Reaction of Tobacco Mosaic Virus with Maleic Anhydride and Some Possible Applications to X-Ray Diffraction Analysis*

L. King and R. N. Perham

ABSTRACT: Each protein subunit of tobacco mosaic virus (TMV) contains an acetylated N-terminal residue and two free ε-amino groups. While both ε-amino groups are free to react with a number of reagents in the isolated protein, only that of lysine-68 reacts readily in the intact virus with acetic anhydride or with methyl picolinimidate. The aim of this work was to react the amino group at lysine-68 of each protein subunit in intact TMV with maleic anhydride, so that the olefinic double bond thereby introduced could be used for the attachment of a heavy atom. It was hoped that the heavy-atom derivative of the virus would then be suitable for structural studies by X-ray diffraction, using the method of isomorphous replacement. The latter method requires that a derivative should contain one major heavy-atom site in high yield. It was therefore required that reaction should occur predominantly at lysine-68 and to an extent approaching 100% modification. Since the X-ray diffraction studies of TMV are performed on the intact virus, it was further necessary that adequate reaction should not result in breakdown of the virus rod. The products of the reaction of the virus with different amounts of maleic anhydride were examined by microelectrophoresis, and conditions that resulted in modification at 79–84% of lysine-68 sites, and at 6–9% of lysine-53 sites, without breakdown of the virus, were established.

At the same time, serine or threonine hydroxyl groups unexpectedly reacted to give 0.7–0.75 maleyl ester group for each maleylamino group formed. Tyrosine hydroxyl groups were shown to react with maleic anhydride to give a derivative that was completely hydrolyzed within minutes to tyrosine and maleic acid. Thus it was not possible subsequently to detect any reaction of the tyrosine residues in TMV that might have occurred. It was also shown that the reactivity of serine or threonine residues is not necessarily a function of their special environment in the tertiary or quarternary structure of the virus, *i.e.*, that, in general, maleyl esters may be detected after maleylation in aqueous solution at pH 8–9 with sufficient excess reagent. Since the maleyl ester bond, but not the maleylamino bond, may be cleaved by treatment with hydroxylamine under mild conditions, and since maleyl ester bonds may be cleaved under the same mildly acidic conditions as maleylamino bonds, the formation of esters is not regarded as a serious disadvantage in the use of maleic anhydride for protein modification. Some of the hydroxyamino acids at the carboxyl-terminal portion of the protein in TMV, residues 142–158, were identified as those reacting with maleic anhydride. The results are discussed with reference to other reactions of the amino groups of TMV, the tertiary structure of the subunit, and the maleylation and succinylation of other proteins.

The X-ray diffraction analysis of the structure of tobacco mosaic virus (TMV) is carried out by examination of oriented specimens of virus gels (Gregory and Holmes, 1965). Using the method of isomorphous replacement, the information thus far obtained (Holmes *et al.*, 1968) has depended on heavy-atom derivatives of TMV containing methylmercury (Franklin and Holmes, 1958), lead (Caspar, 1956), uranyl (Holmes and Leberman, 1963), and osmium (R. Franklin, unpublished results, discussed by Klug and Caspar, 1960). Further derivatives are necessary to improve the present three-dimensional electron density map before the polypeptide backbone of the protein subunit can be traced out. It would be particularly valuable if heavy atoms could be introduced into the virus at chemically defined sites, the three-dimensional position of which might then be identified. Information concerning the microenvironment of certain residues obtained, for example, through chemical investigation, could also help in determining the folding of the protein subunit.

The protein subunit of TMV is unusual in that the N terminus is acetylated and it contains only two lysine residues (positions 53 and 68), the properties of which are rather interesting. Despite the fact that both ε-amino groups are available for reaction with a variety of reagents in the isolated protein (Fraenkel-Conrat and Colloms, 1967; Perham and Richards, 1968), only one, that at position 68, is readily accessible in the intact virus for reaction with acetic anhydride (Fraenkel-Conrat and Colloms, 1967) or with methyl picolinimidate (Perham and Richards, 1968). Since neither group will react in the intact virus with 1-fluoro-2,4-dinitrobenzene (Anderer, 1963) nor with trinitrobenzenesulfonic acid (Scheele and Lauffer, 1969), reaction at position 68 must also be hindered for some reagents.

Maleic anhydride has been shown to react specifically with the amino groups of certain proteins to yield derivatives that are stable above pH 3.5 (Butler *et al.*, 1969). It therefore seemed worthwhile to investigate the reaction of this reagent with TMV for two reasons. First, although exhaustive maleylation degrades TMV (Perham and Richards, 1968), it might be expected that it would be possible to limit reaction predominantly to lysine-68 and the stability of the maleylated virus rod would be of interest in view of the charge change induced at that residue. Secondly, the introduction of the maleyl group at a single chemically defined position could provide a useful heavy-atom site. For example, it might be possible to bind mercurials to the olefinic double bond, since mercuric acetate is known to form such complexes (Lucas *et al.*, 1939).

* From the Medical Research Council Laboratory of Molecular Biology, Cambridge, England, and the Department of Biochemistry, University of Cambridge, Cambridge CB2, 1QW, England. *Received August 31, 1970.*

No attempt has been made to investigate the reaction of maleic anhydride with the RNA in TMV, which could conceivably take place.

Materials and Methods

TMV (Vulgare strain), prepared by the method of Leberman (1966), was stored at 4° at a concentration of about 30 mg/ml in sodium phosphate buffer (pH 7), ionic strength 0.05–0.10. TMV protein was prepared by the acetic acid method of Fraenkel-Conrat (1957) and stored at 4°.

Maleic anhydride, citraconic anhydride, and Bicine (*N,N*-bis-(2-hydroxyethyl)glycine) were obtained from British Drug Houses, Poole, Dorset. [1,4-^{14}C]Maleic anhydride (11.5 mCi/mmole) was purchased from the Radiochemical Centre, Amersham, Bucks., and diluted with nonradioactive maleic anhydride to a specific activity of 0.72 mCi/mmole before use. The radioactive maleic anhydride was dissolved in dioxane containing less than 0.02% water. Reagent grade iodoacetic acid was recrystallized from 60 to 80° petroleum ether before use. Ponceau S was obtained from George T. Gurr, London. All other reagents were reagent grade or better. Deionized water was used throughout.

Low-speed centrifugation was carried out using an MSE 18 centrifuge and high-speed centrifugation using a Spinco Model L centrifuge with a Ti 50 rotor. Analytical ultracentrifugation was performed with a Spinco Model E ultracentrifuge fitted with schlieren optics, using a speed of 27,960 rpm. All centrifugations were performed at 4°.

Spectra were recorded with a Unicam SP 800 spectrophotometer. The concentration of virus solutions was obtained from the scan between 260 and 360 nm. The light-scattering correction was found by extrapolating the line joining the 360- and 320-nm values to 260 nm. This correction was subtracted from the extinction at 260 nm. If the corrected extinction at 260 nm is x, the concentration of virus expressed in milligrams per milliliter of virus protein is given by $0.1x/3.01$, the constant 3.01 having been determined by dry weight analyses, assuming the RNA content of the virus to be 5.1% (Knight and Woody, 1958).

pH measurements were performed with a glass electrode using a Radiometer (Copenhagen) TTT 1c titrator. The pH was maintained during trypsin digestion with the TTT 1c in conjunction with the SBR 2c titrigraph and the SBU 1a syringe buret.

Microelectrophoresis of proteins on cellulose acetate strips was carried out using the Beckman Model R-101 Microzone cell. To obtain satisfactory results for TMV protein with this apparatus, a special buffer was developed containing 0.1 M Tris–0.1 M Bicine–and 8 M urea (R. Leberman, unpublished work). In this buffer, which had an apparent pH of 8.6 (glass electrode), good separation was obtained by carrying out electrophoresis at 250 V for 2 hr. Proteins were detected by staining the strips with ponceau S (0.2%(w/v) in 5% (w/v) trichloroacetic acid) and rinsing with 5% (v/v) acetic acid. The strips were cleared by immersion in glacial acetic acid–water–methanol (20:10:70, v/v) for 30 sec, transferred to glass plates and flattened by a roller, and finally dried for 5 min at 100° (Beckman Instruction Manual RM-IM-3, 1965). The relative intensities of the stained protein bands were determined using a Joyce-Loebl microdensitometer.

To prevent oxidation of the SH group, TMV protein was S carboxymethylated before electrophoresis as follows. The virus solution (0.3 ml containing 10 mg of protein) was placed in a stoppered tube with 0.04 ml of 25 mM dithiothreitol, 0.2 ml of 0.2 M Tris-HCl buffer (pH 8), and 0.5 g of urea. Nitrogen was bubbled through the solution and the tube was stoppered. After a few minutes the opalescence of the virus disappeared, and 0.08 ml of a 50 mM solution of iodoacetic acid was added. The tube was flushed with nitrogen again and left at room temperature for 1 hr in the dark. This mixture of RNA and S-carboxymethylated, denatured protein was applied direct to the cellulose acetate strip for electrophoresis. The presence of the RNA did not interfere with the electrophoresis.

Peptide maps of tryptic digests of TMV protein were prepared as described by Perham and Richards (1968). Paper electrophoresis was carried out at pH 6.5 (Perham and Jones, 1967) and descending chromatography in the system butan-1-ol–acetic acid–water–pyridine (15:3:12:10, v/v). Radioautographs were prepared by leaving the maps tightly in contact with Kodak Autoprocess film for 1 week. The position and intensity of the radioactivity peptides after electrophoresis were also determined using the Packard radiochromatogram scanner, Model 7200.

Peptides were eluted from the paper with 6 M HCl and hydrolyzed for 16–17 hr at 100° before amino acid analysis, which was performed with a Beckman 120C automatic analyzer.

Results

Preliminary Experiments. Maleic anhydride was added in small portions to 250 mg of TMV in phosphate buffer at 5°, the pH being maintained at 7.5–8.5 by the manual addition of KOH. A total of 0.4 g of anhydride was added, corresponding to approximately 280-fold molar excess over protein subunit, and the volume of the solution after reaction was 20–25 ml. The reaction of each portion of anhydride was assumed to be complete when no further decrease in pH could be detected. Addition of all the anhydride could be accomplished in less than 30 min. The solution was dialyzed for 18 hr against water at 4°, and the virus pelleted by centrifugation at 39,000 rpm for 2 hr to separate it from soluble protein that might have formed from breakdown of the virus. The virus pellets were dissolved in a total of approximately 10 ml of water, containing a few drops of 1 M KNO_3 to reduce the viscosity, and given a 10-min centrifugation at 10,000 rpm to remove any insoluble denatured protein. The yield of maleyl-TMV purified in this way showed in fact that breakdown of virus under these conditions was negligible.

A sample of the purified virus was examined by cellulose acetate electrophoresis after carboxymethylation in 8 M urea. The pattern obtained was very similar to that shown in Figure 1, strip 2, sample B except that all seven bands of modified protein could be clearly seen. Reaction of groups other than the two ϵ-amino groups must therefore have occurred. Since TMV contains no histidine or α-amino groups, and the one cysteine residue is known to be inaccessible to reagents of the size of maleic anhydride (Fraenkel-Conrat, 1959; Anderer, 1963), the possible sites of additional reaction are the side chains of serine, threonine, and tyrosine. If, as is most likely, these residues react to form esters, some of the bands should disappear after treatment with hydroxylamine, which cleaves esters to give the alcohol and a hydroxamic acid. Accordingly, a sample of carboxymethylated protein in urea was treated with 0.5 M hydroxylamine hydrochloride for 1 hr at pH 9 and room temperature. Electrophoresis then showed only three bands (Figure 1, strip 2,

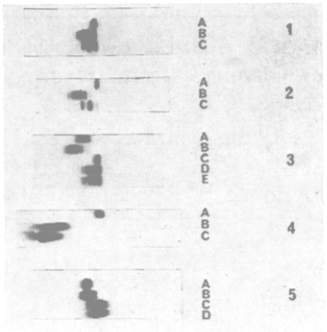

FIGURE 1: Microelectrophoresis at pH 8.3 in 8 M urea on cellulose acetate strips. The strips are aligned so that all control samples (carboxymethylated, unmaleylated TMV protein) are on the same vertical line. Strip 1: reaction of TMV (final concentration, 12–15 mg/ml) with a 50-fold molar excess of maleic anhydride at pH 8–8.5. Control (A), protein from maleyl-TMV (B), protein from hydroxylamine-treated maleyl-TMV (C). Strip 2: reaction of TMV at a final concentration of 20 mg/ml with a 50-fold molar excess of maleic anhydride at pH 8–8.5. A, B, and C are as in Strip 1. Strip 3: maleylation and citraconylation of TMV at pH 8.8–9.3. Conditions of reaction were as for strip 2. No carboxymethylation was carried out in order to show the blurred effect resulting from spurious oxidation of the thiol group. Protein from TMV maleylated at pH 8–8.5 (A), protein from TMV maleylated at pH 8.8–9.3 (B), control (C), protein from TMV citraconylated at pH 8–8.5 (D), and protein from TMV citraconylated at pH 8.8–9.3 (E). Strip 4: maleylation of native and unfolded TMV protein at pH 8–8.5. Control (A), native protein (final concentration, 20 mg/ml) maleylated with 50-fold excess anhydride (B), protein in 8 M urea (final concentration, 10 mg/ml) maleylated with 80-fold excess anhydride to allow for a lower protein concentration than in the case of the native protein (C). Strip 5: unblocking of maleyl-TMV protein by hydroxylamine treatment and by incubation at pH 3.5 and 60° for 18 hr. Protein from maleylation of native protein (A) and protein maleylated in 8 M urea (B), both after hydroxylamine treatment; protein from maleylation of native protein (C) and protein maleylated in 8 M urea (D), both after pH 3.5 treatment. The strip is heavily loaded to ensure detection of any protein unaffected by the unblocking treatments.

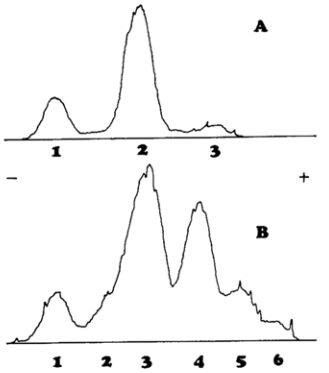

FIGURE 2: Microdensitometer trace of stained and cleared cellulose acetate electrophoresis strips. At pH 8.3 all samples moved toward the anode. The intensities of the peaks were measured in arbitrary units. TMV was maleylated at a final concentration of 12–15 mg/ml with a 50-fold excess of anhydride. Protein from hydroxylamine-treated maleyl-TMV (A), protein from untreated maleyl-TMV (B).

sample C): one faint band of unreacted protein, a strong band assumed to be protein maleylated at position 68 only, and the most acidic band assumed to consist of dimaleylated material. The electrophoretic mobilities showed clearly the charge change of −2 at pH 8.3, resulting from the maleylation of an amino group normally positively charged at this pH to give a negatively charged maleylamino group, and the charge change of −1 at pH 8.3 expected for maleylation of uncharged side chains of hydroxyamino acids.

Optimum Conditions for the Reaction of Maleic Anhydride with Lysine-68 in Intact TMV. By varying the ratio of maleic anhydride to virus, and examining the products by cellulose acetate electrophoresis, conditions for extensive reaction at lysine-68 with little reaction at lysine-53 were obtained. However, under these conditions, electrophoresis showed that esters were formed as described above. Thus in order to obtain the desired specificity, hydroxylamine treatment could not be avoided, although it was found that the esters could be cleaved in the reacted virus, leaving intact TMV containing maleylamino groups only. Electrophoresis showed that treatment of the whole virus with 0.5 M hydroxylamine hydrochloride at pH 8.3–8.4 at room temperature for 18 hr completely broke down the maleyl ester bonds.

In the standard procedure which evolved, 0.083 g of maleic anhydride (50-fold molar excess over protein subunit) was reacted with 300 mg of virus in 15–20 ml of solution. During the reaction the pH changed rapidly. The maleic anhydride, dissolved in dioxane, was therefore introduced gradually by means of a syringe clamped above the virus solution and the pH maintained at 8–8.5. The temperature was kept at 5–10° and the final volume was 20–25 ml. The modified virus was purified by dialysis and centrifugation as described above. The yield of maleyl-TMV before hydroxylamine treatment, calculated spectrophotometrically, was 80%. Similar purification of the unreacted virus also gave a yield of 80%, showing that the losses could be ascribed entirely to the dialysis and centrifugation procedures and not to virus breakdown.

A typical electrophoretic run is shown in Figure 1, strip 1. The transparent electrophoretic strips were examined in a densitometer and typical results are shown in Figure 2. Calculation of the areas under the curves for five separate electrophoreses of the same hydroxylamine-treated preparation (Figure 2A), making the assumptions as to the identity of the bands previously stated, gave values for the reaction at position 68 of between 79 and 84%, and at position 53 of 6–9%. In order to calculate these values, it was assumed that ponceau S, which is an acidic dye, is bound in larger amounts to unreacted protein which contains 13 positive charges, than to monomaleylated protein with 12 positive charges or to dimaleylated protein with 11 positive charges. Thus the areas under the densitometer peaks for dimaleylated (peak 3) and monomaleylated (peak 2) protein were first multiplied by 13/11 and 13/12, respectively, before calculation of percentage modification. (In a separate experiment it was found that increasing the number of negative charges on the protein does not reduce the binding of the dye by protein with a fixed number of positive charges, within the limits of the accuracy of the method.) Similar calculations of the ester content of the material before hydroxylamine treatment (Figure 2B) were made as follows, using 3 runs from the same preparation and neglecting reaction at lysine-53. The sum of the percentage of unreacted protein (peak 1) for maleylated virus before hydroxylamine plus the percentage of material at peak 2 (all of peak 2 disappeared on hydroxylamine treatment) was found to be equal to the percentage of unreacted protein for hydroxylamine-treated virus (calculated from Figure 2A). Thus, peak 3 contained protein reacted at one lysine residue only, and peaks 4, 5, and 6 contained material with one modified lysine residue plus 1, 2, and 3 ester bonds, respectively. It was therefore possible to calculate that 0.7–0.75 maleyl ester bond was formed for each maleyl-

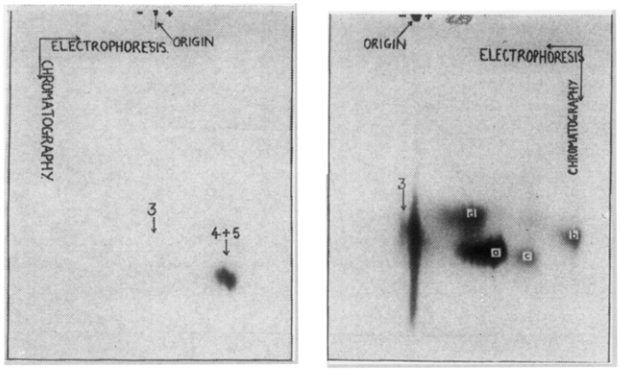

FIGURE 3: Radioautographs of tryptic peptide maps from [1,4-^{14}C]-maleyl-TMV protein. Electrophoresis was carried out at pH 6.5 at 60 V/cm for 1 hr and chromatography in butan-1-ol–acetic acid–water–pyridine (15:3:12:10, v/v) for 17 hr. Tryptic digest from hydroxylamine-treated protein (A, left) and tryptic digest from untreated protein (B, right).

TABLE I: Amino Acid Analysis of the Major Radioactive Tryptic Peptide Obtained from Hydroxylamine-Treated Maleyl-TMV Protein.

Amino Acid	Amt Present (Moles/Mole of Peptide)
Lysine	1.1
Arginine	1.0
Aspartic acid	1.9
Serine	0.9
Proline	0.9
Valine	1.0
Phenylalanine	1.7
Tyrosine	Trace[a]

[a] Low recoveries of tyrosine were always obtained using the conditions described in the text for elution and acid hydrolysis.

amino bond. Errors arising from neglecting maleylation of lysine-53 are small.

Increasing the concentration of all reactants by conducting the identical reaction in a smaller volume increased the intensity and number of bands visible before hydroxylamine treatment, and the intensity of the band corresponding to dimaleylated material after hydroxylamine treatment. Figure 1, strip 2, shows the results of such an experiment in which 300 mg of virus was reacted with 0.083 g of maleic anhydride in a final volume of 15 ml instead of the more usual 20–25 ml.

Examination of maleyl- and control TMV by electron microscopy and analytical ultracentrifugation showed that this limited maleylation and the hydroxylamine treatment did not cause breakdown of the virus into protein or shorter particles. On the other hand, maleylation in hydroxylamine-treated and untreated virus does seem to prevent the end-to-end aggregation that occurs when a control sample is taken through the same purification steps.

Identification of the Reaction Sites in Hydroxylamine-Treated TMV Maleyl-TMV. In order to confirm the assumed nature of the reaction sites in maleyl-TMV, analysis of the radioactive tryptic peptides from [1,4-^{14}C]maleyl-TMV was carried out. The tryptic peptides are numbered according to the scheme of Knight (1963).

Radioactive maleyl-TMV was prepared from 6 μmoles of virus (expressed in terms of protein subunit) by reacting with 250 μmoles of [1,4-^{14}C]maleic anhydride followed by 300 μmoles of cold maleic anhydride in 10 ml of solution. After the usual purification, the [1,4-^{14}C]maleyl-TMV was carboxymethylated in 8 M urea, and the protein precipitated with acetone. The protein was redissolved in 8 M urea containing 0.5 M hydroxylamine hydrochloride at pH 9. After 2 hr at room temperature, the solution was dialyzed against water and then freeze-dried. Tryptic digestion was carried out in 1% ammonium bicarbonate at 37° with a trypsin:protein ratio of 1:100 (w/w). A radioautograph of the peptide map is shown in Figure 3A. The sequences of the two peptides containing the two lysine residues are given in Table III. Modification of lysine-53 would alter the charge of peptide 3 (and hence its position in the peptide map) but not its amino acid analysis, since no cleavage ordinarily occurs at lysine-53 because of the adjacent proline residue. However, modification of lysine-68 would block tryptic cleavage, resulting in the formation of a single tryptic peptide comprising peptides 4 and 5 (Table III). The amino acid composition of the major radioactive peptide is given in Table I, and it can be seen to correspond exactly with the expected analysis for peptides 4 plus 5. The faintly radioactive dual spot running from the origin was assumed to be the result of very slight maleylation occurring in peptide 3, since the latter always has a dual nature (Knight, 1963), and the radioactive dual spot was in the position expected of a maleylated peptide 3. Since no other radioactive spots were found, the assumptions as to the nature of the bands observed in cellulose acetate electrophoresis must be correct. Moreover, since the relative strengths of the radioactive spots for several autoradiographs showed that very little reaction had occurred at lysine-53, it is safe to assume that very little of the electrophoretic band assigned to material maleylated only at position 68 consists of protein maleylated only at position 53. In other words, the calculations as to percentage modification of various residues are not invalidated on this account.

Nature of Hydroxylamine-Labile Bonds Observed in Maleyl-TMV. Although hydroxylamine removed the multiple bands observed on cellulose acetate electrophoresis of maleyl-TMV, and their formation is not therefore a serious disadvantage, it was thought that an investigation of their nature might be of interest. Since the bands were hydroxylamine labile, they must have arisen from maleyl ester formation and were not artefacts, *e.g.*, caused by addition of the single thiol group of TMV across the double bond of maleic acid or maleyl-TMV, or by the addition of one thiol group of dithiothreitol across the double bond of maleyl-TMV with subsequent carboxymethylation of the remaining thiol group. Had such side reactions occurred, they should have been apparent in any event by the appearance of electrophoretic bands with mobilities intermediate between those of the three hydroxylamine-stable bands. It is clear, therefore, that under the conditions used, maleic anhydride is not specific for the amino groups in TMV.

In order to investigate the possibility that a higher pH might render the reaction more specific (Butler *et al.*, 1969, recommended pH 9 for complete and specific blocking of amino groups), TMV was treated with maleic anhydride at pH 8.8–9.3. Since the virus is not stable at high pH for long periods, we had previously avoided pH's above 8.5. Electrophoresis (Figure 1, strip 3) showed that esterification is even

more extensive at pH 9 than at pH 8. The results of the same reaction carried out with citraconic anhydride (Dixon and Perham, 1968) in the same molar ratios are also shown in Figure 1, strip 3. Qualitatively, the results are similar but the reaction appears to proceed more slowly with citraconic anhydride.

The importance of the environment of the hydroxyamino acids was tested by reacting TMV protein with maleic anhydride in the usual way in the presence and absence of 8 M urea. After reaction, the protein samples were dialyzed against water and then precipitated with acetone. They were dissolved, without prior carboxymethylation, in electrophoresis buffer containing 10^{-4} M dithiothreitol. The results of the electrophoresis are shown in Figure 1, strip 4. The extent of maleyl ester formation increases as one proceeds from the intact virus through the isolated, native protein to the protein unfolded in 8 M urea. This indicates that the ester formation in intact TMV cannot be promoted by an hydrophobic environment but is, more likely, a general side reaction during the maleylation of protein amino groups.

The electrophoretic mobility of maleyl protein prepared from native protein was so much greater than that of unmodified protein that the difference could not be accounted for simply by maleylation of the four tyrosine residues per subunit or by the artefacts referred to above. Reaction at serine and/or threonine residues must have occurred. The mobility of the most acidic species in the modified protein prepared from native protein corresponded to reaction of about 50% of all the serine plus threonine residues in the protein. Thus, it is difficult to account for extensive ester formation in the native protein by assuming that the reactive hydroxyl groups are in a hydrophobic environment, since it is very unlikely that 50% of the hydroxyl groups in the protein can be thus situated and at the same time be accessible to maleic anhydride. In separate experiments, *N*-acetyltyrosine (10^{-3} M) was reacted at pH 8 and 5° with a 5-fold molar excess of maleic anhydride. Before pH change had completely ceased, the spectrum was rapidly recorded. At pH 8, *O*-maleyl-*N*-acetyltyrosine has a lower extinction at 275 nm than the same concentration of *N*-acetyltyrosine. The increase in absorption at 275 nm which accompanies the hydrolysis of *O*-maleyl-*N*-acetyltyrosine to *N*-acetyltyrosine and maleic acid can be detected by making rapid scans in quick succession. By this means it was found that hydrolysis was complete in a matter of minutes. Therefore it is clear that *O*-maleyltyrosine residues cannot be contributing to the observed electrophoretic mobility of maleyl-TMV protein, since even if such esters had formed, they would have broken down before electrophoresis took place. In any event, it is unlikely that more than one tyrosine is accessible to maleic anhydride in the intact virus, since it has been shown that only one tyrosine residue, known to be near the surface of the virus, reacts with acetic anhydride and that the ester formed is also not indefinitely stable (Fraenkel-Conrat and Colloms, 1967).

Since the advantage in using maleic anhydride as an amino group blocking reagent rests on the ease of unblocking at pH 3.5, the unblocking of maleyl-TMV protein at this pH was investigated. The maleyl protein samples were reduced and carboxymethylated. Incubating these maleyl protein samples without hydroxylamine treatment, in solution or in suspension at pH 3.5 and 60° for 18 hr, completely unblocked all amino and hydroxyl groups as shown by electrophoresis, while hydroxylamine treatment gave a band corresponding to protein maleylated on both amino groups (Figure 1, strip 5).

TABLE II: Amino Acid Analyses of the Radioactive Tryptic Peptides Obtained from [^{14}C]Maleyl-TMV That Had Not Undergone Hydroxylamine Treatment.[a]

	Amount Present in Moles/Mole of Peptide					
Amino Acid	Calcd for Peptide 4 + 5	Spot a	Spot b	Spot c	Calcd for Peptide 12	Spot d[c]
Aspartic acid	1.8	2	–	–	–	3.5
Threonine	Trace	–	1.7[b]	1.2[b]	2	2.4
Serine	1.1	1	4.2[b]	3.1[b]	6	4.1
Glutamic acid	Trace	–	1.3	1.3	1	2.2
Proline	1.1	1	1.0	0.9	1	2.3
Glycine	Trace	–	2.3	2.0	2	1.9
Alanine	Trace	–	1.0	1.2	1	1.7
Valine	1.0	1	1.0	1.0	1	1.2
Isoleucine	–	–	–	–	–	–
Leucine	Trace	–	0.9	0.9	1	1.0
Phenylalanine	1.8	2	1.1	1.0	1	3.4

[a] Analysis of the acidic and neutral amino acids only were undertaken, since they are sufficient to identify the tryptic peptides of TMV protein. [b] The serine and threonine content of peptide 12 is usually found to be about 60% of the calculated value when peptide hydrolysis is carried out under the conditions described. [c] Since spot d is not a pure tryptic peptide, the analysis given is based on leucine.

The number of hydroxy groups which react in intact TMV is obviously much smaller than the number reacting in the denatured protein. The tertiary and/or quaternary structure of the protein in the intact virus must therefore allow only very few hydroxyl groups to react and it is probable that native conformation is retained. In order to find which tryptic peptides contained the reacted residues and therefore which regions of the protein are accessible to maleic anhydride, a further sample of [1,4-^{14}C]maleyl-TMV was prepared, this time omitting the hydroxylamine treatment. The tryptic digest was not performed in ammonium bicarbonate, since this could conceivably accelerate ester cleavage, but in water maintained at pH 8 with NaOH. The radioautograph obtained from the peptide map is shown in Figure 3B, and the analysis of the spots in Table II. The spots marked b and c contained peptide 12 (Table III). The ratio of the distances moved by

TABLE III: Sequence of Tryptic Peptides 3, 4, 5, and 12 of TMV Protein, as Given by Anderer *et al.* (1965).[a]

Peptide 3: Gln-Phe-Ser-Glu-Val-Trp-Lys-Pro-Ser-Pro-Gln-Val-Thr-Val-Arg
Peptide 4: Phe-Pro-Asp-Ser-Asp-Phe-Lys
Peptide 5: Val-Tyr-Arg
Peptide 12: Ser-Ser-Phe-Glu-Ser-Ser-Ser-Gly-Leu-Val-Trp-Thr-Ser-Gly-Pro-Ala-Thr

[a] Peptide 12 is at the C terminus of the protein. The amide assignments in peptide 3 are those given by Perham and Richards (1968).

spot b and unreacted peptide 12 was found to be 6.3. The molecular weight of peptide 12 is 1694 and the charge −1, and that of peptide 12 maleylated at all 8 hydroxyl groups is 2478, and the charge −9. The theoretical ratio of the mobilities is therefore 6.1 (Offord, 1966). Thus, it was tempting to assume that all 8 hydroxyl groups in the peptide had reacted. However, the electrophoresis results show that not more than 5 hydroxyl groups on any 1 subunit could have reacted since the most acidic band in Figure 1, strip 2, sample B, has a net negative charge of −7 more than the unreacted protein, of which −2 is contributed by the reaction of lysine-68. The calculation of mobilities cannot therefore be valid for such a modification. Spot d did not contain pure peptide and its components have not been positively identified. Spot a contained the peptide 4 + 5. There was a faint spot in the same position as peptide 3 in the radioautograph shown in Figure 3A. The streaking from the origin could mean that the peptides that do not move from the origin on electrophoresis, namely 1 and 10, contain reactive hydroxyl groups, but could also be the result of insoluble material at the origin. These results show that peptide 12 certainly contains reactive hydroxyl groups, and that there may be reactive hydroxyl groups in other portions of the protein chain.

A pH 6.5 electrophoretogram of the tryptic digest was prepared and examined using the radiochromatogram scanner, both immediately after preparation and after 2-months storage of the digest at 4° and pH 8. No loss of maleyl ester containing peptides could be detected.

Discussion

The results described above show that, under suitable conditions, maleic anhydride can react predominantly (79–84%) with the lysine residue at position 68 in the protein subunit with little reaction of the lysine residue at position 53. The specificity of the reaction of maleic anhydride with the amino groups of TMV closely resembles that of acetic anhydride (Fraenkel-Conrat and Colloms, 1967) and methyl picolinimidate (Perham and Richards, 1968). No breakdown of such maleylated virus was detected by electron microscopy and analytical ultracentrifugation, showing that a negative charge can be tolerated at lysine-68. Conformational changes in the virus may have occurred, but this possibility remains to be investigated.

Acetylation of the ε-amino group at position 53 prevents reconstitution of TMV protein with viral RNA (Fraenkel-Conrat and Colloms, 1967), suggesting an important role for that side chain in the self-assembly process, while acetylation of the amino group at lysine-68 does not prevent reconstitution. Moreover, the importance of a positive charge on one or both of the two side-chains at residues 53 and 68 in the aggregation of the virus protein into virus-like rods in the absence of RNA has also been demonstrated (Perham and Richards, 1968). Thus the positive charge at lysine-53 may be very important in the structure of the virus, but it is not necessary that there should be a positive charge at lysine-68. Our present results show that at least 9% of the positive charge at lysine-53 can be converted to negative charge without serious disruption of the virus structure. Since each turn of the protein helix contains approximately 16.33 subunits and the virus rod contains approximately 2000 protein subunits (reviewed by Caspar, 1963), the 6–9% reaction at lysine-53 corresponds to an average of approximately 1 modified subunit per turn of the virus helix. Reaction cannot, therefore, have been restricted to the protein subunits at the end of the virus rod. Moreover, since the particles were not appreciably shortened, reaction presumably occurred randomly along the rod. It is perhaps not surprising that the weakening effect of maleylation of lysine-53 in an average of one subunit per turn of the helix is not in itself sufficient to degrade the virus, since the majority of the forces holding the virus together remain unaffected. However, a chemical modification insufficient to cause degradation may well be sufficient to prevent virus self-assembly. Thus, in view of the inability of protein subunits acetylated at lysine-53 to reconstitute with viral RNA (Fraenkel-Conrat and Colloms, 1967), it seems safe to conclude that subunits maleylated at lysine-53 would be similarly defective. Therefore, while a small percentage of protein subunits maleylated at lysine-53 will not be forced out of the intact virus structure, it is unlikely that they can be induced to go back again if the virus is taken apart and then reconstituted. How high this percentage can be remains to be determined. It is also of interest that maleylation seems to prevent the end-to-end aggregation of the virus rods that occurs when unmodified TMV is taken through the purification steps of dialysis and purification.

The reaction of hydroxyamino acids equivalent to a total of about 0.7 residue/subunit that accompanied maleylation of the amino groups was not negligible. This result was not expected from the results of Butler *et al.* (1969), who found that maleic anhydride was specific for amino groups in the proteins tested. However, since the ester bonds were found to be cleaved by mild treatment of the reacted virus with hydroxylamine, or by incubation at pH 3.5, their formation is not regarded as a serious disadvantage in this work or in the use of maleic anhydride for amino acid sequence studies. However, it is possible that hydroxylamine reacts with the RNA in TMV under the conditions used (Schuster, 1961). Butler *et al.* (1969) recommended a 20-fold molar excess of maleic anhydride over protein amino groups for maleylation under conditions in which the protein presumably has reasonably accessible amino groups. They showed further, in fact, that 90% of the amino groups in chymotrypsinogen could be modified using only 3-fold molar excess of anhydride at a concentration of amino groups that was almost the same as that used in the maleylation of TMV. Freedman *et al.* (1968) observed the formation of hydroxylamine-labile esters during maleylation of immunoglobulins with molar excesses of maleic anhydride over lysine greater than 40-fold, a 40-fold excess being the minimum amount of maleic anhydride that resulted in the complete modification of the lysine residues at the low protein concentration used. In intact TMV the work of Anderer (1963) and Perham and Richards (1968) indicates that the approach to lysine-68 is somewhat sterically hindered and that the limiting size of reagent that can react with this residue is about that of an unsubstituted benzene ring. This probably explains why the complete maleylation of lysine-68 in TMV required a 50-fold molar excess of maleic anhydride rather than the 20-fold excess recommended by Butler *et al.* (1969), with the consequently increased risk of maleyl ester formation. The conclusion, therefore, is that with large excesses of maleic anhydride over lysine residues, serine and threonine will react to give esters stable at least between pH 7 and 9 and that the extent of reaction at the same concentration of amino groups will be different for different proteins for a given excess of reagent, and will depend on the ratio of available amino to available hydroxyl groups. This possibility must be borne in mind when using maleic anhydride and the maximum concentration of protein and the minimum excess of maleic anhy-

dride required to effect complete maleylation should be used.

The behavior of other anhydrides is comparable. For example, Gounaris and Perlmann (1967) modified all 10 lysine residues, 10–13 hydroxyamino acid residues, and 16 tyrosine residues of pepsinogen using 15-fold excess of succinic anhydride based on lysine residues. The tyrosine esters were spontaneously decomposed in 4 hr, but the cleavage of esters formed from the hydroxyamino acids required treatment with 0.7 M hydroxylamine at pH 13. Riordan and Vallee (1964) found that tyrosine as well as lysine residues reacted on succinylation of carboxypeptidase, and that the tyrosine esters decomposed, but the possibility that serine and threonine residues had reacted was not investigated. Colloms (1966), with 15-fold excess acetic anhydride, modified both lysine and two tyrosine residues in TMV protein, and for 40-fold excess anhydride, one more acetyl group, which was not accounted for, was incorporated. The work of Fraenkel-Conrat and Colloms (1967) and Gounaris and Perlmann (1967) suggests that lysine and tyrosine are acylated at the same rates. Gounaris and Perlmann (1967) also found that succinylation of serine and threonine increased sharply after complete modification of lysine and tyrosine, and Freedman *et al.* (1968) found that formation of hydroxylamine-labile groups increased sharply after complete maleylation of lysine residues. Since the serine plus threonine residues exceed the lysine plus tyrosine residues for all the proteins discussed above, serine and threonine residues must be acylated at much lower rates than lysine and tyrosine residues.

These results are in general accord with the studies of Thanassi and Bruice (1966) on the hydrolysis of the monohydrogen phthalate esters, where the mechanism of hydrolysis depends on the pK_a of the conjugate acid of the leaving group. If this is less than approximately 13.5, hydrolysis proceeds with anchimeric assistance of the unprotonated carboxyl group, whereas if it is greater than approximately 13.5, the protonated carboxyl group will be the catalytic species. With maleyl esters, therefore, one would predict that those of tyrosine would be labile at alkaline pH, whereas those of serine and threonine are more likely to break down at acid pH. Such indeed is the case in the present work.

In TMV, reaction of hydroxyamino acids was restricted to very few residues, a maximum of 4 tryptic peptides providing 5 sites under the conditions used. The tryptic peptide positively identified as containing reactive residues toward maleic anhydride is peptide 12, the carboxyl-terminal peptide (Table III). From studies of carboxypeptidase digestion, it is known that the carboxyl-terminal amino acid and the two preceding amino acids project well into the surrounding solution (Harris and Knight, 1952; Tsugita and Fraenkel-Conrat, 1962). It is therefore likely that three of the reactive residues are the carboxyl-terminal threonine-158, serine-154, and threonine-153, especially since at least two of the spots in the radioautograph (Figure 3B) contained peptide 12. It might be useful to discover whether the number of reacting hydroxyamino acids could be reduced by using larger anhydrides. Thus, if an anhydride too large to react at lysine-68 were chosen, it might even be possible to react specifically at one hydroxyamino acid, and perhaps provide a further useful heavy atom site.

Preliminary experiments on the use of maleyl-TMV for binding heavy atoms have indicated that methylmercury is bound to the double bond in the maleyl group. The application of this to the X-ray diffraction study of TMV is under investigation.

Acknowledgments

We thank Mrs. Lynda Simpson for the preparation of TMV and TMV protein, and Mr. K. Edwards for the amino acid analyses.

References

Anderer, F. A. (1963), *Advan. Protein Chem. 18*, 1.
Anderer, F. A., Wittmann-Liebold, B., and Wittmann, H. G. (1965), *Z. Naturforsch. 20B*, 1203.
Butler, P. J. G., Harris, J. I., Hartley, B. S., and Leberman, R. (1969), *Biochem. J 112*, 679.
Colloms, M. (1966), Ph.D. Thesis, University of California, Berkeley, Calif.
Caspar, D. L. D. (1956), *Nature (London) 177*, 928.
Caspar, D. L. D. (1963), *Advan. Protein Chem. 18*, 37.
Dixon, H. B. F., and Perham, R. N. (1968), *Biochem. J. 109*, 312.
Fraenkel-Conrat, H. (1957), *Virology 4*, 1.
Fraenkel-Conrat, H. (1959), *in* Symposium on Sulphur in Proteins, Benesch, R., Benesch, R. E., Boyer, P. D., Klotz, I. M., Middlebrook, W. R., Szent-Györgyi, A. G., and Schwartz, D. R., Ed., New York, N. Y., Academic, p 339.
Fraenkel-Conrat, H., and Colloms, M. (1967), *Biochemistry 6*, 2740.
Franklin, R. E., and Holmes, K. C. (1958), *Acta Crystallogr. 11*, 213.
Freedman, M. H., Grossberg, A. L., and Pressman, D. (1968), *Biochemistry 7*, 1941.
Gounaris, A. D., and Perlmann, G. E. (1967), *J. Biol. Chem. 242*, 2739.
Gregory, J., and Holmes, K. C. (1965), *J. Mol. Biol. 13*, 796.
Harris, J. I., and Knight, C. A. (1952), *Nature (London) 170*, 613.
Holmes, K. C., Klug, .A, and Leberman, R. (1968), Abstracts of IUPAB Meeting, Cambridge, England, June.
Holmes, K. C., and Leberman, R. (1963), *J. Mol. Biol. 6*, 439.
Klug, A., and Caspar, D. L. D. (1960), *Advan. Virus Res. 7*, 225.
Knight, C. A. (1963), Chemistry of Viruses, Vienna, Springer-Verlag, Protoplasmatologia IV Viruses 2, p 52.
Knight, C. A., and Woody, R. B. (1958), *Arch. Biochem. Biophys. 78*, 460.
Leberman, R. (1966), *Virology 30*, 341.
Lucas, H. J., Hepner, F. R., and Winstein, S. (1939), *J. Amer. Chem. Soc. 61*, 3102.
Offord, R. E. (1966), *Nature (London) 211*, 591.
Perham, R. N., and Jones, G. M. T. (1967), *Eur. J. Biochem. 2*, 84.
Perham, R. N., and Richards, F. M. (1968), *J. Mol. Biol. 33*, 795.
Riordan, J. F., and Vallee, B. L. (1964), *Biochemistry 3*, 1768.
Scheele, R. B., and Lauffer, M. A. (1969), *Biochemistry 8*, 3597.
Schuster, H. (1961), *J. Mol. Biol. 3*, 447.
Thanassi, J. W., and Bruice, T. C. (1966), *J. Amer. Chem. Soc. 88*, 747.
Tsugita, A., and Fraenkel-Conrat, H. (1962), *J. Mol. Biol. 4*, 73.