johnsson K. Labeling of fusion proteins with synthetic fluorophores in live cells // *Proc. Natl. Acad. Sci.*— 2004.— T. 101, № 27.— C. 9955–9959.
- [33] Gronemeyer T., Chidley C., Juillerat A., Heinis C., Johnsson K. Directed evolution of O6-alkylguanine-DNA alkyltransferase for applications in protein labeling // *Protein Eng. Des. Sel.*— 2006.— T. 19, № 7.— C. 309–316.
- [34] Juillerat A., Heinis C., Sielaff I., Barnikow J., Jaccard H., Kunz B., Terskikh A., Johnsson K. Engineering substrate specificity of O6-alkylguanine-DNA alkyltransferase for specific protein labeling in living cells // *Chembiochem.*— 2005.— T. 6, № 7.— C. 1263–1269.
- [35] Gautier A., Juillerat A., Heinis C., Corrêa I.R. Jr, Kindermann M., Beaufile F., Johnsson K. An engineered protein tag for multiprotein labeling in living cells // *Chem. Biol.*— 2008.— T. 15, № 2.— C.

- [36] Los G.V., Encell L.P., McDougall M.G., Hartzell D.D., Karassina N., Zimprich C., Wood M.G., Learish R., Ohana R.F., Urh M., Simpson D., Mendez J., Zimmerman K., Otto P., Vidugiris G., Zhu J., Darzins A., Klaubert D.H., Bulleit R.F., Wood K.V. HaloTag: a novel protein labeling technology for cell imaging and protein analysis // *ACS Chem. Biol.*— 2008.— T. 3, № 6.— C. 373–382.
- [37] Drepper T., Eggert T., Circolone F., Heck A., Krauss U., Guterl J.-K., Wendorff M., Losi A., Gärtner W., Jaeger K.-E. Reporter proteins for in vivo fluorescence without oxygen // *Nat. Biotechnol.*— 2007.— T. 25, № 4.— C. 443–445.
- [38] Chapman S., Faulkner C., Kaiserli E., Garcia-Mata C., Savenkov E.I., Roberts A.G., Oparka K.J., Christie J.M. The photoreversible fluorescent protein iLOV outperforms GFP as a reporter of plant virus infection // *Proc. Natl. Acad. Sci.*— 2008.— T. 105, № 50.— C. 20038–20043.
- [39] Shu X., Lev-Ram V., Deerinck T.J., Qi Y., Ramko E.B., Davidson M.W., Jin Y., Ellisman M.H., Tsien R.Y. A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms // *PLoS Biol.*— 2011.— T. 9, № 4.— C. e1001041.
- [40] Mukherjee A., Walker J., Weyant K.B., Schroeder C.M. Characterization of flavin-based fluorescent proteins: an emerging class of fluorescent reporters // *PLoS One.*— 2013.— T. 8, № 5.— C. e64753.
- [41] Wingen M., Potzkei J., Endres S., Casini G., Rupprecht C., Fahlke C., Krauss U., Jaeger K.-E., Drepper T., Gensch T. The photophysics of LOV-based fluorescent proteins--new tools for cell biology // *Photochem. Photobiol. Sci.*— 2014.— T. 13, № 6.— C. 875–883.
- [42] Kumagai A., Ando R., Miyatake H., Greimel P., Kobayashi T., Hirabayashi Y., Shimogori T., Miyawaki A. A bilirubin-inducible fluorescent protein from eel muscle // *Cell.*— 2013.— T. 153, № 7.— C. 1602–1611.
- [43] Shu X., Royant A., Lin M.Z., Aguilera T.A., Lev-Ram V., Steinbach P.A., Tsien R.Y. Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome // *Science.*— 2009.— T. 324, № 5928.— C. 804–807.
- [44] Filonov G.S., Piatkevich K.D., Ting L.-M., Zhang J., Kim K., Verkhusha V.V. Bright and stable near-infrared fluorescent protein for in vivo imaging // *Nat. Biotechnol.*— 2011.— T. 29, № 8.— C. 757–761.
- [45] Auldridge M.E., Satyshur K.A., Anstrom D.M., Forest K.T. Structure-guided engineering enhances a phytochrome-based infrared fluorescent protein // *J. Biol. Chem.*— 2012.— T. 287, № 10.— C.

7000–7009.

- [46] Bhattacharya S., Auldridge M.E., Lehtivuori H., Ihalaenen J.A., Forest K.T. Origins of fluorescence in evolved bacteriophytochromes // *J. Biol. Chem.*– 2014.– T. 289, № 46.– C. 32144–32152.
- [47] Szent-Gyorgyi C., Schmidt B.F., Schmidt B.A., Creeger Y., Fisher G.W., Zakel K.L., Adler S., Fitzpatrick J.A.J., Woolford C.A., Yan Q., Vasilev K.V., Berget P.B., Bruchez M.P., Jarvik J.W., Waggoner A. Fluorogen-activating single-chain antibodies for imaging cell surface proteins // *Nat. Biotechnol.*– 2008.– T. 26, № 2.– C. 235–240.
- [48] Yan Q., Schwartz S.L., Maji S., Huang F., Szent-Gyorgyi C., Lidke D.S., Lidke K.A., Bruchez M.P. Localization microscopy using noncovalent fluorogen activation by genetically encoded fluorogen-activating proteins // *Chemphyschem.*– 2014.– T. 15, № 4.– C. 687–695.
- [49] Zhang M., Chakraborty S.K., Sampath P., Rojas J.J., Hou W., Saurabh S., Thorne S.H., Bruchez M.P., Waggoner A.S. Fluoromodule-based reporter/probes designed for in vivo fluorescence imaging // *J. Clin. Invest.*– 2015.– T. 125, № 10.– C. 3915–3927.
- [50] Plamont M.-A., Billon-Denis E., Maurin S., Gauron C., Pimenta F.M., Specht C.G., Shi J., Quérard J., Pan B., Rossignol J., Moncoq K., Morellet N., Volovitch M., Lescop E., Chen Y., Triller A., Vríz S., Le Saux T., Jullien L., Gautier A. Small fluorescence-activating and absorption-shifting tag for tunable protein imaging in vivo // *Proc. Natl. Acad. Sci.*– 2016.– T. 113, № 3.– C. 497–502.
- [51] McRee D.E., Tainer J.A., Meyer T.E., Van Beeumen J., Cusanovich M.A., Getzoff E.D. Crystallographic structure of a photoreceptor protein at 2.4 Å resolution // *Proc. Natl. Acad. Sci.*– 1989.– T. 86, № 17.– C. 6533–6537.
- [52] Lepthien S., Hoesl M.G., Merkel L., Budisa N. Azatryptophans endow proteins with intrinsic blue fluorescence // *Proc. Natl. Acad. Sci.*– 2008.– T. 105, № 42.– C. 16095–16100.
- [53] Pantoja R., Rodriguez E.A., Dibas M.I., Dougherty D.A., Lester H.A. Single-molecule imaging of a fluorescent unnatural amino acid incorporated into nicotinic receptors // *Biophys. J.*– 2009.– T. 96, № 1.– C. 226–237.
- [54] Liu C.C., Schultz P.G. Adding new chemistries to the genetic code // *Annu. Rev. Biochem.*– 2010.– T. 79.– C. 413–444.
- [55] Partis M.D., Griffiths D.G., Roberts G.C., Beechey R.B. Cross-linking of protein by ω -maleimido alkanoylN-hydroxysuccinimido esters // *J. Protein Chem.*– 1983.– T. 2, № 3.– C. 263–277.

- [56] Lomant A.J., Fairbanks G. Chemical probes of extended biological structures: synthesis and properties of the cleavable protein cross-linking reagent [³⁵S]dithiobis(succinimidyl propionate) // J. Mol. Biol.– 1976.– T. 104, № 1.– C. 243–261.
- [57] Baskin J.M., Bertozzi C.R. Bioorthogonal click chemistry: covalent labeling in living systems // QSAR Comb. Sci.– 2007.– T. 26, № 11-12.– C. 1211–1219.
- [58] Lukinavičius G., Reymond L., D’Este E., Masharina A., Göttfert F., Ta H., Güther A., Fournier M., Rizzo S., Waldmann H., Blaukopf C., Sommer C., Gerlich D.W., Arndt H.-D., Hell S.W., Johnsson K. Fluorogenic probes for live-cell imaging of the cytoskeleton // Nat. Methods.– 2014.– T. 11, № 7.– C. 731–733.
- [59] Myers R.M., Lerman L.S., Maniatis T. A general method for saturation mutagenesis of cloned DNA fragments // Science.– 1985.– T. 229, № 4710.– C. 242–247.
- [60] Lai Y.-P., Huang J., Wang L.-F., Li J., Wu Z.-R. A new approach to random mutagenesis in vitro // Biotechnol. Bioeng.– 2004.– T. 86, № 6.– C. 622–627.
- [61] Miller J.H. Mutagenic specificity of ultraviolet light // J. Mol. Biol.– 1985.– T. 182, № 1.– C. 45–65.
- [62] Scheuermann R., Tam S., Burgers P.M., Lu C., Echols H. Identification of the epsilon-subunit of Escherichia coli DNA polymerase III holoenzyme as the dnaQ gene product: a fidelity subunit for DNA replication // Proc. Natl. Acad. Sci. – 1983.– T. 80, № 23.– C. 7085–7089.
- [63] Cox E.C. Bacterial mutator genes and the control of spontaneous mutation // Annu. Rev. Genet.– 1976.– T. 10.– C. 135–156.
- [64] Maki H., Sekiguchi M. MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis // Nature.– 1992.– T. 355, № 6357.– C. 273–275.
- [65] Greener A., Callahan M., Jerpseth B. An efficient random mutagenesis technique using an E. coli mutator strain // Mol. Biotechnol.– 1997.– T. 7, № 2.– C. 189–195.
- [66] Ravikumar A., Arrieta A., Liu C.C. An orthogonal DNA replication system in yeast // Nat. Chem. Biol.– 2014.– T. 10, № 3.– C. 175–177.
- [67] Crook N., Abatemarco J., Sun J., Wagner J.M., Schmitz A., Alper H.S. In vivo continuous evolution of genes and pathways in yeast // Nat. Commun.– 2016.– T. 7.– C. 13051.
- [68] Hughes S.R., Gibbons W.R., Bang S.S., Pinkelman R., Bischoff K.M., Slininger P.J., Qureshi N., Kurtzman C.P., Liu S., Saha B.C., Jackson J.S., Cotta M.A., Rich J.O., Javers J.E. Random UV-C

mutagenesis of *Scheffersomyces* (formerly *Pichia*) *stipitis* NRRL Y-7124 to improve anaerobic growth on lignocellulosic sugars // J. Ind. Microbiol. Biotechnol.– 2012.– T. 39, № 1.– C. 163–173.

[69] Badran A.H., Liu D.R. Development of potent in vivo mutagenesis plasmids with broad mutational spectra // Nat. Commun.– 2015.– T. 6.– C. 8425.

[70] Eckert K.A., Kunkel T.A. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase // Nucleic Acids Res.– 1990.– T. 18, № 13.– C. 3739–3744.

[71] Beckman R.A., Mildvan A.S., Loeb L.A. On the fidelity of DNA replication: manganese mutagenesis in vitro // Biochemistry.– 1985.– T. 24, № 21.– C. 5810–5817.

[72] Zaccolo M., Williams D.M., Brown D.M., Gherardi E. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues // J. Mol. Biol.– 1996.– T. 255, № 4.– C. 589–603.

[73] Ma X., Ke T., Mao P., Jin X., Ma L., He G. The mutagenic properties of BrdUTP in a random mutagenesis process // Mol. Biol. Rep.– 2008.– T. 35, № 4.– C. 663–667.

[74] Kamiya H., Ito M., Harashima H. Induction of transition and transversion mutations during random mutagenesis PCR by the addition of 2-hydroxy-dATP // Biol. Pharm. Bull.– 2004.– T. 27, № 5.– C. 621–623.

[75] Cadwell R.C., Joyce G.F. Randomization of genes by PCR mutagenesis // PCR Methods Appl.– 1992.– T. 2, № 1.– C. 28–33.

[76] Vanhercke T., Ampe C., Tirry L., Denolf P. Reducing mutational bias in random protein libraries // Anal. Biochem.– 2005.– T. 339, № 1.– C. 9–14.

[77] Wong T.S., Tee K.L., Hauer B., Schwaneberg U. Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution // Nucleic Acids Res.– 2004.– T. 32, № 3.– C. e26.

[78] Gish G., Eckstein F. DNA and RNA sequence determination based on phosphorothioate chemistry // Science.– 1988.– T. 240, № 4858.– C. 1520–1522.

[79] Loakes D. Survey and summary: The applications of universal DNA base analogues // Nucleic Acids Res.– 2001.– T. 29, № 12.– C. 2437–2447.

[80] Steffens D.L., Williams J.G.K. Efficient site-directed saturation mutagenesis using degenerate oligonucleotides // J. Biomol. Tech.– 2007.– T. 18, № 3.– C. 147–149.

[81] Scott J.K., Smith G.P. Searching for peptide ligands with an epitope library // Science.– 1990.– T. 249, № 4967.– C. 386–390.

- [82] Lam K.S., Salmon S.E., Hersh E.M., Hruby V.J., Kazmierski W.M., Knapp R.J. A new type of synthetic peptide library for identifying ligand-binding activity // *Nature*.– 1991.– T. 354, № 6348.– C. 82–84.
- [83] Neuner P., Cortese R., Monaci P. Codon-based mutagenesis using dimer-phosphoramidites // *Nucleic Acids Res.*– 1998.– T. 26, № 5.– C. 1223–1227.
- [84] Virnekäs B., Ge L., Plückthun A., Schneider K.C., Wellnhofer G., Moroney S.E. Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis // *Nucleic Acids Res.*– 1994.– T. 22, № 25.– C. 5600–5607.
- [85] Dennig A., Shivange A.V., Marienhagen J., Schwaneberg U. OmniChange: the sequence independent method for simultaneous site-saturation of five codons // *PLoS One*.– 2011.– T. 6, № 10.– C. e26222.
- [86] Tang L., Gao H., Zhu X., Wang X., Zhou M., Jiang R. Construction of “small-intelligent” focused mutagenesis libraries using well-designed combinatorial degenerate primers // *Biotechniques*.– 2012.– T. 52, № 3.– C. 149–158.
- [87] Taniguchi N., Nakayama S., Kawakami T., Murakami H. Patch cloning method for multiple site-directed and saturation mutagenesis // *BMC Biotechnol.*– 2013.– T. 13.– C. 91.
- [88] Gaytán P., Roldán-Salgado A. Elimination of redundant and stop codons during the chemical synthesis of degenerate oligonucleotides. Combinatorial testing on the chromophore region of the red fluorescent protein mKate // *ACS Synth. Biol.*– 2013.– T. 2, № 8.– C. 453–462.
- [89] Stemmer W.P. Rapid evolution of a protein in vitro by DNA shuffling // *Nature*.– 1994.– T. 370, № 6488.– C. 389–391.
- [90] Coco W.M., Levinson W.E., Crist M.J., Hektor H.J., Darzins A., Pienkos P.T., Squires C.H., Monticello D.J. DNA shuffling method for generating highly recombined genes and evolved enzymes // *Nat. Biotechnol.*– 2001.– T. 19, № 4.– C. 354–359.
- [91] Müller K.M., Stebel S.C., Knall S., Zipf G., Bernauer H.S., Arndt K.M. Nucleotide exchange and excision technology (NExT) DNA shuffling: a robust method for DNA fragmentation and directed evolution // *Nucleic Acids Res.*– 2005.– T. 33, № 13.– C. e117.
- [92] Ness J.E., Kim S., Gottman A., Pak R., Krebber A., Borchert T.V., Govindarajan S., Mundorff E.C., Minshull J. Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently // *Nat. Biotechnol.*– 2002.– T. 20, № 12.– C. 1251–1255.

- [93] Zha D., Eipper A., Reetz M.T. Assembly of designed oligonucleotides as an efficient method for gene recombination: a new tool in directed evolution // *Chembiochem.*– 2003.– T. 4, № 1.– C. 34–39.
- [94] Zhao H., Giver L., Shao Z., Affholter J.A., Arnold F.H. Molecular evolution by staggered extension process (StEP) in vitro recombination // *Nat. Biotechnol.*– 1998.– T. 16, № 3.– C. 258–261.
- [95] Packer M.S., Liu D.R. Methods for the directed evolution of proteins // *Nat. Rev. Genet.*– 2015.– T. 16, № 7.– C. 379–394.
- [96] Sieber V., Martinez C.A., Arnold F.H. Libraries of hybrid proteins from distantly related sequences // *Nat. Biotechnol.*– 2001.– T. 19, № 5.– C. 456–460.
- [97] Ostermeier M., Shim J.H., Benkovic S.J. A combinatorial approach to hybrid enzymes independent of DNA homology // *Nat. Biotechnol.*– 1999.– T. 17, № 12.– C. 1205–1209.
- [98] Bittker J.A., Le B.V., Liu D.R. Nucleic acid evolution and minimization by nonhomologous random recombination // *Nat. Biotechnol.*– 2002.– T. 20, № 10.– C. 1024–1029.
- [99] Romanini D.W., Peralta-Yahya P., Mondol V., Cornish V.W. A heritable recombination system for synthetic Darwinian evolution in yeast // *ACS Synth. Biol.*– 2012.– T. 1, № 12.– C. 602–609.
- [100] Schweitzer B.A., Kool E.T. Aromatic nonpolar nucleosides as hydrophobic isosteres of pyrimidine and purine nucleosides // *J. Org. Chem.*– 1994.– T. 59, № 24.– C. 7238–7242.
- [101] Guckian K.M., Morales J.C., Kool E.T. Structure and base pairing properties of a replicable nonpolar isostere for deoxyadenosine // *J. Org. Chem.*– 1998.– T. 63, № 26.– C. 9652–9656.
- [102] Guckian K.M., Schweitzer B.A., Ren R.X.-F., Sheils C.J., Paris P.L., Tahmassebi D.C., Kool E.T. Experimental measurement of aromatic stacking affinities in the context of duplex DNA // *J. Am. Chem. Soc.*– 1996.– T. 118, № 34.– C. 8182–8183.
- [103] Moran S., Ren R.X., Kool E.T. A thymidine triphosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence selectivity // *Proc. Natl. Acad. Sci.* – 1997.– T. 94, № 20.– C. 10506–10511.
- [104] Kool E.T. Hydrogen bonding, base stacking, and steric effects in DNA replication // *Annu. Rev. Biophys. Biomol. Struct.*– 2001.– T. 30.– C. 1–22.
- [105] Matthew Robert Drake Using chemical synthesis to investigate protein glycosylation: O-mannosylation site specifically modulates the stability and cellulose binding affinity of family 1 carbohydrate binding modules.– University of Colorado, Boulder, 2014.
- [106] Abeywardana T., Pratt M.R. Using chemistry to investigate the molecular consequences of

protein ubiquitylation // *Chembiochem.*— 2014.— T. 15, № 11.— C. 1547–1554.

[107] Boerema D.J., Tereshko V.A., Zhang J., Kent S.B.H. Chemical synthesis and enzymatic properties of RNase A analogues designed to enhance second-step catalytic activity // *Org. Biomol. Chem.*— 2016.— T. 14, № 37.— C. 8804–8814.

[108] Nielsen M.B., Andersen L.H., Rocha-Rinza T. Absorption tuning of the green fluorescent protein chromophore: synthesis and studies of model compounds // *Monatsh. Chem.*— 2011.— T. 142, № 7.— C. 709–715.

[109] Lincke K., Sølling T., Andersen L.H., Klærke B., Rahbek D.B., Rajput J., Oehlenschläger C.B., Petersen M.Å., Nielsen M.B. On the absorption of the phenolate chromophore in the green fluorescent protein—role of individual interactions // *Chem. Commun.*— 2010.— T. 46, № 5.— C. 734–736.

[110] Laskowski R.A., Swindells M.B. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery // *J. Chem. Inf. Model.*— 2011.— T. 51, № 10.— C. 2778–2786.

[111] Wallace A.C., Laskowski R.A., Thornton J.M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions // *Protein Eng.*— 1995.— T. 8, № 2.— C. 127–134.

[112] Anand P., Nagarajan D., Mukherjee S., Chandra N. PLIC: protein-ligand interaction clusters // *Database.*— 2014.— T. 2014, № 0.— C. bau029.

[113] Rarey M., Kramer B., Lengauer T., Klebe G. A fast flexible docking method using an incremental construction algorithm // *J. Mol. Biol.*— 1996.— T. 261, № 3.— C. 470–489.

[114] Friesner R.A., Banks J.L., Murphy R.B., Halgren T.A., Klicic J.J., Mainz D.T., Repasky M.P., Knoll E.H., Shelley M., Perry J.K., Shaw D.E., Francis P., Shenkin P.S. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy // *J. Med. Chem.*— 2004.— T. 47, № 7.— C. 1739–1749.

[115] Halgren T.A., Murphy R.B., Friesner R.A., Beard H.S., Frye L.L., Pollard W.T., Banks J.L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening // *J. Med. Chem.*— 2004.— T. 47, № 7.— C. 1750–1759.

[116] Sousa S.F., Fernandes P.A., Ramos M.J. Protein-ligand docking: current status and future challenges // *Proteins.*— 2006.— T. 65, № 1.— C. 15–26.

[117] Morris G.M., Goodsell D.S., Halliday R.S., Huey R., Hart W.E., Belew R.K., Olson A.J., Others. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function // *J. Comput. Chem.*— 1998.— T. 19, № 14.— C. 1639–1662.

- [118] Morris G.M., Huey R., Lindstrom W., Sanner M.F., Belew R.K., Goodsell D.S., Olson A.J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility // J. Comput. Chem.– 2009.– T. 30, № 16.– C. 2785–2791.
- [119] Trott O., Olson A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading // J. Comput. Chem.– 2010.– T. 31, № 2.– C. 455–461.
- [120] Lyskov S., Chou F.-C., Conchúir S.Ó., Der B.S., Drew K., Kuroda D., Xu J., Weitzner B.D., Renfrew P.D., Sripakdeevong P., Borgo B., Havranek J.J., Kuhlman B., Kortemme T., Bonneau R., Gray J.J., Das R. Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE) // PLoS One.– 2013.– T. 8, № 5.– C. e63906.
- [121] Combs S.A., Deluca S.L., Deluca S.H., Lemmon G.H., Nannemann D.P., Nguyen E.D., Willis J.R., Sheehan J.H., Meiler J. Small-molecule ligand docking into comparative models with Rosetta // Nat. Protoc.– 2013.– T. 8, № 7.– C. 1277–1298.
- [122] DeLuca S., Khar K., Meiler J. Fully flexible docking of medium sized ligand libraries with rosettaligand // PLoS One.– 2015.– T. 10, № 7.– C. e0132508.
- [123] Rauf A., Khan R., Raza M., Khan H., Pervez S., De Feo V., Maione F., Mascolo N. Suppression of inflammatory response by chrysin, a flavone isolated from *Potentilla evestita* Th. Wolf. *In silico* predictive study on its mechanistic effect // Fitoterapia.– 2015.– T. 103.– C. 129–135.
- [124] Dundas J., Ouyang Z., Tseng J., Binkowski A., Turpaz Y., Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues // Nucleic Acids Res.– 2006.– T. 34.– C. W116–W118.
- [125] Binkowski T.A., Naghibzadeh S., Liang J. CASTp: Computed Atlas of Surface Topography of proteins // Nucleic Acids Res.– 2003.– T. 31, № 13.– C. 3352–3355.
- [126] Brylinski M., Skolnick J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation // Proc. Natl. Acad. Sci. – 2008.– T. 105, № 1.– C. 129–134.
- [127] Skolnick J., Brylinski M. FINDSITE: a combined evolution/structure-based approach to protein function prediction // Brief. Bioinform.– 2009.– T. 10, № 4.– C. 378–391.
- [128] Hernandez M., Ghersi D., Sanchez R. SITEHOUND-web: a server for ligand binding site identification in protein structures // Nucleic Acids Res.– 2009.– T. 37.– C. W413–W416.
- [129] Šali A., Blundell T.L. Comparative protein modelling by satisfaction of spatial restraints // J.

Mol. Biol.– 1993.– T. 234, № 3.– C. 779–815.

[130] Arnold K., Bordoli L., Kopp J., Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling // *Bioinformatics.*– 2006.– T. 22, № 2.– C. 195–201.

[131] Reetz M.T., Carballeira J.D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes // *Nat. Protoc.*– 2007.– T. 2, № 4.– C. 891–903.

[132] Guerois R., Nielsen J.E., Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations // *J. Mol. Biol.*– 2002.– T. 320, № 2.– C. 369–387.

[133] Valdar W.S.J. Scoring residue conservation // *Proteins.*– 2002.– T. 48, № 2.– C. 227–241.

[134] Ashkenazy H., Erez E., Martz E., Pupko T., Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids // *Nucleic Acids Res.*– 2010.– T. 38.– C. W529–W533.

[135] Jochens H., Bornscheuer U.T. Natural diversity to guide focused directed evolution // *Chembiochem.*– 2010.– T. 11, № 13.– C. 1861–1866.

[136] Engqvist M.K.M., Nielsen J. ANT: software for generating and evaluating degenerate codons for natural and expanded genetic codes // *ACS Synth. Biol.*– 2015.– T. 4, № 8.– C. 935–938.

[137] Nov Y. When second best is good enough: another probabilistic look at saturation mutagenesis // *Appl. Environ. Microbiol.*– 2012.– T. 78, № 1.– C. 258–262.

[138] Rath A., Davidson A.R. The design of a hyperstable mutant of the Abp1p SH3 domain by sequence alignment analysis // *Protein Sci.*– 2000.– T. 9, № 12.– C. 2457–2469.

[139] Lehmann M., Pasamontes L., Lassen S.F., Wyss M. The consensus concept for thermostability engineering of proteins // *Biochim. Biophys. Acta.*– 2000.– T. 1543, № 2.– C. 408–415.

[140] Cox V.E., Gaucher E.A. Engineering proteins by reconstructing evolutionary adaptive paths // *Methods Mol. Biol.*– 2014.– T. 1179.– C. 353–363.

[141] Voigt C.A., Martinez C., Wang Z.-G., Mayo S.L., Arnold F.H. Protein building blocks preserved by recombination // *Nat. Struct. Biol.*– 2002.– T. 9, № 7.– C. 553–558.

[142] Addington T.A., Mertz R.W., Siegel J.B., Thompson J.M., Fisher A.J., Filkov V., Fleischman N.M., Suen A.A., Zhang C., Toney M.D. Janus: prediction and ranking of mutations required for functional interconversion of enzymes // *J. Mol. Biol.*– 2013.– T. 425, № 8.– C. 1378–1389.

[143] Nosrati G.R., Houk K.N. SABER: a computational method for identifying active sites for new

reactions // *Protein Sci.*– 2012.– T. 21, № 5.– C. 697–706.

[144] Pratter S.M., Konstantinovic C., Di Giuro C.M.L., Leitner E., Kumar D., de Visser S.P., Grogan G., Straganz G.D. Inversion of enantioselectivity of a mononuclear non-heme iron(II)-dependent hydroxylase by tuning the interplay of metal-center geometry and protein structure // *Angew. Chem. Int. Ed Engl.*– 2013.– T. 52, № 37.– C. 9677–9681.

[145] Epstein C.J., Goldberger R.F., Anfinsen C.B. The genetic control of tertiary protein structure: studies with model systems // *Cold Spring Harb. Symp. Quant. Biol.*– 1963.– T. 28, № 0.– C. 439–449.

[146] Koga N., Tatsumi-Koga R., Liu G., Xiao R., Acton T.B., Montelione G.T., Baker D. Principles for designing ideal protein structures // *Nature.*– 2012.– T. 491, № 7423.– C. 222–227.

[147] Kries H., Blomberg R., Hilvert D. De novo enzymes by computational design // *Curr. Opin. Chem. Biol.*– 2013.– T. 17, № 2.– C. 221–228.

[148] Steiner K., Schwab H. Recent advances in rational approaches for enzyme engineering // *Comput. Struct. Biotechnol. J.*– 2012.– T. 2.– C. e201209010.

[149] Kuhlman B., Dantas G., Ireton G.C., Varani G., Stoddard B.L., Baker D. Design of a novel globular protein fold with atomic-level accuracy // *Science.*– 2003.– T. 302, № 5649.– C. 1364–1368.

[150] Tinberg C.E., Khare S.D., Dou J., Doyle L., Nelson J.W., Schena A., Jankowski W., Kalodimos C.G., Johnsson K., Stoddard B.L., Baker D. Computational design of ligand-binding proteins with high affinity and selectivity // *Nature.*– 2013.– T. 501, № 7466.– C. 212–216.

[151] Siegel J.B., Zanghellini A., Lovick H.M., Kiss G., Lambert A.R., St Clair J.L., Gallaher J.L., Hilvert D., Gelb M.H., Stoddard B.L., Houk K.N., Michael F.E., Baker D. Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction // *Science.*– 2010.– T. 329, № 5989.– C. 309–313.

[152] Röthlisberger D., Khersonsky O., Wollacott A.M., Jiang L., DeChancie J., Betker J., Gallaher J.L., Althoff E.A., Zanghellini A., Dym O., Albeck S., Houk K.N., Tawfik D.S., Baker D. Kemp elimination catalysts by computational enzyme design // *Nature.*– 2008.– T. 453, № 7192.– C. 190–195.

[153] Chen C.-Y., Georgiev I., Anderson A.C., Donald B.R. Computational structure-based redesign of enzyme activity // *Proc. Natl. Acad. Sci.* – 2009.– T. 106, № 10.– C. 3764–3769.

[154] Gainza P., Roberts K.E., Georgiev I., Lilien R.H., Keedy D.A., Chen C.-Y., Reza F., Anderson A.C., Richardson D.C., Richardson J.S., Donald B.R. osprey // *Methods in Protein Design.*– 2013.– T. 523.– C. 87–107.

- [155] Ho S.N., Hunt H.D., Horton R.M., Pullen J.K., Pease L.R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction // *Gene*.– 1989.– T. 77, № 1.– C. 51–59.
- [156] Matsumoto A., Itoh T.Q. Self-assembly cloning: a rapid construction method for recombinant molecules from multiple fragments // *Biotechniques*.– 2011.– T. 51, № 1.– C. 55–56.
- [157] Beyer H.M., Gonschorek P., Samodelov S.L., Meier M., Weber W., Zurbriggen M.D. AQUA cloning: a versatile and simple enzyme-free cloning approach // *PLoS One*.– 2015.– T. 10, № 9.– C. e0137652.
- [158] Laemmli U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4 // *Nature*.– 1970.– T. 227, № 5259.– C. 680–685.
- [159] Campanacci V., Nurizzo D., Spinelli S., Valencia C., Tegoni M., Cambillau C. The crystal structure of the Escherichia coli lipocalin Blc suggests a possible role in phospholipid binding // *FEBS Lett*.– 2004.– T. 562, № 1-3.– C. 183–188.
- [160] Schindelin J., Arganda-Carreras I., Frise E., Kaynig V., Longair M., Pietzsch T., Preibisch S., Rueden C., Saalfeld S., Schmid B., Tinevez J.-Y., White D.J., Hartenstein V., Eliceiri K., Tomancak P., Cardona A. Fiji: an open-source platform for biological-image analysis // *Nat. Methods*.– 2012.– T. 9, № 7.– C. 676–682.
- [161] Ovesný M., Křížek P., Borkovec J., Svindrych Z., Hagen G.M. ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging // *Bioinformatics*.– 2014.– T. 30, № 16.– C. 2389–2390.
- [162] Becke A.D. Density-functional thermochemistry. III. The role of exact exchange // *J. Chem. Phys*.– 1993.– T. 98, № 7.– C. 5648–5652.
- [163] Das A.K., Hasegawa J.-Y., Miyahara T., Ehara M., Nakatsuji H. Electronic excitations of the green fluorescent protein chromophore in its protonation states: SAC/SAC-CI study // *J. Comput. Chem*.– 2003.– T. 24, № 12.– C. 1421–1431.
- [164] Ma Y., Sun Q., Li Z., Yu J.-G., Smith S.C. Theoretical studies of chromophore maturation in the wild-type green fluorescent protein: ONIOM(DFT:MM) investigation of the mechanism of cyclization // *J. Phys. Chem. B*.– 2012.– T. 116, № 4.– C. 1426–1436.
- [165] Wanko M., García-Risueño P., Rubio A. Excited states of the green fluorescent protein chromophore: Performance of *ab initio* and semi-empirical methods // *Phys. Status Solidi*.– 2012.– T. 249, № 2.– C. 392–400.

- [166] Neese F. The ORCA program system // *WIREs Comput Mol Sci.*– 2012.– T. 2, № 1.– C. 73–78.
- [167] Nivón L.G., Moretti R., Baker D. A Pareto-optimal refinement method for protein design scaffolds // *PLoS One.*– 2013.– T. 8, № 4.– C. e59004.
- [168] Sarkisyan K.S., Yampolsky I.V., Solntsev K.M., Lukyanov S.A., Lukyanov K.A., Mishin A.S. Tryptophan-based chromophore in fluorescent proteins can be anionic // *Sci. Rep.*– 2012.– T. 2.– C. 608.
- [169] Sarkisyan K.S., Goryashchenko A.S., Lidsky P.V., Gorbachev D.A., Bozhanova N.G., Gorokhovatsky A.Y., Pereverzeva A.R., Ryumina A.P., Zherdeva V.V., Savitsky A.P., Solntsev K.M., Bommarius A.S., Sharonov G.V., Lindquist J.R., Drobizhev M., Hughes T.E., Rebane A., Lukyanov K.A., Mishin A.S. Green fluorescent protein with anionic tryptophan-based chromophore and long fluorescence lifetime // *Biophys. J.*– 2015.– T. 109, № 2.– C. 380–389.
- [170] Sarkisyan K.S., Zlobovskaya O.A., Gorbachev D.A., Bozhanova N.G., Sharonov G.V., Staroverov D.B., Egorov E.S., Ryabova A.V., Solntsev K.M., Mishin A.S., Lukyanov K.A. KillerOrange, a genetically encoded photosensitizer activated by blue and green light // *PLoS One.*– 2015.– T. 10, № 12.– C. e0145287.
- [171] Mérola F., Fredj A., Betolngar D.-B., Ziegler C., Erard M., Pasquier H. Newly engineered cyan fluorescent proteins with enhanced performances for live cell FRET imaging // *Biotechnol. J.*– 2014.– T. 9, № 2.– C. 180–191.
- [172] George Abraham B., Sarkisyan K.S., Mishin A.S., Santala V., Tkachenko N.V., Karp M. Fluorescent protein based FRET pairs with improved dynamic range for fluorescence lifetime measurements // *PLoS One.*– 2015.– T. 10, № 8.– C. e0134436.
- [173] Shaner N.C., Campbell R.E., Steinbach P.A., Giepmans B.N.G., Palmer A.E., Tsien R.Y. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein // *Nat. Biotechnol.*– 2004.– T. 22, № 12.– C. 1567–1572.
- [174] Pletnev S., Shcherbo D., Chudakov D.M., Pletneva N., Merzlyak E.M., Wlodawer A., Dauter Z., Pletnev V. A crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore // *J. Biol. Chem.*– 2008.– T. 283, № 43.– C. 28980–28987.
- [175] Chudakov D.M., Matz M.V., Lukyanov S., Lukyanov K.A. Fluorescent proteins and their applications in imaging living cells and tissues // *Physiol. Rev.*– 2010.– T. 90, № 3.– C. 1103–1163.
- [176] Gaytán P., Roldán-Salgado A. Elimination of redundant and stop codons during the chemical synthesis of degenerate oligonucleotides. Combinatorial testing on the chromophore region of the red fluorescent protein mKate // *ACS Synth. Biol.*– 2013.– T. 2, № 8.– C. 453–462.

- [177] Malo G.D., Pouwels L.J., Wang M., Weichsel A., Montfort W.R., Rizzo M.A., Piston D.W., Wachter R.M. X-ray structure of Cerulean GFP: a tryptophan-based chromophore useful for fluorescence lifetime imaging // *Biochemistry*.– 2007.– T. 46, № 35.– C. 9865–9873.
- [178] Rizzo M.A., Springer G.H., Granada B., Piston D.W. An improved cyan fluorescent protein variant useful for FRET // *Nat. Biotechnol.*– 2004.– T. 22, № 4.– C. 445–449.
- [179] Ivashkin P.E., Lukyanov K.A., Lukyanov S., Yampolsky I.V. A synthetic GFP-like chromophore undergoes base-catalyzed autoxidation into acylimine red form // *J. Org. Chem.*– 2011.– T. 76, № 8.– C. 2782–2791.
- [180] Quillin, M.L., Anstrom, D.M., Shu, X., O’Leary, S., Kallio, K., Chudakov, D.M., Remington, S.J. Kindling Fluorescent protein from *Anemonia sulcata*: dark-state structure at 1.38 Å resolution, // *Biochemistry*.– 2005.– T. 44, № 15.– C. 5774–5787.
- [181] Miyawaki A., Shcherbakova D.M., Verkhusha V.V. Red fluorescent proteins: chromophore formation and cellular applications // *Curr. Opin. Struct. Biol.*– 2012.– T. 22, № 5.– C. 679–688.
- [182] Ma Y., Sun Q., Zhang H., Peng L., Yu J.-G., Smith S.C. The mechanism of cyclization in chromophore maturation of Green Fluorescent Protein: A Theoretical Study // *J. Phys. Chem. B.*– 2010.– T. 114, № 29.– C. 9698–9705.
- [183] Ma Y., Yu J.-G., Sun Q., Li Z., Smith S.C. The mechanism of dehydration in chromophore maturation of wild-type green fluorescent protein: A theoretical study // *Chem. Phys. Lett.*– 2015.– T. 631–632.– C. 42–46.
- [184] Yampolsky I.V., Remington S.J., Martynov V.I., Potapov V.K., Lukyanov S., Lukyanov K.A. Synthesis and properties of the chromophore of the asFP595 chromoprotein from *Anemonia sulcata* // *Biochemistry*.– 2005.– T. 44, № 15.– C. 5788–5793.
- [185] Gross L.A., Baird G.S., Hoffman R.C., Baldrige K.K., Tsien R.Y. The structure of the chromophore within DsRed, a red fluorescent protein from coral // *Proc. Natl. Acad. Sci.* – 2000.– T. 97, № 22.– C. 11990–11995.
- [186] Chudakov D.M., Matz M.V., Lukyanov S., Lukyanov K.A. Fluorescent proteins and their applications in imaging living cells and tissues // *Physiol. Rev.*– 2010.– T. 90, № 3.– C. 1103–1163.
- [187] Voliani V., Bizzarri R., Nifosì R., Abbruzzetti S., Grandi E., Viappiani C., Beltram F. Cis-trans photoisomerization of fluorescent-protein chromophores // *J. Phys. Chem. B.*– 2008.– T. 112, № 34.– C. 10714–10722.

- [188] Dedecker P., De Schryver F.C., Hofkens J. Fluorescent proteins: shine on, you crazy diamond // *J. Am. Chem. Soc.*— 2013.— T. 135, № 7.— C. 2387–2402.
- [189] Pletnev S., Shcherbo D., Chudakov D.M., Pletneva N., Merzlyak E.M., Wlodawer A., Dauter Z., Pletnev V. A Crystallographic study of bright far-red fluorescent protein mKate reveals pH-induced cis-trans isomerization of the chromophore // *J. Biol. Chem.*— 2008.— T. 283, № 43.— C. 28980–28987.
- [190] Andresen M., Stiel A.C., Trowitzsch S., Weber G., Eggeling C., Wahl M.C., Hell S.W., Jakobs S. Structural basis for reversible photoswitching in Dronpa // *Proc. Natl. Acad. Sci.*— 2007.— T. 104, № 32.— C. 13005–13009.
- [191] Yampolsky I.V., Balashova T.A., Lukyanov K.A. Synthesis and spectral and chemical properties of the yellow fluorescent protein zFP538 chromophore // *Biochemistry.*— 2009.— T. 48, № 33.— C. 8077–8082.
- [192] Ando R., Hama H., Yamamoto-Hino M., Mizuno H., Miyawaki A. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein // *Proc. Natl. Acad. Sci.* — 2002.— T. 99, № 20.— C. 12651–12656.
- [193] Mizuno H., Mal T.K., Tong K.I., Ando R., Furuta T., Ikura M., Miyawaki A. Photo-induced peptide cleavage in the green-to-red conversion of a fluorescent protein // *Mol. Cell.*— 2003.— T. 12, № 4.— C. 1051–1058.
- [194] Ivashkin P.E., Yampolsky I.V., Lukyanov K.A. Synthesis and properties of chromophores of fluorescent proteins // *Russ. J. Bioorganic Chem.*— 2009.— T. 35, № 6.— C. 652–669.
- [195] Goedhart J., von Stetten D., Noirclerc-Savoie M., Lelimosin M., Joosen L., Hink M.A., van Weeren L., Gadella T.W.J. Jr, Royant A. Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93% // *Nat. Commun.*— 2012.— T. 3.— C. 751.
- [196] Filonov G.S., Moon J.D., Svensen N., Jaffrey S.R. Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution // *J. Am. Chem. Soc.*— 2014.— T. 136, № 46.— C. 16299–16308.
- [197] Baranov M.S., Lukyanov K.A., Borissova A.O., Shamir J., Kosenkov D., Slipchenko L.V., Tolbert L.M., Yampolsky I.V., Solntsev K.M. Conformationally locked chromophores as models of excited-state proton transfer in fluorescent proteins // *J. Am. Chem. Soc.*— 2012.— T. 134, № 13.— C. 6025–6032.
- [198] Baranov M.S., Solntsev K.M., Baleeva N.S., Mishin A.S., Lukyanov S.A., Lukyanov K.A., Yampolsky I.V. Red-shifted fluorescent aminated derivatives of a conformationally locked GFP

chromophore // Chemistry.– 2014.– T. 20, № 41.– C. 13234–13241.

[199] Baldridge A., Samanta S.R., Jayaraj N., Ramamurthy V., Tolbert L.M. Activation of fluorescent protein chromophores by encapsulation // J. Am. Chem. Soc.– 2010.– T. 132, № 5.– C. 1498–1499.

[200] Baleeva N.S., Myannik K.A., Yampolsky I.V., Baranov M.S. Bioinspired Fluorescent Dyes Based on a Conformationally Locked Chromophore of the Fluorescent Protein Kaede // Eur. J. Org. Chem.– 2015.– T. 2015, № 26.– C. 5716–5721.

[201] Singh S.B., Ondeyka J.G., Herath K.B., Zhang C., Jayasuriya H., Zink D.L., Parthasarathy G., Becker J.W., Wang J., Soisson S.M. Isolation, enzyme-bound structure and antibacterial activity of platencin A1 from *Streptomyces platensis* // Bioorg. Med. Chem. Lett.– 2009.– T. 19, № 16.– C. 4756–4759.

[202] Blom N.S., Tétreault S., Coulombe R., Sygusch J. Novel active site in *Escherichia coli* fructose 1,6-bisphosphate aldolase // Nat. Struct. Biol.– 1996.– T. 3, № 10.– C. 856–862.

[203] Soriano E.V., Rajashankar K.R., Hanes J.W., Bale S., Begley T.P., Ealick S.E. Structural similarities between thiamin-binding protein and thiaminase-I suggest a common ancestor // Biochemistry.– 2008.– T. 47, № 5.– C. 1346–1357.

[204] Teale M., Symersky J., DeLucas L. 3-methyladenine-DNA glycosylase II: the crystal structure of an AlkA-hypoxanthine complex suggests the possibility of product inhibition // Bioconjug. Chem.– 2002.– T. 13, № 3.– C. 403–407.

[205] Flower D.R. The lipocalin protein family: structure and function // Biochem. J.– 1996.– T. 318 (Pt. 1).– C. 1–14.

[206] Campanacci V., Bishop R.E., Blangy S., Tegoni M., Cambillau C. The membrane bound bacterial lipocalin Blc is a functional dimer with binding preference for lysophospholipids // FEBS Lett.– 2006.– T. 580, № 20.– C. 4877–4883.

[207] Schiefner A., Chatwell L., Breustedt D.A., Skerra A. Structural and biochemical analyses reveal a monomeric state of the bacterial lipocalin Blc // Acta Crystallogr. D Biol. Crystallogr.– 2010.– T. 66, № 12.– C. 1308–1315.

[208] Snapp E. Design and use of fluorescent fusion proteins in cell biology // Curr. Protoc. Cell Biol.– 2005.– T. 21.– Unit 21.4.

[209] Cranfill P.J., Sell B.R., Baird M.A., Allen J.R., Lavagnino Z., de Gruiter H.M., Kremers G.-J., Davidson M.W., Ustione A., Piston D.W. Quantitative assessment of fluorescent proteins // Nat.

Methods.– 2016.– T. 13, № 7.– C. 557–562.

[210] Han K.Y., Leslie B.J., Fei J., Zhang J., Ha T. Understanding the photophysics of the spinach-DFHBI RNA aptamer-fluorogen complex to improve live-cell RNA imaging // J. Am. Chem. Soc.– 2013.– T. 135, № 50.– C. 19033–19038.

[211] Galbraith C.G., Galbraith J.A. Super-resolution microscopy at a glance // J. Cell Sci.– 2011.– T. 124, № 10.– C. 1607–1611.

[212] Allen J.R., Ross S.T., Davidson M.W. Sample preparation for single molecule localization microscopy // Phys. Chem. Chem. Phys.– 2013.– T. 15, № 43.– C. 18771–18783.

[213] Phan T.G., Bullen A. Practical intravital two-photon microscopy for immunological research: faster, brighter, deeper // Immunol. Cell Biol.– 2010.– T. 88, № 4.– C. 438–444.