Utilization of Heterologous Alphavirus Junction Sequences as Promoters by Sindbis Virus

JENNIFER M. HERTZ AND HENRY V. HUANG*

*Department of Molecular Microbiology, Washington University School of Medicine,
Box 8230, 660 South Euclid Avenue, St. Louis, Missouri 63110-1093*

Received 10 September 1991/Accepted 24 October 1991

We used Sindbis virus, an alphavirus, as a model to study the evolution of the recognition of viral *cis*-acting sequences. During the life cycle of alphaviruses, a full-length minus-strand RNA is made and serves as a template for both genomic RNA replication and subgenomic mRNA transcription. Transcription initiates at an internal promoter site, the junction sequence, to produce a subgenomic mRNA. The junction sequences of alphaviruses are highly conserved, but they do contain a number of base differences. These could have been essentially neutral mutations during evolution, such that any of the contemporary sequences can be recognized efficiently by any of the alphaviruses. Alternately, the changes could have resulted in significant functional divergence, such that the contemporary viruses can no longer recognize heterologous junction sequences as promoters. To distinguish between these possibilities, we constructed Sindbis virus derivatives with two subgenomic mRNA promoters. One is the wild-type Sindbis virus promoter used for expression of the structural proteins. The other is either the minimal Sindbis virus promoter or the corresponding junction sequences from other alphaviruses, which are placed upstream of the bacterial chloramphenicol acetyltransferase (CAT) gene. RNA analyses were used to determine the relative promoter strengths of the various junction sequences. The results showed that all but two were recognized as promoters by Sindbis virus. CAT enzyme assays were used to measure the accumulation of CAT protein made from mRNAs transcribed by using the heterologous junction sequences as promoters. Most of the viruses expressed amounts of CAT enzyme within 10-fold of each other. The two viruses with junction sequences that were not recognized as promoters did not give significant CAT expression. We conclude that, with respect to Sindbis virus, the junction sequences are functionally conserved; i.e., most of the contemporary nucleotide differences in the junction sequences are neutral or near-neutral mutations. The functional conservation suggests that neither the *cis*-acting sequence nor the cognate binding site of the transcription factor can change independently. This type of coupled evolution between *cis*-acting sequences and their cognate viral protein binding sites may be a general phenomenon. For example, it explains the ubiquitous presence of conserved *cis*-acting sequences in each of the families of RNA viruses. There are implications of this hypothesis for the design of antiviral drugs.

Sindbis virus (SIN) is a positive-strand RNA enveloped virus in the *Alphavirus* genus of the family *Togaviridae*. In the infected cell, the viral genome is translated to produce two polypeptides that are cleaved into four nonstructural proteins. The nonstructural proteins (perhaps in conjunction with host factors) use the genomic RNA as a template to synthesize a minus strand. This minus strand is in turn used as a template by the nonstructural proteins, and possible host factors, for replication and transcription. Replication produces a full-length genomic 49S RNA. Transcription initiates at an internal site, the junction region, to produce a subgenomic 26S mRNA that is colinear with the 3' one-third of the genomic RNA and encodes the viral structural proteins.

The junction region contains a conserved sequence extending from 19 nucleotides (nt) upstream to 5 nt downstream (-19/+5) of the transcription start site of the subgenomic mRNA. This conserved sequence was proposed to be the promoter for subgenomic RNA synthesis of the alphaviruses (37). The junction sequence is not found in naturally occurring defective interfering (DI) genomes, but when it was cloned into a DI genome it produced a subgenomic RNA (28). The DI genome-derived subgenomic RNA initiated at the same position as wild-type virus. Deletion mapping of

the junction region showed that the minimal sequences necessary for transcription of a DI genome-derived subgenomic RNA was -18 or -19 to +5, where +1 is the start site (28). This minimal sequence encompasses the conserved junction sequence (37) and functions properly in both DI genomes and SIN as a promoter, although it is less active than, for example, a -98/+14 promoter (28, 41).

The promoter is probably recognized by viral nonstructural proteins, perhaps in conjunction with host factors. Indeed, all four of the nonstructural proteins, and possibly the polyprotein precursors, are candidate transcription factors. Both in vivo and in vitro evidence indicates that the nsP4 protein is at least part of the viral polymerase (1, 17). Late in infection, the nsP34 fusion protein may be the polymerase for plus-strand synthesis (6, 19). The nsP2 protein is suspected to be involved in transcription, since a temperature-sensitive mutation that maps in nsP2 alters the relative amounts of subgenomic mRNA versus genomic RNA produced in infected cells (18, 22, 45, 46). The subgenomic mRNA, like the genomic RNA, has a 5' m⁷G cap, and a methyltransferase activity that might participate in cap methylation has been mapped to nsP1 (32, 33, 34). It is not known which of the nonstructural proteins are actually transcription factors and whether host factors are also involved. A genetic approach has been used to identify the viral transcription factors. A mutant virus with a 3-nt insertion at position -4 of the promoter made drastically reduced

* Corresponding author.

A. JUNCAT

B. viruses with two promoters

FIG. 1. (A) Structure of the JUNCAT plasmid used as an intermediate for making viruses with two promoters. The plasmid has a polylinker region followed by CAT and the 3' untranslated region (3'UT) of SIN downstream from CAT. The heterologous junction sequences corresponding to the SIN minimal promoter were synthesized as double-stranded oligonucleotides and inserted between the Asp 718 and XbaI restriction sites. The arrow represents the start site for the CAT subgenomic RNA if the junction sequence was active as a promoter. The *Apa*I and *Pvu*I sites were used to construct plasmid clones of viruses with two junction regions (see Materials and Methods). (B) Structures of SIN derivatives with two promoters. The viruses consist of the SIN genome up to the end of the E1 gene. They contain the genes for the SIN nonstructural (NS) and structural (C, E2, E1) proteins, and they include the SIN wild-type promoter for transcription of the subgenomic mRNA encoding the structural proteins. The viruses have the junction sequences of various alphaviruses (5, 23, 37, 48, 51) inserted immediately after the E1 gene. The junction sequences were followed by the bacterial CAT gene and the 3' untranslated region of SIN. The arrows represent the start sites for the subgenomic RNAs. Each SIN derivative was named according to the source of the junction sequences inserted. These sequences (minus strand) are aligned with the RRV junction sequence as the prototype. A dash indicates sequence identity with the RRV sequence. In mpSNRV, the SNRV sequence is the SIN junction sequence but with the U+5 of the RRV junction sequence. Similarly, the ONN* sequence is the ONN junction sequence but with A-19. The SFV-1 junction sequence (51) is identical to the SFV junction sequence (37) but with C-1. The RUB sequence is a rubella virus sequence that resembles the SIN minimal promoter (12, 13). The 5' and 3' sides of the junction sequences are flanked by 4-nt sequences that do not occur in the corresponding positions in any of the alphaviruses. This provides a constant local context and eliminates any fortuitous match of the flanking sequences with any of the alphavirus junction sequences. The mpNO virus does not contain any inserted junction sequences. The junction sequences in the other viruses were inserted into the site indicated by the vertical arrow below the mpNO sequence. The nsP4 termination codon in the EEE and VEE sequences is at positions -8 to -6. The VEE junction sequence is identical to that of EEE.

amounts of subgenomic mRNA (15). The junction regions of 24 revertants of the mutant were sequenced, and all retained the 3-nt insertion (15); thus, they must have had second-site mutations that at least partially suppressed the mutant phenotype. Two of the revertants were able to recognize the mutant promoter, *in trans*, in a DI genome. Therefore, they have suppressor mutations in one or more *trans*-acting viral proteins that are involved in promoter recognition.

SIN is the type species of the alphaviruses, and the other alphaviruses are expected to have transcription mechanisms similar to that of SIN. The alphaviruses have a worldwide distribution (3). SIN is endemic in the Old World. Other alphaviruses whose junction sequences have been determined (5, 8, 23, 37, 48, 51) include Ross River virus (RRV), found in Australia and the South Pacific region; Semliki Forest virus (SFV) and Middelburg virus (MBV), found in Africa; O'Nyong-nyong virus (ONN), found in Africa and Asia; the eastern equine encephalitis virus (EEE), found in the New World and Old World; and the western equine encephalitis virus (WEE) and Venezuelan equine encephalitis virus (VEE), found in the New World. During the evolutionary divergence of the alphaviruses, a number of base changes have been fixed in the conserved junction sequence (Fig. 1B). The functional significance of these differences is not known. In principle, viruses with deleterious mutations that gave inappropriately decreased or increased promoter strengths would have been selected against during evolution, and these deleterious mutations will not be present in the contemporary promoters. However, some of the deleterious mutations could have been retained if their phenotypes were suppressed by coincidentally appearing suppressor mutations. One type of suppressor mutation is a change in the promoter-binding sites of the cognate transcription factors that restores promoter recognition. If the contemporary alphavirus promoters are derived from these kinds of events, then promoter recognition could have diverged so much that it is now essentially virus specific. On the other extreme, suppressor mutations might be rare, and the evolution of the promoter and the binding sites of the transcription factors might have been so constrained that any changes were limited to neutral mutations that did not significantly alter promoter function. In this case, any of the contemporary promoter sequences should be recognized efficiently by any of the alphaviruses. These possibilities may be distinguished by assaying the ability of any given junction sequence to be used as a promoter by different alphaviruses and by assaying for the use of heterologous junction sequences as promoters by any given alphavirus. We describe here the use of the latter approach to show that heterologous junction sequences are used remarkably well as promoters by SIN.

MATERIALS AND METHODS

General recombinant DNA materials and methodology. Restriction enzymes and SP6 DNA-dependent RNA polymerase were obtained from New England BioLabs, Inc., Bethesda Research Laboratories, or Boehringer Mannheim Biochemicals. Radioactive materials were obtained from Amersham Corp. or ICN Pharmaceuticals, Inc. Enzymes were used essentially as recommended by the manufacturers. Plasmids were grown, purified, and analyzed by standard methods, with minor modifications (30).

Plasmid clones. The plasmid JUNCAT was constructed as a convenient intermediate for inserting heterologous alphavirus junction sequences into the SIN genome. The salient

feature of JUNCAT (Fig. 1A) is a polylinker sequence immediately upstream of a module from TRCAT (53). The TRCAT module cloned into JUNCAT starts from the *Xba*I site in the 26S 5' untranslated sequences and consists of the bacterial chloramphenicol acetyltransferase (CAT) gene plus the 3' untranslated region of Toto1000 (from positions 11088 to 11703 [43]) and the poly(A) tail. Also, JUNCAT contains the β -lactamase gene for selection after transformation. The *Asp* 718 and *Xba*I sites of the polylinker were used to clone in synthetic oligonucleotides cDNA sequences. The oligonucleotides consist of junction sequences (-19/+5) of various alphaviruses, corresponding to the minimal promoter of SIN (28), flanked at the 5' and 3' boundaries by 4 nt that are not present at the corresponding positions in any of the alphavirus junction sequences (Fig. 1B), as well as by an *Asp* 718 site and an *Xba*I site 5' and 3' of the junction sequences, respectively. The *Apa*I-to-*Pvu*I (in the β -lactamase gene) fragments of the JUNCAT derivatives were then cloned between the unique *Apa*I site and the *Pvu*I site in the β -lactamase gene of a Toto1101 derivative (16; the *Apa*I site was inserted 1 nt downstream from the stop codon of the E1 structural protein gene). All of the final clones therefore have two junction regions, one that is the SIN promoter upstream of the structural protein genes and a second (-19/+5) that is upstream of the CAT gene. Each SIN derivative was named according to the source of the junction sequences inserted (Fig. 1B). The sequences of the inserted junction regions of all clones were verified by using the dideoxynucleotide chain termination method with Sequenase version 2.0 (United States Biochemical). The primer used corresponds to nt 15 to 32 of the 5' end of the CAT gene, where nt 1 is the initiation codon of CAT.

RNA transcription and transfection. 5'-capped RNA transcripts were synthesized from DNA samples linearized with *Sst*I and transcribed in vitro with SP6 DNA-dependent RNA polymerase (15) at 38°C for 45 min. The template DNA was not removed before transfection. BHK-21 cells in 35-mm dishes (~10⁶ cells) were transfected with various amounts of RNA transcripts (100 ng to 2 μ g) mixed with 8 μ g of lipofectin (Bethesda Research Laboratories) (15). After a 10-min incubation, the cells were washed once with phosphate-buffered saline minus Ca²⁺ and Mg²⁺ (PBS) and once with medium (minimal essential medium with Earle's salts plus 10% fetal calf serum). At 10 h posttransfection, the medium was replaced with fresh medium. Virus-containing medium was collected from 24 to 30 h posttransfection. Virus titers (typically 5 $\times$ 10⁸ to 2 $\times$ 10⁹ PFU/ml) were determined with plaque assays using BHK-21 cells. Viruses from the first passage were used in all the experiments.

RNA analyses. BHK-21 cells were infected by each virus at a multiplicity of infection of 10. Dactinomycin was added immediately after infection. The cells were incubated with [³H]uridine from 4.5 to 6.5 h or 5 to 7 h postinfection, at which time intracellular RNA was isolated (28), and then extracted and precipitated (52). The RNA was denatured with glyoxal and dimethyl sulfoxide and electrophoresed on 1% agarose gels (4). The gels were treated for fluorography (26) and exposed to film for 24 h to 1 week, and the autoradiographs were used as guides to cut out the two subgenomic RNA bands in each sample for quantitation by liquid scintillation counting. The background counts were determined from the average counts of two slices, above and below each band. The molar ratio of CAT mRNA to structural protein mRNA was then calculated by dividing the counts in the CAT mRNA band by the counts in the

structural protein mRNA band and then multiplying the result by 3.62 to correct for the mRNA size difference.

Primer extension analyses. The primer (the same synthetic oligonucleotide used for sequencing; see above) was 5' phosphorylated by using T4 DNA kinase and equal amounts (22 pmol) of [γ -³²P]ATP and the primer. EDTA was added to a final concentration of 2 mM to stop the reaction, and the sample was boiled for 2 min and chilled on ice. RNA samples were added to 0.35 pmol of kinase-treated primer and ethanol precipitated with 20 μ g of glycogen per ml and a final concentration of 0.1 M NaCl. The pellets were redissolved in H₂O, heated to 98°C for 1 min, and annealed at 37°C for 30 min in the presence of 0.5 U of Inhibit-ACE (5 Prime $\rightarrow$ 3 Prime Inc.) and 80 mM KCl, in a final volume of 9 μ l. The annealing mix was then adjusted to 50 mM Tris-HCl (pH 8.3)-3 mM MgCl₂-10 mM dithiothreitol-1 mM deoxynucleotides-200 U of mouse mammary leukemia virus RNase H⁻ reverse transcriptase (Superscript; Bethesda Research Laboratories) in a final volume of 31 μ l. The reaction mixes were incubated for 10 min at 37°C and then 1 h at 42°C. EDTA and glycogen were added at the end of the reaction to final concentrations of 2 mM and 20 μ g/ml, respectively; the samples were extracted once with phenol and once with chloroform and then ethanol precipitated. The pellets were resuspended with 3 μ l of H₂O, and 6 μ l of 80% formamide-dye sample solution was added. The samples were heated at 95°C for 5 min, cooled on ice, and loaded onto a 6% polyacrylamide sequencing gel.

CAT assays. Monolayers of BHK-21 cells were infected at a multiplicity of infection of 10. At 5, 7, and 9 h postinfection, the cells were washed three times with PBS. The cells were scraped off the 35-mm dishes into 1 ml of PBS and transferred to a microcentrifuge tube. The tubes were centrifuged, and the cell pellets were resuspended with 50 μ l of 0.25 M Tris-HCl (pH 8). After 3 freeze-thaw cycles, the tubes were centrifuged and the supernatants were used for the CAT assay (14). Appropriate dilutions of the samples were made to ensure that chloramphenicol acetylation was in the linear range of the assay. One unit of CAT was defined as the amount of enzyme that acetylates 1 μ mol of chloramphenicol per min at 37°C (47). The specific activity of CAT was assumed to be 125 U/mg (47). Where necessary, the protein content was determined by using the Bio-Rad protein concentration kit.

RESULTS

Heterologous junction sequences as promoters for SIN subgenomic mRNA synthesis. SIN derivatives with two promoters were constructed (see Materials and Methods) to determine whether the viral transcription factors can use heterologous junction sequences as promoters. The homologous or heterologous junction sequences were inserted immediately downstream of the E1 gene in a cDNA clone of the SIN genome (Fig. 1A). The inserted sequences (Fig. 1B) were either the SIN minimal -19/+5 promoter sequence (28), the corresponding junction sequences (-19/+5) from other alphaviruses (5, 8, 23, 27, 37, 48, 51), or a sequence from rubella virus (RUB) that resembles the SIN minimal promoter (12, 13). A derivative, mpNO (Fig. 1B), with no exogenous sequences was also made to serve as a negative control. Each SIN derivative was named according to the source of the junction sequence inserted (Fig. 1B). The junction sequence is followed by the bacterial CAT gene. This allows us to use CAT assays to measure the amount of protein that can be expressed when the various junction

FIG. 2. Intracellular viral RNA synthesized by the double-promoter Sindbis virus derivatives. Intracellular viral RNA was labeled with [3 H]uridine and isolated from BHK-21 cells 6.5 to 7 h after infection with each virus. The RNA was denatured and electrophoresed on 1% agarose gels. The gels were treated for fluorography and exposed to film. The lanes are labeled according to the junction sequence in each virus. (A) Genomic RNA and structural protein mRNAs, synthesized by the SIN nonstructural proteins by using the various junction sequences as promoters; 24-h exposure. (B) The CAT subgenomic mRNAs, synthesized by the SIN nonstructural proteins by using the various junction sequences as promoters; 1-week exposure of the same gel. See Fig. 1B for the junction sequences of the viruses. Lane MOCK contains RNA from mock-infected cells.

sequences are used as promoters. The cDNA clones were linearized with *Sst*I and used for runoff transcription with SP6 DNA-dependent RNA polymerase. The capped transcripts were transfected into BHK-21 cells, and virus stocks were collected 24 to 30 h posttransfection. These virus stocks were then used for all experiments described here.

All of the viruses derived from the cDNA clones are expected to make a 5.25-kb mRNA that initiates at the normal SIN promoter. This mRNA is bicistronic, encoding the SIN virus structural proteins followed by the CAT sequences. If the inserted junction sequence is functional as a promoter, the virus is expected to make an additional subgenomic mRNA that is 1.45 kb in length and encodes CAT. All viruses produced genomic RNA and structural protein mRNA of the expected sizes (Fig. 2). Most viruses produced a detectable second subgenomic mRNA whose size was that expected for the CAT mRNA. To facilitate comparisons among the viruses, the amount of the CAT mRNA was normalized to the amount of the structural protein mRNA made by each virus (see Materials and Methods). It should be noted that these molar ratios (Table 1) are results from a 2-h labeling period and do not represent accumulated mRNA.

The SIN minimal promoter of the mpSIN virus gave a molar ratio of 0.12 of CAT mRNA relative to the structural protein mRNA (Table 1; Fig. 2), suggesting that the minimal promoter was used about one-eighth as efficiently as the wild-type promoter. This is due to a combination of the relatively weaker strength of the minimal promoter and a context effect (41). The virus that lacks a minimal promoter, mpNO, gave an apparent mRNA molar ratio of 0.04, which is defined as the background. The junction sequence from RRV gave the most CAT mRNA. Remarkably, it was about twofold stronger than the SIN minimal promoter. The junction sequence of MBV had a promoter strength comparable to that of SIN, while the ONN, EEE, VEE, and SFV

TABLE 1. Relative promoter strengths of different alphavirus junction sequences

Virus	Molar ratio ^a	Relative promoter strength ^b
mpSIN	0.12	1.00
mpRRV	0.21	2.13
mpSNRV	0.16	1.53
mpMBV	0.11	0.88
mpSFV	0.10	0.75
mpONN	0.10	0.75
mpEEE	0.09	0.63
mpVEE	0.08	0.53
mpSFV-1	0.05	0.13
mpRUB	0.05	0.13
mpNO	0.04	0.00

^a Moles of CAT mRNA/mole of structural protein mRNA (see Materials and Methods).

^b Strength of each promoter relative to mpSIN, with the mpNO background subtracted.

junction sequences appeared to be 50 to 90% as active as the minimal promoter of SIN. The mRNA molar ratios of mpSFV-1 and mpRUB viruses were essentially like that of mpNO. The SFV-1 junction sequence (51) is identical to the sequence reported by Ou et al. (37) except that a U is replaced by a C at position -1. This single nucleotide change made the SFV-1 junction sequence much less active as a promoter than that of SFV.

Initiation sites of mRNA transcribed from heterologous junction sequences. The initiation site for the mRNA produced from each junction sequence was mapped by primer extension. All but mpSFV-1 and mpRUB gave major bands of primer extension products that paralleled the amount of CAT mRNA produced by each virus. The pattern of the bands is similar to that of the primer extension products from the 26S mRNA of SIN (15), and they are consistent with mRNA initiation at nucleotide position +1 (Fig. 3). The bands at position -1 are due to reverse transcriptase copying the cap. The RRV, MBV, SFV, and SFV-1 extension products all gave extra bands centered at position -6 that do not correlate with the amount of CAT mRNA made by the corresponding viruses. These bands probably do not represent 5' ends of the CAT mRNA but are most likely reverse transcriptase artifacts caused by local sequence or structure. Consistent with this view, these junction sequences are identical from -2 to -19 (Fig. 1B). Similarly, extension products from the mpRUB virus intracellular RNA are likely to be reverse transcriptase artifacts.

CAT expression with use of heterologous junction sequences as promoters. During evolution, the optimum promoter activity is determined, at least in part, by the amount of viral structural proteins needed for production of progeny virions. Therefore, it is important to measure the amount of protein that can be made when the heterologous junction sequences are used. The viruses were constructed to have the junction sequences upstream of the bacterial CAT gene (Fig. 1). The accumulated CAT enzyme activities were measured at 5, 7, and 9 h postinfection. The mpRRV virus gave the highest levels of CAT expression, from 2.0- to 2.7-fold higher than the levels produced by mpSIN (Fig. 4). The mpSIN, mpMBV, and mpEEE viruses expressed comparable amounts of CAT. Their CAT expression is comparable to or slightly higher (at 9 h) than that of the mpONN, mpSFV, and mpVEE viruses. The ONN and ONN* junction sequences differ only at position -19 (Fig. 1B), G and A, respectively,

FIG. 3. Primer extension analyses of the initiation site of the CAT mRNAs. A primer hybridizing to the CAT sequences was used to prime cDNA synthesis on equal amounts of intracellular RNAs isolated from cells infected with each virus. The primer extension products were electrophoresed on a sequencing gel. The lanes are labeled according to the junction sequence in each virus. The second and third lane show two separate reactions for mpSIN. Lanes labeled G, A, C, and T (two lanes) represent a sequence ladder of the mpSFV cDNA clone, using the same primer to facilitate identification of the initiation sites. The arrow indicates the SIN 26S mRNA start site at position +1. Lane mock contains extension products from mock-infected cells.

and they gave similar CAT expression (results not show). The mpSFV-1 and mpRUB viruses gave almost a 100-fold lower CAT expression than did mpSIN. The very low levels of CAT expressed by mpNO, which has no second pro-

FIG. 4. Comparison of CAT expression from mRNA synthesized with use of the various junction sequences as promoters. The amount of CAT trimers per cell was determined for each virus at 5, 7, and 9 h postinfection. The symbol used to plot the data for each virus is indicated at the right side. For clarity, some of the symbols representing the same time point were displaced horizontally from each other. Error bars indicate ± 1 standard error of the mean.

TABLE 2. CAT expression with use of the hybrid SNRV promoter

Time (h) postinfection ^a	Virus	Mean CAT activity (10^4 U) ^b	Relative CAT expression ^c
5	mpSNRV	3.0 (0.7)	3.0
	mpRRV	4.6 (1.6)	4.6
	mpSIN	1.0 (0.1)	1.0
	mpMBV	1.4 (0.3)	1.4
7	mpSNRV	9.9 (3.2)	2.2
	mpRRV	9.3 (2.9)	2.1
	mpSIN	4.5 (0.7)	1.0
	mpMBV	4.1 (1.2)	0.9
9	mpSNRV	14.0 (0.7)	1.8
	mpRRV	9.5 (3.0)	1.3
	mpSIN	7.6 (1.7)	1.0
	mpMBV	5.0 (1.4)	0.7

^a Extracts of BHK-21 cells infected by each virus were prepared at 5, 7, and 9 h postinfection (see Materials and Methods).

^b CAT expression from four independent experiments was normalized to the protein content of each sample. The standard error of the mean is given in parentheses.

^c CAT expression relative to that of mpSIN at each time point.

moter, could be due to a small amount of translational reinitiation or internal initiation of the bicistronic structural protein mRNA. It could also be due to internal cleavage of the structural protein mRNA to produce translatable CAT mRNA.

Neutral and up mutations. Relative to the SIN junction sequence, the MBV junction sequence has G-for-A substitutions at positions -18 and -9 (Fig. 1B). Since mpSIN and mpMBV gave similar amounts of CAT subgenomic mRNA and CAT expression, these two substitutions do not appreciably affect promoter recognition by the SIN-derived transcription factors. In combination, the changes are therefore neutral mutations. The same substitutions are present in the RRV junction sequence, yet it gave higher levels of CAT subgenomic mRNA and CAT expression. The RRV junction sequence has an additional difference, U at position +5 (U+5), compared with SIN and MBV that might be responsible for its enhanced promoter strength. We constructed a hybrid junction sequence called SNRV that is identical to the SIN junction sequence except for having U+5 (Fig. 1B). By RNA analysis, the SNRV junction sequence was 1.5-fold stronger as a promoter than the SIN minimal promoter (Table 1). In another series of experiments, the mpSNRV virus also gave CAT expression at 5, 7, and 9 h postinfection that was similar to that of the mpRRV virus and higher than that of mpSIN and mpMBV (Table 2). Thus, U+5 is an up mutation for SIN.

DISCUSSION

Use of heterologous junction sequences as promoters by SIN. The heterologous junction sequences, other than that of SFV-1 and RUB, are recognized relatively efficiently by SIN. The results show that they have similar promoter strengths, and there is less than a 10-fold difference in the amount of CAT enzyme produced when they are used as promoters (Fig. 4). The junction sequences have two to seven base changes compared with the SIN minimal promoter. It is clear that the various combinations of base changes have relatively small effects on their use as promoters by SIN. The effect of the individual changes is unknown;

however, the effect of some can be deduced. For example, the MBV and SIN junction sequences differ at positions -9 and -18 (A-to-G changes) but behaved indistinguishably. The combination of the two changes is thus neutral to SIN. Mutagenesis studies showed that G-18 alone is a neutral mutation for SIN (40); thus, it is likely that G-9 by itself is also a neutral mutation. We deduced by the same reasoning, and showed experimentally, that the U+5 of the RRV junction sequence is an up mutation for SIN. It is interesting that SIN does not have U+5. This mutation either never occurred in the SIN lineage or was not retained during evolution. One reason for the latter might be that transcription levels are optimized in relationship to other parts of the viral life cycle, and increased levels of transcription are not necessarily beneficial and may in fact be deleterious. For example, oversynthesis of the subgenomic mRNA may result in decreased levels of genomic RNA synthesis (28).

The junction sequence of SFV-1 and the RUB sequence were very weak promoters for SIN. The low efficiency of the SFV-1 junction sequence is interesting. It differs from the SFV junction sequence reported by Ou et al. (37) only in having C at the -1 position (51). The same U-to-C-1 mutation in a minimal SIN virus promoter resulted in low promoter activity (40). Barring trivial possibilities, and if promoter recognition by SFV is similar to that of SIN, the stock of SFV characterized by Takkinen (51) is potentially quite interesting: it might harbor an unsuspected mutation that suppresses the C-1 mutation. The RUB sequence is used poorly by SIN. Although it resembles the conserved SIN junction sequence (13 identities out of 24 nt [12, 13]), the RUB subgenomic mRNA initiates 17 nt downstream of the sequence (13). Whether the RUB sequence is part of the RUB promoter is yet to be determined.

The promoter activity of the heterologous minimal junction sequences were measured in BHK cells. Their activities in other cell types such as avian and mosquito cells are yet to be determined. For example, some of the nucleotide differences in the junction sequences could be due to selection for optimal promoter activity in specific hosts cells (7). The minimal promoter of SIN is the shortest sequence that is still active as a promoter (28). It is about sixfold less active than a longer promoter, e.g., -98/+14 (28, 41). It would be interesting to determine whether SIN can use the longer junction sequences corresponding to the full-length SIN promoter.

Evolution of promoter recognition. With respect to the SIN transcription factors, most of the heterologous junction sequences, with two to seven base changes, show little functional divergence. One explanation for this is that the transcription factors are promiscuous, able to tolerate considerable sequence changes in the promoter. However, random mutagenesis of the SIN minimal promoter showed that promoters with two or more base changes typically had no or very low activities (40). The alternate explanation is that most of the random mutations that occurred in the junction sequences were deleterious and were eliminated during evolution; those remaining constitute a functionally selected subset of all possible changes. Some of the changes might have been deleterious because they alter nsP4: the promoter overlaps the coding sequences for nsP4. However, the carboxyl terminus of nsP4 is not very strongly conserved. For example, the coding sequences for nsP4 of the New World viruses terminate at position -8, yet they all have seven invariant nucleotides immediately after the termination codon (Fig. 1B). Also, the wobble positions of the nsP4 codons are expected to be mutable in the absence of

other selection, yet they too are well conserved. Other mutations might have been deleterious because of their effects on promoter activity. Some of these can be retained during evolution if, coincidentally, suppressor mutations occur in the binding sites of the transcription factors. These are probably rare, since our results here show that except for the C-1 of SFV-1, the base difference in the alphavirus junction sequences have relatively small effects on promoter function.

The same reasoning applies to the binding site(s) of the viral transcription factors. Mutations that are deleterious for promoter recognition, when they occur, are eliminated during evolution. This suggests that the binding site is also conserved and predicts that other alphaviruses, just like SIN, will be able to recognize efficiently heterologous junction sequences as promoters.

Recognition of viral *cis*-acting sequences by viral proteins. It is unlikely that the functional conservation shown by SIN promoter recognition is unique (see, e.g., reference 25). Indeed, brome mosaic virus and cowpea chlorotic mottle virus are each able to recognize efficiently the *cis*-acting sequences of the other for replication and transcription (38). Similarly, the 3' end of brome mosaic virus can functionally replace the 3' end of tobacco mosaic virus for replication mediated by tobacco mosaic virus replication proteins (21). We believe that in general the viral *cis*-acting RNA sequences and their cognate viral proteins are subjected to mutual constraints from divergence: the *cis*-acting sequence and the cognate binding site cannot change independently. Consistent with this view is the ubiquitous occurrence of conserved, probably *cis*-acting sequences in each family of RNA viruses. They are found at the termini of the genomic RNA, presumably recognized for replication; in internal positions coincident with transcription initiation/termination; and at sites recognized during packaging of the genomic RNA (see, e.g., references 9, 44, 49, and 50). The conservation of these sequences is remarkable: they are frequently more conserved than the protein-coding sequences. For example, while the sequences coding for SIN and SFV nsP4 (probably the viral polymerase [1]) are 65% homologous, the junction sequences are 83% homologous (the junction sequences are 71 to 92% homologous among the alphaviruses). Although the *cis*-acting roles of the conserved sequences of many viruses are yet to be demonstrated, their conservation despite sequence divergence elsewhere in the genome argues strongly for their functional importance.

Implications for antiviral drug design. Viral recognition of *cis*-acting sequences for replication, transcription, and packaging are obviously central aspects of the viral life cycle, and these sequences should be excellent targets for inhibition by antiviral drugs. Virus growth is inhibited when these processes are subjected to interference, e.g., by mutations in the *cis*-acting sequences (2, 10, 11, 15, 24, 31, 35, 36, 42). More relevant to drug design, virus growth can be inhibited when recognition of *cis*-acting sequences is inhibited *in trans*. Inhibition has been obtained in the presence of antisense RNA (39) or with an excess of competing wild-type (50) or mutant (42) *cis*-acting sequences expressed in isolation or as part of DI genomes (29) or satellite RNA (e.g., reference 20). Our results as well as results with other viruses (42, 50) suggest that inhibiting the recognition of *cis*-acting sequences is a general approach for inhibiting virus growth. Since the *cis*-acting sequences are conserved, it may be difficult for the virus to mutate to drug resistance. Also, since the *cis*-acting sequences are recognized at least in part by viral proteins, it may be possible to design drugs

that inhibit the recognition without undesirable effects on host processes. Furthermore, our results suggest that broad-spectrum drugs can be designed to inhibit a given virus that is effective also against closely related viruses.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AI26763.

We are grateful to Chang Hahn for providing us with a clone with an *Apal* site immediately after the E1 termination codon. We thank Chanda McDaniel for technical assistance. We thank Charlie Rice, Sondra Schlesinger, Milton Schlesinger, Eric Schroeter, and the Sindbis virologists at Washington University for discussion and critical reading of the manuscript.

REFERENCES

- Barton, D. J., S. G. Sawicki, and D. L. Sawicki. 1988. Demonstration in vitro of temperature-sensitive elongation of RNA in Sindbis virus mutant *ts6*. *J. Virol.* **62**:3597–3602.
- Bujarski, J. J., P. Ahlquist, T. C. Hall, T. W. Dreher, and P. Kaesberg. 1986. Modulation of replication, aminoacylation and adenylation *in vitro* and infectivity *in vivo* of BMV RNAs containing deletions within the multifunctional 3' end. *EMBO J.* **5**:1769–1774.
- Calisher, C. C., and N. Karabatsos. 1988. Arbovirus serogroups: definition and geographic distribution, p. 19–57. *In* T. P. Monath (ed.), *The arboviruses: epidemiology and ecology*. CRC Press, Inc., Boca Raton, Fla.
- Carmichael, G. G., and G. K. McMaster. 1980. The analysis of nucleic acids in gels using glyoxal and acridine orange. *Methods Enzymol.* **65**:380–391.
- Chang, G. J., and D. W. Trent. 1987. Nucleotide sequence of the genome region encoding the 26S mRNA of eastern equine encephalomyelitis virus and the deduced amino acid sequence of the viral structural proteins. *J. Gen. Virol.* **68**:2129–2142.
- de Groot, R. J., W. R. Hardy, Y. Shirako, and J. H. Strauss. 1990. Cleavage site preferences of Sindbis virus polypeptides containing the nonstructural proteinase: evidence for temporal regulation of polyprotein processing *in vivo*. *EMBO J.* **9**:2631–2637.
- Durbin, R., A. Kane, and V. Stollar. 1991. A mutant of Sindbis virus with altered plaque morphology and a decreased ratio of 26S:49S RNA synthesis in mosquito cells. *Virology* **183**:306–312.
- Faragher, S. G., A. D. Meek, C. M. Rice, and L. Dalgarno. 1988. Genome sequences of a mouse-avirulent and a mouse-virulent strain of Ross River virus. *Virology* **163**:539–526.
- Fields, B. N., D. M. Knipe, R. M. Chanock, M. S. Hirsh, J. L. Melnick, T. P. Monath, and B. Roizman (ed.). 1990. *Fields virology*. Raven Press, New York, New York.
- French, R., and P. Ahlquist. 1987. Intercistronic as well as terminal sequences are required for efficient amplification of brome mosaic virus RNA3. *J. Virol.* **61**:1457–1465.
- French, R., and P. Ahlquist. 1988. Characterization and engineering of sequences controlling *in vivo* synthesis of brome mosaic virus subgenomic RNA. *J. Virol.* **62**:2411–2420.
- Frey, T. K., and L. D. Marr. 1988. Sequence of the region coding for virion proteins C and E2 and the carboxy-terminus of the nonstructural proteins of rubella virus. *Gene* **62**:85–99.
- Frey, T. K., L. D. Marr, A. Sanchez, and R. B. Simmons. 1989. Identification of the 5' end of the rubella virus subgenomic RNA. *Virology* **168**:191–194.
- Gorman, C. M., L. F. Moffat, and B. H. Howard. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**:1044–1051.
- Grakoui, A., R. Levis, R. Raju, H. V. Huang, and C. M. Rice. 1989. A *cis*-acting mutation in the Sindbis virus junction region which affects subgenomic RNA synthesis. *J. Virol.* **63**:5216–5227.
- Hahn, C. S., Y. S. Hahn, T. Braciale, and C. M. Rice. Unpublished data.
- Hahn, Y. S., A. Grakoui, C. M. Rice, E. G. Strauss, and J. H. Strauss. 1989. Mapping of RNA⁻ temperature-sensitive mutants of Sindbis virus: complementation group F mutants have lesions in nsP4. *J. Virol.* **63**:1194–1202.
- Hahn, Y. S., E. G. Strauss, and J. H. Strauss. 1989. Mapping of RNA⁻ temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. *J. Virol.* **63**:1194–1202.
- Hardy, W. R., Y. S. Hahn, R. J. de Groot, E. G. Strauss, and J. H. Strauss. 1990. Synthesis and processing of the nonstructural polyproteins of several temperature-sensitive mutants of Sindbis virus. *Virology* **177**:199–208.
- Harrison, B. D., M. A. Mayo, and D. C. Baulcombe. 1987. Virus resistance in transgenic plants that express cucumber mosaic virus satellite RNA. *Nature (London)* **328**:799–802.
- Ishikawa, M., P. Kroner, P. Ahlquist, and T. Meshi. 1991. Biological activities of hybrid RNAs generated by 3'-end exchanges between tobacco mosaic and brome mosaic viruses. *J. Virol.* **65**:3451–3459.
- Keränen, S., and L. Kääriäinen. 1979. Functional defects of RNA-negative temperature-sensitive mutants of Sindbis and Semliki Forest virus. *J. Virol.* **32**:19–29.
- Kinney, R. M., B. J. Johnson, V. L. Brown, and D. W. Trent. 1986. Nucleotide sequence of the 26 S mRNA of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and deduced sequence of the encoded structural proteins. *Virology* **152**:400–413.
- Kuhn, R. J., Z. Hong, and J. H. Strauss. 1990. Mutagenesis of 3' nontranslated region of Sindbis virus RNA. *J. Virol.* **64**:1465–1476.
- Kuhn, R. J., H. G. M. Niesters, Z. Hong, and J. H. Strauss. 1991. Infectious RNA transcripts from Ross River virus cDNA clones and the construction and characterization of defined chimeras with Sindbis virus. *Virology* **182**:430–441.
- Laskey, R. A., and A. D. Mills. 1975. Quantitative film detection of ³H and ¹⁴C in polyacrylamide gels by fluorography. *Eur. J. Biochem.* **56**:335–341.
- Levinson, R. S., J. H. Strauss, and E. G. Strauss. Complete sequence of the genomic RNA of O'Nyong-nyong virus and its use in the construction of alphavirus phylogenetic trees. *Virology* **175**:110–123.
- Levis, R., S. Schlesinger, and H. V. Huang. 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. *J. Virol.* **64**:1726–1733.
- Levis, R., B. G. Weiss, M. Tsiang, H. Huang, and S. Schlesinger. 1986. Deletion mapping of Sindbis virus DI RNAs derived from cDNAs defines the sequences essential for replication and packaging. *Cell* **44**:137–145.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Marsh, L. E., T. W. Dreher, and T. C. Hall. 1988. Mutational analysis of the core and modulator sequences of the BMV RNA3 subgenomic promoter. *Nucleic Acids Res.* **16**:981–995.
- Mi, S., R. Durbin, H. V. Huang, C. M. Rice, and V. Stollar. 1989. Association of the Sindbis virus RNA methyltransferase activity with the nonstructural protein nsP1. *Virology* **170**:385–391.
- Mi, S., and V. Stollar. 1990. Both amino acid changes in nsP1 of Sindbis virus_{LM21} contribute to and are required for efficient expression of the mutant phenotype. *Virology* **178**:429–434.
- Mi, S., and V. Stollar. 1991. Expression of Sindbis virus nsP1 and methyltransferase activity in *Escherichia coli*. *Virology* **184**:423–427.
- Niesters, H. G. M., and J. H. Strauss. 1990. Defined mutations in the 5' nontranslated sequence of Sindbis virus RNA. *J. Virol.* **64**:4162–4168.
- Niesters, H. G. M., and J. H. Strauss. 1990. Mutagenesis of the conserved 51 nucleotide region of Sindbis virus. *J. Virol.* **64**:1639–1647.
- Ou, J.-H., C. M. Rice, L. Dalgarno, E. G. Strauss, and J. H. Strauss. 1982. Sequence studies of several alphavirus genomic RNAs in the region containing the start of the subgenomic RNA. *Proc. Natl. Acad. Sci. USA* **79**:5235–5239.
- Pacha, R., and P. Ahlquist. 1991. Use of bromovirus RNA3

- hybrids to study template specificity in viral RNA amplification. *J. Virol.* **65**:3693–3703.
39. **Powell, P. A., D. M. Stark, P. R. Sanders, and R. Beachy.** 1989. Protection against tobacco mosaic virus in transgenic plants that express tobacco mosaic virus antisense RNA. *Proc. Natl. Acad. Sci. USA* **86**:6949–6952.
 40. **Raju, R., and H. V. Huang.** Unpublished data.
 41. **Raju, R., and H. V. Huang.** 1991. Analysis of Sindbis virus promoter recognition in vivo, using novel vectors with two subgenomic mRNA promoters. *J. Virol.* **65**:2531–2510.
 42. **Rao, A. L. N., and T. C. Hall.** 1991. Interference in trans with brome mosaic virus replication by RNA-2 bearing aminoacylation-deficient mutants. *Virology* **180**:16–22.
 43. **Rice, C. M., R. Levis, J. H. Strauss, and H. V. Huang.** 1987. Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. *J. Virol.* **61**:3809–3819.
 44. **Rivera, V. M., J. D. Welsh, and J. V. Maizel, Jr.** 1988. Comparative sequence analysis of the 5' noncoding region of the enteroviruses and rhinoviruses. *Virology* **165**:42–53.
 45. **Sawicki, D. L., and S. G. Sawicki.** 1985. Functional analysis of the A complementation group mutants of Sindbis HR virus. *Virology* **144**:20–34.
 46. **Scheele, C. M., and E. R. Pfefferkorn.** 1969. Inhibition of interjacent ribonucleic acid (26S) synthesis in cells infected by Sindbis virus. *J. Virol.* **4**:117–122.
 47. **Shaw, W. V.** 1975. Chloramphenicol acetyltransferase from chloramphenicol-resistant bacteria. *Methods Enzymol.* **53**:737–755.
 48. **Strauss, E. G., R. Levinson, C. M. Rice, J. Dalrymple, and J. H. Strauss.** 1988. Nonstructural proteins nsP3 and nsP4 of Ross River and O'Nyong-nyong viruses: sequence and comparison with those of other alphaviruses. *Virology* **164**:265–274.
 49. **Strauss, E. G., and J. H. Strauss.** 1983. Replication strategies of the single stranded RNA viruses of eukaryotes. *Curr. Top. Microbiol. Immunol.* **105**:1–98.
 50. **Sullenger, B. A., H. F. Gallardo, G. E. Ungers, and E. Gilboa.** 1990. Overexpression of TAR sequences renders cells resistant to human immunodeficiency virus replication. *Cell* **63**:601–608.
 51. **Takkinen, K.** 1986. Complete nucleotide sequence of the non-structural protein genes of Semliki Forest virus. *Nucleic Acids Res.* **14**:5667–5682.
 52. **Weiss, B., D. Goran, R. Cancedda, and S. Schlesinger.** 1974. Defective interfering passages of Sindbis virus: nature of the intracellular defective viral RNA. *J. Virol.* **14**:1189–1198.
 53. **Xiong, C., R. Levis, P. Shen, S. Schlesinger, C. Rice, and H. V. Huang.** 1989. Sindbis virus: an efficient, broad host range vector for gene expression in animal cells. *Science* **243**:1188–1191.