Tetrahedron, 1967, Vol. 23, pp. 2855 to 2861. Pergamon Press Ltd. Printed in Northern Ireland

THE PROTON DISSOCIATION CONSTANT OF PYRROLE, INDOLE AND RELATED COMPOUNDS

G. YAGIL
The Weizmann Institute of Science, Rehovoth, Israel

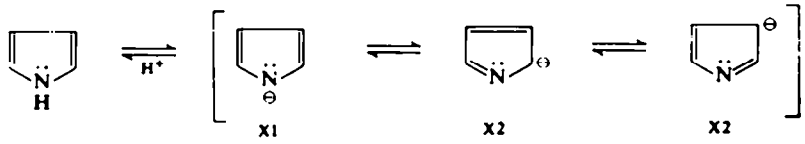
(*Received* 1 *November* 1966)

Abstract—The acid dissociation constants on the H_- and H_- indicator acidity scales of 25 substituted indoles are reported. The pK values range between 12·36 (3-formyl indole) and 19·50 (serotonine). Indole itself has a pK = 16·97. The values obtained for 5 and 2 substituted indoles give a fairly linear correlation with the proper substituent constants, yielding a value of ρ = 1·75 for indole dissociation.

The following additional pK values were found: pyrrole 17·51; pyrazole 14·21; imidazole 14·17; histidine 14·37. pK_a of tryptophan (the indole ring pK) is 16·82.

INTRODUCTION

HETEROCYCLIC compounds which contain an NH group in the ring will dissociate a proton in aqueous hydroxide solutions. This is in contrast to aliphatic amines, which dissociate an hydrogen only in non aqueous media, in which much higher alkalinities are obtainable.[1] The reason for this is that the heterocyclic anion is stabilized by charge distribution over several "resonating" forms, e.g. five for pyrrole:



Few of these dissociation constants have been determined,[2-4] the main reason for this being the severe methodological problems in determining meaningful pK values in concentrated electrolyte solution. One way to overcome these problems (see results) is to report pK values on an indicator acidity scale. Such a scale, based on indole derivatives, has now been constructed for concentrated hydroxides.[5] In this paper, dissociation constants on the H_- and H_- scales for a number of important simple pyrrole and indole derivatives, including such biologically interesting compounds as tryptophan, histidine, tryptamine, are reported. The pK values found, together with the pK values of compounds previously established,[5] are discussed in terms of the Hammett structure-stability ($\rho\sigma$) relationships. This relation is shown to apply to the acid ionization of indoles, extending this relation to the very weak acid range.

[1] cf. K. Bowden, *Chem. Revs* **66**, 119 (1966).
[2] A. Albert, in *Physical Methods in Heterocyclic Chemistry* (Edited by A. R. Katritzky), Vol. 1, p. 1. Acad. Press (1963).
[3] D. D. Perrin, *Dissociation constants of Organic Bases*. Buttersworth, London (1965).
[4] G. B. Barlin and D. D. Perrin, *Quart. Revs.* **20**, 75 (1966).
[5] G. Yagil, *J. Phys. Chem.* in press.

EXPERIMENTAL AND RESULTS

The experimental technique which is employed to determine the dissociation constants is the UV spectrophotometric method. The acid form of most indoles has an absorption peak around 280 mμ (Table 1); this peak is shifted to around 295 mμ upon conversion to base form. An additional absorption peak appears at 310–320 mμ, where the acid form does not absorb; this wavelength is convenient for ionization ratio determination. The single ring compounds like pyrrole and imidazole have no peak above 230 mμ, but a considerable rise in absorbance occurs at 230–250 mμ upon conversion to base, so that ionization ratios can be measured in that region, as has been already done for imidazole by Walba and Isensee.[6,7]

TABLE 1. LIGHT ABSORPTION AND IONIZATION DATA OF PYRROLE AND INDOLE DERIVATIVES T = 25°.

	BH λ_{max}, mμ (ϵ, mM^{-1})	B λ_{max}, mμ (ϵ, mM^{-1})	$C_{1/2}$ M (KOH)	pK
Pyrrole	250[a](0·04)	250[a](0·23)	12·4 ± 0·2	17·51 ± 0·05[d]
Imidazole	235[a](0·03)	235[a](0·50)	1·11 ± 0·05	14·17 ± 0·02[e]
Histidine	235[a](0·17)	235[a](1·67)	1·6 ± 0·1	14·37 ± 0·04
Pyrazole	235[a](0·02)	235[a](0·54)	1·2 ± 0·2	14·21 ± 0·08
Indole	270(5·0)	288·5(4·3)	10·3 ± 0·4	16·97 ± 0·10
Skatole	280(3·6); 287s[b]	350(1·6)	9 ± 1	16·60 ± 0·3
Indoxyl sulfate	277(5·7)	280(3·0);313(2·9)	4·3 ± 0·15	15·23 ± 0·04
5-Methoxyindole 2-carboxylate	290(16·6); 325s	308(15·2)	7·8 ± 0·1	17·03 ± 0·04[c]
Indole 3-carbinol	270(4·6);278(4·6)	290(2·9);315s	8·6 ± 0·1	16·50 ± 0·04
Indole 3-acetate	279(5·6)	285; 312(3·3)	10·0 ± 0·5	16·90 ± 0·15
Gramine	278	—	7 ± 1	16·00 ± 0·3
5-Benzyloxygramine	278	330	10 ± 1	16·90 ± 0·3
Tryptamine	280(5·1)	303	9 ± 0·5	16·60 ± 0·15
Serotonine (5 hydroxy tryptamine)	270(4·7);322(3·4)	310;330	15 ± 1	18·25 ± 0·5[e,f]
5-Hydroxytryptophan	273(5·3);323(4·4)	310; 330	14 ± 1	17·95 ± 0·4[e]
6-Methoxytryptophan	278s; 291(3·3)	298(3·6); 325s	9·3 ± 0·5	16·70 ± 0·15
DL-4-Methyltryptophan	275(5·7)	283; 310	10·0 ± 0·3	16·90 ± 0·10
Salicylate	294(3·57)	250(7·25);302(3·5)	0·26 ± 0·03	13·70 ± 0·05[e,g]

[a] Wavelength employed for measurement, but not λ_{max}.

[b] s-shoulder.

[c] pK is on the H_- scale.

[d] A value of 16·5 is given by McEwen, *J. Am. Chem. Soc.* **58**, 1124 (1936).

[e] Walba and Isensee determined the pK of Imidazole by spectrophotometric measurements at 0·5 M NaCl and found a value of 14·52 (*J. Org. Chem.* **21**, 702 (1956)). The ionization ratios employed in that determination were well below unity.

[f] Another transition(s) at the vicinity of 10, as previously reported, M. M. Rapport, A. A. Green and I. H. Page, *J. Biol. Chem.* **176**, 1243 (1948). See also J. R. Vane, *Brit. J. Pharmacol.* **14**, 87 (1959).

[g] This compound was measured in order to examine to what extent the presently quoted PK values on the H_- scale resemble values obtained by different methods. The careful determination by Z. L. Ernst and J. Menashi leads to pK = 13·59 for salicylate (*Trans. Farad. Soc.* **59**, 230 (1963)).

The hydroxide concentrations which are necessary to convert most compounds measured to base form are relatively high. This presents special problems, because activity coefficients of all species involved, most notably f_{OH^-}, deviate appreciably from unity. Two approaches are available to correct for the contribution of the activity coefficients. The first is to conduct the measurements at several neutral salt concentrations, and to extrapolate to zero ionic strength, but these extrapolations

[6] H. Walba and R. W. Isensee, *J. Org. Chem.* **21**, 702 (1956).

[7] H. Walba and R. W. Isensee, *J. Org. Chem.* **26**, 2789 (1961).

are not always accurate. Moreover, the meaning of ionic strength at these high concentrations is unclear, because the activity coefficients of different electrolytes in concentrated solutions differ widely; the result depends thus on the electrolyte chosen to keep ionic strength constant. The other approach, adopted here, is the indicator acidity approach formulated by Hammett.[8] pK values reported on this scale are thermodynamically valid only if the activity coefficient ratio of acid to base form varies with changing medium in the same way as for the indicators employed for the construction of the scale. Since the H_- scale employed here[5] was constructed with compounds similar to those reported now, the use of the indicator scale seems justified. Indeed the prime reason for examining the compounds reported now was to test their suitability as indicator acids, but they were finally not employed because either enough compounds were available at that hydroxide concentration, or in sufficient solubility or stability precluded their use (the demands of accuracy for pK determination are less stringent).

The full details of the experimental procedure followed and the way ionization ratios $r = \log C_{BH}/C_B-$ are obtained from the experimental data are described in the paper which reports the establishment of the H_- and H_- scales in KOH and NaOH.[5] The pK values reported here are derived from the ionization ratios by the relation[8]

$$H_- = \log r + pK \qquad (1)$$

pK is equal to the value of H_- at the hydroxide concentration $C_{1/2}$ at which $\log r = 0$ (i.e. the acid is half ionized, $C_B = C_{BH}$). To find the pK, log r is plotted vs. $C_{OH}-$, $C_{1/2}$ is read on the plot, and the H_- (or H_-) value corresponding to this $C_{1/2}$ is found in Table 3 of Ref. 5. For acids which already have a negative charge, leading to a double negative charge of the base form, like the indole carboxylates, H_- values are substituted (this is not done when the negative charge is removed from the ring, as in tryptophan).[8] The pK values determined, together with the $C_{1/2}$ values are listed in Table 1. Of course, these pK values are valid only for compounds which behave as an H_- indicator; for those compounds reported here for which an accurate ionization plot was obtained this was actually the case, the slope of log r vs. H_- being within $1{\cdot}0 \pm 0{\cdot}15$.

The first two columns of Table 1 show the position of the absorption maxima λ_{max} at the acid and base form of each compound, and the value of the extinction coefficient ε there. When no value of ε_B- is given, the compound is either not soluble enough to obtain a reliable value, or it is not converted to base form even in the most alkaline solution; a value of ε_B- was assigned to these compounds by extrapolation and the pK value derived is more of an indicative value than of an exact nature. It was felt that even an approximate value is of enough interest with these compounds to have them reported; their approximate nature is indicated by the relative large error of $C_{1/2}$ and pK shown in Table 1. The absorbance measurements of imidazole, pyrazole and histidine had to be conducted on a steep portion of the absorption curve, and the absorbances are sensitive to medium shifts. ε_B- was therefore measured up to 10M KOH, well beyond full conversion to base form, and ε_B- values for each intermediate concentration (0·1–3M KOH) was obtained by linear extrapolation.

DISCUSSION

The heterocyclic compounds discussed here are the weakest acids the dissociation of which, in water, has been studied. The pK of indole, 16·97, is higher by seven units than the pK of phenol. It is therefore of interest to see to what extent are the structure stability relationships, established for the stronger acids such as the phenols or pyridines, still valid. Also, the dissociation in water of an hydrogen from trivalent (in contrast to quaternary) nitrogen has been little studied (Stewart and O'Donnell[9] describe a $\sigma\rho$ relationship for the pK of anilines in water-solvent mixtures).

Several attempts have been made to obtain proper substituent constants for

[8] cf. M. A. Paul and F. A. Long, *Chem. Revs* **57**, 1 (1957); Arnett, *Progress Phys. Org. Chem.* **1**, 223 (1963).

[9] R. Stewart and J. P. O'Donnell, *Canad. J. Chem.* **42**, 1694 (1964).

heterocyclic rings from values for benzene rings,[10–13] but extrapolation to indoles presents some difficulties. It is best to distinguish between the effect of substituents in the 5 position and substituents in the 3 position of the indole ring. The 5 position is essentially *para* to the NH group (except for effects transmitted through the atoms no. 2 and 3). A σ_p^- constant was therefore assigned to the substituents NO_2, CN and CO_2^- which stabilize preferentially the base form by accepting part of the negative charge. For the other substituents: Br, F, MeO and O^-, the use of σ_m seemed more appropriate, because these substituents do not make possible additional resonative forms to the indole or indolate ion (in contrast to benzoic acid). These choices are listed in the last three columns of Table 2, which summarizes the data employed in constructing Fig. 1.

TABLE 2. pK AND σ VALUES OF INDOLE DERIVATIVES EMPLOYED IN FIG. 1.

No.	Substituent	Name	pK		Kind[b]	Source[c]
1	—	Indole	16·97	0		
2	3-CHO	3-Formyl Indole	12·36	1·13	σ_p, an	J
3	3-CO·CH_3	3-Acetyl Indole	12·99	0·89	σ_p, an	J
4	5-NO_2	5-Nitro Indole	14·75	1·26	σ_p^-, an	J
5	5-CN	5-Cyano Indole	15·24	1·00	σ_p^-, an	J
6	5-Br	5-Bromo Indole	16·13	0·39	σ_m	J
7	5-F	5-Fluoro Indole	16·30	0·34	σ_m	J
8	4-F	4-Fluoro Indole	16·30	0·34	σ_m	J
9	3-CH_2·CH·OH, NH_2	L-Tryptophanol	16·91	0·13	$\sigma_I CH_2OH + \sigma_I CH_2NH_2$	C
10	3-CH_2·CH·COO, NH_2	L-Tryptophan	16·82	0·09	$\sigma_I CH_2COO^- + \sigma_I CH_2NH_2$	C
11	4-CH_3 3-CH_2·CH·COO, NH_2	4-Methyl Tryptophan	16·90	0·04	$+\sigma_I$Me	C
12	5-O^- 3-CH_2·CH·COO^-, NH_2	5-Hydroxy Tryptophan	19·20 (17·95)[d]	−0·62	0·09–0·71 ($\sigma_m O^-$)	J
13	3-CH_2·CH_2·NH_2	Tryptamine	16·60	0·08	$\sigma_I CH_2NH_2$	C
14	5-O^- 3-CH_2·CH_2·NH_2	Serotonine	19·50 (18·25)[d]	−0·63	0·08–0·71 ($\sigma_m O^-$)	J
15	3-CH_2·NMe_2	Gramine	16·00	0·08	$\sigma_I CH_2NH_2$	C
16	5-$C_6H_5CH_2O$ 3-CH_2NMe_2	5-Benzyloxy Gramine	16·90	0·03	0·08–0·11 (σ_p^-, ph MeO)	C, B
17	3-CH_2COO^-	Indole 3-Acetic acid	16·90	0·01	σ_I	C
18	3-CH_2OH	Indole 3-Carbinol	16·50	0·05	σ_I	C
19	3-CH_3	Skatole	16·60	−0·05	σ_I	C
20	5-NO_2 2-COO^-	5-NO_2 2-Carboxylate	14·91	1·26–0·03	$\sigma_p^- + \sigma_o$	J, B
21	5-Br 2-COO^-	5-Br 2-Carboxylate	16·10	0·39–0·03	$\sigma_m + \sigma_o$	J, B
22	5-MeO 2-COO	5-MeO 2-Carboxylate	17·03	−0·11–0·03	$\sigma_p^- + \sigma_o$	B, B
23	2-COO^-	Indole 2-Carboxylate	17·13	−0·03	σ_o, pyr	B
24	3-COO	Indole 3-Carboxylate	15·59	0·31	σ_p^-, ph	B
25	5-COO^-	Indole 5-Carboxylate	16·92	0·31	σ_p^-, ph	B

[a] On the H_- scale.
[b] σ_p^-, an — σ_p^- of anilines; σ_p^-, ph — σ_p^- of phenols; σ_o pyr — σ_{ortho} of pyridines.
[c] J—H. H. Jaffe, *Chem. Revs* 53, 191 (1953); C—M. Charton, *J. Org. Chem.* 29, 1222 (1964); B—G. B. Barlin and D. D. Perrin, *Quart. Revs* 20, 75 (1966).

In Fig. 1 the pK values of the various indoles are plotted against the σ constants chosen. It is seen that a straight line can be fairly well drawn through the points for the 5 substituents. The substituent constant for ortho CO_2^- pyridine seems the reasonable choice for the three $2CO_2^-$ derivatives and place these three pK values close to the

[10] Reviewed by H. H. Jaffee and H. Lloyd Jones, *Adv. in heterocyclic Chem.* 3, 209 (1964).
[11] M. J. S. Dewar and P. J. Grisdale, *J. Am. Chem. Soc.* 84, 3539, 3546, 3548 (1962).
[12] M. Charton, *J. Org. Chem.* 30, 3341 (1965).
[13] D. D. Perrin, *J. Chem. Soc.* 5820 (1965).

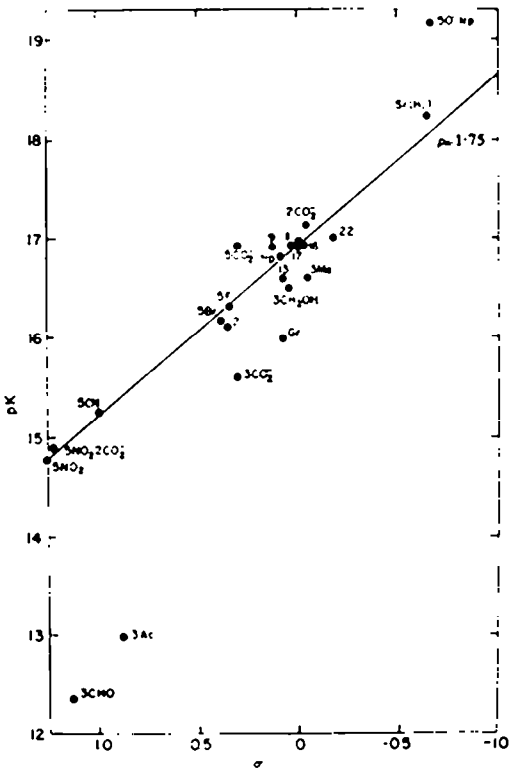


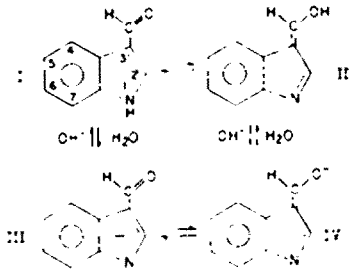
FIG. 1

straight line (Fig. 1). The ρ value found from the slope of the line is $\rho = 1{\cdot}75$. This value is in the range of the values of phenols ($\rho = 2{\cdot}23$)[14] and anilines ($\rho = 2{\cdot}81$).[15] It seems thus that the free energy of indole dissociation is linearly related to that of the stronger acids.

As for position No. 3, this position is not fully aromatic in character. Carbon No. 3 has a negative charge of $-0{\cdot}26$ in the neutral indole molecule.[16] The σ- constants listed by Charton[17] can therefore assigned to most of these derivatives; (others are mostly unavailable; even here some estimates had to be made, as indicated in Table 2). It should be noted that the only small substituent effects are involved for most 3 substituted compounds, so that they cluster near indole and not much can be said about these compounds. The four 3 substituted compounds the pK of which differs considerably from indole are 3-indoxyl sulfate, for which no σ value is available, indole 3-carboxylate, 3-acetyl and 3-formyl indole. To the last three substituents σ_p^- has again been assigned, because they can accept a negative charge (the configuration at this position relative to the nitrogen is different than an ortho substituent in aromatic rings, being trans instead of cis). The pK values of 3-formyl indole and 3-acetyl indole lie two and a half units below their proper place on the line. The spectra of these

[14] A. I. Biggs and R. A. Robinson, *J. Chem. Soc.* 388 (1961).
[15] J. Clark and D. D. Perrin, *Quart. Revs* **18**, 295 (1964).
[16] cf. A. Albert, *Heterocyclic Chemistry* p. 159. Athlone Press London (1959).
[17] M. Charton, *J. Org. Chem.* **29**, 1222 (1964).

two compounds, is however very different from all other indoles (Table 1), even in acid form. It seems likely that the acid form of these compounds is in the "hydroxymethylene" enol form[18] II. The pK is in the proper place if we are dealing here with the ionization of an hydroxy group, activated by an indolynene group. OH^- addition to the CO group is not a likely mode of ionization since the spectra shows that the main difference is in the acid form. There remains 3-CO_2^- indole the pK of which has been well established (it should be also noted that p-CO_2^- lowers also the pK of phenol by 0.56 units[19]). The spectrum of 3-CO_2^- does not suggest an enol form, so that here may be a real indication that a special ρ value should be assigned to the 3 substituted indoles if conventional σ values are used. Considerable differences have also been found in the influence of methyl group in the 1, 2 or 3 position on the acid pK of indoles.[20]



Turning to the compounds other than indoles, let us examine first the effect of adding a benzene ring to the 5-membered heterocyclic ring. Going from pyrrole to indole, the acidity increases by 0·54 units; going from imidazole to benzimidazole (pK = 12·86)[5], and from pyrazole to indazole (pK = 13·80)[5] the acidity increases by 1·31 and 0·41 units respectively. In the case of indazole 0·30 of the rise are due to loss of symmetry (statistical factor) and only 0·11 due to benzene addition. In a qualitative way the rise in acidity is plausible, because the addition of an benzene ring makes possible the contribution of additional structures which assist in the delocalization of the negative charge of the base form, stabilizing this form and lowering the pK. This effect operates differently when a proton is added. Thus the pK is raised by only 0·29 going from pyridine to quinoline, and decreases by 0·19 going to isoquinoline. Going from imidazolium to benzimidazolium the pK is raised from 6·95 to 5·53.[8] It does not seem thus that addition of a benzene ring can be represented by a single substituent constant.

What is the effect of aza substitution? The introduction of an N atom into the 2 position of pyrrole, to give pyrazole, raises the acidity by 3·30 units; introduction into the 3 position to give imidazole, causes a rise of 3·34 units. The same substitutions in indole raises the acidity by 3·17 units (indazole) and 4·11 units (benzimidazole). In each case except indazole, 0·3 of the rise are due to the equivalence of both

[18] A. H. Jackson and A. E. Smith, *J. Chem. Soc.* 5510 (1964). cf. However R. A. Turner, *J. Am. Chem. Soc.* **71**, 3472 (1949).

[19] H. C. Brown, D. H. McDaniel and O. Haffliger, in *Determination of Org. Structures by Physical Methods* (Edited by E. A. Braude and F. L. Nachod), Vol. 1, p. 589. Academic Press, New York (1955).

[20] R. L. Hinman and J. Lang, *J. Am. Chem. Soc.* **86**, 3798 (1964).

nitrogens. To account for the raises in the indole series ($\rho = 1{\cdot}75$) one would have to assign a value of 1·65 for 2-aza substitution and of 2·18 for 3-aza substitution. These figures are well above the values suggested for this kind of substitutions either by Jaffee (Ref. 10, Table 1) or by Perrin.[13] These results, as well as the other results, do not fit the relation $pK = 17{\cdot}0 - 4{\cdot}28 \sum_i \sigma_i$ suggested[4] for heterocyclic N-acids, the main difference being the ρ value found here.

The last topic to be mentioned is the relation between pK_1 and pK_2 of the compounds investigated. Compounds which have two nitrogens in the ring, like imidazole, have a pK_1 corresponding to neutral pH values. This pK concerns proton addition to nitrogen No. 2 which has yet no proton. Bruice and Schmir[21] showed that this pK is linearly correlated with the pK_2 (the pK measured here) for a number of imidazole and benzimidazole derivatives, with a slope of 0·94. A similar correlation by Walba and Isensee yields a slope close to 0·70.[7] The addition of a second proton, involving nitrogen No. 1, will occur only in extremely acid solution, and no information on it is available. Several studies are available on proton addition to the single nitrogen compounds, pyrroles and indoles. Thus the acid $pK(pK_1)$ of pyrrole has been found to be $-3{\cdot}8$,[22-24] leading to a value of $\Delta pK = pK_2 - pK_1 = 21{\cdot}3$. PMR measurements on several methyl substituted[22] pyrroles indicate however, that the proton is added to an α carbon rather than to the NH group. The acid pK of indoles was found[20,25] to range between $+0{\cdot}30$ (1,2-dimethyl indole) and $-7{\cdot}4$ (5-nitro indole). This range indicates relative large ρ values, similar to those of pyridines and quinolines. The difference in ρ prevents the quotation of a single number for the difference in pK_1 an pK_2 of indoles. Some individual numbers, for compounds which were studied in both acid and basic media are: Indole ($\Delta pK = 20{\cdot}5$); 3-methyl indole (21·15), Indole 3 acetic acid (23·05, but not really comparable) and 5 nitro indole (22·15). The difference $pK_2 - pK_1$ lies thus in the range 20–22 units. This figure can however not be quoted to represent the difference between two successive protonations on the same atom, because in acid solutions C-protonation on the negative carbon No. 3 is preferred.[26]

[21] T. C. Bruice and G. L. Schmir, *J. Am. Chem. Soc.* **80**, 148 (1958).
[22] V. Chiang and E. B. Whipple, *J. Am. Chem. Soc.* **85**, 2763 (1963).
[23] see also: N. Naqvi and Q. Fernando, *J. Org. Chem.* **25**, 551 (1960).
[24] R. J. Abraham, E. Bulloch and S. S. Mitra, *Canad. J. Chem.* **37**, 1859 (1959).
[25] see also: G. Berti, A. daSettimo and D. Segnini, *Gazz. Chim. Ital.* **91**, 571 (1961).
[26] R. L. Hinman and E. B. Whipple, *J. Am. Chem. Soc.* **84**, 2534 (1962).