

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М.В. ЛОМОНОСОВА  
Химический факультет

На правах рукописи

**Васильева Ирина Дмитриевна**

**Метод масс-спектрометрического установления первичной структуры  
интактных пептидов амфибий семейства *Ranidae*  
Специальность 1.4.3 – «Органическая химия»**

Диссертация  
на соискание учёной степени  
кандидата химических наук

Научный руководитель:  
доктор химических наук  
Лебедев Альберт Тарасович

Москва – 2022

## ОГЛАВЛЕНИЕ

Введение.....	3
1. Обзор литературы.....	3
1.1. Преимущества масс-спектрометрии с преобразованием Фурье.....	14
1.2. Приборы тандемной масс-спектрометрии.....	16
1.3. Масс-спектрометрия с преобразованием Фурье для <i>de novo</i> секвенирования нетриптических пептидов.....	22
1.3.1. Изомерные аминокислоты.....	22
1.3.2. Масс-спектрометрия с преобразованием Фурье для секвенирования коротких пептидов.....	26
1.3.3. Пептиды с внутримолекулярной дисульфидной связью.....	31
1.4. Комплексные подходы масс-спектрометрии с преобразованием Фурье для секвенирования <i>de novo</i> природных пептидов.....	43
1.5. Методы визуализации и таксономических исследований.....	47
1.6. Программное обеспечение для <i>de novo</i> секвенирования в протеомике.....	49
1.6.1. Поиск по базам данных.....	50
1.6.2. <i>De novo</i> секвенирование.....	50
1.7. Режимы сбора экспериментальных масс-спектрометрических данных.....	55
2. Обсуждение результатов.....	58
2.1. Кожный пептидом остромордой лягушки <i>Rana arvalis</i> (Центральная Словения)....	58
2.2. Сравнительное изучение пептидомов московской и центральнословенской популяций <i>Rana arvalis</i> .....	70
2.3. Кожный пептидом травяной лягушки <i>Rana temporaria</i> (Центральная Словения) ...	73
2.4. Кожный пептидом травяной лягушки <i>Rana temporaria</i> (Архангельская область) ...	82
2.5. Популяционная идентификация особей <i>R. temporaria</i> на основе масс-спектрометрического изучения их кожных пептидомов.....	90
2.6. «Rana box» – раскрытие дисульфидного цикла в интактных пептидах.....	95
2.7. Сравнение результатов ручного и автоматического секвенирования нетриптических пептидов.....	104
3. Экспериментальная часть.....	108
Заключение.....	113
Список сокращений.....	114
Общепринятые обозначения аминокислот.....	116
Список литературы.....	118

## ВВЕДЕНИЕ

**Актуальность темы.** Появившиеся в 80-х годах XX века мягкие методы ионизации (электроспрей и МАЛДИ) позволили изучать строение макромолекул при помощи масс-спектрометрии, что привело к появлению протеомики – постгеномной науки, посвящённой анализу белков и пептидов, в том числе и изучению их функций. Особый вклад в развитие мягких методов ионизации был отмечен Нобелевской премией по химии в 2002 г. Масс-спектрометрия позволила устанавливать первичную структуру белка, то есть секвенировать его, при помощи различных методов фрагментации. Развитие масс-спектрометрических инструментов, особенно появление приборов высокого разрешения (ИЦР ПФ и Orbitrap), практически полностью исключило из использования классический метод секвенирования – деградацию по Эдману. В настоящее время существуют различные способы масс-спектрометрического секвенирования компонентов белковых/пептидных смесей, самый распространённый – «снизу вверх», с применением трипсинолиза и последующим секвенированием коротких пептидов с конструированием на их основе последовательностей исходных белков. Противоположный ему способ – «сверху вниз», где пептиды секвенируют в интактных смесях исключительно методами самой масс-спектрометрии, прибегая порой к небольшим химическим модификациям (ацетилирование *N*-концевых аминогрупп, дериватизация дисульфидных связей). К сожалению, эти модификации далеко не всегда приводят к успешному секвенированию. Они трудоёмки, времязатратны, а более того, они ведут к неизбежным потерям минорных компонентов исследуемых смесей. Понятие *de novo* секвенирование применимо к любому из подходов и говорит о том, что последовательность анализируемой молекулы установлена исключительно из спектральной информации.

Изучение процессов, протекающих в результате химических реакций пептидов в источнике ионизации и ячейке соударений масс-спектрометра, привлекает внимание учёных, заставляя искать новые и актуальные объекты. Отдельный интерес представляют собой пептиды, выделяемые кожными железами амфибий при стрессе. Они обладают разнообразными активностями: антимикробными, антидиабетическими, противоопухолевыми, фунгицидными и т.д., что может сделать их моделями для фармпрепаратов нового поколения. Эти пептиды являют собой сложные объекты для анализа из-за разнообразных посттрансляционных модификаций (ПТМ), обеспечивающих разнообразие молекул и их функций. Одна из самых распространённых ПТМ кожных пептидов амфибий – внутримолекулярная дисульфидная связь, образованная боковыми цепями двух цистеинов и исключаящая фрагментацию внутри образующегося S-S цикла. В подавляющем большинстве протеомных работ для установления последовательностей аминокислот внутри таких циклов прибегают к дериватизации – химической модификации

с целью раскрытия дисульфидной связи, что делает доступными для секвенирования все аминокислоты внутри цикла. Однако использование современных приборов требует очистки анализируемых смесей от реагентов, что зачастую искажает состав. Применение метода «сверху вниз» к интактным образцам без химической модификации, исключительно с помощью доступных на масс-спектрометрах методов фрагментации, с возможностью установить полную аминокислотную последовательность – цель, к которой стремятся многие учёные.

Создание метода секвенирования пептидов «сверху вниз», лишённого всех перечисленных недостатков, является актуальной научной проблемой. В настоящей работе разработан и опробован такой метод для *de novo* секвенирования компонентов кожных секретов амфибий с установлением полных последовательностей интерпретацией спектров, полученных комплексом методов тандемной масс-спектрометрии высокого разрешения без использования химических модификаций.

**Степень разработанности темы.** Секвенирование пептидов – задача трудоёмкая. Существующие автоматические алгоритмы секвенирования, в основном, ориентированы на триптические пептиды (подход «снизу вверх»), имеющие C-концевой лизин/аргинин, дающий выраженную серию у-ионов во фрагментации пептидных связей, активированной соударениями (ДАС). Этот подход предполагает работу с базами белковыми данных для восстановления сиквенсов исходных белков. Если такой белок/пептид отсутствует в базе данных из-за отсутствия геномных данных по конкретному виду амфибий, то для решения задач данной работы метод неприменим. Геномы большинства видов амфибий неизвестны, что делает ручную интерпретацию спектров единственным способом *de novo* секвенирования неизученных видов. Многие группы исследователей практикуют клонирование кДНК, выделенной из кожи амфибий, для получения библиотеки пептидов, потенциально могущих секретироваться, с дальнейшим направленным поиском предсказанных последовательностей в кожных секретах. Но транскриптом не отражает состава пептидома – полного набора всех секретлируемых пептидов: он больше, не учитывает посттрансляционных модификаций, содержит только те классы пептидов, поиск которых был заложен исследователем на стадии геномных экспериментов. Стоит также подчеркнуть, что во всех существующих масс-спектрометрических подходах для установления структур пептидов использовалась только малая часть всех фрагментных ионов тандемных спектров, тогда как большинство этих ионов просто игнорировалось.

В нашей группе долгое время использовали комбинированный метод секвенирования, объединяющий способ «сверху вниз» с простыми химическими модификациями. Из-за потерь минорных компонентов при проведении модификаций и очистки образцов мы впервые

решили отказаться от химического модифицирования и попытаться секвенировать кожные пептиды только методами тандемной масс-спектрометрии высокого разрешения (подход «сверху вниз»).

**Объекты исследования** – это пептидные компоненты кожных секретов особей *Rana arvalis* из центральнословенской популяции, а также особей *Rana temporaria* из центральнословенской и архангельской популяций.

**Методология диссертационного исследования.** Секвенирование компонентов интактных кожных пептидомов проводилось комплексом методов тандемной масс-спектрометрии высокого разрешения в варианте «сверху вниз», разработанным в рамках выполнения данной работы, с ручной интерпретацией полученных масс-спектров. Проведено сравнение результатов ручного и автоматического секвенирования компонентов кожных секретов, в том числе и дериватизованных, на примере программы PEAKS Studio.

**Личный вклад автора** заключался в научной проработке темы диссертации, получении и интерпретации экспериментальных данных, написании диссертации и автореферата к ней. В совместно опубликованных по теме работах соискателю принадлежит существенный вклад.

**Цели и задачи.** Целью данной работы явилось создание надежного метода секвенирования компонентов кожных секретов ранидных амфибий исключительно методами масс-спектрометрии в варианте «сверху вниз». Для этого необходимо было:

- Изучить применимость гибридного MS<sup>3</sup> метода фрагментации EThcD (диссоциация при переносе электрона с дополнительной активацией соударениями при повышенной энергии) к секвенированию пептидных компонентов кожных секретов амфибий семейства *Ranidae*;
- Изучить характер протекания радикального разрыва S-S связей в условиях EThcD фрагментации в пептидах, содержащих одну дисульфидную связь, с целью применения к секвенированию новых кожных пептидов и пептидов с уже известными последовательностями методом «сверху вниз» (без использования химических модификаций);
- Изучить серии фрагментных ионов, не описанные классическими механизмами фрагментации и не используемые в процедурах ручного и автоматического секвенирования, в разных типах тандемных масс-спектров;
- Установить пептидомы трёх ранидных амфибий: *Rana arvalis* (центральная Словения) и *Rana temporaria* (центральная Словения и Архангельск), включая последовательность аминокислот внутри дисульфидных циклов и идентификацию изомерных остатков Leu/Ile;
- Сравнить полученные составы пептидомов с составами других популяций этих двух видов амфибий для выявления потенциальных биомаркеров популяций и вида;

- Сравнить данные, полученные ручной интерпретацией спектров, с результатами автоматического секвенирования с помощью программы PEAKS Studio;

**Научная новизна.** Впервые установление первичной структуры пептидов проводилось без использования химических модификаций, исключительно методами tandemной масс-спектрометрии, доступными на масс-спектрометрах с орбитальной ловушкой: ДАС (диссоциация, активированная соударениями); ДАСПЭ (диссоциация, активированная соударениями при повышенной энергии); ДПЭ (диссоциация под действием переноса электрона) и с помощью MS<sup>3</sup> эксперимента, EThcD.

Показана наибольшая информативность EThcD при секвенировании кожных пептидомов ранидных лягушек методом «сверху вниз». Он отлично зарекомендовал себя в определении последовательностей как коротких, но сложных для секвенирования объектов, темпоринов, так и длинных дисульфидсодержащих пептидов. Комплементарная ДАС/ДАСПЭ фрагментация внутри «Rana box» в условиях EThcD дополнила созданный алгоритм секвенирования «сверху вниз».

Показано, что радикальный разрыв S-S связей в дисульфидсодержащих пептидах протекает для всех изученных семейств пептидов и сопровождается фрагментацией пептидных связей внутри C-концевого S-S цикла («Rana box»). Предложена схема фрагментации дисульфидсодержащих пептидов в условиях EThcD эксперимента, позволяющая установить всю последовательность.

Впервые установлены составы трёх кожных пептидомов ранидных амфибий: *Rana arvalis* (центральная Словения) и *Rana temporaria* (центральная Словения и Архангельск), включая C-концевые последовательности практически всех интактных дисульфидсодержащих пептидов и дифференцированы изомерные остатки Leu/Pe практически во всех 160 обнаруженных пептидах, в том числе 16, описанных впервые. Проведено сравнение пептидомов исследованных популяций с пептидомами других популяций этих двух видов амфибий для выявления потенциальных биомаркеров популяций и вида;

Показана применимость программы автосеквенирования PEAKS Studio для коротких пептидов семейства темпоринов и коротких бревининов 1 (до 24 а.к.), где программно удаётся частично секвенировать линейную часть интактных пептидов или установить полную последовательность части модифицированных пептидов в случае высокой концентрации молекулы и успешно прошедшей реакции дериватизации.

Проведена оценка потенциальной биологической активности новых пептидов из двух пептидомов путем построения их 2D карт и по их физико-химическим характеристикам.

Составлены две таблицы возможных замен для темпоринов (из различных популяций вида *Rana temporaria*) и мелиттин-родственных пептидов, известных на сегодня.

**Практическая значимость работы.** Созданный метод секвенирования природных нетриптических пептидов ранидных амфибий «сверху вниз» исключительно при помощи масс-спектрометрии без модифицирования исходного субстрата может быть использован для установления первичной структуры пептидов других организмов, содержащих одну внутримолекулярную дисульфидную связь. Комплементарность структурных данных, полученных при фрагментациях, протекающих в «Rana box» у дисульфидных пептидов в условиях MS<sup>3</sup> эксперимента EThcD и ДАС/ДАСПЭ, делает обоснованным и более эффективным секвенирование интактных кожных секретов амфибий ранидных лягушек способом «сверху вниз», поскольку устраняет потери их минорных компонентов, неизбежно происходящие при очистке образцов после любой химической модификации. Показанная эффективность секвенирования коротких темпоринов по спектрам EThcD делает этот метод ключевым для применения его к исследованию компонентного состава кожных секретов амфибий по типу «сверху вниз». Параллельный эксперимент с классическими модификациями показал себя значительно менее эффективным. Получены идентификационные номера в международной базе данных UnitProKB для впервые установленных пептидных последовательностей. Продемонстрированные возможности текущей версии программы автоматического секвенирования PEAKS Studio могут быть использованы для получения частичной информации о последовательностях компонентов кожных секретов ранидных лягушек. Показана эффективность продления времени хроматографирования и сбора спектров в режиме зависимых сканирований с временно отложенным сканированием интенсивных ионов (DDA) для секвенирования минорных компонентов секретов. Составленные таблицы возможных замен с выделением консервативных участков в последовательностях темпоринов из кожных секретов *Rana temporaria* и мелиттин-родственных пептидов оказываются полезными при секвенировании новых пептидов этих семейств.

**Публикации.** По полученным в результате выполнения работы данным опубликованы три печатные работы в ведущих зарубежных журналах уровня Q1, индексируемых международными базами данных (Web of Science, Scopus) и рекомендованных для защиты в диссертационном совете МГУ по специальности 1.4.3 – органическая химия, 8 тезисов докладов на российских и международных конференциях.

**Апробация работы.** Основные положения работы доложены на семи отечественных и зарубежных конференциях: на Международной конференции Socratic Symposium 2021 (Любляна, Словения, 2020); на XXVIII-ой Международной научной конференции студентов, аспирантов и молодых ученых "Ломоносов 2021", (Москва, 2021); на IX-ой Всероссийской

конференции Всероссийского масс-спектрометрического общества с международным участием «Масс-спектрометрия и ее прикладные проблемы», (Москва, 2021); Chem2Change, Environmental Chemistry towards Global Change, (Словения, 2022); на Международной научной конференции студентов, аспирантов и молодых учёных «Ломоносов-2022», (Россия, 2022); на XXIV Международной масс-спектрометрической конференции (Маастрихт, Нидерланды, 2022), XII Международной масс-спектрометрической конференции по нефтехимии, экологической и пищевой химии Petromass (Крит, Греция, 2022).

**Объём и структура работы.** Диссертация изложена на 142 страницах. Состоит из введения, литературного обзора, обсуждения результатов, экспериментальной части и заключения. Содержит 15 таблиц, 51 рисунок и 4 схемы. В списке цитируемой литературы содержится 271 ссылка.

**Положения, выносимую на защиту:**

1. Разработанный новый метод «сверху вниз» позволяет устанавливать первичную структуру природных интактных пептидов без каких-либо предварительных химических модификаций.

2. Фрагментация дисульфидсодержащих пептидов с участием радикального центра в условиях EThcD эксперимента позволяет установить последовательность внутри S-S циклов.

3. Комплементарность структурной информации о последовательностях «Rana box», получаемых при фрагментации методами EThcD и ДАС/ДАСПЭ позволяет подтвердить последовательность внутри S-S циклов.

4. Применимость нового метода секвенирования «сверху вниз» показана на трёх не изученных ранее популяциях ранидных амфибий.

5. На изученных вручную образцах применён автоматический алгоритм секвенирования PEAKS Studio, показана его эффективность для получения частичной информации о нетриптических пептидах амфибий семейства *Ranidae*.

## 1. ОБЗОР ЛИТЕРАТУРЫ

Публикации двух последних десятилетий наглядно показали, что пептиды, секретируемые кожными железами и слизистыми оболочками различных позвоночных, представляют собой их основной способ защиты от патогенных микроорганизмов и хищников [1,2]. Такие пептиды были обнаружены у рыб [3,4], рептилий [5], пауков [6], земноводных [7–10], скорпионов [11], млекопитающих [12] и др. Пептиды с теми же функциями известны и для растений [13]. Поскольку природные субстраты представляют собой довольно сложные смеси, важно идентифицировать и классифицировать каждый компонент, чтобы понять его биологическую роль [14].

Антимикробные пептиды (АМП) представляют собой подавляющее большинство защитных пептидов, активных в отношении широкого спектра бактерий. Обычно они многофункциональны, защищают также от вирусов, грибков и даже опухолевых клеток [8,10]. Кроме того, пептиды одного и того же семейства могут продуцироваться разными видами. Например, дефенсины могут секретироваться разными птицами, рептилиями, рыбами, моллюсками, членистоногими, растениями и грибами, млекопитающими и даже человеком [15]. Последовательность регуляторного пептида брадикинина, присутствующего в организме человека, впервые была обнаружена в кожном секрете лягушек [16]. Всесторонний обзор онтогенеза, анатомии, цитологии, нейроэндокринологии и иммунологии кожи амфибий, включая её сходство с кожей человека, был опубликован в 2014 г. И. Хасламом с соавторами [17]. АМП обычно представляют собой небольшие, преимущественно катионные амфипатические  $\alpha$ -спирали. Механизм их взаимодействия с патогенными клетками хорошо изучен [18]. Независимо от способа взаимодействия пептидных  $\alpha$ -спиралей с клеточной мембраной нарушение целостности клеточной стенки сопровождается лизисом. Особо следует отметить, что благодаря уникальному механизму взаимодействия АМП с мембранами патогенных клеток, запуская лизис и неминуемую гибель, становится невозможным развитие резистентности микроорганизмов к таким пептидам. Адаптация к молекулам, выработанным животными в течение миллионов лет, должна потребовать значительного изменения всей структуры патогена, что, согласно современным научным представлениям, крайне маловероятно. Сам факт того, что живые организмы миллионы лет успешно используют эти пептиды, является дополнительным подтверждением эффективности выработанных защитных механизмов. Этот вопрос особенно актуален для борьбы с резистентными внутрибольничными инфекциями, а природные пептиды рассматриваются как потенциальные фармацевтические препараты следующих поколений.

Нейропептиды представляют собой ещё одну важную группу защитных пептидов живых организмов. Они составляют весомую часть кожного секрета лягушек, ядов и токсинов

змей, скорпионов, насекомых, моллюсков-конусов и пауков. Эти молекулы чрезвычайно быстро взаимодействуют с определёнными физиологическими рецепторами, вызывая соответствующую реакцию, поэтому их применение в биохимии и фармакологии весьма перспективно [19].

К сожалению, использование фармацевтических препаратов на основе природных пептидов ограничено их нестабильной выработкой в живых организмах. Существуют также фармакокинетические и фармакодинамические ограничения их использования, в том числе проблема потери активности в зависимости от pH и присутствия солей [20,21]. Таким образом, для эффективного использования природных пептидов в качестве сильнодействующих лекарственных средств будущие исследования должны быть направлены на решение этих проблем.

Земноводные являются наиболее хорошо изученными животными [22,23], хотя их геномы в основном остаются неисследованными. К настоящему времени установлены последовательности примерно 2000 пептидов амфибий, принадлежащих примерно ста семействам [7]. Эти пептиды обладают широкой антибактериальной, противовирусной, противогрибковой и противоопухолевой активностью. Они проявляют нейроактивность, ведут себя как иммуномодуляторы, ингибиторы NO-синтазы, стимуляторы выработки инсулина и т.д. [7,24–27]. В настоящее время антимикробные пептиды часто называют цитолитиками, подчёркивая их особую защитную роль и механизм воздействия на патогенные клетки. Была выдвинута гипотеза, что благодаря цитолитической активности антимикробных пептидов быстрая доставка нейропептидов в нервную и эндокринную системы хищников, вызывающая немедленную боль или рвоту, приводила к спасению лягушки [9]. Стоит подчеркнуть, что при секреции эти пептиды выполняют свои функции в течение секунд, а их избыток быстро и полностью уничтожается амфибией за счёт одновременного высвобождения соответствующих протеаз, расщепляющих зрелые пептиды на неактивные короткие фрагменты [28,29]. Аналогичное наблюдение было сделано в работе [30] при работе с ядом скорпиона *Tityus serrulatus*.

Последовательности многих охарактеризованных пептидов амфибий были определены при помощи комбинации масс-спектральных данных с деградацией по Эдману и/или клонированием кДНК [25,31]. Хотя знание генома очень полезно для установления последовательностей пептидов, оно не даёт информации о многочисленных посттрансляционных модификациях (ПТМ), которые часто определяют конечную функциональную активность этих пептидов [32]. ПТМ включают в себя варианты дисульфидных циклов [33,34], амидирование C-конца, окисление метионина [28], появление в последовательностях гидроксипролина и бромтриптофана [35]. Наличие указанных ПТМ, а

также длина пептидов, делают деградацию по Эдману всё менее полезной, при этом роль масс-спектрометрии быстро возрастает.

Хотя о секвенировании пептидов с помощью масс-спектрометрии впервые сообщалось ещё в 1963 г. [36], оно стало достаточно эффективным при появлении ионизации методом бомбардировки быстрыми атомами (ББА) [37–39]. Однако основные достижения в этой области связаны с появлением новых мягких методов ионизации: ионизации электрораспылением (ИЭР) [40,41] и матрично-активированной лазерной десорбции/ионизации (МАЛДИ) [42–44]. Их активное внедрение в биологические и медицинские исследования совпало с разработкой масс-спектрометра ион-циклотронного резонанса (ИЦР) с внешним источником ионов [45]. ИЭР и МАЛДИ произвели революцию в масс-спектрометрии, позволив эффективно работать с природными макромолекулами, такими как белки, гликаны и ДНК. Секвенирование пептидов в определённом смысле проще по сравнению с другими биополимерами. Пептиды состоят из 20 генетически кодируемых аминокислот, содержащих только атомы С, Н, N, О и S, и ряда других [46,47], причём изомерных аминокислот всего две: изолейцин и лейцин. Секвенирование пептидов с помощью масс-спектрометрии стало настолько популярным, что в последнее время было прекращено производство приборов для альтернативных методов, включая классическую деградацию по Эдману [48].

Граница между массами белков и пептидов весьма условна и составляет 5000 Да. Природные пептиды могут быть довольно длинными. Они могут содержать много основных остатков (Lys и Arg) и различные посттрансляционные модификации. Часто бывает так, что в биологическом/медицинском эксперименте наблюдается белок/пептид, играющий существенную роль в определённом процессе, активирующий или ингибирующий рецептор и т.д. В этом случае желательно идентифицировать этот компонент. Идентификация белков относительно проста, если известны их последовательности. Трипсинолиз и анализ белков методом ВЭЖХ-МС/МС с последующим поиском в базе данных даст ответ даже при детектировании лишь 2-3 триптических пептидов для каждого исходного белка (так называемый подход «снизу вверх»). В настоящее время подавляющее большинство протеомных исследований основано на этом подходе. Однако существуют некоторые проблемы, связанные с его применением: неполный протеолиз белка, присутствие примесных пептидов, автолиз фермента и определение полной последовательности исходного белка, используя триптические фрагменты в качестве «строительных блоков».

В качестве альтернативы существует подход «сверху вниз», позволяющий анализировать весь исходный пептид/белок в его интактной форме. Он может быть гораздо более надёжным, поскольку работает с изучаемой молекулой, получая её структуру при

фрагментации всей молекулы. Метод не имеет упомянутых ограничений подхода «снизу вверх», хотя и требует более мощных масс-спектрометрических приборов [49]

Работая с животными или растениями с неизвестным геномом, исследователи сталкиваются с более сложной ситуацией, поскольку необходимо получить полную аминокислотную последовательность, используя исключительно собственные экспериментальные результаты. Биологически активные пептиды, яды, токсины, гормоны, антитела являются наиболее популярными молекулами в этих исследованиях. Во многих случаях, даже обладая информацией о геноме, приходится иметь дело с многочисленными ПТМ, которые могут значительно усложнить интерпретацию масс-спектров. Без информации о геноме установление аминокислотной последовательности является собой *de novo* секвенирование. Отличный обзор основ *de novo* секвенирования при помощи масс-спектрометрии был опубликован К. Медзиградски и Р. Чакли [50].

Автоматическое *de novo* секвенирование работает достаточно хорошо, когда речь идёт о триптических пептидах, содержащих единственный остаток основной аминокислоты Lys или Arg на C-конце. Первый надёжный автоматический высокопроизводительный метод для *de novo* секвенирования при помощи ИЦР ПФ был разработан в 2005 г. [51]. Он использует высокую точность измерения масс и разрешающую способность ИЦР ПФ, а также по крайней мере два метода тандемной масс-спектрометрии, ДАС и ДЗЭ. Алгоритмы автоматического *de novo* секвенирования становятся менее эффективными, когда исследуются природные длинные нетриптические пептиды, часто имеющие посттрансляционные модификации и принадлежащие организмам с неизвестным геномом. Так, В. Горшков с соавторами [52], изучая трутней медоносных пчел *Apis mellifera* с помощью масс-спектрометра LTQ-FT и ДАС фрагментации, пришли к выводу, что ручное секвенирование, в результате которого были обнаружены 130 новых пептидов, было более эффективным, чем применение алгоритма поиска SEQUEST. Аналогичные выводы были сделаны в работе [53], когда поисковые системы Lutefisk, PEAKS, PepNovo, EigenMS и pNovo/pNovo+ не смогли секвенировать какие-либо пептиды в секрете лягушки *Rana ridibunda*. Напротив, альтернативное ручное секвенирование выявило полные последовательности 12 пептидов. Изучая пептидом тарантула *Acanthoscurria Gomesiana*, Т. Абреу с соавторами [54] пришли к выводу, что, хотя автоматический анализ масс-спектров ускорил процесс, ручная проверка была необходимым шагом для надёжной идентификации пиков ионов с большим зарядом, посттрансляционных модификаций и мутаций. Преимущества ручной интерпретации масс-спектров, по крайней мере, для определения N- и C-концевых участков и ПТМ, также подчёркиваются в исследованиях, посвященных яду моллюска-конуса *Conus textile* [55]. О некоторых проблемах, связанных с применением автоматических алгоритмов для *de novo* секвенирования токсина с

массой 7 кДа из яда техасского кораллового аспида (*Micrurus tener tener*), сообщалось в работе К. Медзиградски и К. Болен [33]. Ручная интерпретация масс-спектральных данных оказалась успешной для определения структуры пептида с тремя дисульфидными связями.

С другой стороны, группа, изучающая антимикробные пептиды американского аллигатора и комодского варана, использовала автоматическое *de novo* секвенирование при помощи программного обеспечения PEAKS и проверяла правильность полученных последовательностей вручную [56–58]. Сравнение последовательностей 8 новых потенциально антимикробных пептидов *Varanus komodoensis*, предсказанных с помощью PEAKS, с последовательностями, определёнными вручную или путем сопоставления с транскриптомом, показало, что некоторые пептиды были идентифицированы на 100% правильно, однако другие имеют менее чем 50% правильно идентифицированных аминокислотных остатков в последовательности [57].

Конечно, будущие версии алгоритмов автоматического секвенирования должны быть более эффективными. Возможный алгоритм должен сочетать информацию из спектров новых пептидов, полученных с использованием различных методов фрагментации, данные об известных последовательностях пептидов, принадлежащих к соответствующим семействам и обладающих определенным сходством, консервативными участками, характерным аминокислотным составом или определёнными посттрансляционными модификациями [34,59]. Все эти моменты так или иначе учитываются при ручной интерпретации спектров. В этом случае применимые условия могут значительно облегчить *de novo* секвенирование новых пептидов. Эффективность этого подхода была продемонстрирована в алгоритме Copovo, разработанном для *de novo* секвенирования пептидов моллюсков-конусов [34].

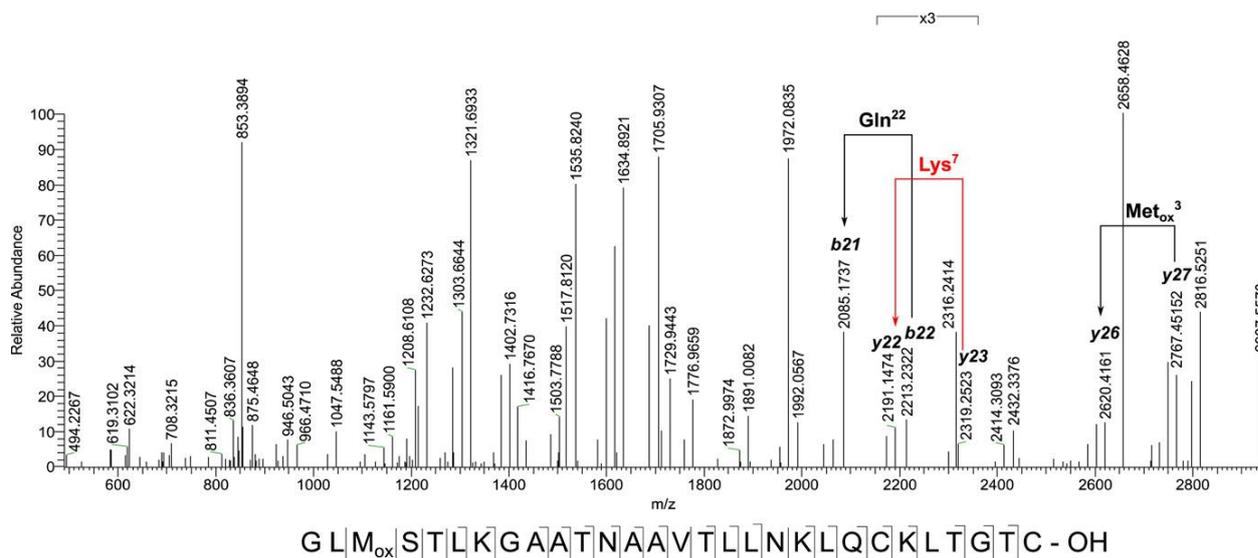
Ионный циклотронный резонанс с преобразованием Фурье (ИЦР ПФ), возможно, является самым мощным методом масс-спектрометрии с чрезвычайно широким набором применений [60]. С момента разработки ИЦР ПФ в 1974 году А. Маршаллом и М. Комисаровым [61] такие приборы обеспечивают широкий диапазон регистрируемых масс, непревзойденную точность их измерения, наивысшую разрешающую способность прибора, высокую чувствительность и многочисленные методы tandemной масс-спектрометрии. Преимущества ИЦР ПФ могут быть использованы для определения первичной структуры новых природных пептидов. Польза такого подхода неоспорима, так как он позволяет проводить наиболее надёжное *de novo* секвенирование важных биополимеров зачастую без какой-либо предварительной информации. Это делает приборы ИЦР ПФ предпочтительными для подобных целей по сравнению с другими масс-спектрометрами. Тем не менее, в самом конце XX века в двух статьях Александра Макарова было сообщено о создании нового масс-анализатора под названием Orbitrap [62,63]. Орбитальная ловушка с преобразованием Фурье

была основана на идее ионной ловушки Кингдона без использования каких-либо магнитных полей. Следует отметить, что орбитальные ловушки в настоящее время демонстрируют сходные с ИЦР характеристики и даже лучшую разрешающую способность для высоких значений  $m/z$  [64]. Последнее связано с тем, что разрешение в приборах ионного циклотронного резонанса обратно пропорционально  $m/z$ , а разрешение в приборах на основе орбитальной ловушки ионов – квадратному корню  $m/z$ . Так, для компактной орбитальной ловушки [65] и прибора ИЦР, оснащенного магнитом 15 Тл, точка равного разрешения находится около  $m/z = 300$  [66]. Принимая во внимание стоимость и простоту эксплуатации приборов, неудивительно, что орбитальные ловушки стали более популярными.

### 1.1. Преимущества масс-спектрометрии с преобразованием Фурье

Непревзойденная разрешающая способность и точность измерения масс – очевидные преимущества приборов с ПФ. Для надёжной идентификации двух фосфорилированных пептидов в работе Ф. Хе с соавторами [67] требовалась точность масс 0.032 и 0.053 ppm при разрешающей способности 552 000, в то время как в работе Р. Боссио с соавторами [68] сообщалось о разрешающей способности 3 300 000 для разделения пары изобарных пептидов RVMRGMR и RSHRGHR. Разрешающая способность 1 000 000 для  $m/z$  2700 была получена в режиме адсорбции для иона антитела с зарядом  $55^+$  (молекулярная масса 150 кДа) при помощи ИЦР ПФ масс-спектрометра с магнитом 21 Тл и ионной ловушкой 4X-ICR. Важно отметить, что эта разрешающая способность была получена всего за 1,5 с [69]. С другой стороны, разрешающая способность 40 000 000 для  $m/z$  609 была получена в режиме амплитуды для ионов резерпина (ММ 609 Да), время накопления данных составляло 300 с, сигнал регистрировали от динамически согласованной ИЦР ионной ловушки и масс-спектрометра ИЦР ПФ с магнитом 7 Тл [70].

Превосходная точность измерения масс обеспечивает надёжное определение элементного состава достаточно сложных биомолекул с массой до нескольких тысяч Дальтон – это как раз диапазон масс природных пептидов. Точное измерение масс фрагментных ионов устраняет проблемы с идентификацией изобарных аминокислотных остатков, таких как лизин (Lys) и глутамин (Glu) или фенилаланин (Phe) и окисленный метионин (Met<sub>ок</sub>). Рисунок 1 демонстрирует, насколько легко можно различить эти изобарные остатки в ДАС спектре окисленного природного пептида бревинина 2Rb из секрета лягушки *Rana ridibunda* [71].



**Рисунок 1.** Идентификация изобарных аминокислотных остатков в ДАС спектре восстановленного с помощью ДТТ бревинина 2Rb с окисленным остатком метионина,  $MH_3^{3+}$ ,  $m/z$  979.857 (гибридный масс-спектрометр LTQ FT 7 Тл, разрешающая способность 100 000 в режиме MS и 50 000 в режиме MS/MS, точность измерения масс 4 ppm).

Сверхвысокая разрешающая способность ИЦР ПФ позволила разделить ионы с разницей масс менее 0,10 ppm в МАЛДИ-ПФ-МС эксперименте при работе со сложными образцами тканей синих крабов *Callinectes sapidus* [72]. Этот подход применялся как дополнительный к МАЛДИ-ВП/ВП эксперименту и помог авторам определить пептиды с  $m/z$  1019.50 и 1019.59, а также  $m/z$  1474.66 и 1474.47. Кроме того, реализована визуализация их распределения в образце ткани.

При работе с фрагментными ионами с массами несколько сотен Да точность в 1 ppm вполне достаточна для определения их элементного и аминокислотного состава, а для однозначного определения элементного состава пептида с молекулярной массой около 1100 Да точность измерения масс должна быть 0.1 ppm [73]. Совмещение информации об элементном составе всей молекулы и нескольких ее фрагментов позволяет получить аминокислотный состав пептида. В 2004 году Б. Шпенглер [74] предложил эффективный подход с использованием ИЦР ПФ, основанный на точном измерении масс. Его первый этап называется анализом состава пептидов и включает определение аминокислотного состава. Когда известен аминокислотный состав пептида, интерпретация его спектра становится намного проще и надёжнее. Шпенглер назвал этот второй этап секвенированием на основе состава (CBS).

Значения масс пептидов не разбросаны случайным образом по всему диапазону масс. Они занимают определенные сегменты со средним значением  $1.000458n$ , где  $n$  — целое число [75]. Эта так называемая «классификация по массе» позволяет различать пептиды, сахара,

липиды и т.д. путём измерения точной массы протонированных молекул [76,77]. Более того, определённое развитие этих представлений дало возможность различать семейства пептидов по кластерам протонированных молекул, полученным с максимальной точностью измерения масс [78]. Подход был основан на построении 2D-карт с нормализованным дефектом масс по оси абсцисс и нормализованным изотопным сдвигом по оси ординат. Оба параметра могут быть легко рассчитаны из данных ИЦР ПФ. Пептиды, относящиеся к одному семейству, занимают определенную область на этом графике, что часто позволяет определить их тип исключительно по их точной молекулярной массе. Более того, эти 2D-карты хорошо визуализируют весь пептидом, позволяя легко и надёжно различать близкородственные виды [79]. Этот момент будет подробно рассмотрен ниже. На самом деле масс-спектрометрию в настоящее время можно рассматривать как мощный инструмент для таксономических исследований: различения видов и даже популяций [59,78,79].

## 1.2. Приборы тандемной масс-спектрометрии

Тандемная масс-спектрометрия становится незаменимой, когда необходимо установить последовательность пептида. В настоящее время существует довольно много методов фрагментации, обладающих своими преимуществами и недостатками. Недавний представительный обзор посвящен методам тандемной масс-спектрометрии для секвенирования пептидов [80].

Приборы ИЦР ПФ позволяют применять несколько методов тандемной масс-спектрометрии. Их несомненным преимуществом является способность улавливать ионы с очень низкими энергиями, а также проводить последовательные масс-спектрометрические эксперименты и масс-спектрометрические эксперименты непосредственно в ионной ячейке. В качестве альтернативы ионы могут быть получены вне ячейки и введены в ИЦР только для проведения точного измерения масс. MS/MS эксперименты являются ключевыми для *de novo* секвенирования, а также для сложных задач, например, дифференцирования изомерных аминокислотных остатков [81]. Метод сохранённой формы волны с обратным преобразованием Фурье (SWIFT), предложенный А. Маршаллом в 1985 г., используется для выделения ионов, представляющих особый интерес для фрагментации [82].

Подробная информация о наиболее популярных методах активации соударениями в ячейке ИЦР представлена в репрезентативном обзоре [83]. Диссоциацию при множественных соударениях (MSA-CID) можно рассматривать как метод медленной активации. Для удовлетворения требований к вакууму применяется импульсный ввод газа для соударений. Затем происходит фрагментация, когда ионы-предшественники многократно возбуждаются. Наиболее популярный подход включает непрерывное вне-резонансное возбуждение (SORI),

когда частота возбуждения немного отличается от частоты иона-предшественника [84]. Так как амплитуда импульса возбуждения невелика, ионы остаются близко к центру ячейки. Последнее довольно важно для более высокого разрешения. Медленное увеличение энергии приводит к преимущественной фрагментации главной цепи с разрывом пептидных связей и образованием *b*- и *y*-серий ионов. Дериватизация может улучшить спектры SORI-CID. Этерификация кислотных аминокислотных остатков удлиняла *b*- и *y*-серии ионов, тогда как дезамидирование с последующим метилированием было полезно для выявления амидных остатков: *C*-концевых, Asn и Gln в МАЛДИ-ИЦР ПФ экспериментах по изучению нейропептидов крабов, оркокининов [85].

Тем не менее, преобладающее число последовательностей пептидов было установлено при помощи фрагментации во внешней ячейке соударений с последующим точным измерением масс в ИЦР или орбитальной ловушке ионов. ДАС остается основным методом фрагментации в протеомных исследованиях, а несколько статей продемонстрировали её применимость для *de novo* секвенирования без каких-либо дополнительных методов. При анализе методом «сверху вниз» яда гадюки *Vipera anatolica* в ДАС спектрах было обнаружено лишь несколько коротких серий *b*- и *y*-ионов [86], а ДАС спектры кожного секрета древесной лягушки *Hyla savignyi* позволили идентифицировать 9 коротких пептидов семейства триптофиллинов [87]. М. Лэнгсдорф с соавторами установили механизм образования внутренних *b*-ионов и элиминирования CO<sub>2</sub> в режиме регистрации отрицательных ионов. Также спектры усложнялись присутствием перестановочных ионов, из-за чего авторы прибегли в следующей работе [88] к дополнительной ДЗЭ фрагментации для установления структуры.

Орбитальные ловушки позволяют использовать фрагментацию, индуцированную соударениями при повышенной энергии (ДАСПЭ, HCD), где столкновения происходят во внешнем мультиполе [89]. Поскольку параметры захвата частиц не влияют на активацию ионов, ДАСПЭ не имеет ограничений в области низких масс в отличие от ДАС, благодаря чему становится возможно работать с фрагментными ионами с низкими массами. Хотя энергия соударений в ДАСПЭ выше, чем в ДАС, она все же меньше 100 эВ. Как и применение ДАС, использование ДАСПЭ приводит в основном к образованию *b*- и *y*-серий ионов и редко используется отдельно для секвенирования пептидов. Так, автоматическое *de novo* секвенирование позволило установить 33 антимикробных пептида в каждом секрете древесной лягушки *Hypsiboas pulchellus* [90]. Последовательности более, чем 150 пептидов в яде африканских гадюк были определены при помощи *de novo* секвенирования исключительно с помощью ДАСПЭ фрагментации [91]. ДАСПЭ становится очень полезной в сочетании с дополнительными методами фрагментации.

При секвенировании дисульфидсодержащих пептидов из кожных секретов лягушек Т. Самгина с соавторами сравнивали спектры ДАС, полученные на приборе ИЦР ПФ с магнитом 7 Тл, LTQ FTICR, и спектры ДАСПЭ, зарегистрированные на приборе с орбитальной ловушкой ионов, LTQ Orbitrap Velos [53]. ДАСПЭ фрагментация обеспечивала более длинные и более выраженные серии *b*- и *y*-ионов, улучшая покрытие сиквенса, хотя большое количество ионов вторичной фрагментации усложняло спектры. Ещё одна положительная особенность ДАСПЭ заключалась в расщеплении пептидных связей на C-концах остатков Pro. Это может быть полезно для обнаружения пропущенных расщеплений сайтов в триптических пептидах. В любом случае, сочетание спектров ДАС и ДАСПЭ определённо повышает надёжность и эффективность *de novo* секвенирования.

Диссоциация, индуцированная столкновениями с поверхностью (ДИП), предполагает очень быстрое возбуждение ионов в одном столкновении. Поскольку для ДИП не требуется газ, это устраняет проблемы, связанные с поддержанием вакуума, и позволяет значительно сократить время сбора данных. В работе Дж. Ласкин с соавторами [92] спектры ДИП были зарегистрированы за 0.2 с, а в экспериментах SORI-CID за 5 с. Как правило, ДИП является отличным методом для исследования механизмов фрагментации ионов пептидов в ИЦР ПФ [93]. Также стоит отметить, что если SORI-CID спектры пептидов с большими массами получаются в основном за счёт фрагментации основной пептидной цепи, то в их спектрах ДИП преобладают фрагментные ионы с малой массой, подобно высокоэнергетическим ДАС [83]. Кроме того, из-за разрывов C-C связей в спектрах могут наблюдаться ионы *d*-, *x*-, *w*- и *v*-серий, предоставляющие дополнительную информацию. К сожалению, из-за малой доступности использование ДИП для *de novo* секвенирования пептидов ограничено.

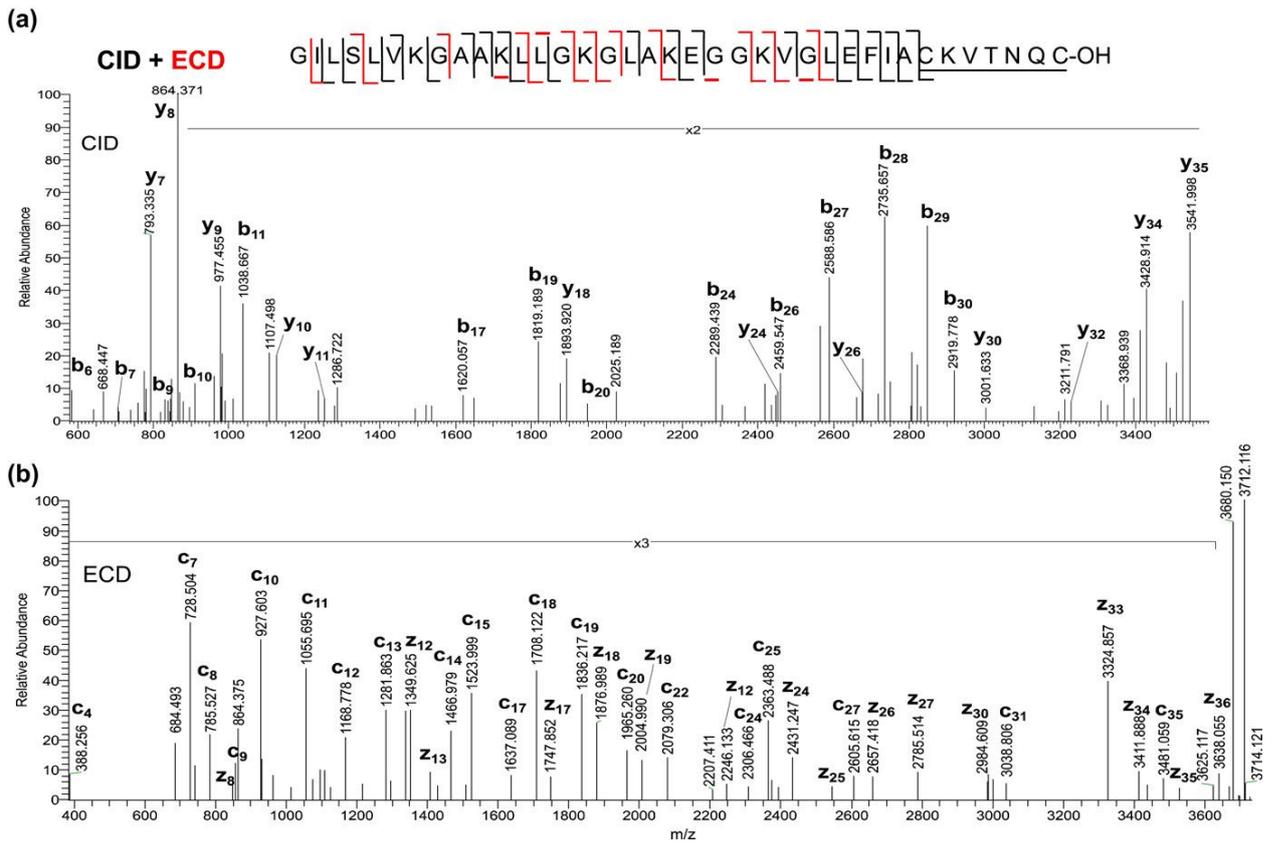
В 1998 г. был предложен новый метод диссоциации ионов пептидов, получивший название диссоциация при захвате электрона (ДЗЭ, ECD) [94]. Этот процесс включал захват низкоэнергетических электронов ( $<1$  эВ) полипротонированными пептидами в ячейке ИЦР [95]. Образующийся катион-радикал гораздо менее стабилен, чем исходный катион, что приводит к фрагментации, контролируемой радикальным центром, с образованием преимущественно *c*- и *z*-ионов. Альтернативный путь фрагментации даёт фрагменты *c*-1 и *z*+1 [96]. Время ДЗЭ фрагментации составляет около  $10^{-12}$  с, а такой процесс может рассматриваться как неэргодический [95], когда внутренняя энергия не успевает распределяться по всем возможным состояниям. Из-за неэргодичности ДЗЭ фрагментация происходит преимущественно по связям N-C<sub>α</sub> пептидной цепи, оставляя нетронутыми многие слабые связи модифицированных остатков [97]. Этот метод позволяет изучать посттрансляционные модификации [98], включая гликозилирование, сульфатирование и фосфорилирование, информация о которых может быть потеряна в условиях ДАС

фрагментации. Поскольку захват электрона уменьшает заряд катиона, метод неприменим к однозарядным протонированным пептидам.

В 2004 г. Дж. Сайка с соавторами [99] применили связанный с ДЗЭ метод, названный диссоциацией при переносе электрона (ДПЭ, ETD), на приборах с ионной ловушкой. ДПЭ основана на ион-ионных реакциях с участием многозарядных ионов пептидов и отрицательно заряженных ионов-реагентов, приводящих к образованию нечётноэлектронных ионов пептидов. Подобно ДЗЭ, в ДПЭ происходит преимущественный разрыв связей N-C<sub>α</sub> с образованием c- и z-серий ионов. Внутренняя энергия образующихся катион-радикалов при ДПЭ ниже, чем при ДЗЭ, из-за энергии связывания электронов в донорных анионах.

Хотя общая эффективность ДЗЭ/ДПЭ ниже, чем эффективность ДАС, иногда они сами по себе позволяют получить ценную информацию об аминокислотной последовательности. Например, при изучении пептидов ядовитых желез паука-волка *Lycosa singoriensis* авторы получили полную последовательность из c- и z-серий ионов при использовании ДЗЭ, а ДАС использовали только для подтверждения [99]. ДПЭ с дополнительной информацией о транскриптоме позволила *de novo* секвенировать сотни антимикробных пептидов из крови американского аллигатора и комодского варана [56–58]. Ранее неизвестные пептиды сельди, сульфаты протамина, были *de novo* секвенированы исключительно с помощью ДПЭ [100]. Недавно были опубликованы два ценных обзора по секвенированию пептидов методами диссоциации, промотируемой электронами [101,102].

Использование ДЗЭ/ДПЭ на полипротонированных пептидах даёт очень ценную информацию о последовательностях, дополняющую ДАС и ДАСПЭ [103]. Хорошим примером являются ДАС и ДЗЭ спектры пептида эскулентина 2Ra, GILSLVKGA AKLLGKGLAKEGGKVGLEFIACKVTNQС–ОН, выделенного из кожного секрета лягушки *Rana ridibunda*, содержащего 37 аминокислотных остатков и C-концевую внутримолекулярную дисульфидную связь. Спектры регистрировали при помощи гибридного масс-спектрометра LTQ FT, оснащённого магнитом 7 Тл (Thermo Fisher Scientific). В спектре ДАС обнаружено расщепление 23 пептидных связей, что позволяет установить только 62% последовательности. В условиях ДЗЭ расщепляется 21 связь (56,7%). Используя комплементарные спектры ДАС и ДЗЭ, удалось получить полную последовательность пептида, за исключением циклического C-концевого фрагмента [71]. Ещё одна особенность, о которой стоит упомянуть, связана с получением комплементарных спектров ДЗЭ с использованием различных зарядовых состояний одной и той же молекулы-предшественника – они могут предоставить дополнительную информацию.



**Рисунок 2.** ДАС (a) и ДЗЭ (b) спектры иона  $MH_5^{5+}$  с  $m/z$  743.229 эскулентина 2Ra с моноизотопной молекулярной массой 3711 Да. Разрешающая способность 100 000 в режиме MS и 50 000 в режиме MS/MS, точность измерения масс 3 ppm.

К. Сасаки с соавторами сравнивали ДАС и ДПЭ методы фрагментации на приборе LTQ Orbitrap XL в различных циклах ВЭЖХ-МС для идентификации эндогенных пептидов из секреторных гранул эндокринной клеточной линии человека [104]. Анализ выявил 795 пептидов в режиме ДАС и 569 пептидов в режиме ДПЭ, пересечение этих групп составило 397 пептидов. Хотя в режиме ДАС было зарегистрировано больше спектров, ДПЭ обеспечила более разнообразную фрагментацию, что позволило идентифицировать пептиды с неинформативными ДАС спектрами. ДПЭ была особенно полезна при дифференцировании изобарных пептидов, имеющих разность масс менее 2 ppm. Комплементарные ДАС и ДПЭ спектры помогли определить сайты фосфорилирования. В работе С. Харта с соавторами ДПЭ в дополнение к ДАС увеличивала степень идентификации пептидов (242 против 74), присутствующих в пептидоме мочи у здоровых женщин во время беременности [105].

В настоящее время для секвенирования пептидов часто используется ряд комплексных методов, основанных на ДЗЭ/ДПЭ. Они улучшают фрагментацию, удлиняют серии ионов, способствуют образованию характеристических ионов и т.д. Их подробное изложение выходит за рамки данного обзора. Однако стоит упомянуть следующие методы: перенос электрона без диссоциации (ETnoD) для изучения пептидно-белковых структур [106];

диссоциация, активируемая соударениями при переносе электрона (ETcaD), чтобы получить как *c/z*-, так и *b/y*-серии ионов путём дополнительной активации продуктов ETnoD, а также оставшихся ионов-предшественников [107]; диссоциация, активированная соударениями при повышенной энергии всех ионов-продуктов после ДПЭ (EThcD) для увеличения покрытия сиквенса [108,109], для работы с гликопептидами [110], для дифференцирования изомерных остатков Leu/Pe [111]; фотоактивация всей реакционной ячейки в реакциях ДПЭ (AI-ETD) для увеличения выхода различных ионов-продуктов [112].

Другой класс эффективных методов фрагментации в тандемной масс-спектрометрии использует фотоны. Один из таких методов использует взаимодействие ионов с ИК-фотонами с энергиями около 0,1 эВ [113]. Ионы, захваченные в ячейке ИЦР, активируются при постоянном облучении низкоэнергетическим (<100 Вт) CO<sub>2</sub>-лазером (10.6 мкм) в течение определенного периода времени, до сотен миллисекунд [114]. Поскольку для возбуждения ионов требуется несколько десятков фотонов, метод получил название инфракрасной мультифотонной диссоциации (ИКМФД). Каждый новый фотон медленно увеличивает энергию иона подобно процессу низкоэнергетической ДАС [115]. ИКМФД приводит к преимущественному образованию *b*- и *y*-серий ионов, однако спектры не идентичны ДАС и могут использоваться в качестве дополнительного МС/МС инструмента. Хотя спектры ДЗЭ и ИКМФД были зарегистрированы для триптических пептидов, они успешно использовались для идентификации сайтов гликозилирования и структур сахаров в гликопротеинах [116]. Комбинация этих методов может быть применима к *de novo* секвенированию.

Другая важная особенность связана с тем, что методы ИКМФД, как и ДЗЭ, не требуют наличия газа для соударений в ячейке ИЦР, что сокращает время анализа.

Диссоциация, протекающая при поглощении инфракрасного излучения абсолютно чёрного тела (BIRD) до сих пор не использовалась для секвенирования нетриптических пептидов, но, тем не менее, это известный, хотя и требующий много времени, метод ИЦР ПФ, полезный для расчёта параметров кинетической диссоциации пептидов [117].

Хотя в подавляющем числе ИЦР ПФ исследований фрагментации пептидов использовался режим регистрации положительных ионов, режим регистрации отрицательных ионов можно использовать для пептидов, содержащих несколько кислотных остатков [118] или посттрансляционные модификации, например, фосфаты или сульфонаты [119]. Благодаря более высокому количеству образующихся отрицательных ионов такие пептиды легче обнаруживаются в смесях. Фрагментация отрицательных ионов пептидов тщательно изучена группой Дж. Бови [120–122]. Хотя ДАС отрицательно заряженных пептидов часто приводит к преимущественным потерям малых молекул: H<sub>2</sub>O, CO<sub>2</sub>, фосфатных групп, SORI-CID иногда оказывается весьма полезной [118]. В то же время методы фрагментации, инициируемой

электроном, показывают более высокую эффективность. Анионы пептидов можно изучать с помощью нескольких соответствующих методов фрагментации: диссоциация отрицательных ионов при переносе электрона (NETD, negative electron transfer dissociation) [123]; диссоциация при отрыве электрона (ДОЭ, EDD), приводящая к расщеплению связей C $\alpha$ -C полидепротонированных пептидов с образованием *a'*- и *x*-серий ионов дополнительно к *c*- и *z*-ионам [124], а также диссоциация при захвате электрона отрицательными ионами (niECD), увеличивающая абсолютное значение заряда иона пептида за счёт захвата электронов с энергией 3–7 эВ и образования *c* и *z*-серий ионов [125]. NETD представляет собой газофазную реакцию электронодефицитных катионов с депротонированными пептидами, приводящую к потере электрона и образованию анион-радикала, который фрагментирует с образованием *a*- и *x*-серий ионов. Наиболее популярным реагентом, выступающим в качестве катиона, является флуорантен, также используется ксенон и некоторые другие молекулы [80,123]. В ДОЭ и niECD лабильные связи посттрансляционных модификаций обычно сохраняются, что позволяет охарактеризовать сульфированные и фосфорилированные пептиды. Оба метода часто обеспечивают лучшее покрытие сиквенса по сравнению с ДАС [80,126]. В работе К. Херсбергера и К. Хоканссона было проведено сравнение трёх упомянутых методов, а также ДАС, для характеристики модельных природных *O*-сульфопептидов [127], однако ни один из них не продемонстрировал очевидных преимуществ, и все они имеют свои специфические недостатки.

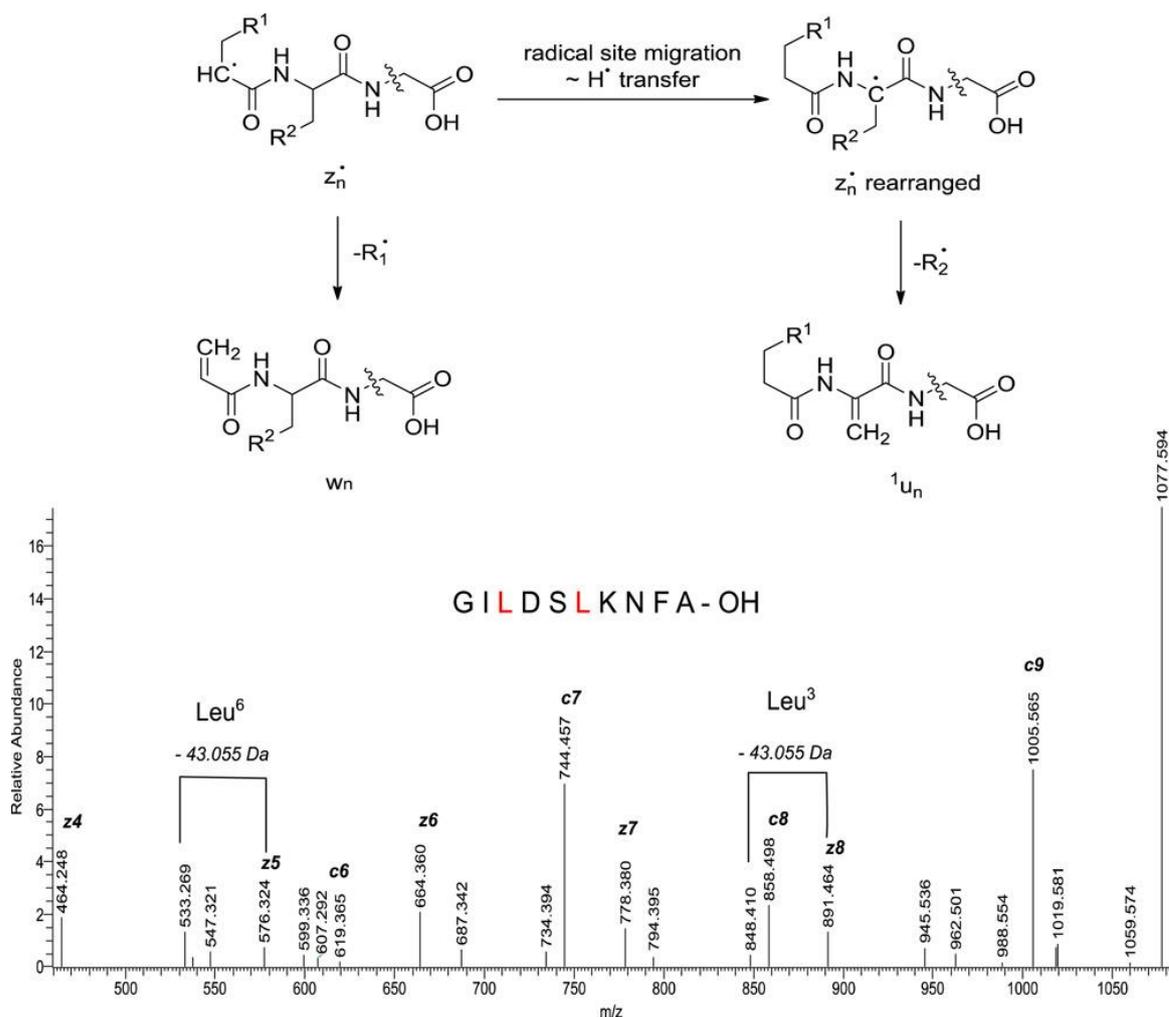
### 1.3. Масс-спектрометрия с преобразованием Фурье для *de novo* секвенирования нетриптических пептидов

#### 1.3.1. Изомерные аминокислоты

ДЗЭ можно эффективно применять для дифференцирования остатков изомерных аминокислот при секвенировании. Группа П. О'Коннора разработала надёжный способ для идентификации присутствия и положения изоаспартата при помощи ДЗЭ и поиска характеристических ионов  $c_n+58.0054$  (C<sub>2</sub>H<sub>2</sub>O<sub>2</sub>) и  $z_n-56.9976$  (C<sub>2</sub>HO<sub>2</sub>), где *n* – положение остатка изоаспартила. Таким же образом можно определять присутствие, но не положение аспарагиновой кислоты, находя ионы  $\{[M+nH]^{(n-1)+}-60\}$  [128]. Позже этот подход был распространён на ДПЭ [129,130]. В другой работе О'Коннор с соавторами дифференцировали  $\alpha$ - и  $\gamma$ -Glu в пептидах человеческого глаза с помощью ДЗЭ [131]. ДЗЭ также оказалась полезной для различения и количественного определения стереоизомерных пептидов (D-аминокислот), если известен их стереоизомерный состав [132,133]. Несколько статей,

цитируемых и обсуждаемых в работе П. Хуртадо и П. О’Коннора [81], касаются применения ДЗЭ для дифференцирования  $\alpha$ -,  $\beta$ - и  $\epsilon$ -пептидов.

Данные об изомерных остатках Leu/Ile можно получить из обычных спектров ДЗЭ [134,135]. На рис. 3 показано образование  $w$ -ионов из соответствующих  $z$ -ионов с радикальным центром, локализованным на  $N$ -концевом остатке Leu/Ile. Остаток лейцина теряет изопропильный радикал (43 Да), а остаток изолейцина – в основном этильный (29 Да) и в некоторой степени метильный (15 Да) радикалы. Стоит также отметить, что остатки алкил-Sus не продуцируют соответствующие  $z$ -ионы, а сразу же претерпевают образование  $w$ -иона как в ДЗЭ, так и в ДПЭ [33,136]. На рис. 3 также представлен фрагмент ДЗЭ спектра  $b_{10}^{2+}$ -иона бревинина 2Ra (GILDSLKNFAKDAAGILLKKASCKLSGQC-OH, ММ 2989.6 Да), полученного в ходе МС<sup>3</sup> эксперимента. Соответствующие ключевые  $z$ - и  $w$ -ионы могут быть использованы для дифференцирования изомерных Leu/Ile в положениях 3 и 6 пептида [135]. Можно предположить, что наблюдаемый пик с  $m/z$  547.321 является  $w$ -ионом остатка Ile, но теоретическая масса этого фрагмента составляет 547.2875, и точность измерения масс не позволяет сделать такого соотнесения.



**Рисунок 3.** Вверху: схема образования  $w$ -ионов из соответствующих  $z$ -ионов. Внизу: фрагмент ДЗЭ спектра иона  $b_{10}^{2+}$  ( $m/z$  539.3001) бревинина 2Ra, полученный в ходе МС<sup>3</sup> эксперимента и

показывающий ключевые  $z$ - и соответствующие  $w$ -ионы для дифференцирования изомерных остатков Leu/Ile в положениях пептида 3 и 6. Разрешающая способность MS 100 000, MS/MS – 25 000. Точность измерения масс 1 ppm.

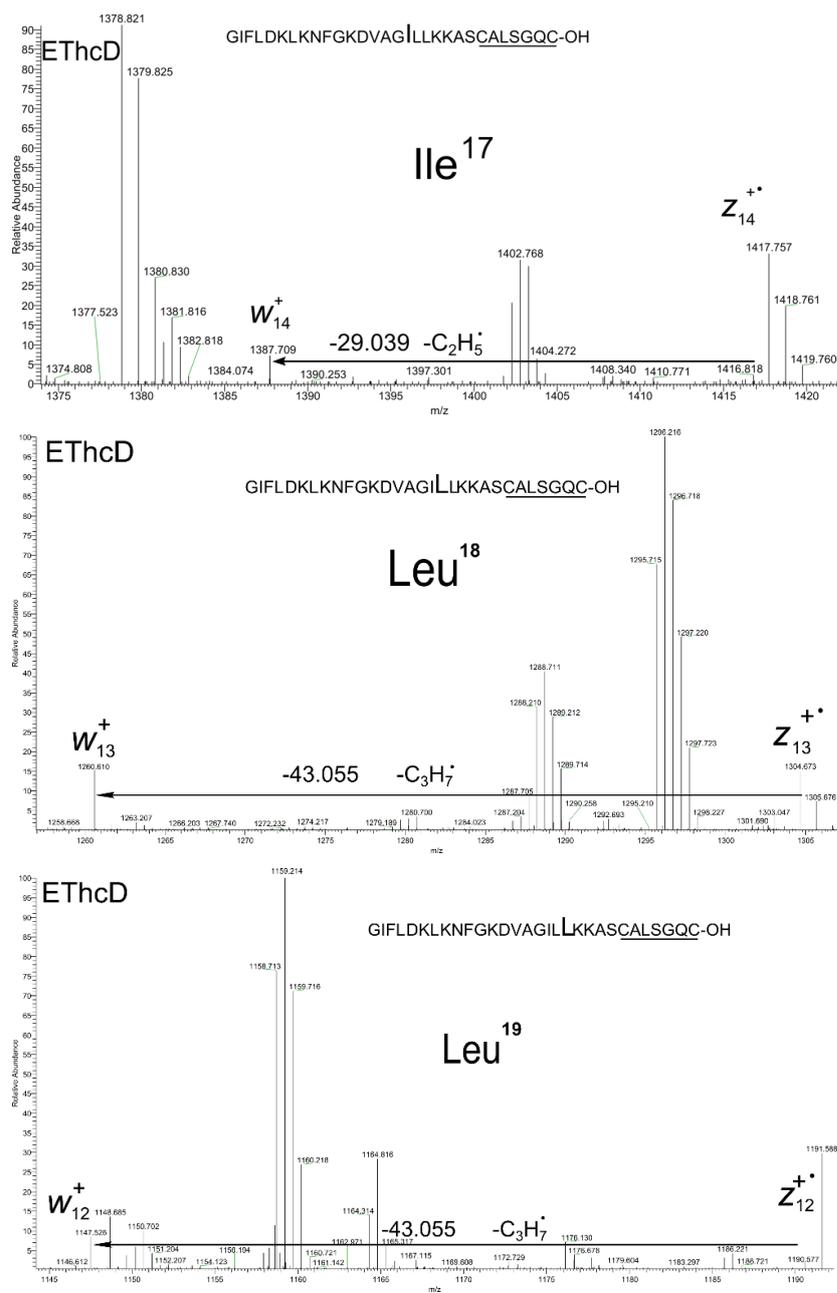
Однако в спектрах не всегда присутствуют нужные  $w$ -ионы. Предложенный Р. Зубаревым с соавторами в 2002 г. метод «горячей» ДЗЭ (HECD) [137,138] заметно повысил эффективность дифференцирования изомерных остатков Leu/Ile за счёт более высокой колебательной энергии исходных ион-радикалов, что приводит к большему количеству разрывов связей [139].

В 2014 г. А. Лебедевым с соавторами был предложен MS<sup>3</sup> подход, включающий последовательное применение ДПЭ и ДАСПЭ для дифференцирования изомерных остатков Leu/Ile в природных нетриптических пептидах [111]. Метод основан на выборе  $z$ -ионов с остатком Leu/Ile на  $N$ -конце с их последующей ДАСПЭ фрагментацией. Он был опробован на пептидах лягушки *Rana ridibunda* и показал 100% эффективность. Спектры ДПЭ-ДАСПЭ обычно содержали только целевые ионы-предшественники и ионы-продукты. Из спектров легко получить информацию для дифференцирования изомерных остатков [111]. Недостатком этого подхода является необходимость отбора каждого целевого  $z$ -иона для дальнейшей ДАСПЭ фрагментации. Позднее это было устранено при помощи метода EThcD, который успешно применялся при анализе как триптических [140,141], так и нетриптических пептидов [142].

Сначала, когда весь диапазон масс был разделён на несколько «полос» примерно по 600 Да, к каждой полосе можно было применить «широкополосную» активацию типа EThcD. Было показано, что подход HECD эффективен для пептидов, содержащих до 7 аминокислотных остатков [140]. Для более длинных  $z$ -ионов ключевые  $w$ -ионы появлялись лишь иногда, и требовалась дополнительная активация  $z$ -ионов. Использование EThcD для всего диапазона масс изучалось в работе Т. Самгиной с соавторами [142]. Наилучшие результаты были получены для родительских ионов с наивысшими возможными зарядовыми состояниями, что позволяло получить весь массив целевых  $z$ -ионов в различных зарядовых состояниях. Хотя спектры были не такими чистыми, как в подходе ДПЭ-ДАСПЭ [111], общий результат был схожим. Степень дифференцирования Leu/Ile была близка к 90%, в частности, было определено 60 из 66 остатков Leu/Ile в длинных (более 20 звеньев) пептидах из кожного секрета лягушки.

Кроме того, при определенных условиях становится возможно контролировать миграцию радикальных центров, в том числе она может быть полностью подавлена [143–145]. Этого можно добиться благодаря минимизации применяемой нормализованной энергии

соударений (NCE) в ДАСПЭ эксперименте. Поскольку реакции, инициируемые радикальным центром, требуют меньше энергии, образование  $w$ -ионов протекает гладко. Любое увеличение NCE запускает конкурирующие процессы, в том числе миграцию радикальных центров с последующим образованием соответствующих  $u$ -ионов (рис. 3). Чем выше NCE, тем ниже интенсивность  $z$ -ионов и выше интенсивность  $u$ -ионов. Это может привести к ошибкам, когда изомерные остатки Leu/Ile занимают соседние положения в последовательности. Минимизируя NCE, можно получить правильные результаты даже для сложных пептидов, содержащих несколько изомерных аминокислотных остатков подряд (рис. 4).



**Рисунок 4.** Фрагменты EThcD спектра иона  $MH_4^{4+}$  с  $m/z$  784,683 бревинина 2Rj, полученного на приборе Orbitrap Elite с разрешающей способностью 120 000, NCE 20, точность измерения

масс 1–3 ppm. Сверху вниз: идентификация  $\text{Leu}^{17}$  за счёт потери  $\text{C}_2\text{H}_5^+$  из иона  $\text{z}_{14}^{++}$ ;  $\text{Leu}^{18}$  за счёт потери  $\text{C}_3\text{H}_7^+$  из иона  $\text{z}_{13}^{++}$ ;  $\text{Leu}^{19}$  за счёт потери  $\text{C}_3\text{H}_7^+$  из иона  $\text{z}_{12}^{++}$ .

Кроме  $w$ -ионов, для определения изомерных остатков можно использовать характеристические  $d$ -ионы. Авторы [146] успешно дифференцировали остатки  $\text{Leu}/\text{Ile}$  в антимикробных пептидах из яда скорпиона *Isometrus maculatus* с помощью высокоэнергетической ДАС и дериватизации пептида по  $N$ -концу при помощи 4-(гуанидинометил)бензойной кислоты (Gmb).

### 1.3.2. Масс-спектрометрия с преобразованием Фурье для секвенирования коротких пептидов

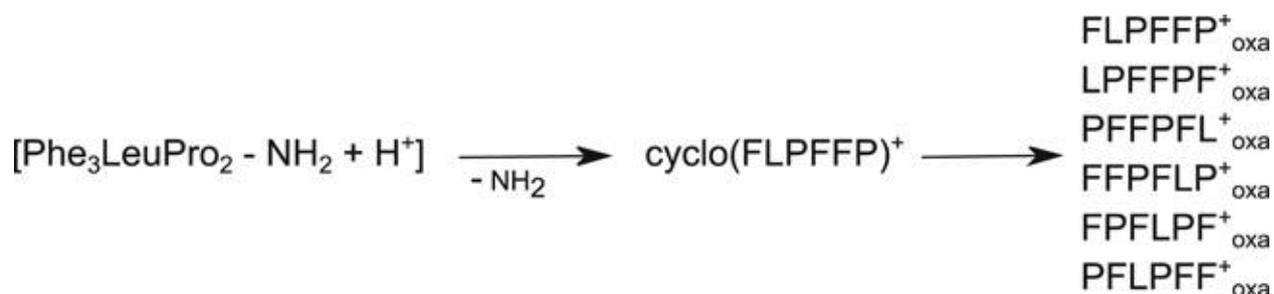
Триптофиллины, брадикинины и темпорины – три основных семейства пептидов лягушек, обладающих небольшой длиной. Триптофиллины (TRNs) – группа небольших пептидов, название которым дал присутствующий в структуре триптофан. Они присутствуют в секретах особей семейств *Hylidae* и у древних лягушек *Ascaphus truei*, что свидетельствует о зарождении экспрессии клонирующих триптофиллины генов ещё на ранних стадиях эволюции. Долгое отсутствие номенклатуры триптофиллинов породило путаницу в их названиях, один и тот же пептид порой имеет два или даже три наименования.

Т. Чен с соавторами [147] предложили классифицировать все известные триптофиллины в зависимости от имеющихся в их структурах консервативных мотивов. Они выделили три группы триптофиллинов: Т-1, 7-8-членные триптофиллины с мотивом, содержащим  $N$ -концевую пару  $\text{Lys-Pro}$ ,  $\text{Trp}^5$  и  $\text{Pro}^7$ , Т-2, содержащие 4-7 аминокислотных звеньев с мотивом  $\text{Pro-Trp}$  и Т-3 – пептиды, содержащие 13 аминокислот, девять из которых консервативны, а пять из них – пролины. Происхождение каждого конкретного триптофиллина отражено двумя первыми буквами в названии пептида. Например, триптофиллин  $\text{Lys-Pro-Hyp-Ala-Trp-Val-Pro}$ , выделенный из секрета толстой квакши (*Pachymedusa dactylos*), обозначен PdТ-1, где две первые буквы указывают вид животного, из чьего секрета был выделен пептид, а номер обозначает класс Т-1. Следующие TRNs из секрета той же квакши будут обозначаться как PdТ1-1; PdТ1-2 и т.д. Структурно схожие с TRNs пептиды, не попадающие ни в один из трех выделенных Ченом классов, называют триптофиллин-родственными или триптофиллин-связанными пептидами.

Физиологическая роль TRNs до конца не выяснена, хотя известно, что многие из них обладают разнообразной нейроактивностью [26].

Первичная структура большинства известных к настоящему времени триптофиллинов за редким исключением была установлена биохимическими методами (секвенированием по Эдману или клонированием кДНК), а не масс-спектрометрическими методами. TRNs

неудобны для масс-спектрометрического исследования, поскольку в их ДАС спектрах часто присутствуют ионы фрагментации циклического перестановочного *b*-иона, образованного по типу «голова к хвосту» и являющегося наиболее интенсивным в спектре. На рисунке 5 показано образование циклического *b*<sub>6</sub> иона триптофиллин-родственного пептида FLPFFP-NH<sub>2</sub> (HaTL-5), выделенного из секрета *Hyla arborea*, с дальнейшим образованием из него шести линейных изомерных исходному пептидов, C-концевая аминокислота в котором находится в оксазолоновой форме[148]:

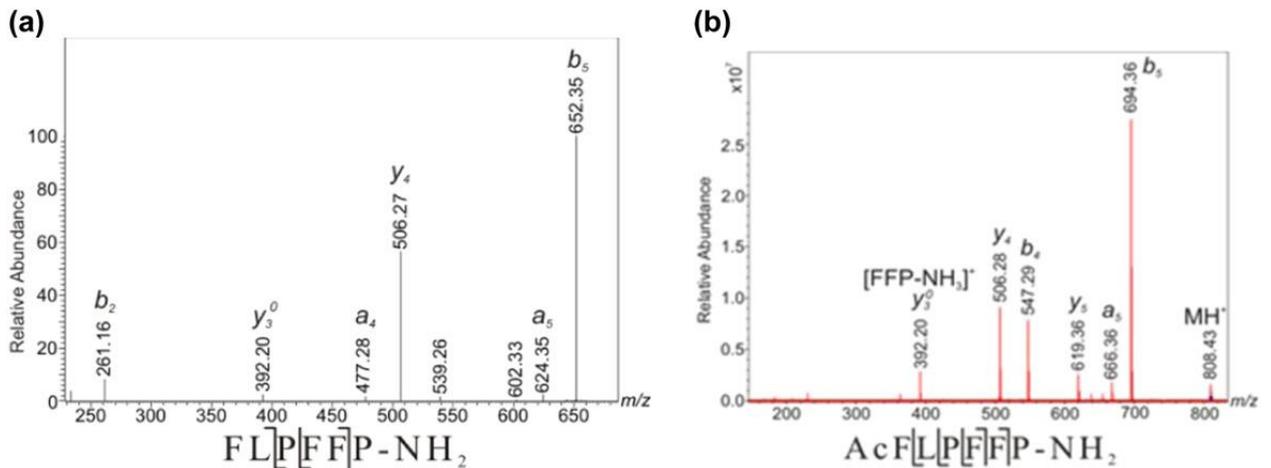


**Рисунок 5.** Циклизация триптофиллин-родственного пептида FLPFFP-NH<sub>2</sub> с образованием шестилинейных изомеров

Все эти изомерные ионы были зарегистрированы в спектре, как и серии их характеристических ионов. Таким образом, получить достоверную последовательность коротких ТРНс масс-спектрометрическим путем – непростая задача.

В работе [148] на пептиде FLPFFP-NH<sub>2</sub>, выделенном из секрета *Hyla arborea*, были изучены три химические N-концевые модификации, способные предотвратить циклизацию родительского иона в газовой фазе масс-спектрометра. Структура пептида FLPFFP-NH<sub>2</sub> (HaTL-5) была получена при помощи деградации по Эдману, а пептид был синтезирован в GenScriptCorporation (США) по стандартной методике, использующей Fmoc-метод и L-аминокислоты. Химические модификации включали в себя ацетилирование, сульфобензоилирование N-концевой аминогруппы и её модификацию в 2,4,6-триметилпиридиниевое производное путём взаимодействия с тетрафторборатом 2,4,6-триметилпиридиния [149], то есть введение фиксированного N-концевого положительного заряда. Все эксперименты проводились на гибридном 7 Тл ИЦР ПФ приборе с линейной ловушкой (LTQ-FT, ThermoElectron) с нано-ЭРИ источником (ProxeonBiosystems). Разрешение составляло 100 000 для МС и 50 000 для МС/МС.

Критериями сравнения трёх процедур были простота модификации NH<sub>2</sub>-группы и её эффективность для МС-секвенирования. Наилучшей модификацией с этих позиций оказалось ацетилирование: реакция проходила при комнатной температуре за час и не требовала дополнительной очистки – избыток реагентов удалялся лиофилизацией образца. На рисунке 6 показаны спектры, демонстрирующие эффективность ацетилирования.



**Рисунок 6.** ЭРИ-ИЦР ПФ ДАС-спектры пептида FLPFFP-NH<sub>2</sub>: (а) интактный пептид, (б) ацетилированный пептид

Триптофиллины и триптофиллин-родственные пептиды из секретов близкородственных видов древесных квакш *Hyla arborea* и *Hyla savignyi* были описаны практически одновременно двумя группами авторов: работы [150] и [87] вышли в одном номере журнала. Экспериментальные данные по *Hyla savignyi* получены на гибридном масс-спектрометре Finnigan LTQ FT Ultra (Thermo Scientific GmbH), с линейной квадрупольной ионной ловушкой и ИЦР ПФ, оборудованным магнитом 6 Тл и нано-ЭРИ. Разрешение на  $m/z$  400 составляло 100 000, точность измерения масс в пределах  $\pm 2.0$  ppm [87,88]. Данные по *Hyla arborea* получены на гибридном масс-спектрометре LTQ-FT Ultra (Thermo Scientific GmbH), с линейной квадрупольной ионной ловушкой и ИЦР ПФ, оборудованным магнитом 7Тл и нанопоточной хроматографической системой 1100 (Agilent). Разрешение составляло 100 000 для МСи 50 000 для МС/МС, как и в работе [149].

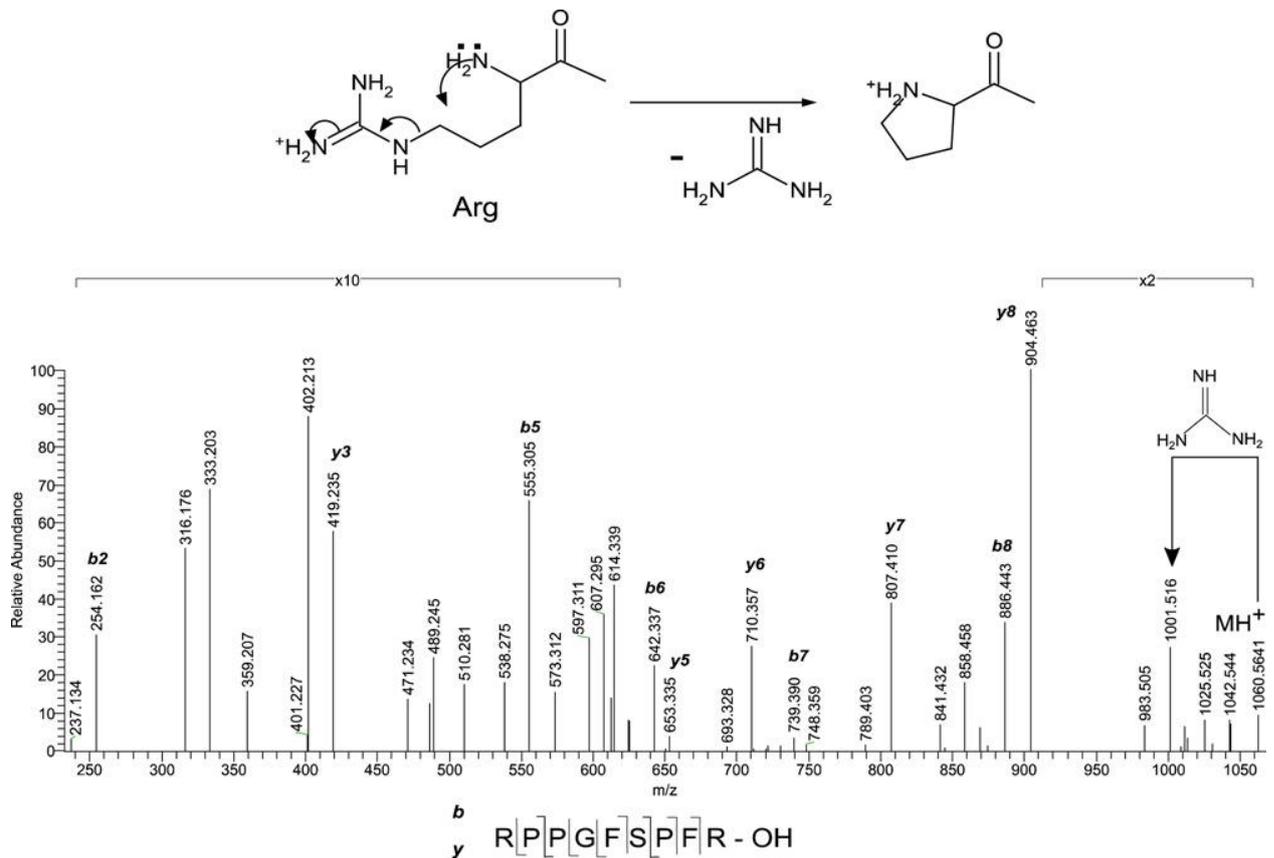
Последовательности пептидов *Hyla savignyi* были определены комбинацией ручной интерпретации суммы спектров ДАС и ДЗЭ. Для подтверждения правильности результатов был проведен синтез всех обнаруженных в секрете пептидов. Их тандемные спектры совпали с аналогичными спектрами исходных пептидов [87,88]. В результате, были установлены последовательности восьми триптофиллинов, содержащих 4-5 аминокислот [87], и трех савиньинов. Савиньин-1 и савиньин-2 – кислотные спейсеры препептидов: DDSEEEVE-OH, НурЕЕVEEERJK-OH, где J – неидентифицированная пара Leu/Ile, Нур – гидроксипролин. Отнесение изомерных Leu/Ile в работах [87,88] не проводилось из-за преимущественного однозарядного состояния протонированных пептидных ионов. Ручная интерпретация сектров пептидов, не содержащих Pro, савиньинон-1 и савиньинон-2, дала хорошие результаты, в то время как секвенирование савиньинон-3 (GJJDPJTGJVGGJJ-NH<sub>2</sub>) с Pro<sup>5</sup> в последовательности было затруднено вторичной фрагментацией интенсивного y<sub>10</sub> иона.

В секрете близкородственной квакши *Hyla arborea schelkownikowi* были установлены последовательности 10 пептидов, лишь один из которых строго соответствовал триптофиллину из семейства T2 – FLPWL-NH<sub>2</sub>, и назван в соответствии с предложенной номенклатурой (HaT2). Остальные пять оказались триптофиллин-родственными пептидами и получили расширение TL: HaTL-1, HaTL-2, HaTL-3, HaTL-4, HaTL-5 [150].

Три пептида в кожных секретах *Hyla arborea schelkownikowi* и *Hyla savignyi* совпали по точной массе. Это два изомера с ММ 674.4023 Да: IIPFW-NH<sub>2</sub> [150] и JJPFW-NH<sub>2</sub> [87]; FLPWL-NH<sub>2</sub> [150] и FJPWJ-NH<sub>2</sub> [87], а также пептид с ММ 1392.83 Да: GJJDPJTGJVGJJ-NH<sub>2</sub> (J-Leu/Pe), названный арбореин-1 [150] и савиггин-3 [87].

Семейство брадикининов, или брадикинин-родственные пептиды (БРП), является модифицированными молекулами пептида, содержащего 9 аминокислотных звеньев, брадикинина, с тремя остатками пролина в последовательности: RPPGFSPFR-OH. Секвенирование пролинсодержащих пептидов связано с некоторыми трудностями из-за вторичной фрагментации  $\gamma$ -ионов, образующихся при разрыве амидных связей N-Pro. Часто в спектрах ДАС или ДАСПЭ наблюдается характерный выброс (59.048 Да) молекулы нейтрального гуанидина (NH<sub>2</sub>-C(NH)-NH<sub>2</sub>) из боковой цепи N-концевого аргинина. Это характерно для всех брадикининов и их C-удлинённых копий (рис. 7). Процесс включает нуклеофильную атаку N-концевой аминогруппы на  $\delta$ -атом углерода боковой цепи аргинина, несущего частичный положительный заряд, за счёт индукционного действия заряженной гуанидиновой группы протонированного аргинина. В результате образуется ион пирролидиния. Очевидно, для N-удлинённых копий брадикинина этот процесс подавлен из-за меньшей нуклеофильности амидного азота [53].

Сорок четыре брадикинин-родственных пептида были обнаружены в секретах пяти ранидных лягушек [151]. Целевой поиск проводили для спектров, зарегистрированных на приборе ИЦР ПФ с магнитом 7 Тл, строя масс-хроматограммы характеристических ионов консервативных фрагментов. Эта методика широко используется в масс-спектрометрии объектов окружающей среды [152]. Так для обнаружения БРП с N-концевыми расширениями в спектрах ДАС использовали  $u_3$ -ионы с  $m/z$  419.2378 – консервативный C-концевой фрагмент, содержащий последовательность PFR [153]. Пептиды с C-концевыми расширениями детектировали по  $c_4$ -иону с  $m/z$  425.2618, присутствующему в спектрах ДЗЭ. Использование двух пар ионов ( $c/z$  и  $b/y$ ) в качестве «меток последовательностей» может значительно увеличить эффективность поиска различных БРП. Следует отметить, что точное измерение масс позволило надёжно идентифицировать изобарный гидроксипролин в пептиде RPHypGFSPFRIAPAST ((Hyp<sup>3</sup>)Br) [135].



**Рисунок 7.** ДАС спектр брадикинина ( $MH_2^{2+}$ ,  $m/z$  530.788), демонстрирующий потерю гуанидина из *N*-концевого остатка аргинина. Разрешение 100 000 в режиме MS и 25 000 в режиме MS/MS. Точность измерения масс 4 ppm [53].

Среди идентифицированных брадикининов [151] наибольший интерес представляют брадикинин птиц, (Thr<sup>6</sup>, Leu<sup>8</sup>)Br, и брадикинин рыб (Arg<sup>0</sup>, Trp<sup>5</sup>, Leu<sup>8</sup>)Br. Первый является агонистом орнитокининового рецептора, одного из двух известных кининовых рецепторов немлекопитающих. Он был выделен и охарактеризован в 1997 г. [154]. До этого брадикинин птиц никогда не обнаруживался в кожных секретах амфибий. Его присутствие подтверждает гипотезу Т. Чена и Ц. Шоу [153,155], касающуюся защитных функций БРП от потенциальных хищников, таких как рыбы и птицы.

Если брадикинины входят в состав кожных секретов всех ранидных лягушек, то темпорины присутствуют преимущественно в секрете бурых лягушек. Название темпорины произошло от названий одного из видов этих лягушек, *Rana temporaria*. Темпорины – короткие (10-17 аминокислотных остатков) пептиды с широкой антимикробной активностью с консервативным Pro<sup>3</sup>. Однако их не так просто секвенировать, используя только ДАС. Значительно более эффективна комбинация комплементарных спектров ДАС и ДЗЭ/ДПЭ [53].

С помощью нано-ИЭР-ИЦР ПФ Зимеринк с соавторами исследовали три коротких пептида из образца яда гадюки *Bothrops moojeni*, проявляющие активность ингибитора ангиотензинпревращающего фермента, ACEiP [156]. Образец яда фракционировали с

помощью ВЭЖХ и тестировали на активность. Активные фракции дополнительно разделяли и анализировали с помощью ИЦР ПФ. Двухзарядные ионы активных пептидов с массой около 1,4 кДа подвергали ДАС. Установленные последовательности – ZKWPPGKVPP, ZKWPRPGPEIPP и ZNWPRPGPEIPP, где Z – пироглутаминовая кислота. Разница между рассчитанной и измеренной массами составляла менее 2 ppm в режиме MS и менее 5 ppm в режиме MS/MS. Структуры, содержащие много остатков пролина в последовательности, включая PP-фрагмент на C-конце, типичны для ACEiP. Несмотря на достаточно высокую точность измерения масс, оставались некоторые неясности. Так, 10-й аминокислотный остаток был определён как изолейцин исключительно на основании гомологии с другими структурами ACEiP. Фрагмент ZN в ZNWPRPGPEIPP изомерен последовательности ZGG, и был выбран исключительно из-за отсутствия в спектрах  $b_2$ -ионов, соответствующих остатку ZGG.

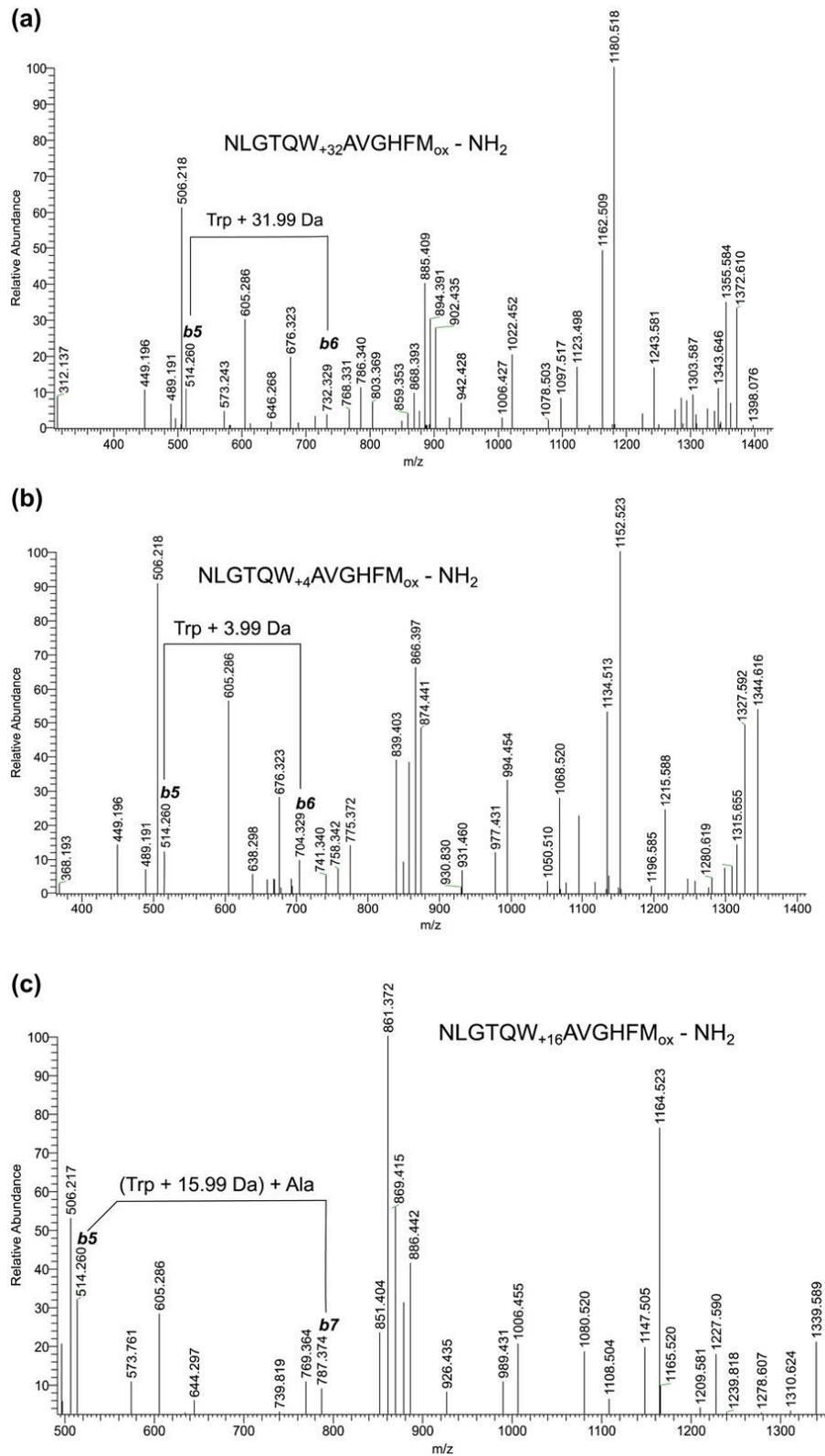
Времяпролётные приборы, приборы на основе ИЦР и орбитальной ловушки использованы для секвенирования коротких пептидов в образце перикардального органа *Homarus americanus* [157]. С помощью Orbitrap авторы обнаружили 50 нейропептидов, при этом только шесть из них наблюдались в эксперименте с ИЦР ПФ. Другие приборы показали несколько лучшие результаты, хотя причина этого в статье не обсуждалась. Она может заключаться в более низкой скорости сбора данных инструментов ИЦР ПФ. Также более высокая скорость регистрации данных в приборах Orbitrap [28] и более эффективный перенос ионов от источника к ловушке позволили секвенировать 76 пептидов против 21 пептида, полученного в ходе анализа на приборе ИЦР ПФ, в образце секрета *Rana temporaria*.

### 1.3.3. Пептиды с внутримолекулярной дисульфидной связью

Довольно часто природные пептиды содержат дисульфидные связи, образованные за счёт взаимодействия боковых цепей двух остатков цистеина в молекуле. Так, пептиды кожного секрета земноводных содержат C-концевой дисульфидный цикл, «Rana box» Размер цикла и общее количество аминокислот в последовательности зависят от семейства пептидов. Наличие такого фрагмента заметно усложняет секвенирование интактных пептидов с помощью масс-спектрометрии, так как связь S-S обычно не расщепляется в условиях ДАС, и последовательность внутри цикла остаётся неизвестной. Использование ДЗЭ может быть более эффективным. Однако успех зависит от зарядового состояния протонированной молекулы и структурных особенностей «Rana box», то есть наличия основных аминокислот, таких как Lys и Arg. Классическая процедура секвенирования дисульфидных пептидов требует раскрытия этой связи либо окислением надкислотами, либо восстановлением с последующим алкилированием.

Наиболее популярный подход включает в себя восстановление S-S связей с помощью дитиотреитола (ДТТ) с последующим карбамидометилированием свободных тиольных групп йодацетамидом (ИАА) [158]. Менее популярными восстановителями являются трис(2-хлорэтил)фосфат, работающий при нейтральных и слабокислых значениях pH [33,159], и  $\beta$ -меркаптоэтанол. Стоит также упомянуть, что *N, N*-диметил-2-хлорэтиламин [160] превращает остатки Cys в их диметилловые аналоги Lys. Ряд других реагентов упоминается в работах [161,162]. Достаточно успешным подходом является окисление S-S связей надмуравьиной кислотой [163]. В работах Т. Самгиной с соавторами используются оба этих подхода для *de novo* секвенирования длинных природных пептидов с S-S связью [135,164–166]. На ИЦР ПФ приборе с магнитом 7 Тл с разрешающей способностью 100 000 в режиме МС и 50 000 в режиме МС/МС были исследованы две серии образцов. При помощи ДАС и ДЗЭ спектров интактных, а также окисленных и восстановленных пептидов, были секвенированы 116 нетриптических пептидов, включая 39 длинных (17-46 аминокислот) пептидов, содержащих S-S-связь, выделенных из секретов четырёх ранидных лягушек (*R. ridibunda*, *R. arvalis*, *R. lessonae* и *R. temporaria*). Последовательности сорока из этих пептидов были установлены впервые.

Окисление надмуравьиной кислотой может привести к побочному окислению остатков метионина и триптофана [163]. Поскольку весь образец секрета подвергается восстановлению или окислению, все пептиды в образце могут быть каким-то образом модифицированы. Однако при использовании точного измерения масс проблема легко решается, так как известны точные массы продуктов [165]: кинуренин (Trp+3.99 Да); гидрокситриптофан (Trp+15.99 Да); *N*-формилкинуренин (Trp+31.99 Да); метионинсульфоксид (Met+15.99 Да); метионинсульфон (Met+31.99 Да) (рис. 8). Несмотря на малое количество информации об этих молекулах в анализируемой смеси, общие фрагменты в спектрах этих пептидов помогают установить взаимосвязь между этими молекулами. Различия масс ионов-предшественников позволяют предположить присутствие остатка Trp, а сдвиги некоторых фрагментных ионов в соответствующих сериях указывают на положение модифицированного остатка.



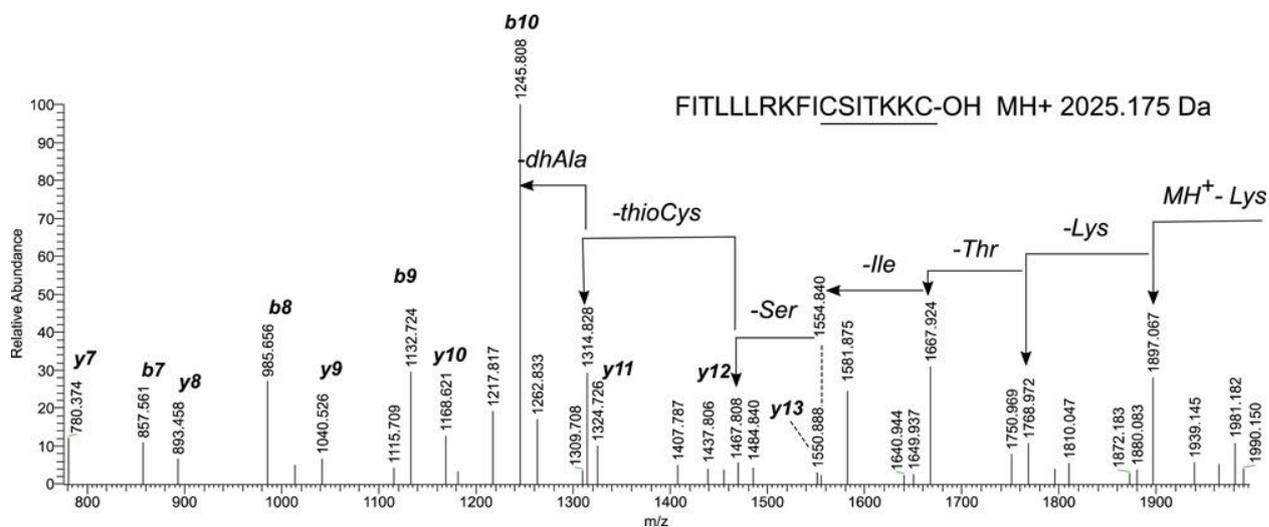
**Рисунок 8.** ДАС спектры бомбезина  $\text{NLGTQWAVGHFM}_{\text{ox}}\text{-NH}_2$  и продуктов окисления из кожного секрета *Rana lessonae* после раскрытия цикла S-S с помощью надкислот: (a) *N*-формилкинуруенин, родительский ион  $\text{MH}_2^{2+}$  с  $m/z$  704.328, (b) кинуруенин, родительский ион  $\text{MH}_2^{2+}$  с  $m/z$  690.328, (c) гидрокситриптофан, родительский ион с  $m/z$  696.330. [165].

Как упоминалось выше, обычно раскрытие внутримолекулярных дисульфидных циклов в ходе протеомного анализа включает восстановление дитиотреитолом с последующей дериватизацией свободных тиольных групп йодацетамидом. Эта модификация не влияет на фрагментацию линейной части пептида. Окисление S-S связи с образованием фрагмента SO<sub>3</sub>H добавляет в структуру пептида два дополнительных протона. Они могут вести себя как мобильные протоны в условиях дефицита протонов. Эта особенность в некоторой степени меняет процессы фрагментации как на линейных, так и на циклических участках исходной молекулы. Применение обеих методик: восстановления и окисления, даёт дополнительную информацию об исходной аминокислотной последовательности [164].

Для обработки свободных тиольных групп цистеинов после восстановления S-S связи может применяться не только йодацетамид. Было проведено сравнение трёх стандартных реагентов: IAA, йодоуксусной кислоты (IAAC) и N-фенилмалеимида (NPM) и девяти новых реагентов, производных NPM и IAA, для обработки SH-групп [167]. Сравнение проводили по их эффективности для секвенирования бревинина 1E (24 а.к., 2675 Да) и бревинина 2Ec (34 а.к., 3517 Да). Две серии образцов (по 12 для каждого пептида) исследовали в режимах ДАС и ДЗЭ. ДАС эксперименты проводились на приборе LTQ-FT Ultra с магнитом 7 Тл (Thermo Fisher Scientific, Бремен, Германия) и наноВЭЖХ. ДЗЭ спектры регистрировали на приборе Apex Ultra Q-FTMS (Bruker Daltonics, Бремен, Германия) с магнитом 12 Тл. Спектры регистрировали в режиме прямого ввода: для бревинина 1E четырёхзарядные ионы фрагментировали при помощи электронов с энергией 1 эВ, длительность импульса составляла 100 мс, для бревинина 2Ec пятизарядные ионы фрагментировали при энергии электронов 1 эВ с длительностью импульса 75 мс. Хотя ни одна из меток не обеспечила 100% покрытие последовательности, для двух пептидов при одном вводе было достигнуто покрытие последовательности более 90%, и, что особенно важно, в обоих случаях было получено 100% покрытие последовательности внутри C-концевого дисульфидного цикла. Этими эффективными метками были йодоуксусная кислота; 2,5-диметилфенилмалеимид; 2,6-диметилфенилмалеимид и N-бензилмалеимид. Наиболее эффективный N-бензилмалеимид продемонстрировал на 20% большее покрытие последовательности, чем традиционный йодацетамид [167]. Положительное влияние новых меток на покрытие сиквенса может быть связано с взаимодействием введённых функциональных групп с другими участками молекулы пептида. Существенным недостатком, ограничивающим их применение для дериватизации, является более низкая растворимость в воде по сравнению с йодацетамидом и необходимость очистки реакционной смеси от используемого реагента. Подходящим растворителем для меток на основе малеимида является смесь вода/ацетонитрил в соотношении 1:1 по объёму.

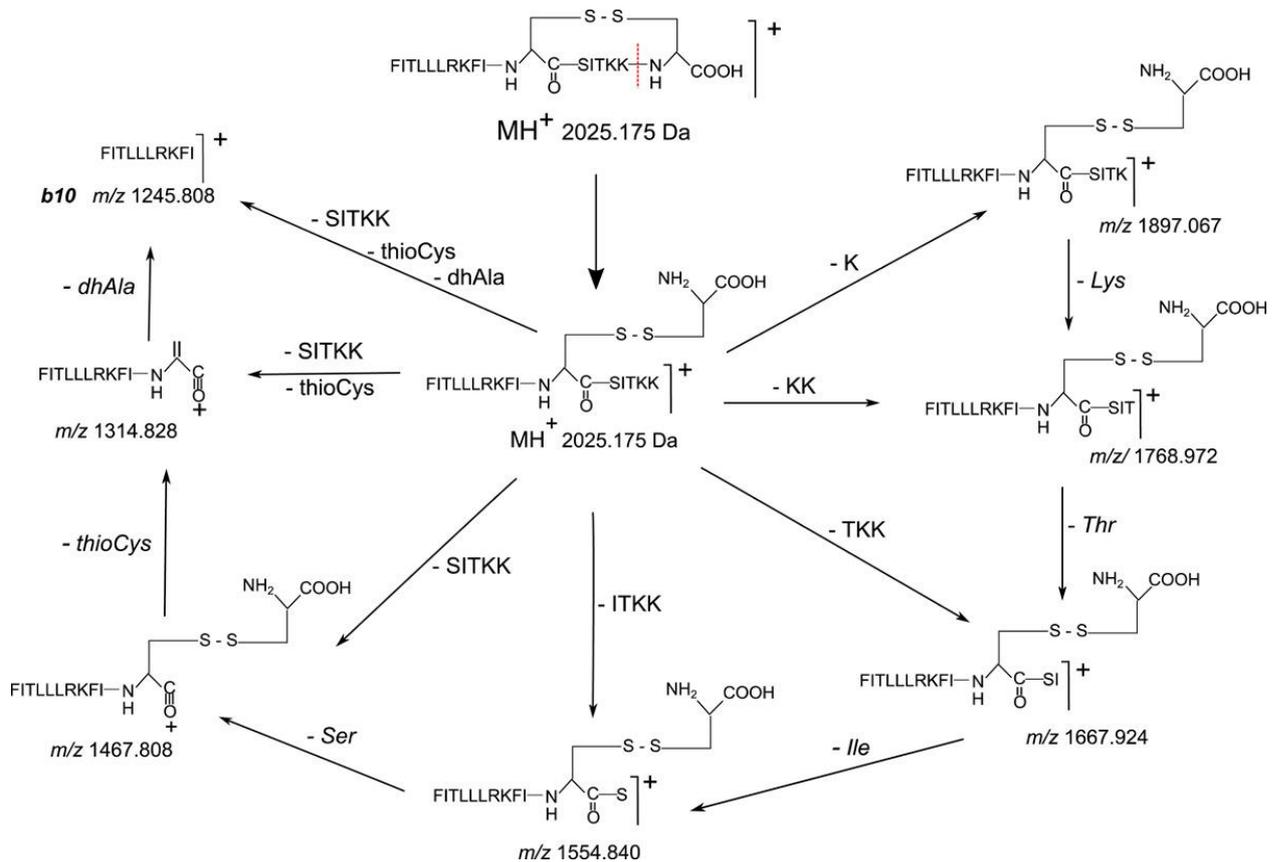
В некоторых случаях последовательность внутри дисульфидных циклов устанавливалась полностью или частично без какой-либо дериватизации исходного пептида. Высокоэнергетическая ДАС в экспериментах МАЛДИ-ВП/ВП обычно обеспечивает большее количество путей фрагментации, в том числе внутри дисульфидных циклов. С. Такур и П. Баларам [168] наблюдали ряд фрагментных ионов, связанных с участком последовательности пептидов моллюсков-конусов, контифанов, внутри интактного C-концевого дисульфидного цикла. Авторы предложили несколько механизмов разрывов связей в цикле. Определённые последовательности внутри дисульфидных циклов могут быть получены с помощью приборов с преобразованием Фурье в режимах ДАС, ДАСПЭ или ДПЭ. Особенно трудным случаем был анализ пептида с молекулярной массой более 7 кДа с тремя дисульфидными связями из яда тexasской коралловой змеи, описанный в работе К. Медзиградски и К. Болен [33]. Полная последовательность пептида, содержащего 60 аминокислотных звеньев, была успешно установлена при помощи ДАС, ДАСПЭ и ДПЭ. В основном с использованием спектров восстановленного пептида, поскольку только 2 C-концевых и 6 N-концевых остатков находились за пределами дисульфидных циклов. Положение S-S связи между остатками Cys<sup>7</sup>-Cys<sup>58</sup> было установлено благодаря сдвигу ионов  $u_3$ - $u_5$  на 1203 Да. В спектре интактного токсина наблюдался ряд пиков, образованных за счёт вторичной фрагментации исходного иона. В ходе этой диссоциации хорошо фрагментировала связь Asp-Pro внутри цикла. Набор этих фрагментов позволил определить два остатка цистеинов, образующих конкретную дисульфидную связь.

Ряд необычных фрагментных ионов, соответствующих аминокислотной последовательности внутри C-концевого дисульфидсодержащего цикла нетриптических пептидов в условиях низкоэнергетической ДАС, наблюдался в ДАС спектрах семейств бревининов 1, бревининов 2, эскулентинов 2 и ранатуеринов 2 [169]. На рисунке 9 показан ДАС спектр дважды протонированного бревинина 1Ta из кожного секрета лягушки *R. temporaria*. Пики этих необычных ионов с высокими значениями  $m/z$  соответствуют последовательной потере Lys; Lys; Thr; Leu/Ile; Ser непосредственно из  $(M+2H)^{2+}$  бревинина 1 Ta и соответствуют последовательности внутри «Rana box» этого пептида: SITKK (рис. 9). Таким образом, после разрыва пептидной связи Lys<sup>16</sup>-Cys<sup>17</sup> внутри C-концевого цикла с сохранённой связью S-S следует фрагментация с участием пептидных связей C-концевых остатков. Так происходит изомеризация исходной 17-звенной протонированной пептидной молекулы бревинина 1Ta FITLLRKFICSITKKC-OH в 16-звенный пептид с остатком цистина в 11-м положении цепи FITLLRKFIC<sup>C</sup>SITKK.



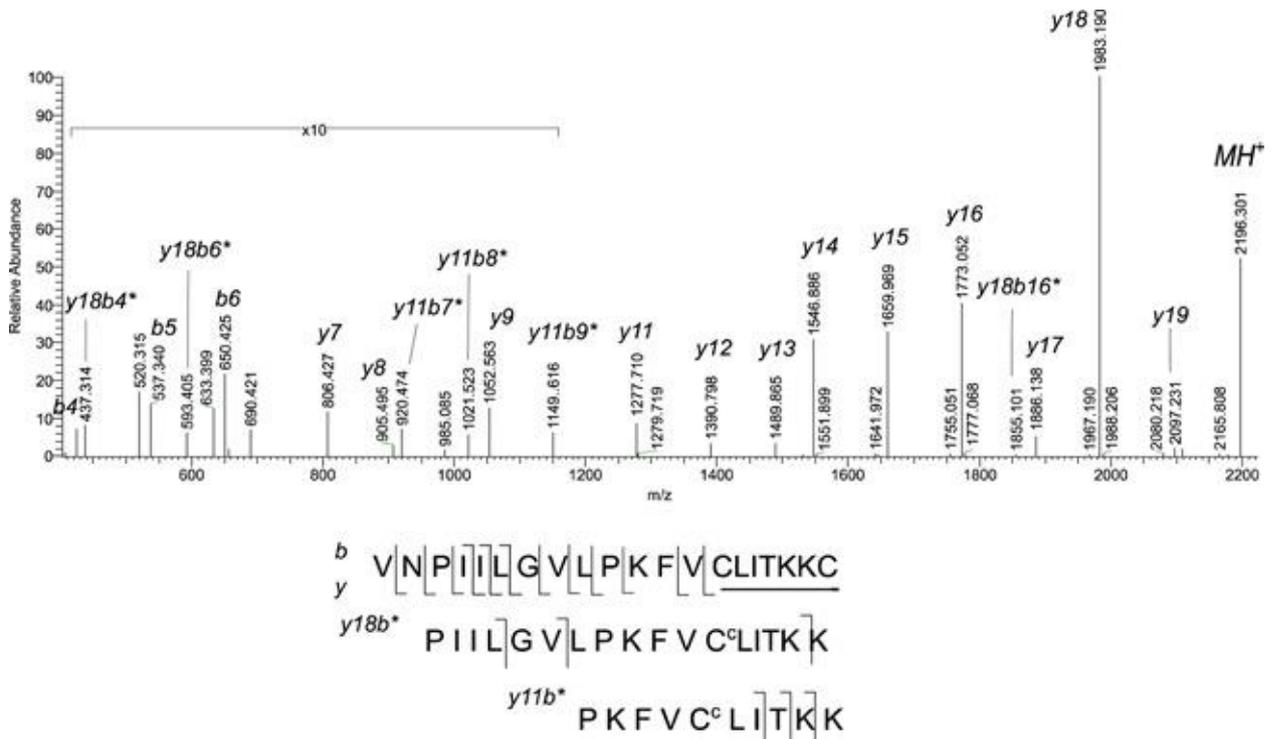
**Рисунок 9.** ДАС спектр  $(MH_2)^{2+}$  бревинина 1Та с  $m/z$  1013,094. Разрешающая способность 100 000 в режиме MS, 25 000 в режиме MS/MS, точность измерения масс 3 ppm [169].

Последующая C-концевая фрагментация образовавшегося изомерного линейного иона не затрагивает дисульфидную связь и в конечном итоге приводит к потере молекулы thioCys и образованию фрагмента дегидроаланина (dhAla). Дальнейшая потеря этого фрагмента приводит к образованию b-иона с  $m/z$  1245.80 (потеря 69,020 Да). Схема 1 иллюстрирует указанные реакции фрагментации. В случае монопротонированных пептидов наблюдается только разрыв связи S-S. Но в случае ионов  $(M+2H)^{2+}$  или  $(M+3H)^{3+}$  может быть расщеплена как связь CO-NH, так и S-S [170]. Бревинин 1Та содержит остаток Arg в линейной части молекулы, протон прочно связан с этим Arg и вряд ли может быть подвижным. Дважды протонированный бревинин 1Та, содержащий только один потенциально подвижный протон, демонстрирует фрагментацию с разрывом амидной связи Lys-Cys на C-конце, так как остаток лизина обладает более высоким сродством к протону, чем остатки, содержащие связь S-S. Наличие двух основных аминокислотных остатков в «Rana box» существенно усиливает эту фрагментацию за счёт увеличения вероятности протонирования по крайней мере одного из них. Полная последовательность внутри дисульфидного цикла была установлена для пептидов, содержащих C-концевой фрагмент Lys-Lys-Cys [169]. Следует отметить, что дефицит подвижных протонов очень важен и наиболее эффективное секвенирование последовательности внутри цикла было осуществлено для менее протонированных молекул пептидов.



**Схема 1.** Пути фрагментации внутри дисульфидного цикла бревинина 1Тa, FITLLLRKFISITKKC-S-OH, в ходе диссоциации, активированной соударениями.

Присутствие остатка Pro в линейной части пептидов, содержащих S-S связь, приводит к преимущественному разрыву связи N-Pro с образованием наиболее интенсивного в спектре у-иона даже в случае дефицита протонов. Однако этот ион может далее фрагментировать по описанному выше механизму с расщеплением пептидных связей внутри дисульфидного цикла. Этот процесс становится особенно выраженным для пептидов с C-концевым фрагментом Lys-Lys-Cys [53,169]. На рис. 10 представлен ДАС спектр бревинина 1Т (ММ 2195.30 Да), содержащего Pro<sup>3</sup> и Pro<sup>10</sup> в последовательности. Выделенные серии ионов отражают *b*- и *y*-ионы, образованные за счёт прямой фрагментации иона бревинина 1Т МН<sub>3</sub><sup>3+</sup>, серию *y<sub>n</sub>b<sup>\*</sup>*-ионов, образованных в ходе вторичной фрагментации с разрывом пептидной связи внутри дисульфидного цикла, а также серии ионов за счёт вторичной фрагментации интенсивных *y*<sub>18</sub>- и *y*<sub>11</sub>-ионов, образующихся при разрыве пептидных связей Pro<sup>3</sup> и Pro<sup>10</sup>. Видно, что раскрытие C-концевого дисульфидного цикла пролинсодержащих пептидов в условиях ДАС происходит только как вторичная фрагментация ионов *y*-серии, что схоже с образованием других внутренних ионов [171]. Такая фрагментация позволяет получить информацию о последовательности внутри дисульфидных циклов интактных пептидов.



**Рисунок 10.** ДАС спектр  $MH_3^{3+}$  иона бревинина 1Т ( $m/z$  732.7748):  $b/y$  – ионы прямой фрагментации  $MH_3^{3+}$ ;  $y18b^*$ ,  $y11b^*$  – ионы вторичной фрагментации ионов  $y18, y11$ ,  $C^c$  – остаток цистина [53].

Ряд исследований дисульфидсодержащих пептидов включал в себя работы о нейропептидах крабов. Гипергликемический гормон ракообразных из *Cancer borealis*, состоящий из 72 аминокислотных остатков с дисульфидными связями, был секвенирован с использованием комбинации подходов «снизу вверх» и «сверху вниз». Были применены МАЛДИ-ИЦР ПФ, *de novo* секвенирование триптических пептидов с помощью ИЭР-Q-ВП и анализ интактных молекул с помощью ВЭЖХ-ИЦР ПФ и различных методов фрагментации [172]. Аналогичная стратегия позволила секвенировать 122 пептида, 42 из них *de novo*, принадлежащих к нейропептидам *Carcinus maenas* [173], и 142 нейропептида *Cancer borealis*, 35 из них *de novo* [174]. Для секвенирования использовали восстановительное метилирование и ацетилирование. Этерификация и дезамидирование с последующим метилированием позволили локализовать кислотные и амидные остатки оркокининов, нейропептидов крабов, изученных с помощью МАЛДИ-ИЦР ПФ [85]. В работе [175] было выполнено *de novo* секвенирование аллатостатинов в образцах *Cancer borealis* и *Homarus americanus*, где в качестве матрицы в экспериментах МАЛДИ использовали ДНВ. Чтобы получить больше информации о фрагментации пептидов, были проведены дополнительные ВП-ВП эксперименты с СНСА и восстанавливающей матрицей DAN. В экспериментах с ВЭЖХ-фракционированными образцами перикардиальных органов *C. Borealis* ион изучаемого пептида продемонстрировал увеличение массы на 2 Да при использовании DAN, что

свидетельствует о наличии одной дисульфидной связи в нём. Для подтверждения *N*-концевой последовательности использовали дериватизацию формальдегидом. Таким же образом исследовали образцы *H. americanus*. Был синтезирован и фрагментирован один из *de novo* секвенированных пептидов, SYWKQCAFNAVSCF-NH<sub>2</sub>, из образца *H. americanus*, его фрагментация подтвердила последовательность пептида. Этот конкретный пептид был также идентифицирован в перикардиальном органе и рото-желудочном ганглии *C. Borealis* при прямом анализе тканей.

Интересный МС подход для идентификации нейропептидов ракообразных, принадлежащих к семейству гипергликемических гормонов (СНН), участвующих в регуляции различных физиологических процессов был разработан в работе Ч. Джиа с соавторами [176]. Пептиды-кандидаты искали в экстрактах тканей с помощью МАЛДИ-ВП/ВП с последующим анализом методом «снизу вверх» при помощи ВЭЖХ-ИЭР-Q-ВП. Поиск по базам данных и *de novo* секвенирование выполняли с помощью анализа методом «сверху вниз», для которого спектры регистрировали на приборе ИЦР ПФ, используя в качестве методов фрагментации ДАС и ДЗЭ, и анализа методом «снизу вверх», в котором спектры регистрировали на приборе Orbitrap Elite, а в качестве методов фрагментации использовали ДАСПЭ и ДПЭ. Шесть нейропептидов ракообразных *Callinectes sapidus* и *Cancer borealis*, относящихся к семейству СНН и обладающих массами 8-9.5 кДа, включая два новых пептида с большим количеством дисульфидных связей и ПТМ, были полностью *de novo* секвенированы без опоры на данные о геноме [176]. Исследование нейропептидов *Panulirus interruptus* с использованием того же подхода позволило секвенировать девять новых пептидов с массами 0.9–8.2 кДа [177]. *N*-концевое диметилирование позволило секвенировать 52 нейропептида в мозге омара [72].

Аналогичным образом исследовали эндогенные нейропептиды с внутримолекулярными дисульфидными связями в мозге крыс [178]. Х. Ю с соавторами идентифицировали 43 пептида, 35 последовательностей были установлены впервые. Эта группа дополнительно разработала методику «сверху вниз» для одновременного анализа микропротеинов и эндогенных пептидов, включающую в себя анализ при помощи МС ПФ и нанокапиллярную ВЭЖХ высокого разрешения [179]. Поиск в базах данных спектров ДАСПЭ и ДПЭ позволил идентифицировать 471 микропротеин и 599 пептидов.

Секвенирование становится особенно сложным, когда пептиды имеют несколько дисульфидных связей. Такие пептиды очень часто содержатся в ядах змей и пауков, а также токсинах моллюсков-конусов. Токсины конусов (*Conus ermineus*) изучались в работе Л. Куинтона с соавторами [180]. Яд разделяли на фракции с помощью катионообменной жидкостной хроматографии, затем каждую фракцию анализировали с помощью нано-ИЭР-ИЦР до и после восстановления дисульфидных связей. Такой подход позволил упростить

анализ за счёт определения количества дисульфидных связей в каждой фракции. SORI-CID применяли к четырём цистеинсодержащим фракциям  $\alpha$ -конотоксина (ЕПА). Авторы использовали автоматизированное *de novo* секвенирование. Сначала были обнаружены две Cys-содержащие последовательности: NWC и NKVC. Во-вторых, были идентифицированы сайты -A+P- и CR на C-конце. Для полного секвенирования пептида были проведены дополнительные эксперименты с синтезированными токсинами и масс-спектрометром Q-ВП. Наконец, аминокислотная последовательность была определена как ZTOGCCWNPACVKNRC-NH<sub>2</sub>. Также предполагалось, что две дисульфидные связи находятся между остатками Cys 1-3 и 2-4, как и для других токсинов этого типа. Эта гипотеза была подтверждена при изучении биологической активности синтетических пептидов.

Образец яда *Atractaspis correctis* был исследован в работе Л. Куинтона с соавторами [181]. Авторы проанализировали интактный и восстановленный образец яда с помощью ИЦП ПФ и обнаружили около 60 молекул в диапазоне масс 600–15000 Да. Половина из них находилась в диапазоне 2000–3500 Да, характерном для сарафототоксинов (SRTX), содержащих 20–30 аминокислотных остатков. Восстановление привело к увеличению масс молекул на 4 Да для двух пептидов, что свидетельствует о наличии двух дисульфидных связей. Сарафототоксины имеют типичные цистеиновые структуры Cys<sup>1</sup>-X-Cys<sup>3</sup> и Cys<sup>11</sup>-X-X-X-Cys<sup>15</sup>. Кроме того, известны, что дисульфидные связи находятся между остатками Cys<sup>1</sup>-Cys<sup>15</sup> и Cys<sup>3</sup>-Cys<sup>11</sup>. ДАС спектры позволили определить первичную структуру сарафототоксина (SRTX-i3). Сравнение с результатами клонирования кДНК позволило обнаружить два других семейства сарафототоксинов (SRTX-i1 и SRTX-i2). Полученная аминокислотная последовательность SRTX-i3 – CSCTDMSDLECMNFCHKDVIWVNRN.

Содержащие большое количество остатков цистеина пептиды улитки *C. monile* изучались в работе С. Наира с соавторами [182]. Анализ фракционированного при помощи ВЭЖХ образца до и после восстановления и алкилирования проводили, используя ДАС, ДЗЭ и ИКМФД фрагментацию, что позволило авторам определить структуру пептида как GNWCSARVCC, где W — бромтриптофан. Присутствие и положение бромтриптофана определяли по необычному изотопному составу. Также в работе был исследован синтетический 6-бромтриптофансодержащий пептид, обнаруживший тот же характер фрагментации. Положение дисульфидных связей изучали при помощи частичного мечения пептида с последующим алкилированием йодацетамидом. Было показано, что цистеины образуют связи между 1-3 и 2-4 остатками.

Интактные и восстановленные с помощью ТСЕР образцы яда *Conus virgo* без предварительного хроматографического разделения исследовали с помощью ВЭЖХ-МС ПФ и МАЛДИ-МС ПФ в работе [183]. В МАЛДИ экспериментах в качестве матрицы использовали

ДНВ и добавляли внутренний стандарт. Эксперименты ИЭР-ИЦР ПФ позволили обнаружить 64 молекулы в диапазоне масс 500-4500 Да. Наиболее интенсивными среди них были пептиды с моноизотопными массами 1328.5135 и 1358.5574 Да в двух различных зарядовых состояниях. Анализ восстановленного образца показал наличие четырех остатков цистеина для обеих молекул. Пептиды были отнесены к  $\alpha$ -конотоксинам и Т-суперсемейству. *De novo* секвенирование восстановленных образцов проводили с помощью тандемной масс-спектрометрии. Хотя двухзарядные ионы были основными пиками в спектрах интактного яда, в спектрах восстановленного образца эти ионы были слабоинтенсивными, что сделало ДЗЭ неприменимой. Установленные последовательности пептидов были определены как ZCC(I/L)T(I/L)PECCRI/L-NH<sub>2</sub> и ZCCPT(I/L)PECCRV-NH<sub>2</sub>. Паттерн Cys-Cys-(Xxx)<sub>5</sub>-Cys-Cys доказывал, что новые конотоксины принадлежали к Т-суперсемейству, обладающему связями между 1-3 и 2-4 остатками цистеина. Однако этот вывод требовал исследования биологической активности. Из МС/МС спектров невозможно было определить локализацию Glu и Pro, поэтому были синтезированы пептиды ZCPTTLPECCRV-NH<sub>2</sub> и ZCCPTLEPCCRV-NH<sub>2</sub>, их тандемные масс-спектры регистрировали в тех же условиях и сравнивали со спектрами природного конотоксина.

Иногда задача по определению положения дисульфидной связи остаётся нерешённой. В работе [184] Л. Куинтон с соавторами изучали яд *Dendroaspis angusticeps*. Его фракционировали с помощью ионообменной хроматографии. Фракцию ВЭЖХ, проявляющую активность ингибитора связывания 3Н-празозина, отделили и анализировали с помощью ИЦР ПФ. Разница масс до и после восстановления составила 8.005 Да, следовательно, в молекуле присутствовали четыре дисульфидные связи. Последовательность исследовали с помощью ДЗЭ иона (M+4H)<sup>4+</sup> с *m/z* 863.37. Пятьдесят остатков были подтверждены при помощи деградации по Эдману, и, наконец, последовательность была определена как LTCVTSKSIIFGITTEDCPDGQNLCKRRHYVVPKIYDSTRGCAATCPIPENYDSIHCCKTDKCNЕ. К сожалению, как и было сказано выше, авторы не определили положения дисульфидных связей.

С помощью EThcD фрагментации на приборе Orbitrap Fusion Lumos была определена последовательность пептида, содержащего 37 аминокислотных остатков, включая 8 цистеинов, с ММ 4109.69 Да [185]. Авторы прибегли к поиску по базам данных. Полученную последовательность подтверждали с помощью трипсинолиза и ВЭЖХ-МС/МС, однако положения дисульфидных связей не были установлены. Новый конотоксин был способен ингибировать рост микобактерий туберкулеза с эффективностью двух препаратов, используемых в настоящее время в клинической практике. В токсине с молекулярной массой

более 7 кДа из яда тexasской коралловой змеи установлено положение только одной из 3 внутримолекулярных S-S связей [33].

Используя комбинацию ВЭЖХ-МС/МС на приборе LTQ Orbitrap Elite и секвенирования по Сэнгеру, Х. Жанг с соавторами [186] идентифицировали 30 предшественников конопептидов, 10 из которых обладали новыми цистеиновыми паттернами: CC-C-C-C-C-C-C-C-C-C-C-C-C-CC-C-C-C-C-C и C-C-C-C-C-CC-C. Первый содержит 20 цистеинов, и надёжное определение положения дисульфидных связей в нём невозможно.

Наиболее эффективным подходом при работе с несколькими дисульфидными связями в новом пептиде может быть частичное восстановление. Хороший пример такого подхода был продемонстрирован в работе П. Ананда с соавторами [187]. Комплементарные ДАС и ДПЭ фрагментации позволили *de novo* секвенировать новый дисульфидсодержащий теретоксин морской улитки *Terebra variegata*. Наиболее сложной задачей для определения структуры таких молекул является установление положения дисульфидных связей. Пептид частично восстанавливали ТСЕР и алкилировали йодацетамидом и *N*-этилмалеимидом (NEM). Такое двойное алкилирование привело к образованию пептидов, в которых содержалось 2, 4 или 6 фрагментов NEM и IAM. Масс-спектрометрия позволила предложить следующие положения дисульфидных связей: Cys<sup>4</sup>-Cys<sup>20</sup>, Cys<sup>5</sup>-Cys<sup>21</sup> и Cys<sup>7</sup>-Cys<sup>16</sup>, что было дополнительно подтверждено методом ЯМР и анализом синтетических пептидов.

Два пептида моллюска *Conus virgo* были изучены в работе [188]. Частичное восстановление и алкилирование *N*-этилмалеимидом с последующим восстановлением ДТТ и алкилированием IAA приводит к образованию пептида с двумя карбамидометилированными и двумя NEM-модифицированными цистеинами. Дальнейший анализ при помощи МС/МС позволил установить положение дисульфидных связей между Cys<sup>2</sup>-Cys<sup>9</sup> и Cys<sup>3</sup>-Cys<sup>10</sup> для обоих исследованных пептидов. Конотоксин Вп5а из яда *Conus bandanus*, содержащий две дисульфидные связи, исследовали в работе [159]. После частичного восстановления и алкилирования авторы получили и разделили 2 пептида, в каждом из которых осталась по одной дисульфидной связи. МС/МС эксперимент для этих молекул доказал, что этот токсин обладает необычным положением дисульфидных связей: Cys<sup>1</sup>-Cys<sup>4</sup>/Cys<sup>2</sup>-Cys<sup>3</sup> вместо Cys<sup>1</sup>-Cys<sup>3</sup>/Cys<sup>2</sup>-Cys<sup>4</sup>, типичного для Т1-конотоксинов. Окончательной установленной последовательностью была NGCCIVRECC, где дифференцирование изомерных остатков Leu/Ile выполняли с помощью деградации по Эдману.

#### 1.4. Комплексные подходы масс-спектрометрии с преобразованием Фурье для секвенирования *de novo* природных пептидов

Поскольку *de novo* секвенирование – действительно сложная задача, наилучшие результаты получаются при совместном использовании всех доступных инструментальных методов и информации из баз данных. Используя комплексные подходы, можно идентифицировать и *de novo* секвенировать большую часть пептидома различных видов. Например, ручное *de novo* секвенирование пептидов лягушек даёт наилучшие результаты при использовании приборов высокого разрешения, нескольких методов фрагментации, различных методов дериватизации, включая введение *N*-концевых меток и модификацию дисульфидных связей, анализ спектров ионов в различных зарядовых состояниях, анализе спектров протеолитических фрагментов зрелых пептидов, информации из баз данных о возможных последовательностях пептидов, а также учёте особенностей, связанных с консервативными участками пептидов определенных семейств [53]. ИЦР ПФ с применением ДАС и ДЗЭ, восстановление и алкилирование цистеинов и ручная интерпретация спектров позволили провести *de novo* секвенирование нескольких пептидов лягушек *Rana ridibunda* и *Rana arvalis*, дифференцировав изобарные и изомерные аминокислотные остатки [134]. Применение дополнительного окисления дисульфидных связей секрета *Rana ridibunda* позволило установить 29 последовательностей вручную [71]. Сочетание окисления и восстановления дисульфидных связей с последующим анализом обоих производных значительно повысило эффективность ручного секвенирования и предоставило дополнительную информацию о структурах молекул. Комбинированный подход позволил обнаружить 17 новых пептидов в секрете *Rana arvalis* [135] и 19 новых пептидов в секрете *Rana lessonae* [165]. Аналогичный подход позволил обнаружить орнитокинин, антагонист рецептора орнитокина, в кожном секрете лягушки *Rana temporaria* [166]. Более высокая скорость сбора данных прибора Orbitrap Velos при той же разрешающей способности по сравнению с ИЦР ПФ позволила секвенировать 76 пептидов у той же бурой лягушки *Rana temporaria* с использованием ДАСПЭ и ДПЭ фрагментации [28]. Для определения структуры семи новых пептидов из кожного секрета *Rana latastei* использовали комбинацию спектров ДАС, ДАСПЭ и ДПЭ полипротонированных пептидов в различных зарядовых состояниях и их протеолитических фрагментов [189].

Спектры ДАС и ДПЭ интактных и триптических пептидов, полученные на приборе LTQ Orbitrap XL, позволили секвенировать два новых пептида в кожном секрете лягушки *Odorrana schmackeri* [190]. Процедура включала восстановление и алкилирование двумя разными алкилирующими агентами. Отнесение ионов разных серий осуществляли путём введения диметильной метки, содержащей тяжёлые и лёгкие изотопы, а для

дифференцирования изомерных Leu/Pe использовали деградацию по Эдману. Несколько новых пептидов, включая четыре сульфатированные формы, были *de novo* секвенированы в секрете лягушки *Pachymedusa dactylos* [191]. Авторы применяли МАЛДИ-ВП/ВП и ИЭР-МС, регистрируя ДАС и ДАСПЭ спектры на приборе Orbitrap XL. В некоторых случаях в ДАСПЭ спектрах не наблюдалось никаких пиков, и требовались дополнительные МС<sup>3</sup>-ДАС эксперименты. Для исследования кожного секрета жабы *Bufo gargarizans* применяли комбинированный анализ методом ВЭЖХ-МС/МС на приборе LTQ Orbitrap Velos Pro, секвенирование РНК и биоинформатический анализ [192]. Анализ ДАС спектров интактных пептидов с зарядовыми состояниями от +2 до +6 при помощи PEAKS, баз данных и гомологического поиска с подтверждением при помощи клонирования кДНК позволил обнаружить 939 пептидов в диапазоне масс 0.7-4 кДа. Обнаруженные пептиды могут проявлять активность в иммунных и метаболических процессах, антиоксидантную активность и антимикробное действие.

Пептиды, содержащиеся в ядах змей обычно имеют большую длину и могут рассматриваться как белки. Яд индонезийской королевской кобры *Ophiophagus hannah* состоит из полипептидов массой 7-25 кДа. В этом случае применялись подходы «снизу вверх» и «сверху вниз». В анализе использовали ДАС и ДАСПЭ фрагментацию, спектры регистрировали на приборе LTQ Orbitrap [193]. Авторы идентифицировали более 30 белков/пептидов, принадлежащих к 10 семействам, в основном,  $\alpha$ -нейротоксинам. Та же группа изучала яды восточноафриканской зелёной мамбы *Dendroaspis angusticeps* и чёрной мамбы *Dendroaspis polylepis* [194]. Комбинация подходов «сверху вниз» и «снизу вверх» позволила идентифицировать 19 белков и 5 натрийуретических пептидов в яде *D. angusticeps* и 11 белков и 4 натрийуретических пептида в яде *D. polylepis*. Аналогичный подход, примененный к яду коричневого паука *Loxosceles intermedia*, позволил обнаружить 190 пептидов/белков [195]. 27 новых нейропептидов были идентифицированы в нематоде *Ascaris suum* при анализе на приборе LTQ Orbitrap Velos с последующим поиском по базам данных [196]. Для получения спектров ДАС, ДАСПЭ и ДПЭ было выполнено четыре цикла анализа в режиме зависимых сканирований с различными параметрами сбора данных.

Последовательности 70 нейропептидов из синусовой железы синего краба были получены при использовании МАЛДИ-МС ПФ, МАЛДИ-ВП/ВП, ИЭР-Q-ВП и МС/МС [197]. Яд паука *Cupiennius salei* исследовали методами МАЛДИ-ВП и ИЭРМС/МС [198] на приборе с орбитальной ловушкой ионов, применяя деградацию по Эдману. Авторы использовали три различных стратегии обессоливания и фракционирования образцов. Пептиды длиной более 25 аминокислот в цепи подвергали предварительному ферментативному расщеплению. Авторы не упомянули метод ПФ как наиболее полезный, однако подход позволил обнаружить 286

компонентов яда. Авторы секвенировали 59 пептидов, разделенных на две группы, различающиеся по биологической активности: короткие пептиды без дисульфидных связей в диапазоне масс 2-4 кДа и Cys-содержащие пептиды. Последовательности 49 пептидов были установлены.

В работе [54] при помощи комбинации ручного и автоматического секвенирования был изучен пептидом тарантула *Acanthoscurria gomesiana*. Авторы анализировали интактные пептиды при помощи ВЭЖХ-МС/МС с помощью масс-спектрометра Q-ВП, а анализ триптических пептидов проводили с помощью масс-спектрометра LTQ Orbitrap Velos с разрешающей способностью 60 000 в режиме МС и 7 500 в режиме ДАСПЭ. Сначала была составлена карта масс интактных пептидов с массами в диапазоне 3000-7000 Да. Было идентифицировано семнадцать известных токсинов, а для трёх пептидов были предложены новые. Была изучена их антимикробная активность в отношении штаммов *E. coli*, *E. cloacae* и *C. albicans*.

Поиск в базах данных и *de novo* секвенирование с помощью спектров ДАС, ДАСПЭ и ДПЭ, полученных на приборе LTQ Orbitrap Velos, позволили идентифицировать 147 пептидов в яде скорпиона *Tityus serrulatus* [30]. Было обнаружено, три фосфорилированных один N-гликозилированный белок. Авторы подчёркивают, что 80% секвенированных молекул представляют собой протеолитические фрагменты исходных пептидов токсинов.

Яд моллюска-конуса *Conus textile* изучался в работе [55]. Авторы получили последовательность 31 пептида, сочетая ВЭЖХ-фракционирование, точное определение масс с помощью МАЛДИ-ВП и ВЭЖХ-МС/МС. Масс-спектрометр LTQ Orbitrap использовался в режимах ДАС и ДПЭ. Авторы отметили бóльшую эффективность ДАС для коротких пептидов и бóльшую эффективность ДПЭ для более длинных пептидов с более высоким зарядом. Чтобы определить количество дисульфидных связей, авторы анализировали разницу масс между восстановленными и алкилированными токсинами и обнаружили молекулы с 1-5 дисульфидными связями. В работе [160] *N,N*-диметил-2-хлорэтиламин использовали вместо йодацетамида для превращения свободных остатков Cys в их диметилловые аналоги Lys. Эта модификация увеличила зарядовые состояния ионов-предшественников и позволила провести полное секвенирование по спектрам ДПЭ. Авторы также объединили *de novo* секвенирование и поиск в библиотеке последовательностей кДНК, чтобы сократить время, необходимое для интерпретации спектров. В дальнейшем исследовании яда моллюска-конуса *Conus textile* использовали спектры ДАС и ДАСПЭ, полученные на приборе LTQ Orbitrap XL [34]. Авторы применили разработанный полуавтоматический алгоритм *Conovo*, основанный на более ранних исследованиях, для уменьшения количества возможных последовательностей. Предлагаемый подход помог обнаружить 60 токсинов, в том числе 13 новых пептидов,

отсутствующих в более ранних исследованиях. При изучении нейропептидов органов перикарда вида *Callinectes sapidus* использовали МАЛДИ-МС ПФ, МАЛДИ-ВП/ВП и ИЭР-Q-ВП. Среди 130 идентифицированных пептидов 44 были новыми [199]. Данные анализа методом ИЦР ПФ и использование ДАС и ДЗЭ позволили на 100% выяснить структуру двух гомологов  $\beta$ -тимозина позвоночных у морского моллюска *Aplysia californica* [200]. Последовательность одного гомолога был определена при поиске по базе данных транскриптома, в то время как другой показал аналогичную потерю брома в ДЗЭ и оказался мутантной формой с заменой His на Gln.

Сорок пять потенциальных антимикробных пептидов были идентифицированы в плазме миссисипского аллигатора с помощью ВЭЖХ-МС/МС анализа на приборе LTQ Orbitrap Elite в режиме ДПЭ. Авторы использовали базу данных экспрессированных тегов последовательностей (EST), NCBI [201]. Поиск при помощи PEAKS обнаружил 5469 EST последовательностей. Кроме того, использовалась база данных транскриптомов миссисипского аллигатора Международной рабочей группы по геному крокодилов [202]. Последовательности целевых антимикробных пептидов, идентифицированные в результате поиска в базе данных и/или секвенированные только путём *de novo* секвенирования, проверялись вручную. В последующем исследовании того же вида количество идентифицированных антимикробных пептидов достигло 340, в том числе 5 новых [58]. Тот же подход был применен к антимикробным пептидам комодского варана *Varanus komodoensis* [57]. Авторы использовали программу PEAKS для поиска последовательностей и подтверждали их ручной интерпретацией спектров. Также успешно использована для проверки информация о транскриптоме. Поскольку авторов интересовали антимикробные пептиды, они ограничили свой поиск ионами с зарядовыми состояниями от +4 до +6 и молекулярными массами 1200-4000 Да. Этот диапазон масс был выбран на основе их более ранних исследований пептидов американских аллигаторов [56,58]. Авторы разработали скрипт Python для идентификации потенциальных антимикробных пептидов и получили список из 48 пептидов, восемь из которых были выбраны для синтеза и оценки их активности. Антимикробную активность этих пептидов исследовали в отношении *P. aeruginosa* и *S. aureus*.

А. Тушар с соавторами исследовали яд муравья *Tetramorium bicarinatum* на приборах LCQ, проводя предварительные эксперименты с низким разрешением, и LTQ Orbitrap Elite для ДАС и ДАСПЭ [203]. *De novo* секвенирование проводили при помощи программного обеспечения PEAKS и данных о транскриптоме. Авторы идентифицировали 37 предшественников пептидов, а последовательности 29 зрелых пептидов были подтверждены масс-спектрометрически.

## 1.5. Методы визуализации и таксономических исследований

Точное измерение масс, доступное для масс-спектрометров с преобразованием Фурье, создало основу для 2D-картирования. Общий метод визуализации данных для сложных смесей пептидов на основе пузырьковых диаграмм обсуждался в [78,79]. Метод был разработан и применён для сравнения пептидомов различных лягушек. Диаграммы построены с использованием нормализованного дефекта масс (NMD) и нормализованного изотопного сдвига (NIS) по осям абсцисс и ординат соответственно. Значения NMD и NIS можно рассчитать по следующим формулам:

$$NMD = 1000 \times \frac{MM-MN}{MM}; NIS = 1000 \times \frac{MA-MM}{MM},$$

где MM – моноизотопная, MN – номинальная, а MA – средняя молекулярные массы [78].

Каждый пептид представлен на диаграмме на рис. 11 в виде пузырька определенного цвета, который определяет семейство пептидов. Положение пузырька зависит от значений NMD и NIS, а его размер пропорционален интенсивности его протонированной молекулы в масс-спектре. Пептиды, принадлежащие к одному семейству, то есть имеющие некоторое сходство в своих последовательностях, занимают соседние положения на графике, демонстрируя некоторую кластеризацию данных. Следовательно, пептидом кожи каждого вида лягушек может быть представлен уникальной диаграммой, демонстрирующей определённые видовые признаки. Двумерное картирование хорошо визуализирует пептидом и может быть успешно использовано в таксономических исследованиях для идентификации близкородственных видов, когда физиологические параметры особи приводят к сомнительным выводам.

На рис. 11 показан пример применения двумерных карт для дифференциации сложного для определения по внешним признакам комплекса лягушек *Pelophylax esculentus*. *R. esculenta* – хорошо известная съедобная лягушка. Она является гибридом двух других представителей комплекса: *R. lessonae* и *R. ridibunda*. Физиологические признаки этих трёх видов очень близки, из-за чего часто возникают ошибки в их идентификации. Но 2D-карты представляют собой прекрасный инструмент для однозначной дифференциации видов [79].

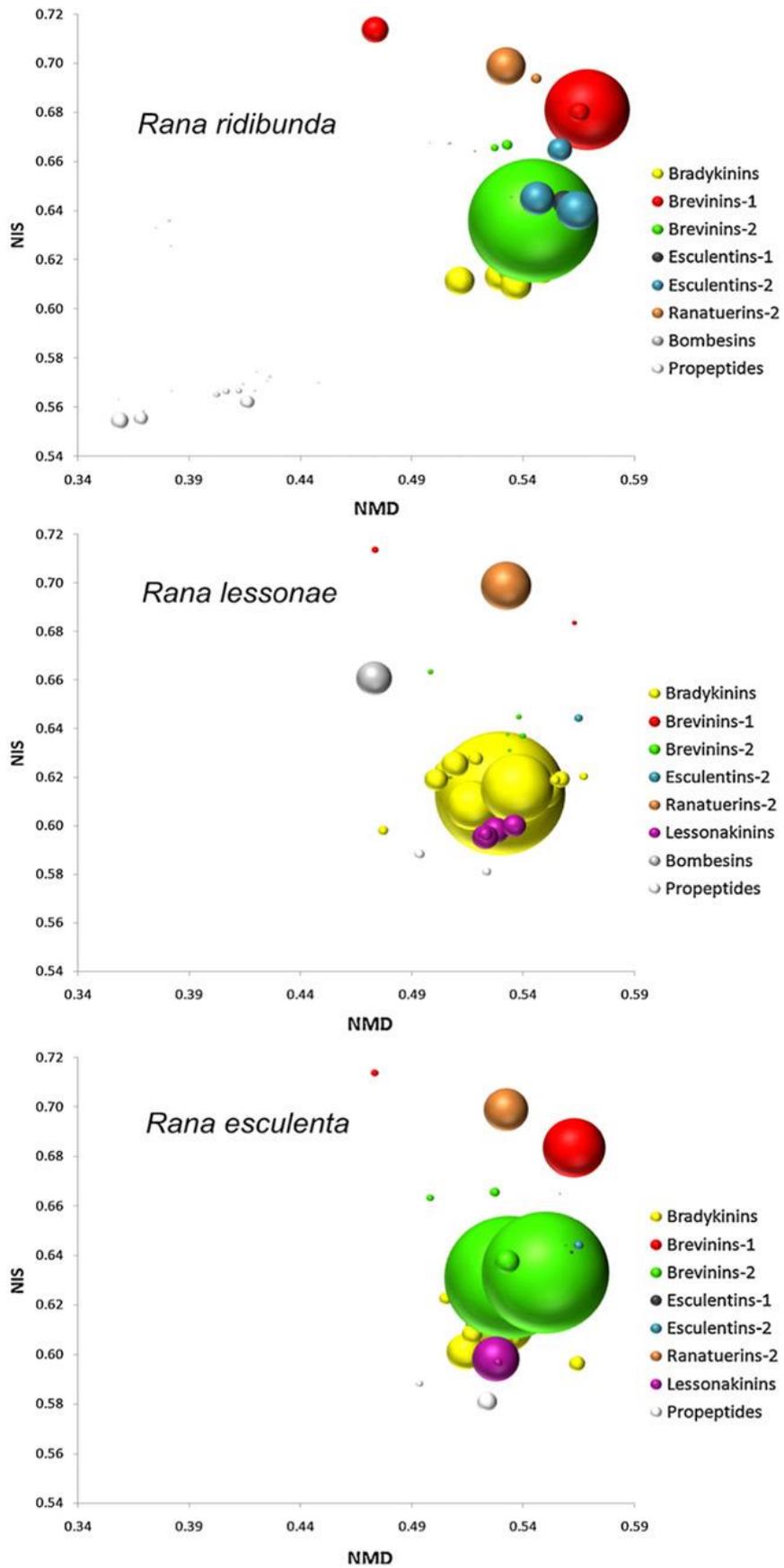
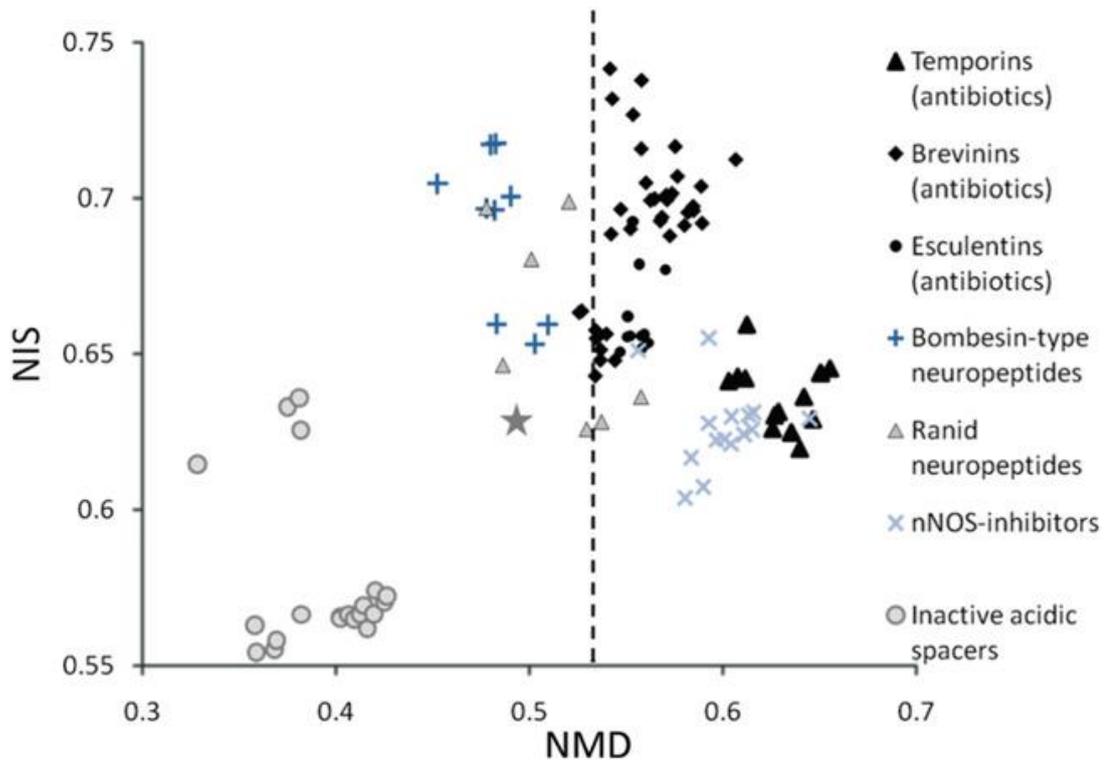


Рисунок 11. 2D-карты трёх родственных видов лягушек [79].

Более того, нанесение точек, соответствующих пептидам с известной биологической активностью, на график в координатах NMD-NIS показывает, что пептиды с одинаковым

типом активности занимают на графике соседние позиции (рис. 12). Это свойство позволяет прогнозировать тип активности новых пептидов, формально используя только информацию об их молекулярных массах [78]. Оценка активности новых *de novo* секвенированных пептидов, выделенных из кожного секрета словенской лягушки из комплекса *Pelophylax esculentus*, была проведена при помощи построения 2D-карт [59].

На самом деле этот метод можно считать даже более интересным, чем популярный QSAR [204], демонстрирующий связь между структурой молекулы и её биологической активностью. Подход NMD-NIS демонстрирует корреляцию между массой (элементным составом) и биологической активностью.



**Рисунок 12.** 2D карта, демонстрирующая кластеризацию пептидов по типу их биологической активности [78]. Звёздочкой отмечен авераджин (Av) –  $C_{4,9384}H_{7,7583}N_{1,3577}O_{1,4773}S_{0,0417}$ , отражающий среднее содержание элементов в остатках аминокислот.

### 1.6. Программное обеспечение для *de novo* секвенирования в протеомике

Как было сказано выше, в современной протеомике существуют различные программные методы идентификации пептидов и белков: протеомные поисковые машины, сопоставляющие экспериментальные масс-спектры с библиотеками или базами данных, и *de novo* секвенирование, в котором последовательность аминокислот определяется из разности фрагментных ионов, напрямую из тандемных масс-спектров без помощи базы данных. Каждый подход обладает как своими преимуществами, так и недостатками [205]: например, в случае *de novo* секвенирования можно идентифицировать пептиды и белки, не содержащиеся

в базах данных, однако для этого требуются масс-спектры высокого разрешения и значительные вычислительные ресурсы. Поиск по базам данных же позволяет упростить обработку данных. Он несколько толерантнее к самим масс-спектрам, но требует существования подходящей белковой библиотеки.

### 1.6.1. Поиск по базам данных

Белковые последовательности доступны для многих организмов: например, открытая база данных UniProt. [206], объединившая в себе несколько предшествующих баз, содержит в себе данные как о уже исследованных белках, так и данные о геномах, которые также могут быть использованы для создания теоретических молекул белка для дальнейшей обработки спектров. Как правило, в качестве используемой библиотеки выбирается база данных, содержащих белковые последовательности интересующего организма.

Все алгоритмы поиска по базам данных работают примерно одинаково: для белков на основании библиотек генерируются теоретические энзиматические фрагменты, а для каждого такого фрагмента – теоретическая масса и фрагментные масс-спектры. Алгоритмы ищут теоретические пептиды, чьи массы и спектры фрагментации соответствуют экспериментальным, в итоге приводя список предполагаемых идентификаций (пептидов-кандидатов). Параметры, по которым проводится поиск в библиотеке включают в себя тип фрагментных ионов, фермент для энзиматического расщепления и точность измерения массы родительского иона. Математические модели, используемые для каждой конкретной ППМ, при этом разнятся [207].

На сегодняшний день существует довольно большое число алгоритмов, наиболее известные объединены в таблице 1.

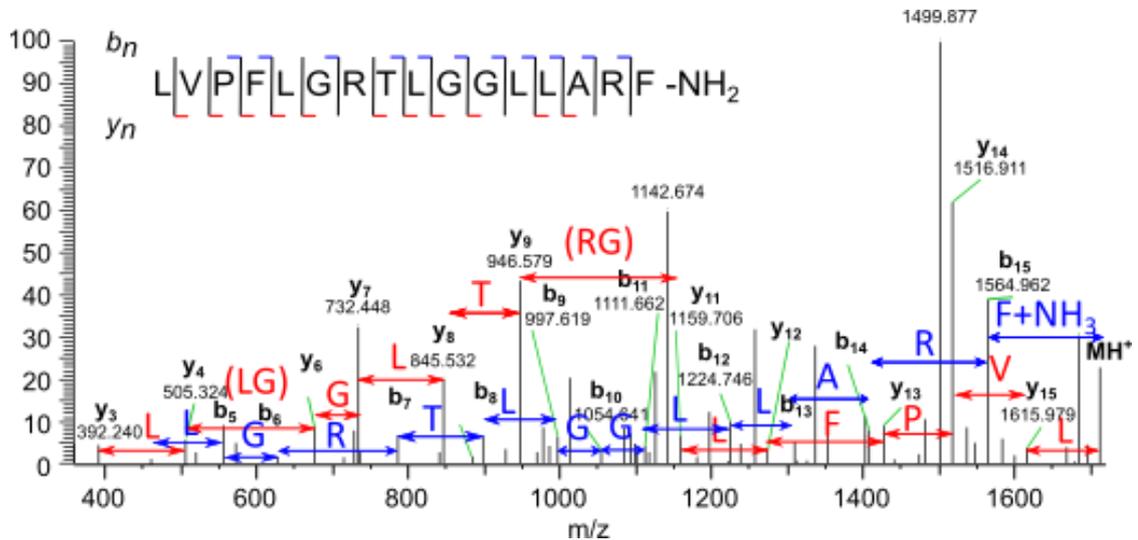
**Таблица 1.** Протеомные поисковые машины[205,208]

Название	Тип лицензии	Год появления
Sequest	Коммерческая	1994 [209]
Mascot	Коммерческая, бесплатный веб-интерфейс	1999 [210]
Phenyx	Коммерческая	2003 [208,211]
X!Tandem	Бесплатная	2003 [212,213]
PEAKS DB	Коммерческая	2003[214]
OMSSA	Бесплатная	2004 [215]
InsPecT	Бесплатная	2005 [216]
ProbID	Бесплатная	2006 [217]

### 1.6.2. *De novo* секвенирование

В случае отсутствия подходящих баз данных *de novo* секвенирование – удачный выбор для интерпретации полученных спектров, поскольку данные не требуют никаких эталонов. Оно состоит в определении аминокислотной последовательности и её модификаций на основе

экспериментально полученного tandemного масс-спектра. Ключевой принцип заключается в расчёте разности масс пар фрагментных ионов и её сравнении с массами остатков двадцати природных аминокислот с учётом того, что лейцин и изолейцин – пара изомерных кислот, и масса их аминокислотных остатков совпадает. В основе многих алгоритмов программного обеспечения лежит построение так называемого «спектрального графа», строящегося на основе tandemного масс-спектра и состоящего из вершин и рёбер, где вершины являются пиками в спектре, а рёбра попарно соединяют вершины, чья разность масс соответствует массе аминокислотных остатков. Оригинальный подход был предложен Бартельсом в 1990 г. Приписываемой спектру последовательностью пептидов будет та, которая позволит сопоставить спектру оптимальный путь в графе. В качестве примера такого графа можно привести масс-спектр темпорина S (рис. 13), где красным и синим выделены графы *b*- и *y*-серий ионов, дополняющие друг друга и в сумме позволяющие определить полную последовательность.



**Рисунок 13** Пример *de novo* секвенирования, темпорин S.

Большие объемы данных значительно осложняют ручное *de novo* секвенирование, которое проводится аналогичным образом. Это способствовало появлению алгоритмов для автоматического секвенирования. Первые подходы к этой задаче, основанные на пошаговом сопоставлении экспериментальных tandemных масс-спектров ионов-предшественников с теоретическими спектрами для каждого такого иона и, к сожалению, подходящие только для коротких аминокислотных последовательностей в силу экспоненциального роста вариантов таких теоретических спектров, были предложены ещё в 80-е годы [37]. На сегодняшний день для *de novo* секвенирования используются подходы на основе новых методов, включая целочисленное [218] и динамическое программирование [221–224], машинное [219–221] и глубокое обучение [222]. В таблице 2 перечислены некоторые доступные на сегодняшний день программные продукты, используемые в них алгоритмы и тип лицензии.

**Таблица 2.** Доступные программные продукты для *de novo* секвенирования [223]:

Название алгоритма	Используемый подход	Тип лицензии	Год появления
Lutefisk	Спектральный граф	Бесплатная	1998 [224]
SeqMS	Спектральный граф	Бесплатная	1998 [225,226]
PEAKS	Динамическое программирование	Коммерческая	2003 [227]
PepNovo	Спектральный граф, динамическое программирование	Бесплатная	2005 [228]
NovoHMM	Скрытая марковская модель	Бесплатная	2005 [229]
DirecTag	Спектральный граф, генерация «тэгов» пептидных последовательностей	Бесплатная	2008 [230]
pNovo+	Спектральный граф, динамическое программирование	Бесплатная	2010 [231,232]
UniNovo	Спектральный граф, динамическое программирование	Бесплатная	2013 [233]
Novor	Спектральный граф, машинное обучение	Бесплатная	2015 [220]
Twister	Секвенирование «сверху вниз», генерация «тэгов» пептидных последовательностей	Бесплатная	2015 [234]
LADS	Машинное обучение	Бесплатная	2016 [235]
UVNovo	Скрытая марковская модель, машинное обучение	Бесплатная	2016 [219]
MRUniNovo	Спектральный граф, динамическое программирование	Бесплатная	2016 [236]
Open-pNovo	Спектральный граф, динамическое программирование	Бесплатная	2017 [237]
DeepNovo	Динамическое программирование, глубокое обучение	Бесплатная	2017 [222]

Дальнейшая часть раздела будет посвящена избранным программным продуктам.

### *Lutefisk*

Данный метод был предложен в 1997 году для обработки тандемных масс-спектров триптических пептидов, полученных при помощи ДАС [224]. Алгоритм использует построение «спектрального графа», который описан выше. Для работы со спектром программа описывает все пики в нём центроидами, то есть учитывает  $m/z$ , соответствующий максимальной интенсивности конкретной массы, и выделяет список значимых фрагментных ионов, определяя локальные максимумы относительно заданного уровня сигнал/шум ( $S/N$ ). На следующей стадии программой выполняется конвертация всех ионов в предположительные *b*- и *y*-ионы по формулам, приведённым в таблице 3.

**Таблица 3.** Уравнения для конвертации спектра ДАС в спектр аминокислотной последовательности [224]

<i>N</i> -концевые ионы	Преобразование в соответствующий <i>b</i> -ион
a	Ион+28
a-17	Ион+45
a-18	Ион+46
b	Ион (без изменений)
b-17	Ион+17
b-18	Ион+18

<b>C-концевые ионы</b>	
y	Пептид-ион+2
y-17	Пептид-ион-15
y-18	Пептид-ион-16

Пептид – молекулярная масса секвенируемого пептида, ион – величина  $m/z$  исследуемого иона в спектре.

Из полученных ионов составляют *N*- и *C*-концевые «доказательные списки» – наборы ионов, которые затем комбинируют в «спектральные последовательности» и определяют возможные аминокислотные последовательности. Полученные последовательности ранжируют на основе интенсивностей фрагментных ионов и подвергают кросс-корреляционному анализу, чтобы получить итоговый список предполагаемых сиквенсов.

### **PEAKS**

Алгоритм PEAKS [223] основан на новой модели динамического программирования [238], предполагающей, что весь спектр состоит из пиков. В случае, если пик в конкретной точке в экспериментальном спектре отсутствует, программа приписывает ему пик с интенсивностью равной нулю. Рассчитанные PEAKS последовательности пептидов соответствуют максимальному числу интенсивных ионов в спектре, причём программа может идентифицировать аминокислоту даже при отсутствии некоторых пиков в спектре благодаря своему алгоритму. Приписывая спектру аминокислотную последовательность, программа рассчитывает «локальную достоверность» для каждой аминокислоты и «среднюю локальную достоверность» для всей определённой последовательности.

В настоящее время доступна уже десятая версия программы – PEAKS Studio X Pro [239]. В ней заложен алгоритм для *de novo* секвенирования DeepNovo, который позволяет получать информацию из спектров, полученных различными методами фрагментации. Первые версии этой программы, как и Lutefisk, более эффективно работали с секвенированием триптических пептидов. В настоящей работе мы тестировали эффективность PEAKS Studio X Pro для *de novo* секвенирования нетриптических кожных пептидов ранидных амфибий.

### **Twister**

Twister – алгоритм, используемый для *de novo* секвенирования белков методом «сверху вниз» [234]. Данный алгоритм основан на применении «тэгов» пептидных последовательностей – алгоритм с максимальной надёжностью определяет участки белка и, комбинируя их, предлагает конечный результат.

На первом этапе спектр деконволюируется, то есть все ионы программно приводятся к моноизотопным аддуктам вида  $MH^+$  или нейтральным частицам вида  $M$ . Деконволюцию обычно проводят программами Xtract [240], MS-Deconv [241] или Hardklör [244]. Затем генерируется набор пептидных «тэгов» фиксированной длины  $k$ , для чего задается очень

высокая точность масс ионов. Полученные «тэги» комбинируются в более длинные аминокислотные последовательности методом построения графа де Брюина, часто используемого в геномике. Найдя оптимальный путь в построенном графе, программа приписывает белку определенную последовательность.

### *DeepNovo*

Данный алгоритм использует в своей основе глубокое обучение [228,245]. Принимая экспериментальный спектр в качестве входных данных, DeepNovo секвенирует пептид, начиная со специального символа «старт», предсказывая каждую последующую аминокислоту с учетом данных предыдущих шагов. Процесс останавливается, если на конкретном шаге выводится специальный символ «конец». Аналогичным образом программа выполняет секвенирование в обратную сторону, в качестве итоговой последовательности выбирается пептид с наилучшей оценкой.

В DeepNovo используются две модели: в первой рассчитывается масса «префикса» – сумма всех аминокислотных остатков и соответствующего конца. Затем к «префиксу» прибавляется каждая возможная аминокислота и рассчитываются массы теоретических *b*- и *y*-ионов, для каждого из которых проводится сравнение с исходным спектром в окне шириной 1 Да. Получаемый профиль ионов затем проходит через свёрточную нейронную сеть (CNN), которая изучает пики и суммирует информацию о фрагментных ионах в спектре. Следующая модель основана на рекуррентной нейронной сети (RNN) и использует долгую краткосрочную память (LSTM). Для этого каждую аминокислоту модель представляет в виде «вектора» и, учитывая «префикс», находит соответствующие «векторы» в спектре.

В итоге данные двух моделей комбинируются, чтобы распределить аминокислоты по вероятностям. Такая обработка гарантирует, что окончательные последовательности, предлагаемые алгоритмом, будут иметь правильную массу пептида.

Большая часть существующего сегодня программного обеспечения имеет дело с автоматическим секвенированием преимущественно триптических пептидов, которые, как правило довольно короткие (до 20 аминокислот) и имеют *C*-концевой заряд (*C*-Lys, *C*-Arg). Это гарантирует интенсивную длинную серию *y*-ионов, что повышает, в итоге, процент покрытия сиквенсов триптических пептидов и, в целом, исходных белков. С нетриптическими пептидами амфибий, осложненными *C*-концевыми внутримолекулярными S-S связями, автосеквенаторы работают не столь эффективно.



вычислительных ресурсов, более дорогих приборов и не годится для ручной интерпретации масс-спектров.

### ***Анализ в режиме зависимых сканирований (DDA)***

В ходе такого типа сбора масс-спектрометрических данных фрагментации подвергаются только выбранные ионы, причём обычно это ионы с наибольшей интенсивностью – «topN». Число выбранных ионов, как правило, не превышает 5-15-ти для одного скана первого порядка ( $MH^+$ ). Для каждого выбранного иона последовательно регистрируется тандемный спектр. Собранный таким образом массив масс-спектрометрических данных проще анализировать, он отлично подходит для *de novo* секвенирования, а также относительного количественного анализа, выполняемого с помощью введения химических меток. Минусом масс-спектрометрического анализа пептидных смесей в режиме зависимых сканирований может быть потеря тандемных спектров минорных компонентов, поскольку в стратегии «topN» прибор в ходе анализа выбирает ионы по интенсивности и последовательно их фрагментирует. Чтобы избежать этого, прибегают к так называемому «динамическому исключению» (dynamic exclusion [243]), в ходе которого массы уже измеренных «topN» помещаются в список ионов с временным запретом регистрировать спектры. По истечению заданного времени, исключенный ион автоматически удаляется из списка, а его тандемные масс-спектры снова регистрируются. Этот прием позволяет зафиксировать спектры минорных компонентов и избежать их потери.

### **Заключительный комментарий**

Масс-спектрометрия с преобразованием Фурье определённо представляет собой мощный инструмент для *de novo* секвенирования природных пептидов. Благодаря непревзойденной разрешающей способности и возможности использования разнообразных методов фрагментации МС ПФ остается наиболее надёжным методом секвенирования, позволяющим проводить анализ методом «сверху вниз» небольших интактных белков или крупных пептидов. Инструменты ИЦР ПФ для секвенирования пептидов постепенно отступают, оставляя место для новых версий приборов на основе орбитальной ловушки, однако новые инструменты ИЦР ПФ, появляющиеся на рынке, могут быть успешно использованы в самых сложных задачах, требующих максимально возможного разрешения. Приборы Orbitrap имеют очень хорошие перспективы: при уже достигнутой разрешающей способности более 1 000 000 они могут работать практически со всеми природными пептидами, так как их массы обычно не превышают 7 кДа. Многочисленные методы тандемной масс-спектрометрии помогают решать самые сложные задачи, связанные с покрытием сиквенса,

дифференцированием изомерных аминокислотных остатков, установлением типов и сайтов ПТМ. Наверняка на каком-то этапе должен появиться эффективный механизм автоматического *de novo* секвенирования, способный учесть всю доступную информацию, полученную в экспериментах.

Важной задачей является и разработка метода масс-спектрометрического секвенирования без применения химических модификаций.

Целью данной работы явилось секвенирование компонентов кожных секретов методами фрагментации орбитальной ловушки для выявления популяционных различий внутри вида амфибий. Для этого необходимо было:

1. Оценить применимость метода секвенирования белков и пептидов «сверху вниз» к кожным секретам ранидных лягушек на примере *Rana arvalis* и *Rana temporaria* из различных популяций.
2. Выявить биомаркеры различных популяций и видов *Rana arvalis* и *Rana temporaria* путем сравнения составов их кожных пептидомов, установленных комплексом методов орбитальной ловушки.
3. Изучить вклад МС<sup>3</sup> метода, EThcD, в секвенирование интактных пептидов с внутримолекулярной С-концевой S-S связью.
4. Провести оценку применимости современных алгоритмов автоматического секвенирования для *de novo* секвенирования интактных пептидов ранидных лягушек путем сравнения с результатами ручной интерпретации масс-спектров.

## 2. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

### 2.1. Кожный пептидом остромордой лягушки *Rana arvalis* (Центральная Словения)

В работе был изучен кожный секрет взрослой особи остромордой лягушки *R. arvalis*, отловленной на территории Центральной Словении (окрестности Любляны). Секвенирование пептидов проводилось ручной интерпретацией совокупности спектров ДАС, ДАСПЭ, ДПЭ и EThcD интактных образцов кожных секретов, полученных на масс-спектрометре Orbitrap Elite ETD. Ручная интерпретация спектров подразумевает суммирование всей полноты структурной информации, полученной при использовании параллельно нескольких методов фрагментации. При этом учитываются данные спектров различной зарядности полипротонированных пептидов и их протеолитических фрагментов, которые, как правило, присутствуют в кожных секретах в больших количествах. Объединение всей полученной спектральной информации по каждому из пептидов работа кропотливая и трудоемкая, требующая особого внимания, но, с другой стороны, ее можно рассматривать и как достоинство ручного секвенирования перед автоматическим [50].

Последовательности пептидов, обнаруженных в кожном секрете особей *R. arvalis* из Центральной Словении, приведены в таблице 4. Для пептидов с известными сиквенсами даны их коды в базе данных UniProtKB (<http://www.uniprot.org>); последовательности новых пептидов отмечены \*. Установленные в работе Leu/Ile выделены жирным шрифтом.

**Таблица 4.** Пептиды, идентифицированные в кожных секретах особей *R. arvalis* из центральнословенской популяции.

№	ММ, Да	Пептид	Последовательность	UniProtKB
1	1809.952	Бревинин 1AVa	FLPLLAASFACTVTKKC-OH	P86159
2	1873.093	Бревинин 1AVb	FVPLLVS <del>K</del> LV <del>C</del> VTKKC-OH	P86160
3*	3369.812	Бревинин 2AV	GLMSVLKKAGKHVAKNVAVSLMDSLKCKISGGC-OH	C0HM56
4	2896.468	Ранатуерин 2AVb	GLMDMVKGAAKNLFASALDTLKCKITGC-OH	P86162
5*	2839.603	Ранатуерин 2AVc	GLLDVVKGAAKNLLATALDKLKCKVTGC-OH	
6	2361.430	FQ-22-1	FVGAALKVLANVLPVVISWIKQ-NH <sub>2</sub>	P86158
7*	2361.426	AK-23-1	ALGGVLKALAKGLPSVISWINQK-NH <sub>2</sub>	C0HM57
8*	1404.870	Темпорин AV	LLPLVGNLLNDLL-NH <sub>2</sub>	C0HM58
9*	1309.833	Темпорин AVa	GLEVLGKILSGIL-NH <sub>2</sub>	C0HM59
10	1059.561	Брадикинин (Br)	RPPGFSPFR-OH	P84895
11	806.408	[des-Arg <sup>1</sup> ] Br	PPGFSPFR-OH	
12	1508.800	[Arg <sup>0</sup> ]RA-11	RRPPGFSPFR <del>I</del> A-OH	
13	1243.683	RA-11-2	RPPGFSPFR <del>I</del> A-OH	P86156
14	1087.581	[des-Arg <sup>1</sup> ] RA-11	PPGFSPFR <del>L</del> A-OH	
15	1340.735	RP-12	RPPGFSPFR <del>I</del> AP-OH	
16	1498.804	RS-14-2	RPPGFSPFR <del>I</del> APAS-OH	P86156
17	1599.852	RT-15	RPPGFSPFR <del>I</del> PAST-OH	
18	1712.936	RL-16	RPPGFSPFR <del>I</del> PASTL-OH	

\* Впервые докладываемые последовательности

Анализ компонентного состава кожного пептидома словенских особей *R. arvalis* проводился с использованием удлиненного времени хроматографирования-120 мин; спектры собирались в режиме зависимых сканирований (DDA) с применением динамического исключения интенсивных ионов с размером списка временно отложенного сканирования 25 ионов. Кроме того, было увеличено время накопления сигнала: для спектров первого порядка оно составило 100 мс для тандемных спектров – 200 мс (см. раздел Экспериментальная часть). Все вместе это позволило не потерять минорные компоненты пептидной смеси.

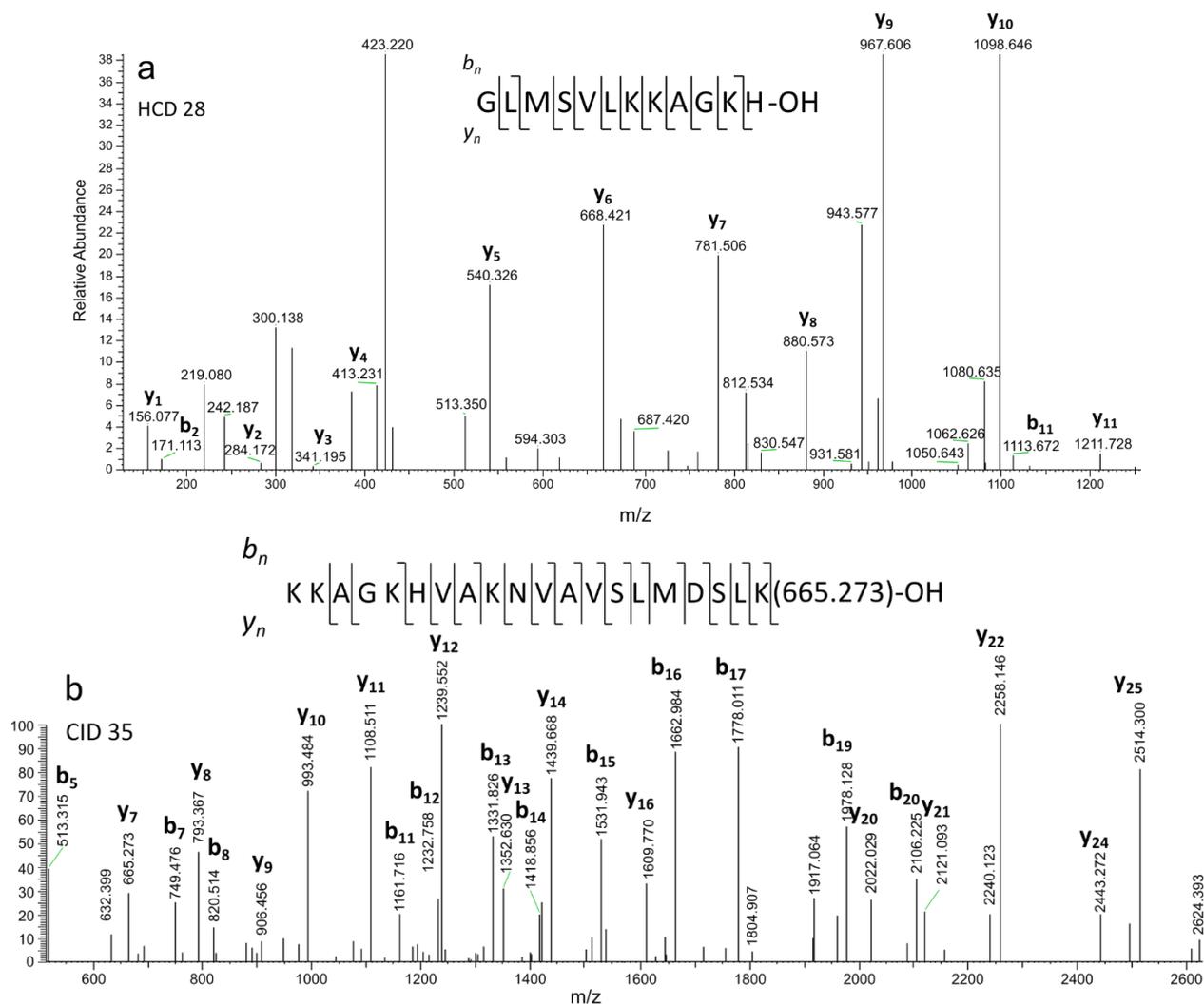
### **Бревинин 2 AV\*, MM 3369.812 Да**

В секрете словенских особей *R. arvalis* обнаружен новый бревинин 2AV, включающий в себя 33 аминокислотных остатка и названный в соответствии с существующей номенклатурой [244]. В секрете он присутствует в небольших количествах. У московских особей *Rana arvalis* он не был детектирован [135]. На рисунке 15 а, б приведены спектры двух частично пересекающихся фрагментов бревинина 2AV, полный спектр которого пептида зафиксирован не был. Спектр ДАСПЭ N-концевого фрагмента (1-12), ион с  $m/z$  423.588<sup>3+</sup>, и спектр ДАС C-концевого фрагмента бревинина 2AV (7-33), ион 924.168<sup>3+</sup>, получены при значениях нормализованной энергии активирующих частиц (NCE) 28 и 35, соответственно.

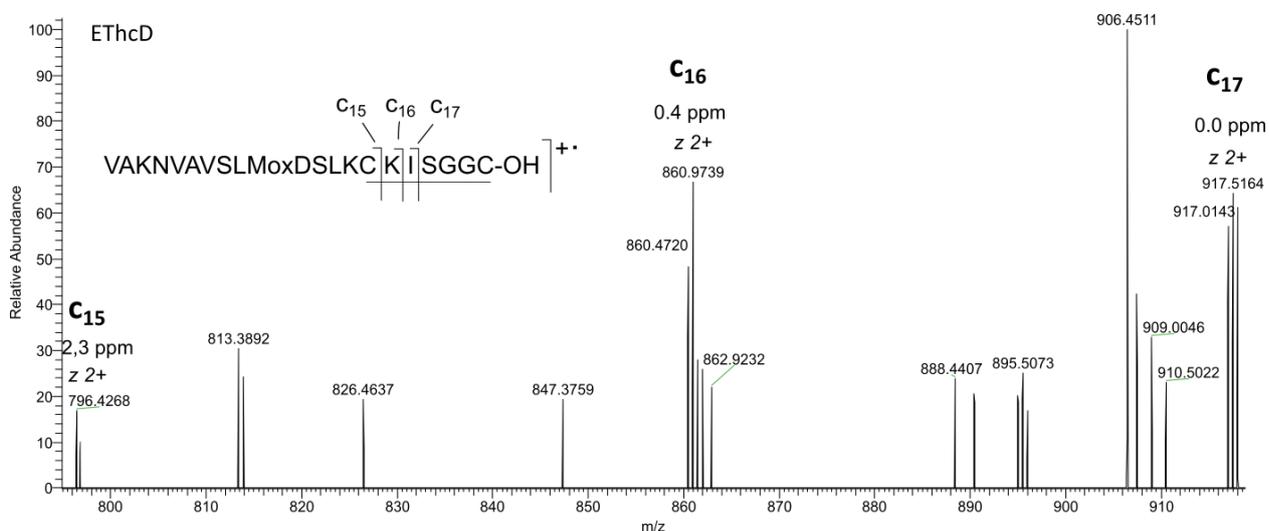
Спектры пересекающихся протеолитических фрагментов дают информацию о всей аминокислотной последовательности бревинина 2AV, за исключением C-концевого фрагмента, внутри «Rana box», которому соответствует  $y_7$ -ион с массой 665.273 Да.

Последовательность внутри «Rana box» была определена по EThcD спектру протеолитического фрагмента бревинина 2AV(13-33) с массой 2136.0726 Да (рис 16). Следует отметить, что C-концевая фрагментация в карбоксамидометилированном бревинине 2AV оказалась чрезвычайно слабой, что характерно для всех бревининов 2 в силу особенностей структур «Rana box» этого пептидного семейства – они, как правило, за редким исключением содержат всего один основной Lys, соседствующий с внутренним цистеином. Поэтому выбранный в работе метод секвенирования этого 33-звенного пептида сверху вниз оказался информативнее. На рис. 16 показана фрагментация только на участке «Rana box» пептида. В результате переноса электрона в EThcD происходит радикальный разрыв дисульфидной связи в протеолитическом фрагменте интактного бревинина 2AV(3-33) с последующей частичной фрагментацией внутри «Rana box» во время дополнительной ДАСПЭ активации. Ионы  $c^{15}$ ,  $c^{16}$ ,  $c^{17}$  спектра EThcD дают следующую последовательность C-концевого фрагмента: SKI(SGGC)-OH<sup>+</sup>. Взаимоположение аминокислот внутри (SGGC) определено из особенностей строения «Rana box» у всех бревининов 2, заключающейся в том, что срединной

аминокислотой 7-ми звенного C-терминального цикла, как правило, является Ser или Thr, а последней аминокислотой Cys.

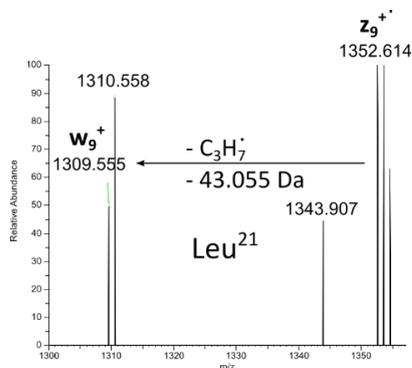


**Рисунок 15.** а – спектр ДАСПЭ N-концевого фрагмента (ММ 1267.743 Да,  $\Delta m = 0.79$  ppm); б – спектр ДАС C-концевого фрагмента, (ММ 2769.481 Да,  $\Delta m = 0.72$  ppm)



**Рисунок 16.** Фрагмент спектра EThcD иона с  $m/z$  713.032<sup>3+</sup> – протеолитического фрагмента бревинина 2AV(3-33) (спектр до экстракции ионов)

Идентификацию в изомерных парах Leu/Ile проводили по характеристическим выбросам радикалов  $C_3H_7^{\cdot}$  (Leu) или  $C_2H_5^{\cdot}$  (Ile) из целевых  $z$ -ионов, содержащих на  $N$ -конце остаток Leu/Ile. На рисунке 17 показан фрагмент спектра EThcD бревинина 2AV (13-33), где видна потеря изопропильного радикала из целевого  $z_9^{+}$ -иона с образованием  $w_9$ -иона. Точность измерения  $w_9$  составляла 3 ppm.



**Рисунок 17.** Участок спектра EThcD протеолитического фрагмента интактного бревинина 2AV (13-33):  $m/z$  713.0320<sup>+</sup>; ETD 50 мс; ДАСПЭ 40 (NCE);  $\Delta m$   $w_9$  = 3 ppm

Суммарная информация ДАС, ДАСПЭ и EThcD спектров даёт следующую последовательность для нового бревинина 2AV:

GLMSVLKKAGKHVAKNVAVSLMDSLKCKISGGC-OH (идентифицированные Leu<sup>21</sup> и Leu<sup>25</sup> выделены жирным шрифтом). Leu<sup>2</sup>, Leu<sup>6</sup>, Ile<sup>29</sup> не были дифференцированы. Отнесение их сделано по структурной аналогии с бревининами 2LT и 2T, выделенными из кожных секретов бурых лягушек *R. latastei* [189] и *R. temporaria* [245] из Словенских популяций:

**Таблица 5.** Последовательности бревининов 2 лягушек европейских популяций

Бревинин 2AV	GLMSVLKKAGKHVAKNVAVSLMDSLK <b>CKISGGC</b> -OH	<i>R. arvalis</i> , Словения
Бревинин 2LT	GLMSVLKKAGKHVAKNVAASLMDSLK <b>CKITGGC</b> -OH	<i>R. latastei</i> , Словения [189]
Бревинин 2T	GLL SGLKKVGVKHVAKNVAV SLMD SLK <b>CKISGDC</b> -OH	<i>R. temporaria</i> , Европа [245]

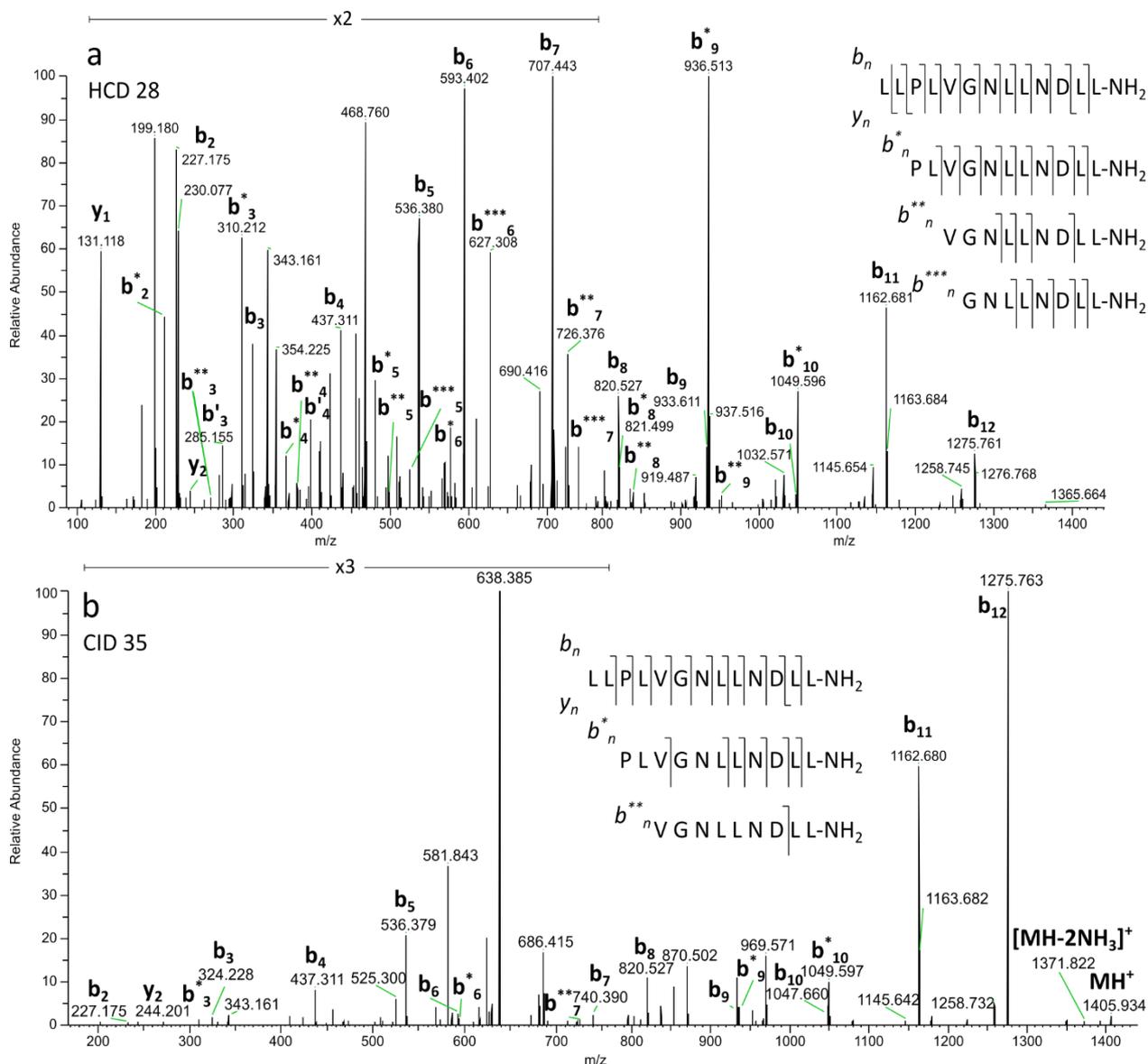
Бревинин 2AV имеет шесть замен по отношению к бревинину 2T (Leu<sup>3</sup>→Met<sup>3</sup>; Gly<sup>5</sup>→Val<sup>5</sup>; Val<sup>9</sup>→Ala<sup>9</sup>; Asp<sup>32</sup>→Gly<sup>32</sup>) и всего лишь две замены в сравнении с бревинином 2LT (Ala<sup>19</sup>→Val<sup>19</sup>; Thr<sup>30</sup>→Ser<sup>30</sup>). Бревинин 2T является антимикробным пептидом широкого спектра действия и уже в микромолярных концентрациях ингибирует рост грамположительных (*Bacillus megaterium*) и грамотрицательных (*Yersinia pseudotuberculosis*; *Escherichia coli*; *Klebsiella pneumoniae*; *Enterobacter agglomerans*) бактерий [246].

**Темпорин AV\*, MM 1404.870 Да**

В секрете словенских особей *R. arvalis* обнаружены два новых темпорина. На рис. 18 а, б приведены ДАСПЭ (NCE 28) и ДАС (NCE 35) спектры темпорина AV, содержащего 13 аминокислотных звеньев, до экстракции.

Темпорины, несмотря на небольшую длину, как правило, дают сложную картину в масс-спектрах из-за вторичной фрагментации интенсивных у-ионов, что затрудняет их *de novo* секвенирование. Ионы вторичной фрагментации отмечены в спектре «\*, \*\*, \*\*\*».

Спектры ДАСПЭ темпоринов, как правило, более многополосные из-за активно протекающих вторичных процессов фрагментации, чем спектры низкоэнергетической активации ДАС. В обоих спектрах на рисунке 18 практически отсутствует серия у-ионов.

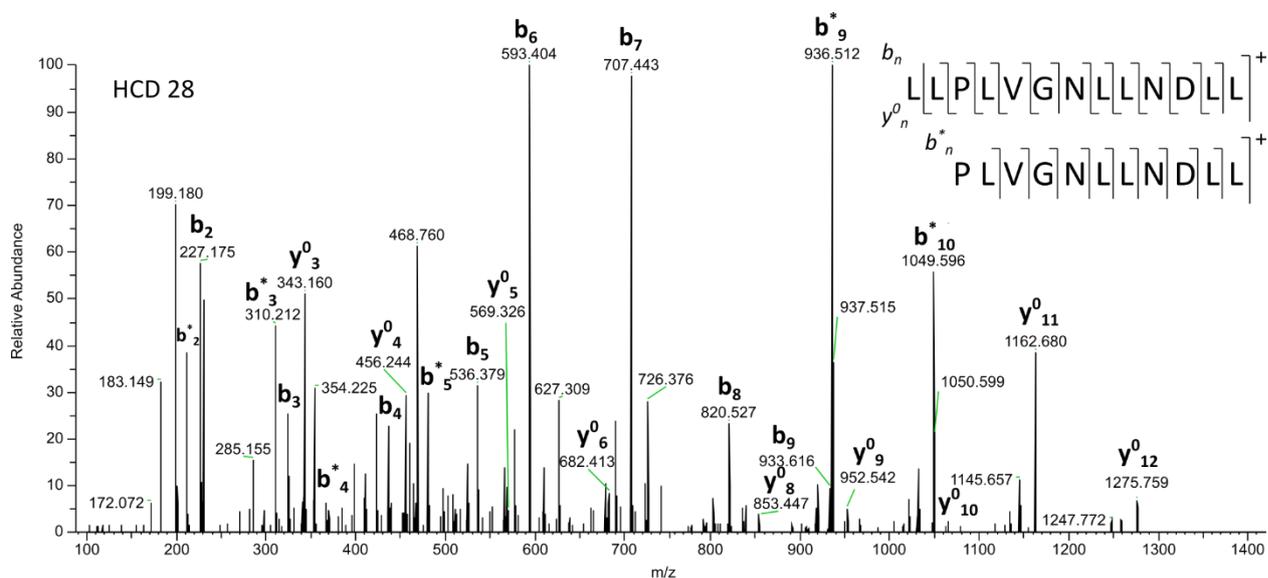


**Рисунок 18.** а – ДАСПЭ (NCE 28), б – ДАС (NCE 35) спектры темпорина AV

При низкоэнергетической активации в спектре ДАС сохраняется протонированная молекула темпорина AV  $MH^+$ , из которого видна потеря двух молекул  $NH_3$  с образованием

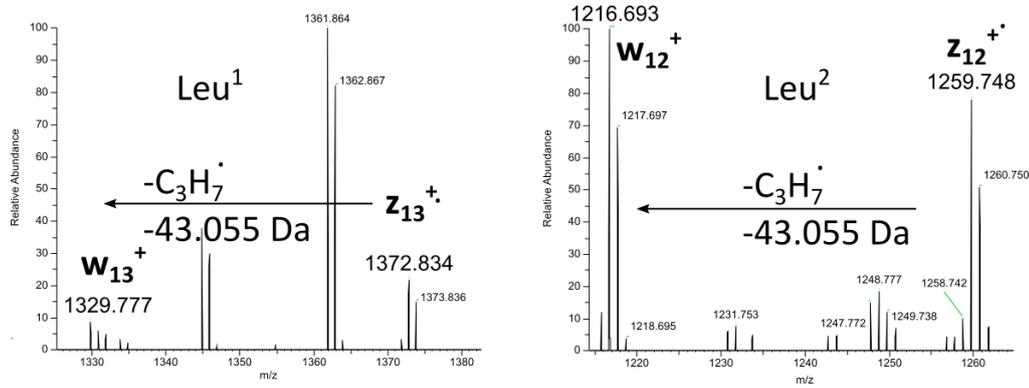
[MH<sup>+</sup> - 2NH<sub>3</sub>]. Элиминирование двух молекул NH<sub>3</sub> происходит, скорее, из боковых цепей двух аспарагинов (Asn<sup>7</sup>, Asn<sup>10</sup>), чем из концевой амидной группы. Косвенно это подтверждается полным отсутствием в спектре y<sup>0</sup><sub>n</sub> ионов (y<sup>0</sup><sub>n</sub> = y<sub>n</sub>-NH<sub>3</sub>). Разрывы амидной связи, образованной с участием Asp<sup>11</sup> (Asp<sup>11</sup> – Leu<sup>12</sup>), проявляются в спектре HCD наиболее интенсивным пиком как в прямой, так и во всех вторичных фрагментациях (Рис. 18a).

В процессе масс-спектрометрического эксперимента протонированные молекулы небольших пептидов могут терять молекулы H<sub>2</sub>O или NH<sub>3</sub> (в зависимости от формы C-конца), что приводит к информативным спектрам, с длинной серией y<sup>0</sup><sub>n</sub> ионов [247]. На рисунке 19 приведен спектр ДАСПЭ иона с m/z 694.928<sup>2+</sup> темпорина AV, потерявшего молекулу аммиака [MH<sup>+</sup>-NH<sub>3</sub>] до экстракции. Точность измерения массы иона составила 2.5 ppm.



**Рисунок 19.** Спектр ДАСПЭ дезамидированного темпорина AV, [MH<sup>+</sup>-NH<sub>3</sub>], (NCE 28), до экстракции ионов

В спектре присутствует практически полный набор *b*-ионов и проходящая до конца серия y<sup>0</sup><sub>n</sub> (y-NH<sub>3</sub>)<sub>n</sub> ионов, что позволяет точно установить *N*-концевую аминокислоту. В спектре ДАСПЭ присутствует также серия вторичных *b*\*-ионов, образовавшихся при распаде y<sup>0</sup><sub>11</sub> (разрыв по N-Pro<sup>3</sup>), включая отсутствующий в прямой фрагментации разрыв Glu<sup>6</sup>-Asn<sup>7</sup>. С учетом всей структурной информации суммы спектров на рис. 18 и 19 мы получаем следующую последовательность темпорина AV: LLPLVGNLLNDLL-NH<sub>2</sub>. *N*-Leu<sup>1</sup> и Leu<sup>2</sup> в структуре темпорина AV были определены по потере *iso*-Pr радикала из целевых ионов z<sub>13</sub> и z<sub>12</sub> в спектре EThcD (показаны в последовательности жирным шрифтом) (Рис. 20). Точность измерения масс w<sub>13</sub> и w<sub>12</sub> ионов составила 1.3 и 1.0 ppm, соответственно. Спектры EThcD были получены фрагментацией ионов MH<sub>2</sub><sup>2+</sup> (m/z 703,443<sup>2+</sup>) и [MH<sub>2</sub> - NH<sub>3</sub>]<sup>2+</sup> (m/z 694.929<sup>2+</sup>) темпорина AV.



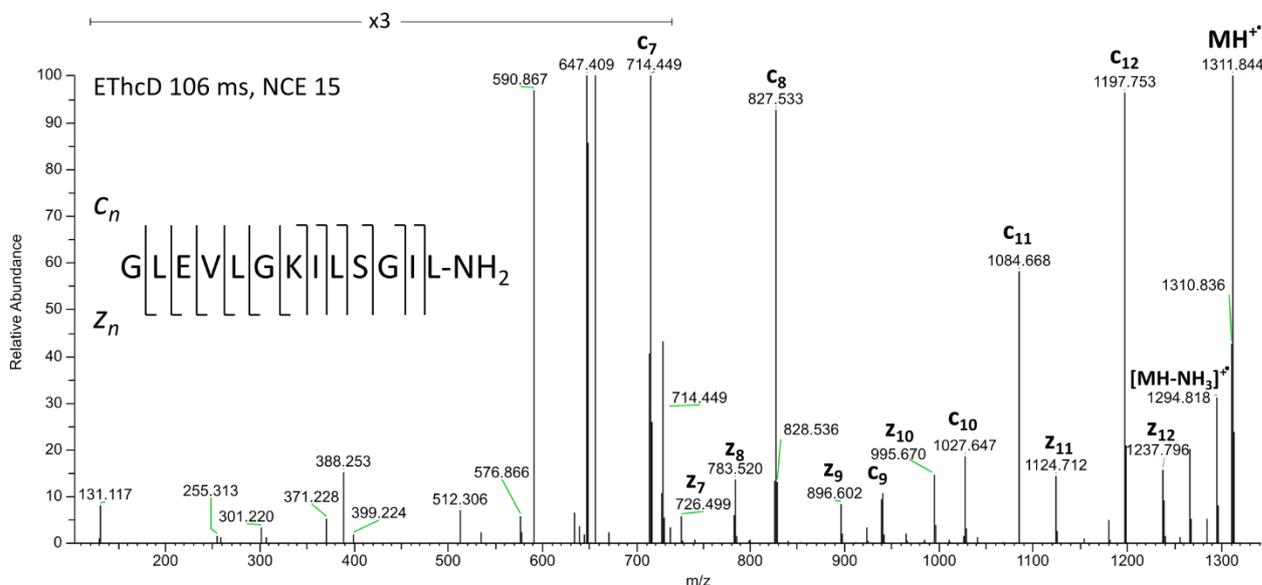
**Рисунок 20.** Фрагменты спектров EThcD для определения Leu<sup>1</sup> (HCD 40) и Leu<sup>2</sup> (HCD 30) в темпорине AV

Темпорин AV структурно близок темпорину 1Ja ILPLVGNLLNDLL-NH<sub>2</sub>, выделенному из кожи японской коричневой лягушки *Rana Japonica*, за исключением замены Leu<sup>1</sup> на Ile<sup>1</sup> [248]. Темпорин 1Ja, как и новый темпорин AV из кожного секрета словенских особей *R. arvalis*, не обладает общим положительным зарядом, но тем не менее, проявляет слабую бактерицидную активность. В концентрации >100 мкМ он ингибирует рост *S. aureus* и *E. coli* [248]. Ранее предполагалось, что темпорины без основных аминокислот в последовательности неактивны. Однако многие из них в концентрациях >60-100 мкМ показывают активность в отношении грамположительных и грамотрицательных бактерий [249]. Например, не имеющие в сиквенсах основных аминокислот темпорины 1PRa и 1PRb способны в концентрациях >100 мкМ ингибировать рост *Pseudomonas aeruginosa*, синегнойной палочки, возбудителя внутрибольничных инфекций, чрезвычайно устойчивых к антибиотикам [250]. А нейтральный темпорин 1Od, выделенный из кожи японской горной коричневой лягушки *Rana ornativentris*, ингибирует рост *S. aureus* на уровне катионных темпоринов: МИК 13 мкМ [249,251].

Таким образом, отсутствие суммарного положительного заряда, считавшегося необходимым для кулоновского взаимодействия пептидов с патогенной мембраной, по-видимому, не является преградой для проявления их антибактериальной активности.

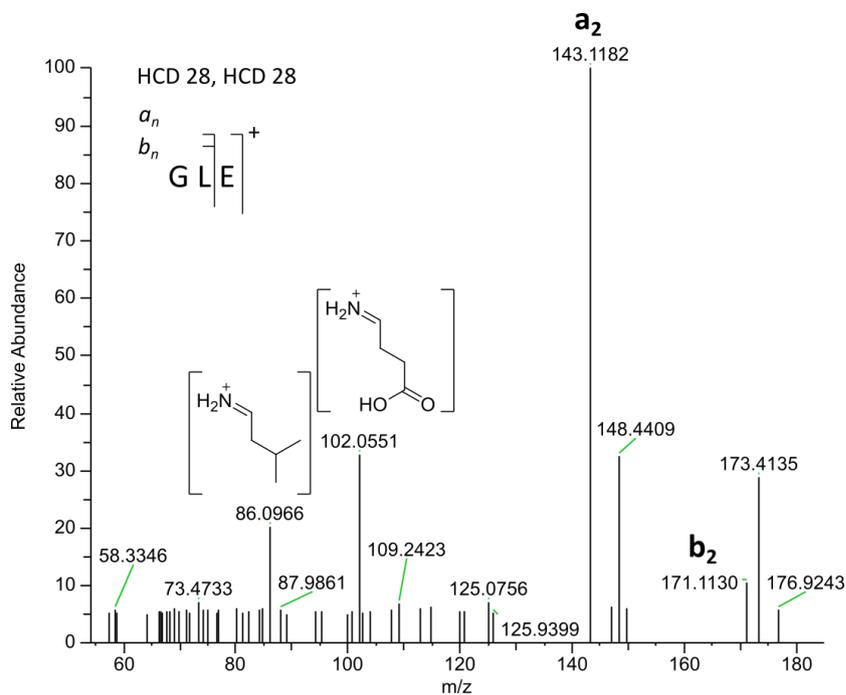
#### **Темпорин AVa\*, MM 1309.833 Да**

На рисунке 21 показан спектр EThcD дипротонированного темпорина AVa с  $m/z$  655.924<sup>2+</sup>, идентифицированного в секрете словенских особей *R. arvalis*. Спектр зарегистрирован при NCE 15. Точность измерения иона составила <1 ppm.



**Рисунок 21.** Спектр EThcD иона  $m/z$  655.924<sup>2+</sup> нового темпорина AVa до экстракции: ДПЭ 106 мс, ДАСПЭ15 NCE.

Серии *c*- и *z*-ионов с высокими значениями  $m/z$  ( $c_7$ - $c_{12}$  и  $z_7$ - $z_{12}$ ) позволяют прочесть всю последовательность темпорина AVa. Присутствие столь необычной для темпоринов глутаминовой кислоты вместо консервативного Pro<sup>3</sup> было подтверждено наличием соответствующего имминиевого иона в MS<sup>3</sup> эксперименте с ионом GLE<sup>+</sup> (рис. 22).

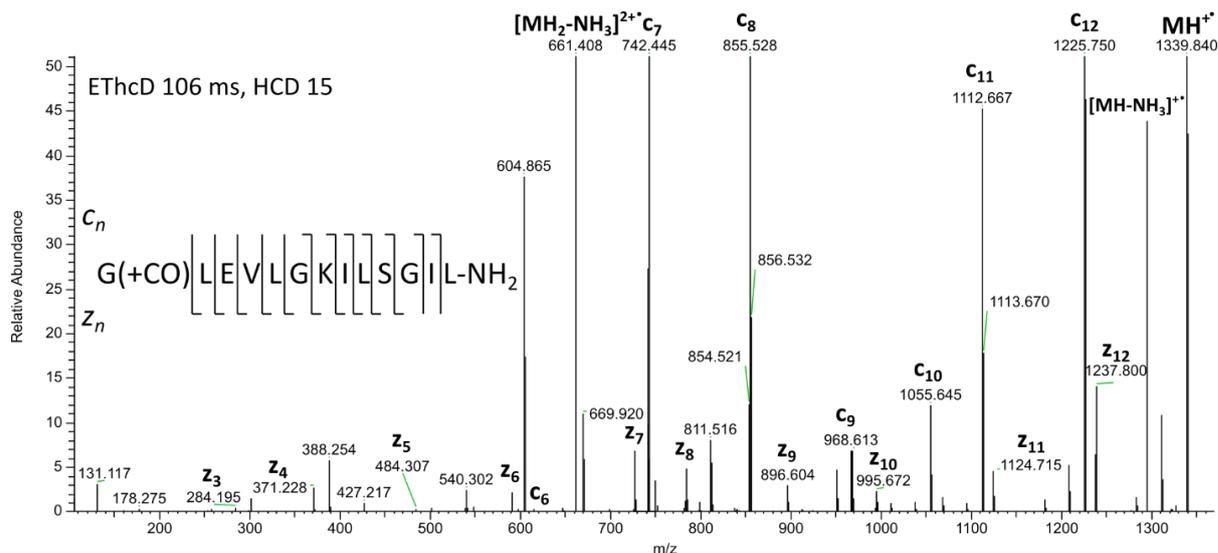


**Рисунок 22.** ДАСПЭ-ДАСПЭ спектр иона  $m/z$  300.1558 темпорина AVa, NCE 28

Спектр получен фрагментацией ДАСПЭ (NCE 28) иона  $b_3$  темпорина AVa, образовавшегося в результате высокоэнергетической активации (ДАСПЭ с NCE 28) протонированного темпорина AVa. В спектре на рис. 22 присутствуют два имминиевых иона с  $m/z$  102.0051 (7.8 ppm) и  $m/z$  86.0966 (3.4 ppm), соответствующих Glu и Leu, соответственно.

Присутствуют также  $b_2/a_2$  ионы с  $m/z$  171.1132 и 143.1182 (точность измерения 2.3 ppm и 2.0 ppm, соответственно), которые свидетельствуют о потере глутаминовой кислоты из иона  $b_3$ .

Замененный на  $\text{Glu}^3$  консервативный  $\text{Pro}^3$  в структуре темпорина AVa следует рассматривать как уникальную замену на сегодняшний день. В литературе описаны два темпорина с заменой  $\text{Pro}^3$  на  $\text{Asp}^3$ : темпорин 1CEa и темпорин 1TGB [252,253], однако сиквенсы этих двух темпоринов были предсказаны клонированием кДНК; в кожных секретах лягушек ранее они не были обнаружены. Темпорин AVa присутствовал в секрете словенских особей и в карбонилированной форме.

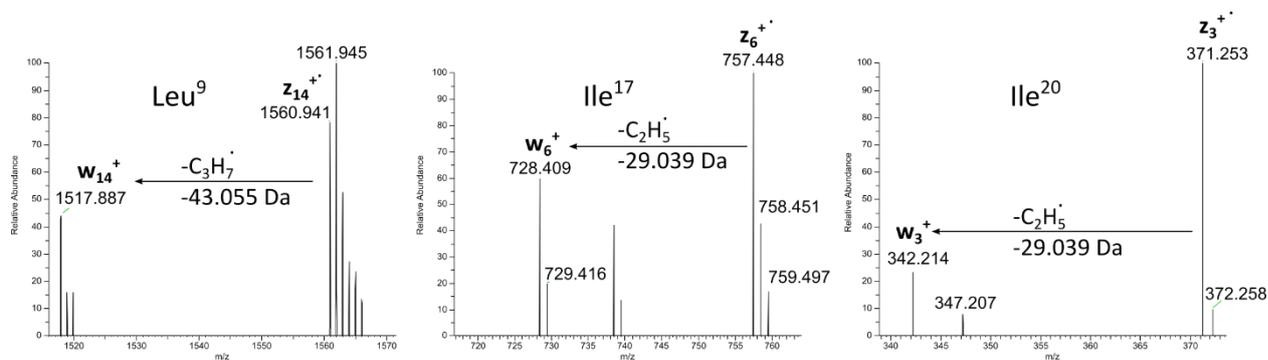


**Рисунок 23.** Спектр EThcD карбонилированного темпорина AVa, до экстракции: ДПЭ 106 мс; ДАСПЭ 15 NCE

Анализ спектра EThcD, приведенного на рис. 23, показал, что карбонилированию подвергся  $N$ -концевой Gly. Более сложной задачей было установление непосредственного места карбонилирования. Сравнение  $b_2$ -ионов в ДАСПЭ спектрах интактного и карбонилированного темпорина AVa (рис. 24), различающихся точно на массу CO, 27.995 Да, ( $m/z$  199.107 и 171.112), показывает, что оба иона легко элиминируют молекулу CO с образованием соответствующих  $a_2$ -ионов, сохраняя ту же разницу масс 27.995, пики  $m/z$  171.112 и 143.117. Хотя оба  $a_2$ -иона всё ещё содержат в своей структуре ещё один фрагмент CO, принадлежащий остатку Gly, он не теряется  $a_2$ -ионом интактного темпорина. Этот факт, вероятно, подтверждает линейную структуру этого  $a_2$ -иона, поскольку циклический ион может легко элиминировать молекулу CO [254]. Напротив,  $a_2$ -ион карбонилированного темпорина отщепляет молекулу CO (рис. 24b). Предполагая линейное строение иона  $a_2$ , можно было бы определить структуру карбонилированного остатка Gly как  $\text{HCO-NH-CH}_2\text{-...}$ . В этом случае  $N$ -концевой CO может быть легко удалён, что сопровождается миграцией водорода от CO к остатку NH.



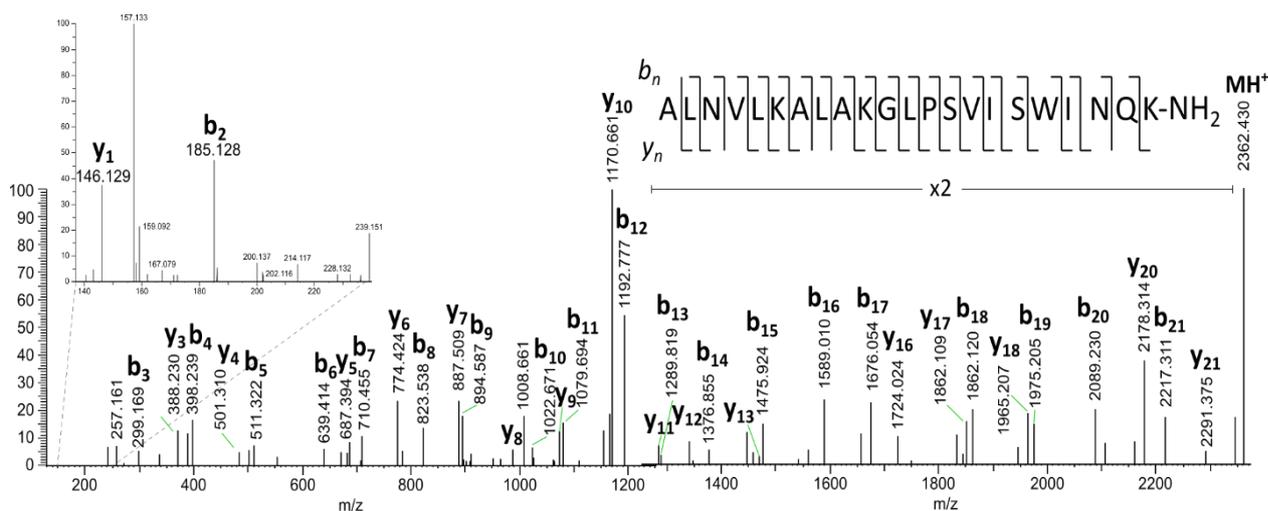
Последовательность FQ-22-1, FVGAALKVLANVLPPVISWIKQ-NH<sub>2</sub>, была установлена в работе [135], когда этот пептид был впервые обнаружен в секретах Московских особей *Rana arvalis*. Отнесение всех пяти его изомерных Leu/Pe проведено в настоящей работе по спектрам EThcD. На рисунке 26 в качестве примера показана идентификация трех изомеров. Точность измерения масс характеристических ионов  $w_{14}$ ,  $w_6$ ,  $w_3$  составила 1.3 ppm, <1 ppm и <1 ppm, соответственно.



**Рисунок 26.** Фрагменты спектра EThcD иона с  $m/z$  788.286<sup>3+</sup> ( $\Delta m$  <2 ppm) FQ-22-1: NCE - 30; ДПЭ – 50 мс

Потеря  $i$ -Pr<sup>•</sup> из  $z_{10}$ -,  $z_{14}$ -,  $z_{17}$ -ионов свидетельствует о присутствии Leu<sup>13</sup>, Leu<sup>9</sup> и Leu<sup>6</sup> в последовательности FQ-22-1, соответственно, а потери C<sub>2</sub>H<sub>5</sub><sup>•</sup> из  $z_3$ - и  $z_6$ -ионов подтверждают в структуре FQ-22-1 Pe<sup>20</sup> и Pe<sup>17</sup> (показаны жирным шрифтом).

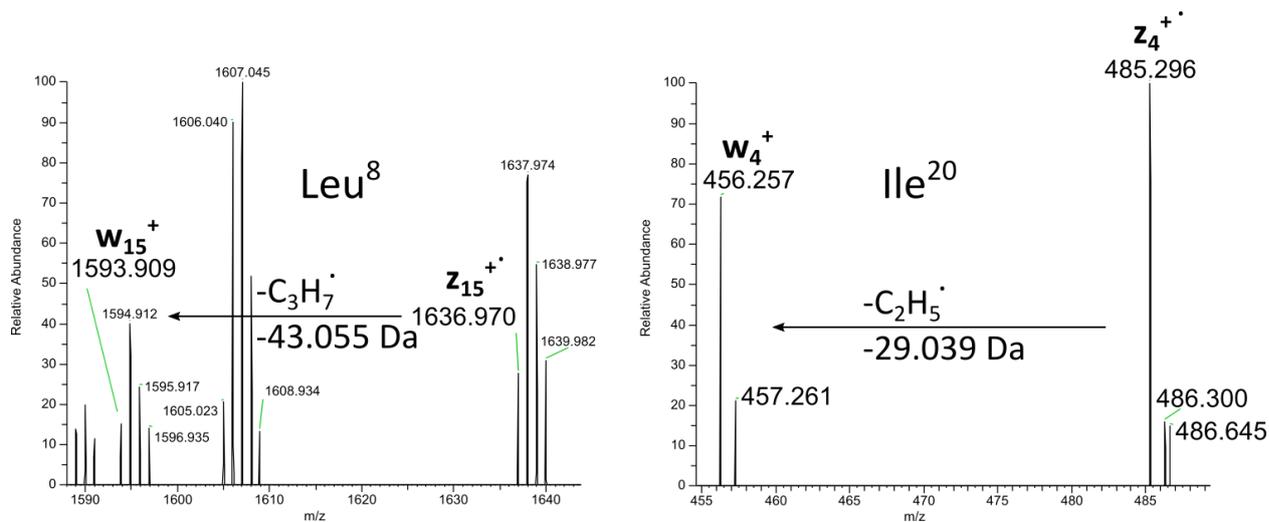
Новый мелиттин-родственный пептид был назван АК-23-1, поскольку АК-23 уже существовал. На рисунке 27 продемонстрирован ДАСПЭ (NCE 28) спектр трижды протонированного пептида АК-23-1 с  $m/z$  788.149<sup>3+</sup>, точность измерения массы <1 ppm.



**Рисунок 27.** Спектр ДАСПЭ иона 788.149<sup>3+</sup> АК-23-1 (NCE 28)

Молекулярные массы обнаруженных мелиттин-родственных пептидов отличаются всего на 0.004 Да, однако хроматографически они разделены почти на 19 минут, поскольку

AK-23-1 имеет 10 замен в последовательности в сравнении с FQ-22-1: ALGGVLKALAKGLPSVISWINQK-NH<sub>2</sub>. Отнесение изомерных Leu/Ile было проведено на основе спектров EThcD, зарегистрированных со ступенчатым увеличением энергии активирующих частиц (NCE 5-40). В качестве примера на рисунке 28 показаны фрагменты спектров EThcD разнозарядных ионов с  $m/z$  788.483<sup>3+</sup>(NCE 40) и  $m/z$  591.615<sup>4+</sup>(NCE 20) полипротонированного пептида АК-23-1, по которым были идентифицированы Leu<sup>8</sup> и Ile<sup>20</sup>. Время экспозиции в ДПЭ составляло 50 мс. Точность измерения масс характеристических  $w_{15}$  и  $w_4$ -ионов равнялась 1.3 ppm и 1 ppm, соответственно.



**Рисунок 28.** Идентификация Leu<sup>8</sup> и Ile<sup>20</sup> в АК-23-1. Слева –  $m/z$  591.615<sup>4+</sup>; справа –  $m/z$  788.483<sup>3+</sup>.

В эксперименте определены все 6 изомерных аминокислот. По характеристическим потерям C<sub>2</sub>H<sub>5</sub><sup>•</sup> из  $z_7$  и  $z_4$  идентифицированы Ile<sup>17</sup> и Ile<sup>20</sup>, а по выбросам C<sub>3</sub>H<sub>7</sub><sup>•</sup> из  $z_{22}$ ,  $z_{18}$ ,  $z_{15}$ ,  $z_{11}$  были установлены Leu<sup>2</sup>, Leu<sup>6</sup>, Leu<sup>9</sup>, Leu<sup>13</sup> в последовательности нового мелиттин-родственного пептида АК-23-1. В экспериментах EThcD по идентификации изомерных аминокислот в паре Leu/Ile мы не наблюдали миграции радикального центра.

Последовательности МРП чрезвычайно изменчивы. Они содержат, как правило, 22-23 аминокислоты, из которых всего три консервативных: Gly<sup>3</sup>, Leu<sup>6</sup>, Leu<sup>9</sup> в положениях последовательности 3, 6 и 9 соответственно. В положениях 14 и 18 происходит единственная замена, а в восьми положениях – две: (2, 8, 10, 12, 13, 17, 19, 21). В таблице 6 суммированы известные на сегодня МРП с указанием года обнаружения, вида лягушки и ссылок на оригинальные работы. Такая таблица возможных замен может оказаться весьма полезной при *de novo* секвенировании новых пептидов этого семейства.

**Таблица 6.** Последовательности существующих мелиттин-родственных пептидов

Вид лягушки	Пептид	Последовательность аминокислот																						
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
<i>Rana temporaria</i> [255]	FQ-22 (1996)	F	I	G	S	A	L	K	V	L	A	G	V	L	P	S	I	V	S	W	V	K	Q	-
<i>Rana tagoii</i> [256]	AR-23 (2003)	A	I	G	S	I	L	G	A	L	A	K	G	L	P	T	L	I	S	W	I	K	N	R
<i>Rana draytonii</i> [257]	RV-23 (2006)	R	I	G	V	L	L	A	R	L	P	K	L	F	S	L	F	K	L	M	G	K	K	V
<i>Rana sakuraii</i> [258]	VR-23 (2007)	V	I	G	S	I	L	G	A	L	A	S	G	L	P	T	L	I	S	W	I	K	N	R
<i>Rana graeca</i> [259]	AK-23 (2019)	A	I	G	A	V	L	K	V	L	A	N	G	L	P	A	L	I	S	F	L	K	E	K
<i>Rana arvalis</i> [135,260]	FQ-22-1 (2009)	F	V	G	A	A	L	K	V	L	A	N	V	L	P	P	V	I	S	W	I	K	Q	-
	AK-23-1 (2021)	A	I	G	G	V	L	K	A	L	A	K	G	L	P	S	V	I	S	W	I	N	Q	K
	LK-22* (2019)	L	L	G	A	A	L	S	A	L	S	S	V	I	P	S	V	I	S	W	F	Q	K	-

\* – последовательность предсказана клонированием кДНК

## 2.2. Сравнительное изучение пептидомов московской и центральнословенской популяций *Rana arvalis*

Пептидомы двух популяций очень близки и содержат одни и те же семейства пептидов: бревинины 1, ранатуерины 2, мелиттин-родственные пептиды и брадикинины. Появление в секрете двух новых семейств, бревинина 2AV и темпоринов, – явление нехарактерное для различных популяций одного вида ранидных лягушек. Это может быть объяснено тем, что все три новых пептида (бревинин 2AV и темпорины AV и AVa) – минорные компоненты кожного секрета словенских лягушек и могли быть зафиксированы благодаря приборным возможностям орбитальной ловушки, продленному времени хроматографирования во время масс-спектрометрического эксперимента, что связано с большим количеством исходных образцов кожных секретов словенских особей, а также отложенному сканированию интенсивных ионов с целью «поймать» минорные компоненты секрета, примененному при анализе словенских образцов.

Таблица 7 позволяет сравнить изменения в компонентном составе кожных пептидомов двух популяций остромордой лягушки: популяцию из Центральной Словении, точнее, окрестностей Любляны, и подмосковную, обитающую в районе звенигородской биостанции МГУ. Столбцы A<sup>S</sup> и A<sup>M</sup> [135] суммируют присутствующие пептиды в секретах словенских и подмосковных особей *R. arvalis*. Для некоторых пептидов указаны их номера в базе данных UniProtKB.

**Таблица 7.** Сравнение составов кожных пептидомов особей из Словенской и Подмосковной популяций *R. arvalis*

№	ММ, Да	Наименование пептида	Последовательность	A <sup>S</sup>	A <sup>M</sup>	UniProtKB, ссылка
1 #	1809.952	Бревинин 1AVa	FLPLLAASFACTVTKKC-OH	+	+	P86159 [135]
2 #	1873.093	Бревинин 1AVb	FVPLLVSKLVCVVTKKC-OH	+	+	P86160 [135]
3 *	3369.812	Бревинин 2 AV	GLMSVLKKAGKHVAKNVAVSLMDSLK CKISGGC-OH	+	-	
4 **	2825.587	Ранатуерин 2AVa	GLLDVVKGAAKNLLASALDKLK CKVTGC-OH	-	+	P86161 [135]
5 #	2896.468	Ранатуерин 2AVb	GLMDMVKGAAKNLFFASALDTLK CKITGC-OH	+	+	P86162 [135]
6 *	2839.603	Ранатуерин 2AVc	GLLDVVKGAAKNLLATALDKLK CKVTGC-OH	+	-	
7 #	2361.430	FQ-22-1	FVGAALKVLANVLPPVISWIKQ-NH <sub>2</sub>	+	+	P86158 [135]
8 *	2361.426	AK-23-1	ALGGVLKALAKGLPSVISWINQK-NH <sub>2</sub>	+	-	
9 *	1404.870	Темпорин AV	LLPLVGNLLNDLL-NH <sub>2</sub>	+	-	
10 *	1309.833	Темпорин AVa	GLEVLGKILSGIL-NH <sub>2</sub>	+	-	
11 #	1059.561	Брадикинин (Br)	RPPGFSPFR-OH	+	+	P84895 [261]
12 **	806.408	[des-Arg <sup>1</sup> , Pro <sup>2</sup> ] Br	PGFSPFR-OH	-	+	
13 #	806.408	[des-Arg <sup>1</sup> ] Br	PPGFSPFR-OH	+	+	
14 **	1187.620	[Arg <sup>0</sup> ]Br	RRPPGFSPFR-OH	-	+	
15 *	1508.800	[Arg <sup>0</sup> ]RA-11	RRPPGFSPFRIA-OH	+	-	
16 #	1243.683	RA-11-2	RPPGFSPFRIA-OH	+	+	
17 *	1087.581	[des-Arg <sup>1</sup> ] RA-11	PPGFSPFRLA-OH	+	-	
18 **	1285.730	RI-11	RPPGFSPFRII-OH	-	+	
19 #	1340.735	RP-12	RPPGFSPFRIAP-OH	+	+	
20 #	1498.804	RS-14-2	RPPGFSPFRIAPAS-OH	+	+	
21 *	1599.852	RT-15	RPPGFSPFRIAPAST-OH	+	-	
22 #	1712.936	RL-16	RPPGFSPFRIAPASTL-OH	+	+	P86156 [71]
23 **	2356.229	RD-21	RPPGFSPFRIAPASTLKR(DE)D-OH	-	+	
24 **	1728.931	[Hyp <sup>3</sup> ]RL-16	RPHypGFSPFRIAPASTL-OH	-	+	
25 **	1225.657	[Glu <sup>5</sup> ]RA-11	RPPGESPFRIA-OH	-	+	

\* – пептид встречается в секрете только у Словенских особей

\*\* – пептид встречается в секрете только у Подмосковных особей

# – пептид встречается в секрете особей в обеих популяциях

Общим для обеих популяций является набор бревининов 1, в который входят бревинин 1AVa и бревинин 1AVb. Вместо ранатуерина 2AVa у словенских особей секретируется его структурный аналог ранатуерин 2AVc, имеющий единственную замену в последовательности: Ser<sup>16</sup>→Thr<sup>16</sup>. Мелиттин-родственный пептид FQ-22-1 также общий для обеих популяций остромордой лягушки.

Пептиды под знаком «\*» характерны лишь для словенских особей, тогда как пептиды под знаком «\*\*» - для особей из подмосковной популяции. Общими для обеих популяций являются те пептиды, которые помечены в таблице знаком «#» – их с осторожностью можно считать маркерами вида *Rana arvalis*.

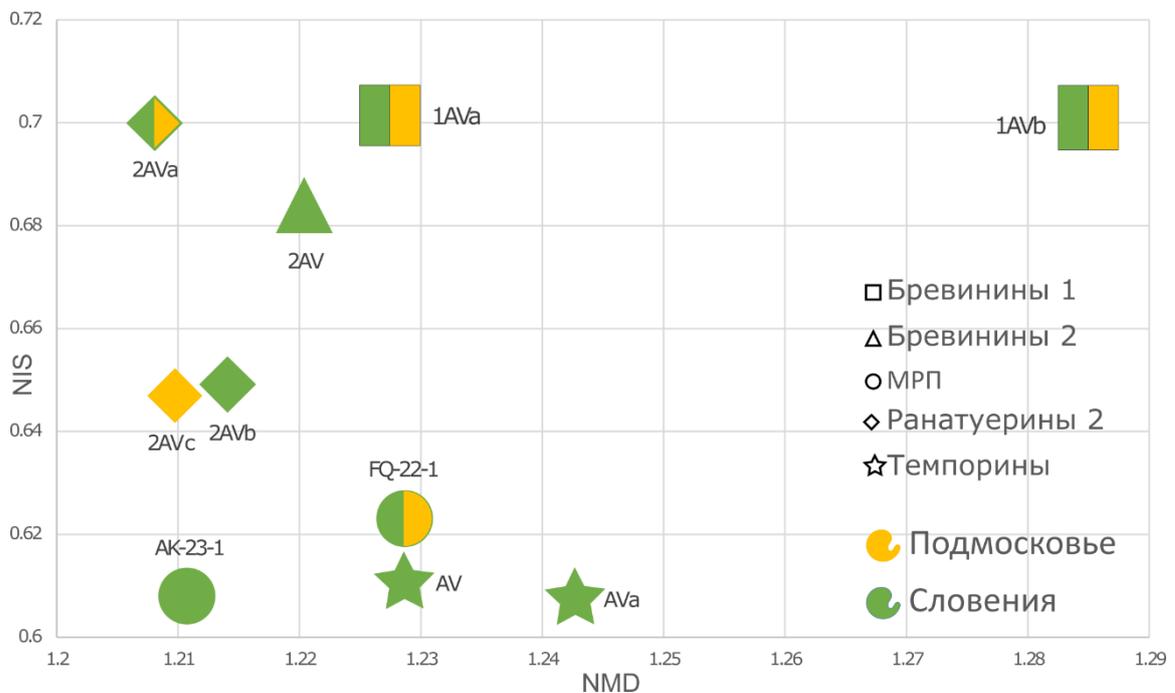
Пять новых пептидов, чьи последовательности докладываются впервые, могут претендовать на роль биомаркеров Словенской популяции *Rana arvalis*: это бревинин 2 AV, ранатуерин 2 AVc, мелиттин-родственный пептид АК-23-1, темпорины AV и AVa. Три

пептида могут претендовать на роль биомаркеров особей из Московской популяции *R. arvalis*: ранатуерин 2AVa, [Hyp<sup>3</sup>]RL-16 и [Glu<sup>5</sup>]RA-11, из них, несомненно, ранатуерин 2AVa – самый значимый претендент.

Все остальные брадикинин-родственные пептиды, обнаруженные в секретах Словенских и Подмосковных особей, являются протеолитическими фрагментами исходного препробрадикинина и лишь отражают путь синтеза брадикинина в организме амфибий.

Построение 2D-карты в координатах NMD-NIS, как описано в соответствующем разделе литературного обзора, позволяет визуализировать популяционные различия у особей остромордых лягушек. На рисунке 29 различными формами обозначены семейства пептидов, обнаруженных в каждом секрете *Rana arvalis*, а разным цветом выделены пептиды, относящиеся к разным популяциям. Если пептид относится к обеим популяциям, то форма маркера окрашена в оба цвета.

Как и было сказано, с помощью 2D-карт можно оценить активность новых пептидов, опираясь на их положение на таком графике. Это даёт основания полагать, что активность нового ранатуерина 2AVc будет крайне схожа с активностью ранатуерина 2AVb. Также такая карта может стать «отпечатком пальца» не только вида, но и популяции – можно увидеть, что ранатуерины 2AVb и 2AVc у Словенской и Подмосковной популяций несколько разнесены. Темпорины AV и AVa, как и мелиттин-родственный пептид АК-23-1, также можно отнести к маркерам Словенской популяции.



**Рисунок 29.** 2D-карта двух популяций *Rana arvalis*

### 2.3. Кожный пептидом травяной лягушки *Rana temporaria* (Центральная Словения)

Список пептидов, идентифицированных в кожном пептидоме особей из окрестностей Любляны, суммирован в таблице 8. Анализ образцов проводился методом «сверху вниз», т.е. без применения трипсинолиза, химических модификаций, исключительно методами самой масс-спектрометрии. Идентифицированные в новых последовательностях изомерные Leu/Pe выделены жирным шрифтом.

**Таблица 8.** Состав кожного пептидома особей *R. temporaria* (Центральная Словения)

№	Пептид	Последовательность	ММ, Да
1	Бревинин 1Т	VNPIILGVLPKFVCLITKKC-OH	2195.294
2	Бревинин 1Та	FITLLLRKFICSITKKC-OH	2024.168
3	Бревинин 1Тб	LVPLFLSKLICFITKKC-OH	1963.1140
4	Бревинин 1Тс*	LVPMFLSKLICFITKKC-OH	1981.096
5	Бревинин 2Т	GLLSGLKKVGVKHAVAKNVAVSLMDSLKCKISGDC-OH	3395.844
6	Бревинин 2Те	GLLDGLKKVGVKHAVAKNVAVSLMDSLKCKISGDC-OH	3423.840
7	Бревинин 2Тf(11-29)	KKFTLNLFHQKCKIGGGC-OH	2132.138
8	FQ-22	FIGSALKVLAGVLPSPVISWVKQ-NH <sub>2</sub>	2310.383
9	Темпорин А	FLPLIGRVLSGIL-NH <sub>2</sub>	1395.896
10	Темпорин В	LLPIVGNLLKSLL-NH <sub>2</sub>	1390.927
11	Темпорин С	LLPILGNLLNGLL-NH <sub>2</sub>	1360.880
12	Темпорин D	LLPIVGNLLNSLL-NH <sub>2</sub>	1376.880
13	Темпорин Е	VLPIIGNLLNSLL-NH <sub>2</sub>	1376.880
14	Темпорин H	LSPNLLKSLL-NH <sub>2</sub>	1095.701
15	Темпорин L	FVQWFSKFLGRIL-NH <sub>2</sub>	1638.939
16	Темпорин N	FXGAXGNAXSRVX-NH <sub>2</sub>	1328.792
17	Темпорин O*	FXGALVNAXRGXX-NH <sub>2</sub>	1354.840
18	Темпорин P*	XVPXXGNXXSGXX-NH <sub>2</sub>	1319.842
19	Темпорин Q*	LVPFXGKTXGGLLARF-NH <sub>2</sub>	1700.046
20	Темпорин R*	XVPFXGRTLGGLLARX-NH <sub>2</sub>	1694.070
21	Темпорин S*	LVPFXGRTXGGLLARF-NH <sub>2</sub>	1728.050
22	Темпорин T*	FXGALVNAXTRVX-NH <sub>2</sub>	1384.851
23	Br	RPPGFSPFR-OH	1059.561
24	[Val <sup>1</sup> ] Br VR-10	VVPPGFSPFR-OH	1101.596
25	Br DR-11	DDRPPGFSPFR-OH	1289.615
26	Br DR-11-1	DVRPPGFSPFR-OH	1273.656
27	Br LR-11-2	LLRPPGFSPFR-OH	1285.729
28	Br LR-10	LRPPGFSPFR-OH	1172.645
29	Br RS-14	RPPGFSPFRIAPAS-OH	1498.804
30	Br RT-15	RPPGFSPFRIAPAST-OH	1599.851
31	Br RL-16	RPPGFSPFRIAPASTL-OH	1713.943
32	Br RA-11	RPPGFSPFRIA-OH	1243.682
33	Br RI-10	RPPGFSPFRI-OH	1172.645
34	Br RP-12	RPPGFSPFRIAP-OH	1340.735
35	Br RD-13	RPPGFSPFRIA VD-OH	1457.777
36	[Thr <sup>6</sup> ] Br	RPPGFTPFR-OH	1073.576
37	[desArg <sup>9</sup> , Thr <sup>6</sup> ] Br	RPPGFTPFR-OH	917.475
38	[Thr <sup>6</sup> ] Br RA-11	RPPGFTPFRIA-OH	1257.698
39	[Thr <sup>6</sup> ] Br RI-10	RPPGFTPFRI-OH	1186.660
40	[Asp <sup>6</sup> ] Br RA-11	RPPGFDPFRIA-OH	1271.677

\* – последовательности докладываются впервые  
неидентифицированные Leu/Pe обозначены в последовательностях как X

*Rana temporaria* – лидер по распространённости среди ранидных лягушек в Европе. Не удивительно, что её кожный секрет подвергся изучению одним из первых. К настоящему

времени он исследован довольно детально, но не исчерпывающе, поскольку полнота его определения зависит от возможностей задействованных приборов и методов [23,28,155,166,255,262,263]. В составе кожного секрета *Rana temporaria* присутствуют пептиды следующих семейств: дисульфидные короткие (17-20 ак) бревинины 1 и длинные бревинины 2 (33 ак) [23,166]; короткие (10-14 ак) амидированные темпорины [23,28]; брадикинин и брадикинин-родственные пептиды (БРП) с большим количеством их *N*- и *C*-расширенных протеолитических копий [151,166,262,263] и мелиттин-родственный пептид (FQ-22) [23].

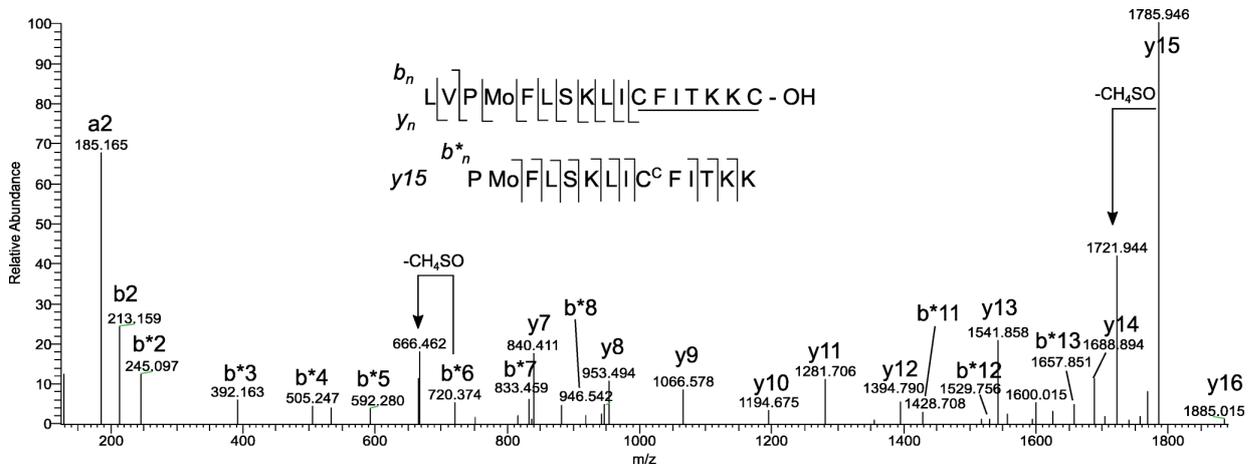
Бревинин 2Т, дисульфидсодержащий пептид, один из основных компонентов секретов особей *Rana temporaria*. Это мощный антибиотик, активный против грамположительных бактерий *Bacillus megaterium* (ЛД 0.2 мкМ) и широкого спектра грамотрицательных бактерий: *Enterobacter agglomerans* (ЛД 2.1 мкМ); *Aeromonas hydrophila* (ЛД 30.0 мкМ); *Klebsiella pneumoniae* (ЛД 0.5 мкМ); *Acinetobacter junii* (ЛД 8.5 мкМ); *Escherichia coli* (ЛД 0.5 мкМ); *Yersinia pseudotuberculosis* (ЛД 0.2 мкМ) [246].

Всего в каждом секрете центральнословенских особей *Rana temporaria* было обнаружено 40 пептидов, 7 из которых оказались новыми. В секретах Словенских лягушек присутствуют 4 коротких дисульфидных пептида из семейства бревининов 1: бревинин 1Т, бревинин 1Та, бревинин 1Тб и новый бревинин 1Тс\*, чья последовательность докладывается впервые. Бревинины 1Т (20 ак) и 1Та (17 ак), пептиды-антибиотики и гемолитики, впервые были выделены из секрета средневропейской (итальянской) лягушки *Rana temporaria* [246].

### ***Бревинин 1Тс\*, ММ 1981.096 Да***

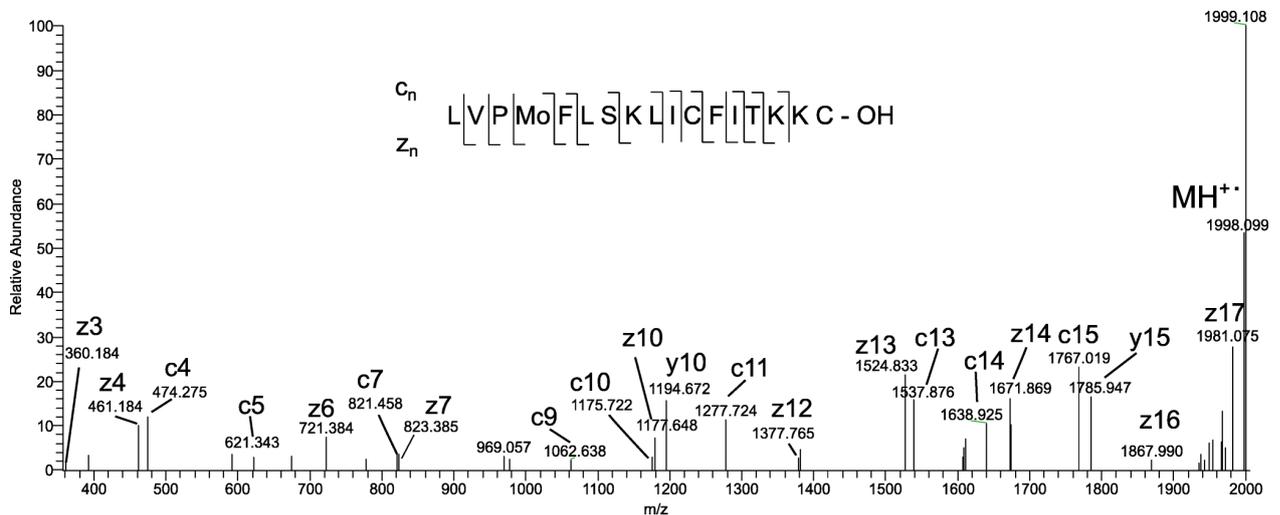
Последовательность нового 17-ти звенного бревинина 1Тс, содержащего *C*-концевой дисульфидный цикл и обнаруженного в секретах Словенских особей *Rana temporaria*, была определена по совокупности спектров ДАСПЭ и EThcD. На рисунке 30 приведен ДАСПЭ спектр трипротонированного бревинина 1Тс, полученного при NCE 28. Точность измерения  $m/z$  666.7033+ составляла 2.3 ppm. Метионин в последовательности интактного бревинина 1Тс\* присутствовал в виде метионинсульfoxида (Met<sub>ox</sub>) – интенсивный ион с  $m/z$  1721.944 соответствует нейтральной потере 63.998 Да из боковой цепи Met<sub>ox</sub>  $y_{15}$ -иона (-CH<sub>4</sub>SO). Спектр ДАСПЭ содержит две серии *b/y* ионов прямой фрагментации MN<sub>3</sub><sup>3+</sup> бревинина 1Тс\* и вторичной фрагментации самого интенсивного иона  $y_{15}$ , образовавшегося при разрыве амидной связи на *N*-конце Pro ( $b_n^*$  на рис. 30). Длинная  $y$ -серия ионов прямой фрагментации ( $y_7$ - $y_{16}$ ) даёт информацию о линейной части бревинина 1Тс\*, а  $y_7$ -ион (840.411 Да) соответствует его дисульфидному 7-звенному циклу. В процессе вторичной фрагментации

иона  $y_{15}$  происходит раскрытие S-S цикла путем разрыва амидной связи Lys<sup>16</sup>-Cys<sup>17</sup> с сохранением самой дисульфидной связи и образованием цистина: Cys<sup>11</sup>Cys<sup>17</sup> [53].



**Рисунок 30.** Спектр ДАСПЭ иона  $666.703^{3+}$  бревинина 1Тс (NCE 28)

В спектре ДАСПЭ на рис. 30 присутствуют  $b^*_{11}$ -,  $b^*_{12}$ - и  $b^*_{13}$ -ионы, которые дают следующую последовательность С-конца бревинина 1Тс: (C<sup>Cys</sup>FI)ТКК. Хотя раскрытие S-S цикла эффективнее протекает в условиях дефицита протонов [53,169], дипротонированного интактного бревинина 1Тс\*, к сожалению, обнаружить не удалось. На рисунке 31 приведен EThcD спектр трёхзарядного иона бревинина 1 Тс\*. Точность измерения  $m/z$   $666.703^{3+}$  составила 0.2 ppm. Время активации ДПЭ составляло 47.16 мс, NCE в ДАСПЭ равнялась 20.



**Рисунок 31.** Спектр EThcD иона  $666.703^{3+}$  бревинина 1Тс (ДПЭ – 47.16 мс, ДАСПЭ – NCE 20)

В условиях EThcD наблюдается раскрытие S-S связи в интактном бревинине 1Тс с активной С-концевой фрагментацией. В результате, длинные серии с- и z-ионов позволяют прочесть всю последовательность бревинина 1Тс, за исключением пары аминокислот (Leu<sup>6</sup>Ser<sup>7</sup>). Таким образом, сумма спектров ДАСПЭ и EThcD даёт полную структурную

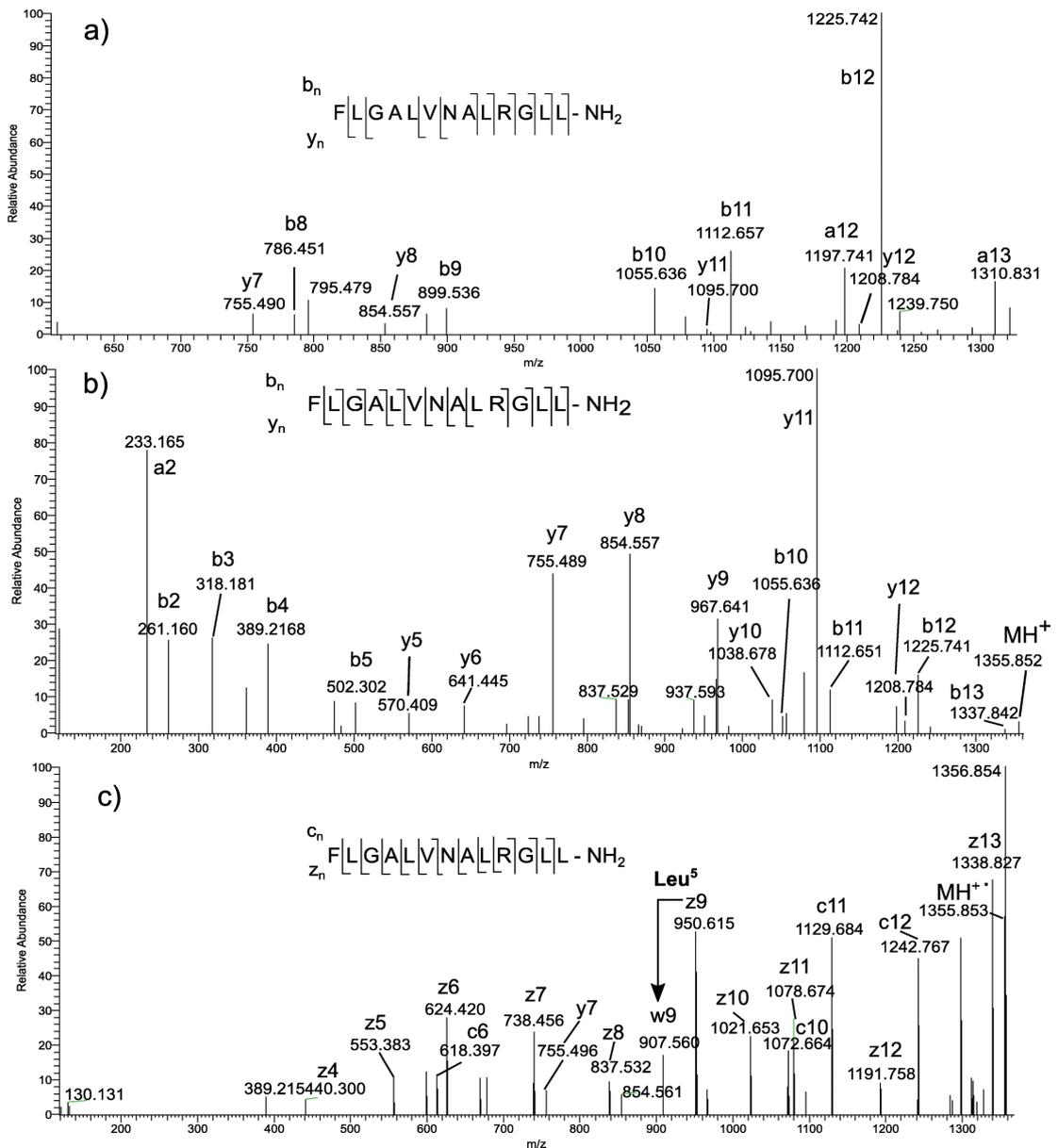
информацию об интактном дисульфидсодержащем бревинине 1Tc\*. Этот пептид отличается всего одной заменой от родственного бревинина 1Tb: Leu<sup>4</sup>→Met<sup>4</sup>.

Leu<sup>9</sup> и Ile<sup>10</sup> в бревинине 1Tc\* отнесены по структурной аналогии с родственными бревинином 1Ta и бревинином 1Tb. Необходимо подчеркнуть, что структура интактного 17-звенного пептида, бревинина 1Tc\*, определена без какого-либо предварительного модифицирования S-S связи, исключительно приборными возможностями орбитальной ловушки, методом секвенирования «сверху вниз». Некоторая заминка возникла с названием этого пептида из-за того, что в работе [246] бревинином 1Tc ошибочно назван бревинин 1T. Учитывая эту неточность и руководствуясь существующей номенклатурой [244], мы назвали новый пептид бревинином 1Tc.

*Rana temporaria* – лидер по содержанию брадикинина RPPGFSPFR-OH в кожных секретах среди ранидных лягушек [26]. В отличие от своего аналога у позвоночных, где брадикинин образуется под действием кинин-калликреиновой системы, кининогены амфибий синтезируются уникальным экзокринным аппаратом кожи в виде неактивных предшественников – препробрадикининов [9]. В результате протеолитического действия эндопротеаз на кининогены в секрете амфибий возникают активный брадикинин и родственные ему пептиды – БРП. БРП воздействуют на гладкую мускулатуру, задействованы в передаче болевого импульса и в ответе организма на воспаление [264]. В секрете словенских особей травяной лягушки присутствует сам брадикинин, набор его N- и C-расширенных копий; а также структурные аналоги брадикинина, имеющие замены аминокислот в самом паттерне брадикинина. Это [Thr<sup>6</sup>]Br с набором C-расширенных копий и [Asp<sup>6</sup>]Br. В секрете словенских особей *Rana temporaria* однако не был обнаружен [Thr<sup>6</sup>, Leu<sup>8</sup>]Br – орнитокинин, агонист орнитокининового рецептора, клонированного Шрёдером с соавторами в 1997 г [154], который наша научная группа ранее обнаруживала в секретах особей Подмосковной популяции [166]. Существует гипотеза, что брадикинин-родственные пептиды амфибий структурно идентичны БРП окружающих их позвоночных, их потенциальных врагов. Таким образом, брадикинин и его структурные аналоги у амфибий выполняют защитную функцию, связываясь с рецепторами напавших хищников, многократно усиливая болевой синдром или рвотный рефлекс. Такая идентичность является причиной природного сохранения консервативности структур брадикининов амфибий [9]. По нашим наблюдениям, число присутствующих в секретах протеолитических расширенных копий самого брадикинина и его структурных аналогов определяется временем протеолитического воздействия эндопротеаз на кининогены, то есть косвенно отражает быстроту и полноту их дезактивации при получении кожного секрета.

**Темпорин О\*, ММ 1354.845 Да**

Один из двух коротких новых темпоринов, состоящих из 13 аминокислотных остатков, обнаруженных в секретах словенских особей, темпорин О\*: FLGALVNALRGLL-NH<sub>2</sub> (ММ 1354.845 Да), чья структура установлена масс-спектрометрическим путем по сумме спектров ДАС, ДАСПЭ и EThcD. Структурно он близок темпорину N, у которого в последовательности также присутствует Arg, а консервативный Pro<sup>3</sup> заменен на Gly<sup>3</sup>. На рисунке 32 приведены ДАС, ДАСПЭ и EThcD спектры дипротонированного темпорина О\* с  $m/z$  678.430<sup>2+</sup>, точность измерения массы которого составила 3.1 ppm.



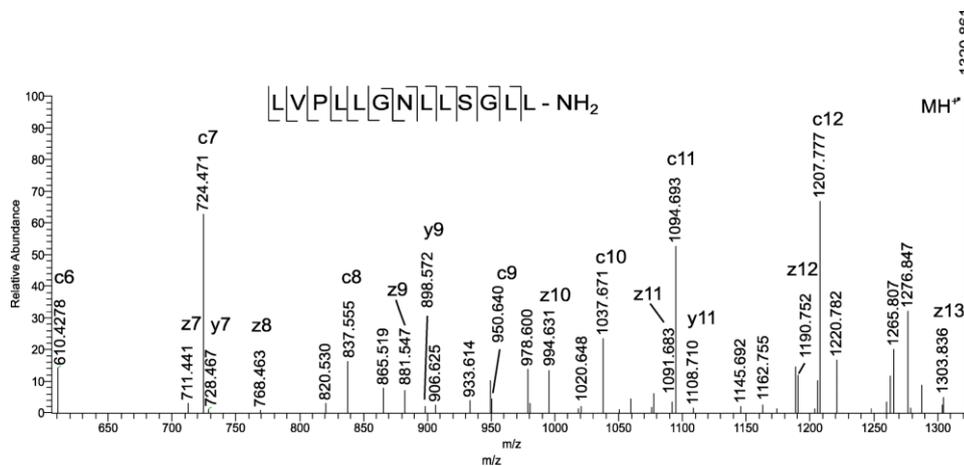
**Рисунок 32.** а – ДАС (NCE 35); б – ДАСПЭ (NCE 28); в – EThcD (ДПЭ – 106,1 мс, ДАСПЭ – NCE 5) спектры иона с  $m/z$  678.430<sup>2+</sup> темпорина О\*

Видно, насколько информативнее в сравнении со спектром низкоэнергетической активации ДАС (рис. 32а) выглядит спектр ДАСПЭ (рис. 32б) темпорина O\*: в нем присутствует больше интенсивных пиков фрагментных ионов, в том числе, и в области низких значений  $m/z$ , а также молекулярный ион темпорина O\*. Сумма спектров ДАС и ДАСПЭ даёт всю последовательность пептида, включая две N-концевые аминокислоты: ион  $y_{12}$  с  $m/z$  1208.784 позволяет уверенно установить разрыв между Phe<sup>1</sup>-Leu<sup>2</sup>. Однако спектр EThcD один позволяет определить полную последовательность темпорина O\*, включая N-концевую аминокислоту, а также идентификацию Leu<sup>5</sup> по характеристической потере изопропильного радикала C<sub>3</sub>H<sub>7</sub>· 43.055 Да из z<sub>9</sub> ( $m/z$  950.616). (Рис. 32с).

Спектр EThcD темпорина O\* оказался более информативным, чем спектр ДПЭ (не показан). Таким образом, в данной работе была отмечена высокая информативность спектров EThcD для *de novo* секвенирования пептидов амфибий.

### Темпорин P\*, MM 1319.854 Да

Ещё один, 13-ти звенный кожный пептид словенских особей *R. temporaria*, темпорин P\*, содержит консервативный Pro<sup>3</sup> и ни одной основной аминокислоты в последовательности. Из-за этой структурной особенности для получения полной последовательности этого пептида было недостаточно суммы двух методов ДАС и ДАСПЭ, даже с учетом протекающих в этих условиях вторичных фрагментаций, поскольку в структуре темпорина P\* есть Pro<sup>3</sup> – т.е. облегченный разрыв амидной связи по N-Pro<sup>3</sup> (связь, образованная третичным атомом азота). Вторичная фрагментация усложняет интерпретацию спектров, но вместе с тем даёт дополнительную структурную информацию о пептиде. Недостающий разрыв в N-концевой паре аминокислот (Leu-Val) удалось получить в спектре EThcD. На рисунке 33 приведен EThcD спектр иона с  $m/z$  660.934<sup>2+</sup> темпорина P (точность измерения массы 0.3 ppm).

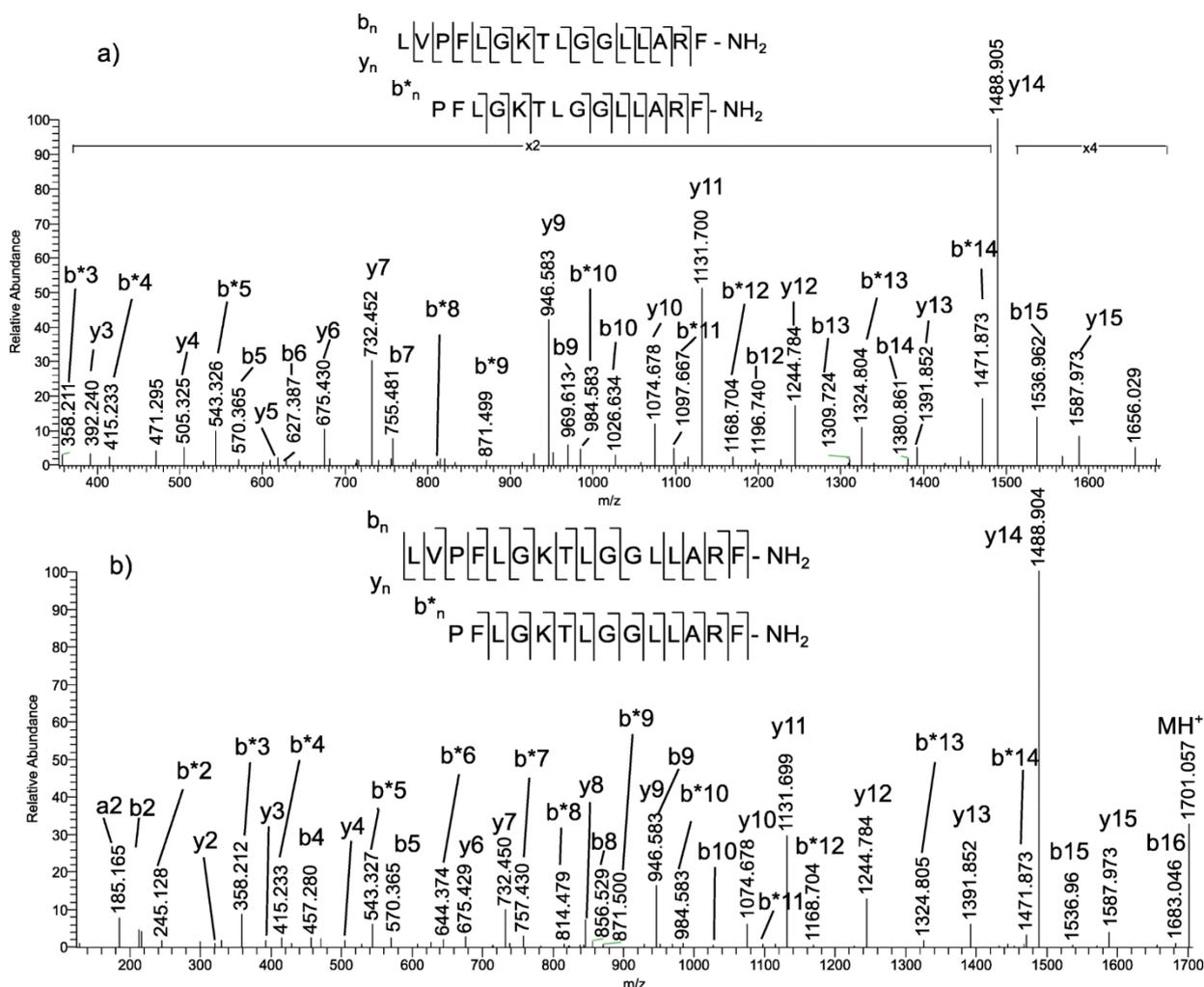


**Рисунок 33.** Спектр EThcD иона с  $m/z$  660.934<sup>2+</sup> темпорина P\* (ДПЭ 106.10 мс, ДАСПЭ NCE 10)

Ион  $z_{12}$  с  $m/z$  1190.752 (0.6 ppm) соответствует NH–CH разрыву пептидной связи между *N*-концевыми Leu<sup>1</sup>-Val<sup>2</sup>. Таким образом, сумма информации спектров (ДАС + ДАСПЭ + EThcD) позволяет прочитать всю последовательность темпорина Р. Идентифицировать изомерные Leu/Ile в темпорине Р не удалось.

### Темпорин Q\*, ММ 1700.046 Да

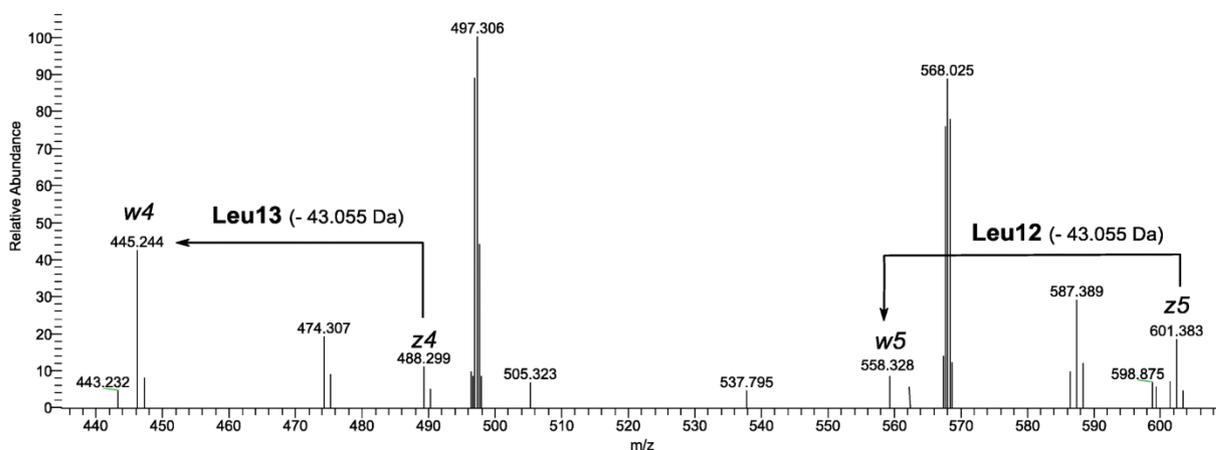
Темпорин Q\* является одним из трёх темпоринов, обнаруженных в секрете словенских особей *Rana temporaria*, имеющих в составе 16 звеньев, в том числе две основные аминокислоты. У темпорина Q\* это Lys и Arg, у темпоринов R\* и S\* – два аргинина. На рис. 34 а, б показаны ДАС и ДАСПЭ спектры двухзарядного иона темпорина Q\* с  $m/z$  851.031<sup>2+</sup> (точность измерения массы иона 1.9 ppm).



**Рисунок 34.** Тандемные масс-спектры темпорина Q\*: а – спектр ДАС иона с  $m/z$  851.031<sup>2+</sup> (NCE 35); б – спектр ДАСПЭ иона с  $m/z$  851.031<sup>2+</sup> (NCE 28)

Структурная информация схематически суммирована на спектрах. Характер фрагментации темпорина Q\* при переходе к высокоэнергетической активации не меняется, о чем свидетельствуют сохраняющиеся соотношения интенсивностей пиков ионов прямой и

вторичной фрагментации в спектрах на рис. 34 а, б. В спектре ДАСПЭ добавляется пик протонированной молекулы и удлиняется серия *b* ионов за счет сохранения в этом методе нижней части масс-спектра. Сумма ДАС и ДАСПЭ спектров позволяет определить всю последовательность темпорина Q\*, включая N-концевую аминокислоту. Идентификация Leu<sup>1</sup>, Leu<sup>12</sup> и Leu<sup>13</sup> проведена с помощью спектра EThcD по соответствующим потерям изопротильных радикалов из целевых z-ионов: z<sub>16</sub> (*m/z* 1684.032), z<sub>5</sub> (*m/z* 601.382) и z<sub>4</sub> (*m/z* 488.299). На рисунке 35 показан фрагмент этого спектра, на котором чётко видны соответствующие потери изопротильных радикалов из целевых z<sub>4</sub> и z<sub>5</sub> ионов, что соответствует Leu<sup>12</sup> и Leu<sup>13</sup>.



**Рисунок 35.** Фрагмент спектра EThcD иона с *m/z* 567.691<sup>3+</sup> темпорина Q\* ( $\Delta m$  0.5 ppm; ДПЭ 47.16 мс; ДАСПЭ NCE 20)

### **Темпорин R\*, ММ 1694.070 Да**

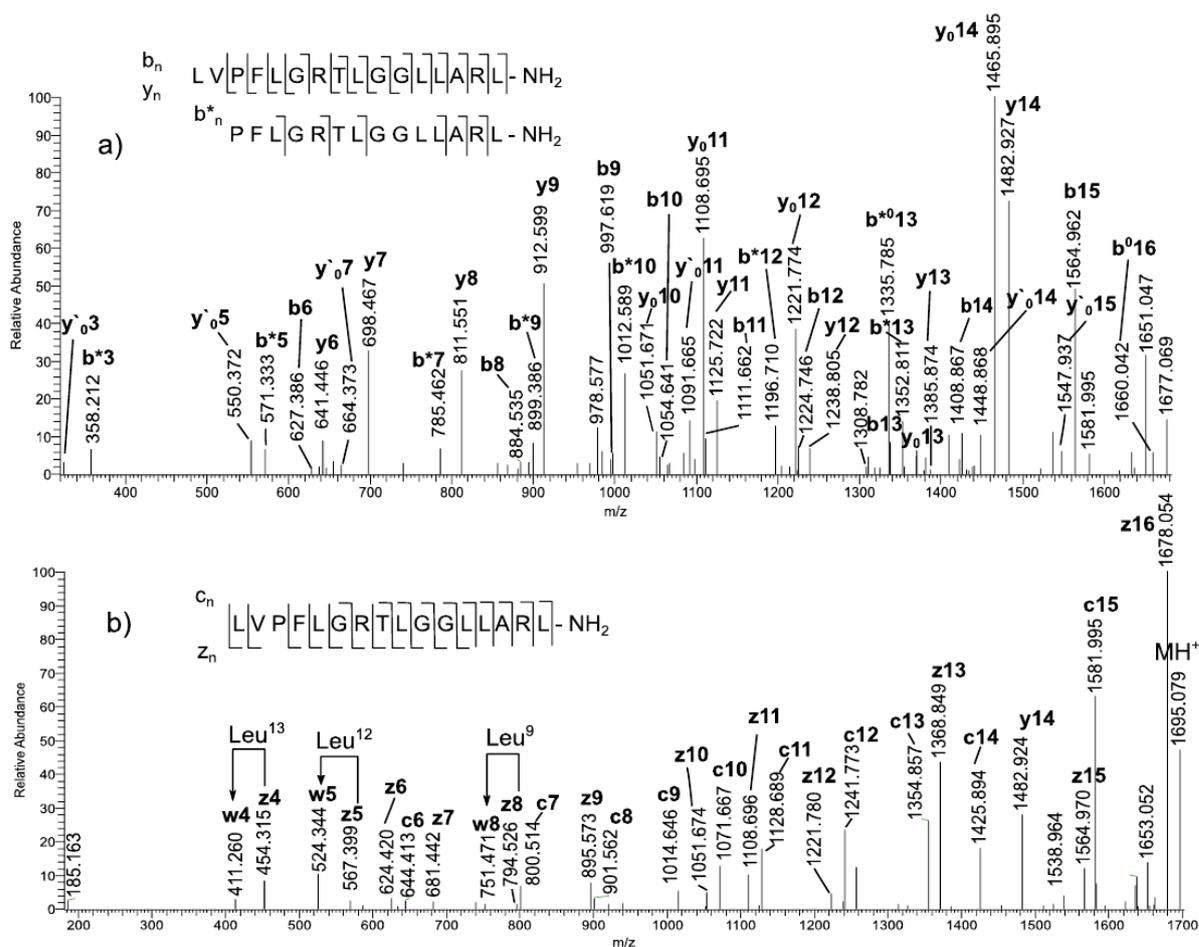
Темпорин R\* и темпорин S\* отличаются от всех известных к настоящему времени темпоринов вида *Rana temporaria*. Это 16-звенные амидные пептиды с двумя аргининами в последовательностях и с сохранившимся консервативным Pro<sup>3</sup>. На рис. 36 приведены спектры ДАС двухзарядного иона темпорина R\* с *m/z* 848.043<sup>2+</sup>, (точность измерения массы составила 0.7 ppm) и EThcD иона с *m/z* 565.698<sup>3+</sup> (0.5 ppm).

Несмотря на то, что ДАС фрагментация двухзарядного иона темпорина R\* протекает в условиях дефицита протона из-за двух аргининов в последовательности (Arg<sup>7</sup>, Arg<sup>15</sup>), в спектре на рис. 37а присутствуют ионы прямой и вторичной (отмечены \*) фрагментации интенсивного y<sub>14</sub> иона (разрыв по N-Pro<sup>3</sup>). Серии *b*- и *y*-ионов прямой и серия *b*\*-ионов вторичной фрагментаций дают структурную информацию о всей последовательности темпорина R\* за исключением взаимоположения в N-концевой паре (Leu-Val).

В спектре на рис. 36а присутствует интенсивный ион с *m/z* 1678.053, соответствующий нейтральной потере молекулы NH<sub>3</sub> из MH<sub>2</sub><sup>2+</sup>-иона темпорина R\*. Большинство аргининсодержащих *y*-ионов в спектре теряют две молекулы NH<sub>3</sub> из-за фрагментации боковых

цепей двух аргининов (помечены  $y_0'$ ). Потеря  $\text{NH}_3$  боковой цепью  $\text{Arg}^{15}$  наблюдается также из  $b_{15}$ - и  $b_{16}$ -ионов ( $b'$ ).

Спектр EThcD помогает уточнить сиквенс темпорина  $R^*$  и идентифицировать три лейцина в его последовательности:  $\text{Leu}^9$ ,  $\text{Leu}^{12}$  и  $\text{Leu}^{13}$  по нейтральным потерям изопропильных радикалов  $\text{C}_3\text{H}_7\cdot$  (43.055 Да) из целевых  $z_8$ ,  $z_5$  и  $z_4$  ионов ( $m/z$  794.526; 567.399; 454.315 соответственно) (рис. 36b). *De novo* секвенирование темпорина  $R^*$  и идентификация  $\text{Leu}^9$ ,  $\text{Leu}^{12}$  и  $\text{Leu}^{13}$  в его последовательности проведены по сумме спектров ДАС и EThcD. Итоговая последовательность темпорина  $R^*$   $\text{XVPFXGRTLGLLLARX-NH}_2$ .



**Рисунок 36.** а – ДАС спектр иона с  $m/z$  848.043<sup>2+</sup> (NCE 35); б – EThcD спектр иона с  $m/z$  565.698<sup>3+</sup> (ДПЭ 50 мс, ДАСПЭ NCE 25) темпорина  $R^*$

### Темпорин $S^*$ , ММ 1728.050 Да

Темпорин  $S^*$  отличается от темпорина  $R^*$  одной заменой  $X^{16} \rightarrow \text{Phe}^{16}$ :  $\text{LVPFXXGRTLXGGLLLARF-NH}_2$ . Он был секвенирован аналогичным образом по сумме спектров ДАС и EThcD двухзарядного и трёхзарядного ионов, соответственно.  $\text{Leu}^1$ ,  $\text{Leu}^{12}$  и  $\text{Leu}^{13}$  идентифицированы по спектру EThcD трёхзарядного иона 577.029<sup>3+</sup> темпорина  $S$  (точность измерения составила 1.7 ppm).

Последовательность минорного *темпорина Т\**, FXGALVNAXTRVX-NH<sub>2</sub>, MM 1384.851 Да, определена при помощи программы PEAKS Studio по сумме спектров ДАС и ДАСПЭ.

## 2.4. Кожный пептидом травяной лягушки *Rana temporaria* (Архангельская область)

В работе был изучен состав кожного пептидома особей *R. temporaria* из окрестностей Архангельска. Обнаруженные в нём пептиды перечислены в таблице 9.

**Таблица 9.** Последовательности пептидов, идентифицированных в кожном пептидоме *R. temporaria* (Архангельская область)

№	Пептид	Последовательность	ММ, Да
1	Бревинин 1Т	VNPIILGVLPKFVCLITKKC-OH	2195.294
2	Бревинин 1Та	FITLLLRKFICSITKKC-OH	2024.168
3	Бревинин 1Тб	LVPLFLSKLICFITKKC-OH	1963.140
4	Бревинин 2Т	GLLSGLKKVVGKHVAKNVAVSLMDSLKCKISGDC-OH	3395.844
5	Бревинин 2Тd	GLWETIKNFGKKFTLNILHNLKCKIGGGC-OH	3229.732
6	FQ-22	FIGSALKVLAGVLPVSVISWVKQ-NH <sub>2</sub>	2310.383
7	Темпорин В	LLPIVGNLLKSLL-NH <sub>2</sub>	1390.927
8	Темпорин С	LLPILGNLLNGLL-NH <sub>2</sub>	1360.880
9	Темпорин D	LLPIVGNLLNSLL-NH <sub>2</sub>	1376.880
10	Темпорин F	FLPLIGKVL SGLL-NH <sub>2</sub>	1367.890
11	Темпорин G	FFPVIGRILNGIL-NH <sub>2</sub>	1456.891
12	Темпорин H	LSPNLLKSLL-NH <sub>2</sub>	1095.701
13	Темпорин K	LLPNLLKSLL-NH <sub>2</sub>	1121.752
14	Темпорин L	FVQWFSKFLGRIL-NH <sub>2</sub>	1638.939
15	Темпорин M	FLPILGKVL SRVL-NH <sub>2</sub>	1452.954
16	Темпорин N	FLGALGNALSRVL-NH <sub>2</sub>	1328.792
17	Темпорин T	FXGALVNAXTRVX-NH <sub>2</sub>	1384.851
18	Темпорин U*	FIGALVHAXTGXX-NH <sub>2</sub>	1322.808
19	Темпорин V*	FIGALVHAXTSXX-NH <sub>2</sub>	1352.820
20	Темпорин W*	FIGALVNALSRVX-NH <sub>2</sub>	1370.839
21	Темпорин X*	FLGALVGALSRVX-NH <sub>2</sub>	1313.818
22	Br	RPPGFSPFR-OH	1059.561
23	Br RS-14	RPPGFSPFRIAPAS-OH	1498.804
24	Br RT-15	RPPGFSPFRIAPAST-OH	1599.851
25	Br RL-16	RPPGFSPFRIAPASTL-OH	1712.936
26	Br RA-11	RPPGFSPFRIA-OH	1243.682
27	Br RI-10	RPPGFSPFRI-OH	1172.645
28	Br RP-12	RPPGFSPFRIAP-OH	1340.735
29	[Thr <sup>6</sup> ] Br	RPPGFTPFR-OH	1073.576
30	[Thr <sup>6</sup> ] Br RA-11	RPPGFTPFRRIA-OH	1257.698
31	[Thr <sup>6</sup> ] Br RI-10	RPPGFTPFRIOH	1186.660

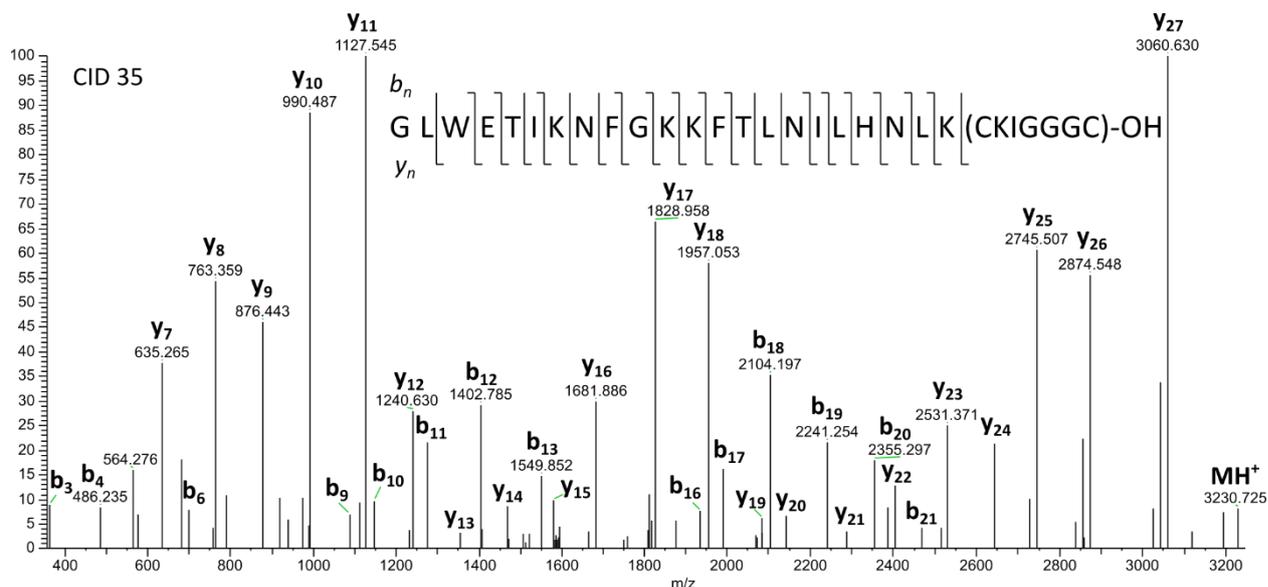
\* – последовательности докладываются впервые

Пептидом архангельских особей травяной лягушки содержит в кожном пептидоме те же семейства пептидов, что и словенские особи: бревинины 1, бревинины 2, темпорины – это все антимикробные пептиды широкого спектра действия. У них так же присутствует мелиттин-родственный пептид FQ-22, мощный цитолитик, как и большинство пептидов из семейства бревининов 1. Брадикинин-родственные пептиды у архангельских лягушек

представлены самим брадикинином и его протеолитическими копиями, которые имеют исключительно *C*-расширение. В секрете присутствует всего один структурный аналог брадикинина, имеющий одну замену в последовательности, Ser<sup>6</sup>→Thr<sup>6</sup>, ([Thr<sup>6</sup>])Br с двумя своими протеолитическими копиями, также с *C*-концевыми расширениями.

### **Бревинин 2Td, ММ 3229.732 Да**

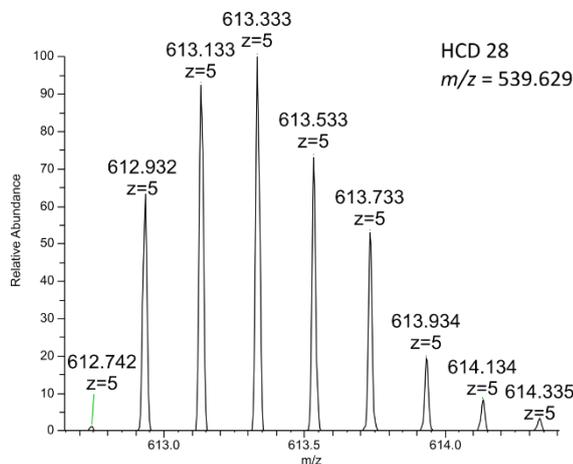
О бревинине 2Td, выделенном из кожи европейской лягушки *R. temporaria*, было сообщено в 1998 году итальянским учёным М. Симмако с коллегами [23]. Чёткого понимания, однако, как была получена его последовательность: клонированием кДНК или выделением из секрета не было. Во всяком случае, Дж. Х. Бови, австралийский ученый, не включил эту последовательность в свой обзор по пептидам амфибий, посчитав, что она не была обнаружена в секрете *R. temporaria* [25]. Транскриптом не равен каждому пептидому, который всегда меньше и который секретируется амфибией в соответствие с окружающими патогенами в каждой конкретной климатической нише, по гипотезе Дж. Теннессена [265,266]. Мы, следовательно, впервые обнаружили бревинин 2Td в секрете архангельских особей *R. temporaria* (рисунок 37).



**Рисунок 37.** Спектр ДАС иона с  $m/z$  808.439<sup>4+</sup> бревинина 2Td (NCE 35)

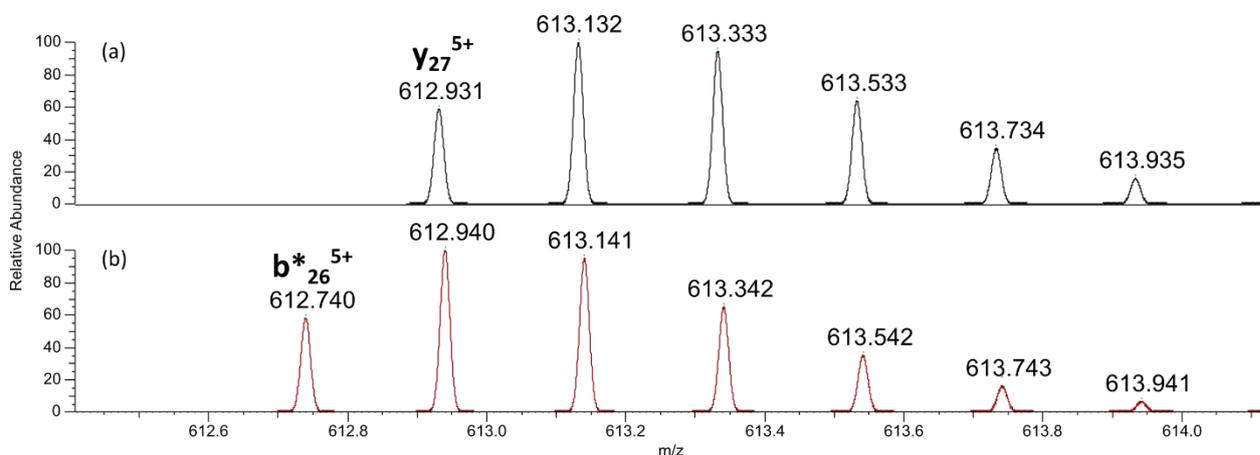
Спектр диссоциации, активированной соударениями, даёт практически полную последовательность этого 33-звенного пептида, за исключением взаимоположения *N*-концевой пары аминокислот и «Rana box». Как видно из спектра, в условиях ДАС в бревинине 2Td не происходит раскрытия дисульфидного цикла по механизму, указанному на схеме 1 литературного обзора. Однако спектр ДАСПЭ выявил потерю из молекулярного иона бревинина 2Td частицы массой 171.064 Да, что соответствует элиминированию трех *C*-

концевых остатков Gly, происходящей при указанной выше фрагментации изомерного молекулярного иона с раскрытым циклом GLWETIKNFGKKFTLNILHNLK<sup>C</sup>KIGGG (см. схему 1 Литературного обзора). Это подтверждается присутствием в спектре ДАСПЭ бревинина 2Td соответствующего  $b_{25}^*$  иона (кластер пятизарядного иона с  $m/z$  мизотопного пика  $612.724^{5+}$  (рис. 38)



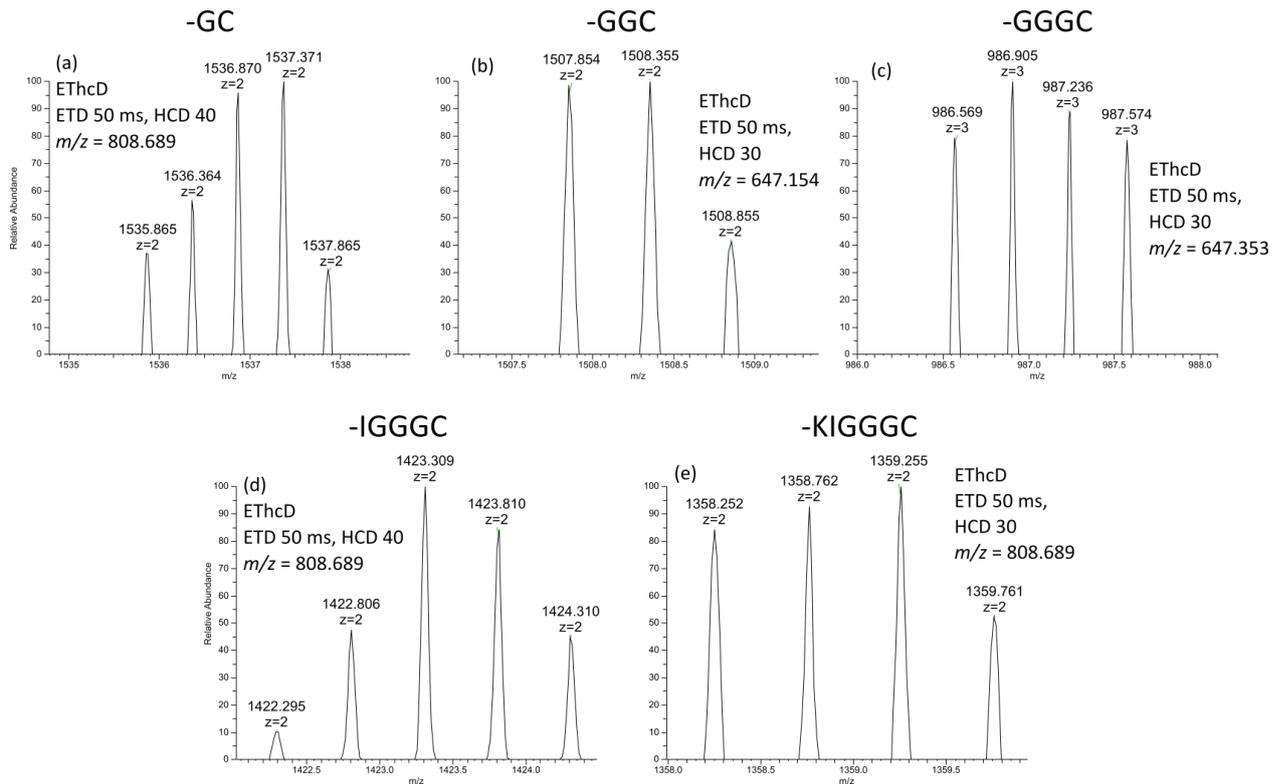
**Рисунок 38.** Кластер иона  $612.742^{5+}$  (точность измерения масс 4 ppm), соответствующий потере GGG из изомерного молекулярного иона бревинина 2Td

Интенсивность кластера довольно высока – это связано с перекрыванием в данном случае пятизарядных кластеров двух ионов –  $y_{27}^{5+}$ , образованного выбросом GL из молекулярного иона бревинина 2Td, и обозначенного выше  $b_{26}^{*5+}$ -иона изомерного пептида с цистином в последовательности – GLWETIKNFGKKFTLNILHNLK<sup>C</sup>KIGGG. Предсказание формы кластера (рис. 39) также подтверждает, что в случае этих ионов пики кластеров перекрываются. Обнаружить потерю GGG становится возможно только благодаря высокой чувствительности и разрешающей способности прибора.



**Рисунок 39.** Изотопная симуляция кластеров ионов: (a) –  $y_{27}^{5+}$ -ион бревинина 2Td (брутто-формула  $C_{140}H_{227}N_{38}O_{35}S_2$ ); (b) –  $b_{26}^{*5+}$ -ион изомерного бревинина 2Td (GLWETIKNFGKKFTLNILHNLK<sup>C</sup>KIGGG, брутто-формула  $C_{142}H_{232}N_{37}O_{34}S_2$ )

Уверенно установить последовательность внутри интактного «Rana box» бревинина 2Td позволяют его спектры EThcD, полученные при NCE 30-40 (рис. 40):



**Рисунок 40.** Фрагменты EThcD спектров бревинина 2Td: (a)  $c_{27}^{2+}$ -ион, точность измерения масс  $\Delta m$  1 ppm; (b)  $c_{26}^{2+}$ -ион,  $\Delta m$  2 ppm; (c)  $c_{25}^{3+}$ -ион,  $\Delta m$  2 ppm; (d)  $c_{24}^{2+}$ -ион,  $\Delta m$  6 ppm; (e)  $c_{23}^{2+}$ -ион,  $\Delta m$  2 ppm

Потери из протонированного бревинина 2Td -(GC), -(GGC), -(GGGC), -(IGGGC) и -(KIGGGC) позволяют установить последовательность «Rana box» как SKIGGGC, что совпадает с ранее предсказанной клонированием кДНК европейской *R. temporaria* последовательностью [23]. Для некоторых кластеров многозарядных ионов, в частности, изображённых на рис. 40b и 40d, не видны низкоинтенсивные моноизотопные пики – это связано с тем, что при регистрации тандемного масс-спектра выбирается наиболее интенсивный пик. Для соединений, содержащих в своём составе более ста атомов углерода, наиболее интенсивным будет уже не моноизотопный пик, и, выбрав такой ион для MS/MS, мы получаем несколько искажённое распределение изотопов. Однако массы измеряются с высокой точностью, что позволяет однозначно устанавливать разность масс между фрагментными ионами, а значит, с высокой достоверностью устанавливать аминокислотные остатки в последовательности, зарядность кластеров и предсказывать массу моноизотопных пиков.

### Темпорин U\*, MM 1322.808 Да

В секрете архангельской травяной лягушки обнаружены четыре новых темпорина. На рисунке 41 приведен спектр EThcD 13-звенного темпорина U\*. Время экспозиции при переносе электрона составляло 50 мс, нормализованная энергия NCE в ДАСПЭ равнялась 40.

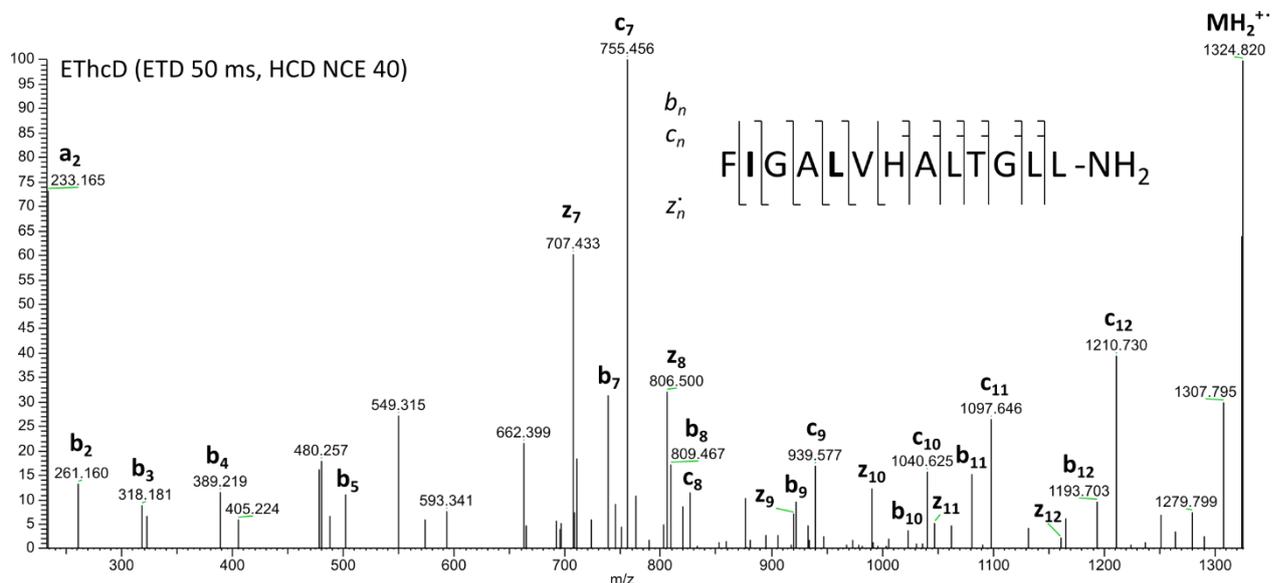


Рисунок 41. Спектр EThcD иона с  $m/z$  662.411<sup>2+</sup> темпорина U\*

В спектре темпорина U\* присутствуют три серии ионов: комплементарные, образующиеся при разрыве связей NH-C $\alpha$ , z<sub>n</sub>/c<sub>n</sub> и b<sub>n</sub>. Их сумма даёт информацию о полной последовательности темпорина U\* и позволяет идентифицировать Pe<sup>2</sup> по потере этильного радикала целевым z<sub>12</sub> ионом и Leu<sup>5</sup> по потере изопропильного радикала целевым z<sub>9</sub>-ионом (рис. 42).

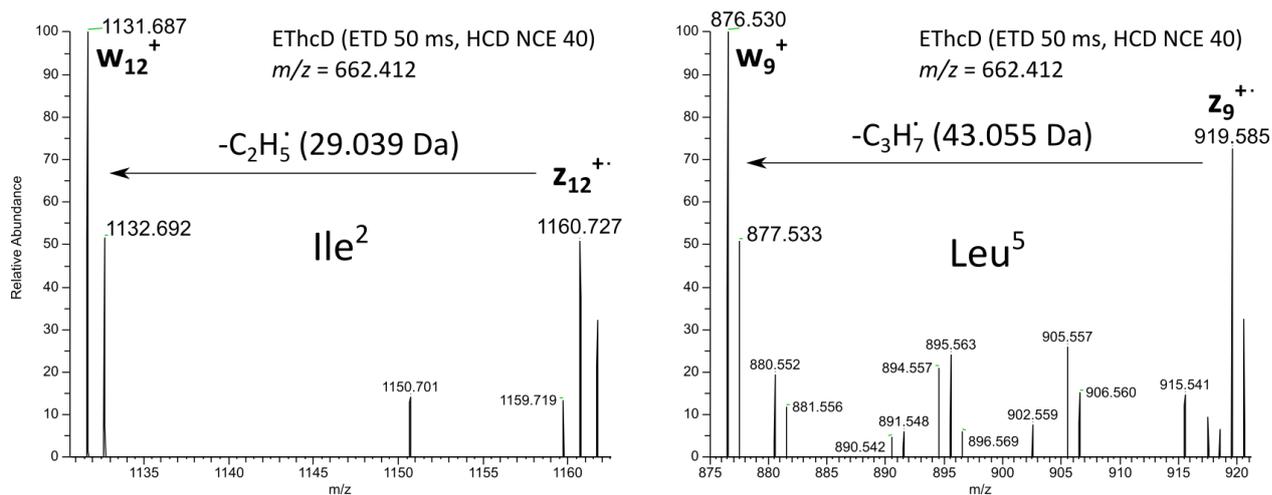
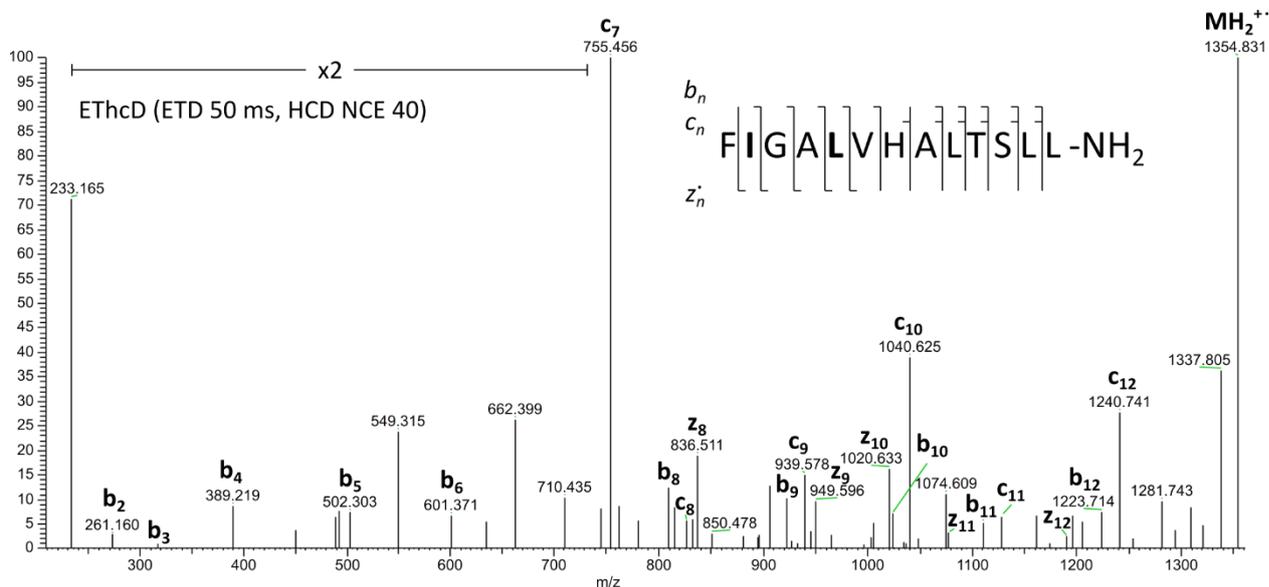


Рисунок 42. Идентификация Pe<sup>2</sup> и Leu<sup>5</sup> в темпорине U\* в спектрах EThcD

Поиск в системе BLAST (UniProt) совпадений с последовательностью темпорина U\* не выявил.

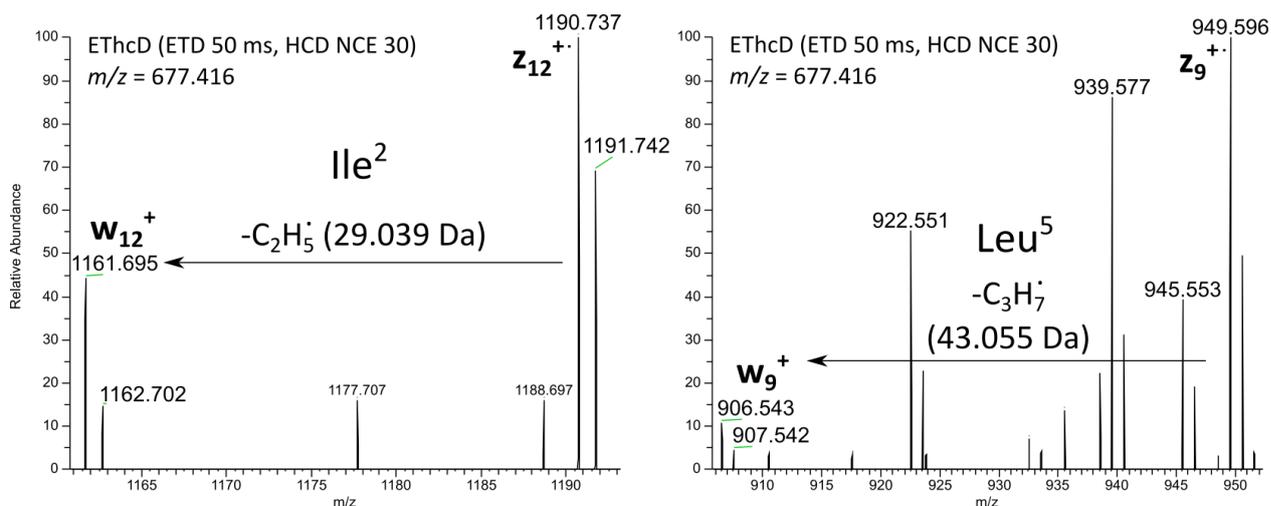
### Темпорин V\*, MM 1352.820 Да

Это второй His-содержащий 13-звенный темпорин, выявленный в секрете архангельских особей *R. temporaria*. На рисунке 43 показан EThcD спектр дипротонированного темпорина V\*. Параметры получения спектра аналогичны указанным выше для темпорина U\*.



**Рисунок 43.** Спектр EThcD иона с  $m/z$  677.417<sup>2+</sup> темпорина V\*: ДПЭ – 50 мс; ДАСПЭ – NCE 40

Три серии образующихся фрагментных ионов в спектре EThcD дают полную последовательность этого короткого пептида. Ile<sup>2</sup> и Leu<sup>5</sup> идентифицированы по характеристическим выбросам этильного и изопропильного радикалов из целевых z<sub>12</sub>- и z<sub>9</sub>-ионов соответственно (рис. 44). Поиск в BLAST (UniProt) не выявил совпадений.



**Рисунок 44.** Идентификация Ile<sup>2</sup> и Leu<sup>5</sup> в темпорине V\* в спектрах EThcD

Темпорины U\* и V\* структурно близки по числу звеньев, по наличию в последовательностях His<sup>7</sup> вместо традиционных у темпоринов в этой позиции Lys<sup>7</sup>/Arg<sup>7</sup>, у

обоих консервативный Pro<sup>3</sup> заменен на Gly<sup>3</sup> и присутствует Thr<sup>10</sup>. Это даёт основание ожидать у этих двух темпоринов схожих биологических активностей.

### Темпорин W\*, MM 1370.839 Да

Последовательность темпорина W\* из кожных секретов архангельских особей *R. temporaria* установлена по его фрагментации в условиях EThcD: время экспозиции при переносе электрона составило 50 мс, нормализованная энергия NCE в ДАСПЭ равнялась 40 (рисунок 45).

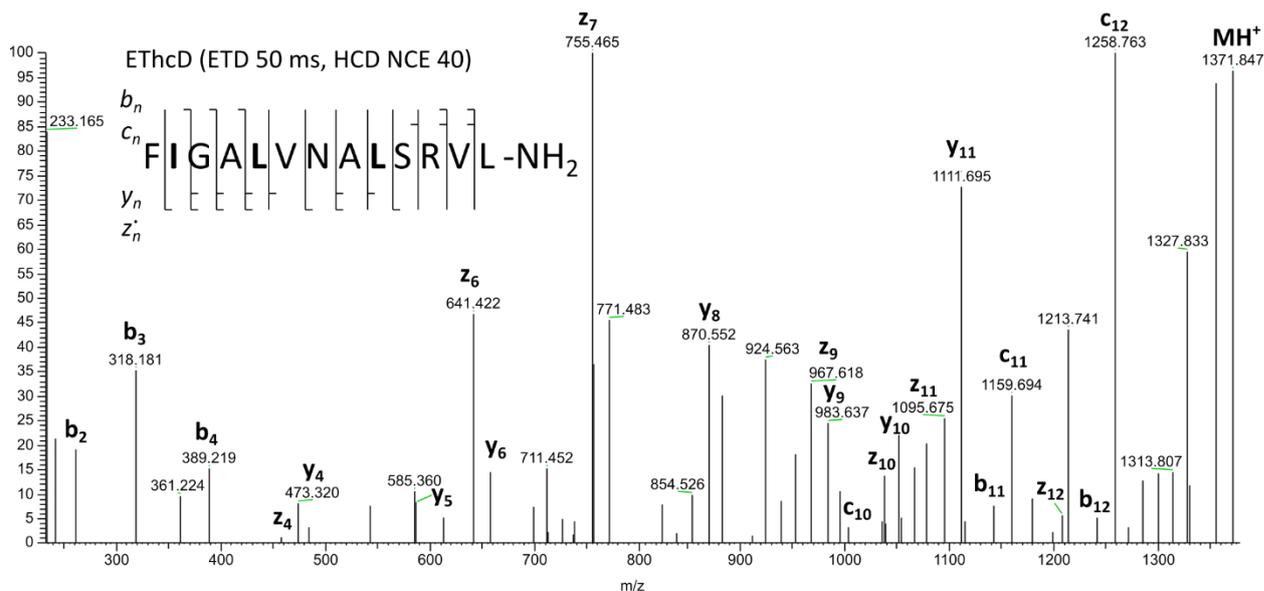


Рисунок 45. Спектр EThcD иона с  $m/z$  686.427<sup>2+</sup> темпорина W\*

EThcD формирует в спектрах две пары комплементарных фрагментных ионов ( $c/z$  и  $b/y$ ), образованных разрывами разных связей в полипептидной цепи: NH-C<sub>α</sub> и CO-NH, соответственно. Это дает возможность получать структурную информацию о полных последовательностях коротких пептидов, особенно, если в них присутствует С-концевой заряд, как в случае темпоринов W\* и X\*.

EThcD позволил также идентифицировать три из четырёх изомерных аминокислот в последовательности темпорина W\* (рис. 46). Ile<sup>2</sup> был определен по потере этильного радикала из целевого z<sub>12</sub><sup>+</sup> иона, тогда как Leu<sup>5</sup> и Leu<sup>9</sup> – по выбросам *iso*-Pr<sup>•</sup> из z<sub>9</sub><sup>+</sup> и z<sub>5</sub><sup>+</sup>.

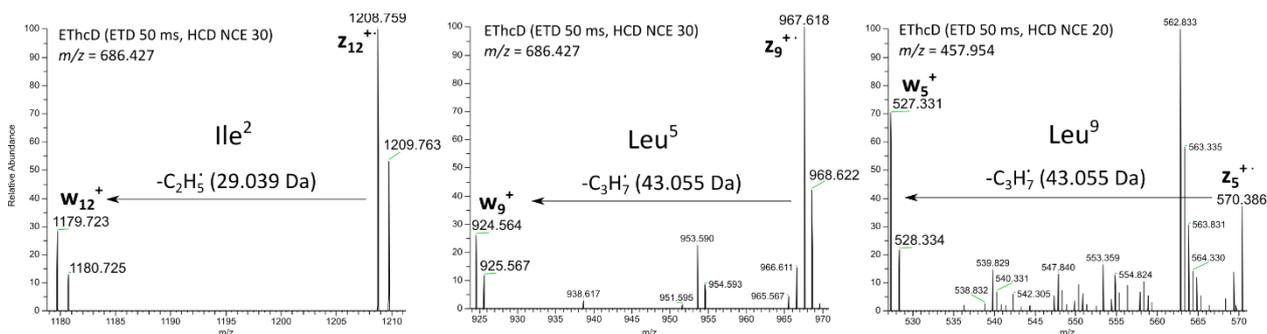


Рисунок 46. Идентификация Ile<sup>2</sup>, Leu<sup>5</sup> и Leu<sup>9</sup> в темпорине W\* по спектрам EThcD

Темпорин W\* является структурным аналогом темпорина T из секрета словенских особей *R. temporaria*, поскольку имеет в сравнении с ним единственную замену: Thr<sup>10</sup>→Ser<sup>10</sup>. По физико-химическим свойствам ее следует рассматривать как равную. BLAST не обнаружил совпадений с сиквенсом темпорина W\*.

### Темпорин X\*, MM 1313.818 Да

Темпорин X\*, ещё один структурный аналог темпорина T, выделенного из секретов словенских травяных лягушек, правда, уже с двумя заменами: Thr<sup>10</sup>→Ser<sup>10</sup> и Asn<sup>7</sup>→Gly<sup>7</sup>. В сравнении с предыдущим темпорином W\* темпорин X\* имеет также две замены: Ile<sup>2</sup>→Leu<sup>2</sup> и Asn<sup>7</sup>→Gly<sup>7</sup>.

На рисунке 47 показан спектр EThcD трёхзарядного молекулярного иона-радикала темпорина X\*: время экспозиции при переносе электрона составило 50 мс, нормализованная энергия NCE в ДАСПЭ равнялась 30.

Две пары комплементарных серий ионов *b/y* и *c/z* дали информацию о всей последовательности темпорина X\*.

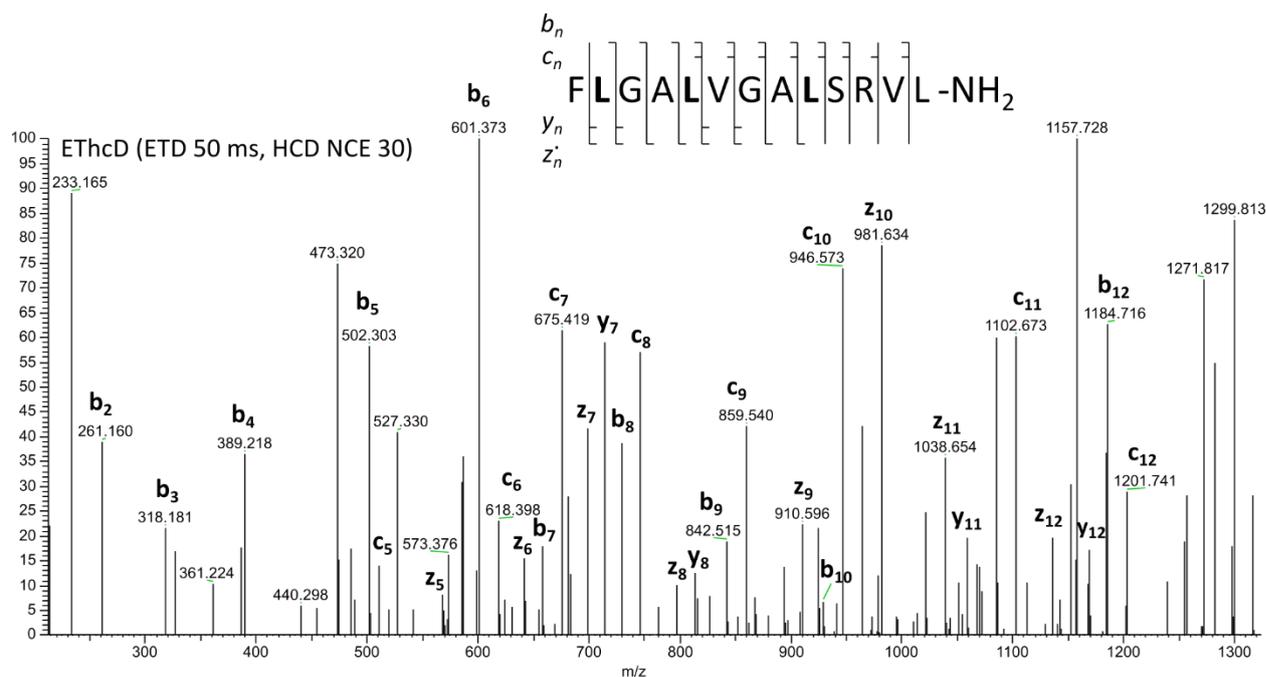
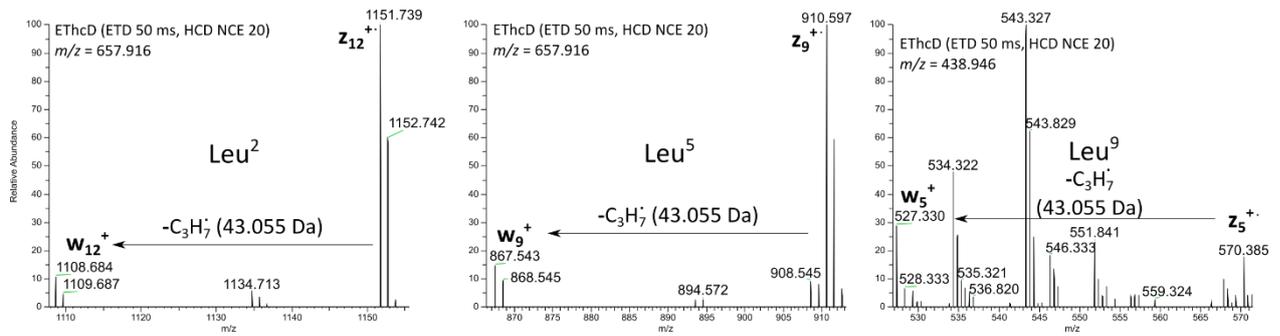


Рисунок 47. Спектр EThcD трижды протонированного темпорина X\* с *m/z* 438.946<sup>3+</sup>

Leu<sup>2</sup>, Leu<sup>5</sup> и Leu<sup>9</sup> в структуре темпорина X\* были определены по потерям *iso*-Pr<sup>•</sup> из целевых *z*<sub>12</sub><sup>+</sup>, *z*<sub>9</sub><sup>+</sup> и *z*<sub>5</sub><sup>+</sup> катион-радикалов с образованием соответствующих характеристических *w*<sub>12</sub><sup>+</sup>, *w*<sub>9</sub><sup>+</sup> и *w*<sub>5</sub><sup>+</sup> ионов (рис. 48). Поиски в BLAST(UniProt) совпадений со структурой темпорина X\* не дали.



**Рисунок 48.** Идентификация  $\text{Leu}^2$ ,  $\text{Leu}^5$  и  $\text{Leu}^9$  в темпорине X\* по спектрам EThcD

Таким образом, спектры EThcD явились опорными при проведении данной работы: (i) для секвенирования дисульфидсодержащих пептидов амфибий с интактной внутримолекулярной S-S связью, обеспечивая ее разрыв и фрагментацию внутри цикла; (ii) для секвенирования коротких темпоринов, часто фрагментирующих с образованием большого числа вторичных ионов, усложняющих интерпретацию спектров; и (iii) для дифференцирования изомерных аминокислот в парах Ile/Leu.

## 2.5. Популяционная идентификация особей *R. temporaria* на основе масс-спектрометрического изучения их кожных пептидомов

В таблице 10 суммированы пептиды, обнаруженные в четырех популяциях *Rana temporaria*: Архангельской (столбец A); Московской (столбец M [28]); Словенской (столбец S) и Итальянской (столбец I [23,255]).

**Таблица 10.** Сравнение составов кожных пептидомов четырех популяций *Rana temporaria*

№	Пептид	Последовательность	ММ, Да	A	M	S	I	UniProt KB, ссылка
1	Бревинин 1Т	VNPIILGVLPK <b>FVCLITKCC</b> -OH	2195.294	+	+	+	+	P82232, [23]
2	Бревинин 1Та	FITLLLRK <b>FICSITKCC</b> -OH	2024.168	+	+	+	+	P82233, [23]
3	Бревинин 1Тб	LVPLFLSKLIC <b>FITKCC</b> -OH	1963.140	+	+	+	-	[166]
4	Бревинин 1Тс	LVPMFLSKLIC <b>FITKCC</b> -OH	1981.096	-	-	+	-	C0HM35,
5	Бревинин 2Т	GLLSGLKKV <b>GKHVAKNVA</b> SLMDSLK <b>CKISGDC</b> -OH	3395.844	+	+	+	+	P82740, [23]
6	Бревинин 2Тd	GLWETIKNFGKK <b>F</b> TLN <b>LHNLKCKIGGGC</b> -OH	3229.732	+	-	-	-	P82235.1, [23]
7	Бревинин 2Тe	GLLDGLKKV <b>GKHVAKNVA</b> SLMDSLK <b>CKISGDC</b> -OH	3423.840	-	+	+	-	
8	Бревинин 2Тf	KK <b>F</b> TLN <b>L</b> FH <b>Q</b> LK <b>CKIGGGC</b> -OH	2132.138	-	+	-	-	
9	MRP-1	FIGSALKVLAGVLP <b>S</b> VIS <b>W</b> VK <b>Q</b> -NH <sub>2</sub>	2310.383	+	+	+	+	P56924, [23]
10	Темпорин А	FLPLIGRVL <b>S</b> GIL-NH <sub>2</sub>	1395.896	-	+	+	+	[23]
11	Темпорин В	LLPIVGNLLK <b>S</b> LL-NH <sub>2</sub>	1390.927	+	+	+	+	P79874, [23]

№	Пептид	Последовательность	ММ, Да	A	M	S	I	UniProt KB, ссылка
12	Темпорин С	LLPILGNLLNGLL-NH <sub>2</sub>	1360.880	+	+	+	+	P56918, [23]
13	Темпорин D	LLPIVGNLLNSLL-NH <sub>2</sub>	1376.880	+	+	+	+	P56919, [23]
14	Темпорин E	VLPIIGNLLNSLL-NH <sub>2</sub>	1376.880	-	+	+	+	[23]
15	Темпорин F	FLPLIGKVLVSGLL-NH <sub>2</sub>	1367.890	+	+	-	+	P56921, [23]
16	Темпорин G	FFPVIGRILNGIL-NH <sub>2</sub>	1456.891	+	+	+	+	P79875, [23]
17	Темпорин H	LSPNLLKSLL-NH <sub>2</sub>	1095.701	+	+	+	+	P79876, [23]
18	Темпорин K	LLPNLLKSLL-NH <sub>2</sub>	1121.752	+	-	-	+	P56923, [23]
19	Темпорин L	FVQWFSKFLGRIL-NH <sub>2</sub>	1638.939	+	+	+	+	P57104, [23]
20	Темпорин M	FLPILGKVLRSVL-NH <sub>2</sub>	1452.954	+	+	+	-	[166]
21	Темпорин N	FLGALGNALSRVL-NH <sub>2</sub>	1328.792	+	+	+	-	[166]
22	Темпорин O	FLGALVNALRGLL-NH <sub>2</sub>	1354.840	-	-	+	-	COHM29
23	Темпорин P	LVPLLGNLLSGLL-NH <sub>2</sub>	1319.842	-	-	+	-	COHM30
24	Темпорин Q	LVPFLGKTLGGLLARF-NH <sub>2</sub>	1700.046	-	-	+	-	COHM31
25	Темпорин R	LVPFLGRTLGLGLARL-NH <sub>2</sub>	1694.070	-	-	+	-	COHM32
26	Темпорин S	LVPFLGRTLGLGLARF-NH <sub>2</sub>	1728.050	-	-	+	-	COHM33
27	Темпорин T	FXGALVNAXTRVX-NH <sub>2</sub>	1384.851	+	-	+	-	COHM34
28	Темпорин U*	FLGAVLHALTGIL-NH <sub>2</sub>	1322.808	+	-	-	-	
29	Темпорин V*	FLGALVHALTSLL-NH <sub>2</sub>	1352.820	+	-	-	-	
30	Темпорин W*	FIGALVNALSRVL-NH <sub>2</sub>	1370.839	+	-	-	-	
31	Темпорин X*	FLGALVGALSRVL-NH <sub>2</sub>	1313.818	+	-	-	-	
32	Брадикинин (Br)	RPPGFSPFR-OH	1059.561	+	+	+	+	
33	[Val <sup>1</sup> ] Br VR-11	VVPPGFSPFR-OH	1101.596	-	+	+		
34	Br KR-11	KVRPPGFSPFR-OH	1286.725	-	+	-		
35	Br HR-11	HVRPPGFSPFR-OH	1295.688	-	+	-		
36	Br PR-11	PGRPPGFSPFR-OH	1213.635	-	+	-		
37	Br NR-11	NVRPPGFSPFR-OH	1272.672	-	+	-		
38	Br AR-11	AVRPPGFSPFR-OH	1229.666	-	+	-		
39	Br DR-11	DDRPPGFSPFR-OH	1289.615	-	+	+		
40	Br DR-11-1	DERPPGFSPFR-OH	1303.630	-	+	-		
41	Br LR-11	LDRPPGESPFROH	1269.646	-	+	-		
42	Br WR-10	WRPPGFSPFR-OH	1245.639	-	+	-		
43	Br HI-13	HVRPPGFSPFRIA-OH	1479.809	-	+	-		
44	Br HS-16	HVRPPGFSPFRIAPAS-OH	1734.931	-	+	-		
45	Br LR-11-1	LVRPPGFSPFR-OH	1271.713	-	+	-		
46	Br DR-11-1	DVRPPGFSPFR-OH	1273.656	-	+	+		
47	Br FR-11	FVRPPGFSPFR-OH	1305.698	-	+	-		
48	Br LR-11-2	LLRPPGFSPFR-OH	1285.729	-	+	+		
49	Br LR-10	LRPPGFSPFR-OH	1172.645	-	+	+		
50	Br FR-11	FLRPPGFSPFR-OH	1319.713	-	+	-		
51	Br GR-12	GLLRPPGFSPFR-OH	1342.750	-	+	-		
52	Br FR-12	FLPRPPGFSPFR-OH	1416.766	-	+	-		
53	Br GR-16	GLLSGLKRPPGFSPFR-OH	1727.983	-	+	-		
54	Br RA-13	RPPGFSPFRIA-VA-OH	1413.787	-	+	-		
55	Br RS-14	RPPGFSPFRIAPAS-OH	1498.804	+	+	+		
56	Br RT-15	RPPGFSPFRIAPAST-OH	1599.851	+	+	+		
57	Br RL-16	RPPGFSPFRIAPASTL-OH	1712.936	+	+	+		
58	Br RA-11	RPPGFSPFRIA-OH	1243.682	+	+	+		
59	Br RI-10	RPPGFSPFRI-OH	1172.645	+	+	+		

№	Пептид	Последовательность	ММ, Да	A	M	S	I	UniProt KB, ссылка
60	Br RP-12	RPPGFSPFRIAP-OH	1340.735	+	+	+		
61	Br RD-13	RPPGFSPFRIAVD-OH	1457.777	-	+	+		
62	[Thr <sup>6</sup> ] Br	RPPGFTPFR-OH	1073.576	+	+	+		
63	[desArg <sup>9</sup> , Thr <sup>6</sup> ] Br	RPPGFTPFR-OH	917.475	-	+	+		
64	[Thr <sup>6</sup> ] Br RA-11	RPPGFTPFRIFA-OH	1257.698	+	+	+		
65	[Thr <sup>6</sup> ] Br RA-13	RPPGFTPFRIVA-OH	1427.803	-	+	-		
67	[Thr <sup>6</sup> ] Br RI-10	RPPGFTPFRIO-OH	1186.660	+	+	+		
66	[Asp <sup>6</sup> ] Br RA-11	RPPGFDPFRIFA-OH	1271.677	-	+	+		
68	[Thr <sup>6</sup> ,Leu <sup>8</sup> ] Br	RPPGFTPLR-OH	1039.592	-	+	-		

По Дж. Теннесену [265–268], изменчивость антимикробных пептидов, наблюдаемая у различных популяций внутри одного вида амфибий, происходит в рамках естественного отбора и диктуется их бактериальным окружением и обитающими там хищниками. На примере четырех популяций травяной лягушки мы можем проверить, как меняется их кожный пептидом в зависимости от места обитания.

Отметим весьма приблизительные географические координаты сравниваемых популяций:

- |                       |                        |                      |
|-----------------------|------------------------|----------------------|
| 1. Архангельск        | 64°33' Северной широты | 40°32' Вост. долготы |
| 2. Звенигород         | 55°44' Северной широты | 36°51' Вост. долготы |
| 3. Люблина            | 46°05' Северной широты | 14°51' Вост. долготы |
| 4. Рим, коммуна Кьети | 42°21' Северной широты | 14°10' Вост. долготы |

Географические координаты итальянской популяции взяты по месту жительства автора М. Симмако [23].

Самая южная популяция – итальянская, самая северная – Архангельская, всего два градуса до Полярного круга (Полярный круг 66°33' Северной широты).

Пептидом итальянской лягушки представлен по работам М. Симмако и не включает структурных аналогов брадикинина с их протеолитическими копиями, расширенными с C- и/или N-концов. Жёлтым цветом в таблице 10 выделены пептиды, встречающиеся в секретах всех популяций: это бревинин 1Т, бревинин 1Та; бревинин 2Т; мелиттин-родственный пептид FQ-22; темпорины В, С, D, G, H, L и брадикинин. Все эти пептиды, за исключением брадикинина, были открыты и охарактеризованы Маурицио Симмако. Их можно предположительно назвать биомаркерами вида травяной лягушки *Rana temporaria*.

Наиболее изменчивым семейством пептидов являются темпорины. М. Симмако в 1998 году секвенировал в секрете европейской лягушке *Rana temporaria* 10 темпоринов,

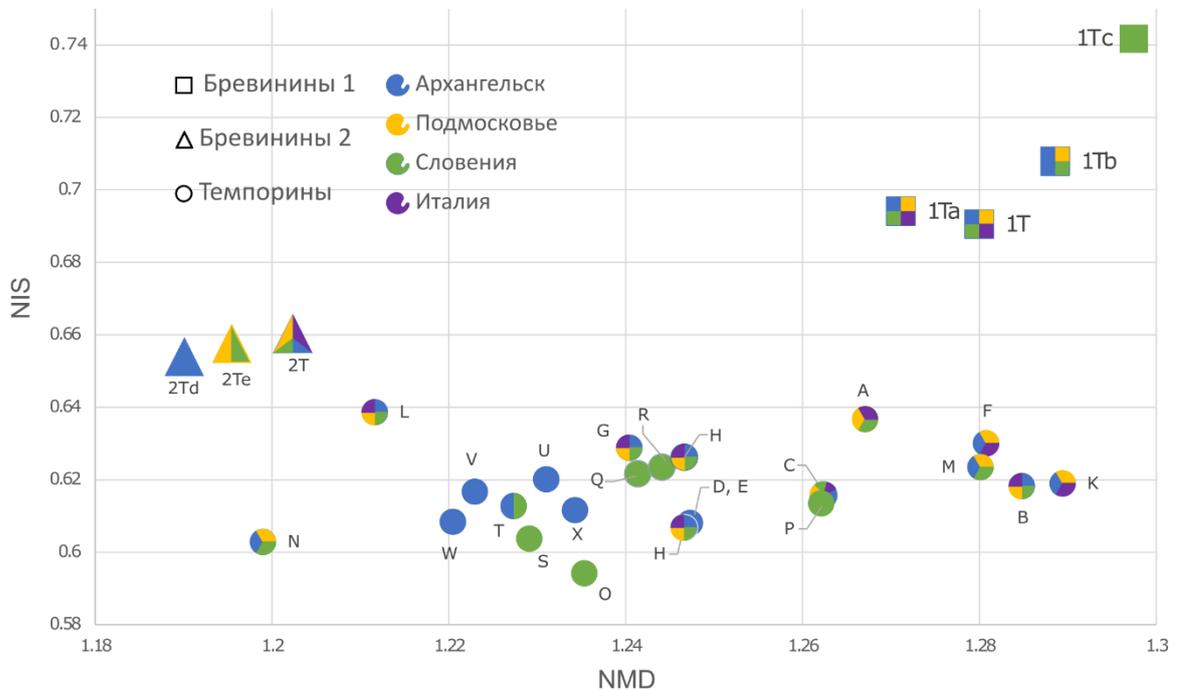
получивших название от А до L. К настоящему времени в таблице 10 насчитывается 22 пептида этого семейства, обнаруженных нашей группой ещё в трех популяциях этого вида.

Темпорины А, Е, F, К, охарактеризованные Симмако, а также темпорины N и M, открытые в московской популяции нами, встречаются в секретах трех популяций из четырех, представленных в таблице 10. У особей из архангельской популяции в секрете присутствуют 15 темпоринов, 4 из которых уникальны: это темпорины U, V, W, X (биомаркеры популяции). В подмосковной популяции – 12 темпоринов, ни одного уникального. У словенских особей 15 темпоринов, 5 уникальных (кандидаты на биомаркеры популяции), у итальянских – 10 темпоринов, ни одного уникального.

Для особей из Подмосковной популяции характерно наличие в секрете протеолитических копий брадикинина с C-концевым и особенно с N-концевыми расширениями. Напротив, у архангельских и словенских лягушек в секретах присутствуют протеолитические копии брадикинина исключительно с C-расширением. Причины этого следует искать в механизме действия протеаз, секретлируемых совместно с секретом.

К уникальными для особей из подмосковной популяции, пожалуй, можно отнести фрагмент бревинина 2Tf (№8 таблицы 10) и [Thr<sup>6</sup>, Leu<sup>8</sup>]Br (№68), являющийся орнитокинином, агонистом орнитокининового рецептора, который был впервые охарактеризован Шрёдером с соавторами [154].

На рисунке 49 показана 2D-карта, объединяющая в себе пептиды всех четырёх популяций *Rana temporaria*.



**Рисунок 49.** 2D-карта четырёх популяций *Rana temporaria*

Поскольку темпорины у всех четырёх популяций травяной лягушки являются наиболее вариативным пептидным семейством, именно за счет них происходит адаптация вида к различным климатическим условиям и меняющемуся патогенному окружению. Наши результаты не противоречат гипотезе Дж. Теннессена, о том, что лягушка сама решает, какой из генетически закодированных в ее хромосомах антимикробных пептидов она будет секретировать в новой экологической нише, поскольку, по мнению Дж. Теннессена, каждый пептид амфибия секретирует не случайно [267].

В таблице 11 суммированы все последовательности известных на сегодняшний день темпоринов, выделенных из кожных секретов *Rana temporaria*.

**Таблица 11.** Таблица структурной аналогии пептидов из семейства темпоринов, выделенных из кожного секрета особей вида *Rana temporaria*.

№	Пептид	Аминокислотная последовательность															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Темпорин А	F	L	P	L	I	G	R	V	L	S	G	I	L			
2	Темпорин В	L	L	P	I	V	G	N	L	L	K	S	L	L			
3	Темпорин С	L	L	P	I	L	G	N	L	L	N	G	L	L			
4	Темпорин D	L	L	P	I	V	G	N	L	L	N	S	L	L			
5	Темпорин E	V	L	P	I	I	G	N	L	L	N	S	L	L			
6	Темпорин F	F	L	P	L	I	G	K	V	L	S	G	L	L			
7	Темпорин G	F	F	P	V	I	G	R	I	L	N	G	I	L			
8	Темпорин H	L	S	P	-	-	-	N	L	L	K	S	L	L			
9	Темпорин K	L	L	P	-	-	-	N	L	L	K	S	L	L			
10	Темпорин L	F	V	Q	W	F	S	K	F	L	G	R	I	L			
11	Темпорин M	F	X	P	I	X	G	K	V	X	S	R	V	X			
12	Темпорин M'	F	X	Нур	I	X	G	K	V	X	S	R	V	X			
13	Темпорин N	F	X	G	A	X	G	N	A	X	S	R	V	X			
14	Темпорин O*	F	X	G	A	L	V	N	A	X	R	G	X	X			
15	Темпорин P*	X	V	P	X	X	G	N	X	X	S	G	X	X			
16	Темпорин Q*	L	V	P	F	X	G	K	T	X	G	G	L	L	A	R	F
17	Темпорин R*	X	V	P	F	X	G	R	T	L	G	G	L	L	A	R	X
18	Темпорин S*	L	V	P	F	X	G	R	T	X	G	G	L	L	A	R	F
19	Темпорин T*	F	X	G	A	X	V	N	A	X	T	R	V	X			
20	Темпорин U*	F	I	G	A	V	X	H	A	L	T	G	X	X			
21	Темпорин V*	F	I	G	A	X	V	H	A	X	T	S	X	X			
22	Темпорин W*	F	I	G	A	L	V	N	A	L	S	R	V	X			
23	Темпорин X*	F	L	G	A	L	V	G	A	L	S	R	V	X			
<b>Число замен от темпорина А</b>		<b>2</b>	<b>3</b>	<b>3</b>	<b>5</b>	<b>3</b>	<b>2</b>	<b>4</b>	<b>5</b>	<b>0</b>	<b>5</b>	<b>2</b>	<b>2</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>2</b>

\* - последовательности установлены в данной работе

Последовательности темпоринов изменчивы, в том числе, и по числу звеньев в цепи. Однако внутри одного вида лягушек, в данном случае *Rana temporaria*, темпорины сохраняют некоторую консервативность в первичных структурах. В таблице 11 звёздочкой (\*) отмечены новые темпорины, чьи последовательности установлены в настоящей работе. В нижней строке таблицы указано число возможных замен в последовательностях темпоринов из *Rana temporaria* относительно темпорина А, первого обнаруженного у *Rana temporaria* пептида этого семейства. Неидентифицированные изомерные Leu/Ile обозначены буквой X.

Темпорины, принадлежащие к виду лягушки *Rana temporaria*, содержат два подсемейства: короткие (10-13 ак) и длинные (16 ак) темпорины; все они амидные. У двух 10-ти звенных темпоринов Н и К имеется пропуск трех аминокислот в положении 4-6 последовательностей. Из таблицы 11 следует, что все темпорины из кожных секретов *Rana temporaria* имеют два консервативных положения в последовательностях (9 и 13), в которых присутствуют лейцины: Leu<sup>9</sup> и Leu<sup>13</sup>. С большой долей вероятности это можно распространить на темпорины М – Q и Т – X. В четырёх положениях последовательностей коротких темпоринов (10-13 ак) возможны только две замены по сравнению с исходным темпорином А: в 1-ом, 6-ом, 11-ом, 12-ом. Эта таблица может оказаться полезной при секвенировании новых темпоринов из различных популяций вида *Rana temporaria*.

## 2.6. «Rana box» – раскрытие дисульфидного цикла в интактных пептидах

Выше было показано, насколько информативным для секвенирования коротких темпоринов является EThcD эксперимент, доступный как метод MS<sup>3</sup> в орбитальной ловушке. Он объединяет в себе протонирование молекул пептидов при электрораспылении с образованием  $MH_n^{n+}$  (ИЭР); дальнейший перенос электрона на полипротонированную молекулу с образованием активного катион-радикала  $MH^{(n-1)+}$  (ДПЭ); и дополнительную активацию его фрагментации диссоциацией при повышенных энергиях (ДАСПЭ), причём энергию мы ступенчато варьировали в интервале нормализованных значений от 5 до 40. Необходимо отметить, что в экспериментах не протекала циклизация коротких пептидов, что позволило избежать дополнительного модифицирования N-концевых аминокислот пептидов. Любое модифицирование, даже самое щадящее, приводит к искажению результата из-за своей неселективности, но, главное, из-за возможной потери в процессе модифицирования и последующей очистки части компонентов, присутствующих в интактном образце в минимальных количествах. Полнота определения кожного пептидома, таким образом, зависит от многих факторов: от процедуры получения кожных секретов; от способа очистки их от деградирующих протеаз и правильного хранения образца; от технических возможностей использованного масс-спектрометра; от режима хроматографирования образцов в MS эксперименте, от вариантов сбора масс-спектрометрических данных и т. д.

Основной сложностью в *de novo* секвенировании кожных секретов ранидных лягушек является установление сиквенса внутри C-концевого дисульфидного цикла. Мы впервые применили подход «сверху вниз» с целью оценить возможности метода EThcD для раскрытия S-S связей в интактных пептидах в дополнение к уже известной фрагментации в условиях низкоэнергетической (ДАС) и высокоэнергетической активации столкновениями (ДАСПЭ) [53,169].

Таблица 12 показывает вклад различных методов фрагментации орбитальной ловушки в секвенирование «Rana box» при ручной интерпретации спектров. В ней выборочно приведены данные по секвенированию дисульфидных пептидов двух видов: *R. temporaria* из центральнословенской популяции и *R. arvalis* из Архангельской области. Спектры при переносе электрона (ДПЭ) отдельно для *R. arvalis* не регистрировались.

Из таблицы 12 видно, что раскрытие дисульфидного цикла в EThcD происходит стабильно и чаще, чем в ДАСПЭ. В наших экспериментах фрагментация внутри интактного «Rana box» в условиях низкоэнергетической активации столкновениями ДАС практически не протекала. Для пептидов из семейства бревининов 1, в силу особенностей строения их «Rana box» (два С-терминальных Lys), оба раскрытия, как правило, происходили: и в ДАСПЭ, и в EThcD. Характерно, что информация, полученная двумя этими методами о сиквенсах внутри интактных «Rana box», была взаимодополняющей, дающей в сумме полную последовательность (бревинин 1Т, бревинин 1Та, бревинин 1Тб). Даже если раскрытия цикла в ДАСПЭ у интактных бревининов 1 не достигались, результаты разрыва S-S связи одного EThcD позволяли установить последовательность внутри «Rana box» полностью (бревинины 1AVa и 1AVb).

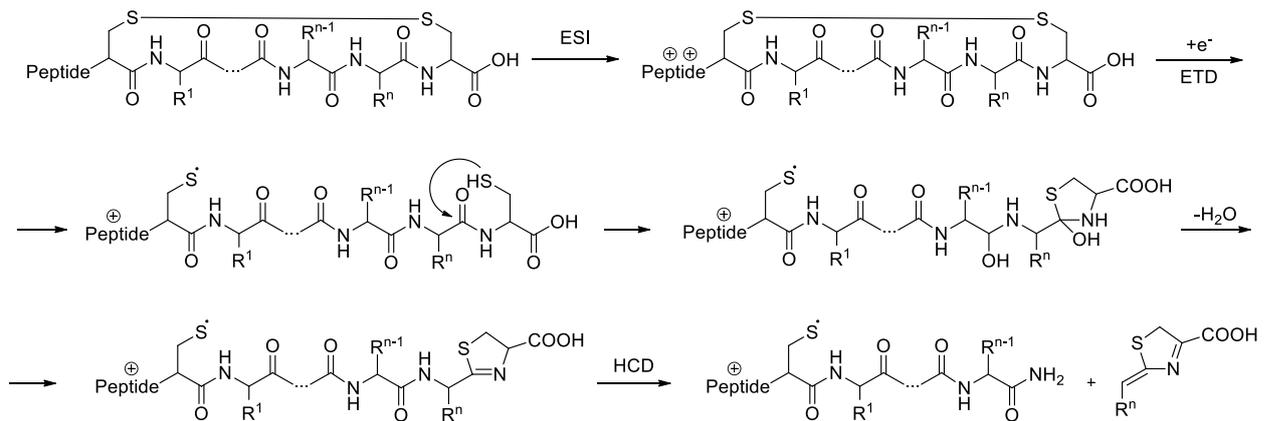
Раскрытие «Rana box» в длинных, 33-звенных бревининах 2Т (*R. temporaria*) и 2AV (*R. arvalis*) происходит, равно как и в их более коротких С-концевых протеолитических фрагментах. Как видно из таблицы 12, в коротких протеолитических фрагментах бревинина 2Т и бревинина 2AV (с зарядом 2<sup>+</sup>) суммой методов ДАСПЭ и EThcD удаётся прочесть последовательность в их интактных «Rana box». Аналогично, в 29-членном бревинине 2Тd в ходе EThcD эксперимента удалось прочитать всю последовательность аминокислот внутри «Rana box», что обсуждалось выше.

И, наконец, дисульфидные ранатуерины, с шестичленным «Rana box», в которых в условиях ДАСПЭ активной фрагментации внутри не происходит. Это объясняется тем, что не удалось обнаружить их более коротких протеолитических фрагментов, с зарядом не выше 2<sup>+</sup>, чтобы добиться активной фрагментации внутри интактного С-концевого цикла. Анализируемые протеолитические фрагменты имели заряд не ниже 4<sup>+</sup> (ранатуерин 2AVb) и не ниже 3<sup>+</sup> (ранатуерина 2AVa). В случае же EThcD раскрытие цикла происходит стабильно, оставляя неразрешённой С-концевую пару аминокислот. Однако её также можно однозначно установить по отсутствию  $b_{n-1}$ -иона, который присутствовал бы в спектре в случае сайта СХ вместо ХС на С-конце пептида.

Таким образом, спектры EThcD оказались по результатам наших исследований крайне информативными. Фрагментация в «Rana box» в условиях EThcD показала себя полностью комплементарной тому раскрытию дисульфидного цикла, который наблюдается в условиях

ДАС/ДАСПЭ. Это позволяет установить сиквенсы в интактных «Rana box» при ручной интерпретации масс-спектров и делает возможным изучение секретов амфибий методом «сверху вниз».

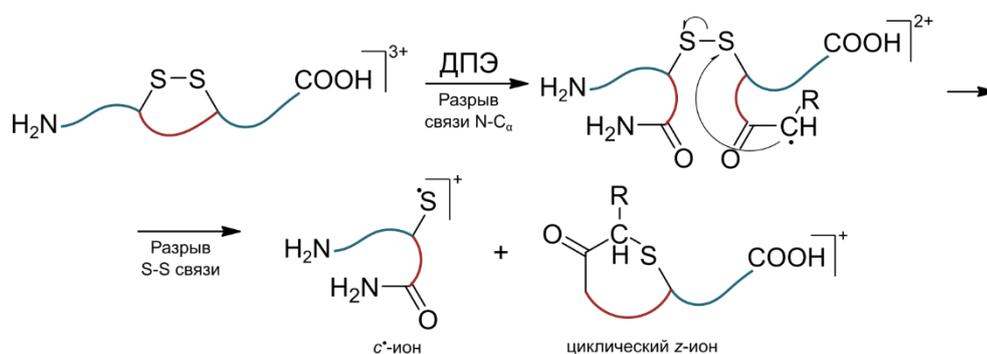
Для объяснения образования *s*-катион-радикалов изначально была предложена схема, изображенная на схеме 2: в ходе ДЗЭ-активации электрон захватывается дисульфидной связью, что вызывает её гомолитический разрыв. Поскольку частица протонноизбыточна, образуется тиольный фрагмент, нуклеофильно атакующий карбоксамидный атом углерода с образованием тиазолидинового фрагмента, который затем теряет молекулу воды. Дополнительная ДАСПЭ активация вызывает разрыв связи N-C $\alpha$ .



**Схема 2.** Первая предложенная схема образования *s*-катион-радикалов в ходе EThcD эксперимента

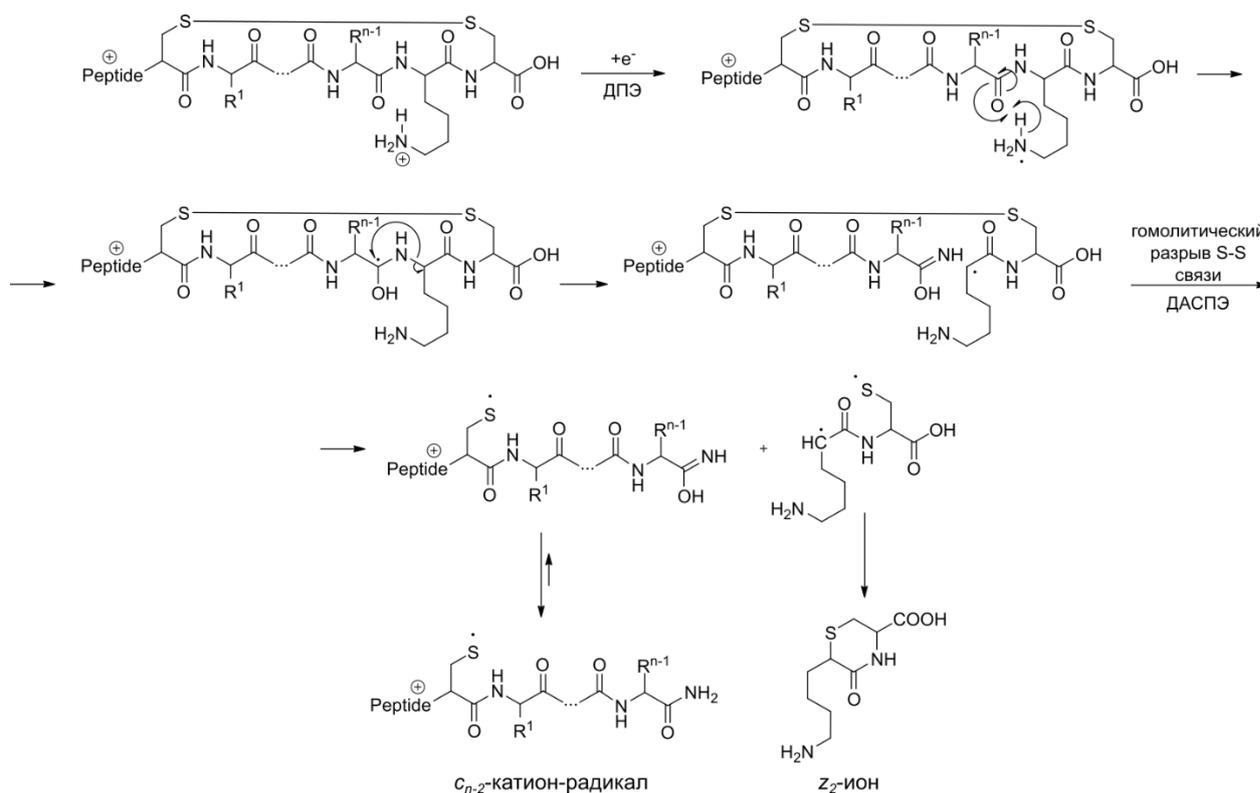
Однако данная схема имеет много недочётов: во-первых, нуклеофильная атака протекает дольше, чем реакции, инициируемые радикальным центром, из-за чего в спектре скорее наблюдались бы другие фрагментные ионы. Кроме того, схема не объясняет образование полной серии ионов с разрывами связей внутри исходного дисульфидного цикла.

Образование циклических ионов и *s*-катион-радикалов было предложено в работе С.Р. Коула с соавторами [269] в рамках следующей схемы фрагментации (схема 3) для 13 коммерчески доступных пептидов с различными размерами дисульфидного цикла (от 5 до 12 аминокислотных звеньев). Предложенная схема подразумевает образование циклического фрагмента в структуре  $z$ -иона. В случае продолжительных последовательностей внутри дисульфидного цикла и разрыве связи N-C $\alpha$  в аминокислоте, находящейся ближе к *N*-концу пептида, можно ожидать, что размер цикла будет слишком большим, что затруднит его образование: например, если  $z$ -ион состоит из трёх остатков аминокислот, включая цистеин, цикл становится уже 9-ти членным. Кроме того, помимо  $z$ -ионов и *s*-катион радикалов авторы наблюдали ионы  $s-33$ ,  $s+32$ ,  $z+33$  и  $z-32$ , которых нам не удалось обнаружить в спектрах даже до деконволюции.



**Схема 3.** Механизм разрыва дисульфидной связи под действием ДПЭ, предложенный в работе [269].

Наиболее вероятная на наш взгляд схема реакции (схема 4) заключается в следующем: в ходе EThcD фрагментации в многозарядных ионах дисульфидсодержащих пептидов электрон захватывается аминокислотным остатком основной аминокислоты внутри цикла с последующим переносом радикального центра на карбоксильный атом углерода, как это предполагает механизм ДЗЭ, предложенный Зубаревым и МакЛафферти [94].



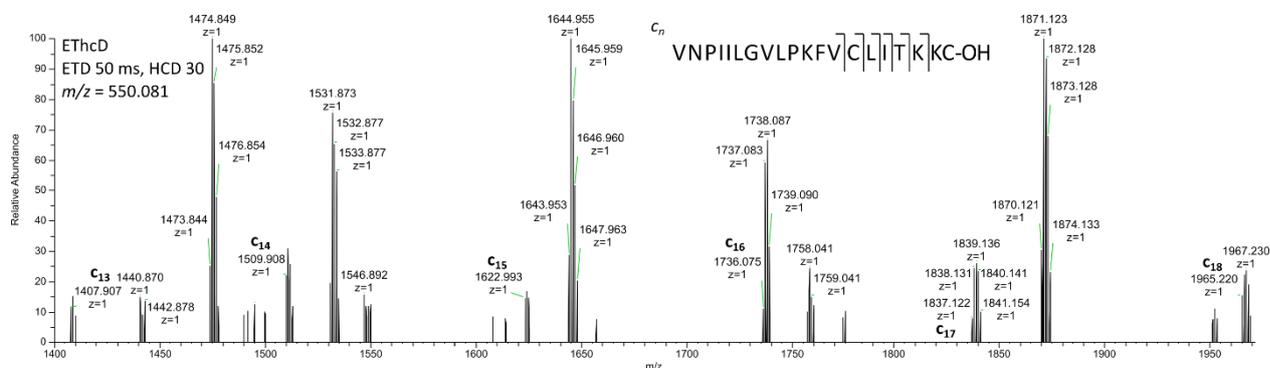
**Схема 4.** Предлагаемый механизм фрагментации дисульфидсодержащих пептидов внутри цикла в условиях EThcD

Возникший радикальный центр инициирует разрыв связи N-C<sub>α</sub>, а последующая ДАСПЭ активация вызывает гомолитический разрыв дисульфидной связи. В результате образуется серия s-ион-радикалов и короткая быстро затухающая серия чётноэлектронных z-ионов, содержащих циклический тиоморфолиноновый фрагмент. Такая схема объясняет появление

нечётноэлектронных  $c$ -ионов и непродолжительность серии  $z$ -катионов – возможно, циклизация иона  $z_3$  будет протекать после 1-3 сигматропного сдвига атома водорода на остаток соседней аминокислоты, но многократная подобная миграция будет менее эффективна. Такая схема объясняет и то, что разность масс между  $c$ -ионом, ещё содержащим остаток внутреннего цистеина и уже не содержащим его, составляет 102 единицы масс, а не 103.

Образование серий  $c$ -катион-радикалов, соответствующих последовательностям «Rana box», происходит не благодаря последовательным выбросам остатков аминокислот из ионов предшествующих поколений, а за счёт миграции радикального центра в раскрывшемся цикле, что в соответствии с существующей в протеомике концепцией мобильного протона [270] можно назвать концепцией мобильного радикального центра.

В EThcD спектрах бревининов в соответствии с предложенным механизмом наблюдается серия  $c$ -катион-радикалов, позволяющих установить всю последовательность внутри «Rana box». На рисунке 50 приведён пример фрагмента такого спектра для бревинина 1Т, чья последовательность была установлена деградацией по Эдману [23] и подтверждена нами в масс-спектрометрических экспериментах. Интересно отметить, что потеря внутреннего цистеина для проанализированных таким образом дисульфидсодержащих пептидов происходит как уменьшение массы на 102.001 Да, то есть радикальный центр остаётся на этом аминокислотном остатке. Дальнейшая серия образующихся классических  $c$ -ионов уже не несёт на себе радикального центра.



**Рисунок 50.** Фрагмент EThcD спектра бревинина 1Т до деконволюции с характеристическими  $c$ -ионами «Rana box»

Оценить вклад информации спектров EThcD в прочтение последовательностей внутри  $C$ -терминальных дисульфидных циклов у интактных пептидов позволяют данные таблицы 13. Правый столбец колонки показывает последовательности интактных дисульфидсодержащих пептидов, полученные с использованием алгоритма автоматического секвенирования PEAKS. Серым цветом выделены совпадающие фрагменты последовательностей при ручном и автоматизированном секвенировании. В таблице 13 суммированы далеко не все полученные данные, что не меняет общей тенденции, наблюдаемой при выполнении настоящей работы.

Из 10-ти интактных дисульфидных пептидов (17-33 аминокислот в последовательностях) программа автоматического секвенирования не определила полностью ни одну последовательность внутри интактных «Rana box» (таблица 13). И это не потому, что в алгоритм программы не заложена фрагментация S-S связи в условиях EThcD. Программа PEAKS смогла установить частичные последовательности внутри C-терминальных интактных циклов трех бревининов 1 из пяти таблицы 13: бревинина 1AVb (5 а.к.), бревинина 1T (4 а.к.) и бревинина 1Ta (5 а.к.). Бревинины 1 не самые сложные для секвенирования объекты - они не содержат более 24-х а.к. и имеют два C-терминальных Lys. У них, как правило, протекают ДАС/ДАСПЭ и EThcD фрагментации в дисульфидных циклах. Два длинных минорных пептида, ранатуерин 2AVb (28 а.к.) и бревинин 2AV (33 а.к.), программа не обнаружила, а 29-ти звенный бревинин 2Td вообще не секвенировала. Таким образом, программа автосеквенирования PEAKS по данным спектров EThcD из 10-ти дисульфидных интактных пептидов частично установили сиквенсы трех «Rana box», не просеквенировала три пептида полностью, а в остальных определила фрагменты линейных частей последовательностей (см. таблицу 13).

Если посмотреть, что добавляет к результатам автоматического секвенирования последовательностей внутри «Rana box» с помощью алгоритма PEAKS химическое модифицирование S-S связи стандартным восстановлением/алкилированием (DTT/IAA), следует обратиться к данным таблицы 14. Сразу отметим принципиальный недостаток модифицирования пептидной смеси с последующей очисткой образца: программа автосеквенатора не находит пять из десяти исходно присутствовавших минорных компонентов пептидома. Однако программа прочитывает полностью последовательности трех бревининов 1: бревинина 1T, бревинина 1Tb и бревинина 1AVa. Без модифицирования дисульфидной связи PEAKS секвенировала эти же пептиды частично. У бревинина 1AVb произошло чрезмерное алкилирование в N-концевом Phe, программа установила его последовательность полностью (см. таблицу 14). Все это бревинины 1, короткие пептиды (до 24 а.к.). В обоих 28-ми звенных ранатуеринах, ранатуерине 2AVb и ранатуерине 2AVc, модификация не прошла – программа находит лишь интактные пептиды. Это ещё один недостаток модифицирования S-S связей: условия модифицирования не универсальны для всех пептидных компонентов, которые имеют разные физико-химические свойства и присутствуют в исходном образце кожного секрета в широком диапазоне концентраций.

**Таблица 12.** Вклад различных методов фрагментации орбитальной ловушки в секвенирование «Rana box» (ручная интерпретация спектров)

Вид лягушки	Пептид	Последовательность	Ручное <i>de novo</i> секвенирование «Rana box»		
			ДАСПЭ	ДПЭ	EThcD Σ (HCD 20, 30, 40)
<i>R. temporaria</i> , Архангельская область	Бревинин 1Т	VNPIILGVLPKFVCLITKKC-OH	(C <sup>c</sup> LL)TKK-OH	-	CLLTK(KC)-OH
	Бревинин 1Та	FITLLLRKFICSITKKC-OH	(C <sup>c</sup> SL)TKKC-OH	-	CSLTK(KC)-OH
	Бревинин 1Тб	LVPLFLSKLICFITKKC-OH	(C <sup>c</sup> FL)TKK-OH	-	CFLT(KC)-OH
	Бревинин 2Т	GLLSGLKKVGGKHVAKNVAVSLMDSLKCKISGDC-OH	-	-	(CKI)SG(DC)-OH
	Протеолитический фрагмент бревинина 2Т	MDSLKCKISGDC-OH	(C <sup>c</sup> KIS)GD-OH	-	CKISG(DC)-OH
	Бревинин 2Тd	GLWETIKNFGKKFTLNILHNLKCKIGGGC-OH	(C <sup>c</sup> KI)GGG-OH	-	CKIGG(GC)-OH
<i>R. arvalis</i> , Центральная Словения	Бревинин 1AVa	FLPLLAASFACTVTKKC-OH	(C <sup>c</sup> TV)TKK	н/р	CTVTK(KC)-OH
	Бревинин 1AVb	FVPLLVSKLVCVVTKKC-OH	(C <sup>c</sup> TV)VKK -	н/р	CTVVK(KC)-OH
	Бревинин 2AV	GLMSVLKKAGKHVAKNVAVSLMDSLKCKISGGC-OH	-	н/р	-
	Протеолитический фрагмент бревинина 2AV	VAVSLM <sub>ox</sub> DSLKCKISGGC-OH	(C <sup>c</sup> KL)SGG-OH	н/р	CKL(SG)(GC)-OH
	Ранатуерин 2AVb	GLMDMVKGAAKNLFFASALDTLKCKITGC-OH	(C <sup>c</sup> KI)(TG)-OH	н/р	CKIT(GC)-OH
	Протеолитический фрагмент ранатуерина 2AVb	VKGAAKNLFFASALDTLKCKITGC-OH	-	н/р	CKIT(GC)-OH
	Ранатуерин 2AVc	GLLDVVKGAAKNLLATALDKLKCKVTGC-OH	-	н/р	CKV(TGC)-OH
	Протеолитический фрагмент ранатуерина 2AVc	NLLATALDKLKCKVTGC-OH	-	н/р	CKVT(GC)-OH

Для *Rana arvalis* спектры ДПЭ отдельно не регистрировали

**Таблица 13.** Вклад спектров EThcD в автоматическое секвенирование интактных дисульфидных пептидов

Вид лягушки	Пептид	Ручное <i>de novo</i> секвенирование, интактный $\Sigma(\text{CID} + \text{HCD} + \text{EThcD})$	<i>De novo</i> секвенирование PEAKS, интактный EThcD
<i>R. arvalis</i> , Словения	Бревинин 1AVa	FLPLLAASFACTVTKKC-OH	FLPLLAASGTPAAPNGTDV(-.98)
	Бревинин 1AVb	FVPLLVS <del>KLVCVVTKKC</del> -OH	FVPLLVS <del>KLVCVVTKQT</del>
	Бревинин 2 AV	GLMSVLKKAGKHVAKNVAVSLMDSLKCKISGGC-OH	Не обнаружен
	Ранатуерин 2AVb	GLMDMVKGAAKNLFASALDTLKCKITGC-OH	Не обнаружен
	Ранатуерин 2AVc	GLLDVVKGAAKNLLATALDKLKCKVTGC-OH	GLLDVVKGAAKNLLATALDKFDRDLKG(-.98)
<i>R. temporaria</i> , Арханг. обл.	Бревинин 1T	VNPIILGVLPKFVCLITKKC-OH	VPNLLLVPLK <del>FVCLLTQMV</del> -NH <sub>2</sub>
	Бревинин 1Ta	FITLLLRKFICSITKKC-OH	FLTLLLRKFLC <del>SLVEKC</del> -NH <sub>2</sub>
	Бревинин 1Tb	LVPLFLSK <del>LICFITKKC</del> -OH	LPVLFLSKGCCLVLTFLV-NH <sub>2</sub>
	Бревинин 2T	GLLSGLK <del>KVGVKHVAKNVAVSLMDSLKCKISGDC</del> -OH	GLLAWK <del>KVGVKHVAQND</del> EDGDHK(+CO)LK <del>KLGEQK</del> -NH <sub>2</sub>
	Бревинин 2Td	GLWETIKNFGK <del>KFTLNILHNLKCKIGGGC</del> -OH	TVRERWNFGK <del>KFVFQKGFVLCQ</del> QK(+CO)T-OH

-серым выделены совпадающие фрагменты последовательностей при ручном и автоматизированном *de novo* секвенировании

**Таблица 14.** Секвенирование внутри «Rana box» ручной интерпретацией суммы спектров и автосеквенатором (интактные и модифицированные секреты)

Вид	Пептид	Секвенирование «Rana box»		
		Ручное интактный пептид Σ (ДАС + ДАСПЭ + EThcD)	Автоматизированное	
			Интактный пептид (ДАС + ДАСПЭ + ДПЭ)	Модифицированный (DTT/IAA) (ДАС + ДАСПЭ + ДПЭ + EThcD)
<i>R. temporaria</i> , Арханг. обл	Бревинин 1Т	VNPIILGVLPKFVCLITKKC-OH	VDPLLLGVNPWWVGLLLHK(+CO)-NH <sub>2</sub>	VNPLLLGVLPKFVC(+57.02)LLTKKC(+57.02)
	Бревинин 1Та	FITLLLRKFICSITKKC-OH	FLTLLK(+CO)KFGTNFLTEL-NH <sub>2</sub>	Не обнаружен (минорный компонент)
	Бревинин 1Тб	LVPLFLSKLICFITKKC-OH	LVPLFLSKWTPPYLPPP-NH <sub>2</sub>	LVPLFLSKLLC(+57.02)FLTKKC(+57.02)
	Бревинин 2Т	GLLSGLKKVVGKHAVAKNVAVSLMDSLK CKISGDC-OH	LGLSGLKKVVGKHAVAKNVAVSLMDSLKM <sub>ox</sub> E EEGL	Не обнаружен модифицированный
	Бревинин 2Тd	GLWETIKNFGKKFTLNILHNLK CKIGGGC-OH	LGWETLKNFGKKFTLNLLHNL TSAEWLG	Не обнаружен (минорный компонент)
<i>R. arvalis</i> , Словения	Бревинин 1AVa	FLPLLAASFACTVTKKC-OH	FLPLLAASFANM <sub>2ox</sub> RLT-OH	FLPLLAASFAC(+57.02)TVTKKC(+57.02)
	Бревинин 1AVb	FVPLLVSKLVCVVTKKC-OH	FVPLLVSKLVACTQRVT-OH	(+57.02)FVPLLVSKLVC(+57.02)VVTKKC(+57.02)
	Бревинин 2 AV	GLMSVLKKAGKHVAKNVAVSLMDSLK CKISGGC-OH	LGMSVLKNAAKHVAKQRVSLMDSLQLCML W-OH	Не обнаружен (минорный компонент)
	Ранатуерин 2AVb	GLMDMVKGAAKNL FASALDTLK CKITGC-OH	LGMDMVKM <sub>ox</sub> LVNEASALLCLKRM <sub>ox</sub> WN- OH	Не обнаружен модифицированный
	Ранатуерин 2AVc	GLLDVVKGAAKNLLATALDKLK CKVTGC-OH	LGLDVVKAQLELLATALDKLKH <sub>ox</sub> FW <sub>ox</sub> S-OH	Не обнаружен модифицированный

- серым выделены совпадающие фрагменты последовательностей при ручном и автоматизированном *de novo* секвенировании

## 2.7. Сравнение результатов ручного и автоматического секвенирования нетриптических пептидов

В таблице 15 суммированы результаты сравнения автоматического (PEAKS) секвенирования и ручной интерпретации масс-спектров интактных пептидов кожных секретов *R. arvalis* и *R. temporaria*. В работе тестируется метод определения последовательностей пептидов «сверху вниз» применительно к кожным пептидомам ранидных лягушек для изучения популяционных различий внутри одного вида амфибий.

Как отмечалось выше, большинство автоматических алгоритмов нацелено на работу с триптическими пептидами, лишенными S-S связей, имеющими C-терминальный положительный заряд (Lys/Arg) и небольшое число звеньев в последовательности. Поскольку алгоритмы программ непрерывно совершенствуются, представлялось интересным протестировать работу распространенного программного продукта PEAKS, как правило, сопровождающего масс-спектрометры, сконструированные на основе орбитальной ловушки (Thermo Scientific). Результаты PEAKS в таблице 15 выданы по сумме данных двух методов: (ДАС+ДАСПЭ), спектры поскольку EThcD регистрировали не для всех образцов. Помимо дисульфидных, в таблице 15 присутствуют также короткие амидированные темпорины, хотя в неё не вошли брадикин со структурными аналогами и их многочисленными протеолитическими копиями.

Аналізу подвергались интактные образцы («сверху вниз»). Итак, ни один S-S – содержащий пептид не был секвенирован полностью. Программа PEAKS удовлетворительно проходит линейную часть таких пептидов, но по сумме двух методов не справляется с «Rana vox». В этих случаях PEAKS начинает злоупотреблять приписыванием несуществующих в реальных последовательностях модификациями, такими как окисление метионина, карбонилирование и так далее. Программа вместе с тем уверенно справилась с секвенированием двух мелиттин-родственных пептидов FQ-22-1 и АК-23-1 в словенской популяции *R. arvalis* (22 и 23 а.к., соответственно). Самым распространенным недостатком является неопределенное взаимоположение в паре аминокислот, особенно, в N-терминальной, а также пропуск разрыва Gly-Gly, вместо которого программа выдает Asn.

Сиквенсы длинных бревининов 2 (например, бревинин 2Т) программа проходит существенно хуже – она лишь фрагментарно определяет участки последовательности этого 33-звенного пептида. К недостаткам также можно отнести пропуск в секвенировании некоторых пептидов, когда алгоритм считает полученный спектр малоинтенсивным. Это, как правило, относится к минорным компонентам анализируемой смеси. Хотя надо

отметить, что благодаря этому автосеквенатору, мы обнаружили свой пропуск в *de novo* секвенировании темпорина Т при анализе спектров центральнословенской *R. temporaria*.

Программа довольно уверенно секвенирует короткие темпорины, несмотря на полное отсутствие в некоторых из них основных аминокислот. Несмотря на малые размеры, это не самые удобные для *de novo* секвенирования пептиды. Их спектры часто осложнены вторичной фрагментацией из-за Pro<sup>3</sup>; отсутствующим разрывом Gly-Lys/Gly-Arg в середине последовательности; повышенного числа Leu/Ile в структуре, что не всегда ведет к активной фрагментации, а также возможной потерей аммиака протонированной молекулой до фрагментации, что приводит к многополосным и трудно интерпретируемым спектрам. Тем не менее из 23-х темпоринов таблицы 16 PEAKS правильно секвенировал 10. Таким образом, воспользоваться результатами автоматического секвенирования может исследователь, давно и скрупулезно работающий с компонентами кожных пептидомов ранидных амфибий и четко представляющий, какие пептидные семейства могут в них присутствовать.

Таблица 15. Сравнение результатов ручного и автоматического секвенирования интактных пептидов с S-S и без S-S связей

Вид лягушки	Пептид	Ручное <i>de novo</i> секвенирование, интактный пептид (ДАС + ДАСПЭ + EThcD)	<i>De novo</i> секвенирование PEAKS, интактный пептид (ДАС + ДАСПЭ)
<i>Rana arvalis</i> , Центральная Словения	Бревинин 1AVa	FLPLLAASFACTVTKKC-OH	FLPLLAASFANM <sub>2ox</sub> RLT-OH
	Бревинин 1AVb	FVPLLVS <del>KL</del> VCVVTKKC-OH	FVPLLVS <del>KL</del> VACTQRVT-OH
	Бревинин 2 AV	GLMSVLK <del>AKG</del> HVAKNVAVSLMDSLKCKISGGC-OH	LGMSVLK <del>NAAK</del> HVAKQRVSLMDSLQ <del>LC</del> M <del>L</del> W-OH
	Ранатуерин 2AVb	GLMDMVKGAAKNL <del>FASAL</del> DTLCKKITGC-OH	LGMDMVKVM <sub>ox</sub> L <del>VNEAS</del> ALLCLKRM <sub>ox</sub> WN-OH
	Ранатуерин 2AVc	GLLDVVKGA <del>AKN</del> LLATALDKLKCKVTGC-OH	LGLDVVKAQLELLATALDKLKH <sub>ox</sub> FW <sub>ox</sub> S-OH
	FQ-22-1	FVGAALKVLANVLP <del>PPVIS</del> WIKQ-NH <sub>2</sub>	VFGAALKVLANVLP <del>VLS</del> WLKQ-NH <sub>2</sub>
	AK-23-1	ALGGVLKALAKGLPSVISWINQK-NH <sub>2</sub>	ALNVLKALAKGLPSVLSWLNQK-NH <sub>2</sub>
	Темпорин AV	LLPLVGNLNDLL-NH <sub>2</sub>	LLPLVGNLNDLLL-NH <sub>2</sub>
	Темпорин AVa	GLEVLGKILSGIL-NH <sub>2</sub>	LGSLLGKLLDGKP-OH
<i>Rana temporaria</i> , Центральная Словения	Бревинин 1T	VNPIILGVLPKFVCLITKKC-OH	VNPLLLGVLPKFNEH <sub>ox</sub> LLPPA-NH <sub>2</sub>
	Бревинин 1Ta	FITLLLRKFICSITKKC-OH	FTLLLRKFLAAM <sub>ox</sub> WRM-NH <sub>2</sub>
	Бревинин 1Tb	LVPLFLSKLICFITKKC-OH	VDM(+15.99)PEEFGDEYFEDVE
	Бревинин 1Tc	LVPMFLSKLICFITKKC-OH	LVPMFLSKLLPHFTQPN-OH
	Бревинин 2T	GLLSGLKKVKGK <del>HVA</del> KNVAVSLMDSLKCKISGDC-OH	LGLSGLKVKKG <del>HVA</del> KRQVSLMDSLKYM <sub>ox</sub> QHE-OH
	Бревинин 2Te	GLLDGLKKVKGK <del>HVA</del> KNVAVSLMDSLKCKISGDC-OH	abs
	MRP FQ-22	FIGSALKVLAGVLP <del>SVIS</del> WVKQ-NH <sub>2</sub>	LASGFLK <del>VLAGVLP</del> SVLSWVKQ-NH <sub>2</sub>
	Темпорин А	FLPLIGRVLSGIL-NH <sub>2</sub>	FLPLLGVRLSGILL-NH <sub>2</sub>
	Темпорин В	LLPIVGNLLKSLL-NH <sub>2</sub>	LLPLVGNLLKSLL-NH <sub>2</sub>
	Темпорин С	LLPIIGNLLNGLL-NH <sub>2</sub>	LLPLIGNLLNGLL-NH <sub>2</sub>
	Темпорин D	LLPIVGNLLNSLL-NH <sub>2</sub>	LLPLVGNLVKSKP-OH
	Темпорин E	VLPIIGNLLNSLL-NH <sub>2</sub>	VLPLIGNLSKVPK-OH
	Темпорин F	FLPLIGKVLSGILL-NH <sub>2</sub>	abs
	Темпорин G	FFPVIGRI <del>LN</del> GIL-NH <sub>2</sub>	FFPVLGRLLNGLL-NH <sub>2</sub>
	Темпорин H	LSPNLLKSLL-NH <sub>2</sub>	abs
	Темпорин K	LLPNLLKSLL-NH <sub>2</sub>	abs
	Темпорин L	FVQWFSKFLGRIL-NH <sub>2</sub>	FVQWFSKFLGRILL-NH <sub>2</sub>
	Темпорин N	FXGAXGNAXSRVX-NH <sub>2</sub>	FLGALGNALSRVL-NH <sub>2</sub>
	Темпорин O	FXGALVNAXRGXX-NH <sub>2</sub>	FLGALVNALRGLL-NH <sub>2</sub>
	Темпорин P	XVPXXGNXXSGXX-NH <sub>2</sub>	VLPLIGNLLSGILL-NH <sub>2</sub>
	Темпорин Q	LVPFXGKTXGGLLARF-NH <sub>2</sub>	LVPFLGKTLGGLLARF-NH <sub>2</sub>
	Темпорин R	XVPFXGRTLGGLLARX-NH <sub>2</sub>	LVPFLGK(+CO)TLGGLLARL-NH <sub>2</sub>
	Темпорин S	LVPFXGRTXGGLLARF-NH <sub>2</sub>	L(+CO)VPFLGKTLGGLLARF-NH <sub>2</sub>
	Темпорин T	abs	FXGALVNAXTRVX-NH <sub>2</sub>

Вид лягушки	Пептид	Ручное <i>de novo</i> секвенирование, интактный пептид (ДАС + ДАСПЭ + EThcD)	<i>De novo</i> секвенирование PEAKS, интактный пептид (ДАС + ДАСПЭ)
<i>Rana temporaria</i> , Архангельская область	Бревинин 1Т	VNPIILGVLPK FVCLITKKC-OH	VDPLLLGVNPWWVGLLLHK(+CO)-NH <sub>2</sub>
	Бревинин 1Та	FITLLLRK FICSITKKC-OH	FLTLLLK(+CO)KFGTNFLTEL-NH <sub>2</sub>
	Бревинин 1Тб	LVPLFLSKLICFITKKC-OH	LVPLFLSKWTPPYLPPP-NH <sub>2</sub>
	Бревинин 2Т	GLLSGLKKVKGKHVAKNVAVSLMDSLKCKISGDC-OH	LGLSGLKKVKGKHVAKNVAVSLMDSLKM <sub>ox</sub> EEEGL
	Бревинин 2Тd	GLWETIKNFGKKFTLNILHNLKCKIGGGC-OH	LGWETLKNFGKKFTLNLLHNL TSAEWLG
	MRP FQ-22	FIGSALKVLAGVLP SIVSWVKQ-NH <sub>2</sub>	LFTPVSLNKGVLLAVLSWVKQ-NH <sub>2</sub>
	Темпорин В	LLPLVGNLLK SLL-NH <sub>2</sub>	LLPLVGNLLK SLL-NH <sub>2</sub>
	Темпорин С	LLPILGNLLN GLL-NH <sub>2</sub>	LLPLLGNLLN GLL-NH <sub>2</sub>
	Темпорин D	LLPIVGNLLN SLL-NH <sub>2</sub>	LLPLLGV SGLN GLL-NH <sub>2</sub>
	Темпорин F	FLPLIGK VLSG LL-NH <sub>2</sub>	FLPLLGK VLSG LL-NH <sub>2</sub>
	Темпорин G	FFPVIGRILN GIL-NH <sub>2</sub>	FFPVLGRLLN GLL-NH <sub>2</sub>
	Темпорин H	LSPNLLK SLL-NH <sub>2</sub>	LSPNLLK SLL-NH <sub>2</sub>
	Темпорин К	LLPNLLK SLL-NH <sub>2</sub>	LLPNLLK SLL-NH <sub>2</sub>
	Темпорин L	FVQWFSKFLGRIL-NH <sub>2</sub>	FVQWFSKFLGRLL-NH <sub>2</sub>
	Темпорин М	FLPLLGK VLSR LV-NH <sub>2</sub>	abs
	Темпорин N	FLGALGNALSRVL-NH <sub>2</sub>	FLQLGNALSRVL-NH <sub>2</sub>
	Темпорин Т	FLGALVNALTRVL-NH <sub>2</sub>	FLQLVNALTRVL-NH <sub>2</sub>
	Темпорин U	FLGAVLHALTGLL-NH <sub>2</sub>	FLQLVHALTGLL-NH <sub>2</sub>
	Темпорин V	FLGALVHALTSLL-NH <sub>2</sub>	FLQHLVALTSLL-NH <sub>2</sub>
	Темпорин W	FIGALVNALSRVL-NH <sub>2</sub>	FLQLVNALSRVL-NH <sub>2</sub>
Темпорин X	FLGALVGALSRVL-NH <sub>2</sub>	LFGALVGALSRVL-NH <sub>2</sub>	

### 3. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

В работе использовали следующие реагенты: ацетонитрил (Sigma-Aldrich, «для ВЭЖХ»), метанол (Fluka, «для ВЭЖХ»), муравьиная кислота (Fluka, «для ВЭЖХ»), йодацетамид (Acros, 98%), дитиотреитол (Acros, 99.4%), гидрокарбонат аммония (Sigma, extra pure), деионизованная вода.

Общая схема эксперимента приведена на рис. 51.

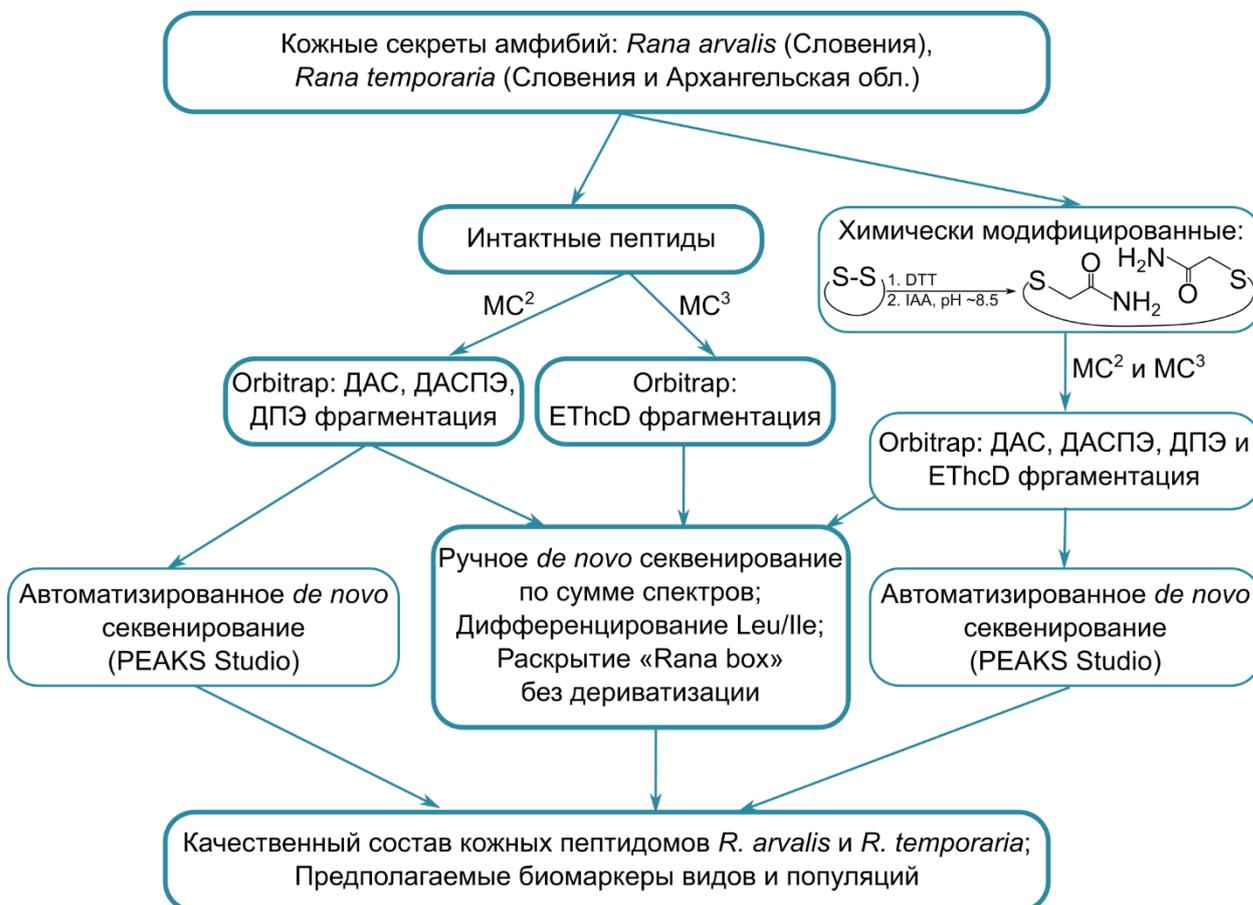


Рисунок 51. Общая схема эксперимента

#### Получение кожных секретов

Для получения кожных секретов перечисленных видов лягушек были отловлены половозрелые особи, которые после электростимуляции были отпущены в места отлова. Две особи *Rana arvalis* (самки) и самец *Rana temporaria* были пойманы в окрестностях Люблины. Для изучения пептидома архангельской лягушки *Rana temporaria* в окрестностях Архангельска был отобран половозрелый самец этого вида. Кожные секреты получали нетравматическим методом электростимуляции кожных желез, описанным в [271]. Метод заключается в стимулировании увлажнённой спины амфибий, на которой расположены железы, импульсными токами малой величины в течение 30-40 секунд с помощью биполярных платиновых электродов лабораторного электростимулятора ЭСЛ-1 (Electro Science Laboratories, King of Prussia, PA, USA).

Применяли следующие параметры стимуляции: напряжение на электродах 10 В; длительность импульсов 3-5 мс (в зависимости от размеров лягушек) при частоте 50 Гц. Секреты смывались 25 мл деионизованной воды в контейнер с равным объёмом метанола. Полученный раствор центрифугировали в течение 15 минут при скорости вращения 3000 об/мин, фильтровали через мембранный фильтр Millex-NV с размерами пор 0.45 мкм (Millipore, Billerica, MA, США), концентрировали на роторном испарителе при температуре 35°C до 1 мл и лиофильно высушивали. Образцы кожных секретов хранили при температуре -26°C.

#### **Очистка кожных секретов с помощью ВЭЖХ**

ВЭЖХ очистка кожных секретов проводилась на обращенно-фазовой колонке C18 (5 мкм, 100 Å, 150x4мм; DrMaisch GmbH, Германия), уравновешенной раствором 10%-ного ацетонитрила в 0.1%-ной водной трифторуксусной кислоте (ТФУК). Объем закальваемой пробы составлял 20 мкл. Элюирование проводили на жидкостном хроматографе «Thermo Separation Products», оборудованном бинарным насосом «Spectra System P2000». Использовали линейный градиент ацетонитрила в диапазоне его концентраций 10 - 70% в 0.1% водной трифторуксусной кислоте за 40 мин при постоянной скорости элюирования 1.0 мл/мин. Хроматографирование осуществляли на рабочей полосе поглощения 214 нм (ультрафиолетовый детектор «Spectra System UV3000»). Элюат собирали в интервале времен удерживания 10 - 40 минут, пробы объединяли, концентрировали на роторном испарителе и лиофильно высушивали.

#### **Очистка кожных секретов с помощью ZipTip перед электрораспылением**

Все спектры были получены на масс-спектрометрах с орбитальной ловушкой, оборудованных наноспреем, что требовало дополнительной очистки всех образцов кожных секретов (интактных и модифицированных). Для очистки использовали наконечники ZipTip (Millipore, micro-C18), которые подготавливали следующим образом: адсорбент однократно промывали 10-ю мкл 50%-го водного раствора ацетонитрила и уравновешивали 10-ю мкл 0.1% водной ТФУ в течение 5 минут. Для образцов модифицированных секретов применяли разбавление пробы равным объемом 0.1%-ной водной ТФУ. Дериватизованные пептиды сорбировали на носитель наконечника ZipTip медленным 10-кратным отбором сбросом пробы, после чего промывали 0.1%-ной водной ТФУ от следов реакционной пробы, как было указано выше. Десорбцию пептидов дериватизованных образцов кожных секретов с носителя ZipTip проводили тремя мкл 50%-го водного ацетонитрила, помещенного в пробирку «Эппендорф» путем 10-кратных последовательных отборов сбросов указанным количеством растворителя через наконечник ZipTip, содержащий адсорбированный пептидный материал.

### **Восстановление и алкилирование дисульфидных связей**

Очищенные и лиофилизированные фракции пептидов растворяли в аммиачном буфере (100 мМ  $\text{NH}_4\text{HCO}_3$ , pH 8.0) в присутствии 10 мМ дитиотреитола. Смесь активно перемешивали при 40°C на термошейкере (Biosan) в течение 30 минут, а затем охлаждали до комнатной температуры. Затем прибавляли раствор йодацетамида до конечной концентрации 10 мМ и выдерживали смесь при комнатной температуре в тёмном месте в течение часа.

### **Окисление дисульфидных связей**

Для получения надмуравьиной кислоты муравьиную кислоту смешивали с 30% раствором перекиси водорода в соотношении 19:1 и выдерживали в течение часа при комнатной температуре. К растворённому в минимальном объёме муравьиной кислоты образцу пептида прибавляли трёхкратный объём полученного раствора надмуравьиной кислоты. Растворы до смешения и во время реакции охлаждали до 0°C и выдерживали при этой температуре в течение часа. После окончания реакции прибавляли пятикратный избыток деионизованной воды и лиофилизировывали полученный субстрат [163].

### **Масс-спектрометрия**

Масс-спектрометрические эксперименты были выполнены на двух масс-спектрометрах с орбитальной ловушкой ионов.

### ***Rana temporaria* и *Rana arvalis*, Словения**

Анализ образцов словенских популяций *Rana arvalis* и *Rana temporaria* проводился на масс-спектрометре Orbitrap Elite ETD (Thermo Scientific, Германия), оборудованном хроматографом Easy nano-LC 1000 (Thermo Scientific, США). Для разделения компонентов использовали подготовленную в лаборатории колонку 75 мкм × 150 мм, заполненную сорбентом Aeris 3.6 мкм WIDEPORЕ XB-C18 (Phenomenex, США). Элюенты для хроматографии: раствор А – 0.1% раствор муравьиной кислоты в деионизованной воде (Milli-Q, Millipore), раствор В – смесь 80% ацетонитрила и 20% раствора А. Разделение проводили при скорости потока 250 нл/мин, градиент – от 5 до 60% раствора В за 120 мин.

Масс-спектры были зарегистрированы в режиме DDA, регистрировали положительные ионы. Разрешение составило 240 000 для иона с  $m/z$  400 в режиме MS и 60 000 для того же иона в режиме MS/MS. Диапазон сканирования  $m/z$  составлял 150-2700 (MS) с последующим выбором пяти наиболее интенсивных ионов для получения их тандемных спектров. При этом применяли динамическое исключение, размер списка составил 25 ионов. Время накопления сигнала для MS спектров составляло 100 мс и 200 мс – для получения MS<sup>2</sup> спектров.

Спектры ДАС, ДАСПЭ и ДПЭ пяти наиболее интенсивных ионов-предшественников в каждом сканировании регистрировались в автоматическом режиме при следующих параметрах:

число повторений записи спектра  $MS^2$  – 2 (repeat count);

размер листа отложенной записи спектров  $MS^2$  – 25 (exclusion list size);

время отложенной повторной записи  $MS^2$  спектров – 200 сек (exclusion duration);

минимальный порог интенсивности пика иона при отборе пяти целевых ионов-предшественников для получения их фрагментных спектров  $MS^2$  – 5000 (minimum signal threshold);

ширина изолирующего окна – 2  $m/z$  (isolation width)

Параметры работы масс-спектрометра в экспериментах были следующие:

входное напряжение на капилляре – 1,6 кВ;

температура капилляра на входе – 200 °С;

нормализованная энергия соударений (NCE) в режимах ДАС и ДАСПЭ – 28 и 35;

пороговая интенсивность ионов для ДПЭ – 300 000 (AGC target for ETD);

время активации в ДПЭ – 100 мс.

Флуорантен использовался для образования анионов, реагирующих с протонированными молекулами-предшественниками пептидов с зарядом  $z \geq 2^+$ .

Для регистрации EThcD спектров использовали трибридный масс-спектрометр Fusion Orbitrap (Thermo Scientific, США), оборудованный хроматографом Easy nano-LC 1000 (Thermo Scientific, США). Для разделения компонентов использовали колонку EASY-Spray (PerMap, Thermo Scientific, США) со следующими параметрами: 50 см, C18, 2 мкм, 100Å. Элюенты для хроматографии: раствор А – 0.1% муравьиной кислоты и 2% ацетонитрила в деионизованной воде (Milli-Q, Millipore), раствор В – 0.1% муравьиной кислоты и 2% деионизованной воды в ацетонитриле. Разделение проводили при скорости потока 300 нл/мин, градиент – 4% раствора В в течение 7 мин, от 4 до 50% раствора В в течение 21 мин, от 50 до 80% раствора В в течение 5 мин, 80% раствора В в течение 5 мин, 95% и от 95 до 4% раствора В в течение 5 мин.

Масс-спектры были зарегистрированы в режиме DDA, регистрировали положительные ионы. Разрешение составило 120 000 для иона в режиме MS и 15 000 в режиме MS/MS. Диапазон сканирования  $m/z$  составлял 300-1700 (MS). Выбор ионов-прекурсоров проводили в режиме «top speed» для ионов с зарядовым состоянием от 2 до 7 и минимальной интенсивностью 50 000.

***Rana temporaria*, Архангельск**

Анализ образцов архангельской популяции *Rana temporaria* проводился на масс-спектрометре Fusion Orbitrap (Thermo Scientific, США), оборудованном хроматографом Easy nano-LC 1000 (Thermo Scientific, США). Для разделения компонентов использовали колонку EASY-Spray (PerMap, Thermo Scientific, США) со следующими параметрами: 50 см, C18, 2 мкм, 100Å. Элюенты для хроматографии: раствор А – 0.1% муравьиной кислоты и 2% ацетонитрила в деионизованной воде (Milli-Q, Millipore), раствор В – 0.1% муравьиной кислоты и 2% деионизованной воды в ацетонитриле. Разделение проводили при скорости потока 300 нл/мин, градиент – 5-60% раствора В в течение 60 мин.

Масс-спектры были зарегистрированы в режиме DDA, регистрировали положительные ионы. Разрешение составило 240 000 для иона в режиме MS и 60 000 в режиме MS/MS) с последующим выбором пяти наиболее интенсивных ионов для получения их тандемных спектров. При этом применяли динамическое исключение (продолжительность 5 с). Время накопления сигнала для MS спектров составляло 100 мс и 200 мс – для получения MS<sup>2</sup> спектров.

**Соблюдение этических норм работы с животными.**

Все работы проводились в соответствии с принципами, изложенными в Приложении А Европейской конвенции по защите позвоночных животных, используемых для экспериментальных и научных целей (ETS №123), принятом в Страсбурге 15 июня 2006 года, а также в соответствии с ГОСТ 33219-2014, разработанным Евро-Азийским Советом по стандартизации, метрологии и сертификации (EASC).

## ЗАКЛЮЧЕНИЕ

В ходе проделанной работы были сделаны следующие выводы:

1. Разработан новый метод установления первичных структур пептидов в варианте «сверху вниз» без использования каких-либо химических модификаций, исключительно методами тандемной масс-спектрометрии, доступными на масс-спектрометрах с орбитальной ловушкой: ДАС, ДАСПЭ, ДПЭ и EThcD. Разработанный метод успешно апробирован на трёх кожных пептидомах амфибий семейства *Ranidae* из разных популяций.

2. Продемонстрирована максимальная информативность EThcD спектров при установлении кожных пептидомов новым методом «сверху вниз», включая длинные дисульфидсрезающие пептиды, новые и установленные ранее с помощью клонирования кДНК или дериватизации дисульфидной связи; и короткие, но сложные для секвенирования темпорины.

3. Предложена схема фрагментации пептидов с одной дисульфидной связью в условиях EThcD эксперимента с образованием *c*-катион-радикалов и чётноэлектронных *z*-ионов, содержащих тиоморфолиноновый фрагмент. Показано, что механизм реализуется для всех изученных семейств дисульфидсодержащих пептидов и сопровождается фрагментацией пептидных связей внутри *C*-концевого S-S цикла («Rana box»).

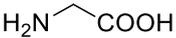
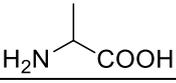
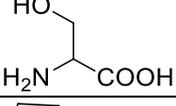
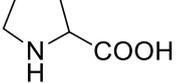
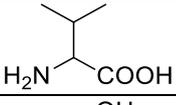
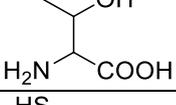
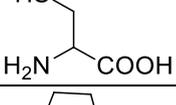
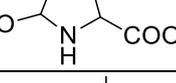
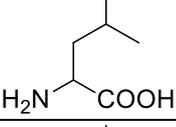
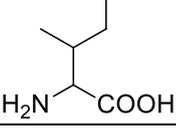
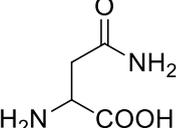
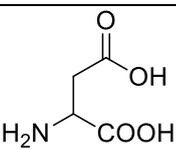
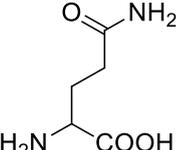
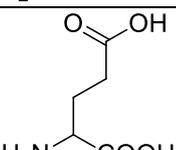
4. Показана применимость программы автосеквенирования PEAKS Studio для коротких пептидов семейства темпоринов и коротких бревининов 1 (до 24 а.к.). Программа позволяет частично секвенировать линейную часть интактных пептидов или установить полную последовательность модифицированных пептидов в случае их высокой концентрации.

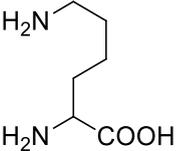
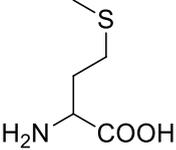
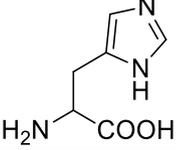
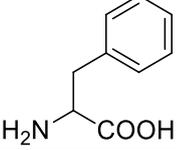
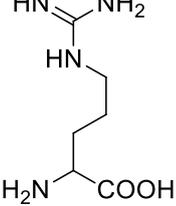
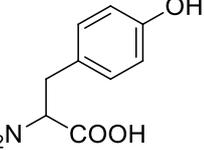
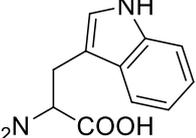
## СПИСОК СОКРАЩЕНИЙ

- Br – брадикинин
- CBS – секвенирование на основе состава (Composition-Based Sequencing)
- CHCA –  $\alpha$ -циано-4-гидроксикоричная кислота
- DAN – 1,5-диаминонафталин
- DHB – 2,5-дигидроксibenзойная кислота
- EST – экспрессированный тег последовательности (Expressed Sequence Tag)
- EThcD – гибридная фрагментация при переносе электрона и активации соударением (Electron-Transfer/ Higher-Energy Collision Dissociation)
- IAA – йодацетамид (Iodoacetamide)
- IAAC – йодоуксусная кислота (Iodoacetic Acid)
- NCE – нормализованная энергия соударений (Normalized Collision Energy)
- NIS – нормализованный изотопный сдвиг (Normalized Isotopic Shift)
- NMD – нормализованный дефект масс (Normalized Mass Defect)
- NPM – *N*-фенилмалеимид
- ppm – миллионная доля (parts per million)
- QSAR – Поиск количественных соотношений структура-свойство (Quantitative Structure-Activity Relationship)
- SORI – непрерывное вне-резонансное возбуждение (Sustained Off-Resonance Irradiation)
- SRTX – сарафототоксин (Sarafototoxin)
- SWIFT – Метод сохранённой формы волны с обратным преобразованием Фурье (Stored Waveform Inverse Fourier Transform)
- ТСЕР – трис(2-карбоксиэтил)фосфин
- ТНPs – триптофиллины
- АМП – антимикробные пептиды
- БРП – брадикинин-родственные пептиды
- ВП – времяпролётный масс-анализатор
- ВЭЖХ – высокоэффективная жидкостная хроматография
- Да – Дальтон
- ДАС (CID) – диссоциация, активированная соударениями (Collision Induced Dissociation)
- ДАСПЭ (HCD) – диссоциация, активированная соударениями при повышенной энергии (Higher-Energy Collision Induced Dissociation)
- ДЗЭ (ECD) – диссоциация при захвате электрона (Electron Capture Dissociation)

ДИП – диссоциация, индуцированная соударением с поверхностью  
ДНК – дезоксирибонуклеиновая кислота  
ДПЭ (ETD) – диссоциация при переносе электрона (Electron Transfer Dissociation)  
ДТТ – дитиореитол  
ИКМФД – инфракрасная мультифотонная диссоциация  
ИЦР – ионный циклотронный резонанс  
ИЭР – ионизация электрораспылением  
кДНК – комплементарная дезоксирибонуклеиновая кислота  
МАЛДИ – матрично-активированная лазерная десорбция/ионизация  
МИК – минимальная ингибирующая концентрация  
ММ – молекулярная масса  
МС – масс-спектрометрии  
МС/МС – тандемная масс-спектрометрия  
ППМ – протеомная поисковая машина  
ПТМ – посттрансляционная модификация  
ПФ – преобразование Фурье  
Тл – Тесла

## ОБЩЕПРИНЯТЫЕ ОБОЗНАЧЕНИЯ АМИНОКИСЛОТ

Аминокислота		Точная масса остатка	Трёхбуквенное обозначение	Однбуквенное обозначение
Глицин		57.021	Gly	G
Аланин		71.037	Ala	A
Серин		87.032	Ser	S
Пролин		97.053	Pro	P
Валин		99.068	Val	V
Треонин		101.048	Thr	T
Цистеин		103.009	Cys	C
Гидроксипролин		113.048	Hyp	Hp
Лейцин		113.084	Leu	L
Изолейцин		113.084	Ile	I
Аспарагин		114.043	Asn	N
Аспарагиновая кислота		115.027	Asp	D
Глутамин		128.059	Gln	Q
Глутаминовая кислота		129.043	Glu	E

Аминокислота	Точная масса остатка	Трёхбуквенное обозначение	Однбуквенное обозначение
Лизин		Lys	K
Метионин		Met	M
Гистидин		His	H
Фенилаланин		Phe	F
Аргинин		Arg	R
Тирозин		Tyr	Y
Триптофан		Trp	W

**СПИСОК ЛИТЕРАТУРЫ**

1. Kastin A. Handbook of Biologically Active Peptides // Handb. Biol. Act. Pept. Elsevier Inc., 2013.
2. Conlon J.M. Host-defense peptides of the skin with therapeutic potential: From hagfish to human // Peptides. Elsevier, 2015. Vol. 67. P. 29–38.
3. Masso-Silva J.A., Diamond G. Antimicrobial Peptides from Fish // Pharm. 2014, Vol. 7, Pages 265-310. Multidisciplinary Digital Publishing Institute, 2014. Vol. 7, № 3. P. 265–310.
4. Najafian L., Babji A.S. A review of fish-derived antioxidant and antimicrobial peptides: Their production, assessment, and applications // Peptides. Elsevier, 2012. Vol. 33, № 1. P. 178–185.
5. van Hoek M.L. Antimicrobial Peptides in Reptiles // Pharm. 2014, Vol. 7, Pages 723-753. Multidisciplinary Digital Publishing Institute, 2014. Vol. 7, № 6. P. 723–753.
6. Klint J.K., Senff S., Rupasinghe D.B., Er S.Y., Herzig V., Nicholson G.M., King G.F. Spider-venom peptides that target voltage-gated sodium channels: Pharmacological tools and potential therapeutic leads // Toxicon. Pergamon, 2012. Vol. 60, № 4. P. 478–491.
7. Xu X., Lai R. The Chemistry and Biological Activities of Peptides from Amphibian Skin Secretions // Chem. Rev. 2015. Vol. 115, № 4. P. 1760–1846.
8. Bowie J.H., Separovic F., Tyler M.J. Host-defense peptides of Australian anurans. Part 2. Structure, activity, mechanism of action, and evolutionary significance // Peptides. Elsevier, 2012. Vol. 37, № 1. P. 174–188.
9. König E., Bininda-Emonds O.R.P., Shaw C. The diversity and evolution of anuran skin peptides // Peptides. 2015. Vol. 63. P. 96–117.
10. Conlon J.M., Mechkarska M., Lukic M.L., Flatt P.R. Potential therapeutic applications of multifunctional host-defense peptides from frog skin as anti-cancer, anti-viral, immunomodulatory, and anti-diabetic agents // Peptides. Elsevier Inc., 2014. Vol. 57. P. 67–77.
11. Harrison P.L., Abdel-Rahman M.A., Miller K., Strong P.N. Antimicrobial peptides from scorpion venoms // Toxicon. Pergamon, 2014. Vol. 88. P. 115–137.
12. Wang G. Human Antimicrobial Peptides and Proteins // Pharm. 2014, Vol. 7, Pages 545-594. Multidisciplinary Digital Publishing Institute, 2014. Vol. 7, № 5. P. 545–594.
13. Tam J.P., Wang S., Wong K.H., Tan W.L. Antimicrobial Peptides from Plants // Pharm. 2015, Vol. 8, Pages 711-757. Multidisciplinary Digital Publishing Institute, 2015. Vol. 8, № 4. P. 711–757.
14. Kardong K. V. Snake Toxins and Venoms: An Evolutionary Perspective // Herpetologica.

- [Herpetologists' League, Allen Press], 1996. Vol. 52, № 1. P. 36–46.
15. Wong J., Xia L., Ng T. A Review of Defensins of Diverse Origins // *Curr. Protein Pept. Sci.* Bentham Science Publishers Ltd., 2007. Vol. 8, № 5. P. 446–459.
  16. Roseghini M., Erspamer G.F., Severini C., Simmaco M. Biogenic amines and active peptides in extracts of the skin of thirty-two European amphibian species // *Comp. Biochem. Physiol. Part C Comp. Pharmacol.* Pergamon, 1989. Vol. 94, № 2. P. 455–460.
  17. Haslam I.S., Roubos E.W., Mangoni M.L., Yoshizato K., Vaudry H., Kloepper J.E., Pattwell D.M., Maderson P.F.A., Paus R. From frog integument to human skin: Dermatological perspectives from frog skin biology // *Biol. Rev.* Blackwell Publishing Ltd, 2014. Vol. 89, № 3. P. 618–655.
  18. Almeida P.F., Pokorny A. Mechanisms of antimicrobial, cytolytic, and cell-penetrating peptides: From kinetics to thermodynamics // *Biochemistry.* American Chemical Society, 2009. Vol. 48, № 34. P. 8083–8093.
  19. Neurotoxins in Neurobiology: Their Actions and Applications. 1st ed. / ed. Dajas F., Tipton K.F. CRC Press, 1994. 196 p.
  20. Koczulla A.R., Bals R. Antimicrobial Peptides // *Drugs.* 2003. Vol. 63, № 4. P. 389–406.
  21. Gordon Y.J., Romanowski E.G., McDermott A.M. A Review of Antimicrobial Peptides and Their Therapeutic Potential as Anti-Infective Drugs // *Curr. Eye Res.* 2005. Vol. 30, № 7. P. 505–515.
  22. Erspamer V., Erspamer G.F., Mazzanti G., Endean R. Active peptides in the skins of one hundred amphibian species from Australia and Papua New Guinea // *Comp. Biochem. Physiol. Part C Comp. Pharmacol.* Pergamon, 1984. Vol. 77, № 1. P. 99–108.
  23. Simmaco M., Mignogna G., Barra D. Antimicrobial peptides from amphibian skin: What do they tell us? // *Biopolymers.* 1998. Vol. 47, № 6. P. 435–450.
  24. Luisa Mangoni M., Di Grazia A., Cappiello F., Casciaro B., Luca V. Naturally Occurring Peptides from *Rana temporaria*: Antimicrobial Properties and More // *Curr. Top. Med. Chem.* 2015. Vol. 16, № 1. P. 54–64.
  25. Pukala T.L., Bowie J.H., Maselli V.M., Musgrave I.F., Tyler M.J. Host-defence peptides from the glandular secretions of amphibians: Structure and activity // *Natural Product Reports.* The Royal Society of Chemistry, 2006. Vol. 23, № 3. P. 368–393.
  26. Xi X., Li B., Chen T., Kwok H. A Review on Bradykinin-Related Peptides Isolated from Amphibian Skin Secretion // *Toxins (Basel).* 2015. Vol. 7, № 3. P. 951–970.
  27. Zasloff M. Antimicrobial peptides of multicellular organisms // *Nature.* 2002. Vol. 415, № 6870. P. 389–395.
  28. Samgina T.Y., Vorontsov E.A., Gorshkov V.A., Hakalehto E., Hanninen O., Zubarev R.A.,

- Lebedev A.T. Composition and Antimicrobial Activity of the Skin Peptidome of Russian Brown Frog *Rana temporaria* // *J. Proteome Res.* 2012. Vol. 11, № 12. P. 6213–6222.
29. Samgina T.Y., Tolpina M.I., Hakalehto E., Artemenko K.A., Bergquist J., Lebedev A.T. Proteolytic degradation and deactivation of amphibian skin peptides obtained by electrical stimulation of their dorsal glands // *Anal. Bioanal. Chem.* 2016. Vol. 408, № 14. P. 3761–3768.
  30. Verano-Braga T., A. A. Dutra A., R. León I., N. Melo-Braga M., Roepstorff P., M. C. Pimenta A., Kjeldsen F. Moving Pieces in a Venomic Puzzle: Unveiling Post-translationally Modified Toxins from *Tityus serrulatus* // *J. Proteome Res.* 2013. Vol. 12, № 7. P. 3460–3470.
  31. Chen X., Wang H., Wang L., Zhou M., Chen T., Shaw C. Identification of Miscellaneous Peptides from the Skin Secretion of the European Edible Frog, *Pelophylax kl. Esculentus* // *Protein J.* Springer New York LLC, 2016. Vol. 35, № 4. P. 291–299.
  32. Vanhoye D., Bruston F., Nicolas P., Amiche M. Antimicrobial peptides from hylid and ranin frogs originated from a 150-million-year-old ancestral precursor with a conserved signal peptide but a hypermutable antimicrobial domain // *Eur. J. Biochem.* John Wiley & Sons, Ltd, 2003. Vol. 270, № 9. P. 2068–2081.
  33. F. Medzihradzsky K., J. Bohlen C. Partial De Novo Sequencing and Unusual CID Fragmentation of a 7 kDa, Disulfide-Bridged Toxin // *J. Am. Soc. Mass Spectrom.* 2012. Vol. 23, № 5. P. 923–934.
  34. Bhatia S., Kil Y.J., Ueberheide B., Chait B.T., Tayo L., Cruz L., Lu B., Yates J.R., Bern M. Constrained de novo sequencing of conotoxins // *J. Proteome Res.* American Chemical Society, 2012. Vol. 11, № 8. P. 4191–4200.
  35. Buczek O., Bulaj G., Olivera B.M. Conotoxins and the posttranslational modification of secreted gene products // *Cell. Mol. Life Sci. C.* 2005 6224. Springer, 2005. Vol. 62, № 24. P. 3067–3079.
  36. BIEMANN K. Mass Spectrometry // <http://dx.doi.org/10.1146/annurev.bi.32.070163.003543>. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 1963. Vol. 32. P. 755–780.
  37. Sakurai T., Matsuo T., Matsuda H., Katakuse I. PAAS 3: A computer program to determine probable sequence of peptides from mass spectrometric data // *Biol. Mass Spectrom.* 1984. Vol. 11, № 8. P. 396–399.
  38. Siegel M.M., Bauman N. An efficient algorithm for sequencing peptides using fast atom bombardment mass spectral data // *Biomed. Environ. Mass Spectrom.* 1988. Vol. 15, № 6. P. 333–343.

39. Zidarov D., Thibault P., Evans M.J., Bertrand M.J. Determination of the primary structure of peptides using fast atom bombardment mass spectrometry // *Biomed. Environ. Mass Spectrom.* John Wiley & Sons, Ltd, 1990. Vol. 19, № 1. P. 13–26.
40. Александров М.Л., Галль Л.Н., Краснов Н.В., Николаев В.И., А. Ш.. Экстракция ионов из растворов при атмосферном давлении – новый метод масс-спектрометрического анализа // *ДАН СССР.* 1984. Vol. 277, № 2. P. 279–383.
41. Yamashita M., Fenn J.B. Electrospray ion source. Another variation on the free-jet theme // *J. Phys. Chem.* 1984. Vol. 88, № 20. P. 4451–4459.
42. Karas M., Bachmann D., Hillenkamp F. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules // *Anal. Chem.* 1985. Vol. 57, № 14. P. 2935–2939.
43. Karas M., Bachmann D., Bahr U., Hillenkamp F. Matrix-assisted ultraviolet laser desorption of non-volatile compounds // *Int. J. Mass Spectrom. Ion Process.* Elsevier, 1987. Vol. 78, № C. P. 53–68.
44. Tanaka K., Waki H., Ido Y., Akita S., Yoshida Y., Yoshida T., Matsuo T. Protein and polymer analyses up to  $m/z$  100 000 by laser ionization time-of-flight mass spectrometry // *Rapid Commun. Mass Spectrom.* 1988. Vol. 2, № 8. P. 151–153.
45. Kofel P., Allemann M., Kellerhals H., Wanczek K.P. External generation of ions in ICR spectrometry // *Int. J. Mass Spectrom. Ion Process.* Elsevier, 1985. Vol. 65, № 1–2. P. 97–103.
46. Wagner I., Musso H. New Naturally Occurring Amino Acids // *Angew. Chemie Int. Ed. English.* 1983. Vol. 22, № 11. P. 816–828.
47. Xie J., Schultz P.G. Adding amino acids to the genetic repertoire // *Curr. Opin. Chem. Biol.* Elsevier Current Trends, 2005. Vol. 9, № 6. P. 548–554.
48. Edman P., Högföldt E., Sillén L.G., Kinell P.-O. Method for Determination of the Amino Acid Sequence in Peptides. // *Acta Chem. Scand.* 1950. Vol. 4. P. 283–293.
49. Bogdanov B., Smith R.D. Proteomics by FTICR mass spectrometry: Top down and bottom up // *Mass Spectrom. Rev.* John Wiley & Sons, Ltd, 2005. Vol. 24, № 2. P. 168–200.
50. Medzihradszky K.F., Chalkley R.J. Lessons in de novo peptide sequencing by tandem mass spectrometry // *Mass Spectrom. Rev.* John Wiley & Sons, Ltd, 2015. Vol. 34, № 1. P. 43–63.
51. Savitski M.M., Nielsen M.L., Zubarev R.A. New Data Base-independent, Sequence Tag-based Scoring of Peptide MS/MS Data Validates Mowse Scores, Recovers Below Threshold Data, Singles Out Modified Peptides, and Assesses the Quality of MS/MS Techniques // *Mol. Cell. Proteomics.* Elsevier, 2005. Vol. 4, № 8. P. 1180–1188.

52. Gorshkov V., Blenau W., Koeniger G., Römpf A., Vilcinskis A., Spengler B. Protein and Peptide Composition of Male Accessory Glands of *Apis mellifera* Drones Investigated by Mass Spectrometry // PLoS One. Public Library of Science, 2015. Vol. 10, № 5. P. e0125068.
53. Samgina T.Y., Vorontsov E.A., Gorshkov V.A., Artemenko K.A., Zubarev R.A., Lebedev A.T. Mass spectrometric de novo sequencing of natural non-tryptic peptides: Comparing peculiarities of collision-induced dissociation (CID) and high energy collision dissociation (HCD) // Rapid Commun. Mass Spectrom. John Wiley and Sons Ltd, 2014. Vol. 28, № 23. P. 2595–2604.
54. Abreu T.F., Sumitomo B.N., Nishiyama M.Y., Oliveira U.C., Souza G.H.M.F., Kitano E.S., Zelanis A., Serrano S.M.T., Junqueira-de-Azevedo I., Silva P.I., Tashima A.K. Peptidomics of *Acanthoscurria gomesiana* spider venom reveals new toxins with potential antimicrobial activity // J. Proteomics. Elsevier, 2017. Vol. 151. P. 232–242.
55. Ueberheide B.M., Fenyö D., Alewood P.F., Chait B.T. Rapid sensitive analysis of cysteine rich peptide venom components // Proc. Natl. Acad. Sci. U. S. A. 2009. Vol. 106, № 17. P. 6910–6915.
56. Bishop B.M., Juba M.L., Devine M.C., Barksdale S.M., Rodriguez C.A., Chung M.C., Russo P.S., Vliet K.A., Schnur J.M., van Hoek M.L. Bioprospecting the American Alligator (*Alligator mississippiensis*) Host Defense Peptidome // PLoS One / ed. Harder J. 2015. Vol. 10, № 2. P. e0117394.
57. Bishop B.M., Juba M.L., Russo P.S., Devine M., Barksdale S.M., Scott S., Settlage R., Michalak P., Gupta K., Vliet K., Schnur J.M., Van Hoek M.L. Discovery of Novel Antimicrobial Peptides from *Varanus komodoensis* (Komodo Dragon) by Large-Scale Analyses and De-Novo-Assisted Sequencing Using Electron-Transfer Dissociation Mass Spectrometry // J. Proteome Res. American Chemical Society, 2017. Vol. 16, № 4. P. 1470–1482.
58. L. Juba M., S. Russo P., Devine M., Barksdale S., Rodriguez C., A. Vliet K., M. Schnur J., L. van Hoek M., M. Bishop B. Large Scale Discovery and De Novo-Assisted Sequencing of Cationic Antimicrobial Peptides (CAMPs) by Microparticle Capture and Electron-Transfer Dissociation (ETD) Mass Spectrometry // J. Proteome Res. 2015. Vol. 14, № 10. P. 4282–4295.
59. Samgina T.Y., Artemenko K.A., Bergquist J., Trebse P., Torkar G., Tolpina M.D., Lebedev A.T. Differentiation of frogs from two populations belonging to the *Pelophylax esculentus* complex by LC-MS/MS comparison of their skin peptidomes // Anal. Bioanal. Chem. 2017. Vol. 409, № 7. P. 1951–1961.

60. Schmitt-Kopplin P., Kanawati B. *Fundamentals and Applications of Fourier Transform Mass Spectrometry*. Elsevier, 2019. P. 780.
61. Comisarow M.B., Marshall A.G. Fourier transform ion cyclotron resonance spectroscopy // *Chem. Phys. Lett.* 1974. Vol. 25, № 2. P. 282–283.
62. Makarov A.A. Mass Spectrometer. US Patent 5,886,346. 1999.
63. Makarov A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis // *Anal. Chem.* 2000. Vol. 72, № 6. P. 1156–1162.
64. Schmidt E.M., Pudenzi M.A., Santos J.M., Angolini C.F.F., Pereira R.C.L., Rocha Y.S., Denisov E., Damoc E., Makarov A., Eberlin M.N. Petroleomics via Orbitrap mass spectrometry with resolving power above 1 000 000 at  $m/z$  200 // *RSC Adv. The Royal Society of Chemistry*, 2018. Vol. 8, № 11. P. 6183–6191.
65. Michalski A., Damoc E., Lange O., Denisov E., Nolting D., Müller M., Viner R., Schwartz J., Remes P., Belford M., Dunyach J.J., Cox J., Horning S., Mann M., Makarov A. Ultra High Resolution Linear Ion Trap Orbitrap Mass Spectrometer (Orbitrap Elite) Facilitates Top Down LC MS/MS and Versatile Peptide Fragmentation Modes // *Mol. Cell. Proteomics*. Elsevier, 2012. Vol. 11, № 3. P. O111.013698.
66. Zubarev R.A., Makarov A. Orbitrap Mass Spectrometry // *Anal. Chem. American Chemical Society*, 2013. Vol. 85, № 11. P. 5288–5296.
67. He F., L. Hendrickson C., G. Marshall A. Baseline Mass Resolution of Peptide Isobars: A Record for Molecular Mass Resolution // *Anal. Chem.* 2000. Vol. 73, № 3. P. 647–650.
68. Bossio R.E., Marshall A.G. Baseline Resolution of Isobaric Phosphorylated and Sulfated Peptides and Nucleotides by Electrospray Ionization FTICR MS: Another Step toward Mass Spectrometry-Based Proteomics // *Anal. Chem. American Chemical Society*, 2002. Vol. 74, № 7. P. 1674–1679.
69. Shaw J.B., Gorshkov M. V., Wu Q., Paša-Tolić L. High Speed Intact Protein Characterization Using 4X Frequency Multiplication, Ion Trap Harmonization, and 21 Tesla FTICR-MS // *Anal. Chem. American Chemical Society*, 2018. Vol. 90, № 9. P. 5557–5562.
70. N. Nikolaev E., N. Vladimirov G., Jertz R., Baykut G. From Supercomputer Modeling to Highest Mass Resolution in FT-ICR // *Mass Spectrom.* 2013. Vol. 2, № Special\_Issue. P. S0010–S0010.
71. Samgina T.Y., Artemenko K.A., Gorshkov V.A., Ogourtsov S. V., Zubarev R.A., Lebedev A.T. De novo sequencing of peptides secreted by the skin glands of the Caucasian Green Frog *Rana ridibunda* // *Rapid Commun. Mass Spectrom.* 2008. Vol. 22, № 22. P. 3517–3525.

72. Ye H., Wang J., Zhang Z., Jia C., Schmerberg C., D. Catherman A., M. Thomas P., L. Kelleher N., Li L. Defining the Neuropeptidome of the Spiny Lobster *Panulirus interruptus* Brain Using a Multidimensional Mass Spectrometry-Based Platform // *J. Proteome Res.* 2015. Vol. 14, № 11. P. 4776–4791.
73. Shaw J.B., Lin T.Y., Leach F.E., Tolmachev A. V., Tolić N., Robinson E.W., Koppelaar D.W., Paša-Tolić L. 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer Greatly Expands Mass Spectrometry Toolbox // *J. Am. Soc. Mass Spectrom.* Springer New York LLC, 2016. Vol. 27, № 12. P. 1929–1936.
74. Spengler B. De novo sequencing, peptide composition analysis, and composition-based sequencing: a new strategy employing accurate mass determination by fourier transform ion cyclotron resonance mass spectrometry // *J. Am. Soc. Mass Spectrom.* No longer published by Elsevier, 2004. Vol. 15, № 5. P. 703–714.
75. Wehofsky M., Hoffmann R., Hubert M., Spengler B. Isotopic Deconvolution of Matrix-Assisted Laser Desorption/Ionization Mass Spectra for Substance-Class Specific Analysis of Complex Samples // *Eur. J. Mass Spectrom.* 2001. Vol. 7, № 1. P. 39–46.
76. Spengler B. Accurate Mass as a Bioinformatic Parameter in Data-to-Knowledge Conversion: Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Peptide De Novo Sequencing // *Eur. J. Mass Spectrom.* 2007. Vol. 13, № 1. P. 83–87.
77. Schlosser A., Lehmann W.D. Patchwork peptide sequencing: Extraction of sequence information from accurate mass data of peptide tandem mass spectra recorded at high resolution // *Proteomics.* 2002. Vol. 2, № 5. P. 524–533.
78. Artemenko K.A., Zubarev A.R., Samgina T.Y., Lebedev A.T., Savitski M.M., Zubarev R.A. Two dimensional mass mapping as a general method of data representation in comprehensive analysis of complex molecular mixtures // *Anal. Chem.* 2009. Vol. 81, № 10. P. 3738–3745.
79. Samgina T.Y., Gorshkov V.A., Artemenko K.A., Vorontsov E.A., Klykov O. V., Ogourtsov S. V., Zubarev R.A., Lebedev A.T. LC–MS/MS with 2D mass mapping of skin secretions' peptides as a reliable tool for interspecies identification inside *Rana esculenta* complex // *Peptides.* 2012. Vol. 34, № 2. P. 296–302.
80. Brodbelt J.S. Ion Activation Methods for Peptides and Proteins // *Anal. Chem.* 2016. Vol. 88, № 1. P. 30–51.
81. Hurtado P.P., O'Connor P.B. Differentiation of isomeric amino acid residues in proteins and peptides using mass spectrometry // *Mass Spectrom. Rev.* 2012. Vol. 31, № 6. P. 609–625.
82. Marshall A.G., Wang T.C.L., Ricca T.L. Tailored excitation for Fourier transform ion

- cyclotron mass spectrometry // *J. Am. Chem. Soc.* 1985. Vol. 107, № 26. P. 7893–7897.
83. Laskin J., Futrell J.H. On the efficiency of energy transfer in collisional activation of small peptides // *J. Chem. Phys. American Institute of Physics AIP*, 2002. Vol. 116, № 10. P. 4302.
84. Gauthier J.W., Trautman T.R., Jacobson D.B. Sustained off-resonance irradiation for collision-activated dissociation involving Fourier transform mass spectrometry. Collision-activated dissociation technique that emulates infrared multiphoton dissociation // *Anal. Chim. Acta. Elsevier*, 1991. Vol. 246, № 1. P. 211–225.
85. Ma M., Kutz-Naber K.K., Li L. Methyl esterification assisted MALDI FTMS characterization of the orcokinin neuropeptide family // *Anal. Chem.* 2007. Vol. 79, № 2. P. 673–681.
86. Göçmen B., Heiss P., Petras D., Nalbantsoy A., Süßmuth R.D. Mass spectrometry guided venom profiling and bioactivity screening of the Anatolian Meadow Viper, *Vipera anatolica* // *Toxicon. Pergamon*, 2015. Vol. 107. P. 163–174.
87. Langsdorf M., Ghassempour A., Römpf A., Spengler B. Characterization of a peptide family from the skin secretion of the middle east tree frog *Hyla savignyi* by composition-based de novo sequencing // *Rapid Commun. Mass Spectrom.* 2010. Vol. 24, № 19. P. 2885–2899.
88. Langsdorf M., Ghassempour A., Römpf A., Spengler B. Isolation and sequence analysis of peptides from the skin secretion of the Middle East tree frog *Hyla savignyi* // *Anal. Bioanal. Chem. Springer*, 2010. Vol. 398, № 7–8. P. 2853–2865.
89. Olsen J. V, Macek B., Lange O., Makarov A., Horning S., Mann M. Higher-energy C-trap dissociation for peptide modification analysis // *Nat. Methods.* 2007. Vol. 4, № 9. P. 709–712.
90. Siano A., Verónica Húmpola M., de Oliveira E., Albericio F., C. Simonetta A., Lajmanovich R., G. Tonarelli G. Antimicrobial Peptides from Skin Secretions of *Hypsiboas pulchellus* (Anura: Hylidae) // *J. Nat. Prod.* 2014. Vol. 77, № 4. P. 831–841.
91. Kodama R.T., Cajado-Carvalho D., Kuniyoshi A.K., Kitano E.S., Tashima A.K., Barna B.F., Takakura A.C., Serrano S.M.T., Dias-Da-Silva W., Tambourgi D. V., Portaro F. V. New proline-rich oligopeptides from the venom of African adders: Insights into the hypotensive effect of the venoms // *Biochim. Biophys. Acta - Gen. Subj. Elsevier*, 2015. Vol. 1850, № 6. P. 1180–1187.
92. Laskin J., V. Denisov E., K. Shukla A., E. Barlow S., H. Futrell J. Surface-Induced Dissociation in a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer: Instrument Design and Evaluation // *Anal. Chem.* 2002. Vol. 74, № 14. P. 3255–3261.
93. Laskin J., Futrell J.H. Collisional activation of peptide ions in FT-ICR mass spectrometry

- // *Mass Spectrom. Rev.* 2003. Vol. 22, № 3. P. 158–181.
94. Zubarev R.A., Kelleher N.L., McLafferty F.W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process // *J. Am. Chem. Soc.* 1998. Vol. 120, № 13. P. 3265–3266.
  95. Zubarev R.A. Reactions of polypeptide ions with electrons in the gas phase // *Mass Spectrom. Rev.* 2003. Vol. 22, № 1. P. 57–77.
  96. Bakken V., Helgaker T., Uggerud E. Models of fragmentations induced by electron attachment to protonated peptides // *Eur. J. Mass Spectrom.* SAGE Publications Sage UK: London, England, 2004. Vol. 10, № 5. P. 625–638.
  97. Zubarev R., Horn D., Fridriksson E., Kelleher N., Kruger N., Lewis M., Carpenter B., McLafferty F. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations // *Anal. Chem.* American Chemical Society, 2000. Vol. 72, № 3. P. 563–573.
  98. Cooper H.J., Håkansson K., Marshall A.G. The role of electron capture dissociation in biomolecular analysis // *Mass Spectrom. Rev.* John Wiley & Sons, Ltd, 2005. Vol. 24, № 2. P. 201–222.
  99. Budnik B.A., Olsen J. V., Egorov T.A., Anisimova V.E., Galkina T.G., Musolyamov A.K., Grishin E. V., Zubarev R.A. De novo sequencing of antimicrobial peptides isolated from the venom glands of the wolf spider *Lycosa singoriensis* // *J. Mass Spectrom.* John Wiley & Sons, Ltd, 2004. Vol. 39, № 2. P. 193–201.
  100. Gucinski A.C., Boyne M.T. Identification of site-specific heterogeneity in peptide drugs using intact mass spectrometry with electron transfer dissociation // *Rapid Commun. Mass Spectrom.* John Wiley and Sons Ltd, 2014. Vol. 28, № 15. P. 1757–1763.
  101. Zhurov K.O., Fornelli L., Wodrich M.D., Laskay Ü.A., Tsybin Y.O. Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis // *Chem. Soc. Rev.* The Royal Society of Chemistry, 2013. Vol. 42, № 12. P. 5014–5030.
  102. Riley N.M., Coon J.J. The Role of Electron Transfer Dissociation in Modern Proteomics // *Analytical Chemistry.* American Chemical Society, 2018. Vol. 90, № 1. P. 40–64.
  103. Savitski M., Nielsen M., Kjeldsen F., Zubarev R. Proteomics-Grade de Novo Sequencing Approach // *J. Proteome Res.* 2005. Vol. 4, № 6. P. 2348–2354.
  104. Sasaki K., Osaki T., Minamino N. Large-scale Identification of Endogenous Secretory Peptides Using Electron Transfer Dissociation Mass Spectrometry // *Mol. Cell. Proteomics.* Elsevier, 2013. Vol. 12, № 3. P. 700–709.
  105. Hart S.R., Kenny L.C., Myers J.E., Baker P.N. Electron transfer dissociation of native

- peptides facilitates enhanced identification of urinary peptides // *Int. J. Mass Spectrom. Elsevier*, 2015. Vol. 391. P. 41–46.
106. Jhingree J.R., Beveridge R., Dickinson E.R., Williams J.P., Brown J.M., Bellina B., Barran P.E. Electron transfer with no dissociation ion mobility–mass spectrometry (ETnoD IM-MS). The effect of charge reduction on protein conformation // *Int. J. Mass Spectrom. Elsevier*, 2017. Vol. 413. P. 43–51.
  107. Swaney D.L., McAlister G.C., Wirtala M., Schwartz J.C., Syka J.E.P., Coon J.J. Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors // *Anal. Chem.* 2007. Vol. 79, № 2. P. 477–485.
  108. Frese C.K., Altelaar A.F.M., Van Den Toorn H., Nolting D., Griep-Raming J., Heck A.J.R., Mohammed S. Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry // *Anal. Chem. American Chemical Society*, 2012. Vol. 84, № 22. P. 9668–9673.
  109. Mommen G.P.M., Frese C.K., Meiring H.D., Gaans-van Den Brink J., De Jong A.P.J.M., Van Els C.A.C.M., Heck A.J.R. Expanding the detectable HLA peptide repertoire using electron-transfer/ higher-energy collision dissociation (EThcD) // *Proc. Natl. Acad. Sci. U. S. A. National Academy of Sciences*, 2014. Vol. 111, № 12. P. 4507–4512.
  110. Marino F., Bern M., P. M. Mommen G., C. Leney A., A. M. van Gaans-van den Brink J., M. J. J. Bonvin A., Becker C., A. C. M. van Els C., J. R. Heck A. Extended O-GlcNAc on HLA Class-I-Bound Peptides // *J. Am. Chem. Soc. American Chemical Society*, 2015. Vol. 137, № 34. P. 10922–10925.
  111. Lebedev A.T., Damoc E., Makarov A.A., Samgina T.Y. Discrimination of Leucine and Isoleucine in Peptides Sequencing with Orbitrap Fusion Mass Spectrometer // *Anal. Chem.* 2014. Vol. 86, № 14. P. 7017–7022.
  112. Gardner M.W., Smith S.I., Ledvina A.R., Madsen J.A., Coon J.J., Schwartz J.C., Stafford G.C., Brodbelt J.S. Infrared multiphoton dissociation of peptide cations in a dual pressure linear ion trap mass spectrometer // *Anal. Chem. American Chemical Society*, 2009. Vol. 81, № 19. P. 8109–8118.
  113. Payne A.H., Glish G.L. Thermally assisted infrared multiphoton photodissociation in a quadrupole ion trap. // *Anal. Chem.* 2001. Vol. 73, № 15. P. 3542–3548.
  114. Sleno L., Volmer D.A. Ion activation methods for tandem mass spectrometry // *J. Mass Spectrom.* 2004. Vol. 39, № 10. P. 1091–1112.
  115. Little D.P., Speir J.P., Senko M.W., O'Connor P.B., McLafferty F.W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing // *Anal. Chem.* 1994. Vol. 66, № 18. P. 2809–2815.

116. Håkansson K., Chalmers M.J., Quinn J.P., McFarland M.A., Hendrickson C.L., Marshall A.G. Combined Electron Capture and Infrared Multiphoton Dissociation for Multistage MS/MS in a Fourier Transform Ion Cyclotron Resonance Mass Spectrometer // *Anal. Chem. American Chemical Society*, 2003. Vol. 75, № 13. P. 3256–3262.
117. D. Price W., D. Schnier P., R. Williams E. Tandem Mass Spectrometry of Large Biomolecule Ions by Blackbody Infrared Radiative Dissociation // *Anal. Chem. Mautner*, 1996. Vol. 68, № 5. P. 859–866.
118. Ewing N.P., Cassady C.J. Dissociation of multiply charged negative ions for hirudin (54–65), fibrinopeptide B, and insulin A (oxidized) // *J. Am. Soc. Mass Spectrom.* No longer published by Elsevier, 2001. Vol. 12, № 1. P. 105–116.
119. Emmett M.R., White F.M., Hendrickson C.L., Shi D.H., Marshall A.G. Application of micro-electrospray liquid chromatography techniques to FT-ICR MS to enable high-sensitivity biological analysis // *J. Am. Soc. Mass Spectrom. Springer-Verlag New York*, 1998. Vol. 9, № 4. P. 333–340.
120. Bowie J.H., Brinkworth C.S., Dua S. Collision-induced fragmentations of the (M-H)- parent anions of underivatized peptides: An aid to structure determination and some unusual negative ion cleavages // *Mass Spectrom. Rev.* 2002. Vol. 21, № 2. P. 87–107.
121. Bilusich D., Bowie J.H. Fragmentations of (M-H) – anions of underivatized peptides. Part 2: Characteristic cleavages of Ser and Cys and of disulfides and other post-translational modifications, together with some unusual internal processes // *Mass Spectrom. Rev.* 2009. Vol. 28, № 1. P. 20–34.
122. Wang T., Nha Tran T.T., Andrezza H.J., Bilusich D., Brinkworth C.S., Bowie J.H. Negative ion cleavages of (M-H)– anions of peptides. Part 3. Post-translational modifications // *Mass Spectrom. Rev. John Wiley and Sons Inc.*, 2018. Vol. 37, № 1. P. 3–21.
123. Huzarska M., Ugalde I., Kaplan D.A., Hartmer R., Easterling M.L., Polfer N.C. Negative electron transfer dissociation of deprotonated phosphopeptide anions: choice of radical cation reagent and competition between electron and proton transfer // *Anal. Chem.* 2010. Vol. 82, № 7. P. 2873–2878.
124. Budnik B.A., Haselmann K.F., Zubarev R.A. Electron detachment dissociation of peptide di-anions: an electron–hole recombination phenomenon // *Chem. Phys. Lett. North-Holland*, 2001. Vol. 342, № 3–4. P. 299–302.
125. Kalli A., Grigorean G., Håkansson K. Electron Induced Dissociation of Singly Deprotonated Peptides // *J. Am. Soc. Mass Spectrom.* 2011. Vol. 22, № 12. P. 2209–2221.
126. Ju Yoo H., Wang N., Zhuang S., Song H., Håkansson K. Negative-Ion Electron Capture

- Dissociation: Radical-Driven Fragmentation of Charge-Increased Gaseous Peptide Anions // *J. Am. Chem. Soc.* 2011. Vol. 133, № 42. P. 16790–16793.
127. Hersberger K.E., Håkansson K. Characterization of O-sulfopeptides by negative ion mode tandem mass spectrometry: Superior performance of negative ion electron capture dissociation // *Anal. Chem.* 2012. Vol. 84, № 15. P. 6370–6377.
128. Cournoyer J.J., Pittman J.L., Ivleva V.B., Fallows E., Waskell L., Costello C.E., O'Connor P.B. Deamidation: Differentiation of aspartyl from isoaspartyl products in peptides by electron capture dissociation // *Protein Sci.* John Wiley & Sons, Ltd, 2005. Vol. 14, № 2. P. 452–463.
129. B. O'Connor P., J. Cournoyer J., J. Pitteri S., A. Chrisman P., A. McLuckey S. Differentiation of aspartic and isoaspartic acids using electron transfer dissociation // *J. Am. Soc. Mass Spectrom.* 2006. Vol. 17, № 1. P. 15–19.
130. P. Sargaeva N., Lin C., B. O'Connor P. Identification of Aspartic and Isoaspartic Acid Residues in Amyloid  $\beta$  Peptides, Including A $\beta$ 1–42, Using Electron–Ion Reactions // *Anal. Chem.* 2009. Vol. 81, № 23. P. 9778–9786.
131. Li X., Lin C., B. O'Connor P. Glutamine Deamidation: Differentiation of Glutamic Acid and  $\gamma$ -Glutamic Acid in Peptides by Electron Capture Dissociation // *Anal. Chem.* 2010. Vol. 82, № 9. P. 3606–3615.
132. Adams C.M., Kjeldsen F., Zubarev R.A., Budnik B.A., Haselmann K.F. Electron capture dissociation distinguishes a single D-amino acid in a protein and probes the tertiary structure // *J. Am. Soc. Mass Spectrom.* Springer-Verlag New York, 2004. Vol. 15, № 7. P. 1087–1098.
133. Adams C.M., Zubarev R.A. Distinguishing and Quantifying Peptides and Proteins Containing d-Amino Acids by Tandem Mass Spectrometry // *Anal. Chem.* American Chemical Society, 2005. Vol. 77, № 14. P. 4571–4580.
134. Samgina T.Y., Artemenko K.A., Gorshkov V.A., Lebedev A.T., Nielsen M.L., Savitski M.M., Zubarev R.A. Electrospray Ionization Tandem Mass Spectrometry Sequencing of Novel Skin Peptides from Ranid Frogs Containing Disulfide Bridges // *Eur. J. Mass Spectrom.* 2007. Vol. 13, № 2. P. 155–163.
135. Samgina T.Y., Artemenko K.A., Gorshkov V.A., Ogourtsov S. V., Zubarev R.A., Lebedev A.T. Mass spectrometric study of peptides secreted by the skin glands of the brown frog *Rana arvalis* from the Moscow region // *Rapid Commun. Mass Spectrom.* John Wiley & Sons, Ltd, 2009. Vol. 23, № 9. P. 1241–1248.
136. Chalkley R.J., Brinkworth C.S., Burlingame A.L. Side-Chain Fragmentation of Alkylated Cysteine Residues in Electron Capture Dissociation Mass Spectrometry // *J. Am. Soc. Mass*

- Spectrom. No longer published by Elsevier, 2006. Vol. 17, № 9. P. 1271–1274.
137. Kjeldsen F., Haselmann K.F., Budnik B.A., Jensen F., Zubarev R.A. Dissociative capture of hot (3–13 eV) electrons by polypeptide polycations: an efficient process accompanied by secondary fragmentation // *Chem. Phys. Lett.* 2002. Vol. 356, № 3–4. P. 201–206.
  138. Kjeldsen F., Haselmann K.F., Sørensen E.S., Zubarev R.A. Distinguishing of Ile/Leu Amino Acid Residues in the PP3 Protein by (Hot) Electron Capture Dissociation in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry // *Anal. Chem.* 2003. Vol. 75, № 6. P. 1267–1274.
  139. Williams J.P., Creese A.J., Roper D.R., Green B.N., Cooper H.J. Hot electron capture dissociation distinguishes leucine from isoleucine in a novel hemoglobin variant, Hb Askew,  $\beta$  54(D5)Val→Ile // *J. Am. Soc. Mass Spectrom.* 2009. Vol. 20, № 9. P. 1707–1713.
  140. Zhokhov S.S., Kovalyov S. V., Samgina T.Y., Lebedev A.T. An EThcD-Based Method for Discrimination of Leucine and Isoleucine Residues in Tryptic Peptides // *J. Am. Soc. Mass Spectrom.* Springer New York LLC, 2017. Vol. 28, № 8. P. 1600–1611.
  141. Kovalyov S. V, Zhokhov S.S., Onoprienko L. V, Vaskovsky B. V, Lebedev A.T. Exploration of doubtful cases of leucine and isoleucine discrimination in mass spectrometric peptide sequencing by electron-transfer and higher-energy collision dissociation-based method // *Eur. J. Mass Spectrom.* 2017. Vol. 23, № 6. P. 376–384.
  142. Samgina T.Y., Kovalev S. V., Tolpina M.D., Trebse P., Torkar G., Lebedev A.T. EThcD Discrimination of Isomeric Leucine/Isoleucine Residues in Sequencing of the Intact Skin Frog Peptides with Intramolecular Disulfide Bond // *J. Am. Soc. Mass Spectrom.* 2018. Vol. 29, № 5. P. 842–852.
  143. Leymarie N., Costello C.E., O'Connor P.B. Electron Capture Dissociation Initiates a Free Radical Reaction Cascade // *J. Am. Chem. Soc.* 2003. Vol. 125, № 29. P. 8949–8958.
  144. Li X., Lin C., Han L., Costello C.E., O'Connor P.B. Charge remote fragmentation in electron capture and electron transfer dissociation // *J. Am. Soc. Mass Spectrom.* 2010. Vol. 21, № 4. P. 646–656.
  145. O'Connor P.B., Lin C., Cournoyer J.J., Pittman J.L., Belyayev M., Budnik B.A. Long-lived electron capture dissociation product ions experience radical migration via hydrogen abstraction // *J. Am. Soc. Mass Spectrom.* 2006. Vol. 17, № 4. P. 576–585.
  146. Miyashita M., Kitanaka A., Yakio M., Yamazaki Y., Nakagawa Y., Miyagawa H. Complete de novo sequencing of antimicrobial peptides in the venom of the scorpion *Isometrus maculatus* // *Toxicon*. Pergamon, 2017. Vol. 139. P. 1–12.
  147. Chen T., Orr D.F., O'Rourke M., McLynn C., Bjourson A.J., McClean S., Hirst D., Rao P.,

- Shaw C. Pachymedusa dancicolor tryptophyllin-1: structural characterization, pharmacological activity and cloning of precursor cDNA // Regul. Pept. Elsevier, 2004. Vol. 117, № 1. P. 25–32.
148. Samgina T.Y., Kovalev S. V., Gorshkov V.A., Artemenko K.A., Poljakov N.B., Lebedev A.T. N-terminal tagging strategy for De Novo sequencing of short peptides by ESI-MS/MS and MALDI-MS/MS // J. Am. Soc. Mass Spectrom. 2010. Vol. 21, № 1. P. 104–111.
149. Zaikin V.G., Halket J.M. Soft ionization mass spectrometry of large molecules // A Handbook of Derivatives for Mass Spectrometry. Charlton, UK: IM Publications LLP, 2009. P. 299–478.
150. Samgina T.Y., Gorshkov V.A., Artemenko K.A., Kovalev S.V., Ogourtsov S.V., Zubarev R.A., Lebedev A.T. Novel natural peptides from Hyla arborea schelkownikowi skin secretion // Rapid Commun. Mass Spectrom. 2010. Vol. 24, № 12. P. 1749–1754.
151. Samgina T.Y., Gorshkov V.A., Vorontsov Y.A., Artemenko K.A., Zubarev R.A., Lebedev A.T. Mass spectrometric study of bradykinin-related peptides (BRPs) from the skin secretion of Russian ranid frogs // Rapid Commun. Mass Spectrom. 2011. Vol. 25, № 7. P. 933–940.
152. Lebedev A.T. Environmental Mass Spectrometry // <http://dx.doi.org/10.1146/annurev-anchem-062012-092604>. Annual Reviews , 2013. Vol. 6. P. 163–189.
153. Chen T., Orr D.F., Bjourson A.J., McClean S., O'Rourke M., Hirst D.G., Rao P., Shaw C. Novel bradykinins and their precursor cDNAs from European yellow-bellied toad (*Bombina variegata*) skin // Eur. J. Biochem. John Wiley & Sons, Ltd, 2002. Vol. 269, № 18. P. 4693–4700.
154. Schroeder C., Beug H., Müller-Esterl W. Cloning and functional characterization of the ornithokinin receptor - Recognition of the major kinin receptor antagonist, HOE140, as a full agonist // J. Biol. Chem. Elsevier, 1997. Vol. 272, № 19. P. 12475–12481.
155. Thompson A.H., Bjourson A.J., Shaw C., McClean S. Bradykinin-related peptides from *Phyllomedusa hypochondrialis azurea*: Mass spectrometric structural characterisation and cloning of precursor cDNAs // Rapid Commun. Mass Spectrom. 2006. Vol. 20, № 24. P. 3780–3788.
156. Siemerink M., Schebb N.H., Liesener A., Perchuc A.M., Schöni R., Wilmer M., Hayen H., Karst U., Vogel M. Development of a fast liquid chromatography/mass spectrometry screening method for angiotensin-converting enzyme (ACE) inhibitors in complex natural mixtures like snake venom // Rapid Commun. Mass Spectrom. 2010. Vol. 24, № 5. P. 687–697.
157. Chen R., Jiang X., Conaway M.C.P., Mohtashemi I., Hui L., Viner R., Li L. Mass spectral

- analysis of neuropeptide expression and distribution in the nervous system of the lobster *Homarus americanus* // *J. Proteome Res.* American Chemical Society, 2010. Vol. 9, № 2. P. 818–832.
158. Allen G. Sequencing of proteins and peptides. North-Holland, 1981. P. 327.
  159. Bao N., Lecaer J.-P., Nghia N.D., Vinh P.T.K. Isolation and structural identification of a new T1-conotoxin with unique disulfide connectivities derived from *Conus bandanus* // *J. Venom. Anim. Toxins Incl. Trop. Dis.* 2020. Vol. 26.
  160. Simon M.D., Chu F., Racki L.R., de la Cruz C.C., Burlingame A.L., Panning B., Narlikar G.J., Shokat K.M. The Site-Specific Installation of Methyl-Lysine Analogs into Recombinant Histones // *Cell.* Cell Press, 2007. Vol. 128, № 5. P. 1003–1012.
  161. Artemenko K.A., Samgina T.Y., Lebedev A.T. De novo sequencing of peptides by mass spectrometry // *Mass-spektrometrija (Rus)*. 2006. Vol. 3, № 4. P. 225–254.
  162. Zaikin V., Halket J. *A Handbook of Derivatives for Mass Spectrometry*. Charlton, UK: IM Publications LLP, 2009. 543 p.
  163. Burlet O., Yang C.-Y., Gaskell S.J. Influence of cysteine to cysteic acid oxidation on the collision-activated decomposition of protonated peptides: Evidence for intraionic interactions // *J. Am. Soc. Mass Spectrom.* 1992. Vol. 3, № 4. P. 337–344.
  164. Samgina T.Y., Artemenko K.A., Gorshkov V.A., Poljakov N.B., Lebedev A.T. Oxidation versus carboxamidomethylation of s-s bond in ranid frog peptides: Pro and contra for de novo MALDI-MS sequencing // *J. Am. Soc. Mass Spectrom.* 2008. Vol. 19, № 4. P. 479–487.
  165. Samgina T.Y., Gorshkov V.A., Vorontsov Y.A., Artemenko K.A., Ogourtsov S. V., Zubarev R.A., Lebedev A.T. Investigation of skin secretory peptidome of *Rana lessonae* frog by mass spectrometry // *J. Anal. Chem.* 2011. Vol. 66, № 13. P. 1298–1306.
  166. Samgina T.Y., Gorshkov V.A., Vorontsov Y.A., Artemenko K.A., Ogourtsov S. V., Zubarev R.A., Lebedev A.T. Mass spectral study of the skin peptide of brown frog *Rana temporaria* from Zvenigorod population // *J. Anal. Chem.* 2011. Vol. 66, № 14. P. 1353–1360.
  167. Samgina T.Y., Vorontsov E.A., Gorshkov V.A., Artemenko K.A., Nifant'ev I.E., Kanawati B., Schmitt-Kopplin P., Zubarev R.A., Lebedev A.T. Novel Cysteine Tags for the Sequencing of Non-Tryptic Disulfide Peptides of Anurans: ESI-MS Study of Fragmentation Efficiency // *J. Am. Soc. Mass Spectrom.* 2011. Vol. 22, № 12. P. 2246–2255.
  168. Thakur S.S., Balaram P. Rapid mass spectral identification of contryphans. Detection of characteristic peptide ions by fragmentation of intact disulfide-bonded peptides in crude venom // *Rapid Commun. Mass Spectrom.* 2007. Vol. 21, № 21. P. 3420–3426.

169. Samgina T.Y., Vorontsov E.A., Gorshkov V.A., Artemenko K.A., Zubarev R.A., Ytterberg J.A., Lebedev A.T. Collision-Induced Dissociation Fragmentation Inside Disulfide C-Terminal Loops of Natural Non-Tryptic Peptides // *J. Am. Soc. Mass Spectrom.* 2013. Vol. 24, № 7. P. 1037–1044.
170. Wells J.M., Stephenson J.L., McLuckey S.A. Charge dependence of protonated insulin decompositions // *Int. J. Mass Spectrom. Elsevier*, 2000. Vol. 203, № 1–3. P. A1–A9.
171. Thorne G.C., Gaskell S.J., Gross M.L. Elucidation of some fragmentations of small peptides using sequential mass spectrometry on a hybrid instrument // *Rapid Commun. Mass Spectrom.* 1989. Vol. 3, № 7. P. 217–221.
172. Ma M., Chen R., Ge Y., He H., G. Marshall A., Li L. Combining Bottom-Up and Top-Down Mass Spectrometric Strategies for De Novo Sequencing of the Crustacean Hyperglycemic Hormone from *Cancer borealis* // *Anal. Chem.* 2008. Vol. 81, № 1. P. 240–247.
173. Ma M., Bors E.K., Dickinson E.S., Kwiatkowski M.A., Sousa G.L., Henry R.P., Smith C.M., Towle D.W., Christie A.E., Li L. Characterization of the *Carcinus maenas* neuropeptidome by mass spectrometry and functional genomics // *Gen. Comp. Endocrinol. Academic Press*, 2009. Vol. 161, № 3. P. 320–334.
174. Ma M., Wang J., Chen R., Li L. Expanding the Crustacean Neuropeptidome Using a Multifaceted Mass Spectrometric Approach // *J. Proteome Res.* 2009. Vol. 8, № 5. P. 2426–2437.
175. Ma M., Szabo T.M., Jia C., Marder E., Li L. Mass spectrometric characterization and physiological actions of novel crustacean C-type allatostatins // *Peptides. Elsevier*, 2009. Vol. 30, № 9. P. 1660–1668.
176. Jia C., Hui L., Cao W., Lietz C.B., Jiang X., Chen R., Catherman A.D., Thomas P.M., Ge Y., Kelleher N.L., Li N. High-definition De Novo Sequencing of Crustacean Hyperglycemic Hormone (CHH)-family Neuropeptides // *Mol. Cell. Proteomics. Elsevier*, 2012. Vol. 11, № 12. P. 1951–1964.
177. Jia C., Lietz C.B., Ye H., Hui L., Yu Q., Yoo S., Li L. A multi-scale strategy for discovery of novel endogenous neuropeptides in the crustacean nervous system // *J. Proteomics. Elsevier*, 2013. Vol. 91. P. 1–12.
178. Yu X., Khani A., Ye X., Petruzzello F., Gao H., Zhang X., Rainer G. High-Efficiency Recognition and Identification of Disulfide Bonded Peptides in Rat Neuropeptidome Using Targeted Electron Transfer Dissociation Tandem Mass Spectrometry // *Anal. Chem. American Chemical Society*, 2015. Vol. 87, № 23. P. 11646–11651.
179. Li W., Petruzzello F., Zhao N., Zhao H., Ye X., Zhang X., Rainer G. Separation and

- identification of mouse brain tissue microproteins using top-down method with high resolution nanocapillary liquid chromatography mass spectrometry // *Proteomics*. Wiley-VCH Verlag, 2017. Vol. 17, № 12.
180. Quinton L., Servent D., Girard E., Molgó J., Le Caer J.P., Malosse C., Haidar E.A., Lecoq A., Gilles N., Chamot-Rooke J. Identification and functional characterization of a novel  $\alpha$ -conotoxin (EIIA) from *Conus ermineus* // *Anal. Bioanal. Chem.* Springer, 2013. Vol. 405, № 15. P. 5341–5351.
181. Quinton L., Le Caer J.-P., Phan G., Ligny-Lemaire C., Bourdais- Jomaro J., Ducancel F., Chamot-Rooke J. Characterization of Toxins within Crude Venoms by Combined Use of Fourier Transform Mass Spectrometry and Cloning // *Anal. Chem.* 2005. Vol. 77, № 20. P. 6630–6639.
182. Nair S.S., Nilsson C.L., Emmett M.R., Schaub T.M., Gowd K.H., Thakur S.S., Krishnan K.S., Balaram P., Marshall A.G. De novo sequencing and disulfide mapping of a bromotryptophan-containing conotoxin by fourier transform ion cyclotron resonance mass spectrometry // *Anal. Chem.* 2006. Vol. 78, № 23. P. 8082–8088.
183. Quinton L., Caer J.P. Le, Vinh J., Gilles N., Chamot-Rooke J. Fourier transform mass spectrometry: A powerful tool for toxin analysis // *Toxicon*. Pergamon, 2006. Vol. 47, № 6. P. 715–726.
184. Quinton L. et al. Isolation and pharmacological characterization of AdTx1, a natural peptide displaying specific insurmountable antagonism of the  $\alpha$  1A- adrenoceptor // *Br. J. Pharmacol.* John Wiley & Sons, Ltd, 2010. Vol. 159, № 2. P. 316–325.
185. Figueroa-Montiel A., Bernáldez J., Jiménez S., Ueberhide B., González L.J., Licea-Navarro A. Antimycobacterial Activity: A New Pharmacological Target for Conotoxins Found in the First Reported Conotoxin from *Conasprella ximenes* // *Toxins* 2018, Vol. 10, Page 51. Multidisciplinary Digital Publishing Institute, 2018. Vol. 10, № 2. P. 51.
186. Zhang H., Fu Y., Wang L., Liang A., Chen S., Xu A. Identifying novel conopeptides from the venom ducts of *Conus litteratus* through integrating transcriptomics and proteomics // *J. Proteomics*. Elsevier, 2019. Vol. 192. P. 346–357.
187. Anand P., Grigoryan A., Bhuiyan M.H., Ueberhide B., Russell V., Quinoñez J., Moy P., Chait B.T., Poget S.F., Holford M. Sample Limited Characterization of a Novel Disulfide-Rich Venom Peptide Toxin from Terebrid Marine Snail *Terebra variegata* // *PLoS One*. Public Library of Science, 2014. Vol. 9, № 4. P. e94122.
188. Kumar Mandal A., Ramakrishnan Santhana Ramasamy M., Sabareesh V., E. Openshaw M., S. Krishnan K., Balaram P. Sequencing of T-superfamily conotoxins from *Conus virgo*: Pyroglutamic acid identification and disulfide arrangement by MALDI mass spectrometry

- // *J. Am. Soc. Mass Spectrom.* 2007. Vol. 18, № 8. P. 1396–1404.
189. Samgina T.Y., Tolpina M.D., Trebse P., Torkar G., Artemenko K.A., Bergquist J., Lebedev A.T. LTQ orbitrap velos in routine de novo sequencing of non-tryptiskin peptides from the frog *rana latastei* with traditional and reliable manual spectra interpretation // *Rapid Commun. Mass Spectrom.* John Wiley and Sons Ltd, 2016. Vol. 30, № 2. P. 265–276.
  190. Evaristo G.P.C., Pinkse M.W.H., Chen T., Wang L., Mohammed S., Heck A.J.R., Mathes I., Lottspeich F., Shaw C., Albar J.P., Verhaert P.D.E.M. De novo sequencing of two novel peptides homologous to calcitonin-like peptides, from skin secretion of the Chinese Frog, *Odorrana schmackeri* // *EuPA Open Proteomics.* Elsevier, 2015. Vol. 8. P. 157–166.
  191. Balderrama G.D., Meneses E.P., Orihuela L.H., Hernández O.V., Franco R.C., Robles V.P., Batista C.V.F. Analysis of sulfated peptides from the skin secretion of the *Pachymedusa dancicolor* frog using IMAC-Ga enrichment and high-resolution mass spectrometry // *Rapid Commun. Mass Spectrom.* John Wiley & Sons, Ltd, 2011. Vol. 25, № 8. P. 1017–1027.
  192. Huo Y., Xv R., Ma H., Zhou J., Xi X., Wu Q., Duan J., Zhou M., Chen T. Identification of <10 KD peptides in the water extraction of *Venenum Bufonis* from *Bufo gargarizans* using Nano LC–MS/MS and De novo sequencing // *J. Pharm. Biomed. Anal.* Elsevier, 2018. Vol. 157. P. 156–164.
  193. Petras D., Heiss P., D. Süssmuth R., J. Calvete J. Venom Proteomics of Indonesian King Cobra, *Ophiophagus hannah*: Integrating Top-Down and Bottom-Up Approaches // *J. Proteome Res.* 2015. Vol. 14, № 6. P. 2539–2556.
  194. Petras D., Heiss P., Harrison R.A., Süssmuth R.D., Calvete J.J. Top-down venomomics of the East African green mamba, *Dendroaspis angusticeps*, and the black mamba, *Dendroaspis polylepis*, highlight the complexity of their toxin arsenals // *J. Proteomics.* Elsevier, 2016. Vol. 146. P. 148–164.
  195. Trevisan-Silva D., Bednaski A. V., Fischer J.S.G., Veiga S.S., Bandeira N., Guthals A., Marchini F.K., Leprevost F. V., Barbosa V.C., Senff-Ribeiro A., Carvalho P.C. A multi-protease, multi-dissociation, bottom-up-to-top-down proteomic view of the *Loxosceles intermedia* venom // *Sci. Data* 2017 41. Nature Publishing Group, 2017. Vol. 4, № 1. P. 1–7.
  196. L. Jarecki J., L. Frey B., M. Smith L., O. Stretton A. Discovery of Neuropeptides in the Nematode *Ascaris suum* by Database Mining and Tandem Mass Spectrometry // *J. Proteome Res.* 2011. Vol. 10, № 7. P. 3098–3106.
  197. Hui L., Cunningham R., Zhang Z., Cao W., Jia C., Li L. Discovery and Characterization of the Crustacean Hyperglycemic Hormone Precursor Related Peptides (CPRP) and Orcokinin Neuropeptides in the Sinus Glands of the Blue Crab *Callinectes sapidus* Using Multiple

- Tandem Mass Spectrometry Techniques // *J. Proteome Res.* 2011. Vol. 10, № 9. P. 4219–4229.
198. Trachsel C., Siegemund D., Kämpfer U., Kopp L.S., Bühr C., Grossmann J., Lüthi C., Cunningham M., Nentwig W., Kuhn-Nentwig L., Schürch S., Schaller J. Multicomponent venom of the spider *Cupiennius salei*: A bioanalytical investigation applying different strategies // *FEBS J.* 2012. Vol. 279, № 15. P. 2683–2694.
  199. Hui L., Xiang F., Zhang Y., Li L. Mass spectrometric elucidation of the neuropeptidome of a crustacean neuroendocrine organ // *Peptides*. Elsevier, 2012. Vol. 36, № 2. P. 230–239.
  200. Romanova E. V., Roth M.J., Rubakhin S.S., Jakubowski J.A., Kelley W.P., Kirk M.D., Kelleher N.L., Sweedler J. V. Identification and characterization of homologues of vertebrate  $\beta$ -thymosin in the marine mollusk *Aplysia californica* // *J. Mass Spectrom.* 2006. Vol. 41, № 8. P. 1030–1040.
  201. <http://www.ncbi.nlm.nih.gov> [Electronic resource].
  202. [www.crocgenomes.org](http://www.crocgenomes.org) [Electronic resource].
  203. Touchard A., Téné N., Chan Tchi Song P., Lefranc B., Leprince J., Treilhou M., Bonnafé E. Deciphering the Molecular Diversity of an Ant Venom Peptidome through a Venomics Approach // *J. Proteome Res.* 2018. Vol. 17, № 10. P. 3503–3516.
  204. Gad S.C. QSAR // *Encyclopedia of Toxicology*. Elsevier, 2014. P. 1–9.
  205. Nesvizhskii A.I. Protein Identification by Tandem Mass Spectrometry and Sequence Database Searching // *Mass Spectrometry Data Analysis in Proteomics*. New Jersey: Humana Press. P. 87–120.
  206. Bairoch A. The Universal Protein Resource (UniProt) // *Nucleic Acids Res.* 2004. Vol. 33, № Database issue. P. D154–D159.
  207. Ivanov M. V., Levitsky L.I., Lobas A.A., Tarasova I.A., Pridatchenko M.L., Zgoda V.G., Moshkovskii S.A., Mitulovic G., Gorshkov M. V. Peptide identification in “shotgun” proteomics using tandem mass spectrometry: Comparison of search engine algorithms // *J. Anal. Chem.* 2015. Vol. 70, № 14. P. 1614–1619.
  208. Kapp E., Schütz F. Overview of Tandem Mass Spectrometry (MS/MS) Database Search Algorithms // *Curr. Protoc. Protein Sci.* 2007. Vol. 49, № 1.
  209. Eng J.K., McCormack A.L., Yates J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database // *J. Am. Soc. Mass Spectrom.* Springer-Verlag New York, 1994. Vol. 5, № 11. P. 976–989.
  210. Perkins D.N., Pappin D.J.C., Creasy D.M., Cottrell J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data // *Electrophoresis*. 1999. Vol. 20, № 18. P. 3551–3567.

211. Colinge J., Masselot A., Giron M., Dessingy T., Magnin J. OLAV: Towards high-throughput tandem mass spectrometry data identification // *Proteomics*. 2003. Vol. 3, № 8. P. 1454–1463.
212. <https://www.thegpm.org/tandem/release.html> [Electronic resource].
213. Craig R., Beavis R.C. A method for reducing the time required to match protein sequences with tandem mass spectra // *Rapid Commun. Mass Spectrom.* 2003. Vol. 17, № 20. P. 2310–2316.
214. Zhang J., Xin L., Shan B., Chen W., Xie M., Yuen D., Zhang W., Zhang Z., Lajoie G.A., Ma B. PEAKS DB: De Novo Sequencing Assisted Database Search for Sensitive and Accurate Peptide Identification // *Mol. Cell. Proteomics*. 2012. Vol. 11, № 4. P. M111.010587.
215. Geer L.Y., Markey S.P., Kowalak J.A., Wagner L., Xu M., Maynard D.M., Yang X., Shi W., Bryant S.H. Open mass spectrometry search algorithm // *J. Proteome Res. American Chemical Society*, 2004. Vol. 3, № 5. P. 958–964.
216. Tanner S., Shu H., Frank A., Wang L.-C., Zandi E., Mumby M., Pevzner P.A., Bafna V. InsPecT: Identification of Posttranslationally Modified Peptides from Tandem Mass Spectra // *Anal. Chem.* 2005. Vol. 77, № 14. P. 4626–4639.
217. Zhang Z., Sun S., Zhu X., Chang S., Liu X., Yu C., Bu D., Chen R. A novel scoring schema for peptide identification by searching protein sequence databases using tandem mass spectrometry data // *BMC Bioinformatics*. 2006. Vol. 7, № 1. P. 222.
218. DiMaggio P.A., Floudas C.A. De Novo Peptide Identification via Tandem Mass Spectrometry and Integer Linear Optimization // *Anal. Chem.* 2007. Vol. 79, № 4. P. 1433–1446.
219. Robotham S.A., Horton A.P., Cannon J.R., Cotham V.C., Marcotte E.M., Brodbelt J.S. UVnovo: A de Novo Sequencing Algorithm Using Single Series of Fragment Ions via Chromophore Tagging and 351 nm Ultraviolet Photodissociation Mass Spectrometry // *Anal. Chem.* 2016. Vol. 88, № 7. P. 3990–3997.
220. Ma B. Novor: Real-Time Peptide de Novo Sequencing Software // *J. Am. Soc. Mass Spectrom.* 2015. Vol. 26, № 11. P. 1885–1894.
221. Yang H., Chi H., Zhou W.-J., Zeng W.-F., Liu C., Wang R.-M., Wang Z.-W., Niu X.-N., Chen Z.-L., He S.-M. pSite: Amino Acid Confidence Evaluation for Quality Control of De Novo Peptide Sequencing and Modification Site Localization // *J. Proteome Res.* 2018. Vol. 17, № 1. P. 119–128.
222. Tran N.H., Zhang X., Xin L., Shan B., Li M. De novo peptide sequencing by deep learning // *Proc. Natl. Acad. Sci.* 2017. Vol. 114, № 31. P. 8247–8252.

223. Muth T., Hartkopf F., Vaudel M., Renard B.Y. A Potential Golden Age to Come-Current Tools, Recent Use Cases, and Future Avenues for De Novo Sequencing in Proteomics // *Proteomics*. 2018. Vol. 18, № 18. P. 1700150.
224. Taylor J.A., Johnson R.S. Sequence database searches viade novo peptide sequencing by tandem mass spectrometry // *Rapid Commun. Mass Spectrom.* 1997. Vol. 11, № 9. P. 1067–1075.
225. Fernandez-de-Cossio J., Gonzalez J., Betancourt L., Besada V., Padron G., Shimonishi Y., Takao T. Automated interpretation of high-energy collision-induced dissociation spectra of singly protonated peptides by ‘seqms’, a software aid for de novo sequencing by tandem mass spectrometry’ // *Rapid Commun. Mass Spectrom.* 1998. Vol. 12, № 23. P. 1867–1878.
226. Fernandez-de-Cossio J., Gonzalez J., Satomi Y., Shima T., Okumura N., Besada V., Betancourt L., Padron G., Shimonishi Y., Takao T. Automated interpretation of low-energy collision-induced dissociation spectra by SeqMS, a software aid for de novo sequencing by tandem mass spectrometry // *Electrophoresis*. 2000. Vol. 21, № 9. P. 1694–1699.
227. Ma B., Zhang K., Hendrie C., Liang C., Li M., Doherty-Kirby A., Lajoie G. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry // *Rapid Commun. Mass Spectrom.* 2003. Vol. 17, № 20. P. 2337–2342.
228. Frank A., Pevzner P. PepNovo: De novo peptide sequencing via probabilistic network modeling // *Anal. Chem. American Chemical Society*, 2005. Vol. 77, № 4. P. 964–973.
229. Fischer B., Roth V., Roos F., Grossmann J., Baginsky S., Widmayer P., Gruissem W., Buhmann J.M. NovoHMM: A Hidden Markov Model for de Novo Peptide Sequencing // *Anal. Chem.* 2005. Vol. 77, № 22. P. 7265–7273.
230. Tabb D.L., Ma Z.-Q., Martin D.B., Ham A.-J.L., Chambers M.C. DirecTag: Accurate Sequence Tags from Peptide MS/MS through Statistical Scoring // *J. Proteome Res.* 2008. Vol. 7, № 9. P. 3838–3846.
231. Chi H., Sun R.-X., Yang B., Song C.-Q., Wang L.-H., Liu C., Fu Y., Yuan Z.-F., Wang H.-P., He S.-M., Dong M.-Q. pNovo: De novo Peptide Sequencing and Identification Using HCD Spectra // *J. Proteome Res.* 2010. Vol. 9, № 5. P. 2713–2724.
232. Chi H., Chen H., He K., Wu L., Yang B., Sun R.-X., Liu J., Zeng W.-F., Song C.-Q., He S.-M., Dong M.-Q. pNovo+: De Novo Peptide Sequencing Using Complementary HCD and ETD Tandem Mass Spectra // *J. Proteome Res.* 2013. Vol. 12, № 2. P. 615–625.
233. Jeong K., Kim S., Pevzner P.A. UniNovo: a universal tool for de novo peptide sequencing // *Bioinformatics*. 2013. Vol. 29, № 16. P. 1953–1962.
234. Vyatkina K., Wu S., Dekker L.J.M., VanDuijn M.M., Liu X., Tolić N., Dvorkin M., Alexandrova S., Luider T.M., Paša-Tolić L., Pevzner P.A. De Novo Sequencing of Peptides

- from Top-Down Tandem Mass Spectra // *J. Proteome Res.* 2015. Vol. 14, № 11. P. 4450–4462.
235. Devabhaktuni A., Elias J.E. Application of de Novo Sequencing to Large-Scale Complex Proteomics Data Sets // *J. Proteome Res.* 2016. Vol. 15, № 3. P. 732–742.
  236. Li C., Chen T., He Q., Zhu Y., Li K. MRUniNovo: an efficient tool for de novo peptide sequencing utilizing the hadoop distributed computing framework // *Bioinformatics.* 2016. P. btw721.
  237. Yang H., Chi H., Zhou W.-J., Zeng W.-F., He K., Liu C., Sun R.-X., He S.-M. Open-pNovo: De Novo Peptide Sequencing with Thousands of Protein Modifications // *J. Proteome Res.* 2017. Vol. 16, № 2. P. 645–654.
  238. Ma B., Zhang K., Liang C. An effective algorithm for peptide de novo sequencing from MS/MS spectra // *J. Comput. Syst. Sci.* 2005. Vol. 70, № 3. P. 418–430.
  239. <https://www.bioinfor.com/peaks-studio/> [Electronic resource].
  240. Horn D.M., Zubarev R.A., McLafferty F.W. Automated reduction and interpretation of // *J. Am. Soc. Mass Spectrom.* 2000. Vol. 11, № 4. P. 320–332.
  241. Liu X., Inbar Y., Dorrestein P.C., Wynne C., Edwards N., Souda P., Whitelegge J.P., Bafna V., Pevzner P.A. Deconvolution and Database Search of Complex Tandem Mass Spectra of Intact Proteins // *Mol. Cell. Proteomics.* 2010. Vol. 9, № 12. P. 2772–2782.
  242. Hu A., Noble W.S., Wolf-Yadlin A. Technical advances in proteomics: new developments in data-independent acquisition // *F1000Research.* 2016. Vol. 5. P. 419.
  243. Kohli B.M., Eng J.K., Nitsch R.M., Konietzko U. An alternative sampling algorithm for use in liquid chromatography/tandem mass spectrometry experiments // *Rapid Commun. Mass Spectrom.* 2005. Vol. 19, № 5. P. 589–596.
  244. Conlon J.M. Reflections on a systematic nomenclature for antimicrobial peptides from the skins of frogs of the family Ranidae // *Peptides.* Elsevier, 2008. Vol. 29, № 10. P. 1815–1819.
  245. Simmaco M., Mignogna G., Barra D., Bossa F. Antimicrobial peptides from skin secretions of *Rana esculenta*. Molecular cloning of cDNAs encoding esculentin and brevinins and isolation of new active peptides. // *J. Biol. Chem.* 1994. Vol. 269, № 16. P. 11956–11961.
  246. Barra D., Simmaco M., Boman H.G., Pasteur-Fondazione Cenci Bolognetti I., di Scienze BiochimichèA Rossi D. Gene-encoded peptide antibiotics and innate immunity // *FEBS Lett.* John Wiley & Sons, Ltd, 1998. Vol. 430, № 1–2. P. 130–134.
  247. Samgina T.Y., Tolpina M.D., Surin A.K., Kovalev S. V., Bosch R.A., Alonso I.P., Garcia F.A., Gonzalez Lopez L.J., Lebedev A.T. Manual mass spectrometry de novo sequencing of the anionic host defense peptides of the Cuban Treefrog *Osteopilus septentrionalis* //

- Rapid Commun. Mass Spectrom. 2021. Vol. 35, № 7.
248. Isaacson T., Soto A.M., Iwamuro S., Knoop F.C., Conlon J.M. Antimicrobial peptides with atypical structural features from the skin of the Japanese brown frog *Rana japonica* // Peptides. Elsevier, 2002. Vol. 23, № 3. P. 419–425.
  249. Romero S.M., Cardillo A.B., Martínez Ceron M.C., Camperi S.A., Giudicessi S.L. Temporins: An Approach of Potential Pharmaceutic Candidates // Surg. Infect. (Larchmt). 2020. Vol. 21, № 4. P. 309–322.
  250. Conlon J.M., Sonnevend Á., Patel M., Al-Dhaheri K., Nielsen P.F., Kolodziejek J., Nowotny N., Iwamuro S., Pál T. A family of brevinin-2 peptides with potent activity against *Pseudomonas aeruginosa* from the skin of the Hokkaido frog, *Rana pirica* // Regul. Pept. Elsevier, 2004. Vol. 118, № 3. P. 135–141.
  251. Kim J.B., Conlon J.M., Iwamuro S., Knoop F.C. Antimicrobial peptides from the skin of the Japanese mountain brown frog, *Rana ornativentris* // J. Pept. Res. 2001. Vol. 58, № 5. P. 349–356.
  252. Shang D., Yu F., Li J., Zheng J., Zhang L., Li Y. Molecular cloning of cDNAs encoding antimicrobial peptide precursors from the skin of the Chinese Brown Frog, *Rana chensinensis* // Zoolog. Sci. Zoological Society of Japan, 2009. Vol. 26, № 3. P. 220–226.
  253. Iwamuro S., Nakamura M., Ohnuma A., Conlon J.M. Molecular cloning and sequence analyses of preprotemporin mRNAs containing premature stop codons from extradermal tissues of *Rana tagoi* // Peptides. Elsevier, 2006. Vol. 27, № 9. P. 2124–2128.
  254. Good D.M., Marin-Vicente C., Zubarev R.A. Are the majority of a2-ions cyclic? // Phys. Chem. Chem. Phys. The Royal Society of Chemistry, 2010. Vol. 12, № 41. P. 13372–13374.
  255. Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L., Barra D. Temporins, Antimicrobial Peptides from the European Red Frog *Rana temporaria* // Eur. J. Biochem. 1996. Vol. 242, № 3. P. 788–792.
  256. Conlon J.M., Sonnevend A., Patel M., Camasamudram V., Nowotny N., Zilahi E., Iwamuro S., Nielsen P.F., Pál T. A melittin-related peptide from the skin of the Japanese frog, *Rana tagoi*, with antimicrobial and cytolytic properties // Biochem. Biophys. Res. Commun. Academic Press, 2003. Vol. 306, № 2. P. 496–500.
  257. Conlon J.M., Al-Ghafari N., Coquet L., Leprince J., Jouenne T., Vaudry H., Davidson C. Evidence from peptidomic analysis of skin secretions that the red-legged frogs, *Rana aurora draytonii* and *Rana aurora aurora*, are distinct species // Peptides. Elsevier, 2006. Vol. 27, № 6. P. 1305–1312.
  258. Suzuki H., Iwamuro S., Ohnuma A., Coquet L., Leprince J., Jouenne T., Vaudry H., Taylor C.K., Abel P.W., Conlon J.M. Expression of genes encoding antimicrobial and bradykinin-

- related peptides in skin of the stream brown frog *Rana sakuraii* // *Peptides*. Elsevier, 2007. Vol. 28, № 3. P. 505–514.
259. Mechkarska M., Kolodziejek J., Musale V., Coquet L., Leprince J., Jouenne T., Nowotny N., Conlon J.M. Peptidomic analysis of the host-defense peptides in skin secretions of *Rana graeca* provides insight into phylogenetic relationships among Eurasian *Rana* species // *Comp. Biochem. Physiol. Part D Genomics Proteomics*. Elsevier, 2019. Vol. 29. P. 228–234.
  260. Rončević T., Krce L., Gerdol M., Pacor S., Benincasa M., Guida F., Aviani I., Čikeš-Čulić V., Pallavicini A., Maravić A., Tossi A. Membrane-active antimicrobial peptide identified in *Rana arvalis* by targeted DNA sequencing // *Biochim. Biophys. Acta - Biomembr.* Elsevier, 2019. Vol. 1861, № 3. P. 651–659.
  261. Brand G.D., Krause F.C., Silva L.P., Leite J.R.S.A., Melo J.A.T., Prates M. V., Pesquero J.B., Santos E.L., Nakaie C.R., Costa-Neto C.M., Bloch C. Bradykinin-related peptides from *Phyllomedusa hypochondrialis* // *Peptides*. Elsevier, 2006. Vol. 27, № 9. P. 2137–2146.
  262. Anastasi A., Erspamer V., Bertaccini G. Occurrence of bradykinin in the skin of *Rana temporaria* // *Comp. Biochem. Physiol.* 1965. Vol. 14, № 1. P. 43–52.
  263. Conlon J.M., Aronsson U. Multiple Bradykinin-Related Peptides From the Skin of the Frog, *Rana temporaria* // *Peptides*. 1997. Vol. 18, № 3. P. 361–365.
  264. Couture R., Harrisson M., Vianna R.M., Cloutier F. Kinin receptors in pain and inflammation // *Eur. J. Pharmacol.* 2001. Vol. 429, № 1–3. P. 161–176.
  265. Tennessen J.A. Enhanced Synonymous Site Divergence in Positively Selected Vertebrate Antimicrobial Peptide Genes // *J. Mol. Evol.* 2005. Vol. 61, № 4. P. 445–455.
  266. Tennessen J.A., Woodhams D.C., Chaurand P., Reinert L.K., Billheimer D., Shyr Y., Caprioli R.M., Blouin M.S., Rollins-Smith L.A. Variations in the expressed antimicrobial peptide repertoire of northern leopard frog (*Rana pipiens*) populations suggest intraspecies differences in resistance to pathogens // *Dev. Comp. Immunol.* 2009. Vol. 33, № 12. P. 1247–1257.
  267. Tennessen J.A., Blouin M.S. Selection for Antimicrobial Peptide Diversity in Frogs Leads to Gene Duplication and Low Allelic Variation // *J. Mol. Evol.* 2007. Vol. 65, № 5. P. 605–615.
  268. Tennessen J.A., Blouin M.S. A revised leopard frog phylogeny allows a more detailed examination of adaptive evolution at ranatuerin-2 antimicrobial peptide loci // *Immunogenetics*. 2010. Vol. 62, № 5. P. 333–343.
  269. Cole S.R., Ma X., Zhang X., Xia Y. Electron Transfer Dissociation (ETD) of Peptides

- Containing Intrachain Disulfide Bonds // *J. Am. Soc. Mass Spectrom.* 2012. Vol. 23, № 2. P. 310–320.
270. Wysocki V.H., Tsaprailis G., Smith L.L., Brexi L.A. Mobile and localized protons: a framework for understanding peptide dissociation. // *J. Mass Spectrom.* 2000. Vol. 35, № 12. P. 1399–1406.
271. Tyler M.J., Stone D.J.M., Bowie J.H. A novel method for the release and collection of dermal, glandular secretions from the skin of frogs // *J. Pharmacol. Toxicol. Methods.* 1992. Vol. 28, № 4. P. 199–200.