



УДК 577.21.5

© 2011

И. Н. Коломиец, М. И. Зарудная, А. Л. Потягайло,  
член-корреспондент НАН Украины Д. Н. Говорун

## Вторичная структура домена PBS вируса иммунодефицита человека субтипов В, С, D и F

*Проведено філогенетичне дослідження вторинної структури домену PBS, який відповідає за ініціацію оберненої транскрипції, для 941 ізоляту ВІЛ-1 субтипів В, С, D і F. Показано, що PBS-сигнал, три додаткових сайти зв'язування з праймером, а також два дуплекси, що впізнаються оберненою транскриптазою, висококонсервативні для більшості ізолятів досліджених субтипів. Вперше виявлено особливості структури PBS-частини, шпильки U5 та нижньої частини домену PBS для ізолятів кожного з досліджених субтипів. Показано зокрема, що шпилька, яка містить PBS-сигнал, має різну довжину (11–24 нт) і стабільність (4–10 ккал/моль), що зумовлено в основному мутаціями в положеннях 85 і 86. Також виявлено, що шпилька U5 має три основні форми, які відрізняються за структурою середньої частини шпильки та різним експонуванням додаткових сайтів зв'язування з праймером.*

Домен PBS, необходимый для инициации обратной транскрипции геномной РНК вируса иммунодефицита человека (ВИЧ-1), содержит PBS-сигнал (primer binding site) длиной 18 нт, образующий дуплекс с 3'-концевой частью праймера tRNA(lys3) [1]. В литературе высказывается предположение о наличии дополнительных мест связывания праймера с доменом, в частности с AAAA и CCCUCA трактами [2–4] или с PAS-сигналом (primer activating signal) [1, 5]. Кроме того, предполагается, что два дуплекса домена PBS, один из которых содержит PAS-сигнал, узнаются обратной транскриптазой [6].

Модели вторичной структуры домена PBS (в основном в составе лидерной области генома ВИЧ-1) представлены в литературе только для изолятов LAI/HXB2 (субтип В), NL4.3 (конструкт субтипа В) и MAL (рекомбинант субтипов А1, D и К). Для домена PBS изолятов субтипа В определены две модели, которые отличаются экспонированием сигналов связывания с праймером и наличием дуплексов, узнаваемых обратной транскриптазой (например, [4, 7]).

Насколько нам известно, в литературе не приводились модели вторичной структуры домена PBS для изолятов самого распространенного в мире субтипа С, а также многих других

субтипов. Следует отметить, что субтип С, наряду с субтипом А, является одним из самых распространенных на территории Украины [8]. Мы впервые исследовали структуру домена PBS для большого пула изолятов ВИЧ-1 разных субтипов (1067 изолятов) и обнаружили, что структура домена субтипов В, С, D и F подобна структуре домена LAI/HXB2/NL4.3, а А/С-содержащих субтипов — MAL. В настоящей работе приведены результаты исследования особенностей вторичной структуры домена PBS для изолятов субтипов В, С, D и F.

**Методы.** Для исследования были взяты все представленные в ГенБанке на конец 2009 г. геномы ВИЧ-1, содержащие последовательность домена PBS. В анализ не включались изоляты ВИЧ-1 с рекуррентной последовательностью домена PBS из одного и того же больного. Вторичную структуру исследуемой области определяли с помощью программы mfold Зукера [9] со следующими параметрами: субоптимальность 5%, window 0. В данной работе первому нуклеотиду домена PBS соответствует 116 нуклеотид в геноме ВИЧ-1 с номером NC\_001802 (RefSec в ГенБанке).

**Домен PBS в изолятах субтипа В.** Предсказание вторичной структуры домена PBS для 441 изолята ВИЧ-1 субтипа В (из разных регионов мира) показало, что в 75% изолятов структура домена (рис. 1, а) подобна представленной для изолятов LAI/HXB2/NL4.3 в первой модели [4]. В этой модели домен PBS можно условно разделить на три части — PBS-часть, которая включает PBS-сигнал, нижняя часть, содержащая дуплекс с PAS-сигналом, и U5-шпилька, предшествующая PBS-сигналу, средняя часть которой содержит дополнительные места связывания с праймером AAAA и CCCUCA, а ее нижний дуплекс (дуплекс U5) узнается обратной транскриптазой.

В 16% исследованных изолятов домен PBS имеет структуру (см. рис. 1, б), подобную представленной во второй модели. В этой модели (см., например, [7] и библиографию в этой работе) дополнительные места связывания с праймером расположены в апикальной области двух разных шпилек, одна из которых (шпилька U5(AB)) включает четыре 5'-концевых нуклеотида PBS-сигнала. Кроме того, в этой модели нет дуплекса U5, узнаваемого обратной транскриптазой. Остальные 9% исследованных нами изолятов субтипа В также имеют домен со шпилькой U5(AB), но с существенно измененной нижней и средней частью домена.

PBS-часть домена в описанных в литературе моделях или не структурирована, или содержит одну шпильку длиной 11 нт (11hPBS), включающую часть сигнала PBS (например, [10]). Мы впервые показали, что PBS-сигнал может быть включен не только в шпильку 11hPBS, но также в шпильки разной длины (11–24 нт) и стабильности (4–10 ккал/моль) с одинаковой верхней частью (11hPBS). Наиболее частые варианты шпилек PBS в изолятах субтипов В, С, D и F показаны на рис. 2. Согласно недавно полученным данным [11], нуклеотид G, предшествующий шпильке 11hPBS, и/или ее 3'-концевой нуклеотид вовлечены в образование самых первых пар оснований при взаимодействии праймера с вирусной РНК. Включение сигнала PBS не только в шпильку 11hPBS, но и в шпильки другой длины и структуры может влиять на процесс образования дуплекса.

В изолятах субтипа В наиболее часто встречаются следующие шпильки PBS: 24hPBS<sup>85U</sup> (30%), 11hPBS (22%), 21hPBS<sup>86C</sup> (10%), 24hPBS (8%), 18hPBS<sup>85G,86C</sup> (8%) и 20hPBS<sup>86G</sup> (5%) (см. рис. 2). Следует отметить, что в 75% изолятов субтипа В со шпилькой U5 в домене PBS PBS-часть содержит наряду со шпилькой PBS вторую короткую шпильку, чаще всего шпильку длиной 8 нт с апикальной петлей GAAA (8hGAAA<sup>98U</sup>, см. рис. 1, а), а в остальных 25% изолятах PBS-часть содержит только шпильку PBS.

Анализ вторичной структуры домена PBS со шпилькой U5 в изолятах исследованных субтипов показал, что эта шпилька имеет три основные формы, которые отличаются по

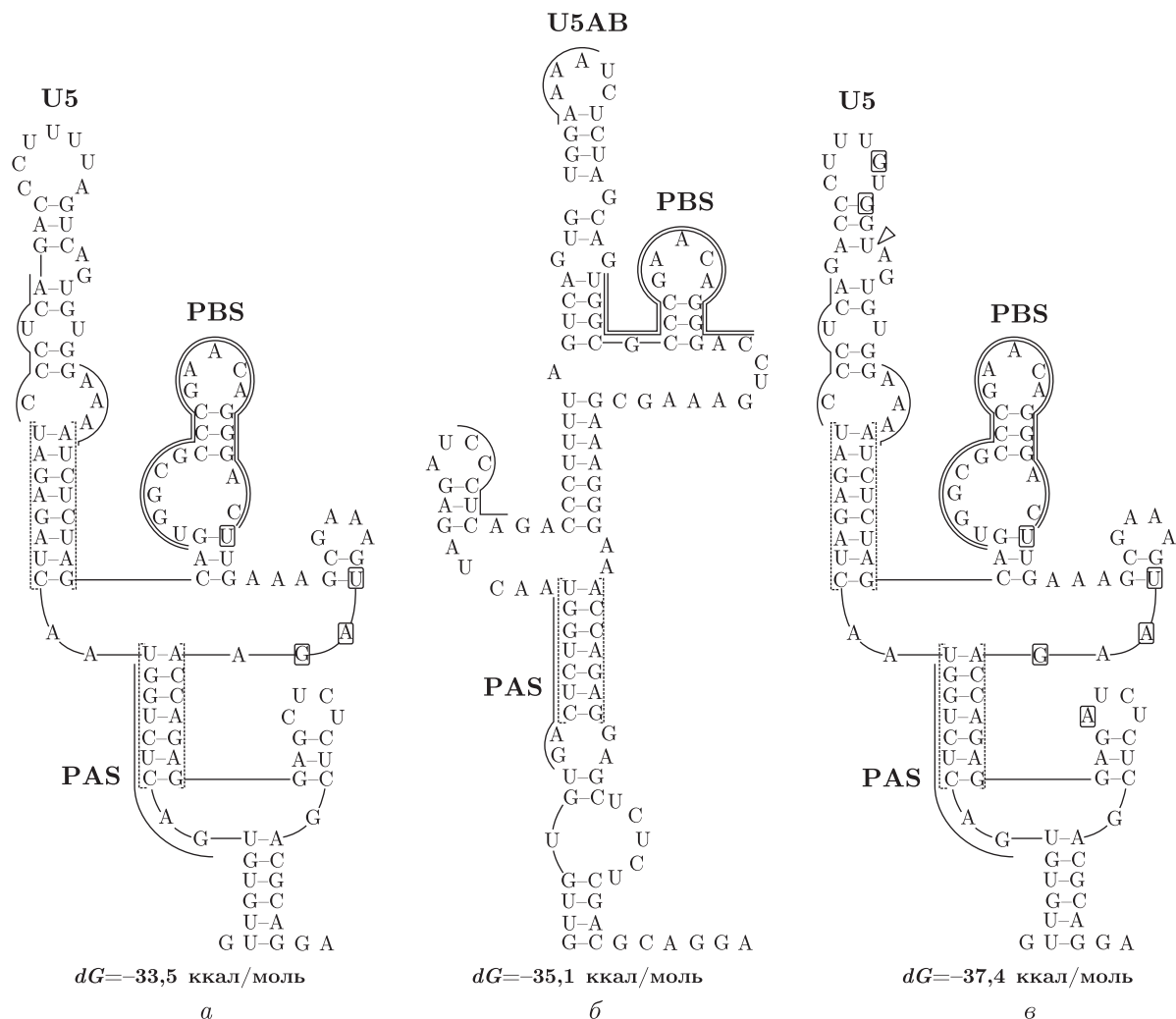


Рис. 1. Оптимальная структура домена PBS геномной РНК ВИЧ-1 изолятов U34603 (*a*), HXB2 (RefSec) (*б*) субтипа В и изолята AY228556 субтипа С (*в*). Замены оснований по сравнению с реферативной последовательностью генома HXB2 (RefSec) отмечены квадратами, делеции — треугольником. Основной PBS-сигнал обозначен двойной линией, дополнительные сайты связывания с праймером — линией, дуплексы, узнаваемые обратной транскриптазой — пунктирной линией

структуре средней части и экспонированию дополнительных сигналов связывания с праймером (рис. 3, *a*; 4, *a*). В зависимости от мутаций и структуры шпильки U5 большинство изолятов субтипа В можно разделить на пять основных групп: к первой группе относятся 32% изолятов со шпилькой U5 без мутаций (см. рис. 3, *a*); ко второй — 21% изолятов с мутациями в апикальной петле и балдже шпильки U5, не приводящими к значительному изменению структуры, например с мутацией 37А (см. рис. 3, *б*); к третьей — 6% изолятов со шпилькой U5, стабилизированной дополнительной G-C парой на конце стебля; к четвертой — 7% изолятов со специфическими изменениями в средней части шпильки U5, обусловленными в основном мутацией 36А (см. рис. 3, *в*); к пятой — 5% изолятов, шпилька U5 которых имеет форму 3 в оптимальной структуре домена PBS, обусловленную в основном мутацией 52А (см. рис. 3, *з*).

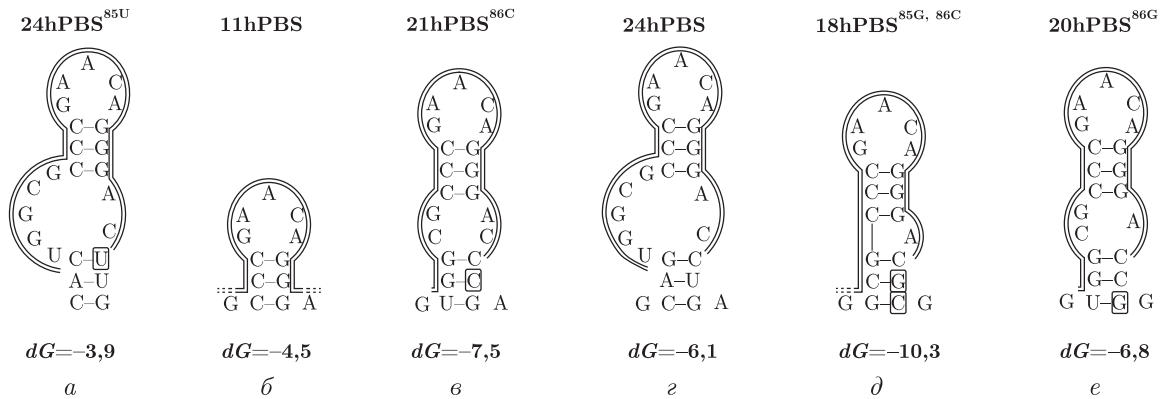


Рис. 2. Варианты шпильки, включающей основной PBS-сигнал — 24hPBS<sup>85U</sup> (*a*), 11hPBS (*б*), 21hPBS<sup>86C</sup> (*в*), 24hPBS (*г*), 18hPBS<sup>85G,86C</sup> (*д*), 20hPBS<sup>86G</sup> (*е*). Замены оснований по сравнению с реферативной последовательностью генома HXB2 (RefSec) отмечены квадратами. Основной PBS-сигнал обозначен двойной линией. Энергия образования шпильки ( $dG$ ) приведена в ккал/моль

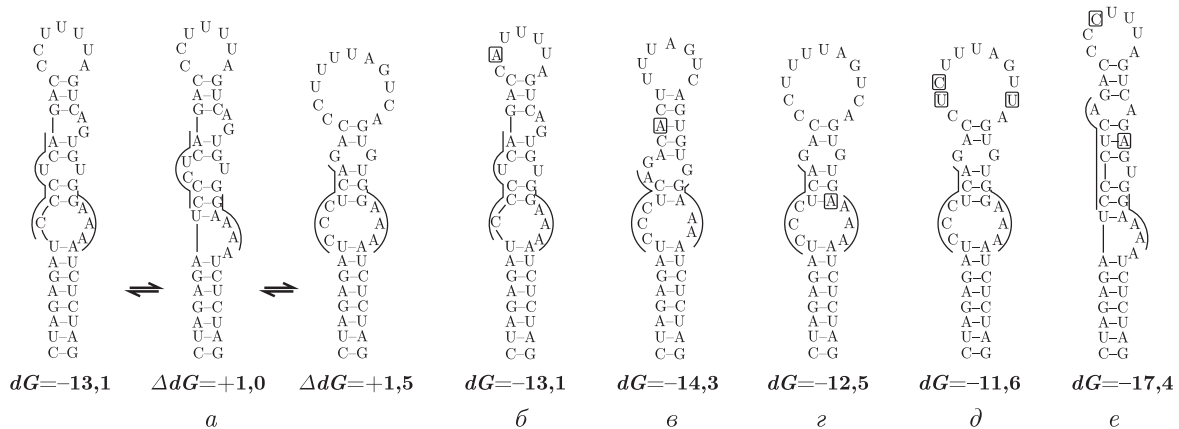


Рис. 3. Варианты и структурные формы шпильки U5 в изолятах субтипов В, D и F — без замен оснований (*a*), с заменой 37А (*б*), с заменой 36А (*в*), с заменой 52А (*г*), с заменами 37U, 38C, 45U (*д*), с заменами 38C, 48А (*е*). Замены оснований по сравнению с реферативной последовательностью генома HXB2 (RefSec) отмечены квадратами. Дополнительные сайты связывания с праймером обозначены линией. Энергия образования основной формы шпильки ( $dG$ ) и увеличение энергии образования неосновных форм относительно основной ( $\Delta dG$ ) приведены в ккал/моль

Наиболее часто встречающейся структурой нижней части домена PBS (BP, bottom part) в изолятах субтипа В является структура без мутаций (35% изолятов). Эта структура встречается преимущественно в форме 1 в доменах со шпилькой U5(AB) (см. рис. 1, б), а в доменах со шпилькой U5 — в форме 1 и 2 (см. рис. 1, а) — примерно с одинаковой частотой. К другим часто встречающимся структурам относятся BP<sup>112A</sup> (10%), BP<sup>105G</sup> (8%) и BP<sup>109A</sup> (7%).

**Домен PBS в изолятах субтипа С.** Исследование вторичной структуры домена PBS в 449 изолятах ВИЧ-1 субтипа С, в основном из Южной и Центрально-Восточной Африки, показало, что домен PBS со шпилькой U5 (см. рис. 1, в) образуется в изолятах этого субтипа значительно чаще (96%), чем в изолятах субтипа В (75%).

PBS-часть домена со шпилькой U5 в изолятах субтипа С чаще, чем в изолятах субтипа В, содержит две шпильки, в ней примерно в 1,5–2 раза чаще встречаются длинные формы



Четвертую группу составляют 9% изолятов, шпилька U5 которых имеет форму 3 (или форму, подобную форме 3) в оптимальной структуре домена PBS, эта форма стабилизируется, например, комбинацией частых мутаций 42G, 45d с 52A или с 46G.

Пятую группу составляют 5% изолятов, шпилька U5 которых имеет форму 2 в оптимальной структуре домена PBS, эта форма стабилизируется, например, комбинацией 42G, 45d с 48C/A (3%).

Интересно, что в малочисленной группе изолятов субтипа С из Северо-Восточной Африки и Бразилии (2%) вместо частых мутаций 42G и 45d U5-шпилька содержит мутации в положении 35/36/37 и 42d, которые стабилизируют шпильку U5 на 1,6 ккал/моль (см. рис. 4, з).

Замена 112C → 112A в нижней части домена PBS субтипа С встречается отдельно (см. рис. 1, в) или в сочетании с другими мутациями в 76% исследованных изолятов. Структуры только с этой заменой (BP<sup>112A</sup>) или с двойной мутацией (BP<sup>109A,112A</sup>) встречаются примерно в 4 раза чаще в изолятах субтипа С, чем субтипа В, причем в изолятах субтипа С встречаются примерно с одинаковой частотой обе формы BP<sup>109A,112A</sup>, а в изолятах субтипа В — только первая форма.

**Домен PBS в изолятах субтипов D и F.** Геномы ВИЧ-1 субтипов D и F представлены в ГенБанке гораздо реже, чем геномы субтипов В и С, поэтому мы смогли изучить вторичную структуру домена PBS лишь для 37 изолятов субтипа D (в основном из Центральной Африки) и 14 изолятов субтипа F1 (в основном из Бразилии). Все изоляты субтипа F имеют домен с U5-шпилькой, а среди изолятов субтипа D, подобно изолятам близкородственного субтипа В, встречаются домены как со шпилькой U5 (73%), так и со шпилькой U5(AB) (24%).

Высокая частота мутаций 85U и 98U и низкая частота мутаций в положении 86 в домене PBS изолятов субтипа D определяет структуру PBS-части в этих изолятах. Во всех изолятах со шпилькой U5 PBS-часть сворачивается с образованием двух шпилек, причем часто встречаются как длинные шпильки PBS — 24hPBS<sup>85U</sup> (30%) и 24hPBS (19%), так и короткая 11hPBS (41%). Наиболее распространенным вариантом второй шпилки, как и в изолятах других субтипов, является 8hGAAA<sup>98U</sup>.

В 30% изолятов субтипа D шпилька U5 содержит только двойную мутацию 37U, 38C, которая не влияет на стабильность трех форм шпилки. Приблизительно в половине изолятов субтипа D шпилька U5 содержит различные комбинации из 2–4 мутаций, которые высоко специфичны для A/G-содержащих субтипов. Одни комбинации приводят к образованию формы 3 шпилки U5 в оптимальной структуре домена (см. рис. 3, д), а другие — к образованию балджа AAA вместо внутренней петли и разрушению верхней пары U-A нижнего дуплекса, как в форме 2 (см. рис. 3, е).

Структура нижней части домена PBS без мутаций встречается в изолятах субтипа D примерно в два раза реже, чем в изолятах субтипа В. Специфическим вариантом по сравнению с изолятами субтипов В и С является структура с двойной мутацией 7d, 8d (BP<sup>7d,8d</sup>, 30% изолятов).

Предсказание вторичной структуры домена PBS в изолятах субтипа F показало, что домен во всех исследованных изолятах содержит шпильку U5, а PBS-часть в этих изолятах сворачивается с образованием двух шпилек. К наиболее часто встречающимся вариантам относятся 18hPBS<sup>85G,86C</sup> (36%), 24hPBS<sup>85U</sup> (29%), 24hPBS (14%) и 11hPBS (7%), первый из этих вариантов встречается в изолятах субтипа F в 3–4 раза чаще, чем в изолятах субтипов В или С.

Шпилька U5 в большинстве изолятов субтипа F (78%) содержит или одиночную мутацию 37A (см. рис. 3, а) или комбинацию 37A с 39C/40C, которые не влияют на стабильность трех форм шпильки. Нижняя часть домена PBS чаще всего сворачивается в структуры ВР<sup>109А,112А</sup> и ВР<sup>112А</sup> так же, как и в изолятах субтипа С.

На основании полученных данных по структуре домена PBS можно заключить, что для изолятов субтипов В и D характерны короткие и длинные формы шпильки PBS в PBS-части и отсутствие мутаций в нижней части домена, а для изолятов субтипов С и F типичны длинные формы шпильки PBS и структура нижней части домена с мутацией 112А. Шпилька U5 домена PBS в изолятах субтипа С имеет более высокую стабильность по сравнению с изолятами субтипов В, D и F из-за высокоспецифических мутаций 42G и 45d, а для шпильки U5 домена PBS в половине изолятов субтипа D характерно наличие мутаций, специфических для А/С-содержащих субтипов.

Таким образом, исследование структуры домена PBS для большого пула изолятов ВИЧ-1 субтипов В, С, D и F показало, что все три структурные части домена (шпилька U5, PBS-часть и нижняя часть) имеют свои особенности для изолятов каждого субтипа, при этом для изолятов всех субтипов PBS-сигнал и три дополнительных сайта связывания с праймером высококонсервативны. Кроме того, оба дуплекса, предположительно узнаваемые обратной транскриптазой, также консервативны для большинства изолятов исследованных субтипов.

*Исследование выполнено при частичной поддержке грантом Украинского научно-технологического центра (проект 4302).*

1. *Abbink T. E., Berkhout B.* HIV-1 reverse transcription initiation: a potential target for novel antivirals? // *Virus Res.* – 2008. – **134**, No 1–2. – P. 4–18.
2. *Goldschmidt V., Paillart J. C., Rigourd M. et al.* Structural variability of the initiation complex of HIV-1 reverse transcription // *J. Biol. Chem.* – 2004. – **279**, No 34. – P. 35923–35931.
3. *Iwatani Y., Rosen A. E., Guo J. et al.* Efficient initiation of HIV-1 reverse transcription in vitro // *Ibid.* – 2003. – **278**, No 16. – P. 14185–14195.
4. *Wilkinson K. A., Gorelick R. J., Vasa S. M. et al.* High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states // *PLoS Biol.* – 2008. – **6**, No 4. – e96.
5. *Saadatmand J., Niu M., Kleiman L., Guo F.* The contribution of the primer activation signal to differences between Gag- and NCp7-facilitated tRNA(Lys3) annealing in HIV-1 // *Virology.* – 2009. – **391**, No 2. – P. 334–341.
6. *Isel C., Westhof E., Massire C. et al.* Structural basis for the specificity of the initiation of HIV-1 reverse transcription // *EMBO J.* – 1999. – **18**, No 4. – P. 1038–1048.
7. *Russell R. S., Liang C., Wainberg M. A.* Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? // *Retrovirology.* – 2004. – **1**. – P. 23.
8. *Іванська Н. В., Трохимчук Т. Ю.* Проблеми мінливості вірусу імунодефіциту людини // *Біополімери і клітина.* – 2004. – **20**, № 3. – С. 171–181.
9. *Zuker M.* Mfold web server for nucleic acid folding and hybridization prediction // *Nucl. Acids Res.* – 2003. – **31**, No 13. – P. 3406–3415.
10. *Henriet S., Richer D., Bernacchi S. et al.* Cooperative and specific binding of Vif to the 5' region of HIV-1 genomic RNA // *J. Mol. Biol.* – 2005. – **354**, No 1. – P. 55–72.
11. *Barraud P., Gaudin C., Dardel F., Tisné C.* New insights into the formation of HIV-1 reverse transcription initiation complex // *Biochimie.* – 2007. – **89**, No 10. – P. 1204–1210.

**I. N. Kolomiets, M. I. Zarudnaya, A. L. Potyahaylo,**  
Corresponding Member of the NAS of Ukraine **D. M. Hovorun**

**The secondary structure of PBS domain in human immunodeficiency virus of B, C, D, and F subtypes**

*Phylogenetic analysis of the secondary structure of PBS domain directing the reverse transcription initiation has been conducted for 941 HIV-1 isolates of B, C, D, and F subtypes. PBS signal and three additional sites for tRNA(lys3) primer binding have been shown to be highly conserved among all subtypes studied, and two duplexes recognized by reverse transcriptase also appeared to be conserved in most isolates studied. The structural peculiarities of PBS part, U5-hairpin, and bottom part of PBS domain have been determined for each subtype studied. In particular, a hairpin with PBS signal was shown to be of different length (11–24 nt) and stability (4–10 kcal/mole) depending mainly on base changes at positions 85 and 86. U5-hairpin was found to have three main forms which differ in the structure of the middle part and in the exposure of additional sites for primer binding.*