Sugar Pucker and Phosphodiester Conformations in Viral Genomes of Filamentous Bacteriophages: fd, If1, IKe, Pf1, Xf, and Pf3†

George J. Thomas, Jr.,*,‡ Betty Prescott,‡ Stanley J. Opella,§ and Loren A. Day∥

Division of Cell Biology and Biophysics, School of Basic Life Sciences, University of Missouri—Kansas City, Kansas City, Missouri 64110-2499, Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104, and Department of Developmental and Structural Biology, The Public Health Research Institute of the City of New York, New York, New York 10015

Received October 29, 1987; Revised Manuscript Received February 2, 1988

ABSTRACT: The laser Raman spectra of filamentous viruses contain discrete bands which are assignable to molecular vibrations of the encapsidated, single-stranded DNA genomes and which are informative of their molecular conformations. Discrimination between Raman bands of the DNA and those of the coat proteins is facilitated by analysis of viruses containing deuterium-labeled amino acids. Specific DNA vibrational assignments are based upon previous studies of A-, B-, and Z-DNA oligonucleotide crystals of known structure [Thomas, G. J., Jr., & Wang, A. H.-J. (1988) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 2, Springer-Verlag, Berlin]. The present results show that canonical DNA structures are absent from six filamentous viruses: fd, If1, IKe, Pf1, Xf, and Pf3. The DNAs in three viruses of symmetry class I (fd, If1, IKe) contain very similar nucleoside sugar puckers and glycosyl torsions, deduced to be C3′-endo/anti. However, nucleoside conformations are not the same among the three class II viruses examined: Pf1 and Xf DNAs contain similar conformers, deduced to be C2′-endo/anti, whereas Pf3 DNA exhibits bands usually associated with C3′-endo/anti conformers. Conformation-sensitive Raman bands of the DNA 3′-C–O–P–O–C-5′ groups show that in all class I viruses and in Pf1 the ssDNA backbones do not contain regularly ordered phosphodiester group geometries, like those found in ordered single- and double-stranded nucleic acids. Xf and Pf3, on the other hand, appear to contain substantially more regularity in 3′-C–O–P–O–C-5′ geometry, by virtue of their Raman marker bands which are similar to those associated with A-type and B-type geometries, respectively. Nonetheless, the nucleoside sugar puckering usually associated with the canonical DNA structures is not present in either of these viruses.

Filamentous bacteriophages are long (800–2000 nm) and narrow (6–7 nm) filaments constructed from a circular single-stranded DNA molecule which encodes the viral genes, several thousand copies of a protein subunit of about 50 amino acids which serves as the major coat protein, and a few copies of one or more minor proteins located at the filament ends. Comprehensive reviews of the molecular biology and morphogenesis of filamentous viruses have been given (Denhardt et al., 1978; Webster & Lopez, 1985). Two classes of filament symmetry have been distinguished on the basis of fiber X-ray diffraction results (Marvin et al., 1974a,b; Makowski & Caspar, 1981). The class I particle, for which the prototype is Ff,[1] contains a helical sheath of protein subunits arranged with fivefold rotational symmetry and an approximate twofold screw axis. If1 and IKe are members of this class. The most extensively studied of the class II virions is Pf1, which contains no rotation axis of symmetry and comprises 27 subunits in a repeat of five turns of the helical filament. Xf and Pf3 are also of class II symmetry (Marvin et al., 1974a; Peterson et al., 1982).

A large body of structural work exists for viruses in both classes, including studies based upon fiber X-ray diffraction, NMR, Raman, IR, fluorescence, and CD spectroscopies. This work has been reviewed recently (Makowski, 1984; Opella et al., 1987; Thomas, 1987; Day et al., 1988). Much is known about the architecture of fd and Pf1 filaments and, by inference, about other members of their respective classes. Some structural characteristics appear to transcend class, for example, the prevalence of α-helical secondary structure in coat protein subunits. Other properties, such as the number of DNA nucleotides packaged per subunit, are not conserved within a class (1.0 in Pf1, 2.1 in Xf, and 2.4 in Pf3). One of the least understood aspects of filamentous virus architecture is the structure of the encapsidated ssDNA molecule, including its modes of self-interaction and interaction with the protein sheath. This is largely a consequence of the paucity of DNA in all the filamentous viruses, ranging from a low of about 6% by weight of the Pf1 virion to a high of about 14% of Pf3. Typically, both the X-ray diffraction and spectroscopic transition intensities contributed by the DNA component of a filamentous virus are diffuse and often obscured by the much more intense contributions from the viral proteins. Nonetheless, with appropriate experimental procedures DNA structural information has been obtained on some of the viruses, and this area has been reviewed recently (Day et al., 1988). With regard to axial nucleotide separations (2.7 Å) and absorbance and CD spectral profiles, the DNA molecules in class I viruses and in Xf bear some similarities to classical DNA structures. On the other hand, the axial nucleotide separation is 2.3 Å in Pf3 and as high as 6 Å in Pf1, and absorbance and CD spectra indicate little base stacking in the latter two viruses.

Solid-state ^{31}P NMR spectra of Pf1 indicate an ordered backbone in which the average phosphodioxy group (PO_2^-)

† This is paer 23 in the series "Studies of Virus Structure by Laser Raman Spectroscopy". Supported by NIH Grants AI11855 (G.J.T.), GM24266 (S.J.O.), and AI09049 (L.A.D.).

* To whom correspondence should be addressed.
‡ University of Missouri—Kansas City.
§ University of Pennsylvania.
∥ The Public Health Research Institute of the City of New York.

[1] The phages fd, f1, and M13 are closely similar members of this prototype.

0006-2960/88/0427-4350$01.50/0

is oriented with its OPO plane approximately perpendicular to the virion axis. This is consistent with a highly extended model for the Pf1 genome (Day et al., 1979). In fd, however, the DNA phosphate groups are randomly oriented (powder pattern) even though they remain immobile on the time scale of the NMR experiments (DiVerdi & Opella, 1981; Cross et al., 1983). The latter result is difficult to reconcile with fiber X-ray experiments which reveal some diffraction from fd DNA (Banner et al., 1981), indicating a pitch of about 27 Å. Apparently there is sufficient regularity to give rise to the fd DNA diffraction, despite the randomness indicated on the time scale of the NMR experiments. It is also possible that the base and sugar moieties do not share the same mobility as the phosphates and thus contribute, albeit weakly, to the fiber diffraction pattern.

In a previous Raman investigation of both class I and class II filamentous viruses, a number of Raman bands could be assigned to the viral DNAs. Many differences were apparent between different viruses, indicating that the DNA structures were not the same (Thomas et al., 1983). However, in the absence of definitive correlations between Raman band patterns and residue structures, the virus-specific conformational properties could not be determined unambiguously in the earlier work. Also, the intensities of Raman scattering contributed by the viral proteins were not accurately determined. These former limitations have now been largely overcome. On the one hand, the availability of specifically deuteriated viruses for Raman spectroscopy has provided a data base upon which definitive assignments have been established for conformation-sensitive bands of the viral DNA. On the other, recent systematic study by X-ray crystallography and Raman spectroscopy of model DNA structures (Benevides et al., 1984, 1986, 1988) has provided a tabulation of Raman spectra–structure correlations for applications to viruses (Thomas et al., 1986a). [For reviews of this field, see Thomas and Wang (1988) and Nishimura and Tsuboi (1986).] The present paper focuses on improvements in the interpretation of the Raman spectra, which lead to deductions about the specific nucleoside sugar puckering and phosphodiester group geometries in the DNA molecules of class I and class II filamentous viruses.

Materials and Methods

Viruses and Model Compounds. The viruses used in this work were prepared in two laboratories by similar protocols. From the Public Health Research Institute (L.A.D.), specimens of normal (nondeuteriated) fd, If1, IKe, Pf1, Xf, and Pf3 were grown and purified for Raman spectroscopy according to described procedures (Thomas et al., 1983). Additional samples of normal fd were obtained from stocks at the University of Pennsylvania (S.J.O.). The following specifically deuteriated fd phages were also prepared at the University of Pennsylvania: (i) fd(3F-d_4), in which the three phenylalanine residues per subunit (F11, F42 and F45) were deuteriated at ortho and meta ring positions, and (ii) fd(2Y-d_4), in which the two tyrosine residues per subunit (Y21 and Y24) were deuteriated at ortho and meta ring positions. All viruses were dissolved in 10 mM tris(hydroxymethyl)aminomethane (Tris) buffer (pH 7.5) containing also 0.15 M NaCl. This corresponds to the same buffer/salt composition employed previously for Raman spectroscopy (Thomas et al., 1983). In control experiments, the Raman spectra of deuteriotyrosine and deuteriophenylalanine were obtained at the same conditions employed for virus.

Each virus was concentrated for Raman spectroscopy to approximately 100 mg/mL by pelleting in the ultracentrifuge. The viscous solution of virus was subsequently transferred to a glass capillary cell (Kimax 34507) which was sealed and thermostated at 20 °C for recording of the Raman spectra. Similar procedures were employed for the deuteriated amino acids.

Raman Spectroscopy. Samples were irradiated with 200 mW of the 514.5-nm line from a Spectra Physics Model 171-18 argon ion laser. The Raman scattering at 90° to the incident laser beam was collected and analyzed on a Spex Ramalog V/VI spectrophotometer under the control of a North Star Horizon-II Model HRZ-2-64K-Q microcomputer (Li et al., 1981). The spectrophotometer was operated with a spectral slit width of 8 cm^{-1}, and data were collected at intervals of 1 cm^{-1} and with integration times of 1.0 s. Each virus was also reexamined with a spectral slit width of 4 cm^{-1} to improve resolution of closely spaced Raman lines (bands) and to search for shoulders on broad or diffuse bands. Generally, the Raman spectrum of each virus was scanned repetitively, and each individual spectrum, as well as the initially computed average spectrum, was displayed to ensure that artifacts were not incorporated into the average. The data shown represent the average of 50–60 scans in the 600–900-cm^{-1} interval.

The computer-averaged Raman spectra were corrected for solvent background by subtraction of the spectrum of buffer. The data in the region 600–850 cm^{-1} are not particularly sensitive to the scaling factor employed in buffer subtraction, as is well-known from the very weak Raman spectra of liquid water and Tris buffer in this interval (Thomas & Kyogoku, 1977; Thomas et al., 1983). Spectra of the buffer showed that it contributed no significant Raman scattering intensity in the 600–850-cm^{-1} interval at the concentration employed. However, in the interval 850–900 cm^{-1}, the buffer contributes somewhat more substantially to the spectrum, and solvent compensation is more problematic. Accordingly, the displayed Raman intensity of the tryptophan line at 875 cm^{-1} varies slightly from one virus sample to another, and this does not necessarily reflect differences in the structure or environment of tryptophan residues. Under Results and Discussion, quantitative assessments and conclusions are confined to bands occurring within the region 600–850 cm^{-1}. Signal averaging, base-line correcting, spectral subtraction, and related routines were performed on-line with data acquisition by the North Star microcomputer.

Spectral deconvolutions were carried out on an IBM-PC/AT microcomputer with an installed numeric data processor (Intel 80287) as described (Thomas & Agard, 1984). The constrained, iterative deconvolution procedure serves to enhance the resolution of poorly resolved bands by compensating for the broadening effects of spectral band-pass and vibrational bandwidth (Thomas, 1985). This is achieved by deconvolution of the observed band shapes with a Gauss–Lorentz function. However, the band assignments and all structural conclusions derived therefrom pertain to the unrefined experimental data shown under Results (Figures 1–3).

Results

Class I Phages: Spectra and Assignments. Figure 1 shows the Raman spectrum of phage fd in the spectral region 600–900 cm^{-1}, where conformation-sensitive bands of the DNA molecule are expected (Thomas et al., 1986a). Similar Raman bands of fd were detected previously, and preliminary assignments have been given (Thomas et al., 1983). Note that the strongest Raman bands in this spectral region are actually due to protein aromatic side chains, phenylalanine (F), tyrosine (Y), and tryptophan (W), as labeled in the figure. Weaker bands of the DNA bases and backbone are also labeled (a,

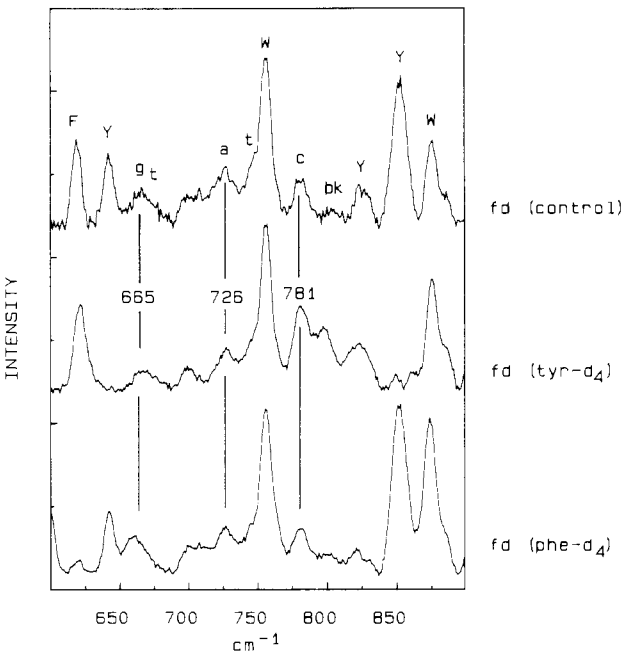

FIGURE 1: Raman spectra in the region 600–900 cm^{-1} of normal fd virus (top), fd containing subunit tyrosines deuteriated at ortho and meta positions (middle), and fd containing subunit phenylalanines deuteriated at ortho and meta positions. Data were obtained on phage at 20 °C and 100 mg/mL in pH 7.5 Tris buffer. Band assignments are indicated by lower-case letters for DNA bases (a, c, g, t) and backbone (bk) and by upper-case letters for protein side chains (F, Y, W), as discussed in the text. The apparent change in relative intensity of the tryptophan band at 875 cm^{-1} is not a consequence of tyrosine or phenylalanine deuteriation but an artifact of solvent–buffer compensation which affects the base line in the region 860–900 cm^{-1} (see text).

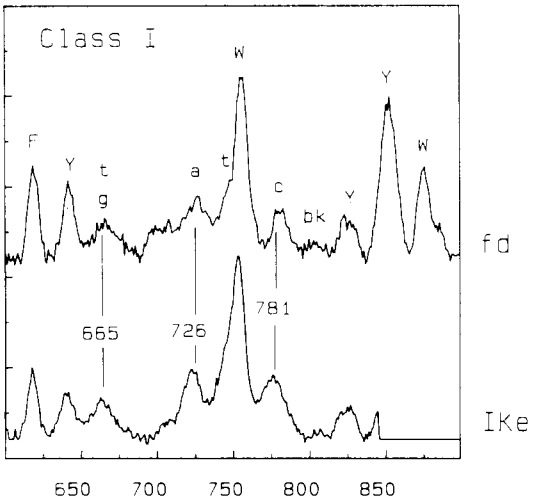

FIGURE 2: Raman spectra in the region 600–900 cm^{-1} of fd and IKe viruses showing principal band assignments. Conditions are as given in Figure 1. For clarity, the IKe spectrum is truncated at 850 cm^{-1} to eliminate strong bands of Y and W residues. [The complete spectra are shown in Thomas et al. (1983).]

t, g, c, and bk). Included in Figure 1 are corresponding Raman data for the specifically deuteriated isotopomers fd(2Y-d_4) and fd(3F-d_4). (Notation is defined above.)

The data of Figure 1 show that the strongest bands of the tyrosine residues, which occur at 640 and 850 cm^{-1} in the spectrum of the normal phage, are eliminated (i.e., shifted outside this spectral region) in the spectrum of fd(2Y-d_4). A lesser tyrosine band, at 827 cm^{-1} in the normal phage, is shifted slightly to lower frequency and is considerably broadened in the spectrum of fd(2Y-d_4), but is not eliminated. These deuteriation shifts for tyrosines in fd are confirmed by the Raman spectrum (not shown) of the isotopically labeled amino acid. The only phenylalanine band of this region occurs at 620 cm^{-1} in normal fd and is shifted outside this region in fd(3F-d_4). This assignment is also confirmed by the Raman spectrum (not shown) of isotopically labeled phenylalanine. The absence of deuteriation shifts in the bands at 755 and 875 cm^{-1} of Figure 1, as well as their high intensities, confirms their assignment to tryptophan (Thomas et al., 1983). No other tryptophan bands are expected in the 600–900-cm^{-1} interval (Takeuchi & Harada, 1986), and none are assigned. The present observations confirm all the protein aromatic assignments and establish the bands at 665, 726, and 781 cm^{-1} as bands of fd DNA, originating from purine and pyrimidine ring vibrations as indicated in Figure 1. Further discussion of the DNA base vibrations has been given by Lord and Thomas (1967). The very weak band at 803 ± 2 cm^{-1} was also assigned previously to fd DNA (Thomas et al., 1983), and that assignment is consistent with the present results. The weak band remaining at 700 cm^{-1} is assigned by default to aliphatic amino acid side chains.

It is important to note that the 827-cm^{-1} band of fd can be assigned with confidence virtually completely to tyrosine, and not to the DNA backbone, for the following reasons: (i) Studies of both single- and double-stranded DNA model structures indicate that any DNA band at this position should exhibit no more than 20% of the intensity of the companion DNA line at 781 cm^{-1} (Thomas et al., 1986a). In the case of fd, the intensity is too large by a factor of 5. (ii) Protein model structure studies show that the tyrosine band at this position must exhibit at least 20% of the intensity of the companion tyrosine line at 850 cm^{-1} (Siamwiza et al., 1975), which is the intensity ratio observed here (top spectrum of Figure 1). In fact, in most proteins the 827-cm^{-1} band intensity is about 40% of the intensity of the 850-cm^{-1} band (Thomas et al., 1986b). Accordingly, within the precision of present measurements (estimated as ±10% for low-intensity bands), there is no margin for a DNA backbone marker band at ca. 827 cm^{-1} in fd. This conclusion applies also to the other filamentous viruses which contain tyrosine, viz., IKe, Pf1, and Xf. It does not apply to Pf3, which contains no tyrosine in its major coat protein yet exhibits a band at 824 cm^{-1} with the intensity expected of a DNA backbone marker. (The significance of this band will be further discussed under Phosphodiester Conformations.) Of additional significance in the spectra of Figures 1 and 2 is the *absence* of a band at 681 cm^{-1} (diagnostic of C2′-endo pucker and *anti*-glycosyl orientation of deoxyguanosine residues) and the *presence* of a band at 665 cm^{-1} (diagnostic of C3′-endo/anti dG). The structural implications of this nucleoside conformation marker are considered under Nucleoside Conformations.

Figure 2 shows a comparison between the Raman spectrum of phage IKe and that of fd. In this illustration, the IKe spectrum has been truncated above 850 cm^{-1} to eliminate the strong tyrosine and tryptophan bands and facilitate comparison of IKe DNA bands with those of fd. [The complete spectrum of IKe has been published previously in Figure 1 of Thomas et al. (1983).] The data of Figure 2 indicate a rigorous band-by-band correspondence between Raman spectra of the two class I phages, especially with regard to the DNA contributions. As in the case of fd, IKe lacks a band at 681 cm^{-1} and exhibits no appreciably intense Raman band between 800 and 850 cm^{-1}. At most a very weak band is present at 804 ± 2 cm^{-1}, and a similarly very weak band could underlie the tyrosine band at 827 cm^{-1}. Identical results have been obtained for another class I phage, If1 (data not shown). In the fol-

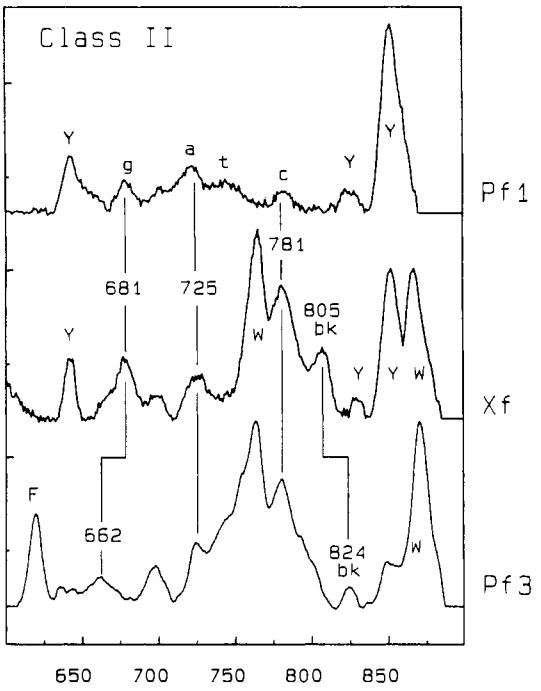

FIGURE 3: Raman spectra in the region 600–900 cm^{-1} of Pf1, Xf, and Pf3 viruses. Conditions are as given in Figure 1. [See also Thomas et al. (1983).]

lowing discussions, the conclusions stated for fd and IKe will apply also for If1.

Class II Phages: Spectra and Assignments. Raman spectra in the region 600–900 cm^{-1} of phages Pf1, Xf, and Pf3 are compared in Figure 3. The spectra are strikingly different from those of class I phages (cf. Figure 2) and from one another. Assignment of most of the bands is straightforward, by analogy with the data shown in Figure 1 and discussed above. In the case of Pf1, which lacks phenylalanine and tryptophan residues, the only interfering bands of the protein are those of tyrosine at 640, 827, and 850 cm^{-1} and those of aliphatic groups which contribute at 700, 725, and 750 cm^{-1}. The latter are usually weak, but in the case of Pf1 they combine with the Raman intensities of adenine at 725 cm^{-1} and thymine at 750 cm^{-1} to produce significantly intense and rather broad bands in the spectrum of Figure 3. The 827-cm^{-1} band, as before, can be attributed entirely to tyrosine residues of Pf1 coat protein. It is remarkable that the spectrum of Pf1 contains so weak a Raman line at 781 cm^{-1}, and no apparent contribution from the DNA backbone between 800 and 850 cm^{-1}. Ordinarily, the cytosine residues of DNA produce a strong band at ca. 781 cm^{-1}, as observed in all other viruses. The extraordinarily low 781-cm^{-1} intensity in Pf1 reflects the DNA composition of this virus (6 wt %) and indicates that the failure to detect a band from the backbone between 800 and 850 cm^{-1} cannot be regarded as conclusive evidence of the absence of such a band. As noted above, such a DNA backbone marker would be expected to exhibit about 20% or less of the 781-cm^{-1} band intensity, which would place its intensity below or close to the limit of detection (i.e., at the noise level) in the Pf1 spectrum.

In the case of Xf (Figure 3, middle), the centers of all Raman bands correlate well with those of Pf1, and the observed intensities are consistent with the DNA composition (12.6 wt %). The Xf virion also lacks phenylalanine in its coat protein but contains tryptophan, which accounts for the strong Raman bands at 760 and 870 cm^{-1}. The strong 805-cm^{-1} band of Xf arises from the DNA backbone and is clearly more intense than any corresponding band in Pf1 or in class I phages. Its structural significance is considered under Phosphodiester Conformations.

The Pf3 virion, which contains 14 wt % DNA, yields the most noise-free Raman spectrum and a number of remarkable differences from both Pf1 and Xf. The absence of tyrosine in the coat protein of this virus compels assignment of the significantly intense Raman band at 824 cm^{-1} to the DNA backbone. We may regard this band as the counterpart of the 805-cm^{-1} band of Xf (and fd, IKe and If1; cf. Figure 2). Unlike Pf1 and Xf, but like fd, IKe, and If1, the Pf3 virus displays a band at 662 cm^{-1} and lacks a band at 681 cm^{-1}. The significance of the prominent 805- and 662-cm^{-1} bands of Pf3 is considered below (see Discussion). There is also evidence of a weak residual shoulder just below 800 cm^{-1} in Pf3, assignable to the DNA (Thomas et al., 1986a). We offer no definitive assignment for the feeble shoulder (at 850 cm^{-1}) to the intense 875-cm^{-1} tryptophan band of Pf3. Although Pf3 DNA could contribute some intensity in this region, we regard it as unlikely that DNA is the sole or major contributor. Since aliphatic amino acid side chains are known to give very weak Raman scattering in this region (Thomas et al., 1983), we presume that their cumulative intensity could account for the very weak shoulder observed (Thomas et al., 1983). It should be kept in mind that the 850-cm^{-1} shoulder in Pf3 is about an order of magnitude weaker than the prominent tyrosine bands near 850 cm^{-1} in other viruses. Thus, feeble contributions from aliphatic side chains in those viruses would not alter the definitive assignments stated earlier.

DISCUSSION

DNA and RNA structures solved by X-ray crystallography have provided a data base for identification of Raman bands which serve as specific structure markers (Benevides et al., 1984, 1986, 1988). Two types of structure markers have emerged from these investigations: (i) Raman bands originating from vibrations in the nucleoside which are sensitive to sugar pucker (torsion angle, δ) and glycosyl torsion (χ) and are termed "nucleoside conformation markers" and (ii) Raman bands originating from vibrations in the phosphodiester group which are sensitive to the backbone torsion angles (α, ζ) and are termed "phosphodiester conformation markers" (Thomas et al., 1986a). These correlations between Raman bands and local geometry are valid, irrespective of the secondary or tertiary conformations of the macromolecule. We discuss here the observed DNA conformation markers of filamentous viruses (Figures 1–3) in relation to the existing data base compiled from model DNA structures.

In the discussion which follows, it is important to keep in mind that each band in the Raman spectrum arises from a vibration localized in a specific type of chemical group and monitors only the local geometry of that group. The nucleoside conformation markers, for example, monitor the conformation of the nucleoside sugar group irrespective of whether the secondary structure is single stranded or double stranded or whether the DNA is bound to protein or is protein free. Similarly, the phosphodiester conformation markers are diagnostic of the phosphodiester group geometry and do not provide information, except by inference, about the secondary or tertiary conformations of the nucleic acid backbone. Therefore, we generally avoid the use of terminology such as A-, B-, or Z-DNA when referring to the structural features of the viral ssDNA genomes monitored by Raman spectroscopy, even though comparison with such structures is sometimes useful.

Nucleoside Conformations. The Raman band most informative of nucleoside conformation is the dG conformation

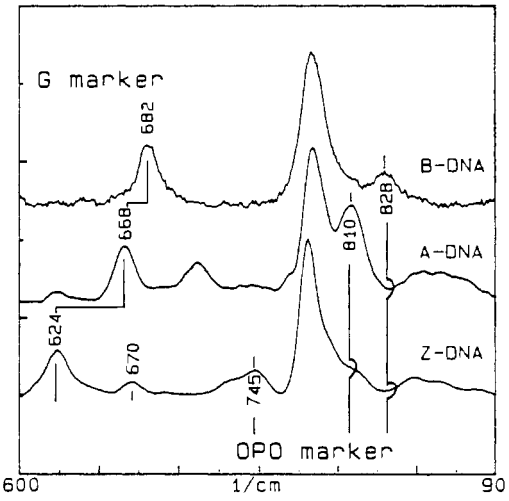

FIGURE 4: Raman spectra in the region 600–900 cm^{-1} of DNA models representative of the three canonical forms of DNA and containing the same base composition. Conformation markers are labeled for guanosine (G) and the backbone (OPO). [Data from Benevides et al. (1984, 1986, 1988) and unpublished work of G. J. Thomas, Jr., and J. M. Benevides.]

marker, which occurs at 625 cm^{-1} for C3′-endo/syn (as occurs in Z-DNA), 668 cm^{-1} for C3′-endo/anti (A-DNA, A-RNA), and 681 cm^{-1} for C2′-endo/anti (B-DNA). Minor deviations from the 681-cm^{-1} band position have also been reported for variants within the B-DNA family (Thomas, 1986; Nishimura & Tsuboi, 1986). The dT conformation is well marked by bands at 665 and 748 cm^{-1} for C2′-endo/anti (B-DNA) and by a band at 777 cm^{-1} for C3′-endo/anti (A-DNa) (Thomas & Benevides, 1985). Although reliable markers also exist for dA and dC residues, they occur outside the 600–900-cm^{-1} interval and, in the case of filamentous viruses, are obscured by much stronger bands of the coat protein. The present analysis thus rests upon interpretation of the bands of dG and dT in the 600–900-cm^{-1} region. For reference, we reproduce in Figure 4 the Raman spectra in the 600–900-cm^{-1} region recently obtained from A-, B-, and Z-DNA models.

Figure 1 (above) demonstrated the absence of a 681-cm^{-1} band in fd, which rules out C2′-endo/anti dG conformers and variants thereof, including C3′-exo, C1′-exo, and O4′-endo/anti. Figure 1 also demonstrated the absence of a 625-cm^{-1} band in fd(2Y-d_4) and fd(3F-d_4), which rules out C3′-endo/syn dG conformers. The band of fd at 665 cm^{-1} thus indicates that only C3′-endo/anti dG conformers are consistent with the fd DNA structure. The low intensity of the 665-cm^{-1} band and the absence of a strong band at 745 cm^{-1} in all spectra of Figure 1 are indications that C2′-endo/anti dT conformers are also absent from fd. On the other hand, the significant intensity of the 781-cm^{-1} band and its low-frequency shoulder (also noticeable in IKe, Figure 2) are consistent with a companion 777-cm^{-1} band from C3′-endo/anti dT conformers. All of the spectral characteristics of fd DNA are thus consistent with C3′-endo/anti conformers of dG and dT (and by extension, dA and dC), and these spectral characteristics are shared by IKe (Figure 2) and If1 (not shown).

The conclusions reached regarding nucleoside conformations in fd, IKe, and If1 apply also to Pf3 but contrast sharply with the data shown for Pf1 and Xf (Figure 3). In accordance with the correlations described above and with the data from model structures shown in Figure 4, it is evident that C2′-endo/anti conformers are alone consistent with the nucleoside conformation markers of Xf and Pf1. The absence of phenylalanine in Pf1 and Xf also leaves a clear Raman window from 600 to 640 cm^{-1} through which no trace of the C3′-endo/syn dG marker can be seen. Further, the contour of the 681-cm^{-1} band in Xf is an exact replica of the profile occurring in B-DNA: The major 681-cm^{-1} peak thus indicates the C2′-endo/anti dG conformation, and the low-frequency shoulder, partially resolved at ca. 670 cm^{-1}, indicates the C2′ endo/anti dT conformation (Thomas, 1986; Thomas & Benevides, 1985).

Taken together, the results of Figures 2 and 3 show that nucleoside conformations in Xf and Pf1 differ markedly from those in Pf3, fd, IKe, and If1.

Phosphodiester Conformations. The ordered DNA backbone in B family structures requires phosphodiester torsion angles [i.e., (α, ζ) of the 3′-C–O–P–O–C-5′ network] within the respective ranges (−60 ± 20°, −120 ± 30°). This conformation, also called gauche$^-$,gauche$^-$ (Arnott et al., 1976), typically generates a Raman band of moderate intensity and large bandwidth near 835 ± 6 cm^{-1} (Erfurth et al., 1972; Thomas et al., 1986a). An illustration is provided by the spectrum of the B-DNA model shown in Figure 4. The greater the range of torsion angles, the greater is the bandwidth of the B backbone marker (Benevides et al., 1988). The B backbone also generates a somewhat more intense band near 790 cm^{-1} which overlaps the 781-cm^{-1} band of dC residues. The effect of the overlapping 781- and 790-cm^{-1} bands is to give the composite band a much greater breadth than other Raman bands in the spectrum and to push the apparent band center close to 785 cm^{-1} (Benevides et al., 1988). This feature is also apparent in the spectrum of the B-DNA model shown in Figure 4.

The ordered DNA backbone in A family structures is far more regular than that of B-DNA, and its narrower range of OPO torsion angles (−65 ± 15°, −65 ± 15°) gives rise to a much sharper and more intense Raman band located at 807 ± 3 cm^{-1} (Erfurth et al., 1972; Thomas & Hartman, 1973; Thomas et al., 1986a). This characteristic is evident in the spectrum of the A-DNA crystal of Figure 4.

We have included the spectrum of a Z-DNA crystal in Figure 4 for comparison with the A- and B-DNA models. The Z-DNA backbone generates a moderately intense backbone marker near 747 cm^{-1}, as well as relatively broad bands which are difficult to resolve from the 781-cm^{-1} band of dC. The absence of syn conformers of dG in significant amount in any viral DNA, as deduced from the nucleoside conformation markers discussed above, makes it unlikely that any backbone contains appreciable amounts of the phosphodiester geometry usually associated with the left-handed Z-type backbone.

Figures 1 and 2 confirm the absence of strong bands near 807 ± 3 cm^{-1} in all class I viruses, thus excluding the phosphodiester geometry characteristic of *regular* A-type structure. The same conclusion applies to Pf1 and Pf3 (Figure 3). The very weak band near 804 ± 2 cm^{-1} in fd, IKe, and Pf1 is much too feeble to indicate extensive A-type conformation in these viruses. On the other hand, the low intensity of this Raman band is consistent with a small fraction of the A-type conformation. By analogy with Figure 4 and data from other model compounds (Thomas & Hartman, 1973), the weak 804-cm^{-1} band would allow for up to 20% (i.e., 10 ± 10%) of the DNA backbone to adopt the A-type conformation in fd, IKe, and Pf1. The large uncertainty in this number reflects the low signal-to-noise ratio observed in the spectra.

Xf is the only virus which clearly exhibits a strong band at 805 cm^{-1}. Both the frequency and intensity of this band are diagnostic of a relatively uniform backbone conformation containing phosphodiester geometry for much of the Xf genome consistent with that found in A-DNA, notwithstanding the presence of C2′-endo/anti nucleoside conformers in this

Table I: Conformational Significance of Raman Bands of DNA in Filamentous Viruses

virus[a]	frequency (cm^{-1})[b]	assignment[c]	conformation proposed[d]
fd, If1, IKe	665	g, (t)	C3′-endo/anti
	726	a	C3′-endo/anti
	745	t	C3′-endo/anti
	781	c	C3′-endo/anti
	804	OPO	?
Pf1	681	g	C2′-endo/anti
	725	a	C2′-endo/anti
	740	t	?
	781	c	C2′-endo/anti
Xf	681	g	C2′-endo/anti
	725	a	C2′-endo/anti
	781	c	C2′-endo/anti
	805	OPO	A genus
Pf3	662	g, (t)	C3′-endo/anti
	725	a	C3′-endo-anti
	745	a	C3′-endo/anti
	781	c	C3′-endo/anti
	824	OPO	B genus

[a] Data on If1 are shown in Thomas et al. (1983). [b] Frequencies are accurate to ±2 cm^{-1} for sharp lines and ±3 cm^{-1} for broad lines or shoulders. [c] Lower-case letters indicate nucleoside contributions. Minor contributors are indicated in parentheses. OPO indicates phosphodiester group conformation marker, dependent on torsions (α, ζ). [d] Nucleoside sugar pucker and glycosyl torsion, where known, are indicated for bands of a, c, g, and t. The weak 804-cm^{-1} band of fd, If1, and IKe is not consistent with phosphodiester geometries, similar to those of either A-DNA or B-DNA. The strong 805-cm^{-1} band of Xf is consistent with a 50% or more of phosphodiester group torsions (α, ζ) like those occurring in A-DNA, and the 824-cm^{-1} band of Pf3 suggests up to 50% of (α, ζ) torsions like those occurring in B-DNA.

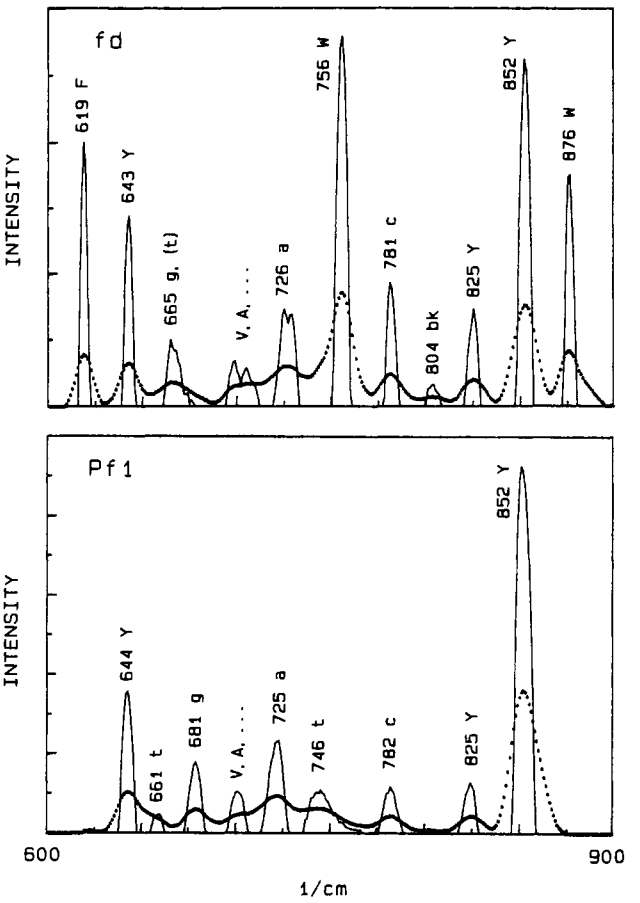

FIGURE 5: Comparison of observed (···) and deconvolved (—) Raman spectra of fd (top) and Pf1 (bottom) in the region 600–900 cm^{-1}. Deconvolution was achieved with a Gauss–Lorentz desmearing function of 17-cm^{-1} half-width (Thomas & Agard, 1984). The complex band near 700 cm^{-1} in each virus has been assigned to aliphatic amino acid side chains (Thomas et al., 1983). Note that an artifact of deconvolution is elimination of the shoulder near 745 cm^{-1} in fd, assigned to thymidine residues (cf. Figure 2).

virus. A reliable estimate of the percentage of A-type backbone structure in Xf DNA requires knowledge of the Raman intensity at 1100 cm^{-1} (Thomas & Hartman, 1973), which unfortunately is obscured by the viral protein. However, if we assume that the intensity of the 781-cm^{-1} band of Xf (Figure 3) accrues only from dC, and serves as a basis for comparison between Figure 3 and Figure 4, we may conclude that approximately 50% of the Xf genome contains nucleotide OPO groups in the same conformation found in A-DNA. [Note that the model A-DNA structure of Figure 4 contains 50% C nucleotides, whereas Xf DNA contains about 25% dC. Accordingly, the 810/783-cm^{-1} band intensity ratio of Figure 4 should be scaled by a factor of 2 for quantitative comparison with the corresponding 805/781 intensity ratio in Xf. See also spectra of calf thymus DNA fibers published elsewhere (Erfurth et al., 1972; Prescott et al., 1984).]

Figures 1–3 also confirm the absence of a moderately intense band near 835 ± 6 cm^{-1} in fd, IKe, Pf1, and Xf, excluding from these viruses the phosphodiester geometry associated with B-type backbone conformations. As noted above, the bands near 827 cm^{-1} in these viruses are assignable in entirety to tyrosines. Further, the centering of the dC marker near 781 cm^{-1} is evidence that no overlapping 790-cm^{-1} band (B backbone) is present in any of these structures. In contrast, the Pf3 spectrum shows a band at 824 cm^{-1}, which cannot be due to tyrosine, is consistent with an origin in the DNA backbone, and is indicative of OPO backbone geometry different from that found both in class I viruses and in Pf1 and Xf. Among all viruses, Pf3 also displays the broadest shoulder on the high-frequency tail of its 781-cm^{-1} band, which most closely approximates the 790-cm^{-1} feature of B-type backbone geometry. The sharpness of the 824-cm^{-1} band in Pf3 and its lower than usual frequency vis-à-vis the standard B-form marker (ca. 835 cm^{-1}) are indications of a narrower range of (α, ζ) torsion angles than found in conventional B-DNA conformations. The band intensity appears to be consistent with a structure involving about 50% or so of the viral genome.

The significant differences between conformation-sensitive Raman bands of prototypical class I (fd) and class II (Pf1) phages are emphasized by resolution enhancement through Fourier deconvolution as shown in Figure 5. The important conformation markers identified for all six viruses and their structural interpretations are summarized in Table I.

Summary and Conclusions

The principal conclusions from this work are the following:

(i) All class I viruses (fd, IKe, If1) contain similar DNA conformations. These have been characterized on the basis of model compound studies as containing the type of nucleoside sugar pucker and glycosyl torsion found in A-DNA, i.e., C3′-endo/anti. This conclusion is consistent with the very similar CD spectra of these viruses (Casadevall & Day, 1983). However, the viral genomes clearly lack the regular phosphodiester backbone geometry found in either A-DNA or B-DNA. Specifically, the backbone torsion angles (α, ζ) of the 3′-C–O–P–O–C-5′ network do not occur uniformly within the gauche$^-$,gauche$^-$ intervals, usually associated with either the A-form (65 ± 15°, 65 ± 15°) or B-form (–60 ± 20°, –120 ± 30°) backbone (Arnott et al., 1976). The data are consistent with only 20% of the phosphodiester groups configured as in A-DNA and an indeterminate but smaller percentage (less than 20%) configured as in B-DNA. In short, the class I viral

DNA backbone is regarded as encompassing a wide range of (α, ζ) conformation angles, the majority of which are not found in canonical DNA secondary structures.

(ii) The class II viruses (Pf1, Xf, Pf3) are dissimilar in their DNA conformations. Pf1 and Xf do share C2′-endo/anti nucleoside conformers, but C3′-endo/anti conformers are indicated for Pf3. The Pf1 DNA backbone is similar to that of class I viral DNAs in phosphodiester geometries, i.e., encompassing a wide range of (α, ζ) values. On the other hand, in Xf DNA, we estimate that 50% of the phosphodiester groups are conformationally equivalent to those found in A-DNA structures. In Pf3 DNA, at least 50% of the phosphodiester groups exhibit a conformation close to that found in B-DNA models. The Raman spectra do not identify unambiguously the remaining OPO conformations for Pf3 or Xf DNA but do exclude regular A- or B-form structures.

(iii) The Raman spectra rule out large amounts of syn nucleoside conformers in DNA of class I viruses and of Pf1 and Xf. Although no compelling evidence for syn conformers can be found in the spectrum of Pf3, a definitive conclusion on this point must await the availability of the specifically deuteriated phenylalanyl isotopomer of Pf3. We note that the so-called "morphogenetic signal" or *mos* sequence of f1 phage DNA, a 75-base segment largely palindromic and rich in alternating dG-dC (Webster & Lopez, 1985), could exist in fd DNA with syn purine nucleosides and phosphodiester geometry like that found in Z-DNA and not be detected by the present experiments.

(iv) The similar DNA conformations in the three class I phages provide circumstantial evidence for similar protein packaging schemes and similar protein–DNA interactions in all class I phages. Conversely, the dissimilarity between class I DNA structures and class II structures suggests different DNA packaging motifs. No class II phage contains the same DNA fingerprint as the class I phages. Thus, although Pf1 and fd exhibit similar phosphodiester conformations, the nucleoside conformers (sugar puckers) are diametrically opposed on the pseudorotational diagram which is used to represent furanose puckering (Saenger, 1984). For Pf3 and fd DNA, in which the furanose puckering schemes are the same, the backbone OPO geometries are radically different.

The present finding of a nonuniform DNA backbone conformation in fd is not inconsistent with ^{31}P NMR results. The immobility in fd DNA dioxy (PO_2^-) groups, as deduced from NMR measurements (Cross et al., 1983), occurs on a time scale several orders of magnitude greater than that of Raman transitions which have a period of the order of 50 fs and does not conflict with either the Raman or fiber X-ray results. A requirement imposed by the Raman results is that the packaged fd genome contains predominantly, if not exclusively, C3′-endo sugar puckering of the nucleosides. Thus, conformational randomness proposed for the fd DNA phosphate groups on the basis of NMR spectral data is evidently not propagated into the nucleoside moieties of the DNA.

References

Arnott, S., Smith, P. J. C., & Chandrasekaran, R. (1976) in *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed. (Fasman, G. D., Ed.) Vol. 2, CRC, Cleveland.

Banner, D. W., Nave, C., & Marvin, D. A. (1981) *Nature (London) 289*, 814–816.

Benevides, J. M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., Rich, A., & Thomas, G. J., Jr. (1984) *Nucleic Acids Res. 12*, 5913–5925.

Benevides, J. M., Wang, A. H.-J., Rich, A., Kyogoku, Y., van der Marel, G. A., van Boom, J. H., & Thomas, G. J., Jr. (1986) *Biochemistry 25*, 41–50.

Benevides, J. M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., & Thomas, G. J., Jr. (1988) *Biochemistry 27*, 931–938.

Casadevall, A., & Day, L. A. (1983) *Biochemistry 22*, 4831–4842.

Cross, T. A., Tsang, P., & Opella, S. J. (1983) *Biochemistry 22*, 721–726.

Day, L. A. (1969) *J. Mol. Biol. 39*, 265–277.

Day, L. A., Wiseman, R. L., & Marzec, C. J. (1979) *Nucleic Acids Res. 7*, 1393–1403.

Day, L. A., Marzec, C. J., Reisberg, S. A., & Casadevall, A. (1988) *Annu. Rev. Biophys. Biophys. Chem.* (in press).

Denhardt, D. T., Dressler, D., & Ray, D. S., Eds. (1978) *The Single Stranded DNA Phages*, Cold Spring Harbor Laboratory Press, New York.

DiVerdi, J. A., & Opella, S. J. (1981) *Biochemistry 20*, 280–284.

Erfurth, S. C., Kiser, E. J., & Peticolas, W. L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 938–941.

Li, Y., Thomas, G. J., Jr., Fuller, M., & King, J. (1981) *Prog. Clin. Biol. Res. 64*, 271–283.

Lord, R. C., & Thomas, G. J., Jr. (1967) *Spectrochim. Acta, Part A 23A*, 2551–2591.

Makowski, L. (1984) *Biol. Macromol. Assem. 1*, 203–253.

Makowski, L., & Caspar, D. L. D. (1981) *J. Mol. Biol. 145*, 611–617.

Marvin, D. A., Wiseman, R. L., & Wachtel, E. J. (1974a) *J. Mol. Biol. 82*, 121–138.

Marvin, D. A., Pigram, W. J., Wiseman, R. L., Wachtel, E. J., & Marvin, F. J. (1974b) *J. Mol. Biol. 88*, 581–600.

Nishimura, Y., & Tsuboi, M. (1986) in *Advances in Spectroscopy* (Clark, R. J. H., & Hester, R. E., Eds.) Vol. 13, pp 177–232, Wiley, New York.

Opella, S. J., Stewart, P. L., & Valentine, K. G. (1987) *Q. Rev. Biophys. 19*, 1–49.

Peterson, C., Dalack, G., Day, L. A., & Winter, W. T. (1982) *J. Mol. Biol. 162*, 877–881.

Prescott, B., Steinmetz, W., & Thomas, G. J., Jr. (1984) *Biopolymers 23*, 235–256.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, Berlin.

Siamwiza, M. N., Lord, R. C., Chen, M. C., Takamatsu, T., Harada, I., Matsuura, H., & Shimanouchi, T. (1975) *Biochemistry 14*, 4870–4876.

Takeuchi, H., & Harada, I. (1986) *Spectrochim. Acta, Part A 42A*, 1069–1075.

Thomas, G. J., Jr. (1985) *Spectrochim. Acta, Part A, 41A*, 217–221.

Thomas, G. J., Jr. (1986) in *Advances in Spectroscopy* (Clark, R. J. H., & Hester, R. E., Eds.) Vol. 13, pp 233–309, Wiley, New York.

Thomas, G. J., Jr. (1987) in *Biological Applications of Raman Spectroscopy* (Spiro, T. G., Ed.) Vol. 1, Chapter 4, pp 135–201, Wiley, New York.

Thomas, G. J., Jr., & Hartman, K. A. (1973) *Biochim. Biophys. Acta 312*, 311–322.

Thomas, G. J., Jr., & Kyogoku, Y. (1977) *Pract. Spectrosc. 1C*, 717–872.

Thomas, G. J., Jr., & Agard, D. A. (1984) *Biophys. J. 46*, 763–768.

Thomas, G. J., Jr., & Benevides, J. M. (1985) *Biopolymers 24*, 1101–1105.

Thomas, G. J., Jr., & Wang, A. H.-J. (1988) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J.,

Eds.) Vol. 2, Springer-Verlag, Berlin.
Thomas, G. J., Jr., Prescott, B., & Day, L. A. (1983) *J. Mol. Biol. 165*, 321–356.
Thomas, G. J., Jr., Prescott, B., & Benevides, J. M. (1986a) *Biomol. Stereodyn. 4*, 227–254.
Thomas, G. J., Jr., Prescott, B., Benevides, J. M., & Weiss, M. A. (1986b) *Biochemistry 25*, 6768–6778.
Webster, R. E., & Lopez, J. (1985) in *Virus Structure and Assembly* (Casjens, S., Ed.) pp 235–267, Jones and Bartlett, Boston.

Chemical and Biological Studies of the Major DNA Adduct of *cis*-Diamminedichloroplatinum(II), *cis*-[$Pt(NH_3)_2${d(GpG)}], Built into a Specific Site in a Viral Genome†

Lisa J. Naser, Ann L. Pinto, Stephen J. Lippard,* and John M. Essigmann*
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Received November 9, 1987; Revised Manuscript Received February 2, 1988

ABSTRACT: A duplex *Escherichia coli* bacteriophage M13 genome was constructed containing a single *cis*-[$Pt(NH_3)_2${d(GpG)}] intrastrand cross-link, the major DNA adduct of the anticancer drug *cis*-diamminedichloroplatinum(II). The duplex dodecamer d(AGAAGGCCTAGA)·d(TCTAGGCCTTCT) was ligated into the *Hin*cII site of M13mp18 to produce an insertion mutant containing a unique *Stu*I restriction enzyme cleavage site. A genome with a 12-base gap in the minus strand was created by hybridizing *Hin*cII-linearized M13mp18 duplex DNA with the single-stranded circular DNA of the 12-base insertion mutant. The dodecamer d(TCTAGGCCTTCT) was synthesized by the solid-phase phosphotriester method and platinated by reaction with *cis*-$[Pt(NH_3)_2(H_2O)_2]^{2+}$ (yield 39%). Characterization by pH-dependent ^1H NMR spectroscopy established that platinum binds to the N7 positions of the adjacent guanosines. The platinated oligonucleotide was phosphorylated in the presence of [γ-^{32}P]ATP with bacteriophage T4 polynucleotide kinase and incorporated into the 12-base gap of the heteroduplex, thus situating the adduct specifically within the *Stu*I site in the minus strand of the genome. Approximately 80% of the gapped duplexes incorporated a dodecanucleotide in the ligation reaction. Of these, approximately half did so with the dodecanucleotide covalently joined to the genome at both 5′ and 3′ termini. The site of incorporation of the dodecamer was mapped to the expected 36-base region delimited by the recognition sites of *Xba*I and *Hind*III. The *cis*-[$Pt(NH_3)_2${d(GpG)}] cross-link completely inhibited *Stu*I cleavage, which was fully restored following incubation of the platinated genome with cyanide to remove platinum as $[Pt(CN)_4]^{2-}$. Gradient denaturing gel electrophoresis of a 289-base-pair fragment encompassing the site of adduction revealed that the presence of the *cis*-[$Pt(NH_3)_2${d(GpG)}] cross-link induces localized weakening of the DNA double helix. In addition, double- and single-stranded genomes, in which the *cis*-[$Pt(NH_3)_2${d(GpG)}] cross-link resides specifically in the plus strand, were constructed. Comparative studies revealed no difference in survival between platinated and unmodified double-stranded genomes. In contrast, survival of the single-stranded platinated genome was only 10–12% that of the corresponding unmodified single-stranded genome, indicating that the solitary *cis*-[$Pt(NH_3)_2${d(GpG)}] cross-link is lethal to the single-stranded bacteriophage.

Most chemicals that bind to DNA form a spectrum of adducts, some or all of which can serve as lethal and/or premutagenic lesions (Miller, 1978). Genotoxicity has been exploited by development of chemotherapeutic agents that kill rapidly dividing cells by binding to DNA and inhibiting replication. One of the most effective antitumor drugs to date is *cis*-diamminedichloroplatinum(II) (*cis*-DDP).[1] This bifunctional electrophilic compound reacts with DNA to form a variety of intra- and interstrand cross-links [for a review, see Sherman and Lippard (1987)]. The principal adducts are the *cis*-[$Pt(NH_3)_2${d(GpG)}] and *cis*-[$Pt(NH_3)_2${d(ApG)}] intrastrand cross-links. Minor products include *cis*-[$Pt(NH_3)_2${d(GpNpG)}] intrastrand cross-links (where N is any intervening nucleotide), interstrand cross-links, and monoadducts. In all of these cases, the N7 atoms of the purine bases have replaced the chloride ligands in the *cis*-DDP coordination plane. The clinically inactive geometric isomer of *cis*-DDP, *trans*-diamminedichloroplatinum(II) (*trans*-DDP), also binds

† This work was supported by National Institutes of Health Research Grants CA40817 and T32ES00597 (to J.M.E.) and CA34992 (to S.J.L.). L.J.N. acknowledges fellowship support from Hazleton Laboratories. The high-field NMR studies were performed at the NMR Facility for Biomolecular Research located at the Francis Bitter National Magnet Laboratory, MIT, which is supported by Grant RR00995 from the Division of Research Resources of the NIH and by the National Science Foundation under Contract C-670.
* Correspondence should be addressed to these authors.

[1] Abbreviations: bp, base pair; DDP, diamminedichloroplatinum(II); ds, double stranded; EthBr, ethidium bromide; EDTA, disodium salt of ethylenediaminetetraacetic acid; IPTG, isopropyl β-D-thiogalactopyranoside; nt, nucleotide; RF, replicative form; ss, single stranded; TAE, 40 mM Tris–acetate (pH 7.5) and 1 mM EDTA; TE, 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside.

0006-2960/88/0427-4357$01.50/0