STRUCTURE AND REPLICATION OF TOBACCO MOSAIC VIRUS

(*Presented by* A. GIERER,
Max-Planck-Institut für Virusforschung, Tübingen, Germany;
contents of paper reported here by ROBLEY C. WILLIAMS)

Tobacco mosaic virus (TMV) is a particularly suitable object for studies on the physical and chemical structure of viruses and on the mutagenic effects of various treatments. The virus may be obtained in gram amounts and readily purified. Its biological activity may be assayed quantitatively by measuring the number of necrotic lesions which it produces on the leaves of suitable tobacco plants; mutants may be detected by noting the characteristics of individual lesions, or the visible expression of symptoms in systemically infected plants. The internal structure of TMV has been determined largely by X-ray analysis, aided and confirmed by electron microscopy. The particle is a nucleoprotein containing about 5 per cent ribonucleic acid (RNA). The protein component consists of 2200 sub-units, presumably chemically identical, arranged in a helical manner to produce the rod-like structure of the virus. Embedded within the protein helix is the RNA, consisting of 6600 nucleotides and appearing to follow the pitch of the helix of the protein sub-units. Each protein sub-unit, of molecular weight 17,500, consists of one chain of 158 amino acids, the complete sequence of which has been independently determined in Tübingen and in Berkeley. There appear to be no simple regularities in the sequential arrangement of the amino acids. Protein sub-units alone (without nucleic acid) are capable of a characteristic reaggregation, under suitable conditions of solution, to form rods which are indistinguishable from those of the intact virus except for a non-uniformity of length.

The RNA component of TMV can be isolated from the virus as a single polymeric entity with a molecular weight of about 2 million. The kinetics of the action of ribonuclease indicate that it exists as a single polynucleotide strand, in contrast to the double-strand character of DNA. In solution the RNA polymer shows evidence of coiling and folding, and this tertiary structure may be reversibly affected by changes of temperature and of ionic strength of the solvent. It appears, then, that the RNA of TMV is a molecule consisting of one linear sequence of 6600 nucleotides, a number far too great to allow a sequential determination of nucleotide arrangement to be made at this time. Some information about sequence is being

gleaned from the products of degradation obtained by specific enzymes. Such degradation fragments have been chromatographed two-dimensionally and show a highly complex pattern which defies any attempt at sequential analysis. However, the pattern of RNA from TMV is characteristic of this material and is easily distinguishable from that of the RNA of the host cell and of other RNA viruses.

The function of the nucleic acid is now generally accepted as being primarily a genetic one. It has been shown, both in Tübingen and in Berkeley, that when RNA is carefully isolated from the virus by any one of several methods it will produce the same symptoms of infection as will the intact, native virus. The level of infectivity, however, is much lower than that of the virus from which it is isolated. It has been shown by meticulous assessment of the amount of protein material associated with the purified RNA that contamination by protein cannot be the source of the infectivity of the RNA. Indeed the infectivity of isolated RNA is not reduced when the level of protein contamination is less than an amount sufficient to provide a single protein sub-unit per RNA molecule. Another bit of evidence for the absence of appreciable contaminant protein is the fact that the purified RNA is not in any way inactivated by anti-serum against the viral protein.

Since the RNA from TMV may be isolated in highly purified form and is infective, it serves as a particularly suitable object for studying the structural basis of the transfer of genetic information. It has been found, in Tübingen and elsewhere, that infection may be initiated by the RNA molecule only in its intact form. Smaller fragments which can be isolated to some extent by centrifugation are inactive. Similarly, treatment by ribonuclease to the extent of causing only a single cleavage of the polynucleotide causes inactivation. On the other hand, the effects of changing the tertiary structure of the molecule are not in themselves inactivating. It appears, then, that the biological activity and genetic specificity are determined solely by the complete sequence of most or all of the 6600 nucleotides within the RNA.

The effects of mutagenic agents upon the viral RNA are helping to clarify the genetic role of the sequence of the nucleotides. The agent most commonly employed is nitrous acid, first shown by Schuster and Schramm to have a mutagenic effect on the RNA from TMV. The chemical action of nitrous acid is known to be the conversion of 3 of the 4 bases of RNA to bases which are similar except for deaminations. One reaction converts cytosine to uracil, both naturally-existing bases. The other 2 reactions convert adenine into hypoxanthine (a close analog of guanine) and guanine into xanthine. The naturally-occurring uracil base is not affected by nitrous acid treatment. The 3 reactions occur at comparable rates and do not cause breakage of the polynucleotide strand. The deamination reaction causes inactivation of the RNA at an exponential rate, indicating that the deamination of a single base can be lethal to the whole RNA polymer. Chemical data correlated with infectivity assay indicate that about 1 in 2 deaminations is, in fact, lethal. The question, then, is whether or not some of the non-lethal deaminations are mutagenic.

Mutants of tobacco mosaic virus can be detected by the symptomology of lesion production and systemic disease, a particularly suitable plant for the study of lesions being Java tobacco. On this plant wild TMV produces diffuse, light chlorotic lesions, whereas RNA which has been treated with nitrous acid will give rise to local necrotic lesions. At low concentration of either virus or RNA the number of lesions is closely proportional to the concentration of the inoculum. With this system for assay Gierer and Mundry have studied in detail the mutagenic action of nitrous acid. Infection with RNA from wild-type TMV yields only about 2 necrotic lesions in 1,000 normal ones, whereas treatment of the virus or of its RNA with nitrous acid results in the production of at least 30 per cent of such necrotic lesions. When nitrous acid is used for a time such that its inactivating effect is small, one finds a linear increase in the number of necrotic lesions with time of treatment. When total infectivity is reduced by only a factor of 2, the number of necrotic lesions reaches 20 times its spontaneous level. The kinetics of the relation of the production of mutant symptoms with the degree of inactivation of the RNA constitute proof that mutants have been produced by the chemical action of the nitrous acid. There is always the certainty that some of the mutants, observed following treatment, have arisen spontaneously and/or have pre-existed in the RNA inoculum. While their quantitative effect cannot be large, there is always the chance that a particular mutant isolated after nitrous acid treatment will not be a chemically-evoked one.

The kinetics of the production of mutants is first-order. In particular, the linear increase of the production of mutants at the beginning of the reaction shows that the deamination of a single nucleotide can be mutagenic. If it is assumed that the efficiency of "plating" is not changed following mutation, the application of target theory to the deamination reaction indicates that there are about 200 nucleotides in the RNA molecule, the deamination of any one of which will lead to a necrotic-lesion mutant. But other symptoms consequent upon mutation have been found, such as host-symptom mutants, and if one includes all the expressions of mutation, it appears that the deamination of any one of about 1000 nucleotides can lead to a mutation.

To summarize the information on target sizes of RNA from TMV: the infecting unit is the entire molecule, i.e., the target size from ribonuclease action is about 6000 nucleotides. This is the same order of size as has been found from radiation inactivation, shown by Lauffer for intact TMV and by Ginoza for its RNA. The target size for inactivation by deamination is about 2000 nucleotides, as shown by Schuster and Schramm. The target size of all deamination mutants, detectable by host symptoms, is about 1000 nucleotides, of which about 200 are accounted for by necrotic lesions.

It is not known which types of deamination are lethal and which are mutagenic. The conversion of cytosine to uracil can be mutagenic, and probably conversion of adenine to hypoxanthine (analogous to guanine) can result in mutation. Vielmetter

and Schuster have shown that guanine deaminations are lethal in bacteriophage DNA. If all the guanine deaminations are lethal in the RNA from TMV, it can be calculated that most of the deaminations of adenine and cytosine will produce viable mutants. But if the deamination of guanine in RNA does not inactivate it, then it would turn out that most of the deaminations of adenine and cytosine are lethal rather than solely mutagenic. Unfortunately it is not known which fraction of base conversions leads to viable mutants and which fraction results in lethality. Both consequences of base conversion seem to be about the same order of size and both together account for most of the nucleotides in the RNA molecule.

What are the chemical consequences of the deamination reaction upon the viral protein found in the progeny generations? Wittmann has analyzed the amino acid composition of the viral protein resulting from mutation by nitrous acid, following reaction conditions that have produced about 10 deaminations per RNA molecule. In many cases it is found that the amino acid composition of the protein of the mutant was like that of the wild-type virus; in about one-third of the cases there was a change. With only one exception the change in amino acid composition was that of the substitution of a single amino acid by another. Such substitution may apparently take place anywhere along the complete peptide chain of the protein. These results lead to interesting consequences. First, it appears that the RNA is poly-functional, in that part of it appears to have the role of a genetic determinant of protein structure whereas other parts appear to have other, but unknown, functions. Second, the deamination of a single nucleotide appears to result in the substitution of only one amino acid by another (the one case observed of 2 substitutions can be accounted for by 2 independent deaminations). This one-to-one relation between single nucleotide deamination and single amino acid substitution is a proof of the postulate of non-overlapping coding schemes, i.e., each amino acid in the protein is determined by an independent group of nucleotides. The target size for any amino acid substitution does not exceed a few nucleotides per amino acid.

The study of effects of specific mutagens, like nitrous acid, is important in finding the coding relation between nucleotide sequence and amino acid sequence. Any hypothesis of coding leads to specific predictions as to which transitions of one amino acid into another can be secured by deamination of a single nucleotide. So far the pattern of changes does not allow the code to be revealed but it does aid in rendering various coding schemes untenable. At present the work with TMV RNA shows compatibility with a triplet code, degenerate or not. It should be emphasized that this approach to the coding problem does not require the use of either sequential analysis of nucleotides nor of genetic mapping.

There are great difficulties encountered in investigating where and how the chemical structures of TMV are produced and assembled in the host cell. Only about one virus particle in a million initiates an infection, it is not possible to infect many cells simultaneously, and the infection spreads by unknown mechanisms from one cell

to the next. But much work has been done on the mechanism of infection, and it is in order to mention some of it which has direct relevance to virus replication. Zech has studied what happens after individual hair cells have been inoculated with TMV. He has found that a few hours following inoculation, but long before intact TMV is detectable, large amounts of nucleic acid are produced in the nucleus (observable by ultraviolet microscopy). Later an excess of RNA appears in the cytoplasm, apparently at the expense of that in the nucleus. Many hours later, intact virus begins to appear. At this time fibrous structures suspectible to digestion by RNAse can be seen in the electron microscope. Zech's findings show that the production of large amounts of RNA is among the first consequences of infection. Whether this material is a precursor of viral RNA, or whether it is the viral RNA itself, is not known. Engler and Schramm have shown that at early stages of infection much more infective RNA may be obtained, by phenol extraction, from the plant tissue than is contained in the intact TMV that is present.

At the more chemical level, Davern and Bonner have reported that 5-fluorouracil, which inhibits DNA synthesis, does not inhibit viral multiplication, thus indicating that the synthesis of viral RNA does not depend on the production of complementary DNA molecules. Staehelin has shown, by pulse-labelling of P^{32}, that RNA is assembled from 5′, rather than 3′, nucleotides. Cleavage of the RNA into 5′ nucleotides leads to excessive labelling of adenine, while cleavage into 3′ nucleotides leads to equal labelling of all bases. The unequal labelling must be due to pool effects of the RNA precursors. It follows that 5′ nucleotides are precursors of RNA synthesis, a result similar to the conclusions drawn from the *in vitro* enzymatic synthesis of nucleic acids, both RNA and DNA.

In summary, it appears that all results on TMV are confirmatory of what is expected from our current ideas on the role of nucleic acid. Many of these experimental confirmations were obtained first with TMV and later generalized to other systems. The basic structure of small viruses (composed of one molecule of nucleic acid and many equivalent protein sub-units), the infectivity of isolated RNA, the evidence in favor of a single-stranded RNA, and the chemical mutation of the isolated genetic material, were all originally discovered in work with TMV and subsequently confirmed for other RNA-containing viruses. At present TMV seems to provide the best system for the study of the chemical effect of mutation on protein structure.

It is well in closing to bring up a few of the principal unsolved problems relating to the function of RNA in the mechanism of infection. Is there a negative template involved in the replication of viral RNA, and if so, is it cellular RNA? Does viral RNA multiply autonomously in the cell without interaction with the chromosomes? Certainly the large target size for mutation suggests autonomous replication of RNA. Does some cellular RNA also self-multiply? What is the reason for the size of viral RNA, of molecular weight about 2 million, which is universally found? It was shown in Tübingen some years ago that most of the cellular RNA is composed of single-

stranded molecules and that the molecular weight of a large part of the microsomal RNA in plant and animal tissues is about 2 million. Is there a relation in size, and in the mechanism of production, of the cellular RNA to the viral RNA? What is the function of those large portions of the RNA of TMV which are apparently not related to the determination of the structure of the protein of the virus, but which are still absolutely necessary for infection? Following inoculation are there early proteins, enzymes, and other substances interacting with the constituents of the host cell before the whole virus is assembled? What is the messenger carrying the genetic information of the viral RNA to the site of protein synthesis? Is the viral RNA itself the messenger, or a sub-unit of it? These are just a few of the questions arising from the state of our current knowledge and ideas on the relation of cells to viruses. The answers to some of them may come from further study of tobacco mosaic virus.

BIBLIOGRAPHY RELEVANT TO PRECEDING PAPER

BAWDEN, F. C. The effect of nitrous acid on tobacco mosaic virus: Mutation or Selection? *Nature, Lond.*, 1959, **184**, 73.

BOEDTKER, H. Some physical properties of infective ribose nucleic acid isolated from tobacco mosaic virus. *Biochim. et Biophysica Acta*, 1959, **32**, 519.

CHEO, P. C., FRIESEN, B. S., and SINSHEIMER, R. L. Biophysical studies on infectious nucleic acid from tobacco mosaic virus. *Proc. Nat. Acad. Sc.*, 1959, **45**, 305.

DOTY, P., BOEDTKER, H., FRESCO, T. R., HASELKORN, R., and LITT, M. Secondary structure in ribonucleic acids. *Proc. Nat. Acad. Sc.*, 1959, **45**, 482.

ENGLER, R., and SCHRAMM, G. Infectious ribonucleic acid as precursor of tobacco mosaic virus. *Nature, Lond.*, 1959, **183**, 1277.

FRAENKEL-CONRAT, H. The role of the nucleic acid in the reconstitution of active tobacco mosaic virus. *J. Am. Chem. Soc.*, 1956, **78**, 882.

FRAENKEL-CONRAT, H., SINGER, B., and WILLIAMS, R. C. Infectivity of viral nucleic acid. *Biochim. et Biophysica Acta*, 1957, **25**, 87.

FRAENKEL-CONRAT, H., and WILLIAMS, R. C. Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components. *Proc. Nat. Acad. Sc.*, 1955, **41**, 695.

FRANKLIN, R. E. Location of the ribonucleic acid in the tobacco mosaic virus particle. *Nature, Lond.*, 1956, **177**, 929.

FRANKLIN, R. E., KLUG, A., and HOLMES, K. C. X-ray diffraction studies of the structure and morphology of tobacco mosaic virus. *Ciba Foundation Symposium on The Nature of Viruses*, 1957, **39**. London: J. and A. Churchill.

FRISCH-NIGGEMEYER, W. Absolute amount of ribonucleic acids in viruses. *Nature, Lond.*, 1956, **178**, 307.

GIERER, A. Structure and biological function of ribonucleic acid from tobacco mosaic virus. *Nature, Lond.*, 1957, **179**, 1297.

GIERER, A. Grösse und Struktur der Ribosenukleinsäure des Tabakmosaikvirus. *Z. Naturforschg.*, 1958, **13b**, 477.

GIERER, A. Die Grösse der biologisch aktiven Einheit der Ribosenukleinsäure des Tabakmosaikvirus. *Z. Naturforschg.*, 1958, **13b**, 485.

GIERER, A. Die Eigenschaften der infektiösen Einheit des Tabakmosaikvirus. "Biochemistry of Virus", *Proc. IVth Int. Congr. Biochem., Vienna*, 1959, p. 58. London: Pergamon Press.

GIERER, A., and MUNDRY, K. W. Production of mutants of tobacco mosaic virus by chemical alteration of its ribonucleic acid *in vitro*. *Nature, Lond.*, 1958, **182**, 1457.

GIERER, A., and SCHRAMM, G. Infectivity of ribonucleic acid from tobacco mosaic virus. *Nature, Lond.*, 1956, **177**, 702.

GIERER, A., and SCHRAMM, G. Die Infektiosität der Ribosenukleinsäure des Tabakmosaikvirus. *Z. Naturforschg.*, 1956, **11b**, 138.

HART, R. G. The nucleic acid fiber of the TMV particle. *Biochim. et Biophysica Acta*, 1958, **28**, 457.

MUNDRY, K. W., and GIERER, A. Die Erzeugung von Mutanten des Tabakmosaikvirus durch chemische Behandlung der Nukleinsäure *in vitro*. *Z. indukt. Abstamm.-u. Vererb.-Lehre*, 1958, **89**, 614.

SCHRAMM, G., SCHUMACHER, G., and ZILLIG, W. An infectious nucleoprotein from tobacco mosaic virus. *Nature, Lond.*, 1955, **175**, 549.

SCHUSTER, H., and SCHRAMM, G. Bestimmung der biologisch wichtigen Einheit der Ribosenukleinsäure des TMV auf chemischen Wege. *Z. Naturforschg.*, 1958, **13b**, 697.

TSUGITA, A., GISH, D. T., YOUNG, J., FRAENKEL-CONRAT, H., KNIGHT, C. A., and STANLEY, W. M. The complete amino acid sequence of the protein of tobacco mosaic virus. *Proc. Nat. Acad. Sc.*, 1960, **46**, 1463.