

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

На правах рукописи

Сачкова Мария Юрьевна

**Двудоменные токсины ядов пауков**

Специальность 02.00.10 – Биоорганическая химия

Диссертация на соискание ученой степени  
кандидата химических наук

Научный руководитель  
к.х.н. Василевский А.А.

Москва – 2014

## Оглавление

<b>Введение .....</b>	<b>5</b>
<b>Обзор литературы .....</b>	<b>7</b>
1. Введение .....	7
2. Состав яда пауков .....	7
2.1 Белковые компоненты яда.....	8
2.2 Пептидные компоненты яда.....	9
2.2.1 Однодоменные токсины .....	9
2.2.2 Двудоменные токсины .....	16
3. Организация белков-предшественников токсинов .....	21
4. Гены токсинов животных .....	23
4.1 Гены токсинов пауков.....	23
4.1.1 Гены пептидных токсинов .....	23
4.1.2 Гены белковых токсинов.....	24
4.2 Гены токсинов скорпионов .....	24
4.3 Гены токсинов моллюсков конусов .....	27
4.4 Гены токсинов змей .....	30
5. Эволюция токсинов белковой/пептидной природы.....	33
5.1 Откуда берутся токсины? .....	33
5.2 Молекулярные механизмы эволюции токсинов .....	35
5.2.1 Конусы .....	36
5.2.2 Змеи .....	37
5.2.3 Скорпионы.....	42
5.2.4 Пауки.....	44
5.2.5 Морские анемоны .....	46
5.3 Заключение .....	47
<b>Материалы и методы.....</b>	<b>49</b>
1. Материалы.....	49
1.1 Неорганические вещества .....	49
1.2 Органические вещества .....	49
1.3 Реактивы для молекулярной биологии: .....	49
1.4 Клеточные линии .....	49
1.5 Питательные среды и их составляющие .....	49

2.	Методы .....	50
2.1	Получение двудоменных токсинов и их производных .....	50
2.1.1	Выделение двудоменных токсинов из цельных ядов.....	50
2.1.2	Химический синтез пептидов .....	51
2.1.3	Получение рекомбинантных пептидов.....	51
2.2	Аналитические методы.....	53
2.2.1	Определение первичной структуры пептидов .....	53
2.2.2	Исследование вторичной структуры пептидов методом кругового дихроизма.....	53
2.2.3	Определение концентрации пептидов .....	54
2.2.4	Масс-спектрометрия.....	54
2.2.5	Методы биоинформатики .....	54
2.3	Изучение функциональной активности пептидов .....	54
2.3.1	Взаимодействие с искусственными мембранами.....	54
2.3.2	Биологическая активность .....	55
2.4	Работа с кДНК и геномной ДНК .....	56
2.4.1	Библиотеки кДНК из ядовитых желез .....	56
2.4.2	Выделение ДНК и РНК из ядовитых желез пауков.....	57
2.4.3	Аmplификация кДНК и фрагментов геномной ДНК и секвенирование.....	57
2.4.4	Анализ нуклеотидных последовательностей .....	58
	<b>Результаты.....</b>	<b>60</b>
1.	Токсины пауков рода <i>Oxyopes</i> .....	60
1.1	Выделение двудоменных токсинов .....	60
1.2	Определение последовательности OtTx .....	61
1.3	Анализ последовательности OtTx .....	61
1.4	Структура белка-предшественника OtTx.....	64
1.5	Получение полноразмерного OtTx 1a и его фрагментов.....	64
1.5.1	Получение OtTx 1a.....	64
1.5.2	Получение OtTx 1a-AMP.....	65
1.5.3	Получение OtTx 1a-ICK .....	65
1.6	Вторичная структура OtTx 1a .....	67
1.7	Биологическая активность OtTx 1a и его производных .....	68
2.	Гены двудоменных токсинов пауков рода <i>Oxyopes</i> .....	69
2.1	Анализ последовательностей кДНК из ядовитых желез пауков <i>Oxyopes</i> .....	69

2.2	Структура генов спайдеринов.....	71
2.3	Молекулярная эволюция двудоменных токсинов <i>Oxyopes</i> .....	72
3.	Выделение новых двудоменных токсинов из яда паука <i>C. punctorium</i> .....	75
4.	Гены двудоменных токсинов <i>C. punctorium</i> .....	76
4.1	Анализ последовательностей кДНК из ядовитых желез паука <i>C. punctorium</i> ....	76
4.2	Структура генов СрТх-подобных токсинов .....	78
4.3	Молекулярная эволюция двудоменных токсинов <i>C. punctorium</i> .....	78
5.	Двудоменные токсины паука <i>L. tarabaevi</i> .....	81
<b>Обсуждение .....</b>		<b>84</b>
1.	Получение рекомбинантных полипептидов .....	84
2.	ОтТх-подобные токсины <i>Oxyopes</i> – новый класс двудоменных токсинов .....	86
2.1	Двудоменная структура токсинов ОтТх.....	86
2.2	N-концевой модуль ОтТх 1a – мощный цитолитический токсин.....	87
3.	Синергизм в основе активности цитоинсектотоксинов <i>L. tarabaevi</i> .....	87
4.	Разнообразие двудоменных токсинов .....	88
4.1	Двудоменные токсины <i>Oxyopes</i> .....	89
4.2	Двудоменные токсины <i>C. punctorium</i> .....	90
5.	Безинтронные гены двудоменных токсинов.....	90
6.	Молекулярная эволюция двудоменных токсинов.....	92
6.1	Анализ типов отбора, действующего на двудоменные токсины.....	92
6.2	Возникновение токсинов типа АМР+ІСК у <i>Oxyopes</i> .....	93
6.3	Возникновение токсинов типа ІСК+ІСК у <i>Cheiracanthium</i> .....	94
<b>Выводы .....</b>		<b>97</b>
<b>Список литературы.....</b>		<b>98</b>

## Введение

Широко известно, что пауки продуцируют яд, состоящий из смеси мощных и селективных токсинов, каждый из которых может иметь свою специфическую биологическую активность. С одной стороны, яд пауков – фактор, несомненно, сыгравший одну из ключевых ролей в эволюции этих беспозвоночных, с другой – он является источником веществ, имеющих важное прикладное значение в качестве лекарственных препаратов или инструментов изучения нервной системы. Эти аргументы обуславливают актуальность исследования разнообразия токсинов пауков. Однако лишь несколько видов, представляющих собой небольшую долю известного биоразнообразия, были изучены подробно, поэтому наше представление о репертуаре токсинов из яда пауков очень ограничено.

Пептидная составляющая наиболее представлена в яде всех изученных видов. Пептиды могут быть двух типов. (А) Нейротоксины обычно богаты дисульфидными связями и в пространстве формируют структуру «цистинового узла» (ИСК – от англ. inhibitory cysteine knot). По-другому такие токсины называют ноттинами. (Б) Цитотоксины обычно являются линейными пептидами, приобретающими  $\alpha$ -спиральную конформацию при взаимодействии с липидными мембранами. Их также часто называют антимикробными пептидами (АМР, от англ. antimicrobial peptide) в связи с тем, что они активны против микроорганизмов.

Разнообразие компонентов яда является эволюционным преимуществом пауков, поэтому они «стремятся» его повышать. Это приводит к формированию так называемых «комбинаторных библиотек» токсинов: яд содержит группы гомологичных пептидов, аминокислотные последовательности которых вариабельны, а общие структурные мотивы консервативны. Однако известны и другие механизмы повышения разнообразия пептидных компонентов яда. Например, существуют дисульфид-богатые токсины с другими типами пространственной укладки. Относительно недавно в яде нескольких пауков были найдены токсины, состоящие из двух модулей (или доменов), каждый из которых соответствует «обычному» токсину. К началу работы над этой диссертацией были известны следующие комбинации доменов: (а) два ИСК домена в составе одного токсина, (б) ИСК домен с длинным линейным С-концевым модулем, обладающим мембранной активностью, (в) два коротких линейных АМР, объединенные в одну молекулу. В принципе, возможен и четвертый вариант комбинации доменов – N-концевой домен типа АМР и С-концевой домен типа ИСК. Двудоменные токсины изучены недостаточно. Так, с одной стороны, неизвест-

ны функциональные роли отдельных модулей большинства токсинов, а с другой, механизмы их молекулярной эволюции.

### **Цель и задачи работы**

Целью данной работы было изучение разнообразия двудоменных (модульных) токсинов пауков. Были поставлены следующие задачи:

1. Поиск новых двудоменных токсинов пауков.
2. Изучение структуры и биологической активности модульных токсинов.
3. Определение структуры генов, кодирующих двудоменные токсины пауков.
4. Выяснение механизмов молекулярной эволюции модульных токсинов.

## Обзор литературы

### 1. Введение

Двудоменные токсины – интересное явление не только в мире токсинов пауков, но и вообще в мире животных токсинов. С одной стороны, они обладают интересными функциональными свойствами, обеспечиваемыми неизученными механизмами, а с другой, представляют собой интересный эволюционный феномен. В «Обзоре литературы» будут рассмотрены те вопросы, которые необходимы для интерпретации результатов этой работы. Во многих случаях это не только информация, посвященная паукам, но и факты из биологии других ядовитых животных.

В главе «Состав ядов пауков» описывалось известное разнообразие токсинов пауков, с особым вниманием к белковым и пептидным токсинам (в т. ч. двудоменным). К настоящему моменту генетика пауков была изучена относительно плохо, поэтому имеется очень мало информации о структуре генов токсинов. Однако у некоторых других ядовитых животных этот вопрос исследован намного лучше. Учитывая параллелизмы в механизмах формирования ядов разных животных, будет рассмотрена структура генов токсинов разных животных (глава «Гены токсинов животных»). Последняя глава Обзора литературы посвящена молекулярной эволюции животных токсинов («Эволюция токсинов белковой/пептидной природы»), позволяющей формировать их широкое разнообразие.

### 2. Состав яда пауков

Пауки – обладатели мощного оружия нападения и защиты – ядов. К настоящему времени уже было опубликовано несколько обзоров, достаточно полно описывающих состав яда пауков и разнообразие их токсинов [1], поэтому здесь приведен лишь краткий обзор на эту тему, необходимый для интерпретации экспериментальных данных.

Компоненты яда имеют разную химическую природу, и в зависимости от их молекулярной массы они подразделяются на низкомолекулярные (<1 кДа), пептиды (<10 кДа) и высокомолекулярные (>10 кДа) вещества – белки. Низкомолекулярные компоненты, входящие в состав яда пауков, относятся к разным группам химических веществ. Это соли, углеводы, аминокислоты, биогенные амины, ацилполиамины и др. [1,2]. Несмотря на высокое разнообразие, в яде каждого паука обычно преобладает определенная группа

компонентов. Так, у пауков семейства Araneidae основным компонентом яда являются низкомолекулярные вещества [3–7], а крупные белки преобладают в ядах пауков семейств Sicariidae [8–10] и Theridiidae [11–13]. Большинство изученных пауков вырабатывает яды с пептидным основным компонентом [1]. Далее будут подробно рассмотрены компоненты яда пауков, имеющие полипептидную природу, поскольку эта диссертация посвящена пептидным токсинам.

## 2.1 Белковые компоненты яда

Белки в составе ядов могут выполнять следующие функции: оказывать прямое токсическое действие [14,15], а также, будучи ферментами, разрушающими тканевые структуры, способствовать распространению токсинов и, кроме того, участвовать в конечных стадиях созревания токсинов [16].

Пауки рода *Latrodectus* (пауки-вдовы) семейства Theridiidae вырабатывают яды, содержащие латротоксины, активные в отношении позвоночных ( $\alpha$ -латротоксин  $\alpha$ -LTX), ракообразных и насекомых [11,17,18]. Латротоксины блокируют проведение нервного импульса, провоцируя истощающий выброс всех известных типов нейромедиаторов из синаптических окончаний жертвы [17,19]. Связавшись с латрофилином или нейрексином Ia,  $\alpha$ -LTX встраивается в мембрану и формирует поры, проницаемые для  $Ca^{2+}$ , и запускает выход нейротрансмиттеров путем экзоцитоза. С другой стороны, связывание  $\alpha$ -LTX с GPCR-рецептором латрофилином приводит к запуску внутриклеточного каскада и выходу ионов кальция из пресинаптических депо [20].

На данный момент изучена первичная структура четырех латротоксинов [21–24]. Это крупные белки (длина более 1000 аминокислотных остатков, а. о.).  $\alpha$ -LTX в активной форме представляет собой тетрамер с каналом и может встраиваться в липидные бислои. Мономер состоит из трех доменов (36 кДа, 76 кДа и 18 кДа), в среднем из которых находятся анкириновые повторы [25].

В состав яда пауков-отшельников семейства Sicariidae входят ферменты, например, фосфатазы, гиалуронидазы, фосфолипазы и различных протеазы. Главные компоненты – фосфолипазы D (или сфингомиелиназы) [26,27]. Сфингомиелиназы D катализируют гидролиз сфингомиелина по  $Mg^{2+}$ -зависимому механизму. Изучена пространственная структура этого фермента из яда паука *Loxosceles laeta*, она относится к укладке типа «ТИМ-бочки» (от англ. TIM barrel). Остатки, ответственные за связывание ионов магния и входящие в состав каталитического сайта, высококонсервативны. Незначительные вариации



последовательности приводят к снижению активности на сфингомиелине, или даже полной ее потере [28].

## **2.2 Пептидные компоненты яда**

### **2.2.1 Однодоменные токсины**

#### **2.2.1.1 Линейные токсины**

Токсины, представляющие собой линейные пептиды, довольно широко распространены в ядах пауков [1,29]. У *Lachesana tarabaevi* (семейство *Zodariidae*) они даже являются основным действующим компонентом яда [30,31]. В настоящее время из ядов нескольких видов пауков (наиболее изучены *Hogna carolinensis*, *Lycosa singoriensis*, *Oxyopes takobius*, *Cupiennius salei*, *L. tarabaevi*) выделено несколько десятков цитолитических пептидов, объединяемых в ~10 – 15 групп.

Обычно линейные пептиды проявляют цитолитическую активность, вызывая некроз тканей жертвы. Они могут вызывать летальный эффект, однако он проявляется только в случае их высокого содержания в яде. Кроме того, известны случаи синергизма между цитолитическими и нейротоксическими компонентами ядов, когда цитолитики, «растворяя» ткани жертвы, делают мишени действия нейротоксинов более доступными [32]. Не стоит забывать об антимикробной активности цитолитических пептидов, превращающей их в «консерванты» для добычи пауков, как было показано для скорпионов [33].

Длина таких пептидов ограничивается ~50 аминокислотными остатками, среди которых значительная доля положительно заряженных. Интересно, что частота встречаемости остатков лизина значительно преобладает над частотой встречаемости остатков аргинина. При pH 7 общий заряд пептида, как правило, попадает в интервал от +3 до +10 [29]. Линейные пептиды в водной среде обычно не упорядочены, однако при контакте с мембранами они приобретают структуру  $\alpha$ -спирали, на поверхности которой находятся пространственно разделенные гидрофильный и гидрофобный кластеры (т.н. амфипатическая  $\alpha$ -спираль). В большинстве случаев амфипатичность проявляется на уровне вторичной структуры. Методом ЯМР установлена пространственная структура нескольких цитолитических пептидов из яда пауков, например, латарцинов (Ltc) 1 и 2а *L. tarabaevi* [34,35], купиеннина 1а из яда *C. salei* [36] и оксиопинина 4а *O. takobius* [37] в условиях, имитирующих мембранное окружение. Такая организация свойственна также цитолитическим пептидам из ядов других организмов, например, для хорошо изученного мелиттина пчелы

[38]. Высокий положительный заряд гидрофильного кластера на поверхности спирали способствует ее связыванию с отрицательно заряженными головками мембранных фосфолипидов, в то время как ее гидрофобная поверхность встраивается внутрь липидного бислоя. Для многих линейных пептидов из яда пауков свойственна пост-трансляционная модификация – амидирование С-концевого аминокислотного остатка [1].

В целом, цитолитические пептиды пауков взаимодействуют с широким спектром мембран и характеризуются низкой специфичностью (хотя в некоторой степени активность антимикробных пептидов может зависеть от липидного состава мембран и мембранного потенциала) [29,39]. В настоящее время разработано несколько моделей взаимодействия цитолитических пептидов с мембранами (наиболее известны модели «доска-бочка», «тороидальной поры» и «ковра»), однако между данными моделями отсутствуют четкие границы, и, возможно, пептиды способны действовать по разным механизмам. В модели «доска-бочка» молекулы пептидов образуют олигомеры, пронизывающие мембрану, что приводит к образованию поры. В модели «ковра» в результате превышения пороговой концентрации встроившегося в мембрану пептида происходит разрушение последней. В модели «тороидальных пор» в образовании поры принимают участие как молекулы пептида, так и липиды мембраны [40].

Поскольку экспериментальная часть нашей работы посвящена, в том числе, токсинам *O. takobius* и *L. tarabaevi*, в качестве примера подробнее рассмотрим цитолитические пептиды, выделенные из их яда.

#### Оксиопинины

Оксиопинины были выделены из яда *O. takobius*. Они проявляют антимикробную, инсектицидную и гемолитическую активность. Было показано, что они разрушают как биологические мембраны, так и искусственные липосомы. При взаимодействии с мембранами клеток насекомых (линия Sf9) наблюдается падение сопротивления на мембране в связи с открытием неселективных ионных каналов. Оксиопинин 1 (48 а. о.) проявляет незначительное сходство с инсектицидным пептидом понерицином L2 из яда муравья и с антимикробным пептидом дермасептином лягушки *Phyllomedusa sauvagii*. Оксиопинины 2a, 2b, 2c и 2d (37 а. о.) высокоомологичны друг другу (минимум 72% консервативных остатков) и так же проявляют некоторое сходство с понерицином L2. С помощью спектроскопии кругового дихроизма (КД) было показано, что вторичная структура оксиопинов среде, имитирующей мембранное окружение –  $\alpha$ -спираль [32].

В отличие от остальных оксиопинов, оксиопинин 4a (Oxt 4a) содержит одну дисульфидную связь на N-конце. Его С-концевая часть образует амфипатическую  $\alpha$ -спираль.

Oxt 4a проявляет мощную цитолитическую и антимикробную активность. Во многом Oxt 4a похож на защитные пептиды амфибий [37].

### Латарцины

Было обнаружено, что яд паука *L. tarabaevi* обладает высокой антимикробной активностью (минимальная ингибирующая концентрация (МИК) на *Escherichia coli* составила 0,05 мг/мл), а затем из него было выделено 7 антимикробных пептидов, объединенных в группу латарцинов (Ltc 1, Ltc 2a, Ltc 3a, Ltc 3b, Ltc 4a, Ltc 4b и Ltc 5). Содержание этих пептидов в яде составляет около 0,1 – 0,5% от суммарного белка. Ltc 3a, Ltc 3b, Ltc 4a, Ltc 4b и Ltc 5 имеют посттрансляционную модификацию – С-концевое амидирование. При анализе библиотеки кДНК из ядовитых желез того же паука было обнаружено еще 5 последовательностей латарцинов (Ltc 2b, Ltc 6a – 6c, и Ltc 7). По первичной структуре латарцины подразделяются на 7 групп. Между некоторыми группами Ltc 1 и Ltc 3a/3b наблюдается значительное сходство (53% идентичных а. о. между Ltc 1 и Ltc 3a/3b, ~50 % идентичных а. о. у Ltc 3a/3b и Ltc 4a/4b).

Латарцины положительно заряжены (pI>10, общий заряд при pH 7 составляет от +5 до +10). Так же, как и большинство других линейных антимикробных пептидов, в водном растворе латарцины не упорядочены, а в окружении, имитирующем мембранное (50% трифторэтанол), они приобретают структуру  $\alpha$ -спирали (рис. 2). Исключением являются Ltc 3a/3b, спирализующиеся уже в водном растворе. В  $\alpha$ -спиральной конформации на поверхности молекул латарцинов образуются положительно заряженный и гидрофобный кластеры.

Латарцины проявляют бактерицидный эффект как на грамположительных, так и на грамотрицательных бактериях (МИК в диапазоне 0,3 – 28 мкМ). Около половины латарцинов не показали действия на дрожжи, остальные же были активны в умеренной степени (МИК = 17 – 35 мкМ). Ltc 2a, Ltc 1 и Ltc 5 показали также гемолитический эффект. Инсектицидной активности в тестах на личинках мясной мухи латарцины не проявили, наблюдался только паралич [30].

Высокая антимикробная активность латарцинов делает их привлекательными в качестве потенциальных антибиотиков. В комбинации с коммерчески доступными антибиотиками наблюдается синергический антибактериальный эффект [41]. Однако токсическое действие латарцинов на эукариотические клетки является ограничением в использовании их в качестве лекарственных препаратов. Одним из способов преодоления этой проблемы является использование не самих пептидов, а генетических векторов, содержащих кодирующие их гены. Латарцины показали свою эффективность против внутриклеточных па-

разитов *Chlamydia trachomatis*. При трансфекции зараженных клеток линии НЕК293 плазмидами, несущими гены латарцинов 1, 2a, 3a, 4b или 5, наблюдалось значительное снижение выживаемости бактерий [42].

Разный уровень гемолитической активности среди латарцинов обусловлен разным уровнем гидрофобности их N-концевого фрагмента (приблизительно 13 первых а. о.) [35]. При точечной замене гидрофобной аминокислоты на гидрофильную в N-концевом фрагменте или удалении гидрофобных N-концевых а. о. гемолитическая активность латарцинов исчезает. Также снижаются цитотоксическое действие на лейкоциты и, в меньшей степени, на линию эритролейкемии K562. Антибактериальная активность при этом остается на прежнем уровне. Данный эффект обусловлен снижением сродства пептидов к цвиттерионным эукариотическим мембранам. В случае замены гидрофильной аминокислоты на гидрофобную наблюдается обратный эффект [43,44]. Механизм связывания Ltc 2a и его производного Ltc 2aG11A (имеющего повышенное сродство к эукариотическим мембранам) с искусственными цвиттерионными мембранами подробно изучен [45,46].

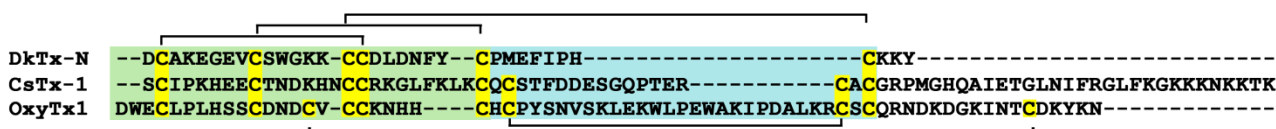
Токсическое действие Ltc 2a на эукариотические клетки происходит в несколько стадий. В результате связывания Ltc 2a с мембраной эритроцита в ней образуются поры диаметром ~2 нм, затем нарушается осмотический баланс, поры реорганизуются в большие отверстия диаметром ~13 нм. В случае клеток K562 пептид вызывает образование «вздутий» на мембране и увеличение клеточного объема, а затем клеточную смерть, при этом образуются поры диаметром ~3,7 нм, более проницаемые для анионов, чем катионов [43].

### 2.2.1.2 Дисульфид-содержащие однодоменные пептиды

Большая часть исследованных пептидов из яда пауков (более 500) содержит дисульфидные связи и проявляет нейротоксическую активность. В отличие от линейных цитолитиков, рассмотренных выше, нейротоксины специфично взаимодействуют со своими рецепторами и оказывают токсический эффект в дозах, на порядки более низких [1].

У большинства дисульфид-содержащих пептидов пауков остатки цистеина организованы в характерные мотивы, благодаря которым полипептидные цепи образуют три типа пространственной укладки [1], однако все они встречаются и среди пептидов других организмов. Цистиновый узел – самая распространенная пространственная укладка среди токсинов пауков (от англ. inhibitor cystine knot – ИСК, или ноттин от англ. knottin). Этот мотив описывается формулой:  $C^1X_{2-7}C^2X_{3-11}C^3X_{0-7}C^4X_{1-17}C^5X_{1-19}C^6$ , где X – любой а. о.,

а дисульфиды замыкаются следующим образом:  $C^1 - C^4$ ,  $C^2 - C^5$ ,  $C^3 - C^6$ . Название данного типа укладки связано с тем, что связь  $C^3 - C^6$  пронизывает кольцо, образованное двумя другими дисульфидными связями и связями основной цепи, что порождает структуру, похожую на узел. Интересно, что такой мотив первичной структуры не всегда приводит к тому, что в пространстве полипептидная цепь формирует цистиновый узел. Изучение большого количества последовательностей пептидов ноттинового типа позволило выделить т.н. основной структурный мотив (англ. principal structural motif – PSM), описываемый формулой  $C^1X_6C^2 \dots C^3C^4$ , где X – любой а. о. Обычно присутствие такого мотива в структуре пептида указывает на наличие нейротоксичности. Для последовательностей, содержащих более шести остатков цистеина, характерен дополнительный структурный мотив (англ., extra structural motif, ESM)  $C^5XC^6 \dots C^7XC^8$ , где X – любой а. о. В данном случае дисульфиды замыкаются следующим образом:  $C^1 - C^4$ ,  $C^2 - C^5$ ,  $C^3 - C^8$ ,  $C^6 - C^7$ . В последовательностях с 10 остатками цистеина образуется еще одна связь ( $C^3 - C^{10}$ ), и, скорее всего, порядок замыкания дисульфидов следующий:  $C^1 - C^5$ ,  $C^2 - C^6$ ,  $C^3 - C^{10}$ ,  $C^4 - C^9$ ,  $C^7 - C^8$  (рис. 1).



**Рисунок 1.** Сравнение аминокислотных последовательностей N-концевого домена токсина DkTx (DkTx-N) и токсинов CsTx-1 и OxyTx1, содержащих по 6, 8 и 10 остатков цистеина (выделены желтым цветом). Горизонтальными скобками над последовательностями показаны дисульфидные связи, составляющие «канонический» мотив цистинового узла, характерный для всех ноттиновых пептидов с 6, 8 или 10 остатками Cys. Скобки под последовательностями соответствуют дисульфидам, имеющим место у токсинов с 8 или 10 Cys. Зеленым цветом выделен главный структурный мотив PSM, а голубым – дополнительный структурный мотив ESM (третий остаток Cys в последовательности OxyTx1 не является его составляющей).

Экспериментально пространственная структура установлена методом ЯМР более чем для 30 пептидов (см. на рис. 2 пространственную структуру ω-агаатоксина Hv1a), и еще для многих пептидов из яда пауков она предполагается, исходя из расположения остатков цистеина вдоль полипептидной цепи. В пространстве пептиды ноттинового типа формируют плотные клубки, в структуре которых преобладают β-тяжи и β-изгибы, α-спиральные элементы встречаются крайне редко. Обычно в структуре присутствует β-шпилька, размеры которой зависят от расстояния между пятым и шестым остатками цистеина, у многих пептидов она дополняется третьим тяжем, образованным N-концевым фрагментом молекулы. При наличии четырех дисульфидов в молекуле одна из связей как правило располагается в β-шпильке. У некоторых δ-атракотоксинов из яда пауков Atrax и

*Hadronyche* (Hexathelidae) в первичной структуре наблюдается триплет ССС, и четвертый дисульфид фиксирует С-концевую область молекулы [47,48].

Другой мотив пространственной укладки, «дисульфид-направленной β-шпильки» (англ. disulfide-directed β-hairpin, DDH), наблюдается у инсектотоксинов к- (или J-) атракотоксинов (J-АСТХ). Их пространственная структура описывается следующим мотивом:  $C^1X_{5-19}C^2X_2G/PX_2C^3X_{6-19}C^4$ , где X – любой а. о., при этом образуются дисульфиды  $C^1 - C^3$ ,  $C^2 - C^4$  [49]. В пространстве DDH-пептиды формируют β-шпильки, стабилизированные двумя дисульфидными мостиками, при этом могут наблюдаться и другие элементы вторичной структуры и дисульфидные мостики. Вероятно, в эволюции мотив DDH послужил основой для появления мотива ИСК [50]. Мотив DDH характерен для хувентоксина-II из яда *Haplopelma schmidti* (Theraphosidae), при этом дисульфиды располагаются следующим образом:  $C^1 - C^3$ ,  $C^2 - C^5$ ,  $C^4 - C^6$ , и ряда других пептидов [49,50].

Мотив Кунитца характерен для пространственной укладки хувентоксина XI из яда *H. schmidti* и его гомологов. Пептид формирует N-концевую  $3_{10}$ -спираль, С-концевую α-спираль и три тяжа антипараллельной β-структуры, а дисульфидные связи у этих молекул образуются таким образом:  $C^1 - C^6$ ,  $C^2 - C^4$ ,  $C^3 - C^5$  [51].

Активность дисульфид-содержащих пептидов из яда пауков крайне разнообразна. Нейротоксины пауков способны взаимодействовать с  $K^+$ ,  $Ca^{2+}$ - и  $Na^+$ -каналами на мембранах электровозбудимых клеток, а также с механо-, термо- и хемовозбудимыми ионотропными рецепторами [1,52–54]. Большинство токсинов пауков взаимодействует с мембранными белками насекомых, однако есть и такие, которые токсичны для млекопитающих. Это связано с гомологией рецепторов у разных таксономических групп, а также с тем, что такие пептиды могут быть средством защиты или нападения на млекопитающих. Кроме того, сродство пептидов пауков к рецепторам млекопитающих может быть случайным [1].

Известно два способа взаимодействия пептидных токсинов пауков с рецепторами. В одном случае связывание происходит напрямую с белковым рецептором из водной среды, в другом же токсин первоначально взаимодействует с мембраной, а только после этого – с рецептором (мембранный катализ) [55–57].

Существуют токсины пауков, воздействующие не только на системы ионного транспорта. Например, лектино-подобный пептид хувенлектин-1 (SHLP-I) из яда *H. schmidti* вызывает агглютинацию эритроцитов [58,59], а псалмопсептооксины I и II (PcФК) из яда *Psalmopoeus cambridgei* (Theraphosidae) обладают антималярийной активностью [60,61]. Молекулярная мишень действия последних неизвестна.

Теперь подробнее остановимся на тех дисульфид-богатых токсинах, которые так или иначе будут упомянуты в разделах «Результаты» и «Обсуждение» этой диссертации.

### Токсины CSTX

CSTX-9 (68 а. о.) является одним из наиболее представленных компонентов яда паука *C. salei*. В тестах на дрозофиле CSTX-9 показал высокую инсектицидную активность ( $LD_{50} = 3,12$  нмоль/г). Аминокислотная последовательность проявляет значительное сходство (53% идентичных остатков) с токсином CSTX-1 из того же яда, который, однако, является двудоменным токсином (см. ниже). Четыре дисульфидные связи этого токсина замыкаются следующим образом:  $C^1 - C^4$ ,  $C^2 - C^5$ ,  $C^3 - C^8$  и  $C^6 - C^7$ , и молекула имеет укладку цистинового узла [62].

CSTX-13 из яда того же паука является нейротоксическим энхансером. Он на 65% повышает действие паралитического токсина CSTX-1 притом, что его собственная инсектицидная активность крайне низкая (50% смертность среди дрозофил наблюдалась при введении 0,05 мкл 309 мкМ раствора CSTX-13 на одну дрозофилу, что в 49 раз выше, чем в случае с CSTX-1 и в 1,5 раза – чем с CSTX-9). Интересно, что содержание CSTX-13 в яде в 7 – 8 раз и в 2,8 раз ниже, чем CSTX-1 и CSTX-9, соответственно. Структура этого токсина необычна – он состоит из двух полипептидных цепей (А и В), соединенных дисульфидными связями. Цепь А состоит из 34 а. о., а цепь В – из 29 а. о., она амидирована на С-конце. Разделение полипептидной цепи токсина на две при сохранении дисульфидного мотива является посттрансляционной модификацией [63].

### Окситоксины

Окситоксины ОхуТх1 и ОхуТх2 из яда паука *Oxyopes lineatus* в низких дозах способны парализовать насекомых (на личинках *Spodoptera litura*  $LD_{50} \sim 5.0$  нмоль/г и  $\sim 6.2$  нмоль/г), их содержание составляет 3.9% и 1.8%, соответственно [64].

Изначально ОхуТх1 был выделен из яда родственного паука *O. takobius* [32], однако ОхуТх2 в его яде не обнаруживается. Длина ОхуТх1 и ОхуТх2 составляет 69 и 55 а. о., соответственно. Окситоксины содержат по 5 внутримолекулярных дисульфидных связей, их С-концевые остатки амидированы. ОхуТх1 является антагонистом потенциал-чувствительных  $Ca^{2+}$ -каналов P/Q-, N- и L- типов, при концентрации 300 нМ он снижает их активность на  $\sim 53\%$ ,  $\sim 65\%$  и  $\sim 63\%$ , соответственно. Однако он не проявляет специфичности к какому-то определенному типу  $Ca^{2+}$ -каналов. Способности связываться и модулировать активность потенциал-зависимых  $Na^+$ -каналов замечено не было. Окситокси-

ны проявляют низкое сходство с другими известными блокаторами  $\text{Ca}^{2+}$ -каналов из яда членистоногих [32,64].

## 2.2.2 Двудоменные токсины

Выше мы рассмотрели разнообразие «классических» токсинов пауков, укладываемые, однако, в рамки нескольких типов пространственных структур. Двудоменные (или модульные) токсины дополнительно повышают разнообразие компонентов яда пауков за счет комбинации свойств «обычных» токсинов в одной молекуле. У двудоменных пептидов пространственная структура каждого из модулей относится к одному из типичных структурных классов токсинов. Так, на данный момент известны молекулы, состоящие из двух линейных («AMP+AMP») или ноттиновых доменов («ICK+ICK»), а также из ноттинового и линейного доменов («ICK+AMP»). В некоторых случаях удается обнаружить гомологию между аминокислотными последовательностями отдельных доменов и известными однодоменными токсинами. Далее мы подробнее рассмотрим каждое из известных на данный момент семейств модульных токсинов.

### 2.2.2.1 Линейные пептиды («AMP+AMP»)

Цитоинсектотоксины (ЦИТ) являются главными компонентами яда паука *L. tarabaevi*: в сумме на них приходится около 20% яда. Непосредственно из яда было выделено 8 токсинов (ЦИТ 1a – 1h), кроме того, в библиотеке кДНК из ядовитых желез паука было обнаружено еще 8 последовательностей, кодирующих пептиды, гомологичные цитоинсектотоксинам [65].

Цитоинсектотоксины «составлены» из двух коротких пептидов, длиной около 30 а. о., соединенных коротким линкером. N-концевой (1 – 34 а. о.) и C-концевой (39 – 69 а. о.) фрагменты по своим свойствам очень похожи на короткие  $\alpha$ -спиральные антимикробные пептиды. Они обладают высоким положительным зарядом при pH 7 (+9 и +7, соответственно) и склонны к формированию  $\alpha$ -спиралей, на поверхности которых располагаются гидрофобные и положительно заряженные кластеры. Последовательность линкера – EEAQ или EEAE (35 – 38 а. о.), что напоминает типичный мотив процессинга белков-предшественников из ядовитых желез паука (т.н. PQM). Наиболее распространенная последовательность сайта PQM у *L. tarabaevi* – EEAR. Вероятно, в эволюции в сайте процессинга двойного предшественника произошла мутация ключевого остатка аргинина, что нарушило созревание пропептида, и образовался «сдвоенный» пептид [65].



СІТ 1а обладает мощной антимикробной активностью как на грамположительных, так и на грамотрицательных бактериях (МИК находится в диапазоне 0,5 – 3,8 мкМ). Кроме того, СІТ 1а проявил цитолитическое действие на эукариотические клетки (полумаксимальная эффективная концентрация (EC<sub>50</sub>) в диапазоне 1 – 6 мкМ). СІТ 1а токсичен для насекомых в низких дозах (ЛД<sub>50</sub> составила ~0,6 нмоль/г и ~2,5 нмоль/г для мясной мухи и ее личинки, соответственно, ~63,3 нмоль/г для тараканов). В сравнении с описанным выше «обычным» линейным токсином Ltc 2a из яда того же паука, цитоинсектотоксины на много более токсичны для насекомых и их клеток. Однако в отношении млекопитающих (мышей) их эффекты одинаковы.

Была показана антихламидийная активность СІТ 1а на ранних стадиях жизненного цикла патогена при трансфекции зараженных клеток НЕК 293 вектором, несущим ген, кодирующий этот пептид [42,66].

#### **2.2.2.2 Дисульфид-содержащие пептиды типа «ІСК+ІСК»**

Из яда двух разных пауков были выделены пептидные токсины, состоящие из двух доменов ноттинового типа – SpTx из яда *Cheiracanthium punctorium* и DkTx из яда *H. schmidtii*.

##### Токсин DkTx

DkTx –компонент яда *H. schmidtii*, вызывающий у жертвы мучительную и пролонгированную боль. Этот токсин специфично и необратимо активирует капсаицин- и термочувствительные каналы TRPV1 млекопитающих. На другие типы каналов (TRP, лиганд-активируемые (5-HT<sub>3</sub>R-A и P2X<sub>2</sub>) и потенциал-зависимые (K<sub>v</sub>1.2, K<sub>v</sub>2.1, K<sub>v</sub>4.3, Ca<sub>v</sub>1.2, Ca<sub>v</sub>3.3 и Na<sub>v</sub>1.7)), многие из которых экспрессируются на чувствительных нейронах, DkTx не действует. Более того, он не активен по отношению к мышам, не экспрессирующим канал TRPV1 [67].

DkTx (79 а. о.) состоит из двух гомологичных ІСК доменов (~67% идентичных а. о.), содержащих по 3 дисульфидные связи. Среди других известных пептидов ноттинового типа нет гомологов DkTx [67].

Высокая эффективность связывания DkTx с рецептором проистекает из его «двухвалентности», подобно антителам. При разделении этого двудоменного токсина на пептиды K1 и K2, соответствующие доменам, наблюдается значительное снижение активности (EC<sub>50</sub> повышается в 50 и в 5 раз для N- и C-концевого доменов, соответственно). Кроме того, их действие не является необратимым. В экспериментах с использованием смеси

пептидов K1 и K2 восстановление активности не происходит ( $EC_{50}$  для смеси равно  $EC_{50}$  для более активного пептида K2). Однако искусственный пептид K1-K1, представляющий собой удвоенный пептид K1, обладает более высокой активностью, чем K1 и, подобно токсину DkTx, вызывает необратимую активацию канала TRPV1. Таким образом, для проявления активности DkTx в полной мере необходимо присутствие двух доменов в составе одной молекулы, которые связываются с каналом, удерживая его в открытом состоянии [67].

При получении рекомбинантного токсина DkTx наблюдались многочисленные трудности. Пептид получали в системе *E. coli* с использованием разных вспомогательных последовательностей и белков-носителей (полигистидиновая последовательность, кетостероид изомеразы (КСИ) и тиоредоксин (Trx)). Наибольший выход наблюдался при получении токсина в нерастворимом виде в составе химеры с КСИ (0,4 – 0,8 мг/л культуры). В этих условиях не происходило замыкание дисульфидных связей, поэтому были подобраны условия для окисления тиольных групп остатков цистеина. В окислительном буфере, ранее применявшемся для фолдинга конотоксинов, выход был крайне низким (около 2%), однако добавление детергентов значительно повысило выход пептида в нативной конформации (до 24 – 30%). Пептиды K1 и K2, полученные путем химического синтеза, так же подвергались аналогичной процедуре рефолдинга, выход составил ~65% и ~30%, соответственно [68].

### Токсин CrTx 1

CrTx 1 – главный компонент яда паука *C. punctoriium*. Этот токсин производит летальное действие на личинок мясной мухи *S. carnaria* ( $LD_{50} = \sim 0,6$  нмоль/г). CrTx 1 (134 а. о.) состоит из двух гомологичных ноттиновых доменов (~28% идентичных а. о.), содержащих по 4 дисульфида. На С-конце пептид амидирован. Аминокислотная последовательность проявляет сходство с последовательностями токсинов LSTX паука *Lycosa singoriensis*, полученными в результате трансляции последовательностей кДНК из его ядовитых желез, а также с токсинами CSTX из яда *C. salei*. Интересно, что С-концевой домен CrTx 1 проявляет больше сходства с CSTX-1 и CSTX-9, а N-концевой – с CSTX-13. В С-концевой части каждого из доменов имеется линейный фрагмент, склонный к формированию  $\alpha$ -спирали [69].

Эволюционное появление CrTx 1, вероятно, как и в случае цитоинсектотоксинов, связано с мутацией в сайте процессинга белка-предшественника. Последовательность линкера между двумя доменами – RKPV, однако при сдвиге рамки считывания в коди-

рующей последовательности ДНК на одну позицию она превращается в AETR – типичную последовательность RQM [69].

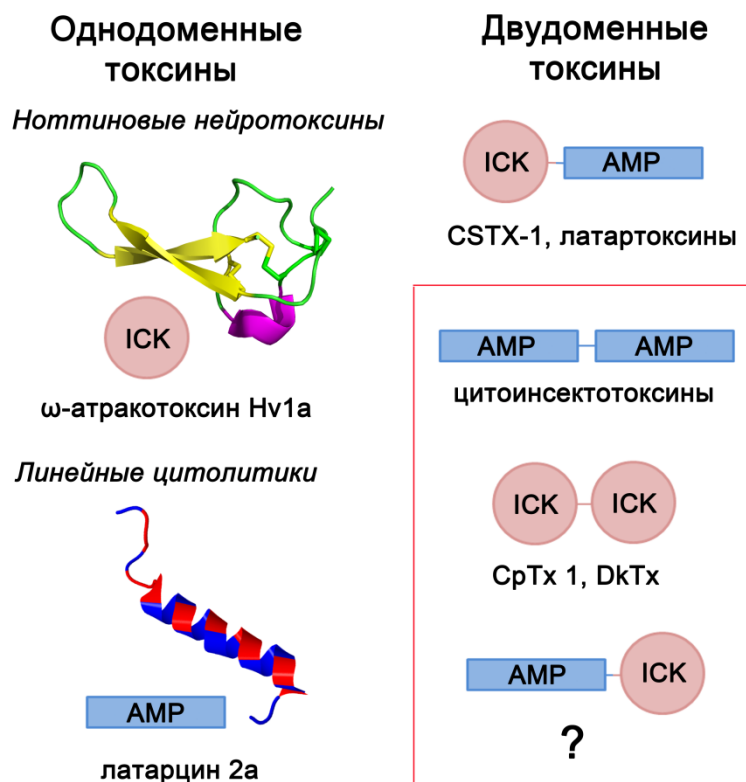
Механизм действия CrTx 1 активно изучался. Было установлено, что токсин проявляет цитолитический эффект на клетки насекомых Sf9 в зависимости от концентрации (5% при 0,5 мкМ и 20% при 10 мкМ). Повреждения мембраны, вызванные действием пептида, сопровождаются значительными изменениями в клеточной морфологии, в том числе, образованием «пузырей» на мембране. Паралитическое действие было изучено на нервно-мышечном препарате мухи. При концентрации 50 – 200 нМ наблюдалась необратимая деполяризация мышечных волокон, приводившая к мышечному сокращению при более высоких концентрациях. Однако, как спонтанные, так и вызванные постсинаптические потенциалы оставались неизменными. Аналогичные результаты были получены на нервно-мышечном препарате лягушки. Ингибирование работы нескольких типов мембранных ион-транспортных белков (потенциал-зависимых Na<sup>+</sup>- и Ca<sup>2+</sup>-каналов, ENaC и Na<sup>+</sup>/K<sup>+</sup>-АТФазы) не привело к изменению эффекта CrTx 1 на нервно-мышечном препарате. Такие результаты говорят в пользу того, что мишенью действия CrTx 1 является цитоплазматическая мембрана, а не какой-либо мембранный белок. Это подтверждается также тем, что токсин способен разрушать искусственные мембраны [69].

### 2.2.2.3 Токсины типа «ICK+AMP»

Токсин CSTX-1 (рис. 2) является мажорным компонентом яда *C. salei*. Он токсичен в отношении насекомых (на мухах семейства Calliphoridae ЛД<sub>50</sub>=1,26 нмоль/г, на дрозофилах 0,35 нмоль/г) и млекопитающих. В первичной структуре выделяется два модуля: N-концевой ноттиновый и C-концевой линейный, богатый остатками лизина. Дальнейшие эксперименты показали важность присутствия обоих доменов в структуре токсина. Было установлено, что CSTX-1 блокирует два типа кальциевых каналов насекомых, экспрессирующихся в дорсальных непарных медианных нейронах (ДНМ): каналы, активируемые средним и низким потенциалом (M-LVA) и активируемые высоким потенциалом (HVA). В тестах на нейронах крысы было показано, что CSTX-1 является селективным блокатором кальциевых каналов L-типа [70,71]. Таким образом, CSTX-1 обладает нейротоксической функцией.

Токсин CSTX-2а является фрагментом CSTX-1, образующимся после удаления C-концевого линейного лизин-богатого фрагмента (CT1-short, 13 а. о.). Для дрозофилы токсичность CSTX-2а в 7 раз ниже токсичности CSTX-1, а при дополнительном удалении по-

ложительно заряженного остатка Arg 61 (в результате чего образуется токсин CSTX-2b) уровень токсичности уменьшается в 190 раз. Токсин CSTX-9, не имеющий цитолитического домена, в 9 раз менее активен против дрозофилы. Активность смеси CSTX-2b и пептида CT1-short была ниже, чем у токсина CSTX-1. CT1-short был менее активен, чем более длинный (30 а. о.) С-концевой фрагмент CT1-long, который проявлял инсектицидную



**Рисунок 2. Разнообразие пептидных токсинов пауков.** Приведены известные пространственные структуры. Левая панель: однодоменные токсины. Пространственная структура ω-атрактоксина Hv1a (PDB 1AXH), нейротоксина с ноттиновым типом пространственной укладки. Показаны боковые цепи остатков цистеина, образующие 3 дисульфидных связи. Цветом выделены элементы вторичной структуры: желтый - β-слой, сиреневый - α-спираль. Пространственная структура латарцина 2a (PDB 2PCO), линейного цитолитика со структурой амфипатической α-спирали. Красным цветом выделены гидрофобные остатки, синим - положительно заряженные. Правая панель: схематичная структура двудоменных токсинов. Рамкой выделены те группы, которым посвящена экспериментальная часть настоящей работы.

активность на дрозофиле в дозе ЛД<sub>50</sub>=82,64 нмоль/г. Кроме того, как полноразмерный токсин CSTX-1, так и CT1-long способны увеличивать неспецифическую ионную проводимость мембран ооцитов (в концентрациях 0,05 – 0,5 мкМ и 5 мкМ, соответственно) [72]. Таким образом, CSTX-1 сочетает в себе свойства нейротоксина и цитолитика, причем первое обеспечивается N-концевым ноттиновым доменом, а второе – С-концевым линейным. Два домена усиливают активность друг друга, причем этот эффект теряется при

«разделении» CSTX-1 на два независимых пептида, соответствующих каждому из доменов. Таким образом, был предложен следующий механизм действия CSTX-1. Катионный С-концевой фрагмент CSTX-1 может служить якорем, взаимодействующим с мембраной, а ингибирование ионных каналов является результатом прямого взаимодействия ноттинового фрагмента с мишенью [73].

Латартоксины (LtTx) из яда *L. tarabaevi* также относятся к рассматриваемому типу двудоменных токсинов [74]. Их аминокислотная последовательность проявляет значительное сходство с токсинами CSTX. Латартоксины образуют две группы (LtTx-1 и LtTx-2), интересной особенностью двух представителей группы LtTx2 является дополнительная дисульфидная связь, соединяющая линейный домен с ноттиновым. Латартоксины проявляют инсектицидную активность против личинок мясной мухи (*S. carnaria*) и сверчков (*Acheta domestica*) (ЛД<sub>50</sub> ~7,3 нмоль/г). В электрофизиологических экспериментах на нервно-мышечном препарате личинки *Calliphora vicina* была показана цитолитическая активность латартоксинов, кроме того, поскольку наблюдалось уменьшение амплитуды вызванных постсинаптических токов, не исключен модулирующий эффект на ионные каналы. Пептид LtTx-1a С, соответствующий С-концевому домену LtTx-1a (21 а. о.), не показал инсектицидной активности вплоть до дозы 88 нмоль/г, однако (в отличие от полного размерного токсина) проявил антибактериальное действие на грамположительные бактерии *A. globiformis* (3,1 мкМ) и *B. subtilis* (6,2 мкМ). Способность С-концевого домена взаимодействовать с мембранами была напрямую показана с помощью его соосаждения с липосомами.

Итак, мы рассмотрели все известные типы пептидных токсинов пауков. На рис. 2 приведена схема, резюмирующая информацию о наиболее распространенных структурно-функциональных типах токсинов: однодоменные цитолитики и нейротоксины токсины, а также их «химеры» – двудоменные токсины с разным сочетанием типов пространственной укладки.

### **3. Организация белков-предшественников токсинов**

подавляющее большинство животных токсинов полипептидной природы синтезируется в виде белков-предшественников. Токсины пауков, скорпионов, моллюсков конусов, морских анемонов и змей являются секретируемыми молекулами, поэтому их предше-

ственники содержат N-концевой сигнальный пептид (или лидерный, препептид), обеспечивающий котрансляционный перенос белка в эндоплазматический ретикулум [1,75–84]. Однако у пауков, конусов и морских анемонов между сигнальным и зрелыми пептидами располагается еще и препептид, который, скорее всего, необходим для правильного фолдинга токсинов и для ингибирования их активности до тех пор, пока они не будут секретированы [1,80,81,85]. Отщепление препептидов происходит посттрансляционно с помощью ограниченного протеолиза по специфическим сайтам процессинга [86].

Теперь мы более подробно рассмотрим биосинтез токсинов пауков.  $\alpha$ -Латротоксин и  $\alpha$ - и  $\delta$ -латроинсектотоксины синтезируются в виде неактивных белков-предшественников, в процессе созревания которых на N-конце полипептидной цепи по специфическому сайту процессинга происходит отщепление пропоследовательности. Кроме того, на C-конце белка происходит отщепление препептида длиной около 200 а. о. [22,24,87,88]. Сфингомиелиназа D также проходит через стадию предшественника, включающего в себя сигнальный пептид, препептид и зрелую цепь [89].

Для белков-предшественников коротких пептидных токсинов характерны специфические мотивы PQM (англ. processing quadruplet motif – четырехчленный мотив процессинга) и iPQM (англ. inverted PQM – обратный PQM), не свойственные другим пептидам. Мотив PQM можно обозначить как  $X_1X_2X_3R$ , где любой  $X_n = E$ . Мотив iPQM является обратным относительно PQM ( $RX_1X_2X_3$ , где любой  $X_n = E$ ) [86,90]. Например, у токсина CrTx 1 сайт процессинга DESR, у цитоинсектотоксинов – EEAR. После сайта процессинга в предшественнике находится зрелая цепь. Среди предшественников линейных цитолитических пептидов встречаются такие, у которых сайт процессинга повторяется более одного раза, что приводит к образованию нескольких зрелых цепей. Так, у *L. tarabaevi* в одном предшественнике могут быть до 5 зрелых цепей, разделенных несколькими вариантами мотива PQM (EAGR, DAGR, EEAR) [30]. Такие предшественники называются сложными.

Кроме отщепления препропоследовательностей многие токсины пауков подвергаются дополнительной посттрансляционной модификации – C-концевому амидированию, при этом происходит отщепление C-концевого остатка глицина и одновременное образование амида на предыдущем аминокислотном остатке [1].

## 4. Гены токсинов животных

### 4.1 Гены токсинов пауков

К настоящему времени не известно ни одной полной последовательности генома паука, поэтому наши данные о генах, кодирующих токсины пауков, ограничиваются отдельными последовательностями, принадлежащим нескольким видам. Здесь мы рассмотрим структуру генов, кодирующих как короткие пептидные нейротоксины, так и относительно крупные белки с ферментативной активностью. Гены коротких линейных цитолитиков изучены не были.

#### 4.1.1 Гены пептидных токсинов

У примитивного аранеоморфного паука *Diguetia canities* была установлена последовательность гена, кодирующего токсин DTX 9.2. Этот инсектицидный пептид содержит 6 остатков цистеина, образующих мотив цистинового узла. Длина транскрибируемого фрагмента составляет 5,5 тысяч пар нуклеотидов (т. п. н.), он содержит 4 интрона, причем первый из них располагается в 5'-нетранслируемой области (НТО), а стартовый кодон – во втором экзоне. Длины экзонов составляют 49 пар нуклеотидов (п. н.) (точка начала транскрипции не была установлена), 77 п. н., 47 п. н., 27 п. н. и 226 п. н. Длины интронов составляют 0,89 – 1,64 т. п. н. Расположение экзонов слабо коррелирует со структурой белка-предшественника: точки сращивания экзонов находятся в последовательностях, кодирующих пропептид, сайт процессинга и зрелый пептид, а также, как уже было сказано, в 5'-НТО. Сплайсинг всех интронов происходит по сайтам /GT и AG/. Вероятно, такая структура гена способствует образованию новых генов путем тасования интронов и, тем самым, повышению разнообразия токсинов. В последовательности хромосомной ДНК в положении –19 находится ТАТА-бокс (ТАТААА), а в положениях –57 и –97 располагаются палиндромные последовательности САСАТGTG и ААGСААТCGСТТ, соответственно, которые, возможно, являются дополнительными промоторными элементами [91].

Описанный выше ген является единственным известным на данный момент примером интрон-богатого гена, кодирующего пептидный токсин паука. У мигаломорфных пауков в последовательности генов, кодирующих δ-гексатоксин *Hadronyche versuta* [92], а также токсины нескольких видов тарантулов (род *Haplopelma*) интронов обнаружено не было. Для паука *H. hainana* была установлена структура генов, кодирующих 52 токсина, принадлежащих к 7 надсемействам [93]. Для токсинов этого паука свойственны 3 мотива расположения остатков цистеина – ICK, DDH и мотив Кунитца. У *H. schmidti* установлена

последовательность гена, кодирующего токсин SHT-I [94], а также последовательности 19 генов, кодирующих хувентоксины (HWTX) трех надсемейств (с мотивами ICK и Кунитца). Интересно, что у этого паука были также обнаружены последовательности, гомологичные генам токсинов из надсемейства HWTX-XI, однако не способные кодировать белки-предшественники токсинов из-за протяженной делеции. Предполагается, что фрагмент экзона гена HWTX-XI образовался из интронной последовательности в результате нарушения механизма сплайсинга [95].

#### **4.1.2 Гены белковых токсинов**

Большинство известных генов коротких пептидных токсинов не имеют интронов. Разнообразие крупных белков-компонентов яда паука гораздо ниже, чем пептидных токсинов, однако среди них наблюдается разнообразие в экзон-интронной структуре. У пауков рода *Loxosceles* гены, кодирующих сфингмиелиназы D, обладают длиной около 6500 п. н. и содержат, по крайней мере, 6 экзонов и 5 интронов [89]. Однако у каракуртов (род *Latrodectus*) гены латротоксинов являются безинтронными и обладают длиной около 3600 – 5300 п. н. [88].

#### **4.2 Гены токсинов скорпионов**

У скорпионов гены токсинов изучены намного более полно, чем у пауков (табл. 1). На примере скорпиона *Leiurus quinquestriatus* было впервые показано, что в геноме одной особи гены токсинов образуют семейства, и то разнообразие, которое до этого наблюдалось на уровне кДНК, полученной из ядовитых желез разных особей, сохраняется в индивидуальном геноме [96]. Недавно была установлена полная последовательность генома *Mesobuthus martensii*. В геноме этого скорпиона обнаружено 116 генов нейротоксинов с известной функцией и их гомологов с предполагаемой функцией, в том числе 61 ген токсинов, мишенью которых являются натриевые каналы, 46 генов токсинов калиевых каналов, 5 – хлоридных каналов и 4 – рианодиновых рецепторов. Среди них 45 кодирует ранее не описанные токсины, а 109 экспрессируются в ядовитых железах. 51 ген нейротоксинов образует 17 кластеров, в каждом из которых содержатся тандемно дублицировавшиеся гены, принадлежащие к одному семейству и обладающих похожей структурой. Интересно, что подобные особенности были обнаружены и в локусе генов защитных пептидов деффензинов, что говорит о сходстве механизмов их эволюции [97].

На данный момент подробно описана организация многих генов нейротоксинов, блокирующих ионные каналы разного типа (см. табл. 1). Все изученные гены токсинов



скорпионов содержат по крайней мере один интрон, располагающийся в области, кодирующей сигнальный пептид. В большинстве случаев он располагается между первым и вторым нуклеотидами глицинового кодона, четвертого от 3'-конца последовательности, кодирующей сигнальный пептид. Все описанные интронные последовательности АТ-богаты (у большинства интронов суммарное содержание нуклеотидов А и Т превышает 70%), начинаются с GT и заканчиваются на AG. 5'-Донорный сайт сплайсинга консервативен, у большинства генов его последовательность G/gtaag (заглавная буква относится к последовательности экзона, строчные – интрона), в то время как 3'-акцепторный сайт сплайсинга более вариабельный. Несмотря на перечисленные общие свойства, последовательности интронов сильно различаются по длине и не проявляют значительного сходства, даже если они входят в состав генов, кодирующих гомологичные токсины. Теперь мы подробнее остановимся на генах токсинов разных видов скорпионов.

Помимо интрона в последовательности, кодирующей сигнальный пептид, некоторые гены токсинов *M. martensii* содержат дополнительный интрон, причем в большинстве случаев он располагается в 5'-НТО. Только у гена, кодирующего токсин VmCa1, интрон находится в области, соответствующей зрелому пептиду [98]. Интересный случай вариабельности длины интрона был обнаружен в случае гена токсина VmKBT. Он кодируется двумя генами, длина интронов которых различается в 6,8 раз (остальные характеристики интронов совпадают). Предполагается, что разная длина интронов связана с механизмами регуляции экспрессии генов: ген с длинным интроном экспрессируется конститутивно, в то время как транскрипция с более короткого гена запускается в ответ на какие-либо стимулы [99]. Интроны, находящиеся в 5'-НТО генов токсинов, действующих на K<sup>+</sup>-каналы, могут влиять на уровень их экспрессии, как было показано на примере VmKcug1a, VmKcug2 и VmKcugx. [100].

У скорпионов рода *Androctonus* были установлены последовательности генов, кодирующих токсины AmmVIII, AmmVIII-related, AmmV, AmmV-related, KTX1, KTX2, AaNI'. Интересно, что хотя токсины и гомологичны друг другу, длина интронов в их генах различается, и последовательности интронов не проявляют сходства друг с другом [101–103]. У гена токсина AaNI' был подробно изучен промотор. Он содержит сайты связывания регуляторных белков MAT- $\alpha$ 2 (ATTTACAT), Pit-1 (WTATTCAT) и IEF1 (GCCATCTG). Кроме того, был обнаружен ТАТА-бокс в положении -29 относительно стартового кодона и два ССААТ-боксы в положениях -140 и -90. Несмотря на отсутствие значимого сходства последовательностей генов AaNI' и представителя семейства скорпионов (семейство двудоменных токсинов скорпионов) опискорпина 3 (Ori-3) [104], их про-

моторы имеют сходную структуру: они включают в себя два модуля (А и В), разделенные спейсером длиной 200 п. н. Модуль А содержит четыре сайта связывания факторов транскрипции ([Oct-1][GATA-1][C/EBPa][Oct-1]), а модуль В – пять сайтов ([C/EBPa][Oct-1][SRF][C/EBPa][Oct-1]). Кроме того, в промоторной последовательности гена Ori-3 были найдены несколько сайтов связывания других факторов транскрипции, ТАТА-бокс, точка начала транскрипции. Интрон-экзонная организация генов Ori-3 и Hs-1 (другого представителя семейства скорпинов) несколько отличается от большинства других генов токсинов. У этих генов интроны располагаются в последовательностях, кодирующих зрелые пептиды, в их 5'-области. Последовательности этих интронов проявляют низкий уровень сходства друг с другом (28%), хотя ближе к сайтам сплайсинга сходство сильно возрастает (до ~80%), что говорит о функциональной важности этих участков. У Ori-3 был также найден интрон в 5'-НТО, не исключено, что Hs-1 тоже имеет еще один интрон в аналогичной области [104,105]. Гены опискорпинов в 3'-НТО содержит четыре альтернативных сигнала полиаденилирования, из-за чего у транскриптов, синтезирующихся с этих генов, наблюдается разная длина 3'-НТО и, по-видимому, разный уровень стабильности, что является механизмом регуляции экспрессии гена [104].

Интроны способствуют повышению разнообразия токсинов. Например, транскрипт VmKT-и *M. martensii* является результатом транс-сплайсинга. VmKT-и является химерой между транскриптами VmKT' и VmKu. Было показано, что VmKT-и не представляет собой единую транскрипционную единицу в геноме скорпиона, однако, VmKT' имеет соответствующий ген. Ген VmKT' содержит интрон (509 п. н.) в области, кодирующей C-концевую часть сигнального пептида. Точка сращивания фрагментов VmKT' и VmKu располагается в области 5' сайта сплайсинга интрона. Вероятно, транскрипт VmKT-и образуется в результате рекомбинации между VmKT' и VmKu на уровне мРНК [106]. У скорпиона *Lychas mucronatus* транскрипты LmTxLP11 и LmVP1.1 обладают идентичными 5'-концевыми последовательностями. LmTxLP11 является удлиненной версией LmVP1.1. Оба транскрипта являются продуктами одного гена и образуются в результате альтернативного сплайсинга. Ген, кодирующий LmTxLP11 и LmVP1.1 содержит четыре интрона. Транскрипт LmVP1.1 образуется в результате использования короткого четвертого экзона, кодирующего только одну аминокислоту и стоп-сигнал и содержащего альтернативный сигнал полиаденилирования. Транскрипт LmTxLP11 длиннее, так как его 3'-концевая область считывается со значительно более длинного пятого экзона. Интроны располагаются во фрагментах, кодирующих сигнальный пептид (ближе к его C-концу) и зрелые пептиды. Эти интроны обладают всеми типичными свойствами: они АТ-богаты (суммарное содер-

жание А+Т составляет 73%, 67% и 68%, соответственно), начинаются с GT и заканчиваются на AG. Последовательности 5'-донорных сайтов сплайсинга G/gtaag, T/gtaag и A/gtacg, соответственно. В первых двух интронах 3'-акцепторные сайты сплайсинга tacag/T и tccag/G, а в третьем имеется два альтернативных сайта: tacag/G (используемый транскриптом LmVP1.1) и cgcag/G (используемый транскриптом LmTxLP11). Длина интронов составляет 88, 737 и 406 п. н. [107].

### 4.3 Гены токсинов моллюсков конусов

Морские брюхоногие моллюски рода *Conus* производят пептидные токсины с разнообразными цистеиновыми мотивами [108]. Структура известных генов конотоксинов различна.

$\alpha$ -Конотоксины Mr1.3, Pu1.1, Lp1.1, Lp1.4, Lp1.6a и Ac1.1b из пяти видов моллюсков-конусов (*C. marmoreus*, *C. leopardus*, *C. achatinus*, *C. pulicarius* и *C. quecrinus*) кодируются генами, состоящими из двух экзонов, разделенными одним интроном. Длина интронов более 700 п. н. (от 791 п. н. у гена Lp1.1 до 1441 п. н. у гена Qc aL-2). Экзон 1 кодирует 5'-НТО, сигнальный пептид и первые несколько аминокислот пропептида, экзон 2 – остальную часть пропептида, зрелый пептид и 3'-НТО. У всех генов интроны располагаются в одной и той же области – в кодоне, кодирующем 6-й а. о. пропептида (в большинстве случаев, аланин). Так же, как интроны генов токсинов скорпионов, интроны *Conus* начинаются и заканчиваются на gt/ag, однако содержание остатков А и Т гораздо ниже (~60%). Во всех генах, кроме Lp1.1, последовательности интронов рядом с донорным и акцепторным сайтами сплайсинга консервативны. В 3'-области эти интроны особенно консервативны: последние 14 нуклеотидов полностью идентичны (-ggttctgggtccag-), а им предшествуют отрезки динуклеотидных повторов GT разной длины. Внутри одного вида интроны более консервативны, но между интронами разных видов наблюдается значительная разница в длинах [109].

Организация генов, кодирующих надсемейство М-конотоксинов (из *C. terebra*, *C. textile* и *C. striatus*) отличается от описанной выше. Эти гены состоят из трех экзонов и двух интронов, причем экзон 1 кодирует сигнальный пептид, экзон 2 – пропептид, экзон 3 – зрелый пептид. Предполагается, что такая структура гена отражается на механизмах эволюции конотоксинов (см. ниже). Для всех интронов характерны сайты сплайсинга GT и AG. Длина каждого из интронов более 3000 п. н. Интересно, что был зафиксирован случай альтернативного сплайсинга, когда транскрипты состояли только из экзонов 1 и 3 [110].

	Токсин	N	Длина, п. о.			% (A+T)			5'сплайс сайт			3' сплайс сайт			
			Расп-е			Расп-е			Расп-е			Расп-е			
			1	2	3	1	2	3	1	2	3	1	2	3	
<i>Mesobutismartensii</i>	BmKBy	1	-	1312	-	-	77	-	-	G/gtaag	-	-	tatag/G	-	[99]
	BmKBT_a	1	-	1529	-	-	76	-	-	G/gtaag	-	-	tatag/G	-	[99]
	BmKBT_b	1	-	225	-	-	79	-	-	G/gtaag	-	-	tatag/G	-	[99]
	BmK αIV	1	-	503	-	-	62	-	-	G/gtaag	-	-	tacag/G	-	[111]
	BmKcugx	1	-	76	-	-	79	-	-	G/gtaag	-	-	ataag/C	-	[100]
	BmKcug1a	2	80	81	-	85	77	-	G/gtaag	G/gtaag	-	tacag/T	aacag/G	-	[100]
	BmKcug2	2	66	87	-	85	75	-	G/gtaag	G/gtgtg	-	ttaag/A	tccag/G	-	[100]
	BmCa1	2	-	72	1074	-	83	71	-	G/gtaag	G/gtaag	-	ttcag/A	tatag/T	[98]
	Mesotoxin	2	228	113	-	82	82	-	-	?	-	-	?	-	[112]
	Martentoxin	2	83	91	-	85	78	-	A/gtaag	A/gtaag	-	taaag/C	tctag/G	-	[113]
Bm-12	1	-	93	-	-	84	-	-	?	-	-	?	-	[114]	
<i>A. mauretanicus</i>	Amm VIII	1	-	166	-	-	74	-	-	G/gtaag	-	-	tacag/G	-	[101]
	Amm VIII-related	1	-	436	-	-	74	-	-	G/gtaag	-	-	tacag/G	-	[101]
	Amm V	1	-	567	-	-	67	-	-	G/gtaag	-	-	tacag/G	-	[115]
	Amm V-related	1	-	413	-	-	75	-	-	T/gacag	-	-	tacag/G	-	[115]
<i>A. australis</i>	KTX2	1	-	87	-	-	86	-	-	A/gtaat	-	-	ttag/T	-	[102]
	KTX1	1	-	87	-	-	78	-	-	A/gtaat	-	-	tctag/T	-	[116]
	AaH I'	1	-	425	-	-	75	-	-	G/gtaag	-	-	tacag/G	-	[103]
<i>T. s.</i>	Ts γ	1	-	475	-	-	73	-	-	G/gtaag	-	-	catag/G	-	[117]
	Ts IV-5	1	-	347	-	-	53	-	-	G/gtaag	-	-	ttag/G	-	[118]
<i>Leivirus quin-questriatus</i>	Lqha-6a	1	-	396	-	-	>80	-	-	G/gtaag	-	-	tacag/G	-	[96]
	Lqha-6b	1	-	399	-	-	>80	-	-	G/gtaag	-	-	tacag/G	-	[96]
	Lqha-6c	1	-	396	-	-	>80	-	-	G/gtaag	-	-	tacag/G	-	[96]
	LqhIT2-53	1	-	307	-	-	>80	-	-	G/gtaag	-	-	tatag /A	-	[96]
	LqhIT2-13	1	-	309	-	-	>80	-	-	G/gtaaa	-	-	tatag /A	-	[96]

	LqhIT1	1	-	578	-	-	>80	-	-	G/gtaag	-	-	caaag/G	-	[96]
	CITx	1	-	95	-	-	>80	-	-	G/gtaag	-	-	tgtag/C	-	[96]
	ChTx	1	-	125	-	-	>80	-	-	G/gtaag	-	-	ttag/T	-	[96]
<i>H. O. c.</i>	Opiscorpine 3	2	1016	-	1160	~50%	-	~50%	G/gtaag	-	A/gtaag	ttcag/G	-	tccag/A	[104]
<i>H. l.</i>	Hs-1	1	???		1073						A/gtaag			Tccag/A	[105]

**Таблица 1. Интрон-экзонная организация известных генов токсинов скорпионов.** *N* – количество интронов в гене; пронумерованные столбцы отражают положение интрона в гене: 1 – 5'-НТО, 2 – сигнальный пептид, 3 – зрелый пептид. *A. australis* – *Androctonus australis*, *O. carinatus* – *Opisththalmus carinatus*, *A. mauretanicus* – *Androctonus mauretanicus*, *H. l.* – *Heterometrus laoticus*. *T. s.* – *Tityus serrulatus*. ? – в литературных источниках соответствующие данные не приводятся, ??? – возможно наличие дополнительного интрона.

#### 4.4 Гены токсинов змей

Большинство коротких пептидных токсинов змей, относящихся к трехпетельным токсинам (3FTX, от англ. three fingered toxins, кардиотоксины, кобротоксины,  $\alpha$ -бунгаротоксины, эрабутоксины,  $\alpha$ -нейротоксины), кодируются генами, построенными из трех экзонов и двух интронов (табл. 2) [76,119–125]. Первый интрон располагается в области, кодирующей С-концевой фрагмент сигнального пептида, второй – в области центральной петли зрелого пептида. Сайты сплайсинга интронов, как и в случае описанных выше других генов токсинов, подчиняются правилу «GT/AG»: интрон начинается с последовательности GT и заканчивается на AG. Интересно, что экзон 1, кодирующий большую часть сигнального пептида, является наиболее консервативным фрагментом гена, а наиболее вариабельны экзоны 2 и 3, кодирующие зрелый пептид. Интрон 1 менее консервативен, чем экзон 1, но более, чем экзон 2; интрон 2, наоборот, более консервативен, чем окружающие его экзоны 2 и 3. У всех известных генов трехпальцевых токсинов длина интрона 1 превышает 1000 п. н. Исключение составляет ген эрабутоксина С, длина интрона 1 которого составляет всего 197 п. н. При сравнении его с генами кардиотоксина 4 и кобротоксина было установлено, что в более длинных интронах имеется вставка, кодирующая малую ядерную РНК. Для многих генов были также изучены промоторные области. Обычно точка начала транскрипции располагается примерно за 29 п. н. от стартового кодона ATG, а ТАТА-бокс – за 25 – 33 п. н. Кроме того, в промоторах нескольких генов имеются сайты связывания факторов регуляции транскрипции NF-1, CACCC, Sp1, GR, LF-A1 и EFlI [122,124].

Теперь рассмотрим гены токсинов, механизм действия которых не связан с взаимодействием с нейрорецепторами. Короткие пептидные токсины змей типа Кунитца (p1pr, от англ. protease inhibitor-like proteins – белки, подобные ингибиторам протеаз, и В-цепи  $\beta$ -бунгаротоксина) и ваприны кодируются генами, состоящими из трех экзонов и двух интронов [77,126]. Гены дизинтегринов отличаются разнообразием экзон-интронной структуры: в их составе можно наблюдать от трех интронов до полного их отсутствия [127]. Это явление связано с интересным эволюционным механизмом их возникновения (см. ниже). Гены фосфолипаз А2 состоят из пяти экзонов и четырех интронов. Первый экзон кодирует 5'-НТО, второй – большую часть сигнального пептида, третий, четвертый и пятый – три части зрелого пептида.  $\beta$ -Бунгаротоксин состоит из двух цепей, А и В, причем цепь А является структурным гомологом фосфолипаз типа А2. Интересно, что гены, кодирующие цепь А, так же имеют разную экзон-интронную структуру. Ген А1 состоит из четырех экзонов и трех интронов, в то время как ген А2 имеет на один интрон и,

соответственно экзон, меньше. При ближайшем рассмотрении выясняется, что при транскрипции с гена A2 происходит пропуск фрагмента, соответствующего первому экзону A1, и синтез мРНК начинается в области, соответствующей концу первого интрона гена A1 [128,129].

Кротамин – короткий пептидный токсин из яда *Crotalus durissus terrificus*, трехмерная структура которого подобна структуре токсинов скорпионов, взаимодействующих с  $\text{Na}^{2+}$ -каналов. Ген, кодирующий кротамин, располагается на хромосоме 2 и содержит 2 интрона [78,130].

	Длина	5' сайт сплайсинга		3' сайт сплайсинга				
		1	2	1	2		1	2
Naja naja atra	кобротоксин b	1258	549	CTTAGgtaag	CCGTGgtaag	tccagGATAC	cacagGAACT	[120]
	кардиотоксин 1	1230	550	CTTAGgtaag	AACGAgtaag	tccagGATAC	cacagTTCCT	[124]
	Кардиотоксин2	1242	551	CTTAGgtaag	AACGGgtaag	tccagGATAC	cacagTTCCT	[124]
	Кардиотоксин3	1226	559	CTTAGgtaag	AAAGGgtaag	tccagGATAC	cacagTTCCT	[124]
	Кардиотоксин 4	1233	553	CTTAGgtaag	AACGGgtaag	tccagGATAC	cacagTTCCT	[119]
	Кардиотоксин N	1239	549	CTTAGgtaag	AAAGGgtaag	tccagGATAC	cacagTTCCT	[124]
	Кардиотоксин 31	1244	548	CTTAGgtaag	AAAGGgtaag	tccagGATAC	cacagTTCCT	[124]
	Кардиотоксин 4N	1231	550	CTTAGgtaag	AACGGgtaag	tccagGATAC	cacagTTCCT	[124]
	Кардиотоксин 10	1236	549	CTTAGgtaag	AAAGGgtaag	tccagGATAC	cacagTTCCT	[124]
	clbp2	1237	552	CTTAGgtaag	ATTGAgtaag	tccagGATAC	cacagAATTT	[124]
	clbp3	1216	549	CTTAGgtaag	ATCTAgtaag	tccagGATAC	tcaagTTCCT	[124]
	кобротоксин	1269	551	CTTAGgtaag	CCGTGgtaag	tccagGATAC	cacagGATAT	[119]
L. s.	эрабутоксин C	197	538	CTTAGgtaag	CCGTGgtaag	tccagGATAC	cacagGAACT	[76]
Bungarus multicinctus	α-bgt(α31)	1793	537	CTTAGgtaag	ATTCTgtaag	tccagGATAT	tgtagGTTCC	[121]
	α-bgt(ν31)	1800	533	CTTAGgtaag	ATTCTgtaag	tccagGATAT	tgtagGTTCC	[121]
	κ1- bgt	1065	537	CTTAGgtaag	ATTCTgtaag	tccagGATAC	tgtagGTTCC	[122]
	κ3- bgt	1210	538	CTTAGgtaag	CTTCTgtaag	tccagGATAC	tgtagGTCAC	[122]
	pilp-1	578	1237					[77]
	pilp-2	579	1244					[77]
	pilp-3	580	1248					[77]
	B1 цепь β-bgt	570	3128	TCATCgtgag	TCTTGgtaag	cacagGGGAT	ttagTGTAT	[131]
	B2 цепь β-bgt	578	3096	TCATCgtgag	TCTTGgtaag	cacagCGGAT	ttagAGTAT	[131]
	B4 цепь β-bgt	580	1639	TCATCgtgag	TCTTGgtaag	cacagCGGAT	ttagAGTAT	[131]
	B5 цепь β-bgt	578	1242	TCATCgtgag	TCTTGgtaag	cacagCATAT	tacagGATGA	[131]
B6 цепь β-bgt	571	1248	TCATCgtgag	TCTTGgtaag	cacagCATAT	ttagAGTAT	[131]	
N. n. s.	ntx-1	1203	553	CTTAGgtaag	CCGTGgtaag	tccagGATAC	cacagGATAT	[132]
S. c. e.	3ftx 1	1091	863	CTCAGgtaag	TCCTGgtaag	tctagGATAC	cacagTACTG	[125]
	3ftx 4	1074	851	CCAGGgtaag	TCCTGgtaag	tctagATACT	cacagTACTG	[125]
	3ftx 5	1045	842					[125]
	3ftx 2	1754	955	GCCAGgtaag	GTCTGgtaag	tctagGATAC	cacagTAATG	[125]
C. d.	кроматамин	900	140	ACCAGgtaag	AAAAAgtaag	catagGGAAT	tatagGGGAA	[78]

**Таблица 2. Интрон-экзонная организация известных генов токсинов змей. 1 – первый интрон, 2 – второй интрон. L. s. – *Laticauda semifasciata*, N. n. s. – *Naja naja sputatrix*, S. c. e. – *Sistrurus catenatus edwardsii*, C. d. t. – *Crotalus durissus terrificus***



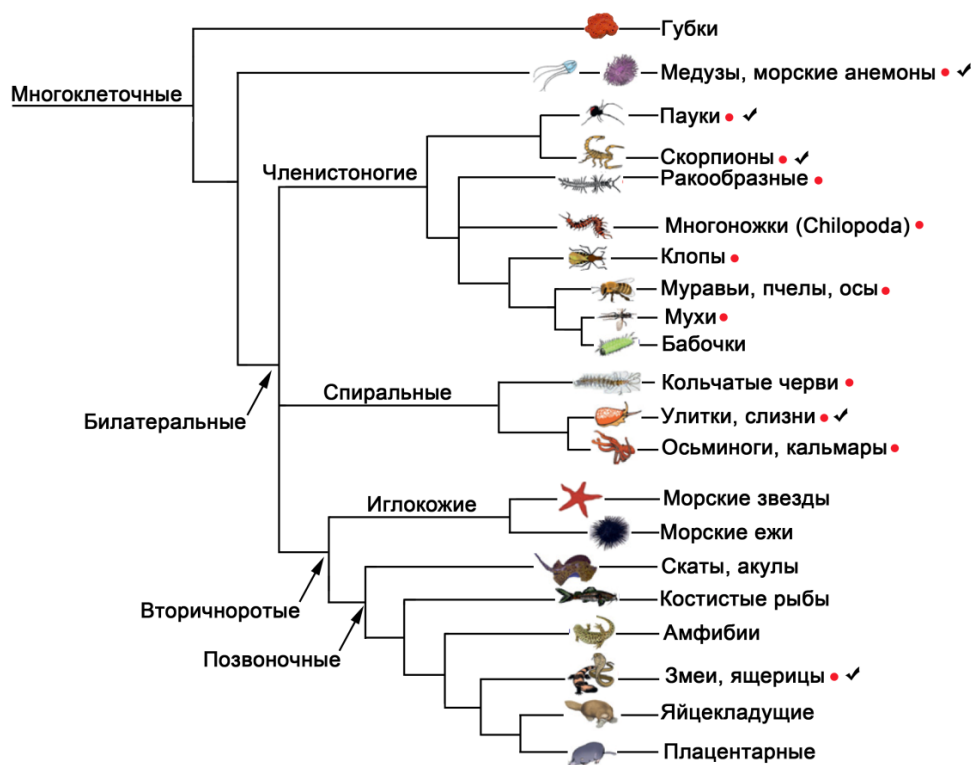
## 5. Эволюция токсинов белковой/пептидной природы

### 5.1 Откуда берутся токсины?

В природе существует большое количество разнообразных ядовитых животных (рис. 3). К настоящему времени относительно подробно изучены яды нескольких групп: среди беспозвоночных – это морские анемоны (Cnidaria), морские улитки конусы (Mollusca), пауки и скорпионы (Arthropoda), а среди позвоночных (Vertebrata) – змеи. Совсем недавно было установлено, что головоногие моллюски также обладают разнообразными по составу ядами. Этих животных объединяет то, что они хищники, и их яды используются для охоты. Интересно, что при таком удаленном родстве состав ядов формируется во многом по похожим принципам – наблюдается их конвергентная эволюция: разные группы организмов независимо «вовлекают» в состав ядов компоненты, действующие на одни и те же мишени (в большинстве случаев, это система гемостаза и нервная система) [133]. 14 типов белков были конвергентно вовлечены в яды двух или более групп животных [134]. Здесь мы остановимся на том, откуда (в эволюционном смысле) появляются белковые/пептидные, компоненты ядов.

Наиболее подробно этот вопрос изучен на примере змей. Было показано, что большинство их токсинов произошло от белков, уже имеющих определенные функции в других тканях. Гены этих белков либо утратили старую функцию и превратились в гены токсинов, либо претерпели дубликации и образовали мультигенные семейства, и уже новые копии взяли на себя функции генов токсинов [135,136]. Этому процессу подверглись следующие белки: семейства ADAM (дизинтегрины и металлопротеиназы), AVIT, LYNX/SLUR, SP1a/Ryanodine, ацетилхолинэстераза, белок C3 системы комплемента, кротазины/ $\beta$ -дефензины, цистатины, эндотелин, факторы V и X, калликреины, ингибиторы протеаз типа Кунитца, L-аминооксидаза, лектин, натрийуретический пептид, фактор роста нервов  $\beta$ , фактор роста эндотелия сосудов, кислый белок сыворотки, фосфолипаза A2, ингибитор лейкопротеиназы. Многие из этих белков влияют на работу кровеносной системы, являются фосфолипазами, протеазами или способны связываться с определенными рецепторами. Как минимум произошло 24 события вовлечения белков из других тканей в состав яда. «Попад» в состав яда, эти белки дивергировали и приобрели новые функции, хотя все они в разной степени сохранили биологическую активность своих предков. Белки семейства LYNX/SLUR дали начало разнообразным пептидным трехпальцевым токсинам. Интересно, что структура генов в большинстве случаев сохранилась [136]. Фосфолипаза A2 вовлекалась в состав ядов чешуйчатых рептилий 4 раза [134,137].

Конвергентная эволюция ядов беспозвоночных прослеживается на примере ядов головоногих, стрекающих, насекомых и скорпионов, в которые независимо были вовлечены фосфолипазы А2. Токсины типа Кунитца в результате независимых событий входили в яды стрекающих, моллюсков конусов, насекомых, скорпионов и дважды в яд пауков [133]. У рыбоядных моллюсков конусов *Conus consors* и *C. adamsonii* в яде были обнаружены гиалуронидазы. Гомологи этих белков были так же зафиксированы в ядах других животных. Установлено, что эти события перехода гиалуронидаз в яд происходили независимо [138]. У головоногих моллюсков (подкласс Coleoidea) так же имеются довольно сложные по составу яды. В их состав входят белки CRISP (от англ. cysteine-rich secretory proteins – цистеин-богатые секреторные белки), антиген 5, белки PR-1 (от англ. pathogenesis-related – связанный с патогенезом), карбоксипептидазы, хитиназы, GON-домен металлопротеаз, гиалуронидазы, пацифастин, фосфолипазы А2, SE-цефалотоксин, сериновые протеазы. Многие из этих белков также используются другими ядовитыми животными, что говорит о конвергентной стратегии «разработки» ядов. Интересно, что яде головоногих хорошо



**Рисунок 3. Схематическое филогенетическое дерево ядовитых животных (по [134,139] с изменениями).** Красными точками отмечены те группы животных, которые используют яды для охоты, галочками – группы, у которых была подробно изучена эволюция токсинов и рассматривающиеся в этом литературном обзоре.

представлены и разнообразны дисульфид-богатые белки (калликреины и пацифастины) [140,141]. У пауков был зафиксирован случай горизонтального переноса генов. Сфинго-

миелиназа D пауков рода *Loxosceles* – единственный известный случай присутствия этого фермента в царстве животных. Скорее всего, паук приобрел его путем горизонтального переноса генов от бактерий рода *Corynebacteria* [142].

Процесс превращения «обычных» белков в токсины у беспозвоночных изучен менее подробно, чем у змей. Короткие дисульфид-богатые пептиды скорпионов, ингибирующие натриевые каналы, скорее всего, произошли от дефензинов с антифунгальной функцией, входящих в систему врожденного иммунитета [143–145]. Ноттиновые токсины играют важную роль в ядах пауков и моллюсков конусов, а также они встречаются в яде скорпионов. Было сделано предположение, что все ноттиновые токсины животных имеют общее эволюционное происхождение [146]. Для многих других коротких пептидных токсинов беспозвоночных, в отличие от дизинтегринов и трехпетельных токсинов змей, предков обнаружено не было.

Изучение ядов разных групп животных показало, что большое количество дисульфидных связей значительно повышает шансы белка превратиться в токсин. Они стабилизируют пространственную структуру белка, освобождая другие аминокислотные остатки от этой функции и позволяя им подвергаться ускоренной эволюции. Такие молекулы эффективно участвуют в «гонке вооружений» и быстро приобретают новые функции [134,136].

## **5.2 Молекулярные механизмы эволюции токсинов**

Неоднократно сообщалось, что токсины в составе ядов разных животных очень разнообразны. Не только яды пауков, но и других беспозвоночных (скорпионов, моллюсков конусов) и позвоночных (змей) содержат т. н. комбинаторные библиотеки пептидов [80,147–149]. Движущая сила эволюции токсинов – «гонка вооружений» между ядовитым животным и его жертвой: хищник «изобретает» разнообразные токсины, эффективно и селективно действующие на жизненно важные мишени жертвы, которая, в свою очередь, «старается» защититься от их действия, вводя мутации в белки-мишени. Вероятно, именно поэтому единственный относительно хорошо изученный яд млекопитающего, утконоса *Ornithorhynchus anatinus*, несколько отличается в эволюционных механизмах [150]: утконос не использует свой яд для добычи пищи. Здесь мы рассмотрим молекулярные механизмы формирования разнообразия токсинов.

Многие токсины эволюционируют по механизму «рождения и смерти» [151]: в результате частых дупликаций образуются новые копии генов токсинов, которые затем дивергируют и приобретают новые функции, сохраняя при этом консервативными ключевые

позиции (во многих случаях, это структурно важные остатки цистеина). Часто наблюдается ускоренная эволюция под действием положительного отбора. На данный момент представляется, что наиболее выражено это явление среди токсинов моллюсков конусов.

### 5.2.1 Конусы

На основании гомологии сигнальных пептидов конотоксины объединяются в надсемейства [108]. Консервативный сигнальный пептид соответствует определенному цистеиновому мотиву в зрелом токсине. Однако в статье [152] сообщается о том, что у *Conus litteratus* обнаружены «химерные» последовательности, у которых сигнальные пептиды надсемейств О и М сочетаются со зрелыми последовательностями с новыми цистеиновыми мотивами.

Для надсемейства А генов конотоксинов четырех близкородственных видов *C. lividus*, *C. diadema*, *C. sanguinolentus* и *C. quercinus* было установлено, что для них характерна высокая скорость «оборота» генов, сопровождающаяся повышенной скоростью эволюции [153]. С момента расхождения *C. lividus*, *C. diadema* и *C. sanguinolentus* надсемейство А претерпело 13 раундов удвоения количества копий генов, причем отдельные локусы подвергались вплоть до 4 дупликаций. Увеличение количества копий генов является результатом, во-первых, мутации дупликации фрагмента геномной ДНК и, во-вторых, закреплением в геноме новых копий гена. Кажется, на протяжении эволюции этих четырех видов генные дупликации происходили с более-менее равномерной скоростью, иногда сопровождаясь повышением скорости дупликаций у отдельных видов. По разным оценкам, скорость дупликаций составляла 0,06 – 0,08 на ген на миллион лет. Эти значения по крайней мере в 2 раза превышают скорость дупликации генов других мультигенных семейств, в т.ч. генов токсинов пауков [154]. Сама по себе (то есть без учета скорости закрепления новых генов в геноме) скорость мутаций дупликаций может быть значительно выше, чем полученные оценки, поскольку новые гены, во-первых, могут теряться, а во-вторых, могут не дивергировать, из-за чего их невозможно отличить друг от друга. Так, у исследованных видов конусов было обнаружено 13 последовательностей псевдогенов, образовавшихся в результате мутаций генов надсемейства А конотоксинов. Предполагается, что повышенной скорости дупликаций способствует расположение генов в областях генома, в которых наблюдается высокая частота неравного кроссинговера [155].

Имеется много свидетельств в пользу того, что после дупликации дивергенция происходит под давлением положительного отбора. На 6 видах конусов (*C. gloriamaris*, *C. magus*, *C. textile*, *C. geographus*, *C. purpurascens*, *C. striatus*) было показано, что в преде-

лах одного надсемейства токсинов можно проследить градиент вариабельности вдоль последовательности белка-предшественника от сигнального пептида через пропептид к зрелому. Эти три фрагмента пропептида кодируются разными экзонами, т. е. фрагменты с разной скоростью эволюции разделяются интронами [83]. У *C. abbreviatus*, *Conus ebraeus* и *C. lividus* во фрагментах генов, кодирующих зрелые пептиды, количество несинонимичных нуклеотидных замен ( $d_N$ ) намного выше, чем синонимичных ( $d_S$ ), что является признаком диверсифицирующего отбора и адаптивной эволюции. Интересно, что в зрелых фрагментах повышается не только  $d_N$ , но и  $d_S$ , возможно, из-за изменения позиции нуклеотида с синонимичной на несинонимичную в процессе эволюции [148]. В работе [84] на *C. arenatus*, *C. pennaceus*, *C. tessulatus* и *C. ventricosus* было установлено, что для зрелого пептида характерны повышенная скорость нуклеотидных замен, причем трансверсии ( $T_v$ ) преобладают над транзициями ( $T_s$ ). Возможно, это объясняется существованием специального механизма повышения частоты мутаций, например, с помощью ДНК-полимеразы, допускающей ошибки с повышенной частотой. Так, ДНК-полимераза V склонна совершать скорее трансверсии, чем транзиции [156,157]. Интересно, что цистеиновые кодоны высококонсервативны, при этом в определенных положениях с большим перевесом преобладает один из возможных кодонов TGT или TGC. В то же время, консервативный остаток аргинина, по которому происходит процессинг при созревании токсина, кодируется всеми аргининовыми кодонами, при этом они используются с той же частотой, как и в среднем среди моллюсков. Таким образом, механизм сохранения остатков цистеина остается не ясным. Для зрелых пептидов характерно соотношение  $d_N/d_S > 1$ , свидетельствующее об адаптивной эволюции.

В работе [153] максимальная скорость несинонимичных замен сразу после дупликации составила 22,9% на миллион лет, что значительно способствовало высокой скорости эволюции этих генов, в результате чего новые копии генов приобрели новые функции. Затем скорость несинонимичных замен заметно снижалась, и средняя скорость составляла 0,9 – 2,7% на миллион лет. В другой публикации [148] сообщается о 1,7 – 4,8% несинонимичных замен на миллион лет.

### 5.2.2 Змеи

Недавно была установлена последовательность генома королевской кобры *Ophiophagus hannah* [135]. Было показано, что гены токсинов змей эволюционируют с помощью нескольких разных механизмов и проявляют совершенно различные скорости дупликации генов и направленного отбора, коррелирующие с их функциональной важно-

стью при охоте на добычу. Были обнаружены многочисленные паттерны дубликации генов и эволюции последовательностей под действием положительного отбора как до, так и после вовлечения их в производство яда. В наиболее активно экспрессирующихся и наиболее функционально важных генных семействах наблюдаются повышенные скорости дубликаций и отбора. Семейство трехпетельных токсинов – наиболее представленное и разнообразное как на уровне транскриптома, так и протеома. Оно подверглось значительной экспансии и демонстрирует высокие уровни положительного отбора и генной дубликации. Семейства фосфолипаз A2, металлопротеиназ змеиного яда и калликреиновых токсинов тоже демонстрируют высокие уровни дубликации генов и признаки положительного отбора. Семейства токсинов с вспомогательными функциями (например, гиалуронидазы, разрушающие ткани жертвы в месте введения яда) демонстрируют более низкий уровень экспрессии, низкий или вообще отсутствующий уровень дубликации генов и отсутствие направленного отбора. Вероятно, такие семейства имеют консервативные функции и не участвуют в «гонке вооружений». Кроме того, подверженность отрицательному отбору и высокую консервативность демонстрируют цитотоксичные трехпетельные токсины [158]. Их функция обеспечивается гидрофобными остатками, занимающими около 40% поверхности. Поскольку взаимодействие цитотоксинов с мембранами неспецифично, они не участвуют в «гонке вооружений» и поддерживают в консервативном состоянии все функционально важные остатки. Относительно слабое влияние положительного отбора наблюдается среди к-нейротоксинов, способных формировать нековалентные гетеродимеры, для чего им необходимо поддерживать поверхности взаимодействия консервативными.

Наблюдается некоторая тенденция к тому, что большие белковые токсины зачастую (но не всегда) эволюционируют медленнее, чем короткие пептидные токсины. Это связано с необходимостью поддержания трехмерной структуры, что достигается за счет консервативности большого количества остатков. Данное явление хорошо иллюстрируется подробно изученным примером RII-дизинтегринов змей, прошедших эволюционный путь от больших белков-ферментов до коротких токсинов.

RII-дизинтегрины – цистеин-богатые полипептиды, широко распространенные в яде гадюк и гремучих змей. Они являются антагонистами рецепторов интегринового типа. RII-дизинтегрины образовались в результате приобретения новой функции дизинтегрин-подобным доменом геморрагической металлопротеиназы RIII (SVMP) из яда змей, ген которой претерпел дубликацию. Далее происходил процесс минимизации гена (потери интронов и кодирующих фрагментов) и белковых структур (некоторых дисульфидных свя-

зей) и дивергенция дизинтегринов. Топология филогенетического дерева дизинтегринов не совпадает с топологией видového дерева, что является признаком того, что это мультигенное семейство подвергалось эволюции по механизму «рождения и смерти», когда появление и исчезновение генных локусов направляется отбором. Остатки цистеина и остатки, входящие в «ядро» молекулы, подвергаются мощному негативному отбору, так как они участвуют в поддержании пространственной структуры дизинтегринов. Дивергенция этих токсинов происходит благодаря мощному положительному отбору среди остатков, расположенных на поверхности молекулы, а также тех, что располагаются в активном центре. Интересно, что постепенное укорочение генов дизинтегринов сопровождалось увеличением скорости аминокислотных замен: в длинных дизинтегринах 18% всех аминокислотных остатков находятся под действием положительного отбора, и это значение возрастает до 23%, 21% и 26% в ряду средних, димерных и коротких дизинтегринов, соответственно [159].

Фосфолипазы A2 проявляют тенденцию к положительному отбору. Это можно заметить при сравнении уровня вариабельности экзонов с уровнем вариабельности интронов. Как уже упоминалось выше, гены, кодирующие фосфолипазу A2, состоят из четырех экзонов и трех интронов. У гадюки *Trimeresurus flavoviridis* уровень сходства интронов гораздо выше, чем у всех экзонов кроме первого, кодирующего сигнальный пептид. Количество нуклеотидных замен на сайт в интронах примерно в четыре раза ниже, чем в экзонах. Поскольку никакой важной функциональной роли у интронов генов фосфолипазы A2 (например, о ней могли бы свидетельствовать элементы вторичной структуры) обнаружено не было, нельзя предположить, что они подвергаются отрицательному отбору. Более вероятно, что скорость эволюции экзонов выше. Действительно, для всех экзонов, кроме кодирующего сигнальный пептид,  $d_N/d_S > 1$ , что свидетельствует о действии положительного отбора [79,160]. Эти выводы подтверждаются при сравнении генов фосфолипаз A2 *T. flavoviridis* и *Trimeresurus gramineus*. Скорость эволюции некодирующих участков генов фосфолипаз была на том же уровне, что и генов ТАТА-боксов связывающих белков, не входящих в состав яда. Это является дополнительным свидетельством в пользу необычно высокой скорости эволюции фосфолипаз [161].

Ситуация, аналогичная только что описанной, наблюдается и среди генов, кодирующих В цепь  $\beta$ -бунгаротоксина: белок-кодирующие фрагменты дивергировали гораздо сильнее, чем интроны (исключение составляет сигнальный пептид), поскольку на них действует положительный отбор [131].

Дополнительно к дупликации генов и положительному отбору в эволюции токсинов могут быть задействованы и другие механизмы. Трехпетельные нейротоксины змей (или  $\alpha$ -нейротоксины) подразделяются на две группы – короткие и длинные, в зависимости от длины их полипептидной цепи. На генах нескольких изоформ эрабутоксина (короткий токсин) и токсина LsIII (длинный токсин) плоскохвостов *Laticauda semifasciata* и *Laticauda laticaudata* было показано, что короткие и длинные  $\alpha$ -токсины дивергировали от общего предка на ранних стадиях эволюции. Гены трехпетельных токсинов содержат два интрона. Нуклеотидные последовательности интрона II высококонсервативны (как минимум, 82% идентичных остатков). В экзоне II, наоборот, наблюдается большое количество мутаций. Гены коротких  $\alpha$ -токсинов имеют удлиненный интрон II, то есть, другими словами, сайт сплайсинга несколько сдвинут ближе к 5'-концу гена у длинных  $\alpha$ -токсинов. Сдвиг сайта сплайсинга произошел из-за точечной мутации, которая обнаруживается вне зависимости от вида змеи и функции кодируемого токсина. Таким образом, эта мутация привела к превращению фрагмента интрона в кодирующий фрагмент и появлению группы длинных  $\alpha$ -токсинов. Дальнейшее накопление мутаций в экзонах, особенно во втором, под давлением положительного отбора привело к расхождению их функций. В отличие от интрона II, интрон I значительно варьирует в своей длине (как это уже описывалось выше в разделе о генах). Интрон I можно разбить на 10 областей, причем разные гены отличаются набором этих областей. Так, интрон I гена эрабутоксина содержит только две области (I и X), а гена токсина LsIII – 6 областей (I, II, V, VII, IX и X). Скорее всего, предковый ген трехпетельных токсинов до момента расхождения видов и функций обладал интроном I, составленным из 8 областей (I – VII и X) [162].

Трехпетельные токсины могут изменять свою длину в эволюции не только путем перехода фрагмента последовательности интрона в экзон, но и путем добавления новых экзонов. Так, ген денмотоксина у змей семейства Colubridae содержит 4 экзона, а не 3, как гены трехпальцевых токсинов семейств Elapidae и Viperidae. В результате в сравнении с трехпальцевыми токсинами змей этих семейств, денмотоксин имеет удлиненный N-концевой сегмент [149,163].

Для объяснения разнообразия трехпальцевых токсинов была предложена гипотеза, получившая название ASSET (от англ. Accelerated Segment Switch in Exons to alter Targeting – ускоренное переключение сегментов для изменения мишени). У змей семейств Elapidae и Viperidae экзоны II и III генов трехпальцевых токсинов (экзон I кодирует консервативный сигнальный пептид) могут подвергаться быстрым изменениям путем ускоренного переключения сегментов, в отличие от соседствующих с ними интронов.



Как уже упоминалось выше, интроны генов  $\alpha$ -токсинов (например, ген эрабутоксина) тоже сегментированы, и уровень сходства между сегментами интронов высокий (>85%). Сегменты экзонов более вариабельны: уровень сходства может быть как низким (12 – 50%), так и высоким (60-100%) [125]. Ускоренное переключение сегментов приводит к изменению свойств поверхности молекулы токсина и изменению его функций. Данное явление наблюдается среди генов трехпетельных токсинов чаще, чем среди генов других семейств токсинов, хотя в некоторой степени оно проявляется и среди генов фосфолипазы A<sub>2</sub>, сериновых протеаз и металлопротеиназ. Гипотеза ASSET не исключает последующего воздействия положительного отбора. Молекулярный механизм ускоренного переключения сегментов остается неясным [164].

Однако недавно появилось сообщение о том, что гипотеза ASSET не может объяснить полного разнообразия трехпальцевых токсинов, и была предложена новая гипотеза RAVER (от англ. Rapid Accumulation of Variations in Exposed Residues – быстрое накопление вариаций среди экспонированных остатков) [158]. Суть этой гипотезы заключается в том, что наиболее активно накопление мутаций происходит в поверхностных областях молекулы токсина под действием положительного отбора. Это подтверждается тем, что в трехпальцевые токсины одновременно имеют высоко консервативные остатки, выполняющие важную функцию поддержания структуры, и подвергающиеся частым заменам экспонированные остатки, участвующие в «гонке вооружений». По мнению авторов гипотезы RAVER, гипотеза ASSET не способна объяснить это важное различие в скорости эволюции разных аминокислотных остатков.

Высокая частота мутаций в генах токсинов змей не всегда приводит к «разработке» наиболее мощного яда. У карликовой медянки *Austrelaps labialis* наблюдается необычно высокая частота инсерций и делеций в генах, кодирующих токсины. Это приводит к тому, что большинство семейств генов токсинов имеют укороченные транскрипты, в то время как гены, кодирующие клеточные белки, такой особенностью не обладают. Скорее всего, такие укороченные транскрипты функционально не активны. Более того, были полностью утрачены семейства металлопротеиназ и CRISP. В результате это может быть причиной того, что яд *A. labialis* обладает в несколько раз более низкой токсичностью, чем другие виды того же рода (*Austrelaps superbis* и *Austrelaps ramsayi*), что, в свою очередь, привело к ухудшению способности добывать пищу и, затем, меньшим размерам *A. labialis* [165]. Другой яркий пример понижения активности яда – мраморная морская змея *Aipysurus eudouxii*, которая перешла на питание икрой рыб, не требующее использования яда. Единственный ген трехпетельного токсина, экспрессирующийся в ее ядовитых железах, пре-

терпел динуклеотидную делецию и, как следствие, сдвиг рамки считывания. В результате продукт этого гена – неактивный укороченный пептид длиной 22 а. о. (вместо 60 – 65 а. о., характерных для активных пептидов этой группы). Однако в результате неиспользования яда он вышел из-под давления естественного отбора, и мутация закрепилась в геноме. По сравнению с другими представителями рода *Aipysurus*, активность яда *A. eydouxi* ниже в 50 – 100 раз [166].

### 5.2.3 Скорпионы

Эволюция токсинов скорпионов во многом подвержена тем же тенденциям, что и эволюция токсинов уже описанных конусов и змей. Она связана с дупликациями генов, скорее всего, по механизму неравного кроссинговера [155], хотя в группе подавляющих (от англ. depressant) токсинов был зафиксирован одиночный случай ретропозиции гена [167]. У скорпионов семейства Buthidae токсины из каждой группы проявляют больше сходства друг с другом, нежели с токсинами других видов, большинство подсемейств токсинов специфичны для определенного вида. Скорее всего, сначала произошло расхождение видов и уже после этого закрепление в геноме дублицированных генов, однако нельзя исключить конверсию генов на ранних стадиях их эволюции [167].

Закрепившиеся в геноме гены в значительной мере дивергировали, однако, как и в случае моллюсков конусов и змей, сигнальный пептид сверхконсервативен и подвержен действию отрицательного отбора. Интересно, что синонимичные сайты тоже консервативны (количество синонимичных замен в сигнальной последовательности ниже, чем в нейтрально эволюционирующем интроне). Кроме того, для сигнальных последовательностей каждой из групп токсинов ( $\alpha$ -токсины и подавляющие токсины) свойственна предпочтительность в использовании определенных кодонов, отличных от тех, что используются для зрелых токсинов. Возможно, это обусловлено каким-либо механизмом регуляции трансляции, различающим частоту использования кодонов, а также регуляцией стабильности мРНК и сплайсинга интронов [167].

Фрагменты генов, кодирующие зрелые токсины скорпионов, подвергаются положительному отбору. Степень диверсификации токсинов в каждой группе различна. Так, блокаторы калиевых каналов более разнообразны, чем высококонсервативные по последовательности и трехмерной структуре модуляторы натриевых каналов. Вероятно, это связано с разной степенью вариабельности их мишеней (у млекопитающих, например, количество разных подтипов калиевого канала в ~10 раз выше, чем натриевого) [167]. Стоит отметить, что у токсинов скорпионов зафиксировать действие положительного отбора

сложнее, чем у, например, моллюсков конусов и змей (была даже выдвинута гипотеза, что дивергенция происходит за счет рекомбинации [80]), и для этого иногда требуются более изощренные методы, чем простой расчет значения  $d_N/d_S$  или сайт-специфические методы (хотя они тоже часто выявляют положительный отбор [168]). Однако путем сравнения большого количества последовательной из транскриптомов ядовитых желез скорпионов семейств Bothriuridae, Buthidae и Urodacidae было установлено, что большинство токсинов с пространственной укладкой типа  $CS\alpha/\beta$  подвергаются эпизодическому влиянию положительного отбора, в то время как большинство токсинов с другими типами пространственной укладки находятся под сильным влиянием отрицательного отбора. С помощью методов MEME и BSRtest были обнаружены сайты и филогенетические ветви, испытывающие эпизодические «скачки» давления положительного отбора, среди  $CS\alpha/\beta$ -токсинов (токсины  $\alpha$ ,  $\beta$ , плезиотипические  $Na_V$ -, длинные  $K_V$ - и  $Cl_V$ -токсины), что отражает динамический характер эволюции яда скорпионов [169].

Было неоднократно показано, что большинство быстро эволюционирующих сайтов располагается на поверхности молекулы. Так, у  $\alpha$ - и  $\beta$ -токсинов скорпионов было показано, что положительный отбор действует на аминокислотные остатки, составляющие их биоактивную поверхность. У  $\alpha$ -токсинов 9 из 11 быстро эволюционирующих сайтов находятся на поверхности молекулы и как минимум на 30% доступны растворителю. В большинстве своем, они находятся вне участков, входящих в состав вторичных структур, только два остатка располагались в  $\beta$ -слое. У  $\beta$ -токсинов четыре из восьми остатков более чем на 30% доступны растворителю и два расположены в  $\alpha$ -спирали. Для  $\alpha$ -токсинов предположение о функциональной важности остатков, находящихся под давлением положительного отбора, подтверждается сравнением с экспериментально полученными данными. 6 из 11 быстро эволюционирующих остатков токсина LqhaIT скорпиона *Leiurus quinquestriatus hebraeus* располагаются в области, необходимой для взаимодействия с мишенью и проявления токсического эффекта, и еще несколько в области токсина, ответственной за его специфичность [170]. В последующих исследованиях эти данные были во многом подтверждены. Эксперименты по направленному мутагенезу на токсинах LqhaIT и Lqh2 из яда того же скорпиона показали, что два остатка, находящихся под действием положительного отбора, важны для активности токсинов и их селективности. Так, замена K41P в токсине LqhaIT привела к значительному снижению его активности на натриевых каналах насекомых и млекопитающих. Замена A39L не только вызвала десятикратное возрастание их активности, но и придала токсину LqhaIT новую способность действовать на канал мозга млекопитающих Nav1.2, а токсину Lqh2 – взаимодействовать с каналом

насекомых DmNav1 [171]. Аналогичные результаты приводятся в [172] и [173]. У токсина MeuTx3B и его гомологов из яда скорпионов рода *Mesobuthus*, ингибиторов калиевых каналов, имеет место несколько другая ситуация. Два сайта, находящихся под влиянием положительного отбора, располагаются в области, близкой к той, что вовлечена во взаимодействие с каналом, однако она далека от функционально важной области, необходимой для ингибирования работы калиевого канала [174].

Однако не все токсины типа CS $\alpha$ / $\beta$  подвержены ускоренной эволюции. Так, липолиптические токсины обнаруживают только 2% таких сайтов. Как и в случае цитотоксинов змей, это, скорее всего, связано с их неспецифическим характером действия. Токсины с типами укладки, отличными от CS $\alpha$ / $\beta$ , подвержены отрицательному отбору. Линейные токсины содержат только 0 – 4% сайтов под действием положительного отбора. Ноттиновые токсины играют скорее вспомогательную роль в яде скорпиона, поэтому, как и в случае «вспомогательных» токсинов змей, на них действует отрицательный отбор. Однако, возможно, роль ноттиновых токсинов недостаточно изучена, и они взаимодействуют с крайне консервативными мишенями, что привело к их выбыванию из «гонки вооружений» [169].

Как и у змей, у скорпионов случайные короткие делеции могут приводить к образованию новых токсинов. В транскриптоме ядовитых желез скорпиона *Androctonus australis* были обнаружены последовательности, гомологичные кодирующим длинные токсины, однако содержавшие делеции ближе к 3'-концу кодирующей последовательности. Это привело к сдвигу рамки считывания и образованию нового стоп-кодона. В результате такие последовательности кодируют укороченные токсины. Предполагается, что такая делеция привела к появлению новой группы коротких токсинов скорпионов (что, в отличие от похожего случая у змей, не привело к снижению активности). В пользу этого говорит и филогенетическая близость этих укороченных последовательностей к коротким токсинам [175].

#### 5.2.4 Пауки

На данный момент накопилось довольно много свидетельств в пользу того, что гены токсинов пауков подвержены частым дупликациям. При изучении последовательностей кДНК токсинов пауков и транскриптомов из ядовитых желез было обнаружено, что токсины формируют семейства, обладающие консервативными препро-последовательностями и варибельными зрелыми пептидами, что свидетельствует о происхождении их от одного предкового гена [90,176–182]. Так, токсины HWTX-I, HWTX-IV,

НWТХ-IIIа и SHL-I паука *H. schmidtii*, скорее всего, произошли от одного предкового гена: сходство их препропептидов превышает 75%, хотя зрелые последовательности очень сильно различаются и имеют разные функции. Токсины НWТХ-II и НWТХ-VII того же паука образовались в результате дупликации другого гена: сходство их препропептидов между собой составляет 95,8%, а зрелых фрагментов – 81,1%, однако эти показатели намного ниже при сравнении их с предыдущей группой [183]. Аналогичные выводы были сделаны и при изучении транскриптома ядовитых желез одной особи *Dolomedes mizhoanus*, что исключает индивидуальную вариабельность [184].

По сравнению с описанными ранее группами, молекулярные механизмы эволюции пептидных токсинов пауков, приводящие к формированию их разнообразия, изучены намного хуже. Единственный пример исследования такого рода – сравнение последовательностей транскриптов, кодирующих паралоги блокатора кальциевых каналов насекомых  $\omega$ -атракотоксина (ACTX) Hv1a пауков *Atrax robustus*, *Hadronyche infensa* и *Hadronyche versuta*. Сигнальные последовательности высоко консервативны, в то время как в пропептидах и зрелых пептидах наблюдается значительная вариабельность. Единственный сильно консервативный участок пропептида – непосредственно предшествующий зрелому пептиду сайт процессинга Arg-Arg. Последовательности зрелых пептидов варьируют очень сильно, но остатки цистеина, формирующие дисульфиды, консервативны. Кроме того, есть небольшое количество других консервативных остатков, некоторые из которых соответствуют функционально важным остаткам  $\omega$ -ACTX-1, определенным с помощью сканирующего мутагенеза. Четыре из шести остатков цистеина (образующие дисульфидные мостики C<sup>1</sup> – C<sup>4</sup> и C<sup>3</sup> – C<sup>6</sup>) проявляют предпочтение в пользу определенного цистеинового кодона. Только для сигнальных пептидов соотношение  $d_N/d_S$  было значительно меньше 1, что свидетельствует об отрицательном отборе. Для пропептидов и зрелых пептидов однозначного вывода сделать не удалось. Однако яд пауков, скорее всего, как и яд скорпионов, содержит большое количество паралогов, появившихся в результате дупликации генов, происходившей независимо у каждого вида. Для филогенетических ветвей, состоящих из близких паралогов  $d_N/d_S$  значительно выше 1, что свидетельствует о том, что расхождение видов сопровождалось адаптивной эволюцией зрелых последовательностей [80].

Среди крупных белковых токсинов пауков ситуация отличается. В соответствии с гипотезой «рождения и смерти», гены сфингомиелиназы D пауков семейства Sicariidae подвергались неоднократным дупликациям: до 0,04 – 0,07 дупликаций на миллион лет, что превышает оценки для многих других мультигенных семейств, но ниже уровня дуп-

ликаций у моллюсков конусов. Дальнейшая эволюция этих генов проходила под воздействием мощного отрицательного отбора. Так, среднее значение соотношения несинонимических нуклеотидных замен к синонимическим для генов фосфолипазы A2 змей на порядок выше, чем для генов сфингомиелиназы D. Однако для некоторых филогенетических ветвей и нескольких аминокислотных остатков влияние положительного отбора все-таки было обнаружено. Все быстро эволюционирующие сайты экспонированы на поверхности молекулы, но располагаются далеко от функциональных сайтов. Сфингомиелиназа D имеет структуру ТИМ-бочки, и наиболее консервативные остатки располагаются на ее поверхности около двух отверстий бочки вблизи активного сайта или в области «мотива втулки». В активном сайте остатки, вовлеченные в связывание фосфата или ионов магния, крайне консервативны (в 98% случаев они идентичны) [154].

Похожая ситуация наблюдается и с другим крупным белковым токсином  $\alpha$ -латротоксином пауков семейства Theridiidae. Сходство нуклеотидных последовательностей генов  $\alpha$ -латротоксина превышает 94% и имеются доказательства того, что их эволюция проходила под действием отрицательного отбора. Однако, благодаря положительному отбору, действующему на уровне физико-химических свойств молекулы и определенных кодонов, а также повышенной частоте несинонимичных замен вдоль филогенетической ветви  $\alpha$ -латротоксинов рода *Latrodectus*, была достигнута высокая степень дивергенции аминокислотных последовательностей. Возможно, на ранних стадиях эволюции *Latrodectus*  $\alpha$ -латротоксин были под влиянием положительного отбора, который затем сменился отрицательным отбором, поддерживающим выгодные для охоты и защиты аллели генов  $\alpha$ -латротоксинов. Долговременный отрицательный отбор свидетельствует в пользу функциональной важности  $\alpha$ -латротоксинов в составе яда [185].

### 5.2.5 Морские анемоны

Токсины морских анемонов также кодируются мультигенными семействами [81]. Однако эволюционные механизмы токсинов морских анемонов несколько отличаются от той картины, которую мы видели для токсинов других животных. Если пептидные токсины описанных выше таксонов, в основном, эволюционируют по механизму «рождения и смерти», то некоторые токсины анемонов используют стратегию «согласованной эволюции».

При анализе последовательности генома морской анемоны *Nematostella vectensis* кроме генов, кодирующие пептиды, гомологичные блокаторам калиевых каналов, было

обнаружено только одно семейство генов нейротоксинов, действие которых направлено на натриевый канал. Все 10 членов одного семейства располагаются в одном участке генома длиной 30 т. п. о. и, что самое удивительное, кодируют один и тот же зрелый продукт (токсин Nv1), причем даже синонимические нуклеотидные замены в этих генах отсутствуют. Последовательности, кодирующие сигнальный пептид и пропептид, а также интрон проявляют некоторую вариабельность. Ретроэлемент, обнаруженный рядом с одним из генов, может быть причиной повышенной частоты дупликаций этих генов [186]. Вероятно, эти гены подвержены согласованной эволюции. Такой механизм был доказан для генов рибосомальной РНК [187] и hsp70 [151] дрозофилы, так же отличающихся низким уровнем разнообразия последовательностей. Вероятно, в согласованной эволюции генов анемоны задействован неравный кроссинговер, как это было показано для рибосомальных генов дрожжей [188]. При анализе разных популяций *N. vectensis* и двух других видов анемоны *Anemonia viridis* и *Actinia equina* выяснилось, что у каждого из видов гены токсинов проявляют больше сходства друг с другом, чем с генами других видов, что так же является свидетельством в пользу согласованной эволюции [189].

Однако, помимо этого, у *A. viridis* и *A. equina* были обнаружены гены, подверженные действию диверсифицирующего отбора. Таким образом, механизмы согласованной эволюции и ускоренной эволюции могут применяться организмом одновременно [189,190]. Мультигенное семейство GS типа Кунитца морской анемоны *Heteractis crispa* эволюционировало путем дупликаций с последующей адаптивной дивергенцией остатка в положении P1 [191].

В эволюции токсинов анемоны участвуют процессы слияния генов (англ. gene fusion) и ретропозиции [192]. *Anemonia viridis* – единственный вид морской анемоны, одновременно производящий токсины типа I и типа III, действующие на Na<sup>+</sup>-каналы. Хотя зрелые токсины совершенно не родственны по последовательности и трехмерной структуре, последовательности, кодирующие сигнальные пептиды в высокой степени сходны. Скорее всего, образование генов этих токсинов связано с явлением слияния генов, принадлежавших разным семействам. Для этого мог быть задействован механизм «негомологичного сращивания концов» [193]. Кроме того, у той же анемоны были обнаружены процессированные псевдогены двух генных семейств, что свидетельствует о ретропозиции [192].

### 5.3 Заключение

Итак, механизмы эволюции пептидных токсинов разных групп животных имеют много общих черт. Зачастую формирование разнообразия токсинов обусловлено много-

кратыми генными дупликациями и последующей диверсификацией зрелых последовательностей под давлением положительного отбора. Во многих случаях, когда его действие не удастся зафиксировать по всей длине зрелого пептида, оно направлено на функционально важные аминокислотные остатки. Наиболее консервативными являются остатки, необходимые для поддержания пространственной структуры токсина. Токсины с неспецифическим механизмом действия (мембраноактивные пептиды) или вспомогательной ролью в яде подвержены отрицательному отбору. С другой стороны, у морских анемонов имеются интересные примеры согласованной эволюции токсинов, когда несколько копий генов токсинов не диверсифицируют и кодируют идентичные последовательности.



## Материалы и методы

### 1. Материалы

#### 1.1 Неорганические вещества

NaCl, HCl (Химмед, Россия); гуанидин гидрохлорид, NaH<sub>2</sub>PO<sub>4</sub>, BrCN (Sigma-Aldrich, США).

#### 1.2 Органические вещества

Ацетонитрил (Криохром, Россия); изопропил-β-D-1-тиогалактопиранозид (ИПТГ) (Fermentas); трис(гидроксиметил) аминотетран (Трис) (ICN, США); додецилсульфат натрия (ДСН), имидазол, этилендиаминтетрауксусная кислота (ЭДТА), 1,4- дитиотреитол, 4-винилпиридин, трифторуксусная кислота (ТФУ), 2,7-бис(2-карбоксиэтил)-5(6)-карбоксифлуоресцеина цетоксиметиловый эфир (БКЭКФ), иодид пропидия (ИП) (Sigma-Aldrich); 1,2-диолеоил-глицеро-3-фосфохолин (ДОФХ), 1,2-диолеоил-глицеро-3-фосфоэтанолламин (ДОФЭ), 1,2-диолеоил-глицеро-3-фосфоглицерол (ДОФГ) (Avanti Polar Lipids, США).

#### 1.3 Реактивы для молекулярной биологии:

Коммерчески доступные наборы Promega (США), Fermentas (Литва), Invitrogen (США), Евроген (Россия), Amersham Biosciences (Великобритания); ферменты рестрикции (Fermentas); TRIzol (Invitrogen); плазмиды pET (Novagen), pAL-TA (Евроген), pBlueScriptSK<sup>+</sup> (Stratagen, США).

#### 1.4 Клеточные линии

Бактериальные клеточные линии: *Escherichia coli* (*E. coli*) BL21(DE3) и Origami B, *Arthrobacter globiformis* ВКМ Ас-1112, *Bacillus subtilis* ВКМ В-501, *Enterococcus faecalis*, штамм ВКМ В-871, *E. coli* DH5a, *Pseudomonas aeruginosa* PAO1, *Staphylococcus aureus* 209 Р. Линии раковых клеток человека: HeLa, A549.

#### 1.5 Питательные среды и их составляющие

Бактотриптон, дрожжевой экстракт (Difco, США), DMEM и Игла (PanEco, Россия), бычья эмбриональная сыворотка (HyClone, США), среда Мюллера-Хинтона (Sigma-Aldrich, США)

## 2. Методы

### 2.1 Получение двудоменных токсинов и их производных

#### 2.1.1 Выделение двудоменных токсинов из цельных ядов.

Яды *O. takobius* и *C. punctorium* заказывали в компании Fauna Laboratories Ltd, Казахстан. Разделение проводили через две стадии – гель-фильтрации и обратнофазовой высокоэффективной жидкостной хроматографии (ОФ-ВЭЖХ). На каждом шаге хроматографический профиль получали путем измерения оптического поглощения при длинах волн 210 и 280 нм.

##### а) Выделение токсинов OtTx

Жидкий яд *O. takobius* (25 мкл) лиофилизовали и растворяли в 100 мкл буфера SEC (150 мМ NaCl, 20 мМ NaH<sub>2</sub>PO<sub>4</sub>). Разделение методом гель-фильтрации осуществляли на колонке TSK 2000SW (7.5 × 600 мм, размер пор 12,5 нм, размер частиц 10 мкм) (Тоyo Soda Manufacturing Co., Япония) при скорости потока 0,5 мл/мин. Доминирующую фракцию затем подвергали дальнейшему разделению методом (ОФ-ВЭЖХ) на колонке Delta-Pak C4 (3,9×150 мм, Waters, США) в линейном градиенте концентрации ацетонитрила (0 – 10% за 10 мин., 10 – 50% за 80 мин. и 50 – 70% за 20 мин.) в 0,1% ТФУ при скорости потока 1 мл/мин. Одну из трех поздно элюирующихся фракций, обладающих антимикробной активностью, подвергали дальнейшему разделению методом ОФ-ВЭЖХ на колонке Jupiter C5 (2×150 мм, Phenomenex, США) в линейном градиенте концентрации ацетонитрила (20 – 60% за 60 мин.) в 0,1% ТФУ при скорости потока 0,3 мл/мин. Окончательную очистку проводили на той же колонке до 95% чистоты.

##### б). Выделение токсинов CpTx

Жидкий яд *C. punctorium* (25 мкл) лиофилизовали и растворяли в 100 мкл 10% ацетонитрила, 0,1% ТФУ. Разделение методом гель-фильтрации проводили на колонке TSK 2000SW в том же растворителе при скорости потока 1 мл/мин. Фракцию, соответствующую диапазону молекулярных масс 7–45 кДа, подвергали дальнейшему разделению методом ОФ-ВЭЖХ на колонке Jupiter C<sub>5</sub> (4,6×150 мм) в линейном градиенте концентрации ацетонитрила (0 – 20% за 10 мин, 20 – 60% за 60 мин, и 60 – 80% за 10 мин) в 0,1 % ТФУ при скорости потока 1 мл/мин.

## 2.1.2 Химический синтез пептидов

Химический синтез пептидов проводил сотрудники лаборатории протеомики ИБХ РАН. N-концевой фрагмент пептида OtTx 1a (OtTx 1a-AMP) синтезировали на пептидном синтезаторе Sygo I (Multi Syn Tech, Германия) методом Fmoc/t-butyl, как это описано в [30]. Степень чистоты пептидов контролировали с помощью ОФ-ВЭЖХ и масс-спектрометрии MALDI.

## 2.1.3 Получение рекомбинантных пептидов

### 2.1.3.1 Получение плазмидных векторов для экспрессии

Последовательность ДНК, кодирующую OtTx 1a с пропептидом (pro-OtTx 1a), конструировали из синтетических олигонуклеотидов (табл. 3) методом полимеразной цепной реакции (ПЦР). Между про-последовательностью и зрелой последовательностью вводился дополнительный метиониновый кодон. Полноразмерную последовательность амплифицировали с помощью прямого праймера f-OtTx 1a, содержащего сайт рестрикции KpnI, и обратный праймер rev-OtTx 1a, содержащий сайт рестрикции XhoI и стоп кодон. Фрагмент ДНК, кодирующий OtTx 1a-ICK, получали с помощью прямого праймера f-OtTx 1a-ICK, содержащего сайт рестрикции BamHI и метиониновый кодон. Продукты ПЦР, кодирующие целевые пептиды, очищали с помощью коммерчески доступных наборов Promega, подвергали рестрикции соответствующими ферментами и лигировали с предварительно линейризованными экспрессионными векторами pET-32b (pro-OtTx 1a) или pET-32a (OtTx 1a-ICK) (Novagen, USA). Используемые вектора содержат последовательности, кодирующие тиоредоксин (Trx) и последовательность из 6 остатков гистидина, что позволяет получать целевые пептиды в составе гибридного белка с тиоредоксином и дополнительным 6-ти гистидиновым «хвостом». Полученные генные конструкции проверяли секвенированием в компании Евроген или ЦКП «Геном».

Название	Последовательность (5'→3')
f-iok-68	CTGCGGTACCACCGGTGACCTGGAAACCGAACTGGAGGCTTCTGAA
2f-iok-68	CTGCAGGAACTGCAGGAAGCTCTAGACCTGATCGGTGAAACCTCTCTGGA
3f-iok-68	ATCTCTCGAAGCTGAGGAGCTCGAAGAAGCCCGTAAATTCAAGTGG
3f-iok-68-M	ATCTCTCGAAGCTGAGGAGCTCGAAGAAGCCCGTATGAAATTCAAGTGG
4f-iok-68	GGCAAGCTTTTCTCCACCGCTAAAAAАCTATACAAGAAGGGTAAGAAACT
5f-iok-68	GTCCAAAAACAAGAACTTCAAGAAAGCTCTGAAATTCGGCAAACAGCTCG
6f-iok-68	CTAAAAACCTGCAGGCTGGTGAAGAGCACGAACCGGGTACTCCAGTTGGT
7f-iok-68	AACAACAАATGCTGGGCTATCGGCACCACTTGCTCTGACGACTGCGACTG
8f-iok-68	CTGTCCGGAACACCACTGCCACTGCCCGGCTGGTAAATGGCTGCCGGGTC

9f-iok-68	TGTTCCGTTGCACCTGCCAGGTTACCGAATCTGACAAAGTTAACAAATGC
1/2rev-iok-68	TTCCTGCAGTTCCTGCAGTTCAGAAGCCTCCAGTTC
2/3rev-iok-68	TCCTCAGCTTCGAGAGATTCCAGAGAGGTTTCACCG
3/4rev-iok-68	GGTGGAGAAAAGCTTGCCCCACTTGAATTTACGGGC
3/4rev-iok-68-M	GGTGGAGAAAAGCTTGCCCCACTTGAATTT <b>CAT</b> ACGG
4/5rev-iok-68	GAAGTTCCTGTTTTGGACAGTTTTCTTACCCTTCTTGT
5/6rev-iok-68	CAGCCTGCAGGTTTTTAGCGAGCTGTTTGCCGAATT
6/7rev-iok-68	AGCCCAGCATTTGTTGTTACCAACTGGAGTACCCGG
7/8rev-iok-68	CAGTGGTGTTCGGACAGCAGTCGCAGTCGTCAGAG
8/9rev-iok-68	GGCAGGTGCAACGGAACAGACCCGGCAGCCATTTAC
rev-iok-68	GCCG <b>CTCGAGCT</b> ATTTCAGCCGGTGGGCATTTGTTAACTTTGTCAGA
f-OtTx 1a-ICK	AAAAAAGGATCC <b>ATGGG</b> TACTCCAGTTGGTAACAA- CAAATGCTGGGCTATCGGCACCACT

**Таблица 3. Синтетические олигонуклеотиды, использованные для конструирования ДНК, кодирующей токсина. Сайты рестрикции (*KpnI* в *f-OtTx 1a*, *XhoI* в *rev-OtTx 1a* и *BamHI* в *f-OtTx 1a-ICK*) подчеркнуты, старт и стоп кодоны выделены жирным, метиониновый кодон – жирным курсивом.**

### 2.1.3.2 Гетерологическая экспрессия

Культуры *E. coli*, штаммы BL21(DE3) или Origami B, трансформировали экспрессионными векторами и растили при активной аэрации при температуре 37°C в среде Luria-Bertani (LB) с добавлением ампициллина (100 мг/мл) до оптической плотности 0,4 – 0,6 при  $\lambda=600$  нм. Экспрессию рекомбинантных генов, кодирующих гибридные белки Ttx-OtTx 1a и Ttx-OtTx 1a-ICK, индуцировали добавлением ИПТГ до концентрации 0,2 – 1 мМ. Культуры инкубировали при аэрации в течение 3 ч при 37°C или ночи при 22°C.

### 2.1.3.3 Очистка гибридных белков

Для выделения гибридных белков использовали металл-хелатную аффинную хроматографию, поскольку гибридные белки содержали последовательность из 6 остатков гистидина. Около 1,5 г клеток ресуспендировали в 15 мл стартового буфера для аффинной хроматографии (300 мМ NaCl, 50 мМ Трис-HCl, pH 8) и подвергали лизису на ультразвуковом дезинтеграторе (Cole Parmer, США). Полученный лизат центрифугировали в течение 15 мин при скорости 15000 об./мин. для удаления нерастворимых частиц. Супернатант наносили на предварительно уравновешенную смолу TALON (Clontech, США) объемом 3 мл, и экстрагировали из него гибридный белок, следуя инструкции производителя: промывали буфером 5 мМ имидазол, 500 мМ NaCl, 5% глицерин, 0.1% тритон X-100, 50 мМ Трис-HCl, pH 8, элюировали буфером 150 мМ имидазол, 300 мМ NaCl, 50 мМ Трис-

HCl, pH 8. В случае белка Trx- OtTx 1a все буферы дополнительно содержали 2М гуанидин гидрохлорид.

#### **2.1.3.4 Процедура рефолдинга**

Trx-OtTx 1a-ICK подвергали окислению кислородом воздуха. Белок доводили до концентрации 0,01 – 0,02 мМ в буфере 100 мМ Трис-HCl (pH 8), 80 мМ NaCl, 50 мМ имидазол и инкубировали при энергичном встряхивании вплоть до 5 дней. Процесс рефолдинга контролировали с помощью ОФ-ВЭЖХ и масс-спектрометрии после расщепления гибридного белка BrCN.

#### **2.1.3.5 Расщепление гибридных белков и очистка целевых продуктов**

К раствору гибридного белка добавляли HCl до концентрации 100 мМ и BrCN в молярном соотношении к остаткам метионина 1:1000, инкубировали в темноте в течение ночи (12 – 15 ч) при комнатной температуре. Очистку рекомбинантных полипептидов проводили методом ОФ-ВЭЖХ на колонке Jupiter C5 (10×250 мм, Phenomenex) в линейном градиенте концентрации ацетонитрила (0 – 80% за 60 мин.) в 0,1% ТФУ при скорости потока 5 мл/мин. Чистоту целевых полипептидов проверяли масс-спектрометрически и с помощью N-концевого секвенирования. Концентрацию измеряли с помощью спектроскопии поглощения.

## **2.2 Аналитические методы**

### **2.2.1 Определение первичной структуры пептидов**

Восстановление дисульфидных связей 1,4- дитиотреитолом и алкилирование тиольных групп 4-винилпиридином проводили согласно методике, приведенной в [194]. N-концевое секвенирование проводили методом пошаговой деградации по Эдману на белковом секвенаторе Procise Model 492 (Applied Biosystems, США) согласно инструкции производителя.

### **2.2.2 Исследование вторичной структуры пептидов методом кругового дихроизма**

Полипептиды растворяли в 50 мМ фосфатном буфере (pH7,2), 50% ТФЭ или 20 мМ ДСН до концентрации 1 мг/мл. Спектры кругового дихроизма снимались сотрудниками лаборатории оптической микроскопии и спектроскопии биомолекул ИБХ РАН на спек-

трополяриметре J-810 (Jasco, Япония) в диапазоне длин волн 190 – 250 нм при 20°C как описано [31].

### **2.2.3 Определение концентрации пептидов**

Концентрации пептидов определяли спектрофотометрически. Спектры оптического поглощения снимали на спектрометре Hitachi U-3210 (Япония). Для расчета использовали формулу  $OD_{280} = \epsilon \times c \times l$ , где  $OD_{280}$  – оптическая плотность раствора при  $\lambda = 280$  нм,  $\epsilon$  – коэффициент молярной экстинкции пептида,  $c$  – концентрация пептида,  $l$  – длина оптического пути.

### **2.2.4 Масс-спектрометрия**

Образцы анализировали методом масс-спектрометрии MALDI на спектрометрах M@LDILR (Micromass, Великобритания) и Ultraflex TOF-TOF (Bruker Daltonik GmbH, Германия), как это описано в [31].

### **2.2.5 Методы биоинформатики**

Предсказание вторичной структуры пептидов проводили с помощью программы Protean (DNASTAR, США), расчет молекулярных масс и  $\epsilon$  – с помощью программы GPMW (Lighthouse Data, Дания).

## **2.3 Изучение функциональной активности пептидов**

### **2.3.1 Взаимодействие с искусственными мембранами**

#### **2.3.1.1 Получение липосом**

Цвиттериевые липосомы получали из 1,2-диолеил-глицеро-3-фосфохолина (ДОФХ), а анионные – из смеси 1,2-диолеил-глицеро-3-фосфоэтаноламина (ДОФЭ) и 1,2-диолеил-глицеро-3-фосфоглицерол (ДОФГ) (3:7). Сухие липиды растворяли в смеси хлороформ/метанол (2:1 по объему) и сушили на роторном испарителе до получения липидной пленки, которую затем лиофилизовали и суспендировали в 50 мМ фосфатном буфере с добавлением 110 мМ NaCl (pH 7,2) до конечной концентрации по липидам 10 мМ. Суспензию обрабатывали ультразвуком и с помощью мини-экструдера (Avanti Polar Lipids) пропускали через поликарбонатный фильтр с диаметром пор 1000 нм (Whatman, Великобритания) 21 раз при комнатной температуре для получения гигантских униламеллярных везикул.

### **2.3.1.2 Тест соосаждения с липосомами**

Полипептиды добавляли к суспензии гигантских униламеллярных липосом до концентрации 5 мкМ. Раствор полипептида в фосфатном буфере использовали в качестве контроля. После инкубации при 37 °С в течение 1 ч смесь центрифугировали 20 мин (20800 g, комнатная температура). Супернатант анализировали ОФ-ВЭЖХ.

### **2.3.2 Биологическая активность**

#### **2.3.2.1 Тесты на антимикробную активность**

Определение значений МИК полипептидов проводили методом двукратных серийных разведений. Клеточные линии *A. globiformis*, *B. subtilis*, *E. faecalis*, *E. coli*, *P. aeruginosa* и *S. aureus*, растили на питательной среде LB. В логарифмической фазе роста культуру разбавляли средой до концентрации  $1 \times 10^5$  колониеобразующих единиц/мл. Пептиды растворяли в 10 мкл деионизованной воды и смешивали с 90 мкл бактериальной культуры. Пептиды, контрольные образцы, не содержавшие пептидов, и образцы контроля стерильности тестировались в двух повторностях. Микробиологические планшеты с образцами инкубировали в течение ночи при 37 °С. Ингибирование бактериального роста определяли измерением оптического поглощения при  $\lambda=620$  нм. За МИК принимали концентрацию пептида, полностью подавлявшую рост бактериальной культуры.

Бактерицидный эффект пептида OtTx 1a-AMP против *E. faecalis* и *S. aureus* оценивался путем высевания бактерий на агаризованную питательную среду после 3 ч инкубации с пептидом в концентрациях, равных МИК, 2×МИК и 4×МИК, и оценки количества образовавшихся через 24 ч колоний.

Для изучения кинетики бактерицидного действия OtTx 1a-AMP культуру *S. aureus* инкубировали в среде с пептидом (1,25 мМ и 2,5 мМ) в течение 0,25, 0,5, 1 и 2 ч, а затем разбавляли в 20 раз чистой средой, высевали на агаризованную среду и через 24 ч подсчитывали колонии.

#### **2.3.2.2 Тест на проницаемость бактериальной мембраны**

Исследования методом конфокальной лазерной сканирующей микроскопии (КЛСМ) проводили сотрудники лаборатории оптической микроскопии и спектроскопии биомолекул ИБХ РАН на конфокальном инвертированном микроскопе (Leica Microsystems, Германия). Для ИП длина возбуждающей волны была 561 нм, а для БКЭКФ – 488 нм. Мощность лазера составила 3 – 10 мкВт на образец. Эмиссию флуоресценции измеря-

ли в промежутке длин волн 650 – 700 нм для ИП и 500 – 550 нм для БКЭЖФ. Культуру *S. aureus* инкубировали с 1 мкМ БКЭЖФ в сбалансированном солевом растворе Хенкса 30 мин. при 37 °С, в течение 5 – 10 мин. обрабатывали 1,25 мкМ ОтТх 1а-АМР и 15 мкМ ИП в среде Мюллера-Хинтона, осаждали клетки центрифугированием (1600 g, 5 мин., комнатная температура) и исследовали КЛСМ. Контрольные клетки подвергали той же процедуре, но без добавления пептида.

### **2.3.2.3 Тест на инсектицидную активность**

Инсектицидную активность исследовали на личинках третьей возрастной стадии мясной мухи *Sarcophaga carnaria*. Пептиды растворяли в физиологическом растворе и по 5 мкл вводили в четвертый сегмент личинок. В качестве контроля использовали физиологический раствор. Летальный и паралитический эффекты фиксировали в течение 24 ч после инъекции.

### **2.3.2.4 Тест на гемолитическую активность**

Капиллярную кровь человека собирали в пробирку с гепарином (10 ед./мл), разбавляли до концентрации  $1 - 0,1 \times 10^7$  клеток/мл средой RPMI-1640, содержащей 10% эмбриональной бычьей сыворотки, и инкубировали с пептидом ОтТх 1а-АМР (0,16–20 мкМ, двукратные разведения) в течение 3 ч при 37°С при аккуратном перемешивании. Выход гемоглобина измеряли как описано в [43].

### **2.3.2.5 Тест на цитотоксичность**

Клеточные линии эпителиоидной карциномы человека HeLa и аденокарциномы легких человека A549 растили при 37 °С, 5% CO<sub>2</sub> в средах DMEM и Игла, соответственно, с добавлением 10% эмбриональной бычьей сыворотки и 2 мМ L-глутамина. Клетки высевали в 96-луночные планшеты до плотности  $1 \times 10^5$  клеток/мл и через 24 ч в лунки добавляли ОтТх 1а-АМР до концентрации 0,16 – 20 мкМ. Через 24 ч степень жизнеспособности клеток оценивали колориметрическим МТТ-методом [195].

## **2.4 Работа с кДНК и геномной ДНК**

### **2.4.1 Библиотеки кДНК из ядовитых желез**

Библиотеки из ядовитых желез пауков *O. takobius* и *S. punctoriun* были получены ранее в тесном сотрудничестве с компанией «Du Pont Agriculture and Nutrition» (США) по методике, описанной в публикации [90]. Ядовитые железы выделяли и гомогенизировали



в жидком азоте, а затем подвергали лизису реагентом TRIzol. Полиаденированную мРНК выделяли на аффинной колонке с олиго-дТ целлюлозой с помощью набора для выделения мРНК. Синтез первой и второй цепей кДНК, присоединение линкеров и клонирование в вектор pBlueScriptSK<sup>+</sup> осуществлялись с помощью коммерчески доступного набора Stratagene. Очистку кДНК проводили с помощью соответствующего набора производства компании Invitrogen. Секвенирование осуществляли с помощью прямого праймера M13 и набора «ABI PRISM Big Dye Terminator Cycle Sequencing Ready reaction kit» и AmpliTaq ДНК-полимеразы на секвенаторе 373 DNA Sequencer (Applied Biosystems).

#### **2.4.2 Выделение ДНК и РНК из ядовитых желез пауков**

Пауки *O. lineatus* были предоставлены А. Федоровым (Fauna Laboratories, Казахстан). Ядовитые железы выделяли и хранили при  $-70\text{ }^{\circ}\text{C}$ . ~100 мг ткани ядовитых желез от нескольких особей использовали для экстракции геномной ДНК и тотальной РНК. Ткань растворяли при  $+55\text{ }^{\circ}\text{C}$  в 500 мкл буфера 50 mM Трис-НСl (рН 8.0), 200 mM NaCl, 100 mM ЭДТА, 1% ДСН, содержащем 1 мг/мл протеиназы К. После добавления 250 мкл 5M NaCl центрифугировали на настольной микроцентрифуге (Eppendorf, Германия) для удаления осадка. К супернатанту добавляли 750 мкл 96% этанола, центрифугировали. Смесь ДНК и РНК экстрагировали из осадка фенолом и хлороформом. После повторного осаждения этанолом смесь ДНК/РНК растворяли в деионизованной воде. Полученную геномную ДНК использовали в ПЦР. Для получения кДНК проводили обратную транскрипцию мРНК с помощью коммерчески доступного набора MINT (Евроген, Россия) согласно методике производителя. Последовательности кДНК были фланкированы последовательностями адапторов, в дальнейшем использовавшихся в реакциях быстрой амплификации концов кДНК (RACE, от англ. rapid amplification of cDNA ends).

#### **2.4.3 Амплификация кДНК и фрагментов геномной ДНК и секвенирование**

Основываясь на последовательностях кДНК из библиотеки ядовитых желез *O. takobius*, были подобраны специфические праймеры (табл. 4) для амплификации кДНК, кодирующей двудоменные токсины *O. lineatus*. 5' и 3'RACE проводили с помощью универсального праймера T7сар и соответствующего специфического праймера. Продукты ПЦР лигировали с вектором рAL-ТА, клонировали и секвенировали. Используя вновь полученные последовательности, подбирали праймеры (табл. 4) для амплификации фрагментов геномной ДНК, кодирующих двудоменные токсины *O. lineatus*.

В случае с *C. punctori*um, праймеры (табл. 4) для амплификации фрагментов геномной ДНК, кодирующих двудоменные токсины, подбирали на основании последовательностей из библиотеки кДНК из ядовитых желез этого паука.

Амплификацию проводили с помощью коммерчески доступного набора для ПЦР (Евроген). Очищенные продукты ПЦР лигировали с вектором рAL-TA и использовали для трансформации компетентных клеток *E. coli*, XL1-Blue. Положительные клоны секвенировали с использованием прямого праймера M13. Для анализа результатов секвенирования и других манипуляций с последовательностями использовали пакет программ Lasergene (DNASTAR, США). Полученные последовательности были зарегистрированы в банке нуклеотидных последовательностей GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>).

	Ген	Матрица	Направление праймера	Последовательность
<i>O. lineatus</i>	<i>OITx1</i>	гДНК	Прямой	CGATTTCAAAATGAAGTTCTCTTTGG
			Обратный	TTGAGCAAGAATTCTATACAACCTG
		кДНК	Прямой	AAGGGTCTAGAGAAAGCAACACC
			Обратный	ATACAACCTGGGAGGTCACAGGTG
	<i>OITx2</i>	гДНК, кДНК	Прямой	AACATCGTTTTCAAAATGAAGATCG
			Обратный	AGTCTACTCAGCCGCAGGTTCCCT
<i>OITx1, OITx1</i>	кДНК	Универсальный праймер T7cap	GTAATACGACTCACTATAGGGCAAGCA GTGGTAACAACGCAGAGT	
<i>C. punctori</i> <i>um</i>	<i>SpTx 1c</i>	гДНК	Прямой	CGATGAAGTTTTTCGTTGTTCTTCAG
			Обратный	GATCATCTATCTCCAAACCAGCG
	<i>cptx 3a</i>	гДНК	Прямой	GATCCATAGCTGTTAATCTAGG
			Обратный	ATCGGAAGCTGAGTTAAGTTGAG

**Таблица 4. Синтетические олигонуклеотидные праймеры, использованные для амплификации кДНК и фрагментов геномной ДНК (гДНК), кодирующих двудоменные токсины *O. lineatus* и *C. punctori*um.**

#### 2.4.4 Анализ нуклеотидных последовательностей

Для трансляции последовательностей кДНК *in silico* использовали программное обеспечение DNASTAR. Сигнальные пептиды определяли с помощью программы SignalP 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>). Последовательности, предшествовавшие сайту процессинга PQM были приняты за пропептиды, следующие после него – за зрелые пептиды. С-концевые положительно-заряженные аминокислоты удаляли (имитируя активность карбоксипептидаз), остаток глицина считали сигналом С-концевого амидирования [86].

Общие и попарные выравнивания последовательностей строили с помощью программы ClustalW [196] в составе пакетов Vector NTI Suite 8 (Life Technologies, США) и MEGA 5 [197]. Поиск возможных сайтов сплайсинга осуществляли с помощью сервиса Fruitfly Splice Predictor ([http://www.fruitfly.org/seq\\_tools/splice.html](http://www.fruitfly.org/seq_tools/splice.html)) [198].

Следующие виды анализа выполнялись с помощью MEGA 5. Филогенетические деревья строили методом ближайшего соседа [199]. Для подсчета доли повторяющихся деревьев, в которых родственные последовательности образовывали одинаковые кластеры (уровень поддержки), применяли пороговый тест (500 повторов). Эволюционные расстояния рассчитывали методом квазикаксимального правдоподобия [200]. Z-тест на отрицательный/положительный отбор проводили методом Нея-Годжобори [201].

Последующий филогенетический анализ выполняли методом максимального правдоподобия с помощью программы CODEML из пакета PAML (версия 4.7a) [202] с графическим интерфейсом PAMLX [203]. Все разрывы последовательностей в выравниваниях исключались из анализа. Для попарных сравнений программу запускали в режиме `runmode=-2` и моделью M0. Для проведения теста отношения правдоподобия (LRT, от англ. likelihood ratio test) программу запускали с `runmode=-2` и пятью моделями M0, M1a, M2a, M7 и M8. Длины ветвей оценивали с помощью модели M0 и использовали во всех последующих анализах. Натуральный логарифм функции максимального правдоподобия ( $\ln L$ ) рассчитывали для каждой из моделей. Тесты LRT составляли для пар M1a/M2a и M7/M8, для каждой из которых рассчитывали  $2\Delta l = 2 \times (\ln L - \ln L_0)$ , где  $L_0$  – модель, не допускающая положительного отбора (M1a и M7). Значения  $2\Delta l$  сравнивали с таблицей значений  $\chi^2$  – для двух степеней свободы (значения  $\chi^2$  для  $p < 0.05$  и  $p < 0.1$  составляют 5.9915 и 4.6052, соответственно).

Поиск гомологичных последовательностей проводили с помощью программы BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) в базах данных нуклеотидных и невырожденных белковых последовательностей. В качестве запросов использовали нуклеотидные и аминокислотные последовательности отдельных доменов двудоменных токсинов.

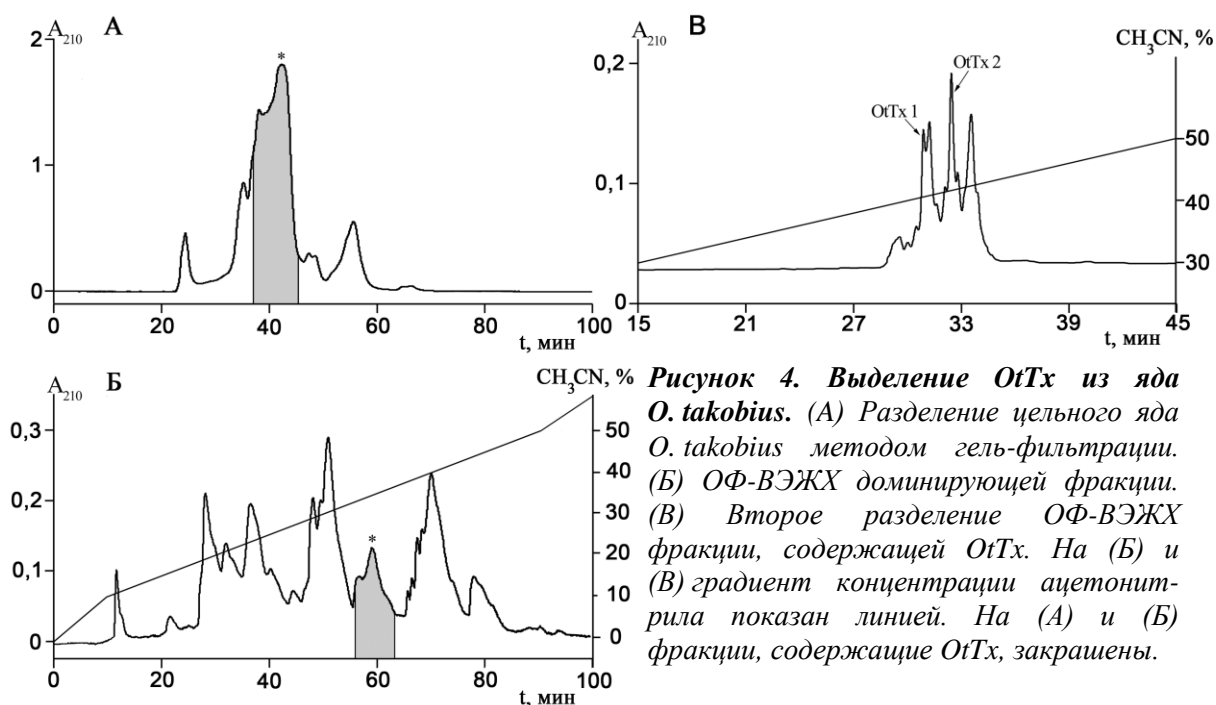
## Результаты

### 1. Токсины пауков рода *Oxyopes*

В этой главе описан новый класс двудоменных токсинов пауков, впервые выделенных из яда *O. takobius* – спайдеринов. Были охарактеризованы их структура, активность, а также гены, кодирующие и механизмы молекулярной эволюции.

#### 1.1 Выделение двудоменных токсинов

Выделение инсектицидных токсинов из цельного яда *O. takobius* проводилось по двустадийной схеме – сначала яд подвергали гель-фильтрации, а затем наиболее представленную пептидную фракцию разделяли методом ОФ-ВЭЖХ. В результате были выявлены новые компоненты с относительно высокой молекулярной массой (12-12,5 кДа). Элюция этих компонентов с обратной фазы (С4 – С5) происходит при концентрации ацетонитрила 35 – 40%. В дальнейшем они оказались эффективны в тестах на инсектицидную активность. С помощью повторной хроматографии два пептида были очищены до состояния 95% чистоты и названы OtTx 1 (12030 Да) и OtTx 2 (12415 Да) (рис. 4).



## 1.2 Определение последовательности OtTx

С помощью N-концевой деградации по Эдману была определена частичная аминокислотная последовательность выделенных токсинов: KFKWGLFSTAKKLYKKGKGLSKNKNFKKALKFGK (OtTx 1) и KFKLPKINWGKLASKAKDVYKKGQKLAKNKNVKKALK (OtTx 2). Ранее в сотрудничестве с компанией Дюпон была получена библиотека кДНК из ядовитых желез паука *O. takobius*. Последовательности кДНК были транслированы *in silico*, и среди них проводился поиск специфических структурных элементов и мотивов, характерных для секретируемых полипептидов и токсинов пауков [86,90]. В результате были определены вероятные аминокислотные последовательности токсинов. Используя полученные частичные N-концевые последовательности в качестве поискового запроса, мы выявили полные последовательности токсинов OtTx 1 и OtTx 2. Кроме последовательностей, фрагмент которых точно совпал с установленными N-концевыми последовательностями, мы также нашли их близкие гомологи. В результате, изоформы токсинов, обнаруженные в яде, были обозначены как OtTx 1a и OtTx 2a, а гомологичные им последовательности как OtTx 1b и OtTx 2b (номера в UniProt P86716 – P86719, в GenBank – JX134894 – JX134897). OtTx 1b отличается от OtTx 1a единичной заменой аланина на треонин в 10 положении. В случае OtTx 2a и OtTx 2b произошла замена фенилаланина на лейцин в положении 4. Расчетные молекулярные массы OtTx 1a (12031 Да) и OtTx 2a (12416 Да) хорошо согласуются со значениями, полученными в результате измерения (см. далее).

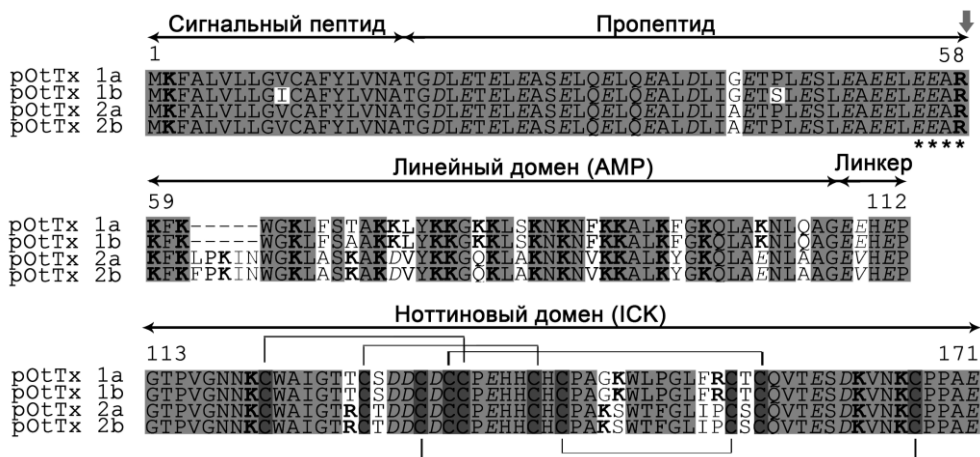
Во всех обнаруженных токсинах содержится по 10 остатков цистеина. Алкилирование нативных OtTx винилпиридином оказалось невозможным, однако при алкилировании после восстановления дисульфидных связей дитиотреитолом молекулярная масса возрастала до значений 13090 Да (OtTx 1a) и 13475 Да (OtTx 2a). Такой результат говорит о том, что все 10 остатков цистеина участвуют в образовании 5 внутримолекулярных дисульфидных связей, и поэтому алкилирование пептидов возможно только после их восстановления.

## 1.3 Анализ последовательности OtTx

OtTx 1 (108 а. о.) и OtTx 2 (113 а. о.) высокогомологичны, их последовательности совпадают на 80% (рис. 5). OtTx 2 длиннее, чем OtTx 1 на 5 аминокислотных остатков. При детальном изучении последовательностей новых токсинов выяснилось, что каждая из них состоит из двух частей (модулей), или доменов. N-Концевая часть (остатки 1 – 44 у

OtTx 1 и 1 – 49 у OtTx 2) не содержит остатков цистеина, в то время как С-концевой домен (остатки 50 – 108 у OtTx 1 и 55 – 113 у OtTx 2) является цистеин-богатым: 10 остатков цистеина составляют 17% его последовательности.

Короткая линкерная последовательность, соединяющая два модуля (EENEP, остатки 45 – 50 у OtTx 1; и EVNEP, остатки 50 – 55 у OtTx 2) напоминает мотив процессинга PQM (в предшественниках OtTx PQM=EEAR). В линкерной последовательности OtTx наблюдается мутировавшая версия PQM, в которой произошла замена аргинина на пролин.



**Рисунок 5. Сравнение последовательностей белков-предшественников токсинов OtTx.** Отрицательно заряженные остатки выделены курсивом, положительно заряженные – жирным шрифтом, остатки цистеина показаны на темно-сером фоне, идентичные остатки – на светло-сером. Области, соответствующие сигнальному пептиду, пропептиду и фрагментам зрелого пептида отмечены стрелками над последовательностями. Сайт процессинга пропептида показан жирной стрелкой, PQM мотив – звездочками. Дисульфидные связи, участвующие в образовании канонического мотива ICK, обозначены горизонтальными скобками над последовательностями, дополнительные дисульфиды – скобками под последовательностями.

В результате поиска гомологичных пептидов в базе данных UniProt было обнаружено два токсина, проявляющих значительное сходство с С-концевым доменом OtTx. Оба этих пептида являются антагонистами потенциал-зависимых кальциевых каналов: окситоксин 1 (OxyTx 1) из яда *O. takobius* (также он был найден в яде *Oxyopes lineatus*) [32] и окситоксин 2 (OxyTx 2) из яда *O. lineatus* [64]. Последовательность С-концевого домена OtTx 1 идентична OxyTx 1 на ~40%, а OxyTx 2 на ~45% (рис. 6). Более того, окситоксины и С-концевой домен OtTx характеризуются общим порядком расположением цистеинов, что указывает на то, что их пространственные структуры также сходны: C1X6C2X3C3XC4C5X4C6XC7XmC8XC9XnC10, где m = 11 для OtTx и OxyTx 2, а для OxyTx 1 m=24, n=9 для OxyTx 2, n =10 для OtTx и n =11 для OxyTx 1. Эта последовательность может быть сведена до структурных мотивов PSM (в нашем случае,

C1X6C2X5C4C5) и ESM (в нашем случае, C6XC7XmC8XC9). В свою очередь, PSM и ESM сводятся к более общему мотиву ICK, который, в нашем случае, выражается как C1X6C2X5C4C5X4C6Xm+4C9. Таким образом, в дополнение к трем дисульфидам C1 – C5, C2 – C6, C4 – C9, формирующим ICK мотив, есть две дополнительные S-S связи C3 – C10 и C7 – C8, причем последняя входит в мотив ESM и наблюдается также у большинства токсинов пауков, содержащих 8 остатков цистеина. Более низкий уровень сходства OtTx (~38% идентичных остатков в случае OtTx 1) обнаруживается с токсинами Tx3-6 [204] и PRTx23C2 [205] (из пауков *Phoneutria nigriventer* и *Phoneutria reidy*, соответственно), которые также являются потенциальными блокаторами кальциевых каналов. Интересно, что в этих токсинах из яда *Phoneutria* наблюдается дополнительная шестая дисульфидная связь. Другая интересная особенность С-концевого домена OtTx – это высокое содержание пролина (~10%).

В открытых базах данных не было обнаружено последовательностей, гомологичных линейному домену OtTx. Этот домен положительно заряжен (заряд +16 и +13 при pH 7 для OtTx 1 и OtTx 2, соответственно; pI ~11) и, согласно расчетам, проявляет склонность к формированию  $\alpha$ -спирали. Положительно заряженные и гидрофобные остатки расположены таким образом, что при спирализации полипептидной цепи они образуют кластеры. Такие свойства характерны для мембраноактивных и антимикробных пептидов из ядов различных пауков, в том числе *O. takobius* [32,37]. Интересно, что некоторые остатки представлены гораздо шире, чем другие. Так, в OtTx 1a доля остатков Lys составляет ~36%, в то время как другие заряженные остатки Asp, Arg, и Glu вообще не представлены; доля Leu и Phe составляют ~14% и ~9%, соответственно, в то время как Ile, Met, и Val отсутствуют.

Таким образом, OtTx образуют новый класс токсинов пауков, объединяющих N-концевой линейный и С-концевой ICK домены.

OtTx 1a-N	<b>KFKW</b> <b>GK</b> L <b>F</b> ST <b>AK</b> KL <b>YK</b> <b>GK</b> KL <b>S</b> KN <b>KN</b> - <b>FK</b> AL <b>K</b> FG <b>KQ</b> L <b>AK</b> N <b>LQ</b> AG	100 %
CIT 1a-N	---G <b>F</b> FN <b>T</b> W <b>K</b> K <b>T</b> <b>K</b> G <b>K</b> AD <b>K</b> IM <b>L</b> --- <b>KK</b> AV <b>K</b> IM <b>V</b> <b>K</b> EG <b>I</b> S <b>K</b> ---	29 %
Oxyopinin 2a	---G <b>K</b> - <b>F</b> SV <b>F</b> G <b>K</b> L <b>L</b> R <b>S</b> T <b>AK</b> V <b>F</b> <b>K</b> GV <b>G</b> K <b>V</b> R <b>K</b> O <b>F</b> K <b>T</b> AS <b>D</b> L <b>D</b> K <b>N</b> O---	32 %
OtTx 1a-C	G <b>T</b> P <b>V</b> G <b>N</b> N <b>K</b> W <b>A</b> I <b>G</b> T <b>T</b> <b>S</b> D <b>D</b> C <b>D</b> C <b>P</b> E <b>H</b> H <b>C</b> H <b>C</b> E <b>A</b> G----- <b>K</b> W <b>L</b> P <b>G</b> L <b>F</b> R <b>T</b> <b>T</b> O <b>V</b> T <b>E</b> S <b>D</b> K <b>V</b> N <b>K</b> L <b>P</b> E <b>A</b> E- <b>---</b>	100 %
Oxytoxin-2	--- <b>A</b> W <b>K</b> L <b>P</b> <b>K</b> D <b>S</b> T <b>G</b> D <b>D</b> C <b>D</b> C <b>E</b> G <b>L</b> H <b>C</b> H <b>C</b> P <b>L</b> R----- <b>N</b> M <b>L</b> P <b>A</b> I <b>L</b> R <b>C</b> S <b>C</b> O <b>S</b> - <b>K</b> D <b>D</b> H <b>I</b> N <b>T</b> O <b>P</b> K <b>Y</b> K <b>S</b>	45 %
Oxytoxin-1	--- <b>D</b> W <b>E</b> L <b>L</b> P <b>L</b> H <b>S</b> S <b>D</b> N <b>D</b> C <b>V</b> <b>C</b> <b>K</b> N <b>H</b> H <b>C</b> H <b>C</b> P <b>S</b> N <b>V</b> S <b>K</b> L <b>E</b> <b>K</b> W <b>L</b> P <b>E</b> W <b>A</b> K <b>I</b> P <b>D</b> A <b>L</b> <b>K</b> R <b>C</b> S <b>C</b> O <b>R</b> N <b>D</b> <b>K</b> D <b>G</b> K <b>I</b> N <b>T</b> O <b>D</b> K <b>Y</b> K <b>N</b>	39 %

**Рисунок 6. Сравнение последовательностей OtTx 1a с другими пептидами из яда пауков.** Вверху N-концевой модуль OtTx 1a (OtTx 1a-N) показан в сравнении с N-концевым фрагментом цитотоксина 1a (CIT 1a-N) и оксиопинином 2a. Внизу: С-концевой модуль токсина (OtTx 1a-C) в сравнении с окситоксинами 1 и 2. Положительно заряженные остатки выделены жирным шрифтом. Идентичные а. о. отмечены светло-серым фоном, консервативные остатки цистеина – темно-серым. Доля остатков, идентичных OtTx 1a, показана справа.

## 1.4 Структура белка-предшественника OtTx

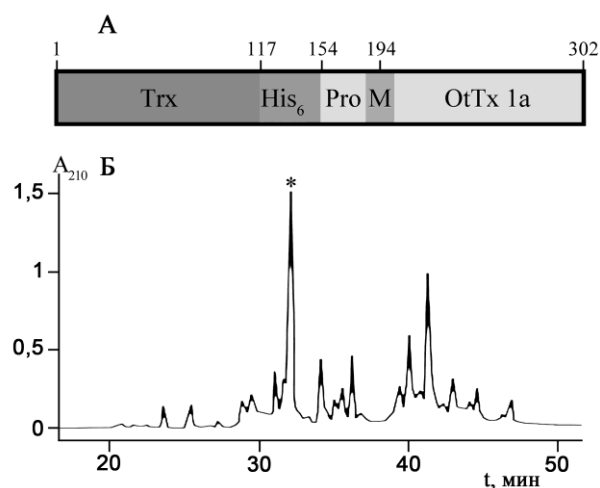
Анализ транслированной библиотеки кДНК из ядовитых желез *O. takobius* позволил выявить последовательности белков-предшественников токсинов OtTx. Предшественник OtTx 1 состоит из 166 остатков и состоит из трех фрагментов. N-концевой сигнальный (пре-) пептид (18 остатков) был определен с помощью программы SignalP. Пропептид (40 остатков) следует после сигнального пептида и заканчивается последовательностью 37EEAR40, являющейся сайтом процессинга PQM и отделяющей его от зрелого пептида. Препропептид – консервативная часть предшественника: между последовательностями предшественников OtTx 1a и OtTx 2a наблюдается только одна аминокислотная замена. Про-последовательность имеет кислотный характер (при pH 7 ее заряд равен -14, pI ~3). Предполагается, что взаимодействуя с положительно заряженным N-концевым линейным доменом зрелой цепи, она снижает его цитолитическую активность.

## 1.5 Получение полноразмерного OtTx 1a и его фрагментов

### 1.5.1 Получение OtTx 1a

Для получения необходимого для функциональных тестов материала была разработана система получения рекомбинантного токсина OtTx 1a в культуре *E. coli*. Поскольку токсин содержит 5 дисульфидных связей, его получали в виде химеры с белком-помощником тиоредоксином (Ttx-OtTx 1a), который *in vivo* способствует формированию правильной пространственной укладки у дисульфид-богатых белков. Для того чтобы снизить возможную мембранолитическую активность N-концевого домена OtTx 1a, в состав химерного белка был введен природный пропептид из белкового предшественника OtTx 1a. В связи с тем, что зрелый OtTx 1a не содержит остатков метионина, использовалось специфическое расщепление с помощью BrCN по остатку метионина, искусственно введенному между последовательностями пропептида и OtTx 1a. Химерный белок Ttx-OtTx 1a также содержал последовательность из 6 остатков гистидина, позволившую очищать белок методом металл-хелатной аффинной хроматографии. Схема химерного белка приведена на рис. 7, А. После расщепления химерного белка BrCN целевой полипептид подвергался очистке с помощью ОФ-ВЭЖХ (рис. 7, Б). Совпадение молекулярных масс, инсектицидной активности и хроматографической подвижности рекомбинантного и нативного токсинов свидетельствовало в пользу их идентичности. Выход составил 0,5 мг/л культуры.





**Рисунок 7. Структура гибридного белка Trx-OtTx 1a (А) и очистка методом ОФ-ВЭЖХ рекомбинантного пептида OtTx 1a после расщепления Trx-OtTx 1a BrCN (Б).** Функциональные элементы Trx-OtTx 1a показаны схематично прямоугольниками. Trx – домен тиоредоксина; His<sub>6</sub> – олигогистидиновая последовательность; М – остаток метионина, введенный непосредственно перед зрелым полипептидом OtTx 1a. Числа над схемой соответствуют номерам а. о. химерного белка. (Б) Фракция, содержащая OtTx 1a, отмечена звездочкой.

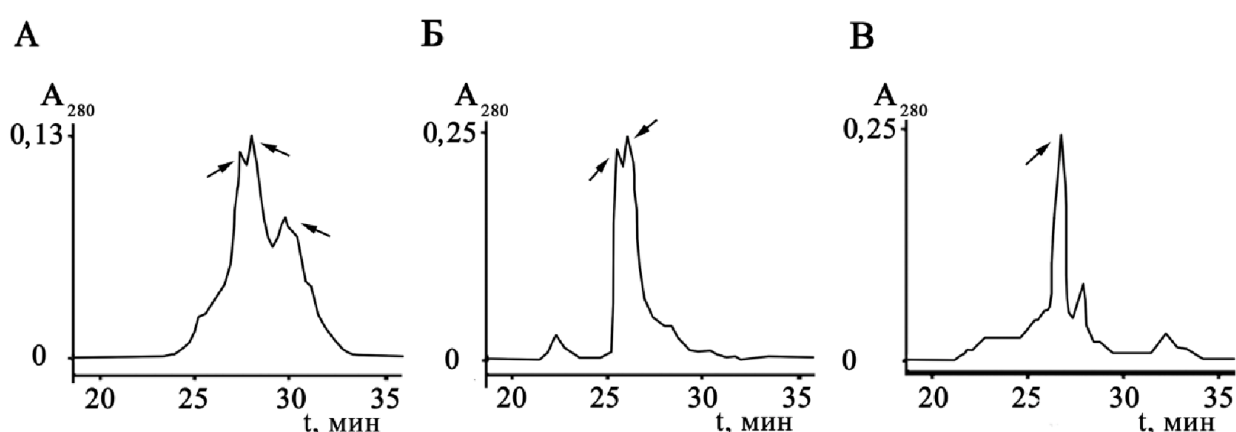
### 1.5.2 Получение OtTx 1a-AMP

Были получены укороченные производные токсина OtTx 1a, соответствующие его отдельным доменам. Не только линкерная последовательность напоминает PQM, но также можно заметить и мутировавшую версию инвертированного мотива PQM (iPQM), который также является сайтом процессинга в сложных предшественниках токсинов [90]. Вероятно, что в нашем случае произошла замена остатка аргинина, необходимого для расщепления предшественника, на глутамин в положении 42. Основываясь на этом, была выведена последовательность, соответствующая «немутировавшей процессированной» версии N-концевого домена OtTx 1a (остатки 1 – 41). Поскольку с большой долей вероятности пептид, соответствующий N-концевому фрагменту OtTx 1a, обладает антимикробной активностью, он был назван OtTx 1a-AMP. OtTx 1a-AMP был синтезирован методами классического химического синтеза в лаборатории протеомики ИБХ РАН.

### 1.5.3 Получение OtTx 1a-ICK

В связи с тем, что в первичной структуре С-концевого домена (остатки 50 – 108) был обнаружен характерный мотив цистинового узла, его назвали OtTx 1a-ICK. Способ получения OtTx 1a-ICK аналогичен тому, как это было описано выше для OtTx 1a. OtTx 1a-ICK получали с помощью генной инженерии в системе *E.coli* в виде химерного белка с тиоредоксином (Trx-OtTx 1a-ICK). В составе этого белка также была последовательность из 6 остатков гистидина для его выделения из клеточного лизата методом аффинной хроматографии. Для расщепления Trx-OtTx 1a-ICK с помощью BrCN между аминокислотными последовательностями тиоредоксина и целевого пептида был введен остаток метионина.

Были проверены разные условия получения Тгх-ОтТх 1а-ІСК: варьировались экспрессионные штаммы *E.coli*, температура и продолжительность инкубации, концентрация индуктора ИПТГ. В результате, в связи с тем, что пептид обладает 5 дисульфидными связями, оптимальным оказался штамм Origami B. Экспрессию проводили при +22° С в течение ночи. Уровень продукции химерного белка Тгх-ОтТх-ІСК был высоким (выход ~100 мг/л). После расщепления гибридного белка на хроматографическом профиле несколько пиков по молекулярной массе их компонентов соответствовали ОтТх 1а-ІСК, что говорит об образовании нескольких конформаций пептида в результате некорректного замыкания дисульфидных связей. Эксперименты по окислению тиольных групп ОтТх 1а-ІСК после расщепления гибридного белка не увенчались успехом, поэтому для получения пептида в правильной конформации процедуре рефолдинга подвергали гибридный белок. Для контроля за ходом реакции отбирали пробы через 1, 2 и 5 суток, проводили расщепление BrCN и анализировали методом ОФ-ВЭЖХ. О правильном замыкании дисульфидных мостиков судили по форме хроматографического профиля. На рис. 8 видно, что в течение нескольких суток происходит постепенное уменьшение количества пиков, соответствующих разным конформациям ОтТх 1а-ІСК, и к пятым суткам остается только один пик. Молекулярная масса полученного таким образом ОтТх 1а-ІСК точно соответствовала расчетной массе с учетом того, что дисульфидные мостики замкнуты. Выход продукта составил 5 мг/л культуры, что значительно превышает выход полноразмерного рекомбинантного токсина, поскольку данный полипептид лишен домена с цитолитической активностью.



**Рисунок 8** Три стадии рефолдинга ОтТх 1а-ІСК. Профиль ОФ-ВЭЖХ после (А) 1 суток, (Б) 2 суток, (В) 5 суток инкубации. Стрелками обозначены разные конформации целевого пептида.

## 1.6 Вторичная структура OtTx 1a

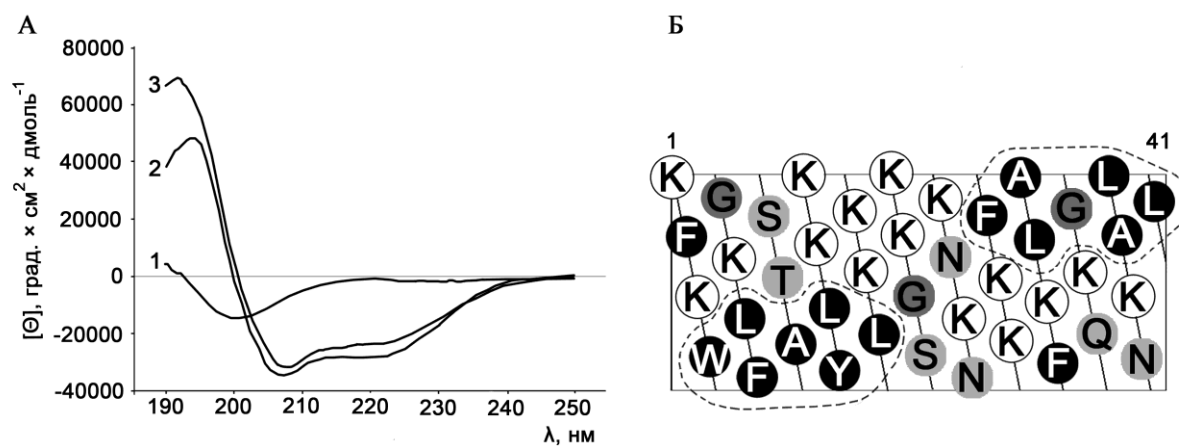
Алгоритмы предсказания вторичной структуры выделяют два домена в составе OtTx 1a. N-концевой фрагмент, состоящий из ~45 аминокислотных остатков, склонен к формированию  $\alpha$ -спирали, в то время как для C-концевого фрагмента характерна высокая склонность к формированию  $\beta$ -поворотов. Линкерная область (остатки 46 – 55), скорее всего, не упорядочена.

Расчеты, полученные методами биоинформатики, подтверждаются при измерении спектров кругового дихроизма (рис. 9). Расчет долей разных типов вторичной структуры в различных растворителях для полноразмерного OtTx 1a, а также для его укороченных производных приведен в табл. 5.

Пептид	Раствор	Вторичная структура, доля (%)			
		$\alpha$ -спираль	$\beta$ -слой	$\beta$ -поворот	неупорядоченная
OtTx 1a	Фосфатный буф.	9,0	29,9	24,0	37,0
	20 мМ ДСН	47,8	7,0	15,0	30,1
	50% ТФЭ	48,4	17,7	19,9	14,1
OtTx 1a-AMP	Фосфатный буф.	7,1	33,7	23,6	35,6
	20 мМ ДСН	81,3	2,0	1,0	15,7
	50% ТФЭ	73,5	1,7	6,2	18,6
OtTx 1a-ICK	Фосфатный буф.	18,6	30,6	21,8	28,9
	20 мМ ДСН	21,5	24,6	22,6	31,3
	50% ТФЭ	17,8	42,2	16,8	23,1

*Таблица 5. Доли типов вторичной структуры OtTx 1a и его производных OtTx 1a-AMP и OtTx 1a-ICK, полученные методом кругового дихроизма*

Высокое содержание  $\beta$ -слоев и  $\beta$ -поворотов в OtTx 1a согласуется с предположением о том, что для этого токсина характерен ноттиновый фолд. То же самое верно для OtTx-ICK, который соответствует C-концевому домену токсина. Более того, в окружении, имитирующем мембранное, вторичная структура OtTx 1a-ICK меняется крайне незначительно. Для OtTx 1a-AMP, соответствующего линейному домену, наоборот, характерна зависимость конформации от характера растворителя. Так, в фосфатном буфере доля  $\alpha$ -спирали ниже 10%, в то время как в растворах ТФУ и ДСН, имитирующих мембранной окружение, содержание  $\alpha$ -спиральной конформации превышает 70%. Соответствующие изменения также наблюдаются в полноразмерном токсине OtTx 1a.



**Рисунок 9.** (А) Спектры кругового дихроизма и (Б) проекция спиральной сети пептида *OtTx 1a-AMP*. (А) Кривая 1 – спектр пептида в 50 мМ фосфатном буфере (рН 7,2), кривая 2 – спектр пептида в 50% ТФЭ, кривая 3 – в 20 мМ ДСН. (Б) Положительно заряженные остатки обозначены белыми кругами, гидрофобные остатки – черными, остатки глицина – темно-серыми, гидрофильные незаряженные остатки – светло-серыми. Гидрофобные кластеры обведены пунктирными линиями.

### 1.7 Биологическая активность *OtTx 1a* и его производных

Токсин *OtTx 1a* проявляет инсектицидную и цитолитическую активности. В тестах на личинках мясной мухи ЛД<sub>50</sub> составила 75 мкг/г (6,3 нмоль/г), а значения минимальной ингибирующей концентрации (МИК) находятся в диапазоне 1 – 10 мкМ (табл. 6). Для выяснения функциональных ролей доменов были протестированы активности *OtTx 1a-AMP* и *OtTx-ICK*.

Было установлено, что *OtTx 1a-AMP* проявляет те же виды активности, что и полноразмерный токсин: ЛД<sub>50</sub> на личинках мясной мухи составил 50 мкг/г (11,1 нмоль/г), а значения МИК для разных видов бактерий попадали в тот же диапазон концентраций. Однако в тех же тестах *OtTx 1a-ICK* активности не показал. Кроме того, было обнаружено, что *OtTx 1a-AMP* способен индуцировать лизис эритроцитов. Его эффективная концентрация, вызывающая 50 %-ный выход гемоглобина, составляет ~8 мкМ. На клеточных линиях A549 и HeLa 50%-ная гибель клеток наблюдается при концентрациях ~20 мкМ и ~15 мкМ, соответственно.

Способность пептида *OtTx 1a-AMP* взаимодействовать с клеточными мембранами подтверждалась с помощью его соосаждения с липосомами. В результате было установлено, что он связывается как с отрицательно заряженными ДОФГ/ДОФЭ липосомами, имитирующими бактериальные мембраны, так и с цвиттерийными ДОФХ липосомами, подобными эукариотическим мембранам.

*Механизм антибактериальной активности OtTx 1a-AMP*

В дальнейших тестах было установлено, что OtTx 1a-AMP не только ингибирует бактериальный рост, но и обладает бактерицидной активностью. Для *S. aureus* и *E. faecalis* значения минимальной бактерицидной концентрации (МБК) и МИК совпадают (табл. 6), бактерицидный эффект OtTx 1a-AMP развивается быстро. Так, инкубация *S. aureus* в течение 15 минут в среде, содержащей пептид в концентрации, равной двукратной МБК, приводит к полной гибели бактериальных клеток. После инкубации *S. aureus* в течение 15 – 30 минут в среде, содержащей 1,25 мкМ OtTx 1a-AMP, при высевании на агаризованную питательную среду вырастает не более 2 – 3 колоний на чашку Петри. После более продолжительной инкубации (более 1 ч) жизнеспособных бактерий не остается.

Бактериальный штамм	Пептиды		
	OtTx 1a	OtTx 1a-AMP	OtTx 1a-ICK
	МИК, мкМ		
<i>A. globiformis</i> ВКМ Ac-1112	0,6 – 1,25	0,6 – 1,25	>25
<i>B. subtilis</i> ВКМ В-501	0,12 – 0,25	0,12 – 0,25	>25
<i>E. faecalis</i> ВКМ В-871	–	5 – 10	–
<i>E. coli</i> DH5	0,12 – 0,25	0,12 – 0,25	>25
<i>P. aeruginosa</i> PAO1	6 – 12	0,37 – 0,75	>25
<i>S. aureus</i> 209P	–	0,6 – 1,25	–

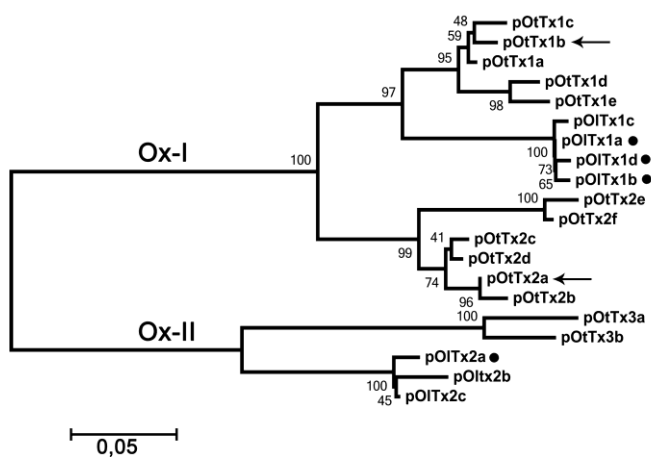
**Таблица 6. Антимикробная активность OtTx 1a и его модулей**

Бактерицидное действие OtTx 1a-AMP связано с образованием дефектов в цитоплазматической мембране. Инкубация клеток *S. aureus* с OtTx 1a-AMP приводила к одновременному вытеканию БКЭКФ из цитоплазмы и проникновению ИП внутрь клетки, которое сопровождалось связыванием с ДНК и усилением флуоресценции. Интактные бактериальные мембраны не пропускают БКЭКФ наружу и предотвращают вход ИП.

## **2. Гены двудоменных токсинов пауков рода *Oxyopes***

### **2.1 Анализ последовательностей кДНК из ядовитых желез пауков *Oxyopes***

Выше были описаны 4 последовательности кДНК, кодирующие спайдерины *O. takobius*. Однако, как оказалось, разнообразие двудоменных токсинов гораздо шире: в дальнейшем были обнаружены 9 новых последовательностей семейства спайдеринов в библиотеке кДНК из ядовитых желез *O. takobius* и 6 новых последовательностей кДНК, кодирующих двудоменные токсины *O. lineatus* (OtTx) (номера в GenBank KF766543 – KF766559, KF766561) (рис. 10).



**Рисунок 10.** Филогенетическое дерево, построенное на основе сравнения последовательностей предшественников спайдеритоксинов. Масштабная линейка соответствует количеству нуклеотидных замен на один сайт вдоль ветви. Уровни поддержки указаны рядом с соответствующими ветвями. Стрелками указаны предшественники токсинов, выделенных из яда *O. takobius*, жирными точками – те, для которых также была установлена последовательность соответствующих фрагментов геномной ДНК.

Все последовательности кДНК были транслированы и обработаны *in silico* (пример для OтTx 2a см. на рис. 11 (как будет показано далее, кодирующие последовательности на уровне кДНК и геномной ДНК совпадают)). В общей сложности были обнаружены 15 новых последовательностей токсинов, относящихся к семейству спайдеритоксинов. Как и многие токсины, описанные ранее [80,84], у спайдеритоксинов разный уровень варибельности сигнальных, про- и зрелых пептидов. Сигнальные пептиды и пропептиды идентичны друг другу, либо проявляют высокий уровень сходства (>74% и >62% идентичных остатков, соответственно), в то время как зрелые фрагменты более варибельны (>42% идентичных а. о.). Зрелые пептиды можно объединить в два подсемейства, различающихся в длину: Oх-I (109 – 114 а. о.) и Oх-II (135 – 142 а. о.). Кроме OтTx 1a, 1b, 2a и 2b, описанных выше, к подсемейству Oх-I были отнесены 10 новых последовательностей: OтTx 1c – 1e, OтTx 2c – 2f, и OтTx 1a – 1c. OтTx 3a и 3b и OтTx 2a – 2c были помещены в подсемейство Oх-II. В пределах каждого подсемейства доля идентичных остатков между токсинами превышает 68%, в то время как между подсемействами этот показатель находится в диапазоне 42 – 50%. Все белки-предшественники группы Oх-I напоминают сложные белки-предшественники с мутировавшим сайтом процессинга между двумя зрелыми цепями, однако ни один из полипептидов Oх-II не содержит последовательностей с высоким уровнем сходства с сайтом процессинга RQM. Разница в длинах между семействами Oх-I и Oх-II обусловлена разницей в длинах N-концевых доменов зрелых полипептидов: они состоят из 47 – 55 и 69 – 73 а. о., соответственно (рис. 12). В линейных доменах обоих подсемейств можно выделить «базовые» последовательности, демонстрирующие умеренное, но значимое сходство друг с другом (как минимум, 30% идентичных остатков). N-концевые домены зрелых пептидов Oх-II обладают «дополнительными» последовательностями, увеличивающими их длину и содержащими несколько повторов Gly-Lys-Ser. Ноттиновые

домены более консервативны (как минимум 59% идентичных а. о.), а у некоторых токсинов они даже идентичны (как на уровне белка, так и на уровне кДНК).

## 2.2 Структура генов спайдеринов

С помощью амплификации и секвенирования фрагментов геномной ДНК была определена последовательность генов, кодирующих токсины О1Тх 1а, О1Тх 1b и О1Тх 2а (рис. 11). Последовательность, кодирующая О1Тх 1d (номер в GenBank KF766560) была найдена только на уровне геномной ДНК, но не кДНК. Как и большинство генов однодоменных токсинов, гены двудоменных токсинов *O. lineatus* оказались безинтронными.

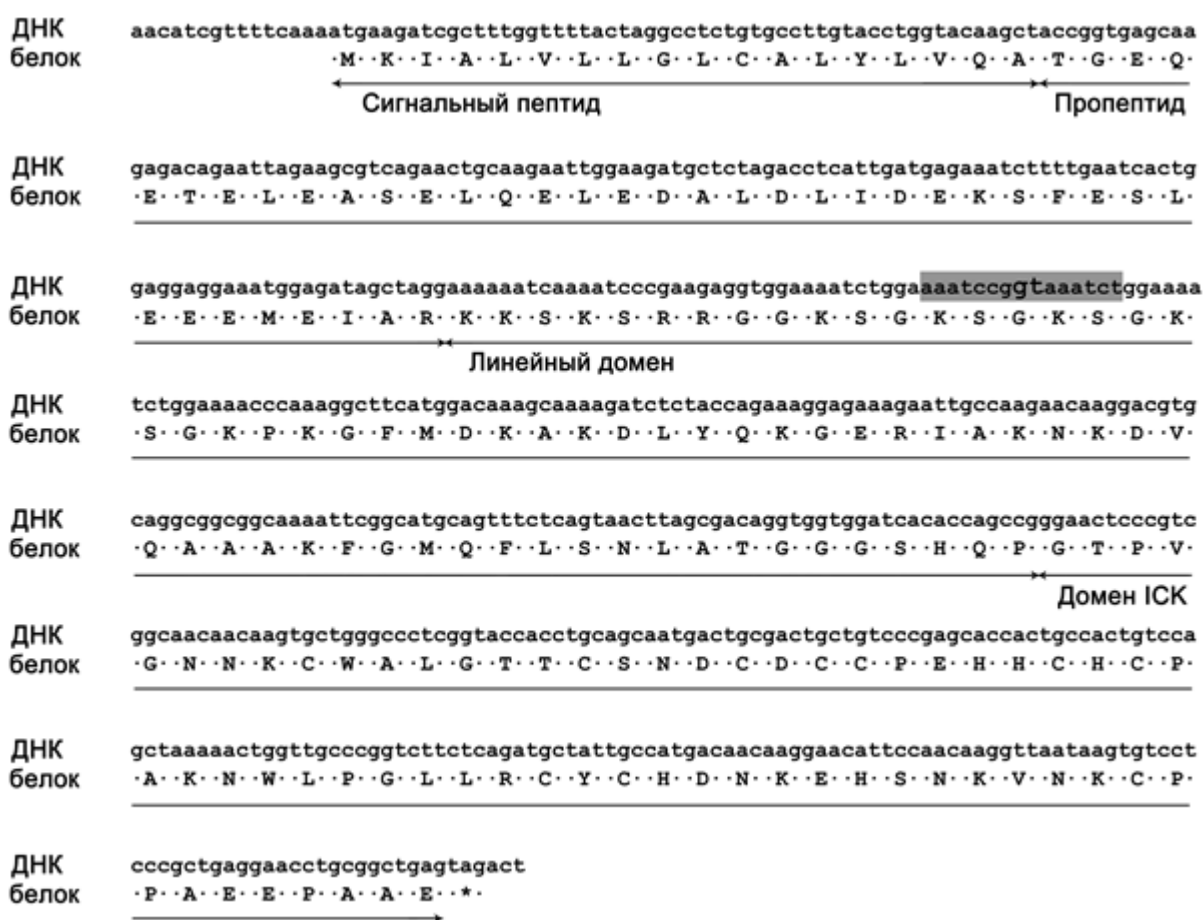
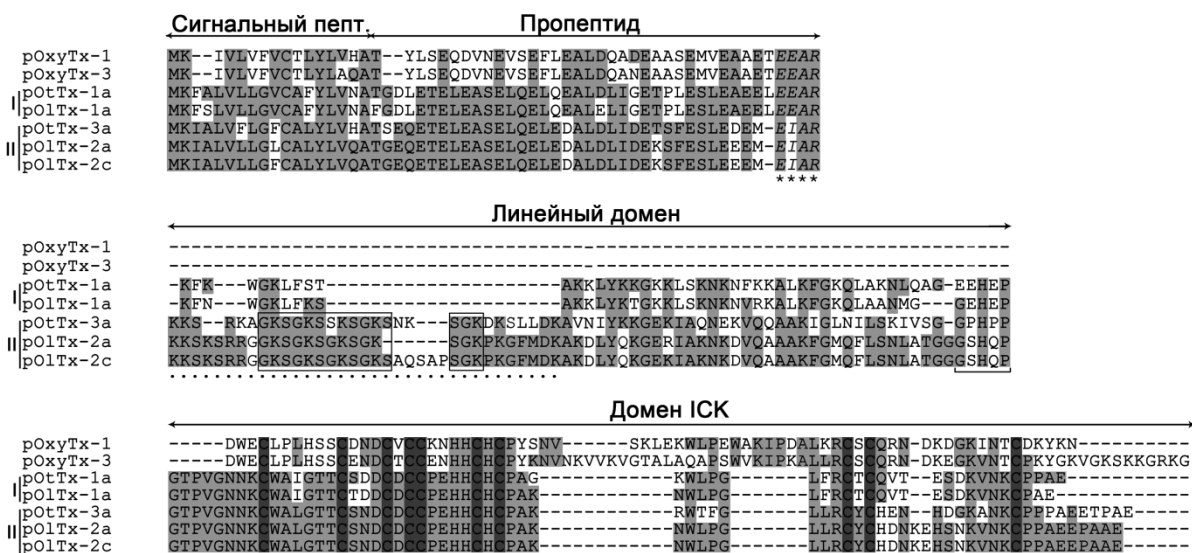


Рисунок 11 Ген, кодирующий О1Тх 2а. Сигнальный пептид, пропептид, линейный и ИСК домены обозначены стрелками. Возможный донорный сайт сплайсинга выделен темно-серым цветом.



**Рисунок 12.** Сравнение белков-предшественников однодоменных ноттиновых токсинов (*OxyTx 1* и *OxyTx 3*) и спайдеринов, принадлежащих к подсемействам *Ox-I* (*OtTx 1a* и *OtTx 1b*) и *Ox-II* (*OtTx 3a*, *OtTx 2a* и *OtTx 2c*). Идентичные остатки выделены светло-серым цветом, консервативные остатки цистеина – темно-серым. Сайты процессинга PQM отмечены курсивом и звездочками, линкерная последовательность между доменами – горизонтальной скобкой. «Дополнительная» последовательность подчеркнута точками, «базовая» последовательность – фрагмент между «дополнительной» и линкерной последовательностями. Повторы Gly-Lys-Ser помещены в прямоугольники.

### 2.3 Молекулярная эволюция двудоменных токсинов *Oxyores*

Был проведен анализ механизмов эволюции фрагментов генов, кодирующих сигнальные пептиды, пропептиды, а также на линейные и ноттиновые домены спайдеринов (всего 20 последовательностей) методом Нея-Годжобори. Были построены выравнивания нуклеотидных последовательностей, кодирующих белки-предшественники, и все позиции, содержащие разрывы, были удалены. Поскольку линейные домены значительно различаются по длине, только их «базовые» фрагменты использовались в анализе. Действие мощного отрицательного отбора было зафиксировано не только сигнальных пептидах и пропептидах, но и на линейных и зрелых доменах (табл. 7). Парное сравнение последовательностей выявило положительный отбор, действующий в парах *OtTx 2a* – *OtTx 2e* и *OtTx 2b* – *OtTx 2e* между их N-концевыми доменами. Последний вывод, однако, не проходит поправки Бонферрони на многократное тестирование, поэтому нельзя исключить, что этот результат является ложноположительным.

Дальнейшее парное сравнение проводили методом максимального правдоподобия, и он подтвердил наши наблюдения. Подавляющее большинство пар N-концевых доменов подвергалось отрицательному отбору:  $\omega \leq 0.3$  для 45,2% пар, и  $0.3 < \omega < 0.9$  для 48,4 % пар. Однако для пар *OtTx 2a* – *OtTx 2e* и *OtTx 2b* – *OtTx 2e* значение  $\omega$  достигло 15,3, и оно было больше 2 для четырех других пар, что можно интерпретировать как признак



	p	$d_S - d_N$	Количество сайтов, использованных в анализе
Сигнальные пептиды	0,039	1,783	18
Пропептиды	0,017	2,149	39
Линейные домены	0,0001	3,951	46
Ноттиновые домены	0,001	3,216	58

**Таблица 7. Результаты Z-теста на отрицательный отбор для сигнальных пептидов, пропептидов, линейных и ноттиновых доменов.**  $d_S$  и  $d_N$  – количество синонимичных и несинонимичных нуклеотидных замен на сайт, соответственно.  $p$  – вероятность того, что нулевая гипотеза о нейтральности ( $d_N = d_S$ ) будет отвергнута в пользу альтернативной гипотезы ( $d_N < d_S$ ). Значения  $p < 0,05$  считали значимыми.

положительного отбора. Аналогичный анализ, проведенный для С-концевых доменов и препро-последовательностей, показал тенденцию к отрицательному отбору без каких-либо указаний на положительный отбор. Таким образом, согласно нашим данным, гены, кодирующие двудоменные токсины *Oxуores*, консервативны, однако положительный отбор может действовать на N-концевой домен. Это было проверено и подтверждено тестами LRT (табл. 8). Оба теста показали, что последовательности, кодирующие зрелые пептиды, подвержены положительному отбору, хотя M7/M8 LRT был менее значим, чем M1a/M2a LRT ( $p < 0,1$  и  $p < 0,05$ , соответственно). Модели M2a и M8, допускающие положительный отбор, показали, что только сайт Leu-14 (нумерация по последовательности OITx 1a) подвержен положительному отбору ( $P > 95\%$ ), и этот сайт относится к N-концевому домену.

Модель	lnL	$2\Delta l$	Сайты под положительным отбором
M1a	-1627,942613	6,361392	14 L
M2a	-1624,761917		
M7	-1624,685734	4,749764	14 L
M8	-1622,310852		

**Таблица 8. Результаты тестов отношения правдоподобия (LRT) для двудоменных токсинов *Oxуores***

Структура двудоменных токсинов выглядит мозаичной, поэтому были предприняты шаги по поиску генов, гомологичных их предкам. Значимого сходства между N-концевым линейным доменом спайделинов и каким-либо другим известным белком обнаружено не было. С-концевой домен ИСК, однако, проявляет значимое сходство с окситоксинами OхуТх 1 и OхуТх 2 [32,64] (~45% идентичных а. о.). В библиотеке кДНК из ядовитых желез *O. takobius* был проведен поиск последовательностей, гомологичных спайдели-

нам. Для этого из нее извлекли последовательности, кодирующие предшественники линейных токсинов (в т. ч. OхyTх 1, номер в GenBank KF766562). Сравнение препропептидов и последовательностей N-концевого домена с помощью программы BLASTP не показало значимого сходства. Однако, последовательности препропептидов и C-концевых доменов двудоменных токсинов показало значимый уровень сходства с соответствующими последовательностями однодоменных ноттиновых токсинов (около 40% сходства) (рис.12).

Интересно, что в целом в составе белков-предшественников последовательности линейных зрелых цепей (в том числе линейные домены OтTх) более вариабельны, чем последовательности ноттиновых зрелых цепей (в т. ч. ноттиновые домены OтTх). Для оценки уровня вариабельности в группе последовательностей можно использовать сумму длин ветвей (СДВ) филогенетического дерева, построенного методом ближайшего соседа. Для каждой группы токсинов (однодоменных линейных, однодоменных 10-ти цистеиновых ноттиновых, двудоменных) из библиотеки кДНК *O. takobius* для последовательностей, кодирующей сигнальные пептиды, пропептиды и зрелые цепи, построили такие филогенетические деревья и рассчитали СДВ (табл.9).

	Сигнальные пептиды	Пропептиды	Зрелые пептиды	
			Линейные	ISK
Линейные токсины	0,67	1,95	4,36	–
ISK токсины	4,53	4,54	–	2,23
OтTх	0,19	0,64	1,25	0,49

**Таблица 9. Суммы длин ветвей филогенетических деревьев, построенных для сигнальных пептидов, пропептидов и зрелых пептидов разных групп токсинов *Oxycorpes*.**

Для фрагментов кДНК, кодирующих линейные зрелые цепи, значение СДВ выше, чем для фрагментов, кодирующих сигнальные пептиды и пропептиды. Наоборот, для фрагментов кДНК, кодирующих ноттиновые зрелые цепи, значение СДВ ниже, чем для фрагментов, кодирующих сигнальные пептиды и пропептиды. Интересно, что эти тенденции, наблюдающиеся среди однодоменных токсинов, сохраняются и при переходе к двудоменности (по крайней мере, в пределах этой библиотеки кДНК). Таким образом, во фрагменте гена, кодирующем линейную последовательность, накапливается больше мутаций, чем в тех, которые кодируют сигнальный пептид, пропептид и последовательность с ISK мотивом. Фрагмент гена, кодирующий ноттиновый домен, относительно фрагмента, кодирующего пропептид, накапливает меньше мутаций.

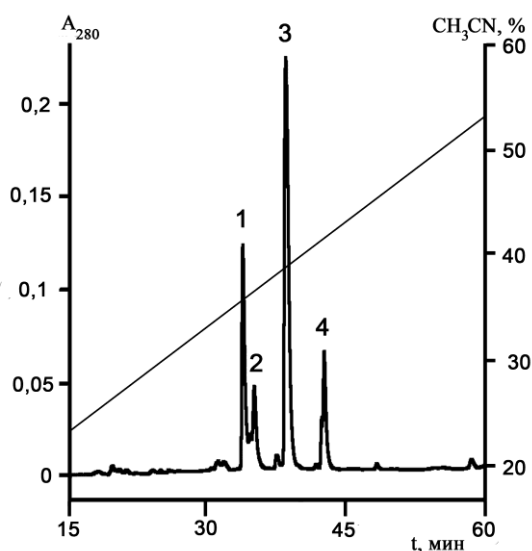
### 3. Выделение новых двудоменных токсинов из яда паука *C. punctoriium*

В яде паука *C. punctoriium* тоже были обнаружены двудоменные токсины, которые, однако, относятся к другому структурному классу – они состоят из двух ноттиновых доменов [69]. Здесь описано выделение новых представителей этого класса.

Лиофилизированный цельный яд *C. punctoriium* фракционировали с помощью двухстадийной хроматографии. Яд растворяли в воде и на первой стадии разделяли методом гель-фильтрации. Пептидная фракция, содержащая компоненты с молекулярной массой около 15 кДа, подвергалась дальнейшему разделению с помощью ОФ-ВЭЖХ. В результате хроматографический профиль содержал 4 пика (рис.13). Ранее уже было опубликовано, что пик 3 соответствует токсину СрТх 1. Все остальные фракции подвергались анализу с помощью масс-спектрометрии МАЛДИ для того, чтобы установить молекулярные массы их компонентов. В результате оказалось, что каждый из пиков соответствовал индивидуальному веществу (табл.10). Для установления частичной N-концевой аминокислотной последовательности выделенных пептидов использовали метод деградации по Эдману (табл.10). Полные аминокислотные последовательности новых пептидов были выведены путем сопоставления частичных N-концевых последовательностей с последовательностями транслированных кДНК (см. далее). Токсины из пиков 1, 2 и 4 были названы СрТх 2а, СрТх 3а и СрТх 4а, соответственно. Так же как СрТх 1, эти токсины обладали летальным и паралитическим действием на личинки мясной мухи (табл. 11). Суммарно токсины СрТх 1–4 составляют по крайней мере 1/3 цельного яда и являются, таким образом, его наиболее представленными компонентами.

№ пика	Измеренные молекулярные массы	N-концевая последовательность	Названия
1	14970	GKKCIERNKECTNDRHGCCR	СрТх 2а
2	14980	TCVPRDGDCTENRKACCRSK	СрТх 3а
3	15100	GKTCIERNKECTNDRHGCCR	СрТх 1а–с
4	15080	ASCTERKHDCTKDRHSCCRG	СрТх 4а

**Таблица 10.** Токсины, идентифицированные в яде *C. punctoriium*. Номера пиков соответствуют нумерации на хроматограмме на рис. 13.



	ЛД <sub>50</sub> , нмоль/г	ПД <sub>50</sub> , нмоль/г
СрТх 2	>3,3	2,2 – 3,3
СрТх 3	2,2 – 3,3	1,7 – 2,2
СрТх 4	2,2 – 3,3	1,7 – 2,2

Таблица 11 (вверху). Результаты теста на инсектотоксичность СрТх на личинках мясной мухи *S. carnaria*. ЛД<sub>50</sub> – 50% летальная доза, ПД<sub>50</sub> – 50% паралитическая доза.

Рисунок 13 (слева). Разделение методом ОФ-ВЭЖХ фракции из яда *C. punctoriom*, содержащей компоненты массой 15 кДа. 1 – СрТх 2а, 2 – СрТх 3а, 3 – СрТх 1а-с и 4 – СрТх 4а.

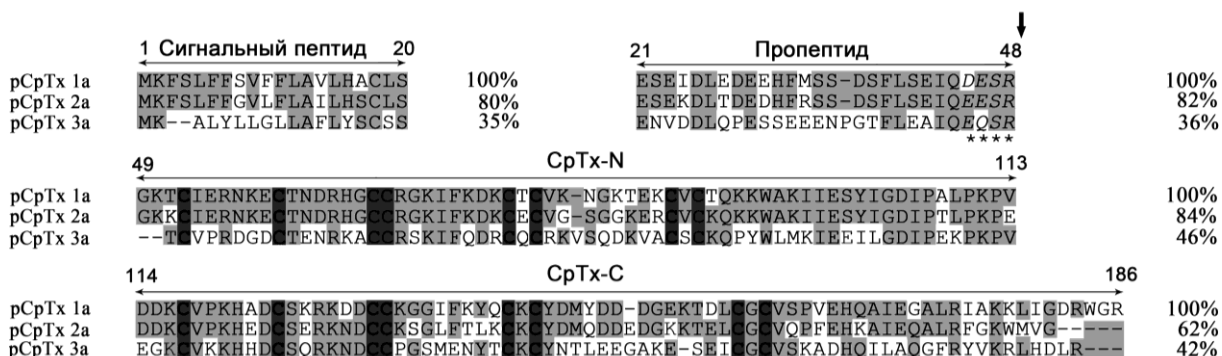
## 4. Гены двудоменных токсинов *C. punctoriom*

### 4.1 Анализ последовательностей кДНК из ядовитых желез паука *C. punctoriom*

Ранее уже сообщалось о трех последовательностях кДНК, кодирующих токсины СрТх 1а, СрТх 1б и СрТх 1с [69]. В этой работе был проведен поиск новых последовательностей двудоменных токсинов среди последовательностей из той же библиотеки кДНК, используя в качестве поискового запроса уже известные последовательности СрТх 1 и вновь установленные частичные N-концевые последовательности СрТх 2 – 4. В результате было обнаружено 12 новых последовательностей, проявляющих сходство с СрТх 1–4, обнаруженными в яде (номера в GenBank KF155268 – KF155271 и KF155273 – KF155280).

Для дальнейшего анализа полученные последовательности кДНК обрабатывались *in silico*. Они были транслированы, а затем были идентифицированы сайты процессинга. Точка отщепления сигнального пептида была установлена с помощью программы SignalP. В качестве сайтов процессинга PQM были приняты четырехаминокислотные последовательности DESR, EESR и EQSR. Последовательности белков-предшественников СрТх 1а, СрТх 2а и СрТх 3а (рСрТх 1а, рСрТх 2а и рСрТх 3а) приведены на рис.14. Исходные нуклеотидные последовательности были разбиты на фрагменты, кодирующие препропептиды и зрелые последовательности для осуществления последующего анализа. Полные аминокислотные последовательности СрТх 2а и СрТх 3а были выведены благодаря совпадению N-концевой последовательности, установленной секвенированием по Эдману, и транслированной и процессированной *in silico* кДНК. Токсин СрТх 2а (расчетная молекулярная

масса 14971 Да) амидирован на С-конце (так же как CrTx 1a – 1c). У CrTx 3a отсутствует С-концевой остаток аргинина белка-предшественника (номера в UniProt C0HJD4 и C0HJD5).

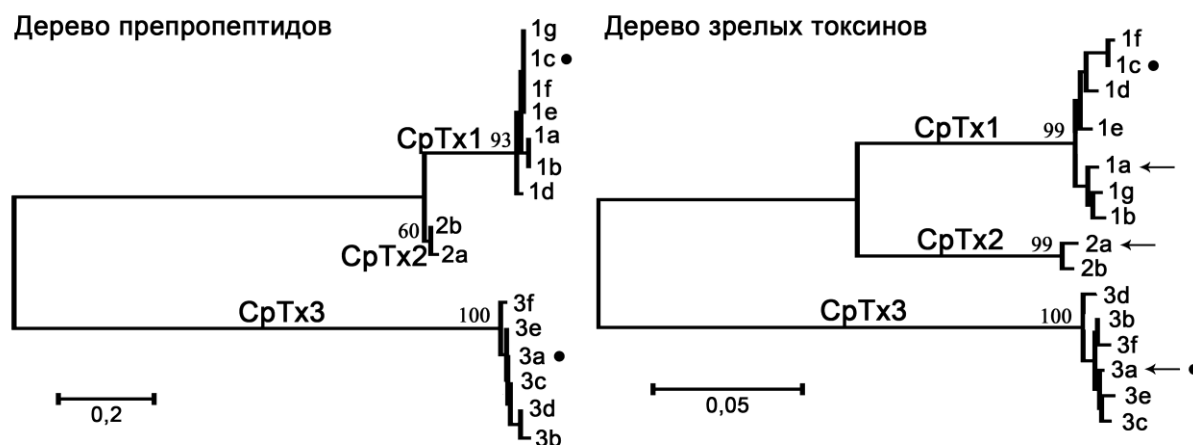


**Рисунок 14** Сравнение последовательностей белков-предшественников токсинов CrTx 1a, CrTx 2a и CrTx 3a. Идентичные остатки выделены серым фоном, остатки цистеина – жирным шрифтом, сайты PQM – курсивом и звездочками. Стрелкой отмечен сайт процессинга. Доли идентичных остатков относительно pCrTx 1a показаны отдельно для сигнальных пептидов, пропептидов и отдельных доменов.

Для изучения разнообразия двудоменных токсинов как на уровне белка, так и кДНК, было предпринято следующее исследование. Методом ближайшего соседа были построены филогенетические деревья для препро-последовательностей, зрелых цепей и фрагментов генов, кодирующих их. Поскольку деревья, построенные на основе последовательностей кДНК и белковых предшественников, совпали, здесь приведены только «кДНК-деревья» (рис.15). Дерево «зрелых токсинов» состоит из трех ветвей, каждая из которых несет по листу, соответствующему одному из идентифицированных в яде токсинов (CrTx 1a – 1c, CrTx 2a, и CrTx 3a). Разница между последовательностями на каждой из ветвей заключается в нескольких точечных нуклеотидных заменах. Топологии «препропептидного» дерева и дерева «зрелых пептидов» в значительной мере совпадают. Таким образом, на уровне последовательностей кДНК были обнаружены три группы двудоменных токсинов *C. punctatorum* CrTx 1, CrTx 2 и CrTx 3. Новые пептиды получили названия CrTx 1d–1g, CrTx 2a, 2b и CrTx 3a–3f.

В качестве примера, иллюстрирующего различия между группами CrTx 1, CrTx 2 и CrTx 3 можно рассмотреть выравнивание белковых предшественников CrTx 1a, CrTx 2a и CrTx 3a (рис.14). Предшественники CrTx 1a и CrTx 2a более сходны друг с другом (76% идентичных а. о.), чем с CrTx 3a (около 42% идентичных а. о.). Препропептиды CrTx 1a и CrTx 2a проявляют высокое сходство (85% идентичных а. о.), в то время как зрелые пептиды более вариабельны (72% идентичных а. о.). Интересно, что вклад по-

следовательностей С-концевых доменов в упомянутую вариабельность более значителен, поскольку между ними доля идентичных а. о. составляет 62%, в то время как между последовательностями N-концевых доменов она достигает 84%. В случае пары CrTx 1 – CrTx 3 зрелые последовательности несколько более сходны (44% идентичных а. о.), чем последовательности препептидов (35%).



**Рисунок 15.** Филогенетические деревья, построенные методом ближайшего соседа на основе сравнения фрагментов кДНК, кодирующих препро-последовательности и зрелые токсины CrTx. Масштабные линейки соответствуют количеству нуклеотидных замен на сайт вдоль каждой ветви. Уровни поддержки для групп CrTx 1, CrTx 2 и CrTx 3 указаны над соответствующими ветвями. Стрелками указаны последовательности, кодирующие токсины, выделенные из яда *C. punctoriun*, жирными точками – для которых также была установлена последовательность соответствующих фрагментов геномной ДНК.

#### 4.2 Структура генов CrTx-подобных токсинов

Фрагменты геномной ДНК паука *C. punctoriun*, кодирующие двудоменные токсины, амплифицировали с помощью специфических праймеров и секвенировали (номера в GenBank KF255415 – KF255417). Сравнение последовательностей кДНК фрагментов геномной ДНК показало, что гены двудоменных токсинов *C. punctoriun* не содержат интронов. Были идентифицированы гены, кодирующие белки-предшественники CrTx 1с, CrTx 3а и CrTx 3г. Аналогично CrTx 1с, CrTx 3а был так же найден в составе яда *C. punctoriun*.

#### 4.3 Молекулярная эволюция двудоменных токсинов *C. punctoriun*

Представляется интересным выяснить, подвергаются ли двудоменные токсины *C. punctoriun* ускоренной эволюции. Для этого с помощью Z-теста был проведен анализ типов отбора на различных участках генов, кодирующих сигнальные пептиды, препепти-

ды и зрелые последовательности белков-предшественников (табл.12). Внутри каждой из групп фрагменты генов, кодирующие сигнальные пептиды и пропептиды, идентичны либо содержат по одной нуклеотидной замене. Фрагменты генов, кодирующие зрелые последовательности, более изменчивы, однако в противоположность ко многим известным токсинам, положительный отбор им не свойственен. В группах СрТх 1 и СрТх 2 наблюдается значимый отрицательный отбор ( $p < 0,05$ ), действующий на зрелых последовательностях. Внутри группы СрТх 3 тенденция к отрицательному отбору сохраняется, хотя уровень значимости снижается ( $p = 0,06$ ) в связи с тем, что в некоторых парах последовательностей (СРТХ 3а – СРТХ 3f, СРТХ 3с – СРТХ 3f) наблюдается положительный отбор ( $p < 0,05$ , однако  $p$  не проходит поправку Бонферрони на многократное тестирование). Сравнение последовательностей, принадлежащих к разным группам показывает, что отрицательный отбор действует как на препропептиды, так и на зрелые цепи. Дальнейший филогенетический анализ методом максимального правдоподобия также подтвердил этот вывод. Для зрелых последовательностей результаты обоих тестов LRT не были значимыми ( $2\Delta l = 0$ ), поэтому нельзя предположить, что на них действует положительный отбор. Таким образом, ускоренная эволюция не свойственна двудоменным токсинам, и отрицательный отбор действует на всех отрезках предшественников токсинов.

Группа токсинов	Сигнальный пептид			Пропептид			Зрелый пептид		
	$p$	$d_S - d_N$	Количество сайтов***	$p$	$d_S - d_N$	Количество сайтов***	$p$	$d_S - d_N$	Количество сайтов***
СрТх 1	Идентичные*			Идентичные*			0,003	2,86	135
СрТх 2	Идентичные*			Идентичные*			0,039	1,78	132
СрТх 3**	Идентичные*			Идентичные*			0,073	1,46	132
СрТх 1а-СрТх 2а	0,161	0,99	18	0,030	1,89	27	0,000	5,96	128
СрТх 1а-СрТх 3а	0,008	2,47		0,001	2,39		0,000	3,44	
СрТх 2а-СрТх 3а	0,012	2,28		0,009	2,36		0,002	3,03	

**Таблица 12. Результаты Z-теста на отрицательный отбор для фрагментов генов, кодирующих сигнальные пептиды, пропептиды и зрелые цепи белков-предшественников СрТх.**  $d_S$  и  $d_N$  – количество синонимичных и несинонимичных нуклеотидных замен на сайт, соответственно.  $p$  – вероятность того, что нулевая гипотеза о нейтральности ( $d_N = d_S$ ) будет отвергнута в пользу альтернативной гипотезы ( $d_N < d_S$ ). Значения  $p < 0,05$  считали значимыми. \*Идентичные либо содержащие одну нуклеотидную замену последовательности. \*\*В некоторых парах был зафиксирован положительный отбор ( $p < 0,05$ ,  $d_N - d_S \sim 2$ ). \*\*\* Количество сайтов, использованных в анализе.

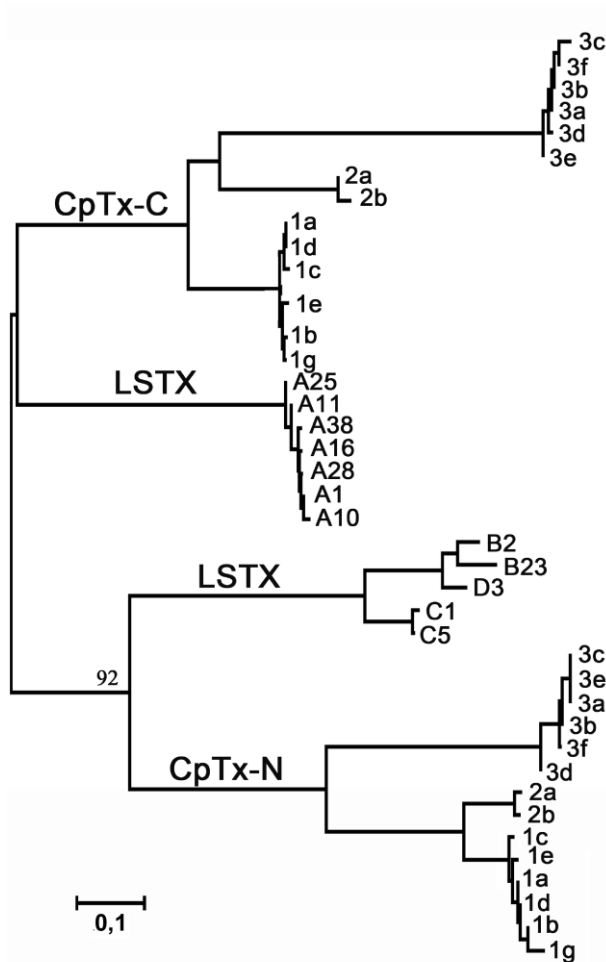
Наиболее очевидной является идея о появлении двудоменных токсинов в результате дубликации предкового гена, кодировавшего ноттиновый пептид. Однако, в случае

СрТх-подобных пептидов это, скорее всего, не так, поскольку их N-концевые (СрТх-N) и С-концевые (СрТх-С) домены проявляют только умеренное сходство (16 – 23%). Таким образом, было сделано предположение о том, что появление СрТх-подобных токсинов связано с эволюционным слиянием отдаленно родственных ноттиновых пептидов. Для подтверждения гипотезы, был предпринят поиск гомологичных последовательностей в библиотеке кДНК из ядовитых желез *C. punctorium* и в базах данных мРНК, находящихся в открытом доступе.

Библиотека кДНК *C.punctorium* содержит последовательности, кодирующие одно-доменные токсины, содержащие тот же характерный мотив из 8 остатков цистеина, как и отдельные домены СрТх-подобных пептидов. Однако эти токсины в значительной мере отличаются как от СрТх-N, так и от СрТх-С (менее 25% идентичных а. о.). Таким образом, было бы неверно рассматривать однодоменные токсины *C.punctorium* в качестве возможных предков СрТх-подобных токсинов.

С помощью программы BLAST в открытых банках кДНК был проведен поиск последовательностей, гомологичных СрТх-N и СрТх-С. В результате были обнаружены несколько последовательностей кДНК из ядовитых желез паука *Lycosa singoriensis* (LSTX). Полученные последовательности были транслированы *in silico*. Белки-предшественники LSTX были разбиты на последовательности препропептидов и зрелых цепей. Были построены выравнивания для фрагментов генов, кодирующие зрелые цепи LSTX, СрТх-N и СрТх-С, а затем методом ближайшего соседа было построено филогенетическое дерево (рис.16). Последовательности СрТх-N образуют общую ветвь с последовательностями LSTXB, LSTXC и LSTXD, в то время как СрТх-С ближе к последовательностям LSTXA. Наивысшая доля идентичных остатков (40%) был зафиксирован в парах последовательностей СрТх 2-N – LSTXC1 и СрТх 2-С – LSTXA1. Препро-последовательность предшественника СрТх 2a идентична соответствующей последовательности LSTXC1 на 31% (в то время как в случае с LSTXA1 это значение составило бы только 21%) (рис.17). Здесь можно предположить, что геном *C. punctorium* также содержит гены гомологичные описанным выше генам *L. singoriensis*, которые по каким-либо причинам не экспрессируются в транскриптом *C. punctorium*. В эволюции в геноме *C. punctorium* могла иметь место «миграция» последовательности, кодирующей зрелый пептид гипотетического гомолога LSTXA1, в ген, кодирующий гипотетически гомолог LSTXC1.





**Рисунок 16.** Филогенетическое дерево, построенное методом ближайшего соседа на основе выравнивания фрагментов кДНК, кодирующих отдельные домены токсинов CpTx (CpTx-N и CpTx-C) и зрелых пептидов LSTX. Использовали только небольшое количество последовательностей, полученных с помощью поиска BLAST в открытых банках нуклеотидных последовательностей, т. к. они в значительной мере сходны и не поменяли бы общей топологии дерева. Уровень поддержки для кластеров LSTX-CpTx показан над соответствующей ветвью.

CpTx 2a-prepro	MKFSLF <sup>g</sup> GVLF <sup>l</sup> LA <sup>l</sup> LH <sup>s</sup> CL <sup>s</sup> SE <sup>s</sup> EKDLT <sup>d</sup> ED <sup>d</sup> HFRSS <sup>d</sup> DF <sup>l</sup> SEI-QEES <sup>r</sup>	100%
LSTXC1-prepro	MKVLV <sup>l</sup> ESV <sup>l</sup> FL <sup>l</sup> TL <sup>l</sup> FSYS <sup>s</sup> TEA----ID <sup>e</sup> FDSDA <sup>e</sup> ED <sup>d</sup> ML <sup>s</sup> LMAN <sup>e</sup> QVR	31%
		****
CpTx 2a-N	GKK <sup>l</sup> IERNKE <sup>c</sup> TNDRHG <sup>c</sup> CRGKIFKDK <sup>c</sup> EE <sup>c</sup> VG-SGG--KER <sup>v</sup> V <sup>c</sup> KOK <sup>k</sup> WAKI <sup>l</sup> ESYIGDIPTLPKPE	100%
LSTXC1-mat	AKA <sup>l</sup> TPRLHD <sup>c</sup> SHDRHS <sup>c</sup> CRGELFKD <sup>v</sup> Y <sup>c</sup> FYPEGEDKTEV <sup>s</sup> SG <sup>l</sup> QOPKSHKY <sup>l</sup> EKVVDKAKTIVG--	40%
CpTx 2a-C	DDK <sup>l</sup> VPKHED <sup>s</sup> SERKND <sup>c</sup> CKSGLF <sup>l</sup> TLK <sup>k</sup> Q <sup>l</sup> YDMQDDE <sup>d</sup> DGKKTEL <sup>l</sup> GC <sup>v</sup> QPFENKAI <sup>e</sup> EQALRF <sup>g</sup> GKWMV <sup>g</sup>	100%
LSTXA1-mat	-KE <sup>l</sup> IPKHHE <sup>s</sup> TSNKHG <sup>c</sup> CRGNF <sup>c</sup> EKYK <sup>l</sup> Q <sup>l</sup> TTVV <sup>t</sup> T-QDGEQ <sup>t</sup> TER <sup>l</sup> FG <sup>l</sup> GTPPHKAA <sup>e</sup> ELV <sup>g</sup> VG <sup>g</sup> FG <sup>g</sup> KKIF <sup>g</sup>	40%

**Рисунок 17.** Сравнение фрагментов белка-предшественника pCpTx 2a с соответствующими фрагментами LSTXC1 и LSTXA1. Препропептиды обозначены как «-prepro», зрелые цепи – «-mat». Идентичные остатки показаны на сером фоне, остатки цистеина – жирным шрифтом, сайты PQM – курсивом и звездочками.

## 5. Двудоменные токсины паука *L. tarabaevi*

Цитоинсектотоксины (ЦИТ) *L. tarabaevi* составлены из двух модулей, каждый из которых склонен к формированию амфипатической  $\alpha$ -спирали и соответствует «обычному» цитотоксину из яда пауков [31].

В этой работе с помощью химического синтеза были получены производные ЦИТ 1a, соответствующие N-концевому (ЦИТ 1a-N, а. о. 1 – 34) и C-концевому (ЦИТ 1a-C, а. о. 39 – 69) модулям. Такое разбиение обусловлено тем, что в последовательность пептидов не была включена линкерная последовательность <sup>35</sup>E<sup>35</sup>E<sup>36</sup>A<sup>37</sup>Q<sup>38</sup>, которая, скорее всего, об-

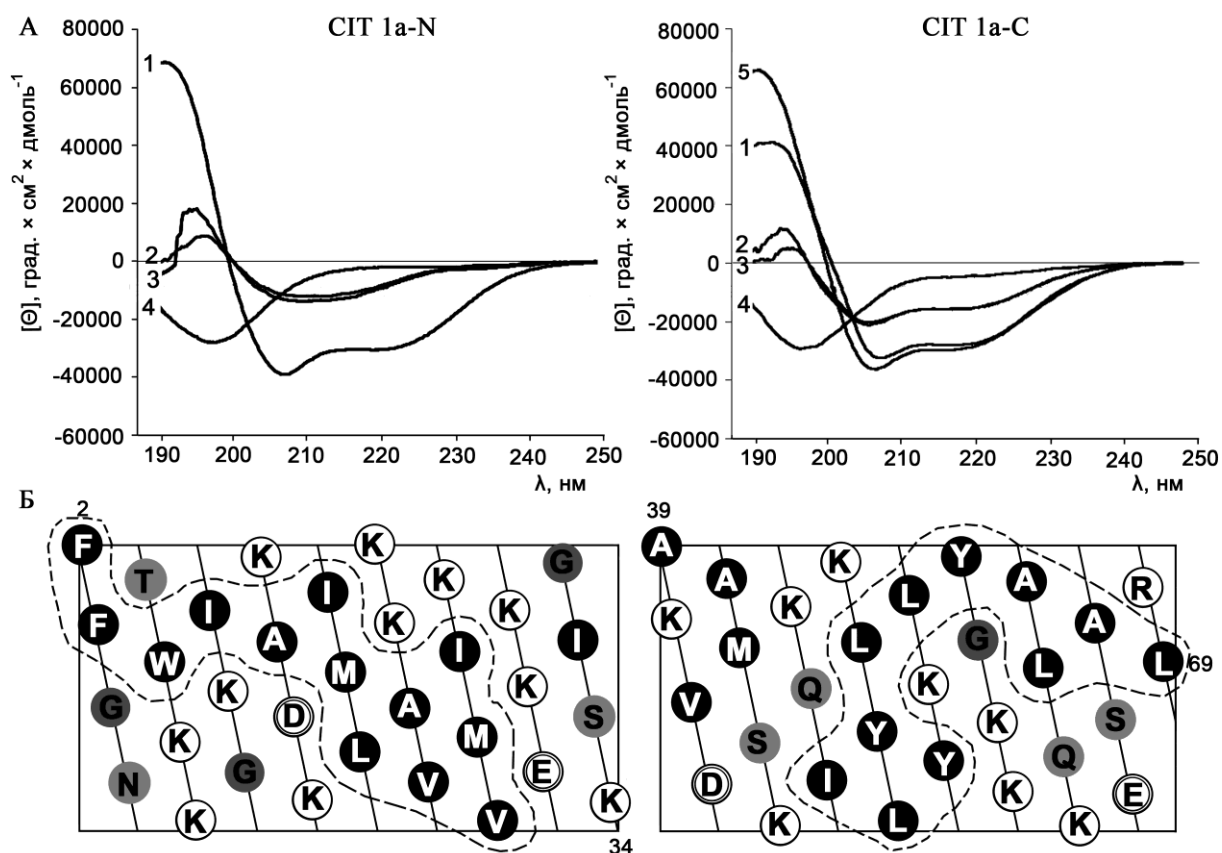
разовалась в результате точечной мутации в сайте процессинга PQM. С помощью методов биоинформатики было показано, что как СІТ 1а, так и СІТ 1а-N и СІТ 1а-С склонны к формированию  $\alpha$ -спиральной пространственной структуры. На рис. 18 приведены проекции спиральной сети для каждого из фрагментов. Были получены спектры кругового дихроизма для СІТ 1а-N и СІТ 1а-С в воде, а также в растворителях, имитирующих мембранное окружение (20 мМ ДСН и 50% ТФЭ), и суспензиях цвиттерионных (ДОФХ) и анионных липосом (ДОФГ/ДОФЭ) (табл. 13). Оказалось, что в воде преобладает доля неупорядоченной вторичной структуры, а в окружении, имитирующем мембранное доминирует  $\alpha$ -спиральная конформация.

Пептид	Раствор	Вторичная структура, доля (%)			
		$\alpha$ -спираль	$\beta$ -слой	$\beta$ -поворот	неупорядоченная
СІТ 1а (по [31])	Вода	13	14	24	<b>49</b>
	ДОФХ	<b>55</b>	5	14	26
	ДОФГ/ДОФЭ	-	-	-	-
	20 мМ ДСН	<b>66</b>	3	8	23
	50% ТФЭ	<b>69</b>	3	7	21
СІТ 1а-N	Вода	7	14	26	<b>54</b>
	ДОФХ	<b>41</b>	17	21	22
	ДОФГ/ДОФЭ	20	28	24	29
	20 мМ ДСН	-	-	-	-
	50% ТФЭ	<b>78</b>	2	2	18
СІТ 1а-С	Вода	8	11	26	<b>56</b>
	ДОФХ	<b>42</b>	4	22	32
	ДОФГ/ДОФЭ	<b>45</b>	7	20	28
	20 мМ ДСН	<b>77</b>	3	0.3	20
	50% ТФЭ	<b>74</b>	1	6	19

*Таблица 13. Доли типов вторичной структуры СІТ 1а и его производных СІТ 1а-N и СІТ 1а-С, полученные методом кругового дихроизма. Жирным шрифтом выделены максимальные значения для каждого из растворителей.*

Антимикробную активность СІТ 1а-N и СІТ 1а-С тестировали в сравнении с полноразмерным токсином СІТ 1а на грамположительных и грамотрицательных бактериях. Оказалось, что МИК СІТ 1а-N значительно ниже, чем СІТ 1а-С, а МИК полноразмерного пептида в несколько раз ниже, чем МИКи каждого из пептидов и их смеси (табл. 14). Таким образом, для проявления высокой антимикробной активности СІТ 1а необходимы оба модуля, причем в их действии наблюдается синергизм. Поскольку СІТ 1а-N проявляет более высокую активность, можно предположить, что именно N-концевой домен является ответственным за цитолитическую активность СІТ 1а, а С-концевой домен играет роль «усилителя».

В тестах на инсектицидную активность на личинках мясной мухи СІТ 1а-Н и СІТ 1а-С оказались неактивны вплоть до доз ~55 нмоль/г,



**Рисунок 18. Вторичная структура пептидов СІТ 1а-Н и СІТ 1а-С.** (А) Спектры кругового дихроизма пептидов СІТ 1а-Н и СІТ 1а-С в разных растворителях: 1 – 50% ТФЭ, 2 – суспензия ДОФГ/ДОФЭ липосом, 3 – суспензия ДОФХ липосом, 4 – вода, 5 – 20 мМ ДСН. (Б) Проекция спиральной сети СІТ 1а-Н и СІТ 1а-С. Положительно заряженные остатки обозначены белыми кругами, гидрофобные остатки – черными, остатки глицина – темно-серыми, гидрофильные незаряженные остатки – светло-серыми, отрицательно заряженные остатки – белыми кругами с двойной линией по краю.

Бактериальный штамм	Полипептиды			
	СІТ 1а, по (Vassilevski <i>et al.</i> 2008)	СІТ 1а-Н	СІТ 1а-С	Эквимоллярная смесь
	МИК, мкМ			
<i>E. coli</i> DH5α	0,6	5 – 10	40 – 80	2,5 – 5
<i>B. subtilis</i>	0,9	1,3 – 2,5	20 – 40	1,3 – 2,5

**Таблица 14. Минимальные ингибирующие концентрации СІТ 1а и его производных СІТ 1а-Н и СІТ 1а-С**

## Обсуждение

### 1. Получение рекомбинантных полипептидов

Большинство рекомбинантных пептидов получают в бактериальных системах, в основном, в культуре *E. coli* [206]. Понятно, что для пептидов, обладающих антимикробными свойствами, этот процесс сопровождается многими трудностями. Линейные катионные пептиды с мембранолитической активностью являются цитотоксичными и, кроме того, они легко подвергаются расщеплению внутриклеточными протеазами. В случае пептидов, содержащих дисульфидные связи, их формирование в цитоплазме бактериальных клеток затруднено, так как в ней поддерживается восстановительный потенциал [207].

Токсин OtTx 1a является «химерой» линейного и дисульфид-богатого пептидов, поэтому получение его рекомбинантного аналога в системе *E.coli* сопровождалось трудностями, свойственными для каждого из типов пептидов. Во-первых, как было показано в этой работе, цитолитический домен в низких концентрациях разрушает бактериальные мембраны. Во-вторых, в первичной структуре этого домена высокое содержание остатков лизина (~36%), поэтому он легко подвергается протеолитическому расщеплению. В-третьих, ноттиновый домен содержит 5 дисульфидных связей, что может значительно затруднить формирование нативной пространственной структуры при получении его в гетерологической системе. Для преодоления этих препятствий использовали комплекс подходов.

При получении рекомбинантных антимикробных пептидов наиболее часто (около 22% случаев) в качестве белка-помощника используется тиоредоксин [206,208], причем как среди линейных мембраноактивных, так и дисульфид-богатых пептидов. Известно, что тиоредоксин повышает растворимость рекомбинантных белков, а также способствует формированию дисульфидных связей в штаммах, мутантных по генам тиоредоксинредуктазы и глутатион редуктазы (например, Origami B) [209,210]. Кроме того, он может использоваться для маскировки мембранолитической активности антимикробных пептидов [85]. Так, ранее в нашей лаборатории уже был достигнут положительный результат при получении линейного цитолитического пептида Ltc2a с использованием тиоредоксина в качестве белка-помощника [211]. Таким образом, тиоредоксин оказался наиболее подходящим белком-помощником при получении токсина OtTx 1a, сочетающего в своей струк-

туре линейный и дисульфид-богатый домены. Согласно полученным результатам, OtTx 1a синтезировался в нативной форме.

Другой интересный подход, использованный в данной работе, это введение природного пропептида OtTx 1a в состав химерного белка Ttx-OtTx 1a. Действие пропептида, с одной стороны, заключается в нейтрализации мембранолитической активности линейного домена, а с другой, в защите самого линейного домена от внутриклеточных протеаз. Эффективность использования пропоследовательностей для защиты рекомбинантных антимикробных пептидов от бактериальных протеаз была показана неоднократно [85].

В случае Ltc 2a его выход при получении в составе гибридного белка с пропептидом не отличался от случая, когда его получали без пропептида. Однако димер Ltc 2a в составе химерного белка без пропептида получить не удалось. Таким образом «нейтрализующего» действия тиоредоксина «было недостаточно» в случае более длинного пептида. Поскольку линейный домен OtTx 1a имеет длину (41 а. о.) и положительный заряд (+16) при pH 7, сравнимые с длиной и зарядом димера Ltc 2a (59 а. о. и +18, соответственно), в состав химерного белка Ttx-OtTx 1a был также введен природный пропептид OtTx 1a. Однако даже использование этих подходов позволило достичь выхода лишь 500 мкг/л, что в 10 раз ниже, чем в случае с OtTx 1a-ICK, не имеющего цитолитического домена.

При получении полноразмерного токсина OtTx 1a происходило правильное замыкание дисульфидных мостиков, и пептид экстрагировался из клеток в нативной конформации притом, что его выход был низким. Что касается пептида OtTx 1a-ICK, то при высоком выходе не происходило правильного замыкания дисульфидных мостиков, и наблюдалось несколько конформаций. Вероятно, образование большого количества разных конформаций OtTx 1a-ICK происходило в связи с накоплением Ttx-OtTx 1a-ICK в цитоплазме клеток до слишком высоких концентраций. Высокие концентрации дисульфидсодержащих пептидов зачастую приводят к их неправильному сворачиванию из-за накопления межмолекулярных интермедиатов [212]. Кроме того, образованию нативной пространственной укладки могут способствовать пропептид и линейный домен OtTx 1a, которые отсутствуют в составе химерного белка Ttx-OtTx 1a-ICK. Тиоредоксин, несомненно, способствовал успешному рефолдингу, так как только в его присутствии процедура рефолдинга приводила к образованию одного продукта.

## 2. **OtTx-подобные токсины *Oxyopes* – новый класс двудоменных токсинов**

### 2.1 **Двудоменная структура токсинов OtTx**

Впервые описанные в этой работе токсины OtTx обладают уникальной модульной архитектурой. N-концевая часть OtTx (OtTx-AMP) соответствует распространенным однодоменным цитотоксинам, лишенным остатков цистеина. Хотя гомология между OtTx-AMP и другими известными белковыми последовательностями не была обнаружена, линейный модуль обладает всеми характерными признаками обычных антимикробных пептидов. Так, OtTx 1a-AMP несет высокий положительный заряд (+16 при pH 7, 39% остатков лизина) и склонен формировать амфипатической  $\alpha$ -спиральной структуры, что было предсказано расчетными методами, а затем подтверждено экспериментально. В окружении, имитирующем мембранное, он приобретает структуру амфифильной  $\alpha$ -спирали с четким разделением гидрофильного и гидрофобного кластеров (рис. 9). Кроме того, как было показано с помощью соосаждения с липосомами, линейные домены проявляют способность связываться с мембранами. Что более важно, OtTx 1a-AMP обладает сильной цитолитической активностью, связанной с его способностью разрушать мембраны (см. ниже).

C-концевая часть OtTx (OtTx-ICK) содержит типичный мотив ИСК, наиболее распространенный среди нейротоксинов пауков. Более того, значимый уровень сходства наблюдается между OtTx-ICK и однодоменными ноттиновыми токсинами из яда того же паука или его близкородственных видов. Цистеиновый мотив (с двумя дополнительными дисульфидами в ИСК мотиве) в точности сохраняется как в OtTx, так и в окситоксинах.

Интересно, что между OtTx и модульными токсинами скорпионов, т. н. скорпинами, имеется сходство [213]. Эти группы токсинов имеют общую архитектуру – N-концевой антимикробный модуль, за которым следует C-концевой нейротоксиновый модуль. В этом случае C-концевой модуль соответствует типичному однодоменному токсину скорпиона с цистеин стабилизированным  $\alpha/\beta$ -мотивом (CS $\alpha\beta$ ). Таким образом, выясняется, что животные с негомологичными ядовитыми железами используют одни и те же эволюционные стратегии – вводят в состав яда модульные токсины с поразительно сходной организацией. Для того чтобы подчеркнуть упомянутое сходство, было предложено название «спайдерины» для OtTx-подобных токсинов (от англ. «spider» – паук).

## 2.2 N-концевой модуль OtTx 1a – мощный цитолитический токсин

Отмеченные выше структурные особенности OtTx 1a-AMP позволили сделать предположение, что он является цитолитическим мембраноактивным пептидом, что и было доказано экспериментально. Он эффективно убивает бактериальные и эукариотические клетки в микромолярных концентрациях, и его неизбирательная активность сравнима с классическим токсином мелиттином из яда пчелы. Были предприняты шаги по выяснению молекулярный механизм цитолитической активности OtTx 1a-AMP. Методом соосаждения с липосомами было напрямую показано, что пептид взаимодействует с мембранами. Далее было показано, что его действие является бактерицидным, а не бактериостатическим, что говорит в пользу способности пептида разрушать мембраны. Последнее было в итоге доказано с помощью конфокальной лазерной сканирующей микроскопии: после добавления OtTx 1a-AMP ИП проникал внутрь бактериальных клеток, в то время как БКЭКФ вытекал наружу. Таким образом, установлено, что именно линейный домен является ответственным за проявление токсичности полноразмерного токсина OtTx 1a. OtTx 1a-AMP наделяет его способностью к мембранной активности, проявляющейся в цитолитическом действии. Кажется, что N-концевой домен является главной функциональной частью OtTx 1a.

В работах по изучению токсинов LtTx и CSTX [71–74,214] показано, что их C-концевые AMP-фрагменты являются мембраноактивными. Было сделано предположение, что у токсинов со структурой ИСК+AMP модуль AMP играет роль мембранного якоря, который может способствовать взаимодействию с рецептором через связывание с мембраной (мембранный катализ) и стабилизировать комплекс токсин-рецептор, предотвращая диссоциацию. Кроме того, модуль AMP придает токсинам способность повреждать мембраны. Поскольку домены того же типа были обнаружены у токсинов класса OtTx, хотя и в другом порядке, возможно, они играют похожую функциональную роль. Однако, относительно OtTx 1a-AMP, токсины со структурой ИСК+AMP обладают гораздо более слабой цитотоксичностью. Более того, как обсуждалось выше, OtTx 1a-AMP – один из сильнейших цитотоксинов своего типа. В составе полноразмерного токсина он несет главную функциональную нагрузку, в то время как у LtTx и CSTX оба модуля равноправны.

## 3. Синергизм в основе активности цитоинсектотоксинов *L. tarabaevi*

Цитоинсектотоксины обладают антимикробной и, что необычно для линейных цитолитиков, высокой инсектицидной активностью [31]. Для обеспечения этой активности

необходимы оба домена, причем они действуют синергично. Отдельные домены вообще не показали инсектицидной активности. В тестах на антибактериальную активность МИК отдельных доменов СІТ 1а значительно выше, чем полноразмерного токсина, причем у СІТ 1а-N она на порядок ниже, чем у СІТ 1а-С. Активность эквимольной смеси остается на уровне активности N-концевого домена. Таким образом, механизм действия СІТ 1а основывается на взаимодействии двух доменов, в результате чего активность пептида выше, чем сумма активностей его фрагментов, причем необходимо, чтобы оба фрагмента были в составе одной молекулы. Синергизм в действии доменов у двудоменных токсинов был также показан у токсина CSTX-1 (ICK+AMP): при удалении линейного домена инсекто-токсичность падает вплоть до 190 раз [72].

#### 4. Разнообразие двудоменных токсинов

К настоящему времени описано 4 класса модульных (или двудоменных) токсинов (табл. 15), выделенных из ядов разных таксономических групп пауков, зачастую находящихся лишь в отдаленном родстве (рис. 19). Интересно, что в яде скорпионов были найдены модульные токсины скорпины, построенные аналогично спайдеринам: они содержат N-концевой антимикробный модуль, за которым следует С-концевой дисульфид-богатый домен. Из яда пауков надсемейства Lycosoidea были выделены токсины с комбинацией доменов «ICK+AMP»: LtTx из *L. tarabaevi* (Zodariidae) [74] и CSTX из *C. salei* (Ctenidae) [73]. В настоящей работе описан новый класс модульных токсинов пауков OtTx с «обратной» комбинацией модулей «AMP+ICK». Кроме того, «гетеро» комбинации дополняются двумя возможными «гомо» вариантами. Цитоинсектотоксины являются комбинацией двух линейных доменов (AMP+AMP) [31]. Токсины типа ICK+ICK были обнаружены в составе ядов двух пауков, состоящих в отдаленном родстве – мигаломорфного *C. guangxiensis* и аранеоморфного *C. punctorium* [67,69]. В этом исследовании было затронуто 3 из четырех известных классов.

N-Концевой модуль	С-Концевой модуль	
	AMP	ICK
AMP	AMP+AMP, цитоинсектотоксины из яда <i>L. tarabaevi</i>	AMP+ICK, OtTx из яда <i>O. takobius</i> ,
ICK	ICK+AMP, LtTx из яда <i>L. tarabaevi</i> , CsTx-1 из яда <i>C. salei</i>	ICK+ICK, CpTx из яда <i>C. punctorium</i> , DkTx из яда <i>H. schmidtii</i>

Таблица 15. Разнообразие модульных токсинов в ядах пауков



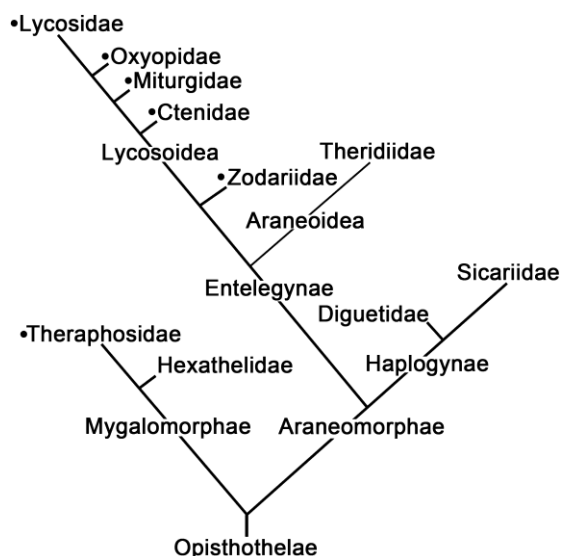


Рисунок 19. Упрощенное филогенетическое дерево семейств пауков, упомянутых в «Обсуждении» (по [238], с изменениями). Семейства, производящие двудоменные токсины, отмечены жирными точками.

#### 4.1 Двудоменные токсины *Oxyopes*

В общей сложности, в данной работе было получено 20 аминокислотных последовательностей двудоменных токсинов пауков рода *Oxyopes* (13 у *O. takobius* и 7 у *O. lineatus*). Два токсина (OtTx 1a и OtTx 2a) были выделены напрямую из яда *O. takobius*, остальные же последовательности токсинов этого паука были установлены в результате анализа библиотеки кДНК из его ядовитых желез. У паука *O. lineatus* 6 аминокислотных последовательностей были установлены на уровне транслированных последовательностей кДНК из ядовитых желез, из которых 3 были также обнаружены при трансляции фрагментов последовательностей геномной ДНК. Двудоменные токсины обоих пауков образуют две группы (Ox-I и Ox-II) в зависимости от последовательности их N-концевого домена.

Линейные домены вносят больший вклад в разнообразие двудоменных токсинов *Oxyopes* (рис. 12). Подобная ситуация наблюдается и среди латартоксина *L. tarabaevi*, структура которых «инвертирована» относительно спайдеринов (ISK+AMP вместо AMP+ISK). Между последовательностями LtTx 1a и LtTx 1b, например, доля идентичных остатков – 64% для ноттиновых и 33% для линейных доменов, а для пары LtTx 2a – LtTx 2с эти значения составляют 73% и 46%, соответственно. Вероятно, более высокий уровень вариабельности линейных доменов обусловлен их мембрано-активной функцией. Для ее обеспечения достаточно определенного распределения заряженных и гидрофобных остатков, позволяющего формировать амфипатическую  $\alpha$ -спираль и не требующего закрепления специфичных а. о. в конкретных положениях. И наоборот, трехмерная структура ноттиновых доменов поддерживается цистеиновым мотивом, а взаимодействие со специфичной белковой мишенью требует более высокой консервативности последовательности.

## 4.2 Двудоменные токсины *C. punctoriium*

Двудоменные токсины *C. punctoriium* довольно разнообразны. Было обнаружено 3 группы СрТх-подобных токсинов на уровне кДНК и белка (СрТх 1, СрТх 2 и СрТх 3), а также на уровне фрагмента N-концевой аминокислотной последовательности зрелого токсина была обнаружена четвертая группа. Каждая из групп представляется результатом недавней дупликации генов, поскольку последовательности в составе каждой из групп различаются небольшим количеством точечных мутаций. Различия между группами более значительны. Интересно, что зрелые последовательности групп СрТх 1 и СрТх 2 различаются в их С-концевых доменах больше, чем в N-концевых. Ранее уже отмечалось [69], что СрТх-N более похожи на нейротоксичный энхансер CsТх-13 из *C. salei*, в то время как СрТх-С проявляют больше сходства с CsТх-1 и CsТх-9, нейротоксинами из того же яда. Таким образом, СрТх-подобные токсины можно рассматривать как химеры, обладающие «самоусиливающей» активностью. В пользу важности N-концевого домена говорит его меньшая вариабельность. Однако, для подтверждения этой гипотезы необходимо доскональное изучение как механизмов активности CsТх-13, так и отдельных доменов СрТх.

## 5. Безинтронные гены двудоменных токсинов

Каждый белок-предшественник токсина паука состоит из сигнального пептида, пропоследовательности и зрелой цепи. К настоящему моменту никакой информации об организации генов, кодирующих предшественники двудоменных токсинов пауков, опубликовано не было. В этой работе было впервые показано, что общим свойством генов двудоменных токсинов пауков, родов *Cheiracanthium* и *Oxyopes*, является отсутствие интронов.

Как описано в разделе «Обзор литературы», большинство изученных к настоящему моменту генов, кодирующих короткие пептидные токсины пауков не содержат интронов [92–94], только паук *D. canities*, считающийся примитивным [215], имеет интроны в генах токсинов [91]. Гены, кодирующие большие белковые компоненты ядов пауков, различаются в их экзон-интронной структуре [89,216]. Среди «не-токсиновых» генов пауков наиболее изученными являются гены, кодирующие белки паутины, среди которых так же наблюдается гетерогенность. Гены *MaSp1*, *MaSp2*, и *AcSp1* паука *Latrodectus hesperus* [217–219] и *NCF-1 Nephila clavipes* [220] состоят из одного огромного экзона, в то время как гены *Flag N. clavipes* богаты интронами [220]. У близких родственников пауков скорпионов

большинство генов токсинов содержат интроны (см. соответствующую главу Обзора литературы).

К настоящему времени не была установлена ни одна последовательность генома паука, поэтому не представляется возможным установить уровень плотности интронов, характерный для пауков. Однако было показано, что плотность интронов у первичноротых (в т. ч. членистоногих) ниже, чем у вторичноротых [221,222]. Таким образом, очень вероятно, что потеря интронов характерна для пауков в целом. Скорее всего, геном пауков характеризуется высокой скоростью оборота некодирующих фрагментов (в т. ч. интронов), поскольку у беспозвоночных она гораздо выше, чем у позвоночных [223], и длина генома сильно варьирует у пауков [224]. Это так же говорит в пользу высокой частоты потери интронов у пауков.

Высокая частота потери интронов может являться эволюционным преимуществом пауков. Поддержание высокого уровня синтеза токсинов в ядовитых железах – один из их приоритетов. Как сообщалось ранее, одним из задействованных механизмов может быть полиплоидизация клеток в тканях ядовитых желез [225]. Другой способ, который мог бы повысить уровень экспрессии генов токсинов, это потеря интронов. Имеются доказательства, что гены, экспрессирующиеся на высоком уровне, проявляют тенденцию к более частой потере интронов и более редкому их приобретению [226]. В последующих исследованиях было бы интересно сравнить количество копий генов токсинов в геномной ДНК из ядовитых желез и из других частей тела паука. Такое исследование помогло бы прояснить, наблюдается ли, наряду с потерей интронов и полиплоидизацией, возможная амплификация копий генов токсинов в ядовитых железах пауков для повышения уровня продукции токсинов.

Отсутствие интронов может также вызывать возрастание частоты мутаций. Согласно модели Л. И. Патрушева, общая частота мутаций, появляющихся в кодирующих областях генома, зависит от соотношения кодирующих и некодирующих последовательностей в геноме [227]. Высокая частота замен в генах токсинов может способствовать разнообразию пептидных токсинов и возникновению новых токсинов пауков. Это наблюдение согласуется с тем, что яды пауков представляют собой комбинаторные библиотеки пептидных компонентов [228]. Кажется, что высокая частота мутаций необходима для появления таких библиотек.

## 6. Молекулярная эволюция двудоменных токсинов

### 6.1 Анализ типов отбора, действующего на двудоменные токсины

Согласно полученным данным, ускоренная эволюция не является ярко выраженной тенденцией двудоменных токсинов. В случае *Oxyopes* анализ на положительный отбор показал, что только линейные домены спайдеринов демонстрируют слабую склонность к ускоренной эволюции. Среди двух тестов LRT только результат теста M1a/M2a был значимым ( $p < 0,05$ ), и только один сайт (Leu-14), расположенный в N-концевом домене (его «базовой» части), подвергался положительному отбору. Ноттиновые домены подвергаются отрицательному отбору. Колебания в длине «дополнительных» последовательностей вносят вклад в вариабельность линейных доменов. У паука *C. punctoriium* зрелые двудоменные токсины тоже подвержены сильному отрицательному отбору.

Тенденция к отрицательному отбору, поддерживающему консервативность двудоменных токсинов, может свидетельствовать об их функциональной важности в составе яда. Примером крайней консервативности важных токсинов является, согласованная эволюция токсинов анемоны *N. vectensis*, действующих на  $\text{Na}^+$ -каналы [186]. В случае спайдеринов, линейный домен OtTx 1 – мощный цитолитик, а возможная мишень ноттинового домена – кальциевые каналы, как и в случае гомологичных ему окситоксинов. Токсины CrTx проявляют высокую инсектицидную активность и так же, как латротоксины (тоже подверженные отрицательному отбору) являются мажорными компонентами ядов: пауки «полагаются» на эффективность этих токсинов и «поддерживают» их консервативность.

К настоящему моменту не было опубликовано данных по анализу типов отбора, действующего на линейные цитолитические токсины пауков. Однако липолитические токсины скорпионов и цитотоксины змей с неспецифическим механизмом действия подвергаются отрицательному отбору [158,169]. Таким образом, невыраженность положительного отбора линейного домена спайдеринов согласуется с общей тенденцией, наблюдаемой среди неспецифических токсинов. С-концевые домены CrTx-подобных токсинов демонстрируют значимое сходство с нейротоксинами CsTx-1 и CsTx-9, однако до сих пор не была показана их нейротоксическая функция, в то время как цитолитические свойства были продемонстрированы [69]. Таким образом, нельзя исключать, что действие отрицательного отбора на CrTx-подобные токсины может быть связано с цитолитической функцией, показанной на примере CrTx 1a.

Среди ноттиновых пептидов пауков быстрая адаптивная эволюция была описана среди паралога  $\omega$ -атракотоксина Hv1a [80]. Как было описано в «Обзоре литературы», большинство коротких цистеин-богатых токсинов моллюсков, змей и скорпионов также

эволюционируют ускоренно. Однако у этих групп ядовитых животных тенденция к положительному отбору токсинов выражена в разной степени. Наиболее подверженными положительному отбору и быстро эволюционирующими являются короткие пептидные нейротоксины моллюсков *Copis* и змей. У скорпионов эта тенденция менее выраженная и можно говорить об эпизодическом положительном отборе [169], а у анемонов наблюдается даже согласованная эволюция нейротоксинов [186]. Скорее всего, это связано с тем, что моллюски *Copis* и ядовитые змеи являются относительно молодыми таксономическими группами, в которых под действием ускоренной эволюции происходит «поиск» новых эффективных токсинов. Ядовитые моллюски конусы появились около 50 млн. лет назад [80], а вовлечение трехпетельных токсинов змей в состав ядов, по разным оценкам, произошло 54 – 100 млн. лет назад [229]. Эволюционный возраст скорпионов составляет ~420 млн. лет [80]. Вероятно, скорпионы, выработав определенный набор эффективных токсинов в начале своего эволюционного пути, «предпочли законсервировать» его, в результате чего их токсины попали под давление отрицательного отбора. Таким образом, действие отрицательного отбора могло «замаскировать» первоначальное действие положительного. Еще более древние морские анемоны, вероятно, обеспечивают консервативность функционально важных токсинов с помощью согласованной эволюции. Наши результаты, показавшие склонность двудоменных токсинов пауков к отрицательному отбору, а также значительный эволюционный возраст пауков (~400 млн. лет [80]) являются еще одним подтверждением этой гипотезы.

## 6.2 Возникновение токсинов типа AMP+ICK у *Oxyopes*

Двудоменные токсины, скорее всего, образовались в результате «вставки» линейного N-концевого домена в «предковый» ноттиновый токсин между пропептидом и зрелым пептидом. Среди транслированных последовательностей кДНК десятицистеиновых токсинов были найдены потенциальные «предковые» последовательности для препропептида и C-концевого домена, в то время как для N-концевого фрагмента никаких гомологов обнаружено не было. Кроме того, не было найдено гомологичные последовательности среди аминокислотных последовательностей, хранящихся в открытых базах данных. Поскольку до сих пор не установлена ни одна последовательность генома паука, в т. ч. и пауков рода *Oxyopes*, невозможно установить, является ли последовательность, кодирующая линейный N-концевой домен, фрагментом, образовавшимся *de novo*, или она имеет гомологи среди других генов. Можно сделать несколько предположений о механизме появления двудоменных токсинов пауков рода *Oxyopes*.

В отличие от пауков, у скорпионов в генах многих токсинов содержится интрон, причем сайты сплайсинга располагаются на границе фрагментов, кодирующих сигнальный и зрелый пептиды (табл. 1 в «Обзоре литературы»). Единственный известный ген токсина паука, содержащий интроны, имеет сайты сплайсинга на границе сигнального пептида и пропептида, а также на границе пропептида и зрелого пептида [91]. Одной из гипотез происхождения линейного домена может быть его возможная эволюция из последовательности интрона, располагавшегося между экзонами, кодирующими пропептид и зрелый пептид. В пользу этого говорит то, что в нуклеотидной последовательности, кодирующей линейный домен, обнаруживается мотив донорного сайта сплайсинга. Возможно, в результате утраты акцепторного сайта сплайсинга удаление интрона стало невозможным, и он эволюционировал в кодирующую последовательность. Учитывая, что у многих скорпионов интроны генов токсинов АТ-богаты (табл. 1), высока вероятность образования лизиновых (АА[ag]), фенилаланиновых (ТТ[tc]), лейциновых (СТ[agtc] и ТТ[ag]) и аспарагиновых (АА[tc]) кодонов. Последовательность фрагмента гена, кодирующая линейный домен, так же является АТ-богатой (А+Т ~60%), что при трансляции дает (на примере *OtTx 1a-AMP*) 39% остатков лизина, лейцина – 15%, фенилаланина – 10%, аспарагина – 7% (в сумме 71%). Переход фрагмента интрона в кодирующую последовательность в результате мутации сайта сплайсинга был, например, зафиксирован среди токсинов змей [162]. Кроме того, образование генов *de novo* в результате «превращения» некодирующих последовательностей в кодирующие – довольно распространенное явление [230].

Другой возможный механизм образования двудоменных токсинов *Oxurores* – инсерция фрагмента ДНК, кодирующего линейный фрагмент, в «предковый» ген однодоменного ноттинового токсина. Вставка могла произойти как в безинтронный ген между фрагментами, кодирующими препропептид и зрелую часть, так и в последовательность возможного интрона.

### **6.3 Возникновение токсинов типа ICK+ICK у *Cheiracanthium***

Согласно литературным источникам, яд паука – комбинаторная библиотека, образующаяся в результате многократных дупликаций генов и их последующей дивергенции [69,231]. Можно было бы предположить, что эволюционное появление двудоменных токсинов произошло в результате частичной дупликации потенциального предкового гена. Это, скорее всего, верно в случае с другим двудоменным токсином *DkTx*, поскольку последовательности его доменов демонстрируют наивысший уровень сходства друг с другом (~50% идентичных а. о.), а не с какой-либо другой ноттиновой последовательностью.

Жингжаотоксин F7-15.33 [232] из яда того же паука обладает ~26% и ~35% а. о., идентичных N- и C-концевым доменам DkTx, соответственно. Это делает последовательность жингжаотоксина F7-15.33 наиболее близкой к последовательности DkTx среди известных последовательностей. Возможно, жингжаотоксин F7-15.33 и DkTx имеют общий предковый токсин, ген которого дублировался и дал начало гену DkTx.

Что касается токсина CrTx, то существуют доказательства в пользу другого механизма: скорее всего, токсины CrTx кодируются химерными генами. Хотя каждый из доменов токсинов CrTx обладает одним и тем же цистеиновым мотивом, их сходство умеренно. Самое интересное то, что они проявляют меньше сходства друг к другу, чем к некоторым другим однодоменным токсинам.

Вероятно, гены, гомологичные генам *lstxa1* и *lstxc* *L. singoriensis*, могли бы быть предковыми для двудоменных токсинов *C. puncturium*. Предшественник LSTXC1 демонстрирует сходство с фрагментом, состоящим из препро-части и N-концевого домена двудоменных токсинов, в то время как зрелый пептид LSTXA1 похож на C-концевой домен. Таким образом, гены, кодирующие белки-предшественники CrTx-подобных токсинов могут быть результатом слияния двух предковых генов за счет одного из двух возможных эволюционных механизмов.

Широко распространенный механизм образования химерных многодоменных белков – перетасовка экзонов [233]. В качестве примеров из мира членистоногих можно привести антимикробный пептид набаецин-3 из паразитической осы *Nasonia vitripennis* [234] и дефензины из разных систематических групп [235], в эволюции которых было задействовано явление перетасовки экзонов. Необходимые условия для доказательства того, что модульный белок образовался в результате перетасовки экзонов, это: (1) гомологичные фрагменты окружены негомологичными последовательностями, и (2) транспозиция обоих фрагментов была опосредована интронной рекомбинацией [233]. Как описано выше, у генов CrTx-подобных токсинов интроны не были обнаружены, так же как и среди остальных генов токсинов пауков (за исключением одного). Таким образом, транспозиция экзона из одного предкового гена однодоменного токсина в другой, приводящая к образованию химерного гена двудоменных токсинов, не кажется вероятной. Однако, принимая во внимание склонность членистоногих к потере интронов (в том числе и после перемещения экзонов [235]), нельзя окончательно отменить гипотезу, что один из возможных путей возникновения двудоменных токсинов – это явление перетасовки экзонов, за которым последовала потеря интрона.

Другой эволюционный механизм кажется более вероятным, поскольку он не требует дополнительных допущений о возможной потере интронов. У млекопитающих химерные транскрипты распространены и образуются в результате т.н. транскрипционно опосредованном слиянии генов. Если за образованием химерных транскриптов следует ретропозиция, может получиться активный химерный ген [236]. Химерный транскрипт VmКТ-и скорпиона *M. martensii* является результатом транс-сплайсинга, который, однако, не подвергался ретропозиции и закреплению в геноме [106]. В нашем случае кодирующие последовательности геномной ДНК полностью совпадают с соответствующими последовательностями кДНК. Следовательно, двудоменные токсины могли образоваться в результате транскрипционно опосредованного слияния генов, описанным у человека.

Последовательности отдельных доменов СрТх 1 проявляют сходство с токсинами CsТх-1 и CsТх-13, блокатором кальциевых каналов и энхансером нейротоксинов, соответственно. CsТх-13 действует синергично с другими CsТх-подобными токсинами, усиливая их активность. Необходимо подробное изучение биологической активности отдельных доменов для того, чтобы выяснить их роль в активности полноразмерного токсина. Если функциональное сходство между доменами СрТх и токсинами CsТх подтвердится, токсины СрТх можно было бы рассматривать как результат слияния генов, кодирующих функционально связанные белки, что широко распространено среди многоклеточных [237].

Ранее выдвигалась гипотеза [69] о том, что в процессе эволюции двудоменные токсины образовались из сложного предшественника в результате сдвига рамки считывания в последовательности, кодирующей сайт процессинга между зрелыми цепями. Это событие привело к нарушению процессинга и, как следствие, образованию двудоменных зрелых пептидов. Однако другое возможное предположение заключается в том, что белок-предшественник двудоменных токсинов – это только промежуточная стадия эволюционного процесса, ведущего к формированию нового сложного белка-предшественника. Для образования сайта процессинга типа PQM необходим сдвиг рамки считывания в последовательности, кодирующей линкер между доменами. В пользу последней гипотезы говорит то, что не один потенциальный сайт процессинга «спрятан» в +1 и +2 рамках считывания в генах, кодирующих СрТх-подобные токсины. Сдвиги рамки могут привести к образованию сайтов процессинга в нескольких точках белка-предшественника, однако только мутация между двумя доменами сможет дать начало сложному предшественнику.



## Выводы

1. В яде паука *Oxyopes takobius* обнаружено новое семейство модульных полипептидных токсинов – спайдеринов. Установлены полные аминокислотные последовательности 4 токсинов, состоящих из 108 – 113 аминокислотных остатков. Спайдерины построены из N-концевого линейного и C-концевого дисульфид-богатого доменов. Они обладают мощной антимикробной, инсектицидной и цитолитической активностью, обусловленной линейным доменом.
2. Яд паука *Cheiracanthium punctorium* является уникальным по молекулярному составу с преобладанием модульных СрТх-подобных токсинов, состоящих из двух дисульфид-богатых доменов. Выделены 3 новых токсина с высокой инсектицидной активностью, для двух из которых определены полные аминокислотные последовательности из 131 аминокислотного остатка.
3. Исследована активность линейных модулей цитоинсектотоксина 1a из яда паука *Lachesana tarabaevi*. Для инсектицидной активности токсина необходимо присутствие обоих модулей, высокая антимикробная активность обеспечивается синергией двух фрагментов.
4. Анализ кДНК из ядовитых желез пауков *O. takobius*, *Oxyopes lineatus* и *C. punctorium* позволил идентифицировать 15 последовательностей спайдеринов и 12 последовательностей СрТх-подобных токсинов. Определены последовательности генов, кодирующих двудоменные токсины *O. lineatus* и *C. punctorium*. Показано, что они не содержат интронов.
5. Анализ механизмов молекулярной эволюции двудоменных токсинов *Oxyopes* и *C. punctorium* показал, что они подвержены действию отрицательного отбора. Предполагается, что в ходе эволюции гены модульных токсинов образовались из генов однодоменных токсинов.

## Список литературы

- 1 Василевский АА, Козлов С, Гришин Е (2009) Молекулярное разнообразие яда пауков. *Успехи биологической химии* **49**, 211–227.
- 2 Rash LD, Hodgson WC (2002) Pharmacology and biochemistry of spider venoms. *Toxicon* **40**, 225–54.
- 3 Гришин Е, Волкова Т, Арсеньев А, Решетова О, Оноприенко В (1986) Структурно-функциональная характеристика аргиопина – блокатора ионных каналов из яда паука *Argiope lobata*. *Биоорганическая химия* **12**, 1121–4.
- 4 Гришин Е, Волкова Т, Арсеньев А (1988) Антагонисты глутаматных рецепторов из яда паука *Argiope lobata*. *Биоорганическая химия* **14**, 883–92.
- 5 Grishin E V, Volkova TM, Arseniev AS (1989) Isolation and structure analysis of components from venom of the spider *Argiope lobata*. *Toxicon* **27**, 541–9.
- 6 Cesar LMM, Mendes MA, Tormena CF, Marques MR, de Souza BM, Saidemberg DM, Bittencourt JC, Palma MS (2005) Isolation and chemical characterization of PwTx-II: a novel alkaloid toxin from the venom of the spider *Parawixia bistriata* (Araneidae, Araneae). *Toxicon* **46**, 786–96.
- 7 Rodrigues MCA, Guizzo R, Gobbo-Neto L, Ward RJ, Lopes NP, dos Santos WF (2004) The biological activity in mammals and insects of the nucleosidic fraction from the spider *Parawixia bistriata*. *Toxicon* **43**, 375–83.
- 8 Pretel F, Gonçalves-de-Andrade RM, Magnoli FC, da Silva MER, Ferreira JMC, van den Berg CW, Tambourgi D V (2005) Analysis of the toxic potential of venom from *Loxosceles adelaida*, a Brazilian brown spider from karstic areas. *Toxicon* **45**, 449–58.
- 9 Silvestre FG, de Castro CS, de Moura JF, Giusta MS, De Maria M, Alvares ESS, Lobato FCF, Assis RA, Gonçalves LA, Gubert IC, Chávez-Olórtegui C, Kalapothakis E (2005) Characterization of the venom from the Brazilian brown spider *Loxosceles similis* Moenkhaus, 1898 (Araneae, Sicariidae). *Toxicon* **46**, 927–36.
- 10 Zobel-Thropp PA, Kerins AE, Binford GJ (2012) Sphingomyelinase D in sicariid spider venom is a potent insecticidal toxin. *Toxicon* **60**, 265–71.
- 11 Ушкарев Ю, Гришин Е (1986) Нейротоксин каракурта и его взаимодействие с рецепторами из мозга крыс. *Биоорганическая химия* **12**, 71–80.
- 12 Usmanov PB, Kazakov I, Kalikulov D, Atakuziev BU, Yukelson LYa, Tashmukhamedov BA (1985) The channel-forming component of the Theridiidae spider venom neurotoxins. *Gen. Physiol. Biophys.* **4**, 185–93.

- 13 Nicholson GM, Gaudins A (2002) Spiders of medical importance in the Asia-Pacific: atracotoxin, latrotoxin and related spider neurotoxins. *Clin. Exp. Pharmacol. Physiol.* **29**, 785–94.
- 14 Scheer H, Madeddu L, Dozio N, Gatti G, Vicentini LM, Meldolesi J (1984) Alpha latrotoxin of black widow spider venom: an interesting neurotoxin and a tool for investigating the process of neurotransmitter release. *J. Physiol. (Paris)*. **79**, 216–21.
- 15 Barbaro KC, Cardoso JL, Eickstedt VR, Mota I (1992) Dermonecrotic and lethal components of *Loxosceles gaucho* spider venom. *Toxicon* **30**, 331–8.
- 16 Feitosa L, Gremski W, Veiga SS, Elias MC, Graner E, Mangili OC, Brentani RR (1998) Detection and characterization of metalloproteinases with gelatinolytic, fibronectinolytic and fibrinogenolytic activities in brown spider (*Loxosceles intermedia*) venom. *Toxicon* **36**, 1039–51.
- 17 Fritz LC, Tzen MC, Mauro A (1980) Different components of black widow spider venom mediate transmitter release at vertebrate and lobster neuromuscular junctions. *Nature* **283**, 486–7.
- 18 Sokolov I V, Ushkarev IA, Grasso A, Grishin E V, Lishko VK [Interaction of black widow spider toxin with bilayer phospholipid membranes]. *Ukr. Biokhim. Zh.* **55**, 179–84.
- 19 Finkelstein A, Rubin L, Tzeng M (1976) Black widow spider venom: effect of purified toxin on lipid bilayer membranes. *Science* **193**, 1009–11.
- 20 Ushkaryov YA, Rohou A, Sugita S (2008) alpha-Latrotoxin and its receptors. *Handb. Exp. Pharmacol.*, 171–206.
- 21 Kiyatkin N, Dulubova I, Grishin E (1993) Cloning and structural analysis of alpha-latroinsectotoxin cDNA. Abundance of ankyrin-like repeats. *Eur. J. Biochem.* **213**, 121–7.
- 22 Dulubova I, Krasnoperov VG, Khvotchev M V, Pluzhnikov KA, Volkova TM, Grishin E V, Vais H, Bell DR, Usherwood PN (1996) Cloning and structure of delta-latroinsectotoxin, a novel insect-specific member of the latrotoxin family: functional expression requires C-terminal truncation. *J. Biol. Chem.* **271**, 7535–43.
- 23 Kiyatkin NI, Dulubova I, Chekhovskaya IA, Grishin E V (1990) Cloning and structure of cDNA encoding alpha-latrotoxin from black widow spider venom. *FEBS Lett.* **270**, 127–31.
- 24 Данилевич В, Лукьянов С, Гришин Е (1999) Клонирование и структура гена, кодирующего  $\alpha$ -латрокрустотоксин в составе ядовитых желез паука каракурта. *Биоорганическая химия* **25**, 537–47.
- 25 Orlova E V, Rahman MA, Gowen B, Volynski KE, Ashton AC, Manser C, van Heel M, Ushkaryov YA (2000) Structure of alpha-latrotoxin oligomers reveals that divalent cation-dependent tetramers form membrane pores. *Nat. Struct. Biol.* **7**, 48–53.
- 26 Babcock JL, Civello DJ, Geren CR (1981) Purification and characterization of a toxin from brown recluse spider (*Loxosceles reclusa*) venom gland extracts. *Toxicon* **19**, 677–89.

- 27 Swanson DL, Vetter RS Loxoscelism. *Clin. Dermatol.* **24**, 213–21.
- 28 Murakami MT, Fernandes-Pedrosa MF, de Andrade S a, Gabdoulkhakov A, Betzel C, Tambourgi D V, Arni RK (2006) Structural insights into the catalytic mechanism of sphingomyelinases D and evolutionary relationship to glycerophosphodiester phosphodiesterases. *Biochem. Biophys. Res. Commun.* **342**, 323–9.
- 29 Kuhn-Nentwig L (2003) Antimicrobial and cytolytic peptides of venomous arthropods. *Cell. Mol. Life Sci.* **60**, 2651–68.
- 30 Kozlov S, Vassilevski A, Feofanov A, Surovoy AY, Karpunin DV, Grishin EV (2006) Latarcins, antimicrobial and cytolytic peptides from the venom of the spider *Lachesana tarabaei* (Zodariidae) that exemplify biomolecular diversity. *J. Biol. Chem.* **281**, 20983–92.
- 31 Vassilevski A, Kozlov S, Samsonova O, Egorova N, Karpunin D, Pluzhnikov K, Feofanov A, Grishin E (2008) Cyto-insectotoxins, a novel class of cytolytic and insecticidal peptides from spider venom. *Biochem. J.* **696**, 687–696.
- 32 Corzo G, Villegas E, Gómez-Lagunas F, Possani LD, Belokoneva OS, Nakajima T (2002) Oxyopinins, large amphipathic peptides isolated from the venom of the wolf spider *Oxyopes kitabensis* with cytolytic properties and positive insecticidal cooperativity with spider neurotoxins. *J. Biol. Chem.* **277**, 23627–37.
- 33 Torres-Larios A, Gurrola GB, Zamudio FZ, Possani LD (2000) Hadrurin, a new antimicrobial peptide from the venom of the scorpion *Hadrurus aztecus*. *Eur. J. Biochem.* **267**, 5023–31.
- 34 Dubovskii P V, Volynsky PE, Polyansky AA, Karpunin D V, Chupin V V, Efremov RG, Arseniev AS (2008) Three-dimensional structure/hydrophobicity of latarcins specifies their mode of membrane activity. *Biochemistry* **47**, 3525–33.
- 35 Polyansky AA, Vassilevski A, Volynsky PE, Vorontsova O V, Samsonova O V, Egorova NS, Krylov NA, Feofanov A, Arseniev AS, Grishin E V, Efremov RG (2009) N-terminal amphipathic helix as a trigger of hemolytic activity in antimicrobial peptides: a case study in latarcins. *FEBS Lett.* **583**, 2425–8.
- 36 Pukala TL, Boland MP, Gehman JD, Kuhn-Nentwig L, Separovic F, Bowie JH (2007) Solution structure and interaction of cupiennin 1a, a spider venom peptide, with phospholipid bilayers. *Biochemistry* **46**, 3576–85.
- 37 Dubovskii P V, Vassilevski A, Samsonova O V, Egorova NS, Kozlov S, Feofanov A, Arseniev AS, Grishin E V (2011) Novel lynx spider toxin shares common molecular architecture with defense peptides from frog skin. *FEBS J.* **278**, 4382–93.
- 38 Raghuraman H, Chattopadhyay A (2007) Melittin: a membrane-active peptide with diverse functions. *Biosci. Rep.* **27**, 189–223.
- 39 Hoskin DW, Ramamoorthy A (2008) Studies on anticancer activities of antimicrobial peptides. *Biochim. Biophys. Acta* **1778**, 357–75.

- 40 Yeaman MR, Yount NY (2003) Mechanisms of antimicrobial peptide action and resistance. *Pharmacol. Rev.* **55**, 27–55.
- 41 Garcia F, Villegas E, Espino-Solis GP, Rodriguez A, Paniagua-Solis JF, Sandoval-Lopez G, Possani LD, Corzo G (2013) Antimicrobial peptides from arachnid venoms and their microbicidal activity in the presence of commercial antibiotics. *J. Antibiot. (Tokyo)*. **66**, 3–10.
- 42 Lazarev VN, Polina NF, Shkarupeta MM, Kostrjukova ES, Vassilevski A, Kozlov S, Grishin E V, Govorun VM (2011) Spider venom peptides for gene therapy of *Chlamydia infection*. *Antimicrob. Agents Chemother.* **55**, 5367–9.
- 43 Vorontsova O V, Egorova NS, Arseniev AS, Feofanov A (2011) Haemolytic and cytotoxic action of laticin Ltc2a. *Biochimie* **93**, 227–41.
- 44 Samsonova O V, Kudryashova KS, Feofanov A (2011) N-terminal moiety of Antimicrobial peptide Ltc1-k increases its toxicity for eukaryotic cells. *Acta Naturae* **3**, 68–78.
- 45 Idiong G, Won A, Ruscito A, Leung BO, Hitchcock AP, Ianoul A (2011) Investigating the effect of a single glycine to alanine substitution on interactions of antimicrobial peptide laticin 2a with a lipid membrane. *Eur. Biophys. J.* **40**, 1087–100.
- 46 Won A, Ruscito A, Ianoul A (2012) Imaging the membrane lytic activity of bioactive peptide laticin 2a. *Biochim. Biophys. Acta* **1818**, 3072–80.
- 47 Pallaghy PK, Alewood D, Alewood PF, Norton RS (1997) Solution structure of robustoxin, the lethal neurotoxin from the funnel-web spider *Atrax robustus*. *FEBS Lett.* **419**, 191–6.
- 48 Fletcher JI, Chapman BE, Mackay JP, Howden ME, King GF (1997) The structure of versutoxin (delta-atracotoxin-Hv1) provides insights into the binding of site 3 neurotoxins to the voltage-gated sodium channel. *Structure* **5**, 1525–35.
- 49 Wang X, Connor M, Smith R, Maciejewski MW, Howden ME, Nicholson GM, Christie MJ, King GF (2000) Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. *Nat. Struct. Biol.* **7**, 505–13.
- 50 Shu Q, Lu S-Y, Gu X-C, Liang S-P (2002) The structure of spider toxin huwentoxin-II with unique disulfide linkage: evidence for structural evolution. *Protein Sci.* **11**, 245–52.
- 51 Peng K, Lin Y, Liang S-P (2006) Nuclear magnetic resonance studies on huwentoxin-XI from the Chinese bird spider *Ornithoctonus huwena*: <sup>15</sup>N labeling and sequence-specific <sup>1</sup>H, <sup>15</sup>N nuclear magnetic resonance assignments. *Acta Biochim. Biophys. Sin. (Shanghai)*. **38**, 457–66.
- 52 Klint JK, Senff S, Rupasinghe DB, Er SY, Herzig V, Nicholson GM, King GF (2012) Spider-venom peptides that target voltage-gated sodium channels: pharmacological tools and potential therapeutic leads. *Toxicon* **60**, 478–91.
- 53 Mouhat S, Andreotti N, Jouirou B, Sabatier J-M (2008) Animal toxins acting on voltage-gated potassium channels. *Curr. Pharm. Des.* **14**, 2503–18.

- 54 Norton RS, McDonough SI (2008) Peptides targeting voltage-gated calcium channels. *Curr. Pharm. Des.* **14**, 2480–91.
- 55 Smith JJ, Alphy S, Seibert AL, Blumenthal KM (2005) Differential phospholipid binding by site 3 and site 4 toxins. Implications for structural variability between voltage-sensitive sodium channel domains. *J. Biol. Chem.* **280**, 11127–33.
- 56 Suchyna TM, Tape SE, Koeppe RE, Andersen OS, Sachs F, Gottlieb PA (2004) Bilayer-dependent inhibition of mechanosensitive channels by neuroactive peptide enantiomers. *Nature* **430**, 235–40.
- 57 Jung HJ, Lee JY, Kim SH, Eu Y-J, Shin SY, Milesco M, Swartz KJ, Kim J II (2005) Solution structure and lipid membrane partitioning of VSTx1, an inhibitor of the KvAP potassium channel. *Biochemistry* **44**, 6015–23.
- 58 Liang SP, Pan X (1995) A lectin-like peptide isolated from the venom of the Chinese bird spider *Selenocosmia huwena*. *Toxicon* **33**, 875–82.
- 59 Lü S, Liang S, Gu X (1999) Three-dimensional structure of *Selenocosmia huwena* lectin-I (SHL-I) from the venom of the spider *Selenocosmia huwena* by 2D-NMR. *J. Protein Chem.* **18**, 609–17.
- 60 Choi S-J, Parent R, Guillaume C, Deregnaucourt C, Delarbre C, Ojcius DM, Montagne J-J, Célrier M-L, Phelipot A, Amiche M, Molgo J, Camadro J-M, Guette C (2004) Isolation and characterization of Psalmopeotoxin I and II: two novel antimalarial peptides from the venom of the tarantula *Psalmopoeus cambridgei*. *FEBS Lett.* **572**, 109–17.
- 61 Pimentel C, Choi S-J, Chagot B, Guette C, Camadro J-M, Darbon H (2006) Solution structure of PcFK1, a spider peptide active against *Plasmodium falciparum*. *Protein Sci.* **15**, 628–34.
- 62 Schalle J, Kämpfer U, Schürch S, Kuhn-Nentwig L, Haerberli S, Nentwig W (2001) CSTX-9, a toxic peptide from the spider *Cupiennius salei*: amino acid sequence, disulphide bridge pattern and comparison with other spider toxins containing the cystine knot structure. *Cell. Mol. Life Sci.* **58**, 1538–45.
- 63 Wullschleger B, Kuhn-Nentwig L, Tromp J, Kämpfer U, Schaller J, Schürch S, Nentwig W (2004) CSTX-13, a highly synergistically acting two-chain neurotoxic enhancer in the venom of the spider *Cupiennius salei* (Ctenidae). *Proc. Natl. Acad. Sci. U. S. A.* **101**, 11251–6.
- 64 Villegas E, Adachi-Akahane S, Bosmans F, Tytgat J, Nakajima T, Corzo G (2008) Biochemical characterization of cysteine-rich peptides from *Oxyopes* sp. venom that block calcium ion channels. *Toxicon* **52**, 228–36.
- 65 Vassilevski A, Kozlov S, Samsonova OVO, Egorova NNS, Karpunin DVD, Pluzhnikov KKA, Feofanov A, Grishin EVE (2008) Cyto-insectotoxins, a novel class of cytolytic and insecticidal peptides from spider venom. *Biochem. J.* **696**, 687–696.

- 66 Lazarev VN, Shkarupeta MM, Polina NF, Kostrjukova ES, Vassilevski A, Kozlov S, Grishin E V, Govorun VM (2013) Antimicrobial peptide from spider venom inhibits *Chlamydia trachomatis* infection at an early stage. *Arch. Microbiol.* **195**, 173–9.
- 67 Bohlen CJ, Priel A, Zhou S, King D, Siemens J, Julius D (2010) A bivalent tarantula toxin activates the capsaicin receptor, TRPV1, by targeting the outer pore domain. *Cell* **141**, 834–845.
- 68 Bae C, Kalia J, Song I, Yu J, Kim HH, Swartz KJ, Kim J II (2012) High yield production and refolding of the double-knot toxin, an activator of TRPV1 channels. *PLoS One* **7**, e51516.
- 69 Vassilevski A, Fedorova IM, Maleeva EE, Korolkova Y V, Efimova SS, Samsonova O V, Schagina L V, Feofanov A, Magazanik LG, Grishin E V (2010) Novel class of spider toxin: active principle from the yellow sac spider *Cheiracanthium punctorium* venom is a unique two-domain polypeptide. *J. Biol. Chem.* **285**, 32293–302.
- 70 Kuhn-Nentwig L, Schaller J, Nentwig W (1994) Purification of toxic peptides and the amino acid sequence of cstx-1 from the multicomponent venom of *Cupiennius salei* (araneae: tenebrionidae). *Toxicon* **32**, 287–302.
- 71 Kubista H, Mafra RA, Chong Y, Nicholson GM, Beirão PSL, Cruz JS, Boehm S, Nentwig W, Kuhn-Nentwig L, Beirao PS (2007) CSTX-1, a toxin from the venom of the hunting spider *Cupiennius salei*, is a selective blocker of L-type calcium channels in mammalian neurons. *Neuropharmacology* **52**, 1650–62.
- 72 Kuhn-Nentwig L, Fedorova IM, Lüscher BP, Kopp LS, Trachsel C, Schaller J, Vu XL, Seebeck T, Streitberger K, Nentwig W, Sigel E, Magazanik LG (2012) A venom-derived neurotoxin, CsTx-1, from the spider *Cupiennius salei* exhibits cytolytic activities. *J. Biol. Chem.* **287**, 25640–9.
- 73 Kuhn-Nentwig L, Schaller J, Kämpfer U, Imboden H, Malli H, Nentwig W (2000) A lysine rich C-terminal tail is directly involved in the toxicity of CSTX-1, a neurotoxic peptide from the venom of the spider *Cupiennius salei*. *Arch. Insect Biochem. Physiol.* **44**, 101–11.
- 74 Kuzmenkov AI, Fedorova IM, Vassilevski A, Grishin E V (2013) Cysteine-rich toxins from *Lachesana tarabaevi* spider venom with amphiphilic C-terminal segments. *Biochim. Biophys. Acta* **1828**, 724–31.
- 75 Pawlak J, Kini RM (2008) Unique gene organization of colubrid three-finger toxins: complete cDNA and gene sequences of denmotoxin, a bird-specific toxin from colubrid snake *Boiga dendrophila* (Mangrove Catsnake). *Biochimie* **90**, 868–77.
- 76 Fuse N, Tsuchiya T, Nonomura Y, Menez a, Tamiya T (1990) Structure of the snake short-chain neurotoxin, erabutoxin c, precursor gene. *Eur. J. Biochem.* **193**, 629–33.
- 77 Chang L, Wang J-J, Cheng Y-C, Chou W-M (2008) Genetic organization of *Bungarus multicinctus* protease inhibitor-like proteins. *Toxicon* **51**, 1490–5.
- 78 Rádis-Baptista G, Kubo T, Oguiura N, Svartman M, Almeida TM., Batistic RF, Oliveira EB, Vianna-Morgante ÂM, Yamane T (2003) Structure and chromosomal localization of the

- gene for crotoamine, a toxin from the South American rattlesnake, *Crotalus durissus terrificus*. *Toxicon* **42**, 747–752.
- 79 Ogawa T, Nakashima K, Oda N, Shimohigashi Y, Ohno M, Hattori S, Hattori M, Sakaki Y, Kihara H (1995) *Trimeresurus flavoviridis* venom gland phospholipase A2 isozymes genes have evolved via accelerated substitutions. *J. Mol. Recognit.* **8**, 40–6.
- 80 Sollod BL, Wilson D, Zhaxybayeva O, Gogarten JP, Drinkwater R, King GF (2005) Were arachnids the first to use combinatorial peptide libraries? *Peptides* **26**, 131–9.
- 81 Frazão B, Vasconcelos V, Antunes A (2012) Sea anemone (Cnidaria, Anthozoa, Actiniaria) toxins: an overview. *Mar. Drugs* **10**, 1812–51.
- 82 Colledge CJ, Hunsperger JP, Imperial JS, Hillyard DR (1992) Precursor structure of omega-conotoxin GVIA determined from a cDNA clone. *Toxicon* **30**, 1111–6.
- 83 Olivera BM, Walker C, Cartier GE, Hooper D, Santos a D, Schoenfeld R, Shetty R, Watkins M, Bandyopadhyay P, Hillyard DR (1999) Speciation of cone snails and interspecific hyperdivergence of their venom peptides. Potential evolutionary significance of introns. *Ann. N. Y. Acad. Sci.* **870**, 223–37.
- 84 Conticello SG, Gilad Y, Avidan N, Ben-Asher E, Levy Z, Fainzilber M (2001) Mechanisms for evolving hypervariability: the case of conopeptides. *Mol. Biol. Evol.* **18**, 120–31.
- 85 Vassilevski A, Kozlov S, Grishin E V (2008) Antimicrobial peptide precursor structures suggest effective production strategies. *Recent Pat. Inflamm. Allergy Drug Discov.* **2**, 58–63.
- 86 Kozlov S, Vassilevski A, Grishin E (2009) Secreted protein and peptide biosynthesis: precursor structures and processing mechanisms. In *Protein Biosynthesis* pp. 225–248.
- 87 Kiyatkin NI, Kulikovskaya IM, Grishin E V, Beadle DJ, King LA (1995) Functional characterization of black widow spider neurotoxins synthesised in insect cells. *Eur. J. Biochem.* **230**, 854–9.
- 88 Данилевич В, Гришин Е (2000) Хромосомные гены нейротоксинов паука каракурта не содержат интронов. *Биоорганическая химия* **26**, 933–9.
- 89 Binford GJ, Cordes MHJ, Wells MA (2005) Sphingomyelinase D from venoms of *Loxosceles* spiders: evolutionary insights from cDNA sequences and gene structure. *Toxicon* **45**, 547–60.
- 90 Kozlov S, Malyavka A, McCutchen B, Lu A, Schepers E, Herrmann R, Grishin E (2005) A novel strategy for the identification of toxinlike structures in spider venom. *Proteins* **59**, 131–40.
- 91 Krapcho KJ, Kral Jr. RM, Vanwagenen BC, Eppler KG, Morgan TK, Kral RM (1995) Characterization and cloning of insecticidal peptides from the primitive weaving spider *Diguetia canities*. *Insect Biochem. Mol. Biol.* **25**, 991–1000.



- 92 Pineda SS, Wilson D, Mattick JS, King GF (2012) The lethal toxin from Australian funnel-web spiders is encoded by an intronless gene. *PLoS One* **7**, e43699.
- 93 Tang X, Zhang Y, Hu W, Xu D, Tao H, Yang X, Li Y, Jiang L, Liang S (2010) Molecular diversification of peptide toxins from the tarantula *Haplopelma hainanum* (*Ornithoctonus hainana*) venom based on transcriptomic, peptidomic, and genomic analyses. *J. Proteome Res.* **9**, 2550–2564.
- 94 Qiao P, Zuo X-PP, Chai Z-FF, Ji Y-HH (2004) The cDNA and genomic DNA organization of a novel toxin SHT-I from spider *Ornithoctonus huwena*. *Acta Biochim. Biophys. Sin. (Shanghai)*. **36**, 656–60.
- 95 Jiang L, Chen J, Peng L, Zhang Y, Xiong X, Liang S (2008) Genomic organization and cloning of novel genes encoding toxin-like peptides of three superfamilies from the spider *Ornithoctonus huwena*. *Peptides* **29**, 1679–84.
- 96 Froy O, Sagiv T, Poreh M, Urbach D, Zilberberg N, Gurevitz M (1999) Dynamic diversification from a putative common ancestor of scorpion toxins affecting sodium, potassium, and chloride channels. *J. Mol. Evol.* **48**, 187–96.
- 97 Cao Z, Yu Y, Wu Y, Hao P, Di Z, He Y, Chen Z, Yang W, Shen Z, He X, Sheng J, Xu X, Pan B, Feng J, Yang X, Hong W, Zhao W, Li Z, Huang K, Li T, Kong Y, Liu H, Jiang D, Zhang B, Hu J, Hu Y, Wang B, Dai J, Yuan B, Feng Y, Huang W, Xing X, Zhao G, Li X, Li Y, Li W (2013) The genome of *Mesobuthus martensii* reveals a unique adaptation model of arthropods. *Nat. Commun.* **4**, 2602.
- 98 Zhijian C, Yun X, Chao D, Shunyi Z, Shijin Y, Yingliang W, Wenxin L (2006) Cloning and characterization of a novel calcium channel toxin-like gene BmCa1 from Chinese scorpion *Mesobuthus martensii* Karsch. *Peptides* **27**, 1235–40.
- 99 Nie Y, Zeng X-C, Luo X, Wu S, Zhang L, Cao H, Zhou J, Zhou L (2012) Tremendous intron length differences of the BmKBT and a novel BmKBT-like peptide genes provide a mechanical basis for the rapid or constitutive expression of the peptides. *Peptides* **37**, 150–6.
- 100 Zeng X-C, Zhang L, Nie Y, Luo X (2012) Identification and molecular characterization of three new K<sup>+</sup>-channel specific toxins from the Chinese scorpion *Mesobuthus martensii* Karsch revealing intronic number polymorphism and alternative splicing in duplicated genes. *Peptides* **34**, 311–23.
- 101 Alami M, Céard B, Legros C, Bougis P, Martin-Eauclaire M-F (2006) Genomic characterisation of the toxin Amm VIII from the scorpion *Androctonus mauretanicus mauretanicus*. *Toxicon* **47**, 531–6.
- 102 Legros C, Bougis P, Martin-Eauclaire M-F (1997) Genomic organization of the KTX2 gene, encoding a 'short' scorpion toxin active on K<sup>+</sup> channels. *FEBS Lett.* **402**, 45–49.
- 103 Delabre M, Pasero P, Marilley M, Bougis P, Biochimie L De, Associte DR, National C, Recherche D (1995) Promoter structure and intron-exon organization of a scorpion  $\alpha$ -toxin gene. *Biochem.* 6729–6736.

- 104 Zhu S, Tytgat J (2004) The scorpine family of defensins: gene structure, alternative polyadenylation and fold recognition. *Cell. Mol. Life Sci.* **61**, 1751–63.
- 105 Uawonggul N, Thammasirirak S, Chaveerach A, Arkaravichien T, Bunyatratchata W, Ruangjirachuporn W, Jeerranaiprepame P, Nakamura T, Matsuda M, Kobayashi M, Hattori S, Daduang S (2007) Purification and characterization of Heteroscorpine-1 (HS-1) toxin from *Heterometrus laoticus* scorpion venom. *Toxicon* **49**, 19–29.
- 106 Zeng X-C, Luo F, Li W-X (2006) Characterization of a novel cDNA encoding a short venom peptide derived from venom gland of scorpion *Buthus martensii* Karsch: trans-splicing may play an important role in the diversification of scorpion venom peptides. *Peptides* **27**, 675–81.
- 107 Ma Y, Zhao R, Li S, Fan S, Wu Y, Liu H, Cao Z, Li W (2009) Characterization of LmTxLP11 and LmVP1.1 transcripts and genomic organizations: alternative splicing contributing to the diversity of scorpion venom peptides. *Toxicon* **53**, 129–34.
- 108 Lewis RJ, Dutertre S, Vetter I, Christie MJ (2012) *Conus* venom peptide pharmacology. *Pharmacol. Rev.* **64**, 259–98.
- 109 Yuan D-D, Han Y-H, Wang C-G, Chi C-W (2007) From the identification of gene organization of alpha conotoxins to the cloning of novel toxins. *Toxicon* **49**, 1135–49.
- 110 Zhou M, Wang L, Wu Y, Zhu X, Feng Y, Chen Z, Li Y, Sun D, Ren Z, Xu A (2013) Characterizing the evolution and functions of the M-superfamily conotoxins. *Toxicon* **76**, 150–9.
- 111 Chai Z-F, Zhu M-M, Bai Z-T, Liu T, Tan M, Pang X-Y, Ji Y-H (2006) Chinese-scorpion (*Buthus martensii* Karsch) toxin BmK alphaIV, a novel modulator of sodium channels: from genomic organization to functional analysis. *Biochem. J.* **399**, 445–53.
- 112 Zhu S, Gao B (2006) Molecular characterization of a possible progenitor sodium channel toxin from the Old World scorpion *Mesobuthus martensii*. *FEBS Lett.* **580**, 5979–87.
- 113 Ji Y-H, Wang W-X, Ye J-G, He L-L, Li Y-J, Yan Y-P, Zhou Z (2003) Martentoxin, a novel K<sup>+</sup>-channel-blocking peptide: purification, cDNA and genomic cloning, and electrophysiological and pharmacological characterization. *J. Neurochem.* **84**, 325–35.
- 114 Wu JJ, Dai L, Lan ZD, Chi CW (2000) The gene cloning and sequencing of Bm-12, a chlorotoxin-like peptide from the scorpion *Buthus martensii* Karsch. *Toxicon* **38**, 661–8.
- 115 Alami M, Ouafik L, Céard B, Legros C, Bougis P, Martin-Eauclaire MF (2001) Characterisation of the gene encoding the alpha-toxin Amm V from the scorpion *Androctonus mauretanicus mauretanicus*. *Toxicon* **39**, 1579–85.
- 116 Laraba-djebaris F, Legros C, Crests M, Cbard B, Romi W, Mansuelle P, Jacquetp G, Rietschoten J Van, Gola M, Pierre E (1994) The Kaliotoxin Family Enlarged. **269**, 32835–32843.

- 117 Becerril B, Corona M, Mejía MC, Martin BM, Lucas S, Bolivar F, Possani LD (1993) The genomic region encoding toxin gamma from the scorpion *Tityus serrulatus* contains an intron. *FEBS Lett.* **335**, 6–8.
- 118 Corona M, Zurita M, Possani LD, Becerril B (1996) Cloning and characterization of the genomic region encoding toxin IV-5 from the scorpion *Tityus serrulatus* Lutz and Mello. *Toxicon* **34**, 251–6.
- 119 Chang L, Lin J, Chou YC, Hong E (1997) Genomic structures of cardiotoxin 4 and cobrotoxin from *Naja naja atra* (Taiwan cobra). *Biochem. Biophys. Res. Commun.* **239**, 756–62.
- 120 Purification CV, Organization G, Chang L, Lin S, Wu B, Lin J (1997) A Novel Neurotoxin, Cobrotoxin 6, from *Naja naja atra* (Taiwan). **1259**, 1252–1259.
- 121 Chang L, Lin S, Huang H, Hsiao M (1999) Genetic organization of alpha-bungarotoxins from *Bungarus multicinctus* (Taiwan banded krait): evidence showing that the production of alpha-bungarotoxin isotoxins is not derived from edited mRNAs. *Nucleic Acids Res.* **27**, 3970–5.
- 122 Chang L, Chung C, Lin J, Hong E (2002) Organization and phylogenetic analysis of kappa-bungarotoxin genes from *Bungarus multicinctus* (Taiwan banded krait). *Genetica* **115**, 213–21.
- 123 Kuch U, Molles BE, Omori-Satoh T, Chanhome L, Samejima Y, Mebs D (2003) Identification of alpha-bungarotoxin (A31) as the major postsynaptic neurotoxin, and complete nucleotide identity of a genomic DNA of *Bungarus candidus* from Java with exons of the *Bungarus multicinctus* alpha-bungarotoxin (A31) gene. *Toxicon* **42**, 381–390.
- 124 Chang L, Lin S-K, Chung C (2004) Molecular cloning and evolution of the genes encoding the precursors of taiwan cobra cardiotoxin and cardiotoxin-like basic protein. *Biochem. Genet.* **42**, 429–40.
- 125 Doley R, Pahari S, Mackessy SP, Kini RM (2008) Accelerated exchange of exon segments in Viperid three-finger toxin genes (*Sistrurus catenatus edwardsii*; Desert Massasauga). *BMC Evol. Biol.* **8**, 196.
- 126 St Pierre L, Earl ST, Filippovich I, Sorokina N, Masci PP, De Jersey J, Lavin MF (2008) Common evolution of waprin and kunitz-like toxin families in Australian venomous snakes. *Cell. Mol. Life Sci.* **65**, 4039–54.
- 127 Bazaá A, Juárez P, Marrakchi N, Bel Lasfer Z, El Ayeb M, Harrison R a, Calvete JJ, Sanz L (2007) Loss of introns along the evolutionary diversification pathway of snake venom disintegrins evidenced by sequence analysis of genomic DNA from *Macrovipera lebetina transmediterranea* and *Echis ocellatus*. *J. Mol. Evol.* **64**, 261–71.
- 128 Kordis D, Gubensek F (1996) Ammodytoxin C gene helps to elucidate the irregular structure of Crotalinae group II phospholipase A2 genes. *Eur. J. Biochem.* **240**, 83–90.

- 129 Chu Y-P, Chang L (2002) The organization of the genes encoding the A chains of beta-bungarotoxins: evidence for the skipping of exon. *Toxicon* **40**, 1437–43.
- 130 Oguiura N, Boni-Mitake M, Rádis-Baptista G (2005) New view on crotoamine, a small basic polypeptide myotoxin from South American rattlesnake venom. *Toxicon* **46**, 363–70.
- 131 Cheng Y-C, Chen K-C, Lin S-K, Chang L (2006) Divergence of genes encoding B chains of beta-bungarotoxins. *Toxicon* **47**, 322–9.
- 132 Afifiyan F, Armugam A, Tan CH, Gopalakrishnakone P, Jeyaseelan K (1999) Postsynaptic alpha-neurotoxin gene of the spitting cobra, *Naja naja sputatrix*: structure, organization, and phylogenetic analysis. *Genome Res.* **9**, 259–66.
- 133 Fry BG, Roelants K, Champagne DE, Scheib H, Tyndall JD a, King GF, Nevalainen TJ, Norman J a, Lewis RJ, Norton RS, Renjifo C, de la Vega RCR (2009) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu. Rev. Genomics Hum. Genet.* **10**, 483–511.
- 134 Casewell NR, Wüster W, Vonk FJ, Harrison R a, Fry BG (2013) Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol. Evol.* **28**, 219–29.
- 135 Vonk FJ, Casewell NR, Henkel C V, Heimberg AM, Jansen HJ, McCleary RJR, Kerckamp HME, Vos RA, Guerreiro I, Calvete JJ, Wüster W, Woods AE, Logan JM, Harrison RA, Castoe TA, de Koning APJ, Pollock DD, Yandell M, Calderon D, Renjifo C, Currier RB, Salgado D, Pla D, Sanz L, Hyder AS, Ribeiro JMC, Arntzen JW, van den Thillart GEEJM, Boetzer M, Pirovano W, Dirks RP, Spaink HP, Duboule D, McGlenn E, Kini RM, Richardson MK (2013) The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20651–6.
- 136 Fry BG (2005) From genome to “venome”: molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. *Genome Res.* **15**, 403–20.
- 137 Fry BG, Scheib H, de L M Junqueira de Azevedo I, Silva DA, Casewell NR (2012) Novel transcripts in the maxillary venom glands of advanced snakes. *Toxicon* **59**, 696–708.
- 138 Violette A, Leonardi A, Piquemal D, Terrat Y (2012) Recruitment of glycosyl hydrolase proteins in a cone snail venomous arsenal: further insights into biomolecular features of *Conus* venoms. *Mar. Drugs*, 258–280.
- 139 Balhara KS, Stolbach A (2014) Marine envenomations. *Emerg. Med. Clin. North Am.* **32**, 223–43.
- 140 Fry BG, Roelants K, Norman J a (2009) Tentacles of venom: toxic protein convergence in the Kingdom Animalia. *J. Mol. Evol.* **68**, 311–21.
- 141 Ruder T, Sunagar K, Undheim E a B, Ali S a, Wai T-C, Low DHW, Jackson TNW, King GF, Antunes A, Fry BG (2013) Molecular phylogeny and evolution of the proteins encoded by coleoid (cuttlefish, octopus, and squid) posterior venom glands. *J. Mol. Evol.* **76**, 192–204.

- 142 Cordes MHJ, Binford GJ (2006) Lateral gene transfer of a dermonecrotic toxin between spiders and bacteria. *Bioinformatics* **22**, 264–8.
- 143 Zhu S, Gao B, Deng M, Yuan Y, Luo L, Peigneur S, Xiao Y, Liang S, Tytgat J (2010) Drosotoxin, a selective inhibitor of tetrodotoxin-resistant sodium channels. *Biochem. Pharmacol.* **80**, 1296–302.
- 144 Zhu S, Peigneur S, Gao B, Umetsu Y, Ohki S, Tytgat J (2014) Experimental conversion of a defensin into a neurotoxin: implications for origin of toxic function. *Mol. Biol. Evol.*
- 145 Wong ESW, Belov K (2012) Venom evolution through gene duplications. *Gene* **496**, 1–7.
- 146 Zhu S, Darbon H, Dyason K, Verdonck F, Tytgat J (2003) Evolutionary origin of inhibitor cysteine knot peptides. *FASEB J.* **17**, 1765–7.
- 147 Escoubas P (2006) Molecular diversification in spider venoms: a web of combinatorial peptide libraries. *Mol. Divers.* **10**, 545–54.
- 148 Duda T, Palumbi S (1999) Molecular genetics of ecological diversification: duplication and rapid evolution of toxin genes of the venomous gastropod *Conus*. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 6820–6823.
- 149 Kini RM, Doley R (2010) Structure, function and evolution of three-finger toxins: mini proteins with multiple targets. *Toxicon* **56**, 855–67.
- 150 Wong ESW, Papenfuss AT, Whittington CM, Warren WC, Belov K (2012) A limited role for gene duplications in the evolution of platypus venom. *Mol. Biol. Evol.* **29**, 167–77.
- 151 Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet.* **39**, 121–52.
- 152 Pi C, Liu J, Peng C, Liu Y, Jiang X, Zhao Y, Tang S, Wang L, Dong M, Chen S, Xu A (2006) Diversity and evolution of conotoxins based on gene expression profiling of *Conus litteratus*. *Genomics* **88**, 809–19.
- 153 Chang D, Duda T (2012) Extensive and continuous duplication facilitates rapid evolution and diversification of gene families. *Mol. Biol. Evol.* **29**, 2019–29.
- 154 Binford GJ, Bodner MR, Cordes MHJ, Baldwin KL, Rynerson MR, Burns SN, Zobel-Thropp P a (2009) Molecular evolution, functional variation, and proposed nomenclature of the gene family that includes sphingomyelinase D in sicariid spider venoms. *Mol. Biol. Evol.* **26**, 547–66.
- 155 Zhang J (2003) Evolution by gene duplication: an update. *Trends Ecol. Evol.* **18**, 292–298.
- 156 Maor-Shoshani a, Reuven NB, Tomer G, Livneh Z (2000) Highly mutagenic replication by DNA polymerase V (UmuC) provides a mechanistic basis for SOS untargeted mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 565–70.

- 157 Tang M, Pham P, Shen X, Taylor JS, O'Donnell M, Woodgate R, Goodman MF (2000) Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature* **404**, 1014–8.
- 158 Sunagar K, Jackson TNW, Undheim E a B, Ali S a, Antunes A, Fry BG (2013) Three-fingered RAVERS: Rapid Accumulation of Variations in Exposed Residues of snake venom toxins. *Toxins (Basel)*. **5**, 2172–208.
- 159 Juarez P, Comas I, Gonzalez-Candelas F, Calvete JJ, Juárez P, González-Candelas F (2008) Evolution of snake venom disintegrins by positive Darwinian selection. *Mol. Biol. Evol.* **25**, 2391–407.
- 160 Nakashima K, Ogawa T, Oda N, Hattori M, Sakaki Y, Kihara H, Ohno M (1993) Accelerated evolution of *Trimeresurus flavoviridis* venom gland phospholipase A2 isozymes. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5964–5968.
- 161 Ogawa T, Nakashima K, Nobuhisa I, Deshimaru M, Shimohigashi Y, Fukumaki Y, Sakaki Y, Hattori S, Ohno M (1996) Accelerated evolution of snake venom phospholipase A2 isozymes for acquisition of diverse physiological functions. *Toxicon* **34**, 1229–36.
- 162 Fujimi T., Nakajyo T, Nishimura E, Ogura E, Tsuchiya T, Tamiya T (2003) Molecular evolution and diversification of snake toxin genes, revealed by analysis of intron sequences. *Gene* **313**, 111–118.
- 163 Pawlak J, Mackessy SP, Fry BG, Bhatia M, Mourier G, Fruchart-Gaillard C, Servent D, Ménez R, Stura E, Ménez A, Kini RM (2006) Denmotoxin, a three-finger toxin from the colubrid snake *Boiga dendrophila* (Mangrove Catsnake) with bird-specific activity. *J. Biol. Chem.* **281**, 29030–41.
- 164 Doley R, Mackessy SP, Kini RM (2009) Role of accelerated segment switch in exons to alter targeting (ASSET) in the molecular evolution of snake venom proteins. *BMC Evol. Biol.* **9**, 146.
- 165 Doley R, Tram NNB, Reza MA, Kini RM (2008) Unusual accelerated rate of deletions and insertions in toxin genes in the venom glands of the pygmy copperhead (*Austrelaps labialis*) from Kangaroo island. *BMC Evol. Biol.* **8**, 70.
- 166 Li M, Fry BG, Kini RM (2005) Eggs-only diet: its implications for the toxin profile changes and ecology of the marbled sea snake (*Aipysurus eydouxii*). *J. Mol. Evol.* **60**, 81–9.
- 167 Kozminsky-Atias A, Bar-Shalom A, Mishmar D, Zilberberg N (2008) Assembling an arsenal, the scorpion way. *BMC Evol. Biol.* **8**, 333.
- 168 Gao B, Zhu S (2012) Accelerated evolution and functional divergence of scorpion short-chain K<sup>+</sup>-channel toxins after speciation. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **163**, 238–45.
- 169 Sunagar K, Undheim EAB, Chan AHC, Koludarov I, Muñoz-Gómez SA, Antunes A, Fry BG (2013) Evolution stings: the origin and diversification of scorpion toxin peptide scaffolds. *Toxins (Basel)*. **5**, 2456–87.

- 170 Zhu S, Bosmans F, Tytgat J (2004) Adaptive evolution of scorpion sodium channel toxins. *J. Mol. Evol.* **58**, 145–53.
- 171 Weinberger H, Moran Y, Gordon D, Turkov M, Kahn R, Gurevitz M (2010) Positions under positive selection – key for selectivity and potency of scorpion alpha-toxins. *Mol. Biol. Evol.* **27**, 1025–34.
- 172 Zhu S, Peigneur S, Gao B, Lu X, Cao C, Tytgat J (2012) Evolutionary diversification of *Mesobuthus*  $\alpha$ -scorpion toxins affecting sodium channels. *Mol. Cell. Proteomics* **11**, M111.012054.
- 173 Tian C, Yuan Y, Zhu S (2008) Positively selected sites of scorpion depressant toxins: possible roles in toxin functional divergence. *Toxicon* **51**, 555–62.
- 174 Gao B, Peigneur S, Dalziel J, Tytgat J, Zhu S (2011) Molecular divergence of two orthologous scorpion toxins affecting potassium channels. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* **159**, 313–21.
- 175 Céard B, Martin-Eauclaire M, Bougis P (2001) Evidence for a position-specific deletion as an evolutionary link between long-and short-chain scorpion toxins. *FEBS Lett.* **494**, 246–248.
- 176 He Q, Duan Z, Yu Y, Liu Z, Liu Z, Liang S (2013) The venom gland transcriptome of *Latrodectus tredecimguttatus* revealed by deep sequencing and cDNA library analysis. *PLoS One* **8**, e81357.
- 177 Duan Z, Cao R, Jiang L, Liang S (2013) A combined de novo protein sequencing and cDNA library approach to the venom analysis of Chinese spider *Araneus ventricosus*. *J. Proteomics* **78**, 416–27.
- 178 Jiang L, Zhang D, Zhang Y, Peng L, Chen J, Liang S (2010) Venomics of the spider *Ornithoctonus huwena* based on transcriptomic versus proteomic analysis. *Comp. Biochem. Physiol. Part D. Genomics Proteomics* **5**, 81–8.
- 179 Chen J, Zhao L, Jiang L, Meng E, Zhang Y, Xiong X, Liang S (2008) Transcriptome analysis revealed novel possible venom components and cellular processes of the tarantula *Chilobrachys jingzhao* venom gland. *Toxicon* **52**, 794–806.
- 180 Fernandes-Pedrosa M de F, Junqueira-de-Azevedo I de LM, Gonçalves-de-Andrade RM, Kobashi LS, Almeida DD, Ho PL, Tambourgi D V (2008) Transcriptome analysis of *Loxosceles laeta* (Araneae, Sicariidae) spider venomous gland using expressed sequence tags. *BMC Genomics* **9**, 279.
- 181 Zhang Y, Chen J, Tang X, Wang F, Jiang L, Xiong X, Wang M, Rong M, Liu Z, Liang S (2010) Transcriptome analysis of the venom glands of the Chinese wolf spider *Lycosa singoriensis*. *Zoology (Jena)*. **113**, 10–8.
- 182 Diego-García E, Peigneur S, Waelkens E, Debaveye S, Tytgat J (2010) Venom components from *Citharischius crawshayi* spider (Family Theraphosidae): exploring transcriptome, venomics, and function. *Cell. Mol. Life Sci.* **67**, 2799–813.

- 183 Diao J, Lin Y, Tang J, Liang S (2003) cDNA sequence analysis of seven peptide toxins from the spider *Selenocosmia huwena*. *Toxicon* **42**, 715–723.
- 184 Jiang L, Liu C, Duan Z, Deng M, Tang X, Liang S (2013) Transcriptome analysis of venom glands from a single fishing spider *Dolomedes mizhoanus*. *Toxicon* **73**, 23–32.
- 185 Garb JE, Hayashi CY (2013) Molecular evolution of  $\alpha$ -latrotoxin, the exceptionally potent vertebrate neurotoxin in black widow spider venom. *Mol. Biol. Evol.* **30**, 999–1014.
- 186 Moran Y, Gurevitz M (2006) When positive selection of neurotoxin genes is missing. The riddle of the sea anemone *Nematostella vectensis*. *FEBS J.* **273**, 3886–92.
- 187 Eickbush TH, Eickbush DG (2007) Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* **175**, 477–85.
- 188 Szostak JW, Wu R (1980) Unequal crossing over in the ribosomal DNA of *Saccharomyces cerevisiae*. *Nature* **284**, 426–30.
- 189 Moran Y, Weinberger H, Sullivan JC, Reitzel AM, Finnerty JR, Gurevitz M (2008) Concerted evolution of sea anemone neurotoxin genes is revealed through analysis of the *Nematostella vectensis* genome. *Mol. Biol. Evol.* **25**, 737–47.
- 190 Nicosia A, Maggio T, Mazzola S, Cuttitta A (2013) Evidence of accelerated evolution and ectodermal-specific expression of presumptive BDS toxin cDNAs from *Anemonia viridis*. *Mar. Drugs* **11**, 4213–31.
- 191 Isaeva MP, Chausova VE, Zelepuga EA, Guzev K V, Tabakmakher VM, Monastyrnaya MM, Kozlovskaya EP (2012) A new multigene superfamily of Kunitz-type protease inhibitors from sea anemone *Heteractis crispa*. *Peptides* **34**, 88–97.
- 192 Moran Y, Weinberger H, Lazarus N, Gur M, Kahn R, Gordon D, Gurevitz M (2009) Fusion and retrotransposition events in the evolution of the sea anemone *Anemonia viridis* neurotoxin genes. *J. Mol. Evol.* **69**, 115–24.
- 193 Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, Kim PM, Palejev D, Carriero NJ, Du L, Taillon BE, Chen Z, Tanzer A, Saunders ACE, Chi J, Yang F, Carter NP, Hurler ME, Weissman SM, Harkins TT, Gerstein MB, Egholm M, Snyder M (2007) Paired-end mapping reveals extensive structural variation in the human genome. *Science* **318**, 420–6.
- 194 Vassilevski A, Kozlov S, Egorov TA, Grishin E V (2010) Purification and characterization of biologically active peptides from spider venoms. *Methods Mol. Biol.* **615**, 87–100.
- 195 Carmichael J, DeGraff WG, Gazdar AF, Minna JD, Mitchell JB (1987) Evaluation of a tetrazolium-based semiautomated colorimetric assay: assessment of radiosensitivity. *Cancer Res.* **47**, 943–6.
- 196 Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–80.



- 197 Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739.
- 198 Reese MG, Eeckman FH, Kulp D, Haussler D (1997) Improved splice site detection in Genie. *J. Comput. Biol.* **4**, 311–23.
- 199 Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–25.
- 200 Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 11030–11035.
- 201 Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 418–26.
- 202 Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591.
- 203 Xu B, Yang Z (2013) PAMLX: A Graphical User Interface for PAML. *Mol. Biol. Evol.* **30**, 2723–4.
- 204 Cordeiro MN, de Figueiredo SG, Valentim A do C, Diniz CR, von Eickstedt VR, Gilroy J, Richardson M (1993) Purification and amino acid sequences of six Tx3 type neurotoxins from the venom of the Brazilian “armed” spider *Phoneutria nigriventer* (Keys). *Toxicon* **31**, 35–42.
- 205 Richardson M, Pimenta AMC, Bemquerer MP, Santoro MM, Beirao PSL, Lima ME, Figueiredo SG, Bloch C, Vasconcelos EAR, Campos FAP, Gomes PC, Cordeiro MN Comparison of the partial proteomes of the venoms of Brazilian spiders of the genus *Phoneutria*. *Comp. Biochem. Physiol. C. Toxicol. Pharmacol.* **142**, 173–87.
- 206 Li Y (2011) Recombinant production of antimicrobial peptides in *Escherichia coli*: a review. *Protein Expr. Purif.* **80**, 260–7.
- 207 Hatahet F, Boyd D, Beckwith J (2014) Disulfide bond formation in prokaryotes: History, diversity and design. *Biochim. Biophys. Acta.*
- 208 Li Y (2009) Carrier proteins for fusion expression of antimicrobial peptides in *Escherichia coli*. *Biotechnol. Appl. Biochem.* **54**, 1–9.
- 209 Berkmen M (2012) Production of disulfide-bonded proteins in *Escherichia coli*. *Protein Expr. Purif.* **82**, 240–51.
- 210 pET System Manual Novagen (2003).
- 211 Shlyapnikov YM, Andreev Y, Kozlov S, Vassilevski A, Grishin E V (2008) Bacterial production of laticin 2a, a potent antimicrobial peptide from spider venom. *Protein Expr. Purif.* **60**, 89–95.

- 212 Bulaj G (2005) Formation of disulfide bonds in proteins and peptides. *Biotechnol. Adv.* **23**, 87–92.
- 213 Conde R, Zamudio FZ, Rodríguez MH, Possani LD (2000) Scorpine, an anti-malaria and anti-bacterial agent purified from scorpion venom. *FEBS Lett.* **471**, 165–8.
- 214 Kuhn-Nentwig L, Schaller J, Nentwig W (1994) Purification of toxic peptides and the amino acid sequence of CSTX-1 from the multicomponent venom of *Cupiennius salei* (Araneae:Ctenidae). *Toxicon* **32**, 287–302.
- 215 Lopez A (1983) Some Observations on the Internal Anatomy of *Diguetia canities* (McCook, 1890) (Araneae, Diguetidae). *J. Arachnol.* **11**, 377–384.
- 216 Данилевич В, Гришин Е (2000) Хромосомные гены нейротоксинов паука каракурта не содержат интронов. *Биоорганическая химия* **26**, 933–939.
- 217 Ayoub NA, Garb JE, Kuelbs A, Hayashi CY (2013) Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). *Mol. Biol. Evol.* **30**, 589–601.
- 218 Ayoub NA, Garb JE, Tinghitella RM, Collin MA, Hayashi CY (2007) Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS One* **2**, e514.
- 219 Tai PL, Hwang GY, Tso IM (2004) Inter-specific sequence conservation and intra-individual sequence variation in a spider silk gene. *Int. J. Biol. Macromol.* **34**, 295–301.
- 220 Beckwitt R, Arcidiacono S, Stote R (1998) Evolution of repetitive proteins: spider silks from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentarius* (Araneidae). *Insect Biochem. Mol. Biol.* **28**, 121–130.
- 221 Jeffares DC, Mourier T, Penny D (2006) The biology of intron gain and loss. *Trends Genet.* **22**, 16–22.
- 222 Roy SW (2006) Intron-rich ancestors. *Trends Genet.* **22**, 468–471.
- 223 Banyai L, Patthy L (2004) Evidence that human genes of modular proteins have retained significantly more ancestral introns than their fly or worm orthologues. *FEBS Lett.* **565**, 127–132.
- 224 Gregory TR, Shorthouse DP (2003) Genome sizes of spiders. *J. Hered.* **94**, 285–290.
- 225 Rasch EM, Connelly BA (2005) Genome size and endonuclear DNA replication in spiders. *J. Morphol.* **265**, 209–214.
- 226 Vinogradov AE (2004) Compactness of human housekeeping genes: selection for economy or genomic design? *Trends Genet.* **20**, 248–253.
- 227 Минкевич И, Патрушев Л (2007) Некодирующие последовательности генома и размер ядра эукариотической клетки как существенные факторы защиты генов от химических мутагенов. *Биоорганическая химия* **33**, 474–477.

- 228 Vassilevski A, Kozlov S, Grishin E V (2009) Molecular diversity of spider venom. *Biochem.* **74**, 1505–34.
- 229 Fry BG, Undheim EAB, Ali SA, Jackson TNW, Debono J, Scheib H, Ruder T, Morgenstern D, Cadwallader L, Whitehead D, Nabuurs R, van der Weerd L, Vidal N, Roelants K, Hendrikx I, Gonzalez SP, Koludarov I, Jones A, King GF, Antunes A, Sunagar K (2013) Squeezers and leaf-cutters: differential diversification and degeneration of the venom system in toxicoferan reptiles. *Mol. Cell. Proteomics* **12**, 1881–99.
- 230 Wu D-D, Zhang Y-P (2013) Evolution and function of de novo originated genes. *Mol. Phylogenet. Evol.* **67**, 541–5.
- 231 Sollod BL, Wilson D, Zhaxybayeva O, Gogarten JP, Drinkwater R, King GF (2005) Were arachnids the first to use combinatorial peptide libraries? *Peptides* **26**, 131–139.
- 232 Liao Z, Cao J, Li S, Yan X, Hu W, He Q, Chen J, Tang J, Xie J, Liang S (2007) Proteomic and peptidomic analysis of the venom from Chinese tarantula *Chilobrachys jingzhao*. *Proteomics* **7**, 1892–1907.
- 233 Patthy L (1999) Genome evolution and the evolution of exon-shuffling—a review. *Gene* **238**, 103–114.
- 234 Zhu K, Gao B, Zhu S (2012) Characterization of a chimeric antimicrobial peptide uncovers evolutionary significance of exon-shuffling. *Biochem. Biophys. Res. Commun.* **428**, 360–364.
- 235 Froy O, Gurevitz M (2003) Arthropod and mollusk defensins—evolution by exon-shuffling. *Trends Genet.* **19**, 684–687.
- 236 Akiva P, Toporik A, Edelheit S, Peretz Y, Diber A, Shemesh R, Novik A, Sorek R (2006) Transcription-mediated gene fusion in the human genome. *Genome Res.* **16**, 30–36.
- 237 Cancherini D V, Franca GS, de Souza SJ (2010) The role of exon shuffling in shaping protein-protein interaction networks. *BMC Genomics* **11 Suppl 5**, S11.
- 238 Kuhn-Nentwig L (2011) Venom composition and strategies in spiders: is everything possible? *Adv. In Insect Phys.* **40**, 1–86.