Correlations in the Ultraviolet Spectra of the Purine and Pyrimidine Bases[1]

Leigh B. Clark and Ignacio Tinoco, Jr.

Contribution from the Department of Chemistry, University of California, Berkeley, California. *Received August 10, 1964*

The spectra of some purine and pyrimidine bases are presented. Correlations found among the bands of these bases and some related compounds suggest that the electronic states of all the bases are simply derived from those of benzene. A classification of the spectral bands is presented which leads to polarization directions and pH dependence consistent with experiment.

Introduction

In order for optical studies (*e.g.*, absorption spectrophotometry, optical rotatory dispersion) of the nucleic acids to be useful for structure determination, it is necessary to understand the optical properties of the isolated nucleic acid bases. A knowledge of the transition energies and polarization directions of the individual bases allows predictions of the optical properties of possible dimers, oligomers, and polymers of nucleotides. These predictions can be compared with experimental measurements to learn about the conformation and structure of these polynucleotides. The absorption and rotation of light by helical oligomers have been treated by assuming values and orientations of the transition moments in each residue.[2,3] The change in light absorption of polynucleotides which occurs on transition from a disordered to an ordered state (hypochromism) has been similarly treated.[4-6]

The examination of the ultraviolet absorption spectra of the purine and pyrimidine bases has been very extensive. Noteworthy is the work of Mason,[7] Stewart and Davidson,[8] and Rosa and Simpson.[9] Aqueous solution spectra to 185 mμ of many of the bases have recently been published by Voet, *et al.*[10]

The present work is an effort to systematize and characterize the spectral bands of the purine and pyrimidine bases which are important constituents in the nucleic acids. A correlation (or assignment scheme) is developed which not only relates the absorption bands in the various compounds but also is consistent with spectral changes due to pH variation.

(1) Supported in part by Public Health Service Research Grant GM 10840 and by an unrestricted grant from Research Corporation.

(2) I. Tinoco, Jr., R. W. Woody, and D. F. Bradley, *J. Chem. Phys.*, **38**, 1317 (1963).

(3) D. F. Bradley, I. Tinoco, Jr., and R. W. Woody, *Biopolymers*, **1**, 239 (1963).

(4) I. Tinoco, Jr., *J. Am. Chem. Soc.*, **82**, 4785 (1960); *J. Chem. Phys.*, **33**, 1332 (1960); **34**, 1067 (1961); *Radiation Res.*, **20**, 133 (1963).

(5) W. Rhodes, *J. Am. Chem. Soc.*, **83**, 3609 (1961).

(6) H. DeVoe and I. Tinoco, Jr., *J. Mol. Biol.*, **4**, 500, 518 (1962).

(7) S. F. Mason, *J. Chem. Soc.*, 219 (1960); 2071 (1954); and references therein.

(8) R. F. Stewart and N. Davidson, *J. Chem. Phys.*, **39**, 255 (1963); *Biopolymers Symp.*, **1**, 465 (1964).

(9) E. Rosa and W. T. Simpson, personal communication.

(10) D. Voet, W. B. Gratzer, R. A. Cox, and P. Doty, *Biopolymers*, 1, 193 (1963).

Experimental

All the spectra were recorded on a Cary Model 15 spectrophotometer utilizing N_2 flushing. Cells of 1-mm. path length were employed for most of the spectral studies. The 9-methyladenine, 9-methylhypoxanthine, and 9-ethylguanine were purchased from Cyclo Chemical Corporation. The acetophenone and benzimidazole came from the Eastman Co., and the remaining compounds were purchased from California Corporation for Biochemical Research. All materials were used directly from the bottle except for the 9-methylhypoxanthine which was recrystallized three times from water, and the pyrimidine which was vacuum distilled three times.

The trimethyl phosphate (TMP) used as a solvent was purchased from Victor Chemical Co. This solvent is highly transparent and was used directly from the bottle without further purification. The spectrophotometer slit width opened to 1.0 mm. at about 181 mμ for a 1-mm. thick sample of this substance. The methylcyclohexane (MCH) was Matheson Spectrograde.

The spectrum of 9-methylhypoxanthine vapor is taken from a study of the vapor spectra of the bases.[11]

Results and Discussion

The spectra of the purine and pyrimidine bases show three or four absorption bands between 180 and 300 mμ in addition to the weak long wave length (290–300 mμ) $n \rightarrow \pi^*$ band in purine and some pyrimidines. These bands appear to be related in the various bases and are presumably $\pi \rightarrow \pi^*$ transitions. The correlations indicate that the spectrum of each base is simply related to the spectrum of pyrimidine, whose electronic states are derived from those of benzene. Thus benzene serves as the parent absorbing chromophore for all the bases. Mason[7,12] has also treated the first π–π^* band of some substituted pyrimidines and purines as being the analog of the 260-mμ benzene band.

The Purines. The similarity of the first two bands of 9-methylhypoxanthine (vapor) to those of benzimidazole and acetophenone is readily apparent from Figure 1. The 280-mμ band of 9-methylhypoxanthine correlates with the 280-mμ band of acetophenone[13] which is itself derived from the forbidden $A_{1g} \rightarrow B_{2u}$ transition of benzene at 260 mμ. The next band of 9-methylhypoxanthine (240 mμ) appears to be derived from the 230-mμ band of acetophenone. The benzene

(11) L. B. Clark, G. Peschel, and I. Tinoco, Jr., to be published.

(12) S. F. Mason, "The Pyrimidines," Interscience Publishers, New York, N. Y., 1962, Chapter 13.

(13) This is the *secondary band* or the ^1L_b band in the nomenclature of L. Doub and J. M. Vandenbelt, *J. Am. Chem. Soc.*, **69**, 2714 (1947), and J. R. Platt, *J. Chem. Phys.*, **17**, 484 (1949), respectively. Similarly the 230-mμ band of acetophenone is called the *first primary band* or the ^1L_a band.

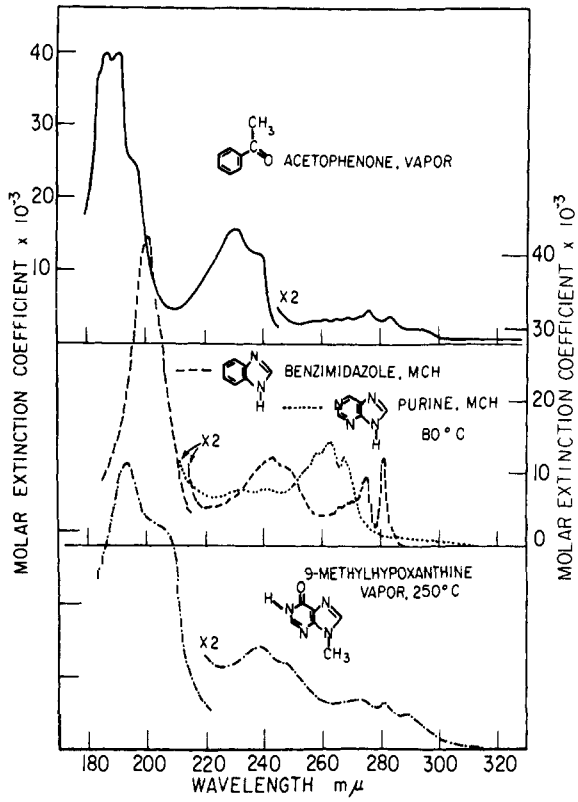


Figure 1. The absorption spectra of 9-methylhypoxanthine, purine, benzimidazole, and acetophenone.

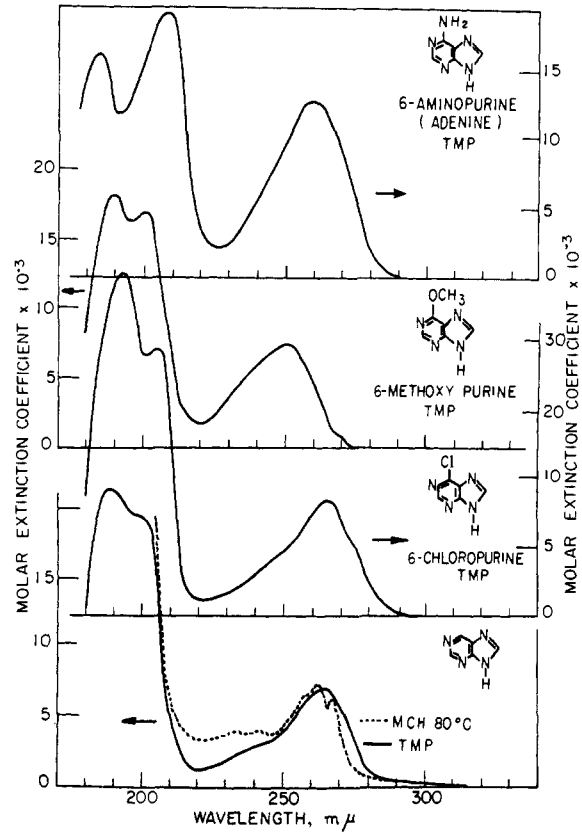


Figure 2. The absorption spectra of purine and some 6-substituted purines, all in trimethyl phosphate solution (TMP) and methylcyclohexane (MCH).

parentage of this band is uncertain, but it probably[14] is largely the $A_{1g} \rightarrow B_{1u}$ transition believed[15] to be at about 203 mμ. Finally, the more intense bands in the 180–210 mμ region are believed to be related to the $A_{1g} \rightarrow E_{1u}$ transition of benzene at about 180 mμ.

Since the intent of this discussion is to relate the absorption bands of the purines and pyrimidines to the transitions of benzene, a simple nomenclature based on the symmetry of the benzene excited states will be adopted. Bands which appear to be derived from the $A_{1g} \rightarrow B_{2u}$ and $A_{1g} \rightarrow B_{1u}$ benzene transitions will be labeled simply B_{2u} and B_{1u}, respectively.

The spectrum of purine (Figure 1) is similar to that of benzimidazole in that there are two bands in the same wave-length region as the first two bands of benzimidazole.[16] Although the intensity distribution is reversed, it would be difficult not to associate the 260- and 235-mμ bands of purine with the B_{2u} and B_{1u} bands, respectively, of the other compounds whose spectra are shown in Figure 1.

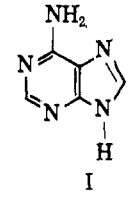
Figure 2 shows the solution spectra of purine and a series of 6-substituted purines in trimethyl phosphate (TMP). The 236-mμ (B_{1u}) band of purine is seen to red shift and to merge with the more intense B_{2u} band as we proceed up the series (–H, –Cl, $-OCH_3$, $-NH_2$) which may be considered as a gradually increasing perturbation on the purine system. Mason[7] has also reached the conclusion that the 260-mμ band of 6-aminopurine (adenine) is composite in nature.

(14) See, however, J. Tanaka and S. Nagakura, *J. Chem. Phys.*, **24**, 311 (1956).

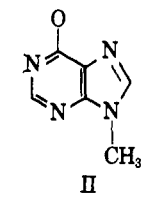
(15) For a discussion of this point, see A. C. Albrecht and W. T. Simpson, *ibid.*, **23**, 1480 (1955).

(16) This is not counting the weak n → π* band of purine which appears as a long tail extending to longer wave lengths.

Additional support for the correctness of the correlation of bands in Figures 1 and 2 comes from the pH dependence of the first two bands of 9-methylhypoxanthine and 9-ethylguanine. The solution spectrum of 9-methylhypoxanthine in TMP (Figure 3) is very similar to the vapor phase spectrum. In neutral aqueous solution, the B_{2u} band blue shifts and appears only as a tail extending from below the more intense B_{1u} band at 249 mμ (Figure 3, pH 6.1). In basic solution (pH 13.8) the two bands merge to give a spectrum which resembles that of adenine, whose structure I is formally similar to that of the anion II. Consis-



I



II

tency with the adenine assignment requires, however, that the intensity distribution between the two bands be reversed in the anion. Evidence for this comes from the pH dependence of the intensity of the first two bands of 9-ethylguanine whose spectrum in TMP (Figure 3) is similar to that of 9-methylhypoxanthine. The aqueous solution spectra show the intensity interchange clearly, for the individual bands may be resolved. A similar intensity interchange is seen to occur for the two bands at about 190 and 210 mμ. The structure of the anion III is formally the same as 2,6-diaminopurine (IV).

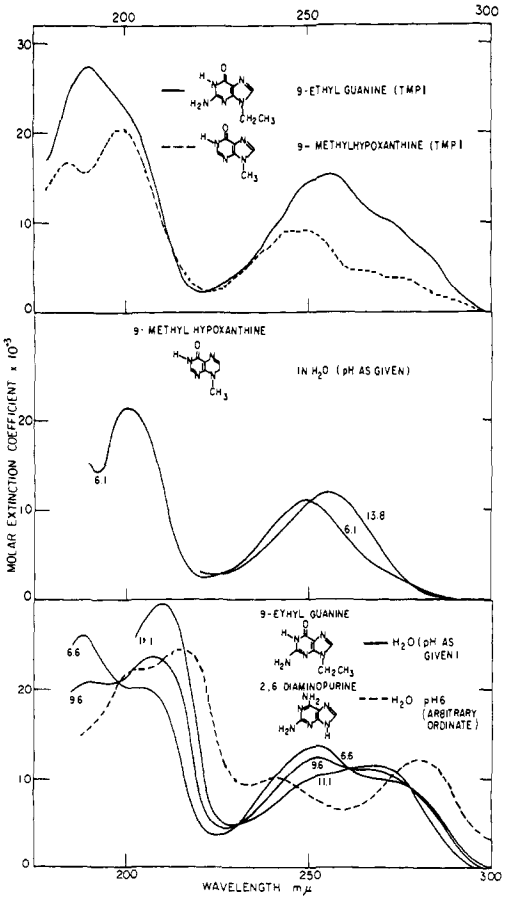
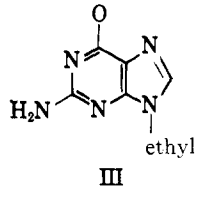
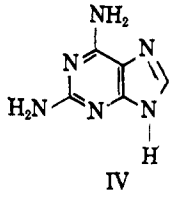


Figure 3. The absorption spectra of 9-ethylguanine, 9-methylhypoxanthine, and 2,6-diaminopurine in trimethyl phosphate and aqueous solution.

The spectrum of IV in H_2O is included in Figure 3 for comparison, and it is seen that the pattern of bands



III



IV

is identical with that for 9-ethylguanine anion III.

The Pyrimidines. The spectrum of purine is compared to that of pyrimidine in Figure 4. The weak bands at about 290–300 mμ in each compound are $n \rightarrow \pi^*$ transitions[7] and will not be discussed. The 260-mμ band (B_{2u}) of purine correlates well, as expected, with the $A_1 \rightarrow B_2$ band of pyrimidine (240 mμ) which is derived from the $A_{1g} \rightarrow B_{2u}$ benzene transition upon lowering the symmetry from D_{6h} to C_{2v}.

It is difficult to assign the pyrimidine band which corresponds to the B_{1u} band of purine. There is a rather weak band ($\epsilon < 1000$) in pyrimidine at about 210 mμ which may correspond to B_{1u}; however, this band is not obviously related to any known band in the spectrum of benzene. The higher energy, more intense, structured band (the onset of which is seen at about 190 mμ) appears to correlate with the structured benzene band at 203 mμ. The 203-mμ band is the second known band in the benzene spectrum and has

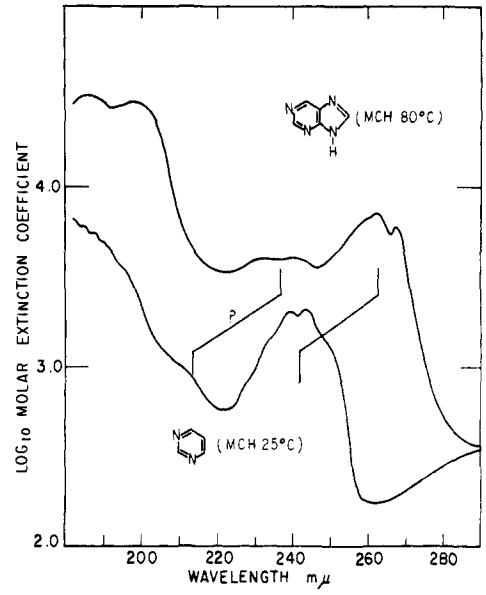


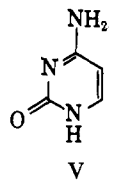
Figure 4. The absorption spectra of pyrimidine and purine in methylcyclohexane; the lines indicate a possible correlation of bands as mentioned in the text.

been assigned as the forbidden $A_{1g} \rightarrow B_{1u}$ transition (which, however, is rather intense owing to borrowing from the nearby allowed $A_{1g} \rightarrow E_{1u}$ band at 180 mμ).

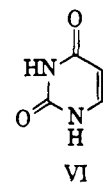
The question of the assignment of the 210-mμ pyrimidine band cannot be resolved at present. For example, the possibility that it is a second $n \rightarrow \pi^*$ transition cannot be ruled out. The fact that the relative intensity increase and energy shift of the B_{2u} band in going from pyrimidine to purine is about the same as those changes between the 210-mμ band of pyrimidine and the 235-mμ band (B_{1u}) of purine supports a correlation between the latter two bands. Such a correlation would suggest that the structured benzene band at 203 mμ is not necessarily the $A_{1g} \rightarrow B_{1u}$ transition, but that it is probably composite in nature.[17]

The spectra of a number of pyrimidine bases are shown in Figure 5. The cytosine (V) spectrum appears to be simply derived from the purine (and also the pyrimidine) spectrum which is included for comparison. Thus the 263-, 235-, 200-, and 187-mμ bands of purine correlate with the 276-, 237-, 204-, and 184-mμ bands of cytosine.

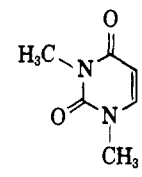
The spectra of uracil (VI) and its methyl derivative 1,3-dimethyluracil (VII) are noteworthy in that the B_{1u} band (237 mμ in cytosine) is absent. However, upon perturbing uracil slightly to form 6-azauracil



V



VI



VII

(VIII), a weak new band appears at about 230 mμ. Furthermore, uracil in basic aqueous solution (pH 14) shows what is probably a band at about 230 mμ. The

(17) The vapor spectrum of hexamethylbenzene similarly shows two absorption bands in the frequency region corresponding to the 203-mμ band of benzene: R. C. Nelson and W. T. Simpson, *J. Chem. Phys.*, 23, 1146 (1955).

Table I. Spectral Data and Assignment of Bands

Compound	Solvent	λ, mμ (ε)[a,b] B_{2u}	B_{1u}	E_{1u}	Other
Acetophenone	Vapor	276 (2.2)	230 (15.6)	190 (40.0)	320 (150), O(nπ*)
6-Aminopurine	TMP	260 (12.6)	...	208 (18.7), 185 (15.8)	...
6-Azauracil	TMP	261 (5.3)	229 (2.4)	195 (6.3), 182 (7.9)	...
Benzimidazole	MCH	275 (4.7)	243 (6.2)	201 42.8)	...
6-Chloropurine	TMP	265 (8.2)	245 (s 3.9)	205 (19.0), 193 (24.2)	...
Cytosine	TMP	277 (7.5)	237 (s 3.5)	204 (11.9), 185 (12.2)	...
2,6-Diaminopurine	H_2O, pH 6	280 (?)	242 (?)	215 (?), 202 (?)	...
1,3-Dimethyluracil	TMP	264 (7.0)	...	206 (6.5), 184 (14.1)	...
	MCH	262 (8.0)	...	203 (7.8), 182 (19.5)	...
9-Ethylguanine	TMP	275 (s 9.4)	256 (15.4)	203 (s 20.), 190 (27.4)	...
	H_2O, pH 6.6	275 (s 9.6)	252 (13.6)	205 (s 20.), 188 (26.0)	...
	pH 9.6	270 (s 10.8)	252 (12.2)	208 (23.7), 190 (21.0)	...
	pH 11.1	269 (11.5)	252 (s 10.2)	210 (29.6)	...
6-Methoxypurine	TMP	251 (7.3)	...	200 (16.7), 190 (18.0)	...
9-Methylhypoxanthine	Vapor	281 (?)	239 (?)	205 (s ?), 193 (?)	...
	TMP	278 (3.8)	247 (9.3)	198 (20.4), 185 (16.8)	...
	H_2O, pH 6.1	260–270 (s)	249 (11.1)	200 (21.5)	...
	pH 13.8	255 (11.8)	...	...	...
Purine	TMP	265 (6.9)	240 (s 3.0)	200 (18.1), 188 (21.1)	290 (0.6), N(nπ*)
Pyrimidine	MCH	242 (2.0)	210 (s 1.0)	190 (6.0)	310 (0.35), N(nπ*)
Uracil	TMP	258 (7.8)	...	203 (8.2), 181 (11.8)	...
	H_2O, pH 7.0	259 (8.1)	...	202 (8.8), 181 (15.5)	...
	pH 14.0	276 (6.4)	230 (s)	...	...

[a] The letter s implies a shoulder. [b] The values given for ε must be multiplied by 10^3 in order to get molar extinction coefficients.

structure of the anion IX now formally resembles the cytosine structure (V). There is no direct evidence for ionization from the N_1 position rather than the N_3 position. However, we write structure IX since it most probably involves more resonance stabilization than the corresponding structure with the N_3 proton ionized. Finally, the preliminary optical rotatory

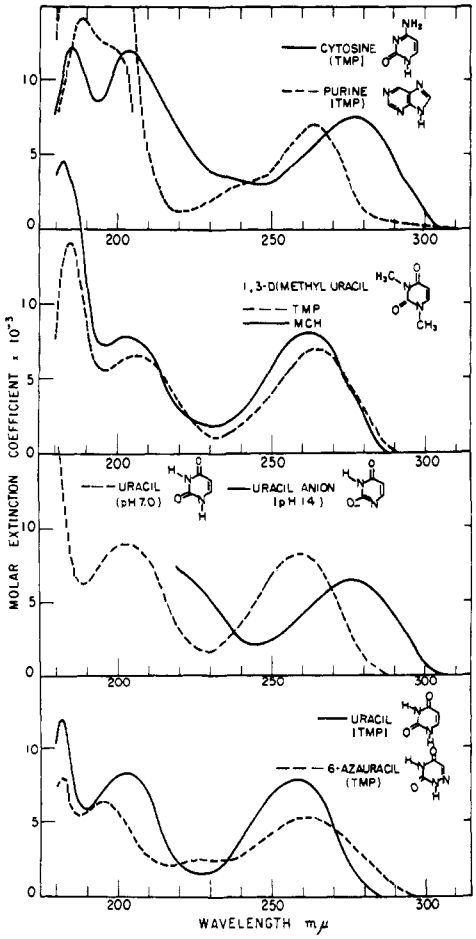
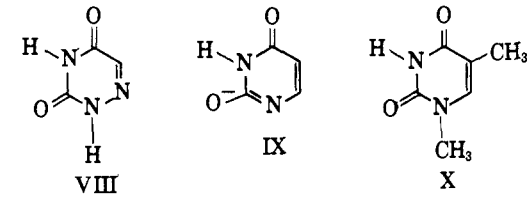


Figure 5. The absorption spectra of some pyrimidine bases; the solvents are noted in the figure.



dispersion curve for neutral (pH 7) uridine suggests the superposition of three Cotton effect curves.[18] Two correspond to the known 256- and 200-mμ absorption bands, and a third is centered at about 230 mμ. Thus, there is evidence for a weak band corresponding to B_{1u} in the spectrum of uracil as well.

The one or two higher energy bands appear to retain their identity for all the compounds in which they occur in the accessible wave-length region. Presumably they are derived from the allowed $A_{1g} \rightarrow E_{1u}$ benzene transition at about 180 mμ. However, the assignment of these bands will not be clear until the questions raised earlier concerning the 210- and 190-mμ bands of pyrimidine can be resolved.

The spectral data and assignment of bands are summarized in Table I.

The Polarization of the Bands. The theory of the intensity changes and polarization directions of the forbidden benzene bands upon substitution has been given by Sklar,[19] Förster,[20] and Platt.[21] The polarization directions of the B_{2u} and B_{1u} bands are given by the vector sum of the individual spectroscopic moments of the various substituents. Although this method is not expected to be completely applicable to the strong perturbations present here, it does constitute

(18) R. Yolles, personal communication.
(19) A. L. Sklar, *J. Chem. Phys.*, 10, 135 (1942); *Rev. Mod. Phys.*, 14, 232 (1942).
(20) T. Förster, *Z. Naturforsch.*, **2a**, 149 (1947).
(21) J. R. Platt, *J. Chem. Phys.*, **19**, 263 (1951).

a basis for making estimates of the polarization directions.

For symmetrical disubstitutions of benzene, the polarization directions are fixed; therefore, for pyrimidine itself the B_{2u} band must be polarized along the line joining the two nitrogens (Figure 6a). The B_{1u} band is polarized perpendicular to this direction. Substitutions which do not change this symmetry, that is, substitutions at the 2- and 5- positions, do not affect the polarization directions.

If we consider that in uracil the two keto groups are the main perturbers of the benzene system, the B_{2u} band is polarized along the line joining these groups (Figure 6b). This is consistent with the experimental result that the 260-mμ band of 1-methylthymine (X) is polarized close (11 ± 2° from C-2 to C-4 and inclined toward the N-1 position)[8,22] to the oxygen–oxygen axis. In this approximation for uracil, substitutions at the 6- position should not change the direction of polarization. Further extrapolation leads one to consider adenine to be nearly symmetrically substituted (Figure 6c). Thus one predicts the B_{2u} band to be polarized along the C-4–C-6 axis; this is 33° from the polarization direction in 9-methyladenine[8,22] (3 ± 3° from the C-4–C-5 axis and inclined toward the N-7 position). The B_{1u} band which is assigned as a weak contributor to the adenine 260-mμ band is polarized perpendicular to the B_{2u} band. There is evidence for such a band in 9-methyladenine both from the polarized absorption[8] and polarized fluorescence[23] spectra. Polarization directions should be mainly unchanged for C-2 substituted derivatives of adenine.

In order to predict the polarization directions for cytosine and guanine, one needs to know the signs and magnitudes of the spectroscopic moments of the substituents. Lacking these requirements we can only suggest that the polarization directions in cytosine and guanine are probably analogous to those in uracil and adenine, respectively.

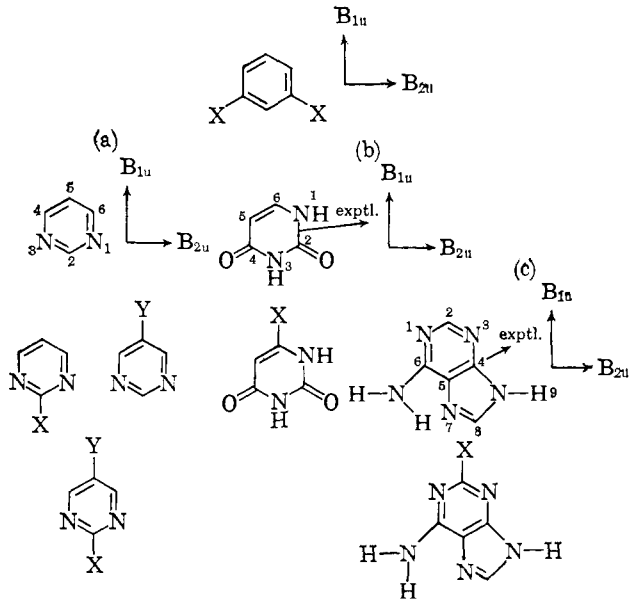


Figure 6. The suggested and experimental directions of transition moments in some pyrimidines and purines; the experimental directions are given in ref. 8 and 22.

Acknowledgment. The authors wish to acknowledge with thanks the helpful discussions with Professor William T. Simpson, Dr. Robert F. Stewart, and Mr. Eugene Rosa. We are especially grateful to the latter for suggesting the use of trimethyl phosphate as a solvent.

(22) R. F. Stewart and L. H. Jensen, *J. Chem. Phys.*, 40, 2071 (1964).
(23) P. R. Callis, E. J. Rosa, and W. T. Simpson, *J. Am. Chem. Soc.*, 86, 2292 (1964).