

20. Кордюм В. А. Эволюция вирусов — попытка нелинейного прогноза. *Биополимеры и клетка*. 2001. Т. 17, № 6. С. 467–486. DOI: <https://doi.org/10.7124/bc.0005D6>.
21. Lukashov V., Goudsmit J. Evolutionary relationships among Astroviridae. *Journal of General Virology*. 2002. Vol. 83. P. 1397–1405. DOI: <https://doi.org/10.1099/0022-1317-83-6-1397>.

CHARACTERIZATION OF SWINE HEPATITIS E VIRUS GENOTYPE 3 OPEN READING FRAMES

Lymanska O. Yu.

National Scientific Center “Institute of Experimental and Clinical Veterinary Medicine”, Kharkiv, Ukraine

The goal of this study was a determination of structural organization features of swine hepatitis E virus genotype 3 open reading frames. Gene selective pressure for genes was estimated using the Tajima's neutrality test as implemented in MEGA 6. Shannon entropy was used for diversity detection at each position of open reading frames. Shannon entropy plots representing quantitative diversity at each amino acid (nucleotide) position of 11 swine HEV-3 isolates aligned sequences were calculated by software BioEdit. For phylogenetic analysis, nucleotide distances were analyzed using the neighbour-joining algorithm. Analysis of nucleotide, non-synonymous and synonymous distances was performed using MEGA6. Parameters of mutational analysis for swine hepatitis E virus genotype 3 open reading frames ORF1 – ORF3 were determined using Tajima's neutrality test. Shannon entropy technique revealed that ORF3 is characterized by the most number of diverse amino acid residues (21,7 %) comparing with ORF1, ORF2 coding regions. Analysis of non-synonymous distances in ORF3 allowed to determine an order of evolutionary events in the history of swine HEV-3 isolates. Avian astroviruses strains were used as control because phylogenetic relationships among them correspond to the evolution history of their host species. Tree topology for swine HEV-3 ORF3 based on non-synonymous distances differs from tree structures based on nucleotide and synonymous distances. Positive values of *D* Tajima's factor for ORF1, ORF2 of swine HEV-3 show on operating positive selection onto these coding regions. Negative value of *D* Tajima's factor for ORF3 coding region indicates onto purifying selection

Keywords: positive selection, Tajima's neutrality test, negative selection, Shannon entropy

УДК 619:616.98:578.828:578.2'21:636.22/.28

DOI 10.36016/VM-2023-109-5

СТРУКТУРНИЙ АНАЛІЗ ПРОТЕЇНІВ ВІДКРИТИХ РАМОК ЗЧИТУВАННЯ ВІРУСУ ІМУНОДЕФІЦИТУ ВЕЛИКОЇ РОГАТОЇ ХУДОБИ

Балак О. К.

Харківський національний медичний університет, Харків, Україна

Лиманська О. Ю.

Національний науковий центр «Інститут експериментальної і клінічної ветеринарної медицини», Харків, Україна, e-mail: olgaliman@ukr.net

Метою роботи було визначення особливостей структурної організації протеїнів ORF2 та ORF3 відкритих рамок зчитування вірусу імунодефіциту великої рогатої худоби (ВІ ВРХ). За допомогою програми ATGgr, яка дозволяє ефективно передбачення кодонів ініціації трансляції з точністю до нуклеотида, визначено ORFs для двох ізолятів ВІ ВРХ з повним геномом. Для передбачення, аналізу вторинної структури та функції протеїнів використовували програму Rhyre2. Для пошуку фрагментів протеїнів в неупорядкованому, або нативно-розгорнутому, стані використовували програму PONDR-FIT. Аналіз амінокислотного складу протеїнів ORF2 та ORF3 ізолятів ВІ ВРХ щодо наявності неполярних, полярних, ароматичних, гідрофобних амінокислотних залишків проведено за допомогою програми PSIPRED. Моделі просторової структури протеїнів отримано за допомогою сервера I-TASSER. Для протеїну ORF3 передбачено 14 % α -спіралей, 17 % β -тяжів, 43 % неупорядкованої структури. Для поліпротеїну Gag, який транскрибується з відкритої рамки зчитування ORF2, визначено 37 % α -спіралей, 0 % β -тяжів, 41 % неупорядкованої структури. Розподіл заряджених амінокислотних залишків характеризує поверхневі властивості протеїнів. Для протеїну ORF2 їхня кількість сягає 23,9 %. Кількість Arg становить 5,2 %, Lys — 8,0 %, Glu — 7,3 %, Asp — 3,4 %. Загальна кількість заряджених амінокислотних залишків ORF3 становить 23,3 %. Кількість Arg становить 12,6 %, Lys —

4,9 %, Glu — 1,9 %, Asp — 3,9 %. З п'яти ORFs для двох ізолятів ВІ ВРХ тільки дві ORFs збігаються за довжиною нуклеотидів (та, отже, відповідних протеїнів). Протеїн ORF3 відноситься до внутрішньо неупорядкованих протеїнів, які не можуть бути стабільно складеними в унікальній тривимірній структурі за фізіологічних умов, а поліпротеїн Gag, який трансклюється з ORF2, — до класу повністю структурованих протеїнів. Вторинна структура обох протеїнів демонструє наявність α спіралей

Ключові слова: поліпротеїн Gag, Lentivirus, Retroviridae

Вірус імунодефіциту великої рогатої худоби (ВІ ВРХ, bovine immunodeficiency virus, BIV), один з представників родини лентівірусів, вперше був виділений у 1969 році від 8-річної корови з персистентним лімфоцитозом, який супроводжувався виснаженням, слабкістю та ураженням центральної нервової системи [1].

Патогенність ВІ ВРХ досі не підтверджено, і це питання залишається дискусійним. Проте зменшення ваги тварини, зниження молочної продуктивності, гіперплазія лімфатичних вузлів, параліч нижніх кінцівок були асоційовані саме з вірусом імунодефіциту ВРХ [2]. ВІ ВРХ може бути причиною імунних розладів та, як наслідок, опортуністичних бактеріальних інфекцій [3]. Застосування серологічних та молекулярно-генетичних методів дозволило виявити ВІВ-інфікованих тварин у багатьох країнах світу, причому рівень інфікованості коливався від 1,4 до 33 % [4–7]. ВІ ВРХ інфікує, в основному, моноцити, макрофаги та лімфоцити, але також був виявлений і в спермі експериментально інфікованих биків [8].

Повний геном ВІ ВРХ в перший [9] і в останній раз (номер L04974 в GenBank, 2006 р.) було секвеновано у США. Таким чином, з 2007 року по теперішній час не було секвеновано жодного повного геному ізоляту ВІ ВРХ, беручи до уваги щорічне зростання інформації в GenBank щодо повних геномів патогенів тварин.

Геном ВІ ВРХ, який належить до роду лентівірусів сімейства Retroviridae, представлений двома копіями одноланцюгової молекули РНК. Для ізоляту НХВ3 (номер NC_001413 в GenBank, який є ідентичним до M32690) довжина РНК становить 8 482 нуклеотиди (н.). Провірусна ДНК ВІ ВРХ довжиною 8 388 п. н. (ізолят R-29, номер L04974 в GenBank) містить гени *gag*, *pol*, *env*, фланковані довгими кінцевими повторами, а також, як вважають, п'ять коротких відкритих рамок зчитування (ORF), які розташовані у центральному регіоні між генами *pol*, *env*, перекриваються з відповідними ORFs та кодують неструктурні регуляторні гени [9]. Генетична та антигенна подібність ВІ ВРХ до вірусів хвороби Джембрана та імунодефіциту людини дозволяє розглядати даний вірус як модель для дослідження патогенезу та оцінки ефективності методів лікування та контролю патогенних лентівірусів [4].

Амінокислотна послідовність детермінує вторинну структуру протеїну, яка, в свою чергу, впливає на формування третинної (просторової) структури. Основними елементами вторинної структури протеїнів є α -спіралі, збагачені аланіном, лейцином, лізином, глутаміном, глутаміновою кислотою, метіоніном, аргініном, та β -тяжи, збагачені валіном, треоніном, фенілаланіном, тирозином [10]. Інформація стосовно вторинної структури протеїнів має важливе практичне значення, враховуючи, наприклад, що вторинна структура вакцинних антигенів — фрагментів повновимірних протеїнів — повинна бути комплементарною до такої у повновимірних молекулах певного збудника.

Первинна структура протеїну визначає також його внутрішню неупорядкованість. Встановлено, що до числа амінокислот, які сприяють неупорядкованості, належать, зокрема гліцин, серін, пролін. Частка неупорядкованих фрагментів в молекулі протеїну, які відіграють певну роль в регуляторних процесах, може бути різною — від послідовності довжиною від декількох амінокислот до повної неупорядкованості [11–13]. Вторинна структура ВІ ВРХ, яка характеризувалася наявністю чотирьох спіральних фрагментів, була прогнозована для протеїну, що утворювався внаслідок антисенсової транскрипції, підтвердженої для представників сімейства Retroviridae [14, 15].

Упакування молекули геномної РНК ретровірусів та її внутрішньомолекулярні взаємодії з протеїнами віріону є необхідними для інфекційності, але до теперішнього часу залишаються недостатньо визначеними конформації протеїнів ВІ ВРХ.

В даній роботі проаналізовано амінокислотний склад протеїнів відкритих рамок зчитування ORF2 та ORF3 двох ізолятів ВІ ВРХ з повним геномом щодо наявності гідрофобних, полярних,

неполярних, ароматичних амінокислотних залишків, побудовано моделі вторинної та тривимірної структур, а також проаналізовано взаємозв'язок між амінокислотою послідовністю та просторовою структурою протеїнів.

Матеріал і методи. Нуклеотидні послідовності ізолятів ВІ ВРХ з повним геномом (таксономічний ідентифікатор (txid) 11657), вірусу гепатиту Е (ВГЕ) генотипу 3 свині та кабана (txid 1678143) отримано через пошук в базі даних GenBank Національного центру біотехнологічної інформації (США).

Програму BioEdit (версія 7.2.5) [16] використано для маніпулювання з нуклеотидними та амінокислотними послідовностями.

За допомогою програми ATGpr (<http://www.atgpr.dbcls.jp>), яка дозволяє ефективно передбачення кодонів ініціації трансляції в послідовностях кДНК та з точністю до нуклеотида визначати відкриті рамки зчитування (ORFs), для двох ізолятів вірусу імунодефіциту ВРХ з повним геномом (NC_001413, L04974) визначено по п'ять ORFs (табл. 1).

Таблиця 1 — Визначені відкриті рамки зчитування (ORFs) для двох ізолятів вірусу імунодефіциту ВРХ з повним геномом. Підкреслено поліпротеїн Gag (довжиною 476 амінокислотних залишків (а. з.), який транскрибується з ORF2, та протеїн (довжиною 103 а. з.), який транскрибується з ORF3.

Ізолят ВІ ВРХ (номер в GenBank)	Положення на геномі	Розмір протеїну, а. з.	Імовірність існування
NC_001413	5415-8126	904	0,34
	316-1743	<u>476</u>	0,33
	5228-5536	<u>103</u>	0,27
	6063-8126	688	0,24
	3816-4739	308	0,21
L04974	5792-8034	761	0,35
	5413-5700	96	0,32
	314-1741	<u>476</u>	0,31
	5226-5531	<u>102</u>	0,27
	1579-4104	842	0,20

Аналіз амінокислотного складу ORF2–ORF3 ізолятів ВІ ВРХ щодо наявності неполярних, полярних, ароматичних, гідрофобних амінокислотних залишків проведено за допомогою програми PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred>).

Третинну структуру, що передбачено, протеїнів ORF2 та ORF3 побудовано за допомогою сервера I-TASSER (Iterative Threading Assembly Refinement, <http://zhanglab.ccmb.med.umich.edu/I-TASSER>) за стандартних параметрів [17].

Для пошуку ослаблених ділянок в протеїнах побудовано графіки внутрішнього безладу для протеїнів ORF3 та ORF2 довжиною 103 та 476 а. з. за допомогою програми PONDR-FIT (<http://original.disprot.org/pondr-fit.php>), яка дозволяє передбачати нативно-розгорнуті ділянки поліпептидного ланцюга. В цій програмі для складання передбачення використовуються набори протеїнів з бази протеїнових структур PDB, в яких методами рентгенівського розсіювання, ЯМР, кругового дихроїзму визначено неструктуровані ділянки протеїнів. Також окремо оцінюється вплив близького розташування N- та C-кінців поліпептидного ланцюга.

Для передбачення, аналізу вторинної структури та функції протеїнів використовували програму Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2>) [18].

Результати досліджень. Як контроль точності визначення відкритих рамок зчитування використано ВГЕ генотипу 3 свині та кабана. Кількість та позиції теоретично визначених нами ORFs та відомих з літератури експериментально визначених ORFs збігаються, що свідчить про надійність програми ATGpr, яку застосували.

Для двох відкритих рамок зчитування, які збігаються для двох ізолятів ВІ ВРХ з повним геномом (NC_001413, L04974), довжиною 103 а. з. (рис. 1) та 476 а. з., проведено біоінформатичний аналіз амінокислотного складу щодо наявності гідрофобних, полярних, неполярних, ароматичних а. з. (табл. 2).

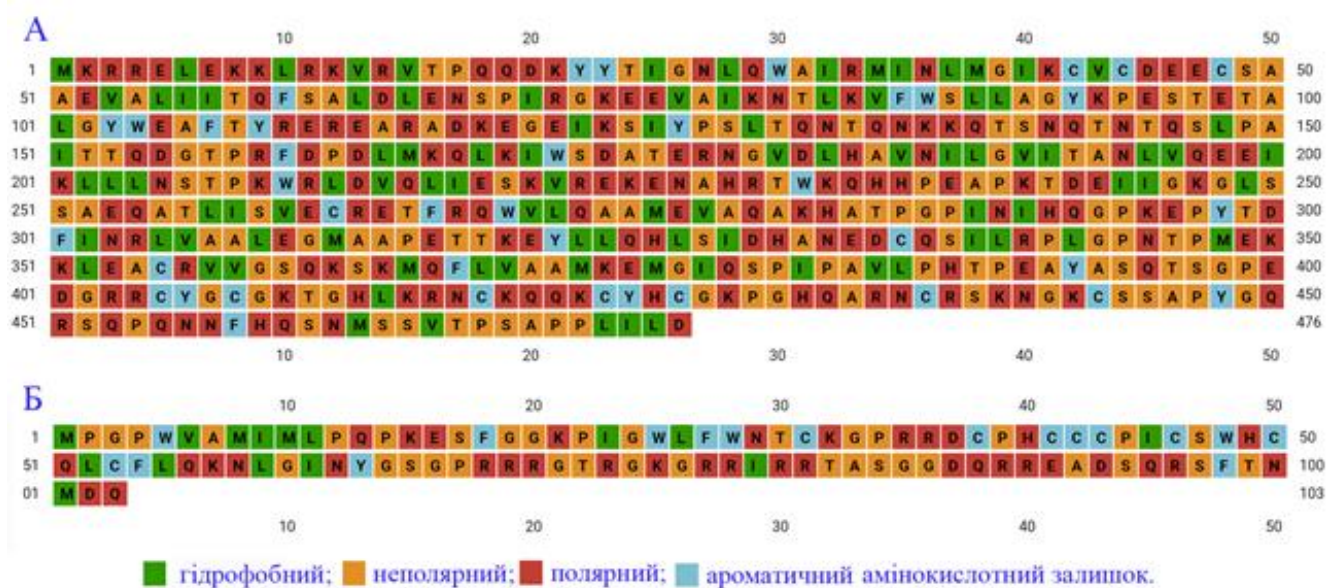


Рис. 1. Аналіз амінокислотного складу протеїнів ORF2 протеїну Gag (A) та ORF3 (Б) ізоляту NC_001413 ВІ ВРХ за допомогою програми PSIPRED. Представлено неполярні, полярні, ароматичні, гідрофобні амінокислотні залишки. R — Arg, K — Lys, E — Glu, D — Asp, C — Cys.

Таблиця 2 — Характеристика первинної структури протеїнів, що транслюються з відкритих рамок зчитування ORF3 та ORF2, довжиною 103/102 та 476 а.з. ізолятів вірусу імунodefіциту ВРХ з повним геномом

Відкрита рамка зчитування/протеїн	Амінокислотні залишки			
	полярні, %	неполярні, %	гідрофобні, %	ароматичні, %
ORF3 NC_001413, 103 а. з.	34,95	33,98	14,56	16,51
ORF3 L04974, 102 а. з.	36,89	31,07	15,53	16,51
ORF2 NC_001413, 476 а. з.	38,66	32,35	20,59	8,40
ORF2 L04974, 476 а.а.	38,66	32,14	20,80	8,40

Для чотирьох протеїнів, що транслюються з ORFs ізолятів ВІ ВРХ NC_001413 (довжиною 103 та 476 а. з.) (рис. 2) та L04974 (довжиною 102 та 476 а. з.), згенеровано та візуалізовано тривимірні структури. Протеїни довжиною 103 а. з. характеризуються значною еластичністю, в той час як протеїни довжиною 476 а. з. мають подібну до невизначеного клубка структуру.

На сервері I-TASSER структурні 3D-моделі створюються шляхом повторної зборки фрагментів, зібраних зі зв'язаних моделей, в яких біологічний сенс кінцевого протеїна впливає із збігання моделі структури з відомими протеїнами в функціональних базах даних.

Розподіл заряджених амінокислотних залишків характеризує поверхневі властивості протеїнів. Для протеїна ORF2 їхня кількість становить 23,9 %. Кількість Arg становить 5,2 %, Lys — 8,0 %, Glu — 7,3 %, Asp — 3,4 %.

Загальна кількість заряджених амінокислотних залишків ORF3 становить 23,3 %. Кількість Arg становить 12,6 %, Lys — 4,9 %, Glu — 1,9 %, Asp — 3,9 %.

З побудованих моделей третинної (рис. 2) та вторинної структури протеїнів, що транслюються з ORF3 ВІ ВРХ (рис. 3) та ORF2 (Gag) (рис. 4), визначено, що поліпротеїн Gag містить 14 α -спіралей, а протеїн ORF3 — 1 α -спіраль.

Протеїни капсиду лентівірусів складаються з N-термінальної області з семи α -спіралей та C-термінального домену, який утворено чотирма α -спіралями [19].

Найбільш вивченим лентівірусом є вірус імунodefіциту людини типу 1 (ВІЛ-1), збудник СНІД. Ретровірусний поліпротеїн Gag, зокрема, бере участь в синтезі вірусних часток, які виходять з плазматичної мембрани інфікованих клітин [20]. В вірусній реплікації та інфекційності ВІЛ-1 задіяно такі домени поліпротеїну Gag як матрикс (p17), капсид (p24), нуклеокапсид (p7), p6, SP1, SP2 [21].

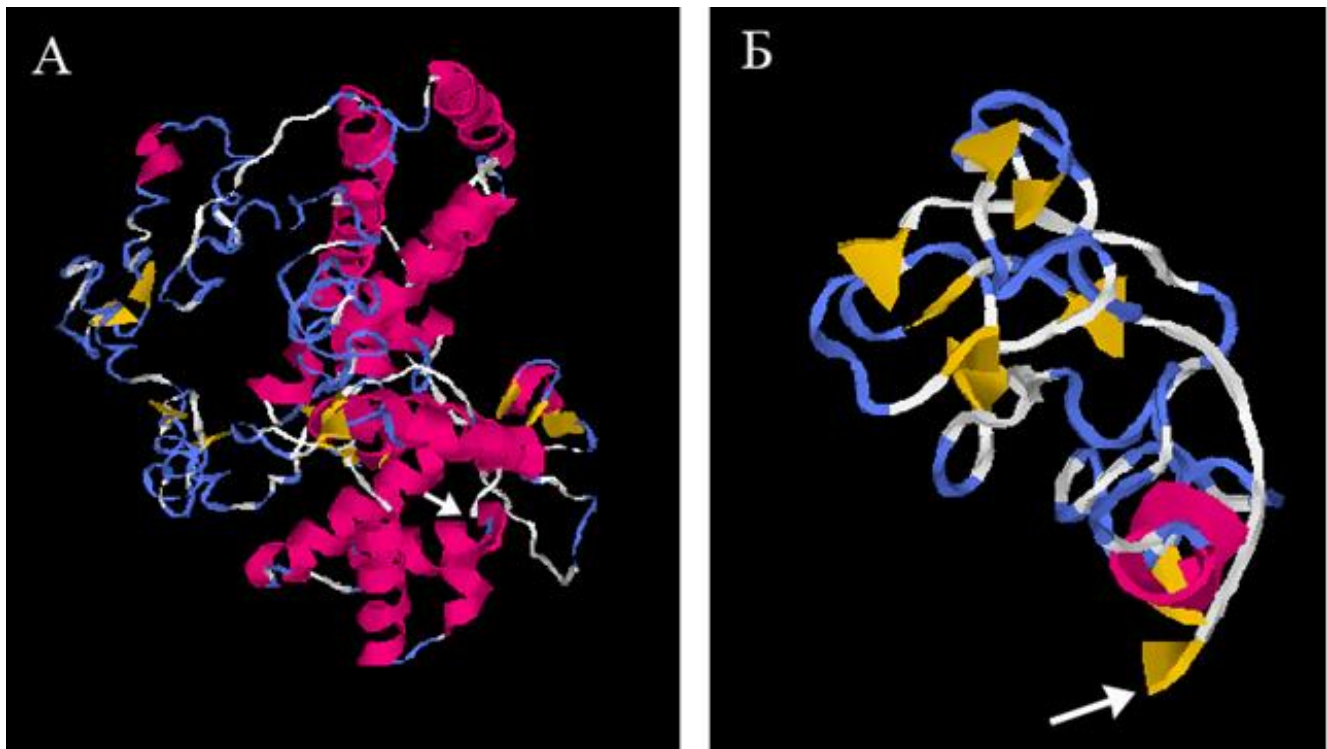


Рис. 2. Модель просторової структури поліпротеїну Gag, що транслюється з відкритої рамки зчитування ORF2 (довжиною 476 а. з.), (А) та протеїну ORF3 (довжиною 103 а. з.) (Б) ізоляту NC_001413 ВІ ВРХ. Аналіз проведено з використанням серверу I-TASSER. Стрілки вказують на N-кінці протеїнів, які є критичними для міжмолекулярної взаємодії.

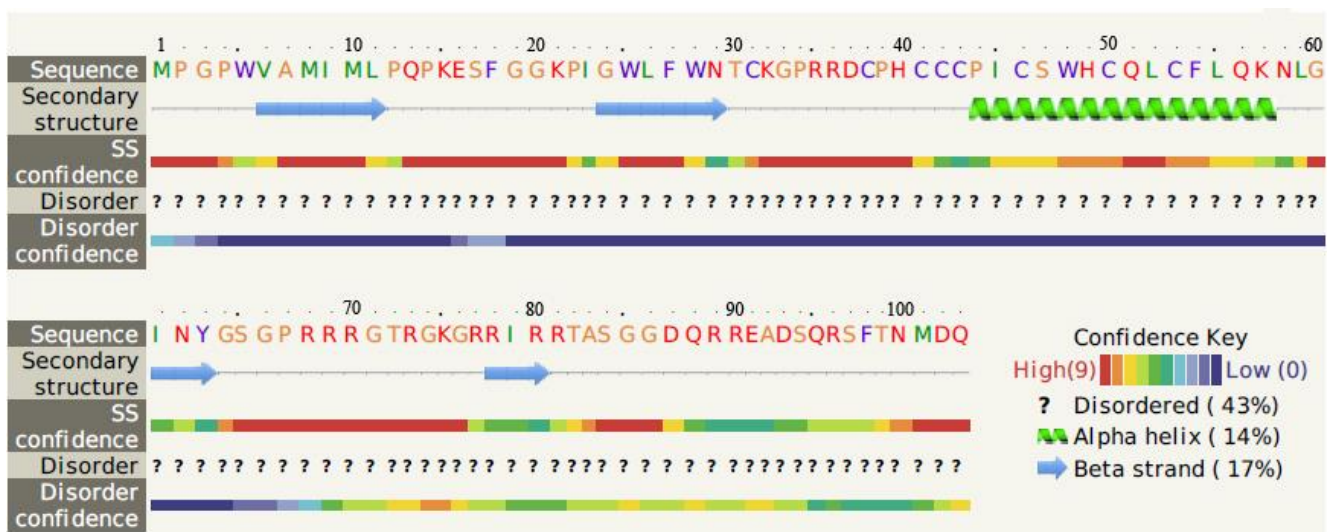


Рис. 3. Вторинна структура протеїну відкритої рамки зчитування ORF3 довжиною 103 а. з. ізоляту ВІ ВРХ з повним геномом NC_001413, яку отримано за допомогою програми Phyre2.

Для протеїну, що транслюється з ORF3 ВІ ВРХ, за допомогою програми Phyre2 передбачено 14 % α-спіралей, 17 % β-тяжів, 43 % неупорядкованої структури (рис. 3). Для поліпротеїну Gag, який транслюється з відкритої рамки зчитування ORF2 ВІ ВРХ, визначено 37 % α-спіралей, 0 % β-тяжів, 41 % неупорядкованої структури (рис. 4).

Отримані результати показують, що протеїни ORF2 та ORF3 ВІ ВРХ є гетерогенними структурами та містять упорядковані ділянки з варіабельною конформаційною стабільністю і внутрішньо неупорядковані фрагменти, які не мають упорядковану структуру.

Додатково структуру протеїнів ORF3 та ORF2 ВІ ВРХ довжиною 103 та 476 а. з. охарактеризовано за допомогою профілів внутрішнього безладу (імовірність амінокислотного залишку перебувати в неупорядкованому, або нативно-розгорнутому, стані) (рис. 5).

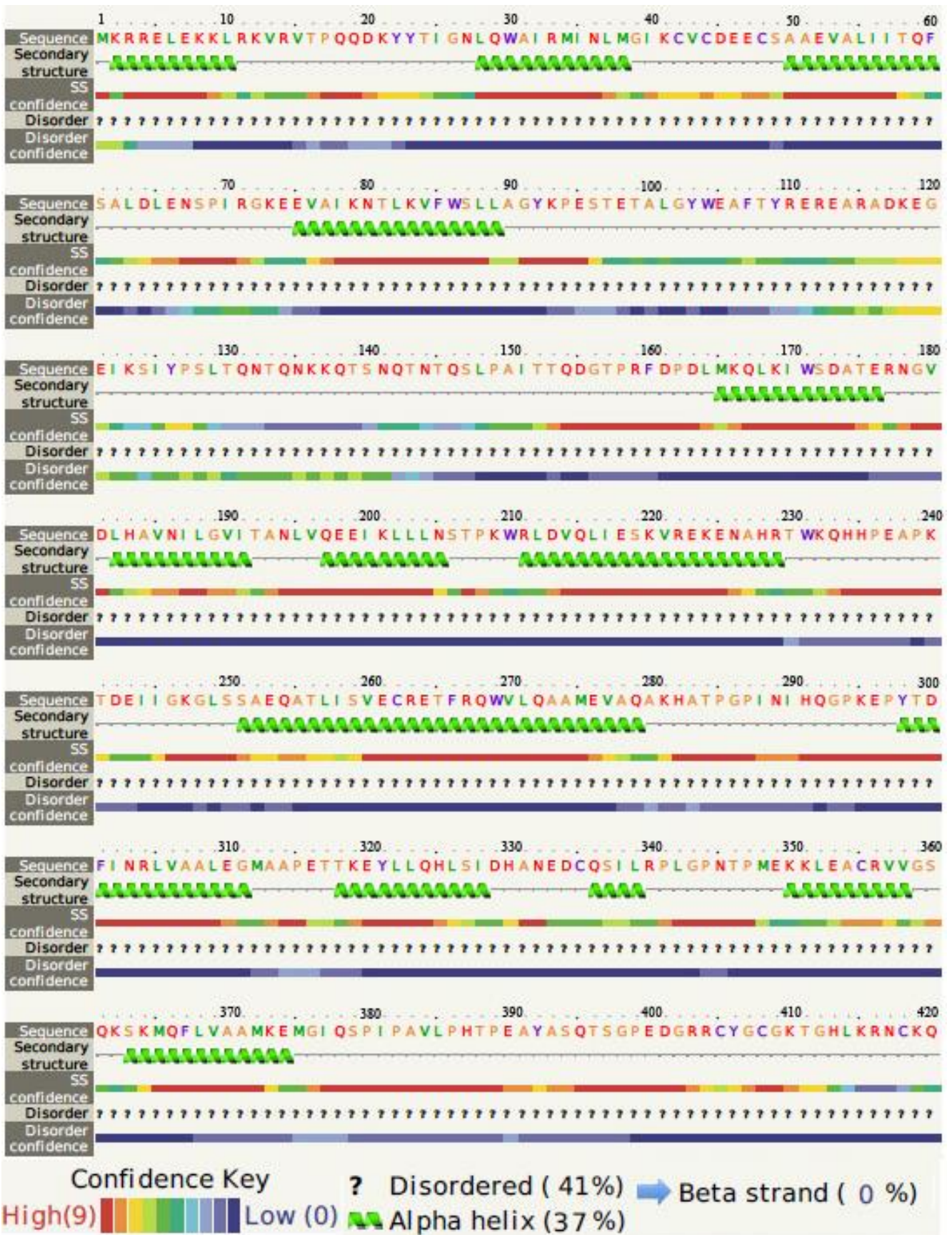


Рис. 4. Вторинна структура поліпротеїну Gag, який транслюється з відкритої рамки читування ORF2, довжиною 476 а. з. ізоляту ВІВРХ з повним геномом NC_001413, що отримано за допомогою програми Phyre2 (показано фрагмент 1–420 а. з.).

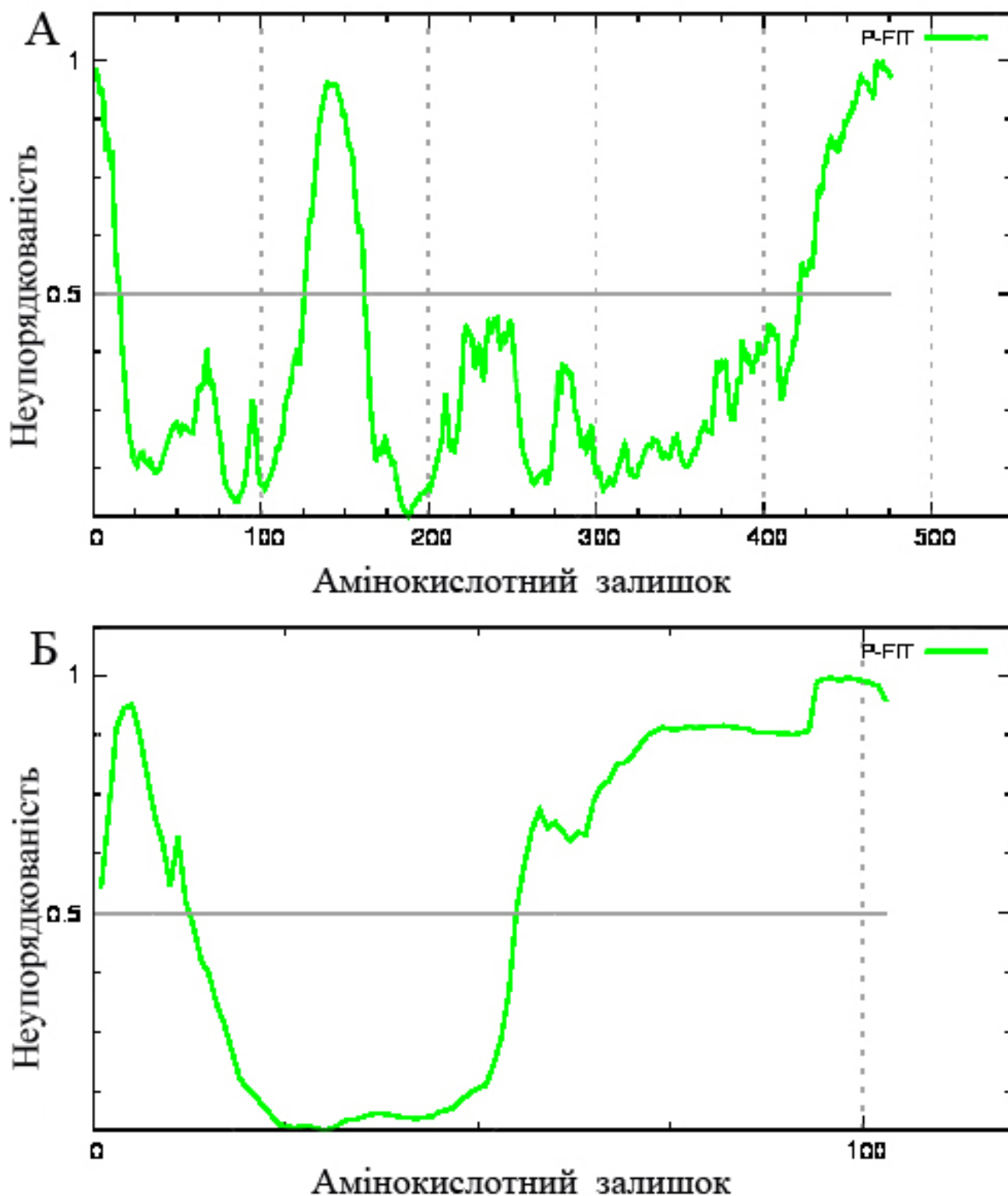


Рис. 5. Графік внутрішнього безладу для поліпротеїну Gag, який транслюється з ORF2 (довжиною 476 а. з.) (А) та протеїну ORF3 (довжиною 103 а. з.) (Б) ізоляту NC_001413 ВІ ВРХ. Імовірність безладу розраховано за допомогою програми DisProt (версія програми PONDR-FIT), в якій графіки представляють профілі внутрішнього безладу. Значення порогу 0,5 відрізняє упорядковану та неупорядковану області вздовж поліпептидного ланцюга (обмежено пунктирними лініями). Области понад пороговим значенням вважаються неупорядкованими.

Висновки. 1. В роботі визначено особливості структурної організації протеїнів двох відкритих рамок зчитування ORF2 (поліпротеїну Gag) та ORF3 вірусу імунодефіциту великої рогатої худоби (ВІ ВРХ).

2. Аналізуючи отримані дані, можна зробити наступні висновки. Протеїн ORF3 відноситься до внутрішньо неупорядкованих протеїнів, які не можуть бути стабільно складеними в унікальній тривимірній структурі за фізіологічних умов. Поліпротеїн Gag, який транскрибується з ORF2, відноситься до класу повністю структурованих протеїнів. Вторинна структура обох протеїнів демонструє наявність α -спіралі: 14 % α -спіралей для ORF3 та 37 % α -спіралей для поліпротеїну Gag.

3. Для оцінки точності передбачення стабільних та неупорядкованих частин поліпептидного ланцюга протеїнів результати роботи програми PONDR-FIT уявляється доцільним порівняти з результатами аналогічних програм.

4. З визначених п'яти ORFs для двох ізолятів ВІ ВРХ з повним геномом (NC_001413 та L04974) тільки дві ORFs збігаються за довжиною нуклеотидів (та, отже, відповідних протеїнів), що підіймає низку запитань щодо наведених в базі даних GenBank відомостей секвенсів зазначених ізолятів ВІ ВРХ з повним геномом.

Перспективи використання отриманих результатів. Відкрита рамка зчитування фактору інфекційності віріону Vif (virion infectivity factor) є консервативною для більшості лентівірусів. Молекули Vif приймають участь у реплікації вірусу за допомогою інактивування антивірусних факторів хазяїна, зокрема, цитидин деамінази АРОВЕС3 [22].

За однакової кількості Vif ВІ ВРХ, Vif ВІЛ, Vif вірус імунодефіциту мавпи (SIV, BIM) при індивідуальній коекспресії з ВІЛ-1 синтез інфекційних віріонів ВІЛ-1 в присутності Vif ВІ ВРХ зменшувався приблизно у 8 разів.

Беручи до уваги, що Vif ВІ ВРХ є потенційним інгібітором реплікації ВІЛ-1, отримані результати може бути використано для молекулярного докінгу комплексу поліпротеїну Gag з Vif ВІ ВРХ, тобто для передбачення найвигіднішої конформації такого комплексу.

Фінансування. Роботу виконано за фінансової підтримки гранту 34.02.01.01Ф Національної академії аграрних наук України.

Список літератури

1. Malmquist W. A., Van der Maaten M. J., Boothe A. D. Isolation, immunodiffusion, immunofluorescence, and electron microscopy of a syncytial virus of lymphosarcomatous and apparently normal cattle. *Cancer Research*. 1969. Vol. 29, No 1. P. 188–200. URL: <https://aacrjournals.org/cancerres/article/29/1/188/477078/Isolation-Immunodiffusion-Immunofluorescence-and>.
2. Passos-Castilho A. M., Marchand C., Archambault D. B23/nucleophosmin interacts with bovine immunodeficiency virus Rev protein and facilitates viral replication. *Virology*. 2018. Vol. 515. P. 158–164. DOI: <https://doi.org/10.1016/j.virol.2017.12.021>.
3. Zhang S. et al. Immune suppression in calves with bovine immunodeficiency virus. *Clinical and Diagnostic Laboratory Immunology*. 1997. Vol. 4. P. 232–235. DOI: <https://doi.org/10.1128/cdli.4.2.232-235.1997>.
4. Bhatia S., Patil S., Sood R. Bovine immunodeficiency virus: a lentiviral infection. *Indian Journal of Virology*. 2013. Vol. 24, No 3. P. 332–341. DOI: <https://doi.org/10.1007/s13337-013-0165-9>.
5. Rodrigues A. P. S. et al. Molecular detection of bovine immunodeficiency virus (BIV) in bovines from the state of Minas Gerais, Brazil. *Arquivo Brasileiro de Medicina Veterinaria e Zootecnia*. 2019. Vol. 71, No 2. P. 711–714. DOI: <https://doi.org/10.1590/1678-4162-10495>.
6. Gonzalez-Fernandez V. D. et al. First evidence of bovine immunodeficiency virus infection in Mexican cattle. *Transboundary and Emerging Diseases*. 2020. Vol. 67, No 5. P. 1768–1775. DOI: <https://doi.org/10.1111/tbed.13530>.
7. Keshavarz H., Mohammadi A., Morovati S. Evidence of bovine immunodeficiency virus: a molecular survey in water buffalo populations of Iran. *Veterinary Medicine and Science*. 2022. Vol. 8. P. 2167–2172. DOI: <https://doi.org/10.1002/vms3.872>.
8. Gradil C. M. et al. Detection of bovine immunodeficiency virus DNA in the blood and semen of experimentally infected bulls. *Veterinary Medicine*. 1999. Vol. 70. P. 21–31. DOI: [https://doi.org/10.1016/s0378-1135\(99\)00130-3](https://doi.org/10.1016/s0378-1135(99)00130-3).
9. Garvey K. J. et al. Nucleotide sequences and genome organization of biologically active proviruses of the bovine immunodeficiency-like virus. *Virology*. 1990. Vol. 175, No 2. P. 391–409. DOI: [https://doi.org/10.1016/0042-6822\(90\)90424-p](https://doi.org/10.1016/0042-6822(90)90424-p).
10. Chou P. Y., Fasman G. D. Prediction of the secondary structure of proteins from their amino acid sequence. *Advances in Enzymology and Related Areas of Molecular Biology*. 1978. Vol. 47. P. 45–48. DOI: <https://doi.org/10.1002/9780470122921.ch2>.
11. Tompa P. Intrinsically unstructured proteins. *Trends in Biochemical Sciences*. 2002. Vol. 27, No 10. P. 527–533. DOI: [https://doi.org/10.1016/s0968-0004\(02\)02169-2](https://doi.org/10.1016/s0968-0004(02)02169-2).

12. Uversky V. Functional roles of transiently and intrinsically disordered regions within proteins. *FEBS Journal*. 2015. Vol. 282, No 7. P. 1182–1189. DOI: <https://doi.org/10.1111/febs.13202>.
13. Brocca S. et al. Liquid-liquid phase separation by intrinsically disordered protein regions of viruses: roles in viral life cycle and control of virus-host interactions. *International Journal of Molecular Sciences*. 2020. Vol. 21, No 23. P. 9045–9075. DOI: <https://doi.org/10.3390/ijms21239045>.
14. Liu B. et al. Evidence for the antisense transcription in the proviral R29-127 strain of bovine immunodeficiency virus. *Sinica*. 2015. Vol. 30, No 3. P. 224–227. DOI: <https://doi.org/10.1007/s12250-015-3559-6>.
15. Rasmussen M. H. et al. Antisense transcription in gammaretroviruses as a mechanism of insertional activation of host genes. *Journal of Virology*. 2010. Vol. 84, No 8. P. 3780–3788. DOI: <https://doi.org/10.1128/jvi.02088-09>.
16. Hall T. A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. 1999. Vol. 41. P. 95–98.
17. Yang J., Zhang Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Research*. 2015. Vol. 43, No W1. P. W174–W181. DOI: <https://doi.org/10.1093/nar/gkv342>.
18. Kelley L. A. et al. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*. 2015. Vol. 10, No 6. P. 846–858. DOI: <https://doi.org/10.1038/nprot.2015.053>.
19. Ovejero C. A., González S. A., Affranchino J. L. The conserved Tyr176/Leu177 motif in the α -helix 9 of the feline immunodeficiency virus capsid protein is critical for Gag particle assembly. *Viruses*. 2019. Vol. 11, No 9. P. 816. DOI: <https://doi.org/10.3390/v11090816>.
20. Garbitt-Hirst R., Kenney S. P., Parent L. J. Genetic evidence for a connection between Rous sarcoma virus Gag nuclear trafficking and genomic RNA packaging. *Journal of Virology*. 2009. Vol. 83, No 13. P. 6790–6797. DOI: <https://doi.org/10.1128/jvi.00101-09>.
21. Marie V., Gordon M. L. The HIV-1 Gag protein displays extensive functional and structural roles in virus replication and infectivity. *International Journal of Molecular Sciences*. 2022. Vol. 23, No 14. P. 7569. DOI: <https://doi.org/10.3390/ijms23147569>.
22. Zheng W. et al. Conserved interaction of lentiviral vif molecules with HIV-1 Gag and differential effects of species-specific vif on virus production. *Journal of Virology*. 2017. Vol. 91, No 7. P. e00064-17. DOI: <https://doi.org/10.1128/jvi.00064-17>.

STRUCTURAL ANALYSIS OF OPEN READING FRAMES OF BOVINE IMMUNODEFICIENCY VIRUS PROTEINS

Balak O. K.

Kharkiv National Medical University, Kharkiv, Ukraine

Lymanska O. Yu.

National Scientific Center “Institute of Experimental and Clinical Veterinary Medicine”, Kharkiv, Ukraine

The goal of this study was determining the structural organization peculiarities of the ORF2 and ORF3 proteins of the bovine immunodeficiency virus (BIV). Five ORFs were determined for two BIV isolates with complete genome using the ATGpr software, which permits effective prediction of translation initiation codons with nucleotide accuracy. Phyre2 software was used to predict, analyze the secondary structure and function of proteins. PONDR-FIT software was used to search for protein fragments in a disordered or natively unfolded state. Analysis of the amino acid composition of ORF2 and ORF3 proteins of BIV isolates regarding the presence of nonpolar, polar, aromatic, and hydrophobic amino acid residues was carried out using PSIPRED software. Models of the 3D-structure of proteins were obtained by I-TASSER server. 14% of α -helices, 17% of β -strands and 43% of disordered structure are predicted for the ORF3 protein. 37% of α -helices, 0% of β -strands, and 41% of disordered structure were determined for Gag polyprotein, which is translated from ORF2. The distribution of charged amino acid residues characterizes the surface properties of proteins. Their number reaches 23.9% for ORF2 protein. The amount of Arg is 5.2%, Lys — 8.0%, Glu — 7.3%, Asp — 3.4%. The total number of charged amino acid residues of ORF3 is 23.3%. The number of Arg is 12.6%, Lys — 4.9%, Glu — 1.9%, Asp — 3.9%. Only two ORFs of five ones coincide in nucleotide length (and, therefore, in length of corresponding proteins) for the two BIV isolates. The ORF3 protein belongs to the intrinsically disordered proteins that cannot be stably folded into a unique three-dimensional structure under physiological conditions, and the Gag polyprotein, which is translated from ORF2, belongs to the class of fully structured proteins. The secondary structure of both proteins shows the presence of α -helices

Keywords: polyprotein Gag, Lentivirus, Retroviridae