In Vitro Characterization of Late Steps of RNA Recombination in Turnip Crinkle Virus

II. The Role of the Priming Stem and Flanking Sequences

Peter D. Nagy and Anne E. Simon[1]

Department of Biochemistry and Molecular Biology, Program in Molecular and Cellular Biology,
University of Massachusetts, Amherst, Massachusetts 01003

Turnip crinkle carmovirus (TCV) has a uniquely high recombination frequency and nonrandom cross-over site distribution among the recombining TCV-associated satellite RNAs. An *in vitro* system has been developed that includes a partially purified TCV replicase preparation (RdRp) and chimeric RNAs that resemble the putative *in vivo* recombination intermediates (Nagy, P. D., Zhang, C., and Simon, A. E., *EMBO J.* 17, 2392–2403, 1998). This system mimics the strand transfer and primer extension steps of recombination events. We characterize in detail three RNA factors that, in addition to the previously characterized motif1-hairpin, can influence the efficient generation of 3′-terminal extension products: (i) a primer binding region, termed the priming stem; (ii) a spacer region; and (iii) a U-rich sequence located 5′ of the motif1-hairpin. The priming stem is formed between the acceptor RNA and the nascent RNA synthesized from the donor RNA template in the recombinants. The stability and location of the priming stem relative to the motif1-hairpin can influence both the efficiency and initiation site of 3′-terminal extension. A short flexible spacer region between the motif1-hairpin and the priming stem can increase the efficiency of 3′-terminal extensions. A U-rich sequence 5′ of the motif1-hairpin facilitates 3′-terminal extensions and its function partly overlaps with that of the spacer region. These RNA factors may also affect the late steps of RNA recombination in TCV. © 1998 Academic Press

Key Words: plant virus; recombination; replication; satellite RNA; template switching.

INTRODUCTION

Genetic RNA recombination can create novel chimeric viruses, repair genomes from mutated (damaged) RNAs, and possibly contribute to the quasispecies nature of RNA viruses (Zimmern, 1988; Lai, 1992; Nagy and Bujarski, 1996; Simon and Nagy, 1996). Mechanistic studies on RNA recombination favor replicase (RNA-dependent RNA polymerase, RdRp)-driven template switching models (King, 1988; Jarvis and Kirkegaard, 1991; Lai, 1992; Nagy and Simon, 1997). These models suggest that recombination events can be divided into three steps: (1) synthesis of the primer on the donor RNA; (2) strand transfer and the binding of the RdRp to the acceptor RNA; and (3) primer elongation on the acceptor RNA (Nagy and Simon, 1997; Nagy *et al.*, 1998). These steps are being studied separately in an *in vitro* system in turnip crinkle virus (TCV) that does not require the amplification of recombination products (Nagy *et al.*, 1998).

TCV is a 4054-base RNA virus that is associated with a number of satellite RNAs (e.g., sat-RNA D, 194 nt, and sat-RNA C, 356 nt) (Simon and Nagy, 1996). Sat-RNA C, which is composed of sequences similar to sat-RNA D at its 5′ end and two 3′ proximal regions from TCV genomic

RNA at its 3′ end, is the product of a natural recombination event (Simon and Howell, 1986). In addition to the naturally occurring recombinant species associated with TCV infections, *de novo* generated recombinants have also been detected. Plants inoculated with TCV genomic RNA, sat-RNA D, and sat-RNA C, which contain deletions in the 5′ region, accumulated a heterogeneous population of recombinant molecules comprising sat-RNA D at the 5′ end and a portion of sat-RNA C at the 3′ end (Cascone *et al.* 1990, 1993). The junction between sat-RNA D and sat-RNA C occurred mainly at position 181 of sat-RNA D (13 bases from the 3′ end of the plus-strand) and one of five consecutive bases in sat-RNA C beginning at position 175 (Cascone *et al.*, 1990, 1993). Just downstream from the recombination site in sat-RNA C minus-strands is a stable hairpin (motif1-hairpin), which was required for recombination (Cascone *et al.*, 1993). In addition to the sat-RNAD/sat-RNA C system, a population of *de novo* recombinants was also detected between sat-RNA D and the TCV genomic RNA. These recombinants were composed (from 5′ to 3′) mainly of full-length sat-RNA D or sat-RNA D truncated by 13 bases joined to variable lengths of TCV 3′-end sequence (Zhang *et al.*, 1991; Carpenter *et al.*, 1995). The location of junction sites in sat-RNA D/TCV recombinants was influenced by the presence of a stable hairpin (motif3-hairpin) in minus-strands of the genomic RNA (Carpenter *et al.*, 1995).

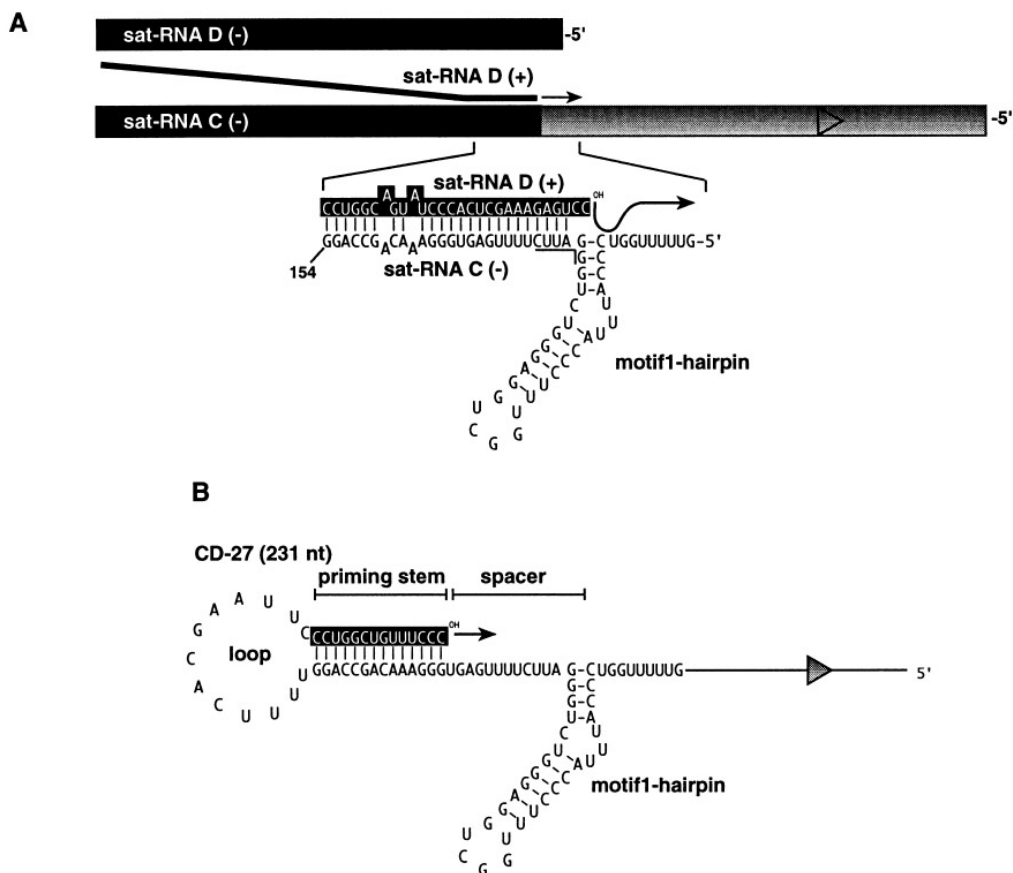[1] To whom reprint requests should be addressed.

FIG. 1. An *in vitro* system to study the late steps of RNA recombination in TCV. (A) Model for *in vivo* RNA recombination between sat-RNA D and sat-RNA C by a replicase-driven template switching mechanism. The recombination intermediates are depicted showing the possible RNA–RNA interaction between the nascent sat-RNA D plus strand, truncated at position −13 at the 3′ end (in *in vivo* experiments, the junctions were located most frequently at positions −13 to −15, as counted from the 3′ end), and sequence 3′ of the motif1-hairpin of sat-RNA C minus strand (the acceptor RNA strand). Sat-RNA D sequence is boxed in black. Underlined nts at the base of the motif1-hairpin indicate the location of the *in vivo* junction site hot-spot in sat-RNA C. In the rectangles representing the sat-RNAs, similar regions are shaded alike. A triangle represents a positive-strand initiation promoter (Guan *et al.,* 1997). (B) An *in vitro* 3′-terminal extension system. Construct CD-27 contains 14 nts from sat-RNA D plus strand (boxed in black), extending from position −27 at the 3′ end, joined by a six-base artificial sequence (GAAUUC) to sat-RNA C minus-strand sequence that includes 33 nt 3′ of the motif1-hairpin, the motif1-hairpin, and the entire sat-RNA C sequence 5′ of the hairpin. Regions referred to in the text as the priming stem, motif1-hairpin, loop, and spacer are shown. Triangle represents the promoter sequence shown in (A).

RNA determinants of RNA–RNA recombination between TCV-associated sat-RNA C and sat-RNA D have been characterized *in vivo* (Cascone *et al.,* 1990, 1993; Nagy *et al.,* 1998). These studies led to a model that suggested that RNA recombinants between sat-RNA D and sat-RNA C are formed by a replicase-mediated strand transfer mechanism where the nascent strand, made on the negative-stranded donor sat-RNA D by the replicase, serves as a primer during reinitiation of RNA synthesis on the negative-stranded sat-RNA C (Fig. 1A). The strand transfer events are guided by the motif1-hairpin on the acceptor sat-RNA C, which is a putative RdRp binding sequence. A simplified two-component *in vitro* system has been developed that uses chimeric RNAs resembling the putative *in vivo* recombination intermediates and a purified TCV RdRp preparation (Nagy *et al.,* 1998; Fig. 1B). The *in vitro* system results in a 3′-terminal extension reaction that is analogous to the

proposed primer extension reaction that occurs *in vivo.* The use of an *in vitro* system made possible the identification of two RNA elements, the motif1-hairpin and the priming stem (the primer landing site on the acceptor RNA), which greatly facilitate the 3′-terminal extension reactions. Competition experiments suggested that the RdRp interacts with the motif1-hairpin, thus the hairpin possibly facilitates the recruitment of the RdRp (Nagy *et al.,* 1998).

Based on replicase-mediated template-switching models, recombination events involve the donor RNA, the nascent RNA that is usually incomplete copy of the donor RNA, the acceptor RNA, and the viral RdRp (Lai, 1992; Nagy and Simon, 1997; Nagy *et al.,* 1998). Depending on the template RNAs, RNA recombination events can be divided into early and late steps. The template during the early steps is the donor RNA, while the template during the late steps is the acceptor RNA. In mechanistic terms,

RNA recombination consists of three steps: (1) synthesis of the nascent primer strand from the donor RNA; (2) strand transfer and binding of the RdRp to the acceptor RNA; and (3) primer elongation on the acceptor RNA (Nagy and Simon, 1997; Nagy et al., 1998).

We have developed an in vitro system in TCV that includes the putative recombination intermediates between sat-RNA C and sat-RNA D and mimics the late steps of the in vivo strand transfer events. This in vitro system was used to further define RNA requirements of 3′-terminal extension reactions. In this report, we extend our previous characterization of the priming stem by showing that its sequence can influence the site of initiation of 3′-terminal extension, while the stability of the priming stem can affect the efficiency of 3′-terminal extension. In addition, a role for pyrimidine-rich sequences flanking the motif1-hairpin in 3′-terminal extension has been elucidated. The pyrimidine-rich sequences may influence RdRp binding directly, or as single-stranded spacers, they may facilitate exposure of the motif1-hairpin structure for binding to the RdRp.

RESULTS AND DISCUSSION

The in vitro system used to study RNA factors involved in 3′-terminal extension

An in vitro system has been developed that mimics RNA recombination between sat-RNA C and sat-RNA D, including the strand transfer and primer elongation steps (Nagy et al., 1998). This system makes use of a partially purified, template-dependent TCV RdRp preparation (Song and Simon, 1994) programmed with a chimeric RNA template that supports efficient 3′-terminal extension (Nagy et al., 1998). This chimeric RNA template (CD-27, 211 nt) consisted of sequences corresponding to the central and 5′ portions of the negative strand of sat-RNA C (including the in vivo recombination hot-spot region and the motif1-hairpin) and a 14-nt segment of the positive-stranded sat-RNA D (boxed in black, Fig. 1B). A 6-nt artificial loop sequence was used to covalently link the sat-RNA D plus-strand fragment to the sat-RNA C minus-strand fragment, as depicted in Fig. 1B, to ensure that an optimal 1:1 molar ratio of interacting sequences were in proximity and in the correct orientation. Sequence complementarity between the 3′-end region of sat-RNA D positive strand and a region 3′ of the motif1-hairpin in the negative strand of sat-RNA C (Fig. 1B) is predicted to form a stable base-paired structure (priming stem). The TCV RdRp starts 3′-terminal extension 12 nt 3′ of the motif1-hairpin in the hybrid construct (Fig. 1B). This "self-primed" 3′-terminal extension reaction generates a product that has a hairpin-like structure with a long double-stranded region and short loop sequence. Overall, the chimeric construct CD-27 resembles the putative in vivo recombination intermediates (Fig. 1A) except that it contains a 13-nt shorter sequence at the 3′ end derived
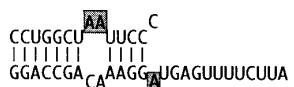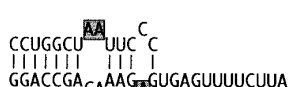
from positive-stranded sat-RNA D. This short priming stem supports 3′-terminal extension about sixfold more efficiently than the longer wt priming stem (Nagy et al., 1998), and therefore was used to characterize the role of mutations in the priming stem on 3′-terminal extension.

The priming stem influences the efficiency and initiation site of 3′-terminal extensions

Previous experiments established that base-pairing within the priming stem plays an important role in 3′-terminal extension reactions (Nagy et al., 1998). Viral and artificial sequences that formed 10 bp or more in the priming stem supported 3′-terminal extension above background level. To further investigate the role of the priming stem in 3′-terminal extension reactions, 6-nt-long portions of the priming stem were inverted from one strand to the other in CD-27. In addition, a single point mutation was introduced during the cloning procedure. The resulting co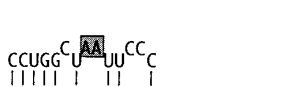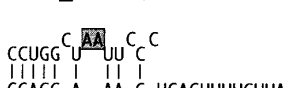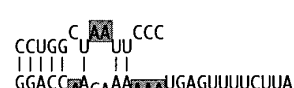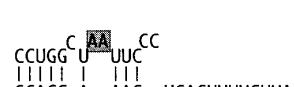nstruct, CD-27/rev (Fig. 2), was tested in a TCV RdRp assay followed by denaturing urea/PAGE. Quantitation of the 3′-terminal extension products by densitometric scanning of exposed films was used to measure the amount of 3′-terminal extension products. Construct CD-27/rev supported 3′-terminal extension as efficiently as wt CD-27 (Fig. 2), confirming our previous observation that the primary sequence of the priming stem is not recognized by the TCV RdRp (Nagy et al., 1998).

The lack of specific sequence requirements in the priming stem permitted the characterization of base-pairing requirements during 3′-terminal extension reactions. This was achieved by introducing mismatch mutations into the 14-bp-long priming stem in CD-27. One, two, and three mutations in the AU rich portion of the priming stem (constructs CD-27/A1, CD-27/A2, CD-27/A3) only slightly decreased the efficiency of 3′-terminal extension reactions (79.0–98.5% activity as compared to the wt CD-27, Fig. 2) without influencing the site of initiation (data not shown). Four mismatch mutations in the same portion of the priming stem resulted in 51% of the level of wt 3′-terminal extension products (construct CD-27/A4; Fig. 2).

Constructs with two mismatch mutations within the AU-rich portion and an additional single mutation that destabilized any of the three 3′-terminal G–C base pairs of the priming stem had less than a 30% reduction in 3′-terminal extension (constructs CD-27/G14, CD-27/G22, CD-27/G23; Fig. 2). This result with CD-27/G14 suggests that base-pairing of the terminal nucleotide is not required for 3′-terminal extension. In contrast, three mismatch mutations that destabilized the 5′ GC-rich portion of the priming stem reduced 3′-terminal extension by almost 75% (construct CD-27/G21; Fig. 2). This demonstrates that the upstream GC-rich region contributes significantly to the stability of the priming stem and consequently the efficiency of 3′-terminal extension. More extensive modifications with five to six mismatch mutations

	Amount of 3'-TX		
	A-site	B-site	C-site
CD-27	100%	<1	<1
CD-27/rev	93.1+12.3	<1	<1
CD-27/A1	98.5+11.1	<1	<1
CD-27/A2	87.0+9.9	<1	<1
CD-27/A3	79.0+13.2	<1	<1
CD-27/A4	51.0+16.2	<1	<1
CD-27/G14	70.9+10.1	<1	<1
CD-27/G22	75.7+9.7	<1	<1
CD-27/G23	82.5+9.3	<1	<1
CD-27/G21	25.8+3.3	<1	<1
CD-27/G11	<1	40.3+11.5	<1
CD-27/G7	<1	7.3+1.8	<1
CD-27/G6	<1	4.7+1.3	3.9+0.9
CD-27/G12	<1	4.5+0.8	3.9+0.6

over the entire priming stem reduced 3′-terminal extension that initiated at the wt position (site A shown in Fig. 3A) by more than 99% (constructs CD-27/11, CD-27/G7, CD-27/G6, and CD-27/G12; Fig. 2).

Curiously, these constructs gave various amounts of 3′-terminal extension products with altered electrophoretic mobilities when compared to CD-27 (Fig. 3B). To characterize these unusual 3′-terminal extension products, all 3′-terminal extension products from the constructs shown in Fig. 2 were treated with single-strand specific S1 nuclease and RNase H in the presence of a deoxyoligonucleotide complementary to the loop sequence of CD-27. A representative denaturing PAGE analysis of the nuclease treated 3′-terminal extension products is shown in Fig. 3B. This analysis led to the identification of two unusual 3′-terminal extension products (indicated by crosses and arrows in Fig 3B; see legend to Fig. 3B for detailed explanation about characterization of the 3′-terminal extension products) for constructs CD-27/G6 and CD-27/G12. Based on the size estimation of the most abundant RNase H and S1 nuclease resistant fragments for these 3′-terminal extension products from denaturing PAGE gel analyses (Figs. 3B and 3C), initiation of 3′-terminal extension was estimated to occur at two novel sites (Fig. 3A). For constructs CD-27/G6 and CD-27/G12, initiation of 3′-terminal extension was estimated to occur within the motif1-hairpin (site B, Fig. 3A) for products that migrated slightly slower (indicated by arrows in Fig. 3B) than the wt CD-27 3′-terminal extension product in denaturing PAGE. The very slow migrating 3′-terminal extension products of constructs CD-27/G6, and CD-27/G12 (shown by crosses in Fig. 3C) initiated from a position approximately 65 nt from the 5′ end. Constructs CD-27/G11 (Fig. 3B) and CD-27/G7 (Fig. 2) generated 3′-terminal extension products that initiated only at the B site with 40.3 and 7.3% efficiency of wt CD-27. We propose that short sequence complementarity between the 3′ ends and these internal sites may have facilitated 3′-terminal extension from sites B and C (Fig. 3A). Overall, these data suggest that base-pairing within the priming stem can influence both the initiation site and the efficiency of 3′-terminal extension in this system. In addition, the above mismatch mutations revealed that the GC-rich portions at the 3′ terminus of the priming stem or the 5′ portion of the priming stem influences 3′-terminal extension to the highest extent. In contrast, the role of base-pairing within the AU-rich portion of the priming stem was less important in 3′-terminal extension.

The finding that 5–10 bp in the priming stem was sufficient to support higher than background level of 3′-terminal extension may be relevant in other systems as well. For example, the formation of short base-paired regions can not only occur between sat-RNA D and sat-RNA C, but it is also possible in the other characterized TCV recombination system between the TCV genomic RNA and sat-RNA D (Nagy and Simon, unpublished data). The role for short base-paired regions in promoting recombination has also been proposed for many recombination systems, including nodaviruses (Li and Ball, 1993), bromoviruses (Nagy and Bujarski, 1995), tombusviruses (White and Morris, 1995), Qβ (Biebricher and Luce, 1992), and Φ6 bacteriophages (Onodera et al., 1993).

Motif1-hairpin facilitates 3′-terminal extension that initiates from the 3′ side of the hairpin to a greater extent than that from the 5′-side

The efficiency of 3′-terminal extension, either initiating at the A site with construct CD-27 (Fig. 3A) or at the natural recombination hot-spot position (Fig. 1A), is influenced by the priming stem and the motif1-hairpin, as previously shown (Nagy et al., 1998; Nagy and Simon, companion paper). Detection of 3′-terminal extension products that initiate at the C site demonstrated that 3′-terminal extension can also occur from the 5′ side of the motif1-hairpin. The amount of 3-TX is, however, very low and approaches the basal level (3–5%) that occurs in the absence of the motif1-hairpin (Nagy et al., 1998; Nagy and Simon, accompanying paper). The reduced efficiency of initiation at the C site may be the result of the less stable priming stem when compared to the more stable priming stem around site A (Fig. 3A). Alternatively or in addition, the motif1-hairpin may be able to facilitate 3′-terminal extension only from a priming stem located on its 3′ side but not from the 5′ side. To examine whether initiation of 3′-terminal extension on the 5′ side of the motif1-hairpin is facilitated by the motif1-hairpin, the initiation of 3′-terminal extension was artificially targeted to 5′ of the motif1-hairpin (construct CD/D site; Fig. 4A). The priming stem in CD/D site (10 bp) is comparable in length to the priming stems of construct CD-27/G23 (Fig. 2) where initiation occurred 3′ of the motif1-hairpin. In addition to the length of the priming stem, the distance of the priming stem from the motif1-hairpin was designed to be similar in these two constructs, except that in CD-27/G23, the priming stem is on the 3′ side of the

FIG. 2. Mismatch mutations in the priming stem can alter the site and efficiency of 3′-terminal extension. All constructs are derivatives of CD-27 (Fig. 1B). Only the sequence of the priming stems with mutations are shown; the remaining parts of the constructs are unchanged. The mutated nucleotides are shown in gray boxes. The percentage of 3′-terminal extension activity for each construct as compared to the wt CD-27 is depicted to the right. Initiation of 3′-terminal extension reaction occurred at different positions for some constructs, as indicated by Site A (corresponding to the wt position, see Fig. 3A), Site B and Site C (Fig. 3A). The efficiency of 3′-terminal extension was analyzed by denaturing PAGE, followed by autoradiography and densitometry. Standard error was calculated from two to three separate experiments.
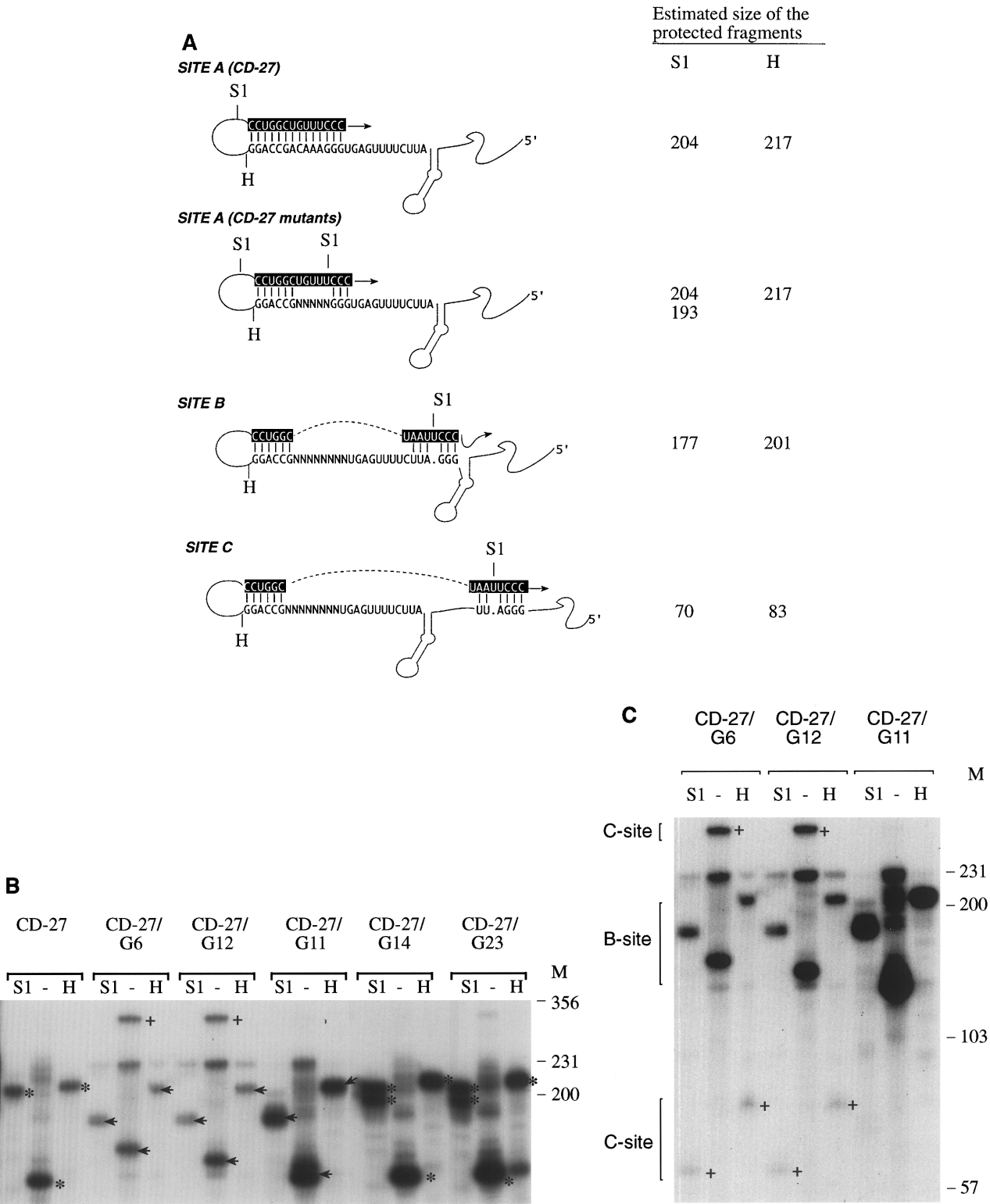
FIG. 3. The priming stem can influence the site and efficiency of 3′-terminal extension. (A) The three proposed initiation sites obtained with constructs shown in Fig. 2. Putative base-pair formation is indicated. Nuclease S1 and RNase H sensitive sites (following hybridization of a loop complementary oligonucleotide, loop 1 oligo, see Materials and Methods) are indicated by letters S1 and H, respectively. (B) A representative experiment showing denaturing gel analysis of radiolabeled 3′-terminal extension products synthesized by *in vitro* transcription with TCV RdRp using the templates from Fig. 2. Control CD-27 was loaded in 1/10th the amount of the other constructs to avoid overexposure. Lanes depicted by S1, −, and H denote products that were treated by S1 nuclease, not treated, or treated with RNase H, respectively. Asterisks depict 3′-terminal extension products that were initiated at Site A [see (A)]. Arrows point to 3′-terminal extension products that were initiated at Site B, while crosses indicate 3′-terminal extension products that were initiated at Site C. Note that in the untreated samples, the 3′-terminal extension products with initiation at
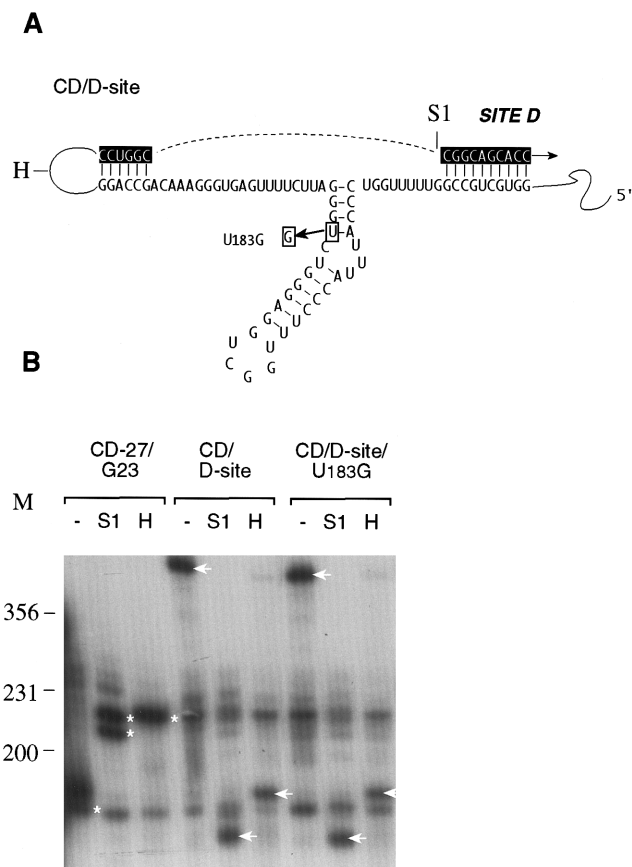
FIG. 4. Initiation of 3′-terminal extension 5′ of the motif1-hairpin occurs inefficiently. (A) Schematic representation of targeting the initiation site of 3′-terminal extension to the 5′ side of the motif1-hairpin is shown. Nuclease S1-and RNase-H-sensitive sites are indicated by letters S1 and H, respectively. (B) A representative experiment showing denaturing gel analysis of radiolabeled 3′-terminal extension products synthesized by in vitro transcription with TCV RdRp using the shown templates. CD/D site/U183G contains a single mutation in the motif1-hairpin, as shown. Control CD-27/G23 (Fig. 2) was loaded in one-tenth the amount of the other reactions to avoid overexposure. Lanes depicted by −, S1, and H denote products that were not treated, treated by S1 nuclease, or treated with RNase H (in the presence of loop 1 oligo, see Materials and Methods), respectively. Asterisks depict the expected 3′-terminal extension products for CD-27/G23 from initiation at the wt A site, while arrows point to 3′-terminal extension products with initiation occurring at Site D. Note that in the untreated sample obtained with CD-27/G23, the hairpin-like RNAs migrate aberrantly (see Fig. 3). Bands that are located in similar positions for the untreated and nuclease-treated samples are double-stranded RdRp products, and they were not characterized further. They became evident due to the substantial overexposure of the gel to visualize 3′-terminal extension products with initiation occurring at Site D. M, single-stranded RNA markers (in bases).

motif1-hairpin while in CD/D site, the priming stem is on the 5′ side of the motif1-hairpin.

Based on our previous results with various priming stems (Nagy et al., 1998; Fig. 2), which showed that 10 bp or longer priming stems can support efficient 3′-terminal extension, the priming stems of constructs CD-27/G23 and CD/D site should support 3′-terminal extension with comparable efficiency. However, 3′-terminal extension obtained with CD/D site resulted in only 5.0% of the level of activity obtained with CD-27/G23. The site of initiation mapped to the targeted region as estimated by sizing S1- and RNase-H-resistant fragments on denaturing PAGE gels (see Fig. 4B). The large drop in 3′-terminal extension activity for construct CD/D site suggests that the motif1-hairpin poorly facilitates 3′-terminal extension that initiates on the 5′ side of the hairpin, if at all.

To examine whether the motif1-hairpin has any effect on 3′-terminal extension from a priming stem located 5′ of the hairpin, a debilitating mutation was introduced into the motif1-hairpin in construct CD/D site. This U-to-G alteration at position 183 resulted in an eightfold reduction in 3′-terminal extension using CD-27 (Nagy and Simon, accompanying paper). The motif1-hairpin mutant U183G when present in construct CD/D site reduced 3′-terminal extension only marginally (to 69.2% of CD/D site levels, Fig. 4). This suggests that the motif1-hairpin plays a limited role in 3′-terminal extension that initiates from the 5′ side of the hairpin. Overall, these experiments demonstrated that the site of 3′-terminal extension can be retargeted 5′ of the motif1-hairpin by the priming stem. In addition, the motif1-hairpin apparently can only greatly facilitate 3′-terminal extension from a priming stem that is located on the 3′ side of the hairpin. This observation may be relevant in vivo, since more than 99% of the sat-RNA D/sat-RNA C recombinants characterized to date have junction sites 3′ of the motif1-hairpin (Cascone et al., 1990, 1993; Nagy et al., 1998).

Effect of the spacer sequence on the efficiency of 3′-terminal extension

Although CD-27 contains only a 14-bp priming stem, it is about sixfold more active in 3′-terminal extension than CD-13 with a 28-bp priming stem (not shown). CD-27 contains a 12-nt-long single-stranded spacer region located between the priming stem and the motif1-hairpin,

Site A have very stable hairpin-like RNAs that migrate much faster than their actual size due to nondenaturation under the conditions used. The 3′-terminal extension products with initiation at Site B have less stable hairpin-like RNAs that, depending on the sequence of a particular mutant within the priming stem, migrate slightly slower than 3′-terminal extension products that initiate at Site A. The 3′-terminal extension products with initiation at Site C migrate according to their actual size since they have hairpin-like RNAs with large loop sequences and short double-stranded regions that are completely denatured under the electrophoresis conditions used. Similarly, all S1- and RNase-H-treated samples are completely denatured under the electrophoresis conditions used, thus they migrate according to their actual size. 3′-terminal extension products obtained with CD-27/G14 and CD-27/G23 generated two S1-nuclease-resistant products that are the consequence of partial cleavage at a mismatched region within the priming stem and a more complete digestion within the loop sequence, as schematically shown in (A). (C) Overexposure of the central portion of the gel shown in panel B to visualize low-molecular-weight 3′-terminal extension products following S1 and RNase H treatment that initiated from Site C (indicated by crosses). M, single-stranded RNA markers (in bases).
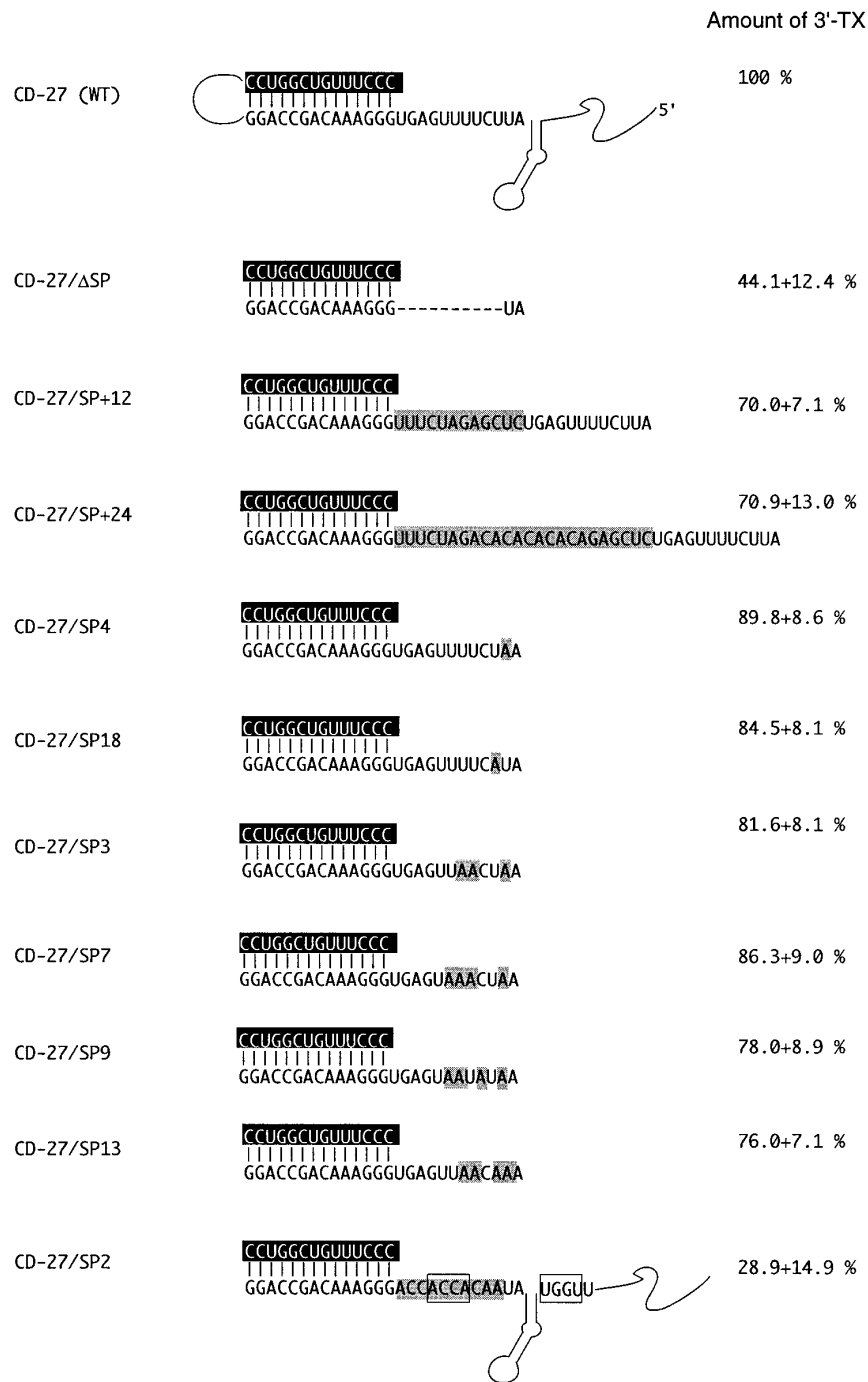
FIG. 5. Effect of sequence modifications within the spacer region on the efficiency of 3'-terminal extension. All the shown constructs are derivatives of CD-27 (Fig. 1B). Only the sequence of the priming stem and the spacer region are shown schematically. Deleted sequences are marked with −, while the inserted sequences and mutations are shown in gray boxes. The percentage of 3'-terminal extension activity for each construct when compared to the wt CD-27 is depicted to the right. Measurement of the relative amounts of 3'-terminal extension products and calculation of standard error were as described in the legend to Fig. 2. The data were normalized based on the number of template-directed radioactive UTP incorporated. The two 4-nt-long sequences (boxed) at the 5' and 3' sides of the motif1-hairpin, respectively, are complementary and may extend the length of the stem of the hairpin in CD-27/SP2.

while the corresponding region is part of the priming stem in CD-13. This suggests that the spacer region in CD-27 may facilitate 3'-terminal extension. To test the effect of the spacer region on 3'-terminal extension of CD-27, we reduced its size to two nucleotides (construct

CD-27/ΔSP, Fig. 5). This resulted in only 44.1% 3'-terminal extension activity (Fig. 5), demonstrating that the sequence and/or structure (i.e., being single stranded) of the spacer region in CD-27 can facilitate 3'-terminal extension. Extension of the length of the spacer region

from 12 nt to 24 and 36 nts in CD-27/SP+12 and CD-27/SP+24, respectively, reduced 3′-terminal extension by 29–30% from the CD-27 level (Fig. 5). This suggests that the presence of a longer spacer has a less detrimental effect on 3′-terminal extension than a short spacer.

To examine whether the sequence of the spacer region can influence 3′-terminal extension, the pyrimidine-rich portion of the spacer region was mutagenized by changing one to four U or C residues to A residues (Fig. 5). Single U-to-A mutations in two different positions within the spacer region (constructs CD-27/SP4 and CD-27/SP18) had only a minor effect on 3′-terminal extension. Three U-to-A mutations within the spacer region (construct CD-27/SP3) reduced 3′-terminal extension by 18%, while four U(C)-to-A mutations in different positions within the spacer region (constructs CD-27/SP7, CD-27/SP9, and CD-27/SP13) reduced 3′-terminal extension by 14–24% (Fig. 5). These data suggest that the sequence of the spacer region can have a moderate effect on 3′-terminal extension. To further examine whether the sequence and/or structure of the spacer region can influence 3′-terminal extension, the spacer region was replaced by a 12-nt unrelated sequence (CD-27/SP2). Four bases of the spacer region in CD-27/SP2 are predicted to base-pair with sequences on the 5′ side of the motif1-hairpin, thus extending the length of the lower stem by 4 bp (shown schematically in Fig. 5). Construct CD-27/SP2 with the heterologous spacer region supported a low level of 3′-terminal extension (29% of the wt level, Fig. 5), suggesting that the sequence and/or structure of the spacer region is important in 3′-terminal extension.

Overall, mutagenesis and deletion experiments have demonstrated that the length, sequence, and structure of the spacer are important parameters in 3′-terminal extension. Constructs with spacers that are 12–36 nt long, single stranded, and pyrimidine (mainly U) rich supported the highest level of 3′-terminal extension. One possible role of the spacer region is to facilitate the proper positioning of the RdRp (possibly bound to the motif1-hairpin, Nagy et al., 1998) and the priming stem.

The role of the spacer region during in vivo recombination is not known. The spacer region is predicted to be part of the priming stem (i.e., base-paired) in the recombination intermediate as shown in Fig. 1A. This suggests that the positioning of the primer stem and the motif1-hairpin may not be optimal during the in vivo events due to the lack of a flexible spacer. This suboptimal positioning of sequences in the recombination intermediate may be the consequence of the early steps that may determine the sat-RNA D 3′-end sequence (i.e., the donor junction site) and the amount of primers synthesized on the donor sat-RNA D. In addition, it is not clear whether the nascent-strand primer together with the RdRp simply falls off the donor template and is then available for capture by the acceptor or if the acceptor can assist in removing (i.e., can displace) the nascent strand and the

RdRp from the donor during the in vivo strand transfer events. One might expect the in vivo system may have different primer requirements than the in vitro system, which does not contain the donor RNA, if the replacement of the donor RNA in the replication complex is assisted by the acceptor RNA.

Effect of a U-rich 5′ flanking sequence on 3′-terminal extension

To analyze the possible role of a U-rich sequence located on the 5′ side of the motif1-hairpin, two new constructs were generated. The first construct was CD-27mini0, which is similar to CD-27 except that the flanking sequence 5′ of the motif1-hairpin was reduced from 148 nts to three G nucleotides (Fig. 6A). The second construct, CD-27mini, was similar to CD-27mini0 except that CD-27mini contains a 20-nt sequence 5′ of the motif1-hairpin that is U rich close to the hairpin (Fig. 6A). CD-27mini0 had 56.1% of the 3′-terminal extension activity obtained with CD-27mini, suggesting a role for the 5′ U-rich flanking sequence in 3′-terminal extension.

To further analyze the role for the 5′ U-rich flanking sequence in 3′-terminal extension, the 3′-terminal extension activity of two new constructs, CD-13mini0 and CD-13mini, was tested (Fig. 6A). CD-13mini0 and CD-13mini are similar to CD-27mini0 and CD-27mini, respectively, except that they contain 28-nt priming stems with the 3′ end representing the in vivo recombination hot-spot sequence (Fig. 6A). The 3′-terminal extension activity of construct CD-13mini0 was only 10.9% of the level of CD-13mini (Fig. 6B). The larger drop in 3′-terminal extension activity for CD-13mini0 than that for CD-27mini0 when compared to CD-13mini and CD-27mini, respectively, might be explained by the presence of the single-stranded spacer region in CD-27mini0, which is base-paired in CD-13mini0. The spacer region is pyrimidine rich and may have some overlapping function with the U-rich 5′ flanking sequence in facilitating 3′-terminal extension. To test this possibility, CD-27mini0/SP9 and CD-27mini0/SP13 were constructed. These constructs contained four pyrimidine-to-A mutations in the spacer region and lacked the 5′ U-rich flanking sequence (Fig. 6A). When compared to CD-27mini, which contains both the pyrimidine-rich spacer and the 5′ U-rich flanking sequence, CD-27mini0/SP9 and CD-27mini0/SP13 exhibited only 11.3 and 10.9% level of 3′-terminal extension activity obtained with CD-27mini, respectively (Fig. 6B). This indicates that 3′ pyrimidine-rich spacer sequence and the 5′ U-rich flanking sequence play significant roles in 3′-terminal extension reactions. Their roles in 3′-terminal extension, however, are overlapping, since separate modifications of either region (constructs CD-27mini0 vs CD-27mini/SP9 and CD-27mini/SP13) had a less detrimental effect on 3′-

A
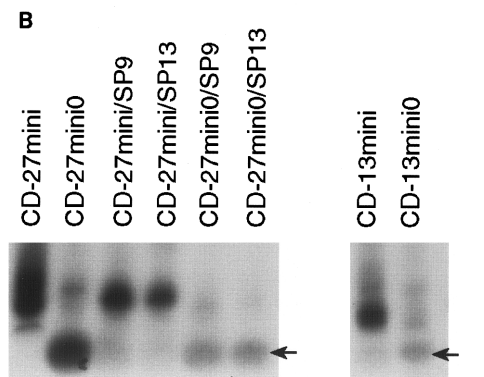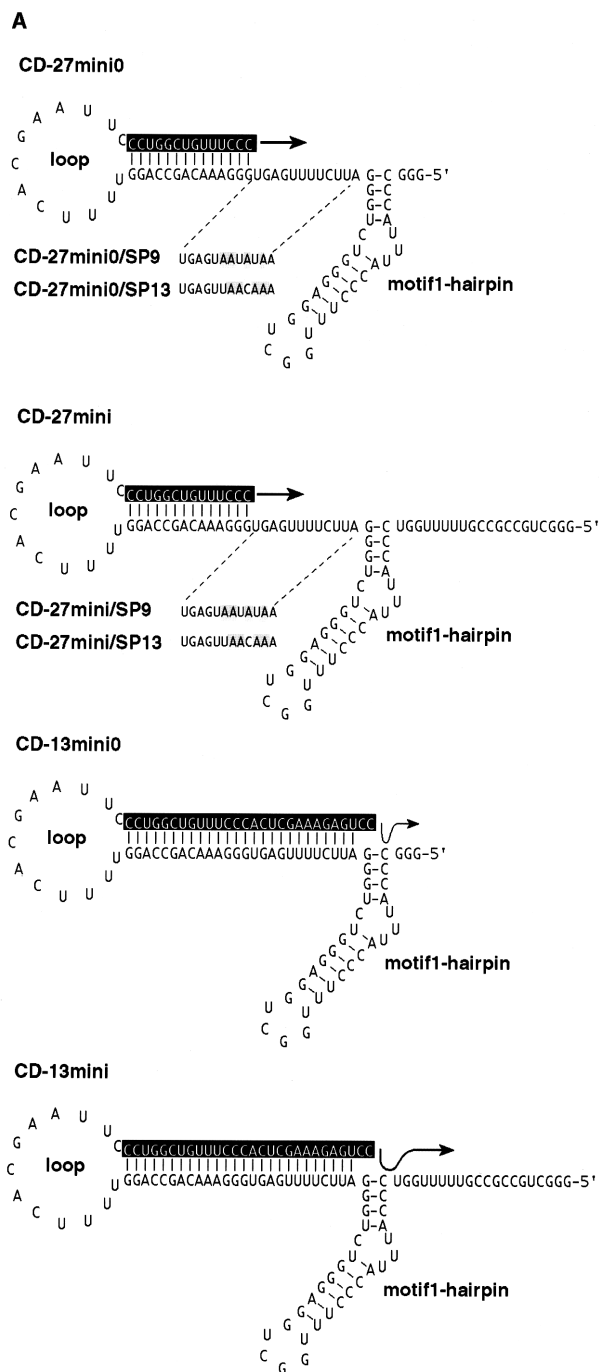
terminal extension than that obtained with the double mutants (constructs CD-27mini0/SP9 and CD-27mini0/SP13). For example, construct CD-27mini0, which contained the pyrimidine-rich spacer but lacked the 5′ U-rich flanking sequence, resulted in 56.1% of 3′-terminal extension activity when compared with CD-27mini (Fig. 6B). In addition, constructs CD-27mini/SP9 and CD-27mini/SP13 (Fig. 6A) with the four pyrimidine-to-A mutations in the spacer region but containing the nonmodified 5′ U-rich flanking sequences, supported 3′-terminal extension at 87.5 and 62.8%, respectively, of the level obtained with CD-27mini.

Taken together, these experiments demonstrate that the lowest 3′-terminal extension activity was observed with constructs that had changes in both the 3′ pyrimidine-rich sequence and the 5′ U-rich flanking sequence. Since these two sequences flanking the motif1-hairpin are single stranded in the minus-stranded sat-RNA C based on enzymatic and chemical structure probing (Carpenter et al., 1995), it is possible that one of the functions of these elements is to expose the motif1-hairpin for RdRp binding. Alternatively, or in addition, it is possible that these regions, in conjunction with the motif1-hairpin, can contribute directly to RdRp binding. Indeed, these sequences were present in the competitor RNAs that were used in increasing amounts to sequester the RdRp and to inhibit 3′-terminal extension of a template RNA in competition experiments (Nagy et al., 1998).

U-rich sequences have been found to play roles in RNA synthesis in other viral systems. For example, the intercistronic region in the RNA3 component of BMV is required for subgenomic RNA synthesis in vivo (French and Ahlquist, 1988) and in vitro (Marsh et al., 1988). These authors proposed that the polyuridylate tract serves as a spacer that allows the RdRp to better access the actual subgenomic RNA promoter. Alternatively, polyuridylate tract may bind to a component (a proposed helicase) essential for RdRp activity (Adkins et al., 1997). These authors suggested that 5- to 8-nt oligouridylate tracts may be required by the BMV RdRp for subgenomic RNA synthesis. In vitro studies with

B

FIG. 6. Effect of a U-rich sequence located 5′ of the motif1-hairpin on the efficiency of 3′-terminal extension. (A) Schematic representation of templates used for the RdRp reactions. CD-27mini0, CD-27mini0/SP9, CD-27mini0/SP13, and CD-13mini0 have similar 5′ deletions. The four U- or C-to-A mutations within the spacer region present in CD-27mini/SP9, CD-27mini0/SP9, CD-27mini/SP13, and CD-27mini0/SP13 are shown. (B) A representative experiment showing denaturing gel analysis of radiolabeled 3′-terminal extension products synthesized by in vitro transcription with TCV RdRp using the shown templates. Arrows point to the 3′-terminal extension product obtained with the mini0 constructs. The PAGE gel shown (right) was obtained by a longer exposure due to the reduced 3′-terminal extension activity of CD-13mini and CD-13mini0 when compared to the CD-27mini and its derivatives (left).
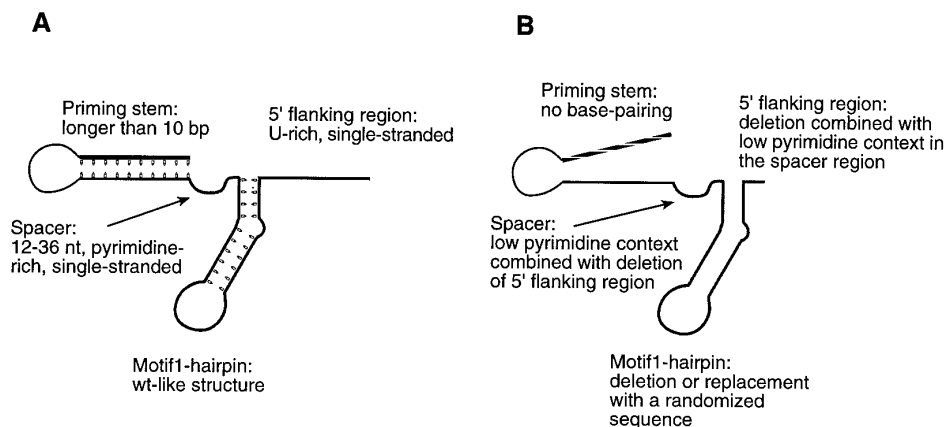
A

Priming stem:
longer than 10 bp

5' flanking region:
U-rich, single-stranded

Spacer:
12-36 nt, pyrimidine-
rich, single-stranded

Motif1-hairpin:
wt-like structure

B

Priming stem:
no base-pairing

5' flanking region:
deletion combined with
low pyrimidine context in
the spacer region

Spacer:
low pyrimidine context
combined with deletion
of 5' flanking region

Motif1-hairpin:
deletion or replacement
with a randomized
sequence

FIG. 7. Schematic representation of the optimal (A) and the least-favored (B) sequences and structures in 3′-terminal extension reactions *in vitro.*

Qβ RdRp also support a role for involvement of pyrimidine-rich regions in RNA synthesis. For example, an *in vitro* selection system for the ability of templates to function in Qβ RdRp reactions selected templates from random sequences that contained 8- to 15-nt sequences with a minimum of 80% pyrimidines (Brown and Gold, 1995). Mutations in the pyrimidine-rich region of one of the selected RNAs reduced its binding activity for Qβ RdRp by 10-fold and abolished its template function (Brown and Gold, 1995).

In summary, this and previous papers (Nagy *et al.,* 1998; Nagy and Simon, accompanying paper) have demonstrated a role for four separate elements in 3′-terminal extension in TCV *in vitro.* These sequences are the motif1-hairpin, the priming stem, a spacer located between the motif1-hairpin and the priming stem, and a U-rich sequence located 5′ of the motif1-hairpin (Fig. 7). The motif1-hairpin may have two functions: (1) to recruit the RdRp and (2) to present the priming stem properly to the catalytic center of the RdRp (Nagy *et al.,* 1998; Nagy and Simon, accompanying paper). The spacer and the 5′ flanking region are proposed to affect RdRp binding both directly and indirectly. The direct effect of the spacer and 5′ flanking sequence on RdRp binding is supported by the observation that they are pyrimidine rich and have overlapping functions. Their indirect role in RdRp binding may be to serve as single-stranded regions to expose the motif1-hairpin. In addition, the spacer region together with the motif1-hairpin may also be involved in proper primer presentation for the RdRp. The priming stem is proposed to function by presenting the template-primer duplex for the RdRp. Priming stems that located 3′ of the motif1-hairpin are favored during 3′-terminal extension. Importantly, priming stems located at a distance from the motif1-hairpin can be used efficiently by the RdRp in 3′-terminal extensions. The observation that the motif1-hairpin, a *cis*-active element (Nagy, Pogany and Simon, unpublished data) can facilitate 3′-terminal extension from a distant site may have implications for formation of recombinants in TCV and other virus systems as well (see Nagy and Simon, accompanying paper).

MATERIALS AND METHODS

RNA template construction

RNA templates were obtained by *in vitro* transcription with T7 RNA polymerase using either PCR amplified DNA templates or purified and *Sma*I-linearized plasmid DNA (Song and Simon, 1994; Nagy *et al.,* 1997). After phenol/chlorophorm extraction, unincorporated nucleotides were removed by repeated ammonium-acetate/isopropanol precipitation (Song and Simon, 1994; Nagy *et al.,* 1997). The obtained RNA transcripts were dissolved in sterile water, and their amount and size were measured by a UV spectrophotometer and 5% polyacrylamide/8 M urea gel (denaturing PAGE) analysis (Song and Simon, 1994; Nagy *et al.,* 1997).

Constructs CD-13mini, CD-27, and CD-27mini have been generated previously (Nagy *et al.,* 1998). Constructs CD-27/A1, CD-27/A2, CD-27/A3, and CD-27/A4 were generated by PCR with primer pairs of hairpin-1 [5′-GGG(A/T)-(A/T)(A/T)(C/T)AGCCAGGGAATTCGTGAAAACCTGGCTG-3′] and T7C3′ (5′-GTAATACGACTCACTATAGGGCAGGCCC-CCCG-3′) using CD-27 as template. After cloning into *Sma*I-digested pUC19, the appropriate constructs were selected by sequence analysis. The resulting constructs were linearized with *Sma*I prior to RNA transcription. Construct CD-27/rev was obtained by PCR with primers T7C3′ and har-rev [5′-GGGAAAGTCGGTGGGAATTCGT-GAAAACC(A/T)(G/C) (G/C) (G/C)ACTTTCCC-ACTC-3′]. The PCR product was cloned into *Sma*I-treated pUC19. CD-27/ΔSP was constructed by PCR with primers har-2 [5′-CAGTGAATTCG-ATATCTCGAGCCTG(G/C) (G/C)TGT-TT(G/C) (G/C) (G/C)ATCCCAGACCCTCC-3′] and T7C3′ using CD-27 as template. The PCR product was digested with *Spe*I and *Eco*RI and ligated into the similarly treated CD-27.

CD-27/SP2 was obtained by PCR and primers T7C3′

and har-D-spacer (5′-ATCGTACGGGAGCGTGAAAACCT-GGCTGTTTCCCTGGTGGTGTTATCCCAGACCCTCC-3′) on CD-27 as template. The PCR product was treated with SpeI and cloned to CD-27/ΔSP that had been digested with SpeI and EcoRV. Construct CD-27/SP+24 was generated by PCR with primers T7C3′ and har-1+30 (CCA-GGGAATTCGTGAAAACCTGGCTGTTTCCCAAAGATCTG-TGTGTGTGTCTCGAGA CTCAAAAGAATCCC-3′) on CD-27 template, followed by digestion with SpeI and EcoRI and ligation into the similarly treated CD-27. Construct CD-27/SP+12 was obtained by digesting CD-27/SP+24 with BgII and XhoI, followed by treatment with Klenow DNA polymerase and recircularization.

Constructs CD-27/SP3, CD-27/SP4, CD-27/SP7, CD-27/SP9, CD-27/SP13, and CD-27/SP18 were generated by PCR with primers T7C3′ and har-1spacerA [5′-CAGG-GAATTCGTGAAAACCTGGCTGTTTCCCACTCA(A/T) (A/T) (A/T)(G/T) (A/T) (A/T)TCCCAG-3′] on CD-27 template. Constructs CD-27/G6, CD-27/G7, CD-27/G11, CD-27/G12, CD-27/G14, CD-27/G21, CD-27/G22, and CD-27/G23 were created by PCR with primers T7C3′ and har-1GC-mod [5′-ATCGTGAATTCGTGAAAACCT(G/T) (G/T) (C/T) TGT-TT(C/T) (C/T) (C/T) ACTCAAAA-GAATCCC-3′] using CD-27/A2 as template. All the above PCR products were digested with SpeI and EcoRI and ligated into a similarly treated CD-27 or CD-27/A2. The appropriate constructs were selected by sequence analysis.

Construct CD/D site was obtained with primers T7C3′ and new-initiation (5′-GGTGCTGCCGAGGCCTTTGGTT-GCCA-GGGAATTCGTGAAAAC-3′) on CD-27 template. CD-27mini0, CD-27mini0/SP9, and CD-27mini0/SP13 were constructed with primers har1/0 (5′-GGGAAACA-GCCAGGGAATTCGTGA-3′) and mini0 (5′-GTAATACGAC-TCACTATAGGTCCCATTTACCCTTTGG-3′) using CD27, CD-27/SP9, and CD-27/SP13, respectively, as templates. CD-13mini0 was obtained with primers mini0 and CD-13 (5′-GGACTCTTTCGAGTGGGATACTGCCAGGGAATTCG-TGA-3′). Constructs CD-27mini/SP9 and CD-27mini/SP13 were generated with primers har1/0 and T7motif1 (5′-GTAATACGACTCACTATAGGGCTGCCGCCGTTTTTGG-3′) on CD-27/SP9 and CD-27/SP13 templates. The PCR products were purified and used for RNA transcription.

3′-terminal extension product analysis

Preparation of template-dependent RdRp from TCV-infected turnip plants, in vitro 3′-terminal extension reactions (similar to transcription reactions, but using different RNA templates), and product analysis was carried out as previously described (Song and Simon, 1994; Nagy et al., 1997, 1998) using a 20-μl RdRp reaction mixture that contained 3 μg of template RNA.

3′-terminal extension products were treated with S1 nuclease or RNase H in the presence of 1 μg of loop1 oligo (5′-AATTCGTGAAAA-3′, complementary to the loop region of CD-27, see Fig. 1B) as described (Nagy et al., 1998) at 37°C for 1 h. After phenol/chlorophorm extraction and ammonium-acetate/isopropanol precipitation, the products were analyzed on a 20-cm-long denaturing 5% PAGE/8 M urea gel, followed by autoradiography and densitometry (Nagy et al., 1997). The data were normalized based on the number of template-directed radioactive UTP incorporated into the 3′-terminal extension products and the molar amount of the template RNA in the RdRp reaction. For some experiments, the gels were stained with ethidium bromide, photographed, and dried, followed by analysis with a phosphorimager as described (Nagy et al., 1997).

REFERENCES

Adkins, S., Siegel, R. W., Sun, J.-H., and Kao, C. C. (1997). Minimal templates directing accurate initiation of subgenomic RNA synthesis in vitro by the brome mosaic virus RNA-dependent RNA polymerase. RNA 3, 634–647.

Biebricher, C. K., and Luce, R. (1992). In vitro recombination and terminal elongation of RNA by Qβ replicase. EMBO J. 11, 5129–5135.

Brown, D., and Gold, L. (1995). Selection and characterization of RNAs replicated by Qβ replicase. Biochemistry 34, 14775–14782.

Carpenter, C. D., Oh, J.-W., Zhang, C., and Simon, A. E. (1995). Involvement of a stem-loop structure in the location of junction sites in viral RNA recombination. J. Mol. Biol. 245, 608–622.

Cascone, P. J., Carpenter, C. D., Li, X. H., and Simon, A. E. (1990). Recombination between satellite RNAs of turnip crinkle virus. EMBO J. 9, 1709–1715.

Cascone, P. J., Haydar, T. F., and Simon, A. E. (1993). Sequences and structures required for recombination between virus-associated RNAs. Science 260, 801–805.

French, R., and Ahlquist, P. (1988). Characterization and engineering of sequences controlling in vivo synthesis of brome mosaic virus subgenomic RNA. J. Virol. 62, 2411–2420.

Guan, H., Song, C., and Simon, A. E. (1997). RNA promoters located on (−)-strands of a subviral RNA associated with turnip crinkle virus. RNA 3, 1401–1412.

Jarvis, T. C., and Kirkegaard, K. (1991). The polymerase in its labyrinth: Mechanisms and implications of RNA recombination. Trends Genet. 7, 186–191.

King, A. M. Q. (1988). Genetic recombination in positive strand RNA viruses. In "RNA Genetics," (E. Domingo, J. J. Holland, and P. Ahlquist, Eds.), Vol II, pp. 149–185. CRC Press, Boca Raton, FL.

Lai, M. C. M. (1992). RNA recombination in animal and plant viruses. Microbiol. Rev. 56, 61–79.

Li, Y., and Ball, L. A. (1993). Nonhomologous RNA recombination during negative-strand synthesis of flock house virus RNA. J. Virol. 67, 3854–3860.

Marsh, L. E., Dreher, T. W., and Hall, T. C. (1988). Mutational analysis of the core and modulator sequences of the BMV RNA3 subgenomic promoter. Nucleic Acids Res. 16, 981–995.

Nagy, P. D., and Bujarski, J. J. (1995). Efficient system of homologous RNA recombination in brome mosaic virus: Sequence and structure requirements and accuracy of crossovers. J. Virol. 69, 131–140.

Nagy, P. D., and Bujarski, J. J. (1996). Homologous RNA recombination in brome mosaic virus: AU-rich sequences decrease the accuracy of crossovers. J. Virol. 70, 415–426.

Nagy, P. D., Carpenter, C. D., and Simon, A. E. (1997). A novel 3′-end repair mechanism in an RNA virus. Proc. Natl. Acad. Sci. USA 94, 1113–1118.

Nagy, P. D., and Simon, A. E. (1997). New insights into the mechanisms of RNA recombination. *Virology* **235**, 1–9.

Nagy, P. D., and Simon, A. E. (1998). *In vitro* characterization of late steps of RNA recombination in turnip crinkle virus. I. Role of the motif1-hairpin structure. *Virology* **249**, 379–392.

Nagy, P. D., Zhang, C., and Simon, A. E. (1998). Dissecting RNA recombination *in vitro*: Role of RNA sequences and the viral replicase. *EMBO J.* **17**, 2392–2403.

Onodera, S., Qiao, X., Gottlieb, P., Strassman, J., Frilander, M., and Mindich, L. (1993). RNA structure and heterologous recombination in the double-stranded RNA bacteriophage ϕ6. *J. Virol.* **67**, 4914–4922.

Simon, A. E., and Howell, S. H. (1986). The virulent satellite RNA of turnip crinkle virus has a major domain homologous to the 3′ end of the helper virus genome. *EMBO J.* **5**, 3423–3428.

Simon, A. E., and Nagy, P. D. (1996). RNA recombination in turnip crinkle virus: Its role in formation of chimeric RNAs, multimers, and in 3′-end repair. *Semin. Virol.* **7**, 373–379.

Song, C., and Simon, A. E. (1994). RNA-dependent RNA polymerase from plants infected with turnip crinkle virus can transcribe (+) and (−)-strands of virus-associated RNAs. *Proc. Natl. Acad. Sci. USA* **91**, 8792–8796.

White, K. A., and Morris, T. J. (1995). RNA determinants of junction site selection in RNA virus recombinants and defective interfering RNAs. *RNA* **1**, 1029–1040.

Zimmern, D. (1988). Evolution of RNA viruses. *In* "RNA Genetics," (J. J. Holland, E. Domingo, and P. Ahlquist, Eds.), Vol. II, pp. 211–240. CRC Press, Boca Raton, FL.

Zhang, C., Cascone, P. J., and Simon, A. E. (1991). Recombination between satellite and genomic RNAs of turnip crinkle virus. *Virology* **184**, 791–794.