November, 1980] Bull. Chem. Soc. Jpn., 53, 3073—3078 (1980)

The Application of Magnetic Circular Dichroism to the Study of the Tautomerism of Cytosine and Isocytosine

Akita Kaito, Masahiro Hatano,* Toru Ueda,** and Susumu Shibuya**
Chemical Research Institute of Non-aqueous Solutions, Tohoku University, Sendai 980
*** Faculty of Pharmaceutical Science, Hokkaido University, Sapporo 060*

(Received March 5, 1980)

The MCD and absorption spectra of cytosine and isocytosine and their *N*-substituted derivatives were measured in the wave-number region of 30000—50000 cm^{-1}. On the basis of the experimental results, the tautomerism of cytosine and isocytosine was discussed. The transition energies, the oscillator strengths, and the Faraday *B* terms of the tautomeric forms of cytosine and isocytosine were calculated within the framework of the intermediate neglect of the differential overlap(INDO) approximation. The calculated results are in good agreement with the experimental data.

The magnetic circular dichroism (MCD) spectra of purine and pyrimidine derivatives, which are the constituents of nucleic acids, have been extensively studied by many investigators.[1–4] Some overlapping absorption bands have been resolved in the MCD spectra, and valuable information about the electronic structure of these compounds has been obtained. Recently, also quantum mechanical calculations of the MCD of some purine derivatives[4,5] and pyrimidine derivatives[6] have been reported.

On the other hand, the minor tautomers of nucleic acid bases are biological interest because they have been considered to be a source of the imperfect base pairing which leads to genetic mutations.[7] The tautomerism of cytosine in solution has been studied by the aid of the electronic absorption spectra.[8–10] Morita and Nagakura[10] have shown that, in an aqueous solution, cytosine takes the [1*H*]-2-keto 4-amino form(I) at room temperature, but with a rise in the temperature the 2-keto 4-imino form(II) coexists with Form I, while cytosine takes Form I and Form II in trimethyl phosphate and in acetonitrile respectively. It has been proposed from the unusual line broadening of the NMR spectra of the H5 proton of cytosine that Form II coexists with Form I to the extent of $15\pm3\%$ at room temperature in a neutral aqueous solution of cytosine.[11] However, Wong[12] has shown that the line broadening is caused by the paramagnetic impurities, and this has raised questions about the existence of 15% of the minor tautomer of cytosine, Form II, at room temperature in a neutral aqueous solution.

The tautomerism of isocytosine has also been extensively studied by measurements of the electronic absorption spectra.[9,10,13] It has been shown that isocytosine takes the [3*H*]-2-amino 4-keto form(IV) in ethanol and in a diethyl ether solution, but in an aqueous solution, the [1*H*]-2-amino 4-keto form(V) coexists with Form IV.[9,10,13]

The tautomeric forms of cytosine and isocytosine in the crystalline state are of interest. X-Ray crystal analysis studies have shown that cytosine takes Form I in crystals,[14] while Forms IV and V coexist at a ratio of 1:1 in the isocytosine crystal.[15]

In this work, we measure the MCD spectra of cytosine, isocytosine, and their derivatives and discuss the tautomerism of cytosine and isocytosine on the basis of the MCD spectra. We also calculate the Faraday *B* terms of three tautomers of cytosine and two tautomers of isocytosine, which are shown in Fig. 1, using wavefunctions obtained by the intermediate neglect of the differential overlap (INDO) procedure.[16,17]

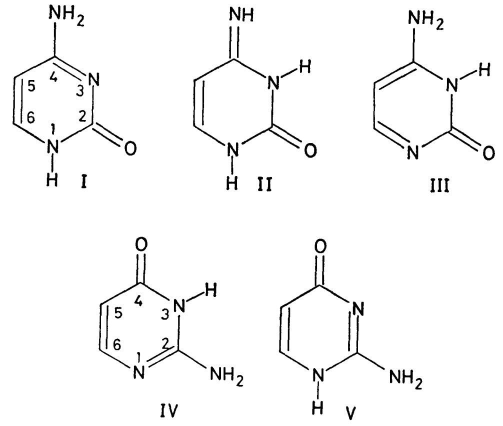

Fig. 1. Tautomeric forms of cytosine and isocytosine.

Experimental

Cytosine and isocytosine were purified by recrystallization from water. 3-Methylcytosine and 3-methylcytidine methosulfate were also recrystallized from a methanol-ethyl acetate solution. 3-Methylisocytosine was prepared according to the procedure by Angier and Curran.[18] Isocytosine was synthesized by the ammonolysis of 2,5′-cyclouridine.[19]

Acetonitrile, ethanol, a Na_2HPO_4–KH_2PO_4 buffer solution(pH 7), and a NaOH–Na_2HPO_4 buffer solution(pH 11) were used as solvents. The pH values of the buffer solutions were controlled so that the neutral form was dominant.

The MCD spectra were recorded with a JASCO J-20A recording circular dichrometer equipped with a 12.5 kG electromagnet. The absorption spectra were measured on a Hitachi EPS-3T recording spectrophotometer.

The experimental values of the Faraday *B* terms were obtained by the gaussian curve-fitting procedure or the method of moments.

Theoretical

The Faraday *B* terms were calculated using wavefunctions obtained by the INDO/S-CI approximation.[16,17] The bonding parameter of the oxygen atom, β_0, was determined so that the transition energies of the $\pi^*\leftarrow\pi$ transitions of some ketones were well reproduced ($\beta_0=-31$ eV). The other parameters in

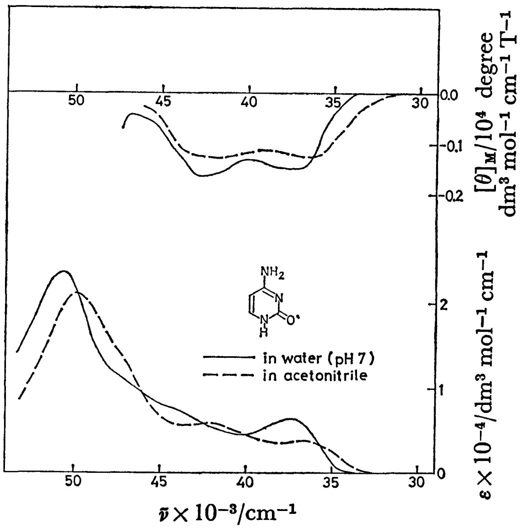

Fig. 2. MCD (upper) and absorption (lower) spectra of cytosine in aqueous solution (pH 7) (——) and in acetonitrile (– – –) at room temperature.

the calculation were described in a previous paper.[5]

The molecular structures of Form I of cytosine and Forms IV and V of isocytosine were taken from the experimental data obtained by X-ray crystal analysis.[14,15] The bond lengths and bond angles of Forms II and III of cytosine were estimated from the experimental values of Forms I, IV, and V.[14,15]

The configuration interaction among singly excited configurations below 12 eV was taken into account.

The Faraday *B* terms were calculated according to the expression derived by Buckingham and Stephens.[20] The procedures of the calculation of the Faraday *B* terms have been reported in the previous paper.[5]

Results and Discussion

The MCD and absorption spectra of cytosine are shown in Fig. 2. The molar ellipticity per unit of the magnetic field, $[\theta]_M$, is expressed by 10^4 degree $dm^3\,mol^{-1}\,cm^{-1}\,T^{-1}$ unit, and the extinction coefficient, ε, by $dm^3\,mol^{-1}\,cm^{-1}$ unit. The spectra in the aqueous solution are in agreement with those measured by Voelter *et al.*[1] In the wavenumber region of 32000—46000 cm^{-1}, two negative MCD bands are observed. In the absorption spectrum, we can observe an intense peak and a shoulder at 50700 cm^{-1} and at about 47000 cm^{-1} respectively, while no distinct MCD band is detected in the wavenumber region of 47000—50000 cm^{-1}. The MCD and absorption spectra in an acetonitrile solution are red-shifted and decrease in intensity relative to the spectra in an aqueous solution.

The MCD, CD, and absorption spectra of 3-methylcytidine methosulfate in an aqueous solution (pH 11) are presented in Fig. 3. 3-Methylcytidine shows a negative MCD band and a positive MCD band at 37400 cm^{-1} and at 43400 cm^{-1} respectively.

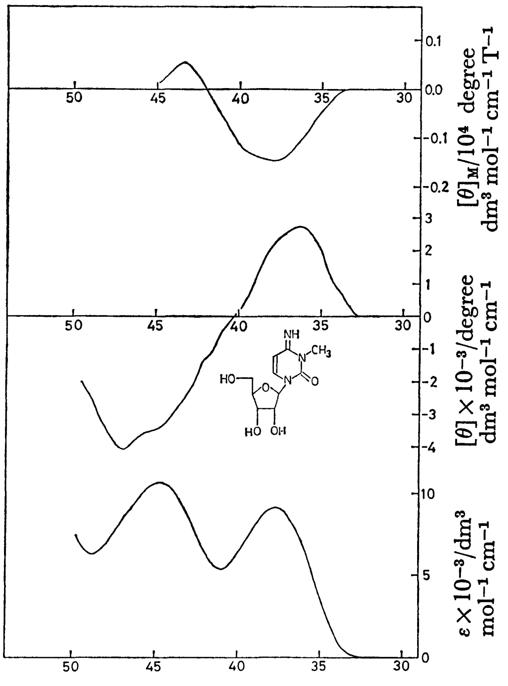

Fig. 3. MCD (top), CD (middle), and absorption (bottom) spectra of 3-methylcytidine in aqueous solution (pH 11) at room temperature.

On the other hand, Morita and Nagakura[10] have concluded that the most stable form of cytosine in acetonitrile is Form II. The MCD and absorption spectra of Form II should resemble those of 3-methylcytidine, because the tautomeric form of 3-methylcytidine is fixed to Form II. However, the sign of the second MCD band of cytosine in acetonitrile is the opposite of that of 3-methylcytidine. In addition, the absorption spectrum of cytosine in acetonitrile is not very similar to that of 3-methylcytidine with respect to the absorption intensity and the position of the absorption peak observed in the wavenumber region higher than 40000 cm^{-1}. The MCD spectrum of cytosine in acetonitrile resembles that of cytosine in an aqueous solution rather than the MCD spectrum of 3-methylcytidine.

These results suggest that the most stable tautomer of cytosine is Form I, not only in a neutral aqueous solution, but also in acetonitrile. We can consider several reasons for the solvent dependence of the MCD and absorption spectra of cytosine: a change in the polarity of the solvent, the specific interaction of the cytosine molecule with the solvent molecules, such as the formation of hydrogen bonding, the association of cytosine molecules, or the contamination of small amount of minor tautomers.

The MCD and absorption spectra of 3-methylcytosine, whose tautomeric form is fixed to Form III,[21] are shown in Fig. 4. Morita and Nagakura[10] have reported that Form III coexists with Forms I and II, to a much smaller extent, in an acetonitrile solution of cytosine at a higher temperature. The sign pattern of the MCD spectrum of 3-methylcytosine is the same as that of 3-methylcytidine (Form II), while the lowest MCD and absorption bands lie at lower wavenumbers than those of cytosine (Form I) and 3-methylcytidine (Form II). The presence of

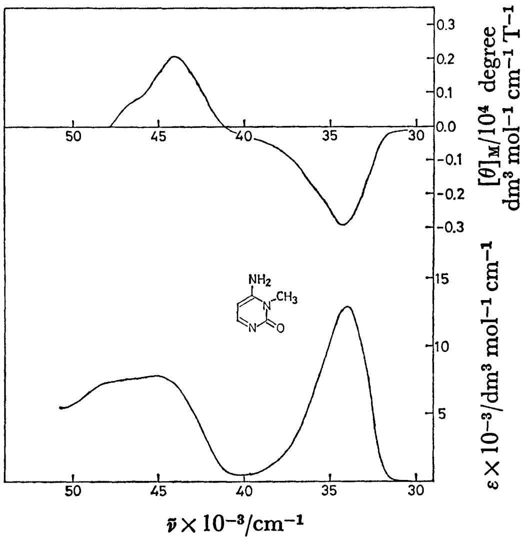

Fig. 4. MCD (upper) and absorption (lower) spectra of 3-methylcytosine in aqueous solution (pH 11) at room temperature.

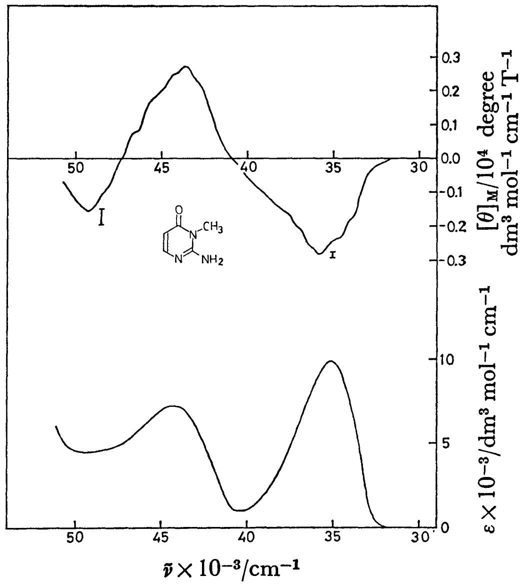

Fig. 5. MCD (upper) and absorption (lower) spectra of 3-methylisocytosine in aqueous solution (pH 7) at room temperature.

Form III in cytosine is not confirmed by the results of this work.

The MCD and absorption spectra of 3-methylisocytosine, whose tautomeric form is fixed to Form IV, are exhibited in Fig. 5. The MCD and absorption spectra of 3-methylisocytosine are very similar to those of 3-methylcytosine, because these two compounds have the same valence-bond structure.

The MCD, CD, and absorption spectra of isocytidine, whose tautomeric form is fixed to Form V, are shown in Fig. 6. Although the lowest negative MCD band almost corresponds to the lowest absorption band at 39100 cm^{-1}, the second negative MCD band lies intermediate between the absorption peak at 48500 cm^{-1} and a shoulder at 43600 cm^{-1}.

The MCD and absorption spectra of isocytosine in an ethanol and in an aqueous solution (pH 7) are shown in Fig. 7. In the ethanol solution, negative and positive MCD bands are observed at 35400 cm^{-1} and at 44500 cm^{-1} respectively. The MCD and absorption spectra of isocytosine in ethanol resemble those of 3-methylisocytosine, suggesting that the prevailing tautomeric form of isocytosine in ethanol is Form IV. In the aqueous solution, the two MCD bands at 35400 cm^{-1} and at 44500 cm^{-1} decrease in intensity, and two negative MCD bands appear at about 38700 cm^{-1} and at 47200 cm^{-1}. These MCD bands appearing in the aqueous solution are analogous to the MCD bands of isocytidine, and they can rea-

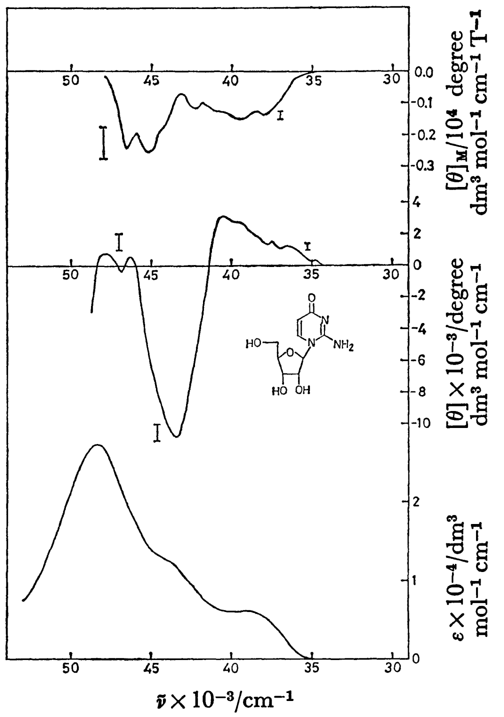

Fig. 6. MCD (top), CD (middle), and absorption (bottom) spectra of isocytidine in aqueous solution (pH 7) at room temperature.

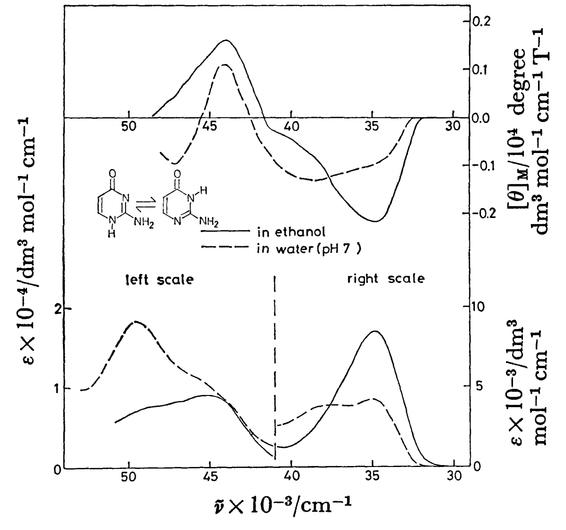

Fig. 7. MCD (upper) and absorption (lower) spectra of isocytosine in aqueous solution (pH 7) (- - -) and in ethanol (——) at room temperature.

TABLE 1. THE CALCULATED AND OBSERVED VALUES OF THE TRANSITION ENERGIES, OSCILLATOR STRENGTH, AND THE FARADAY B TERMS OF FORMS I—V

$\tilde{\nu}_{\mathrm{obsd}} \times 10^{-3}$ / cm^{-1}	f_{obsd}	$B_{\mathrm{obsd}} \times 10^4$ / βD^2 cm	$\tilde{\nu}_{\mathrm{calcd}} \times 10^{-3}$ / cm^{-1}	f_{calcd}	$B_{\mathrm{calcd}} \times 10^4$ / βD^2 cm	
Cytosine in H_2O			Form I			
			29.2	0.000	0.0	A″
			33.1	0.002	−3.2	A″
37.4	0.096	4.1	34.0	0.094	9.2	A′
			42.6	0.003	0.6	A″
42.0	0.100	6.5	43.4	0.090	−1.1	A′
			44.5	0.000	0.7	A″
47.0	0.211		46.9	0.931	0.2	A′
			48.7	0.004	−0.7	A″
52.1	0.639		50.4	0.052	−10.5	A′
3-Methylcytidine			Form II			
			30.6	0.000	0.0	A″
			31.2	0.002	−2.1	A″
37.4	0.17	5.6	37.5	0.305	8.3	A′
			42.6	0.000	5.3	A″
			42.6	0.254	−1.9	A′
45.1	0.39	−1.0	44.6	0.114	−8.4	A′
			47.6	0.018	6.3	A″
			48.6	0.009	−0.1	A″
3-Methylcytosine			Form III			
			27.8	0.000	0.0	A″
34.6	0.198	10.3	31.8	0.254	−23.6	A′
			32.0	0.004	31.8	A″
			43.4	0.002	0.0	A″
			43.7	0.000	0.0	A″
44.1		−4.6	44.8	0.357	−6.5	A′
	0.275		49.3	0.013	13.6	A″
48.5[a]			49.4	0.201	−31.9	A′
			49.6	0.005	6.4	A″
Isocytosine in ethanol			Form IV			
			27.3	0.001	−0.3	A″
35.4	0.169	9.3	33.4	0.264	12.1	A′
			39.2	0.004	−2.1	A″
44.5		−3.9	42.6	0.306	−14.3	A′
			42.8	0.002	−0.3	A″
	0.303		43.4	0.004	−2.5	A″
			44.3	0.019	7.5	A″
			47.7	0.000	−0.2	A″
49.0[a]			49.8	0.188	−8.6	A′
Isocytidine			Form V			
			23.4	0.000	0.0	A″
			33.8	0.002	−0.3	A″
39.1	0.126	5.7	37.3	0.021	4.6	A′
43.6[a]	0.159		41.4	0.387	−29.0	A′
			41.5	0.004	17.4	A″
			44.1	0.003	1.7	A″
45.7[b]		5.2	45.1	0.244	19.0	A′
			46.7	0.012	0.7	A″
			48.4	0.007	0.2	A″
48.5			49.5	0.348	−16.9	A′

a) Wavenumber at the shoulder of the absorption spectrum. b) Wavenumber at the peak of the MCD spectrum.

TABLE 2. THE CALCULATED AND OBSERVED VALUES OF THE TRANSITION ENERGIES, OSCILLATOR STRENGTH, AND THE FARADAY *B* TERMS OF 3-METHYLCYTIDINE AND 3-METHYLISOCYTOSINE

$\dfrac{\tilde{\nu}_{\mathrm{obsd}} \times 10^{-3}}{\mathrm{cm}^{-1}}$	f_{obsd}	$\dfrac{B_{\mathrm{obsd}} \times 10^{4}}{\beta D^2\,\mathrm{cm}}$	$\dfrac{\tilde{\nu}_{\mathrm{calcd}} \times 10^{-3}}{\mathrm{cm}^{-1}}$	f_{calcd}	$\dfrac{B_{\mathrm{calcd}} \times 10^{4}}{\beta D^2\,\mathrm{cm}}$	
3-Methylcytidine			3-Methyl-4-imino-2-pyrimidinone			
			30.5	0.000	−0.2	A″
			30.7	0.002	−1.6	A″
37.4	0.17	5.6	37.2	0.241	6.9	A′
			42.1	0.000	3.4	A″
			42.3	0.188	36.9	A′
45.1	0.39	−1.0	42.6	0.246	−41.1	A′
			45.5	0.007	2.8	A″
			47.8	0.003	−4.8	A″
3-Methylisocytosine			3-Methylisocytosine			
			27.4	0.001	−0.3	A″
35.2	0.174	10.6	33.3	0.238	10.7	A″
			39.3	0.002	−1.5	A″
			41.9	0.005	4.0	A″
44.3	0.254	−6.7	42.3	0.278	−19.0	A′
			42.9	0.000	0.0	A″
			44.1	0.015	5.6	A″
			47.8	0.000	2.2	A″
			48.0	0.199	−10.1	A′

sonably be assigned to the MCD bands of Form V. Thus, the MCD spectrum of isocytosine in an aqueous solution can be well interpreted by assuming that Form V is in equilibrium with Form IV in an aqueous solution of isocytosine.

The calculated transition energies, oscillator strengths, and Faraday *B* terms of Forms I—V are listed in Table 1, along with the experimental values. It should be noted that the positive *B* terms give negative MCD bands. In the ultraviolet region, the present calculation predicts three or four $\pi^* \leftarrow \pi$ transitions for Forms I—V; these transitions can reasonably be assigned to the peaks and the shoulder observed in the absorption and MCD spectra. On the other hand, several transitions with A″ symmetry are predicted below 50000 cm^{-1}. Most of them are characterized by a transition from the nonbonding orbitals of carbonyl oxygen and azanitrogen to the unoccupied π orbitals. Although the lowest $\pi^* \leftarrow \mathrm{n}$ transition of Form V is predicted at a much lower wavenumber than the lowest $\pi^* \leftarrow \pi$, no MCD or absorption band assignable to the lowest $\pi^* \leftarrow \mathrm{n}$ transition is identified.

A large *B* value is predicted for the second $\pi^* \leftarrow \mathrm{n}$ transition of Form III, whose oscillator strength is much smaller than that of the $\pi^* \leftarrow \pi$ transitions. The large *B* term of the second $\pi^* \leftarrow \mathrm{n}$ transition is mainly induced by the magnetic mixing of the second $\pi^* \leftarrow \mathrm{n}$ state with the first $\pi^* \leftarrow \pi$ state, which lies close to the second $\pi^* \leftarrow \mathrm{n}$ state, and this large *B* term is partially cancelled out by the negative *B* term of the first $\pi^* \leftarrow \pi$ transition in the MCD spectrum; the net contribution to the MCD spectrum from these two electronic transitions is about $+12.2 \times 10^{-4}\ \beta D^2$ cm.

The calculated values of the transition energies and the oscillator strengths of the $\pi^* \leftarrow \pi$ transitions of Forms I—V are in good agreement with the experimental values. The magnitudes of the experimental *B* terms of the second MCD band of cytosine are larger than the theoretical *B* values of the electronic transitions predicted in the spectral region of 42000—45000 cm^{-1}. The second $\pi^* \leftarrow \pi$ transition of Form V, whose *B* term is predicted to be negative and large, is not identified in the MCD spectrum of isocytidine. However, the agreement between the calculation and the experiment is satisfactory as a whole.

In order to study the effect of the methylation at the nitrogen atom on the calculated results, we have carried out a similar calculation for 3-methyl-4-imino-2-pyrimidinone (the *N*-methylated derivative of Form II) and for 3-methylisocytosine (the *N*-methylated derivative of Form IV); the results are shown in Table 2. The calculated transition energies of the *N*-methylated compounds tend to be slightly smaller than those of Form II and Form IV themselves. The calculated energy difference between the second and the third $\pi^* \leftarrow \pi$ transitions of Form II becomes smaller upon N-methylation, which increases the magnetic mixing between the second and the third $\pi^* \leftarrow \pi$ states. However, the calculation shows that the spectral features of Form II and Form IV are not much affected by the *N*-methylation.

References

1) W. Voelter, R. Records, E. Bunnenberg, and C. Djerassi, *J. Am. Chem. Soc.*, **90**, 6163 (1968); W. Voelter, G. Barth, R. Records, E. Bunnenberg, and C. Djerassi, *ibid.*, **91**, 6165 (1969); D. L. Elder, E. Bunnenberg, C. Djerassi, M. Ikehara, and W. Voelter, *Tetrahedron Lett.*, **10**, 727 (1970).

2) L. B. Townsend, D. W. Miles, S. J. Manning, and H. Eyring, *J. Heterocyclic Chem.*, **10**, 419 (1970).

3) D. W. Miles, W. H. Inskeep, L. B. Townsend, and

H. Eyring, *Biopolymers*, **11**, 1181 (1972).

4) R. E. Linder, H. Weiler-Feilchenfeld, G. Barth, E. Bunnenberg, and C. Djerassi, *Theoret. Chim. Acta*, **36**, 135 (1974); H. Weiler-Feilchenfeld, R. E. Linder, G. Barth, E. Bunnenberg, and C. Djerassi, *ibid.*, **46**, 79 (1977).

5) A. Kaito and M. Hatano, *Bull. Chem. Soc. Jpn.*, **53**, 3069 (1980).

6) S. L. Wallace, A. Castellan, D. Müller, and J. Michl, *J. Am. Chem. Soc.*, **100**, 6828 (1978); I. Jonas and J. Michl, *ibid.*, **100**, 6834 (1978).

7) J. D. Watson and F. H. Crick, *Cold Spring Harbor Symp. Quant. Biol.*, **18**, 123 (1953).

8) C. Hélène, A. Haug, M. Delbrück, and P. Douzou, *C. R. Acad. Sci. Paris*, **259**, 3385 (1964).

9) C. Hélène and P. Douzou, *C. R. Acad. Sci. Paris*, **259**, 4853 (1964).

10) H. Morita and S. Nagakura, *Theor. Chim. Acta*, **11**, 279 (1968).

11) G. C. Y. Lee, J. H. Prestegrad, and S. I. Chan, *J. Am. Chem. Soc.*, **94**, 951 (1972).

12) Y. P. Wong, *J. Am. Chem. Soc.*, **95**, 3511 (1973).

13) C. Hélène and P. Douzou, *C. R. Acad. Sci. Paris*, **259**, 4387 (1964).

14) D. L. Barker and R. E. Marsh, *Acta Crystallogr.*, **17**, 1581 (1964).

15) B. D. Sharma and J. F. McConnell, *Acta Crystallogr.*, **19**, 797 (1965).

16) J. A. Pople, *J. Chem. Phys.*, **47**, 2026 (1969).

17) J. Ridley and M. Zerner, *Theor. Chim. Acta*, **32**, 111 (1973).

18) R. B. Angier and W. V. Curran, *J. Org. Chem.*, **26**, 1891 (1961).

19) T. Ueda and S. Shibuya, unpublished experimental data.

20) A. D. Buckingham and P. J. Stephens, *Annu. Rev. Phys. Chem.*, **17**, 399 (1966).

21) T. Ueda and J. J. Fox, *J. Am. Chem. Soc.*, **85**, 4024 (1963).