

Minireview

The puzzle of RNA recombination

Alexander B. Chetverin*

Institute of Protein Research, Russian Academy of Sciences, 142292 Pushchino, Russia

Received 23 August 1999

Abstract For more than three decades, RNA recombination remained a puzzle and has only begun to be solved in the last few years. The available data provide evidence for a variety of RNA recombination mechanisms. Non-homologous recombination seems to be the most common for RNA. Recent experiments in both the *in vitro* and the *in vivo* systems indicate that this type of recombination may result from various transesterification reactions which are either performed by RNA molecules themselves or are promoted by some proteins. The high frequency of homologous recombination manifested by some RNA viruses can be easier explained by a replicative template switch.

© 1999 Federation of European Biochemical Societies.

Key words: RNA recombination; RNA virus; Cell-free system; Transesterification reaction; Replicative template switch

1. Introduction

RNA recombination was discovered in the early 60s as an exchange of genetic material between closely related RNA viruses which co-infect the same cell [1,2]. Since then, the formation of recombinant RNAs was observed in infected cells of every kingdom [3]. However, until recently, the molecular mechanism of this process remained an unsolved puzzle. In contrast to DNA recombination, no specific genes were associated with the recombinant formation that precluded dissection of the recombination mechanism by genetic tools; neither could the mechanism be studied biochemically because of the absence of cell-free experimental models. There was no answer even to such basic questions as to whether the recombination is accomplished by viral proteins (e.g. replicase), by host proteins or by RNA molecules themselves. Moreover, since RNA recombination was only observed *in vivo*, the possibility could not be excluded that the recombining species were not RNAs, but their cDNA copies.

Since RNA recombination is a very rare event, it can only be detected if there are means of selective amplification of recombinant molecules. Therefore, for a long time, RNA recombinations were exclusively observed in viral systems, but even in these systems, recombinants were occasionally detected consisting partially or entirely of cellular RNA segments [3–5]. It remained unclear whether any factors accompanying viral infection (e.g. presence of viral replicase) are required for recombinant formation or whether RNA recombination can occur in uninfected cells as well.

Apart from experimental difficulties, progress in this field

was partially retarded by common belief that any type of RNA recombination must be due to a replicative template switch [3,5–8]. A breakthrough has been made during the last 2 years owing to the development of cell-free experimental models. It has appeared that RNAs can indeed recombine without DNA intermediates, that there are a variety of non-replicative mechanisms and that the existence of such mechanisms can be demonstrated in conventional *in vivo* systems.

2. Homologous versus non-homologous recombination

The easiest way to observe RNA recombination is to co-infect a cell with two related viruses which carry non-complementing genetic defects or selective markers, such that no virus multiplication is possible unless the viral genomes exchange their segments [3]. Recombination may be homologous, if both the parents and the progeny are homologous to each other around the cross-over site, or non-homologous, if it occurs otherwise. While homologous recombination preserves the genome structure, non-homologous recombination changes it. Since most of the viral genome comprises coding sequences or those essential for RNA replication and/or packaging into a virion, it is clear that most of the non-homologous recombinants will not survive. Therefore, experiments of this kind mostly reveal the homologous recombination producing an impression that homologous RNA recombination occurs more frequently than the non-homologous one [3]. As will be seen below, such a view is incorrect.

Non-homologous recombination can be observed by crossing deletion variants or fragments of viral RNAs which overlap at non-essential regions and whose recombination may restore the intact genome [9–12]. It can also be observed by using an engineered virus carrying an insert in a non-essential region. Often, such an insert renders the virus viable but confers onto it a small-plaque phenotype. Occasionally, the insert is deleted by intrastrand recombination, resulting in the restoration of the wild-type phenotype, but the deletions are not necessarily precise [13–15]. Spontaneous deletions can also occur in wild-type viruses, resulting in the generation of defective interfering (DI) particles [16,17]. DI particles were first described under the name of ‘inactive viruses’ as early as in 1943 [18], but only in the 70s was it recognized that these are generated by non-homologous RNA recombination. Since the deletions result in a defect or a lack of at least one of the viral genes, DI particles can only propagate if the same cell is co-infected with the intact authentic virus providing the missing proteins *in trans* and are detected as particles contaminating the virus preparation. DI particles have been named so due to their ability to interfere with replication of the helper virus [16,17].

An important characteristic of recombination is its fre-

*Fax: (7) (095) 924-0493. E-mail: alexch@vega.protres.ru

quency. The frequency of homologous recombination greatly varies among different viruses. It is as high as 10^{-5} /nucleotide (nt) in crosses between coronaviruses or picornaviruses belonging to the same serological type [3], but in RNA phages, the frequency is as low as 10^{-11} /nt [19]. The frequency of non-homologous recombination was almost never reported, but it can be roughly estimated from the data presented in some papers. In each studied instance, the estimate gives the same value, about 10^{-5} /nt, i.e. the same order as the frequency of the most efficient homologous recombination. Surprisingly, non-homologous recombination is equally efficient in picornaviruses [13,20] manifesting the highest rate of homologous recombination, in RNA phages [15] producing homologous recombinants a million times less frequently and even in alphaviruses [12] in which homologous recombination was not detected at all [9].

The absence of any correlation in the frequencies of homologous and non-homologous recombination suggests that the mechanisms of these two processes may be entirely different. This conclusion conforms to the observation that some viral polymerase mutations differently affect homologous and non-homologous recombination [21].

3. Recombination in the purified Q β system

Recently, several laboratories reported on the development of cell-free systems for studying RNA recombination [22–26], but only one paper describes a purified system in which the entire process, from mixing the parental strands and up to amplification of recombinant molecules, can be strictly controlled [23]. The reported system utilizes the pure replicase of RNA-containing bacteriophage Q β , a unique RNA-directed RNA polymerase capable to exponentially amplify RNA in vitro [27]. This ability has been employed in the molecular colony technique where single RNA molecules are grown as colonies in an agarose gel containing Q β replicase and rNTPs [28]. Like bacterial colonies, each RNA colony comprises the progeny of a single molecule, i.e. a clone [29].

Not any RNAs can be amplified by Q β replicase, but only those carrying special 5'- and 3'-terminal structures which form the replicase recognition site and whose interaction presumably provides for the single strandedness of the template and the nascent RNA, thereby ensuring many rounds of replication [27]. The reported system [23] used, as recombination substrates, the 5'- and 3'-fragments of RQ135⁻¹ RNA [30], a 134 nt long Q β phage satellite RNA [28] which had arisen by recombination between cellular and viral sequences [30]. The fragments were obtained by breaking the RNA at a natural cross-over site and, to facilitate a replicase switch between the fragments, extending their truncated ends with homologous foreign sequences (Fig. 1). To initiate recombination between the fragments, they were mixed and applied to the replicase-containing agarose. The fragments themselves are not replicable, but recombination between them may produce replicable RNAs resulting in RNA colonies whose number reflects the number of recombinant molecules generated in the system.

RNA colonies did appear, directly demonstrating that RNA molecules can recombine without any DNA intermediates: those simply could not form because there were no deoxyribonucleotides. Recombination occurred under conditions (pH, temperature, salt and Mg²⁺ concentrations) similar to those in the cell. The recombination frequency was propor-

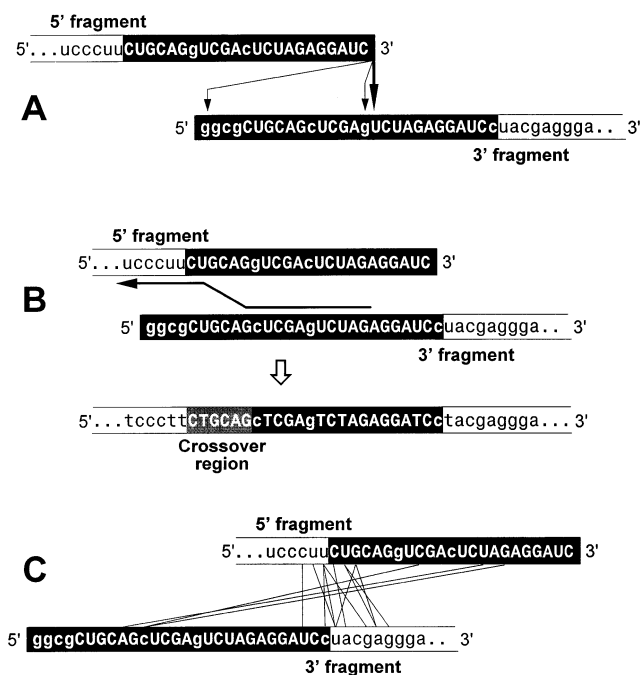


Fig. 1. Joining the parental sequences in the recombinants generated in vitro in the presence of Q β replicase and rNTPs (A), in the presence of AMV reverse transcriptase and dNTP (B) and in the absence of any enzyme and nt (C). Foreign extensions are shown in white letters, capital letters indicate the regions of homology.

tional to the product of the concentrations of the fragments. Extrapolation to the intracellular RNA concentration gives a value of 10^{-5} – 10^{-4} /nt [23], i.e. close to the frequency of non-homologous recombination in RNA phages in vivo [15].

Sequencing of the recombinants has shown that all of them are indeed non-homologous despite the homology of the foreign extensions (Fig. 1A). In contrast, only homologous recombinants were formed in a control experiment, in which the same RNA fragments were incubated in the presence of dNTPs and a reverse transcriptase (Fig. 1B), an enzyme capable of switching between templates [31], implying that the mechanism of RNA recombination in the presence of Q β replicase is different from template switch. The structure of recombinants suggested that they were generated by an attack of the 3'-terminus of the 5'-fragment at the phosphoester bonds within the 3'-fragment (Fig. 1A). Indeed, the recombination did not occur when the 3'-hydroxyl of the 5'-fragment (but not of the 3'-fragment) was eliminated and returned to the original level upon restoration of the hydroxyl (Fig. 2). In contrast to what is claimed in a recent review [8], these modifications did not introduce any bulky or otherwise unusual to RNA groups which could interfere with replicase binding (cf. Fig. 2). Moreover, the fact that similar modifications of a replicable RNA do not affect its replicability [23,32] eliminates the possibility [8] that Q β replicase requires the terminal 3'-hydroxyl for binding or RNA synthesis. Considered together, these results indicate that RNA recombination in the presence of Q β replicase occurs via a transesterification reaction, rather than via a replicative template switch.

The 3'-terminal modifications have recently been used as a tool for exploring the ability of RNAs to recombine in the absence of replicase [33]. In these experiments, a mixture of the RNA fragments was incubated in a salt buffer and, prior

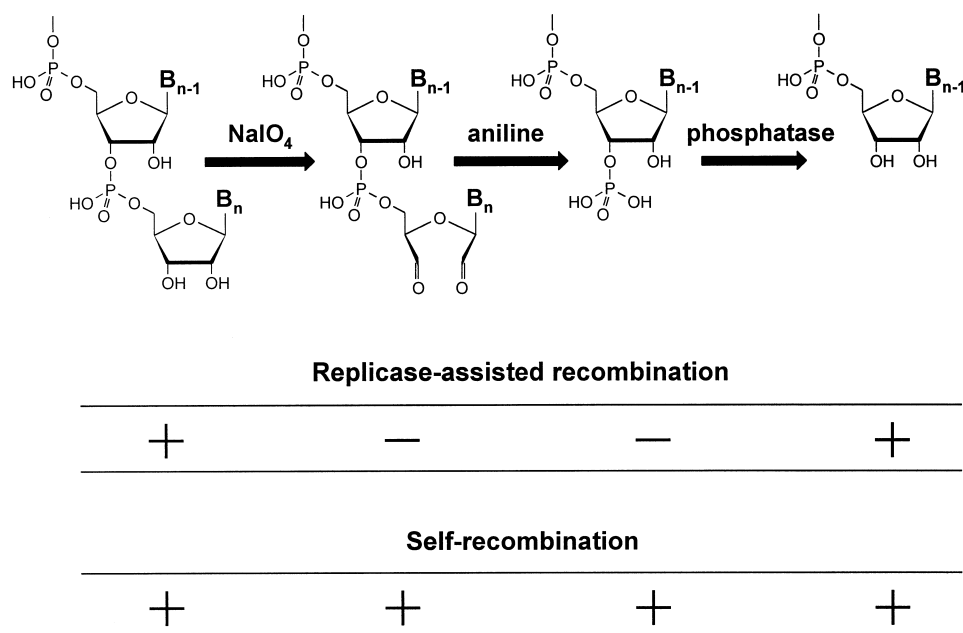


Fig. 2. Terminal modifications of the 5'-fragment and their effects on RNA recombination in the cell-free Q β system. Signs + and – indicate occurrence and absence of the recombination, respectively. Similar modifications of the 3'-fragment do not affect recombination.

to applying to the Q β replicase-containing agarose, it was melted (to destroy any non-covalent complexes) and oxidized with periodate (to prevent further recombination by the above mechanism). Thus, RNA colonies would only grow if recombination had occurred before the oxidation step. It has turned out that RNAs do recombine by themselves. The only requirement is the presence of Mg²⁺ ions at a millimolar concentration. However, such a self-recombination is entirely different from recombination in the presence of Q β replicase. In this case, the fragments recombine by internal segments (Fig. 1C), the 3'-hydroxyls do not participate (Fig. 2) and the reaction is several orders of magnitude slower (Table 1). Moreover, because the chemical step is slow, it becomes rate-limiting and the reaction follows pseudo-first order kinetics. Similar reactions can also occur in *cis*, resulting in deletions of internal RNA segments [33].

A conceivable mechanism of self-recombination includes an attack by 2'-hydroxyls, the attack being either intramolecular, leading to intermediate formation of the 2',3'-cyclic phosphate and 5'-hydroxyl termini which are then cross-ligated, or intermolecular, resulting in a branched structure which is

then copied by Q β replicase to produce the final recombinant molecule. Distribution of cross-over sites revealed no sequence specificity (Fig. 1C) and the recombination rate did not change when the extension sequences were altered [33], indicating that the ability of self-recombination is not associated with some cryptic ribozyme structures. It follows that self-recombination is a general property of RNA and, inasmuch as it requires nothing but RNA itself and Mg²⁺, it must be ubiquitous in nature and involve both viral and cellular RNAs.

Despite its low rate, RNA self-recombination might play an important evolutionary role both in the prebiotic RNA world and in the contemporary DNA world. At a rate of 10^{−9}/nt/h, a new recombinant RNA could arise in a human cell every minute, yielding up to 10²⁰ recombinant molecules during the life span of the human body. Reverse transcription and integration [34] of even a minute fraction of the rearranged sequences would change the genome significantly. Moreover, the reaction rate in the cell milieu might be much higher: cellular proteins might non-specifically enhance self-recombination as they promote the action of ribozymes [35,36].

Thus, the experiments in the cell-free Q β system have demonstrated the existence of at least two different non-replicative mechanisms of RNA recombination, one being performed by RNA itself and the other being somehow promoted by Q β replicase. The mechanism of the replicase action and whether other viral or cellular RNA polymerases can similarly promote reactions between RNA molecules remain to be determined.

4. Template switch versus the breaking and joining mechanism

The generation of a recombinant sequence can be conceived in two different ways: via breaking the parental sequences and joining the resulting fragments or via *de novo* synthesis by replicase which switches to another template after it has copied a portion of the first template.

Table 1
Properties of the two recombination mechanisms in the *in vitro* Q β system

	Self-recombination	Replicase-assisted recombination
Reaction kinetics	Pseudo-first order	Second order
Rate-limiting step	Chemical reaction	Non-covalent interactions
Overall rate ^a , nt ^{−1} h ^{−1}	10 ^{−9}	10 ^{−4} –10 ^{−5}
Nucleophilic attack	By 2'-OH	By 3'-OH
Reacting nt	Internal	3'-Terminal+internal

^aThe rates have been extrapolated to the intracellular phage RNA concentration. Their values can be directly compared with the recombination frequency in RNA phages whose infection cycle takes about 1 h.

The replicative template switch mechanism has been commonly accepted as the only possibility for generating recombinant RNAs since the work of Kirkegaard and Baltimore on homologous recombination in poliovirus [37], which is widely cited as the most strong evidence in favor of this mechanism. They crossed the wild-type virus (guanidine sensitive and temperature resistant) with a double mutant (guanidine resistant and temperature sensitive) and looked for a recombinant progeny that was both guanidine and temperature resistant. A peculiarity of their experiment was that viruses infected the cells in two consecutive steps, so that superinfection was performed under conditions inhibiting the replication of the first virus, the duration of the steps being adjusted to ensure equal yields of the parental viruses by the end of the experiment. An unexpected observation was that the yield of recombinants differed more than 100-fold depending on which restrictive factor was applied at superinfection. Specifically, the number of recombinants was high when the cells were superinfected with the wild-type virus at an elevated temperature and low when the cells were superinfected with the mutant virus in the presence of guanidine. According to the authors, such an asymmetric result could not be produced by a breaking and joining mechanism inasmuch as the parental genomes were equally abundant in each case, but was a consequence of asymmetric action of the restrictive factors on RNA synthesis. By referring to unpublished data, they claimed that while an elevated temperature inhibited the synthesis of only the negative strands of the mutant virus, the synthesis of both the negative and positive strands of the wild-type virus was suppressed in the presence of guanidine. Hence, as required by the template switch mechanism, recombination appeared to depend on the ongoing RNA synthesis.

However, these arguments do not seem to be strong enough. Since the negative strands comprise only 2% of the total viral RNA [38] and are presumably involved in double-stranded replicative intermediates [5], recombination by the breaking and joining mechanism may entirely depend on the availability of the positive strands. Yet, the positive strands synthesized at the first step may not be available during superinfection because of their packaging or involvement into polyribosomes. Hence, recombination by this mechanism may depend on whether the positive strands of each parental virus are synthesized during superinfection and will not occur in the presence of guanidine when synthesis of only the mutant positive strands is allowed. In addition, the possibility cannot be ruled out that the observed asymmetry resulted from a direct suppression of recombination by guanidine.

Thus, the template switch mechanism has never been unequivocally demonstrated for RNA recombination. Nevertheless, it may operate in some viruses, such as picorna- and coronaviruses that manifest an unusually high rate of homologous recombination. Switching between templates was directly demonstrated for retroviral reverse transcriptases [31,39] and for DNA polymerases during PCR [40]. In each of these cases, template switch is possible due to a particular mechanism for dissociating the nascent strand base-paired to its template: the RNA template degradation by the RNase H activity of reverse transcriptase and the heat-induced melting of DNA duplexes during PCR. Obviously, neither of these means is available for RNA viruses. However, some RNA viruses might have evolved special mechanisms to overcome the duplex problem, e.g. the use of ATP-fueled helicases.

While the involvement of the replicative template switch in RNA recombination has yet to be demonstrated, evidence for non-replicative mechanisms of the breaking and joining type is now available from the *in vitro* experiments in the Q β system. The question inevitably arises: to what extent are the conclusions made from *in vitro* data valid for the natural systems [8]? In this regard, recent experiments on recombination between non-replicable and non-translatable fragments of the poliovirus RNA [41] deserve special attention. Upon transfecting appropriate cells, the fragments did produce viable viral genomes, demonstrating that non-replicative recombination can occur *in vivo*. The results show that, similar to self-recombination, the 3'-hydroxyls do not participate in the reaction and suggest an involvement of the 2',3'-cyclic phosphate intermediates. The recombination is non-homologous and, importantly, its frequency is similar to the standard frequency of non-homologous RNA recombination *in vivo*. The results have also revealed recombinational 'hot spots' associated with ribozyme-like RNA structures. Apart from self-recombination, replicase- and ribozyme-assisted recombination, there may exist other non-replicative mechanisms for rearranging the RNA sequences. For example, a mechanism including RNase-catalyzed cleavage and ligation has been proposed [42].

Acknowledgements: I thank Helena Chetverina for helpful discussions. This work was supported by INTAS Grant 97-10348, Grant 99-04-48672 from the Russian Foundation for Basic Research and by an International Research Scholar's award from the Howard Hughes Medical Institute.

References

- [1] Hirst, G.K. (1962) Cold Spring Harbor Symp. Quant. Biol. 27, 303–309.
- [2] Ledinko, N. (1963) Virology 180, 107–119.
- [3] Lai, M.M.C. (1992) Microbiol. Rev. 56, 61–79.
- [4] Chetverin, A.B. (1997) Semin. Virol. 8, 121–129.
- [5] Agol, V.I. (1997) Semin. Virol. 8, 77–84.
- [6] Jarvis, T.C. and Kirkegaard, K. (1991) Trends Genet. 7, 186–191.
- [7] Wimmer, E., Hellen, C.U.T. and Cao, X. (1993) Annu. Rev. Genet. 27, 353–436.
- [8] Nagy, P.D. and Simon, A.E. (1997) Virology 235, 1–9.
- [9] Weiss, B.G. and Schlesinger, S. (1991) J. Virol. 65, 4017–4025.
- [10] Onodera, S., Qiao, X., Gottlieb, P., Strassman, J., Frilander, M. and Mindich, L. (1993) J. Virol. 67, 4914–4922.
- [11] Nagy, P.D. and Bujarski, J.J. (1993) Proc. Natl. Acad. Sci. USA 90, 6390–6394.
- [12] Hajjou, M., Hill, K., Subramaniam, S.V., Hu, J.Y. and Raju, R. (1996) J. Virol. 70, 5153–5164.
- [13] Kuge, S., Kawamura, N. and Nomoto, A. (1989) J. Virol. 63, 1069–1075.
- [14] Pilipenko, E.V., Gmyl, A.P. and Agol, V.I. (1995) Nucleic Acids Res. 23, 1870–1875.
- [15] Olsthoorn, R.C.L. and van Duin, J. (1996) J. Virol. 70, 729–736.
- [16] Holland, J.J. (1985) in: Virology (Fields, B.N., Ed.), pp. 77–99, Raven Press, New York.
- [17] Schlesinger, S. (1988) in: RNA Genetics (Domingo, E., Holland, J.J. and Ahlquist, P., Eds.), Vol. II, pp. 167–185, CRC Press, Boca Raton, FL.
- [18] Henle, W. and Henle, G. (1943) Science 98, 87–89.
- [19] Palasingam, K. and Shaklee, P.N. (1992) J. Virol. 66, 2435–2442.
- [20] Cole, C.N., Smoler, D., Wimmer, E. and Baltimore, D. (1971) J. Virol. 7, 478–485.
- [21] Figlerowicz, M., Nagy, P.D. and Bujarski, J.J. (1997) Proc. Natl. Acad. Sci. USA 94, 2073–2078.
- [22] Qiao, X., Qiao, J. and Mindich, L. (1997) Virology 227, 103–110.
- [23] Chetverin, A.B., Chetverina, H.V., Demidenko, A.A. and Ugarov, V.I. (1997) Cell 88, 503–513.

- [24] Tang, R.S., Barton, D.J., Flanagan, J.B. and Kirkegaard, K. (1997) *RNA* 3, 624–633.
- [25] Duggal, R., Cuconati, A., Gromeier, M. and Wimmer, E. (1997) *Proc. Natl. Acad. Sci. USA* 94, 13786–13791.
- [26] Nagy, P.D., Zhang, C. and Simon, A.E. (1998) *EMBO J.* 17, 2392–2403.
- [27] Chetverin, A.B. and Spirin, A.S. (1995) *Prog. Nucleic Acid Res. Mol. Biol.* 51, 225–270.
- [28] Chetverin, A.B., Chetverina, H.V. and Munishkin, A.V. (1991) *J. Mol. Biol.* 222, 3–9.
- [29] Chetverina, H.V. and Chetverin, A.B. (1993) *Nucleic Acids Res.* 21, 2349–2353.
- [30] Munishkin, A.V., Voronin, L.A., Ugarov, V.I., Bondareva, L.A., Chetverina, H.V. and Chetverin, A.B. (1991) *J. Mol. Biol.* 221, 463–472.
- [31] Negroni, M., Ricchetti, M., Nouvel, P. and Buc, H. (1995) *Proc. Natl. Acad. Sci. USA* 92, 6971–6975.
- [32] Rensing, U. and August, J.T. (1969) *Nature* 224, 853–856.
- [33] Chetverina, H.V., Demidenko, A.A., Ugarov, V.I. and Chetverin, A.B. (1999) *FEBS Lett.* 450, 89–94.
- [34] Brosius, J. and Tiedge, H. (1995) *Virus Genes* 11, 163–179.
- [35] Herschlag, D., Khosla, M., Tsuchihashi, Z. and Karpel, R.L. (1994) *EMBO J.* 13, 2913–2924.
- [36] Coetzee, T., Herschlag, D. and Belfort, M. (1994) *Genes Dev.* 8, 1575–1588.
- [37] Kirkegaard, K. and Baltimore, D. (1986) *Cell* 47, 433–443.
- [38] Novak, J.E. and Kirkegaard, K. (1991) *J. Virol.* 65, 3384–3387.
- [39] Diaz, L. and DeStefano, J.J. (1996) *Nucleic Acids Res.* 24, 3086–3092.
- [40] Frohman, M.A. and Martin, G.R. (1990) in: *PCR Protocols* (Innis, M.A., Ed.), pp. 228–236, Academic Press, New York.
- [41] Gmyl, A.P., Belousov, E.V., Maslova, S.V., Khitrina, E.V., Chetverin, A.B. and Agol, V.I. (1999) *J. Virol.* 73(11), in press.
- [42] Tsagris, M., Tabler, M. and Singer, H.L. (1991) *Nucleic Acids Res.* 19, 1605–1612.