tinct materials across the crater. On the basis of previous Lunar Orbiter and Apollo photogeological information and statistical tests of their spectral properties, we selected small areas (~3 km by 3 km) with both high internal spectral homogeneity and distinct respective spectral behavior to define an end-member basis in the image. Each pixel can then be represented by the fraction of each end-member taken in the linear combination that minimizes the difference between the model and the actual pixel spectrum. It is then possible to produce (i) a standard deviation image revealing the units that are not satisfactorily described by the proposed model and (ii) fraction images relative to each end-member. The end-member fractions are constrained to sum to 1.0, but each fraction may range out of the interval [0,1].

14. With such a ternary model, the residual mean standard (rms) deviation is ≈1% with $s_{rms} < 1\%$ for 51% of the pixel population, and most of the unexplained variance is concentrated in local parts of the image. This case may mean that additional end-members are needed locally. However, the residual is already close enough to the limit expected from our signal-to-noise ratio performance estimate because the rms image mimicks the boundary between the two basic pictures constituting the mosaic with a residual slightly less than 1% (Fig. 4D). This model brackets the fractions of the maximum (45%) of image pixels within the interval [0, 1], the minimum and maximum EJ1, EJ2, and Pk3 fraction values ranging, respectively, between [−1.1, 1.2], [−1.7, 2.4], and [−0.3, 1.7].
15. P. G. Lucey, B. R. Hawke, K. Horton, *Geophys. Res. Lett.* **18**, 2133 (1991).
16. A. L. Sprague *et al.*, *Icarus* **100**, 73 (1992).
17. P. A. Eberhardt, A. Stettler, J. Geiss, N. Grosler, *Moon* **8**, 104 (1973).
18. F. Hörz, R. Grieve, G. Heiken, P. Spudis, A. Binder, in *Lunar Source Book*, G. H. Heiken *et al.*, Eds. (Univ. of Cambridge Press, New York, 1991), pp. 61–120.
19. J. L. Hall, S. C. Solomon, J. W. Head, *J. Geophys. Res.* **86**, 9537 (1981).
20. J. W. Head, *Moon* **12**, 299 (1975).
21. K. A. Howard and H. G. Wilshire, *J. Res. U.S. Geol. Surv.* **3**, 237 (1975).
22. B. R. Hawke and J. W. Head, in *Impact and Explosion Cratering: Planetary and Terrestrial Implications*, D. J. Roddy *et al.*, Eds. (Pergamon, New York, 1977), pp. 815–841.
23. K. A. Howard, *U.S. Geol. Surv. Map I-840* (1975).
24. B. E. Clark, F. P. Fanale, J. W. Salisbury, *Icarus* **97**, 288 (1992).
25. C. F. Chyba, *ibid.* **92**, 217 (1991).
26. M. Settle and J. W. Head, *J. Geophys. Res.* **84**, 3081 (1979).
27. R. J. Pike, *Icarus* **43**, 1 (1980).
28. P. H. Schultz, *Moon Morphology* (University of Texas, Austin, 1976).
29. M. R. Dence, R. A. F. Grieve, P. B. Robertson, in *Impact and Explosion Cratering: Planetary and Terrestrial Implications*, D. J. Roddy *et al.*, Eds. (Pergamon, New York, 1977), pp. 247–275.
30. D. J. Roddy, in *ibid.*, pp. 185–246.
31. C. M. Pieters *et al.*, *J. Geophys. Res.* **90**, 12393 (1985).
32. R. B. Singer, *ibid.* **86**, 7967 (1981).
33. J. L. Whitford-Stark and J. W. Head, *ibid.* **85**, 6579 (1980).
34. R. A. F. Grieve, M. R. Dence, P. B. Robertson, in *Impact and Explosion Cratering: Planetary and Terrestrial Implications*, D. J. Roddy *et al.*, Eds. (Pergamon, New York, 1977), pp. 791–814.
35. R. A. F. Grieve and J. W. Head, *J. Geophys. Res.* **88**, A807 (1983).
36. C. M. Pieters, S. F. Pratt, J. M. Sunshine, *Lunar Planet. Sci.* **XXI**, 962 (1990).
37. E. A. Whitaker, *Proc. Lunar Planet. Sci. Conf.* **12A**, 105 (1981).
38. G. Ryder, *Geophys. Res. Lett.* **18**, 2065 (1991).
39. C. M. Pieters, *ibid.*, p. 2129.
40. This research was supported by the French National Program of Planetology of the Institut National des Sciences de l'Univers and the Centre National de la Recherche Scientifique. We acknowledge the support of the Pic-du-Midi Observatory staff during observations with the 2-m telescope, with special thanks to P. Caumette, J. P. Lafforgue, J. M. Lavie-Cambot, F. Lacassagne, C. Duthu, and P. Dhérété for their dedicated work. We thank C. M. Pieters for providing the F2 relative reflectance spectra, which allowed a comparison between earlier spectroscopic observations and our CCD imaging data. We also appreciate the reviews and constructive comments by J. W. Head and an unidentified reviewer, which improved the discussion significantly.

14 October 1992; accepted 18 February 1993

Sequences and Structures Required for Recombination Between Virus-Associated RNAs

Pamela J. Cascone,* Tarik F. Haydar, Anne E. Simon†

RNA recombination has been described for a number of viruses in the plant and animal kingdoms, but the mechanisms of selection of recombination sites are poorly understood. The nonrandom recombination between two subviral RNAs associated with turnip crinkle virus was used to study the requirement for specific sequences and structures in the generation of recombinant molecules. Single-base mutations that disrupted either the stem or the loop of one of the two computer-predicted stem-loop structures eliminated detectable recombinant molecules. However, recombinants were detected if compensatory mutations were generated that re-formed a stable hairpin structure. These results provide evidence for the necessity of specific structures in the formation of recombinant molecules in this system.

The exchange of genetic material, as has been demonstrated for a growing number of plant and animal viruses, may be important in the evolution, repair, and diversification of viral genomes (*1*). Most of the evidence suggests that viral RNA recombination involves a copy choice mechanism whereby the viral replicase switches from one template to an alternate location on the same or a different template where polymerization of the nascent strand then continues (*2*). An examination of intertypic poliovirus recombinants (*3*) and recombinants between genomic RNAs of the tripartite brome mosaic virus (*4*) has led to the suggestion that local regions of hybridization are the preferred sites for template switching in these systems. However, the availability of many favorable recombination sites in poliovirus and brome mosaic virus genomes and the apparently random nature of RNA recombination in most other viral systems (*5*) have precluded a detailed analysis of how the replicase, nascent strand, and templates interact to generate recombinant molecules.

Turnip crinkle virus (TCV), a monopartite, single-stranded RNA virus, is naturally associated with both defective interfering (DI) RNAs (*6*) and recombinant molecules derived from combining a linear satellite (sat-) RNA and the 3′ region of the viral genomic RNA (*7*, *8*). One of three ~20-nucleotide (nt) consensus sequences, two of which are similar to sequences at or near the 5′ ends of TCV-associated RNAs, is always located at the right side of the crossover junctions (*8*, *9*). These findings led us to propose a replicase-driven, copy choice mechanism for RNA recombination during synthesis of plus-strands (*9*).

An unusual example of aberrant homologous recombination between two TCV subviral RNAs has also been found (*9*). The parental RNAs were sat–RNA D (194 bases), a sat-RNA of unknown origin without appreciable sequence similarity to the viral genomic RNA, and sat–RNA C (356 bases), a molecule originally formed from two recombination events between sat–RNA D and the TCV genome (Fig. 1). Recombinant molecules appear to be precisely joined at one of five bases in sat–RNA C (positions 175 to 179), with 30% of the molecules also containing nontemplate-encoded nucleotides at the crossover point (*9*). The five bases at the junctions of the recombinants are contained within a larger 19-base sequence (motif I) in sat–RNA C (positions 175 to 193) that is similar to the junction sequence in one TCV DI RNA, as well as to a sequence near the 5′ terminus of the TCV genomic RNA.

To determine the role of motif I in recombination between sat–RNA D and sat–RNA C, we constructed a variant of sat–RNA C (deletion of nts 57 to 78) that could only be amplified in plants after recombination with sat–RNA D (*10*). Transcripts containing mutations in and

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

*Present address: Department of Biology, University of California at San Diego, La Jolla, CA 92093.
†To whom correspondence should be addressed.

around motif I, separately and in combination with the 22-nt deletion, were coinoculated with helper virus inoculum (HVI) (*11*) onto 6 to 12 turnip plants. We used direct RNA sequencing (*12*) to confirm that the mutations were faithfully maintained in vivo (unless otherwise noted). Recombinant molecules, detected by Northern (RNA) analysis with sat–RNA D– and sat–RNA C–specific oligonucleotides, were amplified by polymerase chain reaction (PCR) (*13*) and sequenced to locate the crossover sites.

To facilitate the analysis of motif I, we created an Apa I site 13 bases downstream from the motif and an Mlu I site 32 bases upstream from the motif, generating plasmid pCAM (*14*). A derivative of pCAM, pCAMB, was prepared by generation of a unique Bam HI site at the predicted 5′ border of motif I. Inoculation of plants with transcripts of pCAMD or pCAMBD (where D denotes the inclusion of the 22-bp deletion) resulted in the detection of recombinant molecules with the expected junction sites (Table 1).

No recombinant molecules were detected when the 31-bp Apa I–Bam HI fragment (containing motif I) was moved to the Mlu I site or replaced with a 31-bp fragment from the polylinker of pUC8 (*15, 16*). These initial results indicated that motif I was necessary for RNA recombination and that either the location of the signal was important or that sequences critical for motif I recognition extended beyond the Apa I–Bam HI fragment.

To determine the role of the primary sequence of motif I on recombination, we introduced 13 single-, double-, and triple-point mutations in and around motif I in pCAM (Fig. 1 and Table 2). Recombinant molecules with normal crossover sites were detected by Northern blot and PCR analyses when point mutations were introduced at or just upstream from the crossover site (Fig. 2 and Table 1). Recombinant molecules were not detected if the point mutations were 4 to 17 nts downstream from the normal crossover site (*17*). These results demonstrated that alteration of single bases can eliminate the detection of recombinant molecules and confirmed that the sequence required for RNA recombination extended beyond the 19-nt motif predicted by sequence comparisons (*9*).

We therefore constructed alterations at sites distal to the motif (*18*) and tested the altered transcripts for the ability to recom-

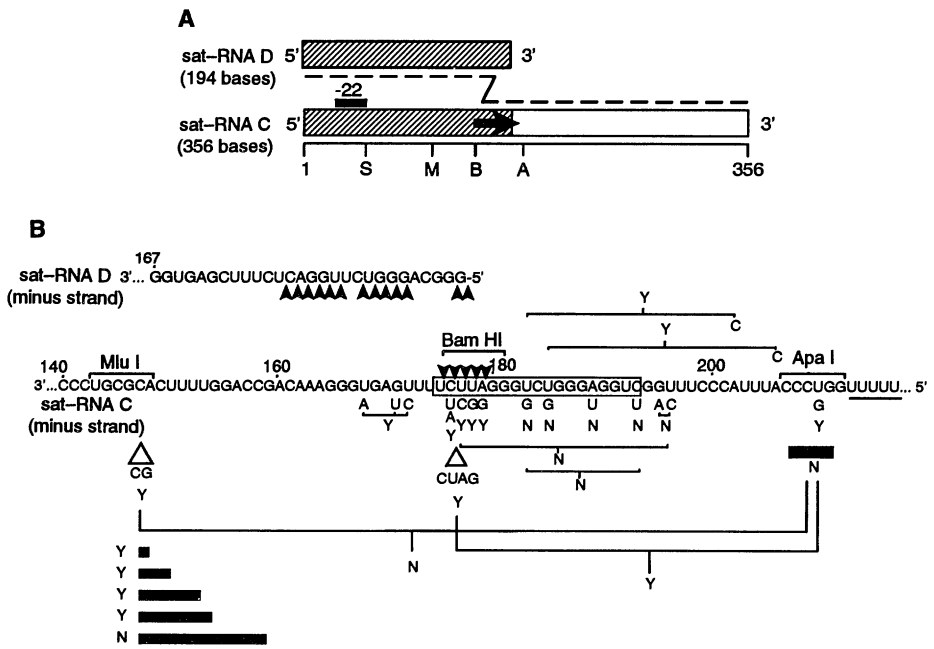

Fig. 1. Summary of mutations constructed in and around motif I and their effects on recombination between sat–RNA C and sat–RNA D. Numbering is from (*7*). (**A**) Schematics of sat–RNA C and sat–RNA D and the location of pertinent restriction sites in the cDNA of sat–RNA C. S, Sna BI; M, Mlu I; B, Bam HI; and A, Apa I. Hatched portions of the two sat-RNAs share 88% similarity; unshaded portion of sat–RNA C is similar to two regions in the TCV genomic RNA. Short arrow represents motif I (positions 175 to 193); short black bar denotes the location of the 22-base deletion. Dashed line indicates the portions of sat–RNA D and sat–RNA C found in recombinant molecules. (**B**) Partial sequences of sat–RNA D and sat–RNA C. The minus strand is presented because recombination is thought to occur during the synthesis of plus strands (*9*). Motif I is boxed. Black arrowheads denote crossover points in recombinant molecules from both the current and other studies (*9*). Open triangles denote insertions, and filled bars indicate deletions. Altered bases shown above the sat–RNA C sequence were compensatory changes that reformed hairpin b (Fig. 3). Alterations connected by lines were in the same construct. Underlined sequence denotes the location of additional U residues found in certain recombinant molecules (Table 2). Y and N below or to the side of the alterations indicate that recombinant RNAs were (Y) or were not (N) detected with transcripts containing the mutations.

Table 1. Location of junctions in recombinants. Inoculated transcripts contained the alterations described in Table 2 as well as a deletion (D) of bases 57 to 78. The location of crossover points in the recombinants was determined after amplification of the RNA by PCR and subsequent cloning. Nontemplate nucleotides at the junctions (NTNs) were not present in either of the two parental sat-RNAs and are presented from the minus strand sequence. No., number of clones; Left, the position in sat–RNA D of the left junction; and Right, the position in sat–RNA C of the right junction.

Plasmid	Left	Right	NTNs	No.
pCAMD	181	176		4
	180	176	AG	3
	189	179		3
	189	179		3
pCAMXD	188	176		1
	182	176		1
	187	178		2
	186	179		1
pCAM79D	181	176	C	3
	181	178		4
pCAM302D	180	173		1
	180	176		1
	194	176		1
	181	178		5
	181	180		1
pCAMBD	181	153		2
	181	178	AA	3
	180	178	AA	2
	181	178		10
pCAM177D	180	144		2
	180	176	AA	2
	180	178	A	1
	185	179		1
	181	179	AUUC	1
	180	179		1
	185	180		9
pCAM55D	180	176	A	2
	193	176	G	1
	180	176	AA	1
	180	178	A	4
pCAMiMD	180	176		1
	181	176		1
	179	177	AAA	1
	180	178	AA	1
	179	179		3
	185	179		1
	185	180		1
	185	180		1
pCAMBdiABD	193	187		4
	183	218		1
pCAM628D	181	176		1
	180	177	AA	6
	180	178	AA	7*
	181	178	U	1
pCAM820D	179	160		2
	181	177		2
	178	178	AAAA	1
	182	178		1
	185	179		1

bine (Fig. 1 and Table 2). A 4-bp deletion at the Apa I site eliminated detectable recombinants, indicating that the region involved in template switching extended at least 15 nts downstream of motif I. Recombinant molecules were detected after a 2-bp insertion at the Mlu I site or a 4-bp insertion at the Bam HI site. Some of the recombinant species derived from transcripts with the Mlu I or Bam HI insertions contained 4 to 13 additional U residues in a region (positions 213 to 217) normally containing only five U residues (*19*). Transcripts containing alterations at both the Apa I and Mlu I sites did not lead to detectable recombinants, whereas recombinant molecules with unusual junctions (positions 187 and 218) were generated in a single plant (out of six) inoculated with transcripts containing alterations at both the Apa I and Bam HI sites.

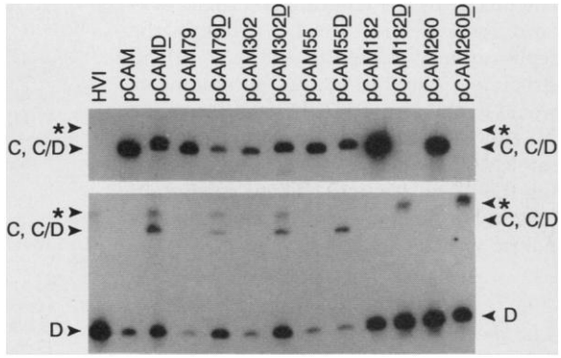

Fig. 2. Example of the Northern blot analysis used for the detection of recombinant molecules. LiCl-insoluble RNA was extracted from leaf tissue 3 weeks after inoculation (*11*) with HVI alone or HVI and transcripts derived from the plasmids indicated above each lane. RNA (3 μg) was subjected to electrophoresis through a denaturing 4% polyacrylamide–50% urea gel and transferred to a Nytran support (Schleicher & Schuell). Identical blots were then hybridized (*6*) with ^{32}P-labeled oligonucleotides complementary to either positions 339 to 356 of sat–RNA C (top) or positions 44 to 59 of sat–RNA D (bottom). Plasmids are defined in Table 2. D denotes that plasmids also contained the 22-base deletion. The sat–RNA D–specific oligonucleotide did not hybridize to the similar region in sat–RNA C because of four mismatched bases. The positions of sat–RNA C (C), sat–RNA D (D), and recombinants between sat–RNA C and sat–RNA D (C/D) are indicated. Asterisk (*) denotes recombinants between sat–RNA D and the 3′ end of the TCV genomic RNA that frequently accumulate in plants inoculated with HVI alone or HVI plus transcripts of sat–RNA C that are uninfectious (*8*).

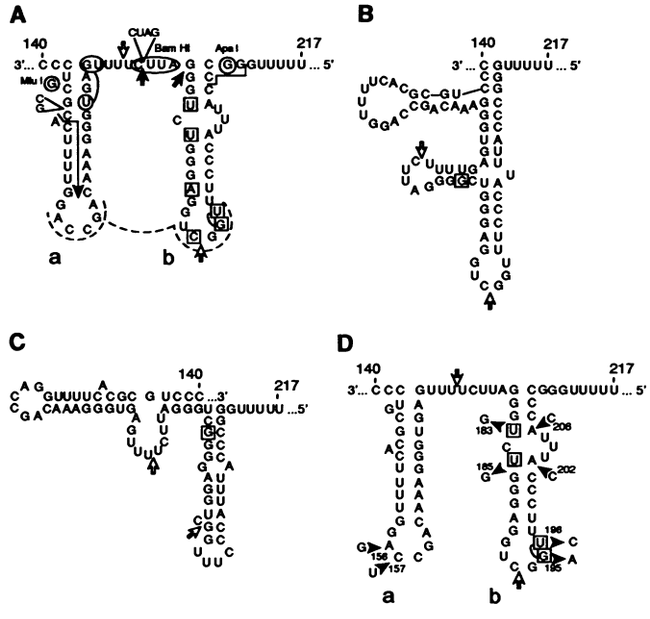

Fig. 3. (**A**) Secondary structure model of the region encompassing 78 bases of the minus strand of sat-RNA derived from pCAM. Numbering is as in Fig. 1. Motif I is situated between the open arrows. Closed arrows encompass the location of normal crossover sites. Dashed lines indicate potential loop-loop interactions. Circled residues allow, and boxed residues abolish, recombination when altered. Bases altered in the same construct are connected by lines. The arrow extending from the Mlu I site indicates the location and direction of the deletions described in Table 2. Bracket encloses the four bases deleted at the Apa I site (pCAMdA) that eliminate detectable recombinants. The inserted bases at the Mlu I site (pCAMiM) and Bam HI site (pCAMBiB) are shown (the Bam HI site was created by changing U^{177} to C). Hairpin loops a and b are indicated. (**B**) Predicted folding of the sat-RNA derived from pCAM182. The boxed residue was altered from a U to a G. (**C**) Predicted folding of the sat-RNA derived from pCAM260. The boxed residue was altered from a U to a G. (**D**) Compensatory changes constructed to repair mismatches in hairpin b or between the two loop regions. Arrows pointing away from the structures indicate alterations that abolish recombination. Arrowheads pointing toward the structures indicate compensatory mutations engineered to repair putative mismatches.

To determine the 5′ border of the sequence required for recombination, we generated deletions of 1 to 12 bp in pCAM from the Mlu I site toward motif I (*10*). The number of plants containing detectable recombinants was substantially reduced if transcripts contained deletions of 7 or 12 bases (Table 2). This result suggested that the left border of the sequence required for recombination extended 16 to 20 residues upstream from the border predicted by sequence comparisons (*9*).

Seventy-eight bases, including the 65 residues between the Mlu I and Apa I sites, were subjected to secondary structure analysis by means of the Zuker and Stiegler algorithm (*20*) (Fig. 3A). In the dual stem-loop structure generated, the normal crossover site is found in a single-stranded region near the base of hairpin b (Fig. 3A). Nucleotides in the two loops could hydrogen bond to form a loop-loop interaction. Sequences corresponding to the Apa I and Bam HI sites are juxtaposed; transcripts with alterations at the Apa I site did not result in detectable recombinants, whereas recombinant molecules were isolated when transcripts contained alterations at both the Apa I and Bam HI sites.

Secondary structure models were generated for each alteration presented in this paper. Transcripts containing changes in which the lowest free energy structure was substantially different from the dual stem-loop structure did not recombine with sat–RNA D at detectable amounts (*16*) (Fig. 3, B and C). We tested the effect of restoring base pairing in hairpin b by generating compensatory mutations in which the original presumptive U:A pairing in two locations was replaced with G:C pairing (Fig. 3D). In both cases, the compensatory changes restored the generation of recombinant molecules (with normal crossover sites) in all inoculated plants (Table 2). These data provide evidence for the existence of hairpin b in vivo and are consistent with our prediction that this structure is involved in template switching between sat–RNA D and sat–RNA C by the replicase.

We also engineered compensatory mutations to test whether mutations in the loop of hairpin b inhibited recombination because of an inability to form loop-loop interactions (Fig. 3D). Sat–RNA C containing the original (A^{195}, C^{196}) and compensatory changes (G^{156}, U^{157}) did not recombine with sat–RNA D at detectable amounts (*21*). Rather than by involvement in RNA tertiary structure, bases in the loop of hairpin b may participate in direct interaction with the replicase.

RNA binding proteins recognize specific primary sequences (*22*) or structures (*23*) of RNA molecules. The most common secondary structural elements recognized by proteins are hairpins, with the strongest interactions generally in loop or bulge loop regions (*24*). Our other findings, that one of three consensus sequences was at the right sides of all junctions and that two of these motifs resembled sequences at the 5′ ends of TCV-associated molecules, suggested that the primary nucleotide sequence was recognized by the enzyme that mediates recombination during synthesis of plus strands. However, based on our findings here, the functional significance of the sequence of motif I (which comprises a portion of hairpin b) in the generation of recombinants is not clear. Converting two A residues to G residues in the motif I sequence did not adversely affect recombination, provided that compensatory mutations re-formed the hairpin. This suggests that the involvement of these two positions in stabilizing the hairpin is more important than the specific primary sequence is.

It is also not known whether the corresponding structure in plus strand sat–RNA C is recognized by the replicase as a site for recombination. To detect recombinant molecules in our in vivo system, they must be functional, which means they must be amplifiable by the replicase and then packaged for systemic spread. Although the replicase may recognize the plus strand structure, the generated recombinants would contain sat–RNA D sequence at the 5′ end and the roughly complementary sat–RNA C sequence at the 3′ end (including the 22-nt deletion). These nearly full-length hairpin molecules would probably not be viable.

Table 2. Effect of mutations in and around motif I on recombination between sat–RNA D and sat–RNA C. Infectivity refers to the number of plants accumulating sat-RNAs visible in denaturing polyacrylamide gels stained with ethidium bromide, relative to the number of plants inoculated with HVI and transcripts containing the alterations. For the recombination assay, alterations shown were combined with the 22-base deletion. In the Northern analysis, the number of plants containing RNA species that hybridized to both sat–RNA D– and sat–RNA C–specific probes is shown relative to the total number of plants inoculated. Y indicates that recombinant molecules were cloned out of at least one of two plant RNA samples after amplification with sat–RNA D– and sat–RNA C–specific oligonucleotides (*13*); N indicates that no PCR products corresponding to recombinant species were detected. Ins, bases inserted; Del, bases deleted; and NT, not tested.

Name	Location of mutations	Infectivity	Recombination assay	
			Northern	PCR
	Base changes			
pCAM	$U^{210} \rightarrow G$	5/6	6/6	Y
pCAMX	$U^{168} \rightarrow A$, $G^{171} \rightarrow U$, $U^{172} \rightarrow C$	6/6	1/6	Y
pCAM79	$C^{176} \rightarrow U$	6/6	5/6	Y
pCAM302	$C^{176} \rightarrow A$	6/6	2/6	Y
pCAMB	$U^{177} \rightarrow C$	6/6	3/6	Y
pCAM177	$U^{178} \rightarrow G$	6/6	5/6	Y
pCAM55	$A^{179} \rightarrow G$	6/6	4/6	Y
pCAM182	$U^{183} \rightarrow G$	6/6	0/6	N
pCAM260	$U^{185} \rightarrow G$	6/6	0/6	N
pCAM118	$A^{189} \rightarrow U$	7/9	0/6	N
pCAM229	$C^{193} \rightarrow U$	5/6	0/6	N
pCAMP	$G^{195} \rightarrow A$, $U^{196} \rightarrow C$	6/6	0/6	N
pCAMBP	$U^{177} \rightarrow C$, $G^{195} \rightarrow A$, $U^{196} \rightarrow C$	6/6	0/6	N
pCAM115	$U^{183} \rightarrow G$, $C^{193} \rightarrow U$	5/9	0/6	N
	Insertions and deletions			
pCAMiM	148-Ins CG	6/6	5/6*	Y
pCAMBiB	177-Ins CUAG	6/6	3/6*	Y
pCAMdA	Del C^{208}–G^{211}	6/6	0/6	N
pCAMBdiAB	177-Ins CUAG; Del C^{208}–G^{211}	6/6	1/6†	Y
pCAMdiAM	148-Ins CG, Del C^{208}–G^{211}	6/6	0/6	N
pCAMd1	Del A^{148}	6/6	6/12	NT
pCAMd3	Del A^{148}–U^{150}	6/6	7/12	NT
pCAMd6	Del A^{148}–U^{153}	6/6	3/12	NT
pCAMd7	Del A^{148}–G^{154}	6/6	1/12	NT
pCAMd12‡	Del A^{148}–G^{159}	6/6	0/12	NT
	Compensatory mutations			
pCAM820	$U^{183} \rightarrow G$, $A^{206} \rightarrow C$	6/6	6/6	Y
pCAM628	$U^{185} \rightarrow G$, $A^{202} \rightarrow C$	6/6	6/6	Y
pCAM456	$G^{195} \rightarrow A$, $U^{196} \rightarrow C$ $A^{156} \rightarrow G$, $C^{157} \rightarrow U$	4/6	0/6	N

*Some recombinant molecules contained 4 to 13 additional U residues at position 213. †Some of the crossover sites were shifted to unusual locations (Table 1). ‡Contains an additional Mlu I linker at the deletion site.

REFERENCES AND NOTES

1. C. R. Pringle, *Virology* **25**, 48 (1965); E. A. Tolskaya, L. A. Romanova, M. S. Kolesnikova, V. I. Agol, *ibid.* **124**, 121 (1983); M. M. C. Lai *et al.*, *J. Virol.* **56**, 449 (1985); G. C. Angenent *et al.*, *Nucleic Acids Res.* **14**, 4673 (1986); J. J. Bujarski and P. Kaesberg, *Nature* **321**, 528 (1986); R. F. Allison, C. Thompson, P. Ahlquist, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1820 (1990); A. L. N. Rao and T. C. Hall, *J. Virol.* **64**, 2437 (1990); M. C. Edwards, I. T. D. Petty, A. O. Jackson, *Virology* **189**, 389 (1992).
2. R. A. Lazzarini, J. D. Keene, M. Schubert, *Cell* **26**, 145 (1981); K. Kirkegaard and D. Baltimore, *ibid.* **47**, 433 (1986); A. M. Q. King, *Nucleic Acids Res.* **16**, 11705 (1988).
3. E. A. Tolskaya *et al.*, *Virology* **161**, 54 (1987).
4. J. J. Bujarski and A. M. Dzianott, *J. Virol.* **65**, 4153 (1991); P. D. Nagy and J. J. Bujarski, *ibid.* **66**, 6824 (1992).
5. M. M. C. Lai, *Microbiol. Rev.* **56**, 61 (1992).
6. X. H. Li, L. A. Heaton, T. J. Morris, A. E. Simon, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 9173 (1989).
7. A. E. Simon and S. H. Howell, *EMBO J.* **5**, 3423 (1986).
8. C. Zhang, P. J. Cascone, A. E. Simon, *Virology* **184**, 791 (1991).
9. P. J. Cascone, C. D. Carpenter, X. H. Li, A. E. Simon, *EMBO J.* **9**, 1709 (1990).
10. Plasmid pPM2-2-47M+, which contains a full-length cDNA of sat–RNA C cloned downstream from an *Escherichia coli* RNA polymerase promoter [A. E. Simon and S. H. Howell, *Virology* **156**, 146 (1987)], was the progenitor plasmid used in this study. Deletions were generated from either the Sna BI site at position 78 in pPM2-2-47M+ or the Mlu I site at position 143 in pCAM using the slow form of Bal 31 as described [A. E. Simon, H. Engel, R. P. Johnson, S. H. Howell, *EMBO J.* **7**, 2645 (1988)]. We determined by sequencing that p2CSdl22 contains a deletion of 22 bp (positions 57 to 78). Mlu I linkers were re-inserted into the deleted segments from pCAM with a described protocol [X. H. Li and A. E. Simon, *J. Virol.* **65**, 4582 (1991)].
11. The HVI contained TCV genomic RNA and sat–RNA D. Inoculation procedures have been described [X. H. Li and A. E. Simon, *Phytopathology* **80**, 238 (1990)].
12. C. D. Carpenter and A. E. Simon, *BioTechniques* **8**, 26 (1990). The oligonucleotide used for sequencing was complementary to positions 196 to 212 in sat–RNA C.
13. PCR amplification was done as described [C. D. Carpenter, P. J. Cascone, A. E. Simon, *Virology* **183**, 586 (1991)] with 2 U of Pyrostase enzyme (Molecular Genetic Resources, Tampa, FL) used for the amplification step. The oligonucleotide used for synthesis of the first strand cDNA was complementary to positions 339 to 356 of sat–RNA C. The oligonucleotide used for second strand synthesis was homologous to positions 100 to 114 of sat–RNA D.
14. We used a gapped-duplex, double-stranded DNA mutagenesis scheme [Y. Morinaga, T. Franceschini, S. Inouye, M. Inouye, *Bio/Technology* **2**, 636 (1984)] or a single-stranded DNA mutagenesis scheme (Bio-Rad Muta-Gene Phagemid In Vitro Mutagenesis Kit) to generate point mutations. To create the Apa I site, we used an oligonucleotide corresponding to positions 199 to 218 in the sat–RNA C cDNA insert in pPM2-2-47M+ (CCATTTACCCGGGTTTTTGC; the altered residue is underlined), generating plasmid pCA. To create the Mlu I site, we used an oligonucleotide corresponding to positions 134 to 153 in pCA (TTCATGCCCTGCGCACTTTTG; the inserted residue is underlined). To generate random point mutations within motif I, we synthesized degenerate oligonucleotides AGTTTTcttagggtcTGGGA (positions 170 to 189) and GGGTCtgggaggtcGGTTTC (positions 180 to 199). These oligonucleotides were prepared such that the positions indicated by lowercase letters were synthesized from pools of nucleotides containing 90% of the complementary residue and 3.3% each of the remaining nucleotides. We generated additional

point mutations with oligonucleotides prepared to contain defined mismatches and create additional unique restriction enzyme sites in pCAM (altered residues are underlined below): Bam HI (GTGAGTTTTCCTAGGGTCTGG, positions 167 to 187); Pvu II (GGAGGTCGACTTCCCATTT, positions 187 to 205); and Xba I (ACAAAGGGAGATCTTTCTTAGGG, positions 160 to 182). To generate the compensatory changes, we used the following oligonucleotides: A^{202} replaced with a C (CAAAGGGGAAATGGGC, positions 195 to 210); A^{206} replaced with a C (AGGGTAAAGGGGCCCAAA, positions 198 to 215); and A^{156} and C^{157} replaced with G and U, respectively (GAAAACCCAGCUGUUU, positions 149 to 164). To combine all mutations in this study with the 22-bp deletion, we digested p2CSdI22 (pPM2-2-47M+ with a deletion of positions 57 to 78) and the various constructs with Pst I and Nco I (Pst I digests at a single site about 1 kb upstream from the cDNA insert and Nco I digests at a single site 22 bp downstream from the Sna BI site in the sat–RNA C cDNA), the fragments were gel purified, and the larger fragment (containing alterations) was ligated to the shorter fragment (containing the *E. coli* RNA polymerase promoter and the 22-bp deletion).

15. Plasmid pCAMB (Table 2) was digested with Apa I and Bam HI, the ends were made flush with the Klenow fragment, and the large plasmid fragment was ligated to a 31-bp segment that had been generated by digestion of pUC8 with Sma I and Hind III, making the ends flush with the Klenow fragment, and gel purifying the small fragment. The Sma I end of the 31-bp fragment was nearest to the corresponding 5′ end of the sat-RNA.
16. P. J. Cascone, thesis, University of Massachusetts (1992).
17. Nucleotide reversion to the wild-type residue was found in nearly all sat-RNAs derived from constructs containing the alteration of U to C at position 196. Furthermore, sat–RNA C species from plants inoculated with HVI and transcripts of pCAM55, in which A^{179} had been changed to a G, revealed that in samples from some plants the G residue had reverted to the original A residue, whereas in other samples G^{182} was converted to an A residue.
18. Insertions and deletions at the Apa I, Mlu I, and Bam HI sites were generated by digestion of either pCAM (alterations at the Apa I and Mlu I sites) or of pCAMB (alterations at the Bam HI site) with the appropriate enzyme, followed by incubation with the Klenow fragment and deoxynucleotides and then by intramolecular religation.
19. It is not clear what caused the addition of U residues. The RNA-dependent RNA polymerase of TCV has been shown to insert additional U residues in a region containing five consecutive U residues during the replication of certain altered sat–RNA C species [C. D. Carpenter, P. J. Cascone, A. E. Simon, *Virology* **183**, 595 (1991)]. This type of phenomenon has also been reported for other systems; for example, *E. coli* RNA polymerase can add extra nucleotides when replicating on poly(U) tracts [L. A. Wagner, R. B. Weiss, R. Driscoll, D. S. Dunn, R. F. Gesteland, *Nucleic Acids Res.* **18**, 3529 (1990)], and in paramyxovirus, additional G residues can be added into a series of G residues [J. A. Southern, B. Precious, P. E. Randall, *Virology* **177**, 388 (1990); S. Vidal, J. Curran, D. Kolakofsky, *J. Virol.* **64**, 239 (1990)].
20. M. Zuker and P. Stiegler, *Nucleic Acids Res.* **9**, 133 (1981).
21. Sequencing the sat-RNAs derived from infection of these transcripts without the 22-base deletion revealed that the two new compensatory changes were maintained; however, as before, A^{195} was preserved, yet C^{196} reverted to the wild-type U residue. Although the generation of C^{196} would have had no effect on the ability of the two loops to interact, C^{196} could base-pair with G^{191}, thereby disrupting the structure of the six-base loop. This could affect the viability of the molecule in an as yet unknown fashion.
22. Z. Qian and J. Wilusz, *Mol. Cell. Biol.* **11**, 5312 (1991); W. Keller, S. Bienroth, K. M. Lang, G. Christofori, *EMBO J.* **10**, 4241 (1991).
23. R. J. Romaniuk *et al.*, *Biochemistry* **26**, 1567 (1987); H. A. Barton, R. S. Eisenstein, A. Bomford, H. N. Munro, *J. Biol. Chem.* **265**, 7000 (1990); H. S. Olsen, P. Nelbock, A. W. Cochrane, C. A. Rosen, *Science* **247**, 845 (1990); B. J. Calnan *et al.*, *ibid.* **252**, 1167 (1991).
24. M. Chastain and I. Tinoco, Jr., *Prog. Nucleic Acid Res.* **41**, 131 (1991).
25. We thank C. Carpenter for technical assistance and critical reading of the manuscript. Supported by NSF grants DMB 9004665 and DMB 9105890 (to A.E.S.) and a Sigma Xi award (to P.J.C.).

8 December 1992; accepted 2 February 1993

Protein-Induced Bending as a Transcriptional Switch

José Pérez-Martín and Manuel Espinosa*

The question of whether protein-induced DNA bending can act as a switch factor when placed upstream of an array of promoters located in tandem was investigated in vivo. The catabolite activating protein binding site of the *fur* operon was replaced by the binding site of the RepA repressor protein, which is able to bend DNA immediately after binding. Appropriately phased induced bending could act as a transcriptional switch factor in vivo.

Gene expression in prokaryotes is frequently controlled by DNA binding proteins that attach upstream of the region proximal to the transcription start site (*1*). Many of these regulatory proteins bend cognate DNA sequences after binding (*2*), but whether the resulting DNA curvature forms part of the mechanism for transcription initiation remains uncertain. Intrinsically bent DNA activates several promoters (*3*, *4*), and curved regions can be exchanged with the binding site of a DNA bending protein (*5*). The catabolite activator protein (CAP) from *Escherichia coli*, which is involved in transcription activation of several operons, bends the DNA after binding to its target (*6*). There are instances of CAP-dependent systems, such as *lac* and *gal* operons, in which upstream intrinsic DNA bends have by themselves the capacity to activate transcription, thus functionally replacing the binding site of CAP (*7*). These findings support the model that upstream DNA bending is an important element in transcription activation (*8*). Recent studies have suggested that, however, contacts between CAP and RNA polymerase are more important than CAP-DNA interactions (*9*), thus making it unclear how bending would affect such protein-protein interactions. To study the effects of protein-induced DNA bending in transcription initiation in vivo, we examined the effect of substituting the CAP site of the *E. coli fur* operon with the heterologous binding site of the DNA bending repressor RepA [recently renamed CopG (*10*)]. Our results support a model in which DNA bending could act as a transcriptional switch: one promoter is activated, whereas an adjacent one is repressed.

The product of the *fur* gene is involved in the control of iron transport and metabolism in *E. coli*, and its synthesis is regulated by CAP (*11*). Like other CAP-activated systems, the *fur* gene is expressed from two tandem overlapping promoters having their −35 and −10 regions located on opposite sides of the DNA helix. Primer extension analysis of *fur* RNA isolated in vivo indicated that under catabolite repression conditions, activity of the downstream promoter P_A (placed at the opposite helix side of the CAP binding sequence) is paralleled by a decrease in transcription from the upstream promoter P_B. In contrast, transcription from P_B (located at the same helix side as the CAP binding site) is stimulated during catabolite activation, whereas transcription from the downstream promoter P_A is repressed (*11*). Transcriptional switches have also been observed in systems like the *gal* operon (*12*), although they may be explained by competition between CAP and RNA polymerase for nearby sites within the same DNA region at the basal promoter. In the case of *fur*, however, target sequences of CAP are significantly distant from the −10 and −35 regions of the basal promoter, P_A (−70 with respect to the transcription start site).

We have proposed (*5*) that when curved regions, or the recognition target of a protein that bends DNA, are placed in phase with a promoter, increases in the DNA curvature would facilitate RNA polymerase-promoter interactions during the initiation of transcription. In contrast, when those sequences are not phased with the promoter, static or induced bends would hinder the formation of a transcriptional

Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Velázquez 144, E-28006 Madrid, Spain.

*To whom correspondence should be addressed.