



УДК 595.443.7-114.52:577.112.5

**ПЕРВИЧНАЯ СТРУКТУРА δ -ЛАТРОИНСЕКТОТОКСИНА
ИЗ ЯДА ПАУКА *Latrodectus mactans tredecimguttatus***© 1996 г. И. Е. Дулубова, М. В. Хвощев*, В. Г. Красноперов, Т. Г. Галкина,
К. А. Плужников, Т. М. Волкова, Е. В. ГришинИнститут биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,
117871, Москва, ГСП-7, ул. Миклухо-Маклая, 16/110

Поступило в редакцию 28.09.95 г.

Осуществлено клонирование структурного гена δ -ЛИТ. Определена нуклеотидная последовательность гена, содержащего открытую рамку считывания длиной 3642 п. о. Выведенная аминокислотная последовательность δ -ЛИТ гомологична последовательностям ранее изученных латротоксинов.

Ключевые слова: нейротоксин, *Latrodectus*, клонирование кДНК.

Яд каракурта содержит семейство высокомолекулярных нейротоксинов, вызывающих массивный выброс нейромедиаторов из пресинаптических окончаний различных животных [1]. Лучше всего изучен α -ЛИТ, специфически действующий только на нервные окончания позвоночных животных [2]. Первичная структура этого токсина, а также α -ЛИТ, действующего на нервные окончания насекомых, была установлена ранее [3, 4]. Информация о строении латротоксинов является важной предпосылкой для детального изучения молекулярных основ механизма действия и специфичности нейротоксинов каракурта. Цель данной работы – определение первичной структуры δ -ЛИТ.

δ -ЛИТ выделяли из цельного яда по методу [5]. N-Концевую последовательность и структуру триптических пептидов определяли с помощью автоматического секвенирования по методу Эдмана.

кДНК получали на основе мРНК, выделенной из ядовитых желез каракурта [6]. Клонирование кДНК осуществляли по тупым концам в плазмидный вектор pSP65 [7]. Скрининг полученной библиотеки кДНК проводили с помощью $5'$ - 32 P-меченого олигонуклеотидного зонда, созданного на основе первых 8 а. о. N-концевой последовательности. Частично вырожденный 23-членный олигонукле-

отид (синтезирован канд. хим. наук Н.С. Быстрым) имел следующее строение:



В результате скрининга были обнаружены 8 клонов, дающих положительный сигнал гибридизации. Выделенные в индивидуальном состоянии клоны анализировали с помощью частичного секвенирования и рестрикционного картирования. На основе полученных данных был сделан вывод о том, что обнаруженные клоны перекрывают фрагмент структурного гена, кодирующий аминокислотную последовательность длиной около 800 а. о. Для получения недостающей информации был проведен повторный скрининг библиотеки кДНК. В качестве специфического зона для гибридизации использовали фрагменты ДНК клонов. В результате скрининга был обнаружен дополнительный клон, содержащий искомую область структурного гена δ -ЛИТ.

На основе ДНК каждого клона получали набор перекрывающихся фрагментов ДНК, содержащих направленные делеции [8]. В результате секвенирования [9] нами была определена нуклеотидная последовательность длиной 4536 п. о., содержащая открытую рамку считывания длиной 3642 п. о. (рисунок). В ее составе обнаружена короткая $5'$ -нетранслируемая область длиной 44 п. о. и фрагмент $3'$ -нетранслируемой области (более 1000 п. о.), который не содержит сигнал полиаденилирования, что согласуется с имеющимися данными об устройстве мРНК, кодирующих латротоксины [3, 4]. Рамка считывания, начинающаяся с кодона АТГ и заканчивающаяся стоп-кодоном

Сокращения: α -ЛИТ – α -латротоксин, α -ЛИТ – α -латроинсектотоксин, δ -ЛИТ – δ -латроинсектотоксин, а. о. – аминокислотный остаток.

* Автор для переписки.

													M	H	-28
g t c a a t t g a	a a c t t t a t g a	t a g g a t t c a c	t t t c t t a t a t	a g a a	a t g	c a t							50		
S	K	E	L	Q	T	I	S	A	A	V	A	R	K	-13	
t c c	a a a	g a a	t t a	c a a	a c t	a t t	t c a	g c a	g c g	g t a	g c a	c g a	a a a	92	
A	V	P	N	T	M	V	I	R	L	K	R	<u>D</u>	<u>E</u>	2	
g c a	g t a	c c c	a a t	a c t	a t g	g t t	a t t	c g g	t t g	a a a	a g a	g a t	g a a	134	
E	D	G	E	M	T	L	E	E	R	Q	A	Q	C	16	
g a a	g a t	g g a	g a a	a t g	a c t	c t a	g a a	g a a	a g a	c a a	g c a	c a a	t g c	176	
K	A	I	E	Y	S	N	S	V	F	G	M	I	A	30	
a a a	g c a	a t a	g a g	t a c	a g c	a a t	t c a	g t t	t t t	g g g	a t g	a t c	g c t	218	
D	V	A	N	D	I	G	S	<u>I</u>	<u>P</u>	<u>V</u>	<u>I</u>	<u>G</u>	<u>E</u>	44	
g a t	g t a	g c t	a a c	g a c	a t c	g g t	t c c	a t t	c c t	g t a	a t t	g g c	g a a	260	
V	V	G	I	V	T	A	P	I	A	I	V	S	<u>H</u>	58	
g t a	g t t	g g c	a t t	g t a	a c t	g c c	c c a	a t t	g c c	a t c	g t a	a g t	c a c	302	
<u>I</u>	<u>T</u>	<u>S</u>	<u>A</u>	<u>G</u>	<u>L</u>	<u>D</u>	<u>I</u>	<u>A</u>	S	T	A	L	D	72	
a t t	a c t	a g c	g c a	g g c	t t g	g a t	a t a	g c t	t c t	a c g	g c a	t t a	g a t	344	
C	D	D	I	P	F	D	E	I	K	E	I	L	E	86	
t g t	g a t	g a t	a t a	c c t	t t t	g a t	g a g	a t t	a a g	g a a	a t a	t t a	g a a	386	
E	R	F	N	E	I	D	R	K	L	D	K	N	T	100	
g a a	a g a	t t c	a a t	g a a	a t a	g a t	a g a	a a g	t t g	g a c	a a g	a a c	a c a	428	
A	A	L	E	E	V	S	K	L	V	S	K	T	F	114	
g c t	g c t	t t g	g a a	g a g	g t c	t c t	a a a	c t g	g t a	a g t	a a a	a c t	t t t	470	
V	T	V	E	K	T	R	N	E	M	N	E	N	F	128	
g t t	a c g	g t g	g a a	a a a	a c a	a g g	a a t	g a a	a t g	a a c	g a a	a a t	t t t	512	
K	L	V	L	E	T	I	E	S	K	E	I	K	S	142	
a a g	c t t	g t t	t t g	g a a	a c t	a t a	g a a	a g c	a a a	g a a	a t a	a a a	t c a	554	
I	V	F	K	I	N	D	F	K	K	F	F	E	K	156	
a t t	g t a	t t c	a a a	a t a	a a t	g a t	t t t	a a a	a a g	t t t	t t t	g a a	a a a	596	
E	R	Q	R	I	K	G	L	P	K	D	R	Y	V	170	
g a a	c g a	c a a	a g a	a t t	a a a	g g t	t t g	c c t	a a a	g a t	a g g	t a t	g t t	638	
A	K	L	L	E	Q	K	G	I	L	G	S	L	K	184	
g c t	a a g	c t t	c t a	g a a	c a a	a a a	g g t	a t t	t t a	g g t	t c t	t t a	a a a	680	
E	V	R	E	P	S	G	N	S	L	S	S	A	L	198	
g a a	g t a	a g a	g a a	c c a	t c t	g g a	a a c	a g t	c t g	a g c	t c c	g c g	t t a	722	
N	E	L	L	D	K	N	N	N	Y	A	I	P	K	212	
a a t	g a a	c t c	t t a	g a c	a a a	a a c	a a c	a a c	t a t	g c c	a t c	c c a	a a a	764	
V	V	D	D	N	K	A	F	Q	<u>A</u>	<u>L</u>	<u>Y</u>	<u>A</u>	<u>L</u>	226	
g t g	g t t	g a t	g a t	a a t	a a g	g c c	t t t	c a g	g c g	c t g	t a t	g c t	t t a	806	
<u>F</u>	<u>Y</u>	<u>G</u>	<u>T</u>	<u>Q</u>	<u>T</u>	<u>Y</u>	<u>A</u>	<u>A</u>	<u>V</u>	<u>M</u>	<u>F</u>	<u>F</u>	<u>L</u>	240	
t t t	t a t	g g a	a c t	c a g	a c t	t a t	g c a	g c c	g t t	a t g	t t t	t t c	t t a	848	
<u>L</u>	E	Q	H	S	Y	L	A	D	Y	Y	Y	Q	K	254	
c t c	g a a	c a a	c a t	t c t	t a t	c t g	g c t	g a t	t a t	t a t	t a c	c a a	a a a	890	
G	D	D	V	N	F	N	A	E	F	N	N	V	A	268	
g g t	g a t	g a t	g t a	a a t	t t t	a a t	g c a	g a a	t t t	a a t	a a t	g t a	g c a	932	
I	I	F	D	D	F	K	S	S	L	T	G	G	D	282	
a t t	a t t	t t t	g a t	g a c	t t t	a a a	t c a	t c a	c t a	a c a	g g a	g g a	g a t	974	
D	G	L	I	D	N	V	I	E	V	L	N	T	V	296	
g a c	g g a	t t a	a t a	g a t	a a t	g t c	a t t	g a g	g t t	c t t	a a c	a c c	g t g	1016	
K	A	L	P	F	I	K	N	A	D	S	K	L	Y	310	
a a a	g c a	t t a	c c a	t t t	a t a	a a g	a a c	g c c	g a c	a g t	a a a	c t a	t a c	1058	
R	E	L	V	T	R	T	K	A	L	E	T	L	K	324	
a g a	g a a	t t a	g t a	a c t	a g a	a c a	a a a	g c t	t t a	g a g	a c t	c t t	a a a	1100	

Нуклеотидная последовательность кодирующей цепи кДНК и аминокислотная последовательность δ -латроинсекто-токсина. Подчеркнуты аминокислотные последовательности, структура которых установлена методами белковой хи-мии. Обведены в рамку гидрофобные последовательности, предположительно образующие трансмембранные сег-менты. Домен с анкириновыми повторами выделен жирным шрифтом. Первая аминокислота каждого анкиринового повтора обведена кругом. Двойной чертой обозначена последовательность PEST-мотива.

N	Q	I	K	T	T	D	L	P	L	I	D	D	I	338
aat	caa	atc	aaa	acg	act	gat	ttg	cct	ctt	ata	gat	gat	ata	1142
P	E	T	L	S	Q	V	N	F	P	N	D	E	N	352
ccc	gaa	act	ttg	tct	caa	gtg	aac	ttt	ccg	aat	gac	gaa	aat	1184
Q	L	P	T	P	I	G	N	W	V	D	G	V	E	366
caa	ttg	cct	aca	cca	ata	gga	aat	tgg	gtt	gat	ggc	gta	gaa	1226
V	R	Y	A	V	Q	Y	E	S	K	G	M	Y	S	380
gtt	agg	tac	gca	gta	cag	tat	gaa	agt	aag	ggc	atg	tat	tcg	1268
K	F	S	E	W	S	E	P	F	T	V	Q	G	N	394
aaa	ttc	agt	gaa	tgg	tct	gaa	cca	ttt	act	gtc	caa	ggg	aac	1310
A	C	P	T	I	K	V	R	V	D	P	K	K	R	408
gct	tgt	ccg	act	ata	aaa	gtt	cgt	gtt	gat	ccg	aaa	aag	aga	1352
N	R	L	I	F	R	K	F	N	S	G	K	P	Q	422
aat	aga	cft	atc	ttt	agg	aag	ttc	aac	tca	gga	aaa	cct	cag	1394
F	A	G	T	M	T	H	S	Q	T	N	F	K	D	436
ttt	gct	gga	acc	atg	act	cat	tca	caa	aca	aat	ttt	aaa	gat	1436
I	H	R	D	L	Y	D	A	A	L	N	I	N	K	450
att	cat	cgt	gat	cta	tac	gat	gca	gcc	tta	aat	att	aat	aag	1478
L	K	A	V	D	E	A	T	T	L	I	E	K	G	464
ttg	aaa	gca	gtg	gat	gaa	gct	aca	act	ttg	att	gaa	aag	ggg	1520
A	D	I	E	A	K	F	D	(N)	D	R	S	A	M	478
gca	gac	ata	gaa	gca	aaa	ttt	gac	aat	gac	aga	agt	gca	atg	1562
H	A	V	A	Y	R	G	N	N	K	I	A	L	R	492
cac	gca	gtt	gca	tat	cga	gga	aat	aac	aaa	ata	gcc	tta	aga	1604
F	L	L	K	N	Q	S	I	D	I	E	L	K	D	506
ttt	ctt	ttg	aaa	aat	caa	tcc	att	gac	atc	gag	tta	aaa	gat	1646
K	(N)	G	F	T	P	L	H	I	A	A	E	A	G	520
aaa	aac	ggc	ttt	act	cct	cta	cac	atc	gca	gct	gaa	gca	ggg	1688
Q	A	G	F	V	K	L	L	I	N	H	G	A	D	534
cag	gca	gga	ttt	gtt	aag	tta	cta	ata	aat	cat	gga	gct	gat	1730
V	N	A	K	T	S	K	(T)	N	L	T	P	L	H	548
gtg	aat	gca	aaa	aca	agt	aag	aca	aat	ttg	aca	cca	tta	cat	1772
L	A	T	R	S	G	F	S	K	T	V	R	N	L	562
ctt	gca	aca	cgt	agt	gga	ttt	tca	aaa	act	gta	aga	aat	tta	1814
L	E	S	P	N	I	K	V	N	E	K	E	D	(D)	576
cta	gaa	agc	cca	aat	aat	aag	gta	aat	gaa	aag	gag	gat	gac	1856
G	F	T	P	L	H	T	A	V	M	S	T	Y	M	590
gga	ttt	aca	cct	ttg	cat	act	gca	gta	atg	agt	act	tat	atg	1898
V	V	D	A	L	L	N	H	P	D	I	D	K	N	604
gtt	gtc	gat	gct	ttg	cta	aat	cat	cca	gac	att	gat	aaa	aat	1940
A	Q	S	T	(S)	G	L	T	P	F	H	L	A	I	618
gcg	cag	tct	acg	tca	gga	ttg	act	cct	ttc	cat	tta	gca	att	1982
I	N	E	S	Q	E	V	A	E	S	L	V	E	S	632
att	aat	gaa	agt	caa	gaa	gtt	gca	gaa	tct	tta	gtg	gaa	agt	2024
N	A	D	L	N	I	Q	D	V	(N)	H	M	A	P	646
aat	gct	gat	cta	aat	att	cag	gat	gtt	aac	cat	atg	gct	cct	2066
I	H	F	A	A	S	M	G	S	I	K	M	L	R	660
att	cat	ttt	gca	gct	tca	atg	ggg	agt	att	aaa	atg	ctt	aga	2108
Y	L	I	S	I	K	D	K	V	S	I	N	S	V	674
tat	ctc	att	tcc	ata	aaa	gat	aaa	gtt	agt	att	aat	tct	gtg	2150
T	E	N	(N)	N	W	T	P	L	H	F	A	I	Y	688
act	gag	aat	aat	aat	tgg	aca	cct	tta	cat	ttt	gct	ata	tat	2192

(Продолжение)

F	K	K	E	D	A	A	K	E	L	L	K	Q	D	702
ttt	aaa	aaa	gaa	gat	gct	gca	aaa	gaa	ttg	ttg	aaa	caa	gat	2234
D	I	N	L	T	I	V	A	D	Ⓞ	N	L	T	V	716
gac	ata	aat	tta	aca	att	gtt	gca	gat	ggt	aat	ctt	acc	gtt	2276
L	H	L	A	V	S	T	G	Q	I	N	I	I	K	730
tta	cat	ctt	gct	gtt	tcg	aca	gga	caa	ata	aat	ata	att	aaa	2318
E	L	L	K	R	G	S	N	I	E	E	K	T	G	744
gaa	tta	ttg	aag	aga	ggc	tcc	aat	ata	gaa	gaa	aaa	act	gga	2360
Ⓞ	G	Y	T	S	L	H	I	A	A	M	R	K	E	758
gaa	gga	tat	aca	tct	ctc	cac	atc	gct	gcg	atg	cga	aag	gag	2402
P	E	I	A	V	V	L	I	E	N	G	A	D	I	772
cca	gag	ata	gct	gtt	gtt	ttg	att	gaa	aac	ggt	gct	gac	ata	2444
E	A	R	S	A	Ⓞ	N	L	T	P	L	H	S	A	786
gaa	gct	cga	tca	gct	gat	aat	tta	aca	cct	tta	cat	tct	gcc	2486
A	K	I	G	R	K	S	T	V	L	Y	L	L	E	800
gca	aaa	ata	gga	agg	aaa	tct	aca	gta	ctt	tac	tta	tta	gaa	2528
K	G	A	D	I	G	A	K	T	A	Ⓞ	G	S	T	814
aaa	gga	gct	gac	att	gga	gct	aaa	aca	gca	gac	ggt	tct	act	2570
A	L	H	L	A	V	S	G	R	K	M	K	T	V	828
gcc	ttg	cat	tta	gct	gta	tct	ggt	cgt	aaa	atg	aaa	act	gtt	2612
E	T	L	L	N	K	G	A	N	L	K	E	Y	D	842
gaa	act	cta	tta	aat	aaa	gga	gca	aat	tta	aaa	gaa	tac	gat	2654
N	Ⓞ	K	Y	L	P	I	H	K	A	I	I	N	D	856
aac	aat	aaa	tat	ttg	cca	ata	cat	aaa	gct	att	att	aat	gat	2696
D	L	D	M	V	R	L	F	L	E	K	D	P	S	870
gac	ctt	gac	atg	gta	cgt	ttg	ttt	ctt	gaa	aaa	gat	ccc	agt	2738
L	K	D	D	E	T	E	Ⓞ	G	R	T	S	I	M	884
ctc	aaa	gat	gat	gaa	aca	gaa	gag	ggt	aga	act	tca	att	atg	2780
L	I	V	Q	K	L	L	L	E	L	Y	N	Y	F	898
tta	att	gtt	cag	aaa	ttg	ctt	ctt	gaa	tta	tat	aac	tat	ttt	2822
I	N	N	Y	A	E	T	L	D	E	E	A	L	F	912
ata	aat	aat	tat	gct	gaa	act	ttg	gat	gaa	gaa	gct	tta	ttc	2864
N	R	L	D	E	Q	G	K	L	E	L	A	Y	I	926
aac	cgc	tta	gat	gaa	caa	ggg	aaa	tta	gag	ctt	gca	tat	atc	2906
F	H	N	K	E	G	D	A	K	E	A	V	K	P	940
ttc	cat	aat	aaa	gaa	ggt	gat	gca	aaa	gag	gct	gtt	aag	cca	2948
T	I	L	V	T	I	K	L	M	E	Y	C	L	K	954
act	atc	ctt	gtt	aca	att	aaa	ctt	atg	gaa	tac	tgc	tta	aaa	2990
K	L	R	E	E	S	G	A	P	E	G	S	F	D	968
aaa	ctt	cgc	gaa	gag	tct	gga	gct	cct	gaa	ggt	agt	ttc	gat	3032
S	P	S	S	K	Q	C	I	S	T	F	S	E	D	982
tct	cca	tct	tca	aag	caa	tgt	att	tct	acc	ttt	tca	gag	gat	3074
E	M	F	R	R	T	L	P	E	I	V	K	E	T	996
gaa	atg	ttt	cgt	cgt	act	tta	ccg	gaa	att	gta	aaa	gaa	acg	3116
N	S	R	Y	L	P	L	K	G	F	S	R	S	L	1010
aac	age	aga	tat	tta	cca	cta	aag	ggc	ttt	tct	cgc	agc	cta	3158
N	K	F	L	P	S	L	K	F	A	E	S	K	N	1024
aat	aag	ttt	ctc	cct	tct	cta	aaa	ttt	gcc	gaa	agt	aag	aat	3200
S	Y	R	S	E	N	F	V	S	N	I	D	S	N	1038
agc	tac	aga	tct	gaa	aat	ttt	gtt	agc	aat	att	gat	tcc	aac	3242
G	A	L	L	L	L	D	V	F	I	R	K	F	T	1052
gga	gca	tta	ctt	tta	ctc	gat	gta	ttt	atc	aga	aag	ttt	atc	3284

(Продолжение)

N	E	K	Y	N	L	T	G	K	E	A	V	P	Y	1066													
a at	g ag	aaa	t ac	aat	t tg	a ct	g ga	aaa	g aa	g ct	g ta	c cc	t at	3326													
L	E	A	K	A	S	S	L	R	I	A	S	K	F	1080													
ctg	g aa	g ca	a ag	g ct	t ca	t ca	t ta	c gt	a tc	g ct	t ct	a aa	t tt	3368													
E	E	L	L	T	E	V	K	G	I	P	A	G	E	1094													
g aa	g aa	c tt	c ta	a ct	g aa	g tt	a aa	g gt	a tt	c cg	g ct	g ga	g ag	3410													
L	I	N	M	A	E	V	S	S	N	I	H	K	A	1108													
cta	att	aat	atg	gcc	g aa	g tg	ag t	t cc	a ac	a ta	c at	a ag	g ca	3452													
I	A	S	G	K	P	V	S	K	V	L	C	S	Y	1122													
att	g ca	ag t	g gt	a ag	c ct	g ta	t ca	a aa	g tc	t ta	t gt	t cg	t at	3494													
L	D	T	F	S	E	L	N	S	Q	Q	M	E	E	1136													
t tg	g at	a cc	t tt	t ct	g aa	t ta	a at	t ct	c aa	c aa	a tg	g aa	g aa	3536													
L	V	N	T	Y	L	S	T	K	P	S	V	I	T	1150													
t ta	g tt	a ac	a ca	t ac	t ta	t cc	a cc	a aa	c ct	t ct	g ta	a tt	a cg	3578													
S	A	S	A	D	Y	Q	K	L	P	N	L	L	T	1164													
t ca	g ca	t ct	g ca	g at	t ac	c ag	a aa	c tt	c tt	a at	t tg	t ta	a ct	3620													
A	T	C	L	E	P	E	R	M	A	Q	L	I	D	1178													
g ca	a ct	t gc	t ta	g aa	c ca	g aa	a ga	a tg	g ct	c aa	c tt	a ta	g at	3662													
V	H	Q	K	M	F	L	R							1186													
g tg	c at	c aa	a ag	a tg	t tt	t ta	c g	t ta	a a	a ta	c c	a tt	c ct	t ct	g t	g t	c at	3711									
c ca	t g	a g	t a	t g	g at	t g	c t	c t	t t	c t	t t	t at	g ca	a at	a t	t t	t t	t t	c c	a c	3761						
t t	t a	a t	g at	t t	a t	t t	c g	g at	t t	c	t	a c	t t	c t	t t	c t	t t	a t	t		3811						
t c	g t	t g	t t	c t	g t	g c	t t	t t	g a	g t	g a	a at	t at	c t	g at	a at	c a	t t	t c	t g	a g	t t	3861				
a g	t t	t c	c t	c t	c ta	a g	g ag	g at	t t	t ta	g g	a at	t ac	a t	g at	t t	t t	t g	g t				3911				
t t	c g	g g	t t	t t	a c	t t	c ta	t at	t t	t t	g t	c c	a aa	g ta	a at	t t	g ta	g g	a g	g g	a at		3961				
t c	t a	a aa	a g	g	a at	g ga	a t	c g	t t	g a	t aa	a at	a	g t	t t	t g	t at	a t	t g	t ac	c t		4011				
t a	t a	a ta	a ta	a	g a	c t	t t	a aa	g ta	a at	t t	t t	t a	a aa	c c	a ag	c at	a a	a g	t at			4061				
a g	a a	g ga	c t	c	g t	a t	g t	t a	t a	t a	t a	t a	t a	a ta	g aa	a at	a ta	a g	t t	g cc	a aa	t g	t c	a t	g	4111	
g at	t c	a a	t c	a	a ca	a at	t ta	a at	a at	t aat	a at	a at	a at	a at	a at	a at	a at	a at	a at	a at	a at	a at	a at	a at	a at	a at	4161
t t	t t	t a	a t	g t	t c	t t	a g	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	t a	4211
g t	t t	a at	t t	t	g c	t at	c t	g c	t	a t	t a	c t	t a	a	a	a	a	a	a	a	a	a	a	a	a	a	4261
t c	t a	a at	a t	t	c t	c g	t a	a c	a t	t at	t t	t a	a aa	g t	g ta	a at	a g	a g	a c	t at	a						4311
t t	t t	t ta	a at	t	a ta	t cc	a t	t t	t	t at	a at	t g	t t	a at	t t	c	t t	c a	a a	a t	t a						4361
c at	t c	c t	t t	t	c c	t g	c at	t at	t at	t t	a c	a at	t c	a at	t a	a at	t g	c c	t at	a t	t t	t t					4411
a a	a g	a c	a at	a	c t	c a	a g	t t	t c	a g	a t	t at	t at	a a	a c	a g	t c	t t	a ta	t t	g at	t c					4461
a a	c g	t a	a at	t	t t	a at	t g	a c	a	a t	t g	c a	a ga	a at	a at	a at	t a	a at	t g	t a							4511
t t	t a	a g	t at	t	t g	a a	c ag	t	t g	a t	t t	t ta	t t	g ca	a at	a t	g	a at	c ta	a at							4561
a t	t t	t c	a a	c	a at	g at	a g	a a																			4581

(Окончание)

ТАА, кодирует белок длиной 1214 а. о. Последовательность, окружающая кодон АТГ, частично отвечает набору статистических правил, установленных для иницирующего кодона эукариотических генов [10]. При анализе выведенной аминокислотной последовательности обнаружены участки, структура которых хорошо согласуется с результатами проведенного нами белкового секвенирования δ -ЛИТ и его фрагментов.

С помощью N-концевого секвенирования нативного токсина было установлено, что первые 28 а. о. предшественника удаляются при созревании δ -ЛИТ. Структура отщепляемой последова-

тельности не характерна для сигнального пептида эукариот (ср. [11]).

По компьютерным данным предсказания вторичной структуры [12], δ -ЛИТ преимущественно образован элементами α -спирали. Несмотря на то что токсин – растворимый белок, в его молекуле обнаружены два гидрофобных участка, потенциально образующих трансмембранные сегменты. Первый участок имеет протяженность 29 а. о. (39-67), второй – 20 а. о. (222-241).

Центральная часть молекулы δ -ЛИТ содержит 13 тандемно расположенных внутримолекулярных

повторов анкиринового типа, обнаруженных также в белках, входящих в состав цитоскелета, а также принимающих участие в регуляции клеточного цикла, дифференцировки и транскрипции [13]. В С-концевой части молекулы токсина расположена область, богатая остатками Pro, Glu и Ser (958-972). Такие последовательности, получившие название PEST-мотива, часто встречаются в белках, подвергающихся внутриклеточной деградации [14].

Таким образом, молекулу токсина можно разделить на три структурных домена: N-концевой домен (1-459), который содержит два трансмембранных сегмента, центральный домен (460-912), состоящий из повторов анкиринового типа, и С-концевой домен (913-1186), в составе которого обнаружен PEST-мотив.

Сравнение аминокислотных последовательностей трех латротоксинов выявило значительную степень гомологии между ними. Степень гомологии для α -ЛИТ и δ -ЛИТ (38%), а также для α -ЛТ и δ -ЛИТ (37%) примерно одинакова. В целом максимальное сходство между токсинами наблюдается в области N-концевых доменов. Это не касается первых 70 N-концевых а. о., которые мало сходны между собой у разных токсинов. Расположение гидрофобных участков, потенциально образующих трансмембранные фрагменты, консервативно у всех трех латротоксинов. В то же время значительная гомология наблюдается только между аминокислотными последовательностями второго участка, тогда как структура первого гидрофобного участка уникальна у каждого токсина. Степень гомологии для δ -ЛИТ и токсинов α -ряда в области N-концевого домена даже выше, чем для α -ЛИТ и α -ЛТ. Таким образом, структура N-концевого домена δ -ЛИТ является промежуточной между соответствующими структурами α -ЛИТ и α -ЛТ.

В области домена с анкириновыми повторами гомология между токсинами существенно ниже (порядка 33 - 34%). δ -ЛИТ содержит только 13 полных анкириновых повторов, тогда как в структуре α -ЛИТ и α -ЛТ обнаружено по 20 копий повторов. Первые 115 а. о. С-концевого домена δ -ЛИТ (913-1027) образуют уникальную последовательность, которая не гомологична ни одному из известных белков. Гомология между δ -ЛИТ и токсинами α -ряда снова наблюдается на С-конце: степень гомологии для последних 159 а. о. δ -ЛИТ (1028-1186) и α -ЛИТ составляет 36%. Эти фрагменты, вероятно, удаляются при созревании обоих токсинов. Значительное сходство между этими участками, скорее всего, объясняется наличием специальных сигналов транспортировки на С-конце латротоксинов.

СПИСОК ЛИТЕРАТУРЫ

1. Grishin E. // Pure & Appl. Chem. 1994. V. 66. P. 783 - 790.
2. Rosental L., Meldolesi J. // J. Pharmacol. Ther. 1989. V. 42. P. 115 - 134.
3. Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V. // FEBS Lett. 1990. V. 270. P. 127 - 131.
4. Kiyatkin N., Dulubova I., Grishin E. // Eur. J. Biochem. 1993. V. 213. P. 121 - 127.
5. Krasnoperov V.G., Shamotienko O.G., Grishin E.V. // J. Natural Toxins. 1992. V. 1. P. 17 - 23.
6. Gubler U., Hoffman B.J. // Gene. 1983. V. 25. P. 263 - 269.
7. Melton D.A. // Nucl. Acids Res. 1984. V. 12. P. 7035 - 7056.
8. Henikoff S. // Gene. 1984. V. 28. P. 351 - 357.
9. Tabor S., Richardson C.C. // Proc. Natl. Acad. Sci. USA. 1987. V. 84. P. 4767 - 4775.
10. Kozak M. // Nucl. Acids Res. 1984. V. 12. P. 857 - 872.
11. Von Heijne G. // Nucl. Acids Res. 1986. V. 14. P. 4683 - 4690.
12. Garnier J., Osguthorpe D.S., Robson B. // J. Mol. Biol. 1978. V. 120. P. 97 - 120.
13. Michaely P., Bennett V. // Trends Cell Biol. 1992. V. 2. P. 127 - 129.
14. Rechsteiner M., Royers S., Rote K. // Trends Biochem. Sci. 1987. V. 12. P. 390 - 394.

The Primary Structure of δ -Latroinsectotoxin from the *Latrodectus mactans tredecimguttatus* Venom

I. E. Dulubova, M. V. Khvoshchev,¹ V. G. Krasnoperov, T. G. Galkina, K. A. Pluzhnikov, T. M. Volkova, and E. V. Grishin

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, V-437, GSP-7, Moscow, 117871 Russia

Abstract—The structural gene of δ -latroinsectotoxin was cloned and its nucleotide sequence was determined. The gene contains an open reading frame of 3642 bp. The deduced amino acid sequence is homologous to the sequences of latrotoxins studied earlier.

Key words: neurotoxin, *Latrodectus*, cDNA cloning.

¹ To whom correspondence should be addressed.