Tetrahedron Vol. 39, No. 17, pp. 2851 to 2861, 1983
Printed in Great Britain.

0040-4020/83 $3.00 + .00

A THEORETICAL STUDY OF THE STRUCTURE, CHARGE DISTRIBUTION AND GAS-PHASE BASICITY OF AZAINDOLES

J. CATALÁN, O. MÓ, P. PÉREZ and M. YÁÑEZ
Departamento de Química Física y Química Cuántica, Centro Coordinado CSIC-UAM, Facultad de Ciencias, C-XIV, Universidad Autónoma de Madrid, Cantoblanco, Madrid, 34, Spain

(*Received UK 30 November* 1982)

Abstract—We have carried out *ab initio* calculations, using a STO-3G minimal basis set, for indazole, benzimidazole, azaindoles and their tautomeric forms. The structural changes which take place in the exchange of the imino proton are fully analized and explained in terms of hybridization effects, valence-shell electron-pairs repulsions and intramolecular non-bonded interactions. These structural changes affect substantially to the endocyclic bond angles which involved both N's. The structure and the charge distribution of the 5-membered ring of azaindoles is very little affected by the position of the aza N. The dicoordinated nitrogen atom is more negative than the tricoordinated one (with the exception of 1H-indazole). As in indoles, C3 of azaindoles is the C atom with the highest electronic density. The prototropic tautomerism in these species is accompanied by a strong migration of charge from the 6 to the 5-membered ring, which results in a considerable increase of the polarity and the basicity of the corresponding tautomeric forms. The predicted PA's are in good agreement with the experimental evidence. C5 is the most reactive C atom of indazole and benzimidazole, although in the former C3 is almost as active as C5. In azaindoles the most basic C atom is C3. Azaindoles are less reactive than indole toward electrophilic agents. Our predicted PA's for N protonation present differences much smaller than those observed in solution. We propose, as a possible explanation, an influence of the tautomeric forms which are more basic in the gas-phase, and more easy to stabilize in aqueous solution, than the normal ones.

In the course of our investigations of the basicity of heterocycles containing fused 5,6-rings, we have reported theoretical studies[1,2] on the electronic structure and gas-phase proton affinities of methylindole derivatives and the N-Mel derivatives of indazole, in an effort to contribute to a better understanding of the experimental behaviour of these compounds.

In the same line of research, indazole, benzimidazole and azaindoles, are a good and convenient model to understand the mutual influence of two non-saturated systems with quite different characteristics. Actually, there are noticeable differences in the basicity, and in general in the reactivity, of this family of compounds, which seem to reveal that the position of the aza-N must have some effects on the electronic structure of these compounds.

On the other hand, all of them and their derivatives, exhibit a quite important biological activity[3]. Some benzimidazole derivatives are basic components[4] of the vitamin B_{12}; a great variety of azaindole derivatives present cardiovascular[5], anti-inflamatory[6], antipyretic and analgesic[7] properties. Some of them seem to suppress the growth of specific tumours cells and other exhibit a quite interesting antidepressant activity.[8,9]

Although, in the literature, there are some all-valence electrons semiempirical studies of some of the compounds considered here,[10–12] the only systematic study carried out on azaindoles has been done using a π-electron semiempirical method[13] that, as we shall discuss later, cannot account for σ-effects which are almost crucial to interpret the behaviour of this kind of molecules.

Moreover, the theoretical treatment of these molecules at the *ab initio* level finds some limitations: the size of these compounds prevents a complete optimization of their geometries, and these have been resolved only for indazole[14] and benzimidazole;[15,16] but in the crystal, where the existence of intermolecular H-bonds produces significant structural deformations which are not likely to occur in the gas-phase. Actually, the few *ab initio* calculations we have found in the literature[17–19] have been performed using "constructed" geometries that, as we have pointed out somewhere else,[2] are not always minimally correct.

Furthermore, these compounds present a prototropic tautomerism which has deserved much attention from the experimental point of view.[20–23] However, the predominance of the normal forms has impeded to know the specific characteristics of the tautomeric forms and their possible influence on the observed properties of azaindoles.

The aim of this paper is to present an *ab initio* study of indazole, benzimidazole and azaindoles (or diazaindenes) and their tautomeric forms, using in all cases, fully optimized INDO geometries, in order to obtain their charge distributions and the intrinsic proton affinity of every ring-position. This should allow us to discuss the influence of the position of the aza-nitrogen on the gas-phase basicity and the reactivity of these compounds. We shall present also a comparative study of the predicted gas-phase basicity and that observed in aqueous solution. An analysis of the possible influence of the tautomers on the behaviour of indazole and the azaindoles is also presented.

COMPUTATIONAL DETAILS

We have carried out the corresponding *ab initio* calculations using a STO-3G minimal basis set. Even at this level, a complete geometry optimization is far beyond our computational capacity. Therefore, we have adopted a geometrical model based in fully optimized INDO geometries and which has been described previously.[1]

The molecular electrostatic potentials evaluated to study the possible approaching paths of a proton to these

molecules in the course of the protonation process, were obtained using the equations developed by Srebrenick *et al*,[24] and which have been implemented by us in the framework of the Gaussian-70 series of programs.

In order to obtain the corresponding charge distributions, we have used the standard density basis sets of the YSP population analysis,[25] mainly because, as we have indicated in previous studies on analogous systems,[1] this partition technique is practically unsensitive to the basis set used to expand the molecular wavefunction and, consequently, small changes should be expected in the charge densities if the basis set is enlarged.

OPTIMIZED GEOMETRIES

The INDO fully optimized geometries of the normal forms and the corresponding tautomers (T) are presented in Tables 1 and 2, respectively. The optimized geometry of 1H-indazole has been already reported by us in a previous publication[2] and it is repeated here to facilitate the discussion which follows on the structural effects of the prototropic tautomerism in these compounds. It should also be pointed out, as an indication, of the goodness of these optimized structures, that the energy obtained for some of the compounds included in the study, is lower than the values reported in the literature,[26] using the same basis set.

As we have indicated in the Introduction, the experimental structure has been resolved only for indazole and benzimidazole, but in the crystal. However, in general, there is a good agreement between our calculated geometrical parameters and the experimental ones. Regarding benzimidazole, the most important discrepancies correspond to the endocyclic bond angles and some of the bond lengths of the 5-membered ring. Nevertheless, it must be taken into account that, in the crystal,[15,16] the molecules are oriented in such a way that each of them is connected to two neighbors via an H bond and, as we shall discuss later, alterations in the coordination of the N atoms have a very important influence on the corresponding endocyclic bond angles. Moreover, it is also possible that, as in the case of indazole,[14] the existence of these intermolecular H-bonds could produce a non-planar structure in the crystal, fact that would affect also to the agreement between theoretical and experimental values. It must be noted that for the structural parameters of the 6-membered ring (much less affected by the H-bonds) this accord is better. It should be noticed that our calculated values confirm the results of Escande *et al.*[16] in the sense that the N1C2 bond length reported by Dik-Edixhoven *et al.*[15] is too short (Table 1).

Although the most dramatic structural changes occur, as we shall discuss later, in going from the normal forms to the tautomers, several aspects of the geometries of these normal forms deserve to be analyzed in some detail.

If one considers indazole and benzimidazole, there is a substantial difference between the bond angles involving the tricoordinated nitrogen atom (C8N1N2 is equal to

Table 1. INDO Fully optimized geometries for indazoles, benzimidazoles, and azaindoles

	INDAZOLE	BENZIMIDAZOLE	4-AZAINDOLE	5-AZAINDOLE	6-AZAINDOLE	7-AZAINDOLE
Bond lengths (Å)[a]						
1-2	1.324	1.378 (1.346)[b]	1.386	1.390	1.385	1.400
2-3	1.333	1.321 (1.311)	1.361	1.359	1.362	1.359
3-9	1.428	1.403 (1.395)	1.438	1.441	1.438	1.440
9-4	1.404	1.397 (1.389)	1.360	1.395	1.401	1.397
4-5	1.376	1.383 (1.386)	1.339	1.339	1.376	1.382
5-6	1.401	1.394 (1.401)	1.393	1.353	1.354	1.391
6-7	1.377	1.383 (1.378)	1.382	1.377	1.339	1.341
7-8	1.401	1.392 (1.401)	1.393	1.397	1.392	1.354
8-1	1.392	1.400 (1.372)	1.399	1.395	1.399	1.395
8-9	1.423	1.421 (1.392)	1.422	1.424	1.424	1.421
1-H	0.975	0.996	0.996	0.996	0.996	1.007
2-H	--	1.089	1.087	1.086	1.086	1.086
3-H	1.087	---	1.084	1.085	1.085	1.085
4-H	1.092	1.091	--	1.096	1.090	1.092
5-H	1.091	1.092	1.095	---	1.095	1.089
6-H	1.092	1.092	1.090	1.095	---	1.095
7-H	1.090	1.090	1.091	1.088	1.093	---
Bond angles (degrees)						
8-1-2	113.6	105.9 (106.6)[b]	108.3	108.6	108.1	107.6
1-2-3	105.1	115.3 (114.0)	110.8	110.1	110.4	110.5
2-3-9	113.1	102.6 (104.2)	106.5	107.3	107.5	107.5
3-9-4	138.9	129.8 (130.0)	130.2	137.6	136.4	138.6
3-9-8	102.8	111.7 (109.5)	107.6	107.0	106.6	106.3
9-4-5	119.5	118.9 (117.8)	115.5	124.3	118.2	119.5
4-5-6	121.0	121.5 (120.9)	125.6	117.3	125.1	119.1
5-6-7	121.9	121.5 (122.3)	119.5	125.4	117.5	125.8
6-7-8	117.0	116.8 (116.1)	116.7	115.4	121.8	112.8
7-8-9	122.4	122.8 (122.4)	120.6	122.2	120.4	127.7
8-1-H	127.5	127.0	125.9	125.9	126.2	126.9
1-2-H	---	119.7	118.3	118.4	118.5	118.4
2-3-H	119.0	---	127.1	127.1	125.8	126.9
9-4-H	120.0	120.7	---	121.3	120.1	120.5
4-5-H	120.2	119.6	114.3	---	121.2	121.1
5-6-H	118.6	119.0	119.6	113.9	---	120.1
6-7-H	121.5	121.7	121.7	122.2	115.7	---

(a) All C–H and NH bond were scaled as indicated in Ref. 1 and therefore they are shorter than the INDO fulloptimized values.

(b) Experimental values taken from Ref. 15.

Table 2. INDO Fully optimized geometries for the tautomeric forms of indazole and azaindoles

	2H-INDAZOLE	4-AZAINDOLE(T)	5-AZAINDOLE(T)	6-AZAINDOLE(T)	7-AZAINDOLE(T)
Bond lengths (Å)[a]					
1-2	1.314	1.340	1.380	1.340	1.384
2-3	1.365	1.410	1.367	1.410	1.369
3-9	1.395	1.386	1.434	1.389	1.434
9-4	1.418	1.382	1.367	1.421	1.375
4-5	1.365	1.368	1.369	1.357	1.405
5-6	1.414	1.366	1.380	1.380	1.368
6-7	1.365	1.407	1.357	1.370	1.367
7-8	1.419	1.373	1.420	1.368	1.383
8-1	1.368	1.394	1.352	1.395	1.348
8-9	1.438	1.457	1.463	1.462	1.453
2-H	1.001	1.094	1.093	1.033	1.093
3-H	1.084	1.083	1.084	1.085	1.084
4-H	1.093	1.001	1.091	1.090	1.093
5-H	1.091	1.090	0.999	1.090	1.088
6-H	1.093	1.089	1.089	0.999	1.090
7-H	1.090	1.093	1.089	1.091	1.042
Bond angles (degrees)					
8-1-2	103.2	101.9	102.7	101.0	99.1
1-2-3	115.3	118.2	116.7	117.0	117.8
2-3-9	106.6	102.0	103.8	104.9	104.8
3-9-4	137.9	134.6	136.8	137.8	140.5
3-9-8	102.7	107.6	104.7	104.1	101.6
9-4-5	119.2	122.0	120.5	120.0	121.5
4-5-6	121.1	120.8	121.9	120.7	119.5
5-6-7	121.9	119.7	120.9	121.9	120.6
6-7-8	118.5	120.8	119.3	120.3	121.8
7-8-9	119.9	118.9	118.8	119.1	118.6
1-2-H	119.0	118.0	117.1	117.9	116.3
2-3-H	119.9	128.1	129.7	126.9	128.7
9-4-H	119.3	119.9	125.7	120.9	120.7
4-5-H	120.5	114.6	119.3	125.5	120.9
5-6-H	118.1	119.4	113.6	118.7	124.4
6-7-H	121.1	118.4	118.9	113.9	118.6

(a) See footnote (a) of Table 1.

113.6° in the first compound, while C8N1C2 is equal to 105.9° in the second one. Similarly the bond angle HN1N2 is equal to 119° in indazole, while the equivalent angle, HN1C2, is equal to 127.1° in benzimidazole). The difference observed in the NH1X2 bond angle is mainly due to non-bonded interactions between the H atom and the σ-lone pair orbital of N2; interaction which takes place in indazole, but which is absent in benzimidazole, explaining why in the former the corresponding bond angle is much smaller than in the latter.

Regarding the C8N1X2 bond angle it has to be taken into account that, in going from indazole to benzimidazole one changes a nitrogen atom by a C–H group, as a consequence the charge density in the N1X2σ-bond diminishes and also the repulsive interaction between this orbital and the one involved in the N1C8 bond, giving rise to a smaller value of this endocyclic bond angle.

In the series of azaindoles, the most important fact is that the geometrical structure of the 5-membered ring is practically unsensitive to the position of the aza-N within the 6-membered ring. As we shall have the opportunity to discuss later, the influence of the position of the aza-N in the 5-membered ring properties is systematically very small.

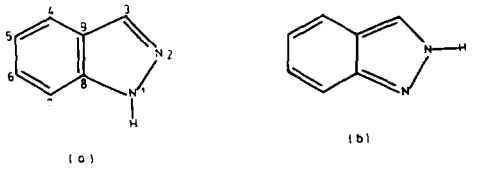
The most important structural changes occur in going from the normal forms to the tautomers. It must be noticed, for instance, that the endocyclic bond angles C8N1N2 closes almost 10°, while N1N2C3 opens almost by the same amount, in going from 1H-indazole to 2H-indazole. Changes are also considerable regarding other geometrical parameters: the bond angle N2C3C9 closes almost 7°, the C8N1 and C3C9 bond lengths decrease 0.023 and 0.033 Å, respectively, while C8C9 and N2C3 bonds lengthen 0.025 and 0.033 Å, respectively.

These geometrical variations can be easily explained using some of the arguments we have presented before in the study of some Me derivatives.[2] We think that for the sake of clarity it is worth to repeat them here.

In going from 1H-indazole to 2H-indazole, we are removing the substituent from N1 and introducing it on N2. Therefore, we are changing from a σ-bonding orbital (that involved in the N1–H bond) to a σ-lone pair orbital. This implies two things: (a) the centre of charge in this orbital is now closer to the nitrogen atom and the repulsion with the other two hybrid orbitals (those involved in the C8N1 and N1N2 bonds) increases; (b) by changing from a σ-bonding orbital to a σ-lone pair orbital, the orbital looses p character and consequently, the other two hybrids gain p character. Both effects contribute simultaneously to decrease the corresponding endocyclic bond angle. For N2, the situation is just the opposite. In this case we change from a σ-lone pair orbital to a σ-bonding orbital. This supposes that: (a) the p character of this orbital increases to improve the overlap with the atomic orbitals of the substituent and, in consequence, the p character of the other two hybrids diminishes and (b) the centre of charge in the new σ-bonding orbital moves away from the N atom, leading to a decrease of the repulsion between this and the other two hybrids. Both effects simultaneously are responsible for the increase of the corresponding endocyclic bond angle (N1N2C3).

The shortening undergone by C8N1 and C9C3 bond lengths and the lengthening found for the C8C9 and

N2C3 bonds are a consequence of the change from a benzenoid-type structure (a) to a quinonoid-one (b).



These changes in the π-electronic structure induced by the tautomeric changeover of the imino proton are evident when one considers the electronic structure of the HOMO (which is a π-orbital) of both species. It is evident that in going from 1H-indazole to 2H-indazole (Fig. 1) there is a migration of π-charge from N2 to N1, C3, C8 and C9, favoring the contribution of the corresponding quinonoid form.

Some systematic structural changes can be observed in azaindoles, as a consequence of the prototropic tautomerism:

(a) The endocyclic bond angles C8N1C2 closes about 6°.

(b) The endocyclic bond angle centered on the N atom of the 6-membered ring opens by about the same amount.

(c) The endocyclic bond angle centered on C2 opens almost 8°.

(d) The endocyclic bond angle centered on C3 closes about 4°.

(e) The N1C2 and N1C8 bonds shorten while the C8C9 bond substantially lengthens.

Changes (a) and (b) are a direct result of going from a σ-bonding to a σ-lone pair orbital and *vice versa*, in a mechanism similar to the one outlined above for indazole. The other structural changes are mainly a consequence of significant variations in the π-electronic distribution of the HOMO's. It is evident (Fig. 2) that in the tautomers there is a certain migration of charge from the 6- to the 5-membered ring.

This charge is accumulated essentially on C3 and N1. As a consequence, the bonding character between N1 and C8 increases, and the corresponding bond length shortens. The decrease in the contribution of the orbitals centered on C2, causes a diminution in the antibonding character between N1 and C2, leading to a shortening of this bond. Finally, in going from the normal forms to the tautomers, the character between C8 and C9 changes from bonding to antibonding, explaining the noticeable increase of the corresponding bond length.

Change (c) obeys to two main effects: (i) the antibonding character between N1 and C3 increases in the tautomer relative to the normal form; (ii) the decrease in the π density on C2 implies an increase in the density of the σ-bonding orbitals and therefore, an increase also in their mutual repulsion. The situation is practically the opposite regarding change (d).

CHARGE DISTRIBUTIONS

We present in Table 3 the YSP charge distribution of those compounds included in this study. The first column, in each case, corresponds to the net charges of the normal form and the second one to those of the tautomer.

Our results indicate that the dicoordinated N atom (which behaves as a π-acceptor) has a greater negative charge than the tricoordinated one (which behaves always as a π-donor) except for indazole, where the tricoordinated N atom is the one which bears the highest negative charge, in agreement with other theoretical calculations.[18] This result is not surprising since, in this particular case, the dicoordinated nitrogen atom occupies position 2 that, as we have shown for indole and methylindoles[1] and it is the case also for azaindoles (Table 3), is always the most positively charged.

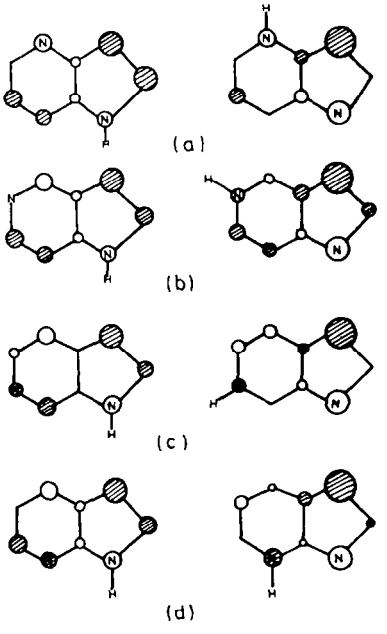


Fig. 2. The highest occupied molecular orbitals of: (a) 4-azaindole, (b) 5-azaindole, (c) 6-azaindole, (d) 7-azaindole and their corresponding tautomeric forms.

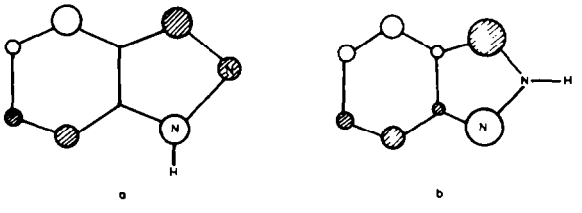


Fig. 1. The highest occupied molecular orbitals of: (a) 1H-indazole, (b) 2H-indazole.

Table 3. YSP Charge densities (in me) for indazole, benzimidazole, azaindoles and their tautomeric forms

Centre	INDAZOLE		BENZIMIDAZOLE	4-AZAINDOLE		5-AZAINDOLE		6-AZAINDOLE		7-AZAINDOLE	
1	-429	-462	-576	-562	-635	-567	-639	-559	-623	-560	-674
2	-392	-335	+408	+239	+263	+234	+228	+242	+247	+231	+230
3	+184	+179	-599	0	- 83	+21.5	- 6	+ 16	- 26	+ 24	- 20
4	+ 98	+ 86	+ 78	-555	-505	+259	+290	+ 53	+ 53	+108	+142
5	+ 58	+ 69	+ 60	+231	+274	-583	-516	+224	+239	+ 24	+ 12
6	+ 87	+ 69	+ 84	+ 50	+ 17	+247	+238	-561	-517	+250	+276
7	+ 50	+ 71	+ 49	+ 70	+122	+ 16	+ 40	+221	+278	-594	-515
8	+228	+ 40	+217	+204	+179	+249	+244	+200	+156	+378	+395
9	+ 39	+196	+185	+216	+244	+ 29	+ 14	+ 78	+ 90	+ 29	+ 34
H(1)	+170	-	+185	+173	--	+173	--	+175	--	+189	--
H(2)	-	+175	- 12	- 10	- 36	- 7	- 32	- 7	- 35	- 5	- 34
H(3)	- 10	0	--	+ 7	- 2	- 2	- 10	- 3	- 12	- 2	- 9
H(4)	- 21	- 25	- 11	--	+169	- 32	- 10	- 14	- 8	- 20	- 15
H(5)	- 23	- 25	- 25	- 32	- 8	--	+169	- 30	- 12	- 11	- 5
H(6)	- 25	- 27	- 26	- 15	- 5	- 32	- 12	--	+169	- 30	- 8
H(7)	- 14	- 12	- 16	- 15	- 11	- 6	+ 4	- 25	0	--	+182

It should be emphasised, however, that the difference between the charges of the two N's is always quite small, in clear contrast with previous theoretical calculations[13] based on π-electron methods, which have predicted differences of the order of 0.4 e^-.

Unfortunately, very little experimental information related to the charge distribution of these compounds has been reported and most of the NMR studies have been carried out on Me derivatives. Nevertheless, for the particular case of indazole, 2H-indazole and 7-azaindole, Black *et al.*,[20] have made an attempt to estimate the corresponding electron distribution from the proton chemical shifts. Our calculated charge densities agree only slightly with those obtained in this experimental estimation, while they follow a trend quite similar to the one obtained by the "variable electronegative SCF" method.[19]

C3, in indazole, bears the highest positive charge, being C5 and C7, in both indazole and benzimidazole, the positions with the highest electron density. These results are in contrast with those obtained using π-electron semiempirical methods which indicate,[13] for instance, that C6 is the most electronegative C atom of benzimidazole.

In the case of azaindoles, C3 has the greatest electron density, similarly to what is found in indoles,[1] and in agreement with the experimental reactivity of these compounds. C2, as in indoles, is the position which presents the greatest positive charge and therefore the more acidic site. Once more these results are in contradiction with those obtained[13] using π-methods, making evident that the inclusion of σ-effects is essential to obtain a reliable charge distribution of these compounds.

It is also interesting to realize that, in the azaindoles family, the charge distribution of the 5-membered ring is very little affected by the position of the aza-N. This fact seems to be corroborated by the experimental behaviour of these compounds in electrophilic substitutions, as we shall discuss in the next section.

If one considers the tautomeric forms, changes in the charge distribution are important in azaindoles, but not too much significant in indazole. 2H-indazole, disregarding those changes affecting N1 and N2, presents very small variations in the charge density of the different centres relative to 1H-indazole.

In azaindoles, however, one can observe that there is a strong migration of charge from the 6- to the 5-membered ring as a consequence of the prototropic tautomerism. This charge is accumulated on N1 and C3, indicating that there is a considerable contribution of the mesomeric forms **1** to **6** (Fig. 3).

It is interesting to note, however, that this migration of charge is not uniform along the series. For 4-azaindoles the charge transferred is 220 me; 331 me for 5-azaindole, 300 for 6-azaindole and 378 me for 7-azaindole. This trend can be understood if one admits that, in some sense, azaindoles could be considered as a particular case of amino-pyridines. Therefore, the migration of charge is favored when the "substituent" is *para* (5-azaindole) and *orto* (7-azaindole) relative to the pyridine N and it is less favored when the "substituent" is *meta* (4- and 6-azaindole). The charge transfer is lower in 4-than in 6-azaindole because, in the former, one of the neighbors of the pyridine N is a C atom (C9) bonded to other two C's and therefore more difficult to polarize than those C's bonded to C and to a H-atom, what is the case for the neighbors of the pyridine N in 6-azaindole.

This migration of charge is confirmed by the strong increment undergone by the polarity of the tautomer relative to the normal form, as indicated by the noticeable increment of the corresponding dipole moments (Table 4). As we shall discuss in the last section this increased polarity of the tautomeric forms could help to explain the basicity in solution exhibited by these compounds.

INTRINSIC PROTON AFFINITIES

The intrinsic gas-phase PA's of each postion was obtained using a very economical procedure[27,28] based on the linear correlation between experimental gas-phase PA's and *ab initio* 1s orbital energies.

The PA's of tricoordinated nitrogens were obtained using eqn (4) of Ref. 28; those of dicoordinated N atoms using eqn (1) of Ref. 29 and the PA's of the ring C's were obtained using eqn (1) of Ref. 28. These calculated PA's are summarized in Table 5.

The first important fact is that all benzazoles included in this study are N bases, in good agreement with the conclusions of the mass analyzed ion kinetic energy spectrometry study of Maquestiau *et al.*[30] and other experimental evidence.[31]

Our results indicate also that benzimidazole is a much stronger base, in the gas-phase, than indazole, in agreement with their basicity in aqueous solution.[32,33]

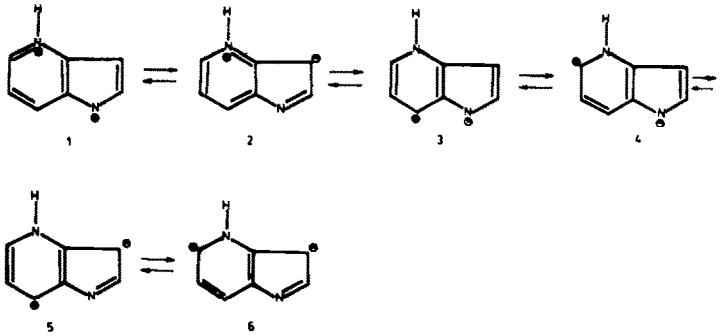


Fig. 3. Mesomeric forms which contribute to the structure of the tautomers of azaindoles. In the figure are presented for the 4-azaindole.

Table 4. Calculated dipole moment (in Debye) for benzazoles and their tautomeric forms

	Normal form (μ_N)	Tautomer (μ_t)	$\mu_N - \mu_t$
Indazole	1.74	2.30	2.26
Benzimidazole	3.19		--
4-azaindole	3.67	5.48	16.6
5-azaindole	3.97	6.36	24.7
6-azaindole	3.30	5.49	19.1
7-azaindole	1.47	4.29	16.1

Table 5. Predicted proton affinities (kcal/mol) of indazole, benzimidazole and azaindoles

Compound	POSITION								
	1	2	3	4	5	6	7	8	9
Indazole	174.9	202.2	172.4	173.7	184.2	175.5	174.7	139.4	181.3
Benzimidazole	182.4	112.5	228.7	183.2	191.5	182.7	180.2	144.6	168.6
4-Azaindole	185.3	144.7	203.3	224.2	163.6	180.1	169.2	145.6	163.6
5-Azaindole	184.6	142.1	196.0	155.6	227.6	158.6	178.0	137.5	183.4
6-Azaindole	184.0	138.8	198.3	179.9	165.9	224.7	154.6	145.3	178.3
7-Azaindole	187.2	145.9	200.6	169.7	186.9	155.6	223.6	125.4	183.7

The sequence predicted in the gas-phase basicity of azaindoles is also coincident with that observe in solution, in the sense that the most basic compound is 5-azaindole (PA = 227.6 kcal/mol), followed by 6- (PA = 224.7 kcal/mol), 4-(PA = 224.2 kcal/mol) and 7-azaindole (PA = 223.6 kcal/mol).

It must be noticed, however, that the differences between gas-phase PA's are very small in contrast with the very different values of the pK's[34] (8.26, 7.95, 6.94 and 4.59, respectively), measured in aqueous solution at 25°. Besides, according to our results, benzimidazole is the strongest base of the whole series, in the gas-phase. This result is not very surprising since position 3 is the most basic position of this kind of compounds[1] but, in aqueous solution benzimidazole behaves as a much weaker base than azaindoles, with the only exception of 7-azaindole. We believe that these discrepancies between our predicted gas-phase basicities and those measured in solution deserve a detailed analysis that we shall carry out at the end of this section.

Regarding the other centres, it is clear that in both, indazole and benzimidazole, the most basic carbon atom is C5. This result is in agreement with experimental evidence, since the reaction of indazole with fuming nitric acid yields 5-NO_2-indazole.[35] Unsubstituted benzimidazoles are attacked by electrophiles preferentially at the 5-position and this compound is nitrated exclusively in the 5-position and on further nitration, a mixture of 5,6-dinitro and 4,6-dinitro derivatives is obtained,[36,37] again in good agreement with our predictions (Table 5), since after C5 the most active positions of benzimidazole toward a proton attack, and in general toward electrophilic reagents, are C4 and C6.

There is also a clear difference between the reactive characteristics of the 5-membered ring of indazole and benzimidazole. C2 in the latter is completely deactivated, while C3 in the former is almost as basic as C5. These results are corroborated by the experience, since no electrophilic substitutions occur on the imidazole ring of benzimidazole,[38] but some electrophilic substitutions of indazole, as halogenation,[35] yield a mixture of 5- and 3-substituted derivatives. Moreover, although no nucleo-

philic substitutions have been observed for unsubstituted benzimidazole,[3c] it is well established that 1-methylbenzimidazole gives the 2-amino derivative when heated with sodium amide.[39]

In the case of azaindoles the most basic C atom is, as in indoles,[1] C3. There are some experimental studies[40] that indicate, for instance, that 5- and 7-azaindole undergo nitration and bromination in C3. Besides, El-Anani *et al.*[41] in a acid catalysed H exchange of azaindoles conclude that 4-, 5- and 7-azaindole exchange the 3-proton, and the observed reactity is also in agreement with our theoretical findings, in the sense that the most reactive species is 4-azaindole, followed by 7- and 5-azaindole (Table 5). The gas-phase PA predicted for this position of azaindoles in systematically lower than that obtained for the same reactive position of indole (210.0 kcal/mol).[1] This result is quite interesting, since due to the fact that the π density on C3 in azaindoles is quite similar to that of indole[42] it has been postulated[42] that the electrophilic substitution of azaindoles must be preceeded by the protonation of the aza N. This protonation results in a decreasing of the π-density at C3 and would explain the lowered activity of azaindoles at the 3-position relative to indole. According to our results C3 is naturally less active toward electrophilic agents in azaindoles than in indole[1] and therefore it is not strictly necessary to assume that a protonation of the system has to take place previous to the electrophilic substitution.

To study the most favored approaching paths of a proton in the protonation process of these compounds, we have evaluated the corresponding molecular electrostatic potentials in a plane parallel to that of the molecule and 1.6 Å above it, since from similar calculations one can expect the local minimum to be found at this distance from the basic centre.

In Fig. 4 we present only the most characteristic cases, which confirm our previous discussion.

In indazole and benzimidazole (Figs. 4a and b, respectively), there is a deep minimum over the dicoordinated N atom, which extends over the six-membered ring, where a shallower local minimum is located on C5. It is also evident that the area over the imidazole ring of benzimidazole is, with the exception of the region around N3, almost repulsive.

In the family of azaindoles the potential maps are very much alike and we shall present only those of 4- and 7-azaindole. The deeper minimum is located on the pyridine-type N and this attractive area embraces the 5-membered ring. In the case of 7-, 6- and 5-azaindole, another local minimum is found on C3.

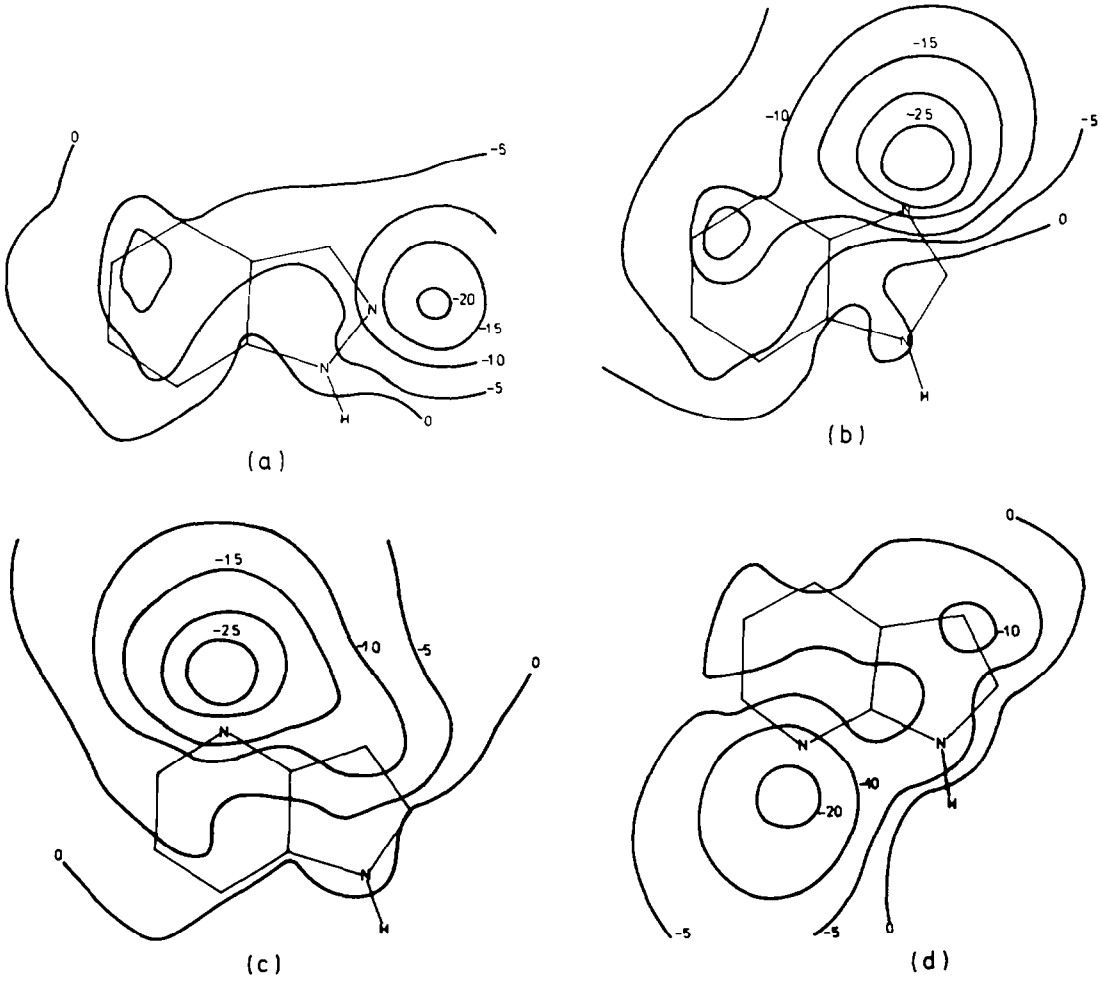


Fig. 4. Electrostatic potential maps: (a) 1H-azaindole, (b) benzimidazole, (c) 4-azaindole, (d) 7-azaindole evaluated in a plane parallel to the plane of the molecule and 1.6 Å above it. All values in kcal/mol.

Let us analyze now the quantitative disagreement between the observed basicity in aqueous solution of unsubstituted benzazoles and our predicted gas-phase basicities. We believe that there are two main possible factors: (a) specific interactions molecule-solvent, which could be different for each compound; (b) the influence of the tautomeric forms on the basicity of these species in solution.

The analysis of point (a) is far beyond the scope of this study and it needs a systematic experimental evaluation of the thermodynamical magnitudes involved in the protonation process. It must be determined whether there is a deep change in the enthalpy of the protonation process, due to some specific solvation effects which should change with the position of the basic centre (in this particular case, the aza-N) and whether it can be expected some significant variation in the entropy term of such a process. However, at least for azaindoles, these entropy changes do not seem quite probable, since it has been shown[43] that, for the particular case of pyridines, there exists a very good linear correlation between the ΔG° for the protonation process in the gas-phase and in aqueous solution.

This kind of study is feasable if one uses the N1-Me derivatives that are, in most cases, stable and eliminate possible interferences from the tautomeric forms.

Actually, it would be not surprising that solvation effects could be different for indazole and benzimidazole, because it is known the existence of a different interaction with solvent, in aqueous media, of imidazole and pyrazole.[44] Obviously, for analogous reason, it is not possible to disregard the possibility of different solvation effects for those two compounds relative to azaindoles. What is clear, according to our results, is that in the case of benzazoles there is no direct correlation between the basicity in solution and in the gas-phase. To determine if solvation effects are, at least one of the factors explaining this absence of correlation is, indeed, a very interesting experimental problem. However, within the series of azaindoles it seems not very likely that solvation effects alone could explain how differences of 0.6 kcal/mol in the gas-phase PA's (7-azaindole vs 4-azaindole) result in a pK difference of 2.4, while variations of 0.5 kcal/mol (6-azaindole vs 4-azaindole) results in a pK difference of 1.0.

Regarding point (b) it is a well established fact[22,23] that, in the case of unsubstituted benzazoles, the normal form is clearly predominant, but it seems also clear that the imino proton certainly exchanges to some extent.[12,20] On the other hand, the N-Me derivatives of some of them (indazole and 7-azaindole) have been isolated under both forms[45,46] and it has been found[45,32] that the tautomeric forms present a higher basicity than the normal ones. For this reason we have considered the possibility that the existence of a prototropic tautomerism in solution could be an important factor to interpret the basicity of benzazoles.

To clarify if such a mechanism has to be taken into account it is necessary to answer two fundamental questions: What is the basicity of the tautomeric forms of the unsubstituted benzazoles, and what is the degree of participation of these forms when we measure the basicity in solution of these compounds.

To answer the first question we have obtained, following the same procedure outlined above, the intrinsic PA of each position of the different tautomers. The corresponding values are presented in Table 6.

As before, the most basic site is the dicoordinated N atom and the most basic C atoms are still C5 and C6 in 2H-indazole and C3 in the azaindole tautomers. However, the intrinsic basicity of these two centres is very much higher than in the normal forms. Such a result corroborates our previous conclusion that the prototropic tautomerism in azaindoles is accompanied by a strong migration of charge from the 6- to the 5-membered ring. The localization of this charge on C3 and N1, through an important contribution of the mesomeric forms of Fig. 3, explains the increased basicity of these two centres in the tautomer.

In Fig. 5 we present, as typical examples, the electrostatic potential maps of 2H-indazole and 5H-pyrrolo-(3,2-c) pyridine.

By comparing Figs. 4(a) and 5(a) it is evident that in 2H-indazole the minimum (now located on N1) is a little deeper than in 1H-indazole, being the rest of the map practically unchanged. However, by comparing Figs. 4(d) and 5(b) one can see that, in both cases, there are two minima (over the dicoordinated nitrogen atom and over C3) but much deeper in the case of 7H-pyrrolo-(2,3-b) pyridine than in 5-azaindole. Also, the area over the 6-membered ring becomes repulsive in the former, as a consequence of the transfer of charge which takes place in the corresponding tautomeric changeover.

Therefore one can, presumably, conclude that if some tautomeric form is present in solution, the basicity

Table 6. Predicted Proton Affinities (kcal/mol) of 2H-indazole and the tautomeric forms of azaindoles

Compound	POSITION								
2H-Indazole	215.0	160.9	150.6	183.4	188.3	190.4	186.4	172.4	186.9
4-Azaindole (T)	249.9	185.1	233.5	167.6	102.6	170.9	152.0	172.9	141.0
5-Azaindole (T)	251.6	196.9	231.4	101.2	166.7	122.3	173.4	170.7	191.6
6-Azaindole (T)	248.1	190.5	236.4	171.1	120.9	167.8	104.9	178.6	188.3
7-Azaindole (T)	249.8	190.9	234.6	147.3	169.8	101.3	167.9	124.1	185.9

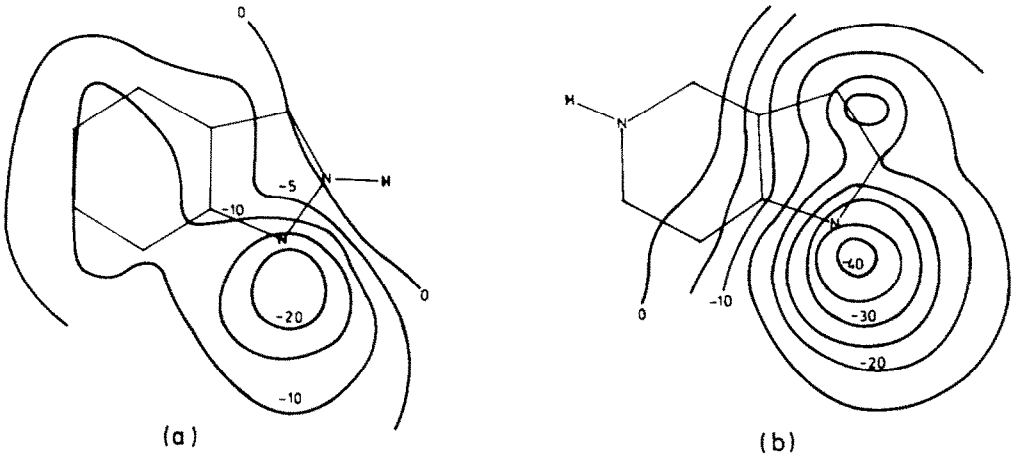


Fig. 5. Electrostatic potential maps for : (a) 2H-indazole, (b) 5H-pyrrolo (3,2-c)-pyridine. All values in kcal/mol.

measured must be higher than that expected if only the normal form was present.

To answer the second question it must be indicated that, in agreement with the experimental evidence, our *ab initio* results indicate that the normal form is always more stable than the tautomeric one (11.7 kcal/mol in the case of indazole and of the order of 23–26 kcal/mol in the case of azaindoles). However, as we have pointed out in the previous section, the dipole moment of the tautomeric form is much greater than that of the normal form. Therefore, in aqueous solution the possible interaction with the solvent should stabilize more the tautomeric than the normal form. The stablization energy increases with the square of the dipole moment; for this reason we have reported in the last column of Table 4 the difference between the square of the dipole moment of the tautomeric and the normal form for each case.

These values suggest that the influence of the tautomeric forms must be maximum in the case of 5- and 6-azaindole and much smaller for 4- and 7-azaindole in clear agreement with the fact that precisely 5- and 6-azaindole are the compounds which exhibit the greater basicity in solution.

This reasoning could also contribute to explain why, in spite that benzimidazole is the strongest base in the gas-phase, it is less basic than 4-, 5- and 6-azaindole in solution. As we have indicated, the exchange (favored in polar media) of the imino proton of azaindoles makes to appear forms which are very basic. Consequently, the measured basicity must be higher than it should be expected from the normal forms alone. However, in the particular case of benzimidazole, if such a prototropic transfer takes place, no change could be observed in the measured basicity, due to the symmetry of this molecule.

CONCLUSIONS

From the results discussed in this paper we can conclude that the most important differences in the structure of the 5-membered ring in indazole and benzimidazole, and the substantial changes which take place in the geometry of benzazoles due to prototropic tautomerism can be easily explained in terms of hybridization effects, valence-shell electron-pair repulsions,[47] and intramolecular non-bonded interactions. The most important geometrical changes take place in the bond lengths and angles related to the two nitrogen atoms. On the other hand, the structure of the 5-membered ring in azaindoles is almost unaffected by the position of the aza-N.

In all cases studied the dicoordinated N atom (π-acceptor) is more negatively charged than the tri-coordinated one (π-donor), except in the case of 1H-indazole, where the aza-N is in position 2, which is very acidic in this type of compounds. The charge distributions calculated using *ab initio* methods do not agree with those obtained with π-methods, which indicates that the inclusion of σ-effects is crucial to describe the charge distribution in these compounds.

The C3 in azaindoles is the atom with the highest electronic density, as in indoles. The charge distribution in the five-membered rig is almost independent on the position of the aza-N.

The prototropic tautomerism in these species is accompanied by strong migration of charge from the 6- to the 5-membered ring. This change accumulates on C3 and N1, resulting in a noticeable increase of the basicity of these two centers and of corresponding dipole moment.

In agreement with experimental evidence, our results indicate that these compounds are N bases, being benzimidazole the most basic of the series and C5 is the most reactive position in both indazole and benzimidazole toward electrophilic reagents. No electrophilic substitution takes place on the 5-membered ring of benzimidazole, while C3 in indazole is almost as basic as C5. The most basic C in azaindoles is C3, just as it is in indoles. Our calculated PA's for C3 in azaindoles indicate that the reactivity of these compounds is lowered relative to indole, in agreement with experiment. Therefore it is unnecessary to postulate a protonation step previous to electrophilic substitution.

Our predicted PA's for the protonation on N in azaindole follow the corresponding PK's in solution, but the differences in basicity between these compounds in the gas phase are much smaller than those observed in solution. This could be explained in two different ways (1) a specific solvation effect yielding a very different value of ΔG for each nitrogen position. (2) the presence of tautomeric forms, which would interact more strongly with the solvent than the normal ones. The

relative amount of each tautomer would also depend on the nitrogen position varying like the polarity of these compounds. We think that the second effect might be the more important in azaindoles. This effect would also explain the low basicity of benzimidazole in solution.

Acknowledgement—All calculations have been performed on the IBM 370/75 Computer at the UAM-IBM Center, Madrid.

REFERENCES

[1]J. Catalán, P. Pérez and M. Yáñez, *Tetrahedron* **38**, 3693 (1982).
[2]J. Catalán, O. Mó, P. Pérez and M. Yáñez, *J. Mol. Struct.* in press.
[3a]R. E. Willette, *Adv. Heterocyclic. Chem.* (Edited by A. R. Katritzky and A. J. Boulton), Vol. 9. Academic Press, New York (1968); [b]L. N. Yakhontov and A. A. Prokopov, *Russian Chem. Rev.* **49**, 428 (1980); [c]P. N. Preston, *Hetero.* (Edited by P. N. Preston), Vol. 40, Part I. Wiley (1981).
[4]R. Bonnett, *Chem. Ref.* **63**, 573 (1963).
[5]US. Pat. 3 363956 (1968). *Chem. Abst.* **69**, 10467 (1969).
[6]E. Bisagni, J. D. Bourzat, C. Labrid, P. Delfort and A. Le-Ridant, *Chim. Theraphic.* **8**, 559 (1973).
[7]Merk and Co. *Netherland Pat.*, 6, 510648, *Chem. Abst.* **65**, 13711 (1966).
[8]US. Pat, 3, 777246 (1968). *Chem. Abst.* **69**, 117131 (1968).
[9]M. R. Bell, J. O. Hoppe, H. E. Lape, D. Wood, A. Arnold and W. H. Sebberis, *Experientia* **23**, 298 (1967).
[10]M. J. S. Dewar and N. Trinajstic, *J. Am. Chem. Soc.* **91**, 6321 (1969).
[11]A. Escande, J. Lapasset, R. Raure, E. J. Vincent and J. Elguero, *Tetrahedron* **30**. 2903 (1974).
[12]P. Mauret, J. P. Fayet and M. Fabre, *Bull. Soc. Chim. Fr.* 1675 (1975).
[13]M. Cireanu, *Rev. Rum. Chim.* **21**, 1007 (1976).
[14]A. Escande and J. Lapasset, *Acta Cryst.* **B30**, 2009 (1974).
[15]C. J. Dik-Edixhoven, H. Schenk and H. van der Meer, *Cryst. Struct. Commun.* **2**, 23 (1973).
[16]A. Escande and J. L. Galigné, *Acta Cryst.* **B30**, 1647 (1974).
[17]M. H. Palmer and S. M. F. Kennedy, *J. Chem. Soc.* Perkin Trans. **2**, 81 (1976).
[18]J. Koller, A. Azman and N. Trinajstic, *Experientia* **32**, 205 (1976).
[19]M. H. Palmer and S. M. F. Kennedy, *J. Molec. Struct.* **43**, 33 (1978). and **43** 203 (1978).
[20]P. J. Black, R. D. Brown and M. L. Heffernan, *Aust. J. Chem.* **20**, 1325 (1967).
[21]L. Stefaniak, *Org. Magn. Res.* **11**, 385 (1978).
[22]I. I. Schuster, C. Dyllick-Brenzinger and J. D. Roberts, *J. Org. Chem.* **44**, 1765 (1979).
[23]A. Fruchier, V. Pelligrin, R. Schimpf and J. Elguero, *Org. Magn. Res.* **18**, 10 (1982).
[24]S. Srebrenick, H. Weinstein and R. Pauncz, *Chem. Phys. Lett.* **20**, 410 (1973).
[25]M. Yáñez, R. F. Stewart and J. A. Pople, *Acta Cryst.* **A34**, 641 (1978); F. Escudero and M. Yáñez, *Molec. Phys.* **45**, 617 (1982).
[26]A. L. Hinde, L. Radom and M. Rasmussen, *Austral. J. Chem.* **32**, 11 (1979).
[27]J. Catalán and M. Yáñez, *J. Chem. Soc.* Perkin Trans **2**, 741 (1979).
[28]J. Catalán and M. Yáñez, *J. Chem. Soc.* Perkin Trans **2**, 1627 (1979).
[29]J. Catalán, O. Mó, P. Pérez and M. Yáñez, *J. Am. Chem. Soc.* **101**, 6520 (1979).
[30]A. Maquestiau, Y. V. Haverbeke, H. Mispreuve, R. Flammang, J. A. Harris, I. Howe and J. H. Beynon, *Org. Mass. Spec.* **15**, 144 (1980).
[31]G. G. Dvoryantzeva, T. N. Ulyanova, N. P. Kostyuchenko, Yu. N. Sheinkar, E. I. Lapan and L. N. Yakhontov, *Tetrahedron Letters*, 2857 (1972).
[32]J. Elguero, A. Fruchier and R. Jacquier, *Bull. Soc. Chim. Fr.* 2619 (1967).
[33]P. Tomasik, R. Zalewski and J. Chodzinski, *Chem. Zvesti* **33**, 105 (1979).
[34]T. K. Adler and A. Albert, *J. Chem. Soc.* 1794 (1960).
[35]W. Von Auwers and J. Kleiner, *J. Phakt. Chem.* **118**, 67 (1928).
[36]G. E. Ficker and D. J. Fry, *J. Chem. Soc.*, 736 (1963).
[37]K. L. Kirk and L. A. Cohen, *J. Org. Chem.* **34**, 384 (1969).
[38]D. Harrison, J. T. Ralph and A. C. B. Smith, *J. Chem. Soc.* 2930 (1963).
[39]A. M. Simonov and N. D. Vitkevich, *Zh. Obshch. Khim.* **30**, 590 (1960); *Chem. Abst.* **84**, 10563d.
[40]L. N. Yakhontov and E. I. Lapan, *Khim. Geterotsikl Soedin.* 1528 (1972); *Chem. Abst.* **78**, 43322, M. M. Robinson, B. L. Robinson and F. P. Butler, *J. Am. Chem. Soc.* **81**, 743 (1959); M. M. Robison and B. L. Robison, *Ibid* **78**, 1247 (1956).
[41]A. El-Anani, S. Clementi, A. R. Katritzky and L. Yakhontov, *J. Chem. Soc.* Perkin Trans. **2**, 1072 (1973).
[42]D. A. Bochvar, A. A. Bogatur'yants, A. V. Tutkevich, L. N. Yakhontov, M. Ya. Uristkya, D. M. Krasnokutskaya and M. V. Rubstov, *Bull. Acad. Sci. USSR, Dir. Chem. Sci.* (Engl. Transl) 327 (1966).
[43]M. Taagepera, K. D. Summerhays, W. J. Hehre, R. D. Topsom, A. Pross, L. Radom and R. W. Taft, *J. Org. Chem.* **46**, 891 (1981).
[44]J. N. Spencer, E. S. Holmboe, M. R. Kirshenbaum, S. W. Barton, K. A. Smith, W. S. Wolbachs, J. F. Powell and C. Chorazy, *Can. J. Chem.* **60**, 1183 (1982); M. J. Blais, O. Enea and G. Berthon, *Thermochim. Acta* **20**, 335 (1977).
[45]M. M. Robison and B. L. Robison, *J. Am. Chem. Soc.* **77**, 6554 (1955).
[46]K. V. Amvers and M. Durberg, *Ber. Dtsch. Chem. Ges*, **53**, 1179 (1920).
[47]A. Domenicano, A. Vaciago and C. A. Coulson, *Acta Cryst.* **B31**, 221 (1975).