Volume 6 Number 12 1979 Nucleic Acids Research

# Atomic resolution analysis of a 2:1 complex of CpG and acridine orange

Andrew H.-J.Wang, Gary J.Quigley and Alexander Rich

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

Received 2 May 1979

ABSTRACT

Cytidylyl-3', 5'-guanosine and acridine orange crystallize in a highly-ordered triclinic lattice which diffracts X-rays to 0.85 Å resolution. The crystal structure has been solved and refined to a residual factor of 9.5%. The two dinucleoside phosphate molecules form an antiparallel double helix with the acridine orange intercalated between them. The two base pairs of the double helical fragment have a twist angle of 10° and it is found to have a C3' *endo*-(3', 5')-C2' *endo* mixed sugar puckering along the nucleotide backbone as has been observed for other simple intercalator complexes. Twenty-five water molecules have been located in the lattice together with a sodium ion. The intercalator double helical fragments form sheets which are held together by van der Waals interactions in one direction and hydrogen bonding interactions in the other. The crystal lattice contains aqueous channels in which sixteen water molecules are hydrogen bonded to the nucleotide, none to the intercalator, five water molecules are coordinated about the sodium ion and four water molecules bind solely to other water molecules. The bases in the base pairs have a dihedral angle of 7 to 8 degrees between them.

## INTRODUCTION

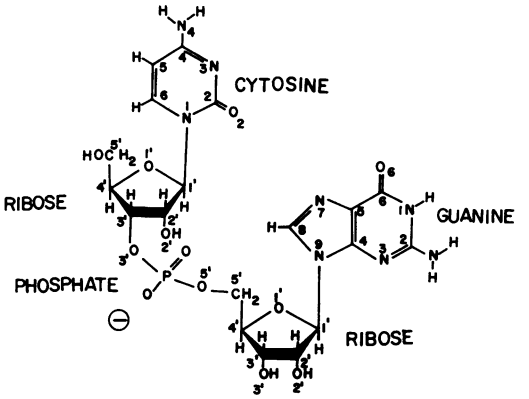
One of the important features of the double helical nucleic acids is their ability to accept planar molecules which can be inserted between the base pairs of the double helix. Many of these intercalators are mutagens or compounds that have other significant biological effects. It is thus of considerable interest to understand the detailed manner in which the intercalator is found within the confines of a double helical structure. The molecular structure of several intercalator compounds has been studied (1-7) and from this body of information we are beginning to understand some of the features of the geometry of intercalation into double helical nucleic acids. Here we report on the crystal structure of a 2:1 complex of cytidylyl-3',5'-

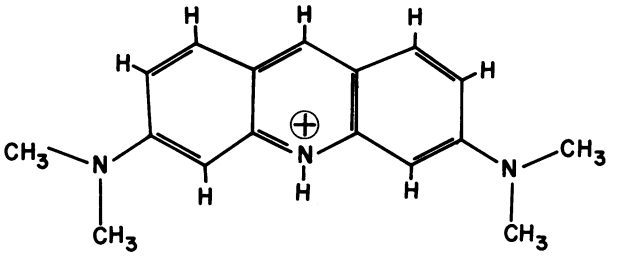
guanosine (CpG) which forms a double helical fragment containing one acridine orange molecule. Unlike most previous intercalator complexes this crystal complex diffracts to 0.8 Å resolution, the structure has been solved and highly refined so that we can determine with great accuracy many organizational details of the complex. This is the first intercalator complex which has two of the dinucleoside phosphate molecules for each intercalator molecule and as such it has a stacking geometry which is different from that seen in the other complexes. In this study, we are also able to resolve the position of 25 water molecules which form an aqueous channel in the crystal lattice. Because it is a 2:1 intercalator complex this is the first structure to have a metal ion which plays an important part in organizing the water molecules in the lattice. Further, the hydrophobic environment of the acridine orange intercalator can be clearly seen both in the interior of the double helical fragment and in the region abutting the aqueous domain. The molecular formula of CpG and acridine orange is shown in Figure 1.

## METHODS

Equimolar amounts of acridine orange (Sigma) and CpG (Collaborative Research) were placed into the depression of a spot plate and sealed with an environment containing 75% 2-methyl-2, 4-pentanediol. The technique was similar to that which has been described previously for forming tRNA crystals. The crystallization solution was set up by Dr. Frances Jurnak. Approximately a year later orange-colored parallelepiped shaped crystals were observed in the dip submerged in a slightly orange-colored mother liquid. A small crystal was washed, then dissolved and its ultra-violet spectrum was measured. Examination of the UV spectrum revealed that both components were present in the crystal. A crystal, measuring 0.2 x 0.2 x 0.4 mm was mounted in a glass capillary and sealed with a droplet of mother liquid inside. Precession camera photography indicated the molecule has crystallized in the triclinic lattice with space group $P1$, $a$ = 12.646 Å, $b$ = 11.792 Å, $c$ = 15.742 Å, $\alpha$ = 92.69°, $\beta$ = 107.02° and $\gamma$ = 90.05°. Three-dimensional X-ray diffraction data were collected on a Picker FACS-1 diffractometer, 6202 reflections




CYTIDYLYL–(3',5')–GUANOSINE



ACRIDINE ORANGE

Figure 1: The molecular formula of CpG and acridine orange.

were registered which were 2 sigma above background, to a resolution of 0.85 Å. The data were corrected for Lorentz and polarization factors as well as for absorption. A Patterson map was calculated which revealed a series of planes stacking at a spacing of 3.4 Å which strongly suggested that the planar intercalator molecule was parallel to the bases. The map also showed the direction of the phosphorus-phosphorus vector. A rotation function search was carried out using model coordinates from the structure of the deoxy CpG-platinum intercalator complex which has been solved recently (1). The program was written by one of us (G.J. Quigley, unpublished result). In the initial stages of the rotation search a 2:2 model was used in which an intercalator molecule was placed inside the double helical fragment as well

as outside it. This yielded a correlation coefficient of 75%. Using this solution, a restrained refinement was carried out to 3 Å resolution using a program of Wayne Hendrikson. An R factor of ~30% was obtained, following which a $2|F_o| - |F_c|$ Fourier map was calculated. This revealed that the complex actually was a 2:1 complex. The model was then readjusted by eliminating the extra intercalator following which successive cycles of full-matrix least squares were carried out, leading ultimately to a residual factor of 9.5%. In the course of this refinement, twenty-four single site water molecules were found and one disordered water molecule was found which occupies two closely adjacent sites about 1 Å apart. A sodium ion was located in the course of the refinement. All atoms were refined anisotropically except for the disordered water molecule. Hydrogen atoms were not included in the refinement.

## RESULTS

The intercalator complex is shown in Figure 2 viewing the molecule almost parallel to the bases. It can be seen that although the nucleotide has a pseudo two-fold axis, the acridine orange is displaced from that axis so that it is closer to one

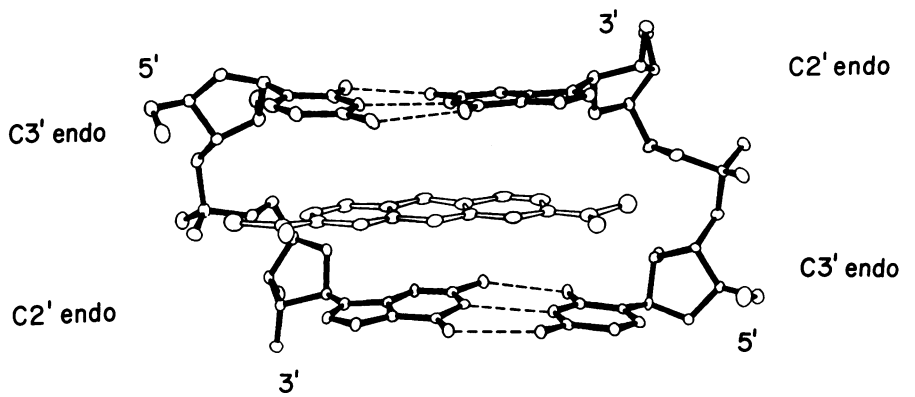


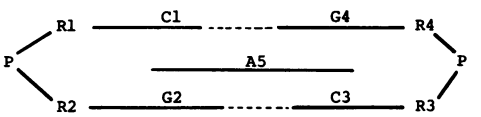

Figure 2: The structure of the 2:1 complex of CpG and acridine orange as viewed almost along the pseudo diad axis in the molecule. The atoms are drawn as anisotropic thermal ellipsoids. The guanine and cytosine residues are hydrogen bonded together in a Watson-Crick base pair.

nucleotide strand than the other. As shown in the Figure, the nucleotide at the 3'-end of the strand has the unusual C2'-*endo* conformation which has been found in all of the simple intercalator complexes. The nucleotide at the 5'-end has the normal C3'-*endo* conformation of ribonucleotides. This view also shows that the bases are close to being co-planar but are not absolutely co-planar. The upper GC base pair in Figure 2 has a dihedral angle of 7° between the two bases while the lower GC base pair has a dihedral angle of 8°. The plane of the acridine orange ring is virtually parallel to the mean plane passing through the two GC base pairs on either side, deviating by only 2° and 3° from the upper and lower planes respectively. Table I lists the dihedral angles between different bases and acridine orange.

A view of the intercalator complex normal to the base pair plane is shown in Figure 3. Here it can be seen that the intercalator has an asymmetric position as mentioned above. Because of this asymmetric position, the intercalator has very good stacking between the cytosine and guanine base on one side of the double helical fragment but much less stacking on the other

TABLE I

Dihedral Angles between Bases and Acridine Orange




| Plane | Plane | Dihedral Angle (degrees) |
|---|---|---|
| C1 | G2 | 5.9 |
| C1 | C3 | 1.1 |
| C1 | G4 | 7.9 |
| C1 | A5 | 3.9 |
| G2 | C3 | 7.0 |
| G2 | G4 | 6.6 |
| G2 | A5 | 3.9 |
| C3 | G4 | 8.3 |
| C3 | A5 | 4.7 |
| G4 | A5 | 4.2 |
| C1, G4 | G2, C3 | 4.5 |
| C1, G4 | A5 | 1.9 |
| G2, C3 | A5 | 2.7 |

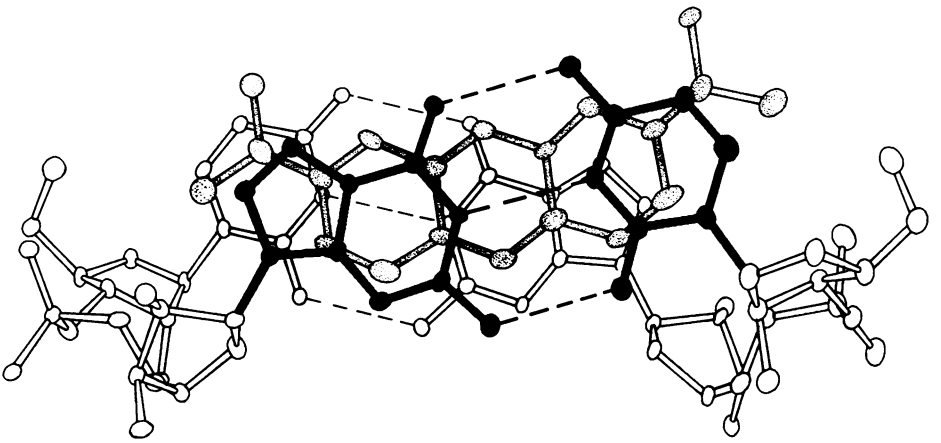


Figure 3: A view of the complex perpendicular to the base pair. The upper base pair is shown in solid black while the acridine orange is shown stippled. The lower base pair is drawn with open bars between the atoms. The size and shape of the atoms represent anisotropic thermal ellipsoids. It can be seen that the acridine orange stacks well on one side of the double helical fragment but poorly on the other where the dimethyl amino group is located between the two bases.

side. Indeed there is very little stacking overlap between the cytosine ring on one side and the outer ring of the acridine orange. Because of this asymmetic rotation, the dimethylamino group protrudes out away from the helix on one side but is largely confined within the helix on the other.

In both Figures 2 and 3 the size and shape of the atoms are represented by thermal ellipsoids. In general it can be seen that the atoms in the stacked region of the molecule have smaller parameters compared to those on the peripheral of the helix.

The intercalator complexes are stacked together to form large sheet-like structures. A view perpendicular to this sheet-like structure is shown in Figure 4. In the vertical direction, the sheets are held together by stacking interactions in which a cytosine base of one complex stacks on a cytosine base of the complex immediately above. In addition, there is an interesting interaction between the ribose O1' and the guanine residue in the adjoining stack. In the horizonal direction, the nucleotide fragments are held together by three hydrogen bonds connecting

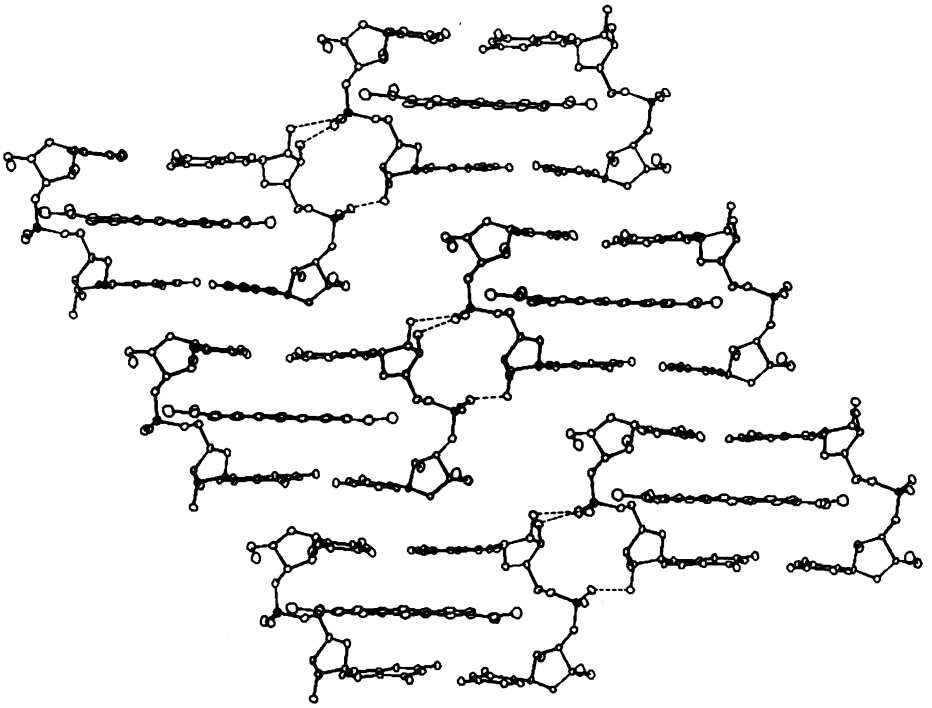


Figure 4: A view of the stacking of helical intercalator complexes to form a sheet-like structure is illustrated here. The structure is viewed along the pseudo diad axis. In the vertical direction the interaction is largely due to stacking between cytosine bases from adjacent helical fragments and an interaction between a guanine base and the ribose O1'. In the horizontal direction three hydrogen bonds are shown which hold the molecules together. The other helical fragments are staggered in such a way that the phosphate groups are separated from each other. The phosphorus atoms are shaded black.

the phosphate-oxygen of one residue with O2' and O3' of the guanosine ribose residue, and an additional interaction between a phosphate-oxygen and an O3' hydroxyl of a second guanosine residue. This type of phosphate-hydroxyl hydrogen bonding interaction occurs seven times in the structure of yeast phenylalanine tRNA (8), and appears to be important in maintaining the tertiary structure of the molecules. In the organization of these sheet structures the nucleotide fragments are staggered relative to each other in such a manner that there is a significant separa-

tion between the phosphate groups on either side of the two columns of intercalator fragments. It is interesting that the positively-charged sodium ion is not found near the negatively-charged phosphate groups. Instead it is coordinated to the O2' hydroxyl of the cytidine residue.

An end view of the packing of these sheets of helical intercalator complexes is shown in Figure 5. We can see the edge of

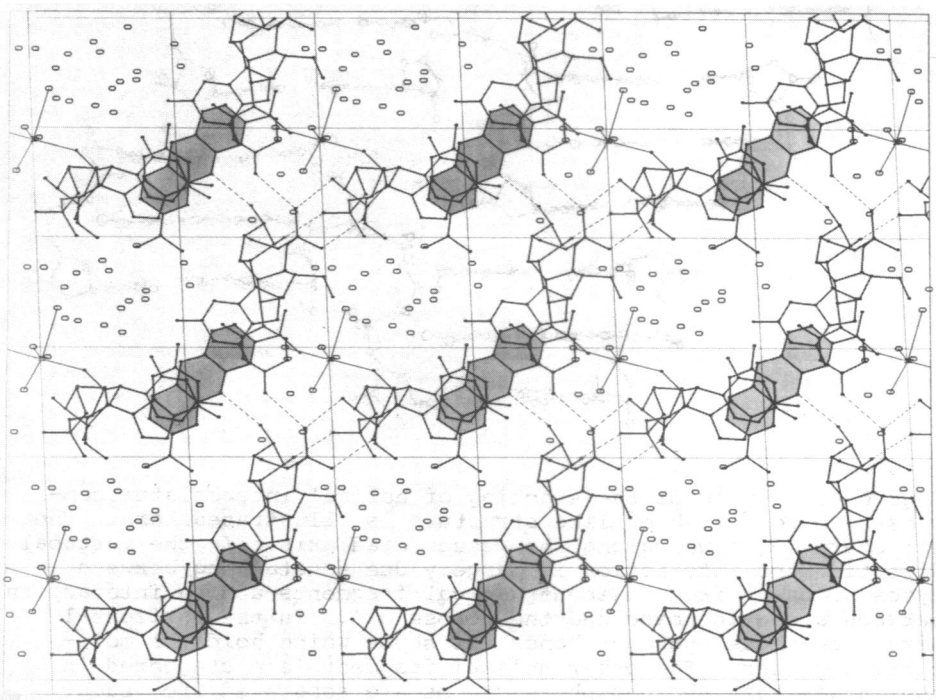


Figure 5: An end view of the structure which shows the way in which the sheets of double helical fragments are organized to form layers. The sheets are running from lower left to upper right; and this represents an end view of the sheet shown in frontal view in Figure 4. It can be seen that two hydrogen bonds connect adjacent sheets. The sheets have somewhat irregular surfaces which in effect enclose the aqueous domains. Water molecules are indicated as small circles. The sodium atom is indicated as a letter M and light lines indicate its coordination. Where the dimethylamino group of acridine orange projects into the aqueous domain, there are no water molecules within 3.5 Å of the methyl group. The grid of light lines indicates the unit cell in the bc projection. The three rings of the acridine orange molecule are shaded for clarity.

the sheet-like intercalator complex stretching diagonally across the Figure. The successive sheets are held together by a hydrogen bond between the cytosine N4 and a phosphate oxygen and a hydrogen bond between the acridine nitrogen atom and O3' of guanosine.

In between the segments where the sheets are held together there are regions in which the sheets bend away from each other in such a way that they form a pocket of water molecules. Twenty-six different water positions are shown together with the sodium ion which is indicated as a letter M. These water molecules form an aqueous channel which extends through the entire crystal and has a diameter of 5 - 6 Ångstroms. Sixteen of the water molecules are hydrogen bonding to the nucleotide, but none of the water molecules is hydrogen bonded to the intercalator. It is interesting to note that where the dimethylamino group of an intercalator protrudes into the aqueous region, there is a considerable hole in the aqueous phase. No water molecule is seen approaching the methyl group in less than 3.5 Å. This is a good example of the ordering of water molecules which occurs when hydrophobic groups intrude into the aqueous phase. The intercalator molecule thus exists almost entirely in a hydrophobic environment, largely tucked in between the two base pairs where it interacts with the $\pi$ electrons of the base pairs. Four water molecules are found inside the aqueous domain where they are hydrogen bonded only with other water molecules. Water constitutes approximately 20% of the crystal contents and it occupies a somewhat irregular shape. It is interesting that the water domain maintains this irregular shape by forming a number of water bonds which are other than the normal tetrahedral interactions that are found in water of the 'ice-like' structures. It is not that the water structure is disordered here, but rather the bond angles are distorted relative to what is seen in larger hydrated regions. An important organizing element in the aqueous domain is the sodium ion which is complexed to six oxygen atoms forming a slightly distorted octahedral complex. Five of these atoms are from water molecules and the sixth is from the O2' hydroxyl group of the cytidine residue.

## DISCUSSION

The bond angles and distances in the dinucleoside phosphate of the acridine orange complex are normal. Because of the highly refined nature of the solution it is worth examining the torsion angles which are listed in Table II in comparison with torsion angles reported from a number of other intercalator complexes. From the data in Table II, it is clear that there is a fair amount of conformational flexibility in the nucleotide backbone among different intercalator complexes.

Acridine orange is an example of an intercalator which slips into a double helix without hydrogen bonding to it. Among these simple intercalators are included ellipticine (5), 9-amino acridine (4), ethidium (2, 3) and the terpyridine platinum intercalator (1). All of these are characterized by a C3'-*endo* pucker at the 5' end and the C2'-*endo* pucker at the 3' end. This generalization appears valid even though one of the intercalator complexes is that of a DNA fragment while the others, including the present study, are RNA fragments. The only exception to this generalization is found in the more complex intercalator proflavine (6,7) which actually hydrogen bonds to the phosphate group and is associated with C3'-*endo* conformation at both 3' and 5' ends.

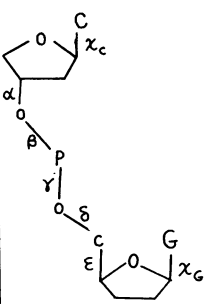
The twist angle between the upper and lower GC base pairs is 10°. This indicates an unwinding angle of 23° for an 11-fold RNA double helix. These results are quite similar to those which have been obtained with previous ribonucleotide intercalator complexes.

Although several intercalator structures have been solved,

TABLE II

Backbone and Glycosidic Torsion Angles for Intercalated Nucleic Acid Fragments

| | α | β | γ | δ | ε | $\chi_C$ | $\chi_G$ | Reference |
|---|---|---|---|---|---|---|---|---|
| CpG-Acridine Orange | 211 | 301 | 288 | 237 | 50 | 9 | 105 | This work |
| | 225 | 298 | 297 | 226 | 40 | 8 | 115 | |
| dCpG-Pt (terpyridine) | 201 | 287 | 282 | 226 | 57 | 32 | 114 | (1) |
| | 194 | 292 | 308 | 217 | 84 | 34 | 117 | |
| iodo CpG-Ethidium | 226 | 281 | 286 | 210 | 72 | 29 | 101 | (3) |
| | 225 | 291 | 291 | 224 | 55 | 24 | 109 | |
| iodo UpA-Ethidium | 207 | 286 | 291 | 236 | 52 | 26 | 99 | (2) |
| | 218 | 302 | 276 | 230 | 70 | 14 | 100 | |
| CpG-Proflavine | 201 | 290 | 289 | 231 | 52 | 17 | 85 | (13) |



very few of them form such highly organized crystal structures that data can be collected to this resolution and refined to such a high degree of certainty. Because of this, the present study reveals several aspects of intercalator structures which are of great interest. This includes the special role played by the solvent molecules surrounding the intercalator complex. It can be seen that the water molecules tend to form ordered arrays with normal tetrahedral bonding where it is feasible but in environments in which the shape is odd, the bonding is not tetrahedral and the water molecules are packed in a somewhat irregular manner. This is the only intercalator complex which has a 2:1 stoichiometry and therefore a metal ion is present. It is interesting that the metal ion is not found complexed to the phosphate group which one might expect on simple electrostatic grounds. Instead, it is bound to an O2' hydroxyl group. In the structure of the double helical fragment ApU (9), a sodium ion was found bound to the carbonyl groups of uracil as well as to the phosphate groups.

It should be noted that there have now been two different structures solved for each of three different intercalators, ethidium (2,3), acridine orange (6) and proflavine (6,7). It is interesting that the manner in which each of these intercalates into a double helical fragment is similar even though the double helical fragment is different. The present structure with acridine orange seems very similar to that reported previously in its complex with iodo CpG (6). These intercalators each have characteristic modes of interacting with the dinucleoside phosphate which differ in a significant way one from the other. For example, the ethidium intercalation is a symmetric one whereas the acridine orange one is asymmetric. The proflavine intercalation is symmetric but it is associated with hydrogen bonding to both phosphate groups on either side.

Because this intercalator complex has a 2:1 stoichiometry the nucleotide base pairs are stacking upon each other. A significant interaction in the present stacking is one in which the ribose oxygen O1' is stacked directly on top of a guanine ring. A similar stacking interaction was found in the structure of the double helical fragments ApU (9) and GpC (10) as well as in the non-helical RNA fragment of UpA (11,12). This type of interaction is also seen in

ribose residue R59 in yeast phenylalanine tRNA where the O1' oxygen is stacked onto adjacent base C48 (8). This suggests that in addition to the well-known base-base stacking stabilization there may be another somewhat weaker but nonetheless significant, interaction associated with the ribose O1' atoms and the unsaturated π electron cloud of purine or pyrimidine rings.

Considerable interest has been associated with the intercalation of planar molecules into the double helix. In the present detailed study we have seen some features which may be of importance in understanding the chemistry of the interaction and the role which they play in biological systems.

ACKNOWLEDGEMENTS

We thank Dr. Frances Jurnak for generously making available to us the results of the crystallization set-up. This research was supported by grants from the N.I.H., N.S.F., N.A.S.A. and The American Cancer Society. A.H.J. Wang is partially supported by the M.I.T. Center for Cancer Research (Grant CA 14051).

REFERENCES

1. Wang, A.H.J., Nathans, J., van der Marel, G., van Boom, J.H. and Rich, A. (1978) Nature 276, 471-474.
2. Tsai, C.C., Jain, S.C. and Sobell, H.M. (1977) J. Mol. Biol. 114, 301-315.
3. Jain, S.C., Tsai, C.C. and Sobell, H.M. (1977) J. Mol. Biol. 114, 317-331.
4. Sakore, T.D., Reddy, B.S. and Sobell, H.M. (1979) J. Mol. Biol. (in press).
5. Jain, S.C., Bhandary, K.K. and Sobell, H.M. (1979) J. Mol. Biol. (in press).
6. Reddy, B.S., Seshadri, T.P., Sakore, T.D. and Sobell, H.M. (1979) J. Mol. Biol. (in press).
7. Neidle, S., et al. (1977) Nature 269, 304-307.
8. Quigley, G.J. and Rich, A. (1976) Science 194, 796-806.
9. Seeman, N.C., Rosenberg, J.M., Suddath, F.L., Kim, J.J.P. and Rich, A. (1976) J. Mol. Biol. 104, 109-144.
10. Rosenberg, J.M., Seeman, N.C., Day, R.O. and Rich, A. (1976) J. Mol. Biol. 104, 145-167.
11. Sussman, J.L., Seeman, N.C., Kim, S.H. and Berman, H.M. (1972) J. Mol. Biol. 66, 403-421.
12. Rubin, J., Brennan, T. and Sundaralingam, M. (1972) Biochem. 11, 3112-3129.
13. Berman, H.M., Neidle, S. and Stodola, R.K. (1978) Proc. Nat. Acad. Sci. 75, 828-832.