

КРАТКОЕ СООБЩЕНИЕ

УДК 577.322.4



Схемы замыкания дисульфидных связей в токсинах пауков семейства Ctenidae. Сравнение с предсказаниями нейросети AlphaFold 2.0

П.А. Миронов , З.О. Шенкарев* 

Институт биоорганической химии имени академиков М.М. Шемякина и Ю.А. Овчинникова, Российская академия наук,
Россия, г. Москва, 117997, ул. Миклухо-Маклая, д. 16/10

*e-mail: zakhar-shenkarev@yandex.ru

Пауки блуждающие, или пауки-бегуны (Ctenidae), имеют многокомпонентные яды, в которых идентифицировано более 500 различных пептидов и белков, названных ктенитоксинами. Основными компонентами яда являются цистеин-богатые пептиды, содержащие мотив ингибиторного цистинового узла (inhibitor cystine knot, ICK). Фармакологическое разнообразие ктенитоксинов позволяет рассматривать некоторые из них как прообразы для создания новых лекарств для лечения хронической боли, болезни Хантингтона, эректильной дисфункции и глаукомы. По расположению остатков цистеина в аминокислотной последовательности ктенитоксины разделяют на 14 групп, содержащих от шести до 14 остатков Cys. В настоящее время определена пространственная структура только одного ктенитоксина ω -CNTX-Pn4a ($\text{Ph}\alpha 1\beta$ или Tx3-6) бразильского странствующего паука *Phoneutria nigriventer*. Еще 10 структурных групп ктенитоксинов имеют гомологию с известными пространственными структурами токсинов пауков других семейств и других белков, а для трех групп структурные гомологи неизвестны. В данной работе предложены возможные схемы формирования дисульфидных связей для всех групп ктенитоксинов. Сравнение полученных схем с предсказаниями программы AlphaFold 2.0 показывает, что эта нейронная сеть не всегда корректно предсказывает структуры цистеин-богатых пептидов, особенно если моделируются структуры зрелых молекул без лидерных последовательностей.

Ключевые слова: нейротоксины пауков, ктенитоксины, ICK-мотив, цистиновый узел, предсказание пространственной структуры, цистеин-богатые пептиды, AlphaFold 2.0

DOI: 10.55959/MSU0137-0952-16-78-3S-3

Введение

Семейство Ктениды (Ctenidae, пауки блуждающие, или бегуны) включает 40 родов и более 400 видов крупных ночных пауков, обитающих в тропиках Старого и Нового Света. К этому семейству относятся пауки рода *Phoneutria*, которые считаются одними из самых ядовитых паукообразных на планете. Несмотря на то, что яд *Phoneutria* обладает высокой нейротоксичностью, его количество, вводимое при укусе, слишком мало, чтобы вызвать смертельный исход у взрослого человека. Укус паука вызывает сильную боль, головокружение, озноб, лихорадку, артериальную гипертензию, тахикардию, мышечные спазмы и приапизм.

Ктениды имеют многокомпонентные яды, в которых методами протеомного и транскриптомного анализа идентифицировано более 500 различных пептидов и белков с массой от 1,2 до 27 кДа, включая нейротоксины [1–4]. Основ-

ным компонентом яда являются цистеин-богатые пептиды ($\text{MW} < 15$ кДа, число Cys ≥ 6), названные ктенитоксинами (ctenitoxins или CNTXs), большинство которых содержит пространственный мотив ингибиторного цистинового узла (inhibitor cystine knot, ICK) [2, 4]. По количеству и расположению остатков Cys ктенитоксины можно разделить на 14 групп (рис. 1) [2–4]. Для ктенитоксинов описано действие на натриевые (NaV, обозначаются префиксами β -, δ - и μ -), калиевые (K-), кальциевые (CaV, ω -) каналы, ионотропные глутаматные рецепторы NMDA-субтипа (Γ -) и каналы TRP [5]. Однако мишени большинства ктенитоксинов до сих пор не установлены (префикс U-).

Яд бразильского странствующего паука *Phoneutria nigriventer* фармакологически чрезвычайно разнообразен [5]. В нем идентифицировано более 100 ктенитоксинов, и некоторые из них могут служить прообразом для создания новых препаратов против хронической боли, болезни Хантингтона,

эректильной дисфункции и глаукомы [4, 5]. Для дизайна новых препаратов на базе существующих токсинов необходимы знания о химической и пространственной структуре компонентов яда.

В настоящее время определена пространственная структура только одного ктенитоксина, ω-CNTX-Pn4a (Phα1β или Tx3-6) из *P. nigriventer* (структурная группа VII, или 12.1) [6]. Phα1β действует одновременно на кальциевые каналы CaV2

и на канал TRPA1 [7], демонстрируя значительный антиноцицептивный эффект в моделях острой и хронической боли. Ктенитоксины десяти структурных групп имеют гомологию с известными пространственными структурами, обнаруженными в других организмах. Три группы (VI, VIII, IX), предположительно, имеют в своей структуре мотивы ICK или DDH (disulfide-directed β-hairpin), однако структурные гомологи для них неизвестны (рис. 1).

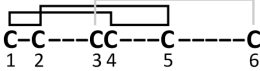
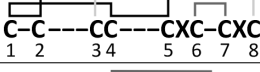
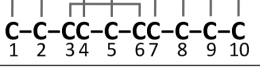
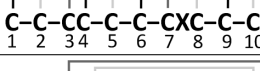
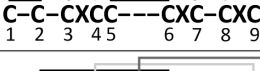
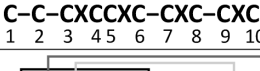
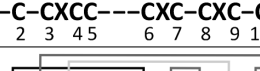
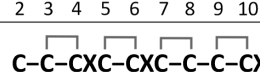
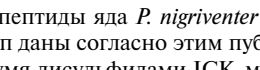
Группа [2,4] [3]	Cys	Схема дисульфидных связей	Фолд / структуры / гомологи в PDB	Актив- ность	S-S схема в UNIPROT/ предсказание AlphaFold 2 без л.п. / с л.п.	Кол-во пептидов [2,4] [3]
I 6.0	6		ICK / 0 / 47	Токс. U	+ / ± / ±	19 123
II 8.0	8		ICK / 0 / 11	Токс. ω, κ, U	+ / ± / +	31 538
III –	10		Элафиноподобный или WAP-домен / 0 / 7	?	+ / ± / ±	6 –
IV –	10		2×DDH, Прокинетицин или домен колипазы / 0 / 13	? не токс.	+ / + / +	10 –
V 10.0	10		ICK / 0 / 1	Токс. Γ, δ, U	+ / – / ±	18 117
– 10.1	10		ICK / 0 / 2	Токс. δ, ω, μ, U	+ / – / –	0 23
VI 12.0	12		ICK / 0 / 0	Токс. ω, U	– / + / +	2 100
VII 12.1	12		ICK / 1 / 0	Токс. ω, U	– / ± / +	5 27
VIII 14.0	14		ICK / 0 / 0	Токс. ω, μ, U	– / ± / ±	11 33
IX –	10		DDH (гр. VI без C1-C5) / 0 / 0	Токс. ω	– / + / +	1 –
X –	12		IGFBP-подобный домен или ингибитор грибковой протеазы-1 / 0 / 1	? Инг. протеаз	– / + / +	9 –
XI –	12		2× Домен тиреоглобулина типа-1 / 0 / 27	? Инг. протеаз	– / + / +	1 –
XII –	12/11		2× Домен тиреоглобулина типа-1 (с/без C1) / 0 / 27	? Инг. протеаз	– / + / +	4/2 –
XIII –	10		Богатый цистеином домен подобный ингибитору трипсина / 0 / 37	? Инг. протеаз	+ / + / +	9 –

Рис. 1. Цистеин-богатые пептиды яда *P. nigriventer* [2, 4] и пептиды содержащие ICK-мотив в ядах Stenidae и родственных семейств [3]. Названия групп даны согласно этим публикациям. На схемах черным цветом выделен цикл, образованный фрагментами основной цепи и двумя дисульфидными ICK-мотива. Светло-серая линия показывает третий дисульфид, проходящий через этот цикл. Остальные дисульфидные связи показаны темно-серым цветом. DDH – мотив disulfide-directed β-hairpin; ICK – мотив ингибиторного цистеинового узла; «структуры» – число известных структур пептидов Stenidae; «гомологи в PDB» – число известных структур гомологичных пептидов с таким расположением дисульфидных связей; «S-S схема в UNIPROT» – наличие правильной аннотации дисульфидных связей в базе данных; «предсказание AlphaFold 2» – наличие правильной схемы дисульфидных связей в структуре предсказанной без лидерной последовательности (л.п.) и с ней.

ICK-мотив обычно содержит β -шпильку, стабилизированную тремя дисульфидными связями со схемой замыкания C1-C4, C2-C5, C3-C6. При этом первые два дисульфида вместе с фрагментами основной цепи пептида, соединяющими цистеины, формируют замкнутый цикл, а третий дисульфид пронизывает этот цикл, образуя «узел» (показано черным и светло-серым цветом на рис. 1 и 2). В токсинах пауков соответствующая первичная структура представлена последовательностью «основного структурного мотива» C1-6-C2-U-C3C4-U-C5-U-C6, где U соответствуют вариабельным последовательностям, а петля C1-C2 чаще всего имеет длину шесть остатков [8]. Большинство ICK-токсинов пауков содержат более шести остатков Cys, и в них часто наблюдается «дополнительный структурный мотив» C5-1-C6-U-C7-1-C8, в котором дисульфидная связь Cys6-Cys7 стабилизирует петлю, расположенную между тяжами β -шпильки ICK-мотива (рис. 1 и 2). Также среди ктениитоксинов встречается мотив DDH, который может рассматриваться как вариант ICK-мотива без первой дисульфидной связи (схема: C2-C5, C3-C6).

Ввиду отсутствия прямых данных о структуре большинства ктениитоксинов большую роль приобретает моделирование по гомологии, с использованием известных пространственных структур, а также расчет структуры *de novo* при помощи таких программ, как Rosetta [9] или PEP-FOLD [10], моделирующих фолдинг белка *in silico* с использованием крупнозернистого представления белковой структуры и специальных силовых полей. Однако применение этих подходов для моделирования структур цистеин-богатых ктениитоксинов встречает ряд трудностей. Как отмечено выше, для многих пептидов этого семейства отсутствуют структурные гомологи, а применение алгоритмов «фолдинга» *in silico* требует априорного знания о системе замыкания дисульфидных связей, которая также во многих случаях неизвестна. В связи с этим особый интерес для предсказания структур цистеин-богатых пептидов вызывает применение глубинных нейронных сетей, опирающихся на известные знания о пространственных структурах, например, AlphaFold 2.0 [11] или RoseTTAFold [12]. Даже в случаях, когда эти алгоритмы не имеют информации о системе замыкания дисульфидных мостиков, они позволяют их уверенно моделировать, пользуясь знаниями о схожих мотивах в других белках. Интересный пример был показан для программы AlphaFold 2.0, которая хорошо предсказала структуру сайта связывания иона цинка в пептидазном домене без априорной информации о наличии такого сайта. При этом в полученной модели ориентация боковых цепей, ответственных за связывание иона, полностью соответствовала экспериментально определенной пространственной структуре комплекса белка с цинком [11].

В данном исследовании на основе экспериментальной структуры Phc1 β мы предлагаем возможные схемы замыкания дисульфидных связей для ктениитоксинов, содержащих ICK и DDH-мотивы и от 6 до 14 цистеинов, и сравниваем полученные схемы с аннотациями дисульфидных связей, приведенными в базе данных UniProt и предсказаниями программы AlphaFold 2.0.

Материалы и методы

Моделирование структур ктениитоксинов. Полные последовательности токсинов и последовательности зрелых пептидов (с лидерными последовательностями и без них) были взяты из базы данных UniProt (www.uniprot.org) и публикаций [2–4]. Расчеты AlphaFold 2.0 выполняли на web-сервере colab.research.google.com [13]. Из групп I–VIII и 10.1 ктениитоксинов было выбрано по 3–5 последовательностей для моделирования пространственной структуры. Из остальных групп (IX–XIII) было взято по одной последовательности для расчетов (таблица). Структуру каждого пептида моделировали как с лидерной последовательностью, так и без нее. В каждом запуске AlphaFold 2.0 было рассчитано по пять моделей. Для определения системы замыкания дисульфидных связей в программе MOLMOL [14] были рассчитаны попарные расстояния между всеми атомами серы в структуре. Дисульфидная связь считалась сформированной, если расстояние между соответствующими атомами было менее 3Å.

Анализ гомологичных структур. Поиск гомологичных белков осуществляли в UniProt, используя данные там аннотации или поиск BLAST. Для определения пептидов, образующих группы, брали данные из дополнительных материалов к статьям [2–4]. Гомологичные пространственные структуры были взяты из базы данных InterPro (www.ebi.ac.uk/interpro/), содержащей классификацию различных белковых семейств и имеющую кросс-ссылки с UNIPROT.

Результаты и обсуждение

Среди известных 14 групп ктениитоксинов, классифицированных по числу и расположению цистеинов в аминокислотной последовательности, можно выделить 8 групп, содержащих ICK или DDH в качестве основных структурных мотивов (рис. 1). Четное число цистеинов (6, 8, 10, 12 или 14) предполагает замыкание всех остатков Cys в дисульфидные связи. Следует отметить, что существует два способа для наименования групп ктениитоксинов. В подходе, предложенном в публикациях [2, 4], группы цистеин-богатых пептидов с MW < 15 кДа из яда *P. nigriventer* нумеруются римскими цифрами. В публикации [3] группы пептидов из ядов пауков семейства Stenidae и родственных ему, содержащих ICK-мотив, нумеруются арабскими цифрами.

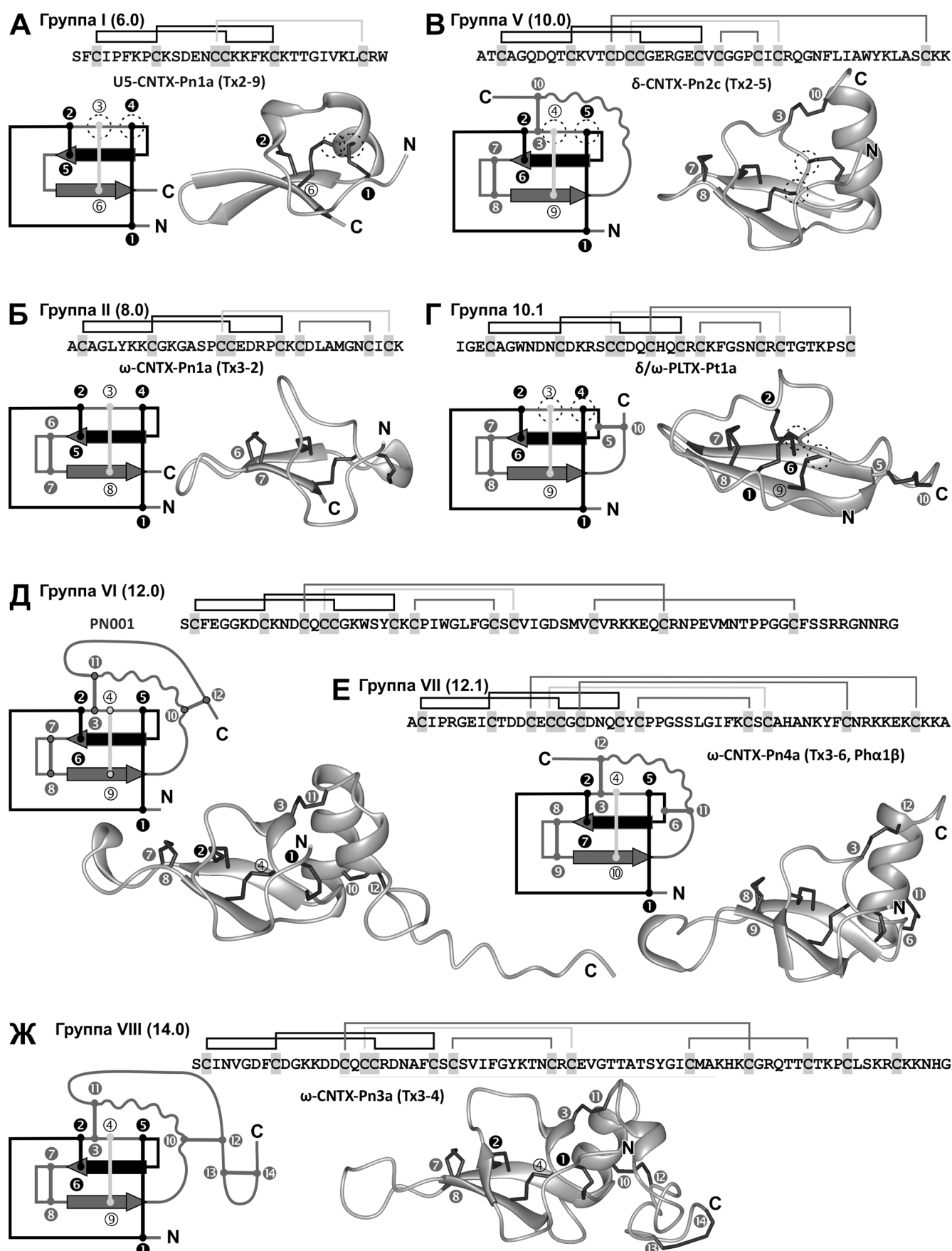


Рис. 2. Схема формирования дисульфидных связей и пространственная структура токсинов из яда *P. nigriventer* и других пауков, содержащих ИСК-мотив. Группы токсинов и схемы дисульфидных связей показаны как на рис. 1. Все приведенные пространственные структуры, кроме Е, предсказаны программой AlphaFold 2.0. Ошибки в предсказании схемы замыкания дисульфидных связей выделены кружками. На панели Е показана экспериментальная структура токсина Ph α 1 β [6].

Таблица

Цистеин-богатые пептиды яда *P. nigriventer* и родственных семейств и результаты расчета их структуры в программе AlphaFold 2.0

Группа	Источник последовательности	BLAST/UniProt Название гомолога	UniProt код гомолога, идентичность	AlphaFold Схема S-S токсин	AlphaFold Схема S-S л.п. + токсин
I 6.0	UP	U5-CNTX-Pn1a	P29426	4+/1–	н.д.
	PN266 [2]	U5-CNTX-Pn1a	P29426 97%	+	5–
	UP	U13-CNTX-Pn1a	P83894	+	н.д.
	PN046 [2]	U13-CNTX-Pn1b	P84017 97%	+	+
	PN086 [2]	U14-CNTX-Pn1a	P83998 100%	+	+
II 8.0	UP	ω-CNTX-Pn1a	O76201	+	+
	UP	α-CNTX-Pn1a	O76200	+	+
	UP	U1-CNTX-Pk1a	P83895	4+/1–	н.д.
	UP	U9-CNTX-Pn1a	P0C2S6	5–	+
III	PN365 [4]	U15-LCTX-Ls1g	B6DD50 47%	5–	1+/4–
	PN346 [2]	U20-LCTX-Ls1d	B6DCY1 44%	+	+
	PN306 [2]	U20-LCTX-Ls1c	B6DD62 40%	+	+
IV	PN123 [2]	U19-CNTX-Pn1a	P83997 93%	+	+
	UP	U19-CNTX-Pn1a	P83997	+	н.д.
	UP	U9-CNTX-Pr1a	P83893	+	н.д.
	PN350 [4]	U3-AATX-Ce1a	Q8MTX1 39%	+	+
V 10.0	UP	U1-CNTX-Pn1a	P61229	5–	+
	UP	γ-CNTX-Pn1a	P59367	3+/2–	+
	UP	δ-CNTX-Pn2c	O76199	4+/1–	+
	UP	δ-CNTX-Pn1a	P59368	2+/3–	4+/1–
– 10.1	#85 DN9510 [3]	δ/ω-PLTX-Pt1a	P0DL38 38%	5–	5–
	#740 DN4598 [3]	U1-PLTX-Pt1d	P36986 55%	5–	5–
	#1157 DN1414 [3]	δ/ω-PLTX-Pt1a	P0DL38 35%	5–	5–
VI 12.0	PN001 [2]	U9-AGTX-Ao1a	Q5Y4U3 64%	+	+
	PN300 [2]	U20-CNTX-Pn1a	P84093 43%	+	+
	UP	U9-AGTX-Ao1a	Q5Y4U3	+	+
VII 12.1	UP	ω-CNTX-Pn4a (Phα1β)	P81792	+	+
	UP	U12-CNTX-Pn1a	P0C2S8	+	+
	UP	U11-CNTX-Pn1a	P0C2S7	+	+
	PN313 [2]	U12-CNTX-Pn1a	P0C2S8 76%	4+/1–	+
VIII 14.0	UP	μ-CNTX-Pn1a	P17727	3+/2–	4+/1–
	UP	ω-CNTX-Pn3a	P81790	4+/1–	3+/2–
	PN319 [2]	ω-CNTX-Pn3a	P81790 96%	4+/1–	3+/2–
IX	PN327 [2]	ω-AGTX-1A	P15969 68%	+	+
X	PN376 [4]	CP8	Q6TUD0 36%	+	+
XI	PN372 [4]	U24-CNTX-Pn1a	P84032 98%	+	+
XII	PN373 [2]	U24-CNTX-Pn1a	P84032 36%	+	н.д.
XIII	PN370 [2]	Venom peptide SjAPI-2	P0DM56 49%	+	+

Примечание: Последовательности пептидов, включая лидерные последовательности (л.п.), были взяты из базы данных UniProt (UP) и из дополнительных материалов к работам Диниз и др. [2], Брюэр и др. [3], Кардосо и др. [4], которые доступны на сайтах изданий. Пептиды, цитируемые по работам Диниз и др. [2] и Кардосо и др. [4], названы как в оригинальных публикациях PNXXX. Для пептидов, цитируемых по работе Брюэр и др. [3], даны ссылки на номера строк в файле с дополнительной информацией (#XXX). Для каждого пептида, не представленного в базе данных UniProt, проведен поиск BLAST для выявления гомологичных пептидов из базы данных. Приведены название ближайшего гомолога, код доступа UniProt и степень идентичности последовательностей зрелых пептидов. В каждом запуске AlphaFold 2.0 было рассчитано пять моделей. Знаком «+» показано число моделей, содержащих правильную схему замыкания дисульфидных связей, знаком «–» – неправильные схемы. «н.д.» – в базе данных или публикации отсутствует информация о лидерной последовательности.

Группа I, или 6.0, включает ктенитоксины с шестью цистеинами, которым соответствует стандартный мотив ICK (рис. 2А). В эту группу входят пептиды, действующие на NaV- и CaV-каналы, например, U5-CNTX-Pn1a (Tx2-9) и ω -CNTX-Pn2a (Tx3-3). Несмотря на сравнительно хорошее описание гомологичных токсинов методами структурной биологии, программа AlphaFold не всегда точно предсказывает их структуру. Так, при моделировании токсина U5-CNTX-Pn1a без лидерной последовательности в одной из пяти моделей получается структура с неправильной схемой дисульфидных связей (C1-C3, C2-C5, C4-C6), не соответствующая ICK-мотиву (рис. 2А), а при моделировании изоформы этого пептида, содержащей замену Phe19Leu (PN266 [2]) с лидерной последовательностью, все пять полученных моделей не соответствуют правильной структуре (таблица).

В группу II, или 8.0, включающую ктенитоксины с восьмью цистеинами, входят пептиды, действующие на каналы Ca_v1, например, ω -CNTX-Pn1a (Tx3-2). Кроме того, к этой группе токсинов относится ω -agatoxin IVA, действующий на каналы Ca_v2.1. В этих токсинах четвертый дополнительный дисульфид стабилизирует удлиненную петлю, выходящую из антипараллельного β -листа ICK-мотива (рис. 2Б). Программа AlphaFold корректно предсказывает структуру пептидов из этой группы при моделировании молекул с лидерными последовательностями. Моделирование зрелых пептидов в некоторых случаях приводит к ошибкам (таблица).

Группа V, или 10.0, включает ктенитоксины с 10 цистеинами, имеющие характерную последовательность C³-1-C⁴-C⁵ (рис. 1). Эти токсины действуют на Na_v-каналы млекопитающих (например, δ -CNTX-Pn2c, или Tx2-5), насекомых, а также на NMDA-рецепторы (например, Γ -CNTX-Pn1a, или Tx4(5-5)). У этой группы токсинов есть структурный гомолог spiderine-1a [15]. В сравнении с предыдущими группами токсины из группы V имеют удлиненный С-конец, который соединяется с остальной частью пептида пятой дисульфидной связью C³-C¹⁰. Согласно нашим данным, в отсутствие лидерных последовательностей программа AlphaFold некорректно предсказывает структуру этих пептидов (рис. 2В, таблица). Этот факт проиллюстрирован в публикации [4], где были получены модели циклических пептидов, содержащих некорректную дисульфидную связь C¹-C¹⁰.

Группа 10.1 включает пептиды с 10 цистеинами, имеющие дополнительный остаток Cys⁵ между основным и дополнительным структурными мотивами (рис. 1). Такие токсины не встречаются в яде *P. nigriventer*, однако они представлены в ядах других пауков семейства Stenidae и родственных им [3]. Например, к этой группе относятся токсины δ/ω -PLTX-Pt1a, μ -HXTX-Mg2a (Magi3) и Tbo-IT2. Известные пространственные

структуры токсинов Magi3 [16] и Tbo-IT2 [17] показывают, что в токсинах из группы 10.1 пятая дисульфидная связь C⁵-C¹⁰ соединяет С-конец молекулы с ICK-мотивом. Согласно нашим данным, как в отсутствие, так и с лидерными последовательностями программа AlphaFold во многих случаях некорректно предсказывает структуру пептидов этой группы (рис. 2Г, таблица).

Примеры групп V (10.0) и 10.1 показывают, что по мере усложнения строения токсинов происходит удлинение С-концевой части пептида, для стабилизации которой возникают дополнительные дисульфидные связи. В соответствии с этим, группы ктенитоксинов VI (12.0), VII (12.1) и VIII (14.0) с 12 и 14 остатками Cys содержат дополнительные дисульфидные связи либо прикрепляющие С-концевую часть пептидов к ICK мотиву, либо стабилизирующие эту С-концевую часть. Сравнение последовательностей этих токсинов с последовательностями токсинов с известными пространственными структурами (spiderine-1a, Magi3, Tbo-IT2 и Ph α 1 β) позволяет предложить топологию дисульфидных связей первого типа (ICK—С-конец) и по остаточному принципу замкнуть связи второго типа (С-конец—С-конец) (рис. 1 и 2Д, Е, Ж). Следует отметить, что в UniProt приведена неправильная аннотация дисульфидных связей для многих из этих пептидов, но в то же время их структура во многих случаях корректно предсказывается программой AlphaFold, особенно если рассматривается полная последовательность пептида. Исключение составляет группа VIII (14.0), в которой при моделировании трех пептидов в каждом случае были получены модели как с правильной, так и с неправильной системой замыкания дисульфидных связей (таблица).

К группе VI (12.0) относятся высокотоксичные молекулы, блокирующие Ca_v1- и Ca_v2-каналы, например, ω -AGTX-Aa3a (ω -agatoxin IIIA) и U20-CNTX-Pn1a (ω -phonetoxin PNTx22C5). К группе VII (12.1) относятся токсины, имеющие низкую токсичность и действующие на каналы Ca_v2, такие как Ph α 1 β [1]. Для ктенитоксинов группы VIII (14.0) описано антагонистическое действие на каналы Ca_v2 (ω -CNTX-Pn3a, Tx3-4) и Na_v1.2 (μ -CNTX-Pn1a, Tx1).

Группа IX содержит единственный ктениотоксин PN327 [2], гомологичный ω -agatoxin-1A. Предшественник зрелого токсина, по видимому, содержит DDH-мотив, схожий с токсинами из группы VI (12.0), но без первой дисульфидной связи Cys¹-Cys⁵ (рис. 1). Впоследствии эта цепь разрезается на два фрагмента, соединенных дисульфидной связью Cys¹⁰-Cys¹² — основной, включающий девять остатков Cys, и минорный трипептид Ser-Pro-Cys [18].

Остальные шесть групп ктенитоксинов (III, IV, X, XI, XII и XIII) имеют пространственные структуры, отличные от ICK- и DDH-моти-

вов, и, по-видимому, содержат нетоксичные молекулы, например, ингибиторы протеаз. Несмотря на то, что для многих из этих пептидов в UniProt отсутствует или приведена неправильная аннотация дисульфидных связей, программа AlphaFold удовлетворительно предсказывает их структуры. Единственным исключением являются пептиды из группы III, имеющие в своей структуре Elafin-like или WAP (whey acidic protein) домен, некоторые из которых не сворачиваются корректно как в отсутствие, так и с лидерными последовательностями.

В результате мы предложили характерные дисульфидные паттерны для групп ктениотоксинов, содержащих ИСК- и DDH-мотивы, используя

данные о пространственной структуре гомологичных токсинов других организмов. Полученные результаты показали, что программа по сворачиванию полипептидных цепей на основе нейронной сети AlphaFold 2.0 не всегда корректно предсказывает структуры цистеин-богатых пептидов, особенно если моделируются структуры зрелых молекул без лидерных последовательностей.

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 22-14-00326). Исследования проведены без использования животных и без привлечения людей в качестве испытуемых. Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Richardson M., Pimenta A.M.C., Bemquerer M.P., Santoro M.M., Beirao P.S.L., Lima M.E., Figueiredo S.G., Bloch C. Jr., Vasconcelos E.A.R., Campos F.A.P., Gomes P.C., Cordeiro M.N. Comparison of the partial proteomes of the venoms of Brazilian spiders of the genus *Phoneutria*. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* 2006;142(3–4):173–187.
2. Diniz M.R.V., Paiva A.L.B., Guerra-Duarte C., Nishiyama M.Y., Mudadu M.A., Oliveira U., Borges M.H., Yates J.R., Junqueira-de-Azevedo I.L. An overview of *Phoneutria nigriventer* spider venom using combined transcriptomic and proteomic approaches. *PLoS One*. 2018;13(8):e0200628.
3. Brewer M.S., Cole T.J. Killer knots: Molecular evolution of inhibitor cystine knot toxins in Wandering spiders (Araneae: Ctenidae). *Toxins*. 2023;15(2):112.
4. Cardoso F.C., Walker A.A., King G.F., Gomez M.V. Holistic profiling of the venom from the Brazilian wandering spider *Phoneutria nigriventer* by combining high-throughput ion channel screens with venomomics. *Front. Mol. Biosci.* 2023;10:1069764.
5. Peigneur S., Lima M.E., Tytgat J. *Phoneutria nigriventer* venom: A pharmacological treasure. *Toxicon*. 2018;151:96–110.
6. Lyukmanova E.N., Mironov P.A., Kulbatskii D.S., Shulepko M.A., Paramonov A.S., Chernaya E.M., Logashina Y.A., Andreev Y.A., Kirpichnikov M.P., Shenkarev Z.O. Recombinant production, NMR solution structure, and membrane interaction of the Ph α 1 β toxin, a TRPA1 modulator from the Brazilian armed spider *Phoneutria nigriventer*. *Toxins*. 2023;15(6):378.
7. Tonello R., Fusi C., Materazzi S., Marone I.M., Logu F., Benemei S., Goncalves M.C., Coppi E., Castro-Junior C.J., Gomez M.V., Geppetti P., Ferreira J., Nassini R. The peptide Ph α 1 β , from spider venom, acts as a TRPA1 channel antagonist with antinociceptive effects in mice. *Br. J. Pharmacol.* 2017;174(1):57–69.
8. Kozlov S., Grishin E. Classification of spider neurotoxins using structural motifs by primary structure features. Single residue distribution analysis and pattern analysis techniques. *Toxicon*. 2005;46(6):672–686.
9. Leaver-Fay A., Tyka M., Lewis S.M., et al. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods in Enzymology. Computer Methods, Part C. Vol. 487*. Eds. M.L. Johnson and L. Brand. Academic Press; 2011:545–574.
10. Shen Y., Maupetit J., Derreumaux P., Tufféry P. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. *J. Chem. Theor. Comput.* 2014;10(10):4745–4758.
11. Jumper, J., Evans, R., Pritzel, A., et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–589.
12. Baek M., DiMaio F., Anishchenko I., et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;373(6557):871–876.
13. Mirdita M., Schütze K., Moriaki Y., Heo L., Ovchinnikov S., Steinegger M. ColabFold: making protein folding accessible to all. *Nat. Methods*. 2022;19(6):679–682.
14. Koradi R., Billeter M., Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 1996;14(1):51–55.
15. Nadezhdin K.D., Romanovskaia D.D., Sachkova M.Y., Oparin P.B., Kovalchuk S.I., Grishin E.V., Arseniev A.S., Vassilevski A.A. Modular toxin from the lynx spider *Oxyopes takobius*: Structure of spiderine domains in solution and membrane-mimicking environment. *Protein Sci.* 2017;26(3):611–616.
16. Titau-Delgado G., Carrillo E., Mendoza A., Mayorga-Flores M., Escobedo-González F.C., Cano-Sánchez P., López-Vera E., Corzo G., Del Rio-Portilla F. Successful refolding and NMR structure of rMagi3: A disulfide-rich insecticidal spider toxin. *Protein Sci.* 2018;27(3):692–701.
17. Korolkova Y., Maleeva E., Mikov A., Lobas A., Solovyeva E., Gorshkov M., Andreev Y., Peigneur S., Tytgat J., Kornilov F., Lushpa V., Mineev K., Kozlov S. New insectotoxin from *Tibellus oblongus* spider venom presents novel adaptation of ICK fold. *Toxins*. 2021;13(1):29.
18. Santos A.D., Imperial J.S., Chaudhary T., Beavis R.C., Chait B.T., Hunsperger J.P., Olivera B.M., Adams M.E., Hillyard D.R. Heterodimeric structure of the spider toxin omega-agatoxin IA revealed by precursor analysis and mass spectrometry. *J. Biol. Chem.* 1992;267(29):20701–20705.

Поступила в редакцию 29.08.2023

После доработки 25.09.2023

Принята в печать 26.09.2023

SHORT COMMUNICATION

Disulfide bond patterns in the toxins of spiders of the Ctenidae family. Comparison with AlphaFold 2.0 predictions

P.A. Mironov , Z.O. Shenkarev* 

*Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
Miklukho-Maklaya 16/10, 117997, Moscow, Russia*

**e-mail: zakhar-shenkarev@yandex.ru*

Wandering spiders (family Ctenidae) have multicomponent venoms in which more than 500 different peptides and proteins, called ctenitoxins, have been identified. The main components of the venom are cysteine-rich peptides containing an inhibitory cystine knot (ICK) motif. The pharmacological diversity of ctenitoxins allows us to consider some of them as prototypes for the development of new drugs for the treatment of chronic pain, Huntington's disease, erectile dysfunction and glaucoma. According to the location of cysteine residues in the amino acid sequence, ctenitoxins are divided into 14 groups containing from 6 to 14 Cys residues. Currently, the spatial structure of only one ctenitoxin, ω -CNTX-Pn4a (Pha1 β or Tx3-6) from the Brazilian wandering spider *Phoneutria nigriventer*, has been determined. Another 10 structural groups of ctenitoxins have homology with the known spatial structures of spider toxins of other families and other proteins, and for three groups the structural homologues are unknown. In this paper, we proposed possible disulfide bonding patterns for all groups of ctenitoxins. A comparison of the obtained schemes with the predictions of the AlphaFold 2.0 program shows that this neural network does not always correctly predict the structures of cysteine-rich peptides, especially if the structures of mature molecules without leader sequences are modeled.

Keywords: *spider neurotoxins, ctenitoxins, ICK motif, cystine knot, spatial structure prediction, cysteine-rich peptides, AlphaFold 2.0*

Funding: The research was funded by Russian Science Foundation, project number 22-14-00326.

Сведения об авторах

Миронов Павел Андреевич — тех.-лаб. лаборатории структурной биологии ионных каналов Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН. Тел: 8-495-335-01-00; e-mail: mironov@nmr.ru; ORCID: <https://orcid.org/0000-0002-7884-3257>

Шенкарев Захар Олегович — докт. физ.-мат. наук, гл. науч. сотр. лаборатории структурной биологии ионных каналов Институт биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН. Тел: 8-495-335-01-00; e-mail: zakhar-shenkarev@yandex.ru; ORCID: <https://orcid.org/0000-0003-1383-3522>