

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА
ФАКУЛЬТЕТ БИОИНЖЕНЕРИИ И БИОИНФОРМАТИКИ

На правах рукописи

Петушкова Анастасия Игоревна

**Структура и специфичность папаин-подобной цистеиновой протеиназы
тритикаина- α**

1.5.3. Молекулярная биология

ДИССЕРТАЦИЯ

на соискание ученой степени
кандидата биологических наук

Научный руководитель:
доктор биологических наук, доцент
Замятнин Андрей Александрович

Москва – 2025

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
1. ОБЗОР ЛИТЕРАТУРЫ	11
1.1 Структура папаин-подобных цистеиновых протеиназ	11
1.1.1 Общая структура	11
1.1.2 Сайт связывания S_2	13
1.1.3 Сайт связывания S_1	17
1.2 Регуляция протеолитической активности папаин-подобных протеиназ	20
1.2.1 Эндогенные ингибиторы	20
1.2.2 Неингибирующее влияние макромолекул на активность протеиназ	22
1.2.3 Влияние ионов на протеолитическую активность.....	23
1.2.4 Зависимость активности от значения pH.....	25
1.3 Папаин-подобные цистеиновые протеиназы в медицине.....	27
1.3.1 Патогенная активность цистеиновых катепсинов человека.....	27
1.3.2 Разработка селективных ингибиторов цистеиновых катепсинов человека	30
1.3.3 Применение растительных протеиназ в медицине.....	35
1.3.4 Цистеиновая протеиназа из пшеницы тритикаин- α в терапии непереносимости глютена.....	38
2. МАТЕРИАЛЫ И МЕТОДЫ	41
2.1 Используемые в работе материалы	41
2.2 Составы буферов и сред	43
2.3 Программное обеспечение	45
2.4 Получение бактериальных клеток-продуцентов	47
2.5 Продукция и очистка протеиназ.....	49
2.6 Скрининг субстратной специфичности тритикаина- α	51
2.7 Измерение протеолитической активности протеиназ	53
2.8 Рентгеноструктурный анализ тритикаина- α	54
2.9 Предсказание структур комплексов тритикаина- α с олигопептидами.....	56

2.10 Выравнивание аминокислотных последовательностей протеиназ.....	56
2.11 Вычисление распределения зарядов на поверхности протеиназ	57
3. РЕЗУЛЬТАТЫ	58
3.1 Продукция каталитического домена в одной полипептидной цепи с продоменом необходима для формирования активного тритикаина- α	58
3.2 Тритикаин- α специфичен к субстратам с гидрофобными и положительно заряженными аминокислотными остатками в P_2 - и P_1 -положениях	62
3.3 Третичная структура тритикаина- α , полученная с помощью метода рентгеноструктурного анализа	65
3.4 Субстрат-связывающая борозда тритикаина- α содержит гидрофобный карман и отрицательно заряженные аминокислотные остатки	69
3.5 Отрицательно заряженные остатки в S_1 - и S_2 -сайтах связывания тритикаина- α отличают фермент от других папаин-подобных протеиназ	75
3.6 Специфичность тритикаина- α к P_1 -положению субстрата зависит от значения pH среды и наличия Glu191 в S_1 -сайте связывания	79
ОБСУЖДЕНИЕ	83
ЗАКЛЮЧЕНИЕ	93
СПИСОК СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ	96
СПИСОК ЛИТЕРАТУРЫ.....	99
ПРИЛОЖЕНИЕ А. Структуры предсказанных комплексов пептидов с тритикаином- α	125
ПРИЛОЖЕНИЕ Б. Множественное выравнивание аминокислотных последовательностей каталитических доменов ППЦП	129
ПРИЛОЖЕНИЕ В. Достоверность различий ферментативной активности тритикаина- α и его мутантов	131

ВВЕДЕНИЕ

Актуальность темы исследования

Протеиназы – это группа ферментов, которые катализируют гидролиз пептидных связей в белках и пептидах, что приводит к их расщеплению до более коротких фрагментов и аминокислот. Долгое время считалось, что ферменты из семейства папаин-подобных цистеиновых протеиназ (ППЦП) выполняют свои функции только в кислой среде лизосомально-эндосомального компартмента и не отличаются высокой субстратной специфичностью. Однако в последнее время было обнаружено, что эти протеиназы выполняют специфические функции внутри различных органелл, а также вне клетки. Растительные ППЦП участвуют в расщеплении запасных белков во время прорастания семян, индукции защитных реакций, процессах старения и регулируемой гибели клеток. ППЦП животных, которые называются цистеиновыми катепсинами, вовлечены в дифференцировку клеток, активацию прогормонов, процессинг антигенов, аутофагию и апоптоз. Также цистеиновые катепсины способствуют метастазированию, ангиогенезу и химиорезистентности злокачественных опухолей. В связи с этим цистеиновые катепсины человека имеют диагностическое значение, а также являются перспективными мишенями для противоопухолевых препаратов. Растительные же ППЦП, наоборот, активно применяют при лечении ряда патологических состояний. Среди них обнаружены ферменты, которые обладают противоопухолевым, антибактериальным, противогрибковым действием, могут влиять на разные этапы свёртывания крови и пищеварения.

Эффективность распознавания субстратов и ингибиторов протеиназой зависит от структуры сайтов связывания в активном центре фермента. Кроме того, эта специфичность подвержена влиянию условий среды, таких как значение pH и наличие/отсутствие других молекул и ионов. Данные о структуре и субстратной специфичности ППЦП могут быть использованы для разработки селективных диагностических зондов, терапевтических ингибиторов, а также проектирования новых ферментов в медицинских целях.

Степень разработанности темы исследования

Активный центр протеиназы находится в субстрат-связывающей борозде, которая представляет собой углубление на поверхности молекулы. В этой борозде для ППЦП было описано пять относительно четко определенных сайтов связывания [1]. Три аминокислотных остатка (а.о.; P_1 , P_2 и P_3) субстрата, расположенные на *N*-концевой стороне от расщепляемой пептидной связи, распознаются соответствующими сайтами S_1 , S_2 и S_3 фермента. Аналогично происходит взаимодействие с другой стороны от связи: а.о. в P_1' - и P_2' -положениях субстрата контактируют с S_1' - и S_2' -сайтами протеиназы, соответственно [1]. В связи с избыточной активностью цистеиновых катепсинов человека в опухолевых тканях и необходимостью разработки селективных зондов и ингибиторов, их субстратная специфичность является наиболее изученной среди ферментов семейства ППЦП.

Многие ППЦП специфичны к основным а.о. в положении P_1 , что в катепсинах В и Z (CtsB и Z) объясняется наличием отрицательно заряженных а.о. в S_1 - и S_1' -сайтах связывания [2–5]. Однако именно сайт S_2 считается основной детерминантой субстратной специфичности ППЦП, что учитывается при разработке зондов и ингибиторов [6]. S_2 -сайт связывания в ППЦП обычно представляет собой гидрофобный карман, который селективен к гидрофобным а.о. в P_2 -положении [2,5,7]. При этом некоторые протеиназы обладают уникальной субстратной специфичностью по этому сайту. Например, CtsK может связывать Pro в P_2 -положении, что определяет его способность расщеплять интактные тройные спирали коллагена [8,9]. Карбоксипептидазы CtsB и Z, а также круципаин из трипаносомы *Trypanosoma cruzi*, в отличие от других ППЦП, содержат отрицательно заряженные а.о. в S_2 -сайте [3,10–12]. Субстратная специфичность сайта связывания S_3 , согласно имеющимся данным, очень низкая [2]. Другая сторона субстрат-связывающей борозды содержит консервативные остатки Trp, которые могут участвовать в гидрофобных и стэкинг-взаимодействиях [3,13,14]. Однако CtsB, L и V также могут связывать полярные а.о. в сайте S_1' [4,10]. Сайт связывания S_2' в эндопептидазах является

гидрофобным, в то время как карбоксипептидазы CtsB и Z содержат в нём положительно заряженные остатки His для взаимодействий с C-концом субстрата [3,10].

Субстратная специфичность и структура растительных ППЦП менее изучены по сравнению с катепсинами человека. Тем не менее множество растительных протеиназ, таких как папаин из папайи *Carica papaya*, бромелаин из ананаса *Ananas comosus* и фицин из фикуса *Ficus insipida*, широко используются в медицине: для лечения онкозаболеваний, расстройств пищеварения и вирусных заболеваний [15–17]. Они входят в состав стоматологических средств и препаратов для лечения ран и ожогов [18–21]. Последние несколько лет проводятся исследования ППЦП, обладающих глютеназной активностью, с целью оценить перспективы их применения в терапии непереносимости глютена и целиакии [22,23]. Одним из таких ферментов является охарактеризованная ранее глютенеза тритикаин- α из пшеницы *Triticum aestivum* [22].

Цели и задачи исследования

Целью данной работы является описание субстратной специфичности ППЦП тритикаина- α из пшеницы *T. aestivum* с идентификацией структурных детерминант этой специфичности. Для достижения цели необходимо решить следующие задачи:

- 1) Установить необходимые условия экспрессии для получения растворимого активного тритикаина- α .
- 2) Охарактеризовать субстратную специфичность тритикаина- α .
- 3) Определить третичную структуру тритикаина- α методом рентгеноструктурного анализа для выявления сайтов связывания в субстрат-связывающей борозде фермента.
- 4) Сравнить структуру сайтов связывания тритикаина- α и других ППЦП.
- 5) Определить влияние значения pH на субстратную специфичность тритикаина- α .

Объект и предмет исследования

Объектом исследования являлась ППЦП тритикаин- α из пшеницы *T. aestivum*. Предметом исследования были определение и характеристика структуры и субстратной специфичности этой рекомбинантной протеиназы.

Научная новизна исследования

В представленной работе была впервые описана субстратная специфичность ППЦП пшеницы тритикаина- α с использованием пептидных субстратов и была впервые получена третичная структура данного фермента. Были идентифицированы а.о., формирующие гидрофобный карман в S_2 -сайте связывания, а также отрицательно заряженные а.о. в S_1 -сайте связывания тритикаина- α , которые могут образовывать водородные и электростатические связи с а.о. в P_1 -положении субстрата. В данной работе была впервые проведена сравнительная характеристика первичных структур тритикаина- α с другими ППЦП в S_1 - и S_2 -сайтах связывания, а также сравнение распределения зарядов в субстрат-связывающей борозде растительных ППЦП. Также была показана зависимость субстратной специфичности тритикаина- α от значений pH в отношении пептидов с различными а.о. в P_1 -положении.

Теоретическая и практическая значимость работы

Теоретическая значимость работы заключается в расширении представлений о субстратной специфичности ППЦП: какие а.о. в активном центре фермента могут на нее влиять и как она зависит от значения pH. Впервые получена третичная структура ППЦП тритикаина- α методом рентгеноструктурного анализа, а также проведена характеристика субстратной специфичности этой протеиназы. Полученные данные могут быть экстраполированы на другие ферменты семейства ППЦП.

Практическая значимость работы заключается в том, что полученные данные и их анализ могут быть использованы для дизайна селективных ингибиторов и зондов ППЦП, которые могут найти применение в медицине. Также полученные данные могут использоваться для редизайна протеиназ с

целью скорректировать их субстратную специфичность для нужд различных сфер промышленности и медицины.

Методология и методы исследования

В работе использовали молекулярно-биологические, биохимические, микробиологические, биофизические и биоинформатические методы исследования.

Наработка генов тритикаина- α из пшеницы (*T. aestivum*) проводилась методом полимеразной цепной реакции (ПЦР). Сайт-направленный мутагенез использовали для внесения точечных мутаций в ген тритикаина- α . Полученные фрагменты дезоксирибонуклеиновых кислот (ДНК) были помещены в экспрессионные векторы с помощью методов молекулярного клонирования. Клетки кишечной палочки *Escherichia coli* были трансформированы полученными плазмидами и затем использованы для продукции целевых белков. Исследуемые ферменты были очищены из бактериальных лизатов при помощи хроматографических методов. Для скрининга субстратной специфичности тритикаина- α использовали библиотеку пептидов, меченных фосфотирозином (pTyr). Измерение каталитической активности исследуемых форм фермента производили с использованием пептидных субстратов, слитых с флуорогенной меткой 7-амино-4-метилкумарин (АМС). Для анализа полученных данных применили модель ферментативной кинетики, описываемую уравнением Михаэлиса-Ментен.

Структура неактивного тритикаина- α была получена методом рентгеноструктурного анализа с использованием источника синхротронного излучения Diamond Light Source, Чилтон, Великобритания. Собранные данные дифракции были обработаны в программе AutoPROC и затем использованы для получения структуры фермента методом молекулярного замещения в программе PHENIX. Кристаллическая структура фермента далее была использована для предсказания возможных расположений пептидов в комплексе с тритикаином- α в программе AlphaFold. Для сопоставления соответствующих участков аминокислотных последовательностей, которые формируют S₁- и S₂-сайты

связывания в различных ППЦП, было использовано множественное выравнивание в программе AliView. Распределения зарядов на поверхности растительных ППЦП были рассчитаны с использованием уравнения Пуассона-Больцмана в программе ProKSim.

Положения, выносимые на защиту

- 1) Для получения активного растворимого тритикаина- α необходима продукция белка в форме полипептида, содержащего продомен и каталитический домен.
- 2) Тритикаин- α проявляет субстратную селективность по отношению к пептидным последовательностям с гидрофобным а.о. в P_2 - и положительно заряженным или полярным незаряженным а.о. в P_1 -положении.
- 3) В субстрат-связывающей борозде тритикаина- α расположены консервативный гидрофобный карман в S_2 -сайте и отрицательно заряженные а.о. Glu191 и Asp289 в сайте связывания S_1 , которые могут участвовать во взаимодействиях с субстратом.
- 4) Субстратная специфичность тритикаина- α к аминокислоте в позиции P_1 субстрата определяется остатками Glu191 и Asp289 в S_1 -сайте фермента и зависит от pH среды.

Степень достоверности и апробация результатов

Диссертационная работа была выполнена с использованием современных методов исследования, соответствующих поставленным целям и задачам. Достоверность приведённых результатов подтверждается их воспроизводимостью при проведении нескольких независимых экспериментов, а также статистической обработкой полученных результатов.

По материалам диссертации было опубликовано шесть статей в рецензируемых научных журналах. Материалы диссертации были представлены и обсуждались на шести конференциях в виде устных и постерных докладов: International Conference "Toolkits for DNA vaccine design, an update" (Москва, Россия, 2016), Sechenov International Biomedical Summit 2017 (Москва, Россия, 2017), Международная научная конференция "XII чтения памяти академика Юрия

Анатольевича Овчинникова" VIII Российский симпозиум "Белки и пептиды" (Москва, Россия, 2017), 45th FEBS Congress (онлайн, 2021), III Объединенный Научный Форум Физиологов, Биохимиков и Молекулярных Биологов (Сочи, Россия, 2022), The 2nd International Electronic Conference on Biomolecules: Biomacromolecules and the Modern World Challenges (онлайн, 2022).

Личный вклад автора

Диссертационная работа написана на основе собственных данных, полученных автором и соавторами с 2016 по 2024 год. Автор работы проводила анализ литературных данных; клонирование генов исследованных белков, их наработку и очистку; измерение ферментативной активности тритикаина- α и его мутантов с последующим анализом полученных данных; подбор условий кристаллизации тритикаина- α и интерпретацию полученной методами рентгеноструктурного анализа и молекулярного замещения электронной плотности с решением модели структуры фермента; анализ предсказанных структур комплексов тритикаина- α с пептидами; выравнивание аминокислотных последовательностей ППЦП; вычисление распределения зарядов на поверхности тритикаина- α и других растительных ППЦП. Соискатель внесла существенный вклад в публикации по теме диссертации и принимала непосредственное участие в их написании, а также представляла результаты исследования на конференциях.

Структура и объем диссертации

Диссертационная работа изложена на 131 странице машинописного текста, включает 28 рисунков и 6 таблиц, и состоит из следующих разделов: введение, обзор литературы, материалы и методы, результаты, обсуждение, заключение, список сокращений и обозначений, список литературы из 294 наименований, приложение А, приложение Б и приложение В.

1. ОБЗОР ЛИТЕРАТУРЫ

1.1 Структура папаин-подобных цистеиновых протеиназ

1.1.1 Общая структура

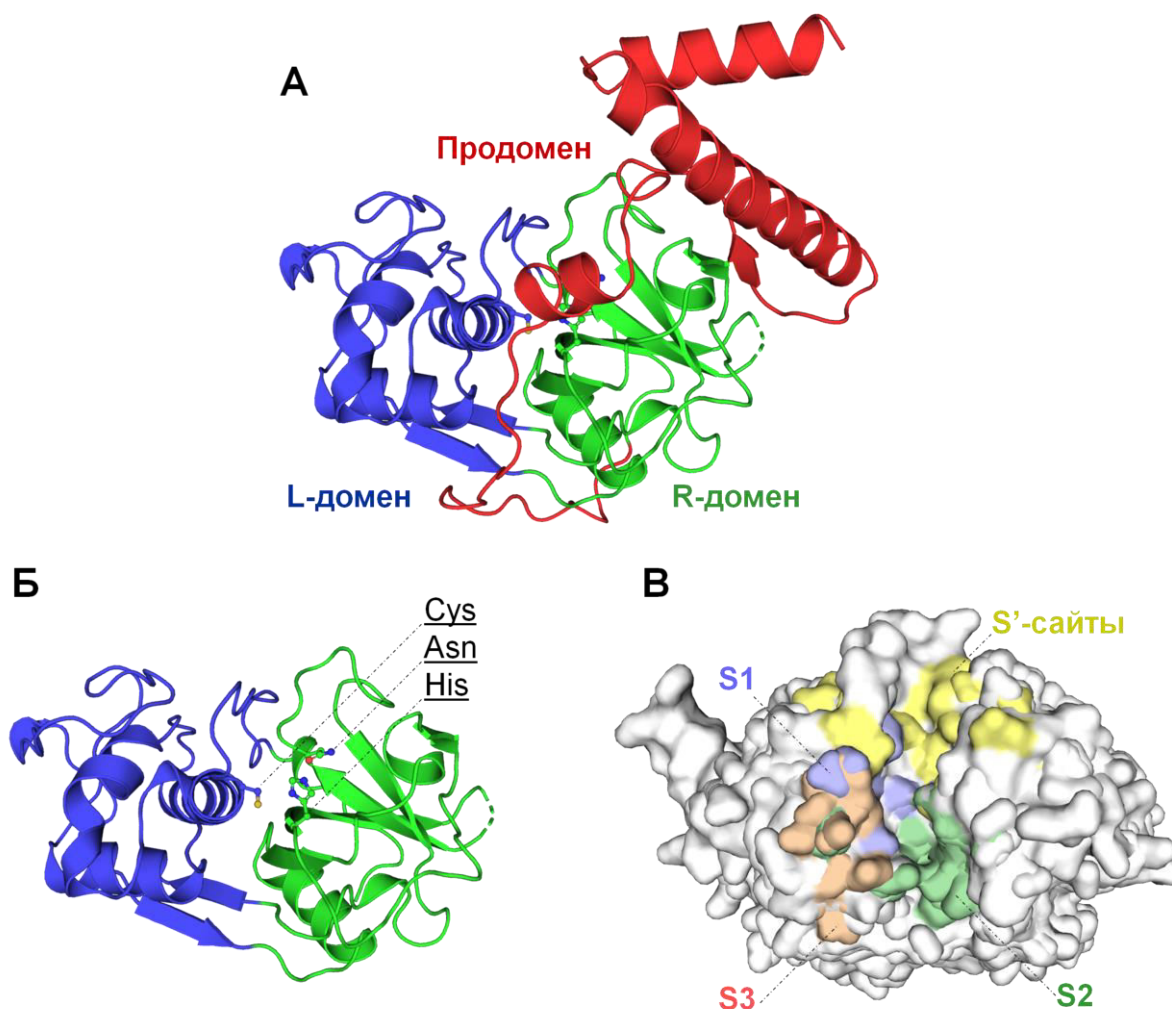


Рисунок 1 – Общие структурные особенности ППЦП [24]. А – Продомен (красный), L-домен (синий) и R-домен (зелёный) профермента CtsL (Protein Data Bank (PDB) ID: 1CJL) [25]. Б – Активная форма CtsL с указанными а.о. каталитической триады (PDB ID: 1CJL) [25]. В – Сайты связывания субстрата S₁ (синий), S₂ (зелёный), S₃ (оранжевый) и S' (жёлтый), указанные на поверхности ППЦП. На рисунке представлены выравненные относительно друг друга структуры CtsB, F, K, L, S, V и Z (PDB ID: 1CSB, 1M6D, 1ATK, 1CJL, 2FQ9, 1FH0 и 1EF7, соответственно) [3,25–30]. Для создания рисунков использовалась программа для визуализации молекул PyMOL [31].

Все ППЦП состоят из каталитического домена, на котором и происходит гидролиз, и продомена, который предотвращает активацию фермента до достижения необходимых условий и/или компартмента (Рисунок 1А). Структура ППЦП может также содержать и другие функциональные домены, но не обязательно. В состав каталитического домена входят два структурных домена: *N*-концевой, состоящий преимущественно из α -спиралей (L-домен), и *C*-концевой, состоящий из β -листа и коротких α -спиралей (R-домен) (Рисунок 1А) [32]. Эти домены образуют V-образную складку, в которой располагается активный сайт, содержащий каталитическую триаду из Cys, His и Asn (Рисунок 1Б) [3,6]. Ещё до взаимодействия с субстратом His депротонирует каталитический Cys. Образовавшийся анион серы атакует карбонильный углерод расщепляемой пептидной связи с образованием тетраэдрического интермедиата [33]. Далее протон, ранее приобретённый His, переносится на аминокгруппу гидролизуемой пептидной связи с высвобождением одного из продуктов реакции. На последнем этапе происходит гидролиз тиоэфирной связи с высвобождением второго продукта и регенерацией протеиназы [33]. В зимогенах ППЦП каталитическая триада закрыта продоменом (Рисунок 1А), который взаимодействует с субстрат-связывающей бороздой (Рисунок 1В). При этом каталитический домен находится в активной конформации [13].

На основе анализа аминокислотных последовательностей цистеиновые катепсины делят на два подсемейства: CtsL-подобные и CtsB-подобные протеиназы [34,35]. Проферменты подсемейства CtsL (CtsF, H, K, L, S, V и W) содержат пропептид с двумя консервативными мотивами: ERFNIN и GNFD. Проферменты подсемейства CtsB лишены либо только мотива GNFD (CtsB), либо их обоих (CtsC, O и Z) [6]. ППЦП растений к тому же имеют консервативные таргетные сигналы и сайты гликозилирования (отсутствующие в катепсинах человека), на основе которых их можно классифицировать. Растительные ППЦП было предложено подразделять на девять групп в зависимости от наличия или отсутствия мотивов ERFNIN/ERFNAQ в продомене, сигнальных последовательностей KDEL или HDEL для нацеливания в эндоплазматический

ретикулум, NPIR последовательности для транспорта в вакуоли, двух Cys в мотиве CCW в активном сайте, гранулин-подобного домена, количества дисульфидных мостиков и сайтов N-гликозилирования [7]. Все протеиназы также можно классифицировать по структуре их субстрат-связывающей борозды. В эндопептидазах CtsF, K, L, O, S, V и W сайты связывания субстрата располагаются по обе стороны от каталитического Cys [36]. У карбоксипептидаз CtsB и Z и у аминопептидаз CtsC и H сайты связывания расположены только по одну сторону от каталитического Cys. Следовательно, эти протеиназы могут связывать либо C-конец, либо N-конец субстрата, а другая сторона борозды занята дополнительными доменами [36]. Оклюзионная петля CtsB и мини-петля CtsZ содержат His для взаимодействия с C-концевой карбоксильной группой субстрата. Мутации His в окклюзионной петле и делеция петли продемонстрировали их важность также и для эндопептидазной активности CtsB [6,37,38]. Мини-цепь CtsH и эксклюзионный домен CtsC несут карбоксильные группы для распознавания положительно заряженного N-конца субстратов [39,40].

В субстрат-связывающей борозде можно выделить сайты связывания для а.о. субстрата: в сторону N-конца (S_1 -, S_2 -, S_3 -сайты связывания и т.д.) и C-конца (S_1' -, S_2' -, S_3' -сайты связывания и т.д.) от гидролизруемой пептидной связи (Рисунок 1В). Специфичность ППЦП определяется а.о., расположенными в этих областях, поэтому далее мы уделим им особое внимание.

1.1.2 Сайт связывания S_2

Сайт связывания S_2 ППЦП считается наиболее важным для распознавания субстратов [41]. Большинство протеиназ этого семейства специфичны к субстратам, содержащим объемные гидрофобные а.о. в P_2 -положении [2,5]. Было показано, что ингибиторы с эпоксидной функциональной группой снижают активность ППЦП эффективнее при наличии большого гидрофобного или ароматического а.о. (Leu, Ile, Phe, норлейцин (Nle), Trp, Tyr, Val) в P_2 по

сравнению с ингибиторами, содержащими малые или гидрофильные а.о. (Ala, Asp, Glu, Lys, Gln, Gly, His, Pro, Ser, Arg, Thr) в P₂-положении [7]. А.о. в S₂-сайте связывания цистеиновых катепсинов, которые могут образовывать связи с субстратами и ингибиторами, показаны на Рисунке 2 и обсуждаются далее в настоящем подразделе.

Гидрофобный S₂-сайт CtsL образован Leu182, Met183, Ala248, Met274 и Gly277 в нижней части [32]. Asp273, Met274, Asp275 и Ala327 ограничивают сайт связывания S₂ по бокам [32]. Несмотря на присутствие Asp273 и Asp275 в S₂-сайте, CtsL предпочитает ароматические а.о., в основном Phe, Trp и Tyr, а также Leu в P₂-положении субстрата [2,42]. S₂-сайт связывания CtsV также специфичен к гидрофобным а.о., предпочитая Phe и Leu, менее эффективно связывая пептиды с Tyr, Val и Ile и хуже всего с Trp (в ряду гидрофобных а.о.) в положении P₂ субстрата [4]. CtsS проявляет специфичность к гидрофобным а.о. Val, Met и Nle в P₂-положении субстрата [6,43]. Карман S₂ в CtsS образован боковой цепью Met185 и метиленовыми группами Gly251 и Gly279 в его нижней части, а боковые цепи Phe184, Val276 и Phe325 формируют его стороны [10]. Gly251, Val276 и Phe325 являются уникальными а.о. в S₂-сайте связывания CtsS [14]. Gly251 в CtsS заменен на Ala в CtsK, L и V, а Val276 заменен на Leu в CtsK и V и Met в CtsL, так что S₂-сайт связывания в CtsS глубже и шире, чем в других катепсинах [10]. Боковая цепь Phe325 в S₂-сайте CtsS может принимать две конформации: для более эффективного связывания Leu или более объемного Phe в P₂-положении субстрата, что является уникальной особенностью CtsS [14]. Замена Phe325Glu в CtsS приводит к снижению сродства к гидрофобным а.о. в P₂-положении субстрата, но в 77 раз увеличивает эффективность гидролиза дипептида Z-RR-MCA [41]. S₂-сайт связывания CtsF также гидрофобный и сформирован Leu337, Pro338, Ala403, Ala432 и Met477 [28]. Ширина кармана S₂ в CtsF ограничена Leu337 с одной стороны и Ile429 и Asp430 – с другой [28].

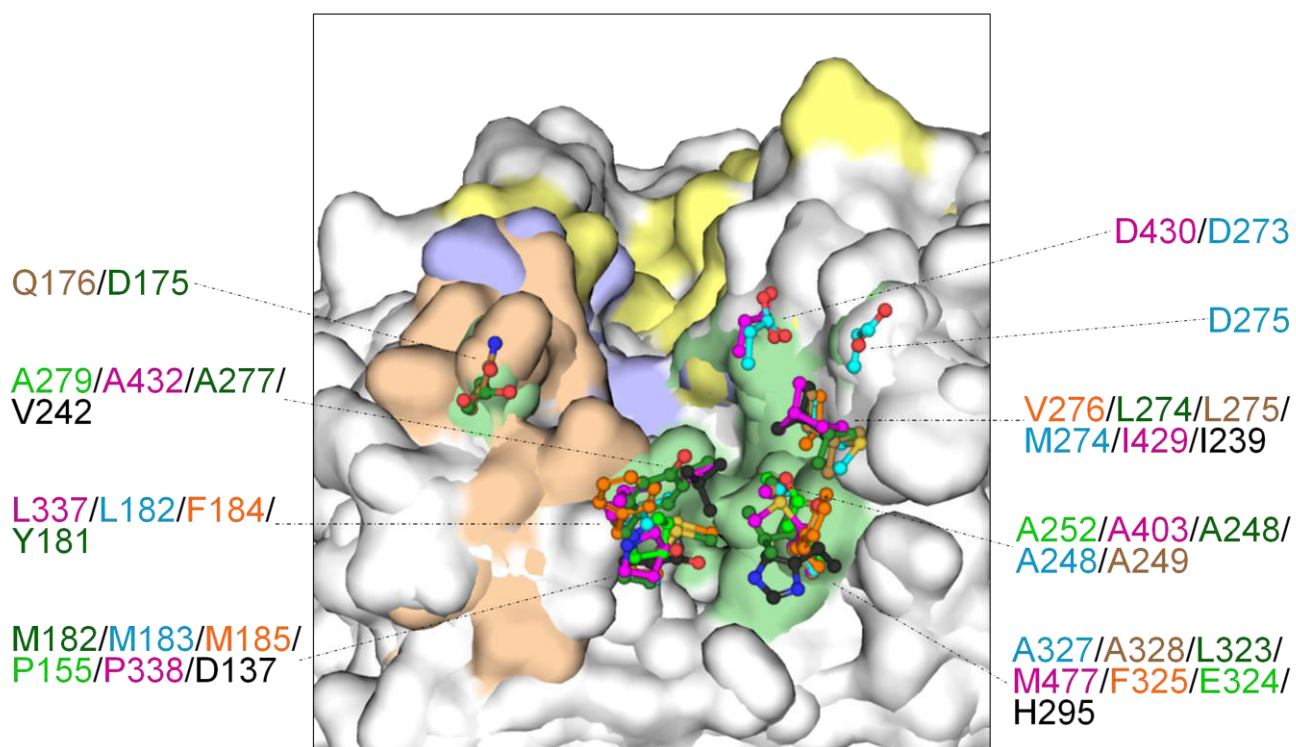


Рисунок 2 – А.о. S_2 -сайта связывания ППЦП [24]. На рисунке показаны поверхности выравненных структур CtsB, F, K, L, S, V и Z (PDB ID: 1CSB, 1M6D, 1ATK, 1CJL, 2FQ9, 1FH0 и 1EF7, соответственно) [3,25–30]. Сайты связывания обозначены теми же цветами, что и на Рисунке 1В. А.о. CtsB показаны светло-зелёным цветом, CtsF – фиолетовым, CtsK – тёмно-зелёным, CtsL – голубым, CtsS – оранжевым, CtsV – бежевым, CtsZ – черным.

Некоторые ППЦП специфичны к Pro в P_2 -положении субстрата и обладают уникальной коллагенолитической активностью. К ним относится CtsK, S_2 -сайт связывания которого состоит из Tyr181, Leu274 и Leu323, которые считаются ответственными за особую специфичность протеиназы [10,44], а также Met182, Ala248 и Ala277 [8]. CtsK активен в отношении субстратов, содержащих как Pro так и Phe в положении P_2 . Тем не менее он преимущественно гидролизует коллаген типа I после Pro в P_2 , в отличие от CtsL, который расщепляет коллаген после Phe в P_2 -положении [45]. Мутации Tyr181Leu и Leu323Ala (как в CtsL) приводили к потере специфичности к Pro и коллагеназной активности CtsK [9,46], тогда как обратные мутации Leu182Tyr и Ala327Leu делали CtsL специфичным к Pro в P_2 -положении субстрата [47]. Искусственно полученная усеченная форма CtsC без эксклюзионного домена (CtsCΔEx) подобно другим ППЦП предпочитает крупные гидрофобные Phe и Leu, но также и Pro в P_2 -положении субстрата [48]. CtsL-подобная протеиназа FhCL2 из печеночной двуустки *Fasciola hepatica* тоже

специфична к Pro в P₂, что опосредовано стэкинг-взаимодействиями между Tyr176 и Pro субстрата (нумерация по FhCL1) [9,49]. Единичная мутация Leu176Tyr в FhCL5 смогла повысить эффективность гидролиза *Tos-GPR-AMC* (Tos – тозил) с Pro в P₂-положении, однако мутации Leu176Tyr и Val267Leu в FhCL1 не привели к значительным изменениям в специфичности фермента [50,51].

Некоторые катепсины также могут связывать полярные а.о. в положении P₂, например CtsB, S₂-сайт которого состоит из Pro155, Ala252, Gly277, Ala279 и Glu324 [10]. Как и другие цистеиновые катепсины, CtsB может взаимодействовать с объемными гидрофобными а.о. в положении P₂ [52]. Но в отличие от большинства катепсинов, сайт S₂ в CtsB содержит отрицательно заряженный Glu324, который может взаимодействовать с положительно заряженными а.о. в P₂-положении субстрата при нейтральном значении pH [5,10,53]. Было показано, что мутация Glu324Gln в CtsB снижала эффективность гидролиза субстратов, содержащих Arg, но не влияла на гидролиз субстратов с Phe в P₂-положении [6,54]. Аналогичная специфичность к субстратам с положительно заряженным а.о. в P₂-положении была обнаружена также и у CtsB-подобных протеиназ из растений [11]. Карбоксипептидаза CtsZ также специфична к полярным а.о. [3]. S₂-сайт связывания этой протеиназы образован Gly135, Asp137, Gly215, Ile239, Val242 и His295. Из них Asp137 уникален для CtsZ и замещен гидрофобным а.о. (Met или Pro) в других цистеиновых катепсинах, а His295 заменен самыми разными кроме положительно-заряженных а.о. в других ППЦП [3]. Glu330 в S₂-сайте связывания круципаина из *T. cruzi* обеспечивает специфичность протеиназы к субстратам Z-RR-AMC и Z-RA-AMC [55]. Мутантная форма с заменой Glu330Ala расщепляет эти субстраты в 300 и 200 раз хуже, а Z-FR-AMC только в два раза хуже по сравнению с диким типом [55].

ППЦП Pd_dinase из бактерии *Parabacteroides distasonis* обладает уникальной диаминопептидазной активностью с сильной специфичностью к N-концевому Gly в положении P₂. Субстрат GR-AMC и ингибитор GR-AOMK (AOMK – ацилоксиметилкетон), а также субстраты с небольшими N-концевыми

Ala, Ser и Thr, эффективно распознаются Pd₂cinase. Однако, в отличие от большинства ППЦП, Pd₂cinase не гидролизует пептиды, содержащие большие гидрофобные а.о. в Р₂-положении [56].

Специфичность большинства ППЦП к гидрофобным а.о. в Р₂-положении субстрата уже давно известна и обоснована наличием гидрофобного кармана в S₂-сайте связывания (Рисунок 2). При этом есть ППЦП (например, CtsK) с уникальной специфичностью к Pro в Р₂-положении и способностью гидролизовать интактные тройные спирали коллагена. А.о., определяющие эту специфичность, были идентифицированы, что позволило провести реинжиниринг CtsL, чтобы он мог расщеплять коллаген подобно CtsK [45]. Также некоторые ППЦП содержат заряженные а.о. в S₂-сайте связывания, которые позволяют им расщеплять субстраты с гидрофильными а.о. Учитывая эти уникальные особенности, можно синтезировать селективные ингибиторы и зонды для ППЦП [3,27].

1.1.3 Сайт связывания S₁

S₁-сайт связывания ППЦП взаимодействует с первым а.о. субстрата от гидролизуемой пептидной связи в сторону N-конца. В папаине сайт S₁ образован Ser154, Cys155 и Gly156 с одной стороны и Gly195, Cys196 и Asn197 с другой, соединенными дисульфидным мостиком Cys155-Cys196 [3]. А.о. в S₁-сайте связывания катепсинов человека, которые могут взаимодействовать с Р₁-положением субстрата, показаны на Рисунке 3. Сайт S₁ CtsB сформирован Gln102, Gly106, Cys108, Gly153 и Gly277 [10]. CtsB предпочитает Arg и Lys в положении Р₁ субстрата среди всех природных а.о., потому что его окклюзионная петля содержит отрицательно заряженный Glu201; однако он также может взаимодействовать с Tyr, Nle, Thr и Gly в Р₁-положении [57]. В то же время субстраты, содержащие искусственные объемные и гидрофобные производные Lys и Cys в Р₁, гидролизуются быстрее, чем субстраты с Arg в Р₁-положении [52]. В CtsF сайт S₁ неглубокий и образован Gly293, Cys333, Met334 и Gly335, среди

которых Met334 уникален в семействе ППЦП [28]. При этом CtsF также специфичен к положительно заряженным а.о. в P₁-положении субстрата [2]. Сайт связывания S₁ в CtsV специфичен к субстратам с Leu, Gln, Arg и Lys в P₁-положении, в то время как CtsL предпочитает Arg и в меньшей степени Lys [4]. Субстраты с отрицательно заряженными и гидрофобными а.о. в P₁-положении плохо гидролизуются CtsL и V, а Pro является наименее предпочтительным [4]. CtsK тоже специфичен к положительно заряженным а.о., а также Gly в P₁-положении субстрата [6,58]. CtsZ эффективнее всего гидролизует субстраты с Arg в P₁-положении из-за боковой цепи Glu133 в сайте S₁ [3]. S₁-сайт связывания CtsS сформирован Gly137, Asn181, Gly182, Arg255 и дисульфидом Cys136-Cys180 [14,59]. Хотя Arg255 вносит положительный заряд в S₁-сайт связывания CtsS, о специфичности этой протеиназы к отрицательно заряженным а.о. в P₁-положении субстрата не сообщалось.

Специфичность к положительно заряженным а.о. в положении P₁ субстрата является обычной и для других ППЦП. Результаты молекулярного докинга продемонстрировали специфичность цистеиновой протеиназы zmCP1 из кукурузы *Zea mays* к Arg в P₁-положении; FR-AMC оказался наилучшим субстратом. С помощью компьютерного моделирования было продемонстрировано, что положительный заряд на Arg в P₁-положении взаимодействует с Gln143 и Cys149, которые формируют оксианионную дыру фермента [5]. Скрининг специфичности CtsL-подобного фермента из клеща *Rhipicephalus microplus* также показал, что эта протеиназа специфична к Leu и Arg в P₁-положении [6,60]. CtsB-подобная цистеиновая протеиназа из *R. sativus*, а также FhCL1 и FhCL2 из *F. hepatica* тоже специфичны к субстратам с положительно заряженным а.о. в P₁-положении [11,61].

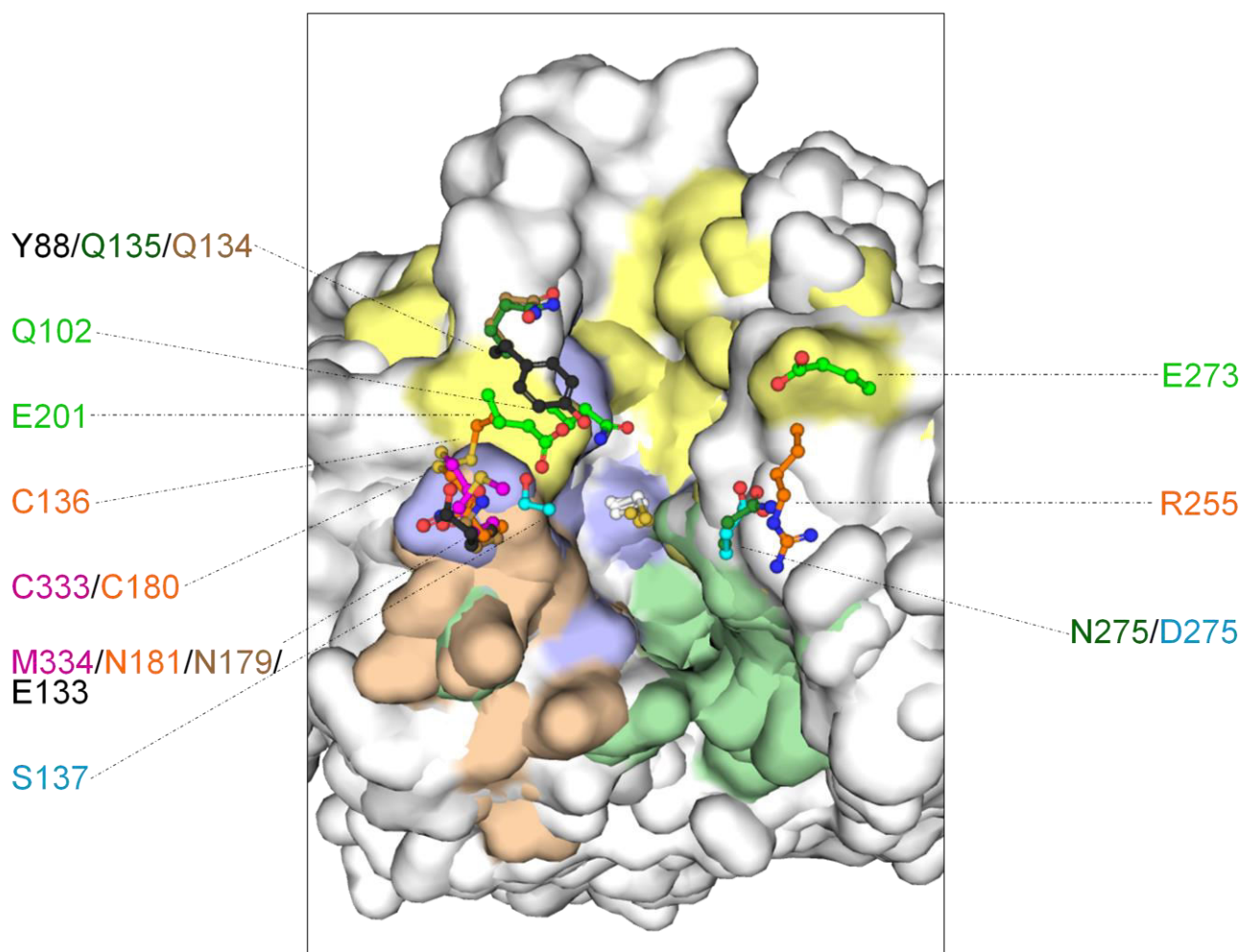


Рисунок 3 – А.о. S_1 -сайта связывания ППЦП [24]. На рисунке показаны поверхности выравненных структур CtsB, F, K, L, S, V и Z (PDB ID: 1CSB, 1M6D, 1ATK, 1CJL, 2FQ9, 1FH0 и 1EF7, соответственно) [3,25–30]. Сайты связывания обозначены теми же цветами, что и на Рисунке 1В. А.о. CtsB показаны светло-зелёным цветом, CtsF – фиолетовым, CtsK – тёмно-зелёным, CtsL – голубым, CtsS – оранжевым, CtsV – бежевым, CtsZ – черным. Также визуализировано местоположение каталитических Cys.

На основе имеющихся данных можно сделать вывод, что ППЦП обладают специфичностью к положительно заряженным а.о. в P_1 -положении субстрата. В некоторых из ППЦП это объясняется наличием в их сайте связывания S_1 а.о., несущих отрицательный заряд (Рисунок 3). Однако считается, что по сравнению с гидрофобным карманом в S_2 -сайте, значимость S_1 -сайта связывания во взаимодействиях ППЦП с субстратом меньше, поэтому многие из клинических ингибиторов цистеиновых катепсинов содержат неполярные заместители в P_1 -положении [62,63]. При этом замена изобутила на более полярный 4-фторлейцин в P_1 -положении ингибитора CtsK L-873724 привела к разработке нового

перспективного клинического ингибитора оданакатиб [64]. Таким образом, исследования селективности S₁-сайта связывания ППЦП могут успешно применяться для разработки новых специфичных ингибиторов.

1.2 Регуляция протеолитической активности папаин-подобных протеиназ

1.2.1 Эндогенные ингибиторы

К эндогенным ингибиторам цистеиновых катепсинов относятся цистатины (включая стефины и кининогены; Таблица 1), тиропины и серпины, которые являются обратимыми конкурентными ингибиторами [6]. Перечисленные ингибиторы взаимодействуют с субстрат-связывающей бороздой, где расположен активный центр фермента. Структурные отличия сайтов связывания ППЦП определяют различия в специфичности ППЦП к разным эндогенным ингибиторам. Например, бычьи стефины А, В и С ингибируют эндопептидазы папаин, CtsL и S более эффективно, чем карбоксипептидазу CtsB [65–67]; а мышинный стефин А ингибирует эндопептидазы CtsL и S сильнее, чем экзопептидазы CtsB, С и Н [68]. Активность CtsZ может подавляться различными ингибиторами цистеиновых протеиназ, однако молекулярное моделирование показало, что боковые цепи His84 и Tyr88 мешают связыванию Val и Ala в последовательности QVVAG стефина В [3]. Бактериальный лейпептин значительно эффективнее ингибирует соответствующие катепсины крыс, чем катепсины человека [8]. Также есть ингибиторы, которые подавляют активность ППЦП в отношении определенного типа субстратов. Например, эластазная активность CtsK и V ингибируется гликозаминогликанами, такими как хондроитинсульфат [69]. А коллагеназная активность CtsK подавляется полиглутаматами и олигонуклеотидами [6,70]. Каталитическая активность многих ППЦП подавляется активными формами кислорода и реактивными формами азота [71]. Папаин, CtsB и К могут быть обратимо ингибированы оксидом азота посредством S-нитрозилирования их каталитического Cys, в то время как

пероксид водорода окисляет тритикаин- α из *T. aestivum*, CtsK и S либо до обратимой формы сульфеновой кислоты, либо до необратимой сульфоновой кислоты [71]. S-метилметантиосульфат (MMTS) может индуцировать образование дисульфидной связи между Cys151 и каталитическим Cys154 в тритикаине- α , тем самым защищая фермент от необратимого окисления [72].

Таблица 1 – Цистатины – наиболее изученные эндогенные ингибиторы цистеиновых катепсинов человека. Полный список известных ингибиторов ППЦП представлен в базе данных MEROPS [73].

Ингибитор	Катепсин	K _i , нМ	Участок узнавания	Источники
Цистатин А	CtsB CtsH CtsL CtsS	8,2 0,31 1,3 0,05	QVVAG	[74,75]
Цистатин В	CtsB CtsH CtsL CtsS	73 0,58 0,23 0,07		
Цистатин С	CtsB CtsH CtsL CtsS	0,25 0,28 <0,005 0,008	QIVAG	
Цистатин D	CtsB CtsH CtsL CtsS	>1000 8,5 25 0,24	QVMAA	[75]
Цистатин SN	CtsB	19	QTVGG	
Цистатин M/E	CtsB CtsL CtsV	32 1,78 0,47	QLVAG	[76,77]
Цистатин F	CtsL	0,31	QIVKG	[78]

Вариабельность ингибиторов ППЦП позволяет регулировать протеолитическую сеть организма. В зависимости от наличия в среде тех или иных ингибиторов одновременно одни ППЦП могут быть активны, а другие – нет. Активность третьих, в то же время, может быть частично ингибирована. Например, CtsK может выполнять функции коллагеназы, в то время как его эластолитическая активность может быть подавлена [69]. Сложность этой

протеолитической сети позволяет поддерживать гомеостаз. Однако есть и другие способы регулирования активности ППЦП, которые будут обсуждаться далее.

1.2.2 Неингибирующее влияние макромолекул на активность протеиназ

Первая макромолекула, которая оказывает влияние на специфичность протеиназы, это её же продомен. Основная роль продомена заключается в подавлении активности ППЦП вплоть до транспортировки в конечный компартмент и/или возникновения необходимых условий для работы фермента. Однако пропептиды также могут способствовать правильной укладке аминокислотной последовательности протеиназы после экспрессии и формировать специфичность ППЦП. При помощи структурных методов, а также изучения каталитической активности и специфичности папаина и его мутантов, было показано, что замены на участке продомена, взаимодействующем с S₂-сайтом зимогена папаина, меняли специфичность зрелого фермента к P₂-положению субстрата [79,80]. По сравнению с диким типом мутанты Ile112Leu были более активны в отношении разветвленных гидрофобных, а Ile112Phe – ароматических а.о. [79,80]

Помимо пропептидов, протеиназная активность цистеиновых катепсинов может регулироваться гликозаминогликанами. Как уже было сказано, они могут ингибировать эластазную активность CtsK и V, однако на коллагеназную активность CtsK эти молекулы действуют по-разному [69,81,82]. В то время как гепарансульфат и гепарин ингибируют коллагеназную активность CtsK, хондроитин и кератансульфат, наоборот, могут её повысить [6]. Для гидролиза тройной спирали коллагена CtsK взаимодействует с хондроитинсульфатом, который связывает и ориентирует несколько молекул протеиназы в комплексе "бусины на нитке" [81,83]. Тем не менее замена а.о. в сайте связывания хондроитина на поверхности CtsK на соответствующие а.о. CtsL не подавляла коллагеназную активность фермента [6]. Также было показано, что

гликозаминогликаны и сульфат декстрана стимулируют автокаталитическую активность цистеиновых катепсинов человека [84,85].

Специфичность ППЦП весьма вариабельна, что может быть связано со структурными различиями этих протеиназ. В дополнение к этому, другие биомолекулы, такие как цистатины и гликозаминогликаны, достаточно тонко регулируют активность ППЦП. Даже продомен в этих протеиназах имеет значение для активности зрелого фермента. Условия среды, в которых протекает реакция, влияют на межмолекулярные взаимодействия и, следовательно, также могут вносить вклад в регуляцию протеолитической активности ППЦП.

1.2.3 Влияние ионов на протеолитическую активность

Наличие или отсутствие определенных ионов может существенно влиять на взаимодействие между молекулами. Ионы могут как экранировать заряды а.о. в субстрат-связывающей борозде протеиназы, тем самым предотвращая образование солевых мостиков, так и увеличивать силу гидрофобных взаимодействий [86,87]. В то время как CtsC нужны галогенид-ионы для активности в отношении субстратов GF-PNA и GF-4MβNA (PNA – p-нитроанилид, 4MβNA – 4-метокси-β-нафтиламид) при pH от 3,6 до 8,2, его усеченная версия CtsCΔEx в них не нуждается [88,89]. Однако оптимальная активность CtsCΔEx по отношению к Z-LR-AMC при pH 5,6 наблюдалась при концентрациях хлорида натрия от 50 до 200 mM [48]. А CtsL и V, наоборот, гидролизovali субстраты Abz-ALR↓SSKQ-EDDnp, Abz-KLR↓SSKQ-EDDnp и Abz-ELR↓SSKQ-EDDnp (Abz – 2-аминобензойная кислота, EDDnp – N-(2,4-динитрофенил)-этилендиамин) хуже в присутствии хлорида натрия [4]. Гидролиз Abz-KLR↓SSKQ-EDDnp под действием CtsV оказался наиболее чувствительным к соли, что может указывать на важность электростатических взаимодействий между положительно заряженным а.о. в P₃-положении субстрата и CtsV при pH 5,5. Хлорид натрия также снижает активность CtsV при pH 5,5 в отношении эластина, но не субстрата Z-FR-MCA [4,90].

Активность растительных протеиназ тоже регулируется ионами. Ранее Matagne et al. удалось выделить отдельные ППЦП из сырого экстракта протеиназ *A. comosus*, обычно называемого "бромелаином", и исследовать их биохимические свойства [91]. Активность всех ППЦП, входящих в состав бромелаина, по отношению к субстратам *Z-RR-AMC* и *Z-FR-AMC* полностью ингибируется 1 мМ Zn^{2+} при pH 6,25, что характерно для ППЦП [92–94]. 10 мМ концентрация ионов Mg^{2+} , Ca^{2+} и Mn^{2+} значительно подавляли активность основных форм бромелаина 1, 2 и 3, форм 1 и 2 ананаина и комозаина [91]. Снижение активности кислого бромелаина формы 1 на 25% наблюдалось в присутствии 10 мМ Mn^{2+} , но не 10 мМ Mg^{2+} или Ca^{2+} . Концентрации от 0,1 до 1 мМ Mg^{2+} , от 0,1 до 0,5 мМ Mn^{2+} и Ca^{2+} повышали активность ананаина формы 1. Свойства ананаина формы 2 были аналогичны форме 1, однако повышение активности наблюдалось при концентрациях от 0,1 до 5 мМ Mg^{2+} и 0,1 мМ Mn^{2+} . Концентрации от 0,1 до 1 мМ Mn^{2+} также повышали активность комозаина [91]. Активность зингибаина из имбиря *Zingiber officinale* в отношении бычьего казеина при pH 7,0 повышается в присутствии 2 мМ Ca^{2+} , но снижается при наличии Zn^{2+} , Mn^{2+} и Mg^{2+} [95]. Молокосвёртывающая активность процераина В из калотрописа *Calotropis procera* повышается при концентрациях Mn^{2+} от 0,01 до 10 мМ, но полностью исчезает при тех же концентрациях Zn^{2+} при pH 7,5 [96]. Казеинолитическая активность микрокарпаина из фикуса *Ficus microcarpa* снижается на 42%, 79%, 85% и 92% ионами Ca^{2+} , Mg^{2+} , Mn^{2+} и Zn^{2+} , соответственно, в концентрации 5 мМ при pH 8,0 [91,97]. Pd_dinase из *P. distasonis* ингибируется избытком Zn^{2+} путем взаимодействия с Cys56, His340 и Asp298 при pH 8,0. В результате петля из а.о. 290-310 закрывает активный центр фермента [56].

За исключением Zn^{2+} , ингибирующего ППЦП, другие ионы могут либо подавлять (как правило, двухвалентные ионы при концентрациях >5 мМ), либо стимулировать активность ППЦП в зависимости от фермента. По-видимому, заряженные а.о. могут играть важную роль в связывании субстрата, тогда как ионы экранируют эти а.о. С другой стороны, большая ионная сила способствует гидрофобным взаимодействиям [86,87].

1.2.4 Зависимость активности от значения pH

От значения pH зависит наличие зарядов на а.о., а следовательно – эффективность внутри- и межмолекулярных взаимодействий [41]. Например, активация CtsB происходит быстрее при кислом pH [6,98], поскольку пропептид CtsB более эффективно взаимодействует с субстрат-связывающей бороздой фермента при pH 6,0, чем при pH 4,0 [99]. При pH 5,5 эластазная активность CtsCΔEx ниже, чем у других катепсинов. Однако при pH 7,4 эластин расщепляется CtsCΔEx более эффективно, чем CtsB, L и V, хотя менее эффективно, чем CtsK и S [48]. Основные формы бромелаина 1 и 2 гидролизуют pGlu-FL-PNA (pGlu – пироглутаминовая кислота) на диапазоне pH от 4,0 до 4,5, а для гидролиза Z-RR-AMC и Z-FR-AMC оптимальным является значение pH 6,0 [91,100]. Оптимальные значения pH для ферментативной активности CtsL-подобной протеиназы Tc-cpl-1 из гельминта *Toxocara canis* в отношении субстратов Z-FR-AMC и Z-RR-AMC составляют 5,0 и 6,5, соответственно [101]. Снижение соотношения эффективности расщепления субстратов Z-FR-AMC к Z-RR-AMC меняется с 16:1 при pH 5,0 до примерно 1:1 при pH 6,5 из-за депротонирования Glu в S₂-сайте фермента [101,102]. По той же причине активность круципаина из *T. cruzi* на диапазоне значений pH от 3,5 до 10,0 растёт в отношении субстратов Z-RR-AMC и Z-RA-AMC, но остается неизменной в отношении Z-FR-AMC [55]. FhCL2 из *F. hepatica* в два раза эффективнее расщепляет Z-PR-AMC при pH 5,5 по сравнению с pH 7,3 [49]. Однако FhCL2 и FhCL3 могут гидролизовать нативный коллаген при нейтральных pH, тогда как CtsK требуются кислые условия среды [9]. Тритикаин-α из *T. aestivum* эффективно расщепляет коллаген только при pH от 3,6 до 4,6 и от 6,0 до 6,5, а глютен – на широком диапазоне pH от 3,0 до 6,5 [22].

Значения pK_a боковых групп а.о. определяют различную специфичность при разных значениях pH. На Рисунке 4 мы можем заметить существенные различия между зарядами на поверхности цистеиновых катепсинов в зависимости от значения pH. Эта зависимость от кислотности среды влияет на сродство к

различным субстратам и может быть использована для разработки специфических ингибиторов, подавляющих ферменты при определенных значениях pH [103]. В частности, можно разработать ингибиторы, которые будут снижать избыточную активность катепсинов человека только при нейтральных pH внеклеточного пространства и цитозоля.

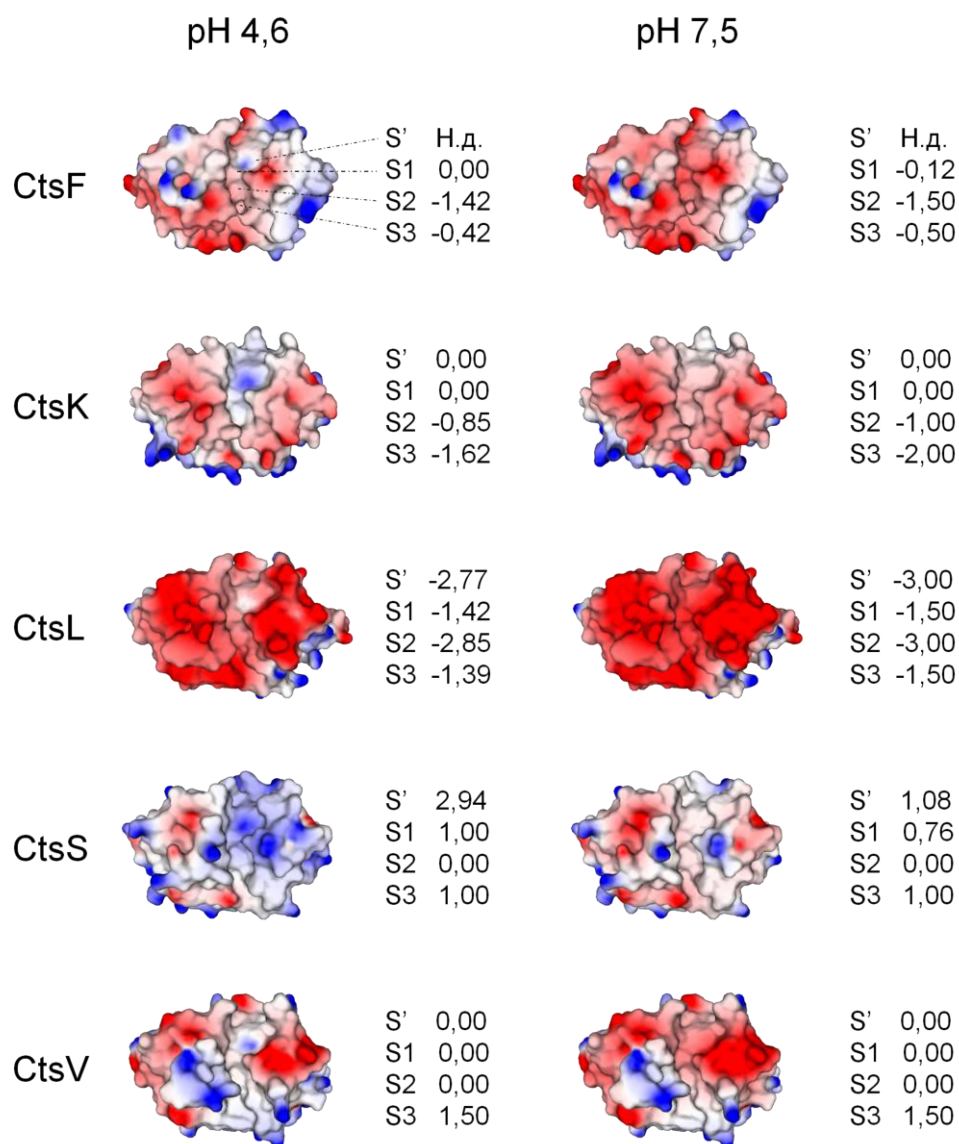


Рисунок 4 – Распределение зарядов на поверхности CtsF, K, L, S и V при pH 4,6 и 7,5 (PDB ID: 1M6D, 4X6H, 2XU3, 6YYR и 7QGW, соответственно) [28,104–107]. Поле электростатического потенциала было смоделировано с помощью программы ProKSim (Protein Kinetics Simulator) [108,109]. Поверхность белка окрашена в зависимости от величины электростатического потенциала от -100 мВ (красный) до +100 мВ (синий). Приблизительные суммарные заряды для сайтов связывания показаны справа от каждой структуры. Их рассчитывали как сумму зарядов а.о., идентифицированных в соответствующих сайтах на момент написания данной работы. Расположение сайтов связывания указано на структуре CtsF и на Рисунке 1. Н.д. – нет данных.

1.3 Папаин-подобные цистеиновые протеиназы в медицине

1.3.1 Патогенная активность цистеиновых катепсинов человека

Микроокружение злокачественных опухолей характеризуется повышенной концентрацией и высокой активностью цистеиновых катепсинов. Помимо самих клеток опухоли, катепсины также секретируются ассоциированными с опухолью Т-клетками, фибробластами, остеокластами, нейтрофилами, тучными клетками, миоэпителиальными и эндотелиальными клетками (Таблица 2) [110]. Считается, что транспорт лизосом из околядерного пространства на периферию клетки во время онкогенеза может приводить к высвобождению катепсинов во внеклеточное пространство, а закисление внеклеточной среды стимулирует активность этих протеиназ [111]. Другой причиной наличия катепсинов во внеклеточном матриксе (ВКМ) может быть снижение экспрессии гена рецептора инсулиноподобного фактора роста 2 (рецептор маннозо-6-фосфата), который вовлечен в транспорт белков к лизосомам [112].

Эндогенные ингибиторы катепсинов, присутствующие в ВКМ, сами могут расщепляться протеолитическими ферментами. Например, было показано, что матриксная металлопротеиназа 2 (ММП2) и аспартатный про-CtsD могут гидролизовать цистатин С [113]. Также ММП расщепляют цистатины Е и М, а про-CtsD активирует про-CtsB [114]. Активированные катепсины в свою очередь способствуют ангиогенезу, пролиферации и метастазированию. Считается, что наибольший вклад в онкогенез вносят CtsB, L, S и Z [115]. Повышение их концентрации в сыворотке крови обычно ассоциировано с плохим прогнозом при онкологических заболеваниях. Вклад того или иного катепсина в патологические процессы может варьироваться в зависимости от типа опухоли и её расположения.

Таблица 2 – Секретия катепсинов в различных типах опухолей [116].

Катепсин	Опухолевые клетки	Ассоциированные с опухолью клетки			
		Макрофаги	Фибробласты	Эндотелиальные клетки	Остеокласты
CtsB	Опухоль головного мозга [117], рак кожи [118], рак молочной железы [119], рак легкого [120], рак желудка [121], рак мочевого пузыря [122], рак шейки матки [123], рак яичника [124], колоректальный рак [125], гепатоцеллюлярная карцинома [126], рак поджелудочной железы [127], рак щитовидной железы [128]	Опухоль головного мозга [129], рак молочной железы [130], рак поджелудочной железы [131]	Рак кожи [60], рак молочной железы [132], колоректальный рак [133], гепатоцеллюлярная карцинома [134], рак поджелудочной железы [131]	Опухоль головного мозга [135], рак молочной железы [136], рак желудка [137]	Рак молочной железы [138], рак щитовидной железы [139]
CtsC	Рак кожи [140]	Н.д.	Рак кожи [140]	Н.д.	Н.д.
CtsF	Рак шейки матки [141]	Н.д.	Н.д.	Н.д.	Н.д.
CtsH	Опухоль головного мозга [142], рак кожи [143], рак предстательной железы [144], колоректальный рак [145]	Н.д.	Н.д.	Н.д.	Н.д.
CtsK	Рак кожи [146], рак молочной железы [147], рак легкого [148], рак желудка [149], рак предстательной железы [150]	Рак легкого [151]	Рак кожи [152], рак молочной железы [153]	Н.д.	Опухоль кости [154], рак молочной железы [155], рак предстательной железы [84], рак почки [156]
CtsL	Опухоль головного мозга [86], рак кожи [87], рак молочной железы [157], рак легкого [158], рак желудка [159], рак яичника [160], колоректальный рак [161], рак поджелудочной железы [162]	Рак поджелудочной железы [131]	Рак кожи [94], рак поджелудочной железы [59]	Н.д.	Н.д.
CtsS	Опухоль головного мозга [163], рак молочной железы [164], рак желудка [165], рак предстательной железы [166], колоректальный рак [167], гепатоцеллюлярная карцинома [168], рак поджелудочной железы [169]	Рак молочной железы [170], рак предстательной железы [171], колоректальный рак [172], гепатоцеллюлярная карцинома [173]	Н.д.	Гепатоцеллюлярная карцинома [174]	Н.д.
CtsV	Рак щитовидной железы [175]	Н.д.	Н.д.	Н.д.	Н.д.

Продолжение Таблицы 2

Катепсин	Опухолевые клетки	Ассоциированные с опухолью клетки			
		Макрофаги	Фибробласты	Эндотелиальные клетки	Остеокласты
CtsZ	Рак желудка [176], рак предстательной железы [177], колоректальный рак [178], гепатоцеллюлярная карцинома [179], рак поджелудочной железы [180]	Рак желудка [181], колоректальный рак [178]	Н.д.	Н.д.	Н.д.

Катепсины расщепляют молекулы клеточной адгезии, такие как интегрины и Е-кадгерины, тем самым разрушая межклеточные контакты опухолевых клеток и способствуя их миграции и метастазированию [115]. Е-кадгерины в основном расщепляются CtsB, L и S, которые также гидролизуют и другие белки ВКМ [110]. Было показано, что CtsB и S расщепляют коллаген, ламинин, фибронектин и тенасцин-С [182]. Кроме того, катепсины разрушают компоненты базальной мембраны, способствуя инвазии опухолевых клеток в кровеносные и лимфатические сосуды, что приводит к появлению метастаз [183]. CtsH способствует миграции опухолевых клеток, расщепляя талин, белок, который регулирует активацию интегрин и клеточную адгезию [184]. CtsK гидролизует коллаген I и остеонектин в костях, что приводит к остеолизу и метастазированию в костную ткань [185]. CtsS играет решающую роль в метастазировании клеток рака молочной железы в головной мозг. Этот фермент расщепляет молекулы адгезии семейства JAM (Junctional Adhesion Molecule), в частности JAM-B, что приводит к нарушению функции гематоэнцефалического барьера. Эта реакция возможна благодаря активности CtsS при нейтральных значениях pH в сосудах головного мозга [186]. CtsZ способствует миграции, адгезии и инвазии опухолевых клеток путем инактивации опухолевого супрессора профилин 1 как в перинуклеарном пространстве, так и на плазматической мембране [177]. Он также активирует ММП2, ММП3 и ММП9, которые расщепляют белки ВКМ в гепатоцеллюлярной карциноме. Также CtsZ способствует эпителиально-мезенхимальному переходу путем активации экспрессии генов мезенхимальных маркеров (фибронектина и виментина) одновременно с подавлением экспрессии

генов эпителиальных маркеров (Е-кадгерина и α -катенина) [179]. Кроме того, зимоген про-CtsZ может взаимодействовать с интегринами, расположенными внутри и вне клетки, что приводит к активации сигнального пути пролиферации FAK/Src (FAK – киназа фокальных контактов; Src – коактиватор стероидных рецепторов), ремоделированию актинового цитоскелета и повышению инвазивности опухолевых клеток [187].

Деградация белков ВКМ катепсинами может приводить к образованию проангиогенных фрагментов, например, в реакции расщепления CtsS ламинина 5 [172]. Катепсины также могут активировать другие ферменты и запускать сигнальные пути, приводящие к экспрессии генов различных факторов роста и цитокинов. Например, CtsL активирует гепараназу, которая в свою очередь регулирует ангиогенез и лимфангиогенез посредством высвобождения факторов роста [188]. CtsS может расщеплять некоторые антиангиогенные факторы, такие как канстатин и аррестин [189], тогда как в результате расщепления CtsL и S коллагена XVIII, наоборот, образуется антиангиогенный пептид эндостатин [190].

Избыточная активность катепсинов в опухолевых клетках часто обусловлена нарушением работы эндогенных ингибиторов цистеиновых протеиназ, цистатинов. Поэтому поиск экзогенных ингибиторов цистеиновых катепсинов является перспективным направлением в разработке новых противоопухолевых средств. Было показано, что ингибиторы катепсинов могут подавлять ангиогенез, инвазию опухолевых клеток и метастазирование.

1.3.2 Разработка селективных ингибиторов цистеиновых катепсинов человека

Активность протеиназ обычно регулируют путем их ингибирования. Ингибиторы ППЦП в основном взаимодействуют с сайтами связывания от S_3 до S_2' [28]. Таким образом, зная структуру и специфичность каждого сайта связывания, можно разработать эффективный ингибитор. E-64 является неселективным ингибитором всех цистеиновых катепсинов кроме CtsC [6,191]. Он был модифицирован для получения специфичного к карбоксипептидазе CtsB

ингибитора путем имитации С-концевого дипептида заменой гуанидинобутиламина на Pro со свободной карбоксильной группой [192,193]. Различные модификации N-конца ингибитора привели к разработке новых малых молекул CA030 и CA074 [192,193]. Для CtsB, H и L K_i ингибитора CA030 составили 4,38, 42900 и 40000 нМ, тогда как K_i ингибитора CA074 получились 1,94, 75000 и 233000 нМ, соответственно [192]. Кристаллическая структура комплекса CtsB с CA030 подтверждает ключевую роль His189 и His190, расположенных в окклюзионной петле CtsB, в распознавании С-концевой карбоксильной группы в сайте связывания S₂' (Рисунок 5) [27]. CtsZ также может быть ингибирован CA074, когда имидазольное кольцо His84 повернуто в положение, эквивалентное положению, занятому His189 в CtsB, что открывает доступ к карману S₂' (Рисунок 5) [3].

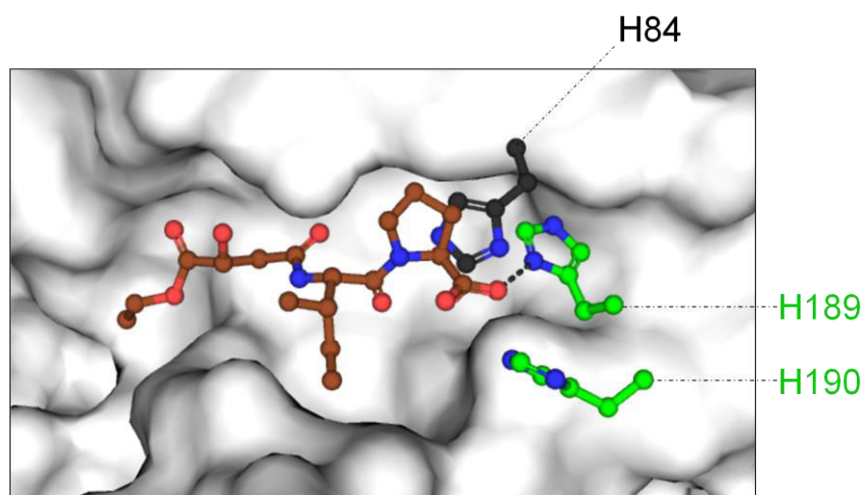


Рисунок 5 – Кристаллическая структура комплекса CA030 (коричневый), связанного с CtsB (серая поверхность; PDB ID: 1CSB) [26,27]. В программе для визуализации молекул PyMOL кристаллическая структура CtsZ (PDB ID: 1EF7) была сопоставлена со структурой CtsB для сравнения локализации а.о., участвующих во взаимодействиях с ингибитором [3,31]. А.о. CtsB окрашены в зелёный цвет, CtsZ – в черный. His84 CtsZ показан в положении, закрывающем карман S₂'. Пунктирная линия указывает возможную нековалентную связь.

Измерение активности цистеиновых катепсинов в присутствии винилсульфоновых ингибиторов с различными заместителями в P₂-положении показали, что наиболее эффективными для CtsS являются ингибиторы, содержащие Leu или Phe в P₂-положении [194]. Они ингибируют CtsS с K_i 13 нМ

и 18 нМ соответственно, при этом показывают менее эффективное ингибирование других цистеиновых катепсинов: соответствующие значения K_i составили 110 нМ и 36000 нМ для CtsK и 120 нМ и 110 нМ для CtsL [194]. Основываясь на кристаллической структуре комплекса был сделан вывод, что хорошее связывание Leu или Phe в S_2 -кармане CtsS достигается за счёт конформационной подвижности Phe325 [14]. В CtsL и V в этом положении находится Ala, а в CtsK – Leu (Рисунок 6). Другой отличительной структурной особенностью является Lys178 в кармане S_3 CtsS. В исследованных комплексах частично отрицательно заряженный кислород морфолинового кольца в P_3 -положении располагается рядом с аминокислотой Lys178 [14]. Положительно заряженный Lys178 заменен кислыми а.о. в CtsK и L и полярным незаряженным Gln в CtsV (Рисунок 6) [14].

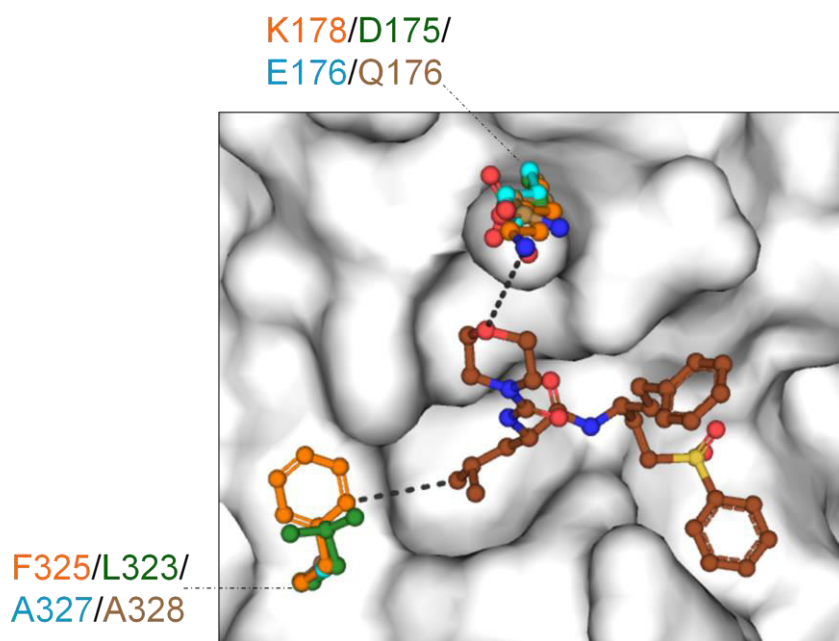


Рисунок 6 – Кристаллическая структура комплекса винилсульфонового ингибитора (коричневый) с Leu в P_2 -положении, связанного с CtsS (серая поверхность; PDB ID: 1NPZ) [14,26]. В программе для визуализации молекул PyMOL кристаллические структуры CtsK, L и V (PDB ID: 1ATK, 1CJL и 1FH0, соответственно) были выравнены со структурой CtsS для сравнения локализации а.о., которые могут участвовать во взаимодействиях с ингибитором [25,29–31]. А.о. CtsS окрашены в оранжевый цвет, CtsK – в зелёный, CtsL – в голубой, CtsV – в бежевый. Пунктирные линии указывают возможные нековалентные связи.

Ковалентный ингибитор N-десметил талассоспирамид С был разработан для ингибирования различных цистеиновых катепсинов [195]. Эксперименты показали, что ингибитор селективен в отношении CtsK и L (IC_{50} 3 нМ и 1 нМ,

соответственно) и менее эффективно ингибирует CtsB и S (IC_{50} 65 нМ и 46 нМ, соответственно) [195]. Комплекс ингибитора с CtsK кристаллизовали и провели его детальный анализ [195]. Макроцикл ингибитора располагается в S_1' -сайте связывания фермента и взаимодействует с Gln133 и Asn275 (Рисунок 7). Карман S_1' CtsL похож на карман CtsK и стерические препятствия для связанной конформации ингибитора отсутствуют [105]. В то же время замена Ser252Arg как в CtsS приводит к стерическим помехам для связывания N-десметил талассоспирамида С (Рисунок 7А) [196], а у CtsB укладка полипептидной цепи в сайте S_1' отличается и является уникальной в семействе катепсинов (Рисунок 7Б) [197]. Эти конформационные различия могут объяснять более низкую эффективность связывания ингибитора с CtsB и S [195].

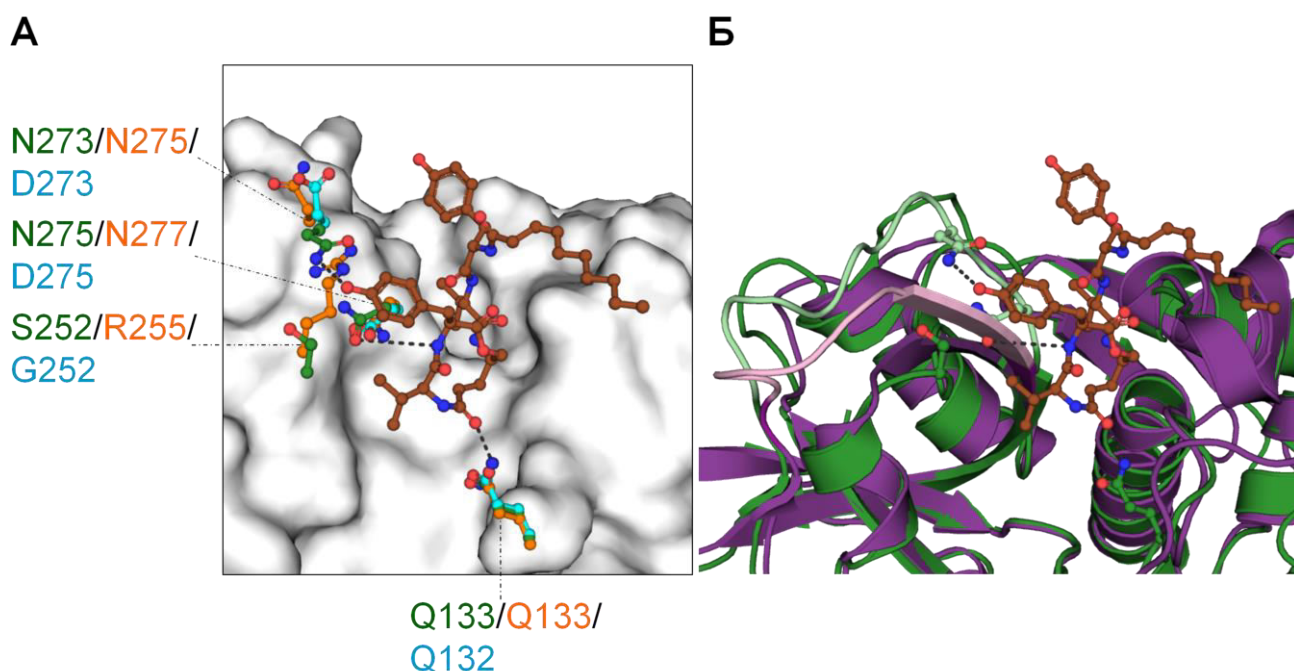


Рисунок 7 – Кристаллическая структура комплекса ковалентного ингибитора N-десметил талассоспирамида С (коричневый), связанного с CtsK (PDB ID: 6HGY) [26,195]. А – CtsK изображен в виде серой поверхности. В PyMOL кристаллические структуры CtsL и S (PDB ID: 1CJL и 2FQ9, соответственно) были выравнены со структурой CtsK для сравнения локализации а.о., участвующих во взаимодействиях с ингибитором [25,31]. А.о. CtsK окрашены в зелёный цвет, CtsL – в голубой, CtsS – в оранжевый. Б – CtsK (зелёный) и CtsB (фиолетовый; PDB ID: 1CSB) были выравнены в PyMOL для сравнения локализации частей молекул, участвующих во взаимодействиях с ингибитором [27,31]. Фрагменты CtsK и B, различающиеся укладкой полипептидных цепей, окрашены в светло-зелёный и розовый цвет, соответственно. Пунктирные линии указывают возможные нековалентные связи.

Так как специфичность CtsB может меняться в зависимости от значения pH, для CtsB смогли разработать специфический pH-зависимый ингибитор [103]. При pH 7,2 CtsB предпочитает положительно заряженные Arg и Lys, а также Nle и Tyr в P₁-положении. В S₂-кармане CtsB расположен Glu324, поэтому он специфичен к положительно заряженным Arg, Lys и His, а также Trp в P₂-положении [57]. При pH 4,6 CtsB предпочитает незаряженные Thr и Gly либо положительно заряженный Arg в P₁, а в P₂-положении CtsB проявляет специфичность либо к кислому Glu, либо к гидрофобному Val. Glu324 может взаимодействовать с Glu субстрата и не обладать селективностью в отношении положительно заряженных а.о., так как при pH 4,6 боковая группа Glu протонирована и не несёт заряд [103]. CtsB показал самое высокое соотношение активностей при pH 7,2/4,6 с субстратом Z-RK-AMC и самое высокое соотношение активностей при pH 4,6/7,2 с субстратом Z-EK-AMC среди исследованных Z-RK-AMC, Z-KK-AMC, Z-KR-AMC, Z-RR-AMC, Z-EK-AMC и Z-ER-AMC. Активность CtsL и V в отношении Z-RK-AMC при pH 7,2 и Z-EK-AMC при pH 4,6 не была детектирована. Замена AMC на AOMK привела к получению пептидных ингибиторов Z-RK-AOMK и Z-EK-AOMK [103]. Z-RK-AOMK продемонстрировал селективное ингибирование CtsB при pH 7,2 по сравнению с pH 4,6 со значениями K_i 130 нМ и 15000 нМ, соответственно. Z-EK-AOMK показал менее эффективное ингибирование CtsB при обоих значениях pH. Z-RK-AOMK ингибировал CtsB с IC₅₀ 20 нМ при pH 7,2, что эффективнее, чем ингибирование CtsC, H, K, S и V с IC₅₀ > 440 нМ. Компьютерное моделирование связывания Z-RK-AOMK с CtsB выявило, что Arg в P₂-положении взаимодействует с карбоксильной группой Glu324 в S₂-сайте связывания. В то же время Lys ингибитора взаимодействует с Asn151 и Glu201 в сайте S₁ фермента [103]. Также было показано, что Z-RK-AOMK может ингибировать CtsB в концентрации 1 мкМ в клетках нейробластомы человека SHSY5Y [103].

Как мы видим, данные о структурных особенностях ППЦП и их влиянии на специфичность позволяют разрабатывать новые эффективные ингибиторы, которые не только обладают селективностью к определенным протеиназам, но и

чувствительны к условиям среды (Таблица 3). Применение этих ингибиторов в медицине может снизить риски побочных эффектов по сравнению с ингибиторами широкого спектра действия.

Таблица 3 – Примеры селективных ингибиторов цистеиновых катепсинов человека [24].

Катепсин	Ингибитор	K _i , нМ	Источник
CtsB	CA030	4,38	[192]
	CA074	1,94	
	<i>Z</i> -RK-AOMK	130 (при pH 7,2) 15000 (при pH 4,6)	[103]
CtsS	4-морфолинкарбонил-Leu-hPhe-Ψ (CH=CH-SO ₂ -фенил)	13	[194]
	4-морфолинкарбонил-Phe-hPhe-Ψ (CH=CH-SO ₂ -фенил)	18	
Катепсин	Ингибитор	IC ₅₀ , нМ	Источник
CtsK	N-десметил талассоспирамид С	3	[195]
CtsL		1	

1.3.3 Применение растительных протеиназ в медицине

В злокачественных опухолях наблюдается избыточная активность цистеиновых катепсинов, и в качестве одного из способов лечения онкозаболеваний предлагается применение ингибиторов катепсинов для подавления этой активности. Однако в некоторых патологических состояниях, наоборот, добавление экзогенных протеиназ может быть необходимым для проведения эффективного лечения (Рисунок 8). Растительные ППЦП папаин и бромелаин активно применяются в медицинских целях и входят в состав некоторых препаратов [198]. Коллагенолитическая активность этих ферментов используется для заживления ран и при лечении ожогов в составе таких препаратов как "Debridase" (папаин), "NexoBrid" (бромелаин) и "Debricin" (фицин) [18–20]. Также папаин применяется в стоматологии: коллагеновые фибриллы могут быть частично разрушены папайновым гелем ("Papacarie" и "Carie Care")

при проведении хемомеханического удаления кариеса [21]. А раствор химопапаина используют для уменьшения размера межпозвонковых грыж [199]. Растительные протеиназы потенциально могут быть использованы и в терапии опухолевых заболеваний. Было показано, что бромелаин подавляет метастазирование и *in vitro* расщепляет рецепторы: интегрин $\alpha 3\beta 1$ и рецептор гиалуронана CD44, которые участвуют в инвазии клеток глиомы [200]. Кроме того, папаин и бромелаин могут снижать количество блокирующих факторов, которые продуцируются опухолевыми клетками для защиты от распознавания иммунной системой [201]. И наоборот, эти протеиназы могут уменьшать количество цитокинов, таких как лимфотоксин- α , способствуя иммуносупрессии [198,201].

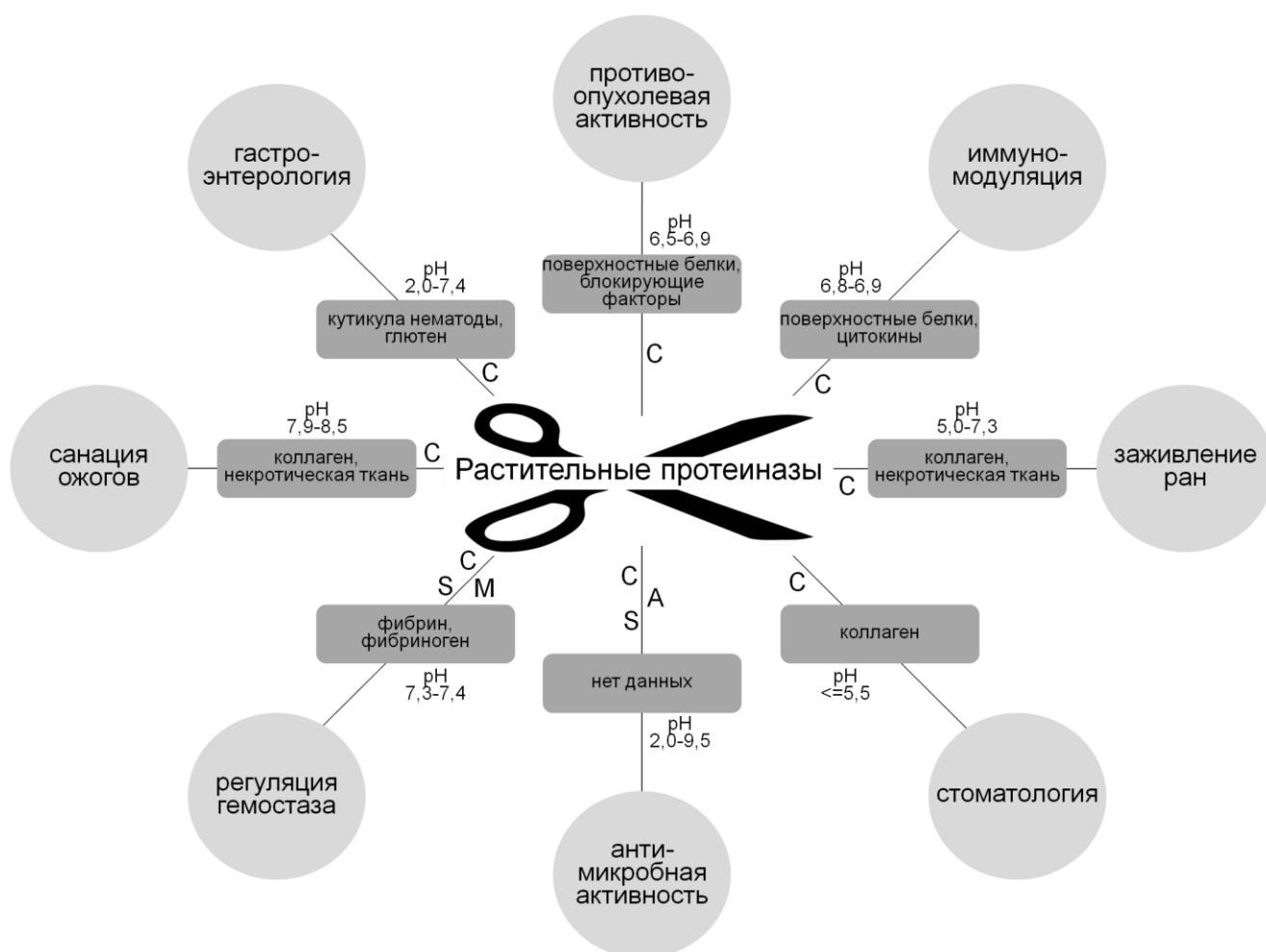


Рисунок 8 – Применение различных семейств растительных протеиназ в медицине [198]. С – цистеиновые протеиназы, S – сериновые протеиназы, А – аспаргатные протеиназы, М – металлопротеиназы.

Было обнаружено, что некоторые растительные протеиназы могут процессировать фибрин и фибриноген и, таким образом, влиять на процессы свёртывания крови. Папаин и пергуларин е I из *Pergularia extensa* могут расщеплять фибриноген подобно тромбину, что приводит к формированию фибриновых сгустков [202,203]. Способность расщеплять фибрин подобно урокиназному активатору плазминогена была обнаружена у бромелаина и сериновых протеиназ из водорослей: кодказы из *Codium fragile*, протеиназы CLP из *Codium latum* и протеиназы CCP из *Costaria costata* [204–207]. Следовательно, эти ферменты можно применять в качестве тромболитических агентов. Особенно многообещающими кажутся сериновые протеиназы из белокопытника *Petasites japonicus* и кодказа, которые, как было показано на мышах, эффективнее защищают от тромбоза, чем урокиназный активатор плазминогена, и при этом не вызывают кровотечений [205,208].

Некоторые растительные протеиназы обладают антибактериальной активностью. Папаин и бромелаин проявляют бактерицидную активность в отношении грамположительных бактерий *Alicyclobacillus sp.* Папаин также известен своей антибактериальной активностью в отношении нескольких штаммов *Bacillus subtilis* и энтеропатогенов, таких как *Enterobacter cloacae*, *E. coli*, *Listeria monocytogenes*, *Salmonella typhimurium*, *Staphylococcus aureus*, *Proteus vulgaris*, *Pseudomonas aeruginosa* и *Klebsiella pneumoniae* [209,210]. Сообщалось также о противогрибковых свойствах двух аспартатных протеиназ StAP1 и StAP3 из *Solanum tuberosum* в отношении спор *Fusarium solani* и цист *Phytophthora infestans* [211,212]. StAP вызывают пермеабиллизацию мембран спор и гиф *F. solani* и цист *P. infestans* [213]. Дальнейшие исследования показали влияние дегликозилирования на противогрибковую активность обеих StAP, которое приводит к более низкой эффективности по сравнению с нативными формами фермента [214].

Некоторые цистеиновые протеиназы растений обладают антигельминтной активностью: они способны разрушать кутикулы нематод при pH 7,4 [215]. Было показано, что фицин, папаин, химопапаин, экстракты латекса молочая и

экстракты, содержащие бромелаин, эффективны против нематод *Heligmosomoides polygyrus* [216]. Папаин и цистеиновые протеиназы из дынного дерева *Carica candamarcensis* в исследованиях на модели язвы желудка у крыс регулировали количество соляной кислоты, этанола, пепсиногена, бикарбоната, простагландинов и т.д., оказывая защитное действие на слизистую желудочно-кишечного тракта [217,218]. Цистеиновая протеиназа P1G10 из *C. candamarcensis* действует как митоген на фибробласты и эпителиальные клетки и стимулирует ангиогенез и заживление ран в моделях язвы желудка и кожи [219]. Бромелаин также приводит к быстрому улучшению симптомов при лечении язвенного колита [220].

Многие растительные протеиназы обладают исключительными характеристиками, такими как высокая стабильность, уникальная субстратная специфичность и широкий диапазон pH для ферментативной активности. Эти ценные свойства делают протеолитические ферменты растительного происхождения перспективными для различных медицинских применений.

1.3.4 Цистеиновая протеиназа из пшеницы тритикаин-α в терапии непереносимости глютена

Одно из направлений медицины, в котором растительные протеиназы могут найти своё применение – это заместительная терапия при целиакии и непереносимости глютена [221]. Глютен накапливается в эндосперме семян злаковых и представляет собой смесь запасных белков глиадинов и глютеинов, которые содержат повторяющиеся последовательности, богатые Pro и Gln [222]. Продукты неполной деградации глютена в процессе пищеварения могут приводить к целиакии – аутоиммунному заболеванию, опосредованному генетической предрасположенностью [223]. К симптомам целиакии относят нарушение всасывания питательных веществ в тонком кишечнике, атрофию ворсинок, гиперплазию крипт и усиление инфильтрации эпителия кишечника лимфоцитами [222]. В данный момент единственным способом терапии целиакии

и непереносимости глютена является строгая безглютеновая диета [224]. Перспективным подходом к лечению целиакии является заместительная терапия с использованием различных протеиназ (глутеназ), способных инактивировать иммуногенные пептиды глютена в желудочно-кишечном тракте человека [225].

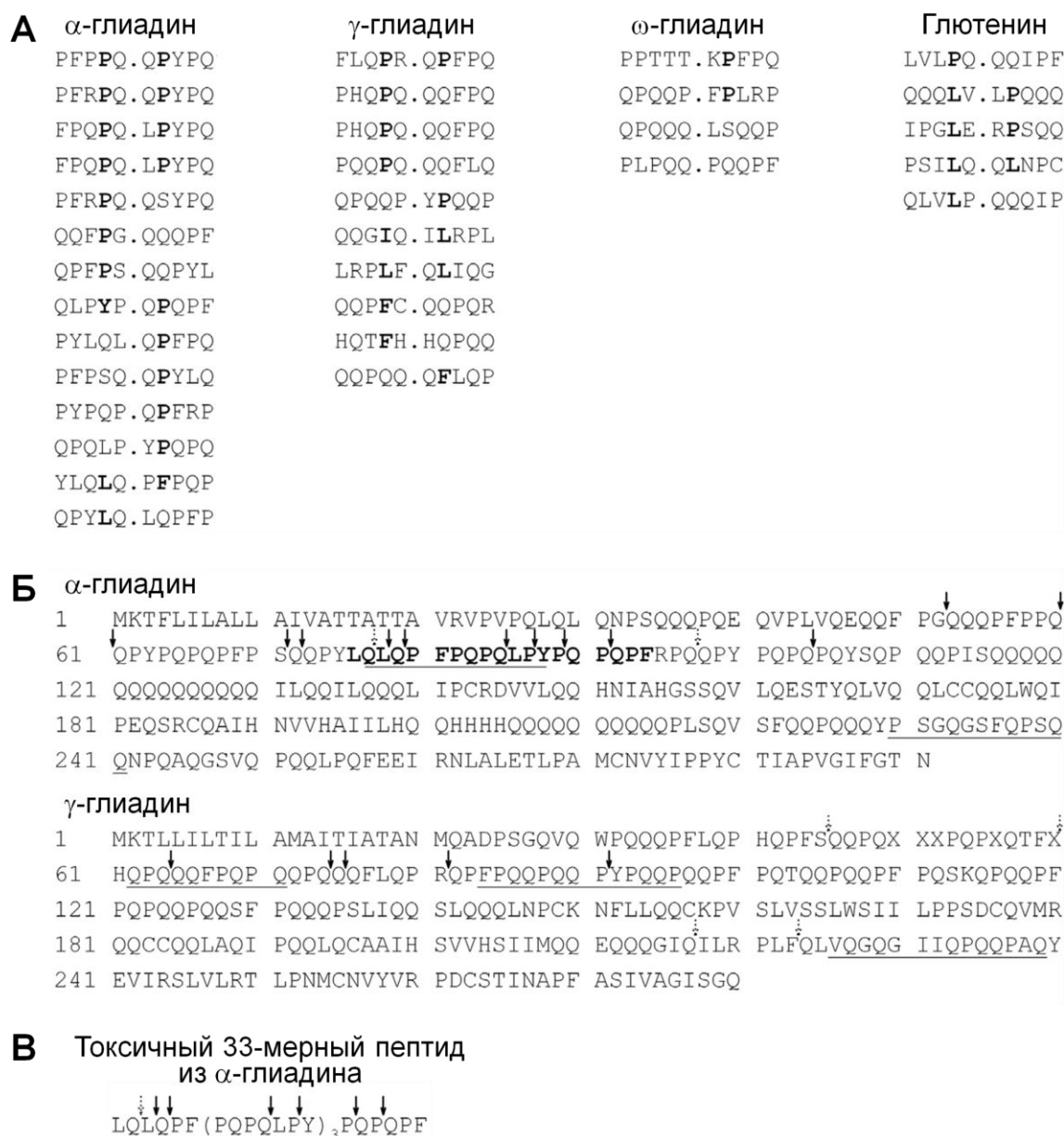


Рисунок 9 – Сайты расщепления тритикаином- α глютеиновых белков пшеницы, выявленные с использованием тандемной масс-спектрометрии MALDI-TOF/TOF [22]. А – Аминокислотные последовательности сайтов расщепления от положения P_5 до P_5' . Полужирным шрифтом выделены гидрофобные и ароматические а.о. в P_2 - и P_2' -положениях. Б – Аминокислотные последовательности глиадин- α и глиадин- γ . Указаны сайты расщепления после 5 мин (пунктирные стрелки) и 15 мин (сплошные стрелки) гидролиза. Подчеркнуты выявленные ранее токсичные пептиды [226,227]. В – Последовательность наиболее иммуногенного пептида из α -глиадина длиной 33 а.о. и сайты расщепления в нём [223].

В 2023 году завершилась вторая фаза клинических испытаний латиглютеназы в качестве препарата для пероральной терапии непереносимости глютена [23]. Латиглютеназа представляет собой смесь из цистеиновой протеиназы В (EP-B 2) ячменя *Hordeum vulgare* и пролилэндопептидазы бактерии *Sphingomonas capsulata* [228,229]. Параллельно разрабатываются ещё несколько альтернативных препаратов [23]. В России, например, было предложено использовать ППЦП тритикаин-α из пшеницы *T. aestivum* в качестве нового средства для терапии целиакии и непереносимости глютена [22,72,230]. Эта протеиназа продуцируется во время прорастания семян пшеницы для гидролиза запасных белков эндосперма, в том числе глютена [231]. Сайты расщепления тритикаин-α глютена пшеницы были идентифицированы методами масс-спектрометрии [22]. Они были обнаружены в большинстве ранее установленных токсичных пептидов, вызывающих воспалительные реакции в кишечнике, включая наиболее иммуногенный пептид длиной 33 а.о. из α-глиадина (Рисунок 9) [22]. В условиях *in vitro* было показано, что тритикаин-α активен на диапазоне значений pH от 3,4 до 6,5 при температуре 37°C и устойчив к пепсин-опосредованному протеолизу, что делает возможным пероральное применение этого фермента для лечения непереносимости глютена [22].

В совокупности эти данные указывают на большой потенциал тритикаина-α в разработке терапевтических средств для лечения целиакии [22,230]. Однако для получения эффективного и безопасного препарата на основе этой протеиназы необходимо провести дальнейшие её исследования, такие как подбор условий для экспрессии, получение трёхмерной структуры, описание субстратной специфичности и регуляции активности тритикаина-α.

2. МАТЕРИАЛЫ И МЕТОДЫ

2.1 Используемые в работе материалы

Реактивы: 2-меркаптоэтанол, бромфеноловый синий, гидроксид натрия, глицерин, диметилсульфоксид (ДМСО), лизоцим, минеральное масло, персульфат аммония, сефадекс G-25, соляная кислота ("Sigma-Aldrich"); 4-(2-гидроксиэтил)-1-пиперазинэтансульфоновая кислота (HEPES), акриламид, ампициллин, ацетат натрия, гидрофосфат натрия, глицин, дигидрофосфат натрия, додецилсульфат натрия (SDS), имидазол, канамицин, мочевиная, Твин 20, уксусная кислота ("PanReas AppliChem"); агар, дрожжевой экстракт, триптон, трис, этилендиаминтетрауксусная кислота (ЭДТА) ("Диаэм"); агароза, бромистый этидий, изопропил- β -D-1-тиогалактопиранозид (ИПТГ), хлорид натрия, *N,N'*-метилен-бисакриламид ("Хеликон"); аденозинтрифосфат, блокирующий буфер SuperBlock T20, маркер молекулярного веса белков, смесь дезоксирибонуклеотидов dNTP Mix, цитрат натрия ("Thermo FS"); буфер для нанесения проб ДНК и рибонуклеиновой кислоты (РНК), маркер молекулярного веса ДНК, олигонуклеотиды (Таблицы 4 и 5) ("Евроген"); вода для ПЦР ("Qiagen"); кумасси R-250 ("ПанЭко"); метанол ("Химмед"); этиловый спирт ("Ферейн"); АМС ("Permpic"); *N,N,N',N'*-тетраметилэтилендиамин (ТЕМЕД) ("Bio-Rad"); Ni-сефароза 6 Fast Flow resin ("Cytiva").

Наборы: набор для измерения концентрации белка Pierce BCA Protein Assay Kit, набор для трансформации TransformAid Bacterial Transformation Kit ("Thermo FS"); набор для очистки ДНК из геля и смесей Cleanup Mini Kit ("Евроген"); микрочипы с мечеными рТур пептидами ("JPT Peptide Technologies"); растворы для кристаллизации BCS (Basic Chemical Space), JCSG-Plus (Joint Center for Structural Genomics) Eco, LMB (Laboratory of Molecular Biology), PACT (pH, Anion, Cation crystallization Trial) premier и SG1 (ShotGun) Eco Screens ("Molecular Dimensions").

Таблица 4 – Олигонуклеотиды, которые были использованы в качестве праймеров для амплификации генов продомена (Triticain- α -ProD), каталитического домена (Triticain- α -CatD) тритикаина- α и гена фермента, включающего в себя оба этих домена (Triticain- α -GM), с расположенной на *N*- или *C*-конце меткой из шести His (6HIS). Сайты рестрикции подчеркнуты.

Генетическая конструкция	Прямой праймер/обратный праймер	Рестриктазы
pQE80L-6HIS-Triticain- α -ProD	ATGGATCCATCGTGTCTGTACGGGGAG/ TATTAAGCTTTTACTCCTCGTTGTCGTCG	<i>Bam</i> HI/ <i>Hind</i> III
pET28a-6HIS-Triticain- α -CatD	TATCATATGCTGCCGGAGACCGTTCG/ TATTAAGCTTTTAGCCCGTCTTCGTCGGG	<i>Nde</i> I/ <i>Hind</i> III
pET28a-6HIS-Triticain- α -GM	TATACATATGTCGATCGTGTCTGTACGG/ TATTAAGCTTTTAGCCCGTCTTCGTCGGG	<i>Nde</i> I/ <i>Hind</i> III
pET28a-Triticain- α -GM-6HIS	TATACCATGGCGGACATGTCTGTCTGTCTG/ TATTCTCGAGGCCCGTCTTCGTCGGGTAG	<i>Nco</i> I/ <i>Xho</i> I
pET15b-6HIS-Triticain- α -GM	TATACATATGTCGATCGTGTCTGTACGG/ TTCTCGAGTTAGCCCGTCTTCGTCGG	<i>Nde</i> I/ <i>Xho</i> I
pET15b-Triticain- α -GM-6HIS	TATACCATGGCGGACATGTCTGTCTGTCTG/ ATTCTCGAGTCAAGTGGTGGTGGTGGTGGGCC CGTCTTCGTCGGGT	<i>Nco</i> I/ <i>Xho</i> I

Таблица 5 – Последовательности олигонуклеотидов, которые были использованы для внесения точечных мутаций в ген Triticain- α -GM. Триплеты, вносящие мутацию, подчеркнуты.

Генетическая конструкция	Олигонуклеотид для мутагенеза
pET15b-6HIS-Triticain- α -GM ^{C154A}	GCGGGAGCGCTTGGGCTTTCTCAGC
pET15b-6HIS-Triticain- α -GM ^{E191N}	TCATACAACAACGGATGCAATGGAG
pET15b-6HIS-Triticain- α -GM ^{E191K}	TCATACAACAAGGATGCAATGGAG

Пептиды: *Ac*-DEV α D-AMC, *Ac*-FESR-AMC, *Ac*-PLVQ-AMC, *Ac*-QLLK-AMC, *Ac*-QLLR-AMC, *Ac*-VLPQ-AMC, Z-FVR-AMC (*Ac* – ацетил) ("Permic"); *Ac*-QLLD-AMC, *Ac*-QLLK-AMC, *Ac*-QLLQ-AMC, *Ac*-QLLR-AMC и AQLLRVGCVLGT ("Пептидные Технологии").

Ферменты: ДНК-лигаза T4 в комплекте с T4 DNA Ligase Buffer, РНКазы А, эндонуклеазы рестрикции FastDigest *Bam*HI, *Hind*III, *Nco*I, *Nde*I, *Xho*I в комплекте с FastDigest Green Buffer ("Thermo FS"); ДНК-полимеразы в Advantage 2 Polymerase Mix в комплекте с Advantage 2 PCR Buffer ("Clontech"); ДНКазы I ("New England BioLabs"); папаин из *C. papaya* ("Merck").

Антитела: моноклональные мышиные IgG1 антитела к pTyr P-Tyr-100 ("Cell Signaling", №9411); поликлональные козы IgG антитела к мышиным

иммуноглобулинам, конъюгированные с флуорофором DyLight 650 ("Thermo FS", №84545).

Плазмиды: pET15b, pET28a и pQE80L ("Merck Millipore"); pQE80L-6HIS-Triticain- α -GM [22].

Штаммы *E. coli*: BL21 (DE3) ("Thermo FS"); Rosetta gami B (DE3) и Rosetta gami 2 (DE3) pLysS ("Novagen").

Расходные материалы: наконечники для автоматических пипеток до 10 мкл ("ULPlust"); до 200 мкл, до 200 мкл для нанесения на гель и до 1000 мкл ("Corning"); до 5 мл и до 10 мл ("Thermo FS"); пробирки типа Эппендорф 0,5 мл PCR clean и 1,5 мл ("Eppendorf"); пробирки конические 15 мл и 50 мл ("Corning"); концентраторы центрифужные Amicon Ultra-4 10 кДа ("Merck"); планшеты черные 96 лунок ("Corning"); планшеты для кристаллизации белка на 96 лунок по 3 сидячие капли, микрорезервуарные полоски по 8 лунок объемом 5 мкл, наконечники-иглы для автоматизированной многоканальной станции дозирования ("SPT Labtech"); пленки для планшет Crystal Clear Sealing Film ("Hampton Research"); чашки Петри 90 мм ("Биомедикал"); мешки для диализа ("Sigma-Aldrich"); фильтровальная бумага ("Красная лента").

2.2 Составы буферов и сред

Электрофорез в агарозном геле

- Трис-ацетатный буфер: 40 mM трис, pH 7,6 (уксусная кислота), 1 mM ЭДТА, 0,5 мг/мл бромистый этидий.
- Агарозный гель: 1,5% агароза (w/v) в трис-ацетатном буфере.

Электрофорез в полиакриламидном геле (ПААГ)

- Разделяющий гель: 375 mM трис, pH 8,8 (соляная кислота), 14% акриламид (w/v), 0,1% SDS (w/v), 0,06% персульфат аммония (w/v), 1 мкл/мл TEMED.

- Концентрирующий гель: 125 мМ трис, рН 6,8 (соляная кислота), 4,5% акриламид (w/v), 0,1% SDS (w/v), 0,06% персульфат аммония (w/v), 1 мкл/мл TEMED.
- Буфер для верхней и нижней камеры: 25 мМ трис, 192 мМ глицин, 0,1% SDS (w/v).
- 4-кратный буфер для образцов: 125 мМ трис, рН 6,8 (соляная кислота), 50% глицерин (v/v), 10% 2-меркаптоэтанол (v/v), 4% SDS (w/v), 0,02% бромфеноловый синий (w/v).
- Краска для гелей: 50% метанол (v/v), 10% уксусная кислота (v/v), 0,04% кумасси R-250 (w/v).
- Раствор для обесцвечивания гелей: 10% метанол (v/v), 10% уксусная кислота (v/v).

Среды для выращивания клеток *E. coli*

- Lysogeny Broth (LB): 1% триптон (w/v), 1% хлорид натрия (w/v), 0,5% дрожжевой экстракт (w/v), рН 7,5 (гидроксид натрия).
- LB-агар: 1,5% агар (w/v) в LB.

Фракционирование

- Буфер 1: 50 мМ трис, рН 7,8 (соляная кислота), 400 мМ хлорид натрия, 1 мМ ЭДТА, 1% Твин 20 (v/v).
- Буфер 2: 50 мМ трис, рН 7,8 (соляная кислота), 400 мМ хлорид натрия, 1 мМ ЭДТА, 2 М мочевины.

Очистка белков

- Буфер А: 50 мМ дигидрофосфат натрия, рН 8,0 (гидроксид натрия), 300 мМ хлорид натрия.
- Буфер Б: 8 М мочевины в буфере А.
- Буфер В: 20 мМ HEPES, рН 8,0 (гидроксид натрия), 100 мМ хлорид натрия.

Отмывочные растворы для пептидного скрининга

- 50 мМ трис, pH 7,2 (соляная кислота), 150 мМ хлорид натрия, 0,1% Твин 20 (v/v).
- 3 мМ цитрат натрия, pH 7,0 (соляная кислота), 30 мМ хлорид натрия, 0,05% Твин 20 (v/v).
- 0,5% SDS (w/v).

Измерение активности с пептидными субстратами

- Буфер с pH 3,6: 15 мМ ацетат натрия, 185 мМ уксусная кислота, 0,5% ДМСО (v/v).
- Буфер с pH 4,6: 98 мМ ацетат натрия, 102 мМ уксусная кислота, 0,5% ДМСО (v/v).
- Буфер с pH 5,6: 182 мМ ацетат натрия, 18 мМ уксусная кислота, 0,5% ДМСО (v/v).
- Буфер с pH 6,5: 137 мМ дигидрофосфат натрия, 63 мМ гидрофосфат натрия, 0,5% ДМСО (v/v).
- Буфер с pH 7,5: 32 мМ дигидрофосфат натрия, 168 мМ гидрофосфат натрия, 0,5% ДМСО (v/v).

2.3 Программное обеспечение

Обработка данных, полученных при измерении ферментативной активности: flexAnalysis для анализа данных, полученных с масс-спектрометра ultrafleXtreme ("Bruker"); Spot-recognition software GenePix Pro для анализа данных с флуоресцентного сканера Axon GenePix 4300SL50 ("Molecular Devices"); CLARIOstar Microplate Reader Software (MARS) для анализа данных с мультимодального ридера CLARIOstar ("BMG LABTECH").

Рентгеноструктурный анализ: AutoPROC, который включает в себя программы X-ray Detector Software (XDS; "Max Planck Institute for Medical Research") [232], POINTLESS, AIMLESS ("The Medical Research Council Laboratory of Molecular Biology") [233,234], The Collaborative Computational Project

Number 4 (CCP4; "Rutherford Appleton Laboratory") [235] и STARANISO [236], для автоматической обработки данных рентгеновской дифракции [237] ("Global Phasing Ltd."); Python-based Hierarchical ENvironment for Integrated Xtallography (PHENIX) для получения структуры методом молекулярного замещения, ее последующего уточнения в электронной плотности и анализа ("Lawrence Berkeley National Laboratory") [238]; MolProbity для построения карты Рамачандрана полученной структуры тритикаина- α ("Duke University") [239]; Structural Classification of Proteins – extended (SCOPe) – база данных белковых мотивов для классификации полученной структуры ("University of California") [240].

Предсказание пептидных комплексов: AlphaFold2 Multimer и ColabFold для предсказания структур комплексов тритикаина- α с пептидами ("European Molecular Biology Laboratory - European Bioinformatics Institute (EMBL-EBI)") [241,242], включая Many-against-Many sequence searching (MMseqs2) для поиска и кластеризации аминокислотных последовательностей [243] и OpenMM7 для релаксации полученных структур в поле Amber [244,245] ("University of California").

Другие программы и базы данных: worldwide PDB (wwPDB) – база данных трёхмерных структур биомолекул, полученных экспериментальными методами [246–248], Pairwise Structure Alignment tool в Research Collaboratory for Structural Bioinformatics (RCSB) PDB для попарного выравнивания трёхмерных белковых структур [249], UniProt Knowledgebase (UniProtKB) – база данных последовательностей и функций белков [250] ("National Institutes of Health"); AlphaFold Protein Structure Database – база данных предсказанных трёхмерных структур биомолекул ("EMBL-EBI") [251]; AliView для множественных выравниваний последовательностей белков и нуклеиновых кислот ("Uppsala University") [252]; ProKSim для расчета зарядов атомов в молекулах и моделирования поля электростатического потенциала ("МГУ им. М.В. Ломоносова") [108,253]; Protein pKa Predictor (PropKa) для вычисления локальных pK_a а.о. полипептидов ("University of Copenhagen") [254]; The PyMOL Molecular

Graphics System для визуализации молекул ("Schrödinger") [31]; SigmaPlot для визуализации и статистического анализа данных ("Systat Software").

2.4 Получение бактериальных клеток-продуцентов

Амплификация нуклеотидных последовательностей продомена (Triticain- α -ProD), каталитического домена (Triticain- α -CatD) и слитой из вышеуказанных доменов формы тритикаина- α (Triticain- α -GM) была произведена методом ПЦР. В качестве матрицы использовалась ранее полученная в лаборатории плаزمиды pQE80L-6HIS-Triticain- α -GM [22]. В качестве прямых и обратных праймеров были использованы олигонуклеотиды, перечисленные в Таблице 4. Каждый праймер фланкирован сайтом рестрикции, кроме того, обратный праймер для конструкции pET15b-Triticain- α -GM-6HIS содержит последовательность, в которой закодирована метка 6HIS. Амплификация проводилась полимеразой из Advantage 2 Polymerase Mix в буфере Advantage 2 PCR Buffer с добавлением смеси дезоксирибонуклеотидов. Сайт-направленный мутагенез проводился по методу мегапраймера с добавлением в смесь ПЦР соответствующего внутреннего мутагенного праймера (Таблица 5) [255]. Мутации производились для замены Glu191 в тритикаине- α на Asn и Lys, а именно кодон 168 GAG в гене Triticain- α -GM был заменен на AAC или AAA, соответственно. Для замены каталитического Cys154 на Ala был заменен кодон 131 TGC на GCT также в реакции ПЦР. Амплифицированные гены были очищены с использованием набора Cleanup Mini Kit.

Полученные амплифицированные гены и соответствующие плазмиды были подвергнуты рестрикции с помощью FastDigest рестриктаз, указанных в Таблице 4, в FastDigest Green Buffer. Реакции проходили в твердотельном термостате при 37°C в течение 30 мин, а затем были остановлены путем инактивации эндонуклеаз рестрикции нагревом смеси до 80°C в течение 10 мин. Проверка эффективности реакции рестрикции производилась с помощью электрофореза в 1,5% агарозном геле с добавлением маркера длин ДНК. После разделения реакционной смеси в

агарозе фрагменты геля, содержащие необходимые продукты, были очищены с помощью набора Cleanup Mini Kit. Полученные фрагменты тритикаина- α были лигированы в соответствующие плазмиды для получения генетических конструкций, указанных в Таблице 4. Лигирование производилось Т4 ДНК-лигазой с аденозинтрифосфатом в качестве кофактора в T4 DNA Ligase Buffer. Реакционная смесь была инкубирована при комнатной температуре в течение 1 часа, после чего реакцию остановили путем инактивации лигазы нагревом до 65°C в течение 10 мин.

Трансформация клеточных культур *E. coli* проводилась с использованием набора TransformAid Bacterial Transformation Kit. Сначала клетки штаммов BL21 (DE3), Rosetta gami B (DE3) и Rosetta gami 2 (DE3) pLysS растили в LB-среде в течение ночи при 37°C и 250 оборотах в минуту (rpm) в шейкере-инкубаторе. Далее 1 мл клеточной культуры был высеян на чашку Петри с LB-агаром для инкубации при 37°C в течение дня в суховоздушном термостате. Для трансформации брались отдельные колонии клеток, которые добавляли в С-среду из набора и инкубировали ночь при 37°C и 250 rpm. На следующий день проводилась сама процедура трансформации методом теплового шока согласно инструкции. Двойная трансформация BL21 (DE3) плазмидами pQE80L-6HIS-Triticain- α -ProD и pET28a-6HIS-Triticain- α -CatD производилась последовательно. Также клетки BL21 (DE3) были трансформированы плазмидами pET28a-Triticain- α -GM-6HIS и pET28a-6HIS-Triticain- α -GM. Клетки Rosetta gami B (DE3) были трансформированы плазмидами pET15b-Triticain- α -GM-6HIS, pET15b-6HIS-Triticain- α -GM, pET15b-6HIS-Triticain- α -GM^{E191N} и pET15b-6HIS-Triticain- α -GM^{E191K}. Клетки Rosetta gami 2 (DE3) pLysS трансформировали плазмидой pET15b-6HIS-Triticain- α -GM^{C154A}. В завершение процедуры трансформации клетки высевали на чашки Петри с 1,5% LB-агаром, содержащим ампициллин в концентрации 100 мкг/мл для конструкций в плазмидах pQE80L и pET15b, канамицин в концентрации 50 мкг/мл для конструкций в плазмиде pET28a или оба антибиотика для двойных трансформантов pQE80L+pET28a. После инкубации выросшие колонии проверяли на наличие вставки методом ПЦР.

Полученные трансформированные клетки наращивали в течение ночи при 37°C и 250 rpm, а затем смешивали с глицерином в пропорции 1:3 для последующего хранения при минус 80°C.

2.5 Продукция и очистка протеиназ

Для экспрессии генов исследуемых белков, небольшое количество соответствующих клеток-продуцентов *E.coli* добавляли в среду LB, содержащую 100 мкг/мл ампицилина для конструкций в плазмидах pQE80L и pET15b, 50 мкг/мл канамицина – для pET28a или оба антибиотика для двойных трансформантов, и инкубировали в течение ночи при 37°C и 250 rpm. Далее ночную культуру разбавляли в 50 раз в LB с соответствующим антибиотиком и инкубировали при 37°C и 250 rpm до достижения оптической плотности (OD₆₀₀) значений от 0,6 до 0,8. Клеточные культуры, экспрессия в которых далее проводилась при температурах 28°C и 18°C, остужали на льду в течение 30 мин. Затем индуцировали экспрессию генов исследуемых белков путем добавления ИПТГ до конечной концентрации 1 mM. Далее клетки инкубировали при 250 rpm и 37°C в течение 3 ч или при более низких температурах: в штаммах BL21 (DE3) при 28°C в течение 15 ч, а в штаммах Rosetta gami B (DE3) и Rosetta gami 2 (DE3) pLysS при 18°C в течение 20 ч. Все дальнейшие работы с клетками и белками проводили на льду или при 4°C.

По окончании экспрессии проводился анализ продукции и растворимости белков. Для этого отбирали клеточную культуру в количестве OD₆₀₀ равном 4, осаждали при помощи центрифугирования в течение 10 мин при 4°C и 10000 g и ресуспендировали осадки в 500 мкл буфера 1. Далее клетки подвергали лизису с использованием ультразвука в течение 2 мин (импульс 5 с через 5 с). Полученную суспензию вновь центрифугировали в течение 30 мин при 4°C и 10000 g и делили на супернатант, содержащий растворимую фракцию, и осадок с нерастворимой фракцией рекомбинантного белка. Осадок дополнительно трижды промывали в

500 мкл буфера 2 и ресуспендировали в 300 мкл буфера 1. Полученные фракции анализировали при помощи электрофореза в ПААГ в денатурирующих условиях.

Для очистки различных форм тритикаина- α (кроме мутанта с заменой Cys154Ala) полученные клеточные культуры после окончания экспрессии центрифугировали в течение 10 мин при 4°C и 10000 g. Далее осадки ресуспендировали в буфере А с добавлением 10 мМ имидазола и подвергали лизису ультразвуком 2 мин (5 с через 5 с). Полученную суспензию центрифугировали в течение 30 мин при 4°C и 10000 g. Далее отбирался полученный осадок или супернатант, в зависимости от того, в какой форме (нерастворимой или растворимой, соответственно), был фермент. Для очистки нерастворимых форм тритикаина- α , полученный после лизиса и центрифугирования осадок растворяли в буфере Б с 10 мМ имидазолом, чтобы затем снова подвергнуть обработке ультразвуком (2 мин, 5 с через 5 с) и центрифугировать в течение 30 мин при 4°C и 10000 g. Полученный супернатант наносили на колонку с Ni-сефарозой, уравновешенную буфером Б с 10 мМ имидазолом, после чего связавшийся белок элюировали буфером Б, содержащим 300 мМ имидазола. Для очистки растворимых форм тритикаина- α , полученный после лизиса и центрифугирования клеточных культур супернатант наносили на колонку с Ni-сефарозой, уравновешенную буфером А с 10 мМ имидазолом, а связавшийся белок элюировали буфером А с 300 мМ имидазолом. В обоих случаях в конце очистки тритикаин- α диализовали против буфера А в течение 24 часов при 4°C.

Для очистки растворимой формы неактивного мутанта тритикаина- α с заменой Cys154Ala клеточные культуры также центрифугировали в течение 10 мин при 4°C и 10000 g. Осажденные клетки затем ресуспендировали в буфере А с 10 мМ имидазолом и 1 мг/мл лизоцима и инкубировали в течение 30 мин, после чего подвергали лизису ультразвуком (2 мин, 5 с через 5 с). К полученному лизату добавляли 10 мкг/мл РНКазы А и 5 мкг/мл ДНКазы I, инкубировали 15 мин и снова лизировали ультразвуком (2 мин, 5 с через 5 с). Далее лизаты центрифугировали в течение 1 часа при 4°C и 150000 g в ультрацентрифуге.

Полученные супернатанты наносили на колонку с Ni-сефарозой, уравновешенную буфером А с 10 мМ имидазолом. Далее колонку промывали от несвязавшегося белка буфером А с 50 мМ имидазолом и элюировали связавшийся белок буфером А с 300 мМ имидазолом. Полученный раствор наносили на колонку с сефадексом G-25, уравновешенную буфером В, после чего элюировали белок несколькими порциями буфера В для получения набора фракций. Концентрация белка в полученных фракциях измерялась по поглощению света с длиной волны 280 нм. Фракции, содержащие белок, собирали в одну пробу для проведения дальнейших исследований.

Очищенные белки далее были сконцентрированы при помощи центрифужных концентраторов с пределом отсечения 10 кДа. Конечная концентрация определялась с использованием набора BCA Protein Assay Kit (анализ Смита с бицинониновой кислотой и ионами меди). Все этапы процедуры очистки контролировались при помощи электрофореза в ПААГ в денатурирующих условиях.

2.6 Скрининг субстратной специфичности тритикаина- α

Скрининг субстратной специфичности тритикаина- α проводился с использованием пептидных микрочипов компании JPT Peptide Technologies, состоящих из пептидов, меченных рТур на С-конце. Библиотека, включающая 3525 олигопептидов, была сформирована на основе известных сайтов расщепления и случайно сгенерированных пептидов. рТур присоединен к С-концу пептида через гидрофильный линкерный фрагмент, а N-конец пептида закреплен на стеклянной подложке также через гидрофильный линкер. Инкубация раствора фермента на микрочипах проводилась в натрий-ацетатном буфере с pH 4 в течение 4 часов при 37°C (Рисунок 10). В качестве контроля проводилась такая же инкубация без добавления фермента.

После проведения инкубации микрочипы обрабатывали отмывочными растворами, а затем инкубировали с раствором первичных мышиных антител к

pTyr в конечной концентрации 1 мкг/мл в блокирующем буфере SuperBlock T20 в течение одного часа при комнатной температуре. Далее микрочипы снова обрабатывали отмывочными буферами, после чего инкубировали с флуоресцентно мечеными вторичными антителами в концентрации 1 мкг/мл в блокирующем буфере SuperBlock T20 в течение 45 мин при комнатной температуре. После очередной обработки отмывочными растворами микрочипы были высушены в центрифуге. Интенсивность флуоресценции измерялась сканером флуоресценции высокого разрешения Axon GenePix 4300SL50 ("Molecular Devices"). Эффективность гидролиза пептидов тритикаином- α определялась по разнице уровня флуоресценции с контролем (Рисунок 10).

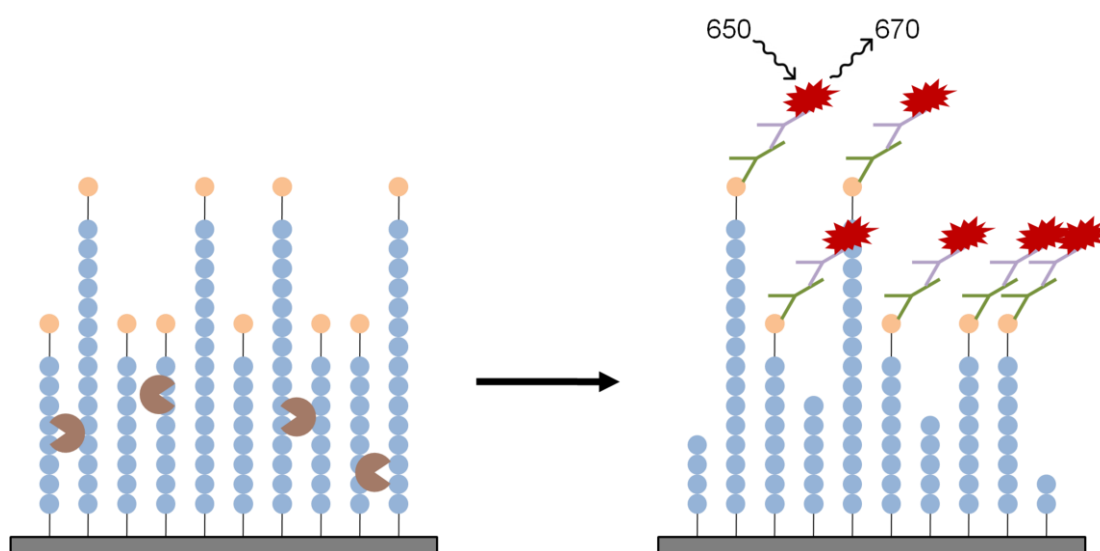


Рисунок 10 – Схематичная иллюстрация скрининга субстратной специфичности протеиназы с использованием меченых pTyr пептидов. А.о. изображены кружками, при этом pTyr – оранжевого цвета; протеиназа – в форме "пакмана"; флуорофор, сцепленный с вторичными антителами, обозначен красным цветом; числа указывают длины волн возбуждения и излучения в нанометрах. Чем выше была эффективность гидролиза, тем меньше оставалось интактных пептидов с меткой pTyr на C-конце, и тем ниже была интенсивность флуоресценции.

Олигопептид AQLLRVGCVLGT, который расщеплялся эффективнее всего (подчеркнута последовательность пептида из скрининга), был синтезирован для определения сайтов расщепления. Пептид инкубировали с тритикаином- α в натрий-ацетатном буфере с pH 4,6 в течение 4 часов при комнатной температуре. Далее сайты расщепления определялись на тандемном времяпролетном масс-спектрометре с матричной лазерной десорбцией/ионизацией (MALDI-TOF/TOF;

ultrafleXtreme "Bruker") в режиме положительных ионов. Моноизотопные массы были измерены на диапазоне 600–5000 m/z с точностью до 30 млн⁻¹. Анализ полученных данных проводился в программе flexAnalysis.

2.7 Измерение протеолитической активности протеиназ

Сравнение активности папаина и форм тритикаина- α с различной доменной организацией было выполнено с использованием флуорогенного пептидного субстрата *Ac*-PLVQ-AMC ("Permic") путем мониторинга флуоресценции высвобожденной флуорогенной метки AMC в зависимости от времени (Рисунок 11). Измерения проводились при комнатной температуре в реакционной смеси объемом 100 мкл, содержащей 20 нМ фермента и 50 мкМ субстрата в буфере с pH 5,6. Изучение специфичности тритикаина- α проводилось в буфере с pH 4,6 с использованием набора флуорогенных пептидных субстратов: *Ac*-DEVD-AMC, *Ac*-FESR-AMC, *Ac*-PLVQ-AMC, *Ac*-QLLK-AMC, *Ac*-QLLR-AMC, *Ac*-VLPQ-AMC и *Z*-FVR-AMC ("Permic"). Флуоресценцию детектировали с использованием флуориметра GloMax-Multi Detection System ("Promega") на длине волны возбуждения 360 нм и длине волны излучения 460 нм. Началом реакции считали момент добавления субстрата, а измерение начинали через 30 с путем замера интенсивности флуоресценции в реакционной смеси каждые 10 с в течение 20 мин.



Рисунок 11 – Схематическая иллюстрация измерения ферментативной активности протеиназ с использованием пептидных субстратов (а.о. изображены кружками) с флуорогенной меткой AMC. Числа указывают длины волн возбуждения и излучения в нанометрах. Гидролиз субстрата приводил к накоплению свободной метки и повышению интенсивности флуоресценции.

Сравнение субстратной специфичности тритикаина- α и его мутантов Glu191Asn и Glu191Lys проводилось путем измерения их ферментативной

активности в реакционном объеме 100 мкл в черных 96-луночных планшетах на флуориметре CLARIOstar ("BMG LABTECH") при возбуждении диапазоном длин волн 350 ± 15 нм и детекции флуоресценции на диапазоне длин волн 440 ± 20 нм. Набор флуорогенных пептидных субстратов для этого эксперимента состоял из *Ac-QLLD-AMC*, *Ac-QLLK-AMC*, *Ac-QLLQ-AMC* и *Ac-QLLR-AMC* ("Пептидные Технологии"). Измерение активности проводилось на диапазоне концентраций субстратов от 0,78125 до 31,25 мкМ в буферах с pH от 3,6 до 7,5 при комнатной температуре. Все использованные в этих измерениях буферы содержали 100 мМ хлорид натрия, 15 мМ 2-меркаптоэтанол и 1 мМ ЭДТА. Реакция запускалась добавлением фермента до конечной концентрации 20 нМ, после чего реакционная смесь перемешивалась, и через 30 с начиналось измерение. Для каждого эксперимента было выполнено три повторности по 100 измерений через каждые 4 с. В качестве контроля проводилось измерение ещё одной повторности, но с буфером вместо фермента.

Обработка полученных данных, построение графиков и их аппроксимация проводились в программе SigmaPlot. Скорость реакции определялась по углу наклона начального отрезка кривых на графике зависимости интенсивности флуоресценции от времени. Полученные единицы измерения были преобразованы в количество гидролизованного субстрата на основе стандартной кривой, полученной в результате измерения флуоресценции известных концентраций свободной метки AMC. Все реакции и измерения повторили три раза. На основе измерений активности протеиназ на диапазоне концентраций субстрата были построены кривые по уравнению Михаэлиса-Ментен, отражающие зависимость начальной скорости реакции от исходной концентрации субстрата.

2.8 Рентгеноструктурный анализ тритикаина- α

Для получения кристаллов очищенного неактивного мутанта тритикаина- α (с заменой каталитического Cys154 на Ala) был проведён скрининг условий кристаллизации с использованием пяти наборов по 96 растворов каждый: BCS,

JCSG-Plus Eco, LMB, PACT premier и SG1 Eco. Кристаллизация проводилась методом "сидячая капля". Для этого автоматизированной многоканальной станцией дозирования Mosquito Crystal ("SPT Labtech") на специальную 96-луночную планшету для кристаллизации были нанесены нанолитровые количества буферов для кристаллизации и раствора белка. В каждой лунке содержалось по четыре раствора: буфер для кристаллизации и три "сидячие" капли со смесью 1 мг/мл белка в буфере В и кристаллизационного буфера в соотношениях 5:3, 1:1 и 3:5. После нанесения всех необходимых компонентов на планшету, она запечатывалась прозрачной пленкой и помещалась в тёмное помещение с постоянной температурой воздуха 4°C. Появление кристаллов в лунках планшеты проверялось ежедневно с использованием бинокля. Полученные кристаллы были извлечены из планшеты, заморожены в жидком азоте и направлены на облучение источником синхротронного излучения Diamond Light Source (beamline I03; Чилтон, Великобритания) при 100К на длине волны 0,96409 Å.

Полученные дифракционные картины были далее автоматически обработаны в программном пакете autoPROC с использованием программы STARANISO. Структура была определена методом молекулярного замещения в программе MRage программного пакета PHENIX. Для этого использовалась часть предсказанной структуры тритикаина-α из базы данных AlphaFold (AF-Q0WXG8-F1) от Met24 до Gly348 (без сигнального пептида и гранулин-подобного домена). Структура "технического" N-концевого участка длиной 20 а.о., который был необходим для продукции и очистки белка, не была решена. Дальнейшие уточнения структуры в электронной плотности для повышения ее точности проводились в программе Refine программного пакета PHENIX. Для полученной структуры была построена карта Рамачандрана в программе MolProbity; также структура была классифицирована согласно базе белковых мотивов SCOPe. Итоговая структура тритикаина-α-GM^{C154A} была депонирована в базе PDB под идентификационным номером 8ZLQ.

2.9 Предсказание структур комплексов тритикаина- α с олигопептидами

Предсказание структур комплексов было выполнено в программе AlphaFold2 Multimer. Кристаллическая структура тритикаина- α -GM^{C154A} и последовательность его каталитического домена (начиная с Ala124) с интактным Cys154 в активном центре были использованы для моделирования комплексов с пептидами DEVSVG, FESRVG, FVRVG, PLVQVG, QLLKVG, QLLRVG и VLPQVG. Множественное выравнивание последовательностей тритикаина- α и структур из базы PDB было построено в ColabFold с помощью MMseqs2. Для каждого комплекса было сгенерировано 40 моделей и ранжировано. Для каждой модели была выполнена минимизация энергии в силовом поле Amber программой OpenMM7. Дальнейший анализ данных был выполнен в программе PyMOL. В частности, были измерены расстояния между карбонильным углеродом третьего а.о. FVRVG или четвертого а.о. других пептидов и серой каталитического Cys154 тритикаина- α . Также были измерены расстояния между азотами и кислородами боковой цепи четвертого а.о. в олигопептиде (третьего в случае FVRVG) и кислородами боковой цепи Glu191 или Asp289 в белке.

2.10 Выравнивание аминокислотных последовательностей протеиназ

Множественное выравнивание аминокислотных последовательностей каталитических доменов ППЦП было выполнено в программе AliView. Последовательности для выравнивания были получены из UniProt KB (идентификаторы указаны в скобках): тритикаин- α (Q0WYG8) из *T. aestivum*, папаин (P00784), химопапаин (P14080) и карикаин (P10056) из *C. papaya*, бромелаин стебля (P14518) и ананаин (P80884) из *A. comosus*, актинидин (P00785) из *Actinidia chinensis*, вигнаин (P12412) из *Vigna mungo*, зингипаин (P82474) из *Z. officinale*, EP-B 2 (P25250) из *H. vulgare*, круципаин (P25779) из *T. cruzi* и цистеиновые катепсины CtsB (P07858), CtsC (P53634), CtsF (Q9UBX1), CtsH (P09668), CtsK (P43235), CtsL (P07711), CtsO (P43234), CtsS (P25774), CtsV

(O60911), CtsW (P56202) и CtsZ (Q9UBR2) из *Homo sapiens*. Выравнивание 10 а.о. до и после сравниваемых а.о. в S₁- и S₂-сайтах связывания тритикаина- α было скорректировано вручную согласно попарному выравниванию структур перечисленных ППЦП в Pairwise Structure Alignment tool в RCSB PDB методом flexible jFATCAT [256].

2.11 Вычисление распределения зарядов на поверхности протеина

Расчёты распределения зарядов на поверхности растительных ППЦП проводились в программе ProKSim для каталитических доменов тритикаина- α ^{C154A} (структура получена в ходе настоящего исследования), папаина из *S. papaya*, ананаина из *A. comosus*, актинидина из *A. chinensis*, зингипаина из *Z. officinale* и EP-B 2 из *H. vulgare* (PDB ID: 2CIO, 6YCB, 2ACT, 1CQD и 2FO5, соответственно) [257–261]. Для этого сначала были рассчитаны значения зарядов атомов при pH 4,6 по уравнению Хендерсона-Хассельбаха с использованием значений pK_a а.о. [262]. Далее для расчёта поля электростатического потенциала была задана кубическая трехмерная матрица со стороной 300 Å с кубическими ячейками со стороной 1 Å с белком в центре этой матрицы. Каждой ячейке было присвоено значение диэлектрической проницаемости (ϵ) равное 2, если ячейка содержит атомы белка, равное 40 – ячейкам, соседним с теми, у которых $\epsilon = 2$, равное 80 – остальным ячейкам. Также для каждой ячейки была задана ионная сила 100 мМ, если ϵ равна 40 или 80, или 0 мМ, если $\epsilon = 2$. На основе зарядов атомов, диэлектрической проницаемости и ионной силы был рассчитан потенциал в ячейках сетки по итерационной формуле, полученной из линеаризованного уравнения Пуассона-Больцмана [263]. Для визуализации полученных данных использовалась программа PyMOL.

3. РЕЗУЛЬТАТЫ

3.1 Продукция каталитического домена в одной полипептидной цепи с продомом необходима для формирования активного тритикаина- α

Для того чтобы исследовать субстратную специфичность тритикаина- α из *T. aestivum*, прежде всего было необходимо подобрать условия для получения активной и растворимой формы этого фермента. Полноразмерный ген тритикаина- α состоит из каталитического домена, в котором закодированы а.о., составляющие активный центр фермента, а также *N*-концевых сигнального пептида и аутоингибирующего продомена и *C*-концевого гранулин-подобного домена (Рисунок 12) [231].

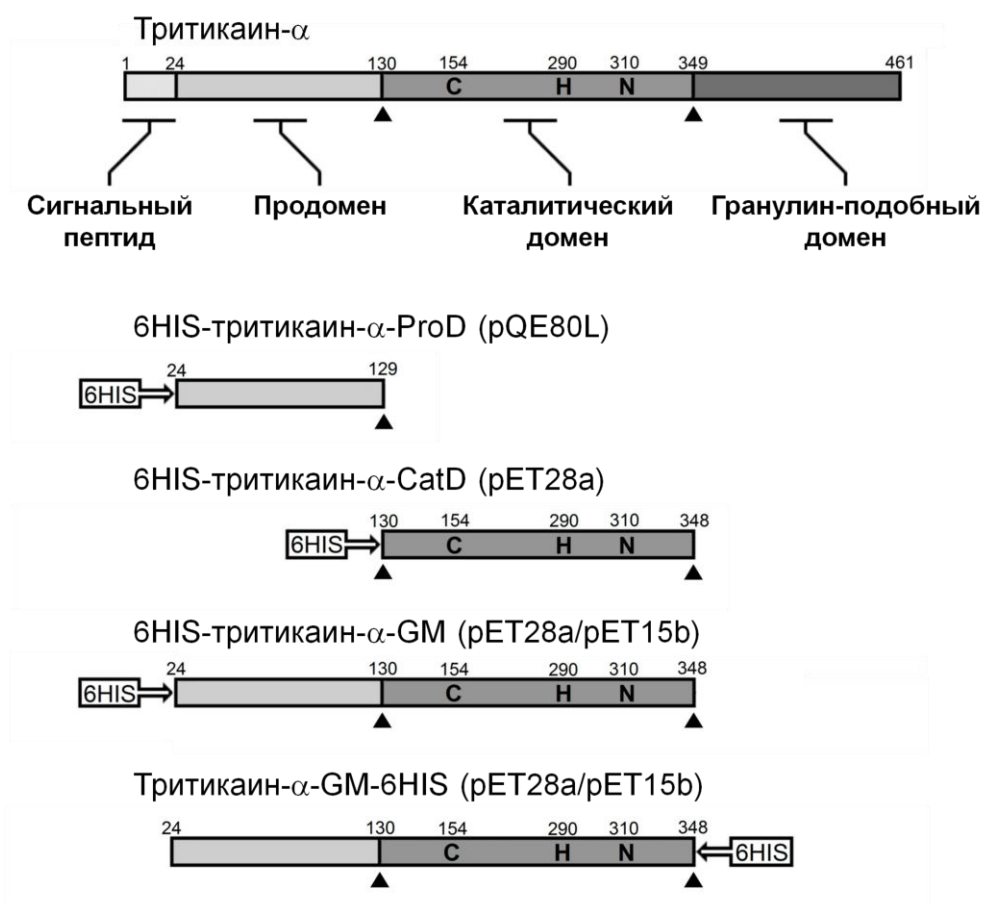


Рисунок 12 – Схема первичной структуры тритикаина- α . Стрелки указывают границы каталитического домена, из которого состоит зрелый фермент. Буквы указывают а.о., составляющие каталитическую триаду фермента. Под структурой полноразмерного тритикаина- α перечислены формы фермента, которые были получены в ходе данного исследования.

Основываясь на литературных данных о том, что для продукции в клетках *E. coli* более предпочтительными являются низкомолекулярные белки [264], сначала была получена генетическая конструкция, содержащая только ген каталитического домена тритикаина- α (6HIS-Triticain- α -CatD, экспрессия продукта с $M_r = 23,57$ кДа; Рисунок 12). Экспрессия в клетках *E. coli* привела к получению $174,60 \pm 4,50$ мг белка из одного литра клеточной культуры. Продукт экспрессии был обнаружен только в нерастворимой форме в тельцах включения, что потребовало дополнительных стадий очистки: денатурации под действием высокой концентрации мочевины с последующим рефолдингом в процессе диализа. Однако даже после рефолдинга белок не проявлял активности в отношении флуорогенного пептидного субстрата *Ac*-PLVQ-AMC (Рисунок 13), сконструированного на основе данных о сайтах расщепления протеиназой глютеиновых белков (Рисунок 9). Полученные результаты указывают на то, что для созревания активной формы тритикаина- α недостаточно полипептида, содержащего только каталитический домен фермента.

Ранее для ряда цистеиновых протеиназ было показано, что продомен может функционировать как шаперон, способствующий правильной укладке зрелого фермента [22,265–267]. Основываясь на этих данных, была сделана генетическая конструкция, содержащая ген продомена тритикаина- α (6HIS-Triticain- α -ProD, $M_r = 12,71$ кДа; Рисунок 12). При экспрессии этого гена в клетках *E. coli* часть продукта была обнаружена в растворимой фракции, что давало возможность продомену способствовать правильной укладке каталитического домена. Однако его коэкспрессия с 6HIS-Triticain- α -CatD в клетках *E. coli* не привела к увеличению растворимости протеиназы. Соответственно, растворимые фракции белка, полученного после совместной экспрессии 6HIS-Triticain- α -ProD с 6HIS-Triticain- α -CatD, не содержали каталитического домена фермента и не проявляли активности в отношении субстрата *Ac*-PLVQ-AMC.

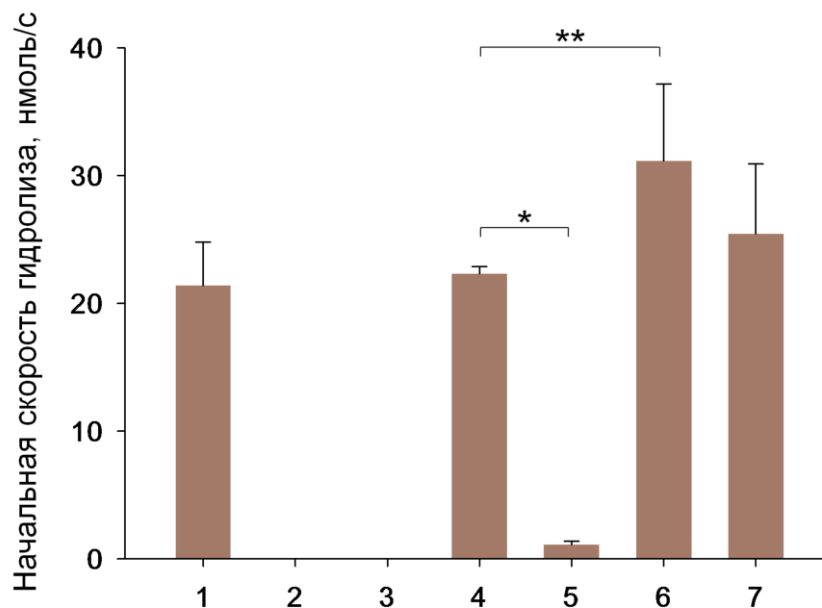


Рисунок 13 – Эффективность расщепления субстрата *Ac*-PLVQ-*AMC* папаином (1) и тритикаином- α различной доменной организации: 6HIS-тритикаин- α -ProD (2), 6HIS-тритикаин- α -CatD (3), 6HIS-тритикаин- α -GM и тритикаин- α -GM-6HIS из телец включения (4 и 5, соответственно; штамм BL21 (DE3)) и из растворимой фракции (6 и 7, соответственно; штамм Rosetta gami B (DE3)). Данные получены на основе трёх независимых измерений для каждого фермента. Статистическая значимость рассчитывалась согласно t-критерию Стьюдента для независимых выборок (* $p < 0,0001$; ** $p < 0,1$).

На основе полученных данных было сделано предположение, что для правильного фолдинга каталитического домена тритикаина- α может требоваться коэкспрессия его гена с генами других доменов фермента в форме единой полипептидной цепи. Чтобы это проверить, была создана конструкция, состоящая из генов продомена и каталитического домена в одной полипептидной цепи, или ген тритикаина- α без *N*-концевого сигнального пептида и без *C*-концевого гранулин-подобного домена (6HIS-Triticain- α -GM, экспрессия продукта с $M_r = 36,27$ кДа; Рисунок 12). Экспрессия этой конструкции в клетках *E. coli* ($61,00 \pm 5,30$ мг/литр клеточной культуры) была менее эффективной, чем экспрессия 6HIS-Triticain- α -CatD, и также приводила к накоплению продуктов экспрессии в тельцах включения, что вновь требовало рефолдинга. Однако полученный белок проявлял протеиназную активность в отношении субстрата *Ac*-PLVQ-*AMC*, сопоставимую с активностью папаина из *C. papaya*, взятого в качестве контрольной ППЦП (Рисунок 13).

Все перечисленные структуры содержали аффинную метку, а именно 6HIS, для одностадийной очистки продукта экспрессии от клеточных лизатов на Ni-сефарозе. Между тем известно, что локализация 6HIS может влиять на эффективность продукции и растворимость рекомбинантных белков, продуцируемых в *E. coli* [268]. Для того чтобы проверить этот эффект на тритикаине- α , была сравнена эффективность экспрессии двух конструкций: с меткой на N-конце или на C-конце (6HIS-Triticain- α -GM и Triticain- α -GM-6HIS, соответственно; Рисунок 12). Оба варианта экспрессировались в клетках *E. coli* BL21 (DE3) исключительно в тельца включения в нерастворимой форме. Примечательно, что уровень экспрессии 6HIS-Triticain- α -GM был в два раза выше уровня экспрессии Triticain- α -GM-6HIS ($61,00 \pm 5,30$ мг белка/литр клеточной культуры против $26,80 \pm 0,91$ мг/л, соответственно). Рефолдинг обеих форм тритикаина- α -GM приводил к созреванию активной протеиназы. Однако удельная протеолитическая активность фермента с C-концевой меткой была крайне низкой, что свидетельствует о том, что лишь незначительная часть белковых молекул с такой последовательностью восстанавливала свою правильную третичную структуру (Рисунок 13).

В ходе дальнейших усилий по улучшению растворимости тритикаина- α -GM был использован штамм Rosetta gami B (DE3) клеток *E. coli* вместо BL21 (DE3). Преимущество штамма Rosetta gami B (DE3) заключается в том, что он содержит шесть кодонов для редких тРНК и несет мутации в генах тиоредоксинредуктазы и глутатионредуктазы, которые помогают образованию внутримолекулярных дисульфидных связей в белках, способствуя фолдингу [269]. В этом штамме тритикаин- α -GM был обнаружен в растворимой фракции. Более того, снижение температуры, при которой проводилась экспрессия, с 37°C до 18°C также повышало долю белка в растворимой форме. Оба варианта протеиназы продуцировались в растворимой форме в клетках *E. coli* Rosetta gami B (DE3) в меньших количествах ($17,30 \pm 2,24$ мг/литр клеточной культуры для 6HIS-Triticain- α -GM и $5,15 \pm 0,50$ мг/л для Triticain- α -GM-6HIS) чем в нерастворимой форме в клетках BL21 (DE3). Однако ферменты, экспрессировавшиеся в растворимой

форме, обладали более высокой ферментативной активностью по сравнению с теми, что экспрессировались в тельца включения (Рисунок 13).

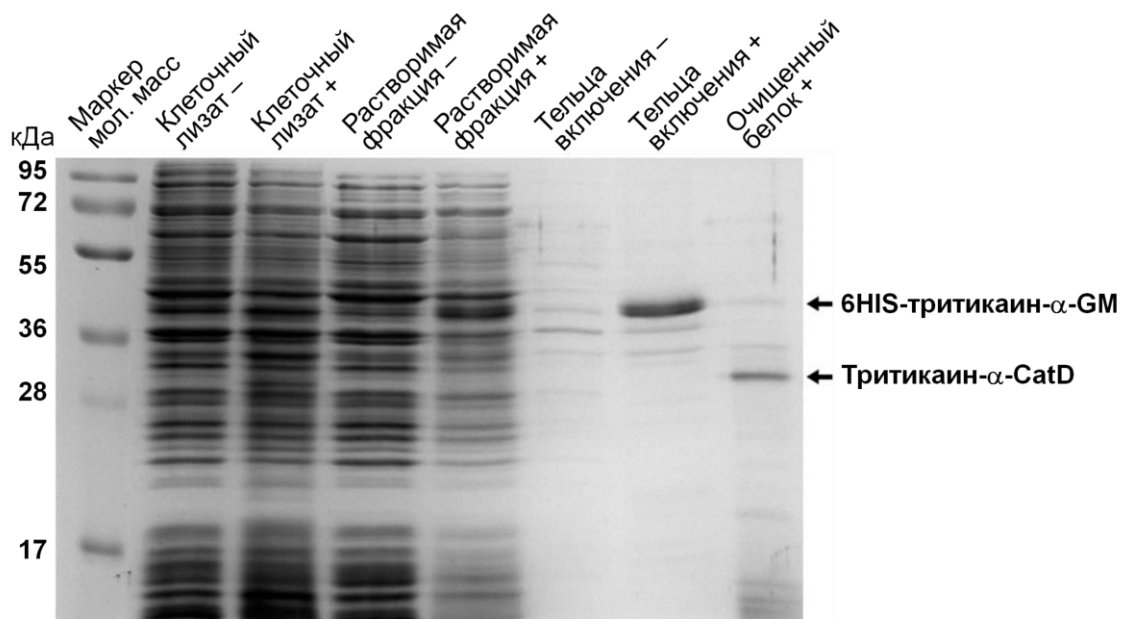


Рисунок 14 – Электрофорез в ПААГ проб, полученных в результате продукции в клетках *E. coli* Rosetta gami B (DE3) и последующей очистки из растворимой фракции 6HIS-тритикаина- α -GM. Препараты, полученные до добавления ИПТГ, обозначены символом "-", после добавления – символом "+".

Таким образом, среди изученных условий оптимальным методом получения каталитически активного фермента оказалась продукция меченого 6HIS по N-концу двухдоменного тритикаина- α -GM в бактериальном штамме Rosetta gami B (DE3) при 18°C с последующей одностадийной аффинной очисткой рекомбинантного продукта (Рисунок 14).

3.2 Тритикаин- α специфичен к субстратам с гидрофобными и положительно заряженными аминокислотными остатками в P₂- и P₁-положениях

После получения активной формы тритикаина- α были проведены исследования субстратной специфичности протеиназы. Для этого в первую очередь был выполнен скрининг активности тритикаина- α в кислой среде в отношении библиотеки из 3525 пептидов длиной 8 или 15 а.о., меченных pTyr. В результате фермент гидролизировал 225 пептидов, из которых пептид с

последовательностью QLLRVGCV расщеплялся эффективнее всего (Рисунок 15). Для определения сайта расщепления этого пептида он был синтезирован и инкубирован с тритикаином- α . Масс-спектрометрический анализ продуктов гидролиза показал, что расщепление происходило преимущественно после Arg (QLLR↓VGCV). Основываясь на полученных данных, для дальнейших исследований были синтезированы флуорогенные пептиды *Ac*-QLLR-AMC и *Ac*-QLLK-AMC. Также, с учётом результатов масс-спектрометрического анализа продуктов расщепления тритикаином- α глютена пшеницы *T. aestivum* (Рисунок 9) [22], были синтезированы субстраты *Ac*-PLVQ-AMC и *Ac*-VLPQ-AMC (Рисунок 16). К перечисленным пептидам были добавлены субстраты других цистеиновых протеиназ: CtsS человека – *Ac*-FVR-AMC, каспазы 3 – *Ac*-DEVD-AMC и метакаспаз типа II – *Ac*-FESR-AMC [84,270,271].

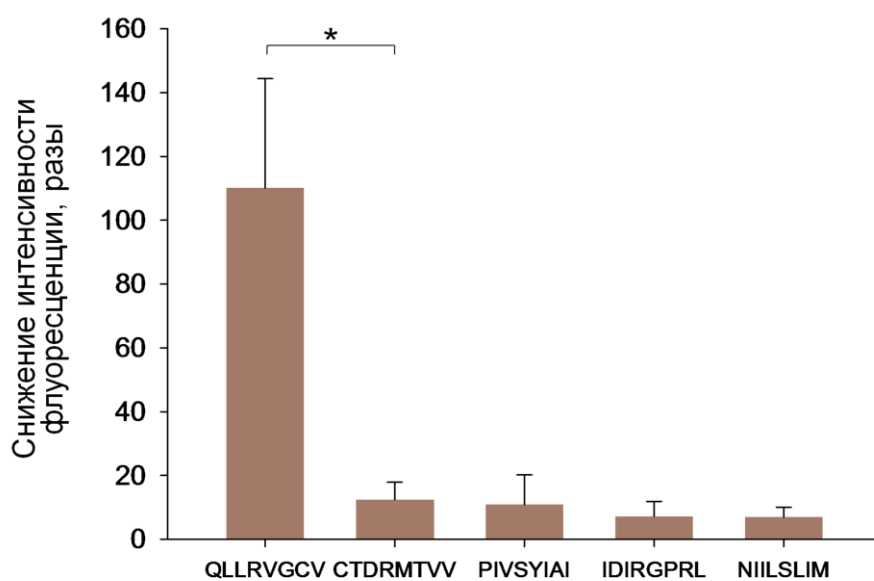


Рисунок 15 – Пять пептидов с самой высокой эффективностью расщепления тритикаином- α из библиотеки, включающей 3525 проверенных в ходе скрининга пептидов. Сравнение проводилось по соотношению интенсивности флуоресценции в тестовом и контрольном образцах (Рисунок 10). Пептид с последовательностью QLLRVGCV гидролизровался значительно лучше всех остальных. Статистическая значимость рассчитывалась согласно t-критерию Стьюдента для независимых выборок (* $p < 0,1$).

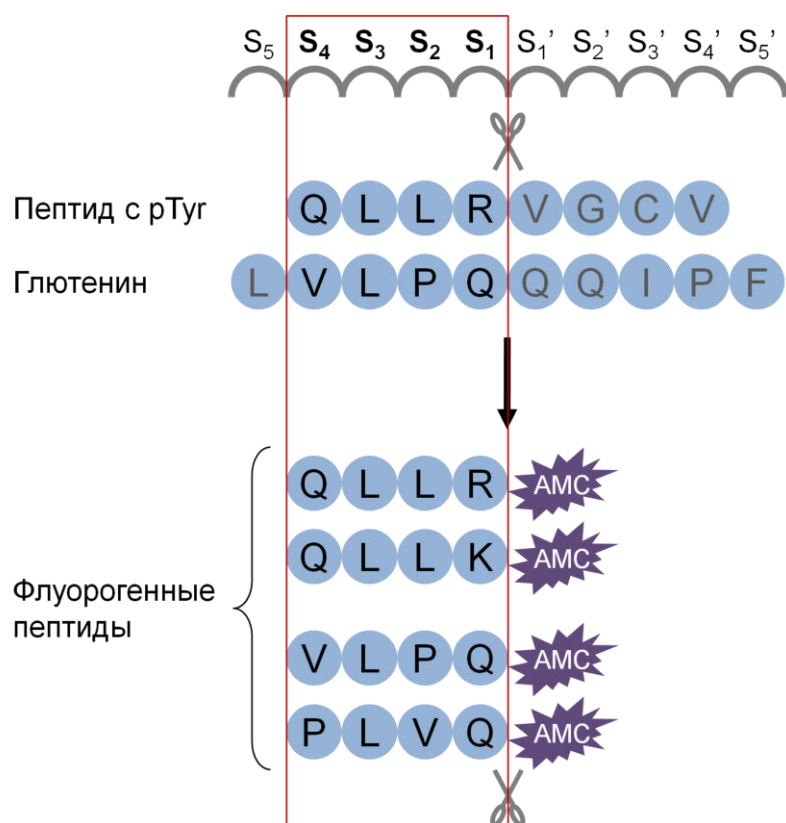


Рисунок 16 – Схематичное изображение подбора пептидных последовательностей флуорогенных субстратов для измерения активности тритикаина- α .

Результаты измерения активности тритикаина- α с перечисленными пептидами при pH 4,6 представлены на Рисунке 17. Наилучшими субстратами для тритикаина- α оказались *Ac*-QLLR-AMC и *Ac*-QLLK-AMC, которые содержат положительно заряженный а.о. в P_1 -положении. Похожий на них субстрат Z-FVR-AMC гидролизует тритикаином- α хуже, что может быть вызвано более короткой длиной этого пептида, из-за которой он может образовывать меньше взаимодействий с а.о. в субстрат-связывающей борозде протеиназы. *Ac*-PLVQ-AMC был третьим по эффективности расщепления тритикаином- α . Интересно, что этот субстрат протеиназа гидролизовала в 20 раз лучше, чем похожий на него *Ac*-VLPQ-AMC. Вероятно, Pro в P_2 -положении ограничивает конформационную подвижность, что может мешать эффективному связыванию субстрата в активном центре фермента. *Ac*-DEVD-AMC, который отличается от других субстратов отрицательно заряженным а.о. в P_1 -положении, расщеплялся тритикаином- α с самой низкой эффективностью. Субстрат *Ac*-FESR-AMC, который единственный из исследованных пептидов содержит гидрофильный а.о. в P_2 -положении,

оказался устойчив к гидролизу тритикаином- α . Это соответствует известным данным о выраженной специфичности большинства ППЦП к гидрофобному а.о. в P_2 -положении субстрата [24,41].

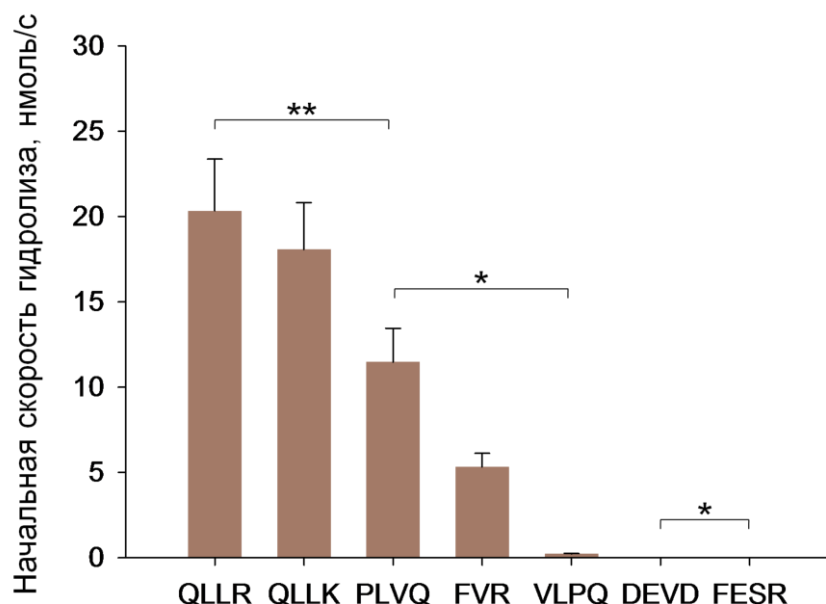


Рисунок 17 – Эффективность расщепления тритикаином- α флуорогенных пептидных субстратов. Каждое измерение выполнено в трёх независимых повторностях. Статистическая значимость рассчитывалась согласно t-критерию Стьюдента для независимых выборок (* $p < 0,001$; ** $p < 0,05$).

Полученные данные указывают на то, что в кислой среде тритикаин- α специфичен к субстратам с гидрофобным а.о. в P_2 - и с положительно заряженным а.о. в P_1 -положении.

3.3 Третичная структура тритикаина- α , полученная с помощью метода рентгеноструктурного анализа

Для выявления структурных особенностей тритикаина- α , определяющих его субстратную специфичность к гидрофобным а.о. в P_2 - и положительно заряженным а.о. в P_1 -положении, было решено получить третичную структуру фермента для последующего анализа. Структуру протеиназы определяли методом рентгеноструктурного анализа, для чего сначала были получены кристаллы белка. Для подбора условий кристаллизации тритикаина- α был использован мутант по

каталитическому Cys154 – 6HIS-тритикаин- α -GM^{C154A}, который не подвержен автокаталитическому расщеплению в отличие от фермента дикого типа (Рисунок 18). Скрининг условий проводился в 480 различных растворах, входящих в состав пяти наборов для кристаллизации: BCS, JCSG-Plus Eco, LMB, PACT premier и SG1 Eco ("Molecular Dimensions"). В результате кристаллы белка выросли в течение 10 дней при 4°C только в одних условиях: 100 mM натрий-ацетатный буфер, pH 4,6, с 3,5% полиэтиленгликолем 4000 (w/v) и 15% глицерином (v/v) в качестве соосадителей (E10 в LMB Screen).

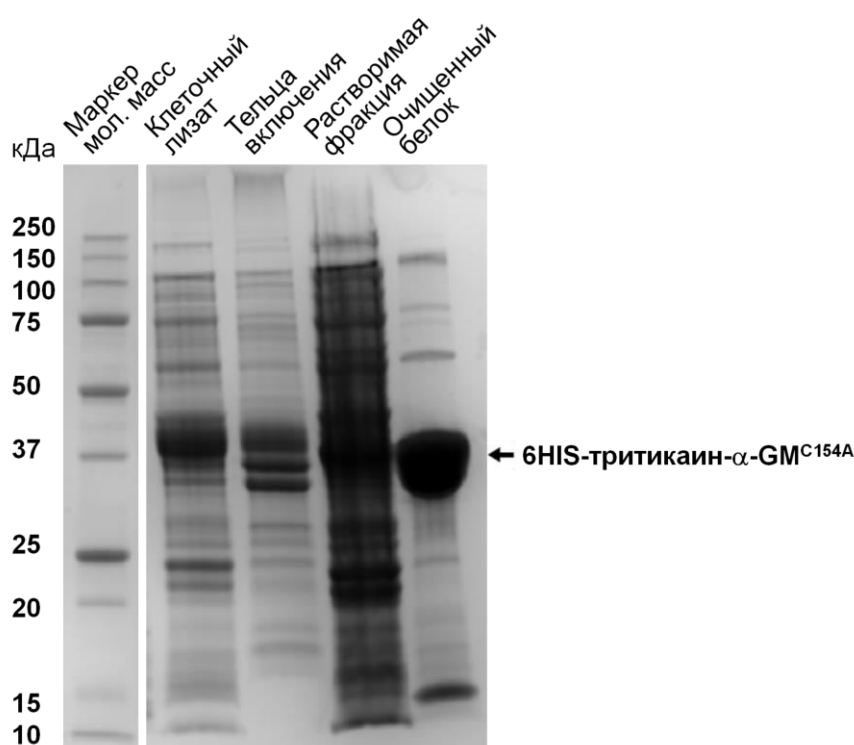


Рисунок 18 – Электрофорез в ПААГ проб, полученных в результате продукции в клетках *E. coli* Rosetta gami 2 (DE3) pLysS и последующей очистки из растворимой фракции 6HIS-тритикаина- α -GM^{C154A}.

В результате трёх независимых экспериментов по продукции, очистке и кристаллизации 6HIS-тритикаина- α -GM^{C154A} были выращены четыре кристалла. Полученные от них на источнике синхротронного излучения картины дифракции были автоматически обработаны в программном пакете AutoPROC. Все кристаллы дифракционировались в низком разрешении: самое высокое разрешение, которое было достигнуто – 3,46 Å (Таблица 6). Из-за низкого разрешения полученных данных, для построения структуры был выбран метод

молекулярного замещения. В качестве модели была использована предсказанная структура тритикаина- α , взятая из базы AlphaFold под идентификационным номером AF-Q0WYG8-F1 [251]. Для проведения молекулярного замещения использовалась часть структуры, состоящая из продомена и каталитического домена (от Met24 до Gly348), что соответствует доменной организации кристаллизованного тритикаина- α -GM^{C154A}. Далее, полученная структура фермента была дополнительно подвергнута нескольким раундам уточнения по электронной плотности.

Полученная в результате структура тритикаина- α -GM^{C154A} подходит под электронную плотность с параметрами $R = 0,246$ и $R_{\text{free}} = 0,267$. Итоговая структура фермента характерна для ППЦП (Рисунок 19А) [73]. *N*-конец продомена состоит из трёх α -спиралей, а *C*-концевой участок закрывает субстрат-связывающую борозду, не формируя вторичных структур. Каталитический домен состоит из двух структурных доменов: *N*-концевого L-домена, сформированного четырьмя α -спиралями и одним антипараллельным β -листом, и *C*-концевого R-домена, в состав которого входят два антипараллельных β -листа и три короткие α -спирали. Между этими структурными доменами расположена V-образная субстрат-связывающая борозда. Она сформирована главным образом самой длинной α -спиралью L-домена, на *N*-конце которой расположен Ala154 (который соответствует каталитическому Cys154 в интактном тритикаине- α), и крупным β -листом из четырёх цепей, в котором располагаются каталитические His290 и Asn310 (Рисунок 19А). Всего в структуру тритикаина- α -GM^{C154A} входит 325 а.о., из которых 306 (95%) находятся в предпочтительных конформациях, а 14 (4%) – в разрешенных (Рисунок 19Б).

Таблица 6 – Параметры собранных данных дифракции кристалла 6HIS-тритикаина- α -GM^{C154A} и полученной по ним структуры белка. В скобках указаны параметры для части данных с наилучшим разрешением.

Собранные данные	
Пространственная группа симметрии	P 65
Параметры ячейки: a, b, c (Å) α, β, γ (°)	126,69, 126,69, 42,36 90, 90, 120
Разрешение (Å)	109,72-3,46 (3,79-3,46)
R_{merge}	0,059 (0,540)
$I / \sigma(I)$	9,4 (1,4)
Полнота данных (%)	78,6 (29,5)
Избыточность	1,9 (1,8)

Полученная структура	
Разрешение (Å)	109,7-3,46 (3,96-3,46)
Количество отражений	4178 (706)
R / R_{free}	0,9410 (1,0021)
Число атомов:	2548
Белка	2548
Лиганда/ионов	0
Воды	0
B -фактор	164,0
Среднеквадратичное отклонение (RMSD):	
Длин связей (Å)	0,23
Углов (°)	0,48

Полученная структура тритикаина- α соответствует структуре цистеиновых протеиназ (d.3) согласно базе данных SCOPe. Она может использоваться для дальнейших исследований структурных особенностей, определяющих субстратную специфичность ППЦП.

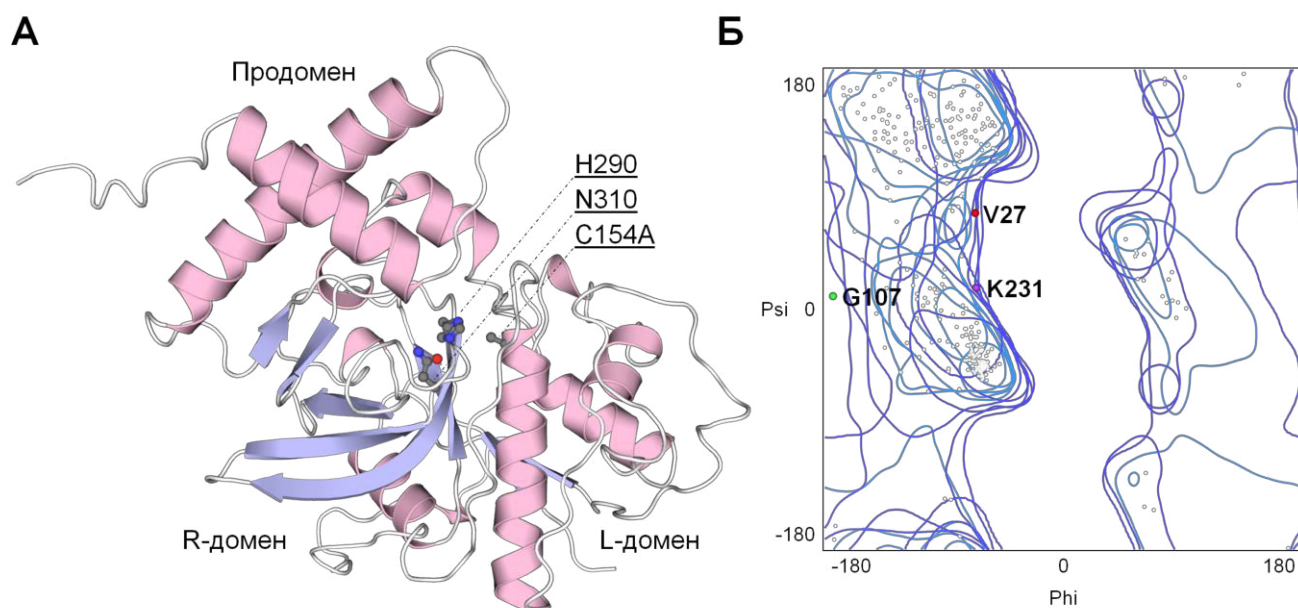


Рисунок 19 – Результаты рентгеноструктурного анализа кристалла белка 6HIS-тритикаина- α -GM^{C154A}. А – Полученная третичная структура фермента (PDB ID: 8ZLQ). Структура повернута так, что R-домен находится слева, а L-домен – справа, чтобы было видно каталитическую триаду. Цветом выделены участки полипептида с разной структурной укладкой. Подписаны структурные домены фермента и а.о., составляющие каталитическую триаду. Б – карта Рамачандрана структуры тритикаина- α -GM^{C154A}, построенная в программе MolProbity. Подписаны а.о., которые, согласно значениям их углов ϕ и ψ , находятся в запрещённой области.

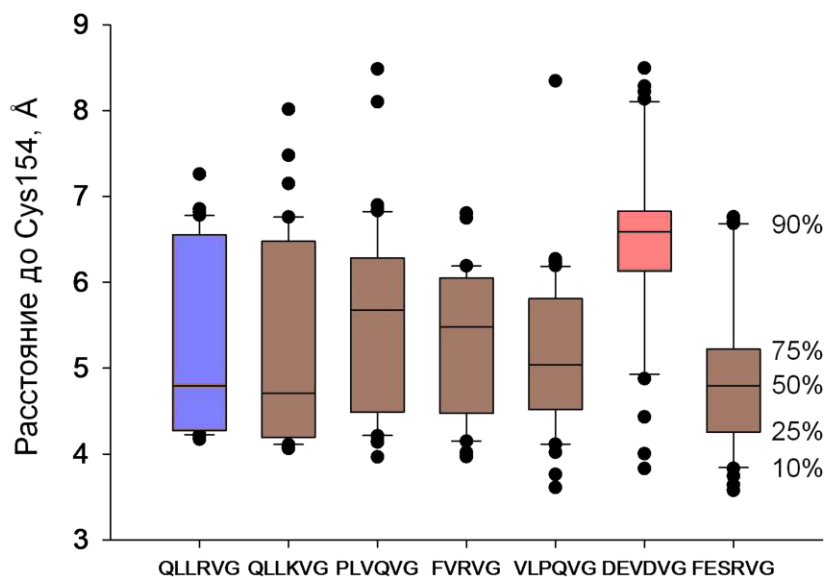
3.4 Субстрат-связывающая борозда тритикаина- α содержит гидрофобный карман и отрицательно заряженные аминокислотные остатки

Чтобы определить структурные особенности, лежащие в основе субстратной специфичности тритикаина- α , описанной выше, кристаллическая структура протеиназы была далее использована для моделирования структур комплексов с олигопептидами. Используемые с этой целью пептиды содержали те же аминокислотные последовательности, что и флуорогенные субстраты, на которых была изучена субстратная специфичность тритикаина- α (Раздел 3.2), а именно: QLLR↓VG, QLLK↓VG, PLVQ↓VG, FVR↓VG, VLPQ↓VG, DEVD↓VG и FESR↓VG (стрелкой указано место гидролиза). Предсказание структур проводилось в программе AlphaFold по последовательности каталитического домена тритикаина- α с интактным каталитическим Cys154. Для каждого пептида

было получено по 40 моделей взаимодействия. Модели сравнили между собой по расстоянию между карбонильным атомом углерода а.о. в P₁-положении пептида и атомом серы Cys154 фермента (Рисунок 20А).

Все пептиды взаимодействовали с субстрат-связывающей бороздой тритикаина-α (Рисунки 20Б и А.1). В лучших моделях комплексов (согласно ранжированию в программе AlphaFold) пептиды QLLR↓VG, QLLK↓VG, PLVQ↓VG и FVR↓VG связывались схожим образом, при котором а.о. в P₂-положении пептида был локализован в гидрофобном кармане S₂ на поверхности тритикаина-α. При этом атакуемый углерод а.о. в P₁-положении QLLR↓VG и QLLK↓VG чаще располагался вблизи серы каталитического Cys154 по сравнению с PLVQ↓VG и FVR↓VG (Рисунок 20А). Остальные пептиды связывались несколько иначе. Например, в лучшей модели комплекса с VLPQ↓VG в сайт S₂ были направлены Leu в P₃- и Gln в P₁-положении пептида, вероятно, из-за конформационных ограничений Pro в P₂ (Рисунок А.1). Пептид DEVD↓VG обычно располагался так, что Asp в P₁ находился далеко от Cys154 (Рисунок 20А). Например, в его лучшей структуре в сайте S₁ находился Gly, что соответствует месту расщепления DEVDVG↓ (Рисунок 20Б). В лучшей структуре комплекса тритикаина-α с пептидом FESR↓VG цепь пептида была развёрнута в субстрат-связывающей борозде (GVR↓SEF) по сравнению с ожидаемым ее положением (Рисунок А.1). При этом Arg в P₁ располагался в сайте связывания S₁, благодаря чему атакуемый атом углерода всё равно находился рядом с серой Cys154 (Рисунок 20А). Таким образом, в полученных структурах положительно заряженный а.о. в P₁-положении QLLR↓VG, QLLK↓VG и FESR↓VG связывался ближе к Cys154 чаще, чем а.о. в P₁ других пептидов из сравнения. В то же время в DEVD↓VG отрицательно заряженный Asp в P₁-положении редко располагался достаточно близко к Cys154 для нуклеофильной атаки (Рисунок 20). Эти структуры подтверждают данные о наличии специфичности тритикаина-α к положительно заряженным по сравнению с отрицательно заряженными а.о. в P₁-положении субстрата (Раздел 3.2).

А



Б

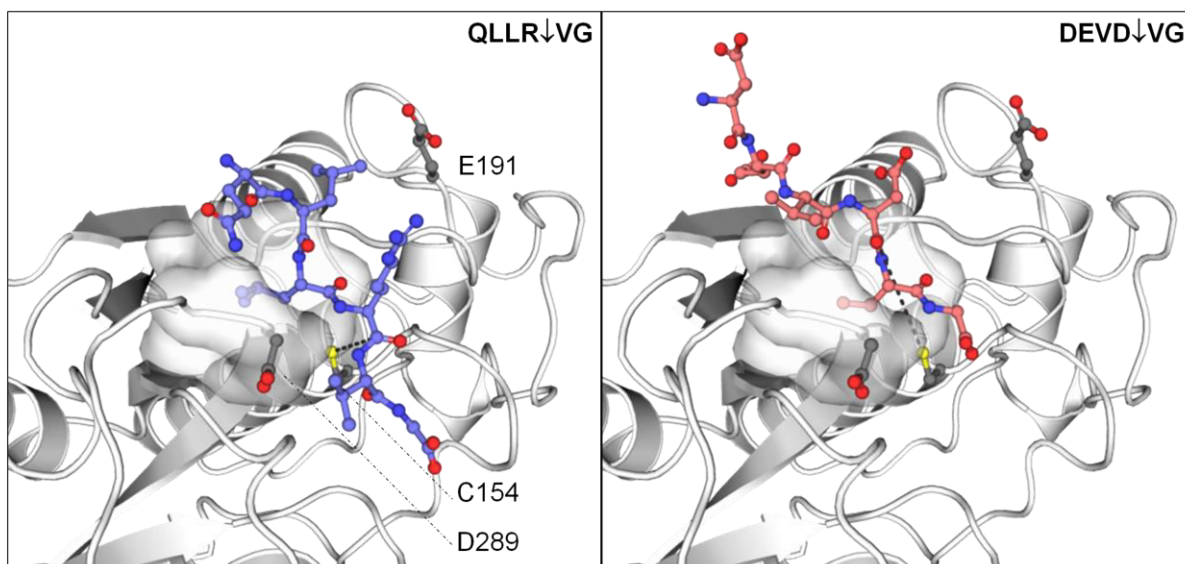
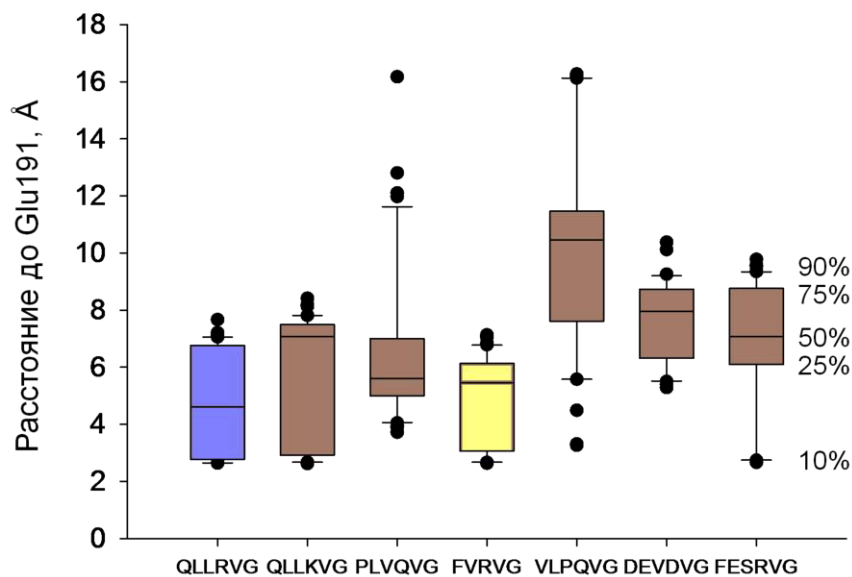


Рисунок 20 – Сравнение предсказанных моделей комплексов олигопептидов, связанных с каталитическим доменом тритикаина- α . А – Расстояния между карбонильным атомом углерода а.о. в P_1 -положении пептида и атомом серы Cys154 фермента в полученных моделях комплексов. Здесь и на Рисунках 21А и 22А справа от графика указаны проценты на примере данных для FESRVG. Б – Лучшие модели комплексов тритикаина- α с пептидами QLLRVG и DEVDVG согласно ранжированию в программе AlphaFold. Гидрофобный карман в S_2 -сайте представлен в виде поверхности. Каталитический Cys154, Glu191 и Asp289 в сайте связывания S_1 фермента выделены и подписаны. Жирной штриховой линией указано расстояние, по которому модели комплексов сравниваются на Рисунке 20А. Структуры для всех пептидов представлены на Рисунке А.1.

А



Б

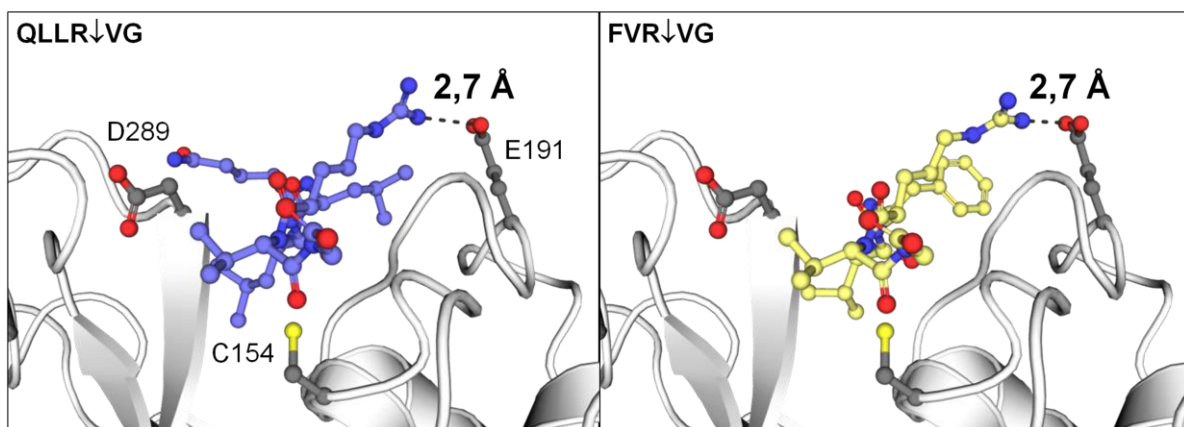
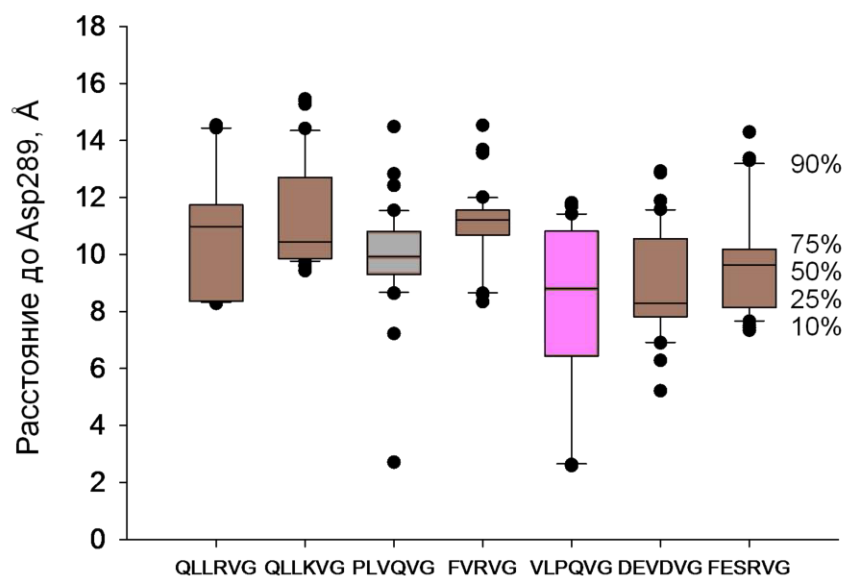


Рисунок 21 – Сравнение предсказанных моделей комплексов олигопептидов, связанных с каталитическим доменом трипсина- α . А – Расстояния между боковыми группами а.о. в P_1 -положении пептида и Glu191 фермента в полученных моделях комплексов. Б – Модели комплексов трипсина- α с пептидами QLLRVG и FVRVG с наименьшим расстоянием между боковыми группами а.о. в P_1 и Glu191 среди структур, в которых расстояние от серы Cys154 до атакуемого атома углерода не превышает 5 Å. Каталитический Cys154, Glu191 и Asp289 в сайте связывания S_1 фермента выделены и подписаны. Штриховой линией указано расстояние, по которому модели комплексов сравниваются на Рисунке 21А. Структуры для всех пептидов представлены на Рисунке А.2.

А



Б

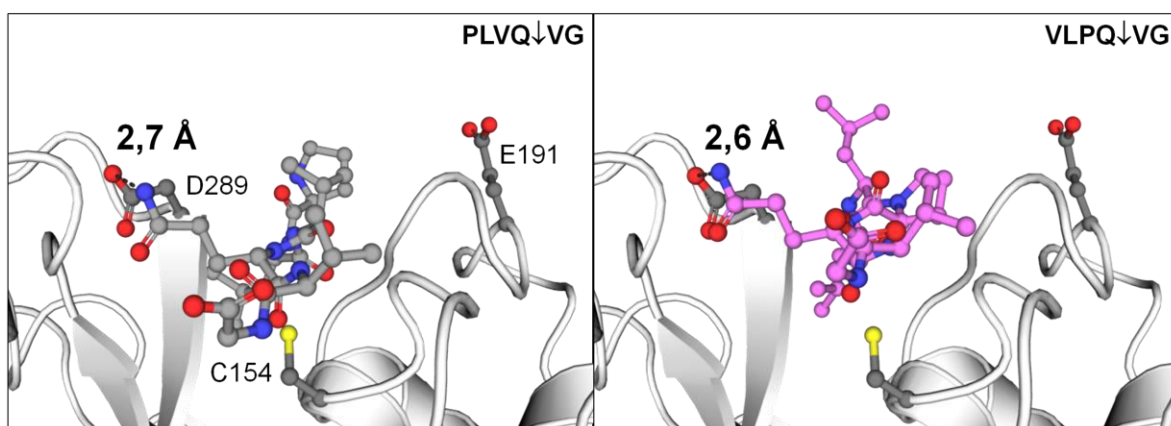


Рисунок 22 – Сравнение предсказанных моделей комплексов олигопептидов, связанных с каталитическим доменом тритикаина- α . А – Расстояния между боковыми группами а.о. в P_1 -положении пептида и Asp289 фермента в полученных моделях комплексов. Б – Модели комплексов тритикаина- α с пептидами PLVQVG и VLPQVG с наименьшим расстоянием между боковыми группами а.о. в P_1 и Asp289 среди структур, в которых расстояние от серы Cys154 до атакуемого атома углерода не превышает 5 Å. Каталитический Cys154, Glu191 и Asp289 в сайте связывания S_1 фермента выделены и подписаны. Штриховой линией указано расстояние, по которому модели комплексов сравниваются на Рисунке 22А. Структуры для всех пептидов представлены на Рисунке А.3.

При визуальном анализе полученных моделей комплексов было обнаружено, что S_1 -сайт связывания на поверхности тритикаина- α фланкирован двумя отрицательно заряженными а.о.: Glu191 и Asp289. Чтобы оценить вероятность участия этих а.о. во взаимодействиях с субстратом, были измерены расстояния между боковыми группами а.о. в P_1 -положении пептида и Glu191 или

Asp289 фермента (Рисунки 21А и 22А). Было обнаружено, что в пептидах с гидрофобным а.о. в P_2 - и положительно заряженным а.о. в P_1 -положении, последний располагался ближе к Glu191 по сравнению с остальными пептидами (Рисунки 21 и А.2). Только среди структур комплексов с QLLR↓VG, QLLK↓VG и FVR↓VG были модели, в которых расстояние между а.о. в P_1 и Glu191 было меньше 3 Å, что необходимо для формирования водородной связи [272], и при этом расстояние между атакуемым атомом углерода и атомом серы Cys154 не превышало 5 Å (Рисунки 21Б и А.2). Наоборот, ни в одной из предсказанных структур комплекса с пептидом DEVD↓VG расстояние от боковой группы а.о. в P_1 до Glu191 или Asp289 не было меньше 5 Å (Рисунки 21А и 22А). В сторону Asp289 были направлены только Gln в P_1 -положении VLPQ↓VG в 12,5% полученных моделей и в одной из структур с PLVQ↓VG (Рисунки 22Б и А.3). Вероятно, этот а.о. менее важен для распознавания заряженных а.о. в P_1 -положении субстрата или ингибитора, но может участвовать в водородных связях.

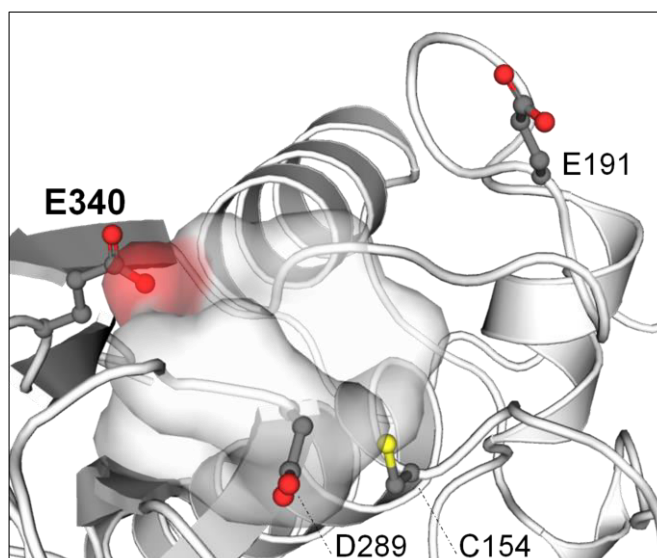


Рисунок 23 – S_2 -сайт в субстрат-связывающей борозде трипсина-α. Гидрофобный карман и Glu340 изображены в виде полупрозрачной поверхности. Каталитический Cys154, а также Glu191 и Asp289 в S_1 - и Glu340 в S_2 -сайтах связывания выделены и подписаны.

Специфичность трипсина-α к гидрофобным а.о. в P_2 -положении субстрата может определяться типичным для семейства ППЦП наличием гидрофобного кармана в S_2 -сайте связывания протеиназы (Рисунок 23). Этот карман

сформирован преимущественно атомами гидрофобных Trp155, Gly196, Leu197, Met198, Ala264, Ile265, Leu288 и Gly291. Примечательно, что на границе S₂-сайта связывания расположен отрицательно заряженный Glu340 (Рисунок 23). Его pK_a = 4,72, согласно расчётам в программе PropKa, поэтому можно предположить, что при pH выше этого значения Glu340 может участвовать в электростатических взаимодействиях с положительно заряженными а.о. в P₂-положении субстрата.

Основываясь на данных, полученных из предсказанных моделей комплексов тритикаина-α с пептидами, можно сделать вывод, что наблюдаемая в кислой среде специфичность тритикаина-α к пептидным субстратам с гидрофобным а.о. в P₂-положении определяется наличием гидрофобного кармана в S₂-сайте связывания фермента, что характерно для семейства ППЦП. В то же время субстратная специфичность к положительно заряженным а.о. в положении P₁ по сравнению с отрицательно заряженными а.о. может быть обусловлена наличием отрицательных зарядов Glu191 и Asp289 в сайте связывания S₁ тритикаина-α.

3.5 Отрицательно заряженные остатки в S₁- и S₂-сайтах связывания тритикаина-α отличают фермент от других папаин-подобных протеиназ

Чтобы определить, насколько консервативны в семействе ППЦП а.о., обнаруженные в составе S₁- и S₂-сайтов связывания тритикаина-α, было проведено выравнивание первичных структур нескольких наиболее изученных ферментов из этой группы. Для этой цели были использованы аминокислотные последовательности растительных ППЦП (папаина, химопапаина и карикаина из *C. papaya*, бромелаина стебля и ананаина из *A. comosus*, актинидина из *A. chinensis*, вигнаина из *V. mungo*, зингипаина из *Z. officinale*, EP-B 2 из *H. vulgare*), круципаина из *T. cruzi* и цистеиновых катепсинов из *H. sapiens* (CtsB, C, F, H, K, L, O, S, V, W и Z) (Рисунки Б.1 и Б.2). Полученное выравнивание аминокислотных последовательностей проверялось на соответствие

выравниванию третичных структур по расположению сравниваемых а.о. в S₁- и S₂-сайтах связывания.

	155	196	197	198	264	265	288	291	340		155	196	197	198	264	265	288	291	340
Тритикаин-α	W	G	L	M	A	I	L	G	E	Катепсин В	W	G	Y	P	A	F	G	A	E
Папаин	W	G	Y	P	V	L	V	A	S	Катепсин С	Y	G	F	P	V	A	T	A	I
Химопапаин	W	G	Y	Q	L	V	L	A	S	Катепсин F	W	G	L	P	A	I	I	A	M
Карикаин	W	G	Y	P	V	V	V	A	S	Катепсин Н	W	G	L	P	A	F	V	A	C
Бромелаин	W	G	W	E	A	V	L	A	D	Катепсин К	W	G	Y	M	A	I	L	A	L
Ананаин	W	G	W	I	A	L	L	A	D	Катепсин L	W	G	L	M	A	I	M	G	A
Актинидин	W	G	Y	I	A	L	I	A	M	Катепсин О	W	G	S	T	I	V	A	A	S
Вигнаин	W	G	L	M	A	I	L	G	M	Катепсин S	W	G	F	M	G	V	V	G	F
Зингипаин	W	G	W	M	T	M	A	A	F	Катепсин V	W	G	F	M	A	M	L	G	A
ЕР-В 2	W	G	L	M	A	V	L	G	E	Катепсин W	W	G	F	V	T	I	V	S	F
Круципаин	W	G	L	M	A	V	L	G	E	Катепсин Z	W	G	N	D	G	I	I	V	H

Рисунок 24 – А.о., составляющие S₂-сайт связывания тритикаина-α, и соответствующие им а.о. в других растительных ППЦП, круципаине и цистеиновых катепсинах человека. Представленные фрагменты взяты из выравниваний аминокислотных последовательностей каталитических доменов соответствующих ферментов (Рисунки Б.1 и Б.2). Нумерация по тритикаину-α. Цветовая схема взята из программы для множественного выравнивания AliView.

Большинство а.о. гидрофобного кармана в S₂-сайте связывания довольно консервативны: Gly196 (нумерация по тритикаину-α) входит в состав S₂-кармана всех использованных в сравнении протеиназ (Рисунок 24). Ile265 заменен на гидрофобные, а Trp155, Leu197, Ala264, Leu288 и Gly291 – на гидрофобные или (значительно реже) на полярные незаряженные а.о. в других протеиназах. На этом фоне выделяется Met198, на месте которого в бромелаине находится Glu68, а в CtsZ – Asp137, которые обеспечивают активность ферментов в отношении субстратов с положительно заряженным а.о. в P₂-положении [3,258]. Glu340 в S₂-сайте связывания тритикаина-α также является переменным. Только ЕР-В 2, круципаин и CtsB также содержат Glu в этой позиции (Рисунок 24). Считается, что этот а.о. отвечает за их особую субстратную специфичность к положительно заряженным а.о. в P₂-положении при нейтральных значениях pH [12,54,103]. На месте Glu340 бромелаин и ананаин содержат отрицательно заряженный Asp, протеиназы из *C. papaya* и CtsO – незаряженный Ser, в то время как у CtsZ в этой позиции находится His295, который вместе с упомянутым выше Asp137

ответственен за субстратную специфичность к полярным а.о. в P₂-положении [3]. В остальных протеиназах на месте Glu340 тритикаина-α находятся гидрофобные а.о. В частности, у CtsK в этой позиции Leu323, который является одним из а.о., определяющих его субстратную специфичность к Pro в P₂-положении[46]. А CtsS содержит Phe325, боковая цепь которого может принимать две конформации для связывания либо Leu, либо более объемного Phe в P₂ ингибитора [14]. Таким образом, гидрофобный карман в S₂-сайте связывания ППЦП консервативен, при этом различия в субстратной специфичности к P₂-положению, как правило, определяются переменными а.о. на месте Met198 и Glu340 тритикаина-α.

	191		289		191		289
Тритикаин-α	E	—	D	Катепсин В	D	—	G
Папаин	Y	—	D	Катепсин С	Q	—	N
Химопапаин	Y	—	D	Катепсин F	K	—	D
Карикаин	H	—	D	Катепсин Н	H	—	N
Бромелаин	Y	—	N	Катепсин К	D	—	N
Ананаин	Y	—	N	Катепсин L	E	—	D
Актинидин	R	—	D	Катепсин О	Y	—	N
Вигнаин	Q	—	N	Катепсин S	K	—	N
Зингипаин	H	—	N	Катепсин V	Q	—	D
ЕР-В 2	D	—	D	Катепсин W	D	—	D
Круципаин	S	—	D	Катепсин Z	G	—	N

Рисунок 25 – А.о., фланкирующие S₁-сайт связывания тритикаина-α, и соответствующие им а.о. в других растительных ППЦП, круципаине и цистеиновых катепсинах человека. Представленные фрагменты взяты из выравниваний аминокислотных последовательностей каталитических доменов соответствующих ферментов (Рисунки Б.1 и Б.2). Нумерация по тритикаину-α. Цветовая схема взята из программы для множественного выравнивания AliView.

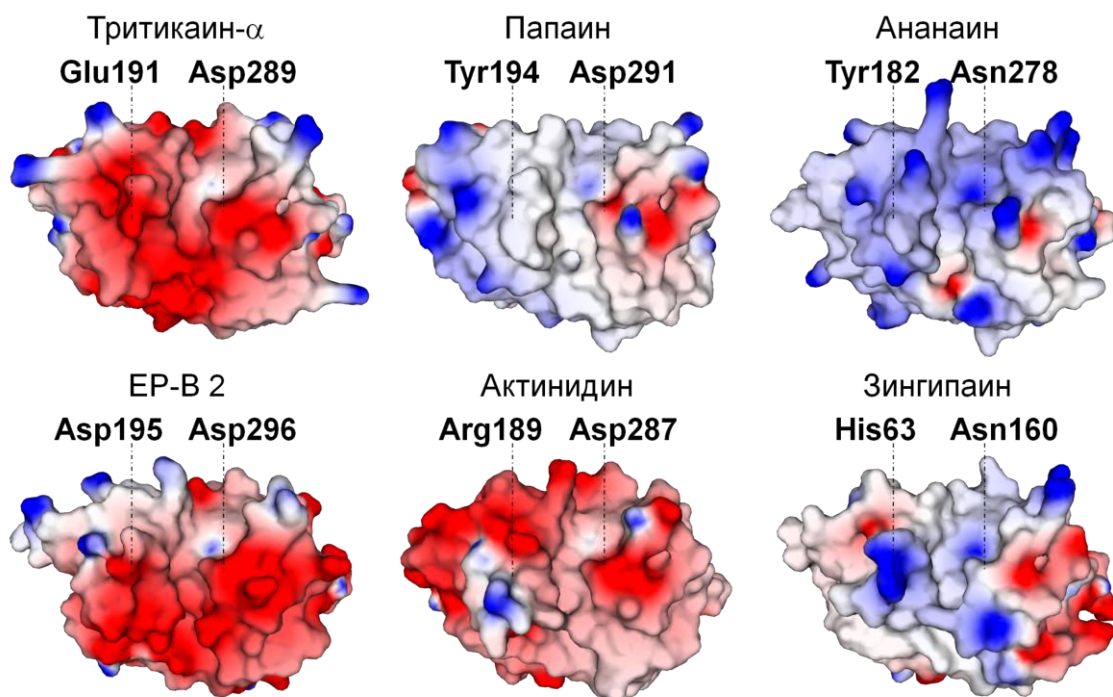


Рисунок 26 – Распределение зарядов на поверхности каталитических доменов тритикаина- α , EP-B 2, папаина, актинидина, ананаина и зингипаина при pH 4,6 (PDB ID: 2CIO, 6YCB, 2ACT, 1CQD и 2FO5, соответственно) [257–261]. Расчёт зарядов атомов и моделирование поля электростатического потенциала были выполнены в программе ProKSim. Поверхность белка окрашена в зависимости от величины электростатического потенциала от -100 мВ (красный) до +100 мВ (синий). Указаны а.о., которые расположены по бокам S_1 -сайта связывания.

Также было проведено сравнение а.о., фланкирующих S_1 -сайт связывания ППЦП (Рисунок 25). Asp289 оказался довольно консервативным: протеиназы из *C. papaya*, актинидин, EP-B 2, круципаин, CtsF, L, V и W также содержат в этой позиции Asp, CtsB – Gly, а остальные протеиназы, использованные в сравнении – полярный незаряженный Asn. Следом за этим а.о. все исследованные протеиназы содержат His из каталитической триады, и далее находятся довольно консервативные гидрофобные а.о. (Рисунки Б.1 и Б.2). Glu191 тритикаина- α , наоборот, вариабелен среди ППЦП: только в CtsL в этой позиции также находится Glu, а в EP-B 2, CtsB, K и W – отрицательно заряженный Asp. Многие протеиназы содержат ароматические His или Tyr в этом положении; также встречаются Arg, Lys, Gln, Ser и Gly (Рисунок 25). Хотя Glu191 вариабелен среди ППЦП, за ним следует консервативный мотив GCxGG, а с *N*-конца он соединен с распространённым среди цистеиновых катепсинов человека Asn (Рисунки Б.1 и

Б.2). Вариабельный Glu191 тритикаина- α в S_1 -сайте связывания ППЦП, возможно, вносит вклад в различия в субстратной специфичности этих протеиназ подобно Glu340 в S_2 -сайте. В случае тритикаина- α отрицательно заряженные Glu191, Asp289 и Glu340, вероятно, влияют на селективность протеиназы, формируя отрицательно заряженную поверхность субстрат-связывающей борозды заметную при сравнении с другими растительными ППЦП (Рисунок 26).

3.6 Специфичность тритикаина- α к P_1 -положению субстрата зависит от значения pH среды и Glu191 в S_1 -сайте

Glu191 и Asp289, pK_a которых равны соответственно 4,57 и 3,09 согласно программе PropKa, могут быть протонированы и незаряжены или нести отрицательный заряд в зависимости от значения pH. В связи с этим было решено определить, как значение pH влияет на специфичность тритикаина- α к а.о. в P_1 -положении субстрата. Для этого были проведены измерения активности фермента в диапазоне pH от 3,6 до 7,5. В качестве субстратов были использованы флуорогенные пептиды *Ac-QLLR-AMC* и *Ac-QLLK-AMC*, в отношении которых тритикаин- α проявлял наибольшую активность (Рисунок 17), а также *Ac-QLLQ-AMC* и *Ac-QLLD-AMC* с незаряженным и отрицательно заряженным полярными а.о. в P_1 -положении, соответственно. В то время как максимальная скорость расщепления субстратов *Ac-QLLR-AMC*, *Ac-QLLK-AMC* и *Ac-QLLQ-AMC* возрастает с повышением pH до 6,5, скорость гидролиза *Ac-QLLD-AMC* падает (Рисунок 27). Более того, при значениях pH 6,5 и 7,5 *Ac-QLLD-AMC* устойчив к расщеплению тритикаином- α . Вероятно, это происходит из-за депротонирования Glu191 и Asp289 при повышении pH, в результате чего они приобретают отрицательный заряд и под действием электростатических сил взаимно отталкиваются с Asp в P_1 -положении субстрата (который также депротонируется). При этом эффективность гидролиза всех субстратов снижается при pH 7,5, что может быть вызвано снижением каталитической активности фермента, который подобно другим ППЦП более активен в кислой среде (Рисунок 27) [22].

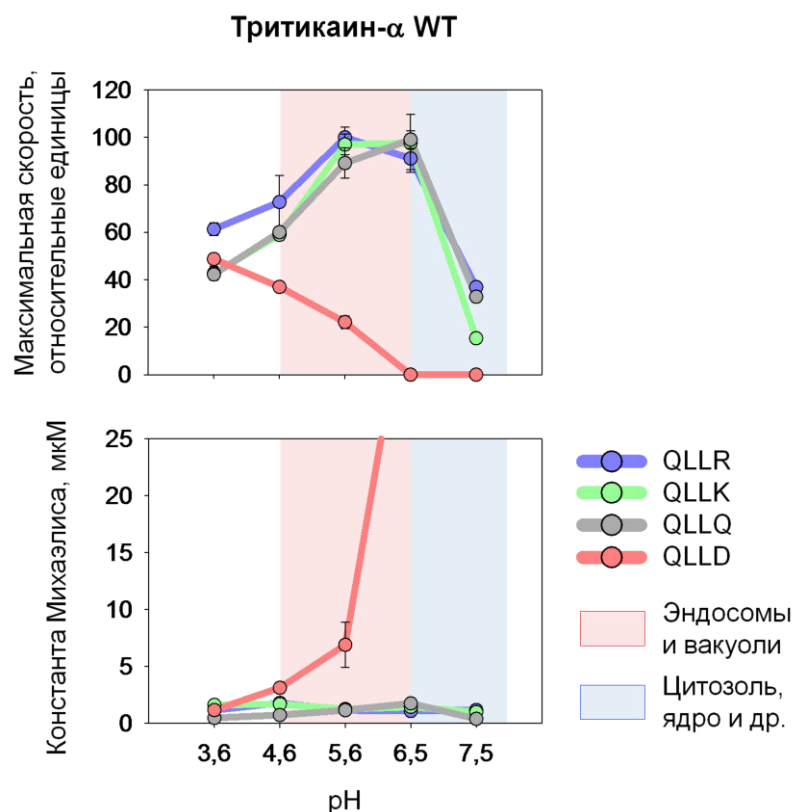


Рисунок 27 – Параметры каталитической активности тритикаина- α в реакциях гидролиза субстратов *Ac*-QLLR-AMC, *Ac*-QLLK-AMC, *Ac*-QLLQ-AMC и *Ac*-QLLD-AMC при значениях pH от 3,6 до 7,5. Фоном обозначены диапазоны pH, характерные для разных компартментов клетки. Каждое значение было получено на основе графиков, построенных по трём независимым измерениям активности тритикаина- α на диапазоне концентраций субстратов. Значения констант Михаэлиса для реакций тритикаина- α с субстратом *Ac*-QLLD-AMC при pH 6,5 и 7,5 показаны как "стремящиеся к бесконечности" для наглядности. Статистическая значимость рассчитывалась согласно t-критерию Стьюдента для независимых выборок (Таблица В.1).

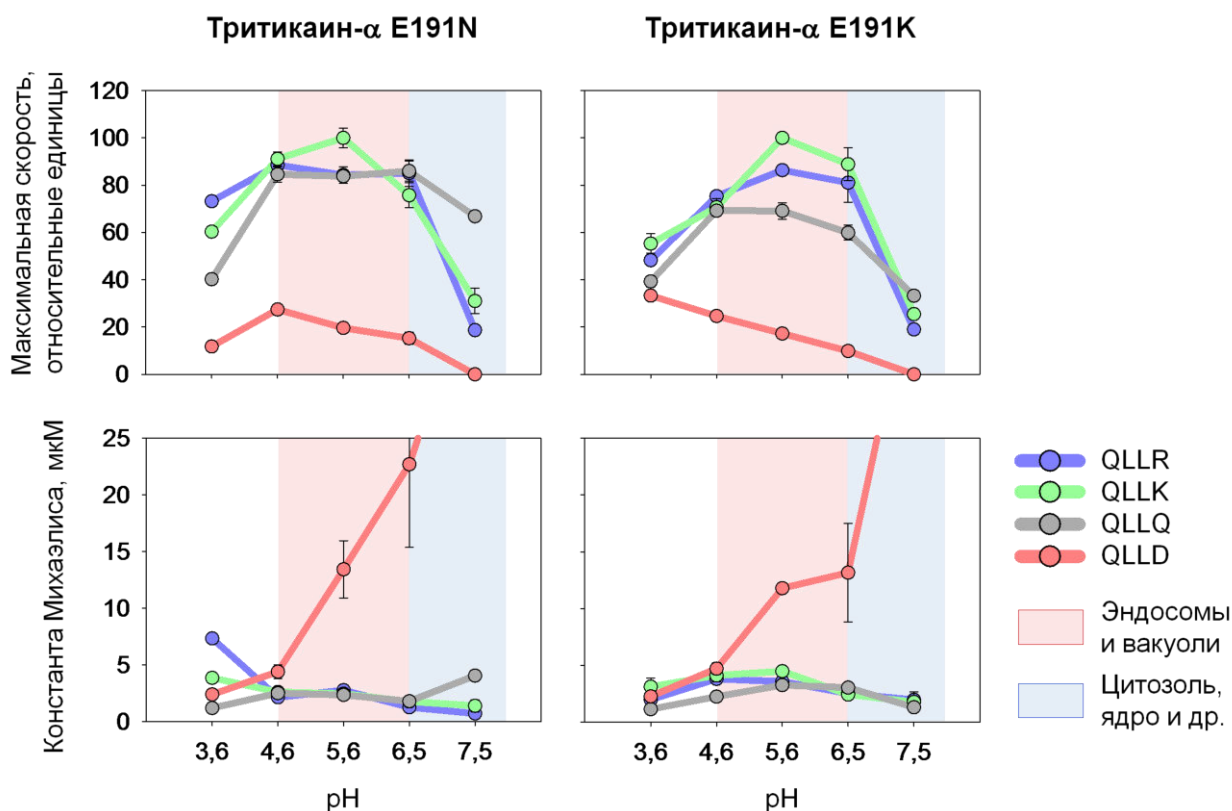


Рисунок 28 – Параметры каталитической активности мутантов тритикаина-α с заменами Glu191Asn (слева) и Glu191Lys (справа) в реакциях гидролиза субстратов *Ac*-QLLR-*AMC*, *Ac*-QLLK-*AMC*, *Ac*-QLLQ-*AMC* и *Ac*-QLLD-*AMC* при значениях pH от 3,6 до 7,5. Фоном обозначены диапазоны pH, характерные для разных компартментов. Каждое значение было получено на основе графиков, построенных по трём независимым измерениям активности мутантных форм тритикаина-α на диапазоне концентраций субстратов. Значения констант Михаэлиса для реакций протеиназ с субстратом *Ac*-QLLD-*AMC* при pH 7,5 показаны как "стремящиеся к бесконечности" для наглядности. Статистическая значимость рассчитывалась согласно t-критерию Стьюдента для независимых выборок (Таблица В.1).

Из пары Glu191 и Asp289 в S₁-сайте связывания тритикаина-α именно Glu191 отличает протеиназу от других ППЦП (Рисунок 25). Поэтому было решено оценить вклад этого а.о. в распознавание Р₁-положения субстрата при разных значениях pH. Для этой цели были получены две мутантные формы тритикаина-α с заменой отрицательно заряженного Glu191 либо на незаряженный Asn, либо на положительно заряженный Lys. Тенденция к увеличению максимальной скорости расщепления субстратов *Ac*-QLLR-*AMC*, *Ac*-QLLK-*AMC* и *Ac*-QLLQ-*AMC* с одновременным снижением скорости расщепления субстрата *Ac*-QLLD-*AMC* при повышении значения pH сохранилась для обоих мутантов

(Рисунок 28). Что интересно, оба мутанта, в отличие от тритикаина- α дикого типа, были способны гидролизовать *Ac-QLLD-AMC* при pH 6,5. Причём самое низкое значение константы Михаэлиса реакции гидролиза *Ac-QLLD-AMC* при этом pH было выявлено для мутанта с заменой Glu191Lys (Рисунок 28). Так как величина константы Михаэлиса обратно пропорциональна константе скорости образования фермент-субстратного комплекса, это может означать, что мутант тритикаина- α с заменой Glu191Lys связывает субстрат *Ac-QLLD-AMC* при pH 6,5 эффективнее, чем дикий тип и мутант с заменой Glu191Asn. Вероятно, Lys191 в S_1 -сайте связывания может образовывать ионные взаимодействия с Asp в P_1 -положении субстрата. При этом константы Михаэлиса для реакций с остальными субстратами всё равно ниже, предположительно, из-за наличия интактного Asp289 в S_1 -сайте связывания.

Полученные данные указывают на то, что специфичность тритикаина- α к субстратам с незаряженными и положительно заряженными а.о. в P_1 -положении по сравнению с субстратами с отрицательно заряженными а.о. становится более выраженной при повышении значений pH до нейтральных. Важный вклад в эту субстратную специфичность вносит Glu191 в S_1 -сайте связывания фермента.

ОБСУЖДЕНИЕ

ППЦП представляют собой большую группу протеолитических ферментов, которые были обнаружены в бактериях, археях, простейших, растениях, животных и вирусах [73]. Цистеиновые катепсины человека относятся к наиболее изученным ППЦП из-за их вовлеченности в процессы онкогенеза [116]. В связи с этим ведется активный поиск ингибиторов этих протеиназ с учетом их структурных особенностей и специфичности [24]. Растительным протеиназам уделяется меньше внимания, хотя и они находят своё применение в медицине [198]. Несмотря на большое количество открытых ППЦП, а именно 232 охарактеризованные протеиназы (при 21368 идентифицированных последовательностях) на момент написания этой работы, распространено мнение, что эти ферменты обладают низкой специфичностью [73]. За исключением некоторых ППЦП с экзопептидазной активностью, протеиназы этого семейства эффективнее всего гидролизуют субстраты с гидрофобным а.о. в P_2 -положении [1,24]. Считается, что специфичные взаимодействия с субстратом происходят также в S_1 - и S_1' -сайтах связывания, а остальные не играют важной роли в распознавании а.о. [1].

Тем не менее даже небольшие отличия в структуре сайтов связывания гомологичных протеиназ могут влиять на эффективность их взаимодействия с субстратами и ингибиторами. Всего два отличающихся а.о. в активном центре CtsS (Phe325 и Lys178) обеспечивают более эффективное ингибирование специфичными винилсульфоновыми ингибиторами с Leu или Phe в P_2 -положении по сравнению с CtsK и L [14,194]. А наличие объемного заряженного Arg255 в субстрат-связывающей борозде CtsS снижало эффективность ингибирования N-десметил талассоспирамидом C в 15 раз по сравнению с CtsK, у которого в этой позиции Ser, и в 46 раз по сравнению с CtsL с Gly в этом положении [195]. Идентификация этих структурных особенностей, влияющих на специфичность, была произведена на основе рентгеноструктурного анализа комплексов и молекулярного докинга ингибиторов в структуры ферментов. Также, используя

данные о структуре различных протеиназ, можно разрабатывать новые ферменты с заданными свойствами. Для повышения каталитической активности ППЦП эрватамина С из *Ervatamia coronaria* его структуру сравнили со структурой эрватамина А – более активной, но менее стабильной протеиназы [273]. Обнаруженные различия в структуре субстрат-связывающей борозды были учтены при создании мутантной формы (Ser32Thr и Ala67Tyr) эрватамина С, которая оказалась активнее исходного фермента в 8 раз [273,274]. CtsK крысы *Rattus norvegicus* обладает более слабой каталитической активностью в отношении субстрата Z-LR-AMC по сравнению с CtsK человека [275]. Увеличить активность этой протеиназы смогли мутации Ser248Ala и Val274Leu (как в CtsK человека) [276]. Точечными мутациями также можно менять специфичность ферментов. Мутантная форма папаина с одной заменой Ser338Glu гидролизовала субстрат Z-RR-AMC в 100 раз эффективнее дикого типа, а аналогичная мутация в CtsS (Phe325Glu) повысила специфичность фермента к Z-RR-MCA в 77 раз [41,277]. Двойной мутант CtsL (Leu182Tyr и Ala327Leu как в CtsK) приобрёл специфичность к Pro в P₂-положении и был способен гидролизовать субстрат Z-GPR-MCA в 100 раз лучше протеиназы дикого типа [47].

Изучение специфичности и картирование сайтов связывания ППЦП используется как для выявления природных субстратов и ингибиторов этих протеиназ, так и в разработке новых ферментов и ингибиторов для дальнейшего применения в медицине. Одной из перспективных протеиназ для терапевтического применения является тритикаин-α из пшеницы *T. aestivum* [231]. Было показано, что тритикаин-α способен расщеплять глютеносодержащие белки, которые выполняют запасающую функцию в семенах зерновых культур [22]. Известно, что продукты неполного расщепления глютена могут вызывать аллергическую реакцию в желудочно-кишечном тракте у людей, страдающих от непереносимости глютена [222]. Тритикаин-α же может гидролизовать глютеносодержащие белки до нетоксичных фрагментов и, следовательно, применяться в заместительной терапии при целиакии [22]. Однако для разработки эффективного препарата на основе тритикаина-α необходимы исследования структуры и

специфичности, которые могут предоставить данные для оптимизации и регуляции активности и специфичности этой протеиназы. Целью данной работы является описание субстратной специфичности ППЦП тритикаина- α из пшеницы *T. aestivum* с идентификацией структурных особенностей его субстрат-связывающей борозды, которые могут формировать специфичность фермента в зависимости от условий среды.

Для достижения цели исследования необходимо было получить активную растворимую форму тритикаина- α , а наиболее распространенным методом получения белков для применения в терапевтических целях является экспрессия генов этих белков в клетках *E. coli* [278]. Однако большинство попыток экспрессии генов ППЦП в бактериях приводили к получению рекомбинантного белка в нерастворимой форме, что требовало последующего проведения рефолдинга *in vitro*, во время которого обычно параллельно происходят два процесса: созревание активного фермента и его автокаталитическая деградация [279]. В ходе проведения представленной работы были предприняты попытки продукции тритикаина- α в форме каталитического домена, тритикаина- α -GM (продомен и каталитический домен в одной полипептидной цепи), а также коэкспрессии генов каталитического и продомена фермента в виде отдельных полипептидов. Только экспрессия гена тритикаина- α -GM приводила к получению активной растворимой протеиназы, что указывает на важную функцию внутримолекулярного шаперона, которую выполняет продомен тритикаина- α . Ранее уже было показано, что продомен участвует в правильном фолдинге некоторых цистеиновых протеиназ [280]. Например, продукция CtsL и S без продомена приводила к получению ферментов с активностью в сотни раз ниже, чем у катепсинов, продуцированных с полноразмерным продоменом [281,282]. Кроме того, на примере папаина было показано, что продомен вовлечен в формирование субстрат-связывающей борозды [80]. Точечные мутации в той части продомена, которая закрывает S₂-сайт связывания в зимогене, приводили к изменению специфичности папаина в отношении а.о. в R₂-положении субстрата [80]. Также в ходе исследований структур растительных ППЦП было

установлено, что все протеиназы этой группы содержат как минимум две дисульфидные связи [7]. Поэтому для улучшения эффективности фолдинга и растворимости тритикаина- α был выбран штамм *E. coli* Rosetta gami B (DE3), который содержит мутации в генах тиоредоксинредуктазы и глутатионредуктазы, что способствует формированию дисульфидных связей в белке [269]. Кроме того, в представленной работе было установлено, что эффективность продукции и активность тритикаина- α были выше для структур с N-концевым 6HIS по сравнению с C-концевым 6HIS [230].

Первые исследования активности тритикаина- α проводились на глютенных белках и коллагене, которые состоят преимущественно из гидрофобных и полярных незаряженных а.о., что ограничивает возможности описания субстратной специфичности фермента [22]. Для дальнейшего исследования активности протеиназы в отношении различных субстратов был проведен пептидный скрининг субстратной специфичности тритикаина- α с использованием библиотеки из 3525 пептидов. Было обнаружено, что исследуемая протеиназа была наиболее активна в отношении субстрата с последовательностью QLLR↓VGCV в положениях $P_4P_3P_2P_1\downarrow P_1'P_2'P_3'P_4'$. Последующее измерение активности тритикаина- α в отношении флуорогенных субстратов подтвердило специфичность фермента в отношении пептидов с гидрофобным а.о. в P_2 - и положительно заряженным а.о. в P_1 -положении. Специфичность большинства ППЦП к гидрофобным а.о. в P_2 -положении является давно установленным фактом [41]. Однако CtsB, который содержит Glu324 в S_2 -сайте связывания, может гидролизовать в том числе и субстраты, содержащие Arg или Lys в P_2 -положении [54]. Замена Glu324Gln приводила к потере этой специфичности. Многие ППЦП предпочитают положительно заряженные и полярные незаряженные а.о. в P_1 субстрата, однако считается, что эта специфичность не так важна как специфичность по P_2 -положению [2]. В случае тритикаина- α прослеживается явное предпочтение к положительно заряженным и полярным незаряженным по сравнению с отрицательно заряженными а.о. в P_1 -положении субстрата.

Распознавание субстрата обеспечивается слабыми взаимодействиями, которые формируются между активным центром протеиназы и пептидом [283]. Чтобы определить, какие взаимодействия с субстратом может формировать тритикаин- α в своём активном центре, в ходе данной работы впервые была получена трёхмерная структура этой цистеиновой протеиназы. Из 480 проверенных растворов кристаллизация неактивного мутанта 6HIS-тритикаина- α -GM^{C154A} происходила только в 100 мМ натрий-ацетатном буфере, pH 4,6, с 3,5% полиэтиленгликолем 4000 (w/v) и 15% глицерином (v/v) в качестве соосадителей. Наилучшее разрешение, с которым происходила дифракция кристаллов, составило 3,46 Å. Такое низкое разрешение могло быть вызвано высокой конформационной подвижностью продомена. Структуры ППЦП с продоменом в среднем дифракционируются с более низким разрешением по сравнению со структурами без продомена [26]. В связи с таким низким разрешением структура тритикаина- α была получена с использованием метода молекулярного замещения на основе предсказанной структуры тритикаина- α AF-Q0WVG8-F1 с последующим уточнением полученной структуры в электронной плотности [284,285]. Итоговая структура тритикаина- α соответствовала электронной плотности с $R = 0,246$ и $R_{\text{free}} = 0,267$ (PDB ID: 8ZLQ). Полученная структура относится к структурам с $\alpha+\beta$ укладкой цистеиновых протеиназ (d.3) согласно базе SCOPe [240].

Полученная структура тритикаина- α была далее использована для предсказания структур комплексов с каждым пептидом из ранее протестированных флуорогенных субстратов для идентификации сайтов связывания на поверхности фермента. На основе анализа полученных моделей были охарактеризованы два субстрат-связывающих сайта S_1 и S_2 , которые объясняют субстратную специфичность фермента. Специфичность тритикаина- α к гидрофобным а.о. в P_2 -положении субстрата, вероятно, обеспечивается наличием гидрофобного S_2 -кармана, образованного Trp155, Gly196, Leu197, Met198, Ala264, Ile265, Leu288 и Gly291. Большинство из этих а.о. либо такие же, либо заменены на гидрофобные а.о. в других ППЦП, что объясняет характерную

специфичность ППЦП к гидрофобным а.о. в P_2 -положении субстрата [2,41]. Сайт связывания S_2 тритикаина- α также содержит отрицательно заряженный Glu340, который заменен гидрофобным или незаряженным полярным а.о. в большинстве других ППЦП. Например, у CtsK в этой позиции Leu323, который является одним из а.о., определяющих его уникальную специфичность к Pro в P_2 -положении субстрата. Мутации Tyr181Leu и Leu323Ala, как в CtsL, приводили к потере специфичности к Pro и коллагенолитической активности CtsK [46]. Однако EP-B 2, круципаин и CtsB также имеют Glu в этой позиции. В этих протеиназах он отвечает за уникальную специфичность к положительно заряженным а.о. в P_2 -положении при нейтральном pH [12,54,103]. Замена Glu324Gln в CtsB снижала эффективность расщепления субстратов с Arg в P_2 , но не оказывала влияния на активность в отношении субстратов с Phe [54]. Бромелаин и ананаин содержат Asp в этой позиции, который также может взаимодействовать с положительно заряженными а.о. в P_2 -положении субстратов [258]. В отличие от других ППЦП у CtsZ на месте Glu340 расположен His295, который определяет специфичность протеиназы к субстратам с длинными гидрофильными боковыми группами в P_2 [3]. Наши результаты показывают, что тритикаин- α не может гидролизовать *Ac-FESR-AMC* с полярным а.о. в P_2 в кислых условиях среды. Однако дальнейшие исследования субстратной специфичности фермента при нейтральных значениях pH, а также использование субстратов с другими полярными а.о. в P_2 -положении могут быть полезны для характеристики специфичности тритикаина- α .

В предсказанных комплексах пептиды с положительно заряженным а.о. в P_1 -положении связывались рядом с каталитическим Cys154 чаще остальных. При этом а.о. в P_1 субстрата располагался между двумя отрицательно заряженными а.о. тритикаина- α : Glu191 и Asp289. Asp289 довольно консервативен: в других ППЦП на его месте либо также расположен Asp, либо полярный незаряженный Asn. При этом ранее для других ППЦП уже было показано, что а.о., соответствующие Asp289 в тритикаине- α , могут контактировать с а.о. связанных пептидов. Например, в недавно полученных кристаллических структурах CtsV с олигопептидами, Asp276 протеиназы либо образует электростатические

взаимодействия с боковой цепью Arg в P₁', либо водородные связи с атомами главной цепи а.о. в P₁-положении [286]. Было обнаружено, что Asp283 в круципаине участвует в водородных и электростатических взаимодействиях с ингибиторами в кристаллических структурах [287]. Его замена на Ser приводила к потере специфичности к Arg в P₁'-положении [288]. В кристаллических структурах соответствующих комплексов карбонильный кислород Asp296 в EP-B 2 из *H. vulgare* и Asn404 в человеческом CtsC взаимодействует с амидной группой а.о. в P₁ их ингибиторов [261,289]. В настоящем исследовании в предсказанных структурах комплексов пептидов с тритикаином-α в сторону Asp289 была направлена боковая группа Gln в P₁-положении PLVQ↓VG и VLPQ↓VG. Таким образом, в ППЦП а.о., соответствующие Asp289 тритикаина-α, преимущественно формируют водородные связи с атомами главной и боковой цепей а.о. в P₁. Однако они также могут участвовать в электростатических взаимодействиях с радикалом а.о. в P₁'-положении субстрата или ингибитора.

В то же время в предсказанных комплексах с тритикаином-α положительно заряженные а.о. в P₁-положении пептидов QLLR↓VG, QLLK↓VG и FVR↓VG зачастую были направлены в сторону Glu191 фермента. Ранее было обнаружено, что Gln176 в CtsV, соответствующий Glu191 тритикаина-α, также взаимодействует с длинными боковыми цепями а.о. в P₃-положении нескольких ингибиторов, сокристаллизованных с ферментом [286]. Это согласуется с результатами, полученными в настоящем исследовании, поскольку Glu191 тритикаина-α располагается близко к а.о. в P₃ во многих из предсказанных комплексов (Приложение А). По-видимому, а.о., соответствующие Glu191 тритикаина-α, несущие длинную боковую цепь, могут участвовать в распознавании а.о. в различных положениях связанного пептида. При этом Glu191, несмотря на расположенную рядом консервативную последовательность, заменен на довольно разнообразные варианты а.о. в различных ППЦП. В частности, у растительных ППЦП Glu191 обычно заменен на нейтральные или положительно-заряженные а.о., что делает S₁-сайт связывания тритикаина-α уникальным среди растительных протеиназ. Особенность тритикаина-α в

сравнении с другими растительными ППЦП видна и по распределению зарядов на поверхности фермента. Поверхность тритикаина- α содержит больше отрицательных зарядов по сравнению с папаином, ананаином, актинидином и зингипаином при pH 4,6, что также может указывать на важный вклад отрицательно заряженных а.о. в субстратную специфичность фермента.

А.о. с противоположными зарядами могут образовывать солевые мостики – вид слабой связи, который включает в себя электростатические и водородные взаимодействия, и потому является самым энергоэффективным среди нековалентных контактов [290]. Однако наличие заряда на а.о. зависит от условий среды. Например, при повышении значения pH происходит депротонирование а.о., которое приводит к приобретению отрицательного заряда Asp и Glu и потере положительного заряда Arg, Lys и His в зависимости от их локального pK_a [291]. Измерение активности тритикаина- α на диапазоне pH от 3,6 до 7,5 показало потерю активности к субстрату с отрицательно заряженным Asp в P_1 -положении (*Ac-QLLD-AMC*) на фоне повышения активности в отношении остальных субстратов (*Ac-QLLR-AMC*, *Ac-QLLK-AMC* и *Ac-QLLQ-AMC*) при повышении pH до 6,5. Так как наличие отрицательно заряженного Glu191 является отличительной особенностью тритикаина- α и в предсказанных комплексах он мог взаимодействовать с положительно-заряженными а.о. в P_1 -положении пептидов, было решено проверить, как влияет именно Glu191 на субстратную специфичность протеиназы. Для этого были сделаны два мутанта тритикаина- α с заменой на незаряженный полярный а.о. Glu191Asn и с заменой на положительно заряженный а.о. Glu191Lys. Оба мутанта были способны гидролизовать *Ac-QLLD-AMC* при pH 6,5, причём мутант с заменой Glu191Lys взаимодействовал с этим субстратом с самой низкой константой Михаэлиса. Однако, как и дикий тип, оба мутанта гидролизovali *Ac-QLLD-AMC* хуже, чем остальные субстраты на всем диапазоне значений pH, что может быть объяснено наличием интактного Asp289 в их S_1 -сайте связывания.

Большинство предыдущих исследований кинетики ППЦП были выполнены при одном (обычно кислом) значении pH. Однако есть небольшое количество

данных, показывающих зависимость специфичности этих ферментов от pH среды. Например, активность круципаина из *T. cruzi* на диапазоне значений pH от 3,5 до 10,0 растёт в отношении субстратов *Z-RR-AMC* и *Z-RA-AMC* из-за наличия Glu330 в S₂-сайте связывания, но не меняется в отношении *Z-FR-AMC* [55]. Соотношение активности Tc-cpl-1 из *T. canis* в отношении субстратов *Z-FR-AMC* и *Z-RR-AMC* тоже меняется с 16:1 при pH 5,0 до 1:1 при pH 6,5 [101,102]. Также известно, что диссоциация продомена с последующей активацией CtsB и L провоцируется низкими значениями pH [98,292]. Эти данные указывают на то, что продомены ППЦП работают как pH-зависимые ингибиторы, наиболее эффективные при нейтральном pH [293]. Кроме того, для CtsB был разработан искусственный pH-чувствительный ингибитор *Z-RK-AOMK* [103]. Он селективно ингибирует CtsB при pH 7,2 по сравнению с pH 4,6 и более эффективен в отношении CtsB по сравнению с CtsC, H, K, V и S. Компьютерное моделирование показало, что Arg в P₂-положении взаимодействует с карбоксильной группой Glu324, тогда как Lys в P₁ взаимодействует с Glu201 и Asn151 [103]. pH-селективные ингибиторы могут быть использованы для предотвращения ферментативной активности при определенных условиях, например патологическая активность человеческих катепсинов в микроокружении опухоли [116]. В случае с тритикаином-α подобные ингибиторы могут быть использованы для подавления аутокаталитической деградации. Например, ранее было показано, что обратимый ингибитор MMTS может служить защитным агентом во время длительной очистки и рефолдинга цистеиновых протеиназ [72].

Полученные в представленной работе результаты описывают структуру и субстратную специфичность протеиназы тритикаина-α, которая является перспективной для применения в заместительной терапии при непереносимости глютена. Эти данные могут быть использованы для повышения эффективности тритикаина-α в требуемых условиях путём внесения модификаций в структуру самого фермента или для улучшения стабильности с помощью селективного pH-зависимого ингибитора. Также полученные данные имеют значимость для

фундаментальных исследований, дополняя имеющуюся информацию о структуре и специфичности ферментов из семейства ППЦП.

ЗАКЛЮЧЕНИЕ

В рамках данной работы был разработан подход к получению активного растворимого тритикаина- α . Только продукция тритикаина- α , состоящего из продомена и каталитического домена в одной полипептидной цепи, приводила к получению активной протеиназы. Эти данные указывают на важную функцию внутримолекулярного шаперона, которую выполняет продомен фермента. Также растворимость и активность удалось повысить за счёт продукции фермента с *N*-концевым 6HIS при 18°C в штамме *E. coli* Rosetta gami B (DE3).

В данном исследовании было обнаружено, что в кислой среде тритикаин- α проявляет специфичность к субстратам с гидрофобным а.о. в P_2 - и с положительно заряженным а.о. в P_1 -положении. Специфичность фермента к положительно заряженным а.о. в P_1 -положении подтвердилась и при моделировании структур комплексов, где атакуемый атом углерода в пептидах QLLR↓VG, QLLK↓VG и FESR↓VG чаще располагался рядом с каталитическим Cys154 в сравнении с остальными пептидами. Однако при измерении активности фермента, *Ac*-FESR-AMC, будучи единственным субстратом с полярным а.о. в P_2 -положении, был устойчив к гидролизу.

В ходе данной работы впервые была получена третичная структура тритикаина- α методом рентгеноструктурного анализа. Структура протеиназы с разрешением 3,46 Å была размещена в базе белковых структур с PDB ID: 8ZLQ и может быть использована для дальнейшего изучения структурных особенностей, определяющих субстратную специфичность этого семейства протеиназ.

В предсказанных структурах комплексов тритикаина- α с пептидами был обнаружен консервативный гидрофобный карман в S_2 -сайте фермента, характерный для ППЦП. При этом S_1 -сайт тритикаина- α фланкирован отрицательно заряженными Glu191 и Asp289. Только боковая группа Gln в P_1 -положении PLVQ↓VG и VLPQ↓VG была направлена в сторону Asp289. В то же время положительно заряженные а.о. в P_1 -положении пептидов QLLR↓VG, QLLK↓VG и FVR↓VG зачастую были направлены в сторону Glu191. Glu191 и

Asp289 в S₁-сайте и Glu340 в S₂-сайте формируют отрицательно заряженную поверхность субстрат-связывающей борозды тритикаина- α , которая отличает его от других ППЦП и может определять его специфичность.

Измерение активности тритикаина- α и его мутантов Glu191Asn и Glu191Lys в диапазоне pH от 3,6 до 7,5 показало, что специфичность тритикаина- α к P₁-положению субстрата зависит от значения pH. Активность всех ферментов в отношении субстрата с отрицательно заряженным Asp в P₁-положении (*Ac-QLLD-AMC*) снижалась на фоне повышения активности в отношении остальных субстратов (*Ac-QLLR-AMC*, *Ac-QLLK-AMC* и *Ac-QLLQ-AMC*) при повышении pH от 3,6 до 6,5. В отличие от дикого типа, оба мутанта тритикаина- α были способны гидролизовать *Ac-QLLD-AMC* при pH 6,5. Эти данные указывают на то, что при нейтральных значениях pH Glu191 фермента важен для специфичности тритикаина- α .

На основе полученных данных может быть разработан pH-зависимый ингибитор для стабилизации фермента в исследовательских и медицинских целях. Также эти данные могут быть экстраполированы на семейство ППЦП и использованы для описания структурных особенностей, определяющих их различия в специфичности в разных условиях среды и между собой. Полученные данные могут быть использованы для разработки диагностических зондов и терапевтических ингибиторов для ППЦП, а также для поиска и дизайна новых протеиназ для применения в энзиматической терапии.

ВЫВОДЫ

- 1) Наличие продомена, ковалентно связанного с каталитическим доменом, у профермента тритикаина- α критически необходимо для созревания активного фермента.
- 2) Тритикаин- α предпочитает субстраты с положительно заряженными и полярными незаряженными аминокислотными остатками в P_1 -положении.
- 3) Определена кристаллическая структура тритикаина- α . Подобно другим папаин-подобным цистеиновым протеиназам, тритикаин- α имеет консервативный гидрофобный карман в S_2 -сайте связывания. Однако тритикаин- α отличается от большинства ферментов семейства наличие отрицательно заряженных Glu191 в S_1 -сайте и Glu340 в S_2 -сайте субстрат-связывающей борозды.
- 4) Субстратная специфичность тритикаина- α более выражена в нейтральных условиях, чем в кислых. В отличие от поведения в кислой среде, при нейтральных значениях pH протеиназа утрачивает способность гидролизовать субстраты, содержащие отрицательно заряженный аспартат в положении P_1 .
- 5) Результаты компьютерного моделирования комплексов фермент-субстрат и экспериментальные данные, полученные с использованием сайт-направленного мутагенеза протеиназы, свидетельствуют о том, что остатки Glu191 и Asp289 в S_1 -сайте тритикаина- α являются ключевыми детерминантами pH-зависимой субстратной специфичности фермента.

СПИСОК СОКРАЩЕНИЙ И ОБОЗНАЧЕНИЙ

4M β NA	–	4-метокси- β -нафтиламид
6HIS	–	аффинная метка из шести гистидинов
A.o.	–	аминокислотный остаток
BKM	–	внеклеточный матрикс
ДМСО	–	диметилсульфоксид
ДНК	–	дезоксирибонуклеиновая кислота
ИПТГ	–	изопропил- β -D-1-тиогалактопиранозид
ММП	–	матриксная металлопротеиназа
Н.д.	–	нет данных
ПААГ	–	полиакриламидный гель
ППЦП	–	папаин-подобная цистеиновая протеиназа
ПЦР	–	полимеразная цепная реакция
РНК	–	рибонуклеиновая кислота
ТЕМЕД	–	<i>N,N,N',N'</i> -тетраметилэтилендиамин
Тритикаин- α -CatD – каталитический домен тритикаина- α		
Тритикаин- α -GM – тритикаин- α , состоящий из продомена и каталитического домена		
Тритикаин- α -ProD – продомен тритикаина- α		
ЭДТА	–	этилендиаминтетрауксусная кислота
Abz	–	2-аминобензойная кислота
Ac	–	ацетил
АМС	–	7-амино-4-метилкумарин
АОМК	–	ацилоксиметилкетон
C-конец	–	карбоксильный конец белка или пептида
Cts	–	катепсин
CtsC Δ Ex	–	искусственная усеченная форма CtsC без эксклюзионного домена
EDDnp	–	<i>N</i> -(2,4-динитрофенил)-этилендиамин

FAK	–	киназа фокальных контактов
HEPES	–	4-(2-гидроксиэтил)-1-пиперазинэтансульфоновая кислота
hPhe	–	гомофенилаланин
IC ₅₀	–	концентрация полумаксимального ингибирования
Ig	–	иммуноглобулин
JAM	–	Junctional Adhesion Molecule; семейство молекул адгезии
K _i	–	константа связывания ингибитора
LB	–	Lysogeny Broth; питательная среда для бактериальных клеток
MALDI-TOF/TOF	–	тандемная времяпролетная масс-спектрометрия с матричной лазерной десорбцией/ионизацией
MCA	–	7-метоксикумарин-4-ацетил
MMTS	–	S-метилметантиосульфонат
M _r	–	молекулярная масса
<i>N</i> -конец	–	амино-конец белка или пептида
Nle	–	норлейцин
OD ₆₀₀	–	оптическая плотность образца при длине волны 600 нм и оптическом пути 1 см
<i>p</i>	–	уровень значимости по <i>t</i> -критерию
P ₄ , P ₃ , P ₂ , P ₁ , P ₁ ', P ₂ ', P ₃ '	–	обозначения а.о. субстрата, согласно порядку их расположения от гидролизуемой пептидной связи в сторону <i>N</i> -конца (P _i) и в сторону <i>C</i> -конца (P _i ')
PDB ID	–	идентификационный номер структуры в Protein Data Bank
pGlu	–	пироглутаминовая кислота
pH	–	мера кислотности водного раствора; pH = -lg[H ⁺]
pK _a	–	показатель константы диссоциации кислоты; pK _a = -lg(K _a)
PNA	–	<i>p</i> -нитроанилид
pTyr	–	фосфотирозин
R	–	фактор достоверности; мера согласованности между кристаллографической моделью и экспериментальным массивом рентгеновских данных, которая рассчитывается на основе ~90% экспериментальных данных

R_{free} – фактор достоверности; мера согласованности между кристаллографической моделью и экспериментальным массивом рентгеновских данных, которая рассчитывается на основе ~10% экспериментальных данных исключенных из расчёта R

R_{merge} – показатель, используемый для оценки качества данных, полученных из дифракционных измерений

RMSD – среднеквадратичное отклонение

rpm – обороты в минуту

$S_4, S_3, S_2, S_1, S_1', S_2', S_3'$ – обозначения сайтов связывания на поверхности протеиназы, в которых располагаются соответствующие а.о. субстрата ($P_4, P_3, P_2, P_1, P_1', P_2', P_3'$)

SDS – додецилсульфат натрия

Src – коактиватор стероидных рецепторов

Tos – тозил

Z – бензилоксикарбонил

Для обозначения а.о. в работе используются трёхбуквенные и однобуквенные сокращения в соответствии с номенклатурой Международного союза теоретической и прикладной химии (IUPAC) [294]:

A – Ala – Аланин	L – Leu – Лейцин
R – Arg – Аргинин	K – Lys – Лизин
N – Asn – Аспарагин	M – Met – Метионин
D – Asp – Аспарагиновая кислота	P – Pro – Пролин
V – Val – Валин	S – Ser – Серин
H – His – Гистидин	Y – Tyr – Тирозин
G – Gly – Глицин	T – Thr – Треонин
Q – Gln – Глутамин	W – Trp – Триптофан
E – Glu – Глутаминовая кислота	F – Phe – Фенилаланин
I – Ile – Изолейцин	C – Cys – Цистеин

СПИСОК ЛИТЕРАТУРЫ

1. Turk D. et al. Revised definition of substrate binding sites of papain-like cysteine proteases // *Biol. Chem.* 1998. Vol. 379, № 2. P. 137–147.
2. Choe Y. et al. Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities // *J. Biol. Chem.* 2006. Vol. 281, № 18. P. 12824–12832.
3. Guncar G. et al. Crystal structure of cathepsin X: a flip-flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease // *Structure*. 2000. Vol. 8, № 3. P. 305–313.
4. Puzer L. et al. Comparative substrate specificity analysis of recombinant human cathepsin V and cathepsin L // *Arch. Biochem. Biophys.* 2004. Vol. 430, № 2. P. 274–283.
5. Liu H. et al. Computational study on substrate specificity of a novel cysteine protease 1 precursor from *Zea mays* // *Int. J. Mol. Sci.* 2014. Vol. 15, № 6. P. 10459–10478.
6. Turk V. et al. Cysteine cathepsins: from structure, function and regulation to new frontiers // *Biochim. Biophys. Acta*. 2012. Vol. 1824, № 1. P. 68–88.
7. Richau K.H. et al. Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics // *Plant Physiol.* 2012. Vol. 158, № 4. P. 1583–1599.
8. Lecaille F. et al. Rat cathepsin K: Enzymatic specificity and regulation of its collagenolytic activity // *Biochim. Biophys. Acta - Proteins Proteomics*. 2020. Vol. 1868, № 2. P. 140318.
9. Stack C., Dalton J.P., Robinson M.W. The phylogeny, structure and function of trematode cysteine proteases, with particular emphasis on the *Fasciola hepatica* cathepsin L family // *Adv. Exp. Med. Biol.* 2011. Vol. 712. P. 116–135.
10. Tomoo K. Development of cathepsin inhibitors and structure-based design of cathepsin B-specific inhibitor // *Curr. Top. Med. Chem.* 2010. Vol. 10, № 7. P. 696–707.

11. Tsuji A. et al. Purification and characterization of cathepsin B-like cysteine protease from cotyledons of daikon radish, *Raphanus sativus* // *FEBS J.* 2008. Vol. 275, № 21. P. 5429–5443.
12. Polticelli F. et al. Probing the cruzain S2 recognition subsite: a kinetic and binding energy calculation study // *Biochemistry.* 2005. Vol. 44, № 8. P. 2781–2789.
13. Groves M.R. et al. Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes // *Proteins.* 1998. Vol. 32, № 4. P. 504–514.
14. Pauly T.A. et al. Specificity determinants of human cathepsin s revealed by crystal structures of complexes // *Biochemistry.* 2003. Vol. 42, № 11. P. 3203–3213.
15. Baker P.J., Numata K. Chemoenzymatic synthesis of poly(L-alanine) in aqueous environment // *Biomacromolecules.* 2012. Vol. 13, № 4. P. 947–951.
16. Tuck S.A. et al. Bronchoscopic thermal vapor ablation in a canine model of emphysema // *Int. J. Chron. Obstruct. Pulmon. Dis.* 2012. Vol. 7. P. 21–31.
17. Hu W. et al. Enzyme catalytic promiscuity: The papain-catalyzed Knoevenagel reaction // *Biochimie.* 2012. Vol. 94, № 3. P. 656–661.
18. Connell J.F., Del Guercio L.R., Rousselot L.M. Debricin; clinical experiences with a new proteolytic enzyme in surgical wounds // *Surg. Gynecol. Obstet.* 1959. Vol. 108, № 1. P. 93–99.
19. Rosenberg L. et al. A novel rapid and selective enzymatic debridement agent for burn wound management: a multi-center RCT // *Burns.* Elsevier Ltd and International Society of Burns Injuries, 2014. Vol. 40, № 3. P. 466–474.
20. Langer V. et al. Enzymatic debridement of large burn wounds with papain-urea: Is it safe? // *Med. J. Armed Forces India.* 2013. Vol. 69, № 2. P. 144–150.
21. Bertassoni L.E., Marshall G.W. Papain-gel degrades intact nonmineralized type I collagen fibrils // *Scanning.* 2009. Vol. 31, № 6. P. 253–258.
22. Savvateeva L. V et al. Glutenase and collagenase activities of wheat cysteine protease Triticain- α : feasibility for enzymatic therapy assays // *Int. J. Biochem. Cell Biol.* 2015. Vol. 62. P. 115–124.
23. Machado M.V. New Developments in Celiac Disease Treatment // *Int. J. Mol. Sci.*

2023. Vol. 24, № 2. P. 945.
24. Petushkova A.I., Savvateeva L. V, Zamyatnin A.A. Structure determinants defining the specificity of papain-like cysteine proteases // *Comput. Struct. Biotechnol. J.* 2022. Vol. 20. P. 6552–6569.
 25. Coulombe R. et al. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment // *EMBO J.* 1996. Vol. 15, № 20. P. 5492–5503.
 26. Berman H.M. et al. The Protein Data Bank // *Nucleic Acids Res.* 2000. Vol. 28, № 1. P. 235–242.
 27. Turk D. et al. Crystal structure of cathepsin B inhibited with CA030 at 2.0-Å resolution: A basis for the design of specific epoxysuccinyl inhibitors // *Biochemistry.* 1995. Vol. 34, № 14. P. 4791–4797.
 28. Somoza J.R., Palmer J.T., Ho J.D. The crystal structure of human cathepsin F and its implications for the development of novel immunomodulators // *J. Mol. Biol.* 2002. Vol. 322, № 3. P. 559–568.
 29. Zhao B. et al. Crystal structure of human osteoclast cathepsin K complex with E-64 // *Nat. Struct. Biol.* 1997. Vol. 4, № 2. P. 109–111.
 30. Somoza J.R. et al. Crystal structure of human cathepsin V // *Biochemistry.* 2000. Vol. 39, № 41. P. 12543–12551.
 31. Schrödinger L. The PyMOL Molecular Graphics System, Version 2.0.
 32. Fujishima A. et al. The crystal structure of human cathepsin L complexed with E-64 // *FEBS Lett.* 1997. Vol. 407, № 1. P. 47–50.
 33. Oanca G. et al. Exploring the catalytic reaction of cysteine proteases // *J. Phys. Chem. B.* 2020. Vol. 124, № 50. P. 11349–11356.
 34. Karrer K.M., Peiffer S.L., DiTomas M.E. Two distinct gene subfamilies within the family of cysteine protease genes // *Proc. Natl. Acad. Sci. U. S. A.* 1993. Vol. 90, № 7. P. 3063–3067.
 35. Cygler M., Mort J.S. Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity // *Biochimie.* 1997. Vol. 79, № 11. P. 645–652.
 36. Turk D., Turk B., Turk V. Papain-like lysosomal cysteine proteases and their

- inhibitors: drug discovery targets? // *Biochem. Soc. Symp.* 2003. № 70. P. 15–30.
37. Illy C. et al. Role of the occluding loop in cathepsin B activity // *J. Biol. Chem.* 1997. Vol. 272, № 2. P. 1197–1202.
38. Nägler D.K. et al. Major increase in endopeptidase activity of human cathepsin B upon removal of occluding loop contacts // *Biochemistry.* 1997. Vol. 36, № 41. P. 12608–12615.
39. Guncar G. et al. Crystal structure of porcine cathepsin H determined at 2.1 Å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function // *Structure.* 1998. Vol. 6, № 1. P. 51–61.
40. Turk D. et al. Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases // *EMBO J.* 2001. Vol. 20, № 23. P. 6570–6582.
41. Brömme D. et al. Engineering the S2 subsite specificity of human cathepsin S to a cathepsin L- and cathepsin B-like specificity // *J. Biol. Chem.* 1994. Vol. 269, № 48. P. 30238–30242.
42. Biniossek M.L. et al. Proteomic identification of protease cleavage sites characterizes prime and non-prime specificity of cysteine cathepsins B, L, and S // *J. Proteome Res.* 2011. Vol. 10, № 12. P. 5363–5373.
43. Lützner N., Kalbacher H. Quantifying cathepsin S activity in antigen presenting cells using a novel specific substrate // *J. Biol. Chem.* 2008. Vol. 283, № 52. P. 36185–36194.
44. Lecaille F. et al. Probing cathepsin K activity with a selective substrate spanning its active site // *Biochem. J.* 2003. Vol. 375, № Pt 2. P. 307–312.
45. Korenč M., Lenarčič B., Novinec M. Human cathepsin L, a papain-like collagenase without proline specificity // *FEBS J.* 2015. Vol. 282, № 22. P. 4328–4340.
46. Lecaille F. et al. Selective inhibition of the collagenolytic activity of human cathepsin K by altering its S2 subsite specificity // *Biochemistry.* 2002. Vol. 41, № 26. P. 8447–8454.
47. Lecaille F. et al. The S2 subsites of cathepsins K and L and their contribution to

- collagen degradation // *Protein Sci.* 2007. Vol. 16, № 4. P. 662–670.
48. Rebernik M., Lenarčič B., Novinec M. The catalytic domain of cathepsin C (dipeptidyl-peptidase I) alone is a fully functional endoprotease // *Protein Expr. Purif.* 2019. Vol. 157. P. 21–27.
 49. Stack C.M. et al. Structural and functional relationships in the virulence-associated cathepsin L proteases of the parasitic liver fluke, *Fasciola hepatica* // *J. Biol. Chem.* 2008. Vol. 283, № 15. P. 9896–9908.
 50. Smooker P.M. et al. A single amino acid substitution affects substrate specificity in cysteine proteinases from *Fasciola hepatica* // *Protein Sci.* 2000. Vol. 9, № 12. P. 2567–2572.
 51. Norbury L.J. et al. Analysis of *Fasciola* cathepsin L5 by S2 subsite substitutions and determination of the P1-P4 specificity reveals an unusual preference // *Biochimie.* 2012. Vol. 94, № 5. P. 1119–1127.
 52. Poreba M. et al. Fluorescent probes towards selective cathepsin B detection and visualization in cancer cells and patient samples // *Chem. Sci.* 2019. Vol. 10, № 36. P. 8461–8477.
 53. Del Nery E. et al. Specificity of cathepsin B to fluorescent substrates containing benzyl side-chain-substituted amino acids at P1 subsite // *J. Protein Chem.* 2000. Vol. 19, № 1. P. 33–38.
 54. Hasnain S. et al. Characterization of cathepsin B specificity by site-directed mutagenesis. Importance of Glu245 in the S2-P2 specificity for arginine and its role in transition state stabilization // *J. Biol. Chem.* 1993. Vol. 268, № 1. P. 235–240.
 55. Zhai X., Meek T.D. Catalytic Mechanism of Cruzain from *Trypanosoma cruzi* As Determined from Solvent Kinetic Isotope Effects of Steady-State and Pre-Steady-State Kinetics // *Biochemistry.* 2018. Vol. 57, № 22. P. 3176–3190.
 56. Xu J.H. et al. A Commensal Dipeptidyl Aminopeptidase with Specificity for N-Terminal Glycine Degrades Human-Produced Antimicrobial Peptides in Vitro // *ACS Chem. Biol.* 2018. Vol. 13, № 9. P. 2513–2521.
 57. Yamamoto A. et al. Substrate specificity of bovine cathepsin B and its inhibition

- by CA074, based on crystal structure refinement of the complex // *J. Biochem.* 2000. Vol. 127, № 4. P. 635–643.
58. Alves M.F.M. et al. S3 to S3' subsite specificity of recombinant human cathepsin K and development of selective internally quenched fluorescent substrates // *Biochem. J.* 2003. Vol. 373, № Pt 3. P. 981–986.
 59. McGrath M.E. et al. Crystal structure of human cathepsin S // *Protein Sci.* 1998. Vol. 7, № 6. P. 1294–1302.
 60. Clara R.O. et al. *Boophilus microplus* cathepsin L-like (BmCL1) cysteine protease: specificity study using a peptide phage display library // *Vet. Parasitol.* 2011. Vol. 181, № 2–4. P. 291–300.
 61. Corvo I. et al. Substrate Specificity of Cysteine Proteases Beyond the S2 Pocket: Mutagenesis and Molecular Dynamics Investigation of *Fasciola hepatica* Cathepsins L // *Front. Mol. Biosci.* 2018. Vol. 5. P. 40.
 62. Korkmaz B. et al. Lung Protection by Cathepsin C Inhibition: A New Hope for COVID-19 and ARDS? // *J. Med. Chem.* 2020. Vol. 63, № 22. P. 13258–13265.
 63. Desmarais S., Massé F., Percival M.D. Pharmacological inhibitors to identify roles of cathepsin K in cell-based studies: a comparison of available tools // *Biol. Chem.* 2009. Vol. 390, № 9. P. 941–948.
 64. Gauthier J.Y. et al. The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K // *Bioorg. Med. Chem. Lett.* 2008. Vol. 18, № 3. P. 923–928.
 65. Turk B. et al. Bovine stefin C, a new member of the stefin family // *J. Biol. Chem.* 1993. Vol. 268, № 10. P. 7323–7329.
 66. Turk B. et al. Identification of bovine stefin A, a novel protein inhibitor of cysteine proteinases // *FEBS Lett.* 1995. Vol. 360, № 2. P. 101–105.
 67. Turk B. et al. Kinetics of inhibition of bovine cathepsin S by bovine stefin B // *FEBS Lett.* 1994. Vol. 339, № 1–2. P. 155–159.
 68. Mihelic M. et al. Mouse stefins A1 and A2 (Stfa1 and Stfa2) differentiate between papain-like endo- and exopeptidases // *FEBS Lett.* 2006. Vol. 580, № 17. P. 4195–4199.
 69. Li Z. et al. Regulation of collagenase activities of human cathepsins by

- glycosaminoglycans // J. Biol. Chem. 2004. Vol. 279, № 7. P. 5470–5479.
70. Selent J. et al. Selective inhibition of the collagenase activity of cathepsin K // J. Biol. Chem. 2007. Vol. 282, № 22. P. 16492–16501.
71. Petushkova A.I., Zamyatnin A.A. Redox-Mediated Post-Translational Modifications of Proteolytic Enzymes and Their Role in Protease Functioning // Biomolecules. 2020. Vol. 10, № 4. P. 650.
72. Makarov V.A. et al. Novel applications of modification of thiol enzymes and redox-regulated proteins using S-methyl methanethiosulfonate (MMTS) // Biochim. Biophys. acta. Proteins proteomics. 2019. Vol. 1867, № 11. P. 140259.
73. Rawlings N.D. et al. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database // Nucleic Acids Res. 2018. Vol. 46, № D1. P. D624–D632.
74. Barrett A.J. The cystatins: a diverse superfamily of cysteine peptidase inhibitors // Biomed. Biochim. Acta. 1986. Vol. 45, № 11–12. P. 1363–1374.
75. Abrahamson M. Cystatins // Methods Enzymol. 1994. Vol. 244. P. 685–700.
76. Ni J. et al. Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins // J. Biol. Chem. 1997. Vol. 272, № 16. P. 10853–10858.
77. Cheng T. et al. Cystatin M/E is a high affinity inhibitor of cathepsin V and cathepsin L by a reactive site that is distinct from the legumain-binding site. A novel clue for the role of cystatin M/E in epidermal cornification // J. Biol. Chem. 2006. Vol. 281, № 23. P. 15893–15899.
78. Ni J. et al. Cystatin F is a glycosylated human low molecular weight cysteine proteinase inhibitor // J. Biol. Chem. 1998. Vol. 273, № 38. P. 24797–24804.
79. Roy S. et al. The structure of a thermostable mutant of pro-papain reveals its activation mechanism // Acta Crystallogr. D. Biol. Crystallogr. 2012. Vol. 68, № Pt 12. P. 1591–1603.
80. Dutta S. et al. Mutation in the Pro-Peptide Region of a Cysteine Protease Leads to Altered Activity and Specificity-A Structural and Biochemical Approach // PLoS One. 2016. Vol. 11, № 6. P. e0158024.

81. Li Z. et al. Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate // *J. Biol. Chem.* 2002. Vol. 277, № 32. P. 28669–28676.
82. Nascimento F.D. et al. Cathepsin X binds to cell surface heparan sulfate proteoglycans // *Arch. Biochem. Biophys.* 2005. Vol. 436, № 2. P. 323–332.
83. Li Z. et al. The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex // *J. Mol. Biol.* 2008. Vol. 383, № 1. P. 78–91.
84. Vasiljeva O. et al. Recombinant human procathepsin S is capable of autocatalytic processing at neutral pH in the presence of glycosaminoglycans // *FEBS Lett.* 2005. Vol. 579, № 5. P. 1285–1290.
85. Caglic D. et al. Glycosaminoglycans facilitate procathepsin B activation through disruption of propeptide-mature enzyme interactions // *J. Biol. Chem.* 2007. Vol. 282, № 45. P. 33076–33085.
86. Laber J.R. et al. Charge Shielding Prevents Aggregation of Supercharged GFP Variants at High Protein Concentration // *Mol. Pharm.* 2017. Vol. 14, № 10. P. 3269–3280.
87. Bogunia M., Makowski M. Influence of Ionic Strength on Hydrophobic Interactions in Water: Dependence on Solute Size and Shape // *J. Phys. Chem. B.* 2020. Vol. 124, № 46. P. 10326–10336.
88. Gorter J., Gruber M. Cathepsin C: an allosteric enzyme // *Biochim. Biophys. Acta.* 1970. Vol. 198, № 3. P. 546–555.
89. Cigic B., Pain R.H. Location of the binding site for chloride ion activation of cathepsin C // *Eur. J. Biochem.* 1999. Vol. 264, № 3. P. 944–951.
90. Brömme D. et al. Human cathepsin V functional expression, tissue distribution, electrostatic surface potential, enzymatic characterization, and chromosomal localization // *Biochemistry.* 1999. Vol. 38, № 8. P. 2377–2385.
91. Matagne A. et al. The proteolytic system of pineapple stems revisited: Purification and characterization of multiple catalytically active forms // *Phytochemistry.* 2017. Vol. 138. P. 29–51.
92. Sluyterman L.A., Wijdenes J. Proton equilibria in the binding of Zn^{2+} and of methylmercuric iodide to papain // *Eur. J. Biochem.* 1976. Vol. 71, № 2. P. 383–

- 391.
93. Franco E., de Araujo Soares R.M., Meza I. Specific and reversible inhibition of *Entamoeba histolytica* cysteine-proteinase activities by Zn^{2+} : implications for adhesion and cell damage // Arch. Med. Res. 1999. Vol. 30, № 2. P. 82–88.
 94. Baeyens-Volant D. et al. A novel form of ficin from *Ficus carica* latex: Purification and characterization // Phytochemistry. 2015. Vol. 117. P. 154–167.
 95. Gagaoua M., Hoggas N., Hafid K. Three phase partitioning of zingibain, a milk-clotting enzyme from *Zingiber officinale* Roscoe rhizomes // Int. J. Biol. Macromol. 2015. Vol. 73. P. 245–252.
 96. Singh A.N., Dubey V.K. Exploring applications of procerain b, a novel protease from *Calotropis procera*, and characterization by N-terminal sequencing as well as peptide mass fingerprinting // Appl. Biochem. Biotechnol. 2011. Vol. 164, № 5. P. 573–580.
 97. Mnif I.H. et al. A cysteine protease isolated from the latex of *Ficus microcarpa*: purification and biochemical characterization // Appl. Biochem. Biotechnol. 2015. Vol. 175, № 3. P. 1732–1744.
 98. Rozman J. et al. Autocatalytic processing of recombinant human procathepsin B is a bimolecular process // FEBS Lett. 1999. Vol. 459, № 3. P. 358–362.
 99. Fox T. et al. Potent slow-binding inhibition of cathepsin B by its propeptide // Biochemistry. 1992. Vol. 31, № 50. P. 12571–12576.
 100. Harrach T. et al. Isolation and partial characterization of basic proteinases from stem bromelain // J. Protein Chem. 1995. Vol. 14, № 1. P. 41–52.
 101. Loukas A., Selzer P.M., Maizels R.M. Characterisation of Tc-cpl-1, a cathepsin L-like cysteine protease from *Toxocara canis* infective larvae // Mol. Biochem. Parasitol. 1998. Vol. 92, № 2. P. 275–289.
 102. LaLonde J.M. et al. Use of papain as a model for the structure-based design of cathepsin K inhibitors: crystal structures of two papain-inhibitor complexes demonstrate binding to S'-subsites // J. Med. Chem. 1998. Vol. 41, № 23. P. 4567–4576.
 103. Yoon M.C. et al. Selective Neutral pH Inhibitor of Cathepsin B Designed Based

- on Cleavage Preferences at Cytosolic and Lysosomal pH Conditions // ACS Chem. Biol. 2021. Vol. 16, № 9. P. 1628–1643.
104. Borišek J. et al. Development of N-(Functionalized benzoyl)-homocycloleucyl-glycinonitriles as Potent Cathepsin K Inhibitors // J. Med. Chem. 2015. Vol. 58, № 17. P. 6928–6937.
105. Hardegger L.A. et al. Systematic investigation of halogen bonding in protein-ligand interactions // Angew. Chem. Int. Ed. Engl. 2011. Vol. 50, № 1. P. 314–318.
106. Schade M. et al. Highly Selective Sub-Nanomolar Cathepsin S Inhibitors by Merging Fragment Binders with Nitrile Inhibitors // J. Med. Chem. 2020. Vol. 63, № 20. P. 11801–11808.
107. Reinke P.Y.A. et al. Calpeptin is a potent cathepsin inhibitor and drug candidate for SARS-CoV-2 infections // Commun. Biol. 2023. Vol. 6, № 1. P. 1058.
108. Kovalenko I.B. et al. The role of electrostatic interactions in the process of diffusional encounter and docking of electron transport proteins // Dokl. Biochem. Biophys. 2016. Vol. 468, № 1. P. 183–186.
109. Fedorov V. et al. Electrostatic Map of the SARS-CoV-2 Virion Specifies Binding Sites of the Antiviral Cationic Photosensitizer // Int. J. Mol. Sci. 2022. Vol. 23, № 13. P. 7304.
110. Gocheva V. et al. IL-4 induces cathepsin protease activity in tumor-associated macrophages to promote cancer growth and invasion // Genes Dev. 2010. Vol. 24, № 3. P. 241–255.
111. Fehrenbacher N. et al. Sensitization to the lysosomal cell death pathway by oncogene-induced down-regulation of lysosome-associated membrane proteins 1 and 2 // Cancer Res. 2008. Vol. 68, № 16. P. 6623–6633.
112. Bannoud N. et al. Cation-dependent mannose-6-phosphate receptor expression and distribution are influenced by estradiol in MCF-7 breast cancer cells // PLoS One. 2018. Vol. 13, № 8. P. 1–17.
113. Laurent-Matha V. et al. Proteolysis of cystatin C by cathepsin D in the breast cancer microenvironment // FASEB J. 2012. Vol. 26, № 12. P. 5172–5181.

114. Dean R.A., Overall C.M. Proteomics discovery of metalloproteinase substrates in the cellular context by iTRAQ labeling reveals a diverse MMP-2 substrate degradome // *Mol. Cell. Proteomics*. 2007. Vol. 6, № 4. P. 611–623.
115. Gocheva V. et al. Distinct roles for cysteine cathepsin genes in multistage tumorigenesis // *Genes Dev*. 2006. Vol. 20, № 5. P. 543–556.
116. Петушкова А.И. et al. Цистеиновые катепсины: перспективы применения в диагностике и терапии злокачественных опухолей // *Биохимия*. 2019. Vol. 84, № 7. P. 953–971.
117. Rao Malla R. et al. Knockdown of cathepsin B and uPAR inhibits CD151 and $\alpha 3 \beta 1$ integrin-mediated cell adhesion and invasion in glioma // *Mol. Carcinog*. 2013. Vol. 52, № 10. P. 777–790.
118. Kawada A. et al. Cathepsin B and D expression in squamous cell carcinoma // *Br. J. Dermatol*. 1996. Vol. 135, № 6. P. 905–910.
119. Yang K.-M. et al. Co-chaperone BAG2 Determines the Pro-oncogenic Role of Cathepsin B in Triple-Negative Breast Cancer Cells // *Cell Rep*. 2017. Vol. 21, № 10. P. 2952–2964.
120. Spiess E. et al. Cathepsin B activity in human lung tumor cell lines: ultrastructural localization, pH sensitivity, and inhibitor status at the cellular level // *J. Histochem. Cytochem*. 1994. Vol. 42, № 7. P. 917–929.
121. Khan A. et al. Cathepsin B expression and its correlation with tumor-associated laminin and tumor progression in gastric cancer // *Arch. Pathol. Lab. Med*. 1998. Vol. 122, № 2. P. 172–177.
122. Dusek P. et al. Diagnostic Efficiency of Serum and Urine Procathepsin B and Cathepsin B in Patients with Carcinoma of the Urinary Bladder // *Clin. Lab*. 2016. Vol. 62, № 9. P. 1709–1715.
123. Wu D. et al. Cathepsin B may be a potential biomarker in cervical cancer // *Histol. Histopathol*. 2012. Vol. 27, № 1. P. 79–87.
124. Nishikawa H. et al. The role of cathepsin B and cystatin C in the mechanisms of invasion by ovarian cancer // *Gynecol. Oncol*. 2004. Vol. 92, № 3. P. 881–886.
125. Bian B. et al. Cathepsin B promotes colorectal tumorigenesis, cell invasion, and

- metastasis // *Mol. Carcinog.* 2016. Vol. 55, № 5. P. 671–687.
126. Ruan J. et al. Over-expression of cathepsin B in hepatocellular carcinomas predicts poor prognosis of HCC patients // *Mol. Cancer.* 2016. Vol. 15. P. 17.
 127. Dumartin L. et al. AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D // *Cancer Res.* 2011. Vol. 71, № 22. P. 7091–7102.
 128. Shuja S. et al. Cathepsin B activity and protein levels in thyroid carcinoma, Graves' disease, and multinodular goiters // *Thyroid.* 1999. Vol. 9, № 6. P. 569–577.
 129. Angelova A. et al. Immunotherapeutic Potential of Oncolytic H-1 Parvovirus: Hints of Glioblastoma Microenvironment Conversion towards Immunogenicity // *Viruses.* 2017. Vol. 9, № 12. P. 382.
 130. Bengsch F. et al. Cell type-dependent pathogenic functions of overexpressed human cathepsin B in murine breast cancer progression // *Oncogene.* 2014. Vol. 33, № 36. P. 4474–4484.
 131. Niedergethmann M. et al. Prognostic impact of cysteine proteases cathepsin B and cathepsin L in pancreatic adenocarcinoma // *Pancreas.* 2004. Vol. 29, № 3. P. 204–211.
 132. Koblinski J.E. et al. Interaction of human breast fibroblasts with collagen I increases secretion of procathepsin B // *J. Biol. Chem.* 2002. Vol. 277, № 35. P. 32220–32227.
 133. Krueger S. et al. Interactions between human colon carcinoma cells, fibroblasts and monocytic cells in coculture—regulation of cathepsin B expression and invasiveness // *Cancer Lett.* 2005. Vol. 223, № 2. P. 313–322.
 134. Mach L. et al. Proteolytic processing and glycosylation of cathepsin B. The role of the primary structure of the latent precursor and of the carbohydrate moiety for cell-type-specific molecular forms of the enzyme // *Biochem. J.* 1992. Vol. 282, № Pt 2. P. 577–582.
 135. Strojnik T. et al. Cathepsin B immunohistochemical staining in tumor and endothelial cells is a new prognostic factor for survival in patients with brain

- tumors // Clin. Cancer Res. 1999. Vol. 5, № 3. P. 559–567.
136. Lah T.T. et al. Cells producing cathepsins D, B, and L in human breast carcinoma and their association with prognosis // Hum. Pathol. 2000. Vol. 31, № 2. P. 149–160.
 137. Liu Y. et al. Cathepsin B on invasion and metastasis of gastric carcinoma // Chin. Med. J. (Engl). 1998. Vol. 111, № 9. P. 784–788.
 138. Ishibashi O. et al. Breast cancer cells express cathepsins B and L but not cathepsins K or H // Cancer Biochem. Biophys. 1999. Vol. 17, № 1–2. P. 69–78.
 139. Gaumann A. et al. The expression of cathepsins in osteoclast-like giant cells of an anaplastic thyroid carcinoma with tracheal perforation // Pathol. Res. Pract. 2001. Vol. 197, № 4. P. 257–262.
 140. Ruffell B. et al. Cathepsin C is a tissue-specific regulator of squamous carcinogenesis // Genes Dev. 2013. Vol. 27, № 19. P. 2086–2098.
 141. Vazquez-Ortiz G. et al. Overexpression of cathepsin f, matrix metalloproteinases 11 and 12 in cervical cancer // BMC Cancer. 2005. Vol. 5. P. 68.
 142. Sivaparvathi M. et al. Expression and the role of cathepsin H in human glioma progression and invasion // Cancer Lett. 1996. Vol. 104, № 1. P. 121–126.
 143. Fröhlich E., Möhrle M., Klessen C. Cathepsins in basal cell carcinomas: activity, immunoreactivity and mRNA staining of cathepsins B, D, H and L // Arch. Dermatol. Res. 2004. Vol. 295, № 10. P. 411–421.
 144. Waghray A. et al. Analysis of a truncated form of cathepsin H in human prostate tumor cells // J. Biol. Chem. 2002. Vol. 277, № 13. P. 11533–11538.
 145. Re E.C. del et al. Alterations in cathepsin H activity and protein patterns in human colorectal carcinomas // Br. J. Cancer. 2000. Vol. 82, № 7. P. 1317–1326.
 146. Ishida M., Kojima F., Okabe H. Cathepsin K expression in basal cell carcinoma // J. Eur. Acad. Dermatology Venereol. 2013. Vol. 27, № 1. P. e128-30.
 147. Littlewood-Evans A.J. et al. The osteoclast-associated protease cathepsin K is expressed in human breast carcinoma // Cancer Res. 1997. Vol. 57, № 23. P. 5386–5390.
 148. Tsai J.-Y. et al. The effect of catalase on migration and invasion of lung cancer

- cells by regulating the activities of cathepsin S, L, and K // *Exp. Cell Res.* 2014. Vol. 323, № 1. P. 28–40.
149. Ren G. et al. Coronin 3 promotes gastric cancer metastasis via the up-regulation of MMP-9 and cathepsin K // *Mol. Cancer.* 2012. Vol. 11. P. 67.
 150. Munari E. et al. Cathepsin K expression in castration-resistant prostate carcinoma: a therapeutical target for patients at risk for bone metastases // *Int. J. Biol. Markers.* 2017. Vol. 32, № 2. P. e243–e247.
 151. Wang R. et al. Tumor-associated macrophages provide a suitable microenvironment for non-small lung cancer invasion and progression // *Lung Cancer.* 2011. Vol. 74, № 2. P. 188–196.
 152. Xie L. et al. Cathepsin K-upregulation in fibroblasts promotes matrigel invasive ability of squamous cell carcinoma cells via tumor-derived IL-1 α // *J. Dermatol. Sci.* 2011. Vol. 61, № 1. P. 45–50.
 153. Andrade S.S. et al. Cathepsin K induces platelet dysfunction and affects cell signaling in breast cancer - molecularly distinct behavior of cathepsin K in breast cancer // *BMC Cancer.* 2016. Vol. 16. P. 173.
 154. Lindeman J.H.N. et al. Cathepsin K is the principal protease in giant cell tumor of bone // *Am. J. Pathol.* 2004. Vol. 165, № 2. P. 593–600.
 155. Duong L.T. et al. Efficacy of a cathepsin K inhibitor in a preclinical model for prevention and treatment of breast cancer bone metastasis // *Mol. Cancer Ther.* 2014. Vol. 13, № 12. P. 2898–2909.
 156. Wang J. et al. Autocrine and paracrine STIP1 signaling promote osteolytic bone metastasis in renal cell carcinoma // *Oncotarget.* 2017. Vol. 8, № 10. P. 17012–17026.
 157. Tholen M. et al. Stress-resistant Translation of Cathepsin L mRNA in Breast Cancer Progression // *J. Biol. Chem.* 2015. Vol. 290, № 25. P. 15758–15769.
 158. Wang L. et al. K-ras mutation promotes ionizing radiation-induced invasion and migration of lung cancer in part via the Cathepsin L/CUX1 pathway // *Exp. Cell Res.* 2018. Vol. 362, № 2. P. 424–435.
 159. Yu S. et al. FOXO3a promotes gastric cancer cell migration and invasion through

- the induction of cathepsin L // *Oncotarget*. 2016. Vol. 7, № 23. P. 34773–34784.
160. Zhang L. et al. Cathepsin L is involved in proliferation and invasion of ovarian cancer cells // *Mol. Med. Rep.* 2015. Vol. 11, № 1. P. 468–474.
161. Tamhane T. et al. Nuclear cathepsin L activity is required for cell cycle progression of colorectal carcinoma cells // *Biochimie*. 2016. Vol. 122. P. 208–218.
162. Brindle N.R. et al. Deficiency for the Cysteine Protease Cathepsin L Impairs Myc-Induced Tumorigenesis in a Mouse Model of Pancreatic Neuroendocrine Cancer // *PLoS One* / ed. Gartel A.L. 2015. Vol. 10, № 4. P. e0120348.
163. Kenig S. et al. Glioblastoma and endothelial cells cross-talk, mediated by SDF-1, enhances tumour invasion and endothelial proliferation by increasing expression of cathepsins B, S, and MMP-9 // *Cancer Lett.* 2010. Vol. 289, № 1. P. 53–61.
164. Gautam J., Bae Y.K., Kim J.-A. Up-regulation of cathepsin S expression by HSP90 and 5-HT7 receptor-dependent serotonin signaling correlates with triple negativity of human breast cancer // *Breast Cancer Res. Treat.* 2017. Vol. 161, № 1. P. 29–40.
165. Yang Y. et al. Cathepsin S mediates gastric cancer cell migration and invasion via a putative network of metastasis-associated proteins // *J. Proteome Res.* 2010. Vol. 9, № 9. P. 4767–4778.
166. Fernandez P.L. et al. Expression of cathepsins B and S in the progression of prostate carcinoma // *Int. J. Cancer*. 2001. Vol. 95, № 1. P. 51–55.
167. Burden R.E. et al. Inhibition of Cathepsin S by Fsn0503 enhances the efficacy of chemotherapy in colorectal carcinomas // *Biochimie*. 2012. Vol. 94, № 2. P. 487–493.
168. Fan Q. et al. Silencing cathepsin S gene expression inhibits growth, invasion and angiogenesis of human hepatocellular carcinoma in vitro // *Biochem. Biophys. Res. Commun.* 2012. Vol. 425, № 4. P. 703–710.
169. Kwok H.F. et al. Antibody targeting of Cathepsin S induces antibody-dependent cellular cytotoxicity // *Mol. Cancer*. 2011. Vol. 10. P. 147.
170. Verdoes M. et al. A nonpeptidic cathepsin S activity-based probe for noninvasive

- optical imaging of tumor-associated macrophages // *Chem. Biol.* 2012. Vol. 19, № 5. P. 619–628.
171. Lindahl C. et al. Increased levels of macrophage-secreted cathepsin S during prostate cancer progression in TRAMP mice and patients // *Cancer Genomics Proteomics*. 2009. Vol. 6, № 3. P. 149–159.
172. Small D.M. et al. Cathepsin S from both tumor and tumor-associated cells promote cancer growth and neovascularization // *Int. J. Cancer*. 2013. Vol. 133, № 9. P. 2102–2112.
173. Yang M. et al. Cathepsin S-mediated autophagic flux in tumor-associated macrophages accelerate tumor development by promoting M2 polarization // *Mol. Cancer*. 2014. Vol. 13. P. 43.
174. Ryschich E. et al. Molecular fingerprinting and autocrine growth regulation of endothelial cells in a murine model of hepatocellular carcinoma // *Cancer Res*. 2006. Vol. 66, № 1. P. 198–211.
175. Tedelind S. et al. Nuclear cysteine cathepsin variants in thyroid carcinoma cells // *Biol. Chem.* 2010. Vol. 391, № 8. P. 923–935.
176. Teller A. et al. Dysregulation of apoptotic signaling pathways by interaction of RPLP0 and cathepsin X/Z in gastric cancer // *Pathol. - Res. Pract.* 2015. Vol. 211, № 1. P. 62–70.
177. Pečar Fonović U. et al. Profilin 1 as a Target for Cathepsin X Activity in Tumor Cells // *PLoS One* / ed. Coles J.A. 2013. Vol. 8, № 1. P. e53918.
178. Jechorek D. et al. Characterization of cathepsin X in colorectal cancer development and progression // *Pathol. - Res. Pract.* 2014. Vol. 210, № 12. P. 822–829.
179. Wang J. et al. Overexpression of Cathepsin Z Contributes to Tumor Metastasis by Inducing Epithelial-Mesenchymal Transition in Hepatocellular Carcinoma // *PLoS One* / ed. Hotchin N.A. 2011. Vol. 6, № 9. P. e24967.
180. Lines K.E. et al. S100P-Binding Protein, S100PBP, Mediates Adhesion through Regulation of Cathepsin Z in Pancreatic Cancer Cells // *Am. J. Pathol.* 2012. Vol. 180, № 4. P. 1485–1494.

181. Krueger S. et al. Up-regulation of cathepsin X in *Helicobacter pylori* gastritis and gastric cancer // J. Pathol. 2005. Vol. 207, № 1. P. 32–42.
182. Guinec N., Dalet-Fumeron V., Pagano M. “In vitro” study of basement membrane degradation by the cysteine proteinases, cathepsins B, B-like and L. Digestion of collagen IV, laminin, fibronectin, and release of gelatinase activities from basement membrane fibronectin // Biol. Chem. Hoppe. Seyler. 1993. Vol. 374, № 12. P. 1135–1146.
183. Chang S. et al. VEGF-A induces angiogenesis by perturbing the cathepsin-cysteine protease inhibitor balance in venules, causing basement membrane degradation & mother vessel formation // Cancer Res. 2009. Vol. 69, № 10. P. 4537–4544.
184. Jevnikar Z. et al. Cathepsin H mediates the processing of talin and regulates migration of prostate cancer cells // J. Biol. Chem. 2013. Vol. 288, № 4. P. 2201–2209.
185. Quintanilla-Dieck M.J. et al. Cathepsin K in melanoma invasion // J. Invest. Dermatol. 2008. Vol. 128, № 9. P. 2281–2288.
186. Sevenich L. et al. Analysis of tumor- and stroma-supplied proteolytic networks reveals a brain metastasis-promoting role for cathepsin S // Nat. Cell Biol. 2014. Vol. 16, № 9. P. 876–888.
187. Akkari L. et al. Distinct functions of macrophage-derived and cancer cell-derived cathepsin Z combine to promote tumor malignancy via interactions with the extracellular matrix // Genes Dev. 2014. Vol. 28, № 19. P. 2134–2150.
188. Abboud-Jarrous G. et al. Cathepsin L is responsible for processing and activation of proheparanase through multiple cleavages of a linker segment // J. Biol. Chem. 2008. Vol. 283, № 26. P. 18167–18176.
189. Wang B. et al. Cathepsin S controls angiogenesis and tumor growth via matrix-derived angiogenic factors // J. Biol. Chem. 2006. Vol. 281, № 9. P. 6020–6029.
190. Veillard F. et al. Cysteine cathepsins S and L modulate anti-angiogenic activities of human endostatin // J. Biol. Chem. 2011. Vol. 286, № 43. P. 37158–37167.
191. Parkes C., Kembhavi A.A., Barrett A.J. Calpain inhibition by peptide epoxides //

- Biochem. J. 1985. Vol. 230, № 2. P. 509–516.
192. Towatari T. et al. Novel epoxysuccinyl peptides A selective inhibitor of cathepsin B, in vivo // FEBS Lett. 1991. Vol. 280, № 2. P. 311–315.
193. Murata M. et al. Novel epoxysuccinyl peptides Selective inhibitors of cathepsin B, in vitro // FEBS Lett. 1991. Vol. 280, № 2. P. 307–310.
194. Brömme D. et al. Peptidyl vinyl sulphones: a new class of potent and selective cysteine protease inhibitors: S2P2 specificity of human cathepsin O2 in comparison with cathepsins S and L // Biochem. J. 1996. Vol. 315 (Pt 1, № Pt 1. P. 85–89.
195. Fournier J. et al. Total Synthesis of Covalent Cysteine Protease Inhibitor N - Desmethyl Thalassospiramide C and Crystallographic Evidence for Its Mode of Action // Org. Lett. 2019. Vol. 21, № 2. P. 508–512.
196. Jadhav P.K. et al. Discovery of Cathepsin S Inhibitor LY3000328 for the Treatment of Abdominal Aortic Aneurysm // ACS Med. Chem. Lett. 2014. Vol. 5, № 10. P. 1138–1142.
197. Greenspan P.D. et al. Identification of dipeptidyl nitriles as potent and selective inhibitors of cathepsin B through structure-based drug design // J. Med. Chem. 2001. Vol. 44, № 26. P. 4524–4534.
198. Balakireva A. V. et al. Trends and Prospects of Plant Proteases in Therapeutics // Curr. Med. Chem. 2019. Vol. 26, № 3. P. 465–486.
199. Simmons J.W., Nordby E.J., Hadjipavlou A.G. Chemonucleolysis: the state of the art // Eur. Spine J. 2001. Vol. 10, № 3. P. 192–202.
200. Grabowska E. et al. Bromelain proteases suppress growth, invasion and lung metastasis of B16F10 mouse melanoma cells // Int. J. Oncol. 1997. Vol. 11, № 2. P. 243–248.
201. Kelly G.S. Bromelain: A literature review and discussion of its therapeutic applications // Altern. Med. Rev. 1996. Vol. 1, № 4. P. 243–257.
202. Shivaprasad H. V. et al. 'Pergularain e I'--a plant cysteine protease with thrombin-like activity from Pergularia extensa latex // Thromb. Res. Elsevier Ltd, 2010. Vol. 125, № 3. P. e100–e105.

203. Doolittle R.F. Clotting of mammalian fibrinogens by papain: a re-examination // Biochemistry. 2014. Vol. 53, № 42. P. 6687–6694.
204. Maurer H.R. Bromelain: biochemistry, pharmacology and medical use // Cell. Mol. Life Sci. 2001. Vol. 58, № 9. P. 1234–1245.
205. Choi J.-H. et al. Thrombolytic, anticoagulant and antiplatelet activities of codiase, a bi-functional fibrinolytic enzyme from *Codium fragile* // Biochimie. 2013. Vol. 95, № 6. P. 1266–1277.
206. Matsubara K. et al. A fibrinolytic enzyme from a marine green alga, *Codium latum* // Phytochemistry. 1999. Vol. 52, № 6. P. 993–999.
207. Kim D.-W. et al. Direct acting anti-thrombotic serine protease from brown seaweed *Costaria costata* // Process Biochem. 2013. Vol. 48, № 2. P. 340–350.
208. Kim D.-W. et al. Purification and characterization of a fibrinolytic enzyme from *Petasites japonicus* // Int. J. Biol. Macromol. 2015. Vol. 72. P. 1159–1167.
209. Eshamah H. et al. Antibacterial effects of natural tenderizing enzymes on different strains of *Escherichia coli* O157:H7 and *Listeria monocytogenes* on beef // Meat Sci. 2014. Vol. 96, № 4. P. 1494–1500.
210. Osato J.A. et al. Antimicrobial and antioxidant activities of unripe papaya // Life Sci. 1993. Vol. 53, № 17. P. 1383–1389.
211. Guevara M. et al. Potato aspartic proteases: Induction, antimicrobial activity and substrate specificity // J. Plant Pathol. 2004. Vol. 86. P. 233–238.
212. Mendieta J.R. et al. Antimicrobial activity of potato aspartic proteases (StAPs) involves membrane permeabilization // Microbiology. 2006. Vol. 152, № Pt 7. P. 2039–2047.
213. Niderman T. et al. Pathogenesis-related PR-1 proteins are antifungal. Isolation and characterization of three 14-kilodalton proteins of tomato and of a basic PR-1 of tobacco with inhibitory activity against *Phytophthora infestans* // Plant Physiol. 1995. Vol. 108, № 1. P. 17–27.
214. Pagano M.R. et al. Roles of glycosylation on the antifungal activity and apoplast accumulation of StAPs (*Solanum tuberosum* aspartic proteases) // Int. J. Biol. Macromol. 2007. Vol. 41, № 5. P. 512–520.

215. Berger J., Asenjo C.F. ANTHELMINTIC ACTIVITY OF CRYSTALLINE PAPAIN // Science. 1940. Vol. 91, № 2364. P. 387–388.
216. Stepek G. et al. Assessment of the anthelmintic effect of natural plant cysteine proteinases against the gastrointestinal nematode, *Heligmosomoides polygyrus*, in vitro // Parasitology. 2005. Vol. 130, № Pt 2. P. 203–211.
217. Chen C.F. et al. Protective effects of *Carica papaya* Linn on the exogenous gastric ulcer in rats // Am. J. Chin. Med. 1981. Vol. 9, № 3. P. 205–212.
218. Mello V.J. et al. The gastric ulcer protective and healing role of cysteine proteinases from *Carica candamarcensis* // Phytomedicine. 2008. Vol. 15, № 4. P. 237–244.
219. Dittz D. et al. Antiangiogenesis, loss of cell adhesion and apoptosis are involved in the antitumoral activity of Proteases from *V. cundinamarcensis* (*C. candamarcensis*) in murine melanoma B16F1 // Int. J. Mol. Sci. 2015. Vol. 16, № 4. P. 7027–7044.
220. Kane S., Goldberg M.J. Use of bromelain for mild ulcerative colitis // Ann. Intern. Med. 2000. Vol. 132, № 8. P. 680.
221. Kurada S., Yadav A., Leffler D.A. Current and novel therapeutic strategies in celiac disease // Expert Rev. Clin. Pharmacol. 2016. Vol. 9, № 9. P. 1211–1223.
222. Balakireva A. V, Zamyatnin A.A. Properties of Gluten Intolerance: Gluten Structure, Evolution, Pathogenicity and Detoxification Capabilities // Nutrients. 2016. Vol. 8, № 10. P. 644.
223. Shan L. et al. Structural basis for gluten intolerance in celiac sprue // Science. 2002. Vol. 297, № 5590. P. 2275–2279.
224. Savvateeva L. V et al. Overview of Celiac Disease in Russia: Regional Data and Estimated Prevalence // J. Immunol. Res. 2017. Vol. 2017, № 2. P. 2314813.
225. Savvateeva L. V, Zamyatnin A.A. Prospects of Developing Medicinal Therapeutic Strategies and Pharmaceutical Design for Effective Gluten Intolerance Treatment // Curr. Pharm. Des. 2016. Vol. 22, № 16. P. 2439–2449.
226. Wieser H., Koehler P. The Biochemical Basis of Celiac Disease // Cereal Chem. 2008. Vol. 85, № 1. P. 1–13.

227. Wieser H., Koehler P. Detoxification of gluten by means of enzymatic treatment // J. AOAC Int. 2012. Vol. 95, № 2. P. 356–363.
228. Lähdeaho M.-L. et al. Glutenase ALV003 attenuates gluten-induced mucosal injury in patients with celiac disease // Gastroenterology. 2014. Vol. 146, № 7. P. 1649–1658.
229. Murray J.A. et al. Latiglutenase Protects the Mucosa and Attenuates Symptom Severity in Patients With Celiac Disease Exposed to a Gluten Challenge // Gastroenterology. 2022. Vol. 163, № 6. P. 1510-1521.e6.
230. Gorokhovets N. V. et al. Rational Design of Recombinant Papain-Like Cysteine Protease: Optimal Domain Structure and Expression Conditions for Wheat-Derived Enzyme Triticain- α // Int. J. Mol. Sci. 2017. Vol. 18, № 7. P. 1395.
231. Kiyosaki T. et al. Wheat cysteine proteases triticain alpha, beta and gamma exhibit mutually distinct responses to gibberellin in germinating seeds // J. Plant Physiol. 2009. Vol. 166, № 1. P. 101–106.
232. Kabsch W. XDS // Acta Crystallogr. D. Biol. Crystallogr. 2010. Vol. 66, № Pt 2. P. 125–132.
233. Evans P. Scaling and assessment of data quality // Acta Crystallogr. D. Biol. Crystallogr. 2006. Vol. 62, № Pt 1. P. 72–82.
234. Evans P.R., Murshudov G.N. How good are my data and what is the resolution? // Acta Crystallogr. D. Biol. Crystallogr. 2013. Vol. 69, № Pt 7. P. 1204–1214.
235. Winn M.D. et al. Overview of the CCP4 suite and current developments // Acta Crystallogr. D. Biol. Crystallogr. 2011. Vol. 67, № Pt 4. P. 235–242.
236. Tickle I.J. et al. STARANISO. Cambridge, United Kingdom: Global Phasing Ltd., 2019.
237. Vonrhein C. et al. Data processing and analysis with the autoPROC toolbox // Acta Crystallogr. D. Biol. Crystallogr. 2011. Vol. 67, № Pt 4. P. 293–302.
238. Liebschner D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix // Acta Crystallogr. Sect. D, Struct. Biol. 2019. Vol. 75, № Pt 10. P. 861–877.
239. Williams C.J. et al. MolProbity: More and better reference data for improved all-

- atom structure validation // Protein Sci. 2018. Vol. 27, № 1. P. 293–315.
240. Fox N.K., Brenner S.E., Chandonia J.-M. SCOPe: Structural Classification of Proteins--extended, integrating SCOP and ASTRAL data and classification of new structures // Nucleic Acids Res. 2014. Vol. 42, № Database issue. P. D304–D309.
241. Jumper J. et al. Highly accurate protein structure prediction with AlphaFold // Nature. 2021. Vol. 596, № 7873. P. 583–589.
242. Mirdita M. et al. ColabFold: making protein folding accessible to all // Nat. Methods. 2022. Vol. 19, № 6. P. 679–682.
243. Mirdita M., Steinegger M., Söding J. MMseqs2 desktop and local web server app for fast, interactive sequence searches // Bioinformatics. 2019. Vol. 35, № 16. P. 2856–2858.
244. Eastman P. et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics // PLoS Comput. Biol. 2017. Vol. 13, № 7. P. e1005659.
245. Case D.A. et al. AmberTools // J. Chem. Inf. Model. 2023. Vol. 63, № 20. P. 6183–6191.
246. Berman H., Henrick K., Nakamura H. Announcing the worldwide Protein Data Bank // Nat. Struct. Biol. 2003. Vol. 10, № 12. P. 980.
247. Berman H. et al. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data // Nucleic Acids Res. 2007. Vol. 35, № Database issue. P. D301-303.
248. wwPDB consortium. Protein Data Bank: the single global archive for 3D macromolecular structure data // Nucleic Acids Res. 2019. Vol. 47, № D1. P. D520–D528.
249. Bittrich S. et al. RCSB protein Data Bank: exploring protein 3D similarities via comprehensive structural alignments // Bioinformatics. 2024. Vol. 40, № 6. P. btae370.
250. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023 // Nucleic Acids Res. 2023. Vol. 51, № D1. P. D523–D531.
251. Varadi M. et al. AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences // Nucleic Acids Res.

2024. Vol. 52, № D1. P. D368–D375.
252. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets // *Bioinformatics*. 2014. Vol. 30, № 22. P. 3276–3278.
253. Хрущев С.С. et al. Моделирование белок-белковых взаимодействий с применением программного комплекса многочастичной броуновской динамики ProKSim // *Компьютерные исследования и моделирование*. 2013. Vol. 5, № 1. P. 47–64.
254. Li H., Robertson A.D., Jensen J.H. Very fast empirical prediction and rationalization of protein pKa values // *Proteins*. 2005. Vol. 61, № 4. P. 704–721.
255. Ke S.H., Madison E.L. Rapid and efficient site-directed mutagenesis by single-tube “megaprimer” PCR method // *Nucleic Acids Res.* 1997. Vol. 25, № 16. P. 3371–3372.
256. Li Z. et al. FATCAT 2.0: towards a better understanding of the structural diversity of proteins // *Nucleic Acids Res.* 2020. Vol. 48, № W1. P. W60–W64.
257. Alpey M.S., Hunter W.N. High-resolution complex of papain with remnants of a cysteine protease inhibitor derived from *Trypanosoma brucei* // *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* 2006. Vol. 62, № Pt 6. P. 504–508.
258. Azarkan M. et al. Structures of the free and inhibitors-bound forms of bromelain and ananain from *Ananas comosus* stem and in vitro study of their cytotoxicity // *Sci. Rep.* 2020. Vol. 10, № 1. P. 19570.
259. Baker E.N., Dodson E.J. Crystallographic refinement of the structure of actinidin at 1.7 Å resolution by fast Fourier least-squares methods // *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.* 1980. Vol. 36, № 4. P. 559–572.
260. Choi K.H., Laursen R.A., Allen K.N. The 2.1 Å structure of a cysteine protease with proline specificity from ginger rhizome, *Zingiber officinale* // *Biochemistry*. 1999. Vol. 38, № 36. P. 11624–11633.
261. Bethune M.T. et al. Heterologous expression, purification, refolding, and structural-functional characterization of EP-B2, a self-activating barley cysteine endoprotease // *Chem. Biol.* 2006. Vol. 13, № 6. P. 637–647.

262. Hills A.G. pH and the Henderson-Hasselbalch equation // *Am. J. Med.* 1973. Vol. 55, № 2. P. 131–133.
263. Fogolari F., Brigo A., Molinari H. The Poisson-Boltzmann equation for biomolecular electrostatics: a tool for structural biology // *J. Mol. Recognit.* 2002. Vol. 15, № 6. P. 377–392.
264. Baneyx F., Mujacic M. Recombinant protein folding and misfolding in *Escherichia coli* // *Nat. Biotechnol.* 2004. Vol. 22, № 11. P. 1399–1408.
265. Paireder M. et al. The death enzyme CP14 is a unique papain-like cysteine proteinase with a pronounced S2 subsite selectivity // *Arch. Biochem. Biophys.* 2016. Vol. 603. P. 110–117.
266. Paireder M. et al. The papain-like cysteine proteinases NbCysP6 and NbCysP7 are highly processive enzymes with substrate specificities complementary to *Nicotiana benthamiana* cathepsin B // *Biochim. Biophys. acta. Proteins proteomics.* 2017. Vol. 1865, № 4. P. 444–452.
267. Yamamoto Y. et al. Proregion of *Bombyx mori* cysteine proteinase functions as an intramolecular chaperone to promote proper folding of the mature enzyme // *Arch. Insect Biochem. Physiol.* 1999. Vol. 42, № 3. P. 167–178.
268. Correa A., Oppezzo P. Overcoming the solubility problem in *E. coli*: available approaches for recombinant protein production // *Methods Mol. Biol.* 2015. Vol. 1258. P. 27–44.
269. Novinec M., Pavšič M., Lenarčič B. A simple and efficient protocol for the production of recombinant cathepsin V and other cysteine cathepsins in soluble form in *Escherichia coli* // *Protein Expr. Purif.* 2012. Vol. 82, № 1. P. 1–5.
270. Яковлев А.А. et al. Катепсин мозга способен расщеплять субстрат каспазы // *Биохимия.* 2008. Vol. 73, № 3. P. 408–413.
271. Bozhkov P. V et al. Cysteine protease mclI-Pa executes programmed cell death during plant embryogenesis // *Proc. Natl. Acad. Sci. U. S. A.* 2005. Vol. 102, № 40. P. 14463–14468.
272. Harris T.K., Mildvan A.S. High-precision measurement of hydrogen bond lengths in proteins by nuclear magnetic resonance methods // *Proteins Struct. Funct.*

- Genet. 1999. Vol. 35, № 3. P. 275–282.
273. Dutta S., Dattagupta J.K., Biswas S. Enhancement of proteolytic activity of a thermostable papain-like protease by structure-based rational design // PLoS One. 2013. Vol. 8, № 5. P. e62619.
274. Ghosh R., Dattagupta J.K., Biswas S. A thermostable cysteine protease precursor from a tropical plant contains an unusual C-terminal propeptide: cDNA cloning, sequence comparison and molecular modeling studies // Biochem. Biophys. Res. Commun. 2007. Vol. 362, № 4. P. 965–970.
275. Marquis R.W. et al. Azepanone-based inhibitors of human and rat cathepsin K // J. Med. Chem. 2001. Vol. 44, № 9. P. 1380–1395.
276. Tada S. et al. Species differences between human and rat in the substrate specificity of cathepsin K // J. Biochem. 2008. Vol. 144, № 4. P. 499–506.
277. Khouri H.E. et al. Engineering of papain: selective alteration of substrate specificity by site-directed mutagenesis // Biochemistry. 1991. Vol. 30, № 37. P. 8929–8936.
278. Craik C.S., Page M.J., Madison E.L. Proteases as therapeutics // Biochem. J. 2011. Vol. 435, № 1. P. 1–16.
279. Brömme D., Nallaseth F.S., Turk B. Production and activation of recombinant papain-like cysteine proteases // Methods. 2004. Vol. 32, № 2. P. 199–206.
280. Wiederanders B., Kaulmann G., Schilling K. Functions of propeptide parts in cysteine proteases // Curr. Protein Pept. Sci. 2003. Vol. 4, № 5. P. 309–326.
281. Ogino T. et al. Function of the propeptide region in recombinant expression of active procathepsin L in Escherichia coli // J. Biochem. 1999. Vol. 126, № 1. P. 78–83.
282. Tobbell D.A. et al. Identification of in vitro folding conditions for procathepsin S and cathepsin S using fractional factorial screens // Protein Expr. Purif. 2002. Vol. 24, № 2. P. 242–254.
283. Blanco A., Blanco G. Enzymes // Medical Biochemistry. Elsevier, 2017. P. 153–175.
284. Than M.E. et al. The 2.0 Å crystal structure and substrate specificity of the

- KDEL-tailed cysteine endopeptidase functioning in programmed cell death of *Ricinus communis* endosperm // *J. Mol. Biol.* 2004. Vol. 336, № 5. P. 1103–1116.
285. Groves M.R. et al. The prosequence of procaricain forms an alpha-helical domain that prevents access to the substrate-binding cleft // *Structure*. 1996. Vol. 4, № 10. P. 1193–1203.
 286. Tušar L. et al. Proteomic data and structure analysis combined reveal interplay of structural rigidity and flexibility on selectivity of cysteine cathepsins // *Commun. Biol.* 2023. Vol. 6, № 1. P. 450.
 287. Santos V.C. et al. The gene repertoire of the main cysteine protease of *Trypanosoma cruzi*, cruzipain, reveals four sub-types with distinct active sites // *Sci. Rep.* 2021. Vol. 11, № 1. P. 18231.
 288. dos Reis F.C.G. et al. The substrate specificity of cruzipain 2, a cysteine protease isoform from *Trypanosoma cruzi* // *FEMS Microbiol. Lett.* 2006. Vol. 259, № 2. P. 215–220.
 289. Mølgaard A. et al. The crystal structure of human dipeptidyl peptidase I (cathepsin C) in complex with the inhibitor Gly-Phe-CHN₂ // *Biochem. J.* 2007. Vol. 401, № 3. P. 645–650.
 290. Xie N.-Z. et al. Exploring Strong Interactions in Proteins with Quantum Chemistry and Examples of Their Applications in Drug Design // *PLoS One*. 2015. Vol. 10, № 9. P. e0137113.
 291. Lunkad R. et al. Role of pK_A in Charge Regulation and Conformation of Various Peptide Sequences // *Polymers (Basel)*. 2021. Vol. 13, № 2. P. 214.
 292. Aich P., Biswas S. Highly Conserved Arg Residue of ERFNIN Motif of Pro-Domain is Important for pH-Induced Zymogen Activation Process in Cysteine Cathepsins K and L // *Cell Biochem. Biophys.* 2018. Vol. 76, № 1–2. P. 219–229.
 293. Pritsch I.C. et al. Regulation of the *Fasciola hepatica* newly excysted juvenile cathepsin L3 (FhCL3) by its propeptide: a proposed “clamp-like” mechanism of binding and inhibition // *BMC Mol. cell Biol.* 2020. Vol. 21. P. 90.
 294. A one-letter notation for amino acid sequences // *Pure Appl. Chem.* 1972. Vol. 31, № 4. P. 641–645.

ПРИЛОЖЕНИЕ А.

Структуры предсказанных комплексов пептидов с тритикаином- α

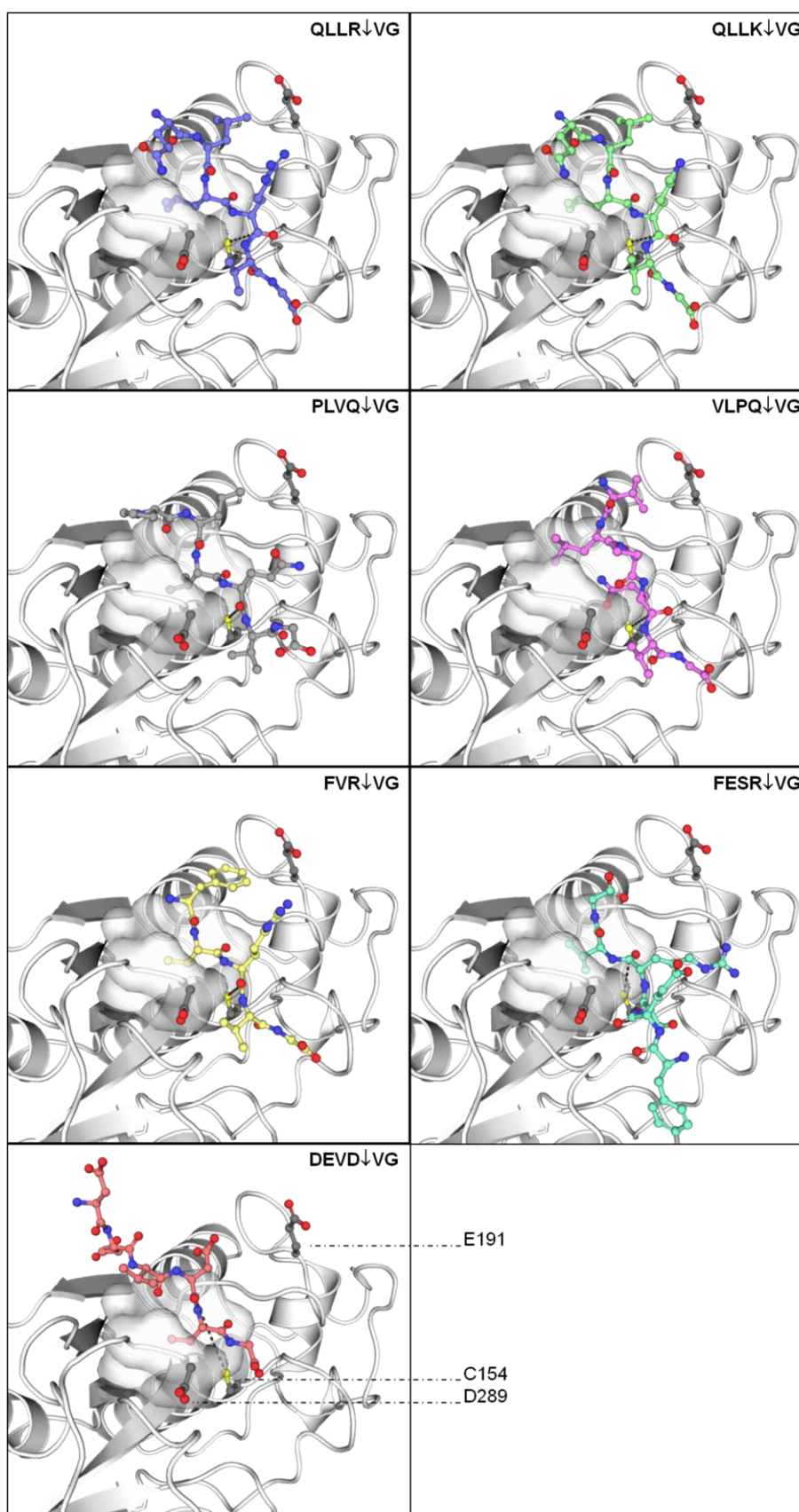


Рисунок А.1 – Лучшие модели комплексов тритикаина- α с пептидами QLRLVG, QLLKVG, PLVQVG, FVRVG, VLPQVG, DEVDBG и FESRVG согласно ранжированию в программе AlphaFold. Гидрофобный карман в S₂-сайте представлен в виде поверхности. Каталитический Cys154, Glu191 и Asp289 в сайте связывания S₁ фермента выделены и подписаны. Пунктирной линией указано расстояние между карбонильным атомом углерода а.о. в Р₁-положении пептида и атомом серы Cys154 фермента (Рисунок 20А).

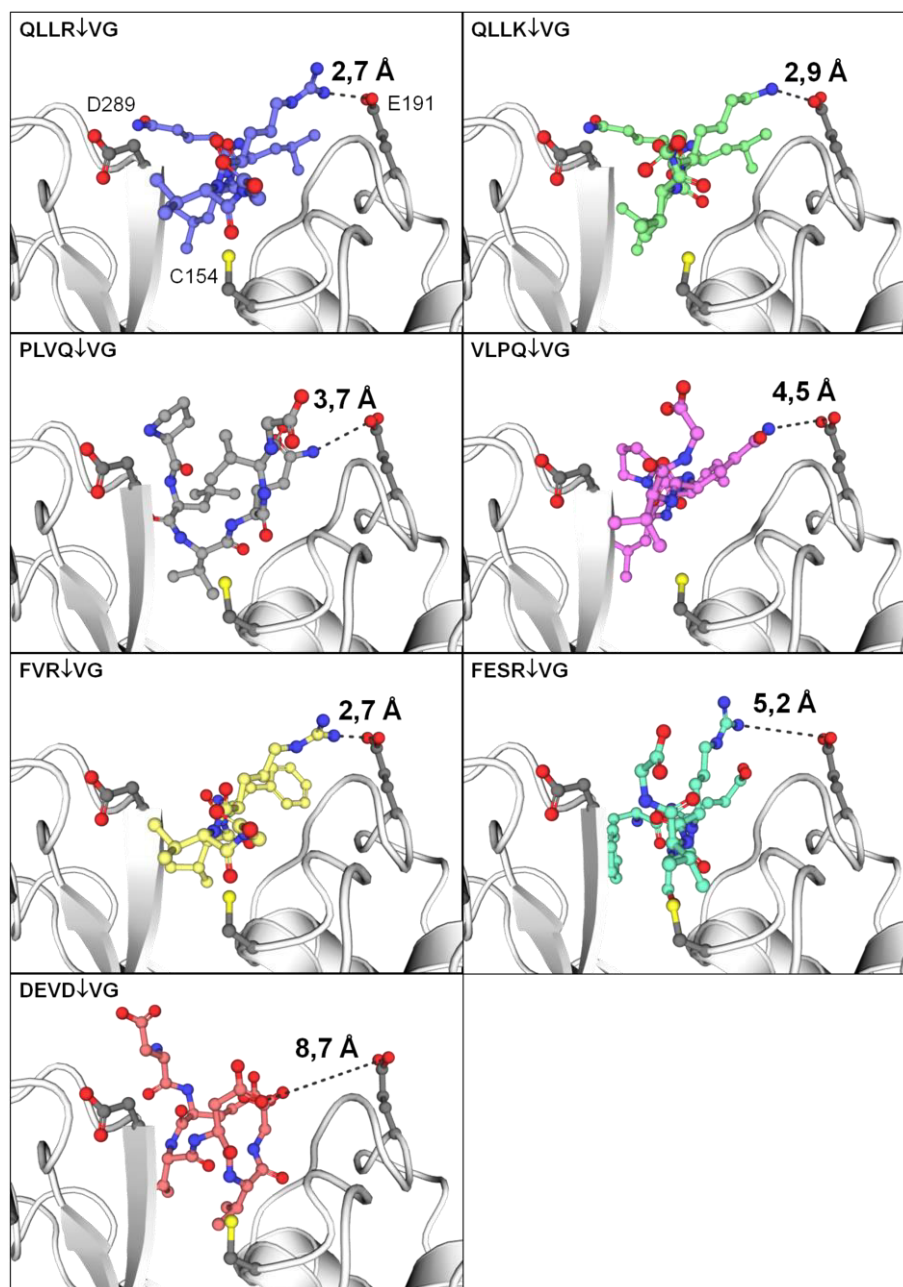


Рисунок A.2 – Модели комплексов трипсина-α с пептидами QLLRVG, QLLKVG, PLVQVG, FVRVG, VLPQVG, DEVDVG и FESRVG с наименьшим расстоянием между боковыми группами а.о. в P₁ и Glu191 среди структур, в которых расстояние от серы Cys154 до атакуемого атома углерода не превышает 5 Å. Каталитический Cys154, Glu191 и Asp289 в сайте связывания S₁ фермента выделены и подписаны. Пунктирной линией указано расстояние между боковыми группами а.о. в P₁-положении пептида и Glu191 фермента (Рисунок 21А).

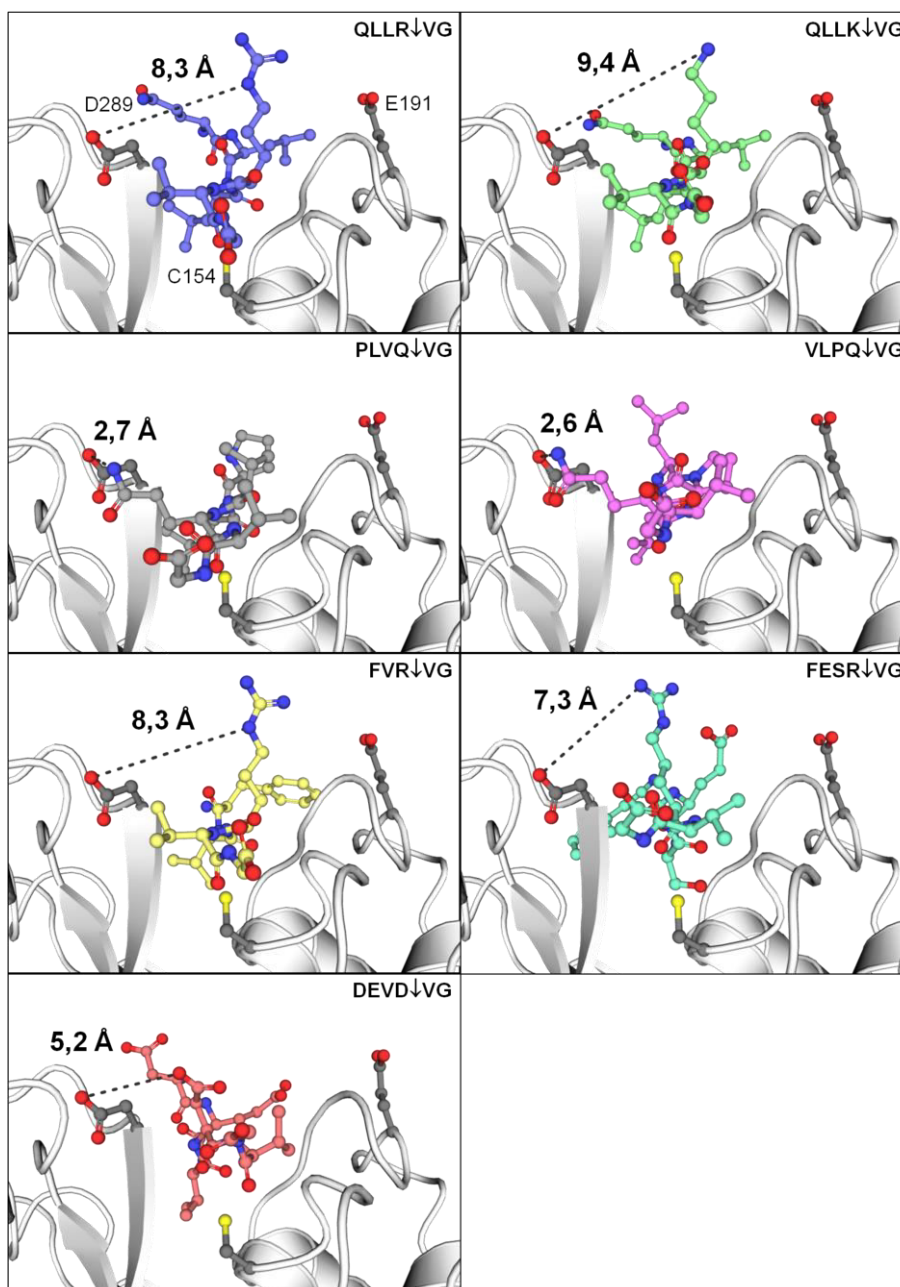


Рисунок А.3 – Модели комплексов тритикаина-α с пептидами QLLRVG, QLLKVG, PLVQVG, FVRVG, VLPQVG, DEVDVG и FESRVG с наименьшим расстоянием между боковыми группами а.о. в P₁ и Asp289 среди структур, в которых расстояние от серы Cys154 до атакуемого атома углерода не превышает 5 Å. Каталитический Cys154, Glu191 и Asp289 в сайте связывания S₁ фермента выделены и подписаны. Пунктирной линией указано расстояние между боковыми группами а.о. в P₁-положении пептида и Asp289 фермента (Рисунок 22А).

ПРИЛОЖЕНИЕ Б.

Множественное выравнивание аминокислотных последовательностей каталитических доменов ППЦП

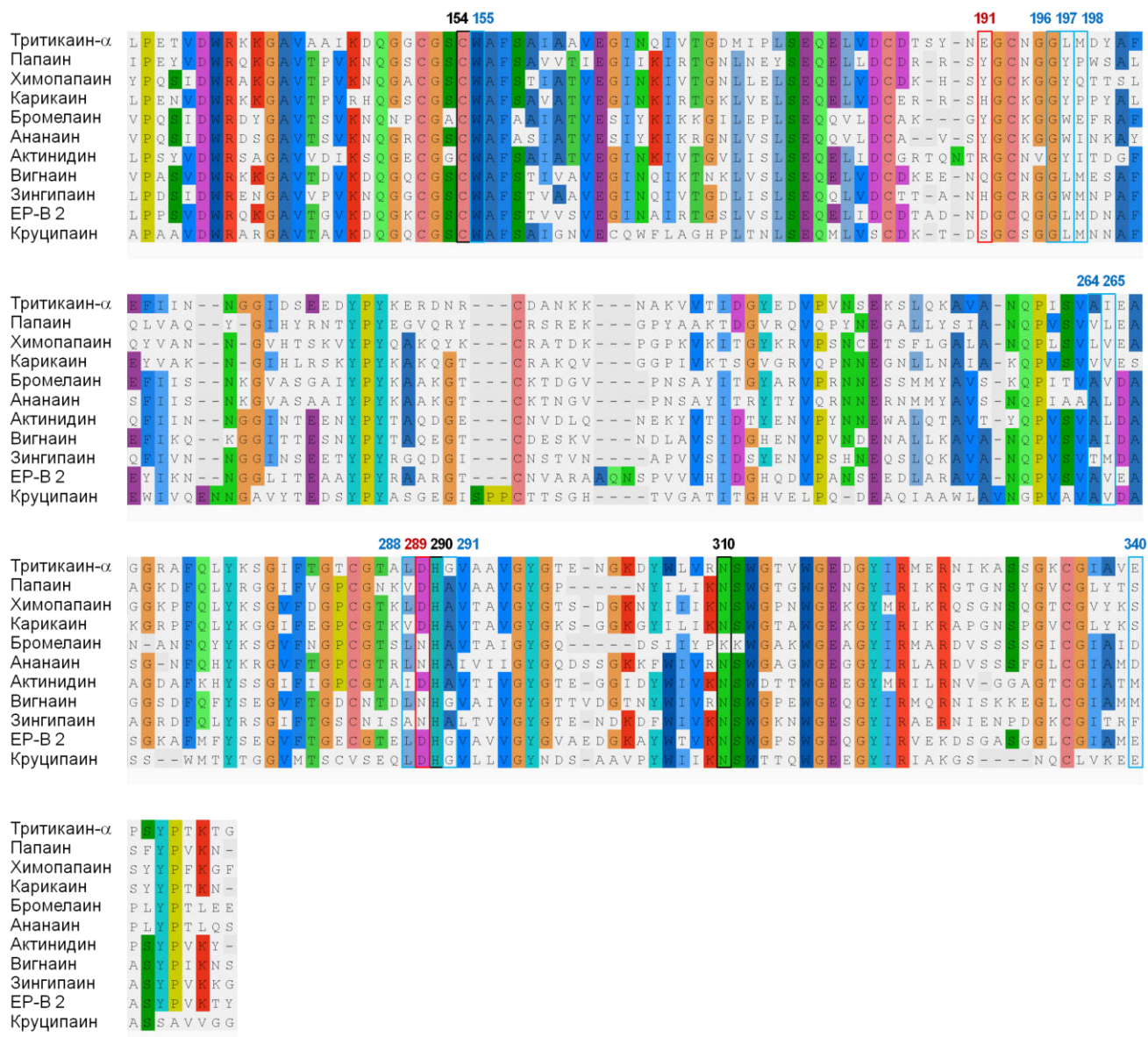


Рисунок Б.1 – Выравнивание аминокислотных последовательностей каталитических доменов тритикаина-α, других растительных ППЦП и круципаина. Синими рамками указаны а.о., которые формируют S₂-сайт связывания (Рисунок 24). Красными рамками выделены а.о., которые фланкируют S₁-сайт связывания (Рисунок 25). Черными рамками выделены а.о., которые составляют каталитическую триаду ППЦП. А.о. пронумерованы согласно их порядку в тритикаине-α. Цветным фоном отмечены консервативные а.о. среди сравниваемых последовательностей; у каждого а.о. свой цвет фона.

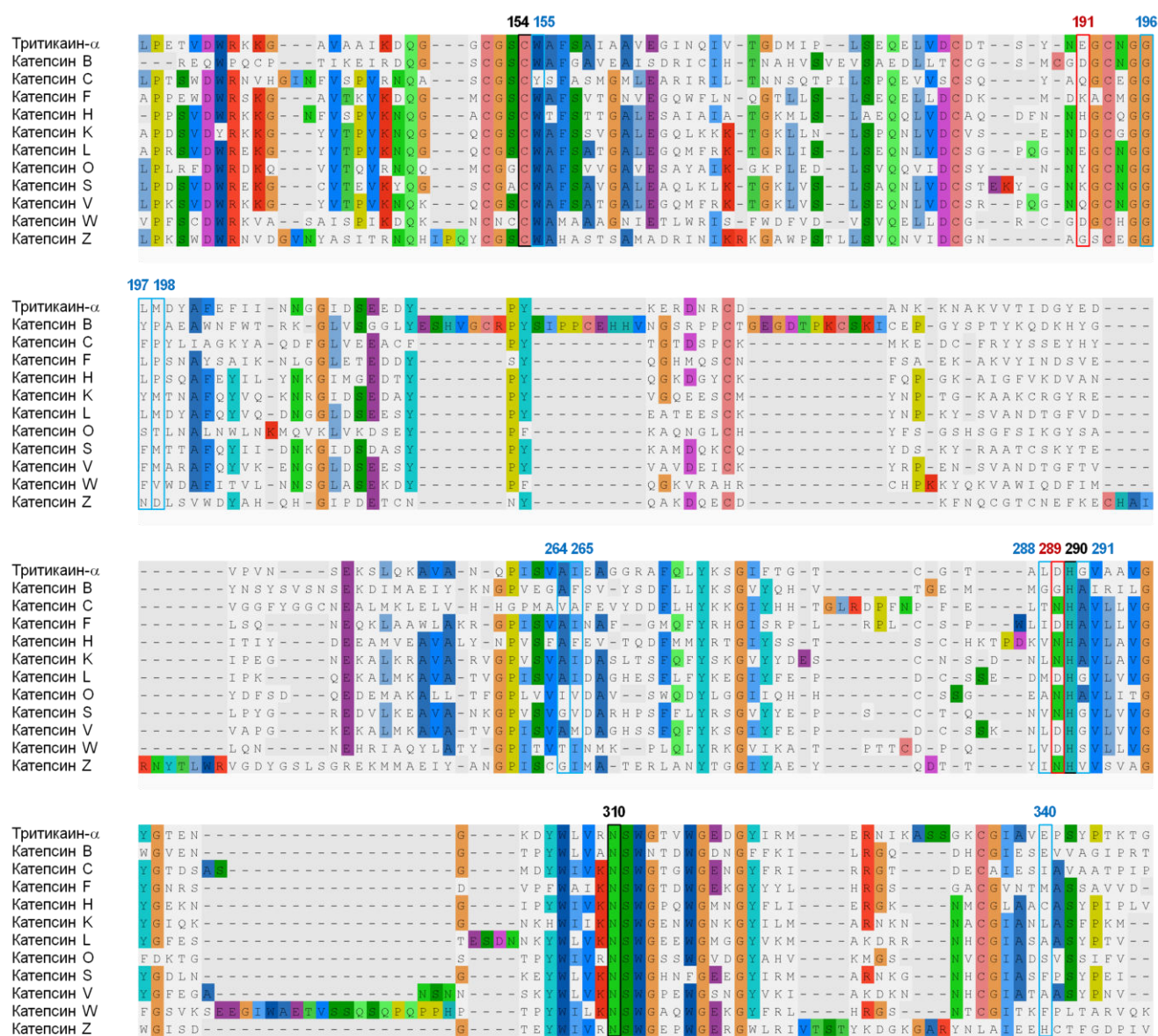


Рисунок Б.2 – Выравнивание аминокислотных последовательностей каталитических доменов тритикаина-α и цистеиновых катеписинов человека. Синими рамками указаны а.о., которые формируют S₂-сайт связывания (Рисунок 24). Красными рамками выделены а.о., которые фланкируют S₁-сайт связывания (Рисунок 25). Черными рамками выделены а.о., которые составляют каталитическую триаду ППЦП. А.о. пронумерованы согласно их порядку в тритикаине-α. Цветным фоном отмечены консервативные а.о. среди сравниваемых последовательностей; у каждого а.о. свой цвет фона.

ПРИЛОЖЕНИЕ В.

Достоверность различий в ферментативной активности тритикаина-α и его мутантов

Таблица В.1 – р-Значения для сопоставления обсуждаемых в Разделе 3.6 величин максимальной скорости и констант Михаэлиса из представленных на Рисунках 27 и 28. Значения сгруппированы по строкам для сравнения между собой. WT – тритикаин-α дикого типа, E191N – тритикаин-α с заменой Glu191Asn, E191K – тритикаин-α с заменой Glu191Lys. Для проведения расчётов константа Михаэлиса реакций, в которых не происходил гидролиз, была принята равной нулю. Статистическая значимость рассчитывалась согласно t-критерию Стьюдента для независимых выборок (* $p \leq 0,0001$; ** $p < 0,001$; *** $p < 0,01$; **** $p < 0,05$). Красным цветом выделены значения $p > 0,05$.

Максимальная скорость							
WT, QLLR, 3,6 и 6,5	0,0007 **	WT, QLLK, 3,6 и 6,5	0,0016 ***	WT, QLLQ, 3,6 и 6,5	0,0001 *	WT, QLLD, 3,6 и 6,5	0,0001 *
WT, 3,6, QLLR и QLLD	0,0027 ***	WT, 3,6, QLLK и QLLD	0,0407 ****	WT, 3,6, QLLQ и QLLD	0,0052 ***		
WT, 6,5, QLLR и QLLD	0,0001 *	WT, 6,5, QLLK и QLLD	0,0002 **	WT, 6,5, QLLQ и QLLD	0,0001 *		
WT, QLLD, 4,6 и 6,5	0,0001 *	E191N, QLLD, 4,6 и 6,5	0,0014 ***	E191K, QLLD, 4,6 и 6,5	0,0001 *		
Константа Михаэлиса							
QLLD, 6,5, E191N и E191K	0,1246						
WT, QLLD, 5,6 и 6,5	0,0038 ***	E191N, QLLD, 5,6 и 6,5	0,1072	E191K, QLLD, 5,6 и 6,5	0,6131		
WT, QLLD, 4,6 и 6,5	0,0001 *	E191N, QLLD, 4,6 и 6,5	0,0125 ****	E191K, QLLD, 4,6 и 6,5	0,0287 ****		