

## Сравнительный анализ первичной и вторичной структуры капсидных белков карлавирусов растений

Т. Ю. Ткаченко, С. С. Витер, Ю. Л. Радавский

Институт биоорганической химии и нефтехимии НАН Украины  
Ул. Мурманская, 1, Киев, 02094, Украина

*Проанализированы первичная и вторичная структуры капсидных белков 23 карлавирусов растений. В результате анализа первичной структуры методом множественного выравнивания выделены семь мотивов, три из которых являются специфичными для капсидных белков карлавирусов. Анализ вторичной структуры этих белков на основе множественного сравнения выявил высокую степень консервативности этой структуры в коровой и С-концевой областях, а также ее значительную переменчивость в N-концевой области. Предложена общая схема вторичной структуры, характерная для капсидных белков карлавирусов растений.*

**Введение.** Карлавирусы (*Carlavirus*) — достаточно гомогенная группа фитовирусов, поражающих ряд важных сельскохозяйственных культур, таких как картофель, горох, люцерна, хмель и др. Вирусные частицы карлавирусов представляют собой слабо изогнутые нити диаметром 12—13 и длиной 610—700 нм. Белковый капсид карлавирусов состоит из однотипных субъединиц, расположенных по спирали и образующих полую трубку. Геномная однонитчатая плюс-РНК уложена спирально в полости белкового капсида и образует с ним нековалентные связи [1].

Несмотря на широкое распространение в природе, группа карлавирусов долгое время оставалась малоизученной. Это обусловлено тем, что большинство представителей данной группы при заражении растений-хозяев вызывают слабые симптомы или латентную инфекцию. Кроме того, низкое содержание в тканях зараженных растений, склонность к агрегации и нестабильность вирусных частиц затрудняют выделение и идентификацию карлавирусов. Благодаря развитию методов молекулярной биологии интерес к представителям этой группы значительно возрос. За последние 12 лет установлены полные нуклеотидные последовательности пяти карлавирусов: М-вируса картофеля [2],

вируса костра голубики [3], латентного вируса хмеля [4], латентного вируса аконита [5], латентного вируса чеснока [6]. Установлены также 3'-концевые нуклеотидные последовательности РНК (в том числе гены капсидных белков) для 18 представителей группы, включая штаммы и изоляты. Было показано, что экспрессия вирусного генома осуществляется путем трансляции с геномной и минимум двух субгеномных РНК [7]. Геном карлавирусов содержит шесть открытых рамок считывания (ОРС) и 3'-концевую поли(А)-последовательность. ОРС5 кодирует капсидный белок, а роль остальных продуктов генома в жизненном цикле вируса окончательно не установлена.

Неизученной остается трехмерная структура капсидного белка карлавирусов. Исследование пространственной организации вирусных структурных белков имеет важное значение для раскрытия механизма взаимодействия белок—белок и белок—РНК, лежащего в основе такого феномена, как самосборка вирусной частицы. Кроме того, данные о структуре вирусного капсида необходимы при изучении биологии и эволюции вирусов, а также для создания чувствительных диагностических тест-систем.

Для большинства вирусов растений пространственная структура капсидного белка не установлена, что связано со сложностью кристаллизации

этих белков, необходимой для последующего проведения рентгеноструктурного анализа (РСА). Для палочковидных фитовирусов методом РСА была установлена структура капсидного белка трех тобамовирусов, в том числе и вируса табачной мозаики (TMV, род *Tobamovirus*) [8—10]. Ряд косвенных данных об особенностях пространственной организации капсидных белков различных фитовирусов получен методами эпитоппного картирования, тритиевой планиграфии, ядерного магнитного резонанса, сайт-направленного мутагенеза. В связи с тем, что на сегодняшний день для большинства белков аминокислотная последовательность является единственной доступной информацией, активно развиваются методы математического анализа первичной структуры белка: сравнение аминокислотных последовательностей путем попарного и множественного выравнивания, предсказание вторичной и третичной структуры, функциональных доменов, антигенных детерминант и др.

В данной работе проведен анализ первичной структуры капсидных белков карлавирусов с использованием метода множественного выравнивания (CLUSTAL W [11]), программы поиска мотивов в белковой базе данных (PROSITE [12]) и метода предсказания вторичной структуры (SOPMA [13]). Метод множественного выравнивания CLUSTAL W позволяет выделить консервативные и переменные участки белков, установить степень гомологии последовательностей и их эволюционно-го родства.

С помощью программы PROSITE осуществляется поиск мотивов в белковых последовательностях, входящих в базы данных. Метод предсказания вторичной структуры SOPMA дает возможность с вероятностью до 73 % установить основные элементы вторичной структуры в полипептидной цепи ( $\alpha$ -спирали,  $\beta$ -складки,  $\beta$ -повороты, участки с нерегулярной структурой) каждого белка. На основании результатов предсказания вторичной структуры для индивидуальных белков проведено множественное сравнение этих структур, в результате которого выведена общая схема организации вторичной структуры, характерная для капсидных белков карлавирусов.

**Материалы и методы.** Поиск первичных структур капсидных белков карлавирусов, использованных в данной работе, был осуществлен в базе данных белковых последовательностей Entrez Proteins, National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov>). В результате поиска обнаружены следующие 23 первичные структуры капсидных белков, подтвержденные публикациями в научных изданиях: М-вируса картофеля (Potato

virus M — PVM) (русский и немецкий штаммы PVM Ru и PVM Germ соответственно) [2, 14], вируса костра голубики (Blueberry scorch virus — BbScV) (штаммы NewJersey-1, NewJersey-2 и Washington-2 — BbScV NJ-1, BbScV NJ-2 и BbScV Wa-2 соответственно) [3], латентного вируса хмеля (Hop latent virus — HLV) [4], латентного вируса аконита (Aconitum latent virus — AcLV) [5], латентного вируса чеснока (Garlic latent virus — GLV) [6], S-вируса картофеля (Potato virus S — PVS) (штаммы Ordinary и Peruvian — PVS O и PVS P соответственно) [15, 16], S-вируса гелениума (Helium virus S — HelVS) [17], бессимптомного вируса лилии (Lily symptomless virus — LSV) (штаммы голландский и корейский — LSV и LSV Kor соответственно) [18, 19], латентного вируса гвоздики (Carnation latent virus — CLV) [20], вируса мозаики тополя (Poplar mosaic virus — PopMV) [21], вируса мозаики хмеля (Hop mosaic virus — HMV) [22], латентного вируса каланхоэ (Kalanchoe latent virus — KLV) [23], обыкновенного латентного вируса чеснока (Garlic common latent virus — GCLV) [24], латентного вируса лука-шалота (Shallot latent virus — SLV) [24], вируса В хризантемы (Chrysanthemum virus B — CVB) [25], латентного вируса лилии (Lily latent virus — LiLV) [26], вируса слабой крапчатости вигны (Cowpea mild mottle virus — CPMMV) (штаммы CPMMV-M и CPMMV-H) [27].

*Множественное выравнивание и поиск мотивов.* Множественное выравнивание первичных структур капсидных белков карлавирусов проводили с использованием программы CLUSTAL W [11] со стандартными параметрами выравнивания.

Программа CLUSTAL W использует метод прогрессивного множественного выравнивания аминокислотных или нуклеотидных последовательностей, включающий три основных этапа: каждая из анализируемых последовательностей попарно выравнивается с остальными для расчета степени их сходства, на основании степени сходства отдельных последовательностей строится дендрограмма, в соответствии с порядком ветвей на которой выполняется прогрессивное множественное выравнивание последовательностей. Значительное увеличение чувствительности метода CLUSTAL W по сравнению с другими подходами было достигнуто путем модификации третьего этапа: при прогрессивном множественном выравнивании последовательностей учитывали не только степень сходства отдельных последовательностей, но и различия в их длине, в профилях гидрофильности, возможную упорядоченность/нерегулярность структуры отдельных участков и другие параметры [11].

Для представления результатов множественного выравнивания нами использован общепринятый однобуквенный код для каждого аминокислотного остатка. При описании мотивов использованы следующие обозначения:

1) идентичные для данной позиции остатки обозначены соответствующим однобуквенным кодом;

2) если в данной позиции находились два—четыре остатка, то эти остатки заключались в квадратные скобки, например, [GT], {MKPT};

3) обозначение X (т. е. любой остаток) использовалось при наличии в данной позиции более четырех остатков;

4) если в ряде последовательностей остаток для данной позиции отсутствует (разрыв при выравнивании), то эта позиция обозначается как остаток/[остатки](0,1), например G(0,1) или [G, T, V](0,1), где «0» обозначалась разрыв, а «1» — наличие остатка (остатков).

Сканирование базы данных известных белковых последовательностей (SWISS-PROT, 121515, TrEMBL, 727844 последовательностей) на наличие мотивов, сходных с выделенными нами после выравнивания капсидных белков карлавирусов, проводили с помощью программы PROSITE (Database of protein families and domains) [12] (<http://cn.expasy.org/prosite/>; <http://cn.expasy.org/tools/scan-prosite/>).

*Предсказание вторичной структуры* капсидных белков карлавирусов проводили по методу SOPMA (self-optimized prediction from multiple alignment) [13, 28] со следующими параметрами: количество конформационных состояний (Number of conformational states) — 4 ( $\alpha$ -спираль,  $\beta$ -складка,  $\beta$ -поворот, нерегулярная структура), порог подобия (Similarity threshold) — 8, ширина окна (Window width) — 17.

Метод SOPMA состоит из следующих этапов: 1) поиск последовательностей, гомологичных анализируемой, в базе белков с известной структурой; 2) анализируемая последовательность и найденные в базе гомологичные последовательности выравниваются методом CLUSTAL W; 3) для каждого из отобранных белков осуществляется предсказание вторичной структуры с использованием алгоритма сходства и специальной процедуры оптимизации параметров предсказания (метод SOPM — self-optimized prediction method); 4) по данным множественного выравнивания последовательностей для каждого остатка в данной позиции усредняется вероятность его нахождения в различных конформационных состояниях ( $\alpha$ -спираль,  $\beta$ -складка,  $\beta$ -поворот, нерегулярная структура); 5) на основании

результатов четвертого этапа каждому остатку анализируемого белка присваивается одно из конформационных состояний, имеющее для данной позиции наибольшую вероятность [13, 28].

Использованные в работе методы компьютерного анализа последовательностей доступны через ExPASy (Expert Protein Analysis System) Molecular Biology Server (<http://www.expasy.org>).

**Результаты и обсуждение.** Во множественном выравнивании были использованы первичные структуры капсидных белков 23 вирусов карлагруппы, в том числе у пяти вирусов анализировали по 2—3 штамма или изолята. В результате проведенного анализа установлено, что большинство капсидных белков карлавирусов имеет высокую степень гомологии друг с другом (рис. 1). Однако при имеющемся сходстве отдельных структурных белков представителей группы существуют и различия, часто коррелирующие со степенью серологического родства между ними [29].

В результате множественного выравнивания аминокислотные остатки капсидных белков были распределены по 333 позициям (рис. 1), номера которых не совпадают с таковыми в индивидуальных последовательностях из-за наличия разрывов в последовательностях (gaps). При дальнейшем обсуждении результатов будут использованы номера позиций, установленные при множественном выравнивании.

Анализ множественного выравнивания показал, что в 34 позициях белки содержат идентичные остатки (отмечены звездочкой, рис. 1), а в 28 позициях — остатки одинаковой химической природы. В 36 позициях остатки идентичны для большинства последовательностей, причем в 14 случаях эти отличия касались капсидного белка PopMV (рис. 1). При исключении последовательности PopMV из множественного выравнивания количество идентичных остатков возросло до 47, а 30 позиций содержали остатки одинаковой химической природы (данные этого выравнивания не приведены).

N-концевые фрагменты капсидных белков (приблизительно 75 позиций) характеризуются высокой вариабельностью (рис. 1). У большинства белков эти фрагменты богаты глицином и пролином, которые препятствуют образованию или деформируют регулярные элементы вторичной структуры, такие как  $\alpha$ -спирали и  $\beta$ -складки. N-концевой участок обогащен также заряженными остатками, увеличивающими его гидрофильность и тем самым повышающими вероятность того, что фрагменты этого участка могут быть экспонированы на поверхности белковой субъединицы. Ранее

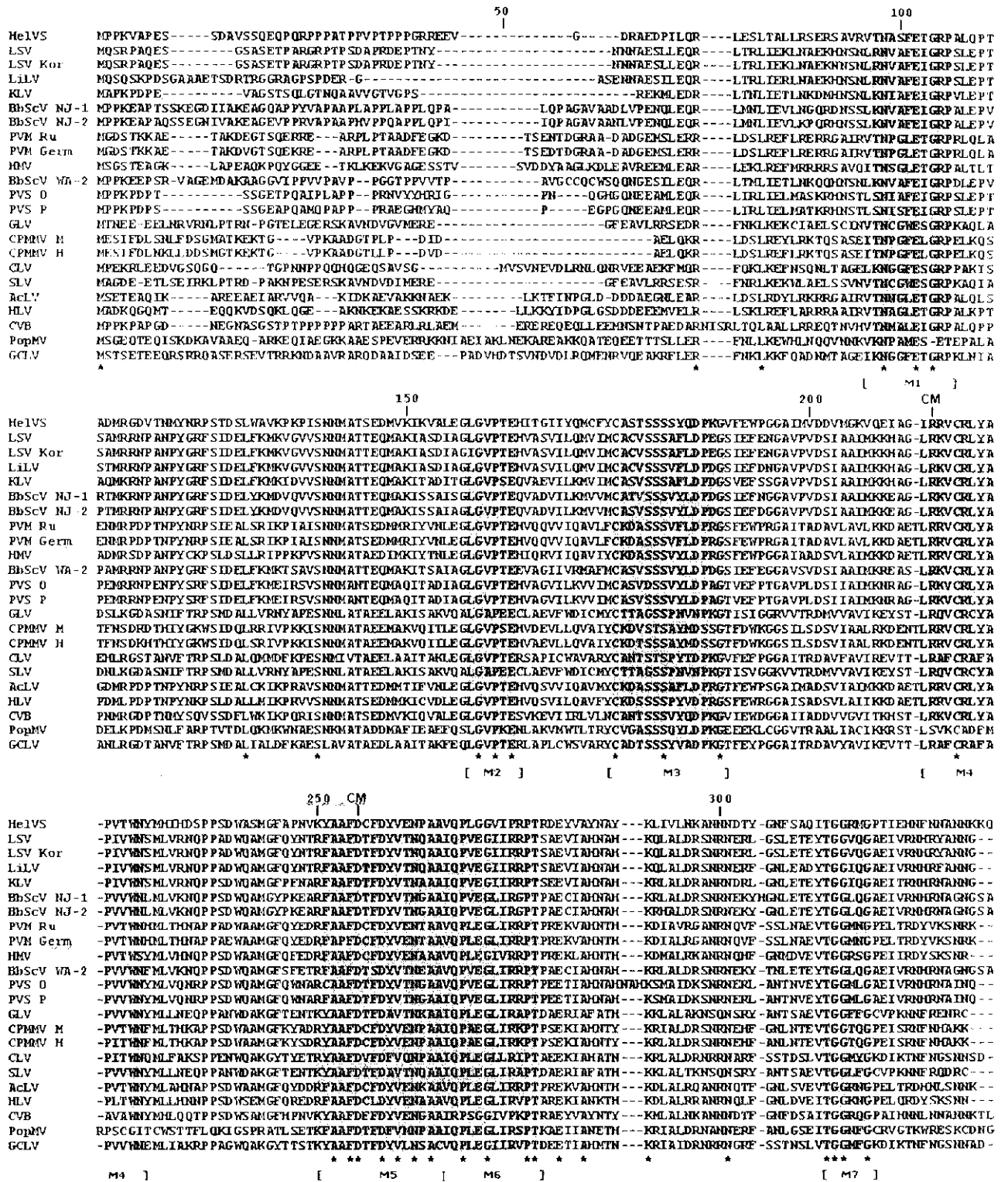


Рис. 1. Множественное выравнивание первичных структур капсидных белков карлавирусов методом CLUSTAL W [11]. Номера позиций указаны с учетом разрывов в последовательностях (-) и не совпадают с таковыми для индивидуальных последовательностей. Консервативные мотивы M1—M7 выделены серым цветом и их границы обозначены квадратными скобками. Позиции, содержащие идентичные аминокислотные остатки, отмечены звездочкой. CM — предполагаемый солевой мостик между Arg<sup>215</sup> и Asp<sup>255</sup>.

установлено, что большинство вирусспецифических эпитопов у представителей палочко- и нитевидных вирусов растений локализованы именно в N-концевой области белка [30—35]. У многих представителей групп *Potyvirus* и *Tobamovirus*, как показано ранее, N- и C-концевые фрагменты капсидных белков экспонированы на поверхности белковой глобулы [8, 36]. Для вируса табачной мозаики, например, с помощью PCA было установлено, что выступающая подвижная часть C-концевого фрагмента значительно длиннее, чем N-концевого [8]. Методом тритиевой планиграфии показано, что C-концевая область капсидного белка X-вируса картофеля (*Potexvirus*) погружена внутрь белковой глобулы, тогда как N-концевая часть экспонирована на ее поверхности [37].

Центральная и C-концевая области капсидных белков карлавирусов содержат консервативные мотивы различной длины, которые либо отделены друг от друга варибельными фрагментами, либо следуют один за другим (рис. 1). Высокая консервативность этих участков может играть важную роль при формировании третичной структуры белковой субъединицы. C-концевой участок насыщен остатками Gly, Asp, Glu, Asn, His, Ser (рис. 1), значительно ограничивающими гидролиз белка клеточными карбоксипептидазами. Следует отметить еще одну особенность, характерную для большинства вирусов со спиральным типом симметрии. На основании анализа пространственной структуры капсидного белка TMV и консервативных мотивов капсидных белков палочко- и нитевидных вирусов растений авторы работы [38] предположили, что консервативные остатки гидрофобного кода Arg<sup>215</sup> и Asp<sup>255</sup> способны образовывать солевой мостик (СМ, рис. 1), который может играть ключевую роль в формировании структуры коровой области белка. В последовательности PopMV вместо Arg<sup>215</sup> находится Ser<sup>215</sup>, однако в позиции 211 находится Arg и можно предположить, что этот остаток принимает участие в образовании солевого мостика.

На основании результатов множественного выравнивания первичных структур капсидных белков 23 карлавирусов нами были условно выделены семь мотивов, шесть из которых характерны для коровой (M1—M6), а один мотив (M7) — для C-концевой области белков.

В соответствии с выбранными нами условными обозначениями (см. «Материалы и методы») выделенные мотивы могут быть представлены следующим образом (рис. 1):

M1 — <sup>96</sup>[VLI]-[TRKS]-N-X-[GAS]-[FLMW]-E-[TIS]-G(0,1)-[RE]-[PT]-<sup>106</sup> (G<sup>104</sup> отсутствует в PopMV);

M2 — <sup>158</sup>[LI]-G-[VA]-P-X-E<sup>163</sup>;

M3 — <sup>176</sup>C-[KTAV]-X-[AVTS]-[SGD]-[ST]-S-X-[YFNV]-X-[DN]-[PS]-X-G<sup>189</sup>;

M4 — <sup>214</sup>[LI]-R-[RKAQ]-[VF]-C-R-[LAC]-[YF]-A-[PA]-[VIL]-[TAV]-W-[NS]<sup>228</sup> (без учета PopMV);

M5 — <sup>250</sup>[RK]-[FYC]-A-[AP]-F-D-[TCFV]-[FLS]-D-[YFA]-V-X-N-X-A-[AC]<sup>265</sup>;

M6 — <sup>266</sup>[IV]-[QR]-P-[ALVS]-[EG]-G-[LIV]-[IVL]-[RP]-X-P-T<sup>278</sup>;

M7 — <sup>313</sup>T-G-G-X-X-G<sup>318</sup>.

Ранее рядом авторов был установлен консервативный мотив, характерный для капсидных белков потекс- и карлавирусов [RK]-[FYW]-A-[GAP]-F-D-X-F-X-X-[LV]-X-X-X-[GAST]-[GAST] [16, 21, 39]. В базах данных белковых последовательностей этот мотив рассматривается как специфический «характерный признак» («signature») капсидных белков потекс- и карлавирусов растений. Он соответствует позициям 250—265 выделенного нами мотива M5 (рис. 1), однако имеет значительные отличия в представлении последовательности мотива, в частности, мотив M5 содержит семь, тогда как мотив «signature» — всего лишь четыре идентичных для всех последовательностей остатков. Остальные мотивы содержат следующее количество идентичных для всех последовательностей остатков: M1—M3 — по три остатка, M4 — пять остатков, M6, M7 — по четыре остатка.

Мотивы M3 и M4 включают в себя консервативный остаток Cys в позициях 176 и 218 соответственно (рис. 1). Для PopMV мотив M4 не является характерным, однако Cys в позиции 218 сохраняется. Консервативность положения цистеина в этих мотивах позволяет предположить, что остатки Cys участвуют в образовании дисульфидных связей либо между собой, либо с другими остатками Cys и, таким образом, могут играть важную роль в формировании пространственной структуры коровой области белка.

Из вышеперечисленных мотивов обращает на себя внимание мотив M7, содержащий последовательность TGG, которая присутствует у всех капсидных белков карлавирусов с известной на данный момент последовательностью. Этот мотив содержит четыре идентичных остатка. В работе [40] было отмечено сходство C-концевых участков капсидных белков карлавирусов и ряда сериновых протеаз. Авторы предположили, что последовательность Thr-Gly-Gly капсидного белка карлавирусов соответствует мотиву Ser-Gly-Gly в пяти клеточных протеазах, в том числе трипсине, химотрипсине и эластазе. Однако впоследствии участие капсидного белка в качестве протеазы в процессинге одного из

геномных продуктов карлавирусов не нашло подтверждения [41] и функция TGG-мотива до сих пор остается неустановленной.

Анализ выведенных мотивов был проведен с использованием программы PROSITE, которая позволяет по заданным мотивам произвести сканирование всех белковых последовательностей, входящих в базы данных, на наличие подобных мотивов в других белках. В результате этого анализа установлено, что мотивы M3, M5 и M6 характерны для капсидных белков карлавирусов, тогда как мотивы M2 и M7 встречаются и в других белках (более 1000 соответствий). Мотив M1 является характерным для всех карлавирусов, однако обнаруживается и в двух других белках: АТФ-связывающий белок FtsE (*Streptococcus lactis*) и белок YFIY (*Bacillus subtilis*). Для мотива M4, который был выведен без учета последовательности PopMV, не найдено соответствий в других белковых последовательностях, кроме как в капсидных белках карлавирусов. Однако этот мотив, выведенный с учетом последовательности PopMV: {LI}-[RS]-X-[VFK]-C-[RA]-X-[YF]-[AM]-R(0,1)-[PA]-[VILS]-[TAVC]-[WG]-[NIS], был обнаружен не только в капсидных белках карлавирусов, но и в последовательности капсидного белка PVX, штамм НВ. Таким образом, в результате анализа выведенных мотивов показано, что три мотива (M3, M5 и M6) являются на данный момент специфичными, тогда как мотивы M1, M2, M4 и M7 — лишь характерными, но не специфичными для капсидных белков карлавирусов.

На следующем этапе анализа капсидных белков карлавирусов проведено предсказание вторичной структуры каждого белка методом SOPMA и множественное сравнение этих структур. Результаты этого анализа представлены на рис. 2, где сходные элементы вторичной структуры условно разделены на девять отдельных участков. Такое представление результатов предсказания позволило отметить особенности, свойственные капсидным белкам карлавирусов: 1) вторичная структура N-концевых областей анализируемых белков переменна, а коровых и C-концевых областей — достаточно консервативна; 2) все капсидные белки склонны к формированию ряда  $\alpha$ -спиралей различной протяженности, соединяющихся участками с нерегулярной структурой (coil), короткими  $\beta$ -складками в сочетании с  $\beta$ -поворотами; 3)  $\alpha$ -спиральные структуры представлены в четырех участках (участки 2, 4, 6, 8) (рис. 2).

Вторичная структура N-концевых участков (участок 1) неоднородна и в отдельных белках представлена нерегулярной структурой либо соче-

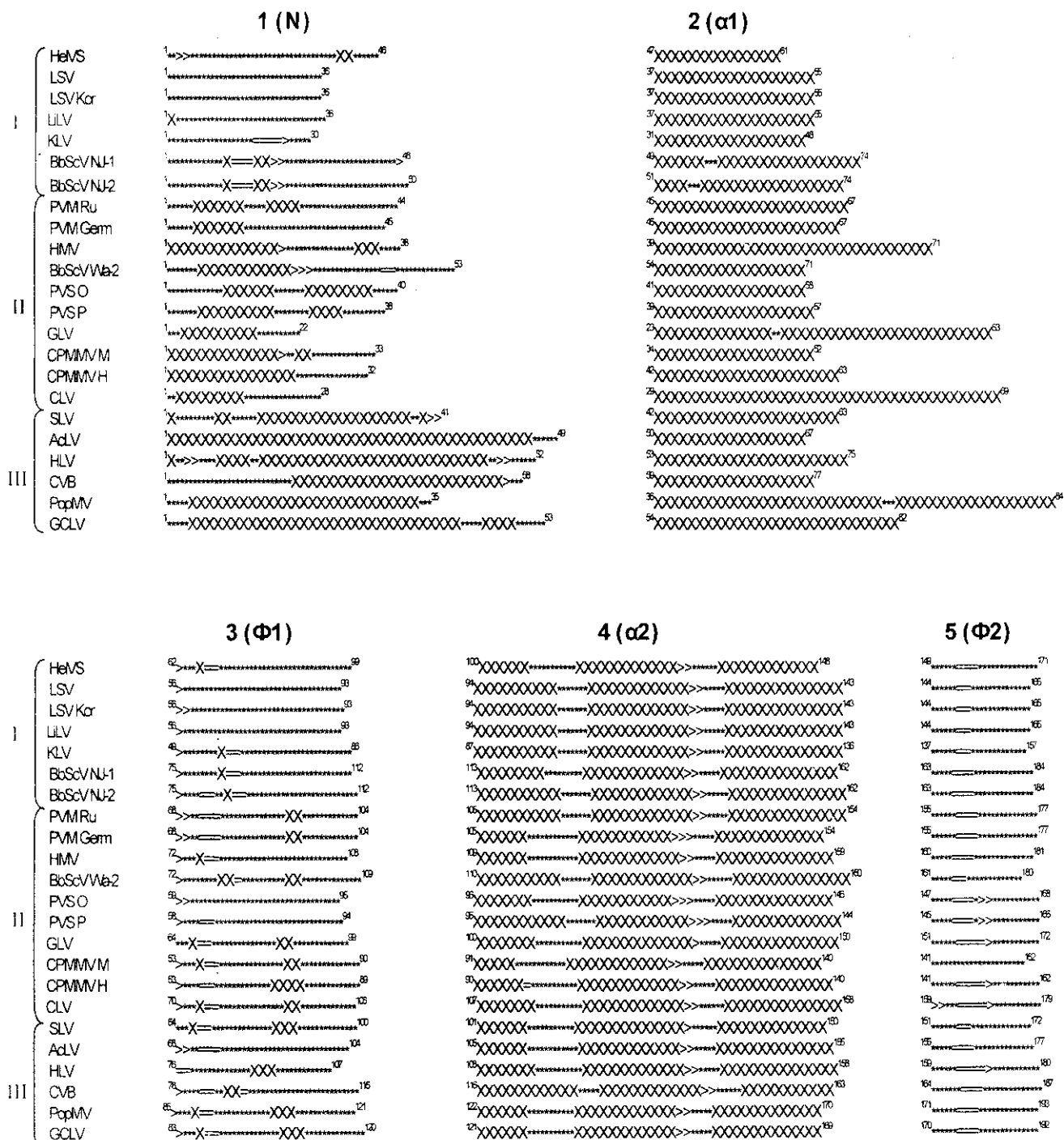
танием последней с короткими  $\alpha$ -спиралями, либо одной и более протяженными  $\alpha$ -спиралями (рис. 2). Исходя из различий в организации участка 1 капсидные белки карлавирусов можно условно разделить на три группы (рис. 2).

Для первой группы (I), к которой можно отнести HelVS, оба штамма LSV, LiLV, KLV и BbScV (штаммы NJ-1 и NJ-2), характерно преобладание нерегулярной структуры. Для N-концевых участков капсидных белков HelVS, KLV и BbScV (штаммы NJ-1 и NJ-2) можно отметить наличие коротких элементов  $\beta$ -складчатой структуры (рис. 2). Во вторую группу (II) входят оба штамма PVM, HNV, BbScV, штамм Wa-2, оба штамма PVS, GLV, оба штамма CPMMV и CLV. Для белков этой группы характерна склонность к формированию одной или двух коротких  $\alpha$ -спиралей в N-концевых участках (рис. 2). Третья группа (III) включает в себя капсидные белки SLV, AcLV, HLV, CVB, PopMV, GCLV, N-концевая часть которых склонна к формированию протяженных  $\alpha$ -спиралей, сочетающихся с разными по длине участками нерегулярной структуры (рис. 2).

Участок 2 представлен  $\alpha$ -спиральной структурой ( $\alpha_1$ ) с разной протяженностью для каждого белка. Для участка 3 характерно преобладание нерегулярной структуры (фланкирующий участок, Ф1). Участок 4 ( $\alpha_2$ ) содержит три  $\alpha$ -спирали, первая из которых ( $\alpha_{2_1}$ ) в большинстве случаев короче двух других. Первая и вторая ( $\alpha_{2_2}$ )  $\alpha$ -спирали соединены участком с нерегулярной структурой, а вторая и третья ( $\alpha_{2_3}$ )  $\alpha$ -спирали связаны между собой участком, образующим  $\beta$ -поворот, переходящий в нерегулярную структуру (рис. 2). Спираль  $\alpha_{2_3}$  включает в себя консервативный остаток Cys<sup>176</sup>, входящий в мотив M3 (рис. 1). Участок 5 (Ф2) представлен нерегулярной структурой, в которой для большинства белков имеют место короткие фрагменты (2—4 остатка)  $\beta$ -складчатого листа (рис. 2). Участок 6 представляет собой протяженную  $\alpha$ -спираль ( $\alpha_3$ ) с практически одинаковой длиной у всех белков. В случае LSV (оба штамма), PVM Germ, CPMMV M и CLV  $\alpha$ -спираль, возможно, имеет небольшую деформацию. Спираль  $\alpha_3$  содержит консервативный остаток Cys<sup>218</sup>, входящий в мотив M4 (рис. 1). В участке 7 (Ф3) представлены все типы вторичной структуры. Для большинства капсидных белков этот участок имеет сходную организацию (рис. 2). N-концевая часть его содержит  $\beta$ -складчатый мотив (2—4 остатка), переходящий во фрагмент с неупорядоченной структурой (10—18 остатков), далее следует  $\beta$ -поворот и короткий  $\beta$ -складчатый фрагмент. Затем

идет короткая  $\alpha$ -спираль и  $\beta$ -поворот в окружении нерегулярной структуры. Исключение составляют капсидный белок PVS O и PopMV, для которых не был предсказан первый из перечисленных  $\beta$ -поворотов, а в случае центральной части этого участка

для PVS O предсказана  $\alpha$ -спираль с небольшими искажениями в ее структуре. Участок 8 ( $\alpha 4$ ) представлен двумя  $\alpha$ -спиралями ( $\alpha 4_1$  и  $\alpha 4_2$ ), соединенными между собой  $\beta$ -складчатой и нерегулярной структурами. В участке 9 (C-концевой участок)



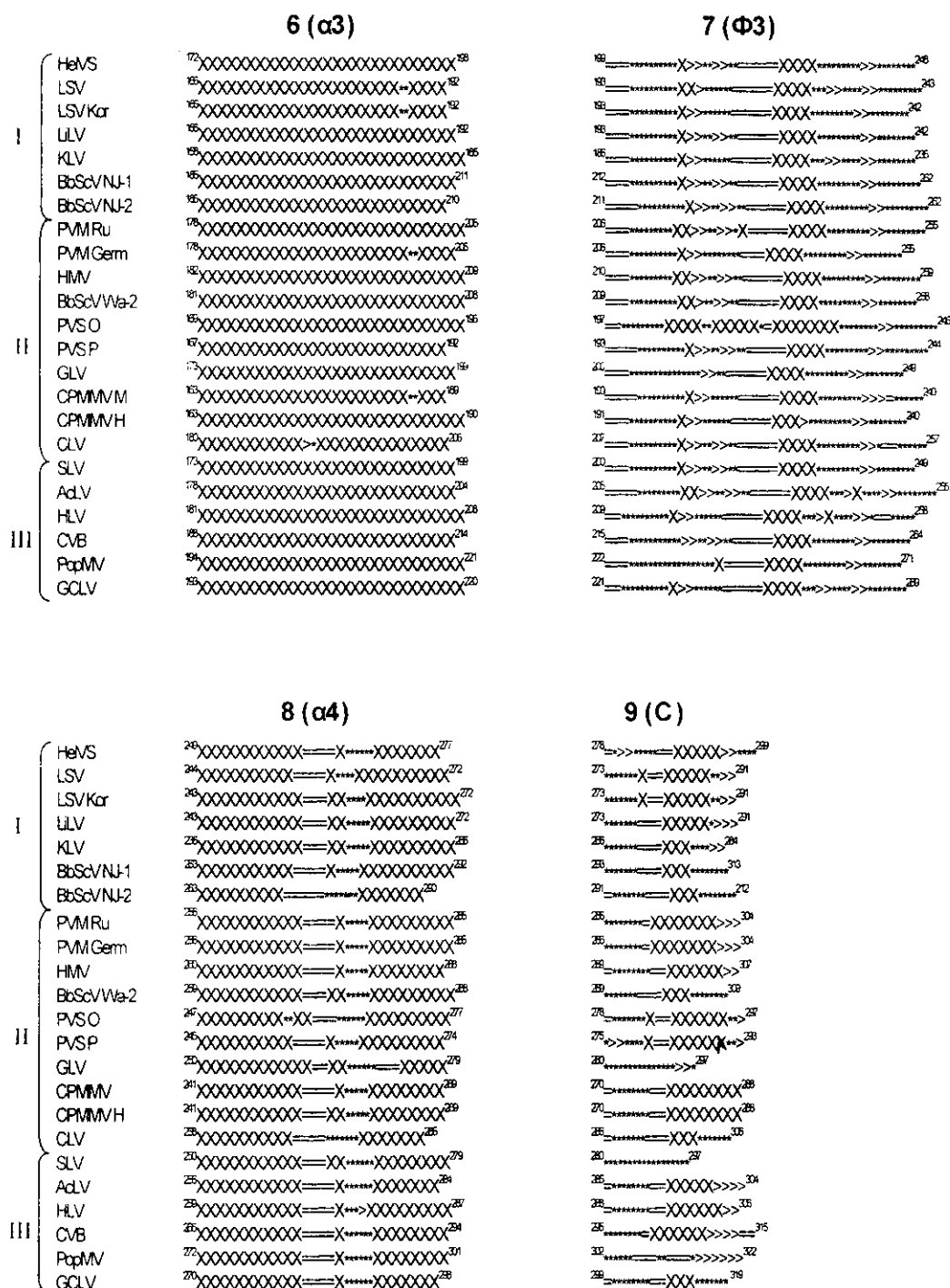


Рис. 2. Множественное сравнение вторичных структур капсидных белков карлавирусов методом SOPMA [13, 28]. Обозначения конформационных состояний: X —  $\alpha$ -спираль,  $\leftarrow$  —  $\beta$ -складка;  $\leftarrow$  —  $\beta$ -поворот, «\*» — нерегулярная структура (coil). Условные группы I, II, III объединяют белки со сходной организацией вторичной структуры N-концевого участка. Номера позиций соответствуют таковым в индивидуальных последовательностях



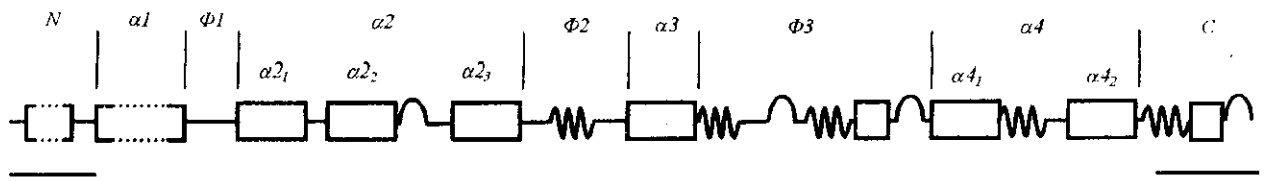


Рис. 3. Общая схема организации вторичной структуры капсидных белков карлавирусов:  $\alpha$ -спирали обозначены прямоугольниками,  $\beta$ -складки — зигзагообразными линиями,  $\beta$ -повороты — дугами, участки нерегулярной структуры — прямыми линиями. Обозначения блоков: N, C-N и C — концевые блоки;  $\alpha 1$ ,  $\alpha 2$ ,  $\alpha 3$ ,  $\alpha 4$  — блоки, содержащие  $\alpha$ -спирали;  $\Phi 1$ ,  $\Phi 2$ ,  $\Phi 3$  — фланкирующие блоки. Пунктирные линии в прямоугольниках обозначают различия в протяженности  $\alpha$ -спиралей для индивидуальных белков. Прямые линии под основной схемой указывают на наличие нерегулярной структуры в N- и C-концевых блоках

представлены все типы вторичной структуры. Для капсидных белков HeIVS, LSV (оба штамма), LiLV, PVM (оба штамма), HNV, PVS (оба штамма), CPMMV (оба штамма), AcLV, HLV и CVB характерно наличие в этом участке короткой  $\alpha$ -спирали. С-концевые остатки (2—6 остатков) ряда капсидных белков карлавирусов — LSV (оба штамма), KLV, PVM (оба штамма), HNV, AcLV, HLV, LiLV, CVB, PopMV и GLV — склонны к образованию  $\beta$ -поворота (рис. 2). N-концевая часть этого участка включает в себя TGG-мотив (мотив M7, рис. 1), для которого характерна нерегулярная вторичная структура.

На основании результатов множественного сравнения вторичной структуры нами предложена общая схема организации вторичной структуры капсидных белков карлавирусов (рис. 3). Эта схема включает в себя N- и C-концевые блоки, четыре  $\alpha$ -спиральных блока ( $\alpha 1$ ,  $\alpha 2$ ,  $\alpha 3$  и  $\alpha 4$ ), соединенных между собой фланкирующими блоками  $\Phi 1$ ,  $\Phi 2$  и  $\Phi 3$ . Следует отметить, что для индивидуальных белков длина N-концевого блока и блока  $\alpha 1$  варьирует, а N-концевому блоку свойственны также различия во вторичной структуре (рис. 2, 3). Остальные блоки, включая C-концевой, характеризуются незначительными вариациями в размерах и вторичной структуре в пределах каждого блока. Ранее в ряде работ [38, 42] отмечена высокая консервативность основных элементов вторичной структуры между представителями палочко- и нитевидных вирусов растений (тобамо-, поти-, потекс- и фурувирусы) при низкой гомологии первичных структур. Авторами высказано предположение о том, что для сохранения функции белка существенным является сохранение основных элементов именно пространственной структуры.

В нашей работе это предположение также нашло свое подтверждение. Первичная структура капсидного белка PopMV имеет низкую гомологию с другими представителями карлагруппы, однако

сохраняет все основные элементы вторичной структуры, характерные для капсидных белков карлавирусов (рис. 2).

Таким образом, на основании результатов анализа первичной и вторичной структуры белков оболочки карлавирусов можно сделать следующие выводы: 1) первичная и вторичная структуры N-концевых участков капсидных белков характеризуются значительной вариабельностью, тогда как коровым и C-концевым областям свойственна высокая степень консервативности; 2) установлены три консервативных мотива (M3, M5 и M6), специфичные для капсидных белков карлавирусов; 3) вторичная структура капсидных белков содержит четыре блока  $\alpha$ -спиралей, соединенных между собой тремя фланкирующими блоками, в которых представлены все типы вторичной структуры; 4) для каждого фланкирующего блока характерна практически строгая последовательность элементов вторичной структуры; 5) несмотря на высокую вариабельность первичной структуры между отдельными белками, вторичная структура остается консервативной для всех капсидных белков карлагруппы.

Ув. читатели!

В статье S. S. Viter, T. Yu. Tkachenko, L. P. Kolomietz, Yu. L. Radavsky «Analysis of antigenic structure of Potato Virus M Ukrainian strains», опубликованной в журнале «Биополимеры и клетка» (2000, Т. 16, № 4, С. 312—319), допущена неточность на стр. 316, вторая колонка. Предложение «The change of lysine side chain positive charge to negative one may result in mimicry of the side chain negative charge of the aspartic acid residue (position 36), which is absent in the <sup>22</sup>Glu/Gly-Arg<sup>44</sup> peptides» следует читать как: «The change of lysine side chain positive charge to negative one may result in mimicry of the side chain negative charge of the aspartic acid residue (position 36), which is absent in the <sup>22</sup>Glu/Gly-Lys<sup>35</sup> peptides». Авторы приносят свои извинения за допущенную неточность.

T. Yu. Tkachenko, S. S. Viter, Yu. L. Radavsky

Comparative analysis of primary and secondary structures of carlaviruses' capsid proteins

#### Summary

The primary and secondary structures of 23 carlaviruses' capsid proteins have been analysed. As a result of the primary structure analysis by multiple alignment seven motifs have been established, three of them being specific for the carlaviruses' capsid proteins. The secondary structure analysis by means of multiple comparison has shown a high degree of secondary structure conservatism in the core and C-terminal regions, and considerable variability in the N-terminus. A general scheme for the secondary structure of the carlaviruses' capsid proteins has been proposed. on the base of multiple comparison results.

Т. Ю. Ткаченко, С. С. Вітер, Ю. Л. Радавський

Порівняльний аналіз первинної та вторинної структури капсидних білків карлавірусів рослин

#### Резюме

Проаналізовано первинну та вторинну структури капсидних білків 23 карлавірусів рослин. У результаті аналізу первинної структури методом множинного вирівнювання виділено сім мотивів, три з яких є специфічними для капсидних білків карлавірусів. Аналіз вторинної структури цих білків на основі множинного порівняння виявив високий ступінь консервативності цієї структури у коровій і С-кінцевій областях, та її значну варіабельність в N-кінцевій області. Запропоновано загальну схему вторинної структури, яка характерна для капсидних білків карлавірусів рослин.

#### СПИСОК ЛИТЕРАТУРЫ

1. Wetter C., Milne R. G. Carlaviruses // Handbook of plant virus infections and comparative diagnosis / Ed. E. Kurstak.— Amsterdam: Elsevier, 1981.—P. 695—730.
2. Zavriev S. K., Kanyuka K. V., Levay K. E. The genome organization of potato virus M RNA // J. Gen. Virol.—1991.—72.—P. 9—14.
3. Cavaleer T. D., Halpern B. T., Lawrence D. M., Podleckis E. V., Martin R. R., Hillman B. I. Nucleotide sequence of the carlavirus associated with blueberry scorch and similar diseases // J. Gen. Virol.—1994.—75.—P. 711—720.
4. Hataya T., Uchino K., Arimoto R., Suda N., Sano T., Shikata E., Uyeda I. Molecular characterization of Hop latent virus and phylogenetic relationships among viruses closely related to carlaviruses // Arch. Virol.—2000.—145.—P. 2503—2524.
5. Fuji S., Yamamoto H., Inoue M., Yamashita K., Fukui Y., Furuya H., Naito H. Complete nucleotide sequence of the genomic RNA of Aconitum latent virus (genus Carlavirus) isolated from *Delphinium* sp // Arch. Virol.—2002.—147.—P. 865—870.
6. Chen J., Chen J., Adams M. J. Molecular characterization of a complex mixture of viruses in garlic with mosaic symptoms in China // Arch. Virol.—2001.—146.—P. 1841—1853.
7. Foster G. D. The structure and expression of the genome of carlaviruses // Res. Virol.—1992.—143.—P. 103—112.
8. Bhyravhatla B., Watowich S., Caspar D. Refined atomic model of the four-layer aggregate of the Tobacco Mosaic Virus coat protein at 2.4-Å resolution // Biophys. J.—1998.—74.—P. 604—615.
9. Wang H., Planchart A., Allen D., Pattanayek R., Stubbs G. Preliminary X-ray diffraction studies of ribgrass mosaic virus // J. Mol. Biol.—1993.—234.—P. 902—904.
10. Wang H., Stubbs G. Structure determination of cucumber green mottle mosaic virus by X-ray fiber diffraction. Significance for the evolution of tobamoviruses // J. Mol. Biol.—1994.—239.—P. 371—384.
11. Thompson J. D., Higgins D. G., Gibson T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice // Nucl. Acids Res.—1994.—22.—P. 4673—4680.
12. Gattiker A., Gasteiger E., Bairoch A. ScanProsite: a reference implementation of a PROSITE scanning tool // Appl. Bioinform.—2002.—1.—P. 107—108.
13. Geourjon C., Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments // Comput. Appl. Biosci.—1995.—11.—P. 681—684.
14. Gramstat A., Courtzopoulos A., Rohde W. The 12 kDa protein of potato virus M displays properties of a nucleic acid-binding regulatory protein // FEBS Lett.—1990.—276.—P. 34—38.
15. Foster G. D., Mills P. R. The 3'-nucleotide sequence of an ordinary strain of potato virus S // Virus Genes.—1992.—6.—P. 213—220.
16. MacKenzie D. J., Tremaine J. H., Stace-Smith R. Organization and intervirial homologies of the 3'-terminal portion of potato virus S RNA // J. Gen. Virol.—1989.—70.—P. 1053—1063.
17. Foster G. D., Millar A. W., Meehan B. M., Mills P. R. Nucleotide sequence of the 3'-terminal region of Helenium virus S RNA // J. Gen. Virol.—1990.—71.—P. 1877—1880.
18. Memelink J., van der Vlugt C. I., Linthorst H. J., Derks A. F., Asjes C. J., Bol J. F. Homologies between the genomes of a carlavirus (lily symptomless virus) and a potexvirus (lily virus X) from lily plants // J. Gen. Virol.—1990.—71.—P. 917—924.
19. Ahn H. I., Ryu J. H., Kim J. K., Lee S. Y., Choi J. K., Park W. M., Ryu K. H. Nucleotide sequence analysis of the 3'-terminal region of two Korean isolates of lily symptomless Carlavirus and expression of the coat protein in *E. coli* // Mol. Cells.—1999.—9.—P. 338—343.
20. Meehan B. M., Mills P. R. Nucleotide sequence of the 3'-terminal region of carnation latent virus // Intervirology.—1991.—32.—P. 262—267.
21. Henderson J., Gibbs M. J., Edwards M. L., Clarke V. A., Gardner K. A., Cooper J. I. Partial nucleotide sequence of poplar mosaic virus RNA confirms its classification as a carlavirus // J. Gen. Virol.—1992.—73.—P. 1887—1890.
22. Hataya T., Arimoto R., Suda N., Uyeda I. Molecular characterization of Hop mosaic virus: its serological and molecular relationships to Hop latent virus // Arch. Virol.—2001.—146.—P. 1935—1948.
23. Nicolaisen M., Nielsen S. L. Analysis of the triple gene block and coat protein sequences of two strains of Kalanchoe latent carlavirus // Virus Genes.—2001.—22.—P. 265—270.
24. Tsuneyoshi T., Matsumi T., Deng T. C., Sako I., Sumi S. Differentiation of Allium carlaviruses isolated from different parts of the world based on the viral coat protein sequence // Arch. Virol.—1998.—143.—P. 1093—1107.
25. Levay K., Zavriev S. Nucleotide sequence and gene organization of the 3'-terminal region of chrysanthemum virus B genomic RNA // J. Gen. Virol.—1991.—72.—P. 2333—2337.
26. Ryu J. H., Park H. W., Park W. M., Lee S. Y., Ryu K. H. Molecular analysis of the 3'-terminal region of lily latent Carlavirus from *Lilium lancifolium* // J. Plant Pathol.—2000.—16.—P. 231—235.
27. Naidu R. A., Gowda S., Satyanarayana T., Boyko V., Reddy A. S., Dawson W. O., Reddy D. V. Evidence that whitefly-

- transmitted cowpea mild mottle virus belongs to the genus *Carlavirus* // Arch. Virol.—1998.—143.—P. 769—780.
28. Combet C., Blanchet C., Geourjon C., Deleage G. NPS@: network protein sequence analysis // Trends Biochem. Sci.—2000.—25.—P. 147—150.
  29. Adams A. N., Barbara D. J., Adams A. N., Barbara D. J. The use of F(abB)<sub>2</sub>-based ELISA to detect serological relationships among carlaviruses // Ann. Appl. Biol.—1982.—101.—P. 495—500.
  30. Hewish D. R., Xiao X. W., Mishra A., Gough K. H., Shukla D. D. Characterisation and epitope analysis of monoclonal antibodies to virions of clover yellow vein and Johnson grass mosaic potyviruses // Arch. Virol.—1993.—133.—P. 127—141.
  31. Shukla D. D., Strike P. M., Tracy S. L., Gough K. H., Ward C. W. The N and C termini of the coat proteins of potyviruses are surface-located and the N terminus contains the major virus specific epitopes // J. Gen. Virol.—1988.—69.—P. 1497—1508.
  32. Commandeur U., Koenig R., Lesemann D.-E., Torrance L., Burgermeister W., Liu Y., Schots A., Alric M., Grassi G. Epitope mapping on fragments of beet necrotic yellow vein virus coat protein // J. Gen. Virol.—1992.—73.—P. 695—700.
  33. Радавский Ю. Л., Витер С. С., Турова И. П., Зайцева Л. С., Ярвекюль Л. В., Саарма М. Ю., Гребенищikov Н. И., Баратова Л. А. Антигенная структура белка оболочки X-вируса картофеля. II. Локализация антигенной(ных) детерминанты в N-концевом участке белка // Биоорг. химия.—1989.—15.—С. 615—619.
  34. Cerovska N., Filigarova M., Subr Z. Partial antigenic characterization of potato virus S (Andean strain) by monoclonal antibodies // Acta Virol.—1996.—40.—P. 23—26.
  35. Viter S. S., Tkachenko T. Yu., Kolomietz L. P., Radavsky Yu. L. Analysis of antigenic structure of Potato Virus M Ukrainian strains // Биополимеры и клетка.—2000.—16.—P. 312—319.
  36. Bloomer A. C., Champness J. N., Bricogne G., Staden R., Klug A. Protein disc of tobacco mosaic virus at 2.8 Å resolution showing the interaction between subunits // Nature (London).—1978.—276.—P. 362—368.
  37. Baratova L. A., Grebenshchikov N. I., Shishkov A. V., Kashirin I. A., Radavsky Yu. L., Jarvekulg L., Saarma M. The topography of the surface of potato virus X: tritium planigraphy and immunological analysis // J. Gen. Virol.—1992.—73.—P. 229—235.
  38. Dolja V. V., Boyko V. P., Agranovsky A. A., Koonin E. V. Phylogeny of capsid proteins of rod-shaped and filamentous RNA plant viruses: two families with distinct patterns of sequence and probably structure conservation // Virology. 1991.—184.—P. 79—86.
  39. Abouhaidar M. G., Lai R. Nucleotide sequence of the 3'-terminal region of clover yellow mosaic virus RNA // J. Gen. Virol.—1989.—70.—P. 1871—1875.
  40. Foster G. D., Mills P. R. Translation of potato virus S RNA *in vitro*: evidence of protein processing // Virus Genes.—1991.—6.—P. 45—52.
  41. Lawrence D. M., Rozanov M. N., Hillman B. I. Autocatalytic processing of the 223-kDa protein of blueberry scorch carlavirus by a papain-like proteinase // Virology.—1995.—207.—P. 127—135.
  42. Pereira L. G., Torrance L., Roberts I. M., Harrison B. D. Antigenic structure of the coat protein of potato mop-top furovirus // Virology.—1994.—203.—P. 277—285.

УДК 577.112.5.083.3  
Надійшла до редакції 25.01.02