Propensity of a picornavirus polymerase to slip on potyvirus-derived transcriptional slippage sites

Hazel Stewart, Allan Olspert, Benjamin G. Butt and Andrew E. Firth

Abstract

The substitution rates of viral polymerases have been studied extensively. However, less is known about the tendency of these enzymes to ‘slip’ during RNA synthesis to produce progeny RNAs with nucleotide insertions or deletions. We recently described the functional utilization of programmed polymerase slippage in the family Potyviridae. This slippage results in either an insertion or a substitution, depending on whether the RNA duplex realigns following the insertion. In this study we investigated whether this phenomenon is a conserved feature of superfamily I viral RdRps, by inserting a range of potyvirus-derived slip-prone sequences into a picornavirus, Theiler’s murine encephalomyelitis virus (TMEV). Deep-sequencing analysis of viral transcripts indicates that the TMEV polymerase ‘slips’ at the sequences U_{n-7} and A_{n-7} to insert additional nucleotides. Such sequences are under-represented within picornaviral genomes, suggesting that slip-prone sequences create a fitness cost. Nonetheless, the TMEV insertional and substitutional spectrum differed from that previously determined for the potyvirus polymerase.

In members of the negative-sense RNA virus families Paramyxoviridae, Rhabdoviridae and Orthomyxoviridae, the RNA-dependent RNA polymerase (RdRp) ‘stutters’ on short poly(U) tracts to polyadenylate the mRNAs [1]. Polymerase stuttering on poly(A) and poly(U) templates is also thought to maintain poly(A) tail length in picornaviruses [2, 3]. Polymerase slippage or slippage can also occur within coding sequences to produce populations of transcripts with altered coding capacity, where nucleotide insertions or deletions allow access to alternative open reading frames. Where subject to purifying selection, this is known as ‘programmed polymerase slippage’ or ‘programmed transcriptional slippage’. Negative-sense RNA viruses in the taxa Ebola virus and Paramyxoviridae have long been known to use polymerase slippage for gene expression [4, 5]. More recently, polymerase slippage was identified in the Potyviridae, the largest family of positive-sense RNA viruses of plants [6, 7]. In these viruses, polymerase slippage occurs at a highly conserved GAAAAAA (GA_6) sequence, giving rise to transcripts with an additional adenylate. Translation of these transcripts allows expression of an essential ‘transframe’-encoded protein, P3N-PIPO. In a subgroup of potyviruses, slippage also occurs at a second GA_6 site, enabling expression of the ‘transframe’-encoded protein, P1N-PISPO [8].

In potyviruses, the single nucleotide insertion rate varies between 0.8 and 2% (P3N-PIPO expression) [6–9] and 5 and 12% (P1N-PISPO expression) [7, 8], presumably influenced by flanking nucleotides. However, a significant number of substitution events are also observed, revealing a mechanism that has been termed ‘to–fro’ slippage, whereby the RdRp is hypothesized to undergo a ‘slip back, template, slip forward’ movement [10]. RNA duplex realignment following templating of the inserted nucleotide leads to the subsequent template nucleotide being ‘skipped’. The resulting transcript maintains the original protein-coding reading frame and length, but possesses a nucleotide substitution at the +7 position, 3’-adjacent to the GA_6 slip site. When the reverse complement of the slippage site is present, substitutions are observed instead at the –1 position. The position of these substitutions allows determination of whether ‘to–fro’ slippage occurs during positive- or negative-sense synthesis. For the potyvirus polymerase, ‘to–fro’ slippage occurs mainly during synthesis of poly(A) rather than poly(U) regardless of the orientation of the GA_6 sequence [10].

Given the evolutionary relatedness of the Picornaviridae and Potyviridae RdRps [11], it is reasonable to envision similar behaviour in both families. Consistent with this idea, bioinformatic analysis of both potyviral and picornaviral genomes revealed under-representation of A_n and U_n (n≥6).
homopolymeric sequences when the functionally utilized potyviral slippage sites were excluded [6] (Fig. 1). Since deleterious effects of Aₙ and Uₙ sequences might also occur at the translational level as a result of ribosomal slippage, the picornavirus analysis was performed in all three reading frames. Selection against Aₙ and Uₙ sequences may reflect a propensity of the RdRp to slip at such sites, leading to a reduction in virus productivity from packaging of the defective transcripts. Slippage events may potentially also lead to more serious in trans antiviral effects, such as dominant negative interference by truncated versions of viral proteins and potential preferential MHC class I antigen presentation of slippage products [12–15].

In this study, we wished to investigate to what extent potyviral slippage sites lead to polymerase slippage in Theiler’s murine encephalomyelitis virus (TMEV), a model picornavirus in the genus Cardiovirus. As with other picornaviruses, TMEV has a polyadenylated positive-sense RNA genome of ~8 kB that encodes a polyprotein which is processed to produce the structural and non-structural viral proteins (Fig. 2a). The 5′ untranslated region (UTR) of ~1 kB contains an internal ribosome entry site. We used an infectious clone with sequence identical to GenBank Accession number X56019.1 except for three nucleotide differences, G2241A, A2390G and G4437A [16]. The wild-type (WT) sequence contains one A₆ tract (in the region encoding 3C), one U₆ tract (in the 5′ UTR) and no A₇ or U₇ tracts.

We inserted candidate polymerase slippage sites into the coding region, rather than the UTRs, so that insertions or deletions would lead to defective genomes that could not amplify without a helper virus. Indeed, as translation beyond the 2A region is required in cis for replication (at least for the related poliovirus) [17], such genomes would not be expected to replicate even in the presence of helper virus. To avoid altering the native viral proteins, we duplicated 24 amino acids of the 2A StopGo sequence via overlap

![Fig. 1. Poly(A) and poly(U) sequences are under-represented in picornavirus genomes. An analysis of 131 Picornaviridae NCBI RefSeqs indicates strong selection against A₇, U₇, A₆ and U₆ sequences. A₅, poly(C), poly(G) and several arbitrary heptanucleotides are included for comparison. Red bars indicate the mean observed frequency per polyprotein ORF of the indicated sequences. Polypeptide ORFs were also randomly shuffled 1000 times while maintaining amino acid sequence and codon bias (as previously described [6] except that here each of the three reading frames was analysed separately). Purple bars indicate mean frequencies per polyprotein ORF in the shuffled sequences. Error bars indicate standard deviations. Values above bars indicate two-tailed z-test p-values showing that the observed counts are statistically different from the expected counts.](image-url)
PCR (TMEV-2SG; Fig. 2a) so that candidate slip-prone sequences could be inserted into restriction sites incorporated between the two StopGos. Translation of the StopGo sequence results in a peptide ending in NPGP that
mediates co-translational polypeptide separation by preventing peptide bond formation between the glycine and final proline. Thus the inserted sequences would be co-translationally excised from the polypeptide with no effect on the amino acid sequences of the flanking 2A and 2B proteins.

Slippage mutant viruses were generated by the ligation of dsDNA linkers into a digested pTMEV-2SG backbone. Clones were designed based upon two potyviral slip-prone mutant viruses were generated by the ligation of on the amino acid sequences of the flanking 2A and 2B translationally excised from the polyprotein with no effect mediating co-translational polypeptide separation by preventing peptide bond formation between the glycine and final proline. Thus the inserted sequences would be co-translationally excised from the polypeptide with no effect on the amino acid sequences of the flanking 2A and 2B proteins.

Slippage mutant viruses were generated by the ligation of dsDNA linkers into a digested pTMEV-2SG backbone. Clones were designed based upon two potyviral slip-prone mutant viruses were generated by the ligation of on the amino acid sequences of the flanking 2A and 2B translationally excised from the polyprotein with no effect mediating co-translational polypeptide separation by pre-...
mixture of wild-type and slippage transcripts that may be packaged with equal efficiency as they emerge from the vesicle. For the other three mutants (GA6 and GA7 in the positive-sense), the ‘Virus 1’, ‘Cell I’ and ‘Virus 2’ samples had decreased levels of insertions compared to the input, indicating that the viral RdRp has a lower tendency to slip on these sequences than the T7 polymerase. Since slippage may also occur during library preparation, these values – which ranged through 0.50±0.12 % (TuMV WT), 1.1±0.16 % (PISPO WT) and 1.6±0.24 % (TuMV+A) – should be considered as upper bounds on the viral RdRp slippage rates.

We also quantified deletional slippage (Fig. 3a, second panel). TuMV+A RC and PISPO+A RC both exhibited high levels of presumably T7-derived deletional slippage (4.8±0.71 and 6.5±1.0 %, respectively); significant rates of slippage on similar U₇ tracts (n=6) by T7 polymerase have been noted previously and are an essential component of efficient transcription termination [18, 21]. This was purged to levels of 0.35±0.15 and 0.23±0.19 % following virus replication. The TuMV+A mutant had similar slippage both in the input (1.8±0.11 %) and following virus replication (1.5±0.26 %), suggesting that both the T7 and viral polymerases had similar deletional slippage propensities at this site. Only low levels of deletional slippage (upper bounds <0.14 %) were observed during replication of GA6 or U₇C slip-site viruses.

Following Olspert et al. [10], we inspected reads for evidence of ‘to-fro’ slippage – that is, a substitution to A or U immediately following or immediately preceding an Aₙ or Uₙ slipp site, respectively (red bars, Fig. 3). Whereas virus infection at low MOI would be expected to purge insertion/deletion mutations, substitutions within the inter-StopGo insert region would likely not be subject to strong selective pressure. Thus substitutions introduced by the T7 polymerase, or during the course of virus growth, would likely be retained and propagated.
At position +7 (i.e. 3'-adjacent to the slip site) of the U₆₋₇C slip sites, there appeared to be a component of C-to-U substitutions (Fig. 3a, third panel, RC mutants, red bars) which was clearly not derived from the input RNA. In contrast, levels of C to not-U substitutions (grey bars) were similar between virus-derived and input RNA. These results indicate that 'to-fro' slippage by the viral RdRp occurs at U₆₋₇C slip sites during positive-sense synthesis. In contrast, for the GA₆₋₇ slip sites, levels of 'to A' substitutions at position +7 were similar between virus-derived and input RNA, indicating that 'to-fro' slippage does not occur to appreciable levels during positive-sense synthesis at GA₆₋₇ slip sites (Fig. 3, 3rd panel, non-RC mutants, red bars). At the −1 position (i.e. 5'-adjacent to the slip site) differences between input and virus-derived RNA were less striking, although the four TuMV mutant viruses showed a possible increase in G to A (TuMV WT, TuMV+A) and C to U (TuMV RC, TuMV+A RC) substitutions compared to input (Fig. 3, 4th panel, red bars).

Due to the background of spurious mutations, Olspert et al. focused on sites with potyvirus RdRp substitution levels ≥0.5% [10]. For the TMEV RdRp, we only observed virus-specific substitutions approaching this level for the TuMV+A RC site (Fig. 3a, third panel). For the TuMV WT slip site (GA₆₋₇), G-to-A substitutions at position +7 were 36-fold lower than previously observed with the potyvirus RdRp (0.015±0.003 versus 0.54±0.07%; and for the TuMV RC slip site (U₆₋₇), C-to-U substitutions at position −1 were 33-fold lower than with the potyvirus RdRp (0.065±0.014 versus 2.2±0.5%) [10].

To summarize, our data indicate that the TMEV RdRp permits insensational slippage at levels of 0.46–4.3% on GA₆₋₇ and U₆₋₇ sequences, with higher levels of slippage occurring when poly(U) is present in the positive strand. The highest levels of 'to-fro' slippage occur at position +7 for the TuMV+A RC slip site, suggesting that slippage on U₆₋₇ sequences may occur predominantly during positive-sense synthesis. Picornaviral RdRps are thought to use RdRp stuttering to maintain genomic poly(A) tail length during replication [2], where the positive-sense sequence is poly(A) rather than poly(U), though slippage may occur during synthesis of either strand, contributing to the final poly(A) tail length. Surrounding RNA structures may contribute to the efficiency of this event; for example, a cis-element of enteroviruses (located within the 3' UTR) is thought to enhance polymerase slippage on the negative-sense poly(U) template to facilitate polyadenylation of the positive-sense viral genome during replication [22]. It is therefore likely that slippage propensity will differ between sites as a result of flanking sequences and/or homopolymer length; however, our study was specifically aimed at potyviral-like slip sites.

It remains possible that picornavirus polymerases may have evolved an increased propensity for slippage in specific genera or species, or that polymerase slippage may be utilized on specific sequences that differ from the potyvirus-derived sequences tested herein. For example, encephalomyocarditis virus (also in the Cardiovirus genus) has a long poly(C) tract in its 5' UTR that is associated with heightened virulence [23]. However there is no direct evidence that polymerase slippage is used during evolution of the length of this tract (a possible alternative is recombination). Our bioinformatic analysis did not reveal any significant under-representation of poly(C) tracts in picornavirus coding regions which suggests that – at least for short, N₆₋₇ tracts – they are less prone to spurious slippage events than poly(A) or poly(U).

Although both potyviruses and cardioviruses possess superfamily I RdRps [11], the tendency of each RdRp to slip upon particular nucleotide sequences appears distinct, with the potyvirus RdRp preferentially slipping during synthesis of GA₆₋₇ regardless of sense. Whereas potyviruses utilize polymerase slippage to access novel ORFs, there are no known cases of this occurring in picornaviruses. As mentioned above, non-programmed polymerase slippage within coding sequences results in defective transcripts which may lead to various negative effects [12–15], and these factors may contribute to the significant under-representation of U₆₋₇ and GA₆₋₇ sequences within picornaviral genomes. This research contributes to our understanding of the mechanisms that shape RNA virus genomic diversity and highlights differences between related viral polymerases, where the potyvirus RdRp may have co-evolved with the expression of the essential P3N-PIPO protein to be specifically tuned to facilitate slippage on the GA₆₋₇ slip sites that potyviruses functionally use for gene expression.

Funding information

This work was supported by a Wellcome Trust grant [106207] and a European Research Council grant [646891] to A.E.F, and an Estonian Research Council grant [PUT1699] and Tallinn University of Technology starting grant [B60] to A.O.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

3. Steil BP, Kempf BJ, Barton DJ. Poly(A) at the 3' end of positive-strand RNA and Vpg-linked poly(U) at the 5' end of negative-strand RNA are reciprocal templates during replication of poliovirus RNA. J Virol 2010;84:2843–2858.
7. Rodamians B, Valli A, Mingot A, San Leon D, Baulcombe D et al. RNA polymerase slippage as a mechanism for the production of


---

**Five reasons to publish your next article with a Microbiology Society journal**

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.