An Analysis of the Lipophilicity of Furazan and Furoxan Derivatives using the CLOGP Algorithm

Rosella Calvino,[a] Alberto Gasco[a,*] and Albert Leo[b,*]
[a] Dipartimento di Scienza e Tecnologia del Farmaco, Via P. Giuria, 9, 10125 Torino, Italy
[b] Department of Chemistry, Pomona College, Claremont, California, 91711, USA

The partition coefficients of a series of furazan and furoxan derivatives have been measured and analysed by means of the CLOGP algorithm. The f_{non} and $f_{non(O)}$ fragmental constants have been evaluated as well as the pairs of σ/ρ constants, which are able to describe the electronic interactions operating in the substituted systems. Analysis of the alkyl derivatives shows that the 'benzyl carbon' requires a large and negative correction factor when linked to either of these two heterocyclic rings rather than to benzene. In contrast, in 1,2,5-thia and 1,2,5-selenadiazole this same carbon behaves similarly to that attached to a phenyl ring.

The CLOGP algorithm is, at the present time, the most widely-used method for calculating octanol–water log P.[1,2] In the field of heterocyclic partition coefficient calculations, however, the algorithm needs considerable improvement and amplification. Recently some lines along which the method could be modified have been proposed.[3]

In this paper we report analysis of the partition coefficients of a series of furazans (1,2,5-oxadiazoles) and furoxans (1,2,5-oxadiazole 2-oxides). A few 1,2,5-X-diazole systems (X = S, Se) are also considered. A few suggestions to improve the CLOGP algorithm are given on the basis of the results obtained.
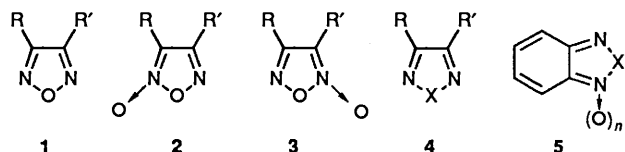
Experimental

Materials and Methods.—All the derivatives have been synthesized and purified according to the methods reported in the literature.[4,5] The log P values reported in Table 1 were obtained with the shake-flask method at room temperature using octanol as the lipidic phase and water as the hydrophilic phase. Each tabulated log P value is the average of four determinations made with different concentrations of solute.

Explanation of Symbols.—In the symbols used in this paper, the SMILES[6] convention is used, *i.e.*, an atom symbol in capitals is aliphatic and in lower case is aromatic; hydrogen is shown as H for both aliphatic and aromatic; f refers to fragment values and F to interaction factors; superscript c is calculated; m is measured, and Bz is benzyl. For example, f_C refers to an aliphatic carbon fragment and f_{non} refers to the furazan fragment where the hetero atoms are considered aromatic. The f values and correction factors actually used for the CLOGP are reported in footnote a of Table 1. Some selected additivity equations derived using these factors are reported in Table 2.

Results and Discussion

The Furazan Series

Calculation of the f_{non} Fragmental Constant.—From the measured log P^m of 1,2,5-oxadiazole 1a, a fragmental constant

$f_{non} = -0.55$ can be easily calculated, eqn. (1). As expected, f_{non} is a hydrophilic fragment.

The Problem of the F^{Bz} Correction.—CLOGP version 3.54 and later provides for a benzyl correction ($F^{Bz} = -0.15$) *i.e.* for aliphatic to benzene carbon bonds.[7] Thus, the expected log P^c for the dimethyl derivative 1b should be 1.15, eqn. (2). However, the measured log P^m_{1b} is 0.705 and therefore eqn. (i) holds.

$$\Delta = \log P^m - \log P^c = -0.45 \qquad (i)$$

This additional large negative factor ($-0.225/Bz$) for a benzyl carbon attached to a hetero ring needs an explanation. It is partially due to the restoration of the β^H of the aromatic π cloud to a value near that of benzene, *i.e.* 0.14.[8] Alternatively, the electron-donating methyl groups could be increasing β_H of the heteroatoms, but the actual location of the H-bond which reduces hydrophobicity in the methyl-substituted furazans has yet to be determined by far IR spectroscopy.[9] A third explanation, which is not in accord with the Kamlet–Taft–Abraham solvatochromic approach,[10] attributes the negative correction to an increase of positive charge on the hydrogens of the methyl substituent.

Whatever the underlying mechanism, if we assume a correction factor $F^{Bz*} = -0.38$ ($-0.15-0.225$) to take into account this fact, not only the calculation for 1b but also those for the others with methyl substituents are greatly improved. For example, log P values of the dialkyl derivatives 1c and 1d are readily predicted if this additional correction factor is used ($\Delta = -0.05, -0.05$), eqns. (3) and (4). Log P for 3-methyl-4-(p-nitrobenzyl)furazan (1e) is reasonably predicted, but it shows a positive deviation ($\Delta = +0.35$), eqn. (5). This point will be discussed further in the section on furoxans. This provides solid support for the f_{non} value derived from the parent 1a, in spite of the difficulties we found in handling this volatile compound.[11]

Extension of Aromaticity.—The CLOGP algorithm accounts for the increase of log P arising from the extension of the aromatic ring system through fusion or direct aromatic substitution by assigning positive correction factors.[1,2]

The fragment value for each carbon atom which either joins or fuses two rings is increased by 0.1 (F_c'). If the fusion atom is also bound to a heteroatom the increase is 0.31 (F_c^*). Thus, the expected log P value for benzofurazan 5a is 1.76, eqn. (6). This figure is in a very good agreement with experimental log P value for 5a ($\Delta = -0.07$). Also, log P for 3-methyl-4-phenylfurazan (1f) is readily predicted ($\Delta = +0.06$), eqn. (7).

Table 1 Partition coefficients

R	R'	Compounds	X	n	log P^m	CLOGP[a]	$\Delta = \log P^m - \log P^c$
H	H	1a			0.16	0.16	0
CH_3	CH_3	1b			0.705	0.70	0
CH_3	C_2H_5	1c			1.18	1.23	−0.05
CH_3	C_3H_7	1d			1.71	1.76	−0.05
CH_3	$CH_2C_6H_4NO_2$	1e			2.21	1.86	+0.35
CH_3	C_6H_5	1f			2.59	2.53	+0.06
CH_3	$CONH_2$	1g			0.305	0.06	+0.24
CH_3	$CONHCH_3$	1h			0.56	0.27	+0.29
CH_3	CO_2CH_3	1i			0.94	1.09	−0.15
CH_3	NH_2	1j			0.09	0.05	+0.04
CH_3	$NHCONHCH_3$	1k			0.30	0.34	−0.04
CH_3	$NHCOCH_3$	1l			0.17	0.40	−0.23
CH_3	$NHCO_2CH_3$	1m			0.42	0.71	−0.29
CH_3	NO_2	1n			0.97	1.04	−0.07
CH_3	SC_6H_5	1o			2.92	3.28	−0.36
CH_3	SOC_6H_5	1p			1.76	1.73	−0.03
CH_3	$SO_2C_2H_5$	1q			0.92	0.85	+0.07
CH_3	$SO_2C_6H_5$	1r			2.36	2.09	+0.27
C_6H_5	C_6H_5	1s			3.83	4.36	−0.53
C_6H_5	CH_2OH[b]	1t			1.915	1.22	+0.70
C_6H_5	NH_2	1u			1.81	1.88	−0.07
C_6H_5	NO_2	1v			2.63	2.87	−0.24
C_6H_5	Cl	1w			3.20	3.37	−0.17
NH_2	NH_2	1x			−0.49	−0.60	+0.11
NH_2	NHC_2H_5	1y			0.51	0.72	−0.21
NH_2	$NHCOCH_3$	1z			−0.23	−0.29	+0.06
NH_2	$NHCOC_6H_5$	1w			1.41	1.20	+0.21
CH_3	CH_3	2a			0.23	0.23	0
CH_3	$CH_2C_6H_5$	2b			2.08	1.65	+0.43
CH_3	C_6H_5	2c			2.17	2.06	+0.11
CH_3	CHO	2d			1.02	0.79	+0.23
CH_3	$COCH_3$	2e			0.80	0.57	+0.23
CH_3	$CONH_2$	2f			0.03	−0.22	+0.25
CH_3	CO_2CH_3	2g			0.56	0.82	−0.22
CH_3	NH_2	2h			−0.17	−0.22	+0.05
CH_3	NO_2	2i			0.78	0.77	+0.01
CH_3	SC_6H_5	2j			2.63	2.92	−0.29
CH_3	SOC_6H_5	2k			1.50	1.56	−0.06
CH_3	$SO_2C_2H_5$	2l			0.65	0.54	+0.11
CH_3	$SO_2C_6H_5$	2m			2.03	1.94	+0.09
CH_3	Cl	2n			1.02	1.16	−0.14
$CH_2C_6H_5$	$CH_2C_6H_5$	2o			3.65	3.07	+0.58
C_6H_5	C_6H_5	2p			3.64	3.89	−0.25
C_6H_5	CH_2OH[b]	2q			1.64	0.75	+0.89
C_6H_5	CHO	2r			2.61	2.61	0
C_6H_5	NH_2	2s			1.47	1.44	+0.03
C_6H_5	Cl	2t			2.82	2.98	−0.16
CH_3	$CH_2C_6H_5$	3b			2.22	1.65	+0.57
CH_3	C_6H_5	3c			2.16	2.06	+0.10
CH_3	CHO	3d			1.29	0.79	+0.50
CH_3	$COCH_3$	3e			0.98	0.50	+0.48
CH_3	$CONH_2$	3f			0.36	−0.29	+0.65
CH_3	CO_2CH_3	3g			0.69	0.82	−0.13
CH_3	NH_2	3h			−0.32	−0.39	+0.07
CH_3	SC_6H_5	3j			2.74	2.83	−0.09
CH_3	SOC_6H_5	3k			1.75	1.44	+0.31
CH_3	$SO_2C_6H_5$	3m			2.26	2.02	+0.24
CH_3	Cl	3p			0.90	1.15	−0.25
C_6H_5	CH_2OH[b]	3q			1.77	0.75	+1.02
C_6H_5	CHO	3r			2.915	2.61	+0.30
C_6H_5	NH_2	3s			1.42	1.44	−0.02
C_6H_5	Cl	3t			2.70	2.98	−0.28
CH_3	CH_3	4a	S		1.28	1.28	0
CH_3	CH_3	4b	Se		0.89	0.89	0
		5a	O	0	1.69	1.76	−0.07
		5b	O	1	1.43	1.29	+0.14
		5c	S	0	1.98	1.88	+0.10
		5d	Se	0	1.64	1.95	−0.31

[a] The reported CLOGP are values obtained using the following f values and correction factors: $f_H = 0.13$, $f_C = 0.195$, $f_c = 0.13$, $f_{NO_2} = -0.030$, $F^{Bz} = -0.15$, $F^{Bz*} = -0.38$, $F_b = -0.12$, $F_c* = 0.31$, $F_c^{\cdot} = 0.10$, $f_{non} = -0.55$, $f_{non(O)} -1.02$, $f_{nsn} = -0.33$, $f_{nsen} = -0.67$, $\sigma_{furazan} = 0.65$, $\rho_{furazan} = 1.30$, $\sigma_{furoxan} = 0.65$, $\rho_{furoxan} = 1.60$. [b] The hydroxymethyl fragments ($-CH_2OH$) on the 3- and 4-positions of pyridine and nitrobenzene are also underpredicted by ca. 0.5 log units, and it appears that a CLOGP change allowing ca. $\frac{2}{3}$ of any F_{elect} to operate through a benzyl carbon would be justified. In the present cases (1t, 2g and 3q) this may also encourage H-bond formation to make the correction even greater.

Table 2 Equations used in the CLOGP calculations and discussed in the text

Equation number	Equation
(1)	$\log P^m{}_{1a} = 2f_H + 2f_c + f_{non}$ $0.16 = 2(0.13) + 2(0.227) + f_{non}$ $(f_{non} = -0.55)$
(2)	$\log P^c{}_{1b} = 6f_H + 2f_C + 2f_c + f_{non} + 2F^{Bz}$ $6(0.227) + 2(0.195) + 2(0.13) - 0.55 + 2(-0.15) = 1.15$
(3)	$\log P^c{}_{1c} = 8f_H + 3f_C + 2f_c + 2F^{Bz*} + F_b + f_{non}$ $8(0.227) + 3(0.195) + 2(0.13) + 2(-0.38) - 0.12 - 0.55 = 1.23$
(4)	$\log P^c{}_{1d} = 10f_H + 4f_C + 2f_c + 2F^{Bz*} + 2F_b + f_{non}$ $10(0.227) + 4(0.195) + 2(0.13) + 2(-0.38) + 2(-0.12) - 0.55 = 1.76$
(5)	$\log P^c{}_{1e} = 9f_H + 2f_C + 8f_c + f_{NO_2} + 2F^{Bz*} + F^{Bz} + F_b + f_{non}$ $9(0.227) + 2(0.195) + 8(0.13) - 0.030 + 2(-0.38) - 0.15 - 0.12 - 0.55 = 1.86$
(6)	$\log P^c{}_{5a} = 4f_H + 6f_c + 2F_c{}^* + f_{non}$ $4(0.227) + 6(0.13) + 2(0.31) - 0.55 = 1.76$
(7)	$\log P^c{}_{1f} = 8f_H + 8f_c + 1f_C + F^{Bz*} + F_c{}^{\cdot} + F_c{}^* + f_{non}$ $8(0.227) + 8(0.13) + 0.195 - 0.38 + 0.1 + 0.31 - 0.55 = 2.53$
(8)	$\log P^m{}_{2a} = 6f_H + 2f_c + 2f_C + 2F^{Bz*} + f_{non(O)}$ $0.23 = 6(0.227) + 2(0.13) + 2(0.195) + 2(-0.38) + f_{non(O)}$ $(f_{non(O)} = -1.02)$
(9)	$\log P^c{}_{5b} = 4f_H + 6f_c + 2F_c{}^* + f_{non(O)}$ $4(0.227) + 6(0.13) + 2(0.31) - 1.02 = 1.29$
(10)	$\log P^c{}_{2c} = \log P^c{}_{3c} = 8f_H + 8f_c + 1f_C + F^{Bz*} + F_c{}^{\cdot} + F_c{}^* + f_{non(O)}$ $8(0.227) + 8(0.13) + 0.195 - 0.38 + 0.1 + 0.31 - 1.02 = 2.06$
(11)	$\log P^c{}_{2p} = 10f_H + 14f_c + 2F_c{}^{\cdot} + 2F_c{}^* + f_{non(O)}$ $10(0.227) + 14(0.13) + 2(0.1) + 2(0.31) - 1.02 = 3.89$
(12)	$\log P^m{}_{5c} = 1.98 = 4f_H + 6f_c + 2F_c{}^* + f_{nsn}$ $f_{nsn} = 1.98 - 4(0.227) - 6(0.13) - 2(0.31) = -0.328$
(13)	$\log P^m{}_{5d} = 1.64 = 4f_H + 6f_c + 2F_C{}^* + f_{nsen}$ $f_{nsen} = 1.64 - 4(0.227) - 6(0.13) - 2(0.31) = -0.668$
(14)	$\log P^c{}_{4a} = 6f_H + 2f_c + 2f_C + 2F^{Bz} + f_{nsn}$ $6(0.227) + 2(0.13) + 2(0.195) + 2(-0.15) + (-0.33) = 1.38$
(15)	$\log P^c{}_{4b} = 6f_H + 2f_c + 2f_C + 2F^{Bz} + f_{nsen}$ $6(0.227) + 2(0.13) + 2(0.195) + 2(-0.15) + (-0.67) = 1.04$

The molecular mechanics program, CONCORD,[12] prefers the phenyl ring at an angle to the furazan ring due to a small degree of steric strain imposed by the adjacent methyl group. However, this effect can be shown to be small and easily overcome by the lower energy of the extended aromatic system when planar.

A different situation occurs for the 3,4-diphenylfurazan (1s), where planarity of the three rings is not possible. In this case there is a large negative Δ value (-0.53), which is logically explained by a resonance decoupling which CLOGP does not consider.

σ/ρ *Interaction in the Furazan System.*—When one or two substituents (not alkyl, arylalkyl or aryl groups) are attached to a heterocyclic aromatic ring, it is necessary to take into account the electronic interactions which operate in the system in order to obtain correct log P calculations. These interactions are treated by assigning to each substituent and to each hetero fragment present in the ring a pair of σ/ρ constants.[1,2]

The σ constant represents the ability of the moiety to behave as 'inducer' in the interaction. For electron-withdrawing substituents, it approximately parallels the classic Hammett σ value and appears to measure the reduction in H-bond acceptor strength of the fragment on which it acts.[2] On the other hand, ρ measures the susceptibility of the H-bonding fragment to electron withdrawal. By an iterative process we found that the best σ/ρ values which reconcile experimental and calculated log P^c values are, for the non fragment, σ = 0.65 and ρ = 1.30.

When two ρ-type substituents are present on the ring (*e.g.* 1x, 1y, *etc.*) the effect is additive, *i.e.* it does not 'age'.[2] Generally speaking the CLOGP algorithm satisfactorily predicts log P values of the furazans in Table 1. However, a few anomalies remain. It is difficult to explain why the CONH$_2$ analogue is underpredicted while CO$_2$Me is overpredicted. This will be discussed further in the section on furoxans.

The Furoxan Series

Furoxans are furazan *N*-oxides. Simple unsymmetrically-substituted furoxans exist as two separate structural isomers, A and B.[4]

A　　　**B**　　　**C**
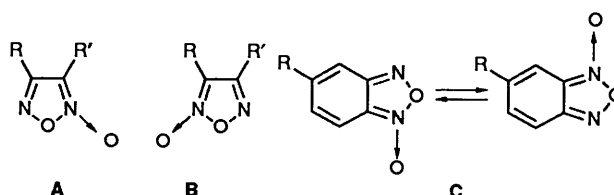
Benzofuroxans **C** present somewhat different problems because they are, in solution, rapidly exchanging systems.[4] No attempt was made to consider separately the two isomeric furoxan series and the analysis of log P was undertaken without considering the 3- and 4-positions as being different.

Calculation of the $f_{non(O)}$ *Fragmental Constant and* F^{Bz*} *Correction.*—In the furoxan series it is impossible to evaluate $f_{non(O)}$ from the measured log P value for the 1,2,5-oxadiazole 2-oxide (furoxan) parent, since because of its instability it is detectable only as transient species.[13] Therefore, we chose the 3,4-dimethylfuroxan (2a) as a reference compound for the calculation of the $f_{non(O)}$ value. It is possible that a different F^{Bz*} value is necessary to take into account a different degree of positive charge on the methyl group when attached to furoxan ring, but since a proliferation of F^{Bz*} correction factors is clumsy for an algorithm like CLOGP, we used the value $F^{Bz*} = -0.38$ as found for the furazan system. Using this approximation the $f_{non(O)}$ fragmental constant value is -1.02, eqn. (8). As expected, $f_{non(O)}$ is more negative than f_{non}.

We can now analyse log P values for derivatives **2b**, **3b** and **2o** which have CH_2Ph attached to the hetero ring. These calculate rather unsatisfactorily and show, as was the case with **1e** in the furazan series, a positive deviation. From coordinates developed from the molecular mechanics program CONCORD one clearly sees that the phenyl ring of a 3- or 4-benzyl substituent is forced close to (and may block solvent access to) the nitrogen of the furazan fragment. This may be the reason it is more hydrophobic than CLOGP calculates.

Extension of Aromaticity.—Benzofuroxan **5b** is the only example in our furoxan series of extension of aromaticity by ring fusion and its log P value can be calculated satisfactorily ($\Delta = +0.14$), eqn. (9). 3-Methyl-4-phenylfuroxan (**2c**) and its 4-methyl isomer **3c** calculate well ($\Delta = +0.11$ and $\Delta = +0.10$, respectively), eqn. (10). In the case of the 3,4-diphenylfuroxan (**2p**) the calculation error is larger, but is still in the acceptable range ($\Delta = -0.25$), eqn. (11).

Like the furazan derivative **1s** the negative deviation could again be explained by resonance decoupling consequent to the loss of planarity between phenyl and furoxan rings. In all phenyl derivatives considered, both in the furazan and the furoxan series, it is probable that resonance decoupling is accompanied by reduction in the hydrophobic SASA. For these reasons speculations on the degree of planarity of the rings in the systems considered are difficult simply on the basis of log P values.

σ/ρ *Interaction in the Furoxan System.*—By an iterative approach we found that the best σ/ρ values for the non(O) fragment which reconcile experimental and calculated log P values of furoxans substituted with groups different from alkyl, aryl and benzyl moieties are $\sigma = 0.65$ and $\rho = 1.60$. On the whole, log P for the furoxans in this study are reasonably well-predicted by the CLOGP program. However, it is not very clear what is responsible for several deviations. The $CONH_2$ analogues in the furoxan series are underpredicted, just as in the furazan series, and the deviation is greater when the amide is adjacent to the 2-oxide moiety (Δ for **3f** = +0.65) than when it is opposite (Δ for **2f** = +0.25). The much more positive log P for **3f** is probably due to a favourable hydrogen bond to the $N \longrightarrow O$ oxygen atom.[14] A few anomalies, *e.g.*, the underprediction of $CONH_2$ analogues and the overprediction of CO_2Me analogues, can be explained by taking into account the fact that the present version of CLOGP uses the same '*ortho* matrix' for six-membered aromatic rings as for five, but the negative correction for 'twist decoupling' and the positive correction for intramolecular H-bonding may well be different owing to the different ring angles.

Thia- and Selena-1,2,5-diazole Derivatives

The unsubstituted 1,2,5-X-diazole rings are difficult to synthesize and purify, and the thia analogue is also very volatile and difficult to handle.[15,16] This makes it difficult to evaluate f_{nsn} and f_{nsen} directly as was done with the oxygen analogue. As was noted previously, the electronegative five-membered

heteroaromatic rings containing two nitrogens or nitrogen(s) and oxygen decrease the hydrophobicity of any 'benzyl' carbons attached to them ($F^{Bz} = -0.15$ to $F^{Bz*} = -0.38$). Since it was already known that a lesser F^{Bz} value applies to thiazole rings, it was a distinct possibility that the alkyl-substituted thia- and selena-diazoles might require an F factor with an intermediate value. For this reason, an alternative fragment evaluation method was attempted; namely, by the way of the benzo-fused systems (**5c** and **5d**) the fragment values $f_{nsn} = -0.328$ and $f_{nsen} = -0.668$ are respectively calculated, eqns. (12) and (13). Assuming the lesser F^{Bz} value, the Δ values for 3,4-dimethyl-1,2,5-thiadiazole (**4a**) and the corresponding seleno analogue (**4b**) are -0.10 and -0.15, respectively, eqns. (14) and (15). It appears, therefore, that the F^{Bz} value for thia- and selenadiazoles is *ca.* -0.21, and therefore much more like the correction for attachment to a phenyl ring than the diazo-aromatics. This should be included in the algorithm of future versions of CLOGP.

Conclusions

This work demonstrates the utility in a program that consistently calculates log P according to a known set of rules. Not only can one predict the hydrophobicity of unmeasured structures, but substantial deviations from measured values can suggest electronic and steric interactions which otherwise may go unnoticed.

References

1 Pomona College, Med. Chem. Project, Claremont, California 91711, USA: THOR MASTERFILE 353; CLOGP version 3.54, 1991.
2 A. J. Leo, in *Comprehensive Medicinal Chemistry*, ed. C. A. Ramsden, vol. 4, Pergamon, Oxford, 1990, p. 295.
3 J. Bradshaw and P. J. Taylor, *Quant. Struct.-Act. Relat.*, 1989, **8**, 279.
4 A. Gasco and A. J. Boulton, *Adv. Heterocycl. Chem.*, 1981, **29**, 251.
5 A. M. Gasco, R. Fruttero, G. Sorba and A. Gasco, *Liebigs Ann. Chem.*, 1991, 1211, and other papers of the series.
6 D. Weininger, *J. Chem. Inf. Comput. Sci.*, 1988, **28**, 31.
7 A. Leo, *J. Chem. Soc., Perkin Trans. 2*, 1983, 825.
8 M. Abraham, G. Whiting, R. Doherty and W. Shuely, *J. Chem. Soc., Perkin Trans. 2*, 1990, 1451.
9 P. G. Green, *Towards a Quantitative Understanding of the Hydrogen Bond*, Harvey Mudd College Lecture Series, Claremont CA; 1992, unpublished.
10 D. Leahy, *J. Pharm. Sci.*, 1986, **75**, 629.
11 R. A. Olofson and J. S. Michelman, *J. Org. Chem.*, 1965, **30**, 1854.
12 R. Pearlman and co-workers, *Concord R Users Manual*, Concord, V. 2.9.3., January 1992, Tripos Associates Inc., St. Louis Mo.
13 C. Grundmann, G. W. Nickel and R. Kumar Bansal, *Ann. Chem.*, 1975, 1029.
14 A. Gasco and A. J. Boulton, *J. Chem. Soc., Perkin Trans. 2*, 1973, 1613.
15 L. M. Weinstock, P. Davis, B. Handelsam and R. Tull, *J. Org. Chem.*, 1967, **32**, 2823.
16 V. Bertini, *Gazz. Chim. Ital.*, 1967, **97**, 1870.