Theoret. chim. Acta (Berl.) 11, 279—295 (1968)

The Electronic Absorption Spectra and the Electronic Structures of Cytosine, Isocytosine, and Their Anions and Cations

HIROSHI MORITA and SABURO NAGAKURA

The Institute for Solid State Physics,The University of Tokyo, Roppongi, Minato-ku Tokyo, Japan

Received April 9, 1968

The electronic absorption spectra of cytosine and isocytosine have been measured in several solvents at room temperature, and the temperature and pH dependencies of the electronic absorption spectra in aqueous solution have also been examined, special attention being paid to the existence of tautomeric forms. Consequently, it is concluded that cytosine in aqueous solution takes 2-keto-6-amino (3H) form (I) at room temperature, but with rising temperature 2-keto-6-imino form (III) coexists with form I, the latter being determined to be more stable by 5.5 kcal/mole than the former, and that cytosine takes form I and form III in trimethyl phosphate and in acetonitrile, respectively. The change in stability of the tautomeric forms by different solvents has been discussed with regard to their calculated dipole moments due to π-electrons and the polarity of the solvents.

The π-electron structures of several tautomeric forms of cytosine (5 tautomeric forms) and isocytosine (3 tautomeric forms) and of their ions have been calculated by combining the configuration interaction with the Pariser-Parr-Pople SCF MO method. The transition energies, oscillator strengths, π-electron densities, π-bond orders and dipole moments have been evaluated. Concerning the transition energy and oscillator strength, the comparison has been made between the theoretical and observed values, showing that the agreement between them is improved in the present calculation compared with previous calculations.

Die Elektronen-Absorptionsspektren von Cytosin und Isocytosin wurden in verschiedenen Lösungsmitteln bei Raumtemperatur gemessen; in wäßriger Lösung wurden auch Temperatur und pH-Abhängigkeiten untersucht, wobei besonders auf die Existenz von Tautomeren geachtet wurde. Es ergab sich, daß Cytosin in wäßriger Lösung bei gewöhnlicher Temperatur in der 2-Keto-6-amino-Form vorliegt, aber bei steigender Temperatur daneben auch die 2-Keto-6-imino auftritt. Letztere dürfte etwa 5,5 Kcal/Mol stabiler sein. Cytosin in Trimethylphosphat und Acetonitril liegt in Form I bzw. Form II vor. Die unterschiedlichen Stabilitäten in verschiedenen Lösungsmitteln wurden in Hinblick auf die berechneten Dipolmomente, die vom π-Elektronensystem und der Polarität des Lösungsmittels herrühren, diskutiert.

Die π-Elektronenstrukturen von fünf tautomeren Formen von Cytosin und von Isocytosin (drei tautomere Formen) sowie von deren Ionen wurden mittels einer Kombination von Konfigurationswechselwirkung und PPP-SCF-MO-Verfahren berechnet. Übergangsenergien, Oszillatorenstärken, π-Elektronendichten und Bindungsordnungen sowie Dipolmomente wurden berechnet, wobei sich bei den beiden ersten Größen Verbesserungen gegenüber früheren Rechnungen ergaben.

Les spectres d'absorption électronique de la cytosine et de l'isocytosine ont été obtenus à température ambiante dans différents solvants; en milieu aqueux la dépendance de la température et du pH a été étudiée, permettant des conclusions sur les problèmes de tautomérie. Il apparaît qu'en milieu aqueux, à température ordinaire, la cytosine se trouve sous la forme 2-céto, 6-amino, mais qu'à température plus élevée elle se trouve aussi sous la forme 2-céto,6-imino. Cette dernière est plus stable d'environ 5,5 Kcal/Mol. Un équilibre analogue se produit dans le triméthylphosphate et l'acétonitrile. Les différentes stabilités dans les différents solvants sont discutées en fonction des moments dipolaires calculés, qui proviennent du système d'électrons π et de la polarité du milieu.

La structure électronique π de cinq formes tautomères de la cytosine et de trois formes tautomères de l'isocytosine ainsi que des ions correspondants a été calculée par une méthode S.C.F. M.O. P.P.P. avec I.C. Les énergies de transition, les forces d'oscillateur, les charges et les indices de liaison des électrons π sont calculés, avec de meilleurs résultats que ceux obtenus précédemment pour les deux premières grandeurs.

Introduction

The electronic structures of the purine and pyrimidine bases of the nucleic acids have recently been studied extensively by many authors from both experimental and theoretical points of view [2, 3, 9, 13, 16, 18, 19, 26]. In a previous paper [27], Tanaka and one of the present authors carried out the semiempirical SCF MO CI calculation on the π-electron structures of adenine and thymine, and derived the satisfactory results on the assignment of their electronic absorption bands. In the present paper, the authors have undertaken to extend the study to cytosine and isocytosine.

A number of spectroscopic studies have already been carried out for cytosine and isocytosine by the aid of the vibrational [7, 17, 25, 28, 29] and electronic absorption spectra [4, 9, 10, 11, 12, 14, 32], special attention being paid to their tautomeric forms. Consequently, cytosine in aqueous solution was concluded to take 2-keto-6-amino form (I) at room temperature. Concerning isocytosine, the situation was found to be more complicated; namely, it takes predominantly 2-amino-6-keto (1 H) isocytosine (VII) in ethanol and in ethyl ether, but in aqueous solution, two tautomeric forms, i.e. 2-amino-6-keto (3H) isocytosine (VI) and form VII exist in a ratio of 48 : 52. Further, the X-ray crystal analysis studies led to the conclusion that cytosine [1] exists in form I in crystals, while the isocytosine crystal consists of the two tautomeric forms, i.e. form VI and form VII in an exact 1:1 ratio [24].

As mentioned above, the results on cytosine seem to be rather simple. However, Hélène and Douzou showed [10, 11] that the electronic absorption spectrum of cytosine changes, though only to a small extent, depending on solvents and temperatures. This fact suggests that form I is changeable to other tautomeric forms in solutions. Berthod and Pullman [2] carried out the Hückel type calculation and concluded that the total energy of form I is lower by several kilocalories than that of form III, although Hoffman and Ladik [13] predicted that form III is more stable than form I.

Under these circumstances, we have undertaken to re-examine the stability of the tautomeric forms of cytosine and isocytosine by measuring the electronic absorption spectra in several solvents, such as water, acetonitrile and trimethyl phosphate (abbreviated hereafter to T.M.P.). In order to clarify the assignments of the observed spectra, the electronic structures of some tautomeric forms of cytosine and isocytosine have been calculated by combining the configuration interaction with the Pariser-Parr-Pople SCF MO method and the evaluated transition energies and oscillator strengths have been compared with the observed values. Furthermore, the electronic structures of the anions and cations of cytosine and isocytosine have also been calculated, the change of the σ-framework through protonation or deprotonation being considered.

Experimental

Materials

Cytosine of *G.R.* grade was purified by vacuum sublimation. Isocytosine was synthesized from guanidine hydrochloride ($NH_2C{=}NHNH_2 \cdot HCl$) and malic acid ($C_3H_5O_3COOH$) according to Caldwell and Kime's method [8] and was purified by repeating three times recrystallization from aqueous solution and finally by vacuum sublimation. Acetonitrile was refluxed over calcium hydride (CaH_2) for 10 hours and then over phosphorus pentoxide for 2 hours. It was finally distilled from CaH_2. T.M.P. of *E.P.* grade was used without further purification.

Measurements. Ultraviolet absorption spectra were measured by a Cary recording spectrophotometer model 14 M, cells of 1 mm light path length being used. Vacuum ultraviolet absorption spectra were measured by a spectrophotometer constructed in our laboratory [30], cells of 0.1 mm light path length being used.

Theoretical

Procedure of Calculation

The π-electron structures of cytosine and isocytosine and their ions were calculated by considering the configuration interaction among various electron configurations constructed from the Pariser-Parr-Pople SCF MO [21, 22], the $\sigma-\pi$ separability approximation being employed and only the π-electrons being taken into accont. The one center Coulomb repulsion integrals $(pp|pp)$ were taken from Orloff and Sinanoğlu's paper [20]. The two center Coulomb repulsion integrals $(pp|qq)$ were calculated by means of the uniformly charged sphere model [21]. The core Coulomb integrals (α_p) were estimated by the following formula derived on the neglection of the Coulomb penetration integrals $(q|pp)$:

$$\alpha_p = -W_{2p}^p - \sum_{q(\neq p)} n_q \times (pp|qq) \tag{1}$$

where n_q is the number of π-electron on atom q, and W_{2p}^p is the first (when $n_p = 1$) or the second (when $n_p = 2$) valence state ionization potential of the $2p\pi$-orbital of atom p. The core resonance integrals (β_{pq}) were estimated by the following formula;

$$\beta_{pq} = -0.869 \times \tfrac{1}{2}(W_{2p}^p + W_{2p}^q) \times S_{pq} \tag{2}$$

where S_{pq} is the overlap integral between the Slater type atomic orbitals at atoms p and q, and the numerical coefficient (i.e. 0.869) was determined so as to reproduce the β_{cc} value for benzene ($\beta_{cc} = -2.39$ eV at $r = 1.39$ Å).

Further, the zero differential overlap approximation was employed. The configuration interaction (C.I.) treatment was made by taking eight singly excited configurations. In actual calculation, the W_{2p} values were determined in such a way that the calculated transition energies agreed well to the observed ones. The W_{2p}^C, W_{2p}^N, $W_{2p}^{N^+}$, and $W_{2p}^{O^+}$ values commonly used for all the tautomeric forms of cytosine and isocytosine are -11.22, $-14.51+0.50$, $-28.72+1.00$, $-33.82+1.00$ eV, respectively, but the W_{2p}^O values used for tautomeric forms are different from one another, as is listed in Table 1.

The geometrical structures of cytosine and isocytosine were taken from the X-ray crystal analysis data by Barker and Marsh [1], and by Sharma and McConnell [24], respectively.

Table 1. *The valence state ionization potential of the oxygen atom (W_{2p}^{Os}) determined parametrically for the tautomeric forms of cytosine and isocytosine (in eV)*

	Cytosine		Isocytosine		
	Amino forms (I) and (II)	Imino form (III)	(3H) form (VI)	(1H) form (VII)	Imino form (VIII)
W_{2p}^{Os}	$-17.73-1.00$	$-17.73+1.00$	$-17.73-2.00$	$-17.73+1.50$	$-17.73+1.00$

The π-electron structures of the anions and cations of cytosine and isocytosine were calculated by assuming the same geometrical structures as those of the respective neutral molecules. The calculation procedures for the ions were almost the same as those mentioned above for the neutral molecules, except for the following points:

a) The two nitrogen atoms in the pyrimidine ring, i.e. N_1 and N_3 are considered to be equivalent, and their valence state ionization potentials (W_{2p}^{N}) are treated as parameters. The values of α_p and β_{pq} with which nitrogen atoms N_1 and N_3 are concerned are determined using Eqs. (1) and (2) as a function of W_{2p}^{N}.

b) The $(pp|qq)$ values of the cytosine ions and the isocytosine ions are estimated to be equal to the arithmetic means of those of amino forms (I) and (II) and of those of amino forms (VI) and (VII), respectively.

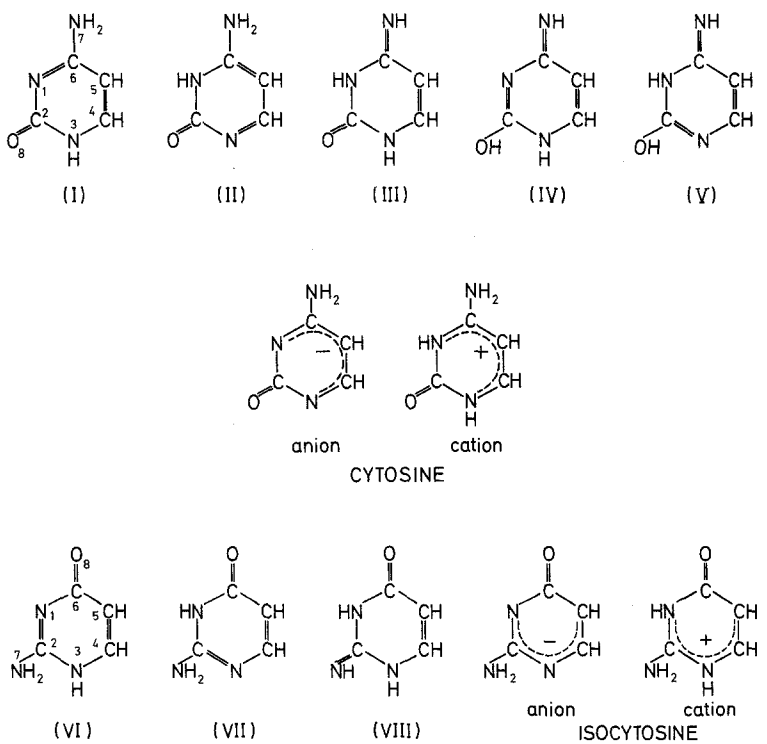

Fig. 1. Tautomeric forms of cytosine and isocytosine, and their anions and cations

c) Each nitrogen atom in the pyrimidine ring is initially assumed to contribute one and a half electrons to the π-electron framework. This assumption is reasonable and necessary for estimating the α_p value by the aid of Eq. (1).

Results and Discussion

Electronic Structure and Spectra of Cytosine

The electronic absorption spectra of cytosine in aqueous solution and in acetonitrile are shown in Fig. 2. The temperature dependence of the absorption spectrum in aqueous solution is shown in Fig. 3a. This figure shows the difference

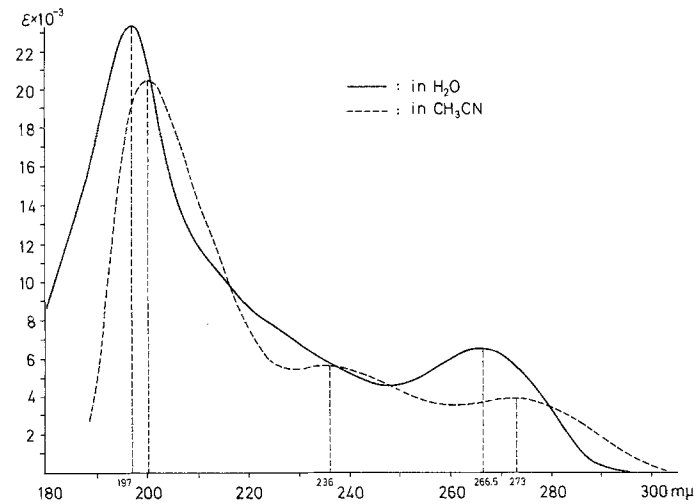

Fig. 2. Electronic absorption spectra of cytosine in aqueous solution and in acetonitrile

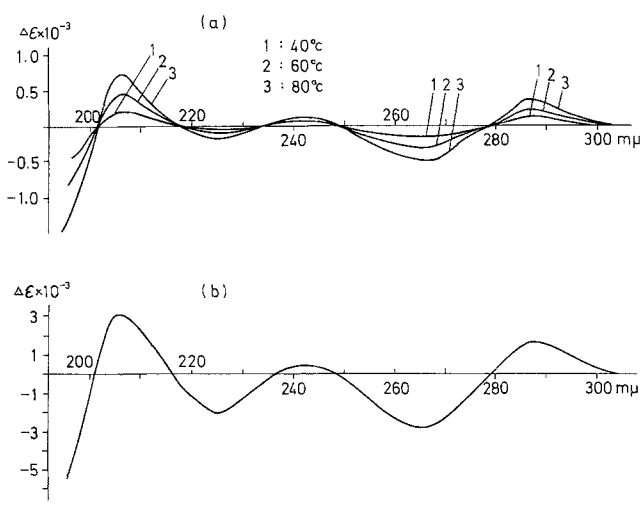

Fig. 3. a $\Delta\varepsilon - \lambda$ curves for cytosine in aqueous solution ($\Delta\varepsilon$ = molar extinction coefficient at respective temperatures − molar extinction coefficient at room temperature) b $\Delta\varepsilon' - \lambda$ curves for cytosine ($\Delta\varepsilon'$ = molar extinction coefficient in acetonitrile − molar extinction coefficient in aqueous solution)

between the absorption spectrum of cytosine of the aqueous solution at high temperature and that at room temperature. Fig. 3a leads to the conclusion that in aqueous solution cytosine exists as an equilibrium mixture between two forms; one of them is predominant at high temperature; the other which shows the absorption peaks at 266.5 and 197.0 mμ and a shoulder at 228 mμ is predominant at room temperature and has already been concluded by Katritzy and Waring [14] and also by Brown and Lyall [4] to be form I. Fig. 3b shows the difference between the absorption spectrum of cytosine in acetonitrile and that of the aqueous solution at room temperature. It is noteworthy that the curves in Figs. 3a and 3b are similar to each other. This means that the form predominant in aqueous solution at high temperature is the same as the stable one in acetonitrile.

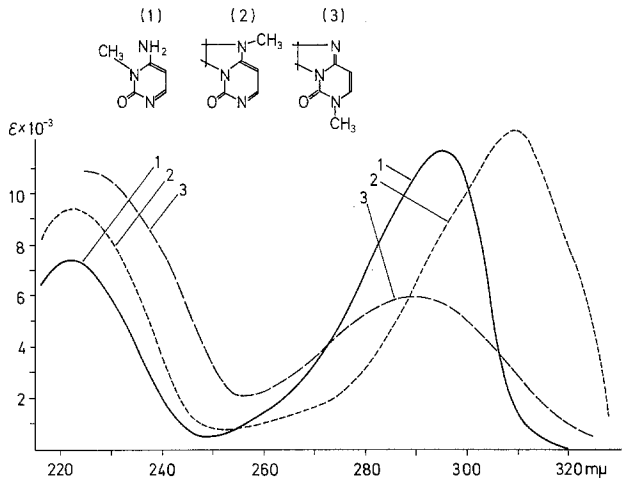

Fig. 4. Electronic absorption spectra of N-methylated cytosines in aqueous solution [Ueda, T., and J. J. Fox: J. Amer. chem. Soc. **85**, 4024 (1963)]

From Fig. 3a, we could determine the value of the equilibrium constant, K = [high temperature form]/[amino form (I)], to be 0.03, 0.07, and 0.12 at 30° C, 50° C and 70° C, respectively. From the temperature dependence of the equilibrium constant, we can obtain the energy and entropy differences between the high temperature form and the amino form (I) to be $\Delta H^\circ = \sim 5.5$ kcal/mole and $\Delta S^\circ = \sim 12$ cal/mole · deg., respectively.

In order to identify the high-temperature form in aqueous solution, we have undertaken to compare the absorption spectrum of the acetonitrile solution of cytosine with those of the aqueous solutions of related substances, which were measured by Ueda and Fox [31] and are shown in Fig. 4. From the similarity of the former absorption curve to curve 3 in Fig. 4 in their peak positions and relative intensities, the stable form of cytosine in acetonitrile and in aqueous solution at high temperature may be concluded to be the 2-keto-6-imino form (III).

In contrast with this result, other authors suggested the equilibrium in aqueous solution between tautomeric forms I and II [14] or among forms I, III, and IV [11]. According to our opinion, however, their suggestions can not be considered to be based on definite experimental or theoretical facts. The present theoretical

study on transition energies, oscillator strengths and dipole moments for the five possible isomers of cytosine also supports our conclusion on the high-temperature form of cytosine in aqueous solution.

Now let us turn to the theoretical results. The transition energies and oscillator strengths calculated for form I in aqueous solution are listed in Table 2, together with the observed values. The theoretical results on the directions of transition moments, the π-electron densities and the π-bond orders are given in Fig. 5.

Kwiatkowski [15], and also Ladik *et al.* [16] have already calculated the transition energies of cytosine by the SCF MO CI method, although their results on the third singlet-singlet transition band are too low (5.63 eV) or too high (6.79 eV), respectively, compared with the observed result (6.29 eV). Furthermore,

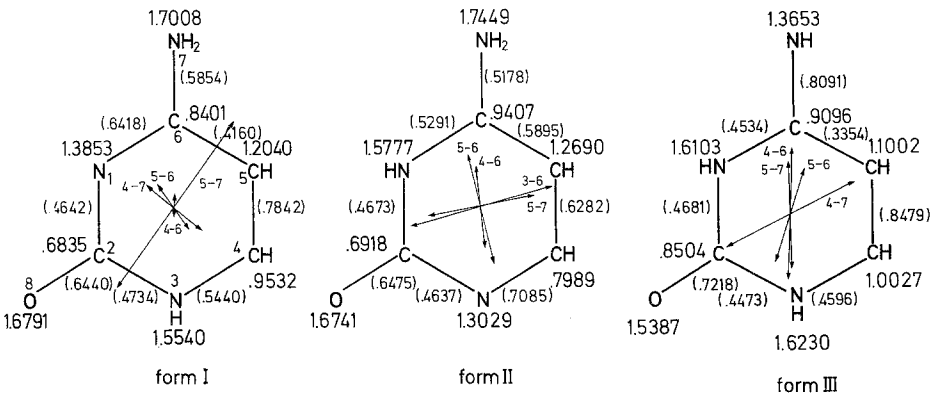

Fig. 5. The π-electron densities, π-bond orders (in parentheses), and directions of the transition moments calculated for forms I, II, and III

Ladik *et al.* [16] obtained a low energy value for the first triplet state. Berthod *et al.* [3] calculated the transition energies using SCF MO with and without CI. Their result without CI is in fairly good agreement with the observed values owing to the appropriate choice of parameters. Nagata *et al.* [18] also calculated the transition energies by the SCF MO method. The comparison of our results with those of the other authors' [3, 15, 16, 18] shows that the agreement on the transition energies and oscillator strengths between the observed and theoretical values, are much improved in our calculation.

The directions of the transition moments calculated by the present authors are almost the same for the first three bands (5–6, 4–6, 5–7) as those by Berthod *et al.* [3] and by Nagata *et al.* [18]. The negative charge at C_5 and the greatest π-bond order at the C_4–C_5 bond are the main features of the cytosine molecule [23a]. These features are well explained by the present study, as is clear in Fig. 5.

In order to check our opinion on the high temperature form in aqueous solution or on the stable form in acetonitrile solution, we calculated the π-electron structures of the tautomeric forms other than form I; namely, 2-keto-6-amino (1H) cytosine (II), form III, 2-hydroxy-6-imino (3H) cytosine (IV), and 2-hydroxy-6-imino (1H) cytosine (V). The calculated transition energies and oscillator strengths are tabulated in Table 2, together with the observed values of cytosine in aceto-

Table 2. *The transition energies (ΔE) and the oscillator strengths (f) calculated for forms I, II, III, IV, and V*

Cytosine in H_2O		Amino form (I)				
$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.651	0.110	4.704	0.075	5–6[c]	3.048	5–6
5.438[a]	0.678[b]	5.479	0.011	4–6	3.487	4–6
6.293		6.432	1.370	5–7	4.409	5–7
		6.681	0.163	4–7	5.597	4–7
		7.700	0.458	3–6	6.482	3–6
1-methyl-cytosine[d]		Amino form (II)				
$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.216	0.18	4.347	0.282	5–6	2.601	5–6
5.585	0.17	5.453	0.322	5–7	3.761	5–7
		6.686	0.227	4–6	4.871	3–6, 5–8
		7.073	0.770	3–6	5.660	4–6, 3–6
		7.476	0.066	5–8, 4–7	6.220	4–6
Cytosine in CH_3CN		Imino form (III)				
$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.542	0.054	4.554	0.224	5–6	2.201	5–6
5.187	0.111	5.445	0.328	5–7, 4–6	3.091	4–6, 4–7
5.793	0.501	5.684	0.497	4–6	3.430	4–7
6.199		6.222	0.698	4–7	5.053	5–7, 4–6
		7.917	0.043	5–8	7.075	3–6, 4–7, 5–8

Hydroxy form (IV)			Hydroxy form (V)		
$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration
3.363	0.014	5–6	4.040	0.131	5–6
4.976	0.523	5–7	4.993	0.616	5–7
6.455	0.055	4–7	6.614	0.429	4–6
6.571	0.103	4–6	6.679	0.038	4–7
7.021	0.001	5–8	7.349	0.248	5–8

[a] shoulder.

[b] Estimated in the wavelength region longer than 180 mμ.

[c] $m-n$ denotes a configuration in which one electron of the m-th molecular orbital is excited to the n-the molecular orbital.

[d] Ueda, T., and J. J. Fox: J. Amer. chem. Soc. **85**, 4026 (1963).

nitrile and 1-methyl cytosine [31] which corresponds to form II. Furthermore, the calculated bond orders and π-electron densities are given in Fig. 5.

The transition energies and oscillator strengths calculated for form III can well explain the observed spectroscopic data of cytosine in CH_3CN. This seems to support our opinion on the stable form of cytosine in acetonitrile. Further evidence is obtained by measuring the temperature dependence of the spectrum of cytosine in acetonitrile solution. In this case, rising up the temperature, we can observe the spectral change which indicates the formation of two species, i.e. form I and to much minor extent form II in addition to form III [1]. This shows that the most stable form in acetonitrile is not form II, but form III.

It is to be noticed that the most stable form of cytosine in acetonitrile is imino form III, while in aqueous solution amino form I is more stable than form III. This result may probably be due to the fact that form I is more stabilized by electrostatic interaction with water molecules than form III is. This is reasonable from the comparison of the polarities of these two tautomeric forms. According to Fig. 5, the $N_1-C_2=O_8$ part in the amino form is highly polarized than that in the imino form is. Owing to this, the dipole moment due to π-electrons of form I turns out to be several times larger than that of form III (see Table 3). This seems to mean that form I is more stabilized in polar solvents, say in water, than form III.

Table 3. *Dipole moments of some tautomeric forms of cytosine (in Debye unit)*

	D_x[a]	D_y[a]	D
Amino form (I) in aqueous solution	+7.437	−0.391	7.447
Amino form (II)	+6.485	6.440	9.139
Imino form (III)	+0.694	−3.350	3.421

[a] The x and y axes are taken to be parallel and perpendicular to the N_3–C_4 direction.

Electronic Structure and Spectra of Isocytosine

The electronic absorption spectra of isocytosine in aqueous, acetonitrile and T.M.P. solutions are shown in Fig. 6. The spectra in acetonitrile and T.M.P. are identical with that in ethyl ether measured by Hélène and Douzou [12]. By an analogy with the case of ethyl ether, it is concluded that isocytosine takes 2-amino-6-keto (1H) isocytosine (VII) in the above two solvents. In aqueous solution, isocytosine is known to consist of two tautomeric forms; i.e., 2-amino-6-keto (3H) isocytosine (VI) and form VII [5, 11, 12]. The former was shown by Hélène and Douzou [11] to be more stable by ~ 1 kcal/mole than the latter in aqueous solution. We determined the corresponding energy difference and the entropy change to be 1.3 kcal/mole and 4.7 cal/mole deg., respectively, from the temperature dependence of the absorption spectrum shown in Fig. 6. This energy difference is much smaller than the value between forms I and III of cytosine in aqueous solution. By analyzing the observed spectrum, the absorption maxima

[1] In this analysis, the spectrum of 1-methyl-cytosine [31] is regarded as that of form II. This may be reasonable because the effect of the methyl substitution upon the absorption spectrum is in general rather small.

of form VI are concluded to appear at 259, 220, and 202.5 mμ [2]. The first maximum wavelength is a little different from the value (255 mμ) determined by Hélène and Douzou [12].

The π-electron structures of form VI, form VII and 2-imino-6-keto isocytosine (VIII) have been calculated by means of the SCF MO CI method, the valence state ionization potentials listed in Table 1 being used. The calculated transition energies and oscillator strenghts are given in Table 4, compared with the observed values. From this table, it is concluded that concerning the transition energies, the calculated values well agree with the observed values except for the

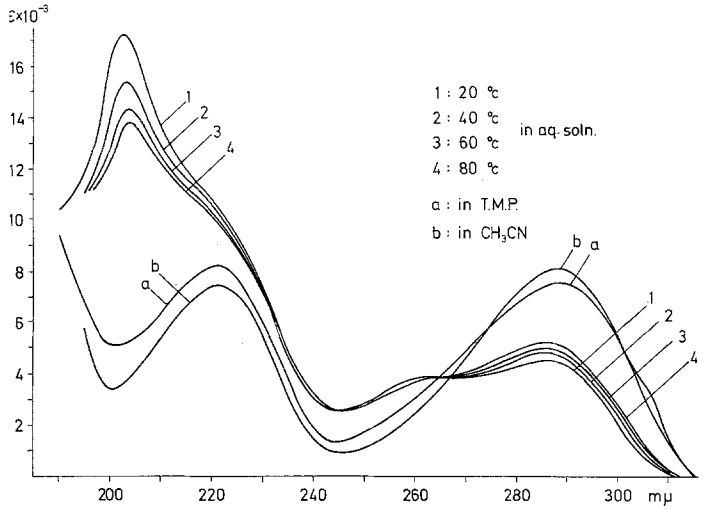

Fig. 6. Electronic absorption spectra of isocytosine under various conditions

third transition of form VII, but the calculated oscillator strengths do not agree quantitatively, though agree qualitatively, with the observed ones.

The values of the π-electron densities, the π-bond orders and the directions of the transition moments of these tautomeric forms of isocytosine are shown in Fig. 7. Pullman and Pullman [23b] calculated the former two quantities of some tautomeric forms of isocytosine by the simple Hückel method. Analogous to the case of cytosine, the greatest π-bond order at C_4–C_5 bond is maintained throughout these three tautomeric forms of isocytosine. Further, for form VII and form VIII, the negative charge appears on C_5 as for cytosine, but this feature does not hold for form VI. Among these three tautomeric forms of isocytosine, only form VI has the highly polarized part; $N_1-C_6=O_8$. The calculated values of the dipole

[2] These values are obtained by subtracting the spectrum of form VII from the spectrum observed with aqueous solution by the aid of the existing ratio at room temperature of form VI to from VII, i.e., 48 : 52 [12]: On that occasion, the 288 and 221 mμ bands of form VII are assumed to shift, respectively, by 2 mμ to shorter wavelengths and by 5 mμ to longer wavelengths in aqueous solution compared with the spectrum in acetonitrile. These shifts are reasonable from the Pariser-Parr-Pople type calculation of form VII. In aqueous solution, the absolute values of W_{2p}^{O} and $W_{2p}^{N_3}$ are thought to be larger than those in acetonitrile because of weak hydrogen bonding with water molecules. The calculated results show that the 288 and 221 mμ bands shift to shorter and longer wavelengths, respectively, with the increasing $|W_{2p}^{O}|$ and $|W_{2p}^{N_3}|$ values.

Table 4. *The calculated and observed values of the transition energies (ΔE), and the oscillator strengths (f) of forms VI, VII, and VIII*

2-amino-6-keto (3H) isocytosine (VI)

$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.787	0.080	4.796	0.026	5–6	3.204	5–6
5.635[a] 6.122	0.837	5.623	0.126	5–7	3.693	5–7 4–6
		6.504	0.852	4–7 4–6	4.538	5–7 4–6 5–8
		6.786	0.769	4–6 5–7 4–7	5.387	4–7
		7.927	0.173	3–6	6.805	3–7 5–8 3–6

2-amino-6-keto (1H) isocytosine (VII)

$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.335	0.166	4.348	0.227	5–6	2.607	5–7
5.486	0.232	5.406	0.532	5–7	3.063	5–6
		6.848	0.380	4–6	4.491	5–8 4–7 3–6
		7.002	0.119	4–7	5.884	4–6
		7.667	0.583	3–6	6.519	3–6 4–7

2-imino-6-keto-isocytosine (VIII)

$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.786	0.312	5–6	2.581	5–7
5.477	0.039	4–6	2.720	5–6
5.647	0.705	5–7	3.465	4–7 5–6
6.082	0.582	4–7	5.190	4–6 5–7 4–7
8.020	0.178	3–6	6.926	3–6

[a] shoulder.

moments of these tautomeric forms are listed in Table 5. The calculated dipole moment of form VI is two times larger than that of form VII. This may be a reason why the former is more stable than the latter in aqueous solution, while the latter is the stable form in acetonitrile.

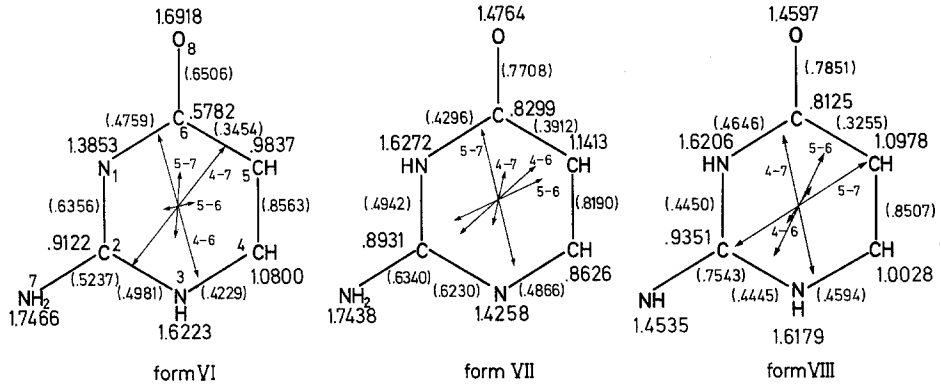

Fig. 7. The π-electron densities, π-bond orders (in parentheses), and directions of the transition moments calculated for forms VI, VII, and VIII

Table 5. *Dipole moments of some tautomeric forms of isocytosine (in Debye unit)*

	D_x[a]	D_y[a]	D
(3H) form (VI)	−6.599	−9.142	11.275
(1H) form (VII)	−6.503	−0.339	6.512
Imino form (VIII)	0.094	−3.944	3.945

[a] The x and y axes are taken to be parallel and perpendicular to the N_3–C_4 direction.

Electronic Structure and Spectra of the Cytosine Ion

The electronic absorption spectra of cytosine in acidic and alkaline aqueous solutions are shown in Fig. 8, and the observed values are listed in Table 6.

The calculated values of the transition energies are shown in Fig. 9, taking the valence state ionization potential in pyrimidine ring, W_{2p}^{N} as a variable. Compared with the observed values, −19.6 eV and −21.1 eV are obtained as the best values of ionization potentials for the nitrogen $2p$ orbitals of the pyrimidine rings

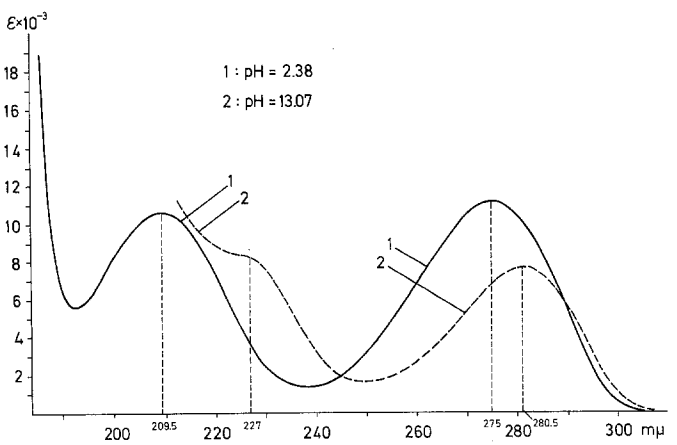

Fig. 8. Electronic absorption spectra of the cytosine anion (at $pH = 13.07$) and cation (at $pH = 2.38$)

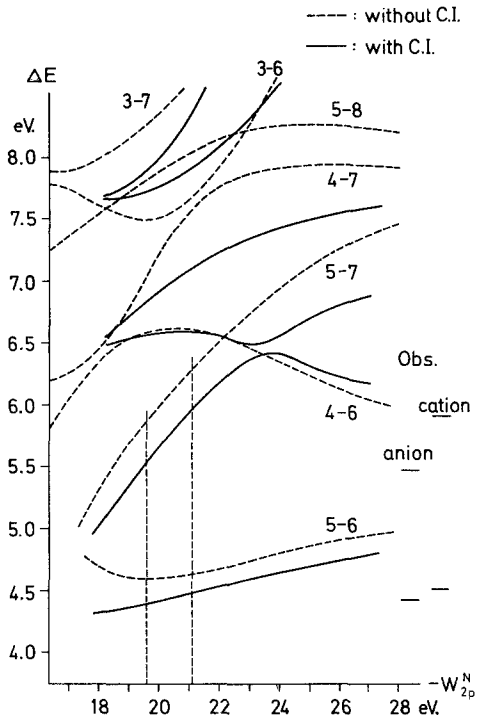

Fig. 9. Dependence of the calculated transition energies of the cytosine anion and cation upon W^{N}_{2p} of the pyrimidine ring (Broken lines represent the transition energies without C.I., and solid lines, with C.I.)

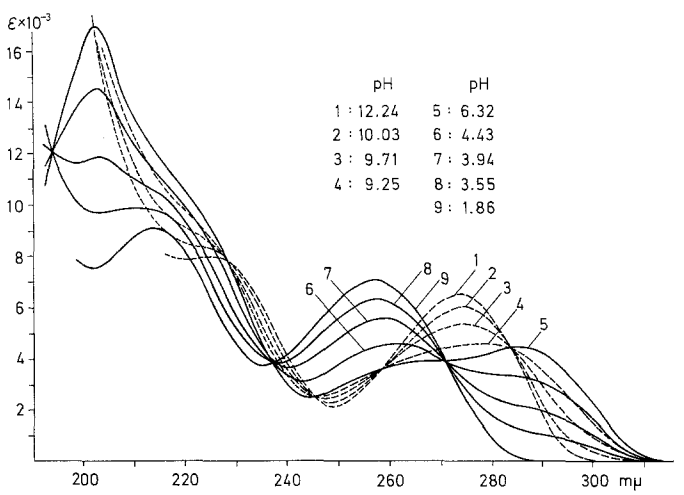

Fig. 10. Electronic absorption spectra of isocytosine in aqueous solution at various *pH* values

of the anion and cation, respectively. Their absolute values are located between the first and second valence state ionization potentials of nitrogen (14.01 eV and 27.72 eV respectively). This is reasonable in view of our initial assumption that the nitrogen atoms of the pyrimidine ring contribute one and a half electrons to the π-electron systems of the cytosine cation and anion. Further, in order to take into

Table 6. *The calculated and observed values of the transition energies (ΔE) and the oscillator strengths (f) of the cytosine anion and cation*

Cytosine anion						
$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.420	0.128	4.415	0.128	5–6	3.108	5–6
5.461	0.124[a]	5.547	0.294	5–7	3.413	5–7
		6.412	0.787	4–6	3.990	4–6
		6.687	0.315	4–7	5.078	4–7 5–8
		7.698	0.436	3–6	6.406	3–6
Cytosine cation						
$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.508	0.219	4.569	0.172	5–6	3.350	5–6
5.917	0.292	5.987	0.351	5–7	3.907	5–7
		6.598	0.627	4–6	4.745	4–6
		7.105	0.277	4–7	5.350	4–7 5–8
		7.923	0.341	3–6	6.512	3–6

[a] Estimated in the wavelength region longer than 220 mμ.

The following W_{2p} values (eV) are used for the above calculation

	Cation	Anion
W_{2p}^{C}	−11.22	−11.22
$W_{2p}^{N_{ring}}$	−14.51 − 6.6	−14.51 − 5.1
$W_{2p}^{N_{amino}}$	−28.72 + 0.5	−28.72 + 1.5
W_{2p}^{O}	−17.73 − 0.2	−17.73 + 0.2

account of a little change of the σ-framework at other hetero-atoms through protonation or through deprotonation, the values of ionization potentials of the nitrogen atom in the amino group (N_7) and of the oxygen atom O_8) are also altered, and the values listed in Table 6 are obtained as the most favourable ones [3]. The calculated transition energies and oscillator strengths are listed in Table 6.

The difference in $W_{2p}^{N_7}$ between the anion and the cation is taken to be 1.0 eV. This difference is somewhat arbitrary, because the transition energies are not sensitive to $W_{2p}^{N_7}$. On the other hand, the transition energies are so sensitive to $W_{2p}^{O_8}$ that the difference in $W_{2p}^{O_8}$ between the anion and the cation must be carefully determined to be 0.4 eV.

Electronic Structure and Spectra of the Isocytosine Ion

The electronic absorption spectra of isocytosine at various *pH* values are shown in Fig. 10. In Fig. 11, the molar extinction coefficients at 257.5 mμ for

[3] We consider that this procedure is more practical than the VESCF method [6].

acidic solutions and at 274 mμ for alkaline solutions are plotted against *pH* values. From this figure, the *pKa* values are determined to be 4.0 and 9.6 which are in good agreement with the values determined by Brown and Teitei potentiometrically [5].

Calculated values of the transition energies are shown in Fig. 12, taking W_{2p}^{N} of the nitrogen atoms in pyrimidine ring as a variable. Comparing the theoretical

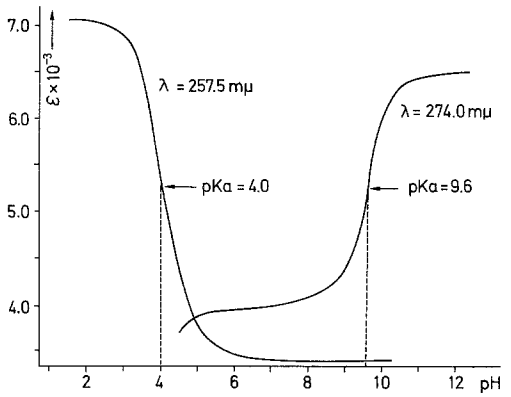

Fig. 11. The *pH* dependence of the molar extinction coefficients of the 274.0 mμ and 257.5 mμ bands for the isocytosine anion and cation, respectively

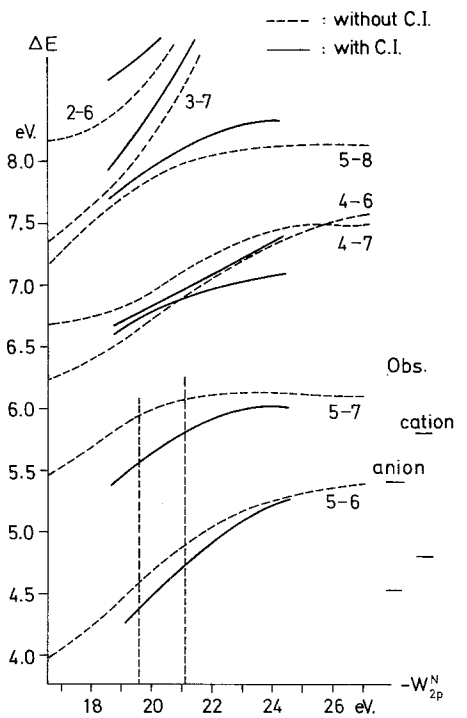

Fig. 12. Dependence of the calculated transition energies of the isocytosine anion and cation upon W_{2p}^{N} of the pyrimidine ring (Broken lines represent the transition energies without C.I., and solid lines, with C.I.)

result with the observed values, -19.6 eV and -21.1 eV are obtained as the best values of the ionization potentials of the nitrogen atoms of the pyrimidine ring for the anion and cation, respectively. Further, just as the case of the cytosine ion, in order to consider a little change of σ-framework at other hetero-atoms through protonation or through deprotonation, the values of the ionization potentials of the nitrogen atom (N_7) in the amino group and of the oxygen atom (O_8) are also altered. The values listed in Table 6, which are entirely the same as those for the cytosine ion, were determined in common in such a way as to give the best agreement with the observed values for the cytosine and isocytosine ions. The calculated transition energies and oscillator strengths are listed in Table 7, together with the observed values.

Table 7. *The calculated and observed values of the transition energies (ΔE) and the oscillator strengths (f) of the isocytosine anion and cation*

Isocytosine anion						
$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.525	0.117	4.394	0.103	5–6	2.973	5–7 5–6
5.395	0.141[a]	5.592	0.312	5–7	3.559	5–6 5–7
		6.579	1.135	4–6 4–7	4.267	4–6
		6.629	0.106	4–7 4–6	4.872	4–7
		7.831	0.081	5–8	6.691	3–7 4–7 5–8
Isocytosine cation						
$\Delta E_{obs.}$ (eV)	$f_{obs.}$	$\Delta E_{calc.}$ (singlet) (eV)	$f_{calc.}$	main configuration	$\Delta E_{calc.}$ (triplet) (eV)	main configuration
4.814	0.149	4.783	0.128	5–6	3.170	5–7 5–6
5.782	0.228[b]	5.843	0.434	5–7	3.954	5–6 5–7
		6.880	0.337	4–7	4.944	4–6
		6.957	0.655	4–6	5.402	4–7
		8.147	0.087	5–8	7.406	3–7 5–8 4–7

[a] Estimated in the wavelength region longer than 218 mμ.
[b] Estimated in the wavelength region longer than 201 mμ.

References

1. Barker, D. L., and R. E. Marsh: Acta Cryst. **17**, 1581 (1964).
2. Berthod, H., and A. Pullman: Biopolymers **2**, 483 (1964).
3. —, C. Giessner-Prettre, and A. Pullman: Theoret. chim. Acta (Berl.) **5**, 53 (1966).

4. Brown, D. J., and J. M. Lyall: Australian J. Chem. **15**, 851 (1962).
5. —, and T. Teitei: Australian J. Chem. **18**, 559 (1965).
6. Brown, R. D., and M. L. Heffernan: Australian J. Chem. **13**, 38 (1960).
7. Brownlie, I. A.: J. chem. Soc. (London) **1950**, 3062.
8. Caldwell, W. J., and H. B. Kime: J. Amer. chem. Soc. **62**, 2365 (1940).
9. Clark, L. B., and I. Tinoco, Jr.: J. Amer. chem. Soc. **87**, 11 (1965).
10. Hélène, C., and P. Douzou: C. R. Acad. Sc. Paris **259**, 3385 (1964).
11. — — C. R. Acad. Sc. Paris **259**, 4853 (1964).
12. — — C. R. Acad. Sc. Paris **259**, 4387 (1964).
13. Hoffman, T. A., and J. Ladik: In: Structure and properties of biomolecules, ed. J. Duchesne, p. 84. New York: Interscience 1964.
14. Katritzy, A. R., and A. J. Waring: Chem. Ind. (London) **1962**, 695; J. chem. Soc. (London) **1963**, 3046.
15. Kwiatkowski, S.: Acta Physica Polonica **24**, 573 (1966).
16. Ladik, J., and K. Appel: Theoret. chim. Acta (Berl.) **4**, 132 (1966).
17. Miles, H.: Biochim. Biophys. Acta **27**, 46 (1958).
18. Nagata, C., A. Imamura, Y. Tagashira, and M. Kodama: Bull. chem. Soc. Japan **38**, 1638 (1965).
19. Nesbet, R. K.: Biopolymer Symp. **1**, 129 (1964).
20. Orloff, M. K., and O. Sinanoğlu: J. chem. Physics **43**, 49 (1965).
21. Pariser, R., and R. G. Parr: J. chem. Physics **21**, 767 (1953).
22. Pople, J. A.: Proc. phys. Soc. (London) A **68**, 81 (1955).
23. Pullman, B.: J. chem. Physics **43**, S 233 (1965).
24. Sharma, B. D., and J. F. McConnell: Acta Cryst. **19**, 797 (1965).
25. Short, L. N., and H. W. Thompson: J. chem. Soc. (London) **1952**, 168.
26. Stewart, R. F., and N. Davidson: J. chem. Physics **39**, 255 (1963).
27. Tanaka, M., and S. Nagakura: Theoret. chim. Acta (Berl.) **6**, 320 (1966).
28. Tanner, E. M.: Spectrochimica Acta **8**, 9 (1956).
29. Thompson, H. M., D. L. Nicholson, and L. N. Short: Disc. Faraday Soc. **9**, 222 (1950).
30. Tsubomura, H. K. Kimura, K. Kaya, J. Tanaka, and S. Nagakura: Bull. chem. Soc. Japan **37**, 418 (1964).
31. Ueda, T., and J. J. Fox: J. Amer. chem. Soc. **85**, 4024 (1963).
32. Wempen, I., and J. J. Fox: J. Amer. chem. Soc. **86**, 2474 (1964).

Professor Dr. S. Nagakura
The Institute for Solid State Physics
The University of Tokyo
Roppongi, Minato-ku
Tokyo, Japan