Higher Order Associations of Adenine and Uracil by Hydrogen Bonding. I. Self-Association of 9-Ethyladenine and 1-Cyclohexyluracil†

Glenn M. Nagel‡ and Sue Hanlon*

ABSTRACT: The self-association processes of 9-ethyladenine (A) and 1-cyclohexyluracil (U) by hydrogen bonding in $CHCl_3$ have been followed simultaneously by near-infrared spectroscopy and vapor pressure osmometry over a concentration range greater than previously examined by others. This approach permits the detection and evaluation of weak complexes of a higher order than the simple bimolecular type with great accuracy. It can also distinguish between open and cyclic structures at the bimolecular level. Between 0.002 and 0.10 *m*, these derivatives form hydrogen-bonded cyclic dimers with a formation constant of 10.0 ± 0.5 m^{-1} for U and 2.0 ± 0.1 m^{-1} for A. In addition, A dimers go on to form trimers (and perhaps higher oligomers) with a stepwise formation constant of 7.1 ± 0.5 m^{-1}. The agreement between the osmotic and the near-infrared spectral data is excellent. This supports the reliability of the results and demonstrates the desirability of employing this dual approach in following weak associations.

Although the association of 9-ethyladenine (A)[1] and 1-cyclohexyluracil (U) has been examined by a variety of techniques (Katz and Penman, 1966; Kyogoku *et al.*, 1967; Hammes and Park, 1968; Binford and Holloway, 1968), the equilibrium constants obtained have been imprecise. Furthermore, the concentration range normally examined in these studies has been too narrow to permit an evaluation of the formation constant of any species other than a simple dimer or an AU complex. The data of Katz and Penman (1966) however, have suggested the presence of higher order self- and mixed associates. This implies that more than one site is available for cyclic AU hydrogen-bonding interaction. Since this possibility has interesting biological implications, we have decided to reinvestigate this interaction using an inherently more accurate approach.

In order to evaluate the formation constants for higher order mixed complexes, it is first necessary to obtain reliable data on the simpler self-associations. The first paper of this series describes the determination of these self-association constants by a novel approach utilizing *two* techniques, near-infrared spectroscopy and vapor pressure osmometry. The former has all of the advantages of a discriminatory spectral method whereas the latter permits direct estimates of the degree of polymerization at any concentration level from changes in a colligative property of an experimental solution. A combination of the two techniques yields inherently more accurate constants and reveals, within limits, the nature of the complexes—open or cyclic—formed. The same approach has been employed in following associations in mixed solutions of A and U. The results of these latter studies are reported in paper II of this series.

Experimental Section

Chloroform (Superior grade), stabilized by 0.02% 2-pentene, was purchased from Matheson Scientific, Inc. Technical grade pentene was purified by fractional distillation. The fraction boiling at 36–37.5°, corresponding to the mixed cis and trans isomers of 2-pentene, was employed in these experiments. D_2O (99.85 mole %) was obtained from Bio-Rad Laboratories.

The base derivatives, 9-ethyladenine (lot R6037) and 1-cyclohexyluracil (lot R6013), were purchased from Cyclo Chemical Corp. 9-Ethyladenine was recrystallized from pentene-stabilized chloroform (either PS-$CHCl_3$ described below or the commercial Matheson product). This recrystallized A as well as U were dried *in vacuo* at 56° and stored in a vacuum desiccator over Drierite until used.

Solutes employed in calibrating the vapor pressure osmometer were biphenyl (mp 69.5–70.5°) and benzil (mp 95–96°) obtained from Eastman Kodak, reagent grade benzophenone (mp 48.1–48.9°) from Fisher Scientific Co., and USP anhydrous caffeine (mp 234–236.5°) from Aldrich Chemical Co. Benzophenone was dried at room temperature *in vacuo* and biphenyl at 50° at atmospheric pressure. The remaining solutes were dried *in vacuo* at 56°. After drying, all solutes were stored in a vacuum desiccator over Drierite. All other chemicals employed were reagent grade.

Preparation of Solvent and Solutions. During the course of this investigation it became apparent to us that special precautions need be observed in the preparation and handling of chloroform solutions if reproducible results were to be obtained. This was especially true for those solutions containing adenine. For this reason, the preparation and manipulation of solutions and solvent are given in detail below.

Both the solvent, pentene-stabilized chloroform, and the solutions of the base derivatives were routinely prepared in a darkroom illuminated with a safelight equipped with a Kodak Series I Wratten red filter. After desiccation over P_2O_5 in amber glass bottles, Matheson Superior grade $CHCl_3$, pentene stabilized, was fractionally distilled in a Vigreaux column of 20 cm in length. The entire still was covered with aluminum foil. The middle fraction, distilling at 61°, was collected and a quantity of purified 2-pentene added to make it 0.04% (v/v)

† From the Department of Biological Chemistry, College of Medicine, University of Illinois, Chicago, Illinois. *Received September 21, 1971.* This investigation was supported by Grants GM15180 and GM00471-10 from the National Institutes of Health, and GB24550 from the National Science Foundation. Taken from the Ph.D. dissertation of G. M. N.

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

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

in pentene. This was subsequently employed as a solvent and is referred to as PS-$CHCl_3$.

Solutions of the base derivatives were prepared on the day of their use by the addition of PS-$CHCl_3$ to preweighed solute samples in amber glass bottles. The solution was then reweighed to obtain the weight of added solvent and concentrations were calculated in molal units. After experimental measurements were performed, a given solution was diluted by adding PS-$CHCl_3$ to a preweighed aliquot of the recovered solution. No more than two dilutions of a parent solution were made. Measurements on both parent solutions and dilution were always made on the same day. All weighing operations employed a Mettler H2OT semimicro analytical balance.

Vapor Pressure Osmometry. A Hewlett-Packard Model 302 vapor pressure osmometer, thermostatted at 25° and equipped with a nonaqueous probe, was employed in these studies. In order to achieve the precision required, it was necessary to stringently standardize the operation procedure suggested by the manufacturers. The details of this procedure will be presented elsewhere (G. M. Nagel and S. Hanlon, in preparation) and we will only briefly summarize as follows. Prior to each measurement, the Wheatstone bridge was nulled with solvent droplets on both thermistor beads. The solution droplet was applied to the sample thermistor at 10-sec intervals for a period of 50 sec. An additional droplet was applied at 55 sec and the final one at 60 sec. The difference in the resistance, ΔR, between sample and reference thermistors was read at 1-min intervals up to 6 min beginning at 2 min after the final solution drop was applied. These values of ΔR were then linearly extrapolated to obtain the value at 0 time, ΔR_{ZT}. A minimum of two determinations of ΔR_{ZT} were made for each solution and for the great majority of solutions, three determinations were made. To minimize solvent evaporation, only a single solution was measured at a time and syringes containing solution were allowed to thermally equilibrate in the instrument for only 20–30 min.

Precautions were taken to minimize exposure of these chloroform solutions to room light by covering the openings of the syringe holders with opaque paper and dropping an opaque cloth over the roof of the instrument chamber. In addition, a no. 66 red glass Klett filter was mounted in front of the illumination port of the chamber.

The osmometer was calibrated at 25° in the concentration range from 2.5×10^{-3} to 1×10^{-1} *m* using four solutes, benzophenone, benzil, biphenyl, and caffeine. The molecular weight range covered with these solutes encompassed the molecular weights of unassociated A and U. Solutions of these four calibrating solutes were prepared in PS-$CHCl_3$ in the fashion previously described. A few solutions were also prepared in the undried commercial solvent as well. No significant difference was observed between the experimental data from solutions in the latter solvent compared to the former. The calibration curve was constructed by plotting the value of ΔR_{ZT} for each solution against its molal concentration, *m*. The data could be adequately represented by three straight lines whose equation was

$$\Delta R_{ZT} = 0.18 + 561\ m \quad (1)$$

between 2.5×10^{-3} to 1.1×10^{-2} *m*, and

$$\Delta R_{ZT} = 0.60 + 498\ m \quad (2)$$

between 1.1×10^{-2} and 6.0×10^{-2} *m*, and

$$\Delta R_{ZT} = 1.37 + 485\ m \quad (3)$$

from 6.0×10^{-2} and 1.0×10^{-1} *m*. The mean deviation of all calibration points from these least-mean-square average lines was less than 1% of ΔR_{ZT}.

This calibration curve together with the values of ΔR_{ZT} for the experimental solutions, permitted an evaluation of the effective molal concentration, m_e, of the experimental solutes, A and U, at a given stoichiometric molal concentration, m_s. The results for the experimental solutes are generally presented in terms of the osmotic coefficient, φ, where

$$\varphi = \frac{m_e}{m_s} \quad (4)$$

Near-Infrared Spectroscopy. Spectra were obtained in the 1.1- to 1.65-μ region with a Cary Model 14CMR spectrophotometer equipped with a 0–0.1–0.2 and 0–1–2 OD slidewire and thermostatted cell adaptors. The temperature of the sample and reference chambers was continuously monitored by a thermistor and bridge assembly manufactured by Yellow Springs Instrument Co. Temperature was controlled to ±0.1° by means of a Haake circulating water bath. Matched pairs of 1.000, 2.000, 5.000, and 10.000 ± 0.002 cm cells from Opticell were employed. Since the solvent was relatively free of absorption bands in the wavelength region of interest, spectra of PS-$CHCl_3$ solutions were read against a reference cell containing pure solvent rather than compensated reference solutions of the type described by Klotz and Franzen (1962).

After subtraction of the appropriate base lines, the spectra observed between 1.100 and 1.650 μ exhibited positive absorption at all wavelengths. Spectral bands were observed between 1.4 and 1.65 μ. Below 1.350 μ, however, these spectra were devoid of structure and showed only a monotonic increase with decreasing wavelength. Such behavior is probably due to unresolved overtones as well as differential scattering and reflection of light by the sample and reference solutions. In order to correct for this effect so that the extinction coefficients of the bands observed above 1.35 μ quantiatively reflected the absorption properties of the groups involved, we consistently translated the base lines upward at 1.330 μ such that they matched the solution spectra between 1.330 and 1.350 μ. This translation process was applied uniformly, as a constant increment amounting to anywhere from 0.01 to 0.02 OD, to the base line throughout the region of the absorption bands of interest. Extinction coefficients were correspondingly calculated as

$$E_\lambda = \frac{OD_\lambda(\text{trans})}{B_0 l} \quad (5)$$

where OD_λ (trans) is the difference in the absorbance of the solution spectra and the *translated* base line at the wavelength, λ. B_0 is the *stoichiometric* molal concentration of the solute species and l is the path length in centimeters.

In principle it would have been better to employ a more accurate correction of the base line obtained from an extrapolation of a log − log plot of the absorbance in regions outside of the absorption bands of interest against wavelength. In practice, however, the value of the translation increment obtained by this latter procedure for the 1.48- to 1.55-μ region differs negligibly from the value obtained by the simple translation at 1.35 μ. This is due to the fact that above 1.2 μ, the nonspecific absorbance is essentially independent of wavelength due to the cancellation of two opposing effects: the light-scattering contribution which would normally result in a

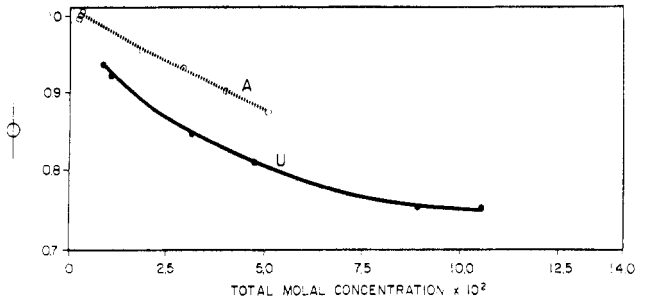

FIGURE 1: The behavior of the osmotic coefficient, ϕ, as a function of the total molal concentration, m_s, of solutes A and U in $CHCl_3$ at 25.0°.

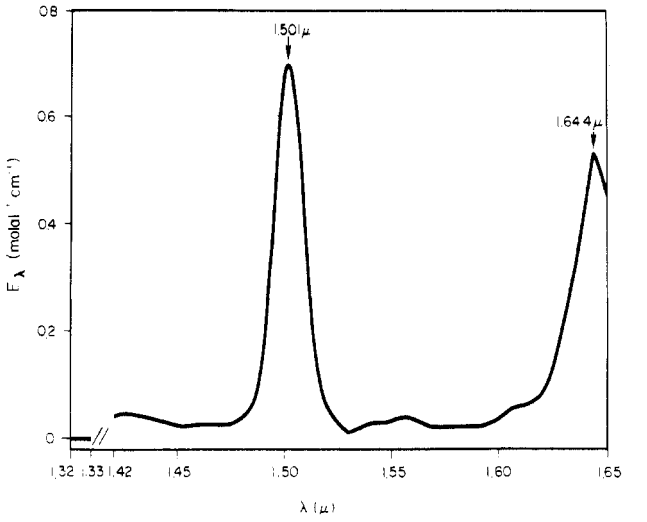

FIGURE 2: The near-infrared spectrum of U. E_λ, calculated by eq 5 in the text, is plotted against wavelength, λ, for a 0.0385 m solution of U in $CHCl_3$ at 25.0°.

decrease in the nonspecific absorbance as the wavelength increases is compensated by an increase in intensity of unresolved overtones (as their order decreases) at longer wavelengths.

Results and Discussion

The results of the vapor pressure osmometry experiments with both A and U showed a continuous decrease in the osmotic coefficient φ, as the stoichiometric concentration of the solute, m_s, increased. This behavior, demonstrated in Figure 1 as a plot of φ *vs.* m_s, is best interpreted in terms of self-association in these solutions since activity effects are not likely to be operative for this concentration range in this particular solvent. At very low concentrations of A, it was possible to obtain a value of $\varphi = 1.00$ within experimental error.[2] This demonstrated that the A preparations were of high purity and stable in the solvent PS-$CHCl_3$. The more dramatic decrease in φ for the U solutions indicates that the self-association of U was more marked. The direct extrapolation of φ to zero concentration was not feasible for these latter solutions.

[2] This was not the case, however, for solutions of A prepared in pure $CHCl_3$ (ethanol and pentene free) where chemical reaction occurs with photoinduced free radicals.

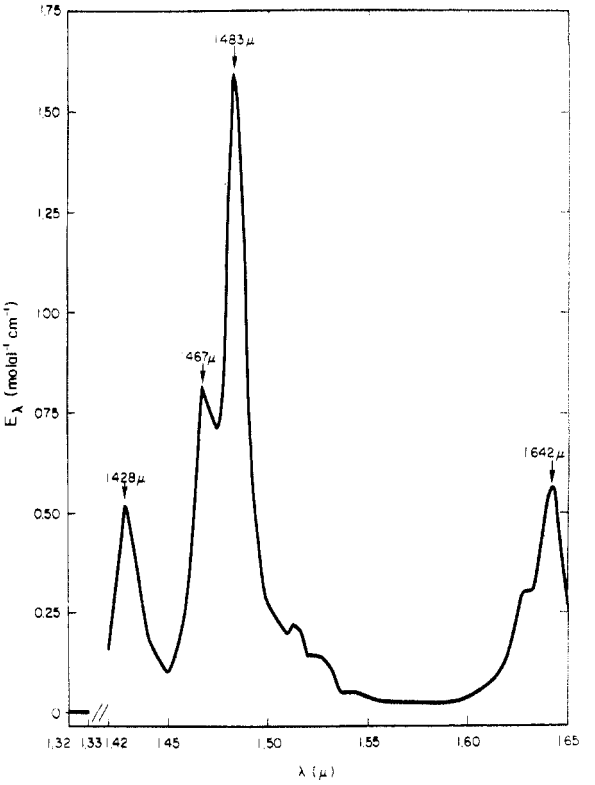

FIGURE 3: The near-infrared spectrum of A. E_λ, calculated by eq 5 in the text, is plotted against wavelength, λ, for a 0.0507 m solution of A in $CHCl_3$ at 25.0°.

TABLE I: Near-Infrared Spectral Characteristics of A and U in $CHCl_3$ at 25°.

Solute	Band Position λ (μ)	Band Position $\bar{\nu}$ (cm^{-1})	$\Delta\bar{\nu}_{1/2}$ (cm^{-1})	E_m (m^{-1} cm^{-1})	Band Assignment
U	1.644	6082		0.530	$2\nu_{C-H}$
U	1.501	6662	71.7	1.04	$2\nu_{N_3-H}$
A	1.642	6090		0.560	$2\nu_{C-H}$
A	1.483	6743	61.0	2.01	$2\nu_{NH_2,sym}$
A	1.467	6817			$(\nu_{NH_2} + ?)$
A	1.428	7003			$2\nu_{NH_2,asym}$

Used in conjunction with the near-infrared data (described below), however, it could be demonstrated that the value of φ for the infinitely dilute U solutions was also 1.00.

Near-infrared spectra of solutions of the two base derivatives were obtained over the same concentration range examined osmotically. Actual spectra of solutions of U and of A are shown in Figures 2 and 3, and their characteristics are summarized in Table I. The intensity of the bands in the 1.5-μ region decreased with increasing concentration but did not change position. The bands near 1.64 μ were invariant with concentration.

Band assignments for these spectra have been made by reference to simpler derivatives as well as to the fundamental spectral region. For both derivatives, the bands near 1.64 μ (1.644 μ (6082 cm^{-1}) in the U spectra and 1.642 μ (6090 cm^{-1})

in the A spectra) have been assigned to the first overtones of CH stretching vibrations. On the basis of their position in the fundamental spectral region (Kyogoku *et al.*, 1967), this is where they would be expected to fall and for the purposes of the present study, a more definitive assignment was not required. It can also be demonstrated that there is no contribution made in this spectral region by NH groups as the position and shapes of these CH bands remained essentially unchanged in D_2O solutions of A and a U model such as α-pyridone.

The concentration invariant character of these CH overtones provided valuable checks on the concentration of the solutions examined spectrophotometrically. Spectra which gave values of E_λ for these bands which differed by more than $\pm 2\%$ from the values reported in Table I were not employed in the calculation of the association constants for the appropriate solution.

The major peak seen in the spectra of U solutions at 1.501 μ (6662 cm^{-1}) is very close to the position of a band in the spectrum of α-pyridone observed at 1.505 μ in dilute solutions. This latter peak is missing in the case of *N*-methylpyridone, whose spectrum is devoid of structure in this region. It is also absent in concentrated solutions of α pyridone under conditions where vapor pressure osmometry reveals that this solute is almost completely dimerized. The ratio, R, between the position of this uracil band (6662 cm^{-1}) and the position of the fundamental N_3H stretching vibration at 3395 cm^{-1} (Kyogoku *et al.*, 1967) is 1.962 which is a value commonly observed for overtone:fundamental ratios of amide NH stretching vibrations in the unassociated state (Hanlon, 1970). Thus, this band has been assigned as the first overtone of the N_3H stretching vibration of uracil in the unhydrogen-bonded or free state.

The spectrum of A in the 1.5-μ region is more complex than that of U. Three actual bands at 1.428 μ (7003 cm^{-1}), 1.467 μ (6817 cm^{-1}), and 1.483 μ (6743 cm^{-1}) together with some minor shoulders can be observed in dilute solution. All of these disappear when the solute is dissolved in D_2O, leaving only the CH overtones at 1.642 μ. Thus, they are all associated with the exocyclic amino group at C_6 in adenine.

Aniline has a set of bands in this spectral region similar in intensity and position, all of which are overtones of the NH_2 functional group in the unassociated state (Whetsel and Lady, 1964; S. Hanlon and E. Major, unpublished). Whetsel and Lady have assigned the high (6693 cm^{-1}) and low (6897 cm^{-1}) wavelength bands to the first overtones of the symmetric and the asymmetric stretching vibrations, respectively. Since the positions of these two vibrations in the fundamental spectral region are 3395 and 3480 cm^{-1} (Bellamy, 1958), the value of R is correspondingly 1.971 for the symmetric stretch and 1.982 for the asymmetric stretch. Using the data of Kyogoku *et al.* (1967) for the band positions of the fundamental stretching vibration of the exocyclic amino group of A in the unassociated state, an R value of 1.974 is calculated for the assignment of the 1.483-μ band as the symmetric stretching overtone. A similar calculation for the assignment of the 1.428-μ band as the asymmetric stretching overtone yields a value of 1.986. The R value of the 1.467-μ band is either too high (1.996) to be an overtone of the symmetric stretching vibration or too low (1.933) to be an overtone of the asymmetric stretching vibration. Since the D_2O studies, however, reveal that it is associated with the NH_2 group, it is probably a combination band involving one of the NH stretching vibrations. Its behavior as a function of concentration parallels that of the 1.428- and 1.483-μ band. Thus, its overlap with the 1.483-μ band does not present problems in employing the latter in spectral analysis.

The extinction coefficient of these NH bands of both A and U decreased with increasing concentration, as might be expected if the association indicated by the osmometry results involved the interaction of NH groups *via* hydrogen bonding. This was also evident in the earlier work of Kyogoku *et al.* (1967). The fundamental spectra of the self-association process (Kyogoku *et al.*, 1967) reveal that the bands due to the hydrogen-bonded NH groups are shifted to longer wavelengths. Hence, their overtones (even if of sufficient intensity) would be expected to fall above 1.51 μ and thus not interfere in the spectral range of interest. The 1.501-μ peak was employed in analysis of the U solutions and the 1.483 μ (because of its higher extinction coefficient) for analysis of the A spectral data. A careful analysis of the half-bandwidth of both peaks as a function of concentration revealed that they were concentration invariant. This demonstrated that there was no significant overlap of any of the fully or partially hydrogen-bonded bands[3] of the NH group and that, furthermore, the extinction coefficient of the 1.467-μ peak in the adenine spectrum changed in a fashion which paralled the change in the extinction coefficient of the 1.483-μ band with concentration. The extinction coefficients, E, at 1.501 and 1.483 μ in separate solutions of U and A, respectively, were thus assumed to be related to the molal concentration of free base derivative, B_1, in the equation

$$\frac{E}{E_m} = \frac{B_1}{B_0} \tag{6}$$

where B_0 is the stoichiometric molal concentration of the base derivative and E_m is the extinction coefficient of the free base monomer or, alternatively, the value of E in infinitely dilute solutions (in units of $m^{-1}\ cm^{-1}$).

As was the case with the osmometry experiments, the association of A was sufficiently weak such that the value of E_m could be obtained experimentally from direct measurements of very dilute A solutions as well as by extrapolation of E *vs.* A_0 to 0 concentration. The value so obtained is given in Table I.

This procedure was not feasible for U, however, as the curve of E *vs.* U_0, changed too rapidly near 0 concentration to permit a reliable extrapolation to be made. The value of E_m, however, for uracil could be evaluated with the aid of the osmotic data. In dilute solutions where the bulk of the base derivative is in the form of unassociated species (or monomer) and associated dimers, a simple linear relationship may be derived linking the osmotic coefficient, φ, and the spectral extinction coefficient, E. Since mass must be conserved in the dimerization reaction, then

$$m_s = B_0 = B_1 + 2B_2 \tag{7}$$

where the symbols, m_s and B_0, have been previously defined, and the subscripts, 1 and 2, indicate monomer and dimer, respectively.

[3] In the case of adenine, the overtone of the free NH of the singly hydrogen bonded exocyclic amino group of adenine would be expected to fall at about 1.46 μ, and, in fact, in the more highly concentrated A solutions, a small shoulder is evident at the lower wavelength side of the 1.467-μ peak. This was sufficiently far removed from the 1.483-μ position, however such that significant overlap did not occur

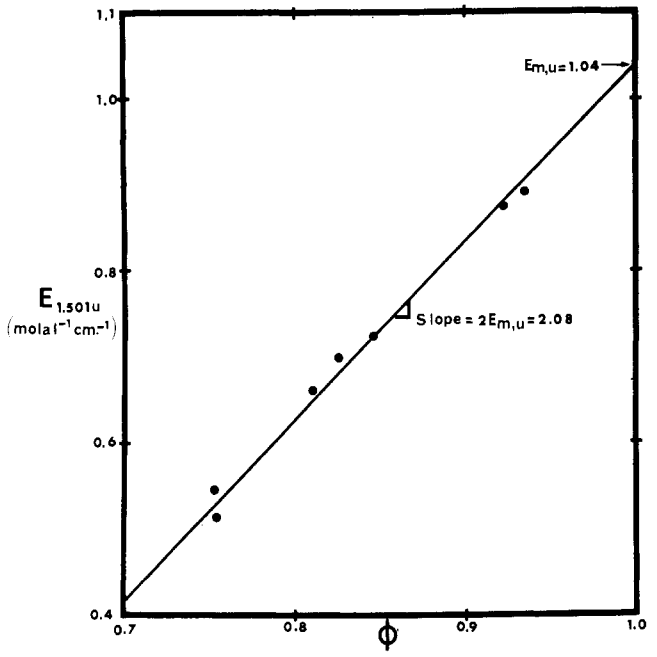

FIGURE 4: Determination of the monomer extinction coefficient, E_m, for U in $CHCl_3$ at 25.0°. Extinction coefficient data for U solutions are plotted against their osmotic coefficients, ϕ, according to eq 10 in text.

The effective molal concentration, m_e, determined by the osmometer, is

$$m_e = B_1 + B_2 \tag{8}$$

and the osmotic coefficient, φ, is given by

$$\varphi = \frac{m_e}{m_s} = \frac{B_1 + B_0}{2B_0} \tag{9}$$

Combining this expression with eq 6 yields

$$E = 2E_m(\varphi) - E_m \tag{10}$$

Thus, a plot of E of a given solution against its osmotic coefficient, φ, should yield a straight line with a slope, $2E_m$, an intercept of $-E_m$ at $\varphi = 0$ and an intercept of $+E_m$ at $\varphi = 1$ in the concentration ranges in which only monomers and dimer exist. If the latter condition is not met or if the solute is not pure, then it is unlikely even in the event that a linear relationship is observed, that the slope and intercept will yield the same value of E_m.

Figure 4 shows such a plot for U in which the agreement between the value of E_m of 1.04 (m^{-1} cm^{-1}) evaluated from both the slope and the intercept at $\varphi = 1.00$ is excellent. A similar plot for A also showed linear behavior but the E_m values from slope and intercept were not consistent nor did they agree with the value determined directly. In the light of subsequent analysis of the association process of A, this could be attributed to the presence of higher order aggregates in the A solution.

Evaluation of the Association Constants. Both the osmometry and the spectral data may be used independently to calculate association constants. In the case of the osmometry data, the dimerization constant, $K_{1,2}$, for the reaction

$$2B_1 \overset{K_{1,2}}{\rightleftharpoons} B_2 \tag{11}$$

was calculated by taking a limit of an equation derived by Schrier (1968)

$$K_{1,2} = \lim_{m_s \to 0} K_{1,2}\,(\text{app}) = \lim_{m_s \to 0}\left[\frac{m_s - m_e}{(2m_e - m_s)^2}\right] \tag{12}$$

where $K_{1,2}$ (app) is the apparent dimerization constant calculated for each m_s using the function on the right-hand side of eq 12. If the association process stops at the dimer stage, the values of $K_{1,2}$ (app) will be invariant with concentration. Values of $K_{1,2}$ (app) which increase with increasing concentration are diagnostic of the existence of higher order aggregation reactions taking place as well.

As demonstrated by the data displayed in the second column of Table II, the dimerization constants calculated in this fashion for U did not vary in any consistent fashion for the concentration range examined and it was concluded that, on the basis of the osmometry data, at least the association process for U stopped at the dimerization stage. The average of the values obtained was 10.1 ± 0.5 m^{-1}. Similar analysis of the A data, however, showed increasing values of $K_{1,2}$ (app) as is demonstrated in Figure 5. Presumably this behavior is due to the presence of higher order aggregates. A least-mean-square extrapolation of these data to infinite dilution to obtain $K_{1,2}$ gave a value of 1.88 m^{-1}.

TABLE II: Values of the Dimerization Constant, $K_{1,2}$ for U in $CHCl_3$ at 25°.

$m_s = U_0$ (m)	Osmometry Results: Calcd by eq (12) (m^{-1})	Spectral Results: Calcd by eq (14) (m^{-1})
0.1056	9.0	19.7
0.08906	10.8	19.4
0.04736	10.3	19.3
0.03840	10.7	19.0
0.03168	10.2	19.9
0.01081	9.9	20.9
0.00896	9.5	19.8
	Av = 10.1 ± 0.5 m^{-1}	Av = 19.8 ± 0.4 m^{-1}
	$K_{1,2}$ = 10.1 ± 0.5 m^{-1}	$K_{1,2}$ = 9.9 ± 0.4 m^{-1}

The dimerization constant was also calculated from the spectral data by applying a relationship derived by Klotz and Franzen (1962)

$$K_{1,2} = \lim_{\alpha \to 0}\left(\frac{\alpha}{1-\alpha}\right)\frac{1}{C_f} \tag{13}$$

where α is the fraction of hydrogen-bonded species and C_f is the concentration of free or nonassociated molecules. This expression was formulated for a simple linear aggregation process for which, in the limit, one hydrogen bond is formed per dimer. If the initial dimer formed is cyclic such that the absorbance of two functional groups is affected (and the molal concentration of bonded groups is $(C_{total} - C_f)/2$ rather than $(C_{total} - C_f)$ one can readily show that the intercept obtained from this extrapolation is $2K_{1,2}$.

If this assumption of cyclic dimerization is incorporated

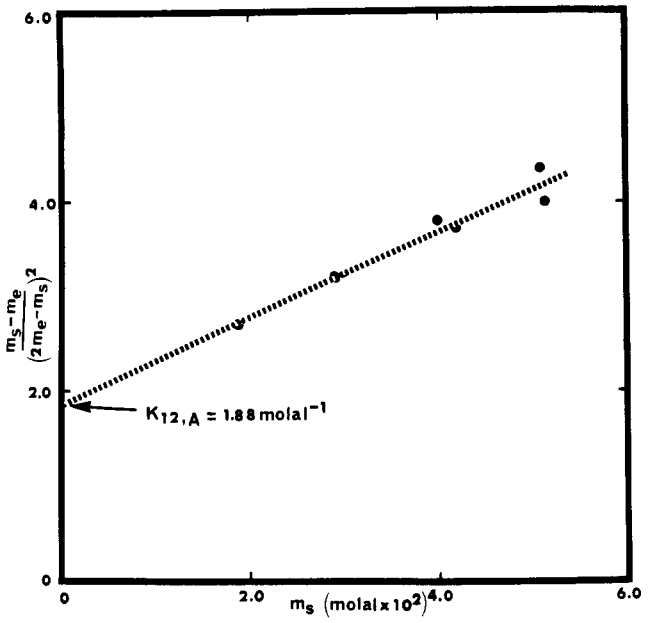

FIGURE 5: Determination of the dimerization constant, $K_{1,2}$, for A in $CHCl_3$ at 25.0° using vapor pressure osmometry results. The data are analyzed according to eq 12 in the text. The dotted line represents the least-mean-square straight line through all points.

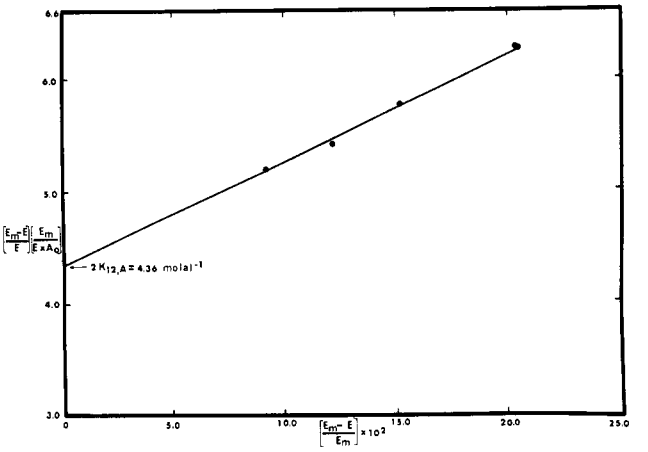

FIGURE 6: Determination of the dimerization constant, $K_{1,2}$, for A in $CHCl_3$ at 25.0° using near-infrared spectral results. The data are analyzed according to eq 14 in the text. The solid line represents the least-mean-square line drawn through all points.

and eq 13 is recast in terms of our experimental variables, we obtain the expression

$$2K_{1,2} = \lim_{\left(\frac{E_m - E}{E_m}\right)\to 0} 2K_{1,2}(\text{app}) = \lim_{\left(\frac{E_m - E}{E_m}\right)\to 0} \left[\left(\frac{E_m - E}{E}\right)\left(\frac{E_m}{EB_0}\right)\right] \quad (14)$$

If the association does not proceed beyond the cyclic dimer state, then a plot of $[(E_m - E)/E][E_m/EB_0]$ *vs.* $[(E_m - E)/E_m]$ will have O slope and will yield the value of $2K_{1,2}$ at any concentration. When higher order complexes are formed whose associative constants are greater than $2K_{1,2}$, a large positive slope should be observed. Agreement between the value of $K_{1,2}$ determined in this fashion with that obtained from the osmometry data will confirm the fact that the dimers formed are cyclic.

The calculated values for U determined in this fashion are assembled in the third column of Table II. They show no systematic variation with concentration. In the light of the osmometry results, we interpret this to mean that the association reaction for U stops at the dimer stage. The average value of $[(E_m - E)/E][E_m/EB_0]$ for U is 19.8 ± 0.4 m which is approximately twice the value of $K_{1,2}$ determined by osmometry. The obvious conclusion is that U forms cyclic dimers in solution with an average $K_{1,2}$ of 10.0 ± 0.5 m^{-1}.

The value of $[(E_m - E)/E][E_m/EB_0]$ in eq 14 increased with increasing concentration of A in solution as is demonstrated by the data shown in Figure 6. This reflects the fact that aggregation occurs beyond the dimer stage for that solute. The intercept at $[(E_m - E)/E_m] = 0$ was 4.36 m^{-1}, again a value approximately twice that from the osmometry results. We have thus concluded that the dimer association of A is also cyclic and is characterized by an average $K_{1,2}$ = 2.0 ± 0.2 m^{-1}.

Both the osmometry and the spectral results reveal that higher order association reactions are occurring in the A solutions. The spectral studies suggest that the constant(s) for the formation of trimers and other aggregates are greater than $K_{1,2}$. The simplest way to analyze these data is to assume that the equilibrium constants for the successive reactions

$$\begin{aligned} B_2 + B &\overset{K_{2,3}}{\rightleftharpoons} B_3 \\ B_3 + B &\overset{K_{3,4}}{\rightleftharpoons} B_4 \\ B_{n-1} + B &\rightleftharpoons B_n \end{aligned} \quad (15)$$

are equal. That is

$$K_{1,2} < K_{2,3} = K_{3,4} = \ldots = \bar{K} \quad (16)$$

For this case, the osmometry data were analyzed by an equation derived by Davies and Thomas (1956a,b)

$$\bar{K} = \left(\frac{1}{m_e}\right) - \left\{\left[2 - \left(\frac{1}{\varphi}\right)\right]\left[\frac{K_{1,2}}{m_s - m_e}\right]\right\} \quad (17)$$

The analysis of the spectral data employed an equation of Coggeshall and Saier (1951), which, cast in our terms, becomes

$$\frac{1}{\bar{K}} = \left[\frac{(1 - \alpha)B_0}{2(K_{1,2} - K_m)}\right]\left[2K_{1,2} - \left(\frac{K_m}{2}\right) + \left(2K_{1,2}K_m + \frac{K_m^2}{4}\right)^{1/2}\right] \quad (18)$$

in which α is as previously defined and

$$K_m = \frac{2(1 - \alpha)^2 B_0}{\alpha} \quad (19)$$

Again, this analysis is appropriate for an association reaction in which one hydrogen bond is formed.

In both the osmometry and the spectral experiments the precision was insufficient to actually demonstrate the existence of any species higher than the trimer. Equations 17 and

TABLE III: Values of the Association Constant for the Formation of Higher Order Species, $\bar{K}_A$, for A in $CHCl_3$ at 25°.

	Osmometry Results	Spectral Results
$m_s = A_0$	Calcd by Eq 17 $(m^{-1})^a$	Calcd by Eq 18 $(m^{-1})^a$
0.05130	6.4	6.3
0.05095	7.2	6.3
0.0420	7.1	
0.04006	7.6	6.3
0.0290	7.6	7.0
0.0187	7.0	8.1
	$\bar{K}_A = 7.3 \pm 0.4$	$\bar{K}_A = 6.8 \pm 0.6$

[a] A value of $K_{1,2} = 2.0\ m^{-1}$ was employed.

TABLE IV: Summary of the Self-Association Characteristics of A and U in $CHCl_3$ at 25°.

		Av Equil Constant		
		This Work		Kyogoku *et al.* (1967)
Process	Constant	m^{-1}	M^{-1}	M^{-1}
$2U \rightleftarrows U_2$	$K_{1,2,U}$	10.0 ± 0.5	6.8	6.1
$2A \rightleftarrows A_2$	$K_{1,2,A}$	2.0 ± 0.2	1.4	3.1
$A_2 + A \rightleftarrows A_3$ $A_3 + A \rightleftarrows A_4$ etc.	$\bar{K}_A$	7.1 ± 0.5	4.8	None

18 may still be employed, however, even though the actual reactions shown in eq 15 stop at the trimer level. In this event, $\bar{K}$ is simply $K_{2,3}$.

The calculation of $\bar{K}$ by these equations is quite sensitive to the value of $K_{1,2}$ employed. In order to compare the $\bar{K}$ values determined by the two methods, it was deemed advisable to use an average value of $K_{1,2} = 2.0$. The values of $\bar{K}$ so calculated for the A association are shown in Table III. The slight trend toward decreasing $\bar{K}$ with increasing concentration is not considered significant. The average value obtained from osmometry was $7.3 \pm 0.4\ m^{-1}$ and from spectrophotometry was $6.8 \pm 0.6\ m^{-1}$.

The agreement between the values of $\bar{K}$ obtained from the two methods does not necessarily imply that the higher order complexes of A are linear. The absorbance at 1.483 μ in the spectrum of A depends upon the freedom of both amino hydrogens and is diminished whether one or both of these hydrogens are hydrogen bonded. When successively higher complexes are formed in the reactions shown in eq 15, the absorbance of only a single NH_2 group—that of the monomer—is altered, no matter whether a linear or a cyclic complex is formed. The former possibility, however, seems unlikely, especially in view of the fact that $\bar{K} > K_{1,2}$. Also, it is theoretically possible to form U_3 complexes by the formation of one hydrogen bond between one of the unused carbonyl oxygen of the U dimer and the NH of another U monomer. The fact that only dimers of U are observed even at high concentrations strongly suggests that reactions involving the formation of a single hydrogen bond do not take place in these systems to a measurable degree.

A summary of the association constants for these various reactions of A and of U is presented in Table IV in both molal^{-1} and molar^{-1} units to facilitate comparison with the values shown in the last column taken from previous studies of Kyogoku *et al.* (1967) in the fundamental region of the infrared. The results of these workers agree very well with ours in the case of the U association, but are in marked disagreement with respect to the A association which they claimed did not proceed measurably beyond the dimer stage. This disagreement, however, can be shown to be a function of the method of data analysis. Kyogoku *et al.* employed a method which relied on a simultaneous determination of both E_m and $K_{1,2}$ by the application of an equation which is appropriate for the case where monomers and dimers only exist in solution. A linear behavior of the particular plot employed was taken as evidence that the assumptions employed were valid. The values of $K_{1,2}$ and $\bar{K}$ for adenine, however, are such that this plot can appear to be linear in the concentration range examined by these authors even though higher order aggregates are present in solutions. If we use a similar analysis for the near-infrared data for A, we obtain a value of $K_{1,2}$ of $4.1 \pm 0.4\ m^{-1}$ or 2.7 M^{-1} in the same concentration range examined by these authors. The value of E_m determined from this plot, however, does not agree with that which was determined directly. As might be expected, however, the plot of the U data was linear throughout the concentration range examined and gave the same values of E_m and $K_{1,2}$ as determined by the previously described methods.

It might be argued that the presence of pentene affects the mode of association as well as the magnitude of the equilibrium constants or, as a two-component solvent, introduces certain complications into the calculation which were ignored. The good agreement between the results of Kyogoku *et al.* (1967) and ours when similar methods of data anaylsis were employed demonstrates that these were not serious considerations. In addition, all of the studies reported in this paper as well as paper II of this series were initially conducted in pure $CHCl_3$ purified by extraction with sulfuric acid and water and distilled over K_2CO_3. The results of such studies yielded equilibrium constants which were reasonably close to the values reported in Table IV, but whose precision was very poor due to reaction of the solutes with photodecomposition products of the solvent.

Conclusion

These results clearly demonstrate that adenine and uracil derivatives form cyclic dimers in solution which, in the case of adenine, go on to form trimers and, perhaps, higher order complexes. It is quite likely that these species above the dimers are also cyclic. The existence of such species has been suggested by Katz and Penman (1966) and Katz (1969) on the basis of nuclear magnetic resonance studies although their data were insufficient to evaluate equilibrium constants.

The average value of $\bar{K}$ determined in our experiments with A is 3.5 times greater than $K_{1,2}$. If we assume that the trimers, for instance, are of the cyclic type analogous to the cyclic dimers, and that both sites, (C_6-NH_2, N_1) and (C_6-NH_2, N_7), are equally effective in forming hydrogen bonds with no interaction between sites, a simplified statistical analysis predicts that the ratio of $K_{2,3}/K_{1,2}$ should be ($^1/_2$). The dimer can be

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.