Optical Rotatory Dispersion, Tryptophan Location, and Base Distribution in Tobacco Mosaic Virus*

Ping-Yao Cheng

ABSTRACT: The optical rotatory dispersion of tobacco mosaic virus in the region of 270–320 mμ shows a peak at 292.6 mμ, followed by a trough at 288.8 mμ and a lower peak at 286.3 mμ, in contrast to the single 288-mμ peak previously reported. Evidence has been presented that the novel feature of twin peaks and a trough represents a combination of the Cotton effects from viral ribonucleic acid and from the aromatic side chain of viral protein, with the peaks at 286.3 and 292.6 mμ belonging to the former and the latter, respectively. The magnitudes of the aromatic Cotton effects of A protein (polymer of about six polypeptide chains), tobacco mosaic virus like protein rods, and tobacco mosaic virus were found to be all different. In addition, the aromatic peak was continuously red shifted, accompanied by a reduction in magnitude, in a period of about 0.5 hr during the polymerization of A protein into tobacco mosaic virus like rods. The difference between the protein rods and tobacco mosaic virus is attributed to the influence of the tobacco mosaic virus ribonucleic acid *in situ* on the aromatic Cotton effect. Thus, from these and other data, it is inferred that in tobacco mosaic virus the residue tryptophan (52) of a polypeptide chain is situated near the ribonucleic acid and at or near the surface of the folded chain. The first Cotton effect of TMV–RNA *in situ* was obtained by subtracting the rotations of the protein rods and the residual aromatic Cotton effect from the dispersion curve of the intact virus. It has the peak, crossover, and trough at 286.3, 271.9, and 260.5 mμ, respectively. The difference between the peak position and the calculated peak wavelength, 288 mμ (by assuming nearest-neighbor interaction and random base sequence), is interpreted to indicate some deviation from randomness in the base sequence of TMV–RNA. From optical rotatory dispersion data of dinucleotides it is further inferred that GA occurs more frequently than AG in the RNA. Both conclusions on the base sequence are supported by the data, inconclusive by themselves, alone, on the small nucleotides present in the hydrolysates of pancreatic ribonuclease and of ribonuclease T1. The use of the first Cotton effect of RNA as a probe for the presence of base pairing and the deviation from randomness of base sequence is discussed.

From X-ray diffraction studies the RNA in TMV is known to exist as a single chain coiled between the turns of the 23-Å-pitch helix of protein subunits. The coil diameter is 80 Å and there are 49 nucleotides in one turn of the RNA helix (Klug and Caspar, 1960). Thus, TMV provides us with a unique model system for a single-stranded RNA helix; it has the best known structure and is free of base-pairing. Optical rotatory dispersion of TMV in the wavelength region of 240–320 mμ has been reported (Simmons and Blout, 1960), but the study was carried out with a relatively insensitive manual spectropolarimeter. We have reinvestigated the optical rotatory dispersion of TMV with a more sensitive, automatic-scanning, recording spectropolarimeter. A new optical rotatory dispersion feature, tryptophan location, and deviation of base sequence from randomness of TMV are reported. In addition, the use of the first Cotton effect of an RNA as a probe for the base sequence and base pairing is discussed.

Materials and Methods

A highly purified TMV preparation was kindly put at the author's disposal by Professor C. A. Knight of the Department of Molecular Biology and Virus Laboratory, University of California, Berkeley. The TMV protein, prepared by the procedure of Fraenkel-Conrat (1957), was also generously supplied by Professor Knight. All optical rotatory dispersion measurements were carried out at 26° with an automatic scanning Cary Model 60 recording spectropolarimeter, using a cell path of 1 cm, unless specified otherwise. Virus or protein concentrations were 0.5 mg/ml or less. The A protein was prepared for optical rotatory dispersion measurements by dissolving powdered TMV protein in water to a concentration slightly higher than 0.5 mg/ml and then adjusting the pH to 8.0 with NaOH. Ultracentrifugal analysis showed the A protein solution to be essentially homogeneous with a sedimentation constant of 4 Svedberg units; this, according to Schramm and Zillig (1955), indicates that it contains mainly hexamers of chemical subunits. This A protein was repolymerized into TMV-like rods by maintaining the protein solution at 0.5 mg/ml and pH 5.9 at 26°; the pH 5.9 protein solution was prepared by dialyzing the aforementioned pH 8.0 A protein solution against 500 volumes

* From the Division of Biology and Medicine, Lawrence Radiation Laboratory, University of California, Livermore, California 94550. *Received May 2, 1968.* This work was performed under the auspices of the U. S. Atomic Energy Commission.

TABLE I: Rotatory Dispersion (270–320 mμ) of TMV under Various Conditions.

Solvent (M)	Major Peak		Trough		Minor Peak	
	Position (mμ)	$[\alpha]^a$ (deg)	Position (mμ)	$[\alpha]^a$ (deg)	Position (mμ)	$[\alpha]^a$ (deg)
NaAc (0.03), pH 5.9[b]	292.6	332	288.8	244	286.3	282
NaAc (0.03) + NaCl (1), pH 5.9[b]	292.6	365	288.8	284	286.3	364
Potassium phosphates (0.1), pH 7.14[b]	292.6	388	288.8	308	286.3	352
NaAc (0.03), pH 5.9[c]	292.6	346	288.8	248	286.3	286

[a] Expressed as $[\alpha]_\lambda - [\alpha]_{310\,m\mu}$. [b] Cell path, 1 cm; solution concentration, 0.5 mg/ml. [c] Cell path, 5 cm; solution concentration, 0.1 mg/ml.

of 0.03 M sodium acetate buffer (pH 5.9) for 3 hr at 5°. The optical rotatory dispersion spectrum of the TMV-like rods was taken after the protein solution had been maintained at 26° for 2.5 hr. The concentrations of TMV and TMV proteins were determined from the optical densities of the solutions, in potassium phosphate buffer at 0.05 μ ionic strength (pH 7.1) in cells of 1-cm path length. The optical density for a solution of 1 mg/ml was taken as 2.7 for the virus at 260 mμ and as 1.2 for the protein at 280 mμ.

Results and Discussion

Optical Rotatory Dispersion of TMV at 270–320 mμ. As shown in Figure 1, the optical rotatory dispersion of TMV at 0.03 M sodium acetate (pH 5.9) had a peak at 292.6 mμ followed by a trough at 288.8 mμ, and then a lower peak at 286.3 mμ. All of the other six RNA viruses studied (P.-Y. Cheng, manuscript in preparation) as well as ribosomes from various species (Blake and Peacocke, 1965; McPhie and Gratzer, 1966; Sarkar *et al.*, 1967) show a single peak in the wavelength region of 270–295 mμ. Thus, the question arises: Do the twin peaks and a trough in Figure 1 represent genuine Cotton effects? As shown in Table I, all three extrema were distinctly observed in two other buffered media: (a) 0.03 M sodium acetate and 1 M NaCl (pH 5.9) and (b) 0.1 M potassium phosphates (pH 7.14). Moreover, the two peaks and the trough occupied the same positions in these three media. None of the three extrema can be artifacts arising from the use of too high absorbancy, since the products of solution absorbancy at 280 mμ and cell thickness (in centimeters) were close to 1.2, *i.e.*, far less than 2 (Urnes and Doty, 1961) in these measurements. The TMV solutions used did not appear noticeably turbid, and the other RNA viruses gave a single RNA peak in the same wavelength region even at solutions of much higher turbidities; thus the turbidity of TMV solutions is not expected to produce an anomaly resembling a Cotton effect. The possibility of an artifact simulating a Cotton effect, produced by aggregation of TMV particles, can be ruled out on the following grounds: (a) As indicated in Table I, in the pH 5.9 medium the specific rotation at each wavelength in the region was unchanged when the concentration was reduced fivefold. (b) TMV particles were found to exist as monomers in two of the media studied first, in the pH 5.9 medium within the concentration range used for optical rotatory dispersion studied (Figure 2); second, in the pH 7.14 medium even at a concentration six times that used for most of the optical rotatory dispersion measurements. Neither was the unusual dispersion curve in Figure 1 caused by instrumental difficulty; it was reproduced with the Cary Model 60 spectropolarimeter in the laboratory of Dr. J. T. Yang. Besides, it was not observed with any other materials that we measured immediately before or after the TMV samples. Based on the foregoing considerations, we conclude that the unusual rotatory feature of TMV represents genuine Cotton effects produced by TMV monomers.

This finding of double Cotton effects is in sharp contrast to the previous report of a single extremum in the 270–320-mμ region, at 288 mμ (Simmons and Blout, 1960). Their material was measured in distilled water, a medium not being used in the present study. However, judging from their published optical rotatory dispersion spectrum, it is most likely that the presence of double Cotton effects was undetected because no rotational

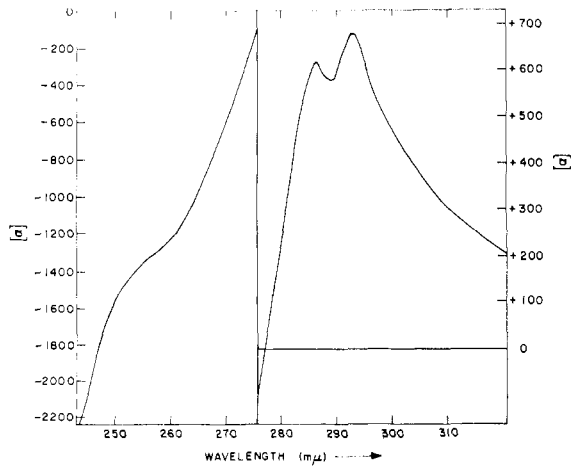

FIGURE 1: Rotatory dispersion (245–320 mμ) of tobacco mosaic virus in 0.03 M sodium acetate (pH 5.9) at 26°. Concentration, 0.5 mg/ml; path length, 1 cm.

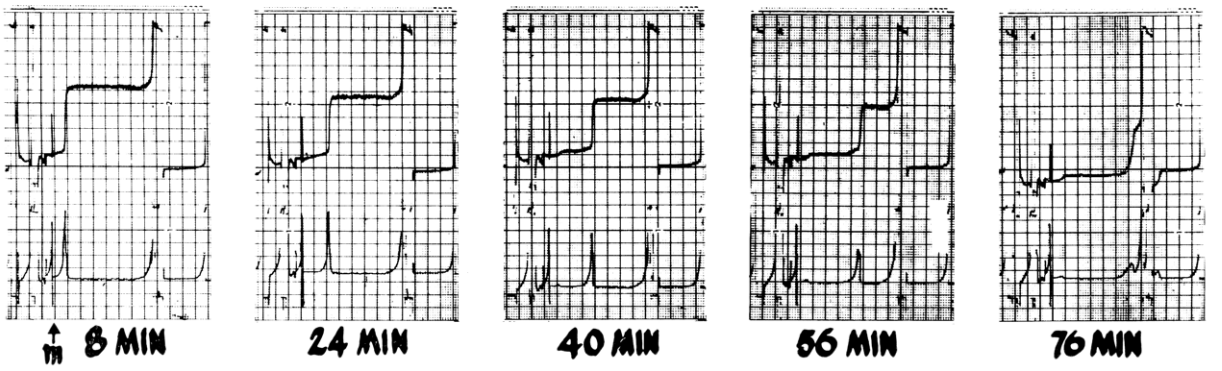

FIGURE 2: Sedimentation velocity patterns of tobacco mosaic virus in 0.03 M sodium acetate (pH 5.9) at 17.2°, concentration, 0.27 mg/ml. All patterns were obtained with a split-beam scanner. The wavelengths of light source were 265 mμ for the first four pictures and 280 mμ for the fifth; the time at which photographs were taken are indicated under each picture. The upper and lower curves in each picture represent plots of c *vs.* x and dc/dx *vs.* x, respectively. Arrow m, meniscus position; centrifugal field to the right. The sedimentation coefficient at 20° and water was calculated to be 187 Svedberg units.

measurement was made at any wavelength between 280 and 300 mμ.

Structural Basis of the Double Cotton Effects. The double Cotton effects observed by us are believed to represent the combined contributions from the viral RNA and viral protein, with the peaks at 286.3 and 292.6 mμ assigned to the former and the latter, respectively. The belief is based on the following considerations: (1) The isolated TMV–RNA in water, which is low in base pairing and thus similar to the viral RNA *in situ*, shows a peak at 285.5 mμ, but no second peak, in the wavelength region considered (Cantor *et al.*, 1966). (2) As shown in Figure 3, the A protein of TMV (aggregates of six polypeptide chains) showed a peak at 292.6 mμ and a trough at about 289 mμ, but no 286-mμ peak. Besides, almost identical peak and trough positions have been reported for carbonic anhydrase by Myers and Edsall (1965), and similar Cotton effects have been observed with several other proteins. (3) Although a large red shift of the RNA peak to the vicinity of 290 mμ has been observed with thermally or pH-denatured RNA, yet the shift was inevitably accompanied by a big drop in rotational magnitude in each case (Samejima and Yang, 1964). Should the 292.6-mμ peak be assigned to TMV–RNA, it would have an α value of 1.9×10^4, considerably larger than the 3.5×10^3 observed for the RNA in water. On the other hand, the rotational magnitude is comparable with those observed for the aromatic Cotton effects of proteins. (4) The existence of double RNA peaks in the wavelength region is neither predicted by theory nor observed experimentally among the many cases studied with various RNA preparations under various conditions. Therefore, the 292.6-mμ peak is considered to be due to TMV protein.

It is well known now that such a protein Cotton effect arises from asymmetrical interactions of the aromatic side chains with the other groups. Since there are three tryptophan, four tyrosine, and eight phenylalanine residues per polypeptide chain of TMV protein, it is natural to inquire into the relative contributions of these aromatic residues to the observed Cotton effect. Simmons *et al.* (1960) suggested that an incipient Cotton effect around 290 mμ observed in their work with the A protein can possibly be assigned to tryptophan or phenylalanine residues. The knowledge accumulated since then seems to indicate that the Cotton effect centered at 291 mμ is dominantly contributed by the ^1L_b–^1A transition of the indole group (Zimmerman and Joop, 1961) on tryptophan, although tyrosine residue can also make minor contributions. The number of tryptophan residues involved cannot be decided from the optical rotatory dispersion data alone.

Effect of Polymerization of TMV Protein into Rod on the Aromatic Cotton Effect. We have obtained two lines of experimental evidence indicating the variation of aromatic Cotton effect with the polymerization. First,

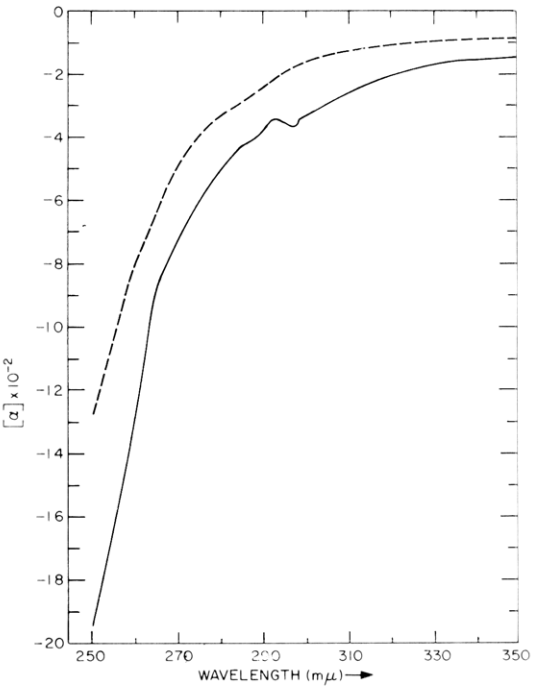

FIGURE 3: Rotatory dispersion (250–350 mμ) of A protein of tobacco mosaic virus and the virus-like protein rods at 26°, (—) A protein: concentration, 0.5 mg/ml (pH 8.2); path length, 1 cm. (---) Protein rods: concentration, 0.5 mg/ml; 0.03 M sodium acetate (pH 5.9); path length, 1 mm.

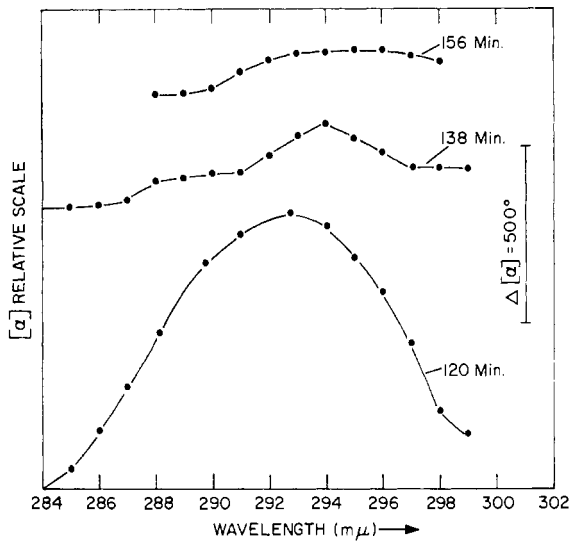

FIGURE 4: Change with time in the aromatic Cotton effect of tobacco mosaic virus protein at pH 5.9. A solution of 0.5 mg/ml of protein at pH 8.2 and low ionic strength was brought to pH 5.9 by dialysis against 0.03 M sodium acetate (pH 5.9) for 3 hr at 5°. The solution was then maintained at 25° to polymerize the protein into rods. The time shown at each curve indicates the number of minutes the solution had been at 25° at the beginning of scanning the curve. Concentration at scanning, 0.247 mg/ml; path length, 5 mm.

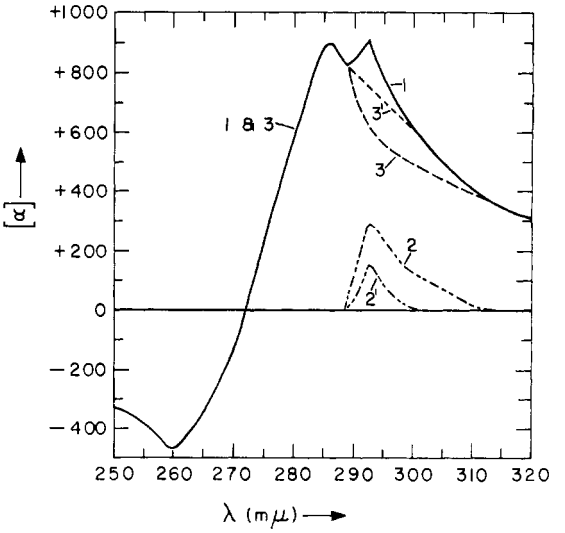

FIGURE 5: Resolution of 250–320-mμ Cotton effect system of tobacco mosaic virus into its components: 1, difference curve obtained by subtracting the rotations of protein rod (Figure 3) from that of the virus (Figure 1); 2, residual aromatic Cotton effect in the difference curve; 3, Cotton effect attributed to the RNA in the virus. Curves 2′ and 3′ are not considered probable, but are the limiting curves of 2 and 3 where the residual aromatic Cotton effect is minimum and the rotation still increases with increasing wavelength at 288.8 mμ.

the Cotton effects for A protein, TMV-like protein rods, and intact TMV all differ in magnitude. At 292.6 mμ the value of α for the Cotton effect was about 290° higher than the RNA curve for virons (Figure 5), but only about 40° higher than the plain curve for A protein, and the protein rods did not exhibit a definite peak at 292 mμ (Figure 3). Secondly, during the process of protein rod formation at pH 5.9, the Cotton effect varied drastically with time within a period of about 0.5 hr; the peak was continuously shifted toward the red with a reduction in magnitude (Figure 4). Since the spectra in Figure 4 were obtained with a solution of 0.3 optical density at 280 mμ in a 0.5-cm cell, any rotatory artifact arising from the use of too strong absorbance is unlikely. Moreover, time dependence of the aromatic Cotton effect was not observed with either A protein or intact virus, even at higher concentrations, when no polymerization was in progress. Thus, the observed strong time dependence of the aromatic Cotton effect is taken to indicate also that the polymerization into TMV-like rod results in alteration of the aromatic Cotton effect. The alteration indicates that polymerization affects either the number, rigidity, and/or other local properties of the optically active structure, such as dielectric constant, presence of a charged group, etc. Evidence has been provided that during the polymerization the protein takes up H^+ ion (Fraenkel-Conrat and Ramachandran, 1959; Ansevin *et al.*, 1964; Scheele and Lauffer, 1967), releases water (Lauffer *et al.*, 1958; Stevens and Lauffer, 1965; Khalil and Lauffer, 1967), and probably has a change in the dielectric constant around carboxyl groups (Ansevin *et al.*, 1964), though these previous studies do not determine whether such environmental changes occur at the vicinity of aromatic amino acids.

Tryptophan Location in TMV. The protein rods look like TMV in the electron microscope. Besides, as is evident from X-ray diffraction patterns, they have the protein subunits arranged in a helix of the same pitch and with the same number of subunits per turn as the intact virus. Moreover, the configuration of the polypeptide chains appear to be unchanged (Franklin, 1955, 1956). Accordingly, the difference in aromatic Cotton effect between the rods and intact virus is attributed to the influence of the viral RNA on the aromatic Cotton effect. Also, one or more residues of tryptophan (much less likely tyrosine) are situated near the RNA. On the basis that the RNA coil has a diameter of 80 Å and is sandwiched between two layers of protein subunits (Klug and Caspar, 1960), the tryptophanyl residue(s) is situated at or near the surface of a protein subunit and is separated from the helical axis of TMV rod at a distance not much different from 40 Å.

Tryptophanyl residue 52 on the TMV polypeptide chain is very probably the optically active aromatic amino acid influenced by the RNA. It can be brought to the vicinity of viral RNA through the bonding of lysine 53 with an RNA phosphate. Such a picture is also in line with two findings indicating that both lysine 53 and the RNA bases are situated in an internal part of TMV. First, lysine 53 is readily acetylated in A protein and only acetylated with difficulty in TMV (Fraenkel-Conrat and Colloms, 1967). Second, the bases in TMV are not accessible to both propylmercuric chloride and mercuric acetate (P.-Y. Cheng, in preparation).

The polymerization of A protein into rods brings some groups from nearby polypeptides into the vicinity of tryptophanyl residue 52 and provides the residue

TABLE II: First Cotton Effect of TMV–RNA under Various Conditions.[a]

	λ_p (mμ)	λ_c (mμ)	λ_t (mμ)	$[\alpha_p] - [\alpha_t]$ ($\times 10^{-4}$)
RNA in TMV				
Present study	286.3	271.9	260.5	2.7
Previous study[b]	288	276	264	2.4
RNA in H_2O, pH 6.9[c]	285.5	271.5	258	0.92
RNA in salt, pH 6.6[c]	282	264	252.5	0.85
RNA (calculated)[c]	288.5	276	261.5	0.79

[a] λ_p, λ_c, and λ_t refer to peak, crossover, and trough wavelengths, respectively. $[\alpha_p]$ and $[\alpha_t]$ are the specific rotations in degrees at λ_p and λ_t. [b] Simmons and Blout (1960). [c] Cantor *et al.* (1966).

with a new environment distinct from that of an aqueous medium. As described in the preceding section, evidence for environmental changes accompanying the polymerization has been provided by Lauffer and his coworkers. Such environmental changes could be responsible for the observed variation of aromatic Cotton effects during the process of polymerization in the absence of viral RNA.

Since each polypeptide chain having three tryptophan residues is attached to three nucleotide residues equally spaced on the RNA coil, it is natural to consider the possibility of one-to-one correspondence between tryptophan and nucleotide. All of the sugar moieties with the possible exception of those at the two ends of the TMV rod have the same distance from, and the same orientation with respect to, the helical axis of TMV particles. The same is true of the phosphate groups. The planes of bases are essentially parallel to the helical axis. With such a structure for the RNA, the one-to-one correspondence concept would anticipate that the interactions among the three pairs of nucleotide–tryptophan are very similar, with three indole rings parallel to each other, whatever moiety on the RNA is responsible for the interaction. Such parallelism has recently been reported from a ultraviolet dichroism study (Schachter *et al.*, 1966). However, this picture requires, in addition, that all three tryptophan residues (17, 52, and 152), widely separated on the polypeptide chain, be at the surface of a subunit and near the RNA coil. Besides being a very stringent constraint for the architecture of TMV particles, the requirement disagrees with the finding that the residues 138–158 are comparatively peripheral on the virus particle (Anderer *et al.*, 1965; Harris and Knight, 1952, 1955; Fraenkel-Conrat and Colloms, 1967; Fraenkel-Conrat and Sherwood, 1967). Therefore, we do not favor this possibility.

First Cotton Effect of Nonbase-Pairing TMV–RNA. Figure 5 shows a difference dispersion spectrum obtained by subtracting the specific rotations of TMV-like protein rods (multiplied by 95%) from those of intact virus, where the dispersion spectra of both virus and protein rods were measured in 0.03 M NaAc (pH 5.9) buffer. As already mentioned in the preceding section, the virus exhibits a greater aromatic Cotton effect than the rods. Consequently the difference spectrum, which would otherwise be attributed to the viral RNA *in situ* alone, contains in addition the residual aromatic Cotton effect. The first Cotton effect of TMV–RNA *in situ* has the peak, trough, and crossover at 286.3, 260.5, and 271.9 mμ, respectively. All of these three wavelengths differ from those of the previously reported difference spectrum obtained by subtracting the rotations of the rods in 0.03 M NaH_2PO_4 from those of intact virus in water (Table II).

Base pairing is impossible for TMV–RNA *in situ*. Thus the present revised first Cotton effect can be compared with that of isolated TMV–RNA with little base pairing. It is expected that the pairing will be inhibited by the coulombic repulsion among phosphate groups and promoted by high ionic strength, which minimizes the repulsion. From Table II, the revised positions of peak, crossover, and trough of the RNA in TMV particles all agree fairly well with the corresponding wavelengths of the RNA in water, and all deviate considerably more from those in the presence of salt.[1] Since it is fair to anticipate that TMV–RNA still contains some residual base pairs even in water, the data support the conclusion that the first Cotton effect of nonbase-pairing TMV–RNA has its peak, trough, and crossing at 286.3, 260.5, and 271.9 mμ, respectively.

Distribution of Base Sequence in TMV–RNA. Table II shows that the first Cotton effect of TMV–RNA in the nonbase-pairing state deviates significantly from that calculated for the RNA by the equation of Cantor and Tinoco (1965). As discussed in the Appendix, the fact that the peak position so calculated (288 mμ) is larger than the experimental wavelength (286.3 mμ) suggests some deviation of base sequence from randomness, with GA probably occurring more frequently on the RNA than AG. Such an indication is supported by the data on the small nucleotides present in the hydrolysates of pancreatic ribonuclease (Rushizky and Knight, 1960) and of ribonuclease T1 (Rushizky *et al.*, 1962). For

[1] As already suggested by Bush and Scheraga (1967), the smaller amplitudes observed with the TMV–RNA in water, in salt, and calculated in comparison with TMV–RNA *in situ* reflect partial base stacking of TMV–RNA in these three cases.

example, the molar concentration ratio of AU/AC is 1.13, which is significantly different from the value of 1.34 (Knight, 1952) for the base ratio of U/C. The molar concentration ratio (GAU + GAC)/(AGU + AGC) was found to be 1.2 from the pancreatic ribonuclease hydrolysate. A similar ratio was estimated for the relative frequencies of GA to AG from the hydrolysate of ribonuclease T1. Only about 61 and 30% of the RNA could be isolated in the form of "soluble" nucleotides after digestion with pancreatic ribonuclease and with ribonuclease T1, respectively; hence the deviation of base distribution from randomness in the hydrolysates could be interpreted as a result of incomplete digestion. However, this interpretation is no longer acceptable in view of the optical rotatory dispersion result obtained with total TMV-RNA. Therefore we conclude that the base sequence of TMV-RNA deviates from randomness, with GA very probably occurring at a higher frequency than AG.

Appendix

λ_p as a Probe for the Randomness of Base Sequence and the Base Pairing.[2] Cantor and Tinoco (1965) estimated $[\phi]$, the molar rotation per residue of a single-stranded polynucleotide for a given wavelength, by the following equation

$$[\phi] = 2\sum_{i=1}^{4}\sum_{j=1}^{4} x_i x_j [\phi_{ij}] - \sum_{i=1}^{4} x_i [\phi_i] \qquad (1)$$

where $[\phi]$ is the molar rotation of dinucleotide of I_pJ, and x_i is the molar fraction of I. Only nearest-neighbor interactions are considered and the bases are assumed to be distributed randomly. To ascertain the validity of the calculation, it is helpful to compare the experimental and the calculated optical rotatory dispersion curves of an RNA with random base distribution and no base pairing. However, TMV-RNA *in situ* is the only RNA so far known definitely to be in the single-stranded state free of base pairing, and, as concluded above, the base sequence deviates from randomness. Nevertheless, we observe that the experimental first Cotton effects of the poly A, poly U, poly C (Cantor *et al.*, 1966), and *Escherichia coli* rRNA (Bush and Scheraga, 1967) in mostly single-stranded state are identical with the dispersion curves so calculated in λ_p, but deviate from the calculated curves in both λ_c and λ_t in each case. Since the problem of base distribution does not exist with homopolyribonucleotides, and the base sequence of the *E. coli* RNA is close to randomness, as indicated by the data on the mono-, di-, and trinucleotides present in the pancreatic ribonuclease hydrolysate of the RNA (Bautz and Hedding, 1964), we feel that the nearest-neighbor calculation is capable of predicting λ_p^0 accurately, if there is no deviation from the randomness. However, the same cannot be said for λ_c^0 and λ_t^0. In other words, the differences in geometric and electrostatic constraints between dinucleotides and a nucleotide polymer already discussed by Cantor *et al.* (1966) evidently seriously affect the accuracy of calculation for λ_c or λ_t but not λ_p. The recent optical rotatory dispersion data for the ribonucleotide sequence $\widehat{A_pA_p}$ in $A_p\widehat{A_pA_p}C_p$ (Inoue *et al.*, 1968) lead to the same conclusion. Accordingly, we take $\lambda_p^0 - \lambda_p^c$ as an indication of the deviation of base sequence from randomness. The converse is not true, however; as evident from the discussion in the next section, many types of the deviation do not affect λ_p by a measurable amount.

To facilitate the evaluation of a deviation from the randomness, the data of Warshaw and Tinoco (1965, 1966) on the λ_p and $[\phi_p]$ of 16 dinucleotides were arranged in decreasing order of λ_p in Table III. It is striking that A_pG and G_pA separate in λ_p by as much as 17 mμ, whereas the separation between two sequence isomers amounts to only 0–3 mμ for the rest of the dinucleotides. Thus, λ_p is shifted to red or blue, depending upon whether A_pG or G_pA is the predominant sequence. On the other hand, an uneven distribution between other sequence isomers usually affects λ_p by an undetectable amount. It can also be estimated from Table III that different combinations of four bases as dimers do not

TABLE III: The Peak of First Cotton Effect of Dinucleotides at 25°.[a]

Dinucleotide	λ_p (mμ)	$[\phi_p] \times 10^{-4}$ (deg)
ApG[b]	291	0.32
GpC[a]	293	0.66
CpG[a]	293	0.66
UpG[b]	293	0.23
GpU[a]	290	0.21
CpC[b]	290	1.07
CpU[b]	288	1.0
UpC[b]	288	0.45
ApC[a]	287	1.05
CpA[b]	289	0.44
UpU[a]	283.5	0.81
ApA[a]	282.5	0.88
ApU[a]	278	0.58
UpA[b]	277	0.11
GpA[b]	277	0.76
GpG[c]	264	6.7

[a] λ_p and $[\phi_p]$ refer to the peak position and the molar rotation at the peak respectively. The data were obtained by Warshaw and Tinoco: (a) 0.01 M phosphate buffer–0.08 M $KClO_4$ (pH 6.9) (1965); (b) the same buffer (pH 7.0) (1966), quoted by Bush and Tinoco (1967); (c) calculated assuming a rigid stacked model and could deviate significantly from the experimental value at room temperature (1966).

[2] λ_p, λ_c, and λ_t designate the wavelengths of the peak, crossover, and trough of first Cotton effect. The superscripts, 0, H_2O, and c to λ denote, respectively, a position in nonbase-pairing state, in water, and calculated by eq 1.

have great differences in λ_p. Considering that four bases are present at equal amounts, the peak positions of (AG) + (UC), (AU) + (GC), and AA + GG + CC + UU are all around 286 mμ. Only the combination (AC) + (GU) has a slightly different λ_p, 288–289 mμ. Thus, a significant deviation of $\lambda_p^0 - \lambda_p^c$ from zero indicates a probable difference in the frequency of occurrence on the RNA chain between A_pG and G_pA.

The λ_p of an RNA can also be used to ascertain the base pairing. As implied in the above discussion, an RNA contains base pairs when its λ_p and λ_p^0 differ. On the other hand, an agreement between λ_p and λ_p^0 indicates a balance between the effects of A–U and G–C pairing on λ_p; nonbase pairing is only a special situation in this case.

Application of the procedure to ascertain the randomness of base sequence and the presence of base pairing calls for a determination of λ_p^0. Neither $\lambda_p^{H_2O}$ nor λ_p^c is always a good approximation for λ_p^0. First, consider $\lambda_p^{H_2O}$. For example, the base distribution of RNA of f2 virus was found to be very close to randomness (Bautz and Hedding, 1964), and yet the $\lambda_p^{H_2O}$ was blue shifted from λ_p^c by as much as 4 mμ (Cantor *et al.*, 1966) indicating a large difference between $\lambda_p^{H_2O}$ and λ_p^0. Most RNA, if not all, still possesses a fair amount of base pairing even in water. Thus, $\lambda_p^{H_2O} - \lambda_p^c$ measures the combined effects on λ_p of base pairing and base sequence. Optical rotatory dispersion alone at the present state of the art is still unable to resolve these two effects. Second, the use of λ_p^c for λ_p^0 is shown to be similarly unreliable by the observed difference between these two quantities, as much as 1.7 mμ, for one RNA (TMV).

References

Anderer, F. A., Wittmann-Liebold, B., and Whitmann, H. G. (1965), *Z. Naturforsch. 20b*, 1203.

Ansevin, A. T., Stevens, C. L., and Lauffer, M. A. (1964), *Biochemistry 3*, 1512.

Bautz, E. K. F., and Hedding, L. (1964), *Biochemistry 3*, 1010.

Blake, A., and Peacocke, A. R. (1965), *Nature 208*, 1319.

Bush, C. A., and Scheraga, H. A. (1967), *Biochemistry 6*, 304.

Bush, C. A., and Tinoco, I., Jr. (1967), *J. Mol. Biol. 23*, 601.

Cantor, C. R., Jaskunas, S. R., and Tinoco, I., Jr. (1966), *J. Mol. Biol. 20*, 39.

Cantor, C. R., and Tinoco, I., Jr. (1965), *J. Mol. Biol. 13*, 65.

Fraenkel-Conrat, H. (1957), *Virology 4*, 1.

Fraenkel-Conrat, H., and Colloms, M. (1967), *Biochemistry 6*, 2740.

Fraenkel-Conrat, H., and Ramachandran, L. K. (1959), *Advan. Protein Chem. 14*, 175.

Fraenkel-Conrat, H., and Sherwood, M. (1967), *Arch. Biochem. Biophys. 120*, 571.

Franklin, R. E. (1955), *Biochim. Biophys. Acta 18*, 313.

Franklin, R. E. (1956), *Nature 177*, 929.

Harris, J. I., and Knight, C. A. (1952), *Nature 170*, 613.

Harris, J. I., and Knight, C. A. (1955), *J. Biol. Chem. 214*, 215.

Inoue, Y., Masuda, M., and Aoyagi, S. (1968), *Biochem. Biophys. Res. Commun. 31*, 577.

Khalil, M. T. M., and Lauffer, M. A. (1967), *Biochemistry 6*, 2474.

Klug, A., and Caspar, D. L. D. (1960), *Advan. Virus Res. 7*, 225.

Knight, C. A. (1952), *J. Biol. Chem. 197*, 241.

Lauffer, M. A., Ansevin, A. T., Cartwright, T. E., and Brinton, C. C., Jr. (1958), *Nature 181*, 1338.

McPhie, P., and Gratzer, W. B. (1966), *Biochemistry 5*, 1310.

Myers, D. V., and Edsall, J. T. (1965), *Proc. Natl. Acad. Sci. U. S. 53*, 169.

Rushizky, G. W., and Knight, C. A. (1960), *Proc. Natl. Acad. Sci. U. S. 46*, 945.

Rushizky, G. W., Sober, H. A., and Knight, C. A. (1962), *Biochim. Biophys. Acta 61*, 56.

Samejima, T., and Yang, J. T. (1964), *Biochemistry 3*, 613.

Sarkar, P. K., Yang, J. T., and Doty, P. (1967), *Biopolymers 5*, 1.

Schacter, E. M., Bendet, I. J., and Lauffer, M. A. (1966), *J. Mol. Biol. 22*, 165.

Scheele, R B., and Lauffer, M. A. (1967), *Biochemistry 6*, 3076.

Schramm, G., and Zillig, W. (1955), *Z. Naturforsch. 106*, 493.

Simmons, N. S., and Blout, E. R. (1960), *Biophys. J. 1*, 55.

Stevens, C. L., and Lauffer, M. A. (1965), *Biochemistry 4*, 31.

Urnes, P., and Doty, P. (1961), *Advan. Protein Chem. 16*, 401.

Warshaw, M. M., and Tinoco, I., Jr. (1965), *J. Mol. Biol. 13*, 54.

Warshaw, M. M., and Tinoco, I., Jr. (1966), *J. Mol. Biol. 20*, 29.

Zimmerman, H., and Joop, N. (1961), *Z. Elektrochem. 65*, 61.