Identification and Characterization of a *cis*-Acting Replication Element (*cre*) Adjacent to the Internal Ribosome Entry Site of Foot-and-Mouth Disease Virus

Peter W. Mason,* Svetlana V. Bezborodova, and Tina M. Henry

Plum Island Animal Disease Center, U.S. Department of Agriculture, Greenport, New York 11944

Received 19 March 2002/Accepted 25 June 2002

Over the last few years, an essential RNA structure known as the *cis*-acting replicative element (*cre*) has been identified within the protein-coding region of several picornaviruses. The *cre*, a stem-loop structure containing a conserved AAACA motif, functions as a template for addition of U residues to the protein primer 3B. By surveying the genomes of representatives of several serotypes of foot-and-mouth disease virus (FMDV), we discovered a putative *cre* in the 5′ untranslated region of the genome (contiguous with the internal ribosome entry site [IRES]). To confirm the role of this putative *cre* in replication, we tested the importance of the AAACA motif and base pairing in the stem in FMDV genome replication. To this end, *cre* mutations were cloned into an FMDV replicon and into synthetic viral genomes. Analyses of the properties of these replicons and genomes revealed the following. (i) Mutations in the AAACA motif severely reduced replication, and all viruses recovered from genomes containing mutated AAACA sequences had reverted to the wild-type sequence. (ii) Mutations in the stem region showed that the ability to form this base-paired structure was important for replication. Although the *cre* was contiguous with the IRES, the mutations we created did not significantly reduce IRES-mediated translation in vivo. Finally, the position of the *cre* at the 5′ end of the genome was shown not to be critical for replication, since functional replicons and viruses lacking the 5′ *cre* could be obtained if a wild-type *cre* was added to the genome following the 3D*pol* coding region. Taken together, these results support the importance of the *cre* in replication and demonstrate that the activity of this essential element does not require localization within the polyprotein-encoding region of the genome.

Foot-and-mouth disease, one of the most important known pathogens of livestock, is caused by a small RNA virus of the family *Picornaviridae*. Foot-and-mouth disease virus (FMDV) is the prototype member of the *Aphthovirus* genus of this family, and although aspects of FMDV replication resemble those of many other picornaviruses, there are notable differences between FMDV and other viruses that include FMDV’s broad host range and several unique genetic features.

The FMDV genome is over 8,300 bases in length and is covalently bound at its 5′ terminus to a 23- to 24-amino-acid genome-linked protein, 3B. The genome encodes three copies of 3B, all of which are apparently utilized (8, 15). No natural isolates of FMDV with fewer than three 3B coding regions have been identified, suggesting that there is a strong selective pressure maintaining this redundancy, since homologous recombination within the FMDV genome (22) should readily eliminate redundant copies of 3B. Despite the observed retention of three 3Bs in all natural FMDVs, viruses lacking one of the 3Bs or with two nonfunctional 3Bs have been derived by genetic engineering (7). Recently, we have observed that viruses lacking two of the three 3Bs can be generated by similar methodology, and these “single-3B” viruses display a reduced ability to replicate in fetal bovine kidney cells, although they grow well in BHK cells and porcine kidney cells (J. M. Pa- checo, T. M. Henry, and P. W. Mason, unpublished data).

Sequenceing of the FMDV genome revealed a 5′ untranslated region (UTR) of over 1,300 bases in length (9), longer than the UTRs found in the genomes of other picornaviruses, such as the enteroviruses (poliovirus, 740 bases [16]) and cardioviruses (encephalomyocarditis virus, 850 bases [6]). The FMDV UTR begins with a 360-base hairpin loop (the S fragment), which is separated from the remainder of the genome by a 150- to 250-nucleotide RNase T1-resistant tract containing approximately 90% C residues, with a small number of U and A residues (28). The poly(C) tract is followed by three to four tandemly repeated pseudoknots of unknown function (5) and the FMDV internal ribosome entry site (IRES) (4, 17).

Pilipenko et al. (30) derived a secondary-structure model for the 5′ end of the FMDV genome by analysis of published nucleotide sequences and chemical and enzymatic treatment of RNAs derived from the same region of the genome of encephalomyocarditis virus (which has an IRES element similar to that of FMDV). This model identified five domains in the 5′ UTR, including several IRES domains that have been shown to be important in binding host cell proteins, such as polyuridylic tract-binding protein (19, 20, 23, 31) and an FMDV IRES-specific *trans*-activating factor, ITAF45 (31). Figure 1 shows a structure of the 5′ end of the FMDV genome of serotype A12, derived from the model of Pilipenko et al. (30).

Several lines of investigation have indicated the importance of the highly structured 5′ end of picornaviral genomes in replication (1–3, 13, 38). However, McKnight and Lemon reported in 1996 the surprising result that an RNA structure within the P1-encoding region of the human rhinovirus 14 genome was required for replication (26). These studies dem-onstrated that this element was able to act only in cis, since

* Corresponding author. Mailing address: PIADC, USDA, ARS, NAA, P.O. Box 848, Greenport, NY 11944-0848. Phone: (631) 323-3177. Fax: (631) 323-2507. E-mail: pwmason@piadc.ars.usda.gov.
replicons lacking this portion of the genome could not replicate in the presence of helper viruses (26).

In a subsequent study, these investigators coined the term cis-acting replicative element (cre) and demonstrated that the cre was a short hairpin loop which contained the sequence AAACA in the unpaired loop that did not require translation to exert its activity (27). Since these reports, there have been reports of similar cre structures in the genomes of other picornaviruses. Surprisingly, these elements, although apparently responsible for the same function, are located in different regions of the genomes of these other viruses: 1B in the case of three different cardioviruses (18), 2C for poliovirus (11, 34), and 2A for human rhinovirus 2 (10). Analyses of cre activity of wild-type and mutant genomes indicate that this essential structure forms in the positive-strand genome, not in the negative-strand copy (11, 27).

Recently, work on poliovirus replication suggested that the cre serves as the template for 3Dpol-mediated uridylylation of 3B (29, 34), a critical step in RNA replication, since the pUpU-3B serves as the primer for RNA synthesis. These studies showed that the cre stimulated the uridylylation of 3B by the picornaviral polymerase (3Dpol). Furthermore, the authors of these studies suggested that the conserved AAACA present in all cres identified to date (see above) serves as a template for addition of the U residues to 3B (29, 34). Since uridylylation of 3B is a requisite for RNA replication, these studies help to

FIG. 1. Two-dimensional representation of a portion of the 5′ UTR of FMDV serotype A12. The structures of domains 2 to 5 are based on the model of Pilipenko et al. (30). The cre (cis-acting replicative element) and its conserved AAACA sequence (boxed) are described in the text.
explain why the cre is essential for replication, but they do not explain why the activity must be supplied in cis.

Here we report the presence of a cre structure in the 5′ UTR of FMDV in a region adjacent to the IRES that differs in location from a putative cre predicted to occur in the 2C region of the FMDV genome (37). Our cre structure was initially identified through its conservation among a number of different serotypes of FMDV, and its functionality was confirmed by examining the properties of viral genomes and replicons containing a panel of deliberately engineered cre mutations. Evaluation of the properties of the mutant genomes and replicons confirmed that the AAACA is absolutely required for FMDV genome replication but suggested that some changes in the stem structure could be tolerated. In addition, we found that replicating genomes and viruses could be recovered from synthetic genomes that contained an inactive (or deleted) cre in the natural 5′ position and a wild-type cre added at the 3′ end of the viral genome, downstream of the last coding region in the viral polypeptide.

**MATERIALS AND METHODS**

Cell lines, plaque assays, and cDNAs. Baby hamster kidney (BHK) cells, strain 21, clone 13 (American Type Culture Collection, Manassas, Va.), were maintained as previously described (33). These cells were used to propagate viruses and to perform plaque assays by standard techniques (33). Two genome-length cDNA clones encoding FMDV genomes for serotype A12 viruses [pRM2, poly(C) tract of 2 bases, and pRMCS35, poly(C) tract of 35 bases] and a heparan-sulfate binding serotype O1 chimaera created from pRMCS5 (pCRM4) have been described previously (33, 35).

RNA isolation, cDNA synthesis, PCR amplification, and sequencing. RNA was isolated from infected cell lysates, culture fluids harvested from infected cells, or individual viral plaques with Trizol (Life Technologies, Gaithersburg, Md.). Viral cDNAs were synthesized with Moloney murine leukemia virus reverse transcriptase (Life Technologies) with random hexamers as primers, and cDNA fragments were amplified by PCR (14) with specific oligonucleotides. PCR-amplified fragments were purified by with the Qiaquick kit (Qiagen, Valencia, Calif.) directly from the amplified fragments were purified by electroporation and from bands excised from agarose gels. Sequence analyses of these fragments were performed with selected primers and asymmetric amplification with Big-Dye terminators (ABI, Foster City, Calif.), followed by resolution on an ABI 3700 sequencer.

Construction of genome-length infectious cDNAs of type A12 containing the β-galactosidase gene. An FMDV replicon was created from the FMDV serotype A12 genome-encoding plasmid pRMCS5 by using the following strategy. A DNA fragment containing the chimeric codon β-galactosidase coding region was amplified by PCR (with Tag polymerase; ABI) from the pβgal-Control plasmid (Clontech, Palo Alto, Calif.) with a forward-sense oligonucleotide containing an EcoRI restriction endonuclease site present in codons 2 and 3 of the Lb form of the FMDV Lb protein, followed by one additional Lb codon and the first six codons of the β-galactosidase coding region, and an antisense oligonucleotide containing an artificial XmaI site (created by silent mutation at codons 17 and 18 for the 18-amino-acid 2A protein), the entire sequence of the 2A protein, a single Gln codon from the C terminus of 1D, and seven C-terminal codons of β-galactosidase. Following endonuclease digestion with EcoRI and XmaI, the fragment was inserted into one sites in a derivative of pRMCS5 containing the artificial XmaI site in 2A, replacing the Lb protein, 1A, 1B, 1C, and 1D coding regions. The resulting plasmid was sequenced through the junction regions, and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside staining of cells transfected with RNA transcribed from this plasmid (see below) confirmed that the replication specified the production of β-galactosidase.

This β-galactosidase-containing plasmid was further modified to contain a 3′ hepatitis delta virus ribozyme similar to one described previously (36). However, in this case the cleavage site was positioned to coincide precisely with the last base of the 15-nucleotide poly(A) tract, and a unique SwaI site was inserted downstream of the ribozyme. Finally, a 5′-end hammerhead ribozyme identical to one used by Herold and Andino [12] to enhance replication by poivivirus replicons] was added to the 5′ end of the FMDV S fragment of our β-galacto-

RESULTS

Analyses of sequence data from the L fragment [poly(C) to poly(A)] of the genomes of SAT2/KEN/3/57 (GenBank accession no. A251473) and O1 Kaufbueren (GenBank accession

Downloaded from http://jvi.asm.org/ on October 14, 2017 by guest
no. X00871) and the complete genomes of C3 Argentina (GenBank accession no. AJ007347), A12 (GenBank accession no. M10975 and L11360), and O/TAW/97 (GenBank accession no. AF308157) revealed the presence of several AAACA motifs in the genomes of these viruses. However, among these genomes, there was only a single AAACA that was common to all, and it was located following the pseudoknot region of the A12 genome [nucleotides 242 to 246 following the poly(C) tract of infectious cDNA plasmid pRMC35]. Furthermore, folding of this segment [nucleotides 220 to 273 following the poly(C) tract] with mFold program version 3.1 (25, 40) produced a stem-loop structure with an exposed loop containing the AAACA for all of these genomes (results not shown).

A schematic diagram depicting the putative cre of the A12 genome is shown in Fig. 1 along with domains 2, 3, 4, and 5 drawn as derived by Pilipenko et al. (30). Interestingly, the descending strand of the stem of the cre element overlaps structural domain 1 of the 5' UTR predicted by Pilipenko et al. (30) based on computer modeling and biochemical analyses of the homologous cardiovirus 5' UTR. The domain 1 defined by Pilipenko et al. (30) has been previously shown to have little influence on the translational activity of the FMDV IRES in vitro (17), but it is often included in published representations of the FMDV IRES.

To analyze the function of the putative cre in RNA replication in the absence of viral replication and transmission, we used the β-galactosidase-expressing replicon RNA shown in Fig. 2. Evaluation of this β-galactosidase replicon in electroporated BHK cells showed that the replicon produced large amounts of β-galactosidase, with a peak in activity detected approximately 12 h postelectroporation (Fig. 3). This β-galactosidase activity was substantially reduced by the addition of guanidine HCl, which blocks FMDV multiplication (32) through inhibition of RNA synthesis, consistent with inhibition
not able to conclude that all of the residual of replication with no effect on translation. However, we were wild-type (WT) and mutant replicons in the pA12-RNAs electroporated into BHK cells and truncated replicon-derived and loss of β-galactosidase-expressing genomes, which results in their detach-
ment from the monolayer and loss ofactivity at 12 h postelectroporation is likely the result ofC35 replicon (results not shown). The fall-off in other experiments showed a slightly better performance by the replicon replicated better than the pA12-gal replicon, butgal replicon, but since we did not determine if the level of guanidine HCl used completely inhibited replication.
In the same experiment, we included a replicon containing a poly(C) tract of only two C’s (pA12-C2-gal), since our wild-type replicon contains a poly(C) tract of only 35 residues and the poly(C) tracts of synthetic FMDV RNAs elongate during passage in culture (33, 39). The C2 replicon replicated to levels similar to those of the replicon containing 35 C’s, suggesting that a very long poly(C) tract is not required for genome replication, consistent with studies showing that poly(C) tract length affects the speci-}

dase activity in BHK cells transfected with plasmid DNAs encoding truncated genomes (Fig. 4). These data clearly show that the mutations in the cre had only modest effects on the translational activity (Fig. 4), demonstrating that these mutations exerted their effects at the level of genome replication.

The same panel of cre mutations were inserted into full-length genomes specifying a cell culture-adapted virus to further evaluate the importance of this segment on genome replication. Recovery of virus from cells electroporated with these genomes showed that by the second or third passage of lysates from electroporated cells, live virus could be recovered from all mutant genomes except for the double mutation (ACCCA, mutant 6; results not shown). Evaluation of the cre sequences from the recovered viruses revealed a reversion to the wild-type AAACA for all point mutants in our panel (mutants 1, 3, 4, and 5; results not shown). Evaluation of the sequences of the cre from the viruses recovered from the first passage of the “stem” mutants 2a and 2b revealed that the mutated bases and their predicted partners in the putative stem-loop structure shown in Fig. 2 were present as mixtures.

To investigate the mixtures of viruses in the first passage recovered from cells transfected with RNA from pCRM4H mutant 2a, the BHK passage 1 material obtained from cells transfected with RNA containing mutation 2a was used as a source of viruses for plaque purification. Following purification, the 5’ UTRs of six plaque-picked viruses were sequenced through a 475-base region commencing 15 bases downstream from the poly(C) tract (encompassing the 4 pseudoknots, the cre, and most of the IRES of these genomes). These sequence analyses revealed that one of the six plaque-purified isolates contained a mixture of genomes, while the remaining five contained one predominant genotype (no mixtures were detectable at any position).

The cre sequences detected in these five plaque isolates are shown in Fig. 5. In the case of plaques 2a.2 and 2a.4, identical sequences were revealed in the cre, although one of these isolates had point mutations within the pseudoknot region of the genome, indicating that they were separate clones. The other four plaque isolates had wild-type sequences throughout the sequenced region except for the substitutions shown in Fig. 5. Evaluation of the structures shown in Fig. 5 suggests that the stem-loop structure is a critical requirement for viral replication. Interestingly, one of the recovered plaques (plaque 2a.3) had four U residues in the mutated region, which does not appear to produce a strongly base-paired stem. However, computer folding (mFold) (25, 40) of a larger RNA sequence, extending from pseudoknot 4 through domain 2, indicated that the favored conformation of the wild-type and all four mutant 2a revertants shown in Fig. 5 formed the cre stem-loop, whereas mutants 2a and 2b (Fig. 2) did not (results not shown).

To determine how fit these mutant plaques were, we performed plaque assays on BHK cells with the same stock of virus that was used to obtain the sequence data shown in Fig. 5. Photographs of these plaques are shown in Fig. 6. All of the revertant viruses produced plaques that were considerably smaller than those obtained with wild-type virus, although mutant 2a.1, predicted to have the most stable structure at the base of the loop (Fig. 5), formed the largest plaques (Fig. 6).

To determine if the 5’ UTR localization of the cre was a requirement for FMDV replication, we undertook a study to
determine if we could translocate the element to the 3' end of the polyprotein. To this end, a wild-type form of the cre element was inserted immediately following the stop codon found at the end of the viral polyprotein (Fig. 7) in replicons and genomes containing either the wild-type 5' UTR cre or a mutant 6 5' UTR cre. Both replicons were able to replicate well compared to wild-type virus (51% and 34% of wild-type activity, respectively, measured as shown in Fig. 4), strongly suggesting that the 3' UTR cre was functioning, since the mutant 6 cre was completely inactive in the replicon context (Fig. 4). Studies with the double cre-containing viral genomes showed similar results. Viruses were readily harvested from the genomes containing the wild-type form of the cre in the 3' UTR and either the wild-type or nonfunctional mutant 6 in the 5' UTR. More importantly, we found that viable viruses with a complete deletion of the cre element could be created (mutant 7 [Fig. 7]) if a cre was added at the start of the 3' UTR (Fig. 7). Viruses recovered from cells transfected with these 3' cre viruses, containing either a double-mutated 5' cre (mutant 6) or no 5' cre (mutant 7), were able to form plaques on BHK cells, although the plaques were considerably smaller than those produced by the wild-type virus (Fig. 8). Sequence analyses of the double cre viruses revealed that both maintained the 3' cre for two passages without any modification. Additionally,

FIG. 5. Predicted structures of the cre of five different plaque isolates recovered from cells transfected with a viral genome encoding cre mutant 2a. Substitutions relative to the transfected RNA are shown in lowercase letters and identified with arrowheads (wild-type and mutant 2a structures are shown at the left for comparison).

FIG. 6. Photographs of plaques of the indicated viruses on BHK cell monolayers. Plaques from the wild-type virus (derived from pCRM4H) were produced by the second passage from the electroporated stock. Plaques from the plaque-picked isolates were produced by the BHK-amplified plaques picked from the first passage postelectroporation (this is the same passage used for the sequence analyses; see text for details).
DISCUSSION

Upon entering a susceptible host cell, the picornavirus genome must be translated by host ribosomes, and the resulting protein products must then be used to produce the negative-sense form of the genome that is needed for subsequent pro-

FIG. 7. Schematic diagram of replicon and infectious FMDV genomes showing the positions of the cre deletion (mutant 7) and the insertion of the cre at the 3′ end of the genome. For this portion of the figure, lowercase letters indicate specific mutations to the genome to facilitate construction, and underlined lowercase letters indicate bases added to the genome to facilitate construction. Other abbreviations and nomenclature are explained in the legend to Fig. 2.

FIG. 8. Photographs of plaques produced by the indicated viruses on BHK cell monolayers. Plaques for all three viruses were produced by virus recovered from the second passage of lysates of cells transfected with synthetic genome-length RNAs by using Lipofectin (see the text).
duction of the positive-stranded genomic RNA. An orchestration of the initial translation and replication activity is likely to be required to ensure efficient viral replication, and hence these viruses must have evolved a method to ensure that translation and replication (which take place in different directions on the infecting genome) do not interfere with each other (3, 13). The fact that picornavirus genomes are infectious in the absence of other viral factors supports the assumption that factors controlling this orchestration are likely to reside in primary or secondary RNA structures. However, it is clear that these structural elements are altered by their interaction with cellular proteins, which are known to influence translation (see the introduction).

The cre has been identified in several picornavirus genomes, but to date it has always been found within the polyprotein-encoding region (see the introduction). However, data from McKnight and Lemon (27) indicated that translation of the sequences encoded by the human rhinovirus 14 cre element was not a requirement for genome replication. Furthermore, McKnight has recently shown that a replicating human rhinovirus 14 genome with the cre moved to the 3’ end of the genome can be obtained (K. McKnight, personal communication). These data argue against a model recently proposed by Barton and coworkers, which suggests that the localization of the cre within the open reading frame contributes to the ability of the virus to control translation and replication by ensuring that the cre does not function to provide uridylylated 3B until the polyprotein-encoding region of the genome is cleared of translating ribosomes (3). Although the translation of the cre may serve as a source of this type of control for other picornaviruses, the same function must be provided by a different mechanism for FMDV, since the FMDV cre element was not a requirement for genome replication. Furthermore, McKnight has recently shown that a replicating human rhinovirus 14 genome with the cre moved to the 3’ end of the genome can be obtained (K. McKnight, personal communication).

On the surface, our ability to translocate the cre element from one end of the genome to the other appears to be a dramatic accomplishment. However, models for genome replication indicate that the 5’ and 3’ ends of the picornavirus genome are noncovalently linked to each other by host cell and viral proteins to facilitate efficient translation and replication (3, 13). Furthermore, Lyons et al. have recently demonstrated that the 5’ end of the poliovirus genome is required in cis for efficient VPg uridylylation and negative-strand RNA synthesis, suggesting that the 5’ and 3’ ends of the genome must interact at this crucial step in the replication cycle (21). If there is a similar requirement for bringing together the 5’ and 3’ ends of the FMDV genome, then our ability to translocate the cre may not result in significant movement of the cre from the three-dimensional context in which it normally functions.

Data obtained in our characterization of mutated forms of genetically engineered replicons, as well as the phenotypes and genotypes of viruses recovered from mutated forms of genetically engineered viral genomes, confirmed that the FMDV cre is required for replication. Interestingly, the cre is found adjacent to the IRES, suggesting that this location could be important for orchestrating the relationship between translation and replication, although this mechanism is apparently not used by other picornaviruses which have the cre located within the polyprotein-encoding portion of their genomes (see above). Although our translation studies indicated a minimal effect of the cre mutations (including a complete deletion) on translation, it is possible that there is a subtle interaction between these contiguous replication and translation elements that is important for efficient viral replication and spread in the animal host. This supposition is consistent with the small-plaque phenotype of viruses in which the cre has been translocated to the 3’ end of the genome.

ACKNOWLEDGMENTS

We thank E. Wimmer, A. Paul, and E. Rieder (SUNY-Stony Brook, Stony Brook, N.Y.) and E. Pilipenko (University of Chicago, Chicago, Ill.) for helpful suggestions in completion of this work, and we thank K. McKnight (Eli Lilly and Co., Indianapolis, Ind.) for supplying unpublished data. We thank E. Rieder and V. O’Donnell (PIADC) for carefully reading the manuscript; B. Bissel, G. Ward, J. Gregory, and S. Boettcher for technical assistance; and C. Beard for construction of the cytomegalovirus-driven T7 plasmid (pCB-T7).

This work was partially supported by the Agricultural Research Service of the USDA (CRIS Project 1940-32000-035-00D) and by a grant from the National Research Initiative Competitive Grants Program of the USDA-CSREES (grant 99-35204-7949).

REFERENCES


