[2] Freyd, P. J., thesis: Functor Theory, Princeton University, 1960.

[3] Kan, D. M., "Adjoint functors," *Trans. Amer. Math. Soc.*, **87**, 294–329 (1958).

[4] Lawvere, F. W., "The convolution ring of a small category," *Notices Amer. Math. Soc.*, **10**, 280 (1963); Errata, *Notices Amer. Math. Soc.*, **10**, 516 (1963).

THE CRYSTAL STRUCTURE OF AN INTERMOLECULAR NUCLEOSIDE COMPLEX: ADENOSINE AND 5-BROMOURIDINE*

BY A. E. V. HASCHEMEYER† AND HENRY M. SOBELL‡

DEPARTMENT OF BIOLOGY, MASSACHUSETTS INSTITUTE OF TECHNOLOGY

The concept of hydrogen-bonding specificity between the purine and pyrimidine bases, adenine and thymine, guanine and cytosine, is fundamental in the present theory of nucleic acid structure and replication. The base-pairing scheme proposed by Watson and Crick[1] in their structural hypothesis for the DNA molecule has gained wide acceptance among biologists and recently has received strong support from three-dimensional Fourier analysis of DNA fiber X-ray diffraction data,[2] and from structure analysis of single crystals containing guanine and cytosine derivatives in an intermolecular complex.[3, 4] On the other hand, Hoogsteen[5] has found a different base-pairing configuration in a crystalline complex of 9-methyl adenine and 1-methyl thymine. Here, the ring nitrogen N3 of thymine hydrogen-bonds to the imidazole nitrogen N7 of adenine instead of bonding to N1 as in the Watson-Crick model. A similar pairing configuration has recently been found in a crystalline complex containing 9-ethyl adenine and 1-methyl uracil.[6] This pairing is of considerable interest since it is thought to occur in the triple-stranded 2:1 complex of polyuridylic acid and polyadenylic acid.[7]

The information derived from these structure investigations has prompted us to investigate other possibilities for cocrystallization of important compounds known to interact in biological systems. The present work describes a single crystal analysis of a nucleoside intermolecular complex between adenosine and 5-bromo-uridine. The presence of the sugars on the purine and pyrimidine bases brings this model system close to the biological systems of interest. Furthermore, the bromine-substituted derivative is of particular interest since the closely related molecule bromodeoxyuridine is a well-known mutagenic agent. The results show the existence of a third type of base-pair configuration.

Methods.—The adenosine-5-bromouridine complex was crystallized in clusters of thin needles by slow evaporation from an aqueous solution containing equimolar quantities of these compounds. Ultraviolet absorption measurements on aqueous solutions made from single crystals confirmed the presence of the two nucleosides in approximately equal proportions. The crystals were found to be orthorhombic, space group $P2_12_12_1$ with $a = 4.80 \pm 0.01$, $b = 15.19 \pm 0.01$, and $c = 31.76 \pm 0.03$ Å; the density determined in benzene-methyliodide solutions was 1.706 ± 0.010 gm/cc. The unit cell contains four asymmetric units, each consisting of an adenosine-bromouridine pair and a water molecule. Equi-inclination Weissenberg photographs were taken about the a axis with filtered CuKα radiation, using the multiple film technique. The intensities were estimated visually and corrected with the appropriate Lorentz-polarization factors. A total of 2,511 reflections were indexed, of which 2,015 were nonzero, representing about 90 per cent of the data accessible in the copper sphere. No correction was made for absorption effects.

Determination of the structure: The bromine atom position was determined from the three-dimensional Patterson function. The heavy atom method was first applied to the centrosymmetric (0kl) projection yielding a trial structure for the bromouridine molecule. Successive Fourier syntheses revealed the complete two-dimensional structure, as shown in Figure 1. A three-dimensional Fourier based on the bromine phases was then easily interpretable, and most of the atomic positions could be determined unambiguously. A second three-dimensional Fourier based on this structure established the complete structure of the nucleoside pair and revealed an extra peak of low electron density. This was tentatively identified as a water molecule.

The structure was refined by the method of least squares using a modified version of the full matrix least squares program for the IBM 7090 computer devised by Busing, Martin, and Levy.[8] The presence of a water molecule associated with the asymmetric unit was confirmed during the course of refinement; variation of the atom multiplier led to a final value of 0.6 water molecules per nucleoside pair, in good agreement with the value of 0.4 obtained from the crystal density. Successive cycles of individual atom isotropic refinement has reduced the over-all residual factor to 12 per cent. The standard deviation for light atom bond lengths is ±0.03 Å based on the standard errors of atomic coordinates calculated from the full normal matrix.[8]

Description of the Structure.—Figure 2 shows the crystal structure viewed down the *a* axis. Presumed hydrogen-bonding contacts are indicated by dashed lines, and the distances are given. Dotted lines indicate other distances of interest. The existence of a strong hydrogen bond between the bases is clearly indicated between N3 of bromouridine and N7 of adenosine with a length of 2.80 Å. A second weak hydrogen bond appears to occur between O2 of bromouridine and N6 of adenosine. The C6-N6-O2 angle of 126° is favorable for hydrogen bonding; however, the distance of 3.10 Å is too long. It is possible that the bond is strained by the interactions of the bromouridine O4 and adenosine O5′ with an adjacent sugar or that the proximity of the water, 2.90Å from O2 of uracil and 2.99 Å from N6 of adenine, perturbs the base pairing.
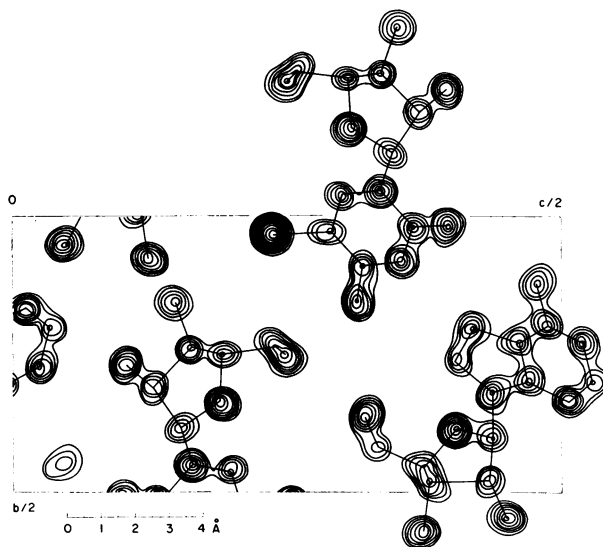

FIG. 1.—(0kl) Fourier projection of the adenosine-5-bromouridine crystal structure. Contours are drawn at equal but arbitrary levels of electron density. The bromine atom contours are drawn for every fourth level of electron density. The extra peak shown in the lower left-hand corner is a water molecule; its low level of electron density is attributed to incomplete occupancy in the crystal lattice.
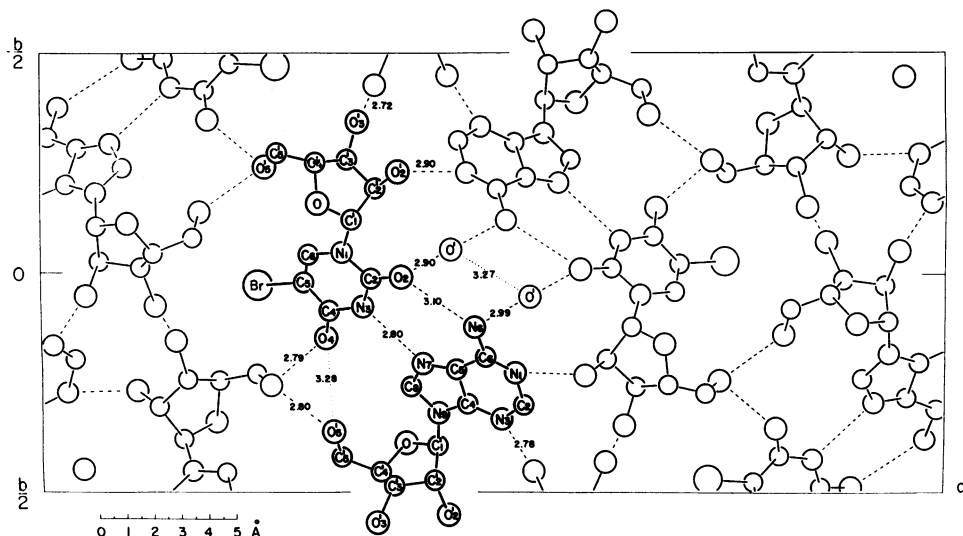
FIG. 2.—Packing diagram of the adenosine-5-bromouridine crystal structure viewed down the *a* axis. Presumed hydrogen-bond contacts are drawn with dashed lines and the distances shown. Dotted lines indicate other contacts of interest.

All hydroxyl groups of the two ribose rings are involved in hydrogen bonding with each other and with the remaining electron donors of the two bases, N1 and N3 of adenine, and O4 of uracil. These bonds are important in maintaining the integrity of the crystal structure along the *b* and *c* axes. The O—H· · ·O distances observed are 2.72, 2.79, and 2.80 Å; the O—H· · ·N distances are 2.78 and 2.90 Å. The bromine atom does not appear to participate in hydrogen bonding; the closest approach is 3.61 Å from C5′ and O5′ of a neighboring adenosine molecule.

Covalent bond distances and angles have been computed and are in good agreement with those found in crystals of adenosine-5′-phosphate,[9] calcium thymidylate,[10] and fluorodeoxyuridine.[11] These results will be reported in detail elsewhere. Least squares planes were determined for the atoms of the bromouracil ring, including the sugar carbon C1′, and for the adenine ring system with its sugar carbon C1′. Both gave excellent fits to least squares planes with maximum deviations of 0.025 Å in the former case and 0.035 Å in the latter with the exception of adenosine C1′ which deviates by 0.05 Å from the adenine plane.

The two ring systems were found to be almost, but not completely, coplanar, with a dihedral angle of 4° 55′ between their least squares planes. The amino nitrogen of adenine lies 0.03 Å above the adenine plane and 0.04 Å below the bromouracil plane. Both ribose sugars are puckered with carbon C3′, deviating by approximately 0.6 Å from the least squares plane formed by the other four members of the ring, as has been observed in other crystals.[9, 10] The angle between the plane of the sugar and that of the base is about 70° in each case; the distance between the glycosidic carbons C1′ of the two nucleosides is 9.65 Å.

The water molecule associated with the asymmetric unit is located near the two-fold rotation axis and is indicated by O′ in Figure 2. It appears to interact weakly with O2 of uracil and N6 of adenine, as indicated by the dashed lines. With
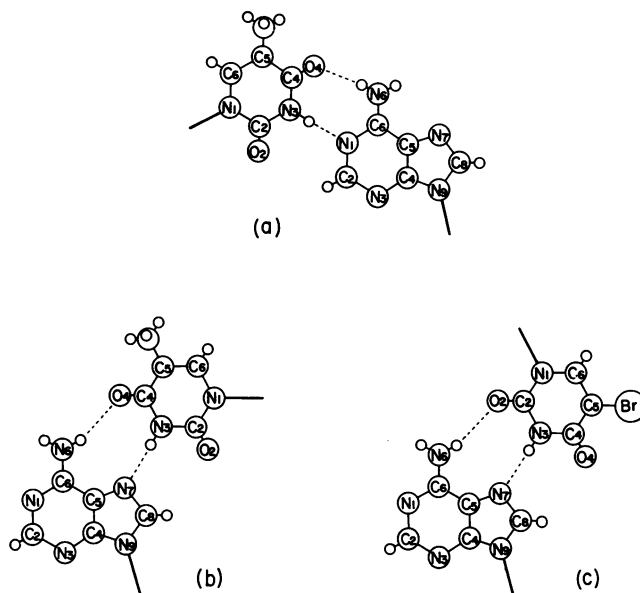
Fig. 3.—Schematic diagram of various base pairings between adenine and thymine. (a) The Watson-Crick pairing hypothesized for DNA. (b) Pairing found by Hoogsteen between 9-methyl adenine and 1-methyl thymine. (c) Pairing found in the present structure determination of the adenosine-5-bromouridine complex.

respect to the bases with which it makes closest contact, it lies 0.07 Å above the adenine plane and 0.24 Å above the bromouracil plane.

Discussion.—An important feature observed in the crystal complexes of nucleic acid components studied thus far is the hydrogen bonding between the purine and pyrimidine bases with resultant formation of a highly planar complex.[3-6] This is illustrated convincingly in the present structure, as shown in Figure 2. Here, in spite of the complexity of the over-all crystal structure, the base pair itself is almost planar, having a dihedral angle of 4°55' between the least squares planes of the individual bases. It is for this reason that we consider the hydrogen-bonded base pair, as observed in the crystal structure, to be an intermolecular complex, which may exist in solution prior to crystallization. Purine and pyrimidine compounds are found to associate extensively in aqueous solutions, although the actual mechanism is unknown.[12]

Figure 3 shows the base pairing arrangement believed to exist in DNA,[1] the pairing found by Hoogsteen,[5] and that found in the present investigation. In both the Watson-Crick and the Hoogsteen pairings the carbonyl oxygen O4 of thymine hydrogen-bonds to the amino group N6 of adenine. In the adenosine-bromouridine crystal complex, however, the amino group of adenine appears to hydrogen-bond to the carbonyl oxygen O2 of bromouracil, although the distance 3.10 Å is rather long, indicating that interaction between these groups is weak. The carbonyl oxygen O4 hydrogen-bonds to a neighboring sugar hydroxyl group.

We are particularly interested in the difference in hydrogen bonding of the carbonyl oxygens of uracil in our crystal compared with that in the Watson-Crick and the Hoogsteen structures. Is this the result of crystal packing forces and

sugar hydroxyl group hydrogen bonding, or does it reflect an alteration in the physical properties of the uracil ring due to the presence of the bromine atom? Another important question is whether or not the base-pair configuration found in the crystal structure truly reflects interactions which exist in solution. If this is so, then the possible influence of the bromine atom should be considered.

The effect of halogen substitution in fully conjugated systems has been studied extensively.[13] Observations on the chemical and physical properties of chlorobenzene, for example, have indicated the importance of resonance effects involving the contribution of a p electron from the halogen atom to the π electron system of the ring. This would tend to increase electronegativity at the ortho- and para-positions. On the other hand, inductive effects, which become more important in systems only partially conjugated, act in the opposite direction. In the case of uracil one might expect the bromine atom to exert primarily an inductive effect on the ring; this would tend to decrease electronegativity of the ortho-carbonyl O4 relative to the para-carbonyl O2. This may explain why the amino group of adenine bonds to O2 rather than to O4 of 5-bromouracil.

A powerful control study, on a complex of adenosine and thymidine, has not been possible as yet, because we have been unable to cocrystallize these compounds. Comparison of a complex of 9-ethyl adenine and 1-methyl-5-bromouracil currently under investigation[14] with the Hoogsteen crystal should be valuable; preliminary results indicate the possibility of the same base-pairing as in adenosine-bromo-uridine. We are inclined to be cautious, however, about making direct comparison of the nucleoside complexes with these systems, as well as with DNA, because of the great differences in their properties in solution in addition to the crystal-packing factors. For example, hydrophobic interactions may play a larger role in determining the base-pair configuration in solution in a system like the Hoogsteen complex, possibly masking effects that are important in nucleoside complexes. The actual occurrence of a water molecule in the adenosine-bromouridine crystal structure and its location provide additional reason to be concerned about the possible role of water in the hydrogen bonding between the bases.

It is of considerable interest that the present structure shows a strong hydrogen bond between N3 of bromouracil and N7 of adenine. The closeness of the contact (2.80 Å) compared with the usual N—H$\cdots$N bond distance of about 2.95 Å[13] suggests that other forces between the bases such as dipole-dipole interactions may be important. This bond also occurs in the adenine-thymine and the adenine-uracil complexes previously studied. It is not clear why it should be favored over the Watson-Crick bonding between uracil N3 and adenine N1. In the present crystal structure N1 is observed to interact with an adjacent sugar hydroxyl group, although in the other reported crystal complexes it shows no hydrogen bonding.

Another interesting feature of the adenosine-bromouridine structure is the presence of a water molecule near the carbonyl oxygen O2 of uracil and the amino nitrogen N6 of adenine. Least squares refinement of the structure has established that statistically there is only 0.6 of a water molecule at this site, indicating incomplete occupancy in the crystal lattice. This is not entirely surprising considering the lack of strong hydrogen bonding by the water molecule. However, the presence of the water molecule, together with the effects of hydrogen bonding by the sugar

hydroxyl groups, may be important in causing distortion of the base-pair configuration and lengthening of the O2-N6 hydrogen-bonding contact.

The results of this structure determination have shown the existence of a different type of base pairing between the adenine and uracil rings in the nucleosides. Whether or not this pairing is in any way related to the phenomenon of mutation is an intriguing question. Preliminary attempts at model building have indicated that it would be difficult, if not impossible, to incorporate this base-pairing configuration into the DNA molecule as it is currently conceived. Further work is proceeding on other nucleoside crystal complexes, with particular attention to their hydrogen-bonding arrangement, the degree of planarity of the base pairs, and the relation of the sugar residues to the bases. Crystal complexes of deoxyguanosine with 5-bromodeoxycytidine[15] and inosine with 5-bromodeoxycytidine[16] are currently being investigated. The former shows the Watson-Crick pairing with the bases approximately coplanar and joined by three hydrogen bonds.

Summary.—The crystal structure of an intermolecular complex containing adenosine and 5-bromouridine has been presented. The results demonstrate the existence of a third type of base pairing which is different from that occurring in DNA as hypothesized by Watson and Crick and that found by Hoogsteen in a crystal complex of 9-methyl adenine and 1-methyl thymine. The significance and possible biological implications of this new structure are discussed briefly.

[1] Watson, J. D., and F. H. C. Crick, *Nature*, **171**, 373 (1953).

[2] Arnott, S., M. H. F. Wilkins, D. A. Marvin, R. Langridge, A. Rich, and L. D. Hamilton, *Abstracts*, Sixth International Congress of the International Union of Crystallography, September. 1963, Rome.

[3] Sobell, H. M., K. Tomita, and A. Rich, these PROCEEDINGS, **49**, 885 (1963).

[4] O'Brien, E. J., *J. Mol. Biol.*, **7**, 107 (1963).

[5] Hoogsteen, K., *Acta Cryst.*, **12**, 822 (1959).

[6] Mathews, F. S., and A. Rich, *J. Mol. Biol.*, in press.

[7] Felsenfeld, G., D. R. Davies, and A. Rich, *J. Am. Chem. Soc.*, **79**, 2023 (1957).

[8] Busing, W. R., K. O. Martin, and H. A. Levy, Oak Ridge National Laboratory, *TM-305* (1962).

[9] Kraut, J., and L. H. Jensen, *Acta Cryst.*, **16**, 76 (1963).

[10] Trueblood, K. N., P. Horn, and V. Luzzati, *Acta Cryst.*, **14**, 965 (1961).

[11] Harris, D. R., and W. M. Macintyre, personal communication.

[12] Ts'o, P. O. P., I. S. Melvin, and A. C. Olson, *Abstracts*, Seventh Annual Biophysical Society Meeting, February, 1963, New York.

[13] Cf. Pauling, L., *The Nature of the Chemical Bond* (Cornell University Press, 1960), 3d ed.

[14] Katz, L.. K. Tomita, and A. Rich, personal communication.

[15] Haschemeyer, A. E. V., A. Rich, and H. M. Sobell, to be published shortly.

[16] Haschemeyer, A. E. V., and H. M. Sobell, work in progress.