



УДК 595.443.7-577.112.5

КЛОНИРОВАНИЕ И СТРУКТУРА ГЕНА, КОДИРУЮЩЕГО α -ЛАТРОКРУСТОТОКСИН В СОСТАВЕ ЯДОВИТЫХ ЖЕЛЕЗ ПАУКА КАРАКУРТА

© 1999 г. В. Н. Данилевич[#], С. А. Лукьянов, Е. В. ГришинИнститут биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН,
117871, Москва, ГСП-7, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 10.10.98 г. Принята к печати 22.03.99 г.

Определена первичная структура гена *crusta* (на уровне геномной ДНК) паука каракурта, кодирующего α -латрокрустотоксин (α -LCT) – высокомолекулярный нейротоксин, специфичный в отношении ракообразных. Общая длина секвенированной ДНК составляет 4693 п.о. Структурная часть хромосомного гена каракурта, кодирующая α -LCT, не содержит интронов. Секвенированная ДНК содержит единственную открытую рамку считывания достаточной длины (4185 п.о.) и кодирует белковый предшественник α -LCT, состоящий из 1395 а.о. Первым аминокислотным остатком белка-предшественника является, как мы предполагаем, Met в положении –10 по отношению к N-концевому остатку Glu1 зрелого токсина. Вычисленная молекулярная масса белка-предшественника (156147 Да) превышает таковую зрелого токсина (~120 кДа) на ~30 кДа. Полученные данные согласуются с представлением о том, что в процессе созревания белок-предшественник претерпевает двойной процессинг – отщепление декапептида в N-концевой части и фрагмента размером ~200 а.о. в C-концевой части. Выявлено наличие у α -LCT ряда несовершенных анкириновых повторов и областей структурной гомологии с ранее изученными латротоксинами; наибольшая гомология (62%) обнаружена с α -латроинсектотоксином (α -LIT).

Ключевые слова: нейротоксин; *Latrodectus*; геномная ДНК; ПЦР-амплификация; нуклеотидная последовательность; выведенная аминокислотная последовательность.

Ядовитые железы паука *Latrodectus mactans tredicimguttatus* вырабатывают целое семейство высокомолекулярных (110–130 кДа) белковых нейротоксинов, стимулирующих секрецию различных нейромедиаторов у позвоночных и беспозвоночных животных [1–4]. Так, α -латротоксин (α -LTX) индуцирует освобождение нейромедиаторов в синапсах позвоночных; α - и δ -латроинсектотоксины (α -LIT и δ -LIT) активны по отношению к насекомым и не токсичны для позвоночных [5, 6]. Еще один представитель семейства – α -латрокрустотоксин (α -LCT) токсичен по отношению к ракообразным и не токсичен для позвоночных и насекомых [7, 8].

Понимание механизма биологической активности токсинов и специфичности их действия невозможно без информации об их структуре. Из всего семейства белковых токсинов, выделенных из яда паука каракурта (около 10), первичная структура определена к настоящему времени лишь у трех представителей: α -LTX [9], α -LIT [10] и δ -LIT [11]. При этом аминокислотные последовательности этих белков были выведены на осно-

вании нуклеотидных последовательностей соответствующих кДНК.

Необходимо отметить, что к моменту выполнения настоящей работы в литературе отсутствовали данные о том, как устроены хромосомные гены каракурта, кодирующие высокомолекулярные нейротоксины. В частности, было неизвестно, содержат ли они интроны. Этот вопрос представляет не только теоретический интерес, но может иметь и практическое значение. В случае отсутствия интронов значительно упрощается задача клонирования и секвенирования соответствующих генов.

Цель настоящей работы – определение первичной структуры (а стало быть и изучение геномной организации) хромосомного гена каракурта, кодирующего синтез α -LCT – одного из наиболее изученных на физиологическом уровне высокомолекулярных нейротоксинов. Определение полной нуклеотидной последовательности кодирующей части гена, кроме того, позволяло установить аминокислотную последовательность соответствующего белка и сравнить его структуру со структурами ранее изученных нейротоксинов данного семейства.

[#] Автор для переписки (e-mail: dan@ibch.siobc.ras.ru; тел.: (095) 336-65-40; факс: 330-73-01).

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Ранее в нашей лаборатории была определена нуклеотидная последовательность внутреннего фрагмента гена *crusta* размером около 1200 п.о. [12]. Для этого данный фрагмент был амплифицирован при помощи ПЦР с использованием двух вырожденных праймеров и кДНК, полученной на основе суммарной мРНК из яда паука. Вырожденные праймеры были сконструированы исходя из аминокислотных последовательностей химотриптического пептидов α -латрокрустотоксина [12]. Полученная информация о нуклеотидной последовательности внутреннего фрагмента гена *crusta* была использована нами для клонирования его недостающих 5'- и 3'-концевых фрагментов с применением метода "прогулки по геному". В настоящее время известно несколько подходов, основанных на технологии ПЦР и пригодных для "прогулки" из известного района геномной ДНК в неизвестный. Мы использовали одну из последних разработок фирмы Clontech [13], которая основана на ранее описанном принципе супрессии ПЦР длинными инвертированными повторами [14]. Согласно описанной процедуре, к концам фрагментов геномной ДНК, расщепленной рестриктазами *SmaI*, *DraI*, *PvuII* и *HpaI*, образующими тупые концы, был пришит специальный адаптер. Этот адаптер образован при отжиге олигонуклеотидов AP1 и AP2 (таблица). Один конец адаптера – двухцепочечный, так что он может лигироваться к двум концам любого фрагмента ДНК, имеющего "тупые" концы. Другой конец – одноцепочечный, и, вследствие этого, не способен вступать в реакцию лигирования. Полученные смеси фрагментов ДНК с пришитыми по концам молекулами адаптера были использованы для ПЦР-амплификации.

Продукты ПЦР – перекрывающиеся фрагменты ДНК различной длины (от 500 до 2800 п.о.), схематически представлены на рис. 1.

В экспериментах по амплификации 5'-концевой области гена *crusta* нами было выделено 5 ПЦР-фрагментов. При этом два фрагмента наибольшего размера (фрагменты S(1) и S(2), длиной ~2200 и ~2800 п.о., соответственно) (рис. 1) были получены в опыте, когда геномную ДНК расщепляли рестриктазой *SmaI*. Как оказалось, меньший из них образовался в результате неспецифического отжига адаптерного праймера AP4 с GC-богатым участком ДНК гена *crusta* (область 614-629). В опытах с фрагментами геномной ДНК, образованными при действии рестриктаз *DraI*, *PvuII* и *HpaI*, было получено по одному ПЦР-фрагменту (на рис. 1 соответствующие фрагменты обозначены D(1), P и H).

При амплификации 3'-концевой области гена *crusta* был выделен лишь один ПЦР-фрагмент длиной около 1200 п.о. (фрагмент D(2)). Этот фрагмент был получен в опыте с геномной ДНК, расщепленной рестриктазой *DraI*.

Кроме того, нами была проведена ПЦР-амплификация с использованием вырожденного праймера PND (соответствующего N-концевой аминокислотной последовательности белка α -LCT) и генспецифического праймера P2. Размер полученного при этом фрагмента ND (рис. 1) составлял ~2350 п.о.

Используя нерасщепленную геномную ДНК и встречные праймеры P1 и P6, мы амплифицировали также фрагмент 2711–3892 гена *crusta*. Этот фрагмент (на рис. 1 он обозначен P1–P6) практически полностью соответствует ранее секвенированной части гена α -LCT [12].

Праймерные и адаптерные олигонуклеотиды, использованные для выделения и клонирования фрагментов гена α -латрокрустотоксина

Олигонуклеотид	Структура олигонуклеотида (5' → 3')*
AP1	GTAATACGACTCACTATAGGGCAGCGTGGTTCGCGGCCGAGGT
AP2	ACCTCGGC
AP3	GTAATACGACTCACTATAGGGC
AP4	AGCGTGGTTCGCGGCCGAGGT
P1	AACCCGAATACTGAGACACCACAG (2711-2734)
P2	GGAAAATGTTGGTTTTGCCTAATG (2808-2831)
P3	TGCTCCTTTGTCCATAAGTAACTG (2830-2854)
P4	ACCACTTTTCATTGCTGCTATGATT (3718-3742)
P5	ATTGGACAATACGATATCGTAAAG (3740-3763)
P6	TACATCTGCTCCSTTTTGTACCAA (3869-3892)
PND	GAAATGTCTAA(A,G)GC(A,T)GA(T,C)CA(A,G)AG

* В скобках указаны координаты участков ДНК гена *crusta*, соответствующие или комплементарные праймерным олигонуклеотидам.

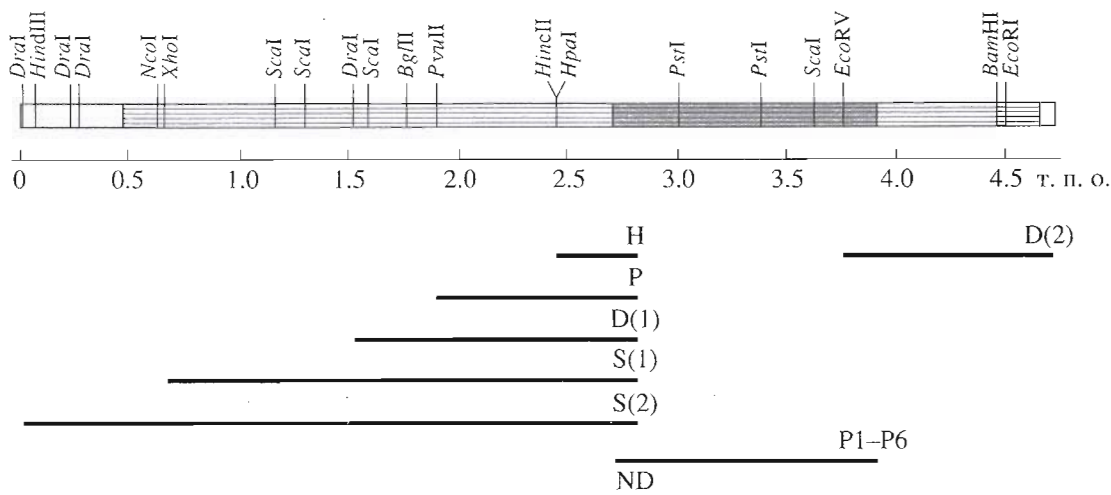


Рис. 1. Рестриктивная карта участка хромосомной ДНК каракурта с геном *crusta* и локализация выделенных с помощью ПЦР-амплификации фрагментов этого гена. Затемненная область – ранее секвенированный фрагмент ДНК. Затрихованная область – кодирующая часть гена *crusta*.

Полученный в итоге набор ПЦР-фрагментов клонировали в плазмиду рBSIIsk(+). ПЦР-фрагменты большой длины, кроме того, были субклонированы в ту же векторную плазмиду. Полученные рекомбинантные плазмиды далее секвенировали. Для секвенирования использовали прямой и обратный универсальные M13-праймеры; праймеры SK и KS, специфичные для рBSIIsk(+); адаптерный праймер AP4, а также генспецифические праймеры, использовавшиеся в ПЦР (таблица). При определении первичной структуры была использована также стратегия шагового секвенирования, основанная на использовании олигонуклеотидных праймеров к уже секвенированным участкам ДНК. С этой целью было синтезировано 7 дополнительных 21–24-звенных праймеров (их структура не представлена).

Общая длина секвенированной нами ДНК – 4693 п.о. (рис. 2). Последовательность содержит открытую рамку считывания длиной 4185 п.о., которая начинается с триплета ATG (кодирует остаток Met) и заканчивается терминирующим кодоном TAA. Последовательность потенциально кодирует белок, состоящий из 1395 а.о. с вычисленной молекулярной массой 156147 Да и значением *pI* 8,2.

Ранее, методами белковой химии была определена структура ряда химотриптических пептидов α -LCT, а также его *N*-концевая аминокислотная последовательность [12]. Выведенная нами структура α -LCT содержит все последовательности химотриптических пептидов (рис. 2). *N*-Концевая последовательность зрелого крестотоксина – EMSKADQC... начинается с аминокислотного остатка Glu под номером 1 (рис. 2). Остаток Met, т.е. начало рамки считывания, находится

проксимальнее остатка Glu1 в положении –10. Других остатков Met проксимальнее *N*-концевой последовательности α -LCT не имеется.

N-Концевая последовательность другого токсина – α -LIT, специфичного к насекомым, имеет сходную структуру – EMSRADQC... [10]. В положении –10 от *N*-концевого аминокислотного остатка этого токсина также находится остаток Met. Более того, прослеживается большая гомология в аминокислотных последовательностях дистальнее метионина (MKGKRVISKR... – у α -LCT и MKGSSAISKR... – у α -LIT). Если трансляция белка действительно начинается с Met (–10), эти декапептиды по причине своей небольшой длины не могут претендовать на роль сигнальных последовательностей, но, вероятно, входят в состав про-последовательностей зрелого белка и отщепляются в процессе созревания токсина. Наличие в положении –2 и –1 последовательности Lys-Arg, узнаваемой специфическими эндопептидазами (расщепляющими связь Arg–X), свидетельствует в пользу посттрансляционного процессинга, происходящего в *N*-концевой части белковой молекулы.

Вычисленная молекулярная масса полипептида, начинающегося с *N*-концевого остатка Glu1 (154962 Да), существенно отличается от молекулярной массы крестотоксина, определенной ранее с помощью SDS-ПААГ-электрофореза (~120 кДа) [7]. Таким образом, α -LCT, по всей видимости, претерпевает процессинг и в *S*-концевом районе во время созревания. Идея о двойном процессинге белков-токсинов из яда паука каракурта была высказана ранее применительно к другим латротоксинам [9–11]. Так, в процессе созревания δ -LIT от его *S*-концевой части отщепляется

1						TCT	TTA	AAA	TAT	ATC	ATA	ATT	CAG	CTA	TTC	TAT	CAA	TAA	TCA	ACA
46	ATT	GAA	GCT	TTC	TTT	TTT	GTT	TAA	CGT	ACA	TCG	AAG	AAG	CAT	TGA	AAT	ATT	CAC	ACA	GGT
106	GGT	ATG	AAT	TAA	ATT	CCA	AAA	GGT	GAA	TAT	CAC	ATT	AAG	TAT	TTC	AAA	TTG	ACA	GTT	GAA
166	GTT	GAT	CCC	TCA	GAC	AAA	ATA	TCA	TTC	CCA	TAA	CAA	TAT	CCG	AAA	CTA	TAA	CTC	CTT	TTT
226	AAA	TGC	CGC	AGA	CAT	TTT	TAA	AAC	TTG	TTC	ATT	GTC	TTG	TCA	AAA	ATT	ATT	TGA	CTA	CAG
286	GTG	ATG	AAA	TTG	TTC	ATG	GTA	AAT	TGC	ATA	TTT	AAG	TCG	AAG	CAA	ATT	TGG	ATG	ATC	ACT
346	ATT	CGT	TCT	GCA	TTA	GCT	GTC	TAT	TGC	CAG	AAT	TAA	GTA	AGT	ATC	TTC	ATT	TTT	CAT	TTC
-28													Val	Ser	Ile	Ile	Phe	His	His	Phe
406	TCG	GCA	AAT	ATT	TTG	GTC	AGG	AAT	AGT	GAA	ATG	AAA	GGA	AAA	CGC	GTA	ATT	TCA	AAA	AGA
-20	Ser	Ala	Asn	Ile	Leu	Val	Arg	Asn	Ser	Glu	<u>Met</u>	<u>Lys</u>	<u>Gly</u>	<u>Lys</u>	<u>Arg</u>	<u>Val</u>	<u>Ile</u>	<u>Ser</u>	<u>Lys</u>	<u>Arg</u>
466	GAA	ATG	TCT	AAG	GCC	GAT	CAA	TGT	ACA	TTT	TTA	TCG	TAT	CAG	TCA	GTT	GCG	TAT	GGT	ACA
1	<u>Glu</u>	<u>Met</u>	<u>Ser</u>	<u>Lys</u>	<u>Ala</u>	<u>Asp</u>	<u>Gln</u>	<u>Thr</u>	<u>Phe</u>	<u>Leu</u>	<u>Ser</u>	<u>Tyr</u>	<u>Gln</u>	<u>Ser</u>	<u>Val</u>	<u>Ala</u>	<u>Tyr</u>	<u>Gly</u>	<u>Thr</u>	
526	TTA	GGC	GAT	GTT	GCA	GGA	GAT	GTT	TCT	TCC	ATT	GAG	GGA	GCG	GAT	TTG	GTC	GCG	ACA	CCG
21	Leu	Gly	Asp	Val	Ala	Gly	Asp	Val	Ser	Ser	Ile	Glu	Gly	Ala	Asp	Leu	Val	Ala	Thr	Pro
586	ATT	GCC	GCA	GGT	GGC	CAC	CTT	GCA	AAA	GGG	GCG	ACG	GAC	GCG	GCC	ATT	ATT	GCC	ATG	GAT
41	Ile	Ala	Ala	Gly	Gly	His	Leu	Ala	Lys	Gly	Ala	Thr	Asp	Ala	Ala	Met	Ile	Ala	Met	Asp
646	TGC	TCG	AGC	ATA	CCA	TTC	GAT	GAG	ATA	AAA	CAG	CAA	TTA	AAT	CAA	AGA	TTC	AAC	GAG	GTG
61	Cys	Ser	Ser	Ile	Pro	Phe	Asp	Glu	Ile	Lys	Gln	Gln	Leu	Asn	Gln	Arg	Phe	<u>Asn</u>	<u>Glu</u>	<u>Val</u>
706	GAT	AAG	AAA	CTG	CAA	AAA	GGC	GCT	GAA	GCA	CTG	GAA	AAC	GTC	CTG	GAA	TGA	GCA	GAG	AAA
81	<u>Asp</u>	<u>Lys</u>	<u>Lys</u>	<u>Leu</u>	<u>Glu</u>	<u>Lys</u>	<u>Gly</u>	<u>Ala</u>	<u>Glu</u>	<u>Ala</u>	<u>Leu</u>	<u>Glu</u>	<u>Asn</u>	<u>Val</u>	<u>Thr</u>	<u>Glu</u>	<u>Leu</u>	<u>Ala</u>	<u>Glu</u>	<u>Lys</u>
766	ACC	TAT	TCT	TCA	GTC	GAA	AAA	ATG	AGA	GTT	GAG	ATG	AGG	GAG	GGC	TTC	AAT	CAC	GTC	ATC
101	<u>Thr</u>	<u>Tyr</u>	<u>Ser</u>	<u>Ser</u>	<u>Val</u>	<u>Glu</u>	<u>Lys</u>	<u>Met</u>	<u>Arg</u>	<u>Val</u>	<u>Glu</u>	<u>Met</u>	<u>Arg</u>	<u>Glu</u>	<u>Gly</u>	<u>Phe</u>	<u>Asn</u>	<u>His</u>	<u>Val</u>	<u>Ile</u>
826	GCC	ACT	ATT	GAA	AAC	GCT	AAT	ACT	AAG	CAA	ATC	ATT	ACA	GGG	ATT	AAT	CAA	ATT	ATA	CAA
121	Ala	Thr	Ile	Glu	Asn	Ala	Asn	Thr	Lys	Gln	Ile	Ile	Thr	Gly	Ile	Asn	Gln	Ile	Ile	Gln
886	TAT	TTC	AAT	GAC	GAA	CGG	GAG	AAT	ATA	AAC	AAC	AGG	CAA	AAG	GAA	GAT	TAT	GTT	GCT	AAA
141	Tyr	Phe	Asn	Asp	Glu	Arg	Glu	Asn	Ile	Asn	Asn	Arg	Gln	Lys	Glu	Asp	Tyr	Val	Ala	Lys
946	CTG	CAA	GAA	CCG	GCA	TCT	GGA	AAT	TTT	TTG	CTT	TAC	CTC	AGA	AAA	TCA	AGA	ACT	TCT	GAG
161	Leu	Gln	Glu	Pro	Ala	Ser	Gly	Asn	Phe	Leu	Leu	Tyr	Leu	Arg	Lys	Ser	Arg	Thr	Ser	Glu
1006	GAT	GGT	AGT	TTG	CAT	AGT	TTG	CTG	TTT	AAG	ATC	ATC	AAC	CAA	GAG	TTG	GCA	ATT	CCT	AAT
181	Asp	Gly	Ser	Leu	His	Ser	Leu	Leu	Phe	<u>Lys</u>	<u>Ile</u>	<u>Ile</u>	<u>Asn</u>	<u>Gln</u>	<u>Glu</u>	<u>Leu</u>	<u>Ala</u>	<u>Ile</u>	<u>Pro</u>	<u>Asn</u>
1066	AAC	GCT	GCT	GAT	AAC	AAT	GCG	ATT	CGA	GCT	CTT	TTC	GCT	TTG	TTT	TAC	GGT	ACA	CAA	ACA
201	<u>Asn</u>	<u>Ala</u>	<u>Ala</u>	<u>Asp</u>	<u>Asn</u>	<u>Ala</u>	<u>Ile</u>	<u>Arg</u>	<u>Ala</u>	<u>Leu</u>	<u>Phe</u>	<u>Ala</u>	<u>Leu</u>	<u>Phe</u>	<u>Tyr</u>	<u>Gly</u>	<u>Thr</u>	<u>Gln</u>	<u>Thr</u>	
1126	TTC	ATT	TCG	ATT	ATG	TTC	TAT	CTC	GTC	AAA	CAG	TAC	TCT	TAT	CTG	GCT	GAT	TAC	CAT	TAC
221	Phe	Ile	Ser	Ile	Met	Phe	Tyr	Leu	Val	Lys	Gln	Tyr	Ser	Tyr	Leu	Ala	Asp	Tyr	His	Tyr
1186	CAA	AAC	GGT	AAT	TTA	GCT	GAA	TTT	AAT	TCA	AAT	TTC	GAT	CAT	ATG	AAA	ACT	GTA	TTC	CAA
241	Gln	Asn	Gly	Asn	Leu	Ala	Glu	Phe	Asn	Ser	Asn	Phe	Asp	His	Met	Lys	Thr	Val	Phe	Gln
1246	GAT	TTC	AAG	TTT	ACT	CTC	ATT	GGT	ATT	AAT	ACA	TCC	AAT	AGT	AAG	CCT	TTG	GTG	AAT	ACA
261	Asp	Phe	Lys	Phe	Thr	Leu	Ile	Gly	Ile	Asn	Thr	Ser	Asn	Ser	Lys	Pro	Leu	Val	Asn	Thr
1306	GTA	CTT	AGC	ATA	ATA	GAA	GAT	GTT	AAA	AAT	AAG	AGA	TTT	ATT	CGA	AAT	TTA	CGA	AGT	AAT
281	Val	Leu	Ser	Ile	Ile	Glu	Asp	Val	Lys	Asn	Lys	Arg	Phe	Ile	Arg	Asn	Leu	Arg	Ser	Asn
1366	TTA	TAT	CAA	AAG	ATA	ATA	AAA	TCA	ACG	AAA	TCG	TTA	TTA	GAT	TTG	AGA	GAA	AAA	ATT	ACC
301	Leu	Tyr	Gln	Lys	Ile	Ile	Lys	Ser	Thr	Lys	Ser	Leu	Leu	Asp	Lys	Arg	Glu	Lys	Ile	Thr
1426	AAA	ATG	GAT	CTT	CCT	ATA	ATT	GAG	GAT	ACG	CCT	AAA	TCC	TCA	GTT	TTA	ATT	AAT	TTT	AGG
321	Lys	Met	<u>Asp</u>	<u>Leu</u>	<u>Pro</u>	<u>Ile</u>	<u>Ile</u>	<u>Glu</u>	<u>Asp</u>	<u>Thr</u>	<u>Pro</u>	<u>Lys</u>	<u>Ser</u>	<u>Ser</u>	<u>Val</u>	<u>Leu</u>	<u>Ile</u>	<u>Asn</u>	<u>Phe</u>	<u>Arg</u>
1486	GAG	AAA	AGT	AGT	TCA	GTA	CCT	CGA	ATT	GAG	ACC	CCA	ATT	TTA	AAA	TGG	ACC	CCT	GGA	ACA
341	<u>Glu</u>	<u>Lys</u>	<u>Ser</u>	<u>Ser</u>	<u>Ser</u>	<u>Val</u>	<u>Pro</u>	<u>Arg</u>	<u>Ile</u>	<u>Glu</u>	<u>Thr</u>	<u>Pro</u>	<u>Ile</u>	<u>Leu</u>	<u>Lys</u>	<u>Trp</u>	<u>Thr</u>	<u>Pro</u>	<u>Gly</u>	<u>Thr</u>
1546	GTA	GTT	AAG	TAC	GCC	ATA	CAA	TAC	GAA	CAA	GAC	GGT	AAG	TAC	TCT	AAA	ATA	AGC	AAG	TGG
361	Val	Val	Lys	Tyr	Ala	Ile	Gln	Tyr	Glu	Gln	Asp	Gly	Lys	Tyr	Ser	Lys	Ile	Ser	Lys	Trp
1606	TCT	AAT	CCA	ATT	ACA	GTT	CAA	AGA	TTG	GCA	AAC	CCA	TAT	ATT	ACC	ATT	GAT	AAA	GAT	CGC
381	Ser	Asn	Pro	Ile	Thr	Val	Gln	Arg	Leu	Ala	Asn	Pro	Tyr	Ile	Thr	Ile	Asp	Lys	Asp	Arg
1666	AGA	AAT	AGG	CTA	GTA	TTT	AGA	CAA	TTT	GGA	AAT	GAG	AAG	CCT	GAA	CTG	ATA	AGT	ATA	TTA
401	Arg	Asn	Arg	Leu	Val	Phe	Arg	Gln	Phe	<u>Gly</u>	<u>Asn</u>	<u>Glu</u>	<u>Lys</u>	<u>Pro</u>	<u>Glu</u>	<u>Ile</u>	<u>Ser</u>	<u>Ile</u>	<u>Leu</u>	
1726	GAT	AGT	TCT	CAA	AAC	GAA	TTT	CGG	GAT	ATT	CAT	CGA	GAT	CTG	TAT	AAC	GCA	GCC	CAA	ATG
421	Asp	Ser	Ser	Gln	Asn	Glu	Phe	Arg	Asp	Ile	His	Arg	Asp	Leu	Tyr	Asn	Ala	Ala	Gln	Met
1786	CCC	TAT	AAG	GAG	ACG	GCT	TTG	GGT	ATT	TGT	CGG	AAA	TTG	ATA	GAT	AGC	GGA	GCA	CAA	GTG
441	Pro	Tyr	Lys	Gly	Thr	Ala	Leu	Gly	Ile	Cys	Arg	Lys	Leu	Ile	Asp	Ser	Gly	Ala	Gln	Val
1846	GGT	GCA	TCA	TTT	GAA	ATG	GGT	AGA	AAA	TCC	ATT	CAT	GCG	TCG	GCA	ACA	GCT	GGA	AAT	GAT
461	Gly	Ala	Ser	Phe	Glu	Met	Gly	Arg	Lys	Ser	Ile	His	Ala	Ser	Ala	Thr	Ala	Gly	Asn	Asp
1906	GAC	GTA	GCA	AGA	CTT	CTT	TTA	GCG	AAA	AAC	AAT	GGT	TTG	CTT	AAT	GTT	CCA	GAC	AAA	AAT
481	Asp	Val	Ala	Arg	Leu	Leu	Leu	Ala	Lys	Asn	Asn	Gly	Leu	Leu	Asn	Val	Pro	Asp	Lys	Asn
1966	GGC	TAC	ACT	CCA	CTT	CAT	ATA	GCT	TCC	GAA	CGT	AAG	AAC	AAC	GAT	TTC	GTT	AAG	TTT	CTC
501	Gly	Tyr	Thr	Pro	Leu	His	Ile	Ala	Ser	<u>Glu</u>	<u>Arg</u>	<u>Lys</u>	<u>Asn</u>	<u>Asp</u>	<u>Phe</u>	<u>Val</u>	<u>Lys</u>	<u>Phe</u>	<u>Leu</u>	
2026	CTT	GAA	AAG	GGG	GCA	GAT	GTC	AAT	GTT	CGT	ACG	TTT	GCA	AAC	GAG	CTA	ACG	CCC	TTA	CAT
520	Leu	Glu	Lys	Gly	Ala	Asp	Val	Asn	Val	Arg	Thr	Phe	Ala	Asn	Glu	Leu	Thr	Pro	Leu	His
2026	TTG	GCT	GCT	CGT	CAA	GAT	TTT	ACG	ATA	ATC	GTT	AAA	ACG	TTG	ATG	GAG	AAA	AGA	GGC	ATC
521	Leu	Ala	Ala	Arg	Gln	Asp	Phe	Thr	Ile	Ile	Val	Lys	Thr	Leu	Met	Glu	Lys	Arg	Gly	Ile
2146	GAT	GTG	AAT	GCG	AAA	GAA	CGA	GCT	GGA	TTC	ACG	CCG	TTA	CAC	CTT	TCG	ATT	ACC	AGC	AAT

Рис. 2. Нуклеотидная последовательность участка хромосомной ДНК каракурта с геном *crusta* и выведенная аминокислотная последовательность кодируемого им предшественника α -LCT. Одной чертой подчеркнуты аминокислотные последовательности химотриптических пептидов, строение которых определено ранее методами белковой химии [12]. Декапептид, предположительно отщепляемый в N-концевой части претоксина, подчеркнут двумя чертами.

561	Asp	Val	Asn	Ala	Lys	Glu	Arg	Ala	Gly	Phe	Thr	Pro	Leu	His	Leu	Ser	Ile	Thr	Ser	Asn
2206	TCC	AGA	GCA	GCC	AGA	ACC	TTA	ATA	AAT	GAA	ACT	CCT	GCC	GGA	ATA	AAC	ATT	AAA	TCA	AAC
581	Ser	Arg	Ala	Ala	Arg	Thr	Leu	Ile	Asn	Glu	Thr	Pro	Ala	Gly	Ile	Asn	Ile	Lys	Ser	Asn
2266	TCT	GGT	CTG	ACA	CCT	CTA	CAT	TTG	GCT	GTA	CTC	CAA	AAT	AAT	CTA	AGT	GCT	GCG	AAA	GTA
601	Ser	Gly	Leu	Thr	Pro	Leu	His	Leu	Ala	Val	Leu	Gln	Asn	Asn	Leu	Ser	Ala	Ala	Lys	Val
2326	TTA	GTT	AAG	AGT	AAT	AAA	AAG	GTA	AAA	TTA	AAC	GAA	ATG	GAT	AAT	AAC	GGT	ATG	ACG	CCT
621	Leu	Val	Lys	Ser	Asn	Lys	Lys	Val	Lys	Leu	Asn	Glu	Met	Asp	Asn	Asn	Gly	Met	Thr	Pro
2386	TTG	CAC	TAC	GCC	TCG	ATG	TTA	GGA	AAC	TTA	GAA	TTT	GTT	AAA	TAT	TTT	ACA	TCC	GAA	CAA
641	Leu	His	Tyr	Ala	Ser	Met	Leu	Gly	Asn	Leu	Glu	Phe	Val	Lys	Tyr	Phe	Thr	Ser	Glu	Gln
2446	GGC	ATC	GAC	GTT	AAC	GCA	AAA	ACA	AAA	GTT	AAA	AAC	TGG	ACC	CCA	TTA	CAT	CTA	GCA	ATT
661	Gly	Ile	Asp	Val	Asn	Ala	Lys	Thr	Lys	Val	Lys	Asn	Trp	Thr	Pro	Leu	His	Leu	Ala	Ile
2506	CTT	TTC	AAA	AAA	TTC	GAT	GTT	GCT	CAA	AGT	TTA	TTA	CAA	GTA	AGA	AAC	ATC	GAT	ATA	AGT
681	Leu	Phe	Lys	Lys	Phe	Asp	Val	Ala	Gln	Ser	Leu	Leu	Gln	Val	Arg	Asn	Ile	Asp	Ile	Ser
2566	ACA	CGA	GCA	GAT	CAG	GCT	ATA	ACA	CCG	TTA	CAT	TTA	GCT	GCG	GCA	ACT	GGA	AAT	TCA	CAG
701	Thr	Arg	Ala	Asp	Gln	Ala	Ile	Thr	Pro	Leu	His	Leu	Ala	Ala	Ala	Thr	Gly	Asn	Ser	Gln
2626	ATA	GTT	AAG	ACT	ATA	CTA	AAT	TCA	GGC	GCA	GTA	GTA	GAT	CAG	GAA	ACA	GCA	AAT	GGT	TTT
721	Ile	Val	Lys	Thr	Ile	Leu	Asn	Ser	Gly	Ala	Val	Val	Asp	Gln	Glu	Ala	Asn	Gly	Thr	Phe
2686	ACA	GCT	CTT	CAT	TTA	GCG	ATA	ATG	AAT	CCA	AAT	ACT	GAA	ACA	CCA	CAG	TTT	CTT	ATC	GCA
741	Thr	Ala	Leu	His	Leu	Ala	Ile	Met	Asn	Pro	Asn	Thr	Glu	Thr	Pro	Gln	Phe	Leu	Ile	Ala
2746	AAA	GGA	GCA	AAT	ATC	AAT	GCA	AAA	ACA	AAT	GAT	GGA	AGT	ACG	CCT	TTA	CAT	TTT	GCT	GCT
761	Lys	Gly	Ala	Asn	Ile	Asn	Ala	Lys	Thr	Asn	Asp	Gly	Ser	Thr	Pro	Leu	His	Phe	Ala	Ala
2806	GCA	TTA	GGC	AAA	ACC	AAC	ATT	TTC	CAG	TTA	CTT	ATG	GAC	AAA	GGA	GCA	AAT	ATA	AAA	GCT
781	Ala	Leu	Gly	Lys	Thr	Asn	Ile	Phe	Gln	Leu	Leu	Met	Asp	Lys	Gly	Ala	Asn	Ile	Lys	Ala
2866	GAA	AAT	TTA	ATT	AAT	CAA	ATG	CCT	ATT	CAT	GAA	GCC	GTT	GTG	AAT	GGG	CAC	CTG	GCA	ATT
801	Glu	Asn	Leu	Ile	Asn	Gln	Met	Pro	Ile	His	Glu	Ala	Val	Val	Asn	Gly	His	Leu	Ala	Ile
2926	GTC	AAA	ATG	CTG	ATT	GAG	CAA	GAT	TCT	TCT	CTT	ATG	AAT	GCG	AAA	AAT	ATG	AGG	GAT	GAA
821	Val	Lys	Met	Leu	Ile	Glu	Gln	Asp	Ser	Ser	Ser	Leu	Met	Asn	Ala	Lys	Asn	Met	Arg	Glu
2986	TAT	CCA	TTT	TAC	CTC	GCT	GCA	GAA	AAA	CGT	TAT	AAA	GAT	GTA	TTT	AAT	TAC	CTT	GAA	AGC
841	Tyr	Pro	Phe	Tyr	Leu	Ala	Ala	Glu	Lys	Arg	Tyr	Lys	Asp	Val	Phe	Asn	Tyr	Leu	Glu	Ser
3046	AAA	GGA	GCT	GAT	GTA	AAT	GAG	AAA	AAT	AAC	GAC	GGA	AAT	ACG	CTT	TTA	CAT	TTG	TTC	TCT
861	Lys	Gly	Ala	Asp	Val	Asn	Glu	Lys	Asn	Asn	Asp	Gly	Asn	Thr	Leu	Leu	His	Leu	Phe	Ser
3106	ATC	AAC	GGG	GAG	GTT	GAG	GTT	GTT	CAG	TTT	CTA	ATT	CAA	AAT	GGT	GCT	GAC	TTT	CGG	CTA
881	Ile	Val	Asn	Gly	Glu	Val	Glu	Val	Gln	Phe	Leu	Ile	Gln	Asn	Gly	Ala	Asn	Gly	Arg	Leu
3166	AGG	AAC	AAG	GAA	AGA	AAG	AGT	TTT	TTC	GAT	CTT	GCC	GTC	GAG	TTC	GGA	CAC	GCC	GGT	ATT
901	Arg	Asn	Lys	Glu	Arg	Lys	Ser	Phe	Phe	Asp	Leu	Ala	Val	Glu	Phe	Gly	His	Ala	Gly	Ile
3226	GTG	GGA	TAT	GCC	ATT	GAA	GAA	AAC	AAG	GTG	GAT	CTC	CAG	GAA	CCT	TAT	CGA	GGG	AAA	ACA
921	Val	Gly	Tyr	Ala	Ile	Glu	Glu	Asn	Lys	Val	Asp	Leu	Gln	Glu	Pro	Tyr	Arg	Gly	Lys	Thr
3286	ATC	CTA	TAT	CAT	GCT	ATT	TGT	GAT	TCT	GTG	AAA	TAC	GAC	AGG	ATA	GAA	GTA	GTG	AGG	TAT
941	Ile	Leu	Tyr	His	Ala	Ile	Cys	Asp	Ser	Val	Lys	Tyr	Asp	Arg	Ile	Glu	Val	Val	Arg	Tyr
3346	TTT	GTC	GAA	ACT	CTT	AAC	GAG	GAC	CAG	TGT	AGT	CCA	CTG	CAG	GAA	GCA	GCA	GCT	TAT	GCT
961	Phe	Val	Glu	Thr	Leu	Asn	Glu	Asp	Gln	Cys	Ser	Pro	Leu	Gln	Glu	Ala	Ala	Ala	Tyr	Ala
3406	CAT	TTA	GAT	TTA	GTG	AAA	TAC	TTT	GTT	CAG	GAG	AGA	GGA	ATA	AAC	CCC	ACT	GCA	TTC	AAT
981	His	Leu	Asp	Leu	Val	Lys	Tyr	Phe	Val	Gln	Glu	Arg	Gly	Ile	Asn	Pro	Thr	Ala	Phe	Asn
3466	AAT	GAC	AAT	CAA	GTG	TCT	CCC	CTA	TGC	ATT	GCA	ATA	GTG	GGT	GCA	CCA	TGT	AGG	TTT	GTA
1001	Asn	Asp	Asn	Gln	Val	Ser	Pro	Leu	Cys	Ile	Ala	Ile	Val	Gly	Ala	Pro	Cys	Gly	Phe	Val
3526	AAA	TCA	TGT	GAT	ACT	CCC	GAA	CGC	TTG	GAT	GTC	GTT	GAG	TAC	CTG	GTG	GAC	AAA	ACG	CCT
1021	Lys	Ser	Cys	Asp	Thr	Pro	Glu	Arg	Leu	Asp	Val	Val	Glu	Tyr	Leu	Val	Asp	Lys	Thr	Pro
3586	GAC	ATA	AAT	AAA	GAA	TGT	GAT	ACA	CAA	CAG	AGT	ACT	CCA	GTC	TCC	AGT	GCA	GTA	TAC	GGC
1041	Asp	Ile	Asn	Lys	Glu	Cys	Asp	Thr	Gln	Ser	Thr	Pro	Val	Ser	Ser	Ala	Val	Tyr	Gly	Gly
3646	AAT	AAA	GTT	TCG	ATT	TTG	AAT	TAT	TTA	ATA	CGA	AAT	GGA	GCT	GAT	CCC	AAT	AAA	AAA	GTT
1061	Asn	Lys	Val	Ser	Ile	Leu	Asn	Tyr	Leu	Ile	Arg	Asn	Gly	Ala	Asp	Pro	Asn	Lys	Lys	Val
3706	AGG	GGA	GAT	CCA	CCA	CTT	TTC	ATT	GCT	GCT	ATG	ATT	GGA	CAA	TAC	GAT	ATC	GTA	AAG	AGT
1081	Arg	Gly	Asp	Pro	Pro	Leu	Phe	Ile	Ala	Ala	Met	Ile	Gly	Gln	Tyr	Asp	Ile	Val	Lys	Ser
3766	TTA	GTA	GAG	CAG	CAT	AAG	ATT	GAC	GTT	AAT	ACA	AGA	AAC	AAA	GAG	CTT	ACA	CCA	TTA	Leu
1101	Leu	Val	Glu	Gln	His	Lys	Ile	Asp	Val	Asn	Thr	Arg	Asn	Lys	Glu	Gln	Phe	Thr	Pro	Leu
3826	CAT	GCA	GCA	GCT	AGT	AAC	GAT	CAC	ATC	GAT	GTG	GTG	AAA	TAC	TTG	ATA	CAA	AAG	GGA	GCC
1121	His	Ala	Ala	Ala	Ser	Asn	Asp	His	Ile	Asp	Val	Val	Lys	Tyr	Leu	Ile	Gln	Lys	Gly	Ala
3886	GAT	GTC	AAT	GCT	AAA	GGT	GAC	GAA	AAC	TTA	AAA	CCC	ATT	GAT	TTA	GCC	GGC	GAA	AAA	TCA
1141	Asp	Val	Asn	Ala	Lys	Gly	Asp	Glu	Asn	Leu	Lys	Pro	Ile	Asp	Leu	Ala	Gly	Tct	Lys	Ser
3946	AAA	GCC	TAC	TTA	AGA	TCA	CTT	GGC	AGG	CGT	TTC	TTC	CGT	AAT	GAG	AGC	CCC	GCT	AAA	TCA
1161	Lys	Ala	Tyr	Leu	Arg	Ser	Leu	Gly	Arg	Arg	Phe	Phe	Arg	Asn	Glu	Ser	Pro	Ser	Lys	Ser
4006	TTT	GAA	ATT	GAT	AAA	TTT	AAC	GCT	ATA	ATG	CCC	GAG	GTT	TCA	ATG	TCC	GGA	AAA	GTA	AGC
1181	Phe	Glu	Ile	Asp	Lys	Phe	Asn	Ala	Ile	Met	Pro	Glu	Val	Ser	Met	Ser	Gly	Lys	Val	Ser

Рис. 2. Продолжение.

4086	CAT	GAC	TCG	AAT	TTT	ATT	CAA	CAC	ATA	TCT	AGT	GGA	ACG	AGA	TCC	AAA	TCA	AAT	TTT	AAT
1201	His	Asp	Ser	Asn	Phe	Ile	Gln	His	Ile	Ser	Ser	Gly	Thr	Arg	Ser	Lys	Ser	Asn	Phe	Asn
4126	TCA	GCG	AAA	AAC	AAA	ATG	TAT	GCA	GAG	AAC	TCT	CAC	GTT	AGA	AGC	ATT	GAT	GTC	AAT	GGA
1221	Ser	Ala	Lys	Asn	Lys	Met	Tyr	Ala	Glu	Asn	Ser	His	Val	Arg	Ser	Ile	Asp	Val	Asn	Gly
4186	GCT	CTT	CTT	TTG	TTG	GAT	TTT	ATG	GTA	AGA	GTT	TTC	AGT	AAT	CGA	AAA	ATG	AAT	TAT	GCT
1241	Ala	Leu	Leu	Leu	Leu	Asp	Phe	Met	Val	Arg	Val	Phe	Ser	Asn	Arg	Lys	Met	Asn	Tyr	Ala
4246	GCA	TCG	ATT	TCT	GGC	ATA	AAG	TCT	CGT	TCA	AAT	TCA	GAA	GCG	CAA	GCT	GAG	GCT	TTG	ATA
1261	Ala	Ser	Ile	Ser	Gly	Ile	Lys	Ser	Arg	Ser	Asn	Ser	Glu	Ala	Gln	Ala	Glu	Ala	Leu	Ile
4306	CTA	ACG	GAA	AGA	TTC	GAA	CAT	CTA	TTG	AAT	GCT	TTG	ATT	GCT	GAC	CAG	AGT	ATC	GAT	TCT
1281	Leu	Thr	Glu	Arg	Phe	Glu	His	Leu	Leu	Asn	Ala	Leu	Ile	Ala	Asp	Gln	Ser	Ile	Asp	Ser
4366	TTA	GAT	TTT	TCA	AAC	GTC	CAT	TCA	AGA	ATA	TAC	AAA	GCT	ATT	ATA	AAT	GGT	AAT	CCA	AAT
1301	Leu	Asp	Phe	Ser	Asn	Val	His	Ser	Arg	Ile	Tyr	Lys	Ala	Ile	Ile	Asn	Gly	Asn	Pro	Asn
4426	GGA	ATT	TCG	GAA	ATG	TTA	TGT	TCT	TAT	GCA	AAA	GAG	TAT	TCT	GAA	TTG	GAT	CCT	GAA	AAA
1321	Gly	Ile	Ser	Glu	Met	Leu	Cys	Ser	Tyr	Ala	Lys	Glu	Tyr	Ser	Glu	Leu	Asp	Pro	Glu	Lys
4486	ATT	GAA	AAA	CTT	CTG	CAA	GAA	TTC	GAG	ACT	TTA	ACT	TTT	ACT	AAG	TCA	TCC	GAA	ATT	CAG
1341	Ile	Glu	Lys	Leu	Leu	Gln	Glu	Phe	Glu	Thr	Leu	Thr	Phe	Thr	Lys	Ser	Ser	Glu	Ile	Gln
4546	ATC	AAT	GAA	AAA	TTC	TCT	CAC	GCT	TTG	TTT	GAA	ACA	TGT	GGA	TTG	AAT	AGA	CCC	ACA	AAC
1361	Ile	Asn	Glu	Lys	Phe	Ser	His	Ala	Leu	Phe	Glu	Thr	Cys	Gly	Leu	Asn	Arg	Pro	Thr	Asn
4606	GTT	TTA	CAA	ATT	AAA	TAA	ACA	AGA	GAT	CGA	AAT	AGA	CAA	TTG	TTG	AAG	AAA	ACT	GAT	TTA
1381	Val	Leu	Gln	Ile	Lys															
4666	ATC	TCC	ACA	TAA	TAA	CTT	AAT	TGA	TTT											

Рис. 2. Окончание.

полипептидный фрагмент длиной около 200 а.о. Тот факт, что идентифицированные химотриптические пептиды α -LCT были найдены в полипептидном фрагменте 1–1133 (рис. 2), но не далее, ко-свенно подтверждает это предположение. Размер отщепляемого С-концевого фрагмента в случае крустотоксина, по нашим оценкам, также составляет около 200 а.о.

Выравнивание аминокислотных последовательностей зрелых белков α -LCT с α -LIT, α -LTX и δ -LIT выявило у них наличие участков структурной гомологии по всей длине полипептидной цепи (рис. 3). Первичные структуры α -LCT и α -LIT [10] обладают наибольшей гомологией, достигающей 62%. Остатки цистеина в молекулах α -LCT и α -LIT занимают практически идентичные позиции, а наиболее заметные отличия в структуре наблюдаются в областях 440–463, 549–559, 689–706 и 997–1022.

Значительная гомология (39%) была выявлена в первичных структурах α -LCT и δ -LIT [11].

Уровень структурной гомологии для аминокислотных последовательностей α -LCT и α -LTX [9] не превышал 26%. Максимальное различие двух токсинов было выявлено в области 997–1022, содержащей большое число остатков цистеина. Эти данные подтверждают предположение о том, что фрагменты латротоксинов с большим числом остатков цистеина, по-видимому, обуславливают специфичность их действия, принимая участие в связывании с соответствующим пресинаптическим рецептором. В то же время данные фрагменты у токсинов α -LCT и α -LIT не характеризуются столь заметными структурными различиями. Вероятно, латротоксины обладают многоточечными участками связывания со своими рецепторами и их таксоспецифичность обеспечивается путем

точечных замен аминокислотных остатков в этих участках.

Ранее, при изучении аминокислотной последовательности других латротоксинов в полипептидной цепи этих белков был выявлен большой район, полностью состоящий из целого ряда несовершенных повторов – так называемых анкириновых [9–11]. В первом приближении они представляют собой фрагменты, состоящие из семи аминокислотных остатков с консенсусной последовательностью TPLN(L/I)A(A/I). Эти повторы встречаются в структуре латротоксинов до 20 раз. В структуре α -LCT мы также выявили 20 повторов анкиринового типа (рис. 3). Как было отмечено ранее [10], эти повторы можно рассматривать как 33- или 34-звенные участки, имеющие консервативную N-концевую часть и вариабельную С-концевую. Анкириновые повторы у α -LCT занимают область от 464 до 1182 а.о.

Из результатов проведенной нами работы следует, что структурная часть гена α -крустотоксина не содержит интронов. Действительно, наличие интрона(ов) внутри структурной части гена с большой вероятностью привело бы к разрыву (укорочению) рамки считывания вследствие появления терминирующего кодона. С другой стороны, процедура выравнивания аминокислотных последовательностей α -LCT и α -LIT, имеющих близкие размеры и значительную гомологию, позволила бы выявить интроны (т. е. вставочные ДНК) даже небольшого размера, которые не нарушают рамку считывания. Таковых выявлено не было.

Отсутствие интронов в достаточно большом структурном эукариотическом гене явилось для нас несколько неожиданным фактом, поскольку,

α-LCT	-1	VSIFIFHSANILVRNSEMKGKRVISKR
α-LIT	-1	ACSSPEVSIHFHFFVYAGSFVKNFKKMKGSSAISKR
α-LTX	-1	RLFSFLTWARDVYVYLYLHLRVSNMISVGEIMERANHSLVRMR
δ-LIT	-1	MHSELQTIISAARVAVKAVPNTWVIRLKR
α-LCT	88	EMS : EMS-----KADQCTESYQSVAYGTLGCVAGVSSII---EGADIVATPIAAGCHEAKGAIDAAAMIADGSSISFEDBEIKQEQENQRFNEWDKKLQKGA :
α-LIT	88	EMS : EMS-----RADQCKLAAAYTAGVETVGNWAADIASI---EGANLVAAPVAAGGHGKGLTAAAMIADGSSISFPEIKEIINKEFKEMGRKLDKNT :
α-LTX	98	EGED--LTLFEKAFICSEFLQOKYVDIASNIGDLSPIVIGKIAGTIAAAMATAHWASGRIDIEQTLGCSDLFPDQIKVEVENRFNEEDRKLDSHS :
δ-LIT	100	DEEDGEMTLERQAQCKAEYSNSVFGMIADVANDHGSIPVIGEVVGIPTAIVSHTSAGLIIASTALDODDIPEBEIKEIIEFERFNEIDRKLDRNT :
α-LCT	188	EALENVTELAARKYSSVERKURVEMREGEFNVHVFATIENTANKQIITCTENQIIQYFENDERENINNRQKEDVAKIQEPASGNFLLYLIRKSRSTSEDCSLSHSLI :
α-LIT	188	EALRHVSKLVSKLSTVEKIRVEMREGEKLVETIENIATKEIVFINKIVQFNNERENINSRQKEEFAKIQEPAPGNFLLYLIRNSRSTSESGTLYSLI :
α-LTX	196	AALEETKLVKESISVVEKTRKQMKRFRDEVMKSIQDAKVSPIESKINNFARVEDTEKERIRGLKENDYILKIEEPN--GILLHFRPESRTPPTDSDLQAPL :
δ-LIT	198	AALEEVSKLVSKLFPVTVEKIRNEMNENFKLVETIESKEIKSIVFKINDFKFEKERQIKGLPKDRYVAKILEQK--GILGSLKEVREPESGNSLSSAL :
α-LCT	286	FKIIN--QELALPNAADNNAIRALFAIFYGQTFISIMFYVAKQYSIADHYHQNGNLAEFNSFQHMKTVDQDFKFTIIGINTSNKPIANVTLSIIE :
α-LIT	285	FRIID--QELALPNNAGNNAIQALYALFYGTFIFISIMFYVAKQYSIADHYVQKGNLSEFNTFQHMKIYQDQKFSHIGINGQ--NPKPLYDEIVLWLN :
α-LTX	294	FSIIE--EGYAVPKSIDDELAFKVLVYALLYGTQYVSVMPFELIQYSFLANHYVEKGYLKYDEYFNSLNWFLDFKSSIVGTGTNNEGILDRVLQVEM :
δ-LIT	294	NELLDKNNYALPKVVDNKAFOALYALFYGTQYVAAVMPFELIQHSYLAQYVYQKGDVNFNAEENVAIIFDFDKSSITGG---DDGILIDNVIEVLN :
α-LCT	386	DVKNRFIRNLSNLYQKIKLTKSTLLDIRKTKMDLPIEDTPKSSVLINERKSSVSPRIETELKWTPTVTVKVAQYEQDQKYSKISKNSNPIITVY :
α-LIT	384	NVKNKSIRNVQNKLFYDLMKQTESLLELKKELANMELPIIDETPLSISISPKERSDDKP--VDPQLKWDKQKVKVAIQEFDQDKFSKISSNSKPVTV :
α-LTX	388	TVKNSEFLGLEKNGVDEMEKINLNFNKKIEEFGKQKUTLSETPENPAQISE--DKDIT--FPEGDWRDQREVRVAVQVASETLFSKISHMSDPSVY :
δ-LIT	391	TVKALPEIKNADSKLYREAVTRTKALETLKNGIKTKTDLPLDDIPELTSQVNEPNDENQLP--FPEIGNVDPGEVRYAVQYVESKGYSKFSEWSEPTVY :
α-LCT	483	QRLAENYI--TIDKDRREPLVFRQNGNEKPELLIISDSSQNEFRDHRDLVNAQMPYKETALGICRKEIIDSGAQVGAQSEMGRKSIHASATAGNDDVA--
α-LIT	481	QHLACPI--SVBKDRRNPLIFRQFGDQIPELVGTLRGQVEFRDHRDLVNAQVVPYAREALSISRTLIQNCANVSEFELGGAHAAASAGNYDVG--
α-LTX	488	REKACPTLRMEVDQTRRNVEVFRKFDSSKQQLVGEITPVLSNEIDEDRDLVNAASNPDSAVGKFTKELYDGANIRALFDHGRVTFVHAAAKSGNDKIMF :
δ-LIT	489	QGNACPTKVRVBPKRNPLIFRKFNKSGKQFAGTMTHSQTNFKPIHRDLVDAALNINKLKAQVDEATTEIEKGADEALFDNDRSAMHAVAVRGNKKI--
α-LCT	581	--RELLAKNGLINVPDKNGYTPRHEASERKNNDPVKFELEKGDVNVRTFANELTPRHEAARQDFTIIVKILMEKRGDYNKARAGFTTPRHLSTSNIS :
α-LIT	579	--EILLNKDINIIEKADKNGYTPRHEAASNKNDVFMELIGNADVVRTKSDLFTPRHEAARRDLTDVQTQLIDITEDJNAQDKSCGFTPEHLISISSTS :
α-LTX	586	GLTFLAK--STEENQPDKYGTPRHYAADSGNAGIVNLIQRGVSINSKTYHFLOTPRHEAARQGFVTFQRIMESPEINERDRDQGFTRHYAIRGCE :
δ-LIT	589	ALRFLKNQSIDFLKDKNGFTRHEAALAEAGQAGVKLINHGADVNAKTSKTNLTPRHEATRSCFSKTVRMLLESPNIKVNEKEKEDGFTPEHTAVMSTY :
α-LCT	680	RAARTLNETPACINIKNSGELTPRHEAVLQNNLSAAKVAVKSNKVKKINENDMGTPRHYASMTGNEFFVYFTS--EQCIDVNAKTKVKVNTPLHLAI :
α-LIT	676	ETAAILFRNTNAVINIKSVGLTPRHEATLQNNLSVSKLIF--AGKAYLNDGANMGTPRHYAYMTGNIEYVDFLLN--QYINIVAAKREKKVTPRHLAI :
α-LTX	684	RILEAFINQISIDVNAKSNGLTPRHEAIKNDWPVASLIL--GSKKVDINAVBENNTLHYAAIIGYLETTKQLINLKE--INAVYSSPGLLSALHYAI :
δ-LIT	687	MVVDALENHRPIDKNAQSTSLTPRHEALINESQEVAESV--ESN--ADENIQVNNHMAPRHFASMSGSIKMLRYLISIKDKVYSINSVTENNNTPLHFAL :

Рис. 3. Сравнение выведенной аминокислотной последовательности α-LCT с последовательностями α-LIT [10], α-LTX [9] и δ-LIT [11]. Гидрофобные участки взяты в рамку. Стрелки R1–R22 указывают начало каждого из 22 несовершенных анкириновых повторов. Заштрихованы консервативные (гомологичные) остатки во всех 4 лаптогокинах.

α-LCT :	LFKKDFVQSLQVNRNIDISTRADQAITPPIHAAATGNSQVKTILNSGAVVDQEFANGFTALHHAIMNPWTETPQFLIAKAGINAKTNDGSTRPLHFAA :	780
α-LIT :	LFKKNDBAERLLSDEMNLRLNETNGGINPHELASATGNKQVLLELLAKNADVTRLTKGFSALHLGIIKNEEIPFIVEKGANVNDKNSGVTPPLHFAA :	776
α-LTX :	LYKHDBYASFLMRSGNWNALKALGGITPPIHLAVIQGRKQILSLMFDIGVNIQEKDEKYPPIHLAAMSYPPELIQIILDQGSFPAKTSNGATPLHLAT :	784
δ-LIT :	YFKKEDAKELKQDPINITIVADGNLTVVHLAVSTQIINIIEKLLKRGSNIEEKQEGYTSIHHAAAMRKEPPIAVVVEIENGADLEARSADNLPPLHSAAS : ↳ R9 ↳ R10 ↳ R11	787
α-LCT :	ALGKTNIFQLIMDKGANIKAEENLINOQMPIHEAVVNGHLAIVKMLIEQDSSLWNAKMRDEYPIYAAEKRYKDVFNYLESKGADVNEKNDNGNTLLHLFS :	880
α-LIT :	GLGKANIPRLLSRGADIKAEEDINSQMPIHEAVSNGHLEIVRIILIEKDPSLMVKNIERNIEYPIYAVEKRYKIDFDYFVSKDAWNEVDHNGNTLLHLFS :	876
α-LTX :	FKKKSQAALILLANNEVMWRTDENGQMPHGAAMTGLLDVAQAIISIDATVVVIEDKNSDTPNLAAQNSHIDVIKVPIDQGADINTRNKKGLAPLAFPS :	884
δ-LIT :	KIGRKSTVLYLLEKAGAD : ↳ R12 ↳ R13 ↳ R14	853
α-LCT :	INGEVEVYQFLIQNGADFRLNKPKSFPDLAVEFGHAGVGYAT-EENKVDLQEPYRGK-----TILYHAICDSVKYDRJEVVRVYFVETLNE-DQ :	969
α-LIT :	STGELEVYQFLMQNGANFRLNKNERKTFPDLAIENGRNLNIVAFV-EKNKVNLOAAHRGK-----TILYHAICDSAKYDKFEIVKYFEKLN-SE :	965
α-LTX :	KKGNLDMVKYLFDKNANVYIADNDGMNPFYAVQNGHLNLVKYAMSEKDKFEWSNTDNRNRRDECPNEECAISHFAVCDVQFDRFEIVKYFYGTLGNFAI :	984
δ-LIT :	INDDLDMVRLFLEKDPFLKDDTEE-----GRTSIMLIVQKLE : ↳ R15 ↳ R16 ↳ R17	904
α-LCT :	CSPLOEAAAYAHILVVKYFVQBERGINPTAFNNDNQVSPLCIALVGAFCGVKSCDTPERLDVVEYISVDKTPDINKECTQQSTPYSSAMYGKNKVISILNYL :	1069
α-LIT :	CNPLHEAAAYAHLLVVKYFVQBERGINPAEFNEENQASPFCTIHGAPCGYSLDCDPPDRLEVEYISDKIPDINGKCDVQENTPTVAIFANKVSIILNYL :	1065
α-LTX :	CGPLHQARYGHLDIVKYLVEBEFLSVDGSKTD--TPLCYA-----SENGHFTVVOXLVSNAGAKVNHDCG-NGMTAIDKAITKNHLQVVOQFL :	1068
δ-LIT :	TLDEEALFNRLD : ↳ R18 ↳ R19	949
α-LCT :	IRNGAD-PNKKVVRGDPPLFIAMIGQYDIVKSLVEQHKIDVNRNK--EQFTPLHAAASNDHIDVVKYLQKQADVNAKGDENLKPIDIAGEKS----KA :	1162
α-LIT :	VGIGAD-PNQVVDGDPPLYIAARQGRFEIVRCLIEVHKVDIVNRNK--ERTALHAAARNDFMDVVKYLVFQGDVNAKAGIDDLRPIDIAGEKA----KA :	1158
α-LTX :	AANGVDFRRKNSRGTPFLTAVAENALDIAEYLIREKQDINENEQNVQDPTALHLAVVYKQLQIKLLIKYGDVTRNAYDKTALDIAIDAKFSNIVE :	1168
δ-LIT :	-----EYCLKLEESGAPGSGFSDSSKQCI STFSEDEM-----FRRLPELIVKETEIN----SR : ↳ R20 ↳ R21 ↳ R22	999
α-LCT :	YERSLGRFPFHNESPKSFEIDKFNAIMPEVSMGKVSHDNSNFIQHISS-GTRSKSNFNSAKNKMYAENSHVRSIDVNGALLILDFMVRVFSNRKMYAA :	1261
α-LIT :	YEQS--SRPLRSGHFSQNEIDSGNTIHGISMSARTND--KLTQQISSKGRSDSN--STEGKMHSENVHVSIDVNGALLILDFMIRVFSKKTNPAP :	1252
α-LTX :	YLKTKSGXPRRE-----YKSSYGERSLL-QTNQISNFI DRKNIEHDHP-LFINADNESSELF SXTASNIDVTGLLILIVLIRYFSKQG--VI- :	1252
δ-LIT :	YELPLKG--FSRS-----LNKFLPSLKTAEKSN-SY-----RSENF-----VSNIDNSGALLILDFVIRKFTNEK--YN- :	1057
α-LCT :	SISGIKRSNSEAQAEAILTEFEHEHINAIADQSIDSDNFVHSRVTYKAIINGNPNGTSEMUCSYAKEXSELDPEKTEKLEKQEFET-----LTFYK :	1355
α-LIT :	YGSRIKTRSAAEAQAEALIMTEFEHENLISGLIGDPIDSDNFVHSKIYKAIMSGRRSVISELCSFAEEYSKINHESTKQLSEFET-----LTTTK :	1346
α-LTX :	-SKESDSADGITQAAALSITEKPEDVINSLNHESAKEQVBLAEPVHGKVAALKSGRNSQIHQICCSLSNISTLKPEDMEKLESYIMNSSHVSLSPEVT :	1351
δ-LIT :	-LTGKEAVPYLEAKASSLRIASKFEELLETEVKGIPAGELINMAEVSSNHHKATASGK--PVSKVICSYLDTFSELSQOMFEVNTYLSLK-----PSVI :	1149
α-LCT :	SSEIQUINEFESHALFETC-GLNRPNTVLOI-----K :	1385
α-LIT :	ASEIHIIEESVYPAPFETC-ELKVNMSVSO-----K :	1376
α-LTX :	DSANEAYGTLHLFGESCLHSD-----GELTKKLM- :	1381
δ-LIT :	TSASADYQXLPNLLTATCLEPERMAQLIDVHQKMLR : 1186	

Рис. 3. Окончание.

как известно [15–20], даже относительно небольшие гены, кодирующие белковые токсины у змей и скорпионов, содержат интроны. У змей гены токсинов состоят из трех экзонов и двух интронов. При этом один интрон встроен в область, соответствующую сигнальному пептиду, второй – находится в пределах собственно структурной части гена [15–17]. У скорпионов гены белковых токсинов содержат по одному интрону, встроеному в участки ДНК, кодирующие сигнальные пептиды [18–20]. В литературе практически отсутствуют данные о структуре генов, кодирующих синтез белковых токсинов у других представителей паукообразных. Поэтому вопрос о типе строения генов белковых токсинов у пауков других видов – безинтронном или интрон-экзонном – остается открытым. Что же касается генов, кодирующих другие высокомолекулярные нейротоксины у каракурта, с большой вероятностью можно предположить, что они также не содержат интронов. Это обстоятельство может сильно облегчить клонирование и изучение структуры соответствующих генов и кодируемых ими белков. В настоящее время мы начали работу по изучению структуры хромосомных генов каракурта, кодирующих α -LIT, δ -LIT и α -LTX.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

В работе использовали: акриламид, Трис, бисакриламид, TEMED (Merck, Германия); агарозу, EDTA, меркаптоэтанол, бромистый этидий, краситель бриллиантовый синий (Sigma, США); Триптон, дрожжевой экстракт (Difco, Англия); персульфат аммония (Serva, Германия). Эндонуклеазы рестрикции *DraI*, *SmaI*, *PvuII*, *HpaI*, *HincII* и *EcoRV* (Fermentas, Литва). Фосфорилаза CIAP (Amersham, США). ДНК-полимераза фага T7 (sequenase 2.0), дезоксинуклеозидтрифосфаты (dNTP) и дидезоксинуклеозидтрифосфаты (ddNTP) (US Biochemicals, США). Термостабильные ДНК-полимеразы *KlenTaq* и *Pfu*, а также полинуклеотидкиназа фага T4 и ДНК-лигаза фага T4 были любезно предоставлены В.М. Крамаровым (лаб. биотехнологии ИБХ РАН).

Структура олигонуклеотидов, использованных в качестве адаптеров и праймеров для ПЦР, представлена в таблице.

Выделение геномной и плазмидных ДНК. В тех случаях, когда не указано особо, использовали общепринятые методики [21].

Геномную ДНК выделяли из замороженных (-70°C) желез паука, используя методику, которая является модификацией гуанидинтиоцианатного метода выделения РНК [22]. Железы паука (~50 шт.) суспендировали в буфере (4 М гуанидинтиоцианат, 25 мМ Na-цитрат, 0.1 М меркаптоэта-

нол, 0.5% саркозил) и гомогенизировали путем многократного пипетирования. К гомогенизированной смеси добавляли равный объем фенола, насыщенного 0.1 М Трис-НСl, pH 8.0, энергично встряхивали, затем добавляли 1/4 объема смеси хлороформа с изоамиловым спиртом (24 : 1) и снова встряхивали. Смесь центрифугировали и отбирали водную фазу, из которой нуклеиновые кислоты осаждали 3 объемами этанола. Осадок промывали 70% этанолом и растворяли в буфере TE (10 мМ Трис-НСl, pH 8.0, 1 мМ ЭДТА). К полученному раствору добавляли два объема 9 М LiCl₂, выдерживали 2 ч при -20°C и центрифугировали. ДНК осаждали из раствора этанолом и после промывки 70% этанолом растворяли в буфере TE.

Плазмидные ДНК для секвенирования выделяли методом щелочной денатурации [23] и очищали на колонках Wizard (Promega).

ПЦР-амплификация фрагментов гена *crusta*. Отдельные перекрывающиеся фрагменты ДНК гена *crusta* получали с помощью ПЦР согласно ранее описанной процедуре [13]. Для этого геномную ДНК подвергали полному расщеплению рестриктазами *SmaI*, *DraI*, *PvuII* и *HpaI* по отдельности (100 нг ДНК на реакцию). Полученные фрагменты ДНК лигировали с олигонуклеотидным дуплексом AP1/AP2 (адаптер) и далее использовали этот образец ДНК в ПЦР-амплификации как описано ниже. Во всех случаях в ПЦР использовали смесь *KlenTaq*-полимеразы и *Pfu*-полимеразы (соотношение активностей 50 : 1). В первом раунде амплификации использовали адаптерный праймер AP3 и генспецифические праймеры P3 (при получении 5'-концевых фрагментов гена *crusta*) и P4 (при получении 3'-концевых фрагментов гена). Условия первого раунда ПЦР: 35–40 циклов; денатурация – 92°C , 30 с; отжиг – 64°C , 50 с; синтез – 72°C , 3 мин. Амплификацию проводили в 50 мкл смеси, содержащей 10 нг двухцепочечной ДНК, по 10 пмоль каждого праймера и буфер (50 мМ Трис-НСl, pH 9.2; 16 мМ (NH₄)₂SO₄; 3 мМ MgCl₂ и 250 мкг/мл БСА).

Полученные в первом раунде амплификации смеси разводили в 100 раз и по 1–2 мкл из разведенной смеси брали для второго раунда ПЦР, в которой использовали адаптерный праймер AP4 и внутренние (гнездовые) генспецифические праймеры P2 и P5, соответственно. Условия второго раунда амплификации были те же, за исключением того, что число циклов составляло 20–25.

Продукты ПЦР анализировали электрофорезом в 1.0–2.0%-ном агарозном геле с применением стандартного Трис-ацетатного буфера.

Фрагменты ДНК, полученные в результате второго раунда ПЦР, очищали методом элюции

из агарозного геля, фосфорилировали полинуклеотидкиназой фага T4 и клонировали в векторную плазмиду pBSIIsk(+) (Stratagene), линейризованную рестриктазой *EcoRV* и дефосфорилированную с помощью щелочной фосфатазы из поджелудочной железы теленка (CIAP). Длинные фрагменты вставочных ДНК у полученных рекомбинантных плазмид подвергали субклонированию, используя рестриктазы *DraI*, *PvuII*, *HincII* и их сочетания.

Однораундовую ПЦР-амплификацию фрагментов геномной ДНК с использованием праймеров P1 и P6, а также P2 и PND проводили в тех же условиях, что и двухраундовую; число циклов составляло 35.

Электротрансформация бактерий. Рекомбинантные ДНК вводили в клетки реципиентного штамма *E. coli* XL-1 Blue (Stratagene) посредством электротрансформации [24, 25]. Селекцию трансформантов проводили на чашках с агаризованной средой LA [21] в присутствии 50 мкг/мл ампициллина.

Секвенирование ДНК. Двухцепочечные ДНК подвергали щелочной денатурации и секвенировали дидезокситерминационным методом [26], используя ДНК-полимеразу фага T7 и [α -³²P]dATP (ИБХ РАН) (первичный анализ рекомбинантных плазмид). Матрицы, соответствующие по своей структуре искомым, (содержащие 5' или 3'-концевые последовательности внутреннего фрагмента гена *crusta* [12]) секвенировали далее на автоматическом секвенаторе (Model 373A, Applied Biosystems, США). Полная структура гена была выведена из последовательностей перекрывающихся фрагментов. Трансляцию нуклеотидной последовательности гена, построение рестриктных карт и сравнение аминокислотных последовательностей (выравнивание) осуществляли с помощью набора компьютерных программ DNASTAR (США).

Авторы выражают благодарность Н.С. Быстрову и В.К. Потапову за синтез олигонуклеотидов, А.В. Липкину и С.А. Козлову за помощь при компьютерной обработке данных и изготовление рисунков, Д.А. Грядуну за установку компьютерной программы Executor и помощь в обработке хроматограмм секвенирования ДНК.

СПИСОК ЛИТЕРАТУРЫ

1. Grasso A. // Biochim. Biophys. Acta. 1976. V. 439. P. 409–412.
2. Ковалевская Г.И., Пашков В.Н., Булгаков О.В., Федорова И.М., Магазаник Л.Г., Гришин Е.В. // Биоорган. химия. 1990. Т. 16. С. 1013–1017.
3. Красноперов В.Г., Шамотиенко О.Г., Гришин Е.В. // Биоорган. химия. 1990. Т. 16. С. 1138–1140.
4. Krasnoperov V.G., Shamotienko O.G., Grishin E.V. // J. Natural Toxins. 1992. V. 1. P. 17–23.
5. Knipper M., Madedu L., Breer Y., Meldolesi J. // Neuroscience. 1986. V. 19. P. 55–62.
6. Магазаник Л.Г., Федорова И.М., Антонов С.М., Ковалевская Г.И., Булгаков О.В., Пашков В.Н., Гришин Е.В. // Биол. мембраны. 1990. Т. 7. С. 660–661.
7. Красноперов В.Г., Шамотиенко О.Г., Гришин Е.В. // Биоорган. химия. 1990. Т. 16. С. 1567–1569.
8. Красноперов В.Г., Шамотиенко О.Г., Гришин Е.В. // Биоорган. химия. 1991. Т. 17. С. 716–718.
9. Kiyatkin N.I., Dulubova I.E., Chekhovskaya I.A., Grishin E.V. // FEBS Lett. 1990. V. 270. P. 127–131.
10. Kiyatkin N., Dulubova I., Grishin E. // Eur. J. Biochem. 1993. V. 213. P. 121–127.
11. Dulubova I.E., Krasnoperov V.G., Khvotchev M.V., Pluzhnikov K.A., Volkova T.M., Grishin E.V., Vais H., Bell D.R., Usherwood P.N. // J. Biol. Chem. 1996. V. 271. P. 7535–7543.
12. Волынский К.Е., Волкова Т.М., Галкина Т.Г., Красноперов В.Г., Плужников К.А., Хвоцев М.В., Гришин Е.В. // Биоорган. химия. 1999. Т. 25. С. 25–30.
13. Siebert P.D., Chenchik A., Kellog D.E., Lukyanov K.A., Lukyanov S.A. // Nucl. Acids Res. 1995. V. 23. P. 1087–1088.
14. Лукьянов С.А., Гурская Н.Г., Лукьянов К.А., Тарабыкин В.С., Свердлов Е.Д. // Биоорган. химия. 1994. Т. 20. С. 701–704.
15. Fuse N., Suchiya T., Nonomura Y., Menez A., Tamiya T. // Eur. J. Biochem. 1990. V. 193. P. 629–633.
16. Chang L.S., Lin J., Chou Y.C., Hong E. // Biochem. Biophys. Res. Commun. 1997. V. 239. P. 756–762.
17. Chang L.S., Chou Y.C., Lin S.R., Wu B.N., Lin J., Hong E., Sun J.Y., Hsiao C.D. // J. Biochem. (Tokyo). 1997. V. 122. P. 1252–1259.
18. Delabre M.L., Pasero P., Marilley M., Bougis P.E. // Biochemistry. 1995. V. 34. P. 6729–6736.
19. Legros C., Bougis P.E., Martin-Eauclaire M.F. // FEBS Lett. 1997. V. 402. P. 45–49.
20. Xiong Y.M., Ling M.N., Wang D.C., Chi C.W. // Toxicol. 1997. V. 35. P. 1025–1031.
21. Maniatis T., Fritsch E.F., Sambrook J. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, N.-Y.: Cold Spring Harbor Lab. Press, 1989.
22. Chirgwin J.M., Przybyla A.E., MacDonald R.J., Rutter W.J. // Biochemistry. 1979. V. 18. P. 5294–5299.
23. Birnboim H. C., Doly J. // Nucl. Acids Res. 1979. V. 7. P. 1531–1532.
24. Hanahan D.J. // Mol. Biol. 1983. V. 166. P. 557–580.
25. Dover W.J., Miller J.F., Ragsdale C.W. // Nucl. Acids Res. 1988. V. 16. P. 6127–6145.
26. Tabor S., Richardson C.C. // Proc. Natl. Acad. Sci. USA. 1987. V. 84. P. 47–67.

Cloning and Structure Determination of the α -Latrocrustoxin Gene from the Black Widow Spider Venom

V. N. Danilevich[#], S. A. Lukyanov, and E. V. Grishin

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

The primary structure of the *crusta* gene encoding α -latrocrustoxin (α -LCT), a high molecular mass neurotoxin specific to crustaceans, was determined in the black widow spider *Latrodectus mactans tredecimguttatus* genome. The total length of the sequenced DNA was 4693 bp. The structural part of the black widow spider chromosome gene encoding α -LCT does not contain introns. The sequenced DNA contains a single extended open reading frame (4185 bp) and encodes a protein precursor of α -LCT, comprising 1395 aa. We assume the Met residue at position -10 relative to the *N*-terminal residue of Glu1 of the mature toxin to be the first one in the protein precursor. The calculated molecular mass of the precursor (156 147 Da) exceeds that of the mature toxin by ~30 kDa. These data are in agreement with the notion that over the course of maturation the protein precursor undergoes double processing—cleavage of a decapeptide from the *N*-terminal part and of a ~200-aa fragment from the *C*-terminal part. α -LCT displayed a number of imperfect ankyrin-like repeats and areas of structural homology with earlier studied latrotoxins; the highest homology degree (62%) was revealed with α -latroinsectotoxin (α -LIT).

Key words: neurotoxin, Latrodectus, genomic DNA, PCR amplification, nucleotide sequence, deduced amino acid sequence

[#] To whom correspondence should be addressed; phone: +7 (095) 336-6540; fax: +7 (095) 330-7301;
e-mail: dan@ibch.siohc.ras.ru.