Finally, a plot of log k_{obs} − log ($[SH^+]/([S] + [SH^+])$) + H_0 against log a_{H_2O} for 1-methoxy-3,5-dihydroxybenzene was fairly linear. The slope of the line is the w-value based on a two-proton process.[6,24] Its value, +2.0, is within the range assigned for reactions in which water acts as a nucleophile in the rate-controlling step.[6b] It is, however, less than for the hydrolysis of 4-(*p*-sulfophenyl)-1-methoxynaphthalene (w was +4.5 in 1–8 molar hydrochloric acid and +8.2 in 1–4 molar perchloric acid) or of 4-(*p*-sulfophenylazo)-1-anisole (w was +3.1 in 1–9 molar hydrochloric acid).[5,6] The corresponding w-plot for 1,3,5-trimethoxybenzene was badly curved for the results at 90°, and the points were widely scattered for the results at 60°, the slopes being generally positive.

(24) In order for a "true" w-value to be obtained, it is necessary that the second protonation follow H_0.

Activation Energy and Entropy.—The rate of hydrolysis of 1-methoxy-3,5-dihydroxybenzene was measured at five temperatures in 63.5% perchloric acid (Table I). The high acid percentage was chosen in order that the ground state be essentially carbon-conjugate acid, SH^+. The value of E_A obtained is 18.5 ± 0.2 kcal./mole and $\Delta S^{\ddagger}$ at 25° is −15.2 ± 0.7 e.u. A large negative $\Delta S^{\ddagger}$ is considered to be characteristic of reactions classified as A-2,[3] but it is not known whether this should apply also to a two-proton process.

Acknowledgment.—The authors gratefully acknowledge the support of this work by the National Science Foundation.

[CONTRIBUTION FROM THE CALIFORNIA INSTITUTE OF TECHNOLOGY, DIVISION OF BIOLOGY, PASADENA, CALIF.]

Interaction and Association of Bases and Nucleosides in Aqueous Solutions[1]

BY PAUL O. P. TS'O,[2] INGELORE S. MELVIN AND ALFRED C. OLSON[3]

RECEIVED OCTOBER 31, 1962

The molal osmotic coefficients (ϕ) of aqueous purine, uridine and cytidine in the concentration range of 0.1 to 1.0 molal, and inosine and caffeine at 0.1 molal, have been determined by thermoelectric measurements of vapor pressure lowering at 25°. The activity coefficients of purine, uridine and cytidine were calculated. The data indicate that these solutes associate extensively in solution and that the association process does not proceed simply to the dimer stage, but continues to form higher polymers. The results are consistent with a set of association processes which have equal equilibrium constants for all successive steps. Equilibrium constants at 25° and standard free energies for association of purine, uridine and cytidine were found to be 2.1, 0.6, 0.9 molal^{-1} and −440, 290, 80 cal./mole, respectively. The interaction of one base with another base was examined by measuring the increase of solubility of adenine-C^{14} or thymine-C^{14} in the presence of varying concentrations of a variety of interactants at 25.5° and 38°. The solubility of adenine or thymine was increased by the addition of soluble purine, nucleosides, pyrimidine or phenol, and was essentially unchanged by the addition of cyclohexanol, adonitol or urea. The total base-C^{14} in solution was assumed to be composed of the free base in solution and the water-soluble base-interactant complex. Equilibrium constants for these association processes have been estimated, and they agree semiquantitatively with the results from measurements of ϕ. The tendency for association and interaction of the free bases and nucleosides in solution can be arranged in the series: purine–purine > purine–pyrimidine > pyrimidine–pyrimidine. The relationship of this finding to the vertical-stacking interaction of the bases in nucleic acids is discussed.

Introduction

In our laboratories, research has been conducted on factors contributing to the formation and stabilization of helices of nucleic acids.[4] It is felt that the attraction and the specificity in the formation of helices by strands of nucleic acids must reside in the interaction of the nitrogenous bases on the polynucleotide chains. Thus it is pertinent to ask whether solutes interact in aqueous solutions of the free bases and nucleosides, and if so, to what extent, with what degree of specificity and with what mechanism? In order to answer these questions, the thermodynamic properties and solution behavior of certain bases and nucleosides in water have been studied. Uncharged compounds at neutral *p*H, such as bases and nucleosides, were chosen in order to avoid the complication due to charge effects.

In the first part of this paper, the osmotic coefficients and the activity coefficients of purine (0.05 to 1.1 m), uridine (0.1 to 0.7 m), cytidine (0.1 to 0.7 m), caffeine (0.1 m) and inosine (0.1 m) at 25° have been determined. The data indicate a high degree of association of the bases and nucleosides in solution. Apparent equilibrium constants for this association are evaluated for purine, uridine and cytidine. In the second part of the paper, the solubilities of adenine and thymine in water at 25.5° in the presence of bases (purine and pyrimidine), nucleosides (uridine and cytidine) and of other compounds of interest are studied. The solubilities of adenine and thymine are greatly enhanced in the presence of these bases and nucleosides, demonstrating that there are interactions between adenine and thymine with these compounds.

Experimental

Materials.—Commercially available compounds of the highest degree of purity were used without further purification. The following A grade (unless specified otherwise) compounds were obtained from California Corporation for Biochemical Research, Los Angeles: adenine, adenine-8-C^{14}, thymine, thymine-2-C^{14}, adonitol, inosine, caffeine (U.S.P. C grade), purine (nitrogen 46.57%, calculated purity 99.85%,[5] containing 0.3% chloride as stated by the manufacturer), cytidine (nitrogen 17.46%[6]) and uridine (nitrogen 11.28%, calculated purity 98.4%[5]). Pyrimidine and adenine were obtained from Nutritional Biochemical Corporation. Some crystalline thymine from Sigma Co., St. Louis, Mo., as well as some crystalline purine from Cyclo Chemical Corporation were also used. Phenol (chromatographic, 88%) was obtained from Mallinckrodt Co., and cyclohexanol (reagent grade) was obtained from Fisher Scientific Co. All other chemicals were reagent grade. Distilled, de-ionized water was used.

Adenine-C^{14} and thymine-C^{14} for solubility studies were prepared by dissolving proper amounts of unlabeled base and base-C^{14} together to give the desired specific activity. Specific activities were determined from counts and optical density measurement (Beckman DK-2 or Cary-11 spectrophotometer). The

(1) The research was supported in part by grants RG-3977, RG-5143 and GM 10316-01 from the National Institutes of Health, U. S. Public Health Service.

(2) On leave of absence from Department of Radiological Sciences, School of Public Health, The Johns Hopkins University, Baltimore, Md.

(3) Western Regional Research Laboratory, Agricultural Research Service, U. S. Department of Agriculture, Albany, Calif. Work conducted under a coöperative agreement with the California Institute of Technology. Mention of trade names, products or manufacturers does not imply endorsement by the U. S. Department of Agriculture over others not mentioned.

(4) P. O. P. Ts'o, G. H. Helmkamp and C. Sander, *Proc. Natl. Acad. Sci. U.S.*, **48**, 686 (1962).

(5) This is calculated on the basis that no other nitrogenous impurities are present.

(6) The nitrogen content exceeds that of the theoretical value of 17.27% based on the chemical formula.

values of 13.5×10^3 and 7.89×10^3 were taken to be the molar extinction coefficients of adenine in water and thymine in 0.01 N HCl, respectively.

Radioactivity Determination.—Radioactivity was assayed with a Tri-Carb liquid scintillation spectrometer (model 314DC) with automatic sample changer, Packard Instrument Co. Aqueous aliquots (0.2 ml.) were counted in glass vials in 20 ml. of scintillation mixture consisting of 10 ml. of absolute ethanol and 10 ml. of reagent grade toluene containing 70.5 mg. of 2,5-diphenyloxazole and 2.0 mg. of 1,4-bis-2-(5-phenyloxazolyl)-benzene.

Solubility Measurements.—Weighed base-C^{14} (5–15 mg.) together with 1 ml. of water was sealed in a 10-ml. glass ampoule. These sealed ampoules were heated in a 70–75° water-bath for 30 minutes. Most of the material was solubilized by this treatment and formed a solution supersaturated with reference to the solution at lower temperature (25.5° or 38°) at which the solubility was measured. A single determination of the solubility required six ampoules. The remaining three were stored at 4° for 2 days, after which they were put in the equilibrating bath. By storing at 4° much of the base precipitated to form an unsaturated solution with respect to the equilibrium solubility at 25.5° or 38°. The ampoules were shaken vigorously in the equilibrating bath for 7 days, after which they were opened and the contents filtered immediately under a slight nitrogen pressure through a fine sintered glass funnel. The change of the solubility of these bases in response to the change of temperature was found to be very slow.

Careful studies of the conditions required to reach equilibrium were made. That equilibrium was established under these conditions is shown by the fact that the same solubility was obtained from the supersaturated samples after heat treatment and from the unsaturated samples after cold storage. The *p*H of these solutions was between 5 and 6.5. No visible sign of biological contamination in any sample was found. For example, there was no increase in turbidity as measured by optical density at 400 mμ for a ribose–adenine solution after immersion in the bath for 11 days.

Measurement of the Osmotic Concentrations of Purine and Nucleosides by Vapor Pressure Lowering.—Vapor pressure lowering was measured by the osmometer model 301A, manufactured by Mechrolab Inc., Mountain View, Calif. The instrument was built according to principles and practices previously described,[7-9] and employs thermistors for the measurement of a temperature differential.[10,11]

Reagent grade sodium chloride was used as standard, and the molal osmotic coefficients (ϕ) were taken from the Appendix in Robinson and Stokes.[12] A calibration curve of the instrument in terms of ΔR (ohm) *vs.* molal osmotic concentration, $\nu m\phi$ (moles per 1000 g. of water) was constructed up to 0.70 molal. The slope of the straight line was found to be 1.25×10^{-2} $\nu m\phi$ per ohm. Variation in measurement of the same solution in 3–4 successive trials was usually less than 4×10^{-2} ohm or 5×10^{-4} $\nu m\phi$. This calibration curve was checked after each set of measurements and has been found to be stable for longer than 8 weeks. Earlier, sucrose was also used as standard and yielded a curve practically identical with that from sodium chloride. The sucrose solution, however, was found to be more unstable during storage and therefore only the sodium chloride was used as standard. Steady-state of the thermal difference was assured by taking 2-, 4- and 6-minute readings which usually did not vary more than 3×10^{-2} ohm.

Chemicals used for preparation of molal solutions were dried in a vacuum oven connected to a Dry Ice trap for 48 hours at 50–60°.

Definitions[13] and Equations Related to Osmotic Coefficient Determinations.—The osmotic concentration is defined as $\nu m\phi$, where ν is the number of species in solution (two for sodium chloride and one for purine, nucleosides, etc.), m is the molality and ϕ is the molal osmotic coefficient.[13]

The relationship between the molal osmotic coefficient, ϕ, and the molal activity coefficient, γ, is given by the Gibbs–Duhem equation and can be expressed in the form

$$\ln \gamma = (\phi - 1) + \int_0^m (\phi - 1) \mathrm{d} \ln m \qquad (1)$$

The relation of ϕ and m was expressed by an empirical polynomial (eq. 2) calculated from data by the method of least squares.[14,15]

$$\phi = 1 + a_1 m + a_2 m^2 + a_3 m^3 + a_4 m^4 \qquad (2)$$

By direct integration of the empirical polynomials (eq. 2) in eq. 1, the following relationship of γ and m can be obtained.[16]

$$\ln \gamma = 2a_1 m + 3/2 a_2 m^2 + 4/3 a_3 m^3 + 5/4 a_4 m^4 \qquad (3)$$

As reported and discussed in later sections of this paper, a large lowering of the osmotic and activity coefficients of purine and nucleoside solutions has been observed. In these cases, it is reasonable to assume that association or polymerization of the purine and the nucleosides occurs as

$$\begin{aligned} B_1 + B_1 &\rightleftharpoons B_2 & K_1 &= B_2/B_1^2 \\ B_2 + B_1 &\rightleftharpoons B_3 & K_2 &= B_3/(B_2)(B_1) \\ &\cdots\cdots & &\cdots\cdots \\ B_n + B_1 &\rightleftharpoons B_{n+1} & K_n &= B_{n+1}/(B_n)(B_1) \end{aligned} \qquad (4)$$

For further simplification, it was also assumed that all the K's are equal.

$$K_1 = K_2 = K_3 \ldots K_n = K \qquad (5)$$

Based on these assumptions Schellman[17] developed the following equation (eq. 6) for the analyses of urea association.

$$K = (1 - \phi)/m\phi^2 \qquad (6)$$

where m is the over-all molal concentration (stoichiometric molality). Alternatively, the data can be treated as if only dimer is formed. In this case, the solution only contains B_1 and B_2, and $K_2, K_3 \ldots K_n$ are all equal to zero. With this assumption and the approach of Schellman,[17] eq. 7 can be derived.

$$K = K_1 = \frac{(1 - \phi)}{m(2\phi - 1)^2} \qquad (7)$$

The calculation of the standard free energy change of each consecutive step of this association is based on eq. 8.

$$\Delta F^0 = -RT \ln K \qquad (8)$$

If eq. 5 and 6 are valid, then the concentration of the solute at various stages of association is related to K and the free solute, B_1, as derived from eq. 4, as

$$\begin{aligned} B_2 &= K(B_1)^2 \\ B_3 &= K^2(B_1)^3 \\ &\cdots\cdots \\ B_n &= K^{n-1}(B_1)^n \end{aligned} \qquad (9)$$

The stoichiometric molality m can be expressed as

$$m = B_1 + 2B_2 + 3B_3 + \ldots + nB_n$$

and from 9

$$\begin{aligned} m &= B_1 + 2KB_1^2 + 3K^2B_1^3 + \ldots + nK^{n-1}B_1^n \qquad (10) \\ &= \frac{B_1}{(1 - KB_1)^2} \end{aligned}$$

Equations 9 and 10 provide all the necessary relationships to compute the population distribution of the solute at various stages of association in solutions of stoichiometric molality, m.

Definitions and Equations Related to Solubility Measurements.—The condition for saturation of a solution is that the chemical potential of the solution is the same in the solid state as in the saturated solution (eq. 11)

$$\bar{F}_{\text{solid}} = \bar{F}^0_B + RT \ln (Sy) \qquad (11)$$

where S is the solubility expressed in the same concentration units as the activity coefficient y (moles per liter). The standard state is that in infinitely dilute solution for all solutes. In the presence of another interacting substance, the solubility of the base may be different but will still be determined by this condition (eq. 11a).

$$\bar{F}_{\text{solid}} = \bar{F}^0_B + RT \ln (S'y') \qquad (11a)$$

From eq. 11 and 11a, the relationship of the activity coefficient ratio of the base in the presence and absence of the interacting substance to the respective solubility ratio can be derived (eq. 12).

$$\ln (S'/S) = \ln (y/y') \qquad (12)$$

Another quantity, solubility increment, $\Delta S/I$, was found to be useful in analyses of the solubility data; $\Delta S/I$ is defined as the

(7) A. V. Hill, *Proc. Roy. Soc.* (London), **A127**, 9 (1930).

(8) E. J. Baldes, *J. Sci. Instr.*, **11**, 223 (1934).

(9) E. J. Baldes, *Biodynamica*, No. 46 (1939).

(10) A. P. Brady, H. Huff and J. W. McBain, *J. Phys. and Colloid Chem.*, **55**, 304 (1951).

(11) H. Huff, J. W. McBain, and A. P. Brady, *ibid.*, **55**, 311 (1951).

(12) R. A. Robinson and R. H. Stokes, "Electrolyte Solutions," Academic Press, Inc., New York, N. Y., 1959, Appendix 8.3.

(13) The convention and symbols are essentially the same as those employed by Robinson and Stokes.[12]

(14) W. E. Milne, "Numerical Calculus," Princeton University Press, Princeton, N. J., 1949, Chapter 9.

(15) The actual calculation was performed by the Computer IBM 7090, in the Computing Center, California Institute of Technology. We would like to acknowledge the assistance of K. Hebert, S. Caine and R. Deverill in setting up the program for the computer.

(16) The evaluation of the activity coefficient γ from eq. 3 was also done by computer.

(17) J. A. Schellman, *Compt. rend. trav. lab. Carlsberg, Sér. Chim.*, **29**, 223 (1956).

increase in solubility of the base (moles/liter), ΔS, per molar concentration of the interactant added to the solution (I). The large increase of solubility or the large lowering of the activity coefficient in the presence of a relatively low concentration (less than one molar) of interactant (I), suggests that water-soluble complexes of the base and I may be formed. It is assumed that the saturating concentration of the base dissolved in the presence of I can be represented by eq. 13

$$[B]_{total} = [B]_{solubility} + [B]_{complex\text{-}I} \qquad (13)$$

where $[B]_{solubility}$ is the concentration of the free base and is equal to S, while $[B]_{complex\text{-}I}$ is the concentration of the base existing in the form of a complex with I. By making additional assumptions, the equilibrium constant of this complex formation can be further analyzed.

Complex formation between the base-C^{14} and the interactant can be expressed as

$$\begin{aligned} I_1 + B &\rightleftharpoons BI_1 & K' &= \frac{BI_1}{[B][I'_1]} \\ I_2 + B &\rightleftharpoons BI_2 & K'_2 &= 2K' \qquad (14) \\ \cdots\cdots\cdots\cdots & & \cdots\cdots\cdots & \\ I_n + B &\rightleftharpoons BI_n & K'_n &= nK' \end{aligned}$$

where B is the concentration or the solubility, S, of the free base, and I_1, $I_2 \ldots$, I_n are the concentrations of the interactant at varying degrees of polymerization. In this formulation, it is assumed that the base can interact with the free form of interactant and the associated form of interactant with the same ease. The total concentration of all the species of $BI_{1,2} \ldots n$ has been defined and measured experimentally as ΔS, which can be expressed in accordance with eq. 14 as

$$\Delta S = \sum_{n=1}^{\infty} BI_n = [B]K'_n \sum_{n=1}^{\infty} nI_n \qquad (15)$$
$$= [B]K'I$$

where I is the stoichiometric molarity of the interactant added. Thus, if eq. 14 is correct, $\Delta S/I$ will be a constant at varying concentrations of interactant since B has a constant value and is measurable. Hence, the association constant, K', can be calculated from eq. 15.

Results

A. Measurements of Osmotic Coefficients.—The molal osmotic coefficients, ϕ, of purine, uridine and

TABLE I

EXPERIMENTAL AND FITTED VALUE[a] OF MOLAL OSMOTIC COEFFICIENTS AT 25°

Experimental		Fitted	
Molal concn.	ϕ	Molal concn.	ϕ
	A. Purine		
0.049	0.909	0.05	0.917
.049	.915		
.098	.847	.10	.849
.098	.847		
		.15	.794
.196	.759	.20	.749
.196	.749		
		.25	.714
.297	.692	.30	.685
.296	.690		
		.35	.662
.400	.646	.40	.643
.398	.645		
		.45	.627
.504	.611	.50	.614
.503	.610		
		.55	.601
		.60	.590
		.65	.578
.685	.567	.70	0.567
.681	.565		
		.75	.555
		.80	.544
		.85	.532
.899	.525	.90	.522
.885	.526		
.904	.524		
		1.00	.505
		1.05	.501
1.114	0.500	1.10	.501
1.107	.502		
	B. Uridine		
		0.05	0.969
0.096	0.937	0.10	0.943
.095	.933		
.094	.952		
.094	.962		
		.15	.921
.194	.910	.20	.901
.192	.896		
.192	.903		
.189	.914		
		0.25	0.883
.293	.860	0.30	0.866
.291	.867		
.292	.861		
.289	.873		
		.35	.849
.393	.839	.40	.833
.395	.838		
.391	.835		
		.45	.817
.503	.800	.50	.801
.501	.804		
.499	.803		
		.55	.786
.579	.768	.60	.773
		.65	.762
.710	.760	.70	.755
.702	.749		
.693	.756		
	C. Cytidine		
		0.05	0.967
0.0955	0.936	.10	.935
.0948	.939		
.0941	.941		
		.15	.905
.194	.888	.20	.876
.189	.878		
.190	.880		
		.25	.850
.291	.833	.30	.826
.292	.822		
.291	.825		
		.35	.804
.397	.791	.40	.785
.395	.786		
.393	.791		
		.45	.768
.503	.750	.50	.752
.502	.749		
.499	.753		
		.55	.738
		.60	.724
		0.65	0.710
0.699	0.695	.70	.695
.695	.697		
.709	.693		
	D. Inosine		
0.100	0.850		
.100	.853		
	E. Caffeine		
.100	.650		
.100	.645		

[a] See "Definitions and equations related to osmotic coefficients determinations" for the calculation of the polynomials. See Table II for the coefficients of the polynomials.

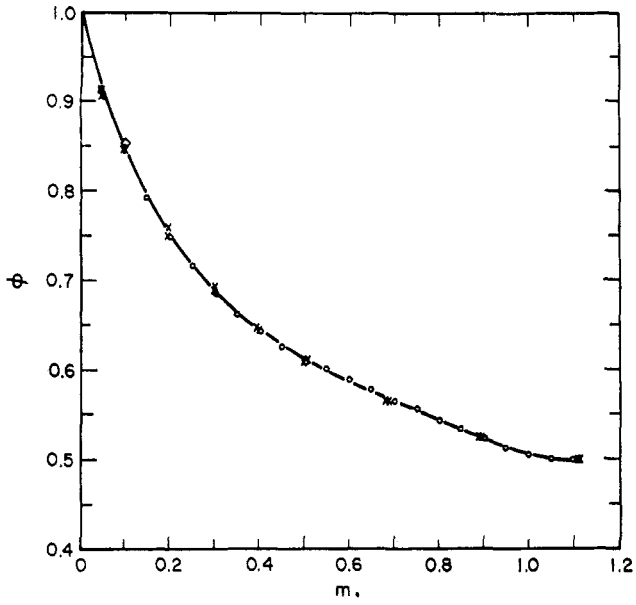

Fig. 1.—The osmotic coefficients (ϕ) of purine at varying molal concentrations together with the osmotic coefficient of inosine ($\Diamond$) and of caffeine (Δ) at 0.1 molal. The experimental ϕ of purine, -X-X-; and the fitted ϕ of purine by computer, -O-O-. All at 25°.

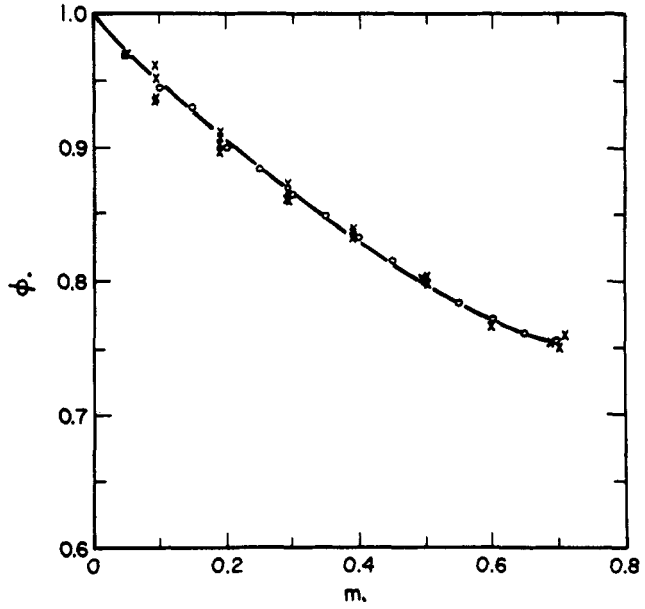

Fig. 2.—The osmotic coefficients (ϕ) of uridine at varying molal concentrations at 25°. The experimental ϕ, -X-X-; and the fitted ϕ by computer, -O-O-.

cytidine at 25° have been determined and are reported in Table I and Fig. 1, 2 and 3. The experimental data for each compound was fitted by a fourth degree polynomial as detailed in the previous section. The fitted values are listed or plotted along with the experimental values in Table I and Fig. 1, 2 and 3. The fitting of these three polynomials to the experimental values is satisfactory. The numerical coefficients of these three polynomials are listed in Table II. In addition, the osmotic coefficients for inosine and caffeine at one concentration, 0.1 *M*, were determined (Table I). The limited range of concentration tested for the inosine and caffeine, and even to a certain extent for cytidine and uridine, is due to the low solubility of these compounds.[18] Molal

(18) The solubility of cytidine and uridine at 25° is close to being 1 molar.

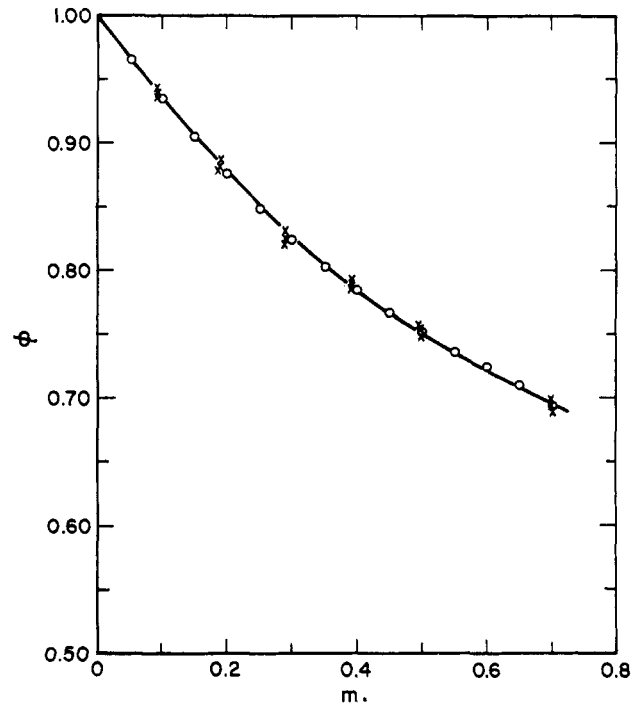

Fig. 3.—The osmotic coefficients (ϕ) of cytidine at varying molal concentrations at 25°. The experimental ϕ, -X-X-; and the fitted ϕ by computer, -O-O-.

activity coefficients for purine, uridine and cytidine at varying concentrations were calculated from the fitted value of osmotic coefficients (eq. 1, 2 and 3) and are reported in Table III. This extensive lowering of activity coefficients of these compounds in solutions as compared with that of other compounds such as amino acids[19] and electrolytes[12] strongly suggests that association or polymerization of these solutes occurs in solution.

Table II

Coefficients[a,b] of the Fitted Polynomials (Eq. 4) in Relating Molal Osmotic Coefficient, ϕ, to the Molal Concentration

	a_1	a_2	a_3	a_4
Purine	−1.8228	+3.4667	−3.3281	+1.1892
Uridine[3]	−0.6734	+1.2714	−2.1523	+1.4218
Cytidine[3]	−0.6627	+0.0430	+1.0306	−0.8978

[a] See the calculation of the coefficients in "Definitions and equations related to osmotic coefficient determinations." [b] An eighth degree polynomial was also fitted for the purine curve and it was found to be no improvement over the fourth degree polynomial equation.

Table III

Molal Activity Coefficients[a] at 25° Computed from the Fitted Osmotic Coefficients[b]

Molal concn.	Purine	Uridine	Cytidine	Molal concn.	Purine	Uridine	Cytidine
0.05	0.844	0.939	0.936	0.60	0.339	0.600	0.537
.10	.728	.888	.878	.65	.324	.582	.518
.15	.641	.845	.824	.70	.311	.568	.499
.20	.575	.808	.776	.75	.297	...	...
.25	.522	.775	.733	.80	.286	...	...
.30	.480	.744	.695	.85	.275	...	...
.35	.446	.716	.661	.90	.264	...	...
.40	.418	.690	.631	.95	.255	...	...
.45	.394	.665	.604	1.00	.247	...	...
.50	.374	.641	.580	1.05	.240	...	...
.55	.355	.620	.558	1.10	.235	...	...

[a] See eq. 3. [b] See Table I.

(19) E. P. B. Smith and P. K. Smith, *J. Biol. Chem.*, **117**, 209 (1937).

TABLE IV

EQUILIBRIUM CONSTANTS, $K = (1 - \phi)/m\phi^2$,[a] COMPUTED FROM THE FITTED OSMOTIC COEFFICIENTS[b]

Molal concn.	Purine	Uridine	Cytidine
0.1	2.1	0.64	0.74
.2	2.2	.61	.81
.3	2.2	.60	.85
.4	2.2	.60	.87
.5	2.1	.62	.87
.6	2.0	.63	.88
.7	1.9	.62	.90
.8	1.9	...	...
.9	2.0	...	...
1.0	1.9	...	...
1.1	1.8	...	...

[a] See eq. 6. [b] See Table I.

TABLE V

EQUILIBRIUM CONSTANTS, $K_1 = (1 - \phi)/m(2\phi - 1)^2$,[a] COMPUTED FROM THE FITTED OSMOTIC COEFFICIENTS[b]

Molal concn.	Purine	Uridine	Cytidine
0.1	3.1	.72	0.86
.2	5.0	.77	1.10
.3	7.7	.84	1.37
.4	10.1	.94	1.66
.5	15.0	1.01	1.95
.6	21.3	1.27	2.29
.7	34.8	1.35	2.86
.8	75.3		
.9	2.87×10^2		
1.0	5.1×10^3		
1.1	1.2×10^5		

[a] See eq. 10. [b] See Table I.

TABLE VI

EQUILIBRIUM CONSTANTS[a] AND THE STANDARD FREE ENERGY CHANGE[b] OF THE ASSOCIATION

	K (Molal^{-1})	ΔF^0
Purine	2.1	−440
Uridine	0.61	+290
Cytidine	0.87	+80

[a] See eq. 9, Table IV, and Fig. 5, 6, 7. [b] See eq. 11.

The mathematical analysis of the association process adopted here is essentially that proposed by Schellman[17] in treating the osmotic coefficient data for urea solutions. Equations were developed to relate the equilibrium constant of the association process to the osmotic coefficient under two sets of assumed conditions: 1, the formation of polymers takes place with the same equilibrium constant at each step (K, eq. 5 and 6); or, 2, only dimer is formed in the solution (K_1, eq. 7). Analyses of the experimental data by graphic methods (Fig. 4) as well as of the fitted data by computation (Table IV and V) were performed in accordance with eq. 6 and 7. The large change of K_1 with increasing concentration especially in the case of purine shows that the assumption of eq. 7 is incorrect (Table V). Thus, it can be concluded that the bases are associated in polymers beyond the dimer stage. On the other hand, the equilibrium constant, K, computed from eq. 6 is relatively constant with concentration (Table IV and Fig. 4). In the case of purine (Fig. 4) the fitting is good except at high concentrations. The diminishing K at high concentrations of purine suggests that K_n for the $n + 1$ degree of association is smaller than K_{n-1} for the n degree of association, which is a very reasonable interpretation of the situation. K computed for uridine (Fig. 4, Table IV) is more scattered than that computed for purine. The reason for this is unknown, except that the experimental values

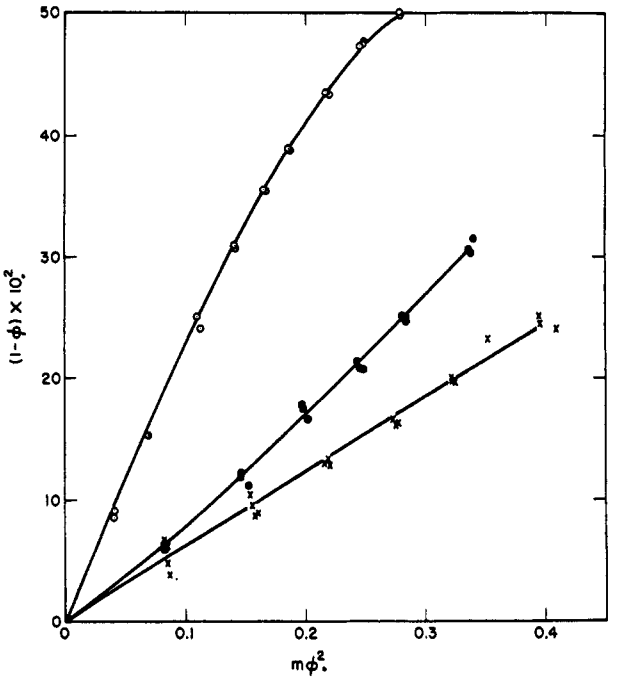

Fig. 4.—The plots of $(1 - \phi)$ *vs.* $m\phi^2$ of purine (-O-O-), of cytidine (-●-●-) and of uridine (-×-×-) in evaluation of K from eq. 6.

of ϕ for uridine are also more variable. In the case of cytidine, a definite, though small, change of the slope (K, eq. 6) can be observed in Fig. 4 as well as in the computed values of K in Table IV. It suggests that K increases slightly in this concentration range as the degree of association becomes higher. Similar results have been observed for associations of amides in benzene.[20] Generally speaking, the consistency of the equilibrium constants computed or fitted in the graph indicate that eq. 7 can be employed successfully as a good approximation of the polymerization process. The equilibrium constants selected from the experimental values are given in Table VI. Standard free energy changes computed from these equilibrium constants are given also in Table VI for the polymerization of purine, uridine and cytidine in aqueous solution.

B. Measurement of Solubility.—The studies of solubility of the sparingly soluble bases in the presence of the highly soluble nucleosides was adopted as an approach to investigate the interaction between various bases and nucleosides.

The solubilities of adenine in the presence of various interactants at 25.5° are reported in Table VII(A), and Fig. 5, and at 38° are reported in Table VII(B) and Fig. 5. The results were analyzed in terms of logarithmic activity ratio, ln y/y (eq. 15), the solubility increment, $\Delta S/I$ and K' (eq. 15). As shown by their low $\Delta S/I$ values in Table VII(A), urea and adonitol are ineffective in raising the solubility of adenine. Cyclohexanol is more effective than either adonitol or urea, but still much less so than phenol or pyrimidine. Purine, the most effective compound tested so far, has a value of $\Delta S/I$ twice as high as those for cytidine and uridine. Pyrimidine and phenol are only half as effective as cytidine or uridine. At 38°, purine is still much more effective in increasing the solubility of adenine than is uridine.

The solubilities of thymine alone and in the presence of various interactants at 25° (I) are reported in Table VIII and Fig. 6. Again purine is more effective than uridine which is more effective than pyrimidine in raising the solubility.

(20) M. Davies and D. K. Thomas, *J. Phys. Chem.*, **60**, 763, 767 (1956).

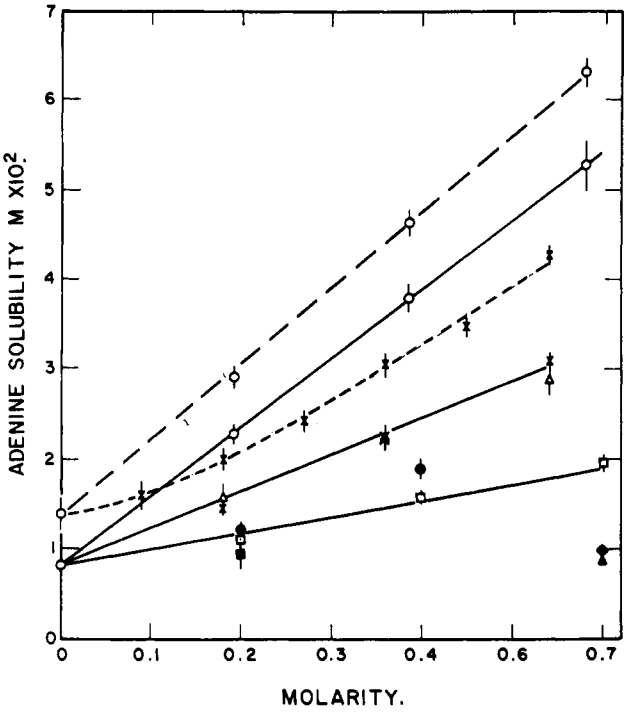

Fig. 5.—The solubility of adenine-C^{14} alone and in the presence of various interacting substances at varying molarity at 25.5° (———) and at 38° (- - - -): purine (○), uridine (×), cytidine (△), pyrimidine (□), phenol (●), cyclohexanol (■), urea (▲), and adonitol (◆). The vertical line across the symbol such as (○) is the magnitude of the standard deviation.

Table VII

Solubility of Adenine in the Presence of Interacting Compounds

Compounds added	Concn., molar	Solubility S or S', molar × 10^3	log S'/S^b = log y/y'	$\Delta S/I^c$ × 10^2	K',[d] molar^{-1}
		(A) 25.5°			
None	...	8.25 ± 0.30[a]	...	...	...
Purine	0.19	22.8	1.1[a]	0.441	7.6 ± 0.7[a]	9.2
	.39	37.9	1.6	.662	7.7	.5	9.3
	.58	52.6	2.9	.804	7.7	.6	9.3
Cytidine	.18	15.6	1.6	.274	4.1	.8	5.0
	.36	22.3	1.5	.431	3.9	.4	4.7
	.54	28.7	1.6	.542	3.8	.5	4.6
Uridine	.18	14.5	0.9	.246	3.5	.5	4.3
	.36	22.3	1.5	.431	3.9	.4	4.7
	.54	30.9	1.2	.574	4.1	.3	4.9
Pyrimidine	.20	11.1	.23	.130	1.5	.2	1.8
	.40	15.8	.58	.283	1.9	.2	2.3
	.60	19.6	.52	.377	1.9	.15	2.3
Phenol	.20	12.0	.50	.161	1.9	.4	2.3
	.40	19.0	.74	.362	2.6	.25	3.1
Cyclohexanol	.20	9.47	.15	.061	0.61	.02	...
Adonitol	.60	9.84	.39	.076	.27	.01	...
Urea	.60	8.88	.36	.033	.11	.01	...
		(B) 38°			
None	...	13.9	0.88[a]	...	...	...
Purine	0.193	29.1	1.33	0.320	7.9	1.6[a]	5.7
	.386	46.3	1.63	.522	8.4	0.9	6.0
	.58	60.9	1.70	.642	8.1	1.6	5.8
Uridine	.09	15.9	1.50	.057	2.2	3.0	1.6
	.18	20.1	1.26	.161	3.5	1.7	2.6
	.27	24.3	1.20	.243	3.9	1.1	2.8
	.36	30.5	1.33	.340	4.6	0.9	3.3
	.45	34.9	1.48	.400	4.7	.6	3.4
	.54	42.7	1.11	.487	5.4	.5	3.9

[a] Standard deviation. [b] See eq. 13. Since the solubility of adenine is so low, the activity coefficient of adenine in saturated solution is likely to be very close to one. [c] See "definitions and equations related to solubility measurements." [d] See eq. 19.

The following discussion concerns three other points which are pertinent to the understanding of these experiments:

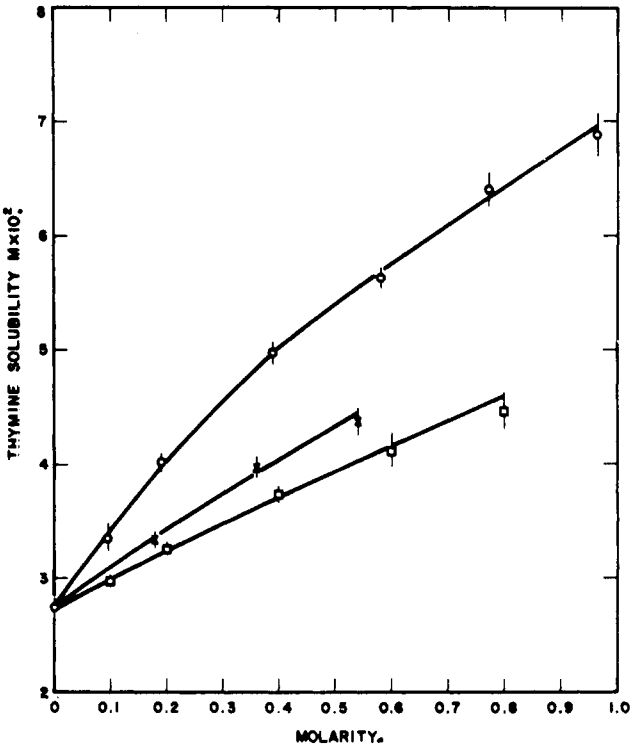

Fig. 6.—The solubility of thymine-C^{14} alone and in the presence of various interacting substances at varying molarity at 25.5°: purine (○), uridine (×) and pyrimidine (□). The vertical line across the symbol such as (○) is the magnitude of the standard deviation.

1. It has been experimentally verified that if the amount of adenine-C^{14} or thymine-C^{14} added to a solution of certain concentration of I is less than the solubility reported in Table VII and VIII, there then will be no solid left in solution and all the added base will be

Table VIII

Solubility of Thymine in the Presence of Interacting Compounds at 25.5°

Compounds added	Concn., molar	Solubility S or S', molar × 10^3	log S'/S^b = log y/y'	$\Delta S/I^c$ × 10^2	K',[d] molar^{-1}
None	...	27.4 ± 0.70[a]	...	...	...
Purine	0.095	33.5	1.20[a]	0.086	6.4 ± 2.0[a]	2.3
	.19	40.2	0.87	.167	6.6	0.80	2.4
	.39	49.8	1.10	.260	5.8	.47	2.1
	.58	56.3	1.03	.314	5.0	.30	1.8
	.77	64.0	1.50	.369	4.7	.30	1.7
	.97	70.7	2.10	.412	4.5	.30	1.6
Uridine	.18	33.4	0.70	.086	3.3	.70	1.2
	.36	39.7	0.95	.161	3.4	.45	1.2
	.54	43.7	1.27	.201	3.0	.37	1.1
Pyrimidine	.10	29.8	0.40	.037	2.4	1.10	0.9
	.20	32.6	.55	.076	2.6	0.60	.9
	.40	37.4	.80	.134	2.5	.38	.9
	.60	41.2	1.50	.176	2.3	.38	.8
	.80	44.6	1.67	.212	2.1	.30	.8

[a] Standard deviation. [b] See eq. 12. Since the solubility of thymine is so low, the activity coefficient of adenine in saturated solution is likely to be very close to unity. [c] See "definitions and equations related to solubility measurements." [d] See eq. 15.

solubilized. For example, if adenine-C^{14} is added to a 0.58 M purine solution in a concentration of 4.5×10^{-2} M (10% less than the solubility reported in Table III, 5.3×10^{-2} M), at the end of the solubility experiment, no solid will be visible in the ampoule and all the adenine-C^{14} can be recovered from the filtered solution. These results indicate that the solubilization effects of interactants, such as purine, do not depend on the presence of the solid phase of the base-C^{14}. Thus, these results rule out the possibility of the formation of solid

solutions or of co-crystallization as the mechanism for the solubilization of the bases by interacting substances.

2. The solubilizing effect of an interacting substance is very large indeed, as compared with that of amino acids or of electrolytes on each other.[12,21] The activity coefficients of adenine in the presence of interactants were lowered to $^1/_2$–$^1/_6$ of the activity coefficient of pure adenine, which is probably very close to unity because of the low concentration. Such a large lowering of the activity coefficient in the relatively dilute solution of interactant (around 0.5 M) justifies the assumption that a water-soluble complex of the base and interactant must be formed.

3. As indicated from Fig. 5 and 6, and Tables VII and VIII, the slopes of these lines or the values of $\Delta S/I$ are usually constant at various concentrations within the experimental variation except in the cases of adenine and uridine at 38° and thymine and purine at 25.5°. $\Delta S/I$ at low concentrations has a very high value of standard deviation owing to the low concentration of interactant, but at high concentration of interactant the standard deviation is around 10%. Though the variations of $\Delta S/I$ in these two exceptions are within the experimental error, the data do suggest a variation of the slope and the $\Delta S/I$ with concentration. Thus, except for these two cases, the relative constancy of $\Delta S/I$ suggests that eq. 15 may be applicable for the calculation of K'.

By addition of HCl, the solubility of adenine is greatly increased. The amount increased is in proportion to the amount of acid added, indicating that adenine hydrochloride is much more soluble than adenine.

Discussion

The osmotic coefficient and activity coefficients of purine, uridine and cytidine reported in this paper show that these compounds associate to a large extent in aqueous solution. The polymers formed in this association involve more than two molecules (dimer) of the solute. The equilibrium constants and the standard free energy changes (Table V) in each step of the association process show that purine has a considerably higher tendency to associate than cytidine or uridine.

The quantitative explanation of the data involving the solubilities of adenine and thymine in the presence of an interactant relies upon the validity of eq. 14 and 15. It is unlikely that either the adenine or thymine are present as polymers, because of their low concentration. However, this possibility cannot be ruled out, because the concentration of base in solution is maintained constant. The relative constancy of $\Delta S/I$ as I is varied, a situation predicted in eq. 15, suggests that this analysis is essentially valid.

The significance of the solubility experiments can be discussed in three main points as: 1. Comparison of experiments done at 25.5° and 38° (Table VIIA and B) indicate a lowering of K' and thus a lowering of the standard free energy change at higher temperature. A negative entropy contribution is therefore demonstrated in these association processes. 2. The K' for thymine in purine solution (Table VIII) is very similar to the K' of adenine in pyrimidine solution, while the K' of adenine in uridine or cytidine solution is slightly higher. There is, therefore, a satisfactory consistency in the measurement of purine–pyrimidine interaction by using either thymine-C^{14} or adenine-C^{14} as the insoluble base in the solubility experiments. 3. The K' of the adenine in purine solution is again much higher than the K' of thymine in uridine solution. This result supports the conclusion drawn from the K-values obtained from osmotic coefficient measurements of purine and uridine solutions. A direct comparison of the values of K and K' derived from these two different methods is not possible without reservation, since the compounds involved are different, *i.e.*, adenine instead of purine, and thymine instead of uridine. With such reservations, however, a direct comparison will show that K' for adenine in purine solution is 9 molar^{-1} while K for purine solution is 2 molal^{-1} (Tables VI and VII), and K' for thymine in uridine is 1.2 molar^{-1} (Table VIII) while K for uridine solution is 0.6 molal^{-1}. These differences are not considered significant since different approximations and compounds are involved.

The substitution of the hydrophilic ribose on the nitrogenous bases does not appear to have a large effect on the association equilibrium constants of these bases, although this substitution has a large influence on solubility. This is supported by the observations: 1. The osmotic coefficients of inosine and purine at 0.1 molal concentration are practically identical. 2. The K' of thymine in uridine solution is also similar to the K' of thymine in pyrimidine solution. 3. The K' of adenine in uridine and cytidine is even slightly higher than that of adenine in pyrimidine solution.

Then, from the association equilibrium constants of purine, pyrimidine and their derivatives, the tendency of interaction and association of these compounds can be arranged in the series: purine–purine > purine–pyrimidine > pyrimidine–pyrimidine.

Only indirect information on the mechanism of the association of these bases and nucleosides can be provided by the present data. It is felt that the mode of association of these free bases and their derivatives in solution is essentially that of vertical-stacking interactions and not of horizontal-hydrogen-bonding interactions. This concept is favored by the following observations and reasoning: 1. The K of purine (2.1) and of pyrimidine nucleosides (0.6–0.9) (Table VI) are considerably higher than those of urea (0.04) evaluated by identical treatment.[17] Thus, the standard free energy changes of the association process for purine is about 1 kcal. more negative than that of urea which already is an excellent hydrogen bonding compound. 2. Pyrimidine, which has only one hydrogen bonding acceptor site, compares favorably with uridine and cytidine in its tendency to interact and to associate with thymine or adenine. On the other hand, cyclohexanol, a saturated ring compound with an OH group for hydrogen bonding, hardly interacts with adenine at all. It should be emphasized that hydrogen bonding is not excluded in this association process, especially when the aggregates can approach each other in close range without the presence of water molecules between them.

Two sets of observations described below may provide some insight as to how significant the contribution of this stacking interaction of bases is to the stability of nucleic acid helices. First, a series of compounds has been tested for effectiveness in lowering the melting temperature of helical poly A and of thymus DNA by measurements of optical rotation.[4] The order of increasing activity was found to be: adonitol, methyl riboside (both negligible) < cyclohexanol < phenol, pyrimidine, uridine, < cytidine, thymidine < purine, adenosine, inosine, deoxyguanosine < caffeine, coumarin, 2,6-dichloro-7-methylpurine. Urea was only $^1/_{50}$ as effective as purine on poly A and $^1/_{10}$ as effective as purine on DNA in lowering their T_m helix-coil transition temperature. Based on these results, it was proposed that the mechanism of the interaction of these bases, nucleosides and related compounds with the

(21) E. J. Cohn, in "Proteins, Amino Acids and Peptides," Ed. by E. J. Cohn and J. T. Edsall, Reinhold Publishing Corp., New York, N. Y., 1943, Chapters 10 and 11.

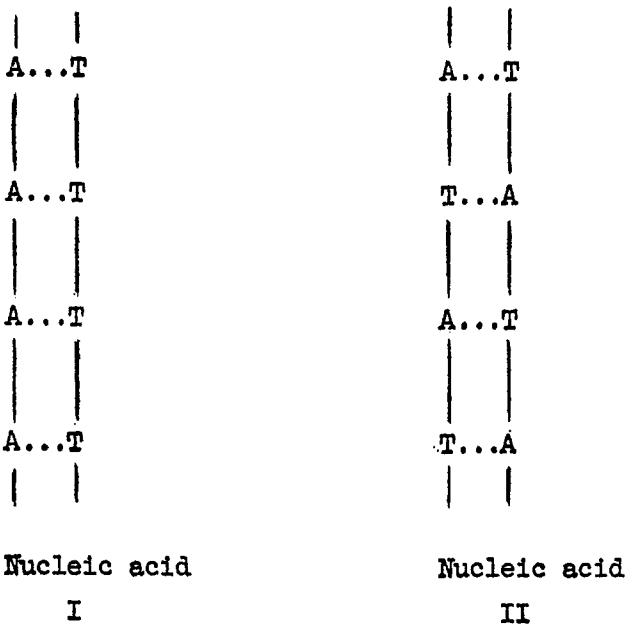

Fig. 7.—The models of two hypothetical nucleic acids having only adenine and thymine in two sets of base sequences.

helical nucleic acid does not necessarily involve hydrogen-bonding but rather hydrophobic and stacking interactions. This interpretation now receives further support from the data in this paper. The tendency of these compounds to associate in water correlates positively with the effectiveness of these compounds in reducing the stability of the helical form of nucleic acids. For example, the preference of the following compounds to associate as indicated by osmotic coefficient and equilibrium constant measurements is: caffeine > purine, inosine > uridine, cytidine. This same relationship holds true also for the effectiveness of these compounds in lowering the T_m, *i.e.*, the stability, of DNA and helical poly A.

Second, the stability of two double-stranded nucleic acids which consist of identical base composition but differ in base sequence may lead to differences in stacking interactions. For example, as illustrated in Fig. 7, nucleic acid I and nucleic acid II both consist of only adenine (A) and thymine (T) of equal amounts as required by the Watson–Crick model. In nucleic acid I, A is stacked on top of A and similarly with T. In nucleic acid II, however, A and T are staggered alternately. Since these two nucleic acids contain the same amount of A and T in composition, the hydrogen bonding contribution to the stability of these two helices should be identical. On the other hand, the A–A stacking interaction presumably is stronger than the A–T interaction which again presumably is stronger than the T–T interaction. It is not certain, however, whether the difference in energy between the A–A and the A–T interaction is exactly half of the difference in energy between the A–A and the T–T interactions. Thus, one may ask which nucleic acid, I or II, has a higher T_m under identical conditions. The answer depends on the relative importance of the contribution of the hydrogen-bonding and of the vertical stacking interactions, as well as on the comparison of the energy of A–T to that of A–A and T–T. If the hydrogen bonding is of importance and if the energy of A–T interaction is smaller than the average energy of A–A and T–T in the double strands, then nucleic acid II may have a lower T_m than nucleic acid I. Though direct verification of this model has not been published, the data of Marmur and Doty[22] did provide an indirect answer to this question. In their plot of guanine plus cytosine (per cent) of DNA composition *vs.* T_m curve (Fig. 4 in ref. 22), the extrapolated T_m of the hypothetical natural dAT DNA was significantly higher than that of the synthetic dAT polynucleotide synthesized by Kornberg's group. It is known that the synthetic dAT polymer has a relatively high molecular weight and a staggered arrangement of adenine and thymine[23] as indicated in nucleic acid II, the most regular type of structure for the staggered arrangement. It is expected that the extrapolated natural dAT DNA will be less regular in its base sequences and should contain a certain percentage of straight runs of adenine and thymine in their base sequence. Thus, the data of Marmur and Doty did suggest that the T_m of nucleic acid I would be higher than that of nucleic acid II. This observation is consistent with the notion that the base vertical stacking interaction may contribute as much as, if not more, substantially to the stability of the helical structure of the nucleic acid than hydrogen bonding. It also suggests that the difference between the A–A and the A–T interactions is not exactly half of the difference in energy between the A–A and the T–T interactions in a helix of double strands.

As early as 1958 it was suggested that hydrogen bonding is probably not the sole source of stability of the DNA helix.[24,25] Subsequent work on properties of nucleic acids in organic solvents suggested that hydrophobic interaction of bases contributes significantly to the stability of the helix.[26–29] Results presented and discussed in this paper substantiate and extend this point of view.

Acknowledgment.—We are grateful to Professor James Bonner for his support in this work and to Professors N. Davidson, R. Mazo and Dr. J. Vinograd for their suggestions concerning the manuscript.

(22) J. Marmur and P. Doty, *J. Mol. Biol.*, **5**, 109 (1962).
(23) J. Josse, A. D. Kaiser and A. Kornberg, *J. Biol. Chem.*, **236**, 864 (1961).
(24) J. M. Sturtevant, S. A. Rice and E. P. Geiduschek, *Discussions Faraday Soc.*, **25**, 138 (1958).
(25) S. A. Rice, A. Wada and E. P. Geiduschek, *ibid.*, **25**, 130 (1958).
(26) T. T. Herskovits, S. J. Singer and E. P. Geiduschek, *Arch. Biochem. Biophys.*, **94**, 99 (1961).
(27) E. P. Geiduschek and T. T. Herskovits, *ibid.*, **95**, 114 (1961).
(28) T. T. Herskovits, *ibid.*, **97**, 433 (1962).
(29) G. K. Helmkamp and P. O. P. Ts'o, *J. Am. Chem. Soc.*, **83**, 138 (1961).