A SPECIFICALLY ENCAPSIDATED FRAGMENT FROM THE RNA OF TOBACCO MOSAIC VIRUS: SEQUENCE HOMOLOGY WITH THE COAT PROTEIN CISTRON

Kenneth E. RICHARDS, H. GUILLEY, G. JONARD and L. HIRTH

Laboratoire des Virus des Plantes, Institut de Biologie Moléculaire et Cellulaire, 15, rue Descartes, Strasbourg 67000, France

Received 11 April 1974

In vitro reconstitution of tobacco mosaic virus (TMV) from its RNA and protein components occurs in a polar fashion, starting at or near the 5′OH end of the RNA chain [1–4]. Initiation of the reconstitution process is thought to involve a highly specific interaction between the 5′ terminus of the RNA and a double disk of TMV protein [4]. This laboratory has recently published the sequence of an RNA fragment containing 103 nucleotides which is derived from TMV RNA by partial T_1 RNase digestion [5]. Among the digestion products this fragment is unique in that it possesses a high affinity for the double disk, becoming rapidly and quantitatively incorporated into the disk structure when the RNA digest and the protein are mixed together. Other properties of the specifically encapsidated RNA fragment (SERF) and its encapsidation product will be described elsewhere [6,7].

In view of its great affinity for the double disk, we have suggested that SERF might correspond to the 5′OH terminal region of the RNA chain. More recent experiments, however, using TMV RNA specifically marked at the 5′-end with radioactive phosphate [6], have rendered this hypothesis implausible. SERF isolated from such RNA by partial T_1 RNase digestion and reaction with disk protein does not contain radioactivity and therefore could not have originated from the 5′-end of the RNA chain. It appears that the true 5′-terminal region is badly degraded by the T_1 RNase digestion step and hence, presumably, is unable to become encapsidated [6].

If it is granted that SERF derives from the interior of the RNA chain rather than its 5′-terminus, then it is conceivable that the sequence might be a portion of a cistron, in particular the coat protein cistron, the only one of the approx. 10 TMV genes for which the amino acid sequence of its product is known. Accordingly we determined the amino acid coding capacity of SERF in the three possible reading frames and compared the sequences derived in this manner to that of coat protein. Such analysis reveals that SERF codes almost exactly for amino acids 96–129 of coat protein and hence must be largely if not completely homologous to the corresponding region of the coat protein cistron. This homology is illustrated in table 1 which lists: (i) the amino acid sequence of TMV wild type coat protein from glu_{95} to val_{130}; (ii) a possible base sequence for the corresponding portion of the coat protein cistron with the triplet codons assigned according to the rules of the genetic code, and (iii) the sequence of SERF. The two base sequences are identical in 98 out of 105 positions, the only differences being the omission of C and CG at nucleotide positions 21 and 88–89 respectively, the insertion of a U at positions 101a, and a sequence inversion at 68–70. We anticipate that these discrepancies, all of which fall within difficult regions of the SERF sequence, will prove to be illusory, resulting from minor errors in sequencing. We conclude, therefore, that SERF is the portion of the coat protein cistron which codes for amino acids 95–129.

It seems likely that SERF interacts with TMV protein because it possesses critical features, either of base sequence or of secondary structure, in common with the initiation site. It will thus be of considerable interest to compare the sequence of SERF to that of the initiation site when the latter becomes available. It will be of interest, furthermore, to determine if the SERF sequence possesses the ability to inter-

Table 1
Homology between the specifically encapsidated RNA fragment (SERF) and the region of the TMV coat protein cistron coding for amino acids 95–130 of the coat protein.

	95 100 105
TMV coat protein	glu val glu asn gln ala asn pro thr thr ala glu thr
Coat protein cistron	GAA–GUU–GAA–AAU–CAG–GCG–AAC–CCC–ACG–ACU–GCC–GAA–ACG
SERF	(G)AA GUU GAA AAU CAG GCG AA__CCC ACG ACU GCC GAA ACG
	1 10 20 30
	110 115 120
	leu asp ala thr arg arg val asp asp ala thr val ala
	UUA–GAU–GCU–ACU–CGU–AGA–GUA–GAC–GAC–GCG–ACG–GUG–GCC
	UUA GAU GCU ACU CGU AGA GUA GAC GAC GAC GCG GUG GCC
	40 50 60 70
	125 130
	ile arg ser ala ile asn asn leu ile val
	AUA–AGG–AGC–GCG–AUA–AAU–AAU–UU A–AUA–GUX
	AUA AGG AGC__G AUA AAU AAU UUUA AUA G
	80 90 100 101a

The base sequence of the coat protein cistron was derived from the amino acid sequence according to the rules of the genetic code. The base in the degenerate third position of each codon was assigned so as to correspond to the base in the analogous position of the SERF.

act with coat protein while it is still part of the intact RNA chain. The RNA of bacteriophage R17 can bind 1–6 mol of its coat protein to an intercistronic region of the RNA molecule near the end of the coat protein cistron [8]; binding at this location is thought to play a role in gene regulation by inhibiting translation of the neighboring RNA synthetase cistron [9,10]. It is possible that SERF may be the site for a similar sort of gene regulation in TMV.

References

[1] Thouvenel, J. C., Guilley, H., Stussi, C. and Hirth, L. (1971) FEBS Letters 16, 204–206.
[2] Ohno, T., Nozu, Y. and Okada, Y. (1971) Virology 44, 510–516.
[3] Stussi, C., Guilley, H., Lebeurier, G. and Hirth, L. (1972) Biochimie 54, 287–296.
[4] Butler, P. J. G. and Klug, A. (1971) Nature, New Biol. 229, 47–50.
[5] Guilley, H., Jonard, J. and Hirth, L. (1974) Biochimie 56, 181–185.
[6] Guilley, H., Jonard, J., Richards, K. E. and Hirth, L., manuscript in preparation.
[7] Jonard, J., Guilley, H. and Hirth, L., manuscript in preparation.
[8] Bernardi, A. and Spahr, P. F. (1972) Proc. Natl. Acad. Sci. U.S. 69, 3033–3037.
[9] Sugiyama, T. and Nakada, U. (1968) J. Mol. Biol. 31, 431–440.
[10] Spahr, P. F., Farber, M. and Gesteland, R. F. (1969) Nature 222, 455–458.