Vol. XXI No. 4 December, 1976

The SPEX Speaker

INDUSTRIES, INC. · 3880 PARK AVENUE · METUCHEN, N. J. 08840 · ☎ (201) 549-7144

RAMAN SPECTROSCOPY OF VIRUSES AND PROTEIN-NUCLEIC ACID INTERACTIONS

G.J. Thomas, Jr.

Department of Chemistry, Southeastern Massachusetts University, North Dartmouth, Massachusetts 02747

The virtues of Raman spectroscopy as a structural probe of biological molecules were recognized long before the advent of modern photoelectric-recording instruments and laser-excitation sources [1]. For example, Edsall and collaborators predated the laser era by several decades in their pioneering Raman studies of the ionization properties of amino acids. This work culminated in the first published Raman spectrum of a biopolymer, lysozyme, a feat accomplished with the Toronto arc source by Garfinkel and Edsall in 1958 [2]. Raman spectra of several carbohydrates were also reported as early as 1942 by Spedding and Stamm [3]. It was not until the late 1960's, however, that the first Raman spectrum of a model nucleic acid was reported, again with excitation by a Toronto arc [4].

In recent years there has been no shortage of innovative applications of laser-Raman experimentation to the solution of problems of biological interest [5]. One example, recently described in the SPEX SPEAKER [6], has been the exploitation of resonance-Raman scattering of visual pigments to learn more about the molecular basis of the visual process. The example of present interest is in the field of virus research where Raman spectroscopy appears to offer promise as a probe of interactions which stabilize the complex architecture of virus particles [7].

Raman spectroscopy of complex biological materials usually goes hand-in-hand with studies of simpler or better characterized model systems — and viruses are no exception. The simplest known viruses are nucleoproteins, that is, they consist of a nucleic acid genome (RNA or DNA) in specific combination with protein. Accordingly these native nucleoproteins yield Raman spectra which are often better understood by analogy with spectra of model nucleoproteins of known structure, as well as by comparison with spectra of the separated viral nucleic acid and protein components. This is the approach we have taken in our work on viruses at S.M.U., a continuing study conducted in collaboration with Prof. K.A.Hartman and his associates at the University of Rhode Island.

This article introduces the reader to some basic structural properties of viruses and model nucleoproteins and their nucleic acid and protein components. The suitability of laser Raman spectroscopy for studies of virus structure and assembly is also examined, and some consideration is given to the limitations of the method.

Composition and Morphology of Viruses

To the spectroscopist a single virus particle (virion) looms as a massive and exceedingly complex structural entity. This can be appreciated by considering one of the simplest and better characterized virions, that of the Tobacco Mosaic Virus (TMV), a portion of which is shown in Fig 1. TMV consists of a single ribonucleic acid molecule (TMV RNA), of molecular weight about 2 million, coated with 2130 protein molecules (hereinafter called the coat protein). Each molecule of coat protein has a molecular weight (MW) of about 17 500. The total mass of the rod-shaped virion thus approximates 4×10^7 amu!

TMV RNA (Fig 2) is itself a heteropolymer containing four different kinds of sub-unit (nucleotide), namely, adenosine monophosphate (AMP), guanosine monophosphate (GMP), cytidine monophosphate (CMP), and uridine monophosphate (UMP). These nucleotides number about 1650, 1400, 1450, and 1750, respectively, in TMV RNA. Finally, there are 18 varieties of repeating unit (amino acid), totaling 158, in each molecule of coat protein [8]. Interactions between TMV RNA and coat protein are believed to play a major role in stabilizing the virion structure. In view of the complexity of this rather simple virion, it is clear that a viable spectroscopic approach to virus structure must differ fundamentally from the protocol usually employed in Raman spectroscopy of small molecules or even of relatively large organic molecules.

In addition to TMV there are a number of other relatively simple and morphologically well-defined viruses of present concern. These include the Turnip Yellow Mosaic Virus (TYMV) and the bacterial viruses (or bacteriophages) MS2, Pf1, and fd.

Like TMV the plant virus TYMV contains RNA and protein yet differs greatly from TMV in molecular organization. The TYMV virion is spheroidal in shape with the TYMV RNA molecule (MW 2

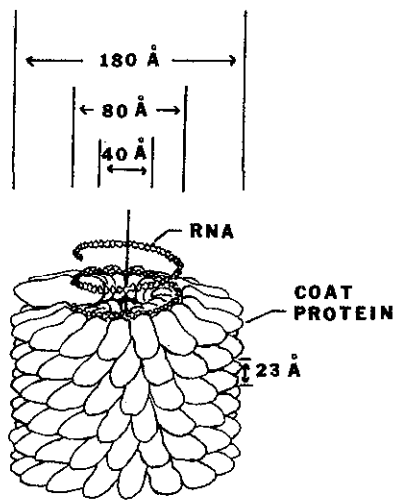


Fig 1 Schematic drawing of a segment (about 1/20th)of the Tobacco Mosaic Virus particle with the coat protein molecules removed from the top two turns to show the encapsulated RNA strand. The helical arrangement of protein molecules about the long axis of the particle has a repeat unit extending 69Å in the axial direction. The repeat contains 49 protein molecules and about 150 nucleotides, which are distributed over three turns of the helix of 23Å pitch. (From data of [8].)

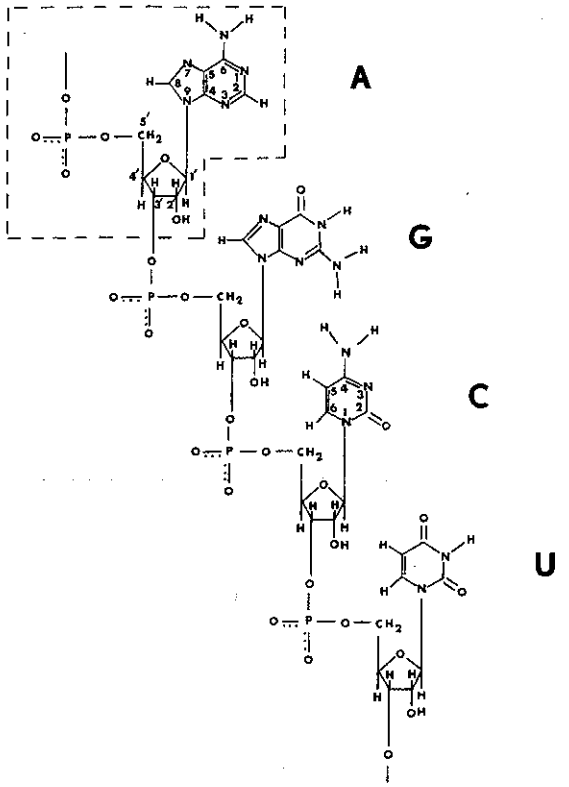


Fig 2 The primary or covalent structure of a segment of the ribonucleic acid (RNA) molecule with the nucleotide sequence -A-G-C-U-. A single nucleotide (A) is indicated by the enclosed area. Many hundreds of nucleotides of each type are present in a typical viral RNA molecule, and the actual sequence of nucleotides is determined by the genetic information content. The numerous C=O, C=N, and C=C double bonds which are present in the heterocyclic bases are responsible in large part for the intense Raman scattering spectrum of RNA.

x 10^6) residing within a shell (capsid) of 180 coat protein molecules (MW 20 000). The capsid is itself a rather stable structure, retaining its integrity even when the encapsulated RNA molecule is removed. There are obviously strong and specific protein-protein interactions to account for such a stable capsid structure; they have yet to be identified, however.

The bacteriophage MS2 is similar to TYMV in both morphology (spheroidal) and composition (one RNA molecule of MW 1.1 x 10^6 plus 180 coat protein molecules of MW 13 750). A recently obtained [9] image reconstruction from electron micrographs allows visualization of the spheroidal RNA bacteriophage as shown in Fig 3 (a).

The bacteriophages Pf1 and fd are members of a family of viruses known as filamentous bacterial (FB) viruses. These, like TMV, are rod-shaped virions with cross sectional diameter (60Å) that is small in comparison to length (about 9000Å in the case of fd and 18 000Å in the case of Pf1). Moreover, they contain DNA instead of RNA as the viral genome. Thus the fd virion consists of a circular and single-stranded DNA molecule (MW 2 x 10^6) sheathed with about 2400 molecules of coat protein (MW 5200). The filamentous shape of fd is revealed by electron microscopy [8] as shown in Fig 3 (b).

In addition to their nucleic acid and coat protein components, many bacteriophages also contain one or a few molecules of a minor protein ("A" protein) which is responsible for attachment of the virion to the cell wall or membrane of the host. The A protein may also enter the host along with the viral genome during infection.

A summary of compositional and morphological properties of these viruses is in Table 1. It is stressed that the entries of Table 1 are among the simplest viruses known, both with respect to molecular composition and structural organization. For comparison with more complex viruses which contain additionally various lipid and carbohydrate components or structural appendages (such as tails), the reader is referred to Knight [8] and Casjens and King [10].

Although viruses have been studied for many years and by a number of physico-chemical methods [8,10], very little is known about their detailed structures and stabilizing interactions. For example, with respect to the bacteriophages mentioned above some unanswered questions of fundamental importance are the following: What is the conformation of the nucleic acid genome (RNA or DNA molecule) in the native virion? What types of interactions are present between subgroups of the nucleic acid genome (between nucleotides for instance)? Do these features of nucleic acid conformation change, and if so how do they change, when the genome extrudes from the native virion, as happens during the process of viral infection? Do specific protein-nucleic acid interactions occur in these viruses, and if so which subgroups of protein and nucleic acid molecules are involved? What are the prevailing ionization states of acidic and basic functional groups in the viral nucleic acid and coat protein molecules? Where models of virion structure have been proposed, are these consistent with the conformational properties of the viral constituents?

Answers to these and related questions are required in order to find effective chemotherapeutic agents which may interfere with virus assembly and thus control viral infection.

Role of Raman Spectroscopy

The rationale for selecting laser-Raman spectroscopy as a probe of virus structure is that the Raman vibrational spectrum contains answers to the important questions posed above, and that such information can be obtained more easily than by other methods currently available. Before proceeding to the discussion of spectra of viruses we consider briefly some advantages of the Raman effect vis-a-vis other physico-chemical techniques.

First, in contrast to methods of x-ray diffraction, Raman spectroscopy offers greater speed, simplicity, and versatility. Potentially the structural information content of x-ray diffraction patterns is considerably greater than that of vibrational spectra when both can be interpreted. Up to the present, however, the resolution of x-ray data of viruses has not been sufficient to disclose details of virus structure at the molecular level. Ordinarily only the gross morphological properties of the capsid and the symmetry of the assembled coat

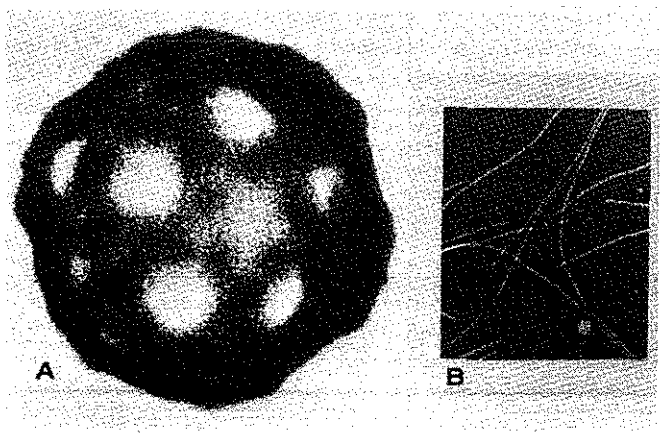


Fig 3 (a) A three-dimensional image reconstruction from electron micrographs of bacteriophage f2, which is very closely related and probably isostructural to bacteriophage MS2. The diameter of the particle is 260Å, and the surface consists of aggregates of coat protein molecules which form a shell or capsid surrounding the enclosed RNA molecule. (From [9].) (b) An electron micrograph of bacteriophage fd. Each filamentous particle has a diameter of 15Å but is too long to be shown in its entirety at this magnification. The coat protein molecules overlap one another in the manner of fish scales and form a sheath which envelops an elongated loop of viral DNA. (From [8].)

Table 1
Properties of Some Simple Plant and Bacterial Viruses

Virus	TMV[a]	TYMV[a]	MS2[a]	fd[b]
Type	plant	plant	bacterial	bacterial
Host	tobacco	turnip	E. coli	P. aeruginosa
Shape	cylindrical	spheroidal	spheroidal	cylindrical
Dimensions (Å)	180 x 3300	300 (diam)	260 (diam)	60 x 9000
Mass (amu)	40×10^6	6.0×10^6	3.6×10^6	15×10^6
Genome				
Type	RNA	RNA	RNA	DNA
Topology	linear strand	linear strand	linear strand	circular strand
Mass (amu)	2×10^6	2×10^6	1.1×10^6	1.8×10^6
Weight % of Virion	5	34	30	12
Fraction AMP	0.28	0.23	0.23	0.24
Fraction GMP	0.23	0.17	0.26	0.20
Fraction CMP	0.22	0.38	0.26	0.22
Fraction UMP (TMP)	0.27	0.22	0.25	0.34
Coat Protein				
Mass (amu)	17 500	20 000	13 750	5200
Number per virion	2130	180	180	2400
A Protein				
Mass (amu)	—	—	35 000	60 000
Number per virion	—	—	1	4
Total Protein				
Weight % of Virion	95	66	70	88

[a]Data of [8].

[b]Data of [11] and citations therein.

protein are revealed. A conspicuous limitation of x-ray studies to date is the lack of information about the secondary structure of the nucleic acid genome and potential sites for its interaction with the coat protein.

Second, Raman spectroscopy imposes no requirement of sample crystalizability. Spectra of satisfactory quality are obtained not only from amorphous solids but also from aqueous suspensions and solutions. The latter, with either H_2O or D_2O as the solvent, are especially important since the chemistry of life takes place in an aqueous medium. Here Raman spectroscopy is also decidedly advantageous over infrared spectroscopy.

Third, in contrast to spectra from other available techniques (such as UV, CD, ORD, and NMR) the Raman spectrum of a virus contains separately resolvable features, assignable to different molecular subgroups of *both the nucleic acid and the protein components*. As will be seen below, this straightforward but important characteristic of the Raman vibrational spectrum renders it far more valuable than other types of spectra in which bands of different virus constituents or subgroups cannot be easily distinguished from one another.

We shall progress to a discussion of the Raman spectra of several viruses and nucleoproteins after considering the information content of Raman spectra of nucleic acids and proteins, the major constituents of viruses.

Raman Spectroscopy of Nucleic Acids

As shown in Fig 2, the building blocks of a nucleic acid are of four main types, each called a nucleotide. The RNA nucleotides, AMP, GMP, CMP, and UMP, each consist of a purine (adenine, A, or guanine, G) or pyrimidine base (cytosine, C, or uracil, U), covalently bonded by a glycosidic linkage to the furanose sugar *ribose*. The ribose moieties are in turn bonded by an ester linkage (5'-carbon) to an orthophosphate group. A second phosphate-ester linkage (3'-carbon) connects the nucleotides in a polymeric chain or "backbone." In DNA the pyrimidine thymine (abbreviated T) replaces U, and deoxyribose replaces ribose. Otherwise, covalent bonding is the same in both RNA and DNA.

The bases, A, G, C, and U (or T), by virtue of their aromaticity, are expected to generate intense Raman scattering. The positions (cm^{-1}), intensities, and origins of their Raman lines have been much discussed for some time [12, 13] as have the Raman-active vibrations of orthophosphate esters [14]. In contrast, Raman lines due to vibrations of ribose or deoxyribose are ordinarily weak (except in the CH stretching region, 2800-3000 cm^{-1}) and therefore contribute little to the Raman spectrum of a given nucleotide in comparison to contributions of the base and phosphate groups [12, 15]. Reflecting optimistically on this situation, we may take heart that the spectra of nucleotides (and therefore of nucleic acids also) are "uncluttered" by Raman lines due to the sugar groups. On the other hand, we may view with some dismay the absence of intense Raman lines that could be potentially helpful in diagnosing the furanose ring structure.

In discussing nucleic acids it is important to keep in mind that the primary or covalent structure (including nucleotide sequence) is not the only factor which determines the three-dimensional configuration and function of the molecule. Also of importance is the secondary structure, the interactions between sub-units of the macromolecule. The typical viral RNA molecule is single-stranded, but where the nucleotide sequence permits, the chain can fold back upon itself and twist to form short hairpin-like helical segments stabilized by Watson-Crick base pairs (hydrogen-bonded AU and GC pairs) [16]. In both paired and unpaired regions of the chain the RNA bases may also be stacked one above another, more or less as steps of a stairwell. The amounts and kinds of base pairing and base stacking constitute the nucleic acid secondary structure, which is undoubtedly of functional significance. We have shown previously [17, 18] that Raman spectroscopy offers one of the best methods available for determining the secondary structure of aqueous RNA. This is because the secondary structure has pronounced effects upon the frequencies and intensities of Raman scattering by certain vibrations of the nucleotide sub-units of RNA, as will be seen below. Many other workers have made additional important contributions to determining the effects of secondary interactions on Raman spectra of nucleic-acid model compounds (for reviews see [5] and [13]).

Perhaps at this point it is appropriate to illustrate with an example. Fig 4 shows Raman spectra of H_2O and D_2O solutions of the purified RNA extracted from bacteriophage MS2. The prominent lines in the spectra are assigned to ring-stretching vibrations of the purine and pyrimidine bases or to phosphorus-oxygen stretching vibrations of the backbone repeat unit. This is accomplished by straightforward analogy with Raman spectra of the nucleotide building blocks [12-15]. The data of Fig 4 also confirm that all external functional groups of the bases (namely, C=O, NH_2, etc...) are in the neutral "keto-amino" tautomeric forms and that the backbone phosphate groups are ionized (PO_2^-).

It will be noted that at the conditions employed in Fig 4 interference from H_2O solvent is negligible, except in the interval 1620-1660 cm^{-1} (the so called "double-bond" region). Because of this fact, vibrational frequencies of the structurally important carbonyl groups (C=O in U, G, and C) are usually sought from spectra of D_2O solutions, since the double-bond region there is unobscured.

In the case of RNA and its analogs the formation of secondary structure can be recognized rather easily by raising the temperature of solution to a point (say, 80 to 90°C) at which such structure is for the most part thermally disrupted. Thereupon the Raman spectrum differs in many respects from that of fully structured RNA (say, below 30°C). We may summarize the major differences between spectra of RNA containing appreciable secondary structure (low T) and RNA essentially devoid of secondary structure (high T) as follows: The ordered secondary structure of RNA produces

(1) intensity losses in several Raman lines due to vibrations of the purine and pyrimidine rings (Raman hypochromism),
(2) intensity increases in certain other Raman lines due also to vibrations of the purine and pyrimidine rings (Raman hyperchromism),
(3) intensity reversals in Raman lines due to external C=O stretching vibrations of U, G, and C, and
(4) a strong and sharp Raman line near 812 ±2 cm^{-1}, due to the ordered RNA backbone, which shifts to 795 cm^{-1} upon disordering of the backbone.

In the absence of a comprehensive theory to adequately explain these effects an empirical approach has been followed. Accordingly, a substantial body of data has been accumulated to correlate the above (and similar observations on DNA) with specific types of nucleic acid secondary interactions [5, 13]. Table 2 summarizes several important Raman lines of RNA, their assignments, and structural significance. This tabulation is based upon a subjective judgment as to which of the currently available data on model compounds are most appropriately transferred to RNA. Therefore some revisions can be expected as additional data become available.

Many of the correlations cited in Table 2 can be visualized by inspection of Fig 4, which shows the temperature dependence of Raman scattering from MS2 RNA. The spectral changes resulting from thermal denaturation of RNA secondary structure can be ascribed specifically to the rupture of hydrogen bonding between A and U and between G and C, to the "melting" of sequences of stacked bases, and to the transition of the RNA chain from a highly ordered backbone geometry (called the A-helix geometry) to a largely disordered one [5, 13].

Fig 4 Raman spectra in the regions 300-1800 cm^{-1} (amplification A=1) and 2800-3100 cm^{-1} (A=1/3) of RNA extracted from MS2 virions. (a) H_2O solutions at 0 and 80°C. (b) D_2O solutions at 0 and 80°C. The RNA concentration in each case is 40$\mu g/\mu l$. Raman frequencies of the prominent lines are listed in cm^{-1}, and assignments are given along the abscissa to RNA base (A, U, G, C), sugar (R), or phosphate (P) groups. The line at 925 cm^{-1}, labeled Ac^-, is due to acetate buffer. See also Fig 2 and [42]. Comparison of (a) and (b) shows the advantage in the double-bond region (1600-1700 cm^{-1} interval) of D_2O over H_2O as a solvent for Raman spectroscopy. On the other hand, the conformationally sensitive frequency of the RNA backbone (near 812 cm^{-1}) is unaffected by the change of solvent but greatly altered by the change of temperature.

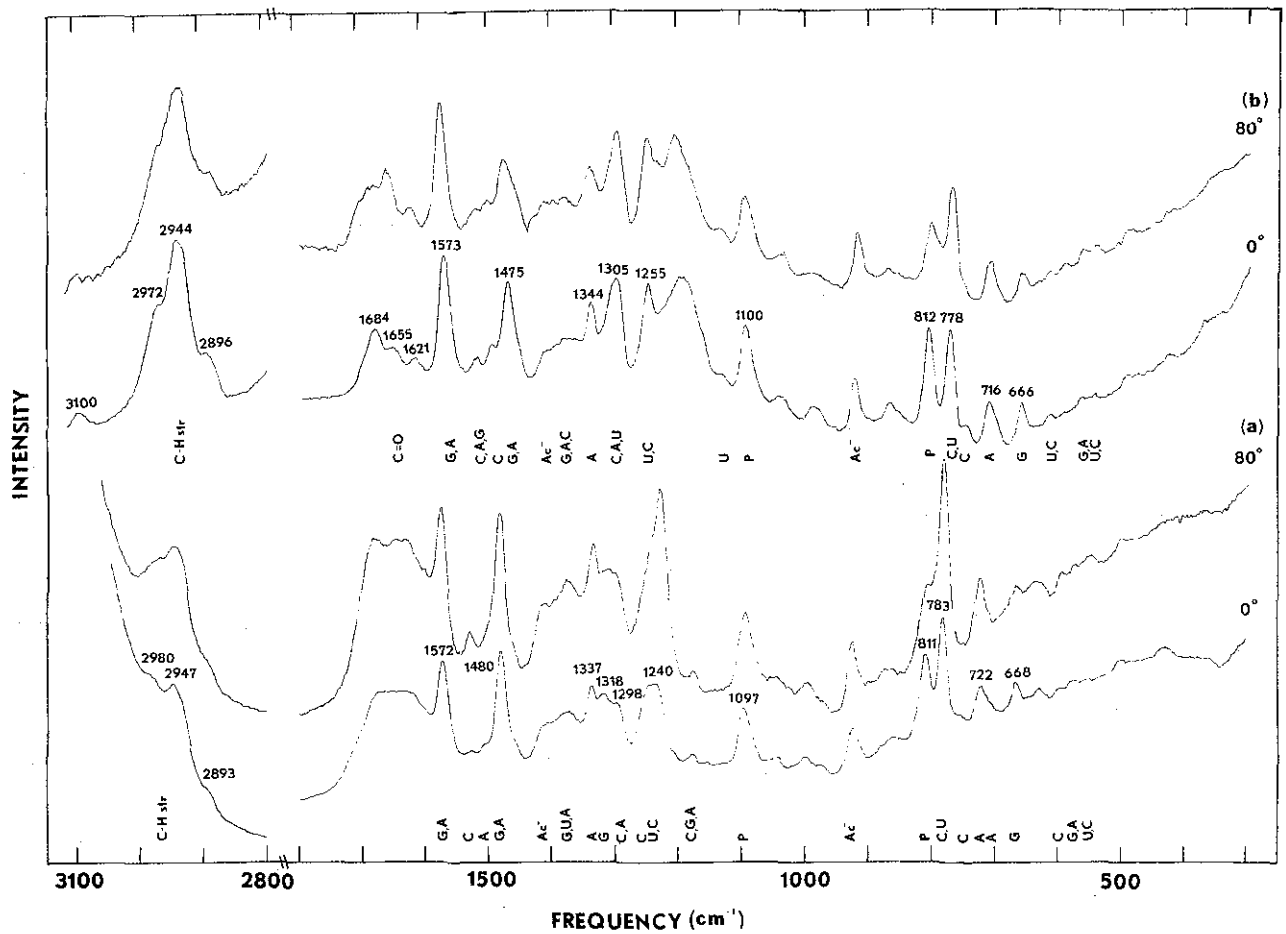


Table 2

Conformation-Sensitive Raman Lines of Aqueous RNA[a]

$\sigma(cm^{-1})$[b]	Residue	Assignment	dI/dT	$d\sigma/dT$	Conformational Significance
670(665)	G	ring str	−	0	base stacking of G
710	R(?)	ring str	−	+	sugar ring pucker
725(715)	A	ring str	+	0	base stacking of A
785(775)	C,U	ring str	+	0	base stacking of C and U
795(795)	P	C-O-P-O-C	+	0	disordered configuration
814(812)	P	C-O-P-O-C str	−	0	gauche^t-gauche^t configuration
1100	P	O—P—O^- str	0	0	useful as internal standard
1238	U,C	ring str, C-N str	+	0	base pairing and stacking of U and C
1301(1305)	A,C,U	ring str	+	0	base pairing and stacking of A
1320	G	ring str	−	+	base stacking of G
1340(1345)	A	ring str	+	0	base stacking of A
1378	G	ring str	−	−	base pairing and stacking of G
1485(1475)	A	ring str	+	0	base stacking of A
1485(1475)	G	ring str	−	0	base stacking of G
1531(1522)	C	ring str	+	+	base pairing and stacking of C
1575(1575)	G,A	ring str	+	+	base pairing and stacking of G and A
(1657)	U,G,C	C=O str	+	0	base pairing of U,G,C
(1684)	U,G,C	C=O str	−	0	base pairing of U,G,C

[a] Abbreviations: A adenine, G guanine, C cytosine, U uracil, R ribose, P phosphate, str stretching, def deformation, dI/dT intensity change with temperature, $d\sigma/dT$ frequency change with temperature. The latter two quantities are denoted as positive, negative, or zero.

[b] Approximate frequencies for H_2O solutions. Values in parentheses refer to approximate frequencies for D_2O solutions. The 1485 cm^{-1} lines of A and G also shift by 20 cm^{-1} to lower frequency upon 8-CH deuteration. (See text and [5] and [13].)

When D_2O is the solvent for RNA or DNA, a further complication arises in the Raman spectrum. This is a consequence of isotopic hydrogen exchange between protons of the nucleotides and deuterons of the solvent. Fig 2 shows that each nucleotide contains several rapidly exchangeable protons, specifically those of NH, NH_2, and OH groups. Since these are exchanged nearly instantaneously by solvent deuterium, they pose no problem to the recording of spectra that are identically reproducible, scan after scan, and completely reversible with respect to temperature changes. The CH proton at ring position 8 of A or G, however, is a slowly exchanging or "pseudoacidic" proton. Consequently, in a nucleic acid the exchange of 8-CH groups (of A and G) requires a considerable "incubation" time lest the spectral changes resulting therefrom be confused with those resulting from alteration of the RNA or DNA secondary structure. We have made a detailed study of the 8-CH exchange reaction in several purine nucleotides in order to distinguish those spectral changes due to denaturation of the nucleic acid secondary structure from those due to 8-CH exchange [19].

The effect of the aforementioned exchange reaction on the Raman spectrum of a nucleotide is by no means insignificant, as can be seen in Fig 5 and Fig 6, which compare Raman spectra of D_2O solutions of non-exchanged (8-CH) and exchanged (8-CD) modifications of AMP as well as of other adenine-containing compounds. The half-life for 8-CH exchange is of the order of hours at 80°C and 10^3 hours at 30°C (Table 3). Hence it is apparent that appreciable exchange (and therefore spectral irreversibility) could take place in the time required to record Raman spectra of D_2O solutions of RNA, DNA, or viruses at 80°C, though not at 30°C.

The pseudo-first-order kinetics and Arrhenius parameters governing 8-CH exchange are conveniently and accurately measured by quantitative Raman spectrophotometry (Table 3). Moreover, the results indicate rather subtle structural differences among different mono- and poly-nucleotides. Further discussion of these findings is given elsewhere [19].

The conformational dependence of Raman scattering from double-stranded DNA has also been studied [20-22]. In aqueous double-stranded DNA all bases are paired (A with T and G with C) as well as stacked, to form the familiar double helix. The secondary structure of aqueous DNA thus differs considerably from that of the aqueous single-stranded RNA discussed above. Moreover, deoxyribose sugars of the DNA backbone exhibit 3'C-exo ring puckering (B-helix geometry) while ribose sugars of RNA exhibit 3'C-endo puckering (A-helix geometry) [23]. As shown by the work of Peticolas and collaborators [22], as well as by work conducted in our laboratory [17], the Raman spectrum offers a convenient means of distinguishing these different nucleic acid geometries (A and B helices) from one another.

Fig 7, for example, shows spectra of eukaryotic DNA in double-stranded (low T) and denatured (high T) states. Thermal denaturation, which converts the double-stranded helix to two separate strands, each devoid of appreciable secondary structure, is seen to produce in the Raman spectrum many of the frequency shifts and intensity changes noted above for RNA denaturation. In addition, the replacement of U by T gives the DNA spectrum a profile somewhat different from that of RNA. Most important for the subsequent discussion, however, is the fact that RNA (Fig 4) and related polyribonucleotides (such as poly(rA) in Fig 5) display a strong Raman line near 812 ±2 cm^{-1} associated with the symmetric stretching vibration of -O-P-O- diester linkages in the RNA backbone. This line is consistently found in model compounds of the A-helix class [17]. Conversely, no line appears at or near this frequency in spectra of DNA (Fig 7) so long as DNA does not

Table 3

Arrhenius Parameters for 8-CH → 8-CD Exchange in Adenine Derivatives[a]

Sample	Temperature Range (°C)	Activation Energy (kcal/mol)	Frequency Factor (hr^{-1})
5'rAMP	20-80	24.2 ±0.6	2.3×10^{14}
cAMP	<50	17.7 ±0.1	9.6×10^{9}
	>50	23.5 ±0.6	8.3×10^{13}
poly(rA)	<60	27.7 ±0.4	1.8×10^{16}
	>60	22.0 ±0.1	3.2×10^{12}

[a] Data from [19].

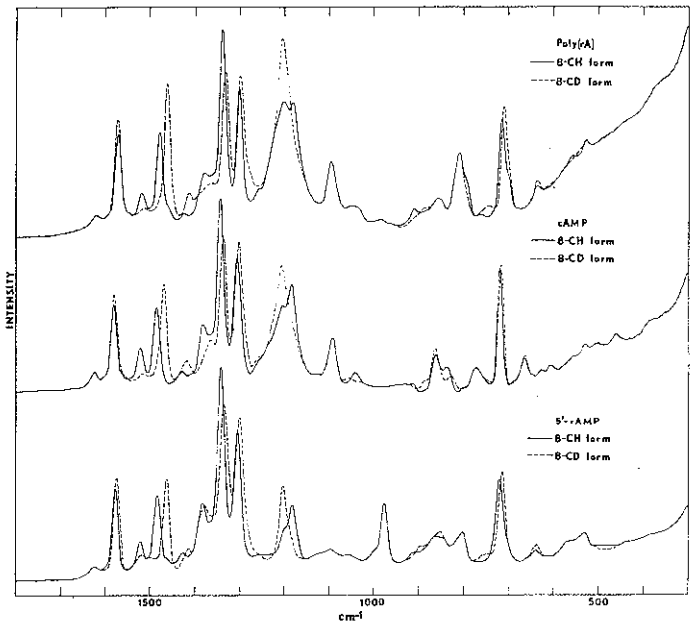


Fig 5 Raman spectra in the region 300-1800 cm^{-1} of D_2O solutions of adenosine-5′-monophosphate (5′-rAMP, bottom), adenosine-3′:5′-monophosphate (cAMP, middle), and polyriboadenylic acid (Poly (rA), top). Spectra of 8-CH and 8-CD forms are shown by the solid and broken curves, respectively. The clearcut shift of the 1485 cm^{-1} line to 1465 cm^{-1} is apparent in each case and provides the most convenient basis for determining the rate of exchange of 8-CH groups by deuterium. The phosphate group frequencies at 980 (bottom) and 1100 cm^{-1} (middle and top) are unaffected by deuteration and serve as marker lines for quantitative calculations. (From [19].)

assume the A-helix geometry. Instead, DNA and other polydeoxyribonucleotides of the B-helix class contain -O-P-O- frequencies outside the 790-820 cm^{-1} interval. As seen in Fig 7, B-DNA gives a line near 830 cm^{-1} which may be associated with its phosphate ester bond stretching motions.

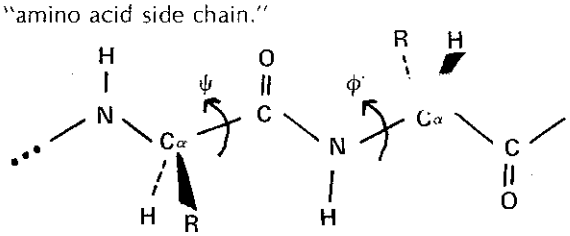
Finally, the intensity of the 812 cm^{-1} line in spectra of native RNAs appears to be directly proportional to the number of RNA nucleotides which exist in regions of ordered A structure and therefore to the number which participate in secondary interactions [17, 18]. Denatured nucleic acids generally exhibit a Raman line near 795-798 cm^{-1} which is also assignable to their -O-P-O- groups (Table 2).

In summary, the position and intensity of the -O-P-O- group frequency provide definite information on the kind and amount of secondary structure in aqueous nucleic acids. The sensitivity of this group frequency to secondary structure may very well originate from coupling of ester -O-P- bond stretching and furanosyl -C-C- and -C-O- bond stretching motions, as has been proposed [15, 24, 25].

Raman Spectroscopy of Proteins

Contributions to the understanding of Raman spectra of proteins have been made by several groups [26-30], and these have been reviewed recently [5, 31].

A correlation between Raman spectra and conformational properties of proteins which is often applicable to viruses depends on Raman group frequencies of amides to ascertain the protein chain conformation. In a typical protein chain the peptide repeating unit (-CO-NH-C_αHR-)$_n$ assumes a configuration defined by the dihedral angles ϕ and ψ, shown below, where R is an aliphatic or aromatic substituent characteristic of the amino acid and often referred to as the "amino acid side chain."

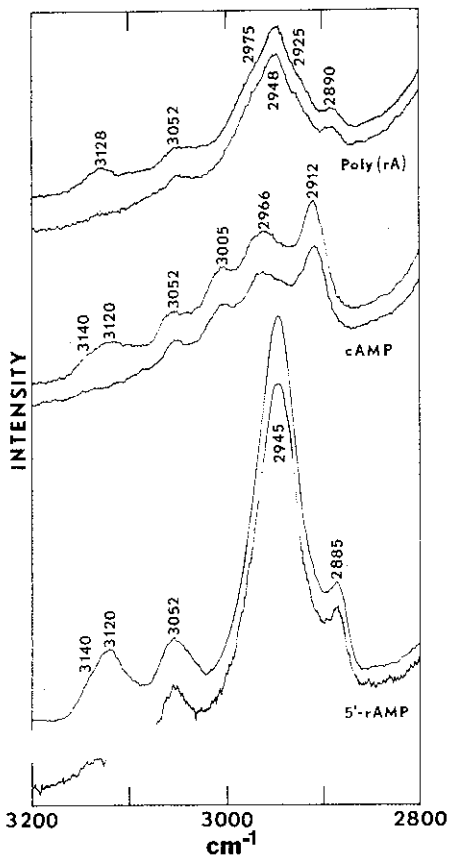


SUBSTITUENTS (R) FOR SELECTED AMINO ACIDS

HO — ⟨O⟩ — CH_2–	: TYROSINE
HS – CH_2 –	: CYSTEINE
HOOC – CH_2–	: ASPARTIC ACID
HOOC – CH_2 –CH_2–	: GLUTAMIC ACID
$H_3\overset{+}{N}$–CH_2–CH_2–CH_2–CH_2–	: LYSINE

Energetically favorable values for the pair (ϕ, ψ) differ according to the identity of the side-chain substituent R at each C_α atom, the solution temperature, ionic strength, pH, and other such factors. Nevertheless there is a relatively small number of preferred ranges within which ϕ and ψ appear to vary rather little, provided the substituents R are defined [32]. The two most important of these conformations for the viral proteins of present interest are the so-called alpha-helix (α) and antiparallel-β-pleated sheet (β) conformations, where, respectively, (ϕ, ψ) is (132°, 123°) and (40°, 315°). In addition, many proteins apparently contain chain segments of irregular structure in which successive peptidyl units or groups of peptidyl units may not exhibit the same configurational geometry. A similar situation can exist for disordered or randomly oriented chain segments in which there is less restriction to rotation about C_α-N and/or C_α-C bonds. These considerations have led to the characterization of protein chain conformation according to three main types, namely α, β, and random-chain structures. (A more comprehensive treatment is given by Dickerson and Geis [33].)

Fig 6 Raman spectra in the region 2800-3200 cm^{-1} of D_2O solutions of adenine derivatives. For each pair of spectra the upper curve corresponds to the 8-CH form and the lower curve to the 8-CD form of the adenine ring in the compound shown. The 8-CH stretching frequency is thus seen to occur at 3120 cm^{-1} in 5′ rAMP and cAMP and at 3128 cm^{-1} in Poly(rA). Note also that the other CH stretching frequencies differ greatly from one compound to another. This is a result of the structurally different sugar groups found in 5′-mono-, 3′:5′-cyclic-, and 3′:5′-poly-nucleotides. (From [19].)



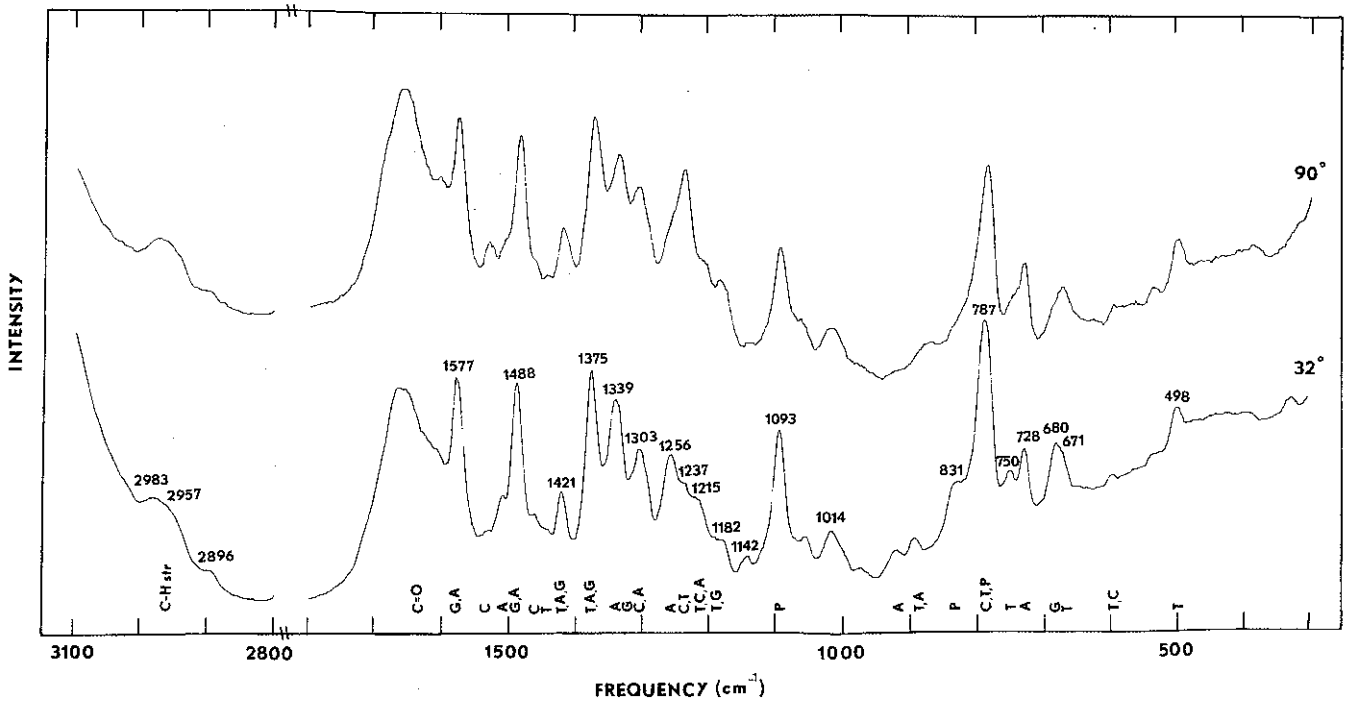


Fig 7 Raman spectra in the regions 300-1800 cm^{-1} (A=1) and 2800-3100 cm^{-1} (A=1/3) of H_2O solutions of a double-stranded helical DNA at 32°C and after denaturation at 90°C. The DNA concentration is 40 $\mu g/\mu l$, and other notation is as described in Fig 4. A major difference between Raman spectra of DNA and RNA is the Raman line of the sugar-phosphate backbone at 835 cm^{-1} in DNA (B-helix geometry) and 812 cm^{-1} in RNA (A-helix geometry). The bases thymine (T) in DNA and uracil (U) in RNA account for most of the other spectral differences between the two nucleic acids. (This DNA was extracted from chicken erythrocyte cell nuclei and provided by Dr. D.E. Olins of the Oak Ridge National Laboratory.)

Studies of model compounds (such as polypeptides and proteins of known structure) show that with few exceptions the Raman-active frequencies of the peptidyl group (called amide I, ~1650-1675 cm^{-1}, and amide III, ~1225-1300 cm^{-1}) fall into the spectral ranges shown in Table 4.

The term "random-chain" merits some further consideration. The model compounds on which the random-chain assignments are based include poly-L-lysine (PLL), for which other evidence suggests an "extended-chain" structure [34]. Moreover, it is questionable whether randomly oriented chain segments can occur to an appreciable extent in compact globular proteins such as the MS2 coat protein. It would appear that the amide I and amide III Raman line positions and intensities do not provide a clear-cut distinction between extended, random, and perhaps other irregular structures as well. In Table 4, and elsewhere, we have cited "random" for consistency with previous usage. A more appropriate label, however, might be "irregular" conformation, meaning one which differs from both α and β conformations and which encompasses a wider range of torsional angles (ϕ, ψ), including the extended-chain conformation of PLL.

Another significant correlation recently established by Siamwiza and others [28] shows that a pair of Raman lines near 830 and 850 cm^{-1}, the so-called "tyrosine doublet," is highly sensitive to interactions of the tyrosine *para*-hydroxyl group. The value of this correlation for diagnosing interactions in viruses can be described as follows. When the higher frequency component (~850 cm^{-1}) of the doublet is more intense than the lower frequency component (~830 cm^{-1}) by a factor of 2.5 or greater, that is, $I_{850}/I_{830}>2.5$, then the *para*-hydroxyl group of tyrosine is the acceptor in a strong hydrogen bond with a highly positive donor (like $-NH_3^+$). When $I_{850}/I_{830}\approx 0.3$, the *p*-OH group is the donor in a strong hydrogen bond with a negative acceptor group (like $-COO^-$). When the *p*-OH group is both donor and acceptor in moderate hydrogen bonding with, for example, H_2O molecules, then $I_{850}/I_{830}\approx 1.2$. The last situation can obviously occur also when there are comparable numbers of tyrosines in both of the preceding states of hydrogen bonding. This correlation is summarized in Table 5A.

Many proteins also contain disulfide (-S-S-) cross links between cysteine residues in separate regions of the chain. Cross-linked cysteine residues, called cystine, are easily distinguished by their

Table 4

Conformation-Sensitive Raman Lines of Peptidyl Groups of Aqueous Proteins[a]

$\sigma(cm^{-1})$	Assignment	Relative Intensity	Conformational Significance
1665-1672	amide I	strong	β-sheet
1660-1670	amide I	strong, broad	random chain
1645-1655	amide I	strong	α-helix
1270-1300	amide III	weak	α-helix
1243-1253	amide III	medium, broad	random chain
1229-1235	amide III	strong	β-sheet

[a] From data of [27] and [31]. Further treatment has been given very recently by Lippert and others, J. Am. Chem. Soc. **98,** 7075 (1976).

Table 5
Conformation-Sensitive Raman Lines of Some Amino Acid Side Chains of Aqueous Proteins

A. The Tyrosine Doublet (~850 and 830 cm^{-1})[a]

Intensity Ratio I_{850}: I_{830}	State of Phenolic Hydroxl Group of Tyrosine
10:4	acceptor of strong H-bonds
10:8	donor and acceptor of moderate H-bonds
3:10	donor of strong H-bonds
7:10	ionized (requiring very high pH)

B. The Disulfide Stretching Frequency (~500-545 cm^{-1})[b]

σ_{SS} (cm^{-1})	Configuration of C_α-C-S-S-C-C_α Network of Cystine
508-513	gauche-gauche-gauche
523-529	trans-gauche-guache or gauche-gauche-trans
542-545	trans-gauche-trans

[a] From data of [28].

[b] From data of [29]. See also [35], however.

-S-S- stretching frequency which is intense in the Raman effect and which occurs in a region relatively free of other group frequencies (500-545 cm^{-1} region) [26]. On the other hand, cysteine is distinguished by its -SH stretching frequency (2500-2600 cm^{-1} region).

When disulfide linkages are present the correlation proposed by Sugeta and others [29] (Table 5B) may be applicable. According to these authors the position of the -S-S- stretching frequency depends upon whether the C_α-C-S-S-C-C_α network is in the all gauche (ggg), trans-gauche-gauche (tgg), or trans-gauche-trans (tgt) configuration. The generality of this correlation has been questioned recently [35], although the value of the -S-S- stretching frequency for qualitative identifications is undisputed.

Other correlations, summarized elsewhere [5, 31], have been proposed for characteristic group frequencies of amino acid side groups of tryptophan, aspartic acid, and glutamic acid.

Experimental Methods

Instrumentation for Raman spectroscopy of biological molecules has been described in detail in various places [5, 13, 31]. In our laboratory we have employed a Spex Ramalog system with sample illuminator designed to accommodate 1.0 mm id capillary cells (Kimax #34507) in the transverse excitation geometry. A uniform and constant temperature in the sample-filled (~10μl) cell is accomplished with a specially designed thermostatable jacket [36]. In studies of macromolecular secondary structure or reaction kinetics it is usually desirable to maintain the temperature of the solution constant to within ±0.5°C for several hours or more.

Since Raman spectra of viral components are usually obtained with 488.0 or 514.5 nm excitation from an argon-ion laser, impurities which might absorb radiation of these wavelengths must be eliminated. The procedures required for such sample purifications are often tedious [5] and will not be discussed here. Care must be exercised, however, to insure that native virions are neither disassembled nor degraded in the course of sample handling procedures. Single-stranded viral nucleic acids are also especially sensitive to chain scissions, and thus their solutions demand special handling to avoid contamination by nucleases.

Raman Spectroscopy of Nucleoproteins

Apart from the requirements of sample purity and optical homogeneity, the success of Raman spectroscopy in nucleoprotein research depends upon two other factors. First, the proportion of nucleic acid to protein in the complex should be such as to allow Raman lines of both components to appear in the spectrum with acceptable signal-to-noise ratios. In practice a molar proportion of roughly 1:2 (nucleotide:peptide unit) is most favorable. The excess of protein in such circumstances compensates for the intrinsically weaker Raman scattering associated with vibrations of amino acid residues as compared with those of nucleotide residues. Fortunately, this condition is met in many viruses of structural interest, including RNA bacteriophages and plant viruses. Second, the Raman lines of nucleic acid and protein components should not overlap appreciably with one another, thus permitting their frequencies and relative intensities to be measured separately and with reasonable accuracy. Of course, it is impossible to achieve a totally interference-free spectrum, but many of the structurally significant Raman lines of RNA and DNA in fact occur in spectral ranges where little or no interference is encountered from protein, and *vice versa*. This will be evident in the discussion which follows. When this condition is not met for a native virus, chemical modification of the virion may be effective in circumventing the problem.

Model Systems

The complexity of viral nucleic acid and protein molecules assures that many different kinds of interactions will be involved in stabilizing the virion structure, including hydrogen bonding, hydrophobic bonding, electrostatic interactions, and so forth. It is helpful therefore to select for Raman study model nucleoproteins in which one or another type of interaction predominates. Here we shall discuss but one such example, the complex formed of double-stranded DNA and poly-L-lysine (PLL) which is believed to be stabilized by electrostatic interaction between the negatively charged $>PO_2^-$ groups of DNA and the positively charged -NH_3^+ groups of lysine side chains [37].

DNA-PLL is of interest for other reasons as well. Considerable controversy has centered around the backbone geometry in PLL-bound DNA, and resolution of this question is of importance because of the compositional similarity between DNA-PLL and DNA-histone complexes of cellular chromatin. Further, circular dichroism measurements have yielded conflicting results regarding the possibility of polymorphism in DNA-PLL depending upon the method of complex formation [38].

Fig 8 shows the Raman spectrum of the complex formed by directly mixing DNA and PLL in high-ionic-strength buffer. The observed spectrum (broken line) is also compared with the sum of spectra of DNA and PLL (solid line), at the same conditions. It is apparent from Fig 8 that the 830 cm^{-1} line of B-DNA is virtually unaltered by complex formation with PLL. The same is true of the 1100 cm^{-1} line ($>PO_2^-$ group vibration). Moreover, no new line appears near 812 cm^{-1} that would suggest the conversion of B-DNA to A-DNA. Likewise, the amide III line of PLL (1240 cm^{-1}) is unchanged as a result of complex formation. On the other hand, the intense Raman lines between 1300 and 1600 cm^{-1}, due to A, G, C, and T ring modes, are greatly increased in intensity by PLL binding. Smaller intensity increases occur also in other lines assigned to DNA base vibrations. These intensity increases represent the elimination of Raman hypochromism.

We have obtained qualitatively similar results for DNA-PLL prepared by the salt-gradient-dialysis method as well as for DNA-PLL prepared in low-ionic-strength media [39]. The important conclusions from these results are:

(1) The structure of the DNA-PLL complex does not depend qualitatively on either the method of complex formation (direct

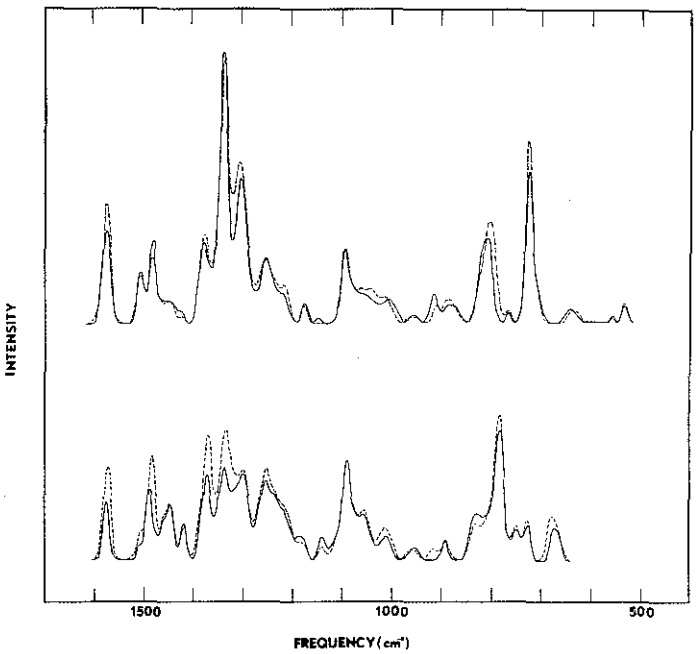


Fig 8 Lower: Raman spectra at 32°C of DNA-PLL (broken curve) and the sum of spectra of DNA and PLL (solid curve) at the same conditions (1.0 M NaCl, pH 7). Upper: Raman spectra at 32°C of poly(rA)-PLL (broken curve) and the sum of spectra of poly(rA) and PLL (solid curve) at the same conditions (0.01 M NaCl, pH 7). (From [39].) Binding of the polypeptide PLL to either DNA or poly(rA) causes certain Raman lines of the nucleotide bases to increase in intensity, suggesting weaker base-base interactions in the complexes. (From [39].)

mixing vs. annealing) or the ionic strength of solution, over the range examined (0.025 to 1.0 M NaCl).

(2) The B-DNA structure is conserved in DNA-PLL.

(3) The extended-chain structure of PLL is conserved in DNA-PLL.

(4) Base stacking interactions in DNA are reduced by the binding of PLL, that is, the double-stranded DNA is converted to a more "open" structure by PLL-binding.

(5) Electrostatic interactions between PO_2^- and NH_3^+, if present in DNA-PLL, do not perturb the frequency or Raman intensity of the 1100 cm^{-1} line (PO_2^- symmetric stretching mode). Such interactions are therefore not as strong as those between PO_2^- and Mg^{2+}, which do alter the 1100 cm^{-1} line [40].

This situation contrasts rather strikingly with the poly(rA)-PLL complex [39], also shown in Fig 8. Here the geometry of the polynucleotide backbone (A-helix) is significantly perturbed by PLL binding, as evidenced by the shift of the -O-P-O- frequency from 812 to 804 cm^{-1}. An even larger shift of the 812 cm^{-1} line is produced by complexing poly-L-arginine to poly(rA). In poly(rA)-PLA the -O-P-O- frequency occurs below 800 cm^{-1} [41].

One interpretation of these results is that binding of lysine-rich and arginine-rich proteins to a nucleic acid substrate occurs without appreciable change in the geometry of the nucleic acid backbone, provided the backbone conformation is of the B-helix type. In the case of nucleic acids of the A-helix type, such binding is accompanied by a significant change in the backbone geometry, and this change is in the direction of a disordered ribonucleotide chain. This hypothesis will be further tested by studies of additional model systems now in progress.

Finally, we may note that in every case so far examined, the $>PO_2^-$ group frequency is unaffected by electrostatic interaction between $>PO_2^-$ and $-NH_3^+$. Therefore the Raman line in question (1100 cm^{-1}) is of little value as a marker of such interaction.

RNA Viruses: MS2 and TYMV

The Raman spectra of MS2 RNA (free of protein) at limits of high and low temperature were shown in Fig 2. From these and related spectra we can plot the intensities of conformationally sensitive Raman lines as a function of the temperature. Several such plots, called RNA melting profiles, are shown in Fig 9. Two types of temperature dependence are revealed.

On the one hand (Fig 9, left), Raman lines originating from ring vibrations of the bases show little change with temperature up to 50°C then exhibit large intensity changes with further increase of temperature. Model compound studies [5, 13] indicate that these effects are due to unstacking of RNA bases. Therefore unstacking is essentially a cooperative phenomenon occuring above 50°C. Note that the intensity changes (dI/dT) are positive for lines of A, C, and U (Raman hypochromism), but negative for lines of G (Raman hyperchromism). The unstacking of pyrimidines also occurs over a narrower temperature interval than that of purines.

On the other hand, Raman lines resulting from carbonyl and phosphate group vibrations change gradually in intensity over the entire temperature range (Fig 9, right). Therefore the rupture of hydrogen bonding between paired bases in hairpin-helical segments of MS2 RNA and the disordering of the RNA backbone are non-cooperative structural transitions. At 30°C 85% of the nucleotide residues of MS2 RNA exist in regions of ordered A-type helical structure. This percentage of ordered structure falls to less than 30% at 80°C.

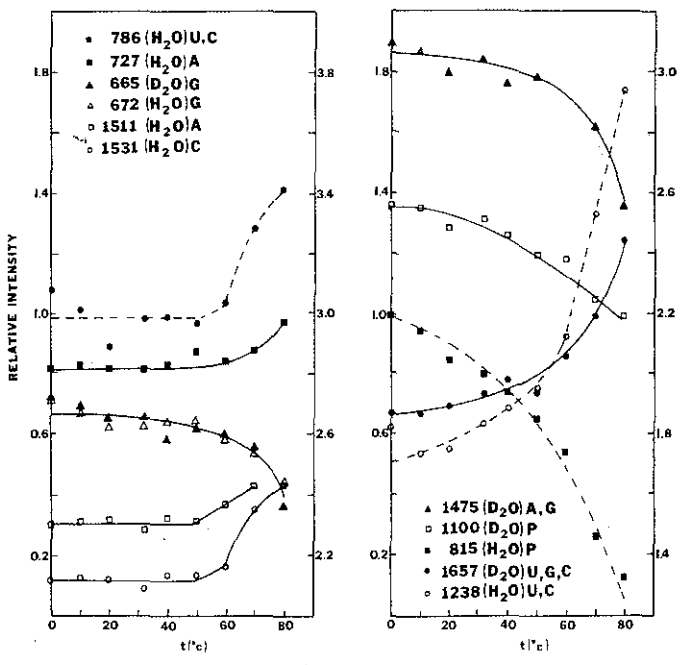
Raman spectra of MS2 capsids (free of RNA) may be analyzed in a similar way [42]. Our results show that:

(1) α, β, and random conformers comprise 0, 60, and 40%, respectively, of the coat protein chains.

(2) The coat protein secondary structure is stable to 55°C, above which temperature the capsids collapse.

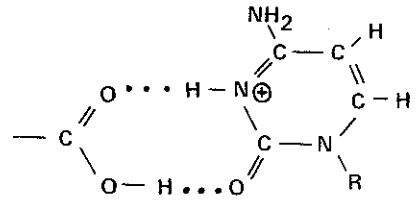
(3) Four tyrosine residues of each coat protein molecule are hydrogen bond donors to negative acceptors (possibly $-COO^-$ groups of aspartic and glutamic acids) of the protein. No tyrosyl groups are accessible for hydrogen bonding with solvent H_2O.

Fig 9 Plots of the peak intensities of selected Raman lines of MS2 RNA vs. temperature. The ordinates (left axis for solid curves and right axis for broken curves) show intensities in arbitrary units after normalization to an internal standard. The group of lines in the left figure display cooperative melting behavior, while those on the right exhibit non-cooperative melting. The keys give the Raman frequencies in cm^{-1}, the solvent employed, and the assignments. See also Table 2. (From [42].)



(4) There are no disulfide bridges either within or between coat protein molecules of the capsid.
(5) Two cysteine residues per coat protein molecule are sufficiently exposed to permit rapid exchange of their sulfhydryl groups with deuterium in D_2O solution.

Comparison of the Raman spectrum of MS2 phage with spectra of the RNA and protein components confirms that the above structural features are maintained also in the native virion [42]. Moreover, neither RNA nor protein of MS2 is altered in structure by increasing the temperature. This result is rather remarkable, since MS2 RNA itself undergoes a substantial change in secondary structure below 50°C, as indicated in the left part of Fig 9. Apparently RNA within the capsid is stabilized by the protein coat, which implies specific RNA-protein interactions. Identification of the interacting groups remains an important future goal of this work. Already the Raman spectra have indicated, however, that interactions between protonated cytosines and carboxyl groups, as postulated by Kaper [43], cannot be a source of stabilization of the MS2 virion. The possibility that protonated cytosines exist in MS2 RNA is eliminated by the Raman data [44].



Raman spectroscopy has also permitted the same conclusion to be made for TYMV [45], which, as Table 1 shows, contains an RNA molecule especially rich in cytosine residues.

DNA Viruses: Pf1 and fd

Raman spectra of Pf1 and fd virions are shown in Fig 10. These virions contain 88 to 95% by weight protein, and thus in each case the Raman scattering from the DNA component is weak by comparison with the intense Raman lines of coat protein. Most of the observed lines are assigned to protein or DNA subgroups as listed along the abscissa of Fig 10.

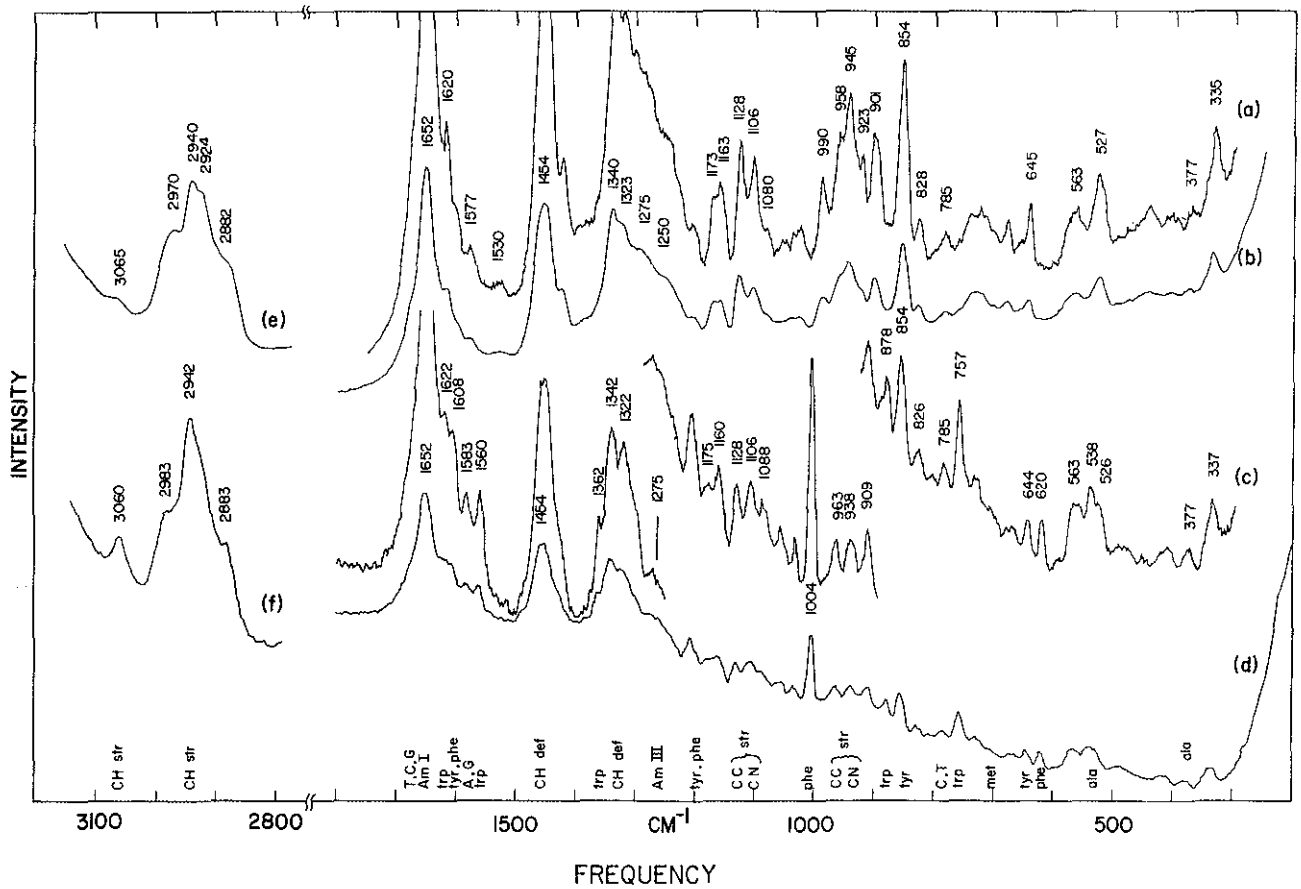
Despite morphological similarities the two FB phages differ greatly in the amino acid compositions of their respective coat proteins, as is evident in the Raman spectra. Note, for example, the presence of phenylalanine (1004 cm^{-1} line) in fd but not in Pf1. (Virologists may do well to take note of the potential here for Raman spectroscopy as a technique for identifying different members of a virus class or mutants of the same viral species.)

Differences in amino acid composition notwithstanding, the amide I (1652 cm^{-1}) and amide III (~1275 cm^{-1} and weak) Raman lines indicate in a clear-cut fashion that the coat protein sheath of the DNA genome consists of uniformly α-helical chains in both Pf1 and fd. Neither β nor irregular structures are present to any appreciable extent.

A conclusion which could not be reached at the time of the original publication of these results [46], but which has been facilitated by the correlation of Siamwiza and others [28], is now evident. The

Fig 10 Raman spectra of FB viruses in 0.05 M NaCl at 32°C and pH 9. (Curve a) Pf1: concentration C=108 μg/μl; excitation wavelength (λ)=488.0 nm; radiant power (P)=100 mW; slit width ($\Delta\sigma$)=5 cm^{-1}; scan rate (r)=50 cm^{-1}/min; rise time (t)=1 sec; amplification (A)=3x. (Curve b) Pf1: $\Delta\sigma$=10 cm^{-1}; t=3 sec; A=1x; other conditions as in (a). (Curve c) fd: C=147 μg/μl; λ=514.5 nm; P= 100 mW; $\Delta\sigma$=5 cm^{-1}; r=50 cm^{-1}/min; t=1 sec; A=3x. (Curve d) fd:$\Delta\sigma$=9 cm^{-1}; t=3 sec; A=1x; other conditions as in (c). (Curve e) Pf1: A=x/3; other conditions as in (b). (Curve f) fd: A=x/3; other conditions as in (d).

Frequencies of prominent lines are in cm^{-1}, and assignments to molecular subgroups are denoted by standard abbreviations. Abbreviations: str, stretching; def, deformation; CH, carbon-hydrogen bond; CC, carbon-carbon bond; CN, carbon-nitrogen bond; A, T, C, and G, adenine, thymine, cytosine, and guanine; ala, alanine; met, methionine; phe, phenylalanine; trp, tryptophan; tyr, tyrosine; Am, amide. This virus contains a small percentage of nucleic acid, and therefore its Raman spectrum is dominated by vibrational frequencies of the coat protein molecules. (From [46].)



tyrosyl OH groups of Pf1 and fd are strongly hydrogen bonded as acceptors to positive donor groups. For both Pf1 and fd the intensity of the higher frequency component (854 cm^{-1}) of the tyrosine doublet greatly exceeds the intensity of the lower frequency component (828 cm^{-1}). The implication of this result is interesting. It suggests that acceptable models of Pf1 and fd structure must allow for close approach of tyrosyl OH groups in one α-helical chain with $-NH_3^+$ (lysine) or NH_2^+ (arginine) groups in the same or adjacent chains, since the latter are the only positive donors available for the indicated hydrogen bonding interaction. This criterion appears to be at least partially fulfilled in a recently proposed model [47].

As mentioned earlier, the Raman scattering of Pf1-DNA or fd-DNA is weak in comparison to that of the respective viral protein. Nevertheless, the data of Fig 10 are informative with respect to the important question of the viral DNA structure. The DNA purines (A and G) and pyrimidines (C and T) are evident by their corresponding ring vibrational frequencies at 1577 and 785 cm^{-1} in Pf1 (Fig 10a). In fd (Fig 10c) these frequencies occur at 1583 and 785 cm^{-1}.

We recall that if the single-stranded DNA molecule assumes a backbone conformation like that of single-stranded RNA (A-helix), then a phosphodiester group frequency should occur near 812 cm^{-1} with Raman intensity comparable to that of the 785 cm^{-1} line of C+T residues [17, 20]. In fact, in Pf1 no Raman line appears between 790 and 825 cm^{-1}, while in fd the line at 801 cm^{-1} is too low in frequency and too weak in intensity to be ascribed to an A-helix mode. Therefore the DNA backbones in Pf1 and fd cannot be in the A-helix conformation. DNA backbones of the B-helix type are, on the other hand, consistent with the Raman data.

FB viruses and related DNA-protein complexes have been under further study in our laboratory, and the results of these studies will be reported shortly. It is clear, however, in every case so far examined, that the DNA molecule which acts as a substrate for protein or polypeptide binding is one which retains the backbone of the native B-DNA conformation. This generalization seems to apply to both double-stranded and single-stranded DNA molecules.

Future Prospects

The application of Raman spectroscopy to biological molecules began in the 1930's with the work of Edsall and collaborators [48], and prospects for Raman spectroscopy in biophysical research were heralded as early as 1940 [1]. We may be, however, just at the beginning of a new realization of the value of Raman spectroscopy for the biosciences. It is certainly true that much more will be learned in the next few years about the structure and assembly of nucleoproteins through laser-Raman spectroscopy in combination with established biochemical procedures.

A new generation of lasers and spectrometers is also now dawning [49]. As these instruments become more generally available they will permit the design of experiments to reveal additional properties of viruses and related biological materials. One area of obvious potential of uv lasers would be to enhance through resonance the Raman scattering from vibrations of nucleic acid sub-units when the native nucleoprotein contains a deficiency of nucleic acid. Pf1, for example, is a virus that contains few aromatic amino acids, and therefore its coat protein would contribute little to a uv-laser-excited spectrum. On the other hand, uv excitation would promote to prominence many of the DNA Raman lines now hidden in the spectrum of Fig 10a.

The methods and approaches discussed here would also find extensive application to other complex biological materials containing nucleic acid and protein molecules.

Acknowledgments

Much of the work described here was supported by grants from the National Science Foundation (GB 41382), the National Institute of Allergy and Infectious Diseases (AI 11855), and the Research Corporation (Brown-Hazen Grant).

REFERENCES

1. J.R. Loofbourow, Rev. Mod. Phys. **12,** 267 (1940)
2. D. Garfinkel, J.T. Edsall, J. Am. Chem. Soc. **80,** 3807 (1958)
3. F.H. Spedding, R.F. Stamm, J. Chem. Phys. **10,** 176 (1942)
4. R.C. Lord, G.J. Thomas, Jr., Dev. Appl. Spectrosc. **6,** 179 (1968)
5. G.J. Thomas, Jr., Y. Kyogoku, in Practical Spectoscopy Series, Vol I, Part C, Ed E.G. Brame, Jr., and J.G. Grasselli, Marcel Dekker, Inc., N.Y. (in press, due 1977)
6. A. Lewis, The Spex Speaker, Vol XXI, No 2 (1976)
7. G.J. Thomas, Jr., Appl. Spectrosc. **30,** 483 (1976)
8. C.A. Knight, Chemistry of Viruses, 2nd edition, Springer-Verlag, N.Y. (1975)
9. R.A. Crowther, L.A. Amos, J.T. Finch, J. Mol. Biol. **98,** 631 (1975)
10. S. Casjens, J. King, Ann. Rev. Biochem. **44,** 555 (1975)
11. S.A. Berkowitz, L.A. Day, J. Mol. Biol. **102,** 531 (1976)
12. R.C. Lord, G.J. Thomas, Jr., Spectrochim. Acta **23A,** 2551 (1967)
13. K.A. Hartman, R.C. Lord, G.J. Thomas, Jr., in Physico-Chemical Properties of Nucleic Acids, Vol 2, Ed J. Duchesne, Academic Press, N.Y. (1973)
14. T. Shimanouchi, M. Tsuboi, Y. Kyogoku, Adv. Chem. Phys. **7,** 435 (1964)
15. G.J. Thomas, Jr., Biochim. Biophys. Acta **213,** 417 (1970)
16. W. Fiers, R. Contreras, F. Duerinck, G. Haegmean, J. Merregaert, W. Min-Jou, A. Raeymakers, G. Volckaert, M. Ysebaert, J. van de Kerckhove, F. Nolf, M. van Montagu, Nature **256,** 273 (1975)
17. G.J. Thomas, Jr., K.A. Hartman, Biochim. Biophys. Acta **312,** 311 (1973)
18. L. Lafleur, J. Rice, G.J. Thomas, Jr., Biopolymers **11,** 2423 (1972)
19. J. Livramento, G.J. Thomas, Jr., J. Am. Chem. Soc. **96,** 6529 (1974); Biochemistry **14,** 5210 (1975)
20. S.C. Erfurth, E.J. Kiser, W.L. Peticolas, Proc. Nat. Acad. Sci. U.S.A. **69,** 938 (1972)
21. S.C. Erfurth, W.L. Peticolas, Biopolymers **14,** 247 (1975)
22. S.C. Erfurth, P.J. Bond, W.L. Peticolas, Biopolymers **14,** 1245 (1975)
23. S. Arnott, Prog. Biophys. Mol. Biol. **6,** 265 (1970)
24. Y. Nishimura, K. Morikawa, M. Tsuboi, Bull. Chem. Soc. Japan **47,** 1043 (1974)
25. E.B. Brown, W.L. Peticolas, Biopolymers **14,** 1250 (1975)
26. R.C. Lord, N.T. Yu, J. Mol. Biol. **50,** 509 (1970); **51,** 203 (1970)
27. M.C. Chen, R.C. Lord, J. Am. Chem. Soc. **96,** 4750 (1974)
28. M.N. Siamwiza, R.C. Lord, M.C. Chen, T. Takamatsu, I. Harada, H. Matsuura, T. Shimanouchi, Biochemistry **14,** 4870 (1975)
29. H. Sugeta, A. Go, T. Miyazawa, Chem. Lett. (Japan) **1972,** 83; Bull. Chem. Soc. Japan **46,** 3407 (1973)
30. P.C. Painter, J.L. Koenig, Biopolymers **15,** 241 (1976)
31. B.G. Frushour, J.L. Koenig, in Advances in Infrared and Raman Spectroscopy, Ed R.J.H. Clark and R.E. Hester, Heyden & Son, Ltd., London (1975)
32. P.Y. Chou, G.D. Fasman, Biochemistry **13,** 211, 222 (1974)
33. R.E. Dickerson, I. Geis, The Structure and Action of Proteins, W.A. Benjamin Co., Reading, Mass. (1970)
34. T.J. Yu, J.L. Lippert, W.L. Peticolas, Biopolymers **12,** 2161 (1973)
35. H.E. van Wart, H.A. Scheraga, J. Phys. Chem. (in press)
36. G.J. Thomas, Jr., J.R. Barylski, Appl. Spectrosc. **24,** 463 (1970)
37. P.H. von Hippel, J.D. McGhee, Ann. Rev. Biochem. **41,** 231 (1972)
38. D. Carroll, Biochemistry **11,** 421, 426 (1972)
39. B. Prescott, C.H. Chou, G.J. Thomas, Jr., J. Phys. Chem. **80,** 1164 (1976)
40. G.J. Thomas, Jr., M.C. Chen, K.A. Hartman, Biochim. Biophys. Acta **324,** 37 (1973)
41. C.H. Chou, G.J. Thomas, Jr. (manuscript in preparation)
42. G.J. Thomas, Jr., B. Prescott, P. McDonald-Ordzie, K.A. Hartman, J. Mol. Biol. **102,** 103 (1976)
43. J.M. Kaper, in RNA Viruses and Ribosomes, North Holland Press, Amsterdam (1972)
44. C.H. Chou, G.J. Thomas, Jr., Biopolymers (in press)
45. T.A. Turano, K.A. Hartman, G.J. Thomas, Jr., J. Phys. Chem. **80,** 1157 (1976)
46. G.J. Thomas, Jr., P. Murphy, Science **188,** 1205 (1975)
47. Y. Nakashima, R.L. Wiseman, W. Konigsberg, D.A. Marvin, Nature **253,** 5486 (1975)
48. J.T. Edsall, J. Chem. Phys. **4,** 1 (1936)
49. J.T. Yardley, Science **190,** 223 (1975)