



УДК 577.214.622

АЛЬТЕРНАТИВНЫЙ СПЛАЙСИНГ пре-мРНК, КОДИРУЮЩЕЙ ЛАТРОФИЛИНПОДОБНЫЙ БЕЛОК (LLP) У *Musca domestica*. ПЕРВИЧНАЯ СТРУКТУРА ЧЕТЫРЕХ СПЛАЙСИРОВАННЫХ ФОРМ мРНК И ИХ БЕЛКОВЫХ ПРОДУКТОВ

© 2005 г. Я. А. Андреев, В. Н. Данилевич[#], Е. В. Гришин

Институт биоорганической химии им. акад. М.М. Шемякина и Ю.А. Овчинникова РАН,
117997 ГСП, Москва, В-437, ул. Миклухо-Маклая, 16/10

Поступила в редакцию 29.07.2004 г. Принята к печати 03.09.2004 г.

Используя вырожденные праймеры к консервативным участкам генов латрофилинподобных белков (LLP), с помощью ПЦР получен внутренний фрагмент ДНК (~2000 п. о.) гомологичного им гена *Musca domestica*. На основании данных секвенирования выделенного фрагмента осуществлен синтез генспецифичных праймеров и в ходе быстрой амплификации 5'- и 3'-концевых последовательностей кДНК (5'- и 3'-RACE) получены все перекрывающиеся фрагменты кДНК гена *llp*, кодирующего LLP домовой мухи. В процессе секвенирования полученных фрагментов кДНК обнаружены четыре альтернативно сплайсированные мРНК. Две длинные мРНК (~6000 нт) отличаются по участкам, кодирующими сигнальные пептиды, и по 5'-концевым нетранслируемым областям. Они кодируют белки большого размера (~1800 а. о.), схожие по доменной организации с латрофилинами млекопитающих. В каждом из них присутствует домен с семью трансмембранными участками, за которым следует протяженный цитоплазматический C-концевой домен. Две другие формы мРНК являются производными двух длинных мРНК; они кодируют сильно укороченные в C-концевой части белки (~900 а. о.). В их состав входит только три трансмембранных участка и короткий уникальный цитоплазматический C-концевой домен (23 а. о.). В работе анализируются ограничения (недостатки) существующих технологий 3'-RACE, выявленные в ходе работы с кДНК большого размера (и в случае альтернативного сплайсинга), и предлагаются пути их преодоления.

Ключевые слова: альтернативный сплайсинг; 3'- и 5'-RACE; ПЦР-амплификация; латрофилинподобный белок; *Musca domestica*; секвенирование; первичная структура белка.

ВВЕДЕНИЕ

Ранее из яда паука караокурта было выделено семейство высокомолекулярных белковых нейротоксинов – латротоксинов (с молекулярной массой от 110 до 130 кДа). В наномолярных концентрациях эти токсины вызывают массированный и истощающий выброс нейромедиаторов из пресинаптических окончаний, что в конечном итоге приводит к блокаде нервного импульса [1, 2]. В настоящее время идентифицировано семь из этих токсинов: α -латротоксин (α -LTX), пять латроинсектотоксинов (α -LIT, β -LIT, γ -LIT, δ -LIT, ε -LIT) и α -латрокrustотоксин (α -LCT), которые селективно действуют на нервные окончания позвоночных животных, насекомых и ракообраз-

ных соответственно [3, 4]. Высокая степень структурной гомологии латротоксинов и практически одинаковый для всех организмов физиологический эффект позволяют предположить существование сходного консервативного механизма, по которому осуществляется воздействие токсина на нервное окончание.

На сегодняшний день наиболее изученным нейротоксином является α -LTX. Известны два его рецептора у позвоночных: латрофилин [5–8] и Ca^{2+} -зависимый receptor нейрексин [4]. Однако функция этих рецепторов и их роль в проявлении токсического действия латротоксина до конца не ясны. При экспрессии в культуре клеток оба рецепторных белка усиливают действие α -LTX, причем укороченные белки без предполагаемых сигнальных доменов оказывают такое же действие [9–12]. Рассматриваемые белки-рецепторы, связываясь с токсином, облегчают формирование ионных каналов, но при этом физически не включаются в структуру канала [13, 14]. С другой стороны, receptor может активироваться связыванием с α -LTX и передавать сигнал в клетку. Подтверж-

Сокращения: LLP – Latrophilin-like Protein (латрофилинподобный белок); *llp* – ген, кодирующий LLP; LTX – латротоксин; LIT – латроинсектотоксин; RACE – Rapid Amplification cDNA Ends (быстрая амплификация концевых последовательностей кДНК); ORF – Open Reading Frame (открытая рамка считывания).

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

дением этому служат эксперименты с мутированным токсином LTX^{N4C}, который не может формировать поры, однако вызывает выброс медиаторов в присутствии внеклеточного Ca²⁺ [15, 16].

К настоящему времени описаны также рецепторы латротоксинов у паразитических нематод *Haemonchus contortus* (HC110-R) и *Caenorhabditis elegans* (B0457.1). Эти латрофилинподобные белки гомологичны друг другу (на уровне 48%) и латрофилину млекопитающих (HC110-R – на 31, B0457.1 – на 29%). Белок HC110-R, найденный как мишень для антигельминтного десипептида PF1022A, показал также связывание с α-LTX [17]. Было также показано, что инактивация методом РНК-интерференции мРНК, кодирующей белок B0457.1, придает *C. elegans* устойчивость к яду каркрута [18].

В то же время рецепторы латротоксинов у насекомых до сих пор достоверно не идентифицированы. Известно, что связывание α-LIT с рецептором саранчи не зависит от концентрации Ca²⁺. Однако добавление в среду двухвалентных катионов (Mg²⁺ или Ca²⁺) способствует индуцированному α-LIT выбросу нейромедиаторов из пресинаптических окончаний, что может также указывать на сходство действия α-LTX и α-LIT [19].

Цель настоящей работы заключалась в поиске, клонировании и определении полной нуклеотидной последовательности мРНК, кодирующую латрофилинподобный белок у *Musca domestica*, и изучении первичной структуры и доменной организации белка LLP – вероятного рецептора латроинсектотоксина.

При проведении данной работы мы столкнулись с несовершенством существующих методик 3'-RACE. Это выражалось, в частности, в том, что в ходе ПЦР с генспецифичным и универсальным праймерами было практически невозможно получить одновременно два отличающихся по размеру фрагмента кДНК, относящихся к двум (или даже к одной) сплайсированным мРНК. В статье обсуждаются методические приемы, которые были использованы для преодоления подобного рода трудностей и возможные молекулярные механизмы этого явления.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В экспериментах по радиолигандному анализу связывания ¹²⁵I-меченого α-LIT с мембранами *M. domestica*, проведенных ранее в нашей лаборатории, было показано, что связывание токсина не зависит от концентрации Ca²⁺ (данные не опубликованы), поэтому мы предположили, что рецепцию α-LIT у домовой мухи осуществляет латрофилинподобный белок. Ближайшим к *M. domestica* видом с секвенированным геном является *Drosophila melanogaster*.

Нами был проведен компьютерный поиск генов, белковые продукты которых гомологичны латрофилину в геноме *D. melanogaster*. При этом было выявлено два гена, кодирующих белки с участками гомологии с латрофилином крысы (*Rattus norvegicus*): ген CG8639 (кодирует белок – AE003836) и ген stan (кодирует белок – AE003828). При сравнении доменной организации этих белков с организацией латрофилина млекопитающих и нематод мы пришли к заключению, что наиболее вероятным гомологом латрофилина млекопитающих у дрозофилы является белок AE003836, кодируемый геном CG8639. Поскольку *M. domestica* в таксономическом отношении довольно близка к *D. melanogaster*, логично было предположить, что латрофилинподобные белки у этих двух насекомых имеют значительную гомологию. В связи с этим белок AE003836 *D. melanogaster* служил референс-стандартом при поиске и идентификации гена llp, кодирующего латрофилинподобный белок (LLP) *M. domestica*.

Получение и секвенирование внутреннего фрагмента кДНК гена llp *M. domestica*

При сравнительном анализе аминокислотных последовательностей латрофилина животных (*R. norvegicus*) и латрофилинподобных белков насекомых (*D. melanogaster*), а также нематод (*C. elegans* и *H. contortus*) были найдены гомологичные участки в их структуре. Эти участки протяженностью от 3–4 до 9–11 а. о. располагались в N-концевой части анализируемых белков (данные не представлены). Нами были выбраны два наиболее длинных участка: один из них располагался в галактозосвязывающем лектиновом домене (GAL), другой – в GPS-домене, содержащем сайт эндопротеолиза. Мы предположили, что эти участки являются консервативными. Расстояние между выбранными участками белка AE003836 *D. melanogaster* составляло около 650 а. о. Структура двух консервативных участков у латрофилина крысы и LLP-белка *D. melanogaster* представлена на рис. 1.

На основании аминокислотной последовательности двух консервативных участков и с учетом вырожденности генетического кода нами был осуществлен дизайн и синтез двух встречных вырожденных олигонуклеотидных праймеров D1 и D3 (структуре праймеров см. в таблице). В ПЦР с вырожденными праймерами D1 и D3 в качестве матрицы использовали как геномную ДНК *M. domestica*, так и препарат первой цепи кДНК, полученный на матрице суммарной РНК из голов домовой мухи. Соответствующий ПЦР-продукт был получен лишь на матрице геномной ДНК мухи. Его размер составлял ~2 т. п. о., что было близко к теоретически ожидаемому, при условии отсутствия больших интронов. Полученный фрагмент ДНК, обозначенный нами как D1–D3, был секвенирован с двух сторон. Как

***** . * .. * . * . * . * . . + . * . * . . * . * . . * . * . . * . * . . * . * .
 m (60) NANYGRTDDKICDaDPFQME-NVQCYLPDAFKIMSQRCCNNRTQCVVAGSDAFPDPCPGTYKYLEVQYDCVP
 dr (45) RANYGRFSITICN-DHGNVEwSVNCMFPKSLSVLNSRCAHQSCGVLAATSMFGDPCPGTHKYLEAHYQCIS
 ***** + .. * .. * . * . * . * . . * . * . . * . * . . * . * . . * . * . . * . * .
 m (803) FWNYSERSmIgWSTQGCRILVESNKHTTCACSHLTNFAVLMAHREIYQGRINELL--LSVITWVGIVIS
 dr (711) FWNYIDHA---WSANGCSLESTNRTHSVCSCNHLTNFAILMDVVDEHQHSLFTMFdgMRIFTYISIGIC

Рис. 1. Выравнивание аминокислотных последовательностей двух консервативных фрагментов латрофилина крысы *R. norvegicus* (m) и латрофилинподобного белка *D. melanogaster* (dr). Серыми прямоугольниками выделены аминокислотные последовательности, к которым были подобраны вырожденные олигонуклеотидные праймеры D1, D2 и D3. Звездочками обозначены идентичные аминокислотные остатки, плюсами – гомологичные, точками – негомологичные аминокислотные остатки.

было выяснено далее, фрагмент D1–D3 *M. domestica* действительно кодирует полипептиды, имеющие высокую степень гомологии с белком AE003836 *D. melanogaster*, т.е. является фрагментом искомого гена (данные не приведены). Наличие инtronов в клонированном фрагменте D1–D3 делало нецелесообразным работу по его полному секвенированию.

Данные по секвенированию 5'- и 3'-концевых участков фрагмента D1–D3 геномной ДНК были

использованы для дизайна и синтеза двух встречных генспецифичных (для *M. domestica*) праймеров – A1 (прямой) и A4 (обратный) (см. таблицу и рис. 2). С помощью этих праймеров и первой цепи кДНК, полученной на основе суммарной РНК из голов мухи, мы амплифицировали внутренний фрагмент кДНК гена *lhp* *M. domestica* размером 1.8 т. п. о. и полностью его секвенировали. Оказалось, что этот фрагмент ДНК *M. domestica* в одной из рамок

Дезоксиолигонуклеотиды, использованные в работе

Олигонуклеотид	Структура олигонуклеотида (5' → 3')	Для чего был использован
D1	CCT TGT CCT GGW ACN CAY AAR TA	ПЦР на геномной ДНК
D2	RTA CTT GTG WGT WCC TGG RCA DGG	ПЦР, 5'-RACE
D3	GC RAA GTT KGT CAR RTG RTT RCA	ПЦР на геномной ДНК
A1	ACC ACA ACT AAT CGT CCC AGT CC	Получение фрагментов кДНК
A2	TGG ACT GGG ACG ATT AGT TGT GG	ПЦР, 5'-RACE
A3	CAA CCA ATC GTA CAC ACA GCA TC	Получение фрагментов кДНК
A4	AAC ACA ACC ATT GGC TGA CCA AGC	Получение кДНК копии гена, получение первой цепи кДНК для 5'- и 3'-RACE
A5	CTC CTC CAA TTC TTC ACC ATC AC	ПЦР, 3'-RACE
A6	TAC TGG ATT GGT TGG CGA AGA CG	ПЦР, 3'-RACE
A7	TGG TAT TGA ACA AAC GGA AAC GAG	ПЦР, 3'-RACE
A8	GAC ATC ATC TGT AAG CTC AAC CC	ПЦР, 3'-RACE
A9	GCC TTT TAT CAC CAA CCA GAC TAC	ПЦР, 3'-RACE
C1	ATC TCG AGA AGC TTA TCT CTC GCT TGA CCC-TTA AGG TTC	Получение длинноразмерных фрагментов кДНК
5long	ATC GAT AAC ACA AGC CAT GGC CG	Получение полноразмерных копий альтернативно сплайсированных вариантов AC и AD
5short	ATA TCA AAG CAT GGC CTT GGT CG	То же, вариантов BC и BD
3long	TTA CTT CTC TAA AGG CTC TAA ATA ACA CC	То же, вариантов AD и BD
3short	TGT TTC GTT CAC AAA CCG TTC AGC	То же, вариантов AC и BC
T-праймер	AAGCAGTGGTAAACAAACGCAGAGTAC(T) ₃₀ N ₋₁ N (N = A, C, G, T; N ₋₁ = A, G, C)	Получение первой цепи кДНК для 5'- и 3'-RACE
TS	AAGCAGTGGTATCAACGCAGAGTACGGCrGrGrG	Получение первой цепи кДНК для 5'-RACE
Heel carrier (T7cap)	GTAATACGACTCACTATAGGGCAAGCAG-TGGTAACAACGCAGAGT	ПЦР с селективной супрессией

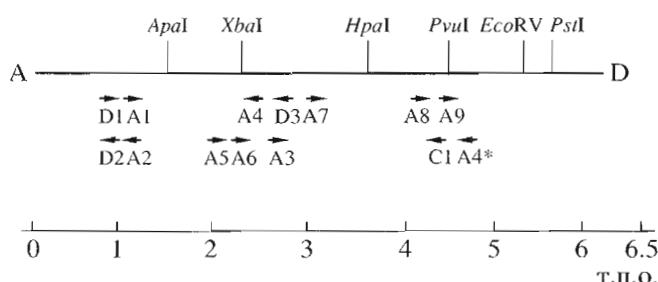


Рис. 2. Физическая карта контига (альтернативно сплайсированного варианта мРНК) AD M. domestica, содержащего ген латрофилиноподобного белка. Показаны уникальные участки узнавания рестриктаз. Стрелками обозначены вырожденные и генспецифичные праймеры. Участок неспецифического отжига праймера A4 отмечен звездочкой.

считывания кодирует полипептидную цепь, имеющую значительную гомологию с белком AE003836 дрозофилы. При этом, на уровне ДНК гомология достаточно низкая (48%), а при сравнении аминокислотных последовательностей достигает 65%.

Завершение работы по получению и секвенированию внутреннего фрагмента кДНК M. domestica позволило нам перейти к следующему этапу работы – выяснению полной первичной структуры мРНК и кодируемого ею белка LLP домовой мухи.

Получение 5'-концевых фрагментов гена llp M. domestica (5'-RACE)

Для получения недостающего 5'-концевого фрагмента кДНК гена, кодирующего LLP, использовали метод 5'-RACE (от Rapid Amplification cDNA Ends – быстрая амплификация концевых последовательностей кДНК), описанный в работе [20]. Первую цепь кДНК получали на основе poly(A⁺)-РНК с использованием генспецифичного праймера A4 и олигонуклеотида TS (таблица).

Далее первая цепь кДНК служила матрицей в ПЦР-амплификации с генспецифичным праймером A2 и универсальным праймером T7cap (другое название – heel carrier [20]) (таблица). Было получено два фрагмента ДНК А и В (рис. 3, дор. 1, 2) длиной 1.1 и 0.8 т. п. о. соответственно. Мы обнаружили, что уровень синтеза фрагмента А сильно зависит от температуры отжига праймеров в ходе термальных циклов. Оптимальной является температура 65.5°C (рис. 3, дор. 1). Повышение температуры отжига в ПЦР всего на 1.3°C резко снижает выход этого фрагмента (рис. 3, дор. 2). Доказательства тому, что указанные ПЦР-фрагменты не являются результатом неспецифического отжига праймеров, были получены в полугнездовой ПЦР. В этом случае разведенную реакционную смесь после первого раунда ПЦР использовали в качестве матрицы для второго раунда ПЦР с праймерами T7cap и D2 (таблица). В этом опыте также получили два фрагмента ДНК А' и В' (не показано), однако они были короче исходных А и В на ~100 п. о., что и следовало ожидать, поскольку последовательность ДНК, соответствующая праймеру D2, расположена проксимальнее последовательности праймера A2 на ~100 п. о. (рис. 2).

В ходе секвенирования полученных фрагментов ДНК было установлено, что оба фрагмента А и В (рис. 4а и 4б соответственно) в 3'-концевой области содержат идентичную нуклеотидную последовательность (267 п. о.), которая в одной из рамок кодирует полипептид, гомологичный соответствующему фрагменту белка AE003836 D. melanogaster. В то же время 5'-концевые участки фрагментов А и В имеют уникальную структуру (их длина 739 и 424 п. о. соответственно). На основании полученных результатов мы сделали заключение, что анализируемые фрагменты кДНК относятся к двум альтернативно сплайсированным мРНК гена, кодирующими LLP M. domestica.

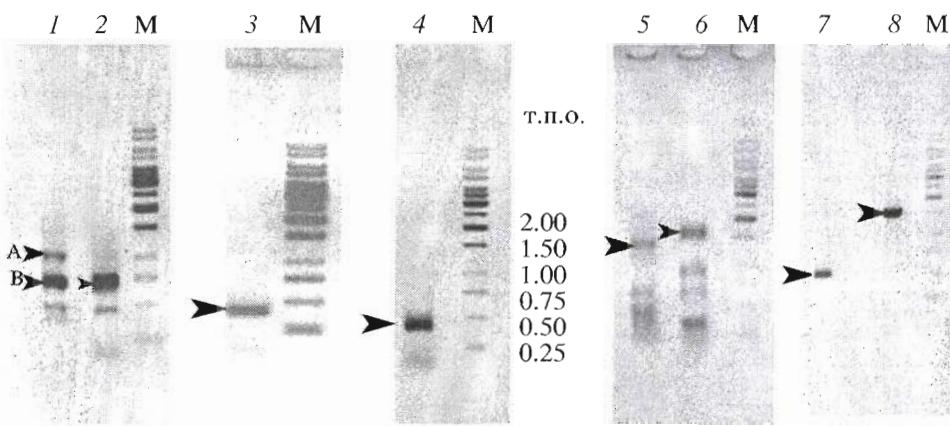


Рис. 3. Электрофорез в 1% агарозном геле фрагментов ДНК гена llp M. domestica, полученных в ходе ПЦР (5'-RACE) с праймерами A2 и T7cap при температуре отжига 65.5°C и 66.8°C (фрагменты А и В) (дор. 1 и 2 соответственно); в полугнездовой ПЦР (3'-RACE) с праймерами A3 и T7cap (дор. 3); в ПЦР с праймерами A7 и T7cap (дор. 4); в ПЦР с праймерами A7 и A4* (дор. 5) и с праймерами A3 и A4* (дор. 6); в ПЦР с праймерами A9 и T7cap и первой цепью кДНК, полученной с poly(A⁺)-РНК (дор. 7) и с суммарной РНК (дор. 8). М – маркеры длины фрагментов ДНК (1 kb DNA ladder, Fermentas).

(а)

старт

601 taacacaagccatgccggcatctacaacaacattattaggcatttggctacagcta
 ·T··Q··A··M··A··G··I··Y··N··N··I··R··H··F··G··L··Q··L··

661 tttgtctcctgctgttacggctaacaatgattaccagtgtgtgcgggtggcatggat
 C··C··L··L··L··L··R··L··T··M··I··T··S··A··V··A·^V··G··M··D··
 SP ↑

721 gcccataatgcaaatcttacggattggtagtccggctctaaatataactgcatacg
 A··A··N··A··N··L··T··D··L··G··S··P·**A**··P··K··Y··Q··T··A··Y··

(б)

1 ggtatatatgattatqaaatqacaacaaaatacaacattccaaagattaccaacgctc
 ·G··I··Y··D··Y··E··Y··D··N··K··I··Q··T··F··Q··R··L··P··T··L··
 ·V··Y··M··I··M··N··M··T··T··K··Y··K··H··S··K··D··Y··Q··R··S··
 ·Y··I··*··L··*··I··*··Q··Q··N··T··N··I··P··K··I··T··N··A··R··

61 gacttgctgtcttcaacttctaaacqtcgcatcataagtataaaaattgagagatcaa
 ·D··L··L··V··F··T··I··L··N··V··A··S··*··V··*··K··L··R··D··Q··
 ·T··C··*··S··S··L··F··*··T··S··H··H··K··Y··K··N··*··E··I··N··
 ·L··A··S··L··H··Y··S··K··R··R··I··J··S··I··K··I··E··R··S··M··

121 tgtgtcatttgataactttacytagtattcattacgtatacatttatataaccatt
 ·C··V··I··C··I··T··F··T··*··Y··S··L··R··I··H··S··I··Y··T··I··
 ·V··S··F··V··*··L··L··X··S··I··H··Y··V··Y··I··L··Y··I··P··F··
 ·C··H··L··Y··N··F··Y··L··V··F··I··T··Y··T··F··Y··I··Y··H··F··

старт

181 ttatataccatataccaagatcccttggtcgtagtaaatttctttcaggta
 ·L··Y··I··P··I··S··K··H··G··L··G··R··V··S··K··F··S··F··Q··V··
 ·Y··I··Y··Q··Y··Q··S··M··A··L··V··V··L··V··N··S··L··F··R··Y··
 ·I··Y··T··N··I··K··A··W··P··W··S··C··*··I··L··F··S··G··T··

старт I

241 cattccaaaaaccatccaacaaccatcaaaactacaaacatcgatcaacaacta**att**
 ·H··S··K··N··Q··S··N··N··H··Q··T··N··Y··K··H··R··Q··Q··L··I··
 ·T··P··K··T··N··P··T··T··I··K··P··T··T··N··I··V··N··N··*··L··
 ·F··Q··K··P··I··Q··Q··P··S··N··Q··L··Q··T··S··S··T··T··N··*··

старт 2-3

301 aattttgtctttat**aatt**tagtattgtatatttacttgcacaccaccacc
 ·N··F··C··P··Y··I··I··S··I··V··I··Y··L··L··V··L··H··T··T··T··
 ·I··F··V··L··I··*··L··V··L··*··Y··I··Y··L··C··C··T··P··P··P··
 ·F··L··S··L··Y··N··*··Y··C··N··I··F··T··C··A··A··H··H··H··H··

361 atcacccaattcaatctttacataaaacttttcaaaattttagct:ctaaatatcaaact
 ·I··T··N··S··I··F··Y··I··N··L··F··K··I··L··**A**··P··K··Y··Q··T··
 ·S··P··I··Q··S··F··T··*··T··F··S··K··F··*··I··L··N··I··K··L··
 ·H··Q··F··N··L··L··H··K··P··F··Q··N··F··S··S··*··I··S··N··C··

Рис. 4. Первичная структура 5'-концевых участков фрагментов А (а) и В (б) кДНК гена *llp* *M. domestica*, полученных в ходе 5'-RACE. Предполагаемые стартовые кодоны подчеркнуты. Соответствующие аминокислотные последовательности приведены в однобуквенном коде в одной (фрагмент А) и в трех (фрагмент В) рамках считывания, звездочками обозначены стоп-кодоны. SP – сигнальный пептид, стрелкой показано место его отщепления. Жирным шрифтом выделена уникальная N-концевая аминокислотная последовательность белка, кодируемого фрагментом А, за ней следует аминокислотная последовательность, общая для фрагментов А и В, она выделена затенением.

*Получение и секвенирование 3'-концевых фрагментов гена llp *M. domestica* (3'-RACE)*

Для установления структуры 3'-концевого фрагмента гена, кодирующего LLP домовой мухи, использовали методологию 3'-RACE и универсаль-

ные олигонуклеотидные праймеры как описано в работе [20]. Первую цепь кДНК для 3'-RACE получали с помощью Т-праймера (см. таблицу), содержащего (dT)₃₀, к 5'-концу которого присоединена уникальная последовательность из 25 нуклеотидов – так называемый адаптер. Теоретически

мы ожидали получить в процессе одного раунда ПЦР недостающий 3'-концевой фрагмент кДНК вплоть до poly(A)-последовательности, однако в эксперименте пришлось провести пять последовательных шагов для того, чтобы получить все перекрывающиеся фрагменты искомого гена. Причиной этого явилось наличие сильно укороченного сплайсированного варианта мРНК в 3'-области, а также то обстоятельство, что использованный при синтезе первой цепи кДНК Т-праймер, в силу особенностей своей структуры, обладал способностью отжигаться и инициировать синтез кДНК в нескольких местах целевой мРНК, а не только в poly(A)-участке.

3'-RACE. Шаг первый – шаг пятый

В ходе ПЦР с генспецифичными праймерами A5, A6 (см. таблицу) и универсальным праймером T7cap был получен набор фрагментов ДНК разной длины в присутствии сильного фона. Для получения индивидуальных фрагментов ДНК проводили второй раунд ПЦР с генспецифичным праймером A3 совместно с T7cap (рис. 2). В этом опыте в качестве матрицы использовали реакционную смесь, полученную в первом раунде ПЦР, разведенную в 50 раз. После 20 термальных циклов был получен ПЦР-продукт размером ~370 п. о. (рис. 3, дор. 3). Секвенирование этого фрагмента показало, что он имеет область перекрывания с фрагментом D1–D3 гена *llp* и в одной из рамок кодирует полипептид, имеющий значительную гомологию с белком AE003836 *D. melanogaster*.

Поскольку в рамке, кодирующей часть белка LLP, терминирующие кодоны отсутствовали, было ясно, что получена лишь часть 3'-концевого фрагмента гена *llp*. Исходя из последовательности фрагмента A3–T7cap, был синтезирован генспецифичный праймер A7 (см. таблицу и рис. 2). Когда праймер A7 был использован в ПЦР с праймером T7cap и первой цепью кДНК в качестве матрицы, то был получен единственный фрагмент ДНК размером ~400 п. о. (рис. 3, дор. 4). Секвенирование этого фрагмента ДНК показало, что он является частью гена *llp*. Фрагмент A7–T7cap в одной из рамок кодирует полипептид размером 54 а. о., который на протяжении 30 а. о. имеет гомологию с соответствующим фрагментом белка AE003836 *D. melanogaster*. Далее идет уникальный участок – полипептид длиной 23 а. о. Кодирующая рамка заканчивается терминирующим кодоном (TGA), за которым идет некодирующая область (293 п. о.).

Из представленных выше результатов создавалось впечатление, что мы получили 3'-концевой фрагмент гена *llp M. domestica* и что этот ген у домовой муки значительно короче, чем его гомолог у *D. melanogaster*. Другое объяснение полученных результатов – альтернативный сплайсинг

пре-мРНК. Иными словами, наряду с короткой должна существовать более длинная альтернативно сплайсированная мРНК. Однако получить с помощью праймеров A3 + T7cap и A7 + T7cap ПЦР-продукт, отличный от описанных выше, не удавалось. Как уже отмечалось выше, наиболее вероятной причиной возникших трудностей явились особенности первичной структуры мРНК исследуемого гена, а также особенности структуры использованного нами в работе универсально-го Т-праймера – его способность предпочтительно отжигаться в нескольких местах целевой мРНК и инициировать синтез кДНК. Вследствие этого в процессе обратной транскрипции мы получали дискретный набор матриц (кДНК) различной длины, укороченных с 3'-конца, которые конкурировали в ПЦР. Радикальным способом преодолеть возникший барьер можно было бы с помощью праймера иной структуры, с меньшим содержанием G + C.

Для получения более длинного сплайсингового варианта мы использовали способность любых достаточно протяженных праймеров неспецифично отжигаться при температуре ревертазной реакции (37–42°C) и служить затравкой для обратной транскриптазы. Таким образом можно получить первую цепь кДНК, у которой на 5'-конце известный праймер, отжегшийся в "случайном" месте. С учетом вышесказанного, первая цепь кДНК была получена с использованием олигонуклеотидного праймера A4 в качестве затравки (выбор праймера A4 был случаен). Место неспецифического отжига этого праймера на мРНК показано на рис. 2. В ПЦР с полученной таким образом матрицей и праймерами A7 + A4* и A3 + A4* были получены фрагменты ДНК длиной 1.3 и 1.5 т. п. о. соответственно (рис. 3, дор. 5 и 6).

Секвенирование фрагмента A7–A4* показало, что он является продолжением фрагмента A3–T7cap и в одной из рамок кодирует полипептид, который имеет гомологию с соответствующей частью белка AE003836 *D. melanogaster*. Другими словами, фрагмент кДНК размером 1.3 т. п. о. является частью гена *llp M. domestica* и входит в состав длинной сплайсированной мРНК этого гена.

Для продвижения в 3'-концевую часть гена *llp* были синтезированы генспецифичные праймеры A8 и A9 (рис. 2 и таблица). В ПЦР в качестве матрицы использовали первую цепь кДНК, полученную с мРНК и Т-праймера в качестве затравки. В ходе однораундовой ПЦР с праймерами A8 + + T7cap и A9 + T7cap нам удалось получить достаточно чистые фрагменты ДНК размером около 700 и 600 п.о. соответственно (фрагмент ДНК A9–T7cap см. на рис. 3, дор. 7). Результаты секвенирования фрагмента A9–T7cap показали, что он является продолжением ПЦР-фрагмента A7–A4* (рис. 2 и 5), но не содержит терминирующего кодона в кодирующей рамке. Попытки получить

более длинный ПЦР-продукт с праймерами A8 + T7cap и A9 + T7cap и кДНК, полученной с мРНК, не приводили к успеху.

ПЦР-продукт значительной длины (~1600 п. о.) с праймерами A9 + T7cap (рис. 3, дор. 8) был генерирован при использовании в качестве матрицы кДНК, полученной с суммарной РНК из голов мух, а не с мРНК. Причина в различиях матричных свойств двух препаратов кДНК не вполне ясна.

Из данных секвенирования следовало, что фрагмент A9-T7cap размером 1600 п. о. включает в себя полностью более короткий фрагмент кДНК (600 п. о.), полученный с той же парой праймеров. При этом смысловая рамка считывания, кодирующая *LLP M. domestica*, заканчивается терминирующим кодоном на расстоянии 115 п. о. от 3'-конца фрагмента. Вслед за терминирующим кодоном идет некодирующая часть гена с poly(A)-последовательностью в конце.

Таким образом были получены все перекрывающиеся фрагменты кДНК гена, кодирующего *LLP M. domestica*.

*Реконструкция полных нуклеотидных последовательностей мРНК, кодирующих *LLP M. domestica**

С помощью компьютерной программы DNA Star (Seqman) из нуклеотидных последовательностей всех перекрывающихся фрагментов кДНК, полученных в ходе 3'- и 5'-RACE, а также внутреннего фрагмента A1-A4 (рис. 5) были выявлены четыре контига (схематически они изображены на рис. 5 и обозначены AC, BC, AD и BD). Все четыре контига имеют область перекрывания (2512 п. о.) с единой рамкой считывания. Вопрос о начале рамок считывания (ORF) у различных контигов будет рассмотрен ниже. Полные нуклеотидные последовательности всех четырех контигов, а также аминокислотные последовательности кодируемых ими белков депонированы в базе данных GenBank, регистрационные номера: AY816749, AY816750, AY816751 и AY816752.

Очевидно, что выявленные контиги соответствуют всем возможным альтернативно сплайсированным мРНК гена *llp*. Как видно из схемы на рис. 5, число альтернативно сплайсированных мРНК может варьировать от двух до четырех. Возникает вопрос, каково же на самом деле число различных мРНК гена *llp* у домовой мухи? Чтобы ответить на этот вопрос, были синтезированы 5'- и 3'-концевые олигонуклеотидные праймеры (5long, 5short, 3long и 3short, см. таблицу). Используя пары встречных праймеров 5long + 3short, а также 5short + 3short и первую цепь кДНК, полученную на матрице суммарной РНК из голов мух, в ходе ПЦР нами были получены фрагменты ДНК, которые по своему размеру (~3 т. п. о.) со-

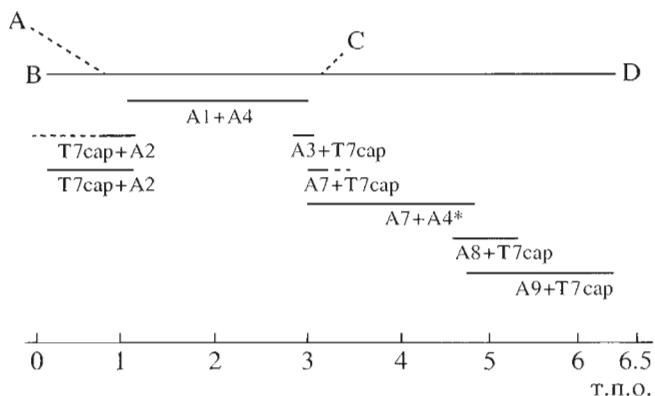


Рис. 5. Схематическое изображение четырех контигов (AD, BD, AC и BC), выявленных в ходе секвенирования перекрывающихся фрагментов кДНК, кодирующих *LLP M. domestica*, полученных в ходе процедур 5'- и 3'-RACE. Все четыре контига имеют общую область перекрывания (участок между разрывами). Показана примерная локализация и размеры перекрывающихся фрагментов кДНК и указаны праймеры, с помощью которых они были амплифицированы. Полные нуклеотидные последовательности четырех контигов представлены в базе данных GenBank, номера: AY816749, AY816750, AY816751 и AY816752.

ответствовали полноразмерным сплайсированным вариантам AC и BC (см. рис. 5). При этом выход продуктов ПЦР в обоих случаях был практически одинаков. Это означает, что в популяции мРНК две короткие альтернативно сплайсированные формы AC и BC присутствуют в одинаковой или близкой концентрации.

Мы попытались получить в ПЦР с парами встречных праймеров 5short + 3long и 5long + 3long фрагменты ДНК, соответствующие длинным альтернативно сплайсированным мРНК типа AD и BD. Однако наши попытки были неудачны. Причиной этого могла быть низкая концентрация целевой мРНК в суммарной РНК, а также недостаточно высокая процессивность обратной транскриптазы, которую мы использовали в работе. Поэтому для выяснения поставленного вопроса были взяты дополнительные генспецифичные праймеры C1 и A1 (см. рис. 2). Праймер C1 отжигается в 3'-концевом участке длинных контигов на расстоянии 1479 п. о. от терминирующего кодона. С помощью этих праймеров мы попытались получить несколько укороченные и перекрывающиеся ПЦР-продукты.

В ПЦР использовали пары встречных праймеров 5long + C1, 5short + C1, а также 3long + A1. В ходе 40 термальных циклов во всех случаях ПЦР-продукты были получены. Их длина соответствовала ожидаемым величинам, т.е. вычисленным на основании известных нуклеотидных последовательностей. Следует отметить, что при этом выход ПЦР-продуктов в случаях с праймерами 5long + C1

и 5short + C1 был практически одинаков. Из полученных данных следует, что в популяции мРНК из голов мух присутствуют обе длинные альтернативно сплайсированные мРНК типа AD и BD в одинаковой концентрации. Таким образом, в организме домовой муки реализуются все четыре альтернативно сплайсированные мРНК: две короткие – AC и BC и две длинные – AD и BD.

Получение полноразмерных кДНК, соответствующих двум длинным альтернативно сплайсированным мРНК

Наличие ПЦР-продукта, амплифицированного с помощью праймеров 3long + A1, позволило нам получить полноразмерные кДНК со структурой AD и BD (рис. 5). Мы установили, что все четыре выявленных нами варианта кДНК (четыре контига) содержат уникальный сайт расщепления рестриктизы *Xba*I (рис. 2). Три очищенных ПЦР-продукта (фрагменты 3long-A1, 5long-C1 и 5short-C1) были обработаны ферментом *Xba*I и лигированы так, чтобы продуктом этой реакции явились две полноразмерные кДНК со структурой AD и BD. Две полученные при этом лигированные смеси были использованы затем в качестве матрицы в ПЦР с парами встречных праймеров 5long + 3long и 5short + 3long. В ходе амплификации в обоих случаях были получены ПЦР-продукты размером более 5 т. п. о., что соответствовало теоретически ожидаемым полноразмерным кДНК, т.е. контигам AD и BD. Их структура была проверена секвенированием.

Начало открытой рамки считывания в структуре длинного 5'-концевого фрагмента кДНК

Для поиска инициирующих (стартовых) кодонов, с которых начинаются рамки считывания латрофилинподобного (ых) белка (ов) у *M. domestica*, был проведен анализ первичной структуры уникальных 5'-концевых фрагментов кДНК, полученных в ходе 5'-RACE. Компьютерная трансляция 5'-концевого фрагмента кДНК длиной 1006 п. о. (фрагмент А, рис. 3, дор. 1) в соответствующей рамке показала, что инициирующим кодоном в данном случае, наиболее вероятно, является метиониновый кодон ATG, расположенный на расстоянии 611 нт от 5'-конца. Инициирующий кодон был выбран с учетом его окружения (контекста). Ранее было установлено, что для инициации трансляции необходимо наличие консенсусной последовательности GCC(A/G)CCAUGG [21, 22], в которой наиболее важны пурины (обычно А) в положении –3 (по отношению к А (+1) в AUG) и G в положении +4. В нашем случае эта последовательность имеет структуру GCCATGG (рис. 4a). Фрагмент ДНК, дистальнее стартового кодона,

кодирует предполагаемый сигнальный пептид длиной 32 а. о. (рис. 4a) и N-концевую часть белка LLP, гомологичную AE003836 *D. melanogaster* (на рис. 4a этот участок белка выделен затемнением). Сигнальный пептид был определен с помощью программы SignalP 3.0.

В 5'-нетранслируемой области (611 п. о.) имеется еще семь ATG-кодонов (на рис. 4a они не показаны), однако все они находятся вне рамки считывания белка LLP. Более того, только один из них находится в окружении, близком к оптимальному для инициации трансляции – AACATGG. ORF в данном случае имеет длину 43 кодона. Длинная нетранслируемая область (611 п. о.), содержащая короткие рамки считывания, возможно, служит для регуляции уровня синтеза латрофилинподобного белка.

Выбор инициирующего кодона в структуре короткого 5'-концевого фрагмента кДНК

5'-Концевой фрагмент кДНК длиной 691 п.о. (фрагмент В, см. рис. 3, дор. 1 и 2) также состоит из 5'-нетранслируемой области и участка, кодирующего N-концевую часть соответствующих LLP (варианты BC и BD). Стартовый кодон в данном случае определить, однако, не удалось, поскольку все пять ATG-кодонов, выявленные внутри анализируемого фрагмента кДНК, находятся вне рамки считывания латрофилинподобного белка. При этом, первый ATG-кодон находится слишком близко к 5'-концу, чтобы быть инициирующим. Еще три кодона не имеют благоприятного окружения (контекста), и только один находится в благоприятном контексте – AGCATGG (на рис. 4б он обозначен как старт). Этот кодон начинает короткую ORF длиной 93 нт, т.е. кодирует пептид размером 31 а. о.

По всей видимости, в качестве инициирующего кодона в данном случае используются альтернативные кодоны. Известно, что в качестве инициирующих могут быть триплеты GUG, UUG, AUU, AUA, CUG, ACG [23]. Исходя из того, что белок LLP трансмембранный и его экстрацеллюлярная часть, скорее всего, подвергается посттрансляционным модификациям, у этого белка должен быть сигнальный пептид. Используя компьютерную программу SignalP 3.0, мы проверили все возможные альтернативные кодоны, находящиеся в одной рамке считывания с рамкой считывания LLP, и выяснили, что только три из них могут инициировать синтез сигнальных пептидов. Это кодоны ATT и дуплет ATAATT, обозначенные на рис. 4б как старт1 и старт2–3. Следует отметить, что все три предлагаемых стартовых кодона не имеют благоприятного окружения, лишь последний имеет А в положении –3. Таким образом, вопрос о стартовом кодоне в структуре короткого 5'-концевого фрагмента кДНК остается открытым.

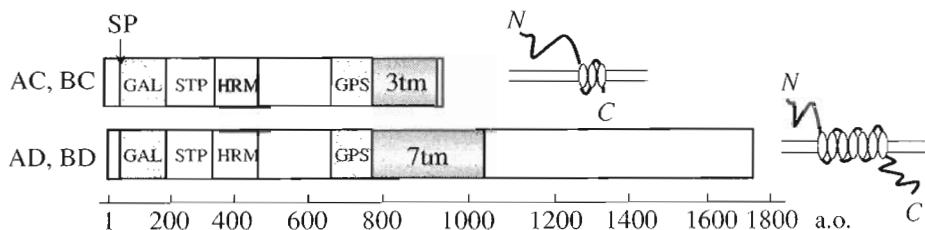


Рис. 6. Предполагаемая доменная организация латрофилинподобных белков *M. domestica*. Сигнальный пептид обозначен как SP, стрелкой показано место его отщепления. GAL – галактозосвязывающий лектиновый домен; STP – участок, богатый серином, треонином, пролином; HRM – мотив, характерный для гормональных рецепторов; GPS – участок с четырьмя консервативными цистеинами, содержащий сайт эндопротеолиза; 7tm – 7-тяжевой трансмембранный домен секретинового семейства; 3tm – укороченный (3-тяжевой) трансмембранный домен.

Характеристика латрофилинподобных белков *M. domestica* и их доменная организация

Компьютерный анализ выведенных аминокислотных последовательностей белков – продуктов трансляции контигов AD и BD (рис. 5) показал, что они обладают всеми консервативными доменами, характерными для латрофилинов [8]. В их структуре присутствуют: галактозосвязывающий лектиновый домен, мотив, характерный для гормональных рецепторов, GPS-домен, 7-тяжевой трансмембранный домен рецепторов секретинового семейства (рис. 6). При этом на уровне аминокислотной последовательности у латрофилинподобного белка *M. domestica* наблюдается гомология с латрофилином крысы (28%) и с латрофилинподобными белками нематод *H. contortus* и *C. elegans* (22%). Однако наибольшая гомология наблюдается с LLP *D. melanogaster* – 64%. Следует отметить, что гомология с латрофилинами позвоночных имеет место только в *N*-концевой части белка, где, как известно, расположен участок связывания токсина. Продукты трансляции контигов AC и BC (рис. 5) также содержат все вышеуказанные консервативные домены, однако трансмембранный домен рецепторов секретинового типа укорочен и содержит только три трансмембранных участка (тяжя) (см. рис. 6). Возможно, в данном случае образуется рецептор, не способный к проведению сигнала.

Функциональная значимость двух разных вариантов *N*-концевой части белков не ясна. Можно предположить, что различные сигнальные пептиды определяют разные посттрансляционные модификации и направление транспорта белка.

С помощью программ NetOGlyc и NetNGlyc в предполагаемом экстрацеллюлярном домене белка LLP *M. domestica* было обнаружено 11 потенциальных сайтов *N*-гликозилирования и четыре сайта *O*-гликозилирования. Это соответствует данным о гликозилировании латрофилина. У латрофилина млекопитающих выявлено семь потенциальных сайтов *N*-гликозилирования [8] и 17 сайтов *O*-гликозилирования (наши данные).

В предполагаемой цитоплазматической *C*-концевой области белка LLP (сплайсинговые варианты AD и BD) присутствуют два протяженных глутаминогорьих участка и один – аспарагиногорьий, а также два потенциальных сайта фосфорилирования – cAMP- и cGMP-зависимых протеинкиназ и два – тирозинкиназы, которые могут играть важную роль в функционировании этого рецептора. У латрофилинов млекопитающих соответствующий участок содержит три потенциальных сайта фосфорилирования (два – cAMP/cGMP-зависимых протеинкиназ и один – тирозинкиназы).

В целом, полученные результаты свидетельствуют о существовании нескольких форм латрофилинподобных белков у домовой мухи и, возможно, у других насекомых. В то же время на основании этих данных не представляется возможным выяснить, являются ли эти белки рецепторами α -ЛТ.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Реактивы. Реактивы отечественного производства “ос. ч” или “ч. д. а.”: этанол, изопропанол, цитрат натрия, NaCl, LiCl. Реактивы импортного производства: Трис, *N*-лаурилсаркозин, Na-соль, глицерин, EDTA, меркаптоэтанол, бромистый этидий, агароза (Sigma, США), гуанидинитиоцианат (Fluka, Швейцария), агар-агар, бактотриптон, дрожжевой экстракт (Difco, США); дезокси-нуклеозидтрифосфаты, набор реагентов для получения первой цепи кДНК (Fermentas, Литва), обратная транскриптаза SuperscriptII (Gibco BRL, США), набор реагентов для выделения плазмид – Wizard (Promega, США), *Taq*-полимераза (ИБХ РАН), Platinum *Taq*-ДНК-полимераза (Gibco BRL), *Pfu*-полимераза (Stratagene, США), набор реагентов для выделения РНК – RNAwiz (Ambion, Канада), набор реактивов для хроматографии на колонке с oligo(dT)-сефарозой (НПАО “Силекс М”, Россия), полинуклеотидкиназа фага T4 (Amersham, США), щелочная фосфатаза из поджелудочной железы теленка (CIAP) (USB, США), ДНК-лигаза фага T4 (ИБХ РАН), рестриктазы *EcoRV*, *NcoI*, *XbaI* (Fermentas, Литва).

Выделение геномной ДНК из *M. domestica*. Исходный материал (личинки или головы взрослых мух ~30 мг) гомогенизировали в 1.5 мл центрифужных пробирках при помощи специального пестика в 300 мкл буфера D (4 М гуанидинийтиоцианат, 25 mM цитрат натрия, 0.5% N-лауроилсарказин, 0.1 M β-меркаптоэтанол). К гомогенату добавляли 300 мкл фенола, насыщенного Трис-HCl pH 8.0, интенсивно перемешивали и инкубировали при 40°C в течение 10 мин. Для разделения фаз к смеси добавляли 125 мкл хлороформа, интенсивно встряхивали на шейкере и центрифугировали при 14000 об/мин на центрифуге Eppendorf (Германия) в течение 2 мин. Водную фазу отбирали и повторяли обработку фенолом и хлороформом. ДНК из водной фазы осаждали 2 объемами 96% этанола. Очистку ДНК от примесей РНК производили с помощью 9 M LiCl. На конечной стадии ДНК растворяли в 50 мкл дистиллированной воды. Концентрацию ДНК определяли электрофоретически.

Суммарную РНК из *M. domestica* выделяли из голов мух при помощи набора реагентов RNAwiz (Ambion, Канада) в соответствии с рекомендациями фирмы-производителя.

Poly(A⁺)-РНК из *M. domestica* выделяли на колонках с oligo(dT)-сефарозой из образцов суммарной РНК при помощи набора реагентов фирмы “Силекс М” согласно протоколу фирмы-производителя.

Синтез первой цепи кДНК. Около 5 мкг суммарной РНК (1–4 мкл) подвергали отжигу с 0.5 мкг универсального или генспецифичного праймера в течение 5 мин при 70°C, затем охлаждали на льду. К полученной смеси добавляли 5-кратный реакционный буфер для обратной транскриптазы (конечная концентрация компонентов 50 mM Трис-HCl, pH 8.3, 50 mM KCl, 4 mM MgCl₂, 10 mM DTT), ингибитор рибонуклеаз RNAguard (Gibco BRL, США) до конечной концентрации 1 ед. акт./мкл и 10 mM смесь dNTP (конечная концентрация каждого компонента 1 mM). Смесь инкубировали в течение 5 мин при 37°C, затем добавляли 200 ед. акт. обратной транскриптазы SuperscriptII. Синтез первой цепи проводили в течение 1 ч при 42°C. Реакцию останавливали прогреванием при 70°C в течение 10 мин. Смесь разводили в три раза деионизованной водой и использовали в ПЦР.

Процедура ПЦР. Реакционная смесь (50 мкл) содержала буфер для ПЦР (50 mM KCl, 10 mM Трис-HCl, pH 8.3, 1.5–2.0 mM MgCl₂), смесь четырех dNTP (0.2 mM каждый), олигонуклеотидные праймеры (10–20 пмоль каждого), 1–10 нг ДНК или первой цепи кДНК и 2.5 ед. акт. термостабильной ДНК-полимеразы. Условия ПЦР: денатурация – 93°C, 30 с; отжиг – 55°–67°C, 30 с; элонгация – 72°C, 1–3 мин. Число циклов обычно составляло 37–40.

Электрофорез продуктов ПЦР проводили в 0.8–1.0% агарозном геле с использованием стан-

дарного Трис-ацетатного буфера (ТАЕ) с бромистым этидием [24].

Лигирование ДНК по тупым концам с векторной плазмидой pBSII-SK(+). Фрагменты ДНК – продукты ПЦР – обрабатывали полинуклеотидкиназой фага T4 в буфере для лигирования [24] при 37°C в течение 1 ч. Полинуклеотидкиназу инактивировали прогревом при 70°C в течение 15 мин. Далее в полученную смесь вносили ДНК векторной плазмиды pBSII-SK(+), предварительно обработанную рестриктазой EcoRV и щелочной фосфатазой, и 5 ед. акт. ДНК лигазы фага T4. Смесь инкубировали при 13.5°C в течение ночи. Для осаждения ДНК к лигированной смеси добавляли 1/10 объема 3 M ацетата натрия, 1 мкл раствора тРНК (препарата суммарной тРНК из *E. coli* в качестве соосадителя) (10 мг/мл) и 2.5 объема 96% этанола. После интенсивного встряхивания смесь выдерживали при –70°C в течение 1 ч и центрифугировали 5 мин при 14000 об/мин. Осадок ДНК промывали 1 мл 80% этанола, подсушивали и растворяли в 10 мкл диеонизованной воды.

Электротрансформация клеток *E. coli*. В кювету установки для электропорации CellJect (Eurogentec, Франция) вносили 40 мкл компетентных клеток *E. coli* TG-1 и к ним добавляли 5 мкл раствора ДНК (лигированной смеси). Параметры импульса выставляли согласно рекомендациям фирмы-производителя прибора.

Препартивное выделение плазмидных ДНК. Плазмидную ДНК выделяли из 5 млочной культуры *E. coli* при помощи набора реагентов Wizard согласно протоколу фирмы-производителя (Promega, США).

Секвенирование ДНК проводили с помощью набора реагентов ABI PRISM® BigDye™ Terminator v. 3.1 с последующим анализом продуктов реакции на автоматическом секвенаторе ДНК ABI PRISM 3100-A (США). Секвенирование осуществлялось в ИМБ РАН в группе А.Б. Полтарауса.

Компьютерный анализ последовательностей ДНК и белков. Процедуру выравнивания нуклеотидных последовательностей ДНК и аминокислотных последовательностей белков проводили с помощью пакета программ DNA Star и программ AliBee – Multiple alignment Release 2.0 (www.genebee.msu.su). Поиск по базам данных гена, кодирующего латрофилинподобный белок проводили, используя программы BLASTP 2.2.2 и BLASTN 2.2.1 (www.ncbi.nlm.nih.gov/BLAST). Анализ доменной организации белков, поиск сигнальных пептидов, потенциальных сайтов гликозилирования и фосфорилирования осуществляли с помощью программ Motif Scan in a Protein Sequence, SignalP, NetOGlyc, NetNGlyc (<http://expasy.org/tools/>), а также программы RPS-BLAST 2.2.2. (www.ncbi.nlm.nih.gov/BLAST).

Работа была поддержана грантами Российского фонда фундаментальных исследований (грант

№ 02-04-49294), Международного фонда CRDF (грант RB2-525), а также грантами по молекулярной и клеточной биологии РАН и Министерства науки и технологий Российской Федерации.

СПИСОК ЛИТЕРАТУРЫ

1. Tzeng M.C., Cohen R.S., Siekevitz P. // Proc. Natl. Acad. Sci. USA. 1978. V. 75. P. 4016–4020.
2. Tzeng M.C., Siekevitz P. // Brain Res. 1978. V. 139. P. 190–196.
3. Frontali N., Ceccarelli B., Gorio A., Mauro A., Siekevitz P., Tzeng M.C., Hurlbut W.P. // J. Cell Biol. 1976. V. 68. P. 462–479.
4. Ushkaryov Y.A., Petrenko A.G., Geppert M., Sudhof T.C. // Science. 1992. V. 257. P. 50–56.
5. Davletov B.A., Shamotienko O.G., Lelianova V.G., Grishin E.V., Ushkaryov Y.A. // J. Biol. Chem. 1996. V. 271. P. 23239–23245.
6. Krasnoperov V.G., Beavis R., Chepurny O.G., Little A.R., Plotnikov A.N., Petrenko A.G. // Biochem. Biophys. Res. Commun. 1996. V. 227. P. 868–875.
7. Krasnoperov V.G., Bittner M.A., Beavis R., Kuang Y., Salnikow K.V., Chepurny O.G., Little A.R., Plotnikov A.N., Wu D., Holz R.W., Petrenko A.G. // Neuron. 1997. V. 18. P. 925–937.
8. Lelianova V.G., Davletov B.A., Sterling A., Rahman M.A., Grishin E.V., Totty N.F., Ushkaryov Y.A. // J. Biol. Chem. 1997. V. 272. P. 21504–21508.
9. Davletov B.A., Meunier F.A., Ashton A.C., Matsushita H., Hirst W.D., Lelianova V.G., Wilkin G.P., Dolly J.O., Ushkaryov Y.A. // EMBO J. 1998. V. 17. P. 3909–3920.
10. Geppert M., Khvotchev M., Krasnoperov V., Goda Y., Missler M., Hammer R.E., Ichtchenko K., Petrenko A.G., Sudhof T.C. // J. Biol. Chem. 1998. V. 273. P. 1705–1710.
11. Ichtchenko K., Khvotchev M., Kiyatkin N., Simpson L., Suga S., Sudhof T.C. // EMBO J. 1998. V. 17. P. 6188–6199.
12. Volynski K.E., Meunier F.A., Lelianova V.G., Dudina E.E., Volkova T.M., Rahman M.A., Manser C., Grishin E.V., Dolly J.O., Ashley R.H., Ushkaryov Y.A. // J. Biol. Chem. 2000. V. 275. P. 41175–41183.
13. Orlova E.V., Rahman M.A., Gowen B., Volynski K.E., Ashton A.C., Manser C., van Heel M., Ushkaryov Y.A. // Nat. Struct. Biol. 2000. V. 7. P. 48–53.
14. Khvotchev M., Sudhof T.C. // EMBO J. 2000. V. 19. P. 3250–3262.
15. Capogna M., Volynski K.E., Emptage N.J., Ushkaryov Y.A. // J. Neurosci. 2003. V. 23. P. 4044–4053.
16. Volynski K.E., Capogna M., Ashton A.C., Thomson D., Orlova E.V., Manser C.F., Ribchester R.R., Ushkaryov Y.A. // J. Biol. Chem. 2003. V. 278. P. 31058–31066.
17. Saeger B., Schmitt-Wrede H.-S., Dehnhardt M., Benten W.P.M., Krucken J., von Samson-Himmelstjerna G., Wiegand H., Wunderlich F. // FASEB J. 2001. V. 15. P. 1332–1334.
18. Mee C.J., Tomlinson S.R., Perestenko P.V., De Pomerai D., Duce I.R., Usherwood P.N., Bell D.R. // Biochem. J. 2004. V. 378. P. 185–191.
19. Magazanik L.G., Fedorova I.M., Kovalevskaya G.I., Pashev V.N., Bulgakov O.V., Grishin E.V. // Neuroscience. 1992. V. 46. P. 181–188.
20. Matz M., Shagin D., Bogdanova E., Britanova O., Lukyanov S., Diatchenko L., Chenchik A. // Nucleic Acids Res. 1999. V. 27. P. 1558–1560.
21. Kozak M. // Cell. 1986. V. 44. P. 283–292.
22. Kozak M. // J. Mol. Biol. 1987. V. 196. P. 947–950.
23. Kozak M. // Mol. Cell. Biol. 1989. V. 9. P. 5073–5080.
24. Maniatis T., Fritsch E.F., Sambrook J. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, N.Y.; Cold Spring Harbor Lab. Press, 1989.

Alternative Splicing of pre-mRNA Encoding the *Musca domestica* Latrophilin-Like Protein: Primary Structures of Four Spliced Forms of mRNA and Their Protein Products

Ya. A. Andreev, V. N. Danilevich*, and E. V. Grishin

*Phone: +7 (095) 336-6540; fax: +7 (095) 336-7301; e-mail: dan@mail.ibch.ru
Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences,
ul. Miklukho-Maklaya 16/10, Moscow, 117997 Russia

An internal DNA fragment (~2000 bp) homologous to the conserved regions of genes encoding latrophilin-like proteins (LLPs) was obtained by the PCR technique using degenerate primers to these gene regions. The gene-specific primers were synthesized based on the results of sequencing of the isolated fragment, and all overlapping cDNA fragments of the *llp* gene encoding the *Musca domestica* LLP were obtained by the rapid amplification of cDNA 5'- and 3'-ends (5'- and 3'-RACE). Four alternatively spliced mRNAs were found while sequencing the obtained cDNA fragments. Two long mRNAs (~6000 nt) differ in the structures of both the sites encoding signal peptides and 5'-terminal untranslated regions. They encode large proteins (~1800 aa), whose domain organization is similar to that of mammalian latrophilins. Each deduced protein contains a domain with seven transmembrane regions followed by an extended cytoplasmic C-terminal domain. Two other mRNA forms are derived from these long mRNAs; they encode proteins severely truncated at their C-termini (~900 aa). They are composed of only three transmembrane regions and a short unique cytoplasmic C-terminal domain (23 aa). The limitations and drawbacks of the existing 3'-RACE techniques found during study of the long alternatively spliced cDNAs are analyzed, and ways for overcoming these difficulties are proposed. The English version of the paper: *Russian Journal of Bioorganic Chemistry*, 2005, vol. 31, no. 2; see also <http://www.maik.ru>.

Key words: alternative splicing, latrophilin-like protein, *Musca domestica*, PCR amplification, protein primary structure, 3'- and 5'-RACE, sequencing