formed in four ways but can dissociate in only one whereas the trimer, which can also form in four ways, can dissociate in two ways. Without considering any of the higher order possibilities, this analysis lead to the following formulation.

$$K_{1,2} = 4K$$
$$K_{2,3} = 2K \quad (20)$$

where K is the intrinsic equilibrium constant for the formation of one cyclic set of hydrogen bonds. On the basis of these considerations, the actual ratio of $\bar{K}/K_{1,2}$ is seven times greater than predicted on a statistical basis alone.

Although this is a particularly high ratio, amounting to a $\Delta G°$ of some 1.2 kcal, this peculiar behavior is not unique to this purine system. It has also been observed in the self-association of alcohols and some phenols in CCl_4 (Coggeshall and Saier, 1951) as well as a variety of amides in benzene (Davies and Thomas, 1956a,b). These latter authors have generally attributed this effect to a differential and unfavorable entropy of formation for the dimer compared to the higher oligomers. There may also be an additional solvation contribution to both the entropy and the enthalpy of association in this current case where the aggregation of the planar aromatic rings of A provide large surface areas for Van der Waals interactions with the highly polarizable solvent, $CHCl_3$.

The present studies have demonstrated that near-infrared spectroscopy is a useful and discriminatory method capable of yielding data in purine and pyrimidine association studies of accuracy comparable to that obtained by a colligative property method such as vapor pressure osmometry. The two techniques may also be used in a complimentary fashion in cases where reliance upon one alone might be unfeasible. The information obtained from such studies also permits an evaluation of whether the complex formed at the dimer level is open or cyclic.

References

Bellamy, L. J. (1958), The Infrared Spectra of Complex Molecules, New York, N. Y., Wiley.

Binford, J. S., and Holloway, D. M. (1968), *J. Mol. Biol. 31*, 91.

Coggeshall, N. D., and Saier, E. L. (1951), *J. Amer. Chem. Soc. 73*, 5414.

Davies, M., and Thomas, D. K. (1956a), *J. Phys. Chem. 60*, 763.

Davies, M., and Thomas, D. K. (1956b), *J. Phys. Chem. 60*, 767.

Hammes, G. G., and Park, A. C. (1968), *J. Amer. Chem. Soc. 90*, 4151.

Hanlon, S. (1970), *in* Spectroscopic Approaches to Biomolecular Conformation, Urry, D. W., Ed., Chicago, Ill., American Medical Association Press, p 161.

Katz, L. (1969), *J. Mol. Biol. 44*, 279.

Katz, L., and Penman, S. (1966), *J. Mol. Biol. 15*, 220.

Klotz, I. M., and Franzen, J. S. (1962), *J. Amer. Chem. Soc. 84*, 3461.

Kyogoku, Y., Lord, R. C., and Rich, A. (1967), *J. Amer. Chem. Soc. 89*, 496.

Schrier, E. E. (1968), *J. Chem. Ed. 45*, 176.

Whetsel, K. B., and Lady, J. H. (1964), *J. Phys. Chem. 68*, 1010.

Higher Order Associations of Adenine and Uracil by Hydrogen Bonding. II. Formation of Complexes in Mixed Solutions of 9-Ethyladenine and 1-Cyclohexyluracil†

Glenn M. Nagel‡ and Sue Hanlon*

ABSTRACT: The interactions of 9-ethyladenine (A) and 1-cyclohexyluracil (U) by hydrogen bonding in $CHCl_3$ at 25° have been followed by the same techniques, near-infrared spectroscopy and vapor pressure osmometry, employed in the preceding study (paper I) of the self-associations of these solutes. The greater accuracy of this dual approach permits an evaluation of the formation constant for the principal trimolecular complex, AU_2, from the bimolecular complex, AU. The formation constant, K_{AU}, for the reaction, A + U $\rightleftarrows$ AU is 110 ± 9 m^{-1}, while for the reaction, AU + U $\rightleftarrows$ AU_2, is 21 ± 4 m^{-1}. The ratio, K_{AU}/K_{AU_2} = 5.2 ± 1.4, is sufficiently close to the predicted statistical value of 4 to justify the conclusion that both the Hoogsteen and the Watson–Crick sets of hydrogen-bonding sites on adenine (C_6-NH_2, N_7 and C_6-NH_2, N_1) have equal affinity for uracil and can be occupied simultaneously without substantially altering the individual site affinity. This result has interesting biological implications for interactions at the polynucleotide level.

Despite the recent questions raised by Donohue (1969, 1970), it seems highly probable that the specific arrangements of hydrogen bonds proposed by Watson and Crick (1953) for the complementary associations of the purine and pyrimidine partners in DNA are the ones which do indeed exist in the native double-helical structure. As Donohue has pointed out,

† From the Department of Biological Chemistry, College of Medicine, University of Illinois, Chicago, Illinois. *Received September 21, 1971.* This investigation was supported by Grants GM 15180 and GM 00471-10 from the National Institutes of Health, and GB 24550 from the National Science Foundation. Taken in part from the Ph.D. dissertation of G. M. N.

‡ Current address: Department of Molecular Biology, University of California, Berkeley, Calif.

however, alternate arrangements of these complementary bases are possible (Donohue, 1956; Donohue and Trueblood, 1960) and, in fact, when polymer constraints are relaxed, many of the low molecular weight purine and pyrimidine derivatives cocrystallize in these alternate arrangements. (For a summary of the adenine–uracil and adenine–thymine structures, see Sakore *et al.*, 1969.) In solution, the association constants for the association of uracil with various adenine derivatives in which one of the two sets of hydrogen-bonding sites is blocked support the fact that both sites can be utilized equally in hydrogen bonding to uracil (Katz, 1969).

Since alternate base arrangements and multiple interaction sites on purines and pyrimidines have interesting biological implications, we felt that an examination of the relative chemical affinity of a given nucleic acid base for its complementary partner in each set of the alternate arrangements was warranted.

An approach to this problem which we felt had a reasonable chance of success was to evaluate the formation constants of higher order mixed complexes in solutions containing the complementary bases and compare their values to those which would be predicted on a statistical basis. For example, if the two sets of hydrogen-bonding sites on 9-ethyladenine (A),[1] the Watson–Crick and the Hoogsteen sites, have equal affinity for 1-cyclohexyluracil (U) and there is no interaction between these sites, then the formation constant, K_{AU}, in reaction 1 should be four times as large as the constant,

$$A + U \underset{}{\overset{K_{AU}}{\rightleftharpoons}} AU \qquad (1)$$

K_{AU_2}, in reaction 2. This statistical ratio of four is arrived at

$$AU + U \overset{K_{AU_2}}{\rightleftharpoons} AU_2 \qquad (2)$$

by arguments similar to those of Klotz (1953) for the relationship between the successive binding constants for the process in which a macromolecule binds multiple ligands at sites of inherently equal affinity. In this present case, adenine is analogous to the macromolecule with two sites. The further assumption has been made that the occupancy of one set of sites by uracil does not exert a further selective effect as to which of the carbonyl oxygens will be utilized in the interaction of the second uracil with adenine.

In order to determine the weak formation constants, K_{AU} and K_{AU_2}, we have employed the same two techniques, vapor pressure osmometry and near-infrared spectroscopy, employed in the self-association studies of 9-ethyladenine and 1-cyclohexyluracil reported in paper I (Nagel and Hanlon, 1972). The precision and accuracy of these techniques, as revealed by the former studies, indicated that they would be ideally suited for the present determinations.

Experimental Section

The compounds employed and the details of the experimental methods, calibration and purification procedures, and instrumentation are identical with that reported in the preceding (paper I) of this series (Nagel and Hanlon, 1972), and will be only briefly recapitulated here. 9-Ethyladenine (A) and 1-cyclohexyluracil (U) were obtained from Cyclo Chemical Corp. 9-Ethyladenine was purified by recrystallization. All solutes were dried and stored over a desiccant prior to use. All solutions were prepared on a molal basis in pentene-stabilized chloroform (PS-$CHCl_3$) whose preparation has been previously described in paper I. One solution of 9-ethyladenine in DMF was prepared for near-infrared spectroscopy and run against a balanced reference solution containing CCl_4 in the same solvent (Klotz and Franzen, 1962). The DMF and CCl_4 were reagent grade materials which were dried over a desiccant and redistilled prior to use.

Weighings were made on a Mettler Model H20T semimicro balance. Near-infrared spectra were run at the stated temperature in a Cary Model 14 CMR spectrophotometer equipped with thermostatted cell adapters. Temperature of the cell compartments was monitored by a telethermometer bridge and probe manufactured by Yellow Springs Instrument Co. Osmometry measurements were made in the manner previously described in paper I with a Hewlett-Packard Model 302 vapor pressure osmometer.

Steiner Analysis of Osmometry Data. In order to obtain the concentrations of unassociated species in mixed solutions from the osmometry data, the method of Steiner (1968) was employed. Since the analysis involves a number of computational steps, a brief summary of the procedure is given below.

The basic equation in the Steiner analysis (Steiner, 1968) appropriate for mixed solutions containing A and U is

$$\theta = \ln X_A + \beta \ln X_U = (1 + \beta)\int_0^{m_e}\left(\frac{\varphi - 1}{m_e}\right)dm_e + \ln\left(\frac{1}{1+\beta}\right) + \beta \ln\left(\frac{\beta}{1+\beta}\right) \qquad (3)$$

where β is the ratio of the stoichiometric base concentrations, U_0 to A_0, m_e is the effective molal concentration of all species in solution, φ is the osmotic coefficient ($\varphi = m_e/m_s$), and X_A and X_U are the mole fractions of free A (A_1) and free U (U_1) in solution, respectively, based on effective molal concentrations, *i.e.*

$$X_A = \frac{A_1}{m_e} \qquad (4)$$

$$X_U = \frac{U_1}{m_e} \qquad (5)$$

Using this analytical approach, a series of mixed solutions of A and U were prepared at three different values of β. The experimental data for a given set of solutions at constant β were plotted as $1/\varphi$ *vs.* m_e. The data were smoothed and the value of $1/\varphi$ corresponding to various values of m_e were taken from this curve in the range between $m_e = 0.08$ and $= 0.0064$ molal. The parameter, $(\varphi - 1)/m_e$, was calculated for each φ, m_e pair and plots of $[(\varphi - 1)/m_e]$ *vs.* m_e were made for each set of solutions. A smooth curve was drawn through the points and extrapolated to $m_e = 0$. The function, θ, in eq 3 above was evaluated for various values of m_e from the area under each of the three plots and the appropriate β values. The values of θ so obtained were then plotted against β for fixed values of m_e in order to obtain the derivative, $(\partial\theta/\partial\beta)_{m_e}$, which Steiner (1968) has shown to be

$$\left(\frac{\partial\theta}{\partial\beta}\right)_{m_e} = \ln X_U \qquad (6)$$

[1] Abbreviations used are: 9-ethyladenine, A; 1-cyclohexyluracil, U; pentene-stabilized chloroform, PS-$CHCl_3$; dimethylformamide, DMF.

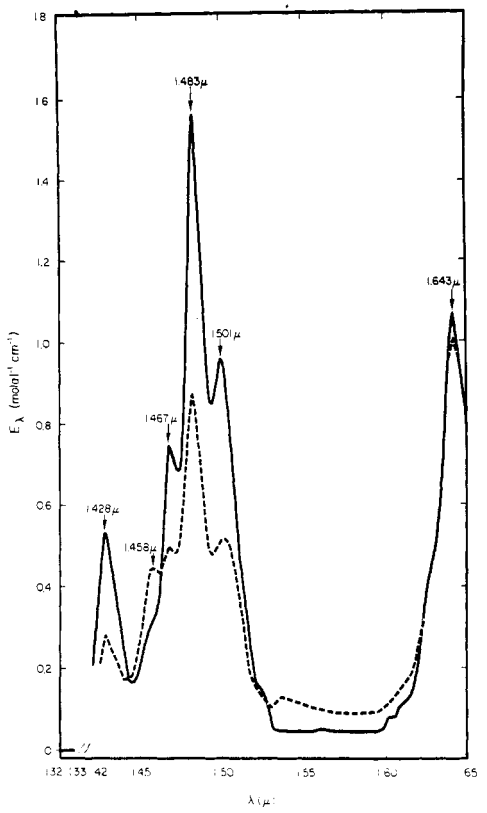

FIGURE 1: Spectra of mixed solutions of A and U at $\beta = 1.0$ in $CHCl_3$ at 25°. The extinction coefficients, E_λ, calculated on the basis of the stoichiometric molal concentration of a single base derivative, are plotted against wavelength, λ, in microns, for a solution in which $A_0 = U_0 = 5.64 \times 10^{-3}$ *m* (——) and 6.85×10^{-2} *m* (-----).

The value of $(\partial\theta/\partial\beta)_{m_0}$ for unique values of β was determined by taking the tangent to each curve at $\beta = 0.5$, 1.0, and 2.0. In order to minimize errors, an average of five separate determinations for each slope was obtained. The average derivation from the mean was 1–4% of $\partial\theta/\partial\beta$. The values of X_U and X_A were then calculated from eq 6 and 3 and the concentrations of the free bases, U_1 and A_1, from eq 5 and 4.

Results and Discussion

Near-Infrared Spectral Results. Near-infrared spectra of 1:1 molal mixtures of A and U are shown in Figure 1 for two different stoichiometric concentrations of these bases. The peaks displayed in the 1.5-μ region have been previously assigned to the first overtones of the N-H stretching vibrations of A and U (Nagel and Hanlon, 1972). With the exception of the 1.458-μ band (which has been assigned to the first overtone of the stretching vibration of the free NH in the singly hydrogen-bonded NH_2 group of adenine) the intensities of the peaks in the 1.5-μ region decrease with increasing concentration, reflecting the disappearance of the free-base species with increasing degree of association.

The analysis of the free base concentration employed the band at 1.501 μ, corresponding to the first overtone of the NH stretching vibration of U (Nagel and Hanlon, 1972) and at 1.483 μ, corresponding to the first overtone of the symmetric stretching vibration of the NH_2 group of adenine (Nagel and Hanlon, 1972). It is evident from these spectra that significant overlap of these two bands occurs. We have attempted to correct for this overlap by using absorption

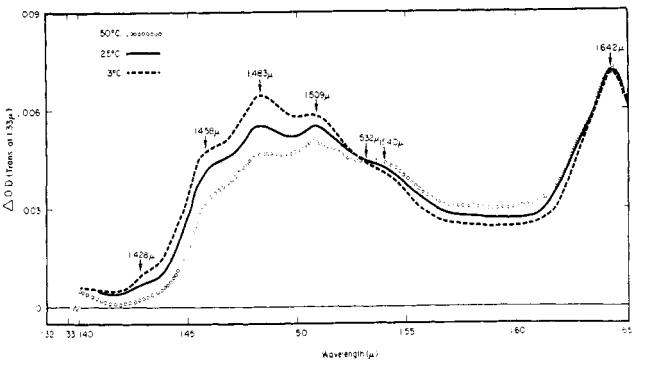

FIGURE 2: Spectra of A in dimethylformamide at various temperatures. The absorbance of the solution corrected for nonspecific absorption effects, ΔOD (trans at 1.33 μ), is plotted against wavelength in microns for a 0.134 M solution of A in DMF at 3° (-----), 25° (——), and 50° (oooo).

ratios of the individual bands (eq 7 and 8) determined from

$$k = (E_{1.483\,\mu}/E_{1.501\,\mu}) \text{ for U} \tag{7}$$

$$j = (E_{1.501\,\mu}/E_{1.483\,\mu}) \text{ for A} \tag{8}$$

the spectral properties of the separate A and U solutions, and then substituting these ratios together with the observed absorbances of the solution at 1.483 and 1.501 μ ($OD_{1.483\,\mu}$ and $OD_{1.501\,\mu}$) and the appropriate monomer extinction coefficients, E_A and E_U, at 1.483 and 1.501 μ, respectively, into the equations

$$A_1 = \frac{OD_{A\,1.483\,\mu}}{lE_A} = \frac{(OD_{1.483\,\mu} - kOD_{1.501\,\mu})}{(1 - kj)lE_A} \tag{9}$$

$$U_1 = \frac{OD_{U\,1.501\,\mu}}{lE_U} = \frac{(OD_{1.501\,\mu} - jOD_{1.483\,\mu})}{(1 - kj)lE_U} \tag{10}$$

where $OD_{A\,1.483\,\mu}$ and $OD_{U\,1.501\,\mu}$ correspond to the absorbance of the mixed solution at the indicated wavelength corrected for the overlap of the other interfering band and l is the path length of solution. These expressions for the free-base concentrations are valid only if the complexes formed are cyclic.

The ratio k, for U was 0.064 ± 0.004 and showed no concentration dependence. The ratio, j, for the A spectra was concentration dependent, decreasing by 10% from the highest to the lowest concentration. In order to correct for this effect, the calculated values for this ratio were plotted against the stoichiometric concentration of A and were extrapolated to infinite dilution to yield a value of 0.15 ± 0.02.

The half-band width of the 1.483-μ peak was considerably increased over the value found in the spectra of A in solution alone. This increase could not be accounted for by an overlap of any of the other bands observed in the mixed solutions. In the light of this effect as well as the observed concentration dependence of k, it appeared quite likely that there was one or more hidden absorption bands due to overtones of the hydrogen-bonded mixed complexes.

In order to explore this possibility, the spectrum of A was examined as a function of temperature in the amide solvent, DMF, which would provide a high concentration of carbonyl oxygens for hydrogen bonding to the exocyclic amino groups of adenine. The results of this experiment are shown in Figure 2 as a plot of the absorbance *vs.* wavelength. (The spectra were corrected for nonspecific absorption and scattering by sub-

TABLE I: Band Assignments in the Near-Infrared Region for Hydrogen-Bonded Species of Adenine.

Band Positions				
Near Infrared				
μ	cm^{-1}	Fundamental[b] (cm^{-1})	R[a]	Assignment
1.458	6859	3485	1.968	2ν (H—N—H···O=C) ↔
1.509	6627	3327	1.992	2ν (H—N—H···O=C) ↔
1.540	6494	3256	1.994	2ν (C=O···H—N—H···O=C) ↔ ↔

[a] $R = \bar{\nu}$ (near ir)/$\bar{\nu}$ (fundamental) (as defined in text). [b] Taken from Kyogoku *et al.* (1967).

tracting the measured optical density at 1.33 μ from the spectrum in the wavelength region displayed.) These spectra show unresolved absorption bands which appear to be centered at 1.428 (7003), 1.458 (6859), 1.483 (6743), 1.509 (6627), 1.540 (6494), and 1.642 μ (6090 cm^{-1}). There is also considerable absorption in the 1.570- to 1.640-μ region although no maximum is apparent. The positions and other characteristics of some of these bands are summarized in Table I.

The absorption band centered at 1.642 μ is similar in position to that observed in the A spectra in PS-$CHCl_3$ and is undoubtedly due to the first overtones of the CH vibrations. Part of the absorbance at 1.483 μ, by similar reasoning, has been assigned to the first overtone of the symmetric stretching vibration of the free NH_2 group of adenine but in view of the very small absorbance in the region of the asymmetric stretching vibration at 1.428 μ, the major share of the 1.483-μ absorbance is due to the overlap of the broad bands at 1.458 and 1.509 μ.

The assignments for the bands at 1.458, 1.509, and 1.540 μ as well as the absorbance near 1.57 μ can be made by reference to the fundamental spectra of mixed solutions of A and U obtained by Kyogoku *et al.* (1967) as well as the temperature dependence of the spectra in Figure 2. In the following discussion we have made the assumption that the vibrational properties of the A···DMF species are approximately those of the A::U hydrogen-bonded species and that the nature of the solvent does not exert a marked influence on band positions. This latter assumption is supported by the fact that the position of the free NH and CH overtones are essentially the same in DMF as in PS-$CHCl_3$.

The fundamental infrared spectra of A and U mixtures show two bands at 3485 and 3327 cm^{-1} (Kyogoku *et al.*, 1967). The band at 3485 cm^{-1} has been assigned to the free NH vibration of the amino group of adenine which is singly hydrogen bonded and the one at 3327 cm^{-1} to the NH of the same group which is hydrogen bonded to the carbonyl oxygen of uracil. The ratios, R, given in Table I, of the frequencies of the near-infrared bands at 1.458 μ (6859 cm^{-1}) and 1.509 μ (6627 cm^{-1}) to those on the fundamental region reveal that the former are the first overtones of these fundamental bands. The higher ratio, R, of 1.99 for the overtone–fundamental relationship of the hydrogen-bonded NH is consistent with that observed for the other types of hydrogen-bonded NH groups (Hanlon, 1970) and is attributable to the fact that hydrogen-bond formation generally reduces the anharmonicity of a stretching vibration (Pimentel and McClellan, 1960). The assignment of the 1.458-μ band is further supported by the fact that its position corresponds exactly to what is predicted on the basis of the positions of the first overtones of the symmetric and the asymmetric stretching vibrations of the free NH_2 group of adenine (Bellamy, 1958).

As Figure 2 demonstrates, the lowering of the temperature of the solution decreased the intensity of the bands at 1.428 1.458, 1.483, and 1.509 μ but had the opposite effect on the band at 1.540 μ and the absorbance in the 1.570-μ region. The spectra displayed have an isosbestic point at 1.532 μ which was also exhibited by other spectra in the same experiment at intermediate temperatures. The simplest explanation of this phenomenon is that the lowering of the temperature of this A solution results in the conversion of the singly hydrogen-bonded NH_2 of A (whose absorption bands fall at 1.458 and 1.509 μ) to the doubly hydrogen bonded amino species with an absorption band at 1.540 μ and a broad band somewhere around 1.570 or 1.580 μ. It should be noted the Kyogoku *et al.* (1967) observed a well-defined band at 3256 cm^{-1} and a very diffuse but strong band from 3230 cm^{-1} to below 3170 cm^{-1} in the fundamental spectra of mixed solutions of A and U in $CHCl_3$. They attributed these to hydrogen-bonded NH bands of adenine but were not more specific. The ratio, R, for the near-infrared absorptions at 1.540 and 1.570 μ to these fundamental bands is 1.99. Hence we have concluded that these near-infrared absorptions are first overtones of the fundamental bands so that the latter are also due to the NH_2 group of adenine in which both hydrogens are hydrogen bonded to carbonyl oxygens. In the light of this assignment we may also conclude that the mixed solutions of A and U examined by Kyogoku *et al.* (1967) contained AU_2 species.

The presence of bands due to associated adenine species complicates the straight forward calculation of the free base concentration from the overtone spectral data at 1.483 and 1.501 μ. This could be shown to be the case by preliminary calculations of the apparent equilibrium constant, K_{AU}(app) by substituting the concentration of free A and U obtained by the application of eq 9 and 10 into eq 11 and 12 in which

$$K_{AU}(\text{app}) = \frac{U_0 - \sum_i iU_i}{A_1U_1} \quad (11)$$

$$K_{AU}(\text{app}) = \frac{A_0 - \sum_i iA_i}{A_1U_1} \quad (12)$$

TABLE II: Determination of K_{AU} by Extrapolation of Near-Infrared Spectra and Vapor Pressure Osmometry Results to Infinite Dilution.[a]

	Spectral Results[b]		Osmometry Results	
$\beta = U_0/A_0$	U Data Eq 15a	A Data Eq 15b	U Data Eq 15a	A Data Eq 15b
0.5	105	109		110
1.0	111	92	114	102
2.0	119	89	132	
Average	104 ± 9		115 ± 9	

[a] K_{AU} (m^{-1}) in $CHCl_3$ at 25°. [b] The index of aggregation, i, was limited to 1.

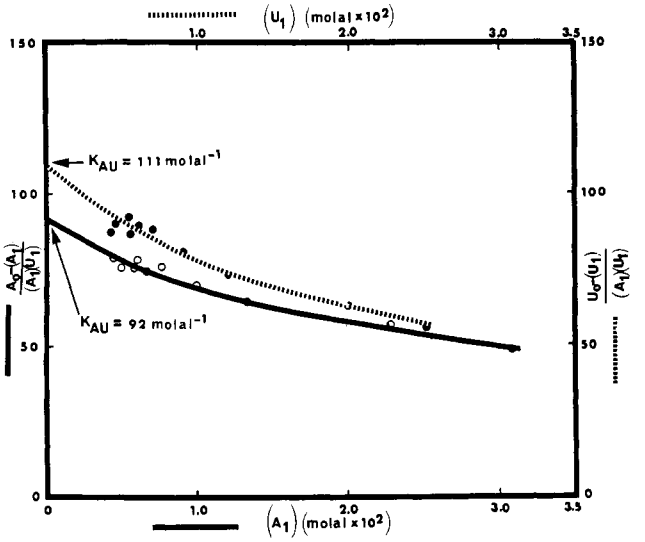

FIGURE 3: Determination of K_{AU} from the spectral data of mixed solutions of A and U in $CHCl_3$ at 25°. The dashed curve represents the values of K that were calculated for the U data by eq 11 for $i = 1$ and plotted on the right-hand ordinate against the free-U concentration on the upper abscissa. The solid line represents the values of K that were calculated for the A data by eq 12 for $i = 1$ and plotted on the left-hand ordinate against the free A concentration on the lower abscissa.

the subscript "0" indicates the total stoichiometric molal concentration of the given base derivative and the subscript "i" is the index of aggregation. On the basis of the previous study (Nagel and Hanlon, 1972), we limited i to 2 for U and 4 for A as the association of U does not proceed beyond the dimer stage and the association constants are so small for A that concentrations of aggregates higher than the tetramer are negligible in the concentration range covered in these experiments. The concentration of the self-associates ($i > 1$) was calculated from the free-base concentration by application of the appropriate equilibrium expressions.

$$U_2 = K_{1,2U}(U_1)^2 \tag{13}$$

$$A_i = K_{1,2A}\bar{K}^{i-2}(A_i)^i \tag{14}$$

$K_{1,2U}$, $K_{1,2A}$, and $\bar{K}$ were taken from our previous study (Nagel and Hanlon, 1972). Equations 11 and 12 attribute all mixed association to AU complex formation. If higher order species exist, the values of K_{AU}(app) should normally increase with increasing concentration. The true K_{AU} should thus be the limit in infinitely dilute solutions.

$$\lim_{U_1 \to 0} [K_{AU}(\text{app}), \text{eq } 11] = K_{AU} \tag{15a}$$

$$\lim_{A_1 \to 0} [K_{AU}(\text{app}), \text{eq } 12] = K_{AU} \tag{15b}$$

The equilibrium constant K_{AU}(app) so calculated, decreased with increasing concentration. This is to be expected since the absorbance at 1.483 and 1.501 μ included not only the free-base concentration but the hydrogen-bonded NH contributions as well, thus giving rise to a spuriously high absorbance at these wavelengths.

As the concentration of solutes in these mixed solutions decreases, however, eq 9 and 10 should become increasingly more accurate. In the limit of 0 concentration we may furthermore neglect the self-associates since these preliminary calculations as well as the data of Kyogoku *et al.* (1967) have shown that the magnitude of K_{AU} is much greater than the self-association constants. Proceeding on these assunptions, we have calculated values of K_{AU}(app) using eq 11 and 12 with the restriction that $i = 1$ (*i.e.*, concentrations of self-associates have been neglected). The "apparent" free-base concentrations, A_1 and U_1, were obtained from the spectral data by applying eq 9 and 10. The values of K_{AU}(app) obtained from eq 11 and 12 by this method were plotted against the appropriate apparent free-base concentration and extrapolated to zero concentration. A typical plot is shown in Figure 3 for the A and U data in a mixed solution at $\beta = 1.0$. Similar behavior was observed at other β values. The extrapolated results for the various solutions examined are given in Table II. By this procedure an average value of $K_{AU} = 104 \pm 9$ m^{-1} was obtained. The agreement between the various values is considered to be very good when one considers that the magnitude of the curvature near zero concentration is dependent upon the relative weights of factors which will be different for each value of β, namely, the contribution of the self-associates, the higher order mixed association reactions and the type and magnitude of the absorption bands of the hydrogen-bonded A species.

Vapor Pressure Osmometry. The vapor pressure osmometry data from mixed solutions at different ratios are shown in Figure 4. The concentrations, A_1 and U_1, calculated by the method of Steiner (1968) (as described in the experimental section) were 30–40% lower at the upper end of the concentration scale than those calculated from the spectral data.

The greater accuracy in the estimation of the free A_1 and U_1 concentrations in the osmometry experiments permitted a determination of K_{AU} by the inherently more accurate extrapolation of K_{AU}(app) calculated from eq 11 and 12 in which $i = 2$ for U and $i = 4$ for A self-association. The typical concentration behavior of K_{AU}(app) as a function of the free-base concentration is demonstrated for the A data in Figure 5 for a mixed solution at $\beta = 0.5$. Similar results were obtained at other β values. The extrapolated values of K_{AU}(app) for this solution set as well as the others examined are reported in the last two columns of Table II. The average value of $K_{AU} = 115 \pm 9$ m^{-1}, obtained from the osmometry results is in excellent agreement with that obtained from the spectral data. This fact demonstrates that the AU complex formed is cyclic. The average of the two methods is given in

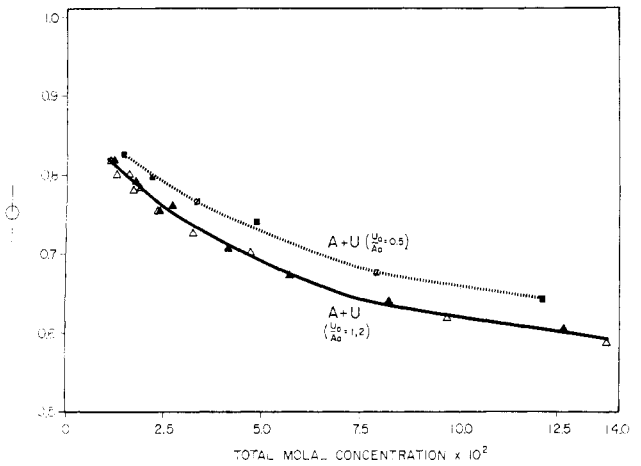

FIGURE 4: Behavior of the osmotic coefficient, φ, as a function of total molal concentration in mixed solutions of A and U in $CHCl_3$ at 25° at β ratios of 0.5 (■), 1.0 (△), and 2.0 (▲).

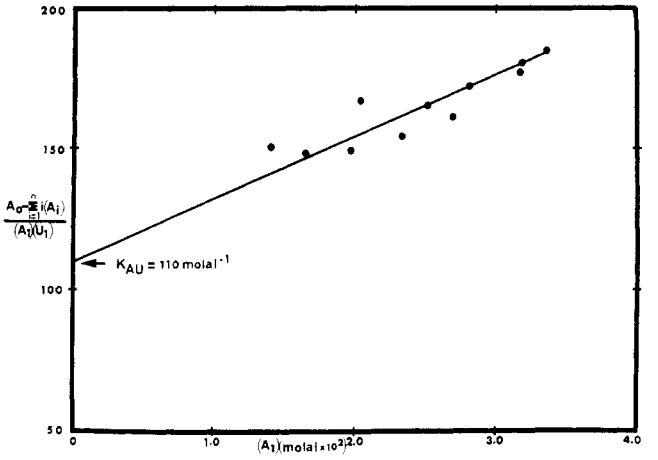

FIGURE 5: Determination of K_{AU} from osmometry data in mixed A and U solution in $CHCl_3$ at 25°. Values of K_{AU} were calculated from eq 12 for $i = 4$ by substituting the free-A concentration obtained from the Steiner analysis of the osmometry data and plotted on the ordinate against the free-A concentration on the abscissa.

Table III together with the value determined by Kyogoku *et al.* (1967).

Higher Order Complexes. The increase in K_{AU} with increasing concentration of solute observed in the vapor pressure osmometry experiments is best explained by the presence of higher order mixed complexes. Since U self-association had been previously shown to stop at the dimer (Nagel and Hanlon, 1972), the most likely higher order complexes are the AU_2, the A_2U, and perhaps, A_2U_2 and A_3U. The precision of our measurements and the concentration limits imposed by the solubility of A and U in $CHCl_3$ were insufficient to evaluate the formation constants for all of these possible complexes. We could, however, extract the values of K_{AU_2} and an upper limit for K_{A_2U} from these experiments by employing the following analytical approach.

The conservation of mass in mixed solutions requires that

$$U_0 = \sum_i iU_i + AU + 2(AU_2) + A_2U + A_3U + 2(A_2U_2) \quad (16)$$

$$A_0 = \sum_i iA_i + AU + AU_2 + 2(A_2U) + 3(A_3U) + 2(A_2U_2) \quad (17)$$

Subtraction of eq 17 from 16 and subsequent rearrangement gives

$$(U_0 - \sum_i iU_i) - (A_0 - \sum_i iA_i) = AU_2 - A_2U - 2(A_3U) \quad (18)$$

Since each of these higher order species can be formally represented as being derived from a reaction of a given free base with the AU complex, eq 18 may be rewritten as

$$\frac{(U_0 - \sum_i iU_i) - (A_0 - \sum_i iA_i)}{A_1U_1} = K_{AU}K_{AU_2}U_1 - K_{AU}K_{A_2U}A_1 - 2K_{AU}K_{A_2U}K_{A_3U}(A_1)^2 \quad (19)$$

If the concentration of A_2U and A_3U are small compared to AU_2, then a plot of the function on the left-hand side of eq 19 against the free U concentration should be linear and go through the origin. The value of K_{AU_2} can be evaluated from the slope if the value of K_{AU} is known.

A plot of this type for the osmometry data obtained in the set of solutions for which $\beta = 2.0$ (where AU_2 formation should be maximal) is shown in Figure 6. Despite the scatter, the least mean square behavior indicates a polynomial of degree 1 with an intercept of 0.3 m^{-1} and a slope of 2630 m^{-2}. Using an average K_{AU} of 110 m^{-1}, the value of K_{AU_2}, evaluated from the slope, is 24 m^{-1}.

Another method of obtaining K_{AU_2} is to subtract eq 20

$$m_e = \sum_i U_i + \sum_i A_i + AU + AU_2 + A_2U + A_3U + A_2U_2 \quad (20)$$

for the effective molality, m_e, from eq 16 to yield

$$(U_0 - \sum_i iU_i) - (m_e - \sum_i U_i - \sum_i A_i) = AU_2 + A_2U_2 \quad (21)$$

With the appropriate substitutions and rearrangements, eq 21 becomes

$$\frac{(U_0 - \sum_i iU_i) - (m_e - \sum_i U_i - \sum_i A_i)}{A_1U_1} = K_{AU}K_{AU_2}U_1 + K_{AU}K_{A_2U_2}A_1U_1 \quad (22)$$

If A_2U_2 is small compared to AU_2 in the concentration range examined, then a plot of the left-hand side of eq 22 *vs.* U_1 should again be linear, pass through the origin, and yield K_{AU_2} from the slope. Such a plot for the set of solutions at $\beta = 2.0$ is shown in Figure 7. The value of K_{AU_2} determined in this fashion is 17 m^{-1} for $K_{AU} = 110$ m^{-1}. The average value obtained from the two determinations of K_{AU_2} is 21 ± 4 m^{-1} as reported in Table III.

If K_{AU} is known, the upper limit of K_{A_2U} may be estimated from eq 23 and 24.

Subtraction of eq 20 from 17 yields

$$(A_0 - \sum_i iA_i) - (m_e - \sum_i A_i - \sum_i U_i) = A_2U + 2(A_3U) + A_2U_2 \quad (23)$$

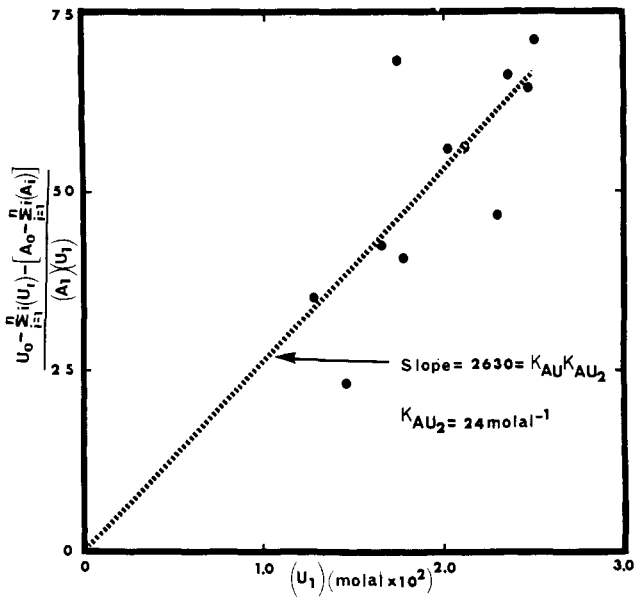

FIGURE 6: Determination of K_{AU_2} in mixed solutions of A and U in $CHCl_3$ at 25° by eq 19. Data from the osmometry experiments are plotted on the ordinate as the function in the left-hand side of eq 19 against the free-U concentration on the abscissa for solutions of A and U at $\beta = 2.0$.

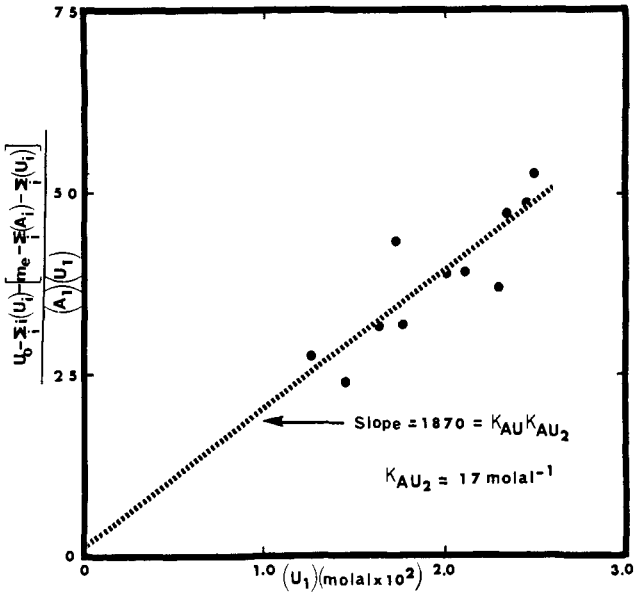

FIGURE 7: Determination of K_{AU_2} in mixed solutions of A and U in $CHCl_3$ at 25° by eq 22. Data from the osmometry expriments are plotted on the ordinate as the function in the left-hand side of eq 22 against the free U concentration on the abscissa for solutions of A and U at $\beta = 2.0$.

Substitution of the appropriate formal equilibrium constants and rearrangement gives

$$\left[\frac{(A_0 - \sum_i iA_i) - (m_e - \sum_i A_i - \sum_i U_i)}{(A_1)^2U_1}\right] = (K_{AU}K_{A_2U} + 2K_{A_3U}K_{A_2U}K_{AU}A_1 + K_{A_2U_2}K_{AU}{}^2U_1) \quad (24)$$

If none of the complexes in eq 23 existed, the left-hand side of eq 24 should be 0 at all values of A_1 and U_1. Calculations showed that it was slightly positive but the scatter of the value was too great to detect any systematic trend with increasing concentration of free A and U. If the assumption is made that the positive value is due entirely to the presence of the A_2U complex, an upper limit for the association constant for A_2U formation may be calculated as 6 m^{-1} from the data of the $\beta = 0.5$ and 1.0 solutions where composition would favor the formation of complexes rich in A. Using this maximal estimate of K_{A_2U}, the concentration of A_2U complex in the solution at $\beta = 2.0$ is calculated to be less than 10% of the concentration of the AU_2 complexes. The original assumption that the only higher order species present in significant amounts is the AU_2 complex in these $\beta = 2.0$ solutions is thus a reasonable one.

Conclusion

The ratio K_{AU}/K_{AU_2} determined in these experiments is 5.2 ± 1.4. This value is sufficiently close to the statistical value of 4 to justify the conclusion that the two hydrogen bonding sites on adenine have equal affinity for uracil. Although it might be argued that this conclusion could have been anticipated from the fact that many base derivatives cocrystallize in the Hoogsteen arrangement and that poly(A + 2U) is a stable complex, the steric properties of polymers and crystals frequently impose their own requirements on packing arrangements which are not always completely compatible with the intrinsic chemical affinities of the principle interacting functional groups of the molecule. We have demonstrated in this study that even in the absence of polymeric and crystalline constraints, the intrinsic affinity of the Watson–Crick sites on adenine matches that of the Hoogsteen sites, as far as the hydrogen bonding interaction with uracil in $CHCl_3$ is concerned. One may also conclude that the intrinsic affinity of one site on adenine is not substantially impaired by the occupation of the alternate site. We do not feel that these conclusions would be markedly altered if the more pertinent enthalpic values for the associations had been obtained rather than free energies, as determined in these present studies. There is no reason to believe that the behavior of adenine toward thymine would be different in either regard, nor would the ratio of these affinities of the two sets of sites be altered by interactions in aqueous solvents.

As a corollary, we can conclude that since the hydrogen bonding affinity of the Watson–Crick and the Hoogsteen sites in adenine is equal, then the recognition properties of the latter match those of the former. The Hoogsteen sites could then readily participate in a variety of interactions at the chromosomal level without the necessity of DNA strand un-

TABLE III: Summary of the Formation Constants for the Mixed Complexes of A and U in $CHCl_3$ at 25°.

Constant	Process	Av, This work		Av, Kyogoku *et al.* (1967)
		m^{-1}	M^{-1}	(M^{-1})
K_{AU}	A + U ⇄ AU	110 ± 9	74 ± 6	103 ± 36
K_{AU_2}	AU + U ⇄ AU_2	21 ± 4	14 ± 3	
K_{A_2U}	AU + A ⇄ A_2U	*Ca.* 6	*Ca.* 4	
	$\frac{K_{AU}}{K_{AU_2}}$ = 5.2 ± 1			

winding. Perhaps the runs of dihydrouracil in chromosomal RNA (Huang and Bonner, 1965) are anchored to a matching set of runs of poly(dA) in the DNA genome *via* interactions at the adenine Hoogsteen site. If the excess functional groups of the other bases are also reactive toward their complimentary counterparts this suggests an interesting mechanism of transcription using the intact but perhaps slightly distorted Watson and Crick duplex as the template.

We have also demonstrated in these experiments, the feasibility and advantages of employing two techniques of a different nature for the study of a very complex set of weak interactions. For such cases, utilization of a single technique can lead to serious errors in the interpretation of data and the value of the associations constants. Since the mixed association of A and U is occasionally used as a standard to test other methods, we suggest that the more accurate values of the association constants presented in this and the preceding paper be employed. In addition, these values also permit an accurate evaluation of the more pertinent enthalpies of the various associative processes by calorimetric methods.

References

Bellamy, L. J. (1958), The Infrared Spectra of Complex Molecules, New York, N. Y., John Wiley and Sons.

Donohue, J. (1956), *Proc. Nat. Acad. Sci. U. S. 42*, 60.

Donohue, J. (1969), *Science 165*, 1091.

Donohue, J. (1970), *Science 167*, 1700.

Donohue, J., and Trueblood, K. N. (1960), *J. Mol. Biol. 2*, 363.

Hanlon, S. (1970), *in* Spectroscopic Approaches to Biomolecular Conformation, Urry, D. W., Ed., Chicago, Ill., American Medical Association Press, p 161.

Huang, R. C., and Bonner, J. (1965), *Proc. Nat. Acad. Sci. U. S. 54*, 960.

Katz, L. (1969), *J. Mol. Biol. 44*, 279.

Klotz, I. M. (1953), *Protein 1*, 748.

Klotz, I. M., and Franzen, J. S. (1962), *J. Amer. Chem. Soc. 84*, 3461.

Kyogoku, Y., Lord, R. C., and Rich, A. (1967), *J. Amer. Chem. Soc. 89*, 496.

Nagel, G., and Hanlon, S. (1972), *Biochemistry 11*, 816.

Pimentel, G. C., and McClellan, A. L. (1960), The Hydrogen Bond, San Francisco, Calif., W. H. Freeman and Co.

Sakore, T. O., Tavale, S. S., and Sobell, H. M. (1969), *J. Mol. Biol. 43*, 361.

Steiner, R. F. (1968), *Biochemistry 7*, 2201.

Watson, J. D., and Crick, H. F. C. (1953), *Nature (London) 171*, 737.

Studies on the Conformation of Purine Nucleosides and Their 5′-Phosphates†

Morio Ikehara,* Seiichi Uesugi, and Katsumi Yoshida

ABSTRACT: Circular dichroism of various purine nucleosides and their 5′-monophosphates was measured. From the relative rotational strength and the sign of the Cotton effect following conclusions were reached. (1) Purine nucleosides having bulky substituents in 8 position, such as bromo, 2-hydroxypropyl, and methylmercapto group, have syn conformations. (2) Introduction of a phosphate to the 5′-hydroxyl group of anti-type nucleosides, such as adenosine, 8,2′- and 8,3′-*S*-cycloadenosine, does not change the circular dichroism (CD) curve in the B-band region. (3) Introduction of a phosphate to the 5′-hydroxyl group of syn-type nucleosides such as 8-substituted purine nucleosides caused a drastic change of CD curves over a wide range of wavelength. (4) Both in syn- and anti-type purine nucleosides, introduction of the 5′-phosphate caused a significant change in the Cotton band of 200–220 nm.

Recently, optical rotatory dispersion (ORD) and circular dichroism (CD) of purine and pyrimidine nucleosides were extensively investigated for the elucidation of the conformation of nucleosides, nucleotides, and polynucleotides (Yang and Samejima, 1969; Miles *et al.*, 1969). A theoretical study for the calculation of rotational strength of nucleosides has been reported in order to interpret the effect of substituents to the base (Miles *et al.*, 1967) and the sugar moiety (Miles *et al.*, 1968). According to this theory, if the torsion angle (Donohue and Trueblood, 1960) of a nucleoside is fixed, substituents on the base and the sugar moiety give a change in the direction of the transition moment of the base chromophore. The sign of the rotational strength is determined by whether the polarization angle exceeds a critical value or not. If the torsion angle changes, the anisotropic effect of the sugar moiety will affect the sign of the rotational strength.

We measured CD of a variety of purine nucleosides and their 5′-monophosphates, especially those having substituents in 8 position, in order to obtain information about the interaction of the base and the phosphate moiety. Among 8-substituted purine nucleosides, some compounds have cancerostatic activity (Bloch *et al.*, 1966) and relationship of the structure and the function is of interest. From the results of CD measurements, the following conclusions were drawn. (1) Purine nucleosides having bulky substituents in the 8 position exist in syn conformation. (2) Introduction of a phosphate group to the 5′-OH of anti-type nucleosides does not change

† From the Faculty of Pharmaceutical Sciences, Osaka University, Toyonaka, Osaka, Japan. *Received June 11, 1971.* This work was supported by Grant-Aid for Scientific Research from the Ministry of Education.