

УДК 577.21

## МАЛ, ДА УДАЛ. РАЗНООБРАЗИЕ ФУНКЦИЙ МАЛЫХ БЕЛКОВ И ПЕПТИДОВ – ПРОДУКТОВ ТРАНСЛЯЦИИ КОРОТКИХ РАМОК СЧИТЫВАНИЯ

### Обзор

© 2021 П.В. Сергиев<sup>1,2,3\*</sup>, М.П. Рубцова<sup>1\*</sup>

<sup>1</sup> Московский государственный университет им. М.В. Ломоносова, химический факультет, 119991 Москва, Россия; электронная почта: petya@genebee.msu.ru, mprubtsova@gmail.com

<sup>2</sup> Центр наук о жизни Сколковского института науки и технологий, 143025 дер. Сколково, Россия

<sup>3</sup> Институт функциональной геномики Московского государственного университета им. М.В. Ломоносова, 119991 Москва, Россия

Поступила в редакцию 27.04.2021

После доработки 27.05.2021

Принята к публикации 27.05.2021

Функционирование клетки является сложно регулируемым процессом. Долгие годы исследования в областях протеомики и функциональной геномики были сфокусированы на роли белков в жизнедеятельности клетки. Развитие науки привело к пониманию того, что короткие открытые рамки считывания, ранее считавшиеся нефункциональными, в действительности выполняют разнообразные функции. В составе полицистронных мРНК короткие рамки считывания часто регулируют их стабильность и эффективность трансляции основной рамки считывания. Совершенствование методов протеомного анализа позволило идентифицировать продукты трансляции коротких рамок считывания в количествах, позволяющих предположить существование функциональной роли у синтезируемых пептидов и коротких белков. Их изучение открывает новый уровень регуляции функционирования клетки и ее адаптации к изменяющимся условиям. Настоящий обзор посвящен анализу функций недавно открытых пептидов и коротких белков.

**КЛЮЧЕВЫЕ СЛОВА:** пептиды, малые рамки считывания, длинные некодирующие РНК, альтернативные белки.

DOI: 10.31857/S0320972521090098

### ВВЕДЕНИЕ

Разнообразие белков предопределяет функционирование живых существ, а для таких сложных организмов, как млекопитающие, даже простая каталогизация протеома является нетривиальным и далеко не законченным процессом. В результате развития системной биологии стало очевидным, что транскрипции подвергается много больше участков генома, чем предполагалось ранее [1, 2]. С возникновением методов высокопроизводительного секвенирования РНК [3] описание транскриптов достигло небывалых ранее масштабов. Поскольку многие из выявленных транскриптов не содержали протяженных рамок считывания, их в большинстве своем аннотировали как длинные некодирующие РНК (см. [4, 5] для обзора). Короткие рамки считывания, обнаруживаемые в этих тран-

скриптах, на первых порах считали случайностью и игнорировали. С развитием методов системного исследования биосинтеза белка в клетках, в первую очередь рибосомного профилирования [6], а также в результате совершенствования методов протеомного анализа [7–10] научное сообщество с удивлением обнаружило, что и трансляция транскриптома не ограничена каноническими длинными рамками считывания, начинающимися непременно с инициаторного AUG-кодона. Результатом такой «всеобщей» трансляции являются короткие белки и пептиды (sPEP), продукты трансляции т.н. коротких открытых рамок считывания (кОРС), а также длинные белки, необычные продукты трансляции известных мРНК, часто называемые альтернативными белками (alt-PROT), или некодирующих РНК.

Не акцентируя чрезмерно внимание на современных методах выявления кодирующих последовательностей в геномах и транскриптомах, авторы рекомендуют заинтересованным в методологической стороне вопроса читателям обра-

Принятые сокращения: НТО – нетранслируемая область; ОРС – открытая рамка считывания.

\* Адресат для корреспонденции.

таться к нескольким обзорным статьям на эту тему [11, 12]. В данном обзоре мы остановимся на результатах исследований последних лет, посвященных поиску функциональной роли малых белков.

### ГДЕ НАХОДЯТ НОВЫЕ ТРАНСЛИРУЕМЫЕ ОРС?

Где же закодированы неизвестные ранее белки, обнаруживаемые с помощью современных компьютерных и экспериментальных методов, и в чем их особенности, позволившие им оставаться в тени столько времени? Во-первых, большая часть этих белков, хотя и не все, имеет небольшие размеры, до 100 аминокислотных остатков (а.о.). Поскольку в любой некодирующей последовательности можно встретить случайно присутствующие нефункциональные короткие рамки считывания, исследователю достаточно сложно выявить ту небольшую долю рамок считывания, которые транслируются с образованием функциональных коротких белков и пептидов. Во-вторых, новые белки образуются при трансляции, которая начинается с неканонических инициаторных кодонов. Так, помимо AUG, который остается наиболее распространенным инициаторным кодоном (с него начинается трансляция в 64% случаев, по данным [13], в 43% – по данным [9]), у эукариот трансляция часто начинается с CUG (в 3% – по данным [9]; в 23% – [13]), GUG (в 6% – [9]; в 9% – [13]) и UUG (в 4% – [13]) кодонов. Довольно большой разброс частот использования различных кодонов в качестве инициаторных объясняется использованием разных выборок ОРС, содержащих разные соотношения канонических и альтернативных белков, а также, возможно, различным процентом включения ложноположительных результатов. Подробное сопоставление обсуждаемых частот использования кодонов для инициации трансляции можно найти в отдельной работе [14].

В-третьих, неизвестные ранее белки и пептиды могут быть закодированы в мРНК ранее аннотированных генов так, что их рамки считывания располагаются в 5'-нетранслируемой области (5'-НТО) или 3'-НТО, а также в рамке считывания, частично или полностью перекрывающейся, но не совпадающей с ранее аннотированной. В этой группе наиболее часто встречаются ОРС в 5'-НТО – по разным данным, от 40 [9] до 68% [13] всех неаннотированных ранее ОРС. Такой высокий процент может объясняться тем, что сам факт считывания 5'-ОРС, даже если продукт этого процесса нефункционален,

может служить механизмом регуляции трансляции «основной» рамки считывания. Этот процесс хорошо изучен для мРНК *GCN4* дрожжей, содержащей несколько 5'-ОРС, наличие которых обеспечивает селективную трансляцию этой мРНК в условиях аминокислотного голодания [15]. В дальнейших исследованиях было показано, что подобный тип регуляции очень распространен [16]. В случае 5'-ОРС *MIEF1*, *DDIT3*, *FBXO9*, *HMG2* и *HAUS6* обнаружено взаимодействие продуктов трансляции «основной» рамки считывания и ОРС, расположенной ближе к 5'-концу соответствующих мРНК [13]. В этих случаях, возможно, расположение двух рамок на одной мРНК необходимо для координированного синтеза соответствующих белков, как это происходит в случае бактериальных оперонов. При этом остается открытым вопрос о механизмах эффективной трансляции обеих рамок считывания в системе трансляции эукариот.

Можно предположить, что повышенная частота использования отличающихся от AUG кодонов в качестве инициаторных для 5'-ОРС имеет «механистическое» объяснение. Использование AUG-кодона в данном случае приведет к стабильному ингибированию трансляции основной рамки. Против такого варианта, если он возникнет в ходе случайной мутации, будет действовать отбор, связанный с функционированием основной ОРС. Напротив, использование инициаторного кодона, который может с некоторой вероятностью пропускаться сканирующим комплексом, будет закрепляться отбором, поскольку способствует синтезу обоих белковых продуктов, закодированных в мРНК [17].

Значительная доля выявленных функциональных ОРС, по некоторым данным – 24% [13], находится в транскриптах, аннотированных ранее как некодирующие. При этом новые транслируемые ОРС обнаруживаются как в транскриптах, функция которых ранее не была известной [18–24], так и в РНК, имеющих другую охарактеризованную функцию, например, белок TERP, закодированный в теломеразной РНК [25], Humanin – в митохондриальной 16S рРНК [26], E-пептид – в бактериальной 23S рРНК [27]. Описаны случаи, когда новые функции имеет и длинная некодирующая РНК, и кодируемый ею белок. Например, РНК *AW112010* кодирует небольшой белок, стимулирующий синтез IL-12p40. Мутантные мыши с нарушением синтеза пептида, закодированного в *AW112010*, демонстрируют большую скорость распространения сальмонелл и большую бактериоциемия, а также ускоренную смертность от сальмонеллеза [22]. В последующей работе обнаружена самостоятельная роль РНК *AW112010*,

закрывающаяся в привлечении KDM5A и последующем деметилировании остатков H3K4me3 гена, кодирующего IL-10 [28]. Сходным образом, после обнаружения функциональной роли пептида SPAR (другое название – SPAAR), закодированного в *LINC00961*, заключающейся в подавлении активности mTORC1-киназного комплекса [24] (см. более подробное описание ниже), выявлено, что *LINC00961* как РНК имеет отдельную функцию в ангиогенезе, причем ее активность противоположна таковой для пептида SPAR [29].

Транслируемые ОРС находят в таких необычных РНК, как кольцевые (circRNA), образующиеся в результате ошибочного сплайсинга [30–32] (см. обзор [33]), и примикроРНК (pri-miRNA) [34]. В случае кольцевых РНК встает вопрос о механизме инициации трансляции, поскольку классическая модель сканирования, очевидно, для них не подходит. Авторы соответствующих работ объясняют взаимодействие рибосом с кольцевыми РНК-участками внутренней посадки рибосом (IRES, internal ribosome entry site) и привлечением инициаторных факторов модифицированными остатками m<sup>6</sup>A; предложенные модели требуют дальнейшей тщательной проверки (см. рассуждения на эту тему в [35]). Таким образом, разнообразие кодирующих РНК значительно выходит за рамки наших обычных представлений. Это, с одной стороны, ошеломляет, а с другой стороны, вызывает определенную долю скепсиса.

В чем же может заключаться функциональная роль продуктов трансляции недавно открытых ОРС, малых и альтернативных белков? Хотя общего ответа на этот вопрос нет, мы рассмотрим несколько интересных частных случаев, описанных в современной научной литературе.

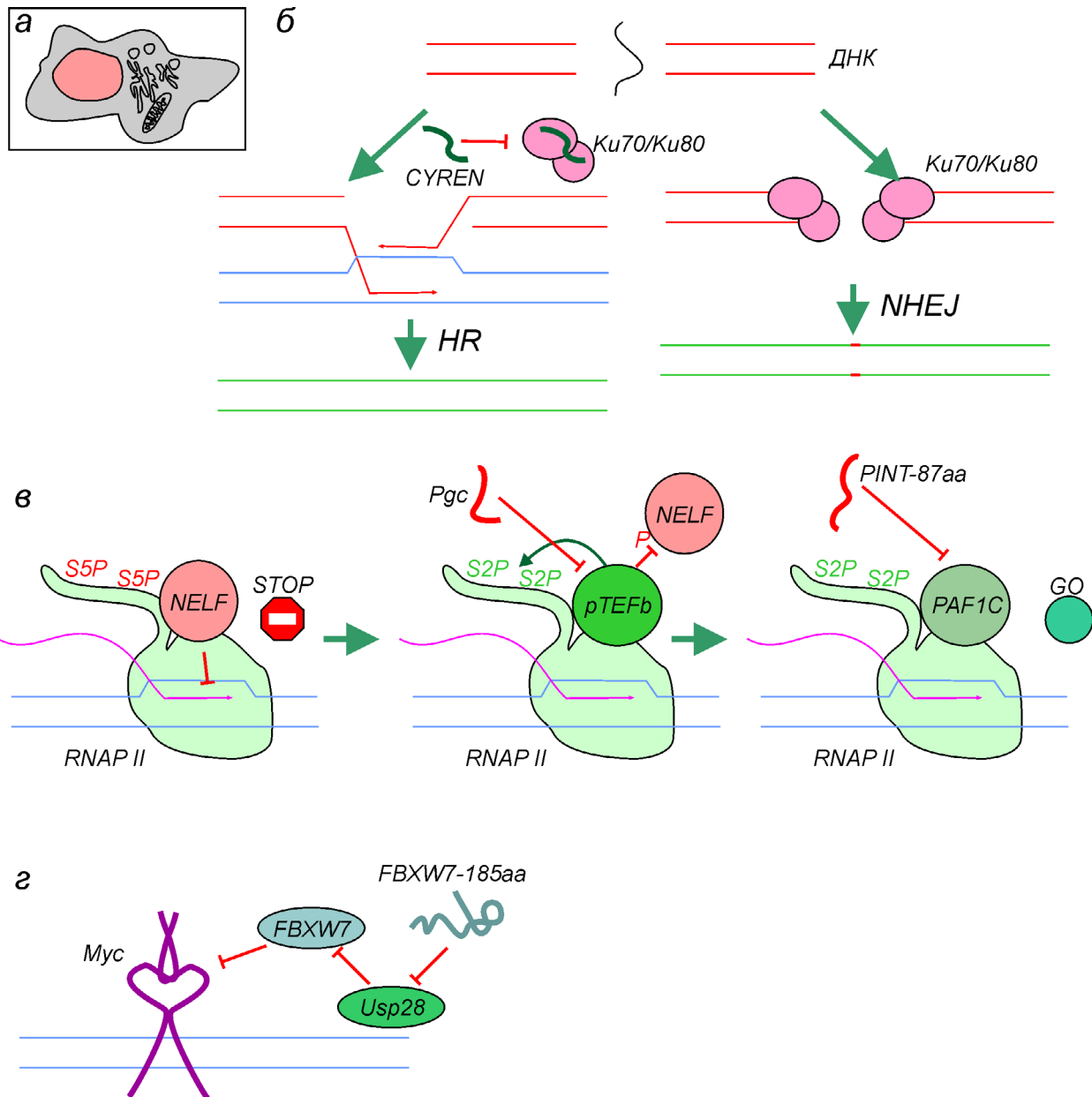
### РОЛЬ МАЛЫХ БЕЛКОВ В БИОСИНТЕЗЕ ДНК И РНК

Практически во всех экспериментах по исследованию функциональной роли вновь открываемых малых белков определяют их внутриклеточную локализацию. Например, в работе Chen et al. [13] локализацию малых белков изучали с помощью их гибридов с пептидным «довеском» зеленого флуоресцентного белка GFP11, который при взаимодействии с оставшейся частью GFP1-10 образует функциональный флуорофор, детекцию которого можно осуществить при помощи микроскопии. Так, например, обнаружена ядерная локализация продукта трансляции 5'-ОРС *HMG2*. Хотя функциональная роль этого пептида не определена, в науч-

ной литературе имеются сведения о других, более тщательно исследованных малых белках, чья функция в ядре связана с репликацией и репарацией ДНК, а также с процессом транскрипции (рис. 1, а).

Белок SHPRH-146aa, закодированный в кольцевой РНК, проявляет онкосупрессорные свойства, способствуя деградации необходимого для пролиферации ядерного белка PCNA, обеспечивающего процессивность репликативных ДНК-полимераз [30]. AltMRV11 длиной 134 а.о. закодирован в 3'-НТО гена *MRV11*. Продукт трансляции этой рамки имеет ядерную локализацию и взаимодействует с BRCA1 [36], белком-онкосупрессором, участвующим в репарации разрывов ДНК. Интересно, что ранее это взаимодействие определили с помощью дрожжевой двугибридной системы, но посчитали ложноположительным [37]. Кроме AltMRV11, в регуляции репарации двуцепочечных разрывов принимает участие пептид CYREN-2 [38], по сути, являющийся короткой изоформой ранее аннотированного белка CYREN/MRI (модулятор ретровирусной инфекции) [39], образующейся в результате альтернативного сплайсинга. Обе формы CYREN взаимодействуют с Ku70/Ku80-субъединицами ДНК-зависимой протеинкиназы, участвующей в репарации двуцепочечных разрывов ДНК по механизму негомологичного соединения концов (NHEJ, nonhomologous end joining). Известно, что работа NHEJ может вызывать мутации, но не имеет альтернативы в G1-фазе клеточного цикла, когда ДНК еще не реплицирована с образованием сестринских хроматид. В поздней S- и G2-фазах, когда двуцепочечные разрывы ДНК могут быть исправлены с помощью более точного механизма гомологичной рекомбинации (HR, homologous recombination), активность NHEJ подавляется, в том числе и с помощью прямого взаимодействия CYREN и Ku70/Ku80 (рис. 1, б). При инактивации CYREN повышается частота слияния теломерных участков хромосом [38], а инактивация CYREN/MRI вместе с одним из основных участников системы NHEJ – XLF – приводит к эмбриональной летальности мышей из-за полного блокирования системы NHEJ [40].

Согласно нескольким недавним публикациям, продукты трансляции малых рамок считывания участвуют в регуляции транскрипции. Например, пептид Pgc, состоящий из 71 а.о., подавляет транскрипцию в предшественниках зародышевых клеток дрозофилы, взаимодействуя с киназой pTEFb (рис. 1, в). pTEFb осуществляет фосфорилирование белка NELF, а также остатка Ser2 C-концевого домена РНК-полимеразы II [41]. Подобное фосфорилирование приво-



**Рис. 1.** Функционирование малых белков в ядре, участие в процессах репарации ДНК и транскрипции. *а* – Схема строения эукариотической клетки. Цветом выделено ядро, в котором осуществляется функционирование показанных на рисунке белков. *б* – Пептид CYREN ингибирует репарацию двуцепочечных разрывов ДНК по пути негомологичного соединения концов ДНК (NHEJ, nonhomologous end joining), подавляя активность Ku70/Ku80 (на рисунке HR – homologous recombination). *в* – Остановка РНК-полимеразы II, вызванная белком NELF и обусловленная фосфорилированием остатков Ser5 (S5P) гептааминокислотных повторов С-концевого домена (CTD) RNAP II (слева). Киназный комплекс рTEFb фосфорилирует NELF и остатки Ser2 (S2P) гептааминокислотных повторов CTD RNAP II, что обуславливает связывание PAF1C и переход к процессивной элонгации (по центру). Пептид Pgc ингибирует рTEFb, а пептид PINT-87aa подавляет активность PAF1C, что приводит к репрессии транскрипции (справа). *г* – Влияние белка FBXW7-185 aa на стабильность транскрипционного фактора Мус. FBXW7-185 aa конкурирует с FBXW7 за взаимодействие с Usp28-деубиквитиназой, что приводит к повышению активности FBXW7 и дестабилизации Мус

дит к снятию ингибирования и увеличивает эффективность транскрипции. После действия рTEFb активность РНК-полимеразы II увеличивается под действием другого фактора элонгации, PAF1c [42]. Пептид PINT-87aa, закодиро-

ванный в кольцевой РНК, проявляет онкосупрессорные свойства, взаимодействуя с PAF1c (рис. 1, *в*) и замедляя таким образом транскрипцию онкогенов [31]. Еще одним малым белком, взаимодействующим с транскрипционным фак-

тором и проявляющим онкосупрессорные свойства, является FBXW7-185 aa [43]. FBXW7-185 aa конкурирует с полноразмерным FBXW7 за взаимодействие с Usp28-деубиквитиназой, стабилизирующей транскрипционный фактор Мус. Таким образом, действие FBXW7-185 aa противоположно активности FBXW7 и приводит к дестабилизации Мус (рис. 1, з).

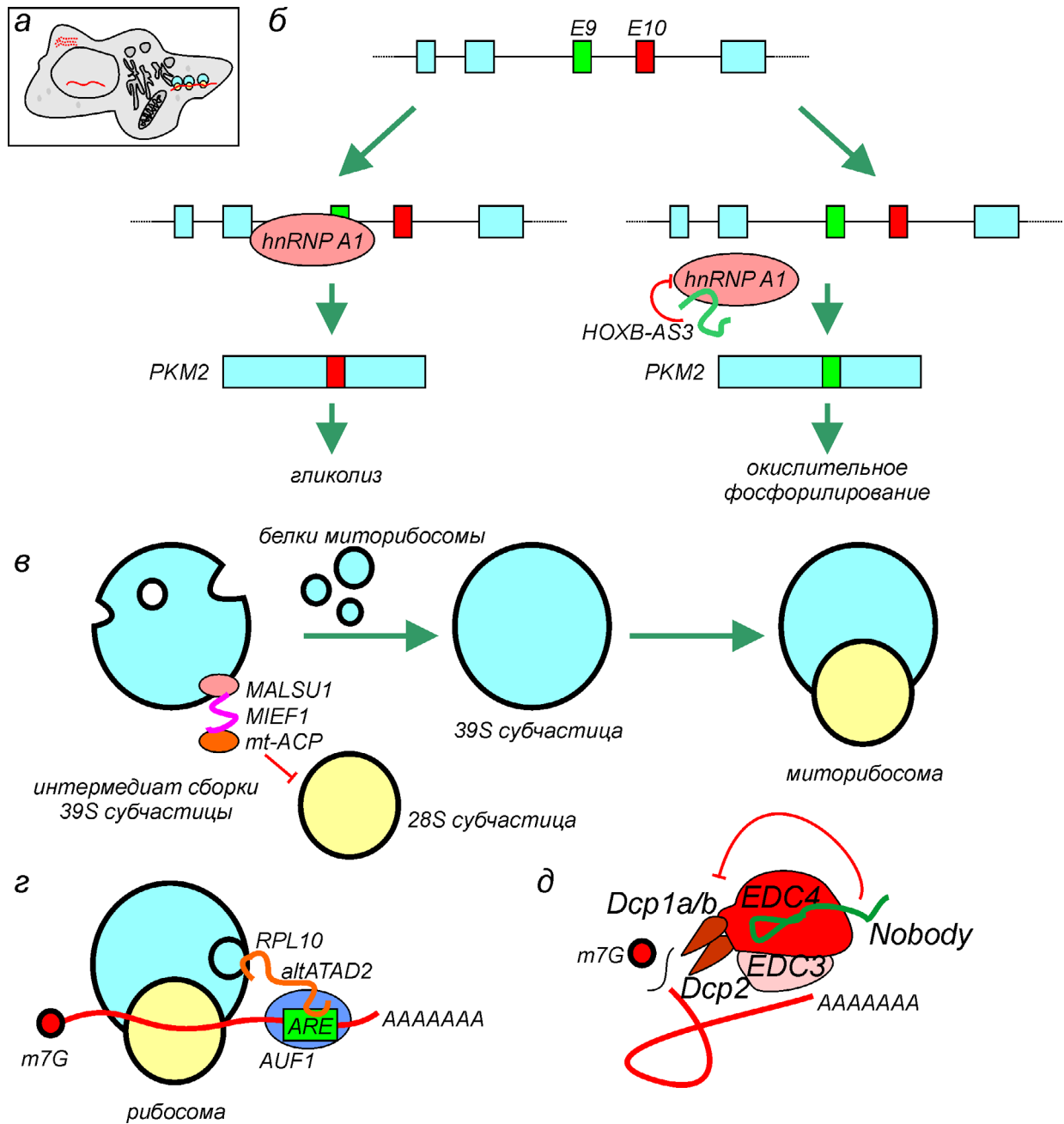
### РОЛЬ МАЛЫХ БЕЛКОВ В СОЗРЕВАНИИ РНК, СИНТЕЗЕ И СВРАЧИВАНИИ БЕЛКОВ

В научной литературе можно найти множество примеров регуляции созревания мРНК и трансляции с помощью малых белков (рис. 2, а). Некоторые из этих исследований демонстрируют пока довольно поверхностные данные. Так, описано взаимодействие белков altSETD1B и altLINC00624 с аминоксил-тРНК синтетазами [44], однако функциональные последствия этих взаимодействий исследованы не были. Значительно более подробно исследованы функции других малых белков в процессинге РНК и биосинтезе белка. Пептид, состоящий из 53 а.о., закодированный в транскрипте *HOXB-AS3*, взаимодействует с hnRNP A1 и рядом других факторов процессинга РНК [45]. Экспрессия *HOXB-AS3* падает в клетках раковых опухолей, например, при раке прямой кишки. Подобное снижение экспрессии предвещает плохой прогноз для пациентов. Инактивация рамки считывания при сохранении большей части РНК *HOXB-AS3* приводит к повышению колониеобразования, стимуляции роста и миграции клеток. Эктопическая экспрессия данного гена в клетках рака прямой кишки, соответственно, наоборот, замедляет пролиферацию клеток. Онкосупрессорная природа *HOXB-AS3* объясняется его антагонизмом с hnRNP A1. Пептид *HOXB-AS3* «отменяет» включение экзона 10 в мРНК пируваткиназы, обусловленное активностью hnRNP A1. Фетальная сплайс-форма пируваткиназы содержит экзон 10 и участвует в аэробном гликолизе, в то время как сплайс-форма, содержащая экзон 9, экспрессируется в тканях взрослого организма и участвует в окислительном фосфорилировании. Раковые клетки, возвращаясь к гликолитическому метаболизму в результате эффекта Варбурга, начинают экспрессировать фетальную форму пируваткиназы [46], что обусловлено активностью hnRNP A1 [47]. Пептид *HOXB-AS3* подавляет эффект Варбурга и рост клеток опухоли, являясь антагонистом hnRNP A1 и способствуя экспрессии пируваткиназы с включенным экзон 9 [45] (рис. 2, б).

Помимо участия в созревании мРНК, малые белки принимают участие в созревании рибосом, одним из них является пептид MIEF1-MP/L0R8F8 длиной 70 а.о. Этот пептид закодирован в 5'-ОПС гена белка MID51. «Основной» продукт трансляции этой мРНК, MID51, локализуется во внешней мембране митохондрий и регулирует их слияние [48]. Функциональность соответствующей 5'-ОПС, *MIEF1-MP/L0R8F8*, предсказана на основе данных рибосомного профилирования и анализа консервативности последовательности [16]. Дальнейшее исследование показало, что *MIEF1-MP* взаимодействует с миторибосомами и способствует митохондриальной трансляции [49]. Суперэкспрессия *MIEF1-MP* повышает, а нокдаун – снижает эффективность митохондриальной трансляции практически всех белков, синтезируемых миторибосомами. Примечательно, что уровень трансляции *MIEF1* превышает таковой для основной аннотированной рамки этого гена, белка MID51 [50]. Пространственное расположение пептида MIEF1-MP/L0R8F8 в структуре интермедиата сборки миторибосомы определено в лаборатории В. Рамакришнана с помощью криоэлектронной микроскопии [51]. В этой структуре MIEF1-MP соединяет гомолог бактериального фактора сборки рибосом RsfS, MALSU1, и митохондриальный носитель жирных кислот mt-ACP. Предположительно, эта белковая пирамидка предотвращает преждевременное взаимодействие интермедиата сборки большой субъединицы миторибосомы с малой субъединицей (рис. 2, в).

Взаимодействие белка altATAD2 длиной 139 а.о. с цитоплазматическим рибосомным белком RPL10 обнаружено с помощью пришивок с последующей масс-спектрометрической идентификацией [52]. Помимо взаимодействия с рибосомным белком, обнаружен контакт altATAD2 с AUF1/hnRNP D, узнающим AU-богатые последовательности РНК. Как предполагают исследователи, altATAD2 может служить посредником в привлечении рибосом на мРНК, имеющие AU-богатые нуклеотидные последовательности, хотя молекулярный механизм функционирования altATAD2 пока так и остается гипотетическим (рис. 2, г).

Дегградация мРНК, несущих AU-богатые последовательности и другие дестабилизирующие сигналы, происходит в специализированных компартментах цитоплазмы, называемых Р-тельцами (P-bodies). В этих структурах локализован продукт трансляции *LINC01420* – пептид Nobody длиной 68 а.о. Анализ белок-белковых взаимодействий выявил его ассоциацию с аппаратом декэпирования, причем прямое взаи-



**Рис. 2.** Функционирование малых белков в процессах созревания, трансляции и деградации РНК в клетке. *а* – Схема локализации процессов созревания, трансляции и деградации РНК в клетке. *б* – Регулируемый альтернативный сплайсинг пре-мРНК пируваткиназы. В эмбриональных и раковых клетках повышенная концентрация hnRNP A1 подавляет включение экзона 9 и способствует включению экзона 10, что приводит к синтезу PKM2-изоформы. В большинстве клеток взрослого организма пептид HOXB-AS3 подавляет активность hnRNP A1, что приводит к включению экзона 9, но не 10 – и появлению изоформы PKM1, способствующей окислительному фосфорилированию. *в* – Пептид MIEF1 вместе с MALSU1 и mt-ACP взаимодействует с интермедиатом сборки большой субчастицы (39S) митохондриальной рибосомы, чтобы не допустить преждевременного взаимодействия с малой (28S) субчастицей. *г* – Белок altATAD2 взаимодействует с AUF1, узнающим мРНК с AU-богатыми последовательностями (ARE), и рибосомным белком RPL10 (на рисунке m7G – 5'-кэп). *д* – Пептид Nobody напрямую взаимодействует с EDC4 и опосредованно – с другими компонентами комплекса, декэпирующего мРНК, в результате чего подавляет их активность

модействие Nobody с EDC4 важно для опосредованного сближения этого пептида с EDC3, Dcp1A, Dcp1B и Dcp2. При этом Nobody является отрицательным регулятором сборки Р-телца,

которые исчезают при его суперэкспрессии (рис. 2, д). Одновременно с исчезновением Р-телца при суперэкспрессии Nobody в клетке возрастает количество мРНК, являющихся субстратом

нонсенс-опосредованной деградации (NMD, nonsense-mediated mRNA decay) [53].

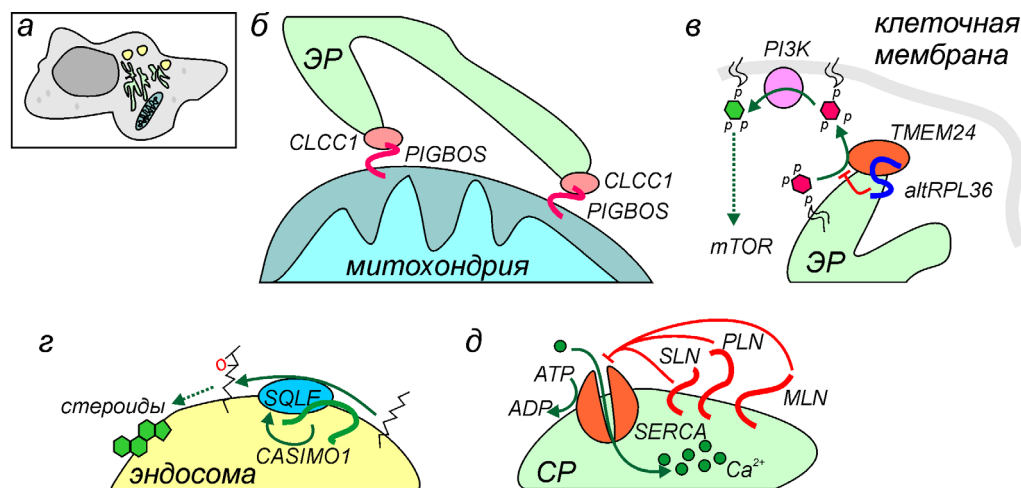
После синтеза белки часто становятся субстратами шаперонов, которые способствуют формированию правильной пространственной структуры. Роль малых белков в этом процессе также имеет существенное значение. Так, недавно определена функциональная роль пептида, закодированного в 5'-ОПС мРНК *ASNSD1*, названного ASDURF [54]. Как оказалось, этот пептид длиной 96 а.о. является компонентом комплекса шаперонов PAQсомы, похожего на комплекс префолдинов. Наличие ASDURF и его взаимодействие с другими компонентами PAQсомы объясняет существовавшее ранее противоречие в предполагаемом белковом составе PAQсомы – тогда были известны лишь 5 субъединиц этого комплекса, при том, что сходные структуры содержат, как правило, 6 субъединиц. Как компонент PAQсомы, ASDURF принимает участие в формировании структур мРНК и компонентов сигнального пути mTOR.

### МАЛЫЕ БЕЛКИ В СИСТЕМЕ ВЕЗИКУЛЯРНОГО ТРАНСПОРТА

Значительная доля малых белков и пептидов имеет трансмембранные сегменты и встраивает-

ся в различные мембранные структуры клетки и/или взаимодействует с мембранными белковыми комплексами (рис. 3, а). Например, в ходе скринингового исследования функциональной роли и белковых партнеров малых белков выявлено, что пептид, закодированный в *LINC00998*, совыделяется с группой белков эндоплазматического ретикулума, а инактивация этого гена приводит к изменению, по-видимому, компенсаторному, экспрессии генов, имеющих отношение к процессу гликозилирования, также происходящему в данном клеточном компартменте [13]. Согласно той же работе [13], пептид длиной 70 а.о., закодированный в гене *RP11-84A1.3*, локализуется в цитоплазматической мембране и взаимодействует с группой белков той же локализации.

Пептид длиной 54 а.о. закодирован в РНК, транскрибирующейся с цепи ДНК, комплементарной гену *PIBG*, в связи с чем получил курьезное название PIGBOS [21]. Исследование этого пептида показало, что он локализуется на внешней мембране митохондрии и взаимодействует с белком CLCC1, расположенном в эндоплазматическом ретикулуме (ЭР), в местах контакта ЭР с митохондриями. Взаимодействие с CLCC1 важно для функционирования PIGBOS. Нокаут *PIGBOS*, как и нокаут *CLCC1*, приводит к усугублению стресса ЭР, ответу развернутого белка (UPR, unfolded protein response) при обработке



**Рис. 3.** Функционирование малых белков в системе везикулярного транспорта. *а* – Схема клетки с выделением органелл, показанных на следующих панелях. *б* – Взаимодействие пептида PIGBOS, локализованного во внешней митохондриальной мембране в местах контакта эндоплазматического ретикулума (ЭР) и митохондрии с белком ЭР CLCC1. Данное взаимодействие предотвращает стресс ЭР (UPR, unfolded protein response). *в* – Белок altRPL36, локализованный в местах контакта ЭР и цитоплазматической мембраны, взаимодействует с TMEM24 и ингибирует перенос PI(4,5)P2 из ЭР на внешнюю мембрану. Таким образом, это ингибирование приводит к активации mTOR-киназы через образование PI(3,4,5)P3 (фосфатидилинозитолтрифосфата) киназой PI3K. *г* – Расположенный в мембране эндосом пептид CASIMO1 стимулирует сквален-эпоксидазу SQLE, способствуя повышенному синтезу различных стероидных соединений. *д* – Пептиды фосфоламбан (PLN), сарколипин (SLN) и миорегулин (MLN), локализованные в саркоплазматическом ретикулуме (CP) различных типов мышц, подавляют активность АТФ-зависимой кальциевой помпы SERCA

туникамицином и усилению апоптоза, вызванного этим стрессом (рис. 3, б).

Белок altRPL36 закодирован в мРНК рибосомного белка RPL36, причем область, кодирующая альтернативный белок, начинается раньше, а заканчивается позже «основной» ОРС. Поскольку эти кодирующие области сдвинуты друг относительно друга со сдвигом рамки считывания (на некротное трем число нуклеотидов), соответствующие аминокислотные последовательности не похожи друг на друга. В отличие от RPL36, являющегося компонентом рибосомы, altRPL36 взаимодействует с трансмембранным белком ЭР TMEM24, локализованным в местах контактов ЭР и цитоплазматической мембраны. AltRPL36 ингибирует активность TMEM24, направленную на транспорт фосфатидилинозитолдифосфата, PI(4,5)P<sub>2</sub>, из ЭР во внешнюю мембрану. Функция как altRPL36, так и TMEM24 состоит в регулировании фосфорилирования. Инактивация altRPL36 приводит к активации TMEM24, увеличению концентрации субстрата киназы PI3K, PI(4,5)P<sub>2</sub>, и, таким образом, к активации mTOR сигнализации (рис. 3, в).

В эндосомах локализован пептид длиной 83 а.о., CASIMO1, который является продуктом трансляции рамки считывания, находящейся в небольшой РНК NR\_029455. Экспрессия этого гена возрастает в клетках рака молочной железы с высоким уровнем экспрессии ядерных рецепторов стероидных гормонов ER/PR [55]. Инактивация кодирующего CASIMO1 гена замедляла, а повышенная экспрессия — ускоряла рост клеток. Эксперименты по коиммунопреципитации продемонстрировали, что CASIMO1 взаимодействует с другим онкогенным белком, сквален-эпоксидазой SQLE, и стимулирует ее активность. Повышенную экспрессию SQLE также наблюдают в клетках рака молочной железы. По всей видимости, функция CASIMO1 заключается в стимуляции активности сквален-эпоксидазы (рис. 3, г).

Группа малых белков встречается в специализированном типе ЭР, саркоплазматическом ретикулуме, клеток различных типов мышц. Пептиды фосфоламбан (PLN) [56, 57] длиной 52 а.о. и сарколипин (SLN) [58] длиной 31 а.о. присутствуют в сердечной мышце и медленных мышцах соответственно. Их активность направлена на ингибирование АТФ-зависимой кальциевой помпы саркоплазматического ретикулума SERCA (для обзора см., например, [59]), работа которой способствует расслаблению мышц после сокращения. SERCA также может генерировать тепло при ингибировании транспорта ионов кальция без снижения ее АТФазной активности. SLN, взаимодействуя с SERCA, может

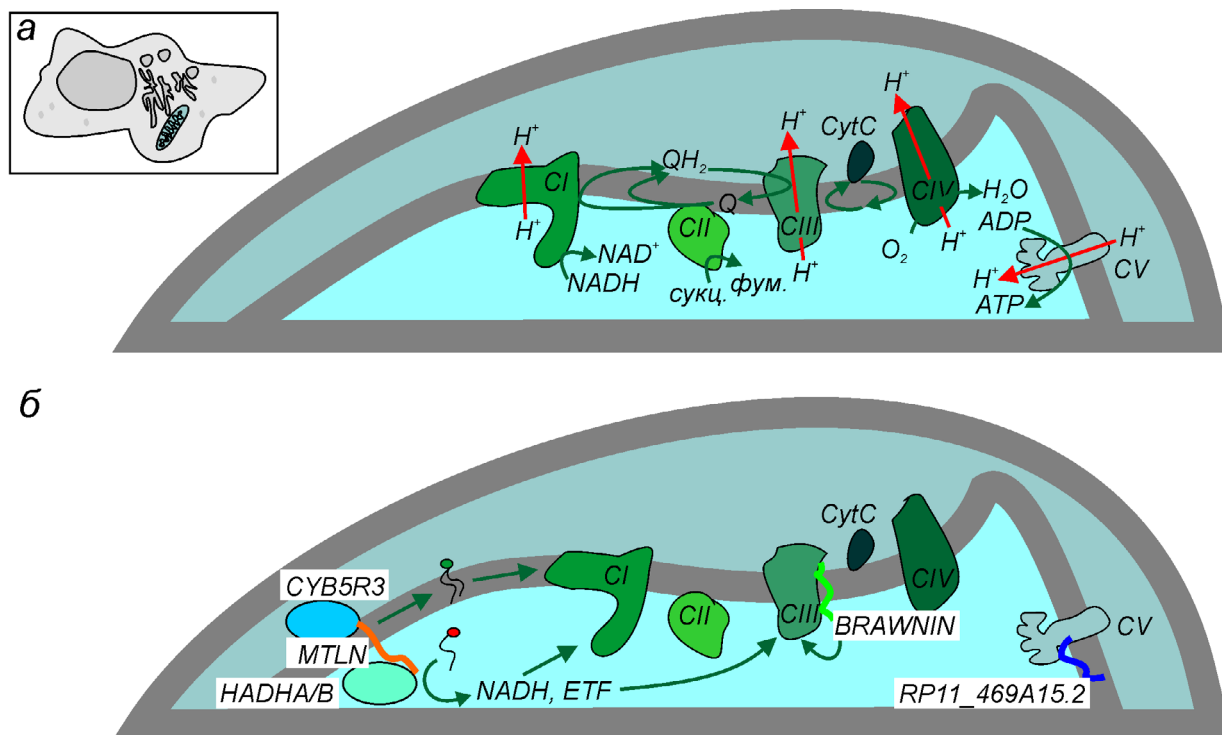
способствовать такого рода термогенезу [60]. Миорегулин (MLN), пептид длиной 46 а.о., взаимодействует с SERCA в мышцах всех типов. Так же как сарколипин и фосфоламбан, миорегулин содержит трансмембранную альфа-спираль в N-концевой части молекулы. Взаимодействие миорегулина с насосом SERCA отрицательно регулирует активность последнего. Инактивация гена миорегулина приводит к более эффективному закачиванию ионов кальция в саркоплазматический ретикулум, что стимулирует работу мышц [23] (рис. 3, д).

Помимо функционирования пептидов и малых белков в различных компартментах системы везикулярного транспорта, известны примеры пептидов/малых белков, экспортируемых через эту систему во внеклеточное пространство. Так, пептид Humanin длиной 24 а.о. закодирован в гене 16S митохондриальной рРНК. Учитывая, что митохондриальная рРНК содержит множество гипотетических рамок считывания, предшествующих рамке, кодирующей Humanin, маловероятно, чтобы считывание этой рамки считывания происходило в ходе трансляции в митохондриях. Возможно, экспрессия этого гена может осуществляться с копий митохондриальной ДНК, присутствующих в ядерном геноме в виде псевдогенов. Согласно оригинальной статье [26], продукт экспрессии этого гена секретируется в межклеточную среду и подавляет токсичность бета-амилоида, основного агента патогенеза болезни Альцгеймера, в отношении культуры нейронов. В последующие годы, несмотря на открытые вопросы о механизме экспрессии, у пептида обнаружили ряд других «чудодейственных» свойств в отношении остеопороза, сердечно-сосудистых заболеваний, диабета и рака, наряду с первоначально обнаруженной активностью в отношении нейродегенеративных заболеваний (для обзора см. [61]).

## МАЛЫЕ БЕЛКИ МИТОХОНДРИЙ

В протеоме митохондрий (рис. 4, а) значительную долю (до 5%) составляют малые белки длиной до 100 а.о. [62]. Выше мы уже обсуждали функциональную роль недавно открытого митохондриального белка MIEF1/LOR8F8, участвующего в сборке митохондриальных рибосом [49, 51]. Более широко в научной литературе представлены недавно открытые пептиды, влияющие на работу комплексов дыхательной цепи митохондрий. Например, выполненное в лаборатории Дж. Вейсмана исследование множества малых ОРС выявило митохондриальную локализацию продукта трансляции РНК





**Рис. 4.** Функционирование малых белков в мембранах митохондрий. *а* – Схема клетки с указанием митохондрий и схема работы окислительного фосфорилирования. Показаны комплексы дыхательной цепи I–IV (CI–CIV), комплекс V (CV) АТФ-синтазы, цепь переноса восстановительных эквивалентов и направление переноса протонов. *б* – Функционирование недавно открытых митохондриальных пептидов. Показано взаимодействие RP11\_469A15.2 с цитохром C оксидазой, стимуляция цитохром C редуктазы пептидом BRAWNIN и взаимодействие пептида MTLN с цитохром B5 редуктазой 3 (CYB5R3) и компонентами трифункционального комплекса окисления жирных кислот HADHA/B. Активность MTLN влияет на липидный состав мембран и эффективность дыхания

*RP11\_469A15.2*, ранее считавшейся некодирующей [13]. Соответствующий пептид длиной 62 а.о. взаимодействует с комплексом IV дыхательной цепи, цитохром C оксидазой. Молекулярный механизм функционирования данного пептида остается пока неизвестным (рис. 4, б).

Более подробно исследована функция еще одного недавно открытого митохондриального пептида, получившего название BRAWNIN. Этот пептид состоит из 71 а.о., консервативен у позвоночных и локализуется во внутренней мембране митохондрий [63]. Экспрессия гена этого пептида находится под контролем основного транскрипционного фактора, отвечающего за биогенез митохондрий, PGC1 $\alpha$ , и регулируется киназой AMPK, активирующейся в условиях нехватки АТФ. У рыб *Danio rerio* с инактивированным геном, кодирующим BRAWNIN, наблюдается лактазный ацидоз, характерный для многих митохондриальных заболеваний. Как показали опыты по коиммунопреципитации, BRAWNIN взаимодействует с комплексом III (СIII) дыхательной цепи, цитохром C редуктазой. Фенотипическое проявление инактивации гена этого пептида, согласно анализу метаболо-

ма, соответствует дефектам функционирования СIII (рис. 4б).

В нашей лаборатории [18], как и в нескольких других [19, 20], была исследована функция малого белка митохондрий миторегулина (MTLN), закодированного в транскрипте, ранее аннотированном как длинная некодирующая РНК *LINC00116*. Присутствие MTLN продемонстрировано в клетках человека [19, 20], мыши [18, 64] и рыбы *D. rerio* [65]. Последовательность этого белка длиной 56 а.о. весьма консервативна у позвоночных и содержит трансмембранный сегмент, закоривающий MTLN во внутренней митохондриальной мембране [20]. Данные о партнерах MTLN, полученные в разных лабораториях, несколько отличаются друг от друга. Так, в работах Makarewicz et al. [20] и Friesen et al. [66] продемонстрировано взаимодействие MTLN с компонентами комплекса  $\beta$ -окисления жирных кислот HADHA и HADHB, а в последней работе – также и взаимодействие с АТФ-синтазой. Наши эксперименты по коиммунопреципитации выявили взаимодействие MTLN с цитохром B5 редуктазой, CYB5R3. Практически все исследова-

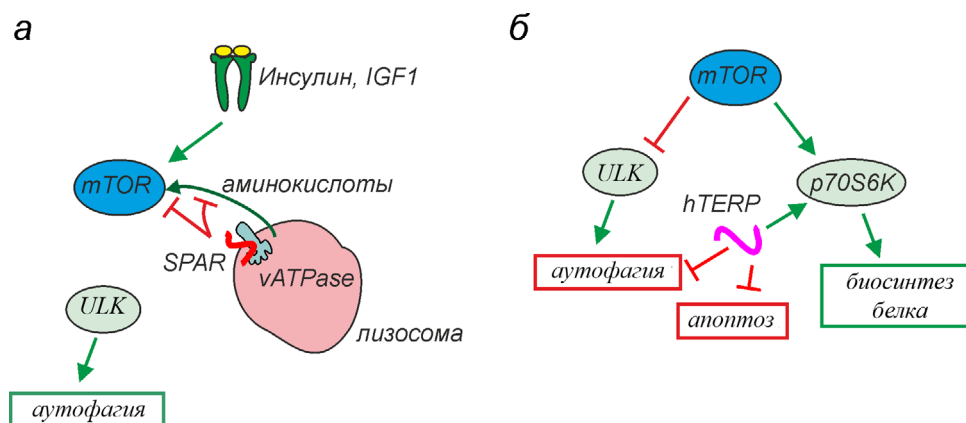
тели сходятся в том, что MTLN способствует работе комплексов респираторной цепи, однако, опять же, нюансы этого влияния видятся ими по-разному. В нашей работе [18] и работе Lin et al. [64] было продемонстрировано снижение эффективности работы комплекса I (CI) дыхательной цепи, NADH-дегидрогеназы. По результатам других исследований было выдвинуто предположение, что наблюдаемый при инактивации MTLN дефект функционирования дыхательной цепи проявляется только при окислении пальмитоилкарнитина, субстрата системы  $\beta$ -окисления жирных кислот [19, 20, 66]. Хотя мы и наблюдали снижение активности CI комплекса дыхательной цепи, это явление проявлялось только на уровне пермеабилizованных клеток и выделенных митохондрий, но не очищенного CI. Поэтому мы сделали вывод о том, что MTLN не является ни компонентом CI, ни фактором его сборки. Поскольку известно, что работа CI сильно зависит от липидного состава мембран, а партнер MTLN, CYB5R3, участвует в метаболизме липидов, мы изучили изменения липидного состава клеток с инактивированным геном, кодирующим MTLN. Оказалось, что отсутствие функционального митогенрегулятора действительно снижало количество необходимого для работы CI фосфатидилхолина и повышало содержание в клетке триглицеридов. Таким образом, функциональная роль MTLN связана, на наш взгляд, с формированием липидного состава клеточных мембран посредством взаимодействия с CYB5R3. В свою очередь, правильная работа данной системы обеспечивает работу комплекса CI дыхательной цепи за счет благоприятного липидного окружения (рис. 4, б).

#### МАЛЫЕ БЕЛКИ В РЕГУЛЯЦИИ МЕТАБОЛИЗМА

Несколько малых белков вовлечено в сигнальные пути, регулирующие метаболизм. Ключевая роль в этих сигнальных путях отводится киназным комплексам mTOR, стимуляция которых приводит к более активному биосинтезу белков, а ингибирование — к усилению аутофагии, регулируемого распаду компонентов клетки (см. [67] в качестве актуального обзора). Выше упомянуто, что белок altRPL36 является негативным регулятором mTOR-пути, поскольку ингибирует белок TMEM24, который транспортирует PI(4,5)P2. В свою очередь, PI(4,5)P2 является субстратом киназы PI3K, активирующей mTOR через несколько белков-посредников [7].

В РНК *LINC00961*, изначально аннотированной как длинная некодирующая, обнаружена ОРС, соответствующая пептиду длиной 75–90 а.о. в зависимости от используемого старт-кодона. Этот пептид, названный SPAR, имеет трансмембранный домен, с помощью которого он взаимодействует с мембраной позднего эндосом и лизосом [24]. Аффинное совыделение показало, что SPAR взаимодействует с v-АТФазой, протонным насосом, отвечающим за понижение рН в эндосомах и лизосомах. Оказалось, что SPAR ингибирует активацию mTORC1 киназного комплекса, привлекаемого к мембране лизосомы комплексом v-АТФазы, Ragulator и GTPаз семейства Rag при экспорте аминокислот из лизосомы в цитозоль. Этот эффект проявлялся при увеличении концентрации аминокислот после голодания. В клетках с избыточной экспрессией *LINC00961*, кодирующей SPAR, mTORC1 киназный комплекс не мог связаться с мембраной лизосом и оставался диффузно локализованным в цитозоле. Активация mTORC1 при помощи внешнего сигнального пути от рецепторных тирозинкиназ через PI3K не зависит от SPAR [24]. Поскольку известно, что активация mTOR-пути способствует восстановлению мышц после повреждения, у мышей с инактивированным геном *5430416009Rik*, кодирующим SPAR, была протестирована способность мышц к регенерации после повреждения кардиотоксином. Оказалось, что мышцы мышей без SPAR восстанавливаются быстрее, что вполне укладывается в модель гиперстимуляции mTOR-пути в отсутствие данного ингибиторного пептида [24]. Интересно отметить, что в последующих работах также обнаружены отдельные, не связанные с кодированием SPAR, функции РНК *LINC00961* [29] (рис. 5, а).

Регулятором метаболического ответа оказался малый белок hTERP (human TERP), продукт трансляции предшественника теломеразной РНК человека [25]. В процессе транскрипции гена теломеразной РНК, регулируемой мульти-субъединичным комплексом Integrator [68], синтезируется удлиненный предшественник, содержащий открытую рамку считывания, трансляция которой приводит к образованию белка длиной 121 а.о. Оказалось, что белок hTERP способствует выживанию клеток в условиях индукции апоптоза, а мутации в нем приводят к нарушениям протекания аутофагии (рис. 5, б). В отсутствие белка hTERP происходит активация киназы p70S6K1 (неопубликованные данные), что согласуется с данными, полученными при изучении мышей, в организме которых произведена делеция гена теломеразной РНК [69].



**Рис. 5.** Функционирование малых белков в регуляции аутофагии и биосинтеза белка. *а* – Пептид SPAR, взаимодействующий с протонной помпой лизосомы, подавляет активацию киназы mTOR аминокислотами, экспортируемыми из лизосомы. При этом SPAR не влияет на активацию mTOR через сигнальный путь от рецепторных тирозинкиназ. *б* – Белок hTERP, закодированный в теломеразной РНК, участвует в регуляции переключения метаболизма клетки, воздействуя на аутофагию и биосинтез белка, а также подавляет апоптоз (подробнее – в тексте)

Участие белка hTERP в модуляции киназных каскадов, регулирующих метаболизм клеток, приоткрывает неизведанный уровень стимуляции пролиферации, ассоциированный как правило с активацией теломеразного комплекса, одним из основных компонентов которого является теломеразная РНК. Известно, что стимуляция клеток иммунной системы, онкотрансформация клеток и другие процессы, требующие повышенной скорости пролиферации и, следовательно, ресурсов, сопровождаются переключением метаболической программы, а также повышением активности теломеразы. Для активации теломеразного комплекса требуется повышенная экспрессия гена обратной транскриптазы, а также эффективный процессинг теломеразной РНК с образованием формы, ассоциированной с теломеразным комплексом. Можно предположить, что выбор одного из двух путей экспрессии гена теломеразной РНК должен зависеть от метаболического состояния клеток и скорости их пролиферации, а также влиять на эти процессы. Таким образом, белок hTERP может функционировать либо как регулятор метаболической программы клетки, либо как сенсор переключения окислительного фосфорилирования и гликолиза. Роль белка TERP в регуляции метаболизма клеток требует дальнейшего изучения.

## ЗАКЛЮЧЕНИЕ

Малый размер, низкая консервативность последовательностей и незначительные количе-

ства продуктов трансляции коротких рамок считывания затрудняют их идентификацию и изучение функциональной роли. Развитие методов исследования позволило открыть новый уровень регуляции внутриклеточных процессов при помощи пептидов и малых белков. Исследования последних лет демонстрируют участие продуктов трансляции малых рамок считывания во всех процессах функционирования клеток, а истинное количество кодируемых ими белковых продуктов трудно оценить. Постоянно появляющиеся публикации, демонстрирующие разнообразные функции продуктов трансляции коротких рамок считывания, приоткрывают до сих пор нераскрытый мощный регуляторный потенциал «серой материи» протеома. Исследования, направленные на идентификацию и выяснение роли компонентов «альтернативного» протеома, абсолютно необходимы для получения полной картины регуляторных событий, лежащих в основе как поддержания гомеостаза клетки, так и развития патологических процессов.

**Финансирование.** Работа выполнена при финансовой поддержке Российского научного фонда (грант № 21-64-00006).

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

**Соблюдение этических норм.** Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

## СПИСОК ЛИТЕРАТУРЫ

1. Numata, K., Kanai, A., Saito, R., Kondo, S., Adachi, J., et al. (2003) Identification of putative noncoding RNAs among the RIKEN mouse full-length cDNA collection, *Genome Res.*, **13**, 1301-1306.
2. Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., et al. (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs, *Nature*, **420**, 563-573.
3. Wilhelm, B. T., Marguerat, S., Watt, S., Schubert, F., Wood, V., et al. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution, *Nature*, **453**, 1239-1243.
4. Ulitsky, I., and Bartel, D. P. (2013) lincRNAs: Genomics, evolution, and mechanisms, *Cell*, **154**, 26-46.
5. Kopp, F., and Mendell, J. T. (2018) Functional classification and experimental dissection of long noncoding RNAs, *Cell*, **172**, 393-407.
6. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., and Weissman, J. S. (2009) Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling, *Science*, **324**, 218-223.
7. Cao, X., Khitun, A., Na, Z., Dumitrescu, D. G., Kubica, M., et al. (2020) Comparative proteomic profiling of unannotated microproteins and alternative proteins in human cell lines, *J. Proteome Res.*, **19**, 3418-3426.
8. Ma, J., Ward, C. C., Jungreis, I., Slavoff, S. A., Schwaib, A. G., et al. (2014) Discovery of human sORF-encoded polypeptides (SEPs) in cell lines and tissue, *J. Proteome Res.*, **13**, 1757-1765.
9. Slavoff, S. A., Mitchell, A. J., Schwaib, A. G., Cabili, M. N., Ma, J., et al. (2013) Peptidomic discovery of short open reading frame-encoded peptides in human cells, *Nat. Chem. Biol.*, **9**, 59-64.
10. Cardon, T., Hervé, F., Delcourt, V., Roucou, X., Salzet, M., et al. (2020) Optimized sample preparation workflow for improved identification of ghost proteins, *Anal. Chem.*, **92**, 1122-1129.
11. Cardon, T., Fournier, I., and Salzet, M. (2021) Shedding light on the ghost proteome, *Trends Biochem. Sci.*, **46**, 239-250.
12. Chugunova, A., Navalayeu, T., Dontsova, O., and Sergiev, P. (2018) Mining for small translated ORFs, *J. Proteome Res.*, **17**, 1-11.
13. Chen, J., Brunner, A.-D., Cogan, J. Z., Nuñez, J. K., et al. (2020) Pervasive functional translation of noncanonical human open reading frames, *Science*, **367**, 1140-1146.
14. Cao, X., and Slavoff, S. A. (2020) Non-AUG start codons: expanding and regulating the small and alternative ORFeome, *Exper. Cell Res.*, **391**, 111973.
15. Hinnebusch, A. G. (1997) Translational regulation of yeast GCN4, *J. Biol. Chem.*, **272**, 21661-21664.
16. Andreev, D. E., O'Connor, P. B., Fahey, C., Kenny, E. M., Terenin, I. M., et al. (2015) Translation of 5' leaders is pervasive in genes resistant to eIF2 repression, *eLife*, **4**, e03971.
17. Michel, A. M., Andreev, D. E., and Baranov, P. V. (2014) Computational approach for calculating the probability of eukaryotic translation initiation from ribo-seq data that takes into account leaky scanning, *BMC Bioinformatics*, **15**, 380.
18. Chugunova, A., Loseva, E., Mazin, P., Mitina, A., Navalayeu, T., et al. (2019) LINC00116 codes for a mitochondrial peptide linking respiration and lipid metabolism, *Proc. Natl. Acad. Sci. USA*, **116**, 4940-4945.
19. Stein, C. S., Jadiya, P., Zhang, X., McLendon, J. M., Abouassaly, G. M., et al. (2018) Mitoregulin: a lincRNA-encoded microprotein that supports Mitochondrial super-complexes and respiratory efficiency, *Cell Rep.*, **23**, 3710-3720.e8.
20. Makarewich, C. A., Baskin, K. K., Munir, A. Z., Bezprozvannaya, S., Sharma, G., et al. (2018) MOXI is a mitochondrial micropeptide that enhances fatty acid  $\beta$ -oxidation, *Cell Rep.*, **23**, 3701-3709.
21. Chu, Q., Martinez, T. F., Novak, S. W., Donaldson, C. J., Tan, D., et al. (2019) Regulation of the ER stress response by a mitochondrial microprotein, *Nat. Commun.*, **10**, 4883.
22. Jackson, R., Kroehling, L., Khitun, A., Bailis, W., Jarret, A., et al. (2018) The translation of non-canonical open reading frames controls mucosal immunity, *Nature*, **564**, 434-438.
23. Anderson, D. M., Anderson, K. M., Chang, C.-L., Makarewich, C. A., Nelson, B. R., et al. (2015) A micropeptide encoded by a putative long noncoding rna regulates muscle performance, *Cell*, **160**, 595-606.
24. Matsumoto, A., Pasut, A., Matsumoto, M., Yamashita, R., Fung, J., et al. (2017) mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide, *Nature*, **541**, 228-232.
25. Rubtsova, M., Naraykina, Y., Vasilkova, D., Meerson, M., Zvereva, M., et al. (2018) Protein encoded in human telomerase RNA is involved in cell protective pathways, *Nucleic Acids Res.*, **46**, 8966-8977.
26. Hashimoto, Y., Niikura, T., Tajima, H., Yasukawa, T., Sudo, H., et al. (2001) A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and A, *Proc. Natl. Acad. Sci. USA*, **98**, 6336-6341.
27. Tenson, T., DeBlasio, A., and Mankin, A. (1996) A functional peptide encoded in the Escherichia coli 23S rRNA, *Proc. Natl. Acad. Sci. USA*, **93**, 5641-5646.
28. Yang, X., Bam, M., Becker, W., Nagarkatti, P. S., and Nagarkatti, M. (2020) Long Noncoding RNA AW112010 Promotes the differentiation of inflammatory T Cells by suppressing IL-10 expression through histone demethylation, *J. Immunol.*, **205**, 987-993.
29. Spencer, H. L., Sanders, R., Boulberdaa, M., Meloni, M., Cochrane, A., et al. (2020) The LINC00961 transcript and its encoded micropeptide, small regulatory polypeptide of amino acid response, regulate endothelial cell function, *Cardiovasc. Res.*, **116**, 1981-1994.
30. Zhang, M., Huang, N., Yang, X., Luo, J., Yan, S., et al. (2018) A novel protein encoded by the circular form of the SHPRH gene suppresses glioma tumorigenesis, *Oncogene*, **37**, 1805-1814.
31. Zhang, M., Zhao, K., Xu, X., Yang, Y., Yan, S., et al. (2018) A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma, *Nat. Commun.*, **9**, 4475.
32. Yin, H., Shen, X., Zhao, J., Cao, X., He, H., et al. (2020) Circular RNA CircFAM188B encodes a protein that regulates proliferation and differentiation of chicken skeletal muscle satellite cells, *Front. Cell Dev. Biol.*, **8**, 522588.
33. Wang, J., Zhu, S., Meng, N., He, Y., Lu, R., and Yan, G.-R. (2019) ncRNA-encoded peptides or proteins and cancer, *Mol. Ther.*, **27**, 1718-1725.
34. Fang, J., Morsalin, S., Rao, V. N., and Reddy, E. S. P. (2017) Decoding of non-coding DNA and non-coding RNA: Pri-micro RNA-encoded novel peptides regulate migration of cancer cells, *J. Pharmaceut. Sci. Pharmacol.*, **3**, 23-27.
35. Shi, Y., Jia, X., and Xu, J. (2020) The new function of circRNA: translation, *Clin. Transl. Oncol.*, **22**, 2162-2169.
36. Vanderperre, B., Lucier, J.-F., Bissonnette, C., Motard, J., Tremblay, G., et al. (2013) Direct detection of alternative

- open reading frames translation products in human significantly expands the proteome, *PLoS One*, **8**, e70698.
37. Liu, Y., Woods, N. T., Kim, D., Sweet, M., Monteiro, A. N. A., and Karchin, R. (2011) Yeast two-hybrid junk sequences contain selected linear motifs, *Nucleic Acids Res.*, **39**, e128-e128.
  38. Arnoult, N., Correia, A., Ma, J., Merlo, A., Garcia-Gomez, S., et al. (2017) Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN, *Nature*, **549**, 548-552.
  39. Agarwal, S., Harada, J., Schreifels, J., Lech, P., Nikolai, B., et al. (2006) Isolation, characterization, and genetic complementation of a cellular mutant resistant to retroviral infection, *Proc. Natl. Acad. Sci. USA*, **103**, 15933-15938.
  40. Hung, P. J., Johnson, B., Chen, B.-R., Byrum, A. K., Bredemeyer, A. L., et al. (2018) MRI is a DNA damage response adaptor during classical non-homologous end joining, *Mol. Cell*, **71**, 332-342.e8.
  41. Hanyu-Nakamura, K., Sonobe-Nojima, H., Tanigawa, A., Lasko, P., and Nakamura, A. (2008) Drosophila Pgc protein inhibits P-TEFb recruitment to chromatin in primordial germ cells, *Nature*, **451**, 730-733.
  42. Hou, L., Wang, Y., Liu, Y., Zhang, N., Shamovsky, I., et al. (2019) Paf1C regulates RNA polymerase II progression by modulating elongation rate, *Proc. Natl. Acad. Sci. USA*, **116**, 14583-14592.
  43. Yang, Y., Gao, X., Zhang, M., Yan, S., Sun, C., et al. (2018) Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis, *JNCI: J. Natl. Cancer Institute*, **110**, 304-315.
  44. Cardon, T., Franck, J., Coyaud, E., Laurent, E. M. N., Damato, M., et al. (2020) Alternative proteins are functional regulators in cell reprogramming by PKA activation, *Nucleic Acids Res.*, **48**, 7864-7882.
  45. Huang, J.-Z., Chen, M., Chen, D., Gao, X.-C., Zhu, S., et al. (2017) A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth, *Mol. Cell*, **68**, 171-184.e6.
  46. Christofk, H. R., Vander Heiden, M. G., Harris, M. H., Ramanathan, A., Gerszten, R. E., et al. (2008) The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth, *Nature*, **452**, 230-233.
  47. Chen, M., Zhang, J., and Manley, J. L. (2010) Turning on a fuel switch of cancer: hnRNP proteins regulate alternative splicing of pyruvate kinase mRNA: Figure 1. *Cancer Res.*, **70**, 8977-8980.
  48. Palmer, C. S., Osellame, L. D., Laine, D., Koutsopoulos, O. S., Frazier, A. E., and Ryan, M. T. (2011) MiD49 and MiD51, new components of the mitochondrial fission machinery, *EMBO Rep.*, **12**, 565-573.
  49. Rathore, A., Chu, Q., Tan, D., Martinez, T. F., Donaldson, C. J., et al. (2018) MIEF1 microprotein regulates mitochondrial translation, *Biochemistry*, **57**, 5564-5575.
  50. Delcourt, V., Brunelle, M., Roy, A. V., Jacques, J.-F., Salzet, M., et al. (2018) The protein coded by a short open reading frame, not by the annotated coding sequence, is the main gene product of the dual-coding gene MIEF1, *Mol. Cell. Proteomics*, **17**, 2402-2411.
  51. Brown, A., Rathore, S., Kimanius, D., Aibara, S., Bai, X., et al. (2017) Structures of the human mitochondrial ribosome in native states of assembly, *Nat. Struct. Mol. Biol.*, **24**, 866-869.
  52. Cardon, T., Salzet, M., Franck, J., and Fournier, I. (2019) Nuclei of HeLa cells interactomes unravel a network of ghost proteins involved in proteins translation. *Biochim. Biophys. Acta Gen. Subjects*, **1863**, 1458-1470.
  53. D'Lima, N. G., Ma, J., Winkler, L., Chu, Q., Loh, K. H., et al. (2017) A human microprotein that interacts with the mRNA decapping complex, *Nat. Chem. Biol.*, **13**, 174-180.
  54. Cloutier, P., Poitras, C., Faubert, D., Bouchard, A., Blanchette, M., et al. (2020) Upstream ORF-encoded ASDURF is a novel prefoldin-like subunit of the PAQosome, *J. Proteome Res.*, **19**, 18-27.
  55. Polycarpou-Schwarz, M., Groß, M., Mestdagh, P., Schott, J., Grund, S. E., et al. (2018) The cancer-associated microprotein CASIMO1 controls cell proliferation and interacts with squalene epoxidase modulating lipid droplet formation, *Oncogene*, **37**, 4750-4768.
  56. Kranias, E. G., and Hajjar, R. J. (2012) Modulation of cardiac contractility by the phospholamban/SERCA2a regulome, *Circ. Res.*, **110**, 1646-1660.
  57. MacLennan, D. H., and Kranias, E. G. (2003) Phospholamban: a crucial regulator of cardiac contractility, *Nat. Rev. Mol. Cell Biol.*, **4**, 566-577.
  58. Tupling, A. R., Bombardier, E., Gupta, S. C., Hussain, D., Vigna, C., et al. (2011) Enhanced Ca<sup>2+</sup> transport and muscle relaxation in skeletal muscle from sarcolipin-null mice, *Am. J. Physiol. Cell Physiol.*, **301**, C841-C849.
  59. Gorski, P. A., Ceholski, D. K., and Young, H. S. (2017) Structure-function relationship of the SERCA Pump and its regulation by phospholamban and sarcolipin, in *Membrane Dynamics and Calcium Signaling*, vol. 981, Springer International Publishing, Cham, pp. 77-119.
  60. Bal, N. C., and Periasamy, M. (2020) Uncoupling of sarcoendoplasmic reticulum calcium ATPase pump activity by sarcolipin as the basis for muscle non-shivering thermogenesis, *Phil. Trans. R. Soc. B*, **375**, 20190135.
  61. Hazafa, A., Batool, A., Ahmad, S., Amjad, M., Chaudhry, S. N., et al. (2021) Humanin: a mitochondrial-derived peptide in the treatment of apoptosis-related diseases, *Life Sci.*, **264**, 118679.
  62. Calvo, S. E., Clauser, K. R., and Mootha, V. K. (2016) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins, *Nucleic Acids Res.*, **44**, D1251-D1257.
  63. Zhang, S., Reljić, B., Liang, C., Kerouanton, B., Francisco, J. C., et al. (2020) Mitochondrial peptide BRAWNIN is essential for vertebrate respiratory complex III assembly, *Nat. Commun.*, **11**, 1312.
  64. Lin, Y.-F., Xiao, M.-H., Chen, H.-X., Meng, Y., Zhao, N., et al. (2019) A novel mitochondrial micropeptide MPM enhances mitochondrial respiratory activity and promotes myogenic differentiation, *Cell Death Dis.*, **10**, 528, doi: 10.1038/s41419-019-1767-y.
  65. Wang, L., Fan, J., Han, L., Qi, H., Wang, Y., et al. (2020) The micropeptide LEMP plays an evolutionarily conserved role in myogenesis, *Cell Death Dis.*, **11**, 357, doi: 10.1038/s41419-020-2570-5.
  66. Friesen, M., Warren, C. R., Yu, H., Toyohara, T., Ding, Q., et al. (2020) Mitoregulin controls  $\beta$ -oxidation in human and mouse adipocytes, *Stem Cell Rep.*, **14**, 590-602.
  67. Liu, G. Y., and Sabatini, D. M. (2020) mTOR at the nexus of nutrition, growth, ageing and disease, *Nat. Rev. Mol. Cell Biol.*, **21**, 183-203.
  68. Rubtsova, M. P., Vasilkova, D. P., Moshareva, M. A., Malyavko, A. N., Meerson, M. B., et al. (2019) Integrator is a key component of human telomerase RNA biogenesis, *Sci. Rep.*, **9**, 1701.
  69. Ferrara-Romeo, I., Martinez, P., Saraswati, S., Whittemore, K., Graña-Castro, O., et al. (2020) The mTOR pathway is necessary for survival of mice with short telomeres, *Nat. Commun.*, **11**, 1168.

**LITTLE BUT LOUD. THE DIVERSITY OF FUNCTIONS  
OF SMALL PROTEINS AND PEPTIDES –  
TRANSLATIONAL PRODUCTS OF SHORT READING FRAMES**

**Review**

**P. V. Sergiev<sup>1,2,3\*</sup> and M. P. Rubtsova<sup>1\*</sup>**

<sup>1</sup> *Lomonosov Moscow State University, Faculty of Chemistry, 119991 Moscow, Russia;  
e-mail: petya@genebee.msu.ru, mprubtsova@gmail.com*

<sup>2</sup> *Center of Life Sciences, Skolkovo Institute of Sciences and Technology, 143025 Skolkovo, Russia*

<sup>3</sup> *Institute of Functional Genomics, Lomonosov Moscow State University, Moscow, 119991 Moscow, Russia*

Cell functioning is a tightly regulated process. For many years, research in the fields of proteomics and functional genomics has been focused on the role of proteins in cell functioning. The advances in science have led to the uncovering that short open reading frames, previously considered non-functional, serve a variety of functions. Short reading frames in polycistronic mRNAs often regulate their stability and translation efficiency of the main reading frame. The improvement of proteomic analysis methods has made it possible to identify the products of translation of short reading frames in quantities that suggest the existence of functional role of those peptides and short proteins. Studies demonstrating their role unravel a new level of the regulation of cell functioning and its adaptation to changing conditions. This review is devoted to the analysis of functions of recently discovered peptides and short proteins.

*Keywords:* peptides, short open reading frame, long noncoding RNAs, alternative proteins