A “Slide-back” Mechanism for the Initiation of Protein-primed RNA Synthesis by the RNA Polymerase of Poliovirus*

Received for publication, July 11, 2003, and in revised form, August 20, 2003 Published, JBC Papers in Press, August 22, 2003, DOI 10.1074/jbc.M307441200

Aniko V. Pauli, Jiang Yin, JoAnn Mugavero, Elizabeth Rieder, Ying Liu, and Eckard Wimmer

From the Department of Molecular Genetics and Microbiology, School of Medicine, Stony Brook University, Stony Brook, New York 11790

Poliovirus RNA replication is initiated when a molecule of UMP is covalently linked to the hydroxyl group of a tyrosine in the terminal protein VPg. This reaction can be reproduced in vitro with an assay that utilizes two purified viral proteins, RNA polymerase 3Dpol and viral protein 3CDpro, synthetic VPg, UTP, and Mg2+. The template for the reaction is either poliovirus RNA or transcripts of a small RNA hairpin, termed cre(2C), located in the coding sequence of protein 2CATPase. The products of the reaction are VPgpU and VPgpUpU, the primers used by 3Dpol for RNA synthesis. With mutant template RNAs in this assay we determined the precise initiation site. Our results indicate that 1) 3Dpol does not possess strict specificity toward the nucleotide it links to VPg, 2) A-5 of the conserved 5AAA7 sequence in the loop is the template nucleotide for the linkage of both the first and second UMPs to VPg, 3) VPgpUpU is synthesized by a “slide-back” mechanism, and 4) A-6 provides specificity to the reaction during the slide-back step and also modulates the uridylylation reaction. In additional experiments we determined the effect of mutations in the 5AAA7 sequence of cre(2C) on viral growth, RNA replication, and on the activity of the 2CATPase protein. Furthermore, we observed that the spacing between G-1 and A-5 and the size of the loop affect the yield but not the nature of the VPg-linked products.

One of the mechanisms to initiate the replication of viral genomes involves the use of terminal proteins as primers for DNA or RNA synthesis (1, 2). In the genome these proteins are covalently attached to the 5′ end of the DNA/RNA strands. During the initiation step they provide the hydroxyl group required by the polymerase to start the synthesis of long chains. In this process the very first step is the linkage of a nucleotide to the hydroxyl group of an amino acid in the terminal protein catalyzed by the DNA/RNA polymerase. Interestingly, the same polymerases are also responsible for the elongation of DNA/RNA strands. Thus, these enzymes are structurally adapted for using either a protein or a nucleic acid primer.

Poliovirus (PV), a member of the Picornaviridae virus family, possesses a linear plus strand RNA genome about 7500 nt long (3). It contains a long 5′-NTR, a single open reading frame, and a short 3′-NTR with a poly(A) tail (Fig. 1A, Ref. 4). The 5′-terminal U (UMP) of the RNA is covalently linked to the hydroxyl group of a tyrosine in the terminal protein VPg (5, 6). The replication of this RNA requires all of the nonstructural proteins of the virus and also several cis-replicating elements (for review, see Refs. 2 and 7). The viral proteins most directly involved in RNA synthesis are the template- and primer-dependent RNA polymerase 3Dpol (8, 9), the terminal protein VPg (9, 10), and 3CDpro (10, 11), a protease (12, 13), and an important RNA-binding protein (14). The RNA polymerase has two distinct types of activities. First, it can elongate RNA chains on a suitable template (8). Second, it can covalently link a nucleotide to the hydroxyl group of tyrosine in VPg (9). The cis-replicating elements include a cloverleaf like structure in the 5′-NTR (14) and a heteropolymeric region in the 3′-NTR with the poly(A) tail (15). In addition, an important RNA hairpin (cre(2C)) is located in the PV-coding sequence of protein 2CATPase (16). This RNA element is similar in size and shape to rhinoviral cre elements such as HRV2 cre(2A) (17) and HRV14 cre(VP1) (18). All of these hairpins contain a conserved sequence of 5GXXAAXXXXXX14 (19, 20) either in one large loop or in a loop and a nearby bulge of the stem. Mutational analysis of this sequence indicated that the first two A residues of the 5AAA7 sequence are essential for viral viability, whereas changes in the other nucleotides resulted in growth defects (11, 19, 20). Using an in vitro assay system the function of these A residues has been recently identified for poliovirus (10, 11), HRV2 (17), and HRV14 (19), as the template for the linkage of U to VPg and for the elongation of the resulting VPgpU precursor into VPgpUpU. The synthesis of VPgpUpU is catalyzed by 3Dpol and requires the presence of polymerase 3CDpro and either full-length viral RNA or transcripts of cre RNA (10); VPg + 3Dpol + 3CDpro + UTP + viral RNA (or cre RNA) + Mg2+ = VPgpU + VPgpUpU. Because both plus- and minus-strand RNAs are VPg-linked in vitro (21) it is generally accepted that VPgpUpU is the primer used by PV 3Dpol for the synthesis of both RNA strands. Recent studies suggest that the VPgpUpU made on the cre(2C) is used only for plus but not for minus strand RNA synthesis. Although it is also not yet known where the VPg-linked precursors are made for minus-strand synthesis, the poly(A) tail has been proposed as a possible site (9, 22, 23).

Based on the polarity of the 5′-5AAA7-3′ sequence one would expect that A-6 is the template nucleotide for the linkage of U to VPg. However, our previous studies with an 6A→C mutant of PV cre(2C) in the in vitro uridylylation reaction showed that the primary product was VPgpU, which was only poorly elongated to the dinucleotide stage (11). This observation suggested...
the possibility that A-5 rather than A-6 was the template nucleotide for the attachment of U to VPg. To determine the possibility that A-5 rather than A-6 was the template nucleotide for the attachment of U to VPg, we tested the effect of the triple A mutations described in Table I. The parental dual-cre PV contains an A—C mutant cre at the native position and a wt cre between the cloverleaf and the internal ribosomal entry site (IRES) inserted into a Nhel site (Fig. 1C). The wt cre at the Nhel site was replaced by mutant cre carrying one of the triple A mutations.

**Enzymes—Poliovirus 3DPol** was expressed in *Escherichia coli* from plasmid pT7T-3D and purified as described before (32). Poliovirus 3DPol (3Cpro, 3CDpol, H40A) containing a C-terminal His tag was also expressed in *E. coli* and purified as described before (10).

**VPg Uridylation Assay**—The uridylation of VPg was measured as described before (10) except that the UTP concentration was increased to 25 μM. When indicated [α-32P]UTP (0.75 μCi, 3000 Ci/mmol, PerkinElmer Life Sciences) was replaced by an equal amount of unlabeled UTP, CTp, or ATP, and unlabeled UTP was replaced by 25 μM unlabeled GTP, CTP, or ATP, respectively. The reaction products were quantitated by a PhosphorImager (Molecular Dynamics, Storm 860) by measuring the amount of [α-32P]UIMP incorporated into product.

**In Vitro Transcription and Translation of mRNAs—Plasmid DNAs were linearized with EcoRI and translated with phage T7 RNA polymerase (31). The RNAs were purified and translated in HeLa cell-free extracts at 30 °C in the presence of Tran*32P*-label (ICN Biochemicals) (33). Samples of the translation reactions were analyzed on SDS-polyacrylamide (12.5%) gels.

**RNA Transfection and Plaque Assays—RNA transfectants were introduced into HeLa R19 cells with the DEAE-dextran method, and virus titers were determined by plaque assays on HeLa R19 monolayers (33).**

**Recombinant Repair (RT)/PCR—Single plasmids isolated from virus stocks were grown in HeLa R19 cells. The total cellular RNA was harvested and used as templates for the synthesis of a new fragment, which was cut with XhoI (nt 4433) and MluI (nt 4831) and cloned into similarly restricted pT7PVVM. cDNA fragments corresponding to the PV1 cre(2C) sequence (nt 4445—4504) were obtained by PCR amplification using oligonucleotides previously described (11).

**pPV1A(2C)/PVatt(S) Mutants**—The parental dual-cre PV contains an A—C mutant cre at the native position and a wt cre between P1 and P2, inserted into a ScaI site (Fig. 1B; Ref. 20). The 5A—C mutation was replaced by one of the triple A mutations.

**pPILA(2C)/PVatt(N) Mutants**—The parental dual-cre replicon (20) contains the luciferase gene in place of the P1 domain of the poliovirus genome, a 5A—C mutant cre in the last coding region, and a wt cre between the cloverleaf and the internal ribosomal entry site (IRES) inserted into a Nhel site (Fig. 1C). The wt cre at the Nhel site was replaced by mutant cre carrying one of the triple A mutations.

**Materials and Methods**

**Plasmids—**All poliovirus cDNA sequences were derived from plasmid pT7PVVM (31). The sequences listed for plasmids and oligonucleotides refer to the full-length poliovirus sequence (3). Nucleotide changes in oligonucleotides are underlined (Table I). The sequences listed for plasmids and oligonucleotides refer to the full-length poliovirus sequence (3). Nucleotide changes in oligonucleotides are underlined (Table I). The sequences listed for plasmids and oligonucleotides refer to the full-length poliovirus sequence (3). Nucleotide changes in oligonucleotides are underlined (Table I). The sequences listed for plasmids and oligonucleotides refer to the full-length poliovirus sequence (3). Nucleotide changes in oligonucleotides are underlined (Table I).
its mutant derivatives were transfected into HeLa R19 cells. Their growth was monitored by the appearance of CPE, and the genotypes of the resulting viruses were determined by DNA sequencing after RT/PCR.

**PV RNA Replication Assays**—The replication of poliovirus RNA was assayed by measurement of luciferase expression in HeLa R19 cells transfected with RNA derived from the parental pBl/5A–C/PVv(NT) plasmid (Fig. 1C) or its mutant derivatives. Luciferase assays were carried out as described before (20).

**Computer-based Prediction of RNA Structures**—RNA secondary structure was predicted by using the MFOLD program designed by M. Zucker (www.bioinfo.rpi.edu/applications/mfold).

**RESULTS**

**Effect of Nucleotide Substitutions in the 5'AAA7 Sequence of cre(2C) Elements on Replication Phenotypes**—Based on mutational analyses of the HRV14cre(VP1) (19) and poliovirus cre(2C) (20) and in comparison with other known (17) or putative entero- and rhinovirus cre elements, a common 5'XXX-AARXXXXX14 motif has been proposed that is essential for VPg uridylylation in vitro. For poliovirus this motif is 5'GXXX-AAXXXXXX14 (Fig. 1A) in which the 5'AAA7 triplet is likely to serve as template for VPg-uridylylation (10, 11). To decipher the mechanism of this reaction, we carried out a complete mutational analysis involving every other possible base substitution of the three consecutive adenosine residues and analyzed their effect in the context of viral genomes, dual-cre viral genomes, or luciferase-expressing dual-cre replicons.

The results of these experiments are summarized in Fig. 2. With one exception (7A → G) all nucleotide substitutions lead to a change in the amino acid sequence of 2CATPase (Fig. 2, column B). The mutation 5A → U introduces a termination codon (Fig. 2, column B) and, thus, has been analyzed only with a replicon (see “Results”). Virus growth (Fig. 2, column C) could be impared not only by reduced uridylylation per se but by imperfect proteolytic processing of the polyprotein or by debilitated 2CATPase function. Impaired protein synthesis is highly unlikely since the full-length transcript RNAs of wt and mutated plasmid cDNAs yielded nearly identical translation and processing patterns in the HeLa cell-free translation system of Molla et al. (33) (data not shown). The possibility of impaired 2CATPase function will be addressed later.

As shown in Fig. 2, column C, nearly all of the mutations lead to replication phenotypes. The most severe is the 5A → C substitution, which we consider lethal because we have never been able to isolate revertants. Two of the mutations led to a change in the amino acid sequence of 2CATPase (Fig. 2, column B) and, thus, has been analyzed only with a replicon (see “Results”). Virus growth (Fig. 2, column C) could be impaired not only by reduced uridylylation per se but by imperfect proteolytic processing of the polyprotein or by debilitated 2CATPase function. Impaired protein synthesis is highly unlikely since the full-length transcript RNAs of wild-type and mutated plasmid cDNAs yielded nearly identical translation and processing patterns in the HeLa cell-free translation system of Molla et al. (33) (data not shown). The possibility of impaired 2CATPase function will be addressed later.

To separate cre function from the function of 2CATPase, we repeated the analysis with a dual-cre viral genome, as described previously (20). In this genome, a second cre element was inserted into the open reading frame between P1 and P2 (Fig. 1B), and the extra cre-specific amino acid sequence was cleaved out by 2Apro (20). The genomes that we have analyzed, therefore, contain an endogenous cre with mutations as indicated in Fig. 2, column A, and a wt cre(2C) in the SacI site (Fig. 1B). The replication phenotypes of these viral genomes are summarized in Fig. 2, column D. Our results indicate that the positively charged lysine encoded by the 5'AAA7 sequence can be functionally fully replaced in the 2CATPase protein by either a Gln or a Ghu and to a smaller extent by Asn. The other substitutions appear to exert a detrimental effect. These data indicate that the lethal growth phenotype of the 5A → C mutation is not related to a defect in 2CATPase function, and therefore, A-5 of the conserved uridylylation motif is important for cre function.

The result that the lethal 5A → C mutation can be fully complemented in a dual-cre genome (20) has allowed us to analyze the effect of mutations in the 5'AAA7 triplet directly on RNA synthesis. For this purpose, we have made use of a dual-cre PV replicon carrying a luciferase gene (Fig. 1C). The endogenous cre(2C) in this replicon was inactivated by the 5A → C mutation, and a second cre was inserted at a NheI site into the 5'-NTR of the genome (Fig. 1C, Ref. 20). This second cre element that is highly efficient in intragenomic complementation of the inactivated endogenous cre(2C) (20) was then mutated as depicted in Fig. 2, column A. RNAs were transcribed from the corresponding plasmid RNAs and transfected into HeLa cells, and luciferase activity was measured as a measure of RNA replication (20). The data, expressed as % of wt luciferase activity, are shown in Fig. 2, column E. It is evident that none of the second cre elements carrying a mutation in A-5 or A-6 was able to rescue the endogenous, inactivated 5A → C cre(2C). Mutations in A-7, on the other hand, could still rescue the debilitated endogenous cre(2C), although to different extents. These results indicate an essential role for A-5 and A-6 and a less important role for A-7 of the 5'GXXAXAXXXXA14 motif in cre(2C) function.

**A-5 Is the Template Nucleotide for the Linkage of Uridylic Acid to VPg—** Genetic studies and in vitro uridylylation assays with 5A → C and 6A → C cre(2C) variants suggest that the 5'AAA7 dinucleotide serves as template for VPgPpU synthesis (11). Based on the polarity of the 5'-AAA7-3' sequence, we expected that A-6 is the nucleotide in the template for the linkage of U to VPg following by reading A-5 to elongate VPgPpU into VPgPpUpU. To determine the precise initiation site we adapted a method from studies of DNA polymerases, which catalyze protein-primed DNA synthesis. Although the terminal protein in the genomes of these viruses is always linked to a single type of nucleotide, these DNA polymerases are able to use any of the four dNTPs to nucleotidylate their terminal proteins (26–30). Therefore, we reasoned that if 3Dpol resembles these DNA polymerases, the relaxed nucleotide specificity of 3Dpol would offer a simple and direct way to test for the precise initiation site in cre(2C). Accordingly, we analyzed the mutant cre(2C) RNA templates to support nucleotidylation of VPg in the presence of either one of the four NTPs (Fig. 3, A–D). Surprisingly, when A-6 was replaced by G, C, or U, there was no change in the nucleotide specificity of the reaction (Fig. 3, A–C, lane 9), and the only products observed were VPgPpU and VPgPpUpU. In contrast the same substitutions at the A-5 position had a very different effect. UMP was no longer incorporated into product (Fig. 3, A–C, compare lane 1 with lane 5), and the nucleotide complementarity to the first position was now the one preferentially bound to VPg. Thus, 5A → G yielded VPgPpC (Fig. 3A, lane 6), 5A → C yielded VPgPSG (Fig. 3B, lane 7), and 5A → U yielded VPgPA (Fig. 3C, lane 8).

These results indicate that initiation is strictly template-dependent and that 3Dpol does accept as substrate a nucleotide in which the base is different from U depending on the nature of the 5'AAA7 template. The efficiency of linkage of C, G, or A to VPg is only about 52, 22, and 15%, respectively, of that of U (Fig. 3, A–C). Since of the three mutant templates only 5A → G

---

1 A. V. Paul, E. Rieder, J. Yin, and E. Wimmer, unpublished results.
is predicted to form an altered loop, a structural change cannot be the sole reason for the relatively low yield of nucleotidylylate products. To explore the possibility that these results were due to a defect in the stimulation of the reaction by 3CDpro on the mutant templates, we have tested nucleotidylation addition both in the presence and absence of 3CDpro (data not shown). The stimulation by 3CDpro observed on both wt and mutant templates was about 30-fold, an observation indicating that mutations at the A-5 position of the template have no inhibitory effect on the enhancement of the nucleotidylation reaction by 3CDpro. It is likely that the efficiency of the linkage of a mononucleotide to VPg is determined primarily by the nature of the base. Our results indicate that pyrimidines are preferentially linked to VPg over purines.

It should be noted that when VPg is linked to G or A no elongation of VPgpG (Fig. 3B, lane 7) or VPgpA (Fig. 3C, lane 8) to the dinucleotide level is observed. In case of VPgpC, only about 20% of the product is VPgpCpC (Fig. 3A, lane 6). These
**Addition of the Second U**

It has been well established that the **cre(2C)** sequence, located between P1 and P2, plays a crucial role in VPgpU uridylylation. This is apparent considering that the synthesis of VPgpNpN, although the hairpin structure of the **GG** mutant was predicted to be severely altered (Fig. 3D). As discussed above, the reason for this is most likely related to the preference of 3Dpol for pyrimidines over purines in the synthesis of VPgpNpN, although other possibilities cannot be excluded.

A-7 of the **AA** Sequence Is Not a Template Nucleotide, but It Weakly Enhances the Uridylation of VPg—As we have shown above, the substitution of A-7 with C or U resulted in impaired viral growth phenotypes, which were due to a defect both in RNA replication and in 2CATPase function (Fig. 2, columns E and D, respectively). To try to correlate the RNA replication defect with impaired VPgp uridylylation, we tested the A-7 variant **cre(2C)** RNAs as templates in the in vitro uridylylation reaction. As shown in Fig. 3E, the synthesis of VPgpUpU proceeded with all the mutant **cre** RNAs. The yield of product relative to wt **cre(2C)** template with 7A→G, 7A→C, and 7A→U templates was 92, 43, and 64%, respectively (Fig. 3E, compare lane 1 with lanes 2, 6, and 10, respectively).

**Elongation of VPgpU on cre(2C)** RNA Is Aborted after the Addition of the Second U—It has been well established that in vitro on the **cre(2C)** template PV 3Dpol synthesizes only two products, VPgpU and VPgpUpU. It is not yet known how the

---

### Table: Effect of mutations in the **AA** sequence of PV **cre(2C)** elements on replication phenotypes

| **A**       | **B**       | **C** | **D** | **E** |
|-------------|-------------|-------|-------|-------|
| **cre(2C)** sequence | Amino acid | 2CATPase | Introncomplementation of **cre(2C)** variants by wt **cre(2C)** (dual **cre** virus) | Introncomplementation of **cre(2C)** by **cre(2C)** variants (dual **cre** replicon) |
| A6A6A6     | Lys         | ++    | ++     | 100   |
| C6A6A6     | Gln         | −−−    | ++     | 0.1   |
| G6A6A6     | Glu         | q.i.   | ++     | 0.1   |
| U6A6A6     | Ter         | NA     | NA     | 0.1   |
| A6C6A6     | Thr         | q.i.   | +      | 0.1   |
| A6G6A6     | Arg         | q.i.   | +      | 0.2   |
| A6U6A6     | Ile         | q.i.   | +      | 0.1   |
| A6A6C6     | Asn         | +      | ++     | 20    |
| A6A6G6     | Lys         | ++     | ++     | 120   |
| A6A6U6     | Asn         | ++     | ++     | 4.8   |
| C6C6A6     | Pro         | −      | ND     | ND    |
| G6G6A6     | Gly         | −      | ND     | ND    |

---

Fig. 2. Effect of mutations in the **AA** sequence of PV **cre(2C)** elements on replication phenotypes. Column **A**, mutations in the endogenous **cre(2C)** (Fig. 1A) or in the second **cre(2C)** of dual-creat viruses and replicons. Column **B**, amino acid changes in 2CATPase resulting from mutations in the endogenous **cre(2C)**. Column **C**, growth phenotypes of viruses carrying mutations in the endogenous **cre(2C)**. ++, wt growth, CPE at 24 h post-transfection; −−−, CPE at 48 h, revertants isolated from transfection supernatants; +++, CPE at 72 h, revertants isolated from transfection supernatants; q.i., quasi-infected, indicating that all replicating progeny viruses reverted to the wt sequences; *a*, Rieder et al. (11). Column **D**, to test for a possible defect in 2CATPase function due to the mutations we used dual-creat viral genomes (Fig. 1B). The endogenous **cre(2C)** contained one of the mutations indicated in column **A**, whereas the rescuing **cre(2C)**, located between P1 and P2 (20), had wt **cre(2C)** sequence. The virus progeny was tested for the presence of revertants in the transfection supernatants. ++++, wt growth, no revertants; ++++, viable virus, contains >50% revertants to wt; +++, viable virus, contains 25–50% revertants to wt. wt replication (+++) indicates that the mutation in 2CATPase had no significant effect on enzyme function. Column **E**, the effect of mutations on RNA replication was measured using a dual-creat replicon carrying a luciferase gene (Fig. 1C). In these constructs (20) the endogenous **cre(2C)** was inactivated (5A→C) for uridylylation and replication, and the rescuing **cre** element, mapping to the 5′-NTR, carried the mutations indicated in column **A**. The data are expressed as % of wt luciferase activity. Wild type luciferase activity is taken as 100%. ND, not done; NA, not applicable.

---

### Discussion

Results strongly contrast with the elongation reaction of VPgpU, where 80% of the total uridylylation product is VPgpUpU (Fig. 3A–C, lane 1).

A-6 Is Not Used as the Template for VPg Uridylylation but It Modulates the Reaction; the Slide-back Mechanism—The **AA** mutant **cre** RNAs (6A→G, 6A→C, 6A→U) were inactive in directing specific nucleotidyltransferase of VP when tested with the complementary NTP as donor nucleotide (Fig. 3A–C, lanes 10–12, respectively). If some incorporation of NMD did occur, the values were essentially the same as those obtained for nucleotide misincorporation by the polymerase on a wt **cre(2C)** template (see for example Fig. 3A, compare lane 10 with lanes 2–4). Although A-6 cannot serve as template for VPgpU synthesis, this nucleotide nevertheless plays a crucial role in VPg uridylylation. This is apparent considering that the synthesis of VPgpU on A6A6A6 mutant template RNAs is very inefficient (Fig. 3A–C, compare lane 1 with lane 9). Both A-5 and A-6 of the template are critical for optimal synthesis of VPgpUpU and, thus, for initiation of RNA synthesis. These results are consistent with a slide-back mechanism by which VPgpU is synthesized on the A-5 template nucleotide and then relocated to base pair with A-6 followed by the A-5-dependent addition of the second U (see “Discussion”).

Having two identical nucleotides at positions 5 and 6 is by itself not sufficient for efficient synthesis of VPgpNpN. When **AA** were changed to **CC** or **GG**, only trace amounts of VPgpGpG or VPgpCpC, respectively, were synthesized (Fig. 3D, lanes 3 and 5). **UU** mutants have not been tested. Of the two double mutants **GG** yielded more product than **CC** only, the hairpin structure of the **GG** mutant was predicted to be severely altered (Fig. 3D). As discussed above, the reason for this is most likely related to the preference of 3Dpol for pyrimidines over purines in the synthesis of VPgpNpN, although other possibilities cannot be excluded.
FIG. 3. Specificity of VPg nucleotidylylation in vitro on wt and mutant cre(2C) templates. The standard uridylylation assays were carried out with either wt or mutant cre(2C) RNA templates ("Materials and Methods"). The reactions contained one of the four $[^{32}P]NTP$s, and the same NTP was present unlabeled at 25 mM. The autoradiography of the reaction products is shown, and the quantitation of the data is displayed below. The structures of the loop and part of the stem of the wt and mutant cre(2C) RNAs are shown on the top of the figure. The conserved A residues are shown in bold, and mutations are indicated in italics. A, comparison of wt cre(2C) with 5A→G and 6A→G templates. B, comparison of wt cre(2C) with 5A→C and 6A→C templates. C, comparison of wt cre(2C) with 5A→U and 6A→U templates. D, comparison of wt cre(2C) with 5A→C/6A→C and 5A→G/6A→G templates. E, comparison of wt cre(2C) with 7A→C, 7A→G, and 7A→U templates.
The conserved motif is important for cre(2C) function. VPg uridylylation was analyzed with wt and mutant cre(2C) templates (“Materials and Methods”). The autoradiography of the reaction products is shown, and the quantitation of the data is displayed below. Sequences and structures of the loop and part of the stem of cre(2C) are shown on top of the figure. The conserved motif in cre(2C) is shown in bold face, and mutations are indicated in italics.

As shown in Fig. 4 the yield of product in the in vitro reaction was much reduced compared with that obtained with the wt cre(2C) or the 6A→C template (Fig. 4, compare lane 6 with lanes 1 and 5, respectively). These observations confirmed that the two essential As can be located only at the fifth and sixth positions of the conserved motif for optimal VPg-uridylylation. A poliovirus RNA containing the 4C→A/6A→C mutations when transfected into HeLa cells yielded a small plaque virus that contained a 6C→A reversion but retained the 4C→A mutation (data not shown).

A Small Loop Size in cre(2C) Is Detrimental for VPg Uridylylation—We have shown above that a large loop consisting of 14 nt in the 4C→A cre(2C) mutant is fully functional as the template in the VPg uridylylation reaction. This observation is consistent with previous findings that the large loop (14 nt) of HRV14 cre(VP1) is an excellent template for PV 3Dpol in vitro (10, 19). Recently we tested a double mutant of PV cre(2C) in vitro in which the 3G→A/11C→U mutations reduced the predicted size of the loop to 5 nt and relocated A-5 to the stem (20). This mutant RNA, which resembles the cre(2A) RNA of HRV2 (17), retained about 25% of the template activity of the wt cre(2C) (data not shown). Although these results suggested that a small loop is detrimental to the template function of cre(2C), the possibility remained that the defect was rather due to the hydrogen-bonded state of A-5 in the stem, which was predicted for the structure of the 3G→A/11C→U mutant. To distinguish between these two possibilities we constructed a 4C→G mutant in which the predicted loop size was reduced to 5 nt by hydrogen bonding of G-4 with C-10, but A-5 was no longer part of the stem (Fig. 4). In the in vitro assay this mutant also had about 25% of the wt template activity (Fig. 4, compare lane 1 with lane 3), about the same as that of the 3G→A/11C→U double mutant. These results showed that A-5 is functional when located either in the loop or in the upper part of the stem. On the other hand a small loop size is detrimental for VPg-uridylylation even if the correct spacing between G-1 and A-5 is retained. This conclusion was supported by the genotype of a revertant we obtained from a poliovirus carrying the 4C→G mutation. Interestingly, in this revertant
FIG. 5. The proposed slide back mechanism for the uridylylation of VPg. Viral protein 3CDpol binds to the PV cre(2C) RNA and enhances the binding of the 3Dpol−VPg complex to the loop. During the initiation step U is linked to the hydroxyl group of tyrosine in VPg using A-5 as the template nucleotide. VPgpU then slides back to hydrogen bond with A-6 followed by the addition of a second U again on the A-5 template nucleotide. Nucleotides of the conserved motif are shown in bold, and nucleotides involved in the slide back are boxed. The figure was adapted from Paul (2).

the original 4C→G mutation was retained, but now it acquired a new mutation (10C→A), which led to the formation of an enlarged (14-nt) loop.

DISCUSSION

We have previously shown that the inability of the 5A→C and 6A→C cre(2C) mutants to support the uridylylation of VPg in vitro is correlated with their lethal and quasi-infectious growth phenotypes, respectively (11). We have now confirmed and extended these results by testing both in vitro and in vivo mutants containing all the possible nucleotide substitutions in the triple A sequence of the loop. We observed that no nucleotide changes at the A-5 and A-6 positions of PV cre(2C) are tolerated either in vitro or in vivo VPg uridylylation. Substitution of A-7 with pyrimidines resulted in reduced yields of products in vitro and in moderately defective growth in vivo. Of all the mutants tested the 7A→G was the only one similar to the wt. Our results are in full agreement with studies of Yang et al. (19), who used the HRV14 cre(VP1) RNA as template for PV 3Dpol. Measurements of RNA replication with a dual cre luciferase replicon independent of 2C\textsuperscript{ATPase} function fully supported these conclusions. Using dual-cre polioviruses (20) we also show that the lysine encoding the 5AAA7 sequence in the loop can be functionally exchanged in 2C\textsuperscript{ATPase} RNA and UTP as the nucleotide donor. However, it is known that although the genomes of some DNA viruses are also linked to terminal proteins, their DNA polymerases are able to use any of the four dNTPs to uridylylate their terminal proteins (26–30). These observations suggested to us the possibility that if 3Dpol resembles these DNA polymerases, its relaxed nucleotide specificity would offer a simple and direct way to test for the precise initiation site on the cre(2C) template. Our previous studies (11) and those presented here show that both A-5 and A-6 of the cre(2C) loop are essential for viral growth, RNA replication, and VPg uridylylation, suggesting that these nucleotides are directing VPgpUpU synthesis. Because RNA polymerases always synthesize chains in the 5′→3′ direction, we expected A-6 to serve as the template for the linkage of the first U to VPg followed by the addition of the second U on the A-5 nucleotide. To address this question we changed either A-5 or A-6 to C, G, or U and tested the mutant cre(2C) RNAs as templates in vitro using one of the four labeled NTPs as nucleotide donors. Our results clearly showed that the RNA polymerase was able to link G, C, and A to VPg on mutant templates that contained C, G, or U at the first position (A-5) of the triple A sequence, respectively. The nucleotidylylated peptides, VPgpN (N is C, G, or A), could not be efficiently elongated into a dinucleotide VPgpNpN or VPg-pNpU. In contrast, alteration of the sixth position did not produce any change in the nucleotide specificity of the initiation reaction but drastically reduced the yield of products, particularly that of the VPg-linked dinucleotides. Although A-7 is fully conserved in all of the cre elements studied thus far, the alteration of this position leads only to a modest reduction in the yield of VPgpU and VPgpUpU, and only traces of other nucleotidylate products can be observed. The effect of substitutions at this position is more detrimental when a purine to pyrimidine change is involved than what is observed with an A to G transition.

These results are consistent with a slide-back mechanism for the synthesis of VPgpUpU. According to the basic model for the initiation reaction, 3CDpol interacts with a complex of 3Dpol and VPg (2, 10, 36–38) and enhances their binding to the cre(2C) structure. In the first step a U is linked to the hydroxyl group of tyrosine in VPg on the A-5 template nucleotide (Fig. 5). This is followed by a slide-back of VPgpU to hydrogen bond with A-6. In the final step the second U is added again using A-5 as the template nucleotide. Accordingly to this mechanism
Poliovirus 3Dpol Phage Φ29 DNA polymerase

| Template            | 5' A A A A C 3' | 5' A C T T T 3' |
|---------------------|------------------|------------------|
| Protein-priming     | 5' A A A A C 3' | 5' A C T T T 3' |
| "Slide-back"        | 5' A A A A C 3' | 5' A C T T T 3' |
| Addition            | 5' A A A A C 3' | 5' A C T T T 3' |

Fig. 6. Comparison of the protein priming reactions catalyzed by the polymerases of poliovirus and bacteriophage Φ29. Poliovirus 3Dpol uses A-5 in the cre(2C) RNA as template for uridylylation of VPg, whereas the DNA polymerase of phage Φ29 uses the penultimate T (T₉) at the end of its linear DNA strand as template for the deoxyuridylylation of its terminal protein (27). A slide-back step follows during which the nucleotidylylated proteins are aligned with A-6 and T₉, respectively. The second nucleotide is added in the 5' → 3' direction. The figure was adapted from Paul (2). Nucleotides involved in the slide-back step are shown in bold, and those involved at each step are boxed.

The function of A-5 is to serve as template for the linkage of both the first and second U to VPg. Our data suggest that A-6 has at least two functions. First, it provides specificity to the reaction during the slide-back step, ensuring that VPgpN (N is A, C, or G) is not elongated, since N cannot efficiently or not at all hydrogen-bond with A-6. Second, A-6 modulates the nucleotidylylation reaction by enhancing either template recognition or the formation of the initiation complex on the cre(2C), resulting in increased yields of nucleotidylylated products. The latter function of A-6 is supported by the observation that only a very small amount of VPgpU is formed when this nucleotide is substituted with C, G, or U, although A-5 is unchanged. To a lesser extent A-7 also has a negative influence on the yield of uridylylated products.

The slide-back mechanism used by 3Dpol for the synthesis of VPgpUpU is similar to the mechanism by which the DNA polymerase of phage Φ29 links two A residues to its terminal protein TP (Fig. 6). The two reactions, however, do differ in their template requirement for the nucleotidylylation reaction. Although the internal cre element serves as template for 3Dpol (10), phage Φ29 DNA polymerase uses sequences at the 3' end of its partially opened double-stranded DNA (27). As a consequence of the difference between the location of the template nucleotides, the elongation of TP-dA by phage Φ29 DNA polymerase yields first TP-dAdA and then continues until the complementary DNA strand is finished. The elongation of VPgpU by PV 3Dpol, however, is aborted after the addition of the second U. This is not surprising since further copying of the loop sequence in cre(2C) would result in VPgpUpU linked to a G and then a C, etc., which would not be able to hydrogen-bond with the 3' end sequence of minus strands (3'-AAUUU-5') to prime plus strand RNA synthesis. It is interesting to note that even if poly(A) is the template, VPgpU elongation is to a large extent aborted at the stage of VPgpUpU (see below).

It is generally assumed that the function of the slide-back step in the protein-priming reaction catalyzed by the viral DNA polymerases is to ensure the integrity of the viral DNA strands and to provide specificity to the reaction. One might speculate that 3Dpol uses the slide-back mechanism to overcome the problem created by the lack of strict nucleotide specificity of the enzyme in the nucleotidylylation reaction it catalyzes. Because 3Dpol does not possess a proofreading activity it would not be able to remove the incorrect nucleotide from VPgpN (N is C, G, or A). The slide-back step would ensure that this mutant precursor is not elongated into a dinucleotide. The incorrect base-pairing of VPgpN with A-6 would most likely result in its dissociation from the replication complex.

It should be noted that although the primary cre-dependent mechanism of VPgpUpU synthesis involves the slide-back step, there appears to be an alternate, minor reaction leading to the elongation of the VPg-linked mononucleotide. This is suggested by two lines of evidence. First, when using an A-6 mutant cre(2C) template a small fraction of the product is elongated into a dinucleotide, although VPgpU cannot slide-back and hydrogen-bond with the mutant nucleotide at the sixth position. Second, on a 5A→G template a minor fraction of the product is elongated into VPgpUpC. Again, the addition of the second C must occur without a slide back of VPgpC to the sixth position, where an A is located. Therefore, we must conclude that in these cases one or more alternate mechanisms are used by the polymerase in the synthesis of the dinucleotide. For example, the addition of the second nucleotide, a pyrimidine, to VPgpN might occur in a non-templated fashion. The existence of variant mechanisms for protein-primed initiation has also been observed with some DNA polymerases such as phage Φ29 (39), Cp-1 (30), and GA-1 (29).

As we discussed above the protein-priming reaction catalyzed by PV 3Dpol also resembles the nucleotidylylation of the RT of hepatitis B virus in that both of these enzymes use an internal RNA hairpin as template for the reaction (34). However, the two enzymes differ in the mechanism they use for the addition of the second nucleotide to the protein-NMP precursor. Hepatitis B virus RT does not and cannot use a slide-back mechanism for the addition of the second nucleotide since the first two template nucleotides on the ε RNA are not identical. In addition, the two polymerases differ in that the length of the nucleotide linked to the protein by hepatitis B virus RT but not by PV 3Dpol depends on RNA sequence and structure in the ε hairpin. When the bulge in the ε RNA template is enlarged by insertions, the RT can synthesize a nucleotidylylated oligonucleotide, which is longer than the usual 4 nt (40). However, these longer protein-linked oligonucleotides appear to be defective for the template switch before minus strand DNA synthesis.

The poliovirus polymerase is not able to elongate VPgpUpU into a VPgpUpUpU even when C₄ of cre(2C) is substituted with an A. Therefore the termination of elongation by PV 3Dpol is likely due to some structural characteristics of the polymerase that are a prerequisite of the nucleotidylylation reaction. Such abortive synthesis of VPgpUpU also occurs in vitro on a poly(A) template (9), in crude replication complexes isolated from poliovirus-infected cells (41, 42), and in poliovirus-infected HeLa cells in vitro (43). It is interesting to note that picornaviral RNA polymerases are mechanistically similar to DNA-dependent RNA polymerases in that during the initiation phase RNA synthesis is aborted. It has been recently reported that T7 RNA polymerase also starts transcription by repeated abortive initiation events that produce short (2–6 nt long) RNA fragments (44). During transition from initiation to the elongation phase the enzyme undergoes major structural changes, which then result in its release from the promoter followed by its translocation along the DNA template.

Murray and Barton (22) and Morasco et al. (23) have recently provided in vitro evidence suggesting that poliovirus minus strand RNA synthesis, although VPg-dependent, is cre-independent. Plus strand RNA synthesis, on the other hand, is cre-dependent. Does this imply two distinct mechanisms in the initiation of viral RNA synthesis? We consider it highly likely that in both scenarios the primer for 3Dpol is VPgpUpU. It is not yet known where the primer is made for minus strand synthesis, but one possible site is the poly(A) tail (9, 22, 23).
Whether the primer synthesized on poly(A) (9) results from a slide-back mechanism, however, is unknown.

In additional experiments we continued our studies on the effects of changes in cre(2C) structure on VPg uridylylation. Previous studies (19, 20) have shown that the large loop of HRV14 cre(VP1) is an excellent template for uridylylation. In agreement with previous studies showing that the large loop of HRV14 cre(VP1) results from a slide-back mechanism, however, is unknown. Whether the primer synthesized on poly(A) (9) results from a cre-VPg-uridylylation. In agreement with previous studies showing that the large loop of HRV14 cre(VP1) is an excellent template for uridylylation. The relocation of A-5 from the loop of HRV14 (45). In contrast to a large loop, a small loop in cre(2C) containing only five nucleotides does not function well as the template for uridylylation. The relocation of A-5 from the loop of HRV14 (45). In contrast to a large loop, a small loop in cre(2C) containing only five nucleotides does not function well as the template for uridylylation. The relocation of A-5 from the loop of HRV14 (45). In contrast to a large loop, a small loop in cre(2C) containing only five nucleotides does not function well as the template for uridylylation.

Our previous studies of the mechanism of VPg uridylylation reaction catalyzed by the RNA polymerase of HRV2 (17), a member of the rhinovirus genus in the Picornaviridae, is in agreement with the results presented in this paper. The fact that all known picornaval internal cre elements contain a conserved AAA sequence in a loop or bulge of a hairpin suggests that this mechanism will be a common feature in the replication of all members of this virus family.

Acknowledgments—We thank J. H. van Boom for providing synthetic VPg and K. Kirkegaard for a gift of plasmid pTST-3D.

REFERENCES

1. Salas, M., Miller, J. T., Leis, J., and DePamphilis, M. L. (1996) in DNA Replication in Eukaryotic Cells (DePamphilis, M. L., ed) pp. 131–176, Cold Spring Laboratory Press, Cold Spring Harbor, NY
2. Paul, A. V. (2002) in Molecular Biology of Picornaviruses (Semler, B. L., and Wimmer, E., eds) pp. 227–246, American Society for Microbiology, Washington, D. C.
3. Kitanura, N., Semler, B. L., Rothberg, P. O., Larsen, G. R., Adler, C. J., Dorner, A. J., Emini, E. A., HanecaK, R., Lee, J. J., van der Werf, S., Anderson, C. W., and Wimmer, E. (1981) Nature 291, 547–553
4. Wimmer, E., Hellen, C. U. T., and Cao, X. (1993) Annu. Rev. Genet. 27, 353–436
5. Ambros, V. (1978) J. Biol. Chem. 253, 5263–5266
6. Rothberg, P. G., Harris, T. R., Nomoto, A., and Wimmer, E. (1978) Proc. Natl. Acad. Sci. U. S. A. 75, 4868–4872
7. Agol, V. I., Paul, A. V., and Wimmer, E. (1999) Virus Res. 62, 129–147
8. Flanegan, J. B., and Baltimore, D. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 3677–3680
9. Paul, A. V., van Boom, J. H., Filippov, D., and Wimmer, E. (1998) Nature 393, 280–284
10. Paul, A. V., Rieder, E, Kim, D. W., van Boom, J. H., and Wimmer, E. (2000) J. Virol. 74, 10359–10370
11. Rieder, E., Paul, A. V., Kim, D. W., van Boom, J. H., and Wimmer, E. (2000) J. Virol. 74, 10371–10380
12. Jore, J., de Geus, B., Jackson, R. J., Pouwels, P. H., and Enger-Valk, B. E. (1988) J. Gen. Virol. 69, 1627–1636
13. Ypma-Wong, E. F., Pypuc, K. V., Maslova, S. V., Melders, W. J. G., Slot, H. J., and Agol, V. I. (1996) EMBO J. 15, 5428–5436
14. Goodfellow, I., Chausathy, Y., Richardson, A., Meredith, J., Almond, J. W., Barclay, W., and Evans, D. J. (2000) J. Virol. 74, 4590–4600
15. Gerber, K., Wimmer, E., and Paul, A. V. (2001) J. Virol. 75, 10979–10990
16. McKnight, K. L., and Lemon, S. M. (1998) RNA (N. Y.) 4, 1569–1584
17. Yang, Y., Rijnbrand, R., McKnight, K. L., Wimmer, E., Paul, A., Martin, A., and Lemon, S. M. (2002) J. Virol. 76, 7485–7494
18. Yin, J., Paul, A. V., Wimmer, E., and Rieder, E. (2003) J. Virol. 77, 5152–5166
19. Nomoto, A., Detjen, B., Pozzatti, R., and Wimmer, E. (1977) Nature 268, 208–213
20. Murray, K. E., and Barton, D. J. (2003) J. Virol. 77, 4739–4750
21. Morasco, J. B., Sharma, N., Parilla, J., and Flanegan, J. B. (2003) J. Virol. 77, 5136–5144
22. Lortet, P. E., Escricou, N., Rueille, J., and Michelis, T. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 11560–11565
23. Mason, P. W., Bezerodeva, S. V., and Henry, T. M. (2002) J. Virol. 76, 9686–9694
24. King, A. J., and van der Vliet, P. C. (1994) EMBO J. 13, 5786–5792
25. Mendez, J., Blanco, L., Esteban, J. A., Bernado, A., and Salas, M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 9579–9583
26. Caldentey, J., Blanco, L., Bamford, D. H., and Salas, M. (1993) Nucleic Acids Res. 21, 3725–3730
27. Blanca, B., Blanco, L., and Salas, M. (1996) J. Mol. Biol. 264, 453–464
28. Martin, A. C., Blanco, L., Garcia, P., Salas, M., and Mendez, J. (1996) J. Mol. Biol. 260, 369–377
29. Van der Werf, S., Bradley, J., Wimmer, E., Studier, F. W., and Dunn, J. J. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2339–2344
30. Pata, J. D., Schultz, S. C., and Kirkegaard, K. (1995) RNA (N. Y.) 1, 466–477
31. Molla, A., Paul, A. V., and Wimmer, E. (1991) Science 254, 1647–1651
32. Seeger, C., and Mason, W. S. (1996) in DNA Replication in Eukaryotic Cells (DePamphilis, M. L., ed) pp. 815–831, Cold Spring Laboratory Press, Cold Spring Harbor, NY
33. Wimmer, E. (1982) Cell 28, 199–201
34. Xiang, W., Cuconati, A., Hope, D., Kirkegaard, K., and Wimmer, E. (1998) J. Virol. 72, 6732–6741
35. Lyle, J. M., Clewell, A., Richardson, K., Richards, O. C., Hope, D. A., Schultz, S. C., and Kirkegaard, K. (2002) J. Biol. Chem. 277, 16324–16331
36. Paul, A. V., Peters, J., Mugavero, J., Yin, J., van Boom, J. H., and Wimmer, E. (2003) J. Virol. 77, 891–904
37. Gonzalez-Huix, V., Salas, M., and Hermoso, J. M. (2000) J. Biol. Chem. 275, 46547–46553
38. Jordan, H., and Loeb, D. D. (1997) J. Virol. 71, 5345–5354
39. Takegami, T., Kuhn, R. J., Anderson, C. W., and Wimmer, E. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 7447–7451
40. Toyoda, H., Yang, C. F., Takeda, N., Nomoto, A., and Wimmer, E. (1987) J. Virol. 61, 2816–2822
41. Crawford, N. M., and Baltimore, D. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 7452–7455
42. Yin, Y. W., and Steitz, T. A. (2002) Science 298, 1387–1399
43. Goodfellow, I. G., Kerrigan, D., and Evans, D. J. (2003) RNA (N. Y.) 9, 124–137