

РОЛЬ НУКЛЕОКАПСИДА В МУЛЬТИМЕРИЗАЦИИ СТРУКТУРНОГО БЕЛКА GAG РЕТРОВИРУСА *gypsy*

© 2012 г. Б. В. Сёмин^{1,2*}, О. Г. Леонова², Т. А. Тренделева³, Р. А. Звягильская³,
Ю. В. Ильин², В. И. Попенко²

¹Всероссийский научно-исследовательский институт экспериментальной ветеринарии им. Я.П. Коваленко
Российской академии сельскохозяйственных наук, Москва, 109391

²Институт молекулярной биологии им. В.А.Энгельгардта Российской академии наук, Москва, 119991

³Институт биохимии им. А.Н.Баха Российской академии наук, Москва, 119071

Поступила в редакцию 09.06.2011 г.

Принята к печати 08.09.2011 г.

Структурный белок Gag ретровируса дрозофилы *gypsy* состоит из капсидного и нуклеокапсидного доменов. Белок Gag образует вирусоподобные частицы в бактериальной клетке; кроме того, капсидный домен Gag тоже способен мультимеризоваться, причем без участия нуклеокапсидов. Однако агрегаты, образуемые капсидом, гетерогенны по размеру и гораздо менее организованы, чем частицы, сформированные полноразмерным Gag *gypsy*. Нуклеокапсид существенно влияет на организацию частиц, и эту функцию определяет последовательность аминокислотных остатков на его N-конце (проксимальная часть нуклеокапсидов). Формирование частиц эффективно происходит в присутствии РНК или однонитевых олигонуклеотидов произвольной последовательности.

Ключевые слова: *gypsy*, ретровирус, ретротранспозон, вирусоподобная частица, структурный белок, Gag, бактериальная система экспрессии.

EFFECT OF NUCLEOCAPSID ON MULTIMERIZATION OF *gypsy* STRUCTURAL PROTEIN GAG, by B. V. Syomin^{1, 2*}, O. G. Leonova², T. A. Trendeleva³, R. A. Zvyagilskaya³, Yu. V. Ilyin², V. I. Popenko² (¹Kovalenko All-Russian Institute of Experimental Veterinary Medicine, Moscow, 109391 Russia, *e-mail: bsyomin@yandex.ru; ²Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia; ³Bach Institute of Biochemistry, Russian Academy of Sciences, 119071 Moscow, Russia). The structural protein (Gag) of the *gypsy Drosophila* retrovirus lacks matrix, but contains capsid and nucleocapsid domains. The Gag forms virus-like particles in a bacterial cell; besides, its capsid alone is able to form aggregates. However, aggregates assembled from the capsid were variable in size and displayed much less organization than particles formed by the whole Gag. The nucleocapsid exerts influence on the organization and structure of particles, and this function is directed by sequence of amino acid residues at its N-terminus (a nucleocapsid proximal part). The particle assembling occurs in the presence of any RNAs or single stranded DNA oligonucleotides.

Keywords: *gypsy*, retrovirus, retrotransposon, virus-like particle, structural protein, Gag, bacterial system of expression.

Эндогенные ретровирусы и ретротранспозоны, будучи интегральной частью эукариотического генома, как правило, занимают около 10%, а в некоторых случаях до 50% генома [1]. Их способность к транспозициям и горизонтальному переносу [2–4] представляет один из основных факторов генетической нестабильности и пластичности генома, что в рамках отдельного организма может привести к развитию ряда заболеваний, в том числе и у человека [5].

Ретровирус дрозофилы *gypsy* – удобная модель для изучения биологии одной из групп эндоген-

ных ретровирусов – *Metaviridae*, или *Ty3/gypsy* группы, сейчас называемой также эрантирвирусами [6, 7]. Интерес к изучению *gypsy* обусловлен его довольно высокой частотой транспозиций, амплифицированностью в геноме и способностью к межвидовому переносу [3]. Природу *gypsy* можно рассматривать двояко. С одной стороны, это ретротранспозон, активно экспрессирующийся в геноме ряда видов плодовой мушки [6, 8], и в этой связи представляет собой один из факторов, обеспечивающих вариативность генома за счет своих транспозиций. С другой стороны, это инфекционный ретровирус, о чем свидетельствует его

* Эл. почта: bsyomin@yandex.ru

структурная организация: *gypsy* содержит между длинными концевыми повторами гены *gag*, *pol* и *env*, свойственные ретровирусам, и, главное, частицы, образуемые этим элементом, обеспечивают его горизонтальный транспорт [3, 9, 10].

Как и все ретровирусы и ретротранспозоны, *gypsy* кодирует структурный белок Gag. Особенность Gag *gypsy* заключается в том, что его аминокислотная последовательность не содержит известных канонических мотивов, характерных для Gag всех подсемейств ретровирусов: лентивирусов, онковирусов и спумавирусов [6, 11]. Более того, нет свидетельств, позволяющих предполагать, что Gag *gypsy* процессируется на матриксный, капсидный и нуклеокапсидный пептиды [11, 12], как это присуще структурным белкам большинства ретровирусов позвоночных [13]. Выделить в Gag *gypsy* соответствующие домены возможно лишь виртуально. Дополняет картину то, что в нуклеокапсидной части этого белка отсутствуют классические последовательности “цинковые пальцы” (или так называемые ССНС-мотивы), обычно обеспечивающие связывание структурного белка с РНК ретровируса [14, 15]. Вместе с тем, такая организация Gag не ограничивается исключительно ретровирусом *gypsy*, а характерна для ряда ретротранспозонов дрозофилы и дрожжей [6]. Кроме того, структурные белки спумавирусов млекопитающих [16] и ряда эндогенных ретровирусов, обнаруженных также и у человека (например, HERVL [1]), лишены ССНС-мотивов подобно Gag *gypsy*.

Несмотря на необычную организацию, Gag *gypsy*, как и все Gag, отвечает за образование вирусной частицы. Более того, он самодостаточен для мультимеризации в глобулярные частицы без участия каких-либо факторов эукариотической клетки, поскольку, экспрессированный в бактериях, способен образовывать частицы [11]. Вместе с тем, практически ничего не известно о том, какие домены аминокислотной последовательности ответственны за способность этого типа Gag мультимеризоваться в вирусоподобную частицу. Прояснить этот вопрос важно не только с точки зрения понимания механизма функционирования отдельно взятого белка, но и того, каким образом несхожие по аминокислотной последовательности белки могут выполнять тождественные функции.

Традиционный подход, используемый для выявления роли отдельных участков структурного белка ретровируса в формировании вирусоподобной частицы, — экспрессия делеционных мутантов Gag в бактериальной системе [17–19]. Gag *gypsy* состоит только из двух доменов: капсида и нуклеокапсида [11]; причем нуклеокапсидный домен был определен по его классическому свойству — отжигать тРНК к вирусной РНК [20]. Вме-

сте с тем, из концепции, согласно которой мультимеризация Gag начинается со взаимодействия нуклеокапсида с РНК, следует, что нуклеокапсид играет ключевую роль в формировании частицы ретровируса [14, 15, 21, 22]. В этой работе мы экспериментально оценили вклад нуклеокапсида в формирование частицы *gypsy*. Для этого сравнили способность к мультимеризации полноразмерной формы белка с двумя его мутантными формами: в одном случае из белка Gag нуклеокапсид удалили полностью, а в другом — только С-концевую часть нуклеокапсида, содержащую мотив, определяющий большой положительный заряд белка и распространенный в белках, связывающихся с РНК.

Нами показано, что для эффективной мультимеризации белка необходима нуклеиновая кислота. Однако нуклеокапсид нельзя считать единственным необходимым фактором мультимеризации Gag *gypsy*. Капсидная часть белка также способна мультимеризоваться в присутствии нуклеиновой кислоты. Нуклеокапсид существенно влияет на организацию и структуру частиц, и эту функцию определяет N-концевая часть нуклеокапсида *gypsy*. Мультимеризация всех рассмотренных в работе белков (полноразмерный Gag, капсид Gag, нуклеокапсид Gag) эффективно происходит в присутствии РНК или однонитевых олигонуклеотидов, при этом для образования частиц не требуется никакой специфичной последовательности нуклеиновых кислот.

ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Бактериальные линии. Линия клеток *Escherichia coli* X1-Blue (“Stratogene”, Великобритания) использована для клонирования, а линия *E. coli* BL21 (DE3) (“Novagen”, США) — для экспрессии белка.

Создание конструкций. Молекулярное клонирование выполнено по стандартным методикам [23]. Плазмиды, предназначенные для экспрессии в бактериях белка Gag *gypsy* и его делеционных мутантов, получены на основе плазмиды BG1, состоящей из вектора pBlueBac4.5/V5-His TOPO с клонированным ДНК-фрагментом геномной копии Gag *gypsy*, синтезированной полимеразой PfuI [24]. При получении конструкции “1” для экспрессии в бактериях полноразмерной формы Gag NcoI-HindIII фрагмент плазмиды BG1 клонировали в сайты NcoI-HindIII вектора pET23d(+). Конструкцию “1ΔNC” получили в результате клонирования фрагмента ДНК HindII-NcoI из конструкции BG1 в вектор pET23d(+) по сайтам MluI (3'-конец затуплен фрагментом Кленова) и NcoI.

Конструкцию “1ΔRRM” получили в результате клонирования в вектор pET23d(+) по сайтам

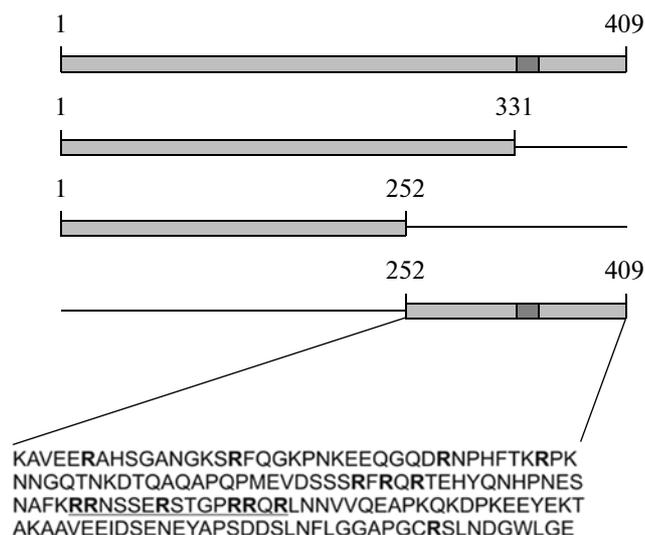


Рис. 1. Схематическое изображение использованных в работе белков, экспрессированных в бактериях. Сверху вниз изображены “1”, “1ΔRRM”, “1ΔNC” и “NC” соответственно. Последовательность аминокислотных остатков нуклеокапсида приведена на схеме. Положительно заряженная последовательность, содержащая парные остатки аргинина и серина, подчеркнута и названа SR-мотив, на схемах белков она представлена темным прямоугольником.

DraII (затуплен фрагментом Кленова) и NcoI фрагмента ДНК HindII-NcoI из конструкции BG1. Плазмиду, позволяющую экспрессировать нуклеокапсид Gag *gypsy* (“NC”), получили, выделив StyI-HindIII 3'-концевой фрагмент Gag из плазмиды BG1 и достроив концы фрагмента до тупых фрагментом Кленова. Полученный фрагмент лигировали в сайт NheI плазмиды pET23d(+), достроенный до тупого. Созданные конструкции были способны экспрессировать в бактериальной клетке мутантные (делетированные) формы Gag *gypsy*, схематическое изображение которых представлено на рис. 1.

Экспрессия белка. Каждую из полученных конструкций трансформировали в клетки *E. coli* линии BL21(DE3). 5 мл ночной культуры трансформированных клеток разбавляли (1 : 100) свежей культуральной средой (LB/amp) и инкубировали при 37°C до достижения оптической плотности 0.8 (при длине волны 600 нм). Ген T7 РНК-полимеразы под контролем *lac*-промотора постоянно слабо экспрессируется даже в отсутствии индуктора (“Novagen”). Чтобы увеличить выход водорастворимой порции Gag (см. ниже) клетки с трансформированными в них плазмидами инкубировали при 4°C в течение 72 ч без изопропилтиогалактозида (ИПТГ) в присутствии ампицилина (50 мкг/мл). Электрофорез белков проводили по методике Лемли [23].

Выделение вирусоподобных частиц из бактериальных клеток. Клетки осаждали центрифугированием (12000 × g), промывали фосфатно-солевым буфером и ресуспендировали в буфере А (20 мМ Трис-НСl, рН 7.8, 20 мМ NaCl, 2 мМ PMSF, 0.25% Triton X-100), гомогенизировали в гомогенизаторе Даунса и разрушали ультразвуком при помощи дезинтегратора “Bandtlin sonopuls” (Германия). Полученный лизат осветляли центрифугированием при 36000 × g в течение 15 мин при 4°C, супернатант наслаивали на 3 мл 30%-ного раствора сахарозы в буфере А и центрифугировали 2 ч при 180000 × g в роторе SW41 на центрифуге Optima TL-100 (“Beckman-Coulter”, США). Осадок ресуспендировали в 500 мкл буфера А, гомогенизировали в гомогенизаторе Даунса и вновь фракционировали, используя метод ультрацентрифугирования в ступенчатом градиенте 20/30/70%-ных концентраций сахарозы, взятых в соотношении 500 : 250 : 190 мкл (ротор TLS55, 81000 × g, 68 мин, 4°C). Из пробирки последовательно отбирали 196 мкл алиquotы и, таким образом, получали 8 фракций, которые затем анализировали методом электрофореза в ПААГ [23].

Очистка мономеров белка. Лизаты бактериальных клеток, после экспрессии полноразмерного Gag или его делетированных форм, центрифугировали при 12000 × g в течение 15 мин. Образовавшийся осадок трижды промывали раствором 0.5%-ного Triton X-100 с 10 мМ ЭДТА и оставляли на 1 ч при комнатной температуре в 8 М мочевины, содержащей 0.2 М NaCl, 0.1 М Трис-НСl, рН 6.8. Получившийся раствор центрифугировали при 280000 × g в течение 20 мин. Осадок после центрифугирования не содержал Gag, т.е. агрегаты изучаемых белков полностью растворялись в присутствии 8 М мочевины. Поскольку клонирование выполнено так, что к С-концу всех экспрессируемых белков подшита последовательность из 6 остатков гистидина (полигистидиновый тракт), полноразмерную форму Gag или его делетированные формы выделяли из супернатанта методом аффинной хроматографии на Ni-NTA-агарозе по методике производителя (“Pharmacia”). Полученный таким образом препарат затем дополнительно очищали гель-фильтрацией на колонке Superose 12 (“Pharmacia”) в системе FPLC.

Связывание меченой ДНК с белками, иммобилизованными на фильтре, проводили в буфере следующего состава: 20 мМ Трис-НСl, рН 7.5, 50 мМ NaCl и 5%-ного обезжиренного молока – по методике, описанной нами ранее [25]. ДНК метили Р³² при помощи “рассеянной затравки” (“Amersham”) по методике производителя. Для получения меченой РНК *in vitro* плазмиду pGEMgyp [26] линеаризовали рестриктазой BamHI и проводили реакцию синтеза РНК SP6 полимеразой (“Amersham”) по методике производителя. В ка-

честве предшественника использовали [P^{32}]-меченый АТФ в количестве 100 мкКи на одну реакцию синтеза.

Получение частиц *in vitro*. Концентрацию мочевины в растворах очищенного белка постепенно снижали (от 6 до 0.25 М) при диализе в буфере следующего состава: 50 мМ Трис-НСl, рН 6.8; 150 мМ NaCl, 5 мМ MgCl₂, 5%-ный глицерин.

Электронная микроскопия. Образцы (5 мкл) наносили на свежеионизированные коллодиево-угольные пленки-подложки, окрашивали 2%-ным уранилацетатом и анализировали на электронном микроскопе JEM-100CX (80 кВ; "JEOL", Япония).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

На первом этапе работы исследована способность белка Gag *gypsy* образовывать вирусоподобные частицы без участия нуклеокапсида. Такой подход основан на представлении, что образование вирусоподобной частицы инициируется при взаимодействии нуклеокапсида с РНК [14, 15, 21, 22]. Не так давно на С-конце Gag *gypsy* выявили нуклеокапсид [20]. Нуклеотидная последовательность, кодирующая нуклеокапсидный домен, делетирована в созданной нами конструкции "1ΔNC". Сравнив экспрессию целевых продуктов, синтезируемых конструкциями "1ΔNC" (без нуклеокапсида) и "1" (полноразмерный Gag), оценили способность этих белков формировать частицы, используя для этого тест, примененный нами ранее [11]. Клетки с экспрессированными белками разрушали ультразвуком, отбирали осветленный лизат и ультрацентрифугировали, как принято при очистке вируса. Образовавшийся осадок растворяли и наслаивали на трехступенчатый градиент увеличивающихся концентраций сахарозы 20/30/70%. Как показано ранее для белка конструкции "1" [11], в условиях ультрацентрифугирования (180000 × g, 40 мин) частицы проходят через 20- и 30%-ную сахарозу и задерживаются на границе между 30- и 70%-ной сахарозой. Вместе с тем, белок без нуклеокапсида – конструкция "1ΔNC" – также аккумулировался на 30–70%-ной интерфазе, и этот результат стал отправной точкой для дальнейшей работы.

Электронно-микроскопическое исследование препаратов интерфазы 30- и 70%-ной сахарозы выявило существенное различие между структурами, образуемыми белками "1" и "1ΔNC" (рис. 2а,б). Если Gag *gypsy* (продукт конструкции "1") формирует достаточно однородные глобулярные частицы со средним размером 27 нм (рис. 2а), то его делеционный мутант, с удаленным нуклеокапсидом, образует глобулярные агрегаты, размеры которых варьируют от 3 до 50 нм (рис. 2б). Несмотря на то, что "1ΔNC" формирует неоднородные агрегаты,

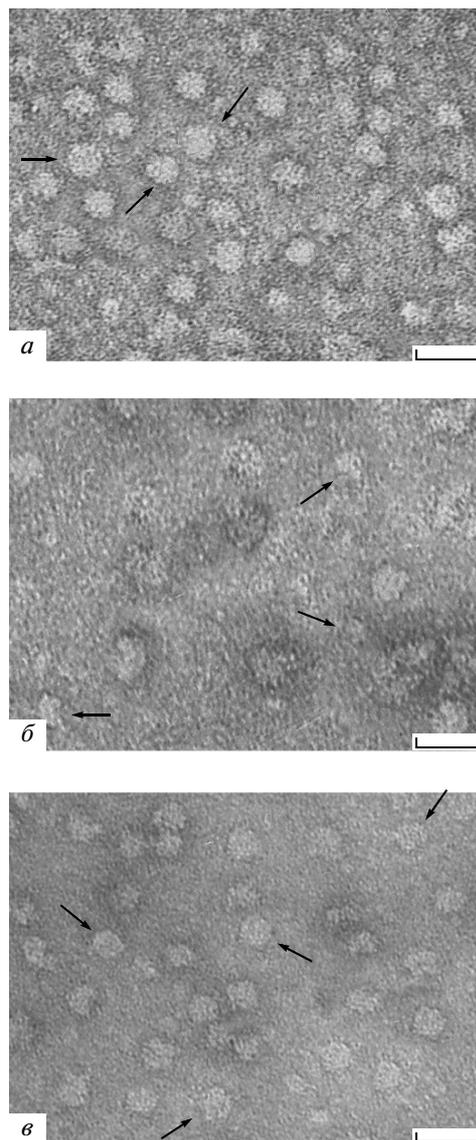


Рис. 2. Электронная микроскопия частиц, очищенных из бактериальных клеток, после экспрессии в них белка "1" (а), "1ΔNC" (б), "1ΔRRM" (в). Стрелками отмечены примеры наиболее характерных для каждого препарата частиц. Масштабный отрезок составляет 50 нм.

полученный результат свидетельствует о способности Gag мультимеризоваться без участия нуклеокапсида.

Ранее нами показано, что структурный белок Gag *gypsy* имеет высокое сродство к нуклеиновым кислотам. Константа связывания белка равна $4 \times 10^8 \text{ M}^{-1}$ [25]. Это значение хорошо коррелирует с константами связывания классически организованных Gag. Поскольку пока нет альтернативного объяснения механизма инициации мультимеризации Gag, кроме как через взаимодействие с нуклеиновой кислотой [14, 15, 21, 22], нами экспериментально проверено, только ли нуклеокапсид

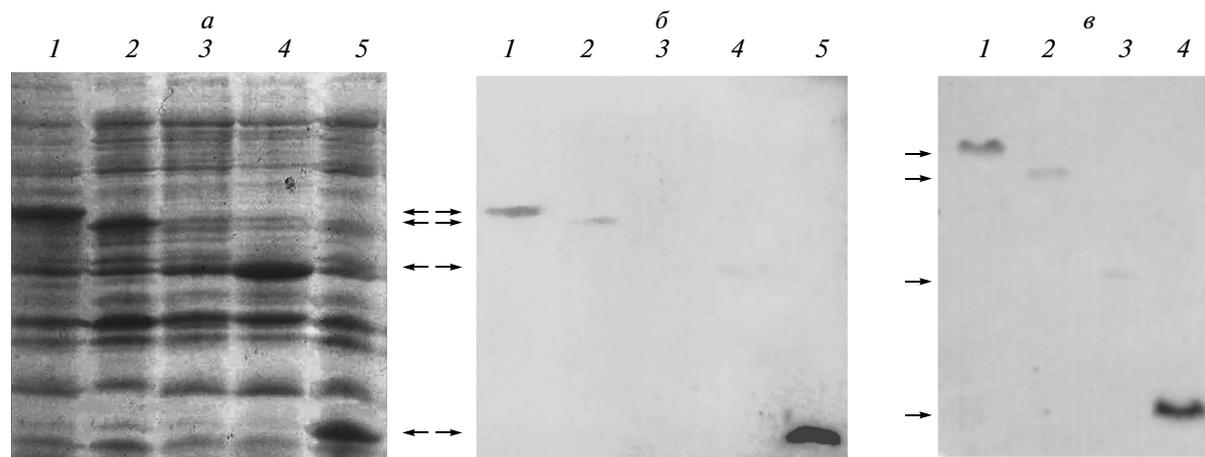


Рис. 3. Тест на связывание с нуклеиновыми кислотами белков “1”, “ Δ RRM”, “ Δ NC” и “NC”. *a* – Электрофоретический анализ лизатов клеток в 12.5%-ном ПААГ. Гель окрашен коллоидным кумасси G-250. 1, 2, 4, 5 – Бактериальные лизаты после экспрессии белков “1”, “ Δ RRM”, “ Δ NC” и “NC” соответственно; 3 – контроль. *б* – Препарат, аналогичный изображенному на рисунке *a*, после переноса на фильтр и связывания с ДНК, меченой [P^{32}] методом “рассеянной затравки”. *в* – Связывание пептидов бактериальных лизатов после экспрессии белков “1”, “ Δ RRM”, “ Δ NC” и “NC” (1–4 соответственно) и фракционирования в 12.5%-ном ПААГ с [P^{32}]-меченой (+)-цепью РНК *gypsy*, синтезируемой *in vitro* при помощи SP6 РНК-полимеразы с плазмиды PGEMgyp.

определяет сродство Gag к нуклеиновым кислотам. Для этого применен метод лиганд-блотинга [25]. Метод основан на связывании иммобилизованных на мембране белков с мечеными нуклеиновыми кислотами. Этот метод был успешно опробован при идентификации структурных белков Tyl1 в лизатах клеток *Saccharomyces cerevisiae* [27] и ранее использован нами для детекции Gag *gypsy* [25].

Конструкции, кодирующие Gag *gypsy* (“1”), мутантную форму Gag без нуклеокапсида (“ Δ NC”) и нуклеокапсид (“NC”), экспрессировали в клетках *E. coli* BL21(DE3). После экспрессии клетки лизировали и подбирали объемы аликвот из лизатов таким образом, чтобы в них были одинаковые массы экспрессированных белков. Затем компоненты лизатов фракционировали с помощью электрофореза в 10%-ном SDS-ПААГ (рис. 3*a*) и переносили на нитроцеллюлозную мембрану электроблотингом и связывали с денатурированной лидерной последовательностью ДНК *gypsy*, меченой P^{32} с помощью метода рассеянной затравки. Результат представлен на рис. 3*б*. Полноразмерная форма белка “1” и нуклеокапсид “NC” эффективно связываются с нуклеиновой кислотой. Разница в эффективности связывания между “1” и “NC” усилена тем, что при равенстве условий иммобилизации (по массе белка) количество нуклеокапсида на фильтре в 2.4 раза больше (соответственно больше центров связывания), чем полноразмерной формы Gag. Кроме того, нельзя исключить возможность более эффективного переноса на мембрану низкомолекулярных полипептидов по сравнению с вы-

сокомолекулярными. Обращает на себя внимание тот факт, что deletированные формы Gag, прежде всего “ Δ NC”, также проявляют слабое сродство к нуклеиновым кислотам (рис. 3*в*). И это, скорее всего, не артефакт, связанный с большим количеством “ Δ NC” на мембране, поскольку фракционированные белки бактериального лизата, взятые в сравнимых количествах (см. рис. 3*a*) и иммобилизованные таким же образом, не проявляли никакого сродства к нуклеиновым кислотам. По-видимому, как и в случае нуклеокапсида, связывание deletированного по нему полипептида с нуклеиновыми кислотами обусловлено положительно заряженными участками с большим содержанием остатков лизина (их осталось 15) и аргинина (16), причем оставшаяся, капсидная, часть Gag образует глобулярные агрегаты (рис. 2*б*). Таким образом, нуклеокапсид не есть исключительная область белка, обеспечивающая связывание с нуклеиновой кислотой, но его присутствие необходимо для образования однородных глобулярных частиц (рис. 2). Из этого можно предположить, что нуклеокапсид обеспечивает специфичность взаимодействия мономеров при сборке частицы.

Нуклеокапсид любого из известных Gag находится на N-конце белка и представлен областью «низкой сложности», которую достаточно просто определить, используя, например, возможности сервера DisProt VL3-N predictor [<http://www.ist.temple.edu/disprot/>]. Именно по этим элементарным признакам идентифицирован нуклеокапсид Gag *gypsy* [20]. Анализируя эту область, мы обратили внимание на участок, расположенный между 333 и 348 аминокислотными остатками белка

(рис. 1), который, наряду со значительным положительным зарядом, содержит последовательность, похожую на SR-мотив [28], определяемую наличием пары остатков Arg и Ser. Этот мотив характерен для белков, специфически связывающихся с нуклеиновыми кислотами. Поскольку последовательность, определяющая РНК-распознающий мотив (RRM) нуклеокапсида Gag *gypsy*, не известна, мы экспериментально оценили вклад вышеуказанного участка в мультимеризацию Gag. Для этого создали делеционный мутант Gag *gypsy* “1ΔRRM” и проанализировали структуры, которые он образует в бактериальной клетке. Использовали описанный выше тест: из водорастворимой части лизата ультрацентрифугированием выделяли фракцию частиц, которую затем также ультрацентрифугированием анализировали с помощью ступенчатого градиента сахарозы. В результате показано, что Gag *gypsy*, у которого удалено больше половины нуклеокапсида на С-конце (белок “1ΔRRM”), способен не менее эффективно, чем полноразмерная форма Gag формировать частицы в бактериальной клетке. Частицы, образуемые “1ΔRRM”, морфологически мало отличались от таковых для полноразмерного белка (рис. 2в). Таким образом, результаты сравнения мультимерных структур, образуемых Gag и двумя его делеционными мутантами, свидетельствуют о том, что нуклеокапсид необходим для построения глобулярных частиц, и эта функция может быть обеспечена его N-концевым (проксимальным) участком.

Ранее нами показано, что основная масса экспрессируемого в бактериях Gag *gypsy* образует макромолекулярные агрегаты – тельца включения – внутри бактериальной клетки [11]. Сходным образом ведут себя и мутанты Gag: “1ΔNC” и “1ΔRRM”. Такое поведение белков довольно типично при сверхэкспрессии в бактериальной клетке [29]. Белки, собранные в тельца включения, аккумулируются в осадке при центрифугировании бактериального лизата, т.е. в водонерастворимой форме [11, 29]. Этот материал и взят нами для моделирования *in vitro* вирусоподобных частиц *gypsy*, т.е. мы рассмотрели, способны ли делеционные мутанты Gag, образующие макромолекулярные агрегаты, формировать частицы *in vitro*.

Стратегию дальнейших экспериментов вырабатывали на основании следующих результатов, полученных для классически организованных Gag: во-первых, для мультимеризации структурного белка необходимы нуклеиновые кислоты и, во-вторых, последовательности, которая лучше других стимулирует этот процесс, не выявлено. Так, добавление любой гетерологичной РНК к чистым препаратам экспрессированных бактериально структурных белков HIV, M-PMV и RSV приводит к их быстрой мультимеризации и образованию частиц *in vitro* [13, 30, 31].

Водонерастворимую фракцию, полученную в виде осадков в результате центрифугирования (при 12000 × g) лизатов бактериальных клеток после экспрессии в них белков, растворяли в 8 М мочеvine и очищали с помощью аффинной хроматографии на никель-агарозе. Затем для полного удаления из препарата нуклеиновых кислот проводили гель-фильтрацию в системе FPLC. Таким образом, получили чистые бактериально экспрессированные денатурированные мономеры Gag (“1”) и два его делеционных мутанта (“1ΔRRM” и “1ΔNC”).

Далее искали ответ на следующие вопросы: 1) возможен ли рефолдинг белков, 2) способны ли ренатурированные белки образовывать вирусоподобные частицы и 3) необходима ли для образования частиц нуклеиновая кислота? В описываемом ниже тесте использовали два типа нуклеиновых кислот: раствор полиаденилированной РНК из клеток *D. melanogaster* (линия 67j25D, клетки этой линии использовали ранее в работах [3] и [26]), который получали по приведенной нами ранее методике [3], и 22-членный олигонуклеотид, состоящий из повторяющейся пары GT. Нуклеотидный состав (GT)₂₂ выбран на том основании, что именно такой олигонуклеотид использовали при тестировании *in vitro* мультимеризации Gag, имеющих ССНС-мотив, и размер нуклеотидной заправки в 22 нуклеотида обеспечивал эффективную мультимеризацию мономеров [18, 22, 32].

Отобрав по три аликвоты из каждого очищенного белкового препарата (концентрация белка 0.3 мг/мл), добавляли к одной из них раствор полиаденилированной РНК, к другой – олигонуклеотид, а третью использовали как контроль. Масса добавленных нуклеиновых кислот составляла примерно 6% от массы белка, поскольку, по литературным данным, на примере структурных белков HIV [33] и RSV [13, 32], такое соотношение было найдено наиболее эффективным. С помощью диализа в препаратах снижали концентрацию мочеvine от 8 до 0.25 М, наслаивали на ступенчатый градиент 20/30/70%-ной сахарозы и центрифугировали при 180000 × g в течение 45 мин – как и в предыдущих экспериментах по выделению вирусоподобных частиц. Затем для каждого образца анализировали фракцию, соответствующую интерфазе 30- и 70%-ной сахарозы (рис. 4). Если нуклеиновую кислоту в образец не добавляли, то белка в рассматриваемой зоне практически не было, т.е. макромолекулярные структуры не образовывались. Вместе с тем, основная масса белка во всех препаратах, куда была добавлена РНК, находилась на границе 30- и 70%-ой сахарозы. При добавлении олигонуклеотида полноразмерная форма белка, также как и в присутствии РНК, эффективно аккумулировалась в “зоне частиц” градиента. Обращает на себя

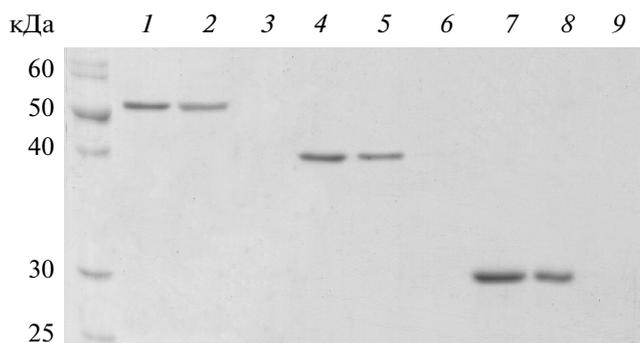


Рис. 4. Анализ мультимеризации белков в присутствии нуклеиновой кислоты. Электрофорез в 12,5%-ном SDS-ПААГ фракций интерфазы 30/70%-ной сахарозы: 1–3 – “1”, 4–6 – “1ΔRRM”, 7–9 – “1ΔNC”. В образцы 1, 4, 7 препаратов очищенного белка добавляли РНК; в образцы 2, 5, 8 – 22-членный дезоксирибонуклеотид; образцы 3, 6, 9 не содержали нуклеиновой кислоты.

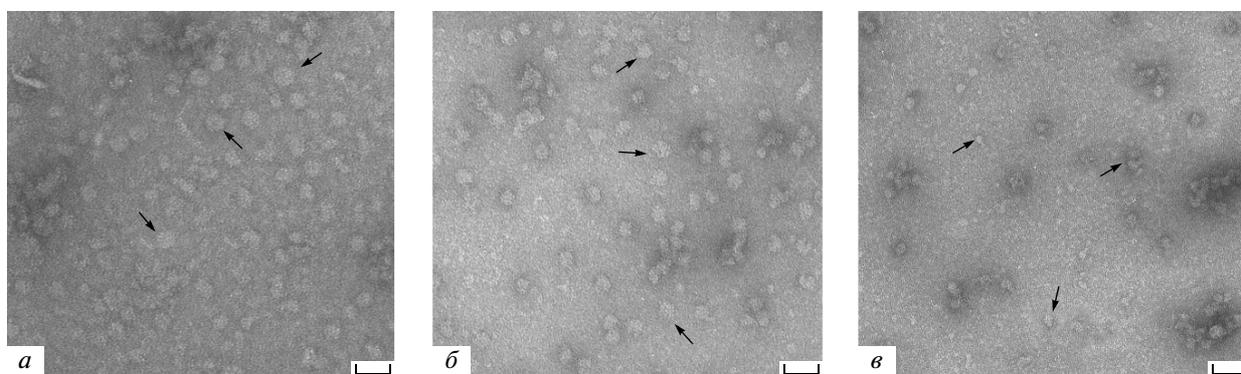


Рис. 5. Электронная микрофотография препаратов частиц, образованных *in vitro* из изначально денатурированного белка в присутствии РНК. На микрофотографиях представлен материал интерфазы 30/70%-ной сахарозы, полученный в результате ультрацентрифугирования препаратов, содержащих белки конструкций “1” (а), “1ΔRRM” (б) или “1ΔNC” (в). Стрелками отмечены примеры наиболее характерных для препарата частиц. Масштабный отрезок составляет 50 нм.

внимание тот факт, что белок с делетированным нуклеокапсидом (“1ΔNC”) также проявлял способность эффективно мультимеризоваться в присутствии нуклеиновых кислот.

Электронно-микроскопическое исследование образцов, полученных из препаратов интерфазной фракции, показало, что все рассматриваемые белки способны эффективно мультимеризоваться при добавлении нуклеиновой кислоты. Для каждого белка частицы, образуемые при добавлении РНК, схожи с частицами, образуемыми при добавлении олигодезоксирибонуклеотида, и подобны тем, которые выделены из бактериальных клеток (рис. 5). Средний диаметр частиц в образцах белков конструкций “1” и “1ΔRRM” составлял 27 нм. Кроме того, в тех образцах белка “1”, где добавлена РНК, иногда встречались тубулярные (цилиндрические) структуры (рис. 5а); их образование может быть обусловлено “наматыванием” Gag на длинные молекулы РНК, во всяком случае, такой механизм обсуждался ранее [18]. В

образцах “1ΔRRM” и “1ΔNC” таких структур не зарегистрировано. Отметим, что, несмотря на эффективную мультимеризацию в присутствии нуклеиновой кислоты, белок без нуклеокапсида (“1ΔNC”) образовывал глобулы, размер которых варьировал в широких пределах (диаметр от 5 до 20 нм), и которые проявляли склонность к образованию агрегатов (рис. 5в).

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

На основании результатов, полученных как при изучении экспрессии Gag *gypsy*, так и структурных белков других ретровирусов [13, 30, 31, 34], можно предположить, что независимо от аминокислотной последовательности структурного белка его экспрессия в гетерологичной системе приводит к образованию вирусоподобных частиц.

К настоящему времени структура Gag *gypsy* детально не изучена, но сравнительный биоинформатический анализ белка позволяет считать, что в

нем отсутствует домен, соответствующий матриксу, и, таким образом, он состоит только из капсида и нуклеокапсида [11], и в этом отношении Gag *gypsy* больше напоминает не Gag ретровирусов позвоночных, а структурные белки ретротранспозонов дрожжей Ty1 и Ty3 [34].

В этой работе показано, что мультимеризация капсида может происходить без участия нуклеокапсида; и этот факт согласуется с результатами авторов, которые изучали способность делеционных мутантов структурных белков других ретровирусов и ретротранспозонов формировать частицы *in vitro* [34–37]. Кроме того, полученные нами результаты хорошо согласуются с наблюдением, что добавление нуклеиновой кислоты значительно ускоряет мультимеризацию белка с мотивом “цинковые пальцы” [14, 22].

Экспрессия капсидной части структурного белка (“1ΔNC”) в бактериях, как и моделирование самосборки такого полипептида *in vitro*, приводит к образованию неоднородных глобулярных агрегатов. Из этого можно предположить, что нуклеокапсид Gag *gypsy* не есть необходимое условие для мультимеризации белка. Однако несомненно его роль в организации и симметрии частиц, которые образуются через взаимодействие белка с нуклеиновой кислотой. Нами показано, что мультимеризация мономеров белков происходит как при добавлении РНК, так и ДНК. Причем нуклеотидный состав добавляемых нуклеиновых кислот не важен – такой вывод сделан нами на основании результатов, приведенных в предыдущем разделе и полученных в дополнительных экспериментах, где к препаратам белка добавляли различные нуклеиновые кислоты. Например, экстракт РНК из бактериальных клеток также вызывал эффективную мультимеризацию белков. Эти результаты хорошо согласуются с данными, полученными на структурных белках, имеющих “цинковые пальцы”: Gag RSV [13, 22] и Gag Ty3 [34].

Нуклеокапсид любого из известных Gag находится на N-конце белка и представлен областью “низкой сложности” [38]. Именно по этим простейшим признакам ранее был определен нуклеокапсид Gag *gypsy*, а затем экспериментально доказана одна из его функций – способность отжигать тРНК к РНК *gypsy* [20]. Для обеспечения связывания с РНК нуклеокапсид должен содержать последовательность или последовательности, распознающие РНК – “РНК-распознающий мотив” (RRM). Хорошо известно, что центр RRM в классически организованных структурных белках ретровирусов представлен одним или двумя мотивами, образованными из остатков цистеина и гистидина (CX₂CX₄NX₄C), так называемыми CCHC-мотивами “цинковые пальцы” [39]. В Gag *gypsy* такой мотив отсутствует. Более того, сравнение аминокислотной последовательности нук-

леокапсида *gypsy* с последовательностями известных нуклеокапсидов, представленных в базах данных [<http://blast.ncbi.nlm.nih.gov/>], не выявило никакого значимого сходства. В первичной структуре нуклеокапсида Gag *gypsy* не найдено и других канонических RRM: ни [RK]-G-[FY]-[GA]-[FY]-[ILV]-X-[FY] или [ILV]-[FY]-[ILV]-XN-L [40], ни последовательностей, обогащенных пролином, которые также могут определять связывание с РНК [41]. Вместе с тем, в первичной структуре нуклеокапсида *gypsy* мы выделяем область, которую можно связать с серин-аргининовым RRM (SR-RRM) [28, 42, 43], – с 333-го по 348-й аминокислотный остаток (рис. 1). Хотя белки, имеющие SR-RRM, изучают уже более 20 лет, не установлен общий аминокислотный консенсус этого РНК-распознающего мотива. Замечено, что в SR-мотивах часто присутствуют остатки серина и аргинина, расположенные парами [28, 44], что характерно и для рассматриваемой области нуклеокапсида (рис. 1). Кроме того, в этой же области нуклеокапсида находится большое количество положительно заряженных аминокислотных остатков, которые и определяют высокий положительный заряд белка. Обычно SR-RRM находится рядом с другими, вышеупомянутыми, типами RRM [28, 43–45], но это не относится к нуклеокапсиду Gag *gypsy*. Таким образом, даже с точки зрения функционирования SR-RRM Gag *gypsy* организован нестандартно. Вместе с тем, область 333–348 аминокислотной последовательности Gag имеет сходство с участками ряда белков, специфически связывающихся с нуклеиновыми кислотами (таблица). Такой результат наводит на мысль, что указанная область нуклеокапсида может быть основным, если не единственным, участком специфического связывания Gag с РНК *gypsy*, что в эукариотической клетке дрожифилы инициирует образование частиц ретротровируса.

Чтобы оценить, насколько важна область с 333 по 348 аминокислотные остатки нуклеокапсида Gag при экспрессии белка в бактериях, мы удалили указанную область вместе с C-концом нуклеокапсида (белок “1ΔRRM”). Показано, что делетированный белок эффективно образует глобулярные частицы того же размера, что и полноразмерная форма. Значит, проксимальной части нуклеокапсида достаточно для того, чтобы обеспечить построение однородных глобулярных частиц в бактериях и *in vitro*. Вместе с тем, несмотря на то, что частицы, образуемые в гетерологичной бактериальной системе, напоминают частицы, образуемые в клетках дрожифилы [11], отсутствие какой либо определенности в последовательности нуклеиновых кислот при сборке бактериально полученных белков в частицу, указывает на ограниченный характер действия нуклеокапсида в этом случае. То есть за рамками модельных экспериментов, представленных в работе, остались механизмы специфиче-

Сравнение участка аминокислотной последовательности Gag *gypsy* (остатки 333–348) со специфическими последовательностями белков, связывающих нуклеиновые кислоты

№	Идентичность/сходство, %	Последовательность		Ссылка
1	100.0/100.0	333	KRRNSSERSTGPRRQR	348 Gag <i>gypsy</i>
2	64.7/76.5	39	KRRSGSERGSTVTRRER	55 ref XP_003263241.1
3	58.8/70.6	131	KRTSSSERVSPGGRRE	146 sp O00423.3 EMAL1_HUMAN
4	62.5/68.8	457	DRRNDRERS-PRRER	470 ref ZP_03841440.1
5	62.5/68.8	462	DRRNDRERS-PRRER	475 ref YP_002153108.1
6	56.2/68.8	5	GRRNRRRN-GPRRAR	19 sp P17966.1 CAPSD_BYDVM
7	56.2/68.8	904	VVRNSSERSLGPFNKQ	919 gb EFW96029.1
8	50.0/68.8	1044	RDRSGSQSTSRRRQR	1759 sp Q8BTI8.3 SRRM2_MOUSE
9	43.8/68.8	1051	SRRSSDRSSSRERAK	1066 sp Q9P1Y6.3 PHRF1_HUMAN
10	52.6/68.4	138	AKMDSSQPERSSRRPR	156 sp Q9H160.2 ING2_HUMAN
11	50.0/62.5	78	RRRKSSERKVGHRQTY	93 sp Q0SM76.1 RL21_BORAP
12	43.8/62.5	241	RSRSPKRRSPRRER	257 sp Q8NAV1.1 PR38A_HUMAN
13	56.2/56.2	260	VRRRSRSRSPGRRRHR	274 sp Q6PG31.1 RNPS1_DANRE
14	56.2/56.2	439	SGGRGTNPRRQ	451 ref NP_738196.1
15	56.2/56.2	1013	PRR-ERDDGPRRDR	1025 sp B3P211.1 EIF3A_DROER
16	50.0/56.2	590	ARRNSSERTLSPGLPS	605 sp P57058.1 HUNK_HUMAN
17	37.5/56.2	204	VEEGLERTTGPRLRQ	219 pdb 3H38 A
18	55.0/55.0	203	RRGSSPPPWERSTQPRR	219 ref ZP_07299269.1
19	50.0/50.0	1000	SHGRSRERSTSARRSR	1015 sp Q8TFG6.1 PPK18_SCHPO

ского белок-нуклеинового узнавания, которые имеют место при репликации вируса в хозяйских клетках.

Подводя итог, отметим, что в результате проделанной работы показано, что мультимеризация капсидной части белка Gag *gypsy* может происходить без участия нуклеокапсида, хотя последний существенно влияет на структуру образуемых частиц. Показано, что для формирования белком Gag *gypsy* однородных глобулярных частиц в гетерологической системе важна только N-концевая часть нуклеокапсида. Добавление нуклеиновых кислот, как РНК, так и ДНК, инициирует образование частиц мономерами Gag и его делеционных мутантами, рассмотренными в работе. Не выявлено ни какой-либо определенной нуклеотидной последовательности, ни оптимального размера нуклеиновой кислоты, запускающей мультимеризацию бактериально экспрессированных белков.

Авторы благодарят Т. Русанову и М. Привалову за помощь при манипуляциях с бактериями, очистке белков и нуклеиновых кислот.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований: проекты 08-04-00227-а и 09-04-01731-а.

СПИСОК ЛИТЕРАТУРЫ

- Jern P., Sperber G.O., Blomberg J. 2005. Use of Endogenous Retroviral Sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy. *Retrovirology*, 2, doi:10.1186/1742-4690-2-50.
- Gao F., Bailes E., Robertson D.L., Chen Y., Rodenburg C.M., Michael S.F., Cummins L.B., Arthur L.O., Peeters M., Shaw G.M., Sharp P.M., Hahn B.H. 1999. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature*, 397, 436–441.
- Syomin B.V., Fedorova L.I., Surkov S.A., Ilyin Y.V. 2001. The endogenous *Drosophila melanogaster* retrovirus *gypsy* can propagate in *Drosophila hydei* cells. *Mol. Gen. Genet.* 264, 588–594.
- Yohn C.T., Jiang Z., McGrath S.D., Hayden K.E., Khaitovich P., Johnson M.E., Eichler M.Y., McPherson J.D., Zhao S., Paabo S., Eichler E.E. 2005. Lineage-specific expansions of retroviral insertions within the genomes of African great apes but not humans and orangutans. *PLoS Biol.* 3, 1–11.
- International human genome sequencing consortium. 2001. Initial sequencing and analysis of the human genome. *Nature*, 409, 860–921.
- Сёмин Б.В., Ильин Ю.В. 2003. Эрантивиромы *Drosophila*. *Генетика*, 39, 657–663.
- Boeke J.D., Eickbush T.H., Sandmeyer S.B., Voytas D.F. 2000. Metaviridae. In: *Virus Taxonomy: ICTV VIIth Report*. Ed. Murphy F.A. N.Y.: Springer-Verlag, pp.123–135.
- Ludwig A., Valente V.L., Loreto E.L. 2008. Multiple invasions of Errantivirus in the genus *Drosophila*. *Insect. Mol. Biol.* 17, 113–124.
- Song S.U., Gerasimova T., Kurkulos M., Boeke J.D., Corces V.G. 1994. An Env-like protein encoded by a *Drosophila* retroelement: evidence that *gypsy* is an infectious retrovirus. *Gen. Dev.* 8, 2046–2057.
- Pelisson A., Mejlumian L., Robert V., Terzian C., Bucheton A. 2002. *Drosophila* germline invasion by the endogenous retrovirus *gypsy*: involvement of the viral *env* gene. *Insect. Biochem. Mol. Biol.* 160, 201–209.

11. Сёмин Б.В., Иванова Л.А., Попенко В.И., Ильин Ю.В. 2011. Структурный белок Gag ретровируса *D. melanogaster gypsy* формирует вирусоподобные частицы в бактериальной клетке. *Молекуляр. биология*. **45**, 517–523.
12. Сёмин Б.В., Ильин Ю.В. 2006. Детекция структурного белка (Gag) эндогенного ретровируса МДГ4 (gypsy) в культивируемых клетках. *Докл. Акад. Наук*. **408**, 1–3.
13. Campbell S., Vogt V.M. 1995. Self-assembly *in vitro* of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. *J. Virol.* **69**, 6487–6497.
14. Adamson C.S., Jones I.M. 2004. The molecular basis of HIV capsid assembly five years of progress. *Rev. Med. Virol.* **14**, 107–121.
15. Scarlata S., Carter C. 2003. Role of HIV-1 Gag domains in viral assembly. *Biochim. Biophys. Acta Biomembr.* **1614**, 62–72.
16. Linial M.L. 1999. Foamy viruses are unconventional retroviruses. *J. Virol.* **73**, 1747–1755.
17. Rumlova-Klikova M., Hunter E., Nermut M.V., Pichova I., Ruml T. 2000. Analysis of Mason-Pfizer monkey virus Gag domains required for capsid assembly in bacteria: role of the N-terminal proline residue of CA in directing particle shape. *J. Virol.* **74**, 8452–8459.
18. Yu F., Joshi S.M., Ma Y.M., Kingston R.L., Simon M.N., Vogt V.M. 2001. Characterization of Rous sarcoma virus Gag particles assembled *in vitro*. *J. Virol.* **75**, 2753–2764.
19. Morikawa Y., Goto T., Momose F. 2004. Human immunodeficiency virus type 1 Gag assembly through assembly intermediates. *J. Biol. Chem.* **279**, 31964–31972.
20. Gabus C., Ivanyi-Nagy R., Depollier J., Bucheton A., Pelisson A., Darlix J.-L. 2006. Characterization of a nucleocapsid-like region and of two distinct primer tRNA^{Lys,2} binding sites in the endogenous retrovirus Gypsy. *Nucl. Acids Res.* **34**, 5764–5777.
21. Taylor G.M., Ma L., Vogt V.M., Post C.B. 2010. NMR relaxation studies of an RNA-binding segment of the rous sarcoma virus gag polyprotein in free and bound states: a model for autoinhibition of assembly. *Biochemistry*. **49**, 4006–4017.
22. Johnson M.C., Scobie H.M., Ma Y.M., Vogt V.M. 2002. Nucleic acid-independent retrovirus assembly can be driven by dimerization. *J. Virol.* **76**, 11177–11185.
23. Маниатис Т., Фрич Э., Сэмбрук Дж. 1984. *Молекулярное клонирование*. М.: Мир.
24. Сёмин Б.В., Pelisson A., Ильин Ю.В., Bucheton A. 2004. Экспрессия структурного белка Gag ретровируса gypsy рекомбинантным бакуловирусом в культуре клеток *Spodoptera frugiperda*. *Докл. Акад. Наук*. **398**, 702–704.
25. Сёмин Б.В., Турапов О.А., Степанов А.С., Ильин Ю.В. 1999. Связывание с нуклеиновыми кислотами белка, кодируемого первой открытой рамкой считывания ретротранспозона gypsy (МДГ4). *Молекуляр. биология*. **33**, 423–427.
26. Syomin B.V., Kandror K.V., Semakin A.B., Tsuprun V.L., Stepanov A.S. 1993. Presence of the gypsy (MDG4) retrotransposon in extracellular virus-like particles. *FEBS Lett.* **323**, 285–288.
27. Mellor J., Fulton A.M., Dobson M.J., Roberts N.A., Wilson W., Kingsman A.J., Kingsman S.M. 1985. The Ty transposon of *Saccharomyces cerevisiae* determines the synthesis of at least three proteins. *Nucl. Acids Res.* **13**, 6249–6262.
28. Shepard P., Hertel K.J. 2009. SR protein family. *Genome Biol.* **10**, 242.
29. Burgess R.R. 2009. Refolding solubilized inclusion body proteins. *Methods Enzymol.* **463**, 259–282.
30. Strambio-De-Castillia C., Hunter E. 1992. Mutational analysis of the major homology region of Mason-Pfizer monkey virus by use of saturation mutagenesis. *J. Virol.* **66**, 7021–7032.
31. Morikawa Y., Goto T., Sano K. 1999. *In vitro* assembly of human immunodeficiency virus type 1 Gag protein. *J. Biol. Chem.* **274**, 27997–28002.
32. Ma Y.M., Vogt V.M. 2002. Rous sarcoma virus Gag protein-oligonucleotide interaction suggests a critical role for protein dimer formation in assembly. *J. Virol.* **76**, 5452–5462.
33. Parker S.D., Wall J.S., Hunter E. 2001. Analysis of Mason-Pfizer monkey virus Gag particles by scanning transmission electron microscopy. *J. Virol.* **75**, 9543–9548.
34. Larsen L.S., Kuznetsov Y., McPherson A., Hatfield G.W., Sandmeyer S. 2008. TY3 GAG3 protein forms ordered particles in *Escherichia coli*. *Virology*. **370**, 223–227.
35. Muriaux D., Mirro J., Harvin D., Rein A. 2001. RNA is a structural element in retrovirus particles. *Proc. Natl. Acad. Sci. USA*. **98**, 5246–5251.
36. Sandmeyer S.B., Clemens K.A. 2010. Function of a retrotransposon nucleocapsid protein. *RNA Biol.* **7**, 642–654.
37. Rumlova-Klikova M., Hunter E., Nermut M.V., Pichova I., Ruml T. 2000. Analysis of Mason-Pfizer monkey virus Gag domains required for capsid assembly in bacteria: role of the N-terminal proline residue of CA in directing particle shape. *J. Virol.* **74**, 8452–8459.
38. Obradovic Z., Peng K., Vucetic S., Radivojac P., Brown C.J., Dunker A.K. 2003. Predicting intrinsic disorder from amino acid sequence. *Proteins*. **53**, 566–572.
39. Mely Y., Piemont E., Sorinas-Jimeno M., de Rocquigny H., Jullian N., Morellet N., Roques, B.P., Gerard D. 1993. Structural and dynamic characterization of the aromatic amino acids of the human immunodeficiency virus type I nucleocapsid protein zinc fingers and their involvement in heterologous tRNA(Phe) binding: a steady-state and time-resolved fluorescence study. *Biophys. J.* **65**, 1513–1522.
40. Kielkopf C.L., Lücke S., Green M.R. 2004. U2AF homology motifs: protein recognition in the RRM world. *Genes Dev.* **18**, 1513–1526.
41. Grainger R.J., Beggs J.D. 2005. Prp8 protein: At the heart of the spliceosome. *RNA*. **11**, 533–557.
42. Chou T.B., Zachar Z., Bingham P.M. 1987. Developmental expression of a regulatory gene is programmed at the level of splicing. *EMBO J.* **6**, 4095–4104.
43. Amrein H., Gorman M., Nothiger R. 1988. The sex-determining gene tra-2 of *Drosophila* encodes a putative RNA binding protein. *Cell*. **55**, 1025–1035.
44. Hargous Y., Hautbergue G.M., Tintaru A.M., Skrisovska L., Golovanov A.P., Stevenin J., Lian L.Y., Wilson S.A., Allain F.H. 2006. Molecular basis of RNA recognition and TAP binding by the SR proteins SRp20 and 9G8. *EMBO J.* **25**, 5126–5137.
45. Zahler A.M., Lane W.S., Stolk J.A., Roth M.B. 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev.* **6**, 837–847.