Volume 7 Number 6 1979 Nucleic Acids Research

Structure models for DNA in filamentous viruses with phosphates near the center

Loren A.Day, Robert L.Wiseman and Christopher J.Marzec

Department of Biochemistry, The Public Health Research Institute of The City of New York, Inc., 455 First Avenue, New York, NY 10016, USA

Received 10 October 1979

ABSTRACT

DNA structure models deduced from X-ray and physicochemical data for Pf1, Xf, and fd viruses have two antiparallel chains wound in helices of ∿15 Å pitch with the phosphates near the structure axes and the bases directed outward. The models, which differ for each virus, are used to interpret ultraviolet absorbance and fluorescence data in terms of DNA-protein interactions.

INTRODUCTION

Each of the filamentous bacterial viruses Pf1, Xf and fd contains a circular single-stranded DNA packed in a sheath of protein subunits. The X-ray fiber diffraction patterns for Pf1 and Xf indicate that each protein sheath has subunits in a one start helix of ∿15 Å pitch (1-3). Diffraction patterns for fd (4,5) have been interpreted as indicative of a perturbed one start helix of ∿15 Å pitch (5) or as indicative of a five start helix (3). Despite similarities in subunit packing in all three viruses, differences in DNA packing are pronounced. For example, Pf1 is 2.0μ long and Xf is 1.0μ long, but both of their DNAs contain ∿7400 nucleotides, and chemical data indicate almost exactly one nucleotide per subunit in Pf1 and two nucleotides per subunit in Xf (6,7). On the other hand, fd is 0.9μ long with 6400 nucleotides, and it has 2.3 or 2.4 nucleotides per subunit (8-12). The amino acid sequences give subunit molecular weights of 4609 for Pf1, 4343 for Xf and 5240 for fd and show the basic groups clustered in the carboxy-terminal portions, acidic groups in the amino-terminal portions, and nonpolar residues in the central portions of all three subunits but there are significant differences in the details of the sequences (13-15).

Herein we describe some geometrical considerations which constrain the packing of the DNAs in these viruses. It is assumed that the single-stranded circular DNA molecule in each virion extends from one end to the other and back again so that there are two antiparallel chains. It is also assumed that

the DNA structures are helical.

TWO CHAIN DNA STRUCTURES WITH ∿15 Å PITCH

Consider first Pfl. If there were exactly one nucleotide per subunit (6,7), one would expect that the antiparallel chains would each lie in a helix of the same pitch and helical sense as the one start helix of protein subunits. This could occur if a given subunit were to interact with a nucleotide from the "down" chain and the next subunit were to interact with a nucleotide from the "up" chain. Following this "alternate interaction" concept, we use parameters of the helix of subunits to deduce parameters for the DNA helix. It is known that there are close to 5.4 protein subunits per turn of the protein helix (2,3,6,7). The "alternate interaction" concept then implies 2.7 nucleotides per turn of each DNA chain and thus a rotation per nucleotide of 360°/2.7 = 133°. The axial rise per nucleotide in each DNA chain would be twice the rise per subunit and it varies from 5.4 Å to 5.6 Å depending on experimental conditions (1-3). A direct experimental value for the axial rise per nucleotide is obtained by dividing the virus length by half the number of nucleotides, 2.0μ/3700 = 5.4 Å (6). The two parameters, the axial rise per nucleotide and the rotation per nucleotide, together with the maximum distance of 7.1 Å between neighboring phosphates in one chain (16), give 2.5 Å for the maximum distance of the phosphates from the structure axis. The simple geometrical relationships are shown in Figure 1.

The stereochemical feasibility of DNA helices with ∿15 Å pitch and axial nucleotide translations of 5 Å to 6 Å was first explored by mani-

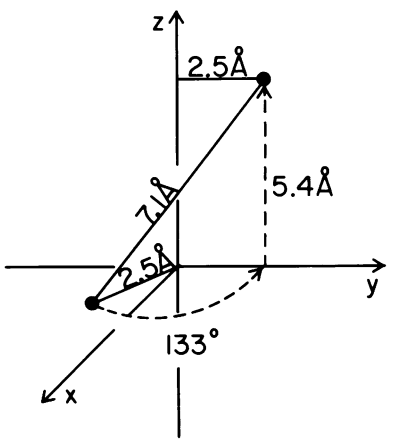

Figure 1. Geometrical relationship between the axial nucleotide translation of 5.4 Å, the rotation per nucleotide of 133°, and the interphosphate distance of 7.1 Å used to calculate the maximum distance of 2.5 Å between the phosphates and the structure axis of Pfl.

pulations with space-filling, Corey-Pauling-Koltun (CPK) atomic models (17). It proved quite easy to obtain conformations which satisfied the constraints. All had the phosphates near the axis and the bases directed outward. The bases on the outside could be unstacked or arranged in small "stacks" consisting of one base from each chain. Figure 2A shows a 15 Å pitch model with bases unstacked for Pf1 DNA.

In Xf the axial rise per nucleotide is 1.0μ/3700 = 2.7 Å, and for 5.4 nucleotides per turn, the rotation per nucleotide is 360°/5.4 = 67°. These parameters yield a maximum distance of phosphates from the axis of Xf of 6.0 Å. As in the case of Pf1, we could satisfy the constraints only with phosphates directed toward the axis. Many particular structures are possible, since the phosphates can be closer to the axis than 6.0 Å, and the bases can be stacked or unstacked and have various orientations relative to the axis. The general type of structure envisaged for the DNA in Xf is shown in Figure 2B.

The noninteger number of nucleotides/subunit in fd makes the deduction of the general features of the DNA structure in fd less direct. The axial rise per nucleotide in fd is 0.90μ/3200 = 2.8 Å. The X-ray fiber diffraction patterns (4,5) indicate an approximate repeat of 32 Å and show meridional intensity at 16 Å, and, as mentioned above, both a one start helix and a five start helix of protein subunits have been proposed to explain the patterns (3,5). In the course of working with CPK models and helix lattice diagrams we have found DNA models for fd in which ridges and grooves of the DNA and protein structures would mesh. In contrast to Pf1 and Xf, the sense and pitch of the DNA and protein helices could differ. The models have two to three turns in 32 Å, again with phosphates in and bases out. The general type of structure is the same as shown in Figure 2B. Such structures with between 5 and 6 nucleotides per turn of each chain provide five grooves into which α-helical segments of the protein subunits fit. Figure 3 is provided to help make this idea clear.

We have also shown the feasibility of various structures with trial calculations using the linked atom least squares program (LALS) developed and described by Smith and Arnott (18), creating helical DNA structure models with two antiparallel chains of 15 Å pitch and 2.7 or 5.4 residues per turn. Backbone dihedral angles describing the particular conformations of two of them are listed in Table 1, simply as examples. Models with either right or lefthanded helical senses and with either C3'-exo or C3'-endo sugar conformations, all having low contacts between nonbonded

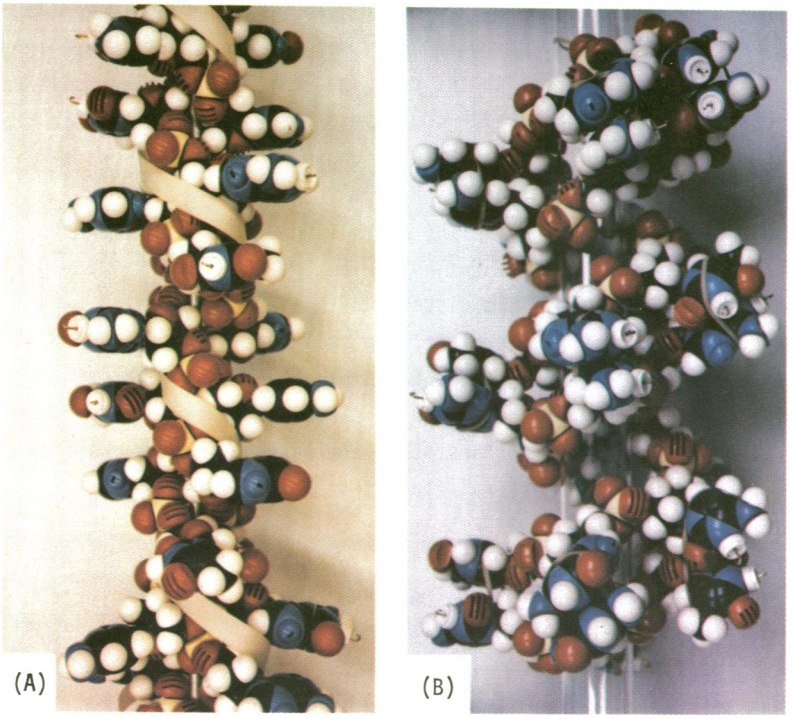

Figure 2. CPK models of a Pf1 type DNA structure (A) and an Xf type DNA structure (B), each wound from a circle of 30 nucleotides. The pitch of each helix is approximately 15 Å. Paper tape follows the Pf1 helix (A) and rubber bands hold the bases in stacks in the Xf helix (B).

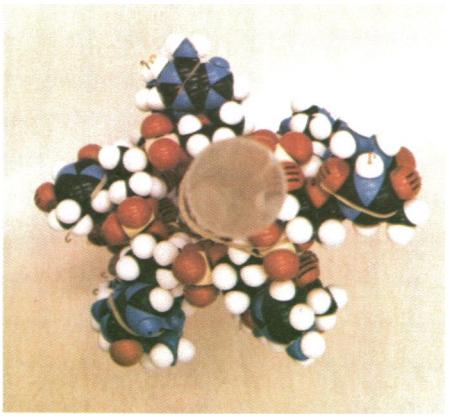

Figure 3. End of view of two turns of a helix of the type shown in side view in Figure 2B. The helix has been wound on a hollow plastic cylinder of 3 Å scale radius.

Table 1. Dihedral angles of two possible DNA structures with 15 Å pitch.

Axial nucleotide translation	ω (φ')	φ (ω')	ψ (ω)	θ (φ)	ξ (ψ)
5.55 Å (Pf1)	-139.8°	-36.6°	-156.3°	-108.5°	174.1°
2.78 Å (Xf)	-154.8°	-40.7°	-131.9°	-58.0°	-94.3°

Both helices have lefthanded sense, C3'-exo sugar conformations, and "anti" N-glycosidic bond rotations. Two common dihedral designations are used.

atoms, have been found using LALS, but at present we have no sure criteria for choosing among them.

The contributions of phosphate charges to the electrostatic energy for models having phosphates near the center deserve consideration. The net charge on each of these viruses at neutral pH is negative due to several acidic amino acids in the amino-terminal portions of the subunits which are probably farthest from the DNA. Amino acids which can bear positive charges are in the carboxy-terminal portions of the subunits (Figure 4), where they are available to neutralize both the negative charges of the phosphates and also the negative charges on the terminal carboxylate group. Figure 4 shows that there are enough positive charges in the carboxy-terminal region to accomplish this. The LALS DNA models for fd and Xf have the phosphates at substantially greater radii than the radii near 2.5 Å typically found for the Pf1 models, so the problem is most serious for Pf1. However, even with the phosphorus atom at a radius of 2.5 Å, the charge bearing oxygens of the phosphate group can be another 1 to 1.5 Å further out, reducing the problem somewhat. In Pf1, the primary amino on the long side-chain of lys_{45} could be very close to a phosphate group, since the suggested DNA

Pf1. (1) --ala$_{36}$-gly-leu$_{38}$-ile-tyr$_{40}$-ser-met-leu-arg$^{+}_{44}$-lys$^{+}_{46}$-ala-COO^{-}

Xf. (2) --ile$_{34}$-lys$^{+}_{36}$-val-tyr-lys$^{+}_{38}$-trp-val$_{40}$-arg$^{+}_{42}$-arg$^{+}_{44}$-ala-met-COO^{-}

fd. (2.3) --lys$^{+}_{40}$-leu-phe$_{42}$-lys$^{+}_{44}$-lys$^{+}_{46}$-phe-thr-ser-lys$^{+}_{48}$-ala-ser$_{50}$-COO^{-}

Figure 4. The carboxy-terminal portions of the subunits (13–15). Numbers of nucleotides/subunit are in parentheses. Aromatic residues are underlined.

structures are quite open. The Xf and fd type DNA structures are also open, allowing easy access to the phosphate from the outside. These considerations, of course, do not preclude the possibility of phosphate charge neutralization through metal ions.

SOME ABSORBANCE AND FLUORESCENCE PROPERTIES OF THE VIRUSES

We now present some absorbance and fluorescence spectra and examine them qualitatively in the context of the DNA models and amino acid sequences given above, since the spectra depend on interactions between the bases of the DNAs and the aromatic chromophores of the proteins. The interactions are mediated by energy transfer mechanisms, relative positions of the bases and chromophores, local solvent environments, and temperature. It is certainly not possible to proceed from such spectral data to well-determined virus structures. We show, instead, that the structure models offered above are consistent with the observed spectra.

The absorbance spectra (Figure 5) were used to calculate the apparent absorbance per nucleotide at 260 nm by subtracting estimated tyrosyl, tryptophyl, and phenylalanyl absorbance contributions, as well as light scattering contributions at this wavelength (Table 2). The molar nucleotide absorbance coefficient obtained by this simple analysis for Pfl DNA is higher than values for nucleotides in base-base stacked structures, whereas those obtained for Xf and fd DNAs are similar to values for base-base stacked structures.

In Figure 6 fluorescence emission spectra for each virus are compared

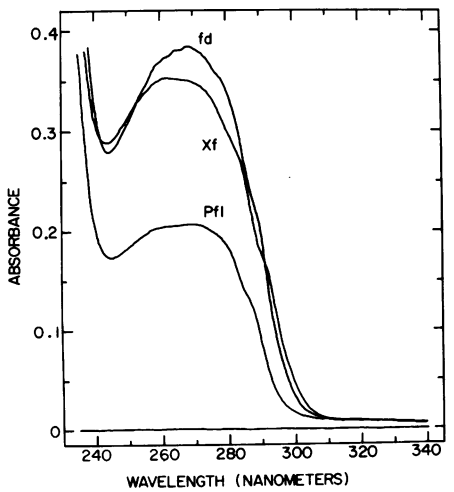

Figure 5. Absorbance spectra of virus solutions prepared to contain 1 mg/ml in 0.15 M KCl, 0.02M Na phosphate, pH 7 on the basis of available extinction coefficients (6,7,12). Optical cells of 1 mm light path were used.

Table 2. Extinction Coefficients at 260 nm

Virus	Nucleotides per subunit	Aromatic residues in each subunit	Molar ext. coeff. per subunit + nucleotides (observed)	Molar ext. coeff. per subunit (assigned)	Molar ext. coeff. per. nucleotide (calculated)[a]
Pf1	1.0	Tyr_{25},Tyr_{40}	9200	1200	8000
Xf	2.0	Tyr_{37} Trp_{39}	16600	4400	6100
fd	2.3	Tyr_{21},Tyr_{24} Trp_{26} Phe_{11},Phe_{42},Phe_{45}	20900	5400	6700

[a]Values calculated without light scattering corrections are 8900, 6500, and 7200, for Pf1, Xf and fd respectively. Handbook values were used for absorbancies at 260 nm for the aromatic protein groups.

with spectra for isolated protein subunits and spectra for model compounds. In Pf1 virus the only aromatic groups are the two tyrosines and one nucleotide per subunit. Absorbance studies have shown that deprotonation of one tyrosine per subunit is reversible and has a pK of 11.3 in 0.5 M KCl, but that deprotonation of more than one leads to irreversible disruption of the structure (7). We have found that about 90% of the tyrosyl fluorescence is characterized by a pK of 11.2 (Table 3) and conclude, from the correspon-

Table 3. Fluorometric Titration of Pf1 Virus

Preparation	Quench %	Apparent pK
1	88	11.1
2	88	11.2
3	83	11.2
4	90	11.2
1 (reverse)	89	11.2
N-acetyltyrosinamide	100	9.7

The fluorescence excitation was at 282 nm and emission was at 304 nm. Titrations were begun near neutral pH at 1.0×10^{-5} M tyrosine in 0.5 M KCl. Successive increments of KOH decreased the fluorescence. The reverse titration was begun at pH 11.9 in 0.5 M KCl and successive increments of HCl increased the fluorescence. The temperature was 26°C.

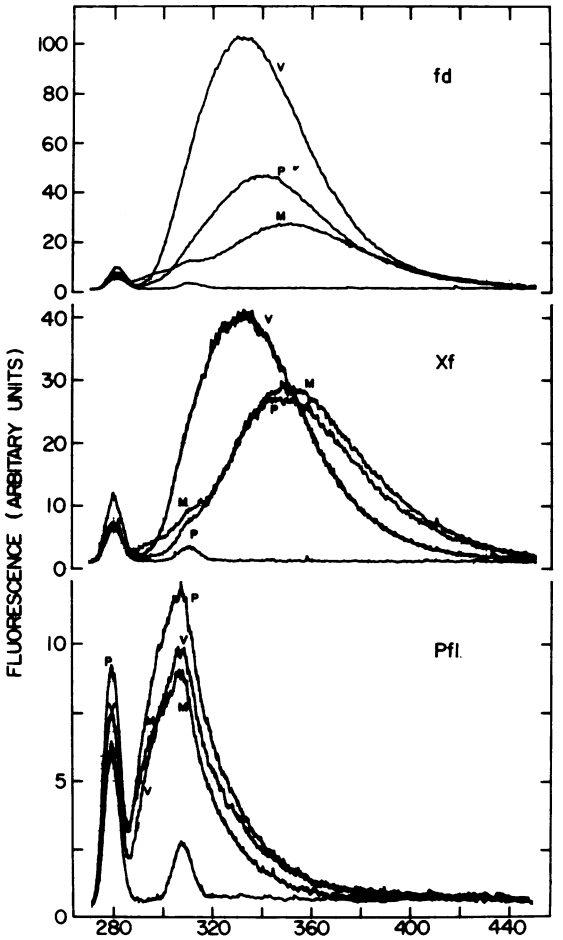

Figure 6. Fluorescence emission spectra of viruses (V) compared to spectra of their isolated protein subunits (P) or their constituent fluorophores in model compounds (M), L-tyrosyl-glycine and L-tryptophyl-glycine. Excitation was at 280nm. Viruses and model compounds were in 0.15 M NaCl, 0.02 M sodium phosphate, pH 7, and the isolated proteins were in this buffer containing between 0.1 and 0.4 M guanidine hydrochloride depending on dilution of stocks in 8 M GuHCl required. The concentration of subunits was $1X10^{-6}$ M for (V) and (P), and the concentrations of model compounds were equivalent to 10^{-6} M in subunits. The instrumental gain was different for each virus system as indicated by the ordinate scale. Rayleigh and Raman scattering peaks are at 280nm and 310nm, respectively. The source lamp and monochromator from a Cary 60 provided the incident radiation and a McPherson-Heath EU-700 monochromator and Hamamatsu R292 photomultiplier with appropriate electronics were used to record the uncorrected emission spectra.

dence in pK values, that 90% of the fluorescence derives from the one reversibly titratable tyrosine. Evidently its fluorescence is enhanced relative to model compounds whereas the fluorescence from the other tyrosine is quenched (Fig. 6). Enhancement can occur if a fluorophore is rigidly constrained in a nonpolar environment and quenching can occur through base-tyrosine stacking, through hydrogen bonding, or through the presence of a nearby anion (19-23). In Pf1, tyr_{40} is close enough to the DNA to interact directly with it. We think that base-tyrosine stacking in Pf1 is a good explanation of the high molar absorbance per nucleotide, the disruption of the virus on irreversible deprotonation of both tyrosines, and the low fluorescence from the one tyrosine, probably tyr_{40}, that is not reversibly titrated in 0.5 M KCl.

The fluorescence spectra of fd and Xf have maxima near 332 nm, typical of fluorescence from tryptophan in hydrophobic environments (20). Contributions to the spectra from tyrosines are too weak to be clearly evident, although energy transfer between tyrosines and tryptophans could occur in each virus. Based on a quantum yield of 0.13 for tryptophan (19), we estimate a high quantum yield of 0.31 for trp_{26} in each subunit in the fd virion and a quantum yield of 0.11 for trp_{39} in each subunit in the Xf virion. Since base-tryptophan stacking is expected to be accompanied by fluorescence quenching (22), not enhancement, and a nearby anion, such as phosphate, could also cause quenching, we conclude that there are no DNA-tryptophan interactions in fd. On the other hand trp_{39} and tyr_{37} in Xf probably lie near enough to the DNA to participate in direct interactions with it. This could explain the $\sim$3 fold difference in quantum yield between trp_{39} of Xf and trp_{25} of fd. In both viruses the low molar nucleotide absorbance at 260 nm indicates base-base stacking, and our DNA models for Xf and fd easily allow base-base stacking with the bases on the outside. There is sufficient space in the models (Figure 2B) to allow the DNA to interact with phe_{42} and phe_{45} in the case of fd, and to interact with tyr_{37} and trp_{39} in the case of Xf.

DISCUSSION

In Pf1 and Xf, DNA structure models with phosphates directed toward the axes follow, as necessary deductions, from straightforward interpretations of the Pf1 and Xf diffraction patterns in terms of helices of 15 Å pitch for the subunits, from DNA-protein helix matchings suggested by the integer numbers of nucleotides per subunit, and from physicochemical data which yield the

nucleotide axial translations. In the case of fd, DNA models with phosphates on the inside follow from our hypothesis that the DNA ridges and grooves mesh with the ridges and grooves of a five start helix of protein subunits. It is possible, of course, that future experiments will show the pitch of the DNA helix in one, or even all, of these viruses to be other than ∿15 Å. This could occur if the matches of DNA and protein structures are achieved other than as described here, or if the protein structures are significantly different than currently thought. Although the detailed interpretation of the X-ray fiber diffraction pattern is still in progress, the extraction of 15 Å pitches for the protein subunit helices in Pfl and Xf is straightforward (1-3) and the five start helix for fd seems well supported (3,10). Accepting the experimental parameters used above and placing our faith in a simple matching of DNA and protein structures, we propose that DNA structures of the two general types that we have described do in fact occur in filamentous viruses containing circular single-stranded DNA.

The models can be tested by comparisons with X-ray diffraction data when higher resolution is achieved. For example, our Pfl DNA models call for peaks of electron density of approximately 0.8 electrons per $Å^3$ at a radius of about 2.5 Å, somewhat higher than electron densities in this region estimated from diffraction patterns limited to 8 Å resolution (3). In addition, a consequence of alternate interactions in Pfl is the existence of two kinds of subunit, one interacting with the "up" chain and the other interacting with the "down" chain, whereas in Xf there need be only one kind of subunit interacting with two nucleotides, one nucleotide from each chain. The X-ray patterns of these two viruses should reflect this difference through the presence of intensity on half layer lines for Pfl but not for Xf. A consistent difference of this sort in the diffraction patterns would provide indirect support for the models. The models can be further tested and refined by direct determinations of DNA helix parameters by neutron diffraction using D_2O-H_2O contrast matching techniques. Also, ^{31}P nuclear magnetic resonance studies might ultimately give accurate phosphorous-phosphorous distances in all three structures. Finally, high resolution electron microscopy might reveal a central hole down the axes of Xf and fd, and, through suitable metal ion staining, helical parameters of the DNAs in all three viruses.

ACKNOWLEDGMENTS

We are grateful to Drs. Peter J. Campbell Smith and Struther Arnott for

providing us with LALS and to Dr. William Winter for discussions and for providing us with test data for the program. Financial support was through grant AI 09049 from the U.S. Public Health Service.

REFERENCES

1. Marvin, D. A., Wiseman, R. L. and Wachtel, E. J. (1974) J. Mol. Biol. 82, 121-138.
2. Wachtel, E. J., Marvin, F. J. and Marvin, D. A. (1976) J. Mol. Biol. 107, 379-383.
3. Makowski, L. and Caspar, D. L. D. (1978) in The Single Stranded DNA Phages (D. Denhardt, D. Dressler and D. S. Ray, editors), pp. 627-643, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
4. Marvin, D. A. (1966) J. Mol. Biol. 15, 8-17.
5. Marvin, D. A., Pigram, W. J., Wiseman, R. L., Wachtel, E. J. and Marvin, F. J. (1974) J. Mol. Biol. 88, 581-600.
6. Wiseman, R. L. and Day, L. A. (1977) J. Mol. Biol. 116, 607-611.
7. Day, L. A. and Wiseman, R. L. (1978) in The Single Stranded DNA Phages (D. Denhardt, D. Dressler and D. S. Ray, editors), pp. 605-625, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
8. Marvin, D. A. and Hohn, B. (1969) Bacteriol. Rev. 33, 172-209.
9. Frank, H. and Day, L. A. (1970) Virology 42, 144-154.
10. Newman, J., Swinney, H. L. and Day, L. A. (1977) J. Mol. Biol. 116, 593-603.
11. Beck, E., Sommer, R., Auerswald, E. A., Kurz, Ch., Zink, B., Osterburg, G. & Schaller, H., and Sugimoto, K., Sugisaki, H., Okamoto, T. & Takanami, M. (1978) Nucl. Acids. Res. 5, 4495-4503.
12. Berkowitz, S. A. and Day, L. A. (1976) J. Mol. Biol. 102, 531-547.
13. Nakashima, Y. and Konigsberg, W. (1974) J. Mol. Biol. 88, 598-600.
14. Nakashima, Y., Wiseman, R. L., Konigsberg, W. and Marvin, D. A. (1975) Nature 253, 68-71.
15. Frangione, B., Nakashima, Y., Konigsberg, W. and Wiseman, R. L. (1978) FEBS Letters 96, 381-384.
16. Franklin, R. F. and Gosling, R. G. (1953) Nature 171, 740-741.
17. Koltun, W. L. (1965) Biopolymers 3, 665-679.
18. Smith, P. J. C. and Arnott, S. (1978) Acta Cryst. A34, 3-11.
19. Longworth, J. W. (1971) in Excited States of Proteins and Nucleic Acids (R. F. Steiner and I. Weinryb, editors), pp. 319-484, Plenum Press, New York.
20. Burstein, E. A., Vedenkina, N. S. and Ivkova, M. N. (1973) Photochem. Photobiol. 18, 263-279.
21. Brun, F., Toulme, J. J. and Helene, C. (1975) Biochemistry 14, 558-563.
22. Mayer, R., Toulme, F., Montenay-Garestier, Th. and Helene, C. (1979) J. Biol. Chem. 254, 75-82.
23. Montenay-Garestier, Th. (1975) Photochem. Photobiol. 22, 3-6.