σ_I Values for Heterocycles†

Peter J. Taylor* and Alan R. Wait
I.C.I. Pharmaceuticals Division, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4TG

Charton's relation, that the pK_a values of substituted guanidines are solely a function of σ_I, has been vindicated in a comprehensive study involving many more compounds. The resulting equation has been used, in conjunction with the pK_a values of a range of guanidinoheterocycles, to derive the σ_I values for a number of heteroaryl substituents. Trends among the data are discussed; some apparent anomalies are noted and explained.

In recent years, the dual substituent parameter (DSP) approach has emerged as a powerful tool for the elucidation of chemical reaction mechanisms and intramolecular, *e.g.* spectroscopic, interactions.[1–4] Essentially the DSP approach consists in analysing the data in question by means of a combination of field (σ_I) and one of several possible sets of resonance (σ_R) substituent constants; this has the advantage over use of the familiar Hammett values, σ_m or σ_p, that no explicit blend of field and resonance forces need be assumed. Equation (1) has become accepted as the defining relation;[3,5] since many σ_p values are known,[6] or in general are easy to obtain, a virtue of (1) is that in

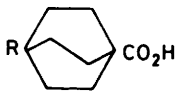
$$\sigma_p = \sigma_I + \sigma_R \quad (1)$$

such cases, if σ_I or σ_R can be established by some means, the other one then follows.

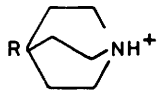
Use of the DSP approach clearly requires the existence of a large established pool of substituent σ_I and σ_R values, and several recent compilations exist.[3,5,6] For a number of reasons[1,3,5,7] it is much more difficult to find processes that depend wholly on resonance interactions than processes that depend wholly on field effects, so that, in general, σ_I has to be established for a substituent ahead of σ_R; in that sense, σ_I may be regarded as the primary variable. Charton's compilation,[5] one of the most comprehensive and internally consistent to date, lists over 250 σ_I values. From this compilation, σ_I values for heterocycles *considered as substituents* are almost wholly absent.

The classical method of establishing σ_I values is through pK_a measurement for substituted bicyclo[2.2.2]octane-1-carboxylic acids (**1**) or quinuclidinium cations (**2**) from which resonance interactions are known to be absent.[3,5,7] Few such measurements exist, and none for heterocycles. Useful secondary standards include pK_a values for substituted acetic acids RCH_2CO_2H, provided that steric and hydrogen-bonding interactions can be excluded.[5] This is likely to be a particular problem for heterocycles, especially those here to be discussed; in the closely related case of the heterocyclic carboxylic acids, it has been noted[8] that unusual changes in ΔH and ΔS on ionisation make these pK_a values unsuitable for the deduction of σ constants. Up to the present, there has seemed to exist no clearly reliable method for measurement of the σ_I values for heteroaryl substituents.

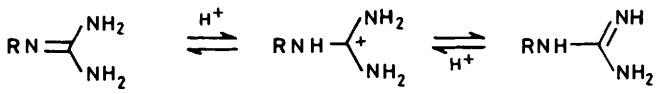
† Since the 'inductive' is now established as wholly or predominantly a field effect, a number of authors have recently advocated replacement of the symbol σ_I by σ_F. We have chosen to stay with the older nomenclature, not only for reasons of familiarity, but also because we believe that there is a potentially useful distinction to be drawn between σ_F as a theoretical or gas-phase quantity and σ_I when modified by solvation, as is certainly the case here.



(1)



(2)



(3a) (3b) (3c)

Some 20 years ago, Charton[9] established equation (2) as the relation which best fits the pK_a values of substituted

$$pK_a = 14.20 - 24.09\ \sigma_I \quad (2)$$
$$(n\ 8,\ R^2\ 98.0\%,\ s\ 0.80)$$

guanidinium cations (**3**). Our H_2-receptor antagonist programme[10] has concentrated largely on guanidinoheterocycles[11] some of which, if equation (2) is valid, are suitable for establishing heteroaryl σ_I values. A specially attractive feature of equation (2) in this respect is its high ρ_I value, which satisfies Charton's criterion[5] that even quite large errors in the measured variable (pK_a) would lead, in the absence of systematic variations, to σ_I values of acceptable accuracy. Before that desirable aim can be achieved, however, it is necessary to re-validate Charton's relation in the light of more recent data and a careful consideration of possible perturbing factors. We start by performing that re-validation, and go on to consider the heteroaryl σ_I values to which the revised regression equation gives rise.

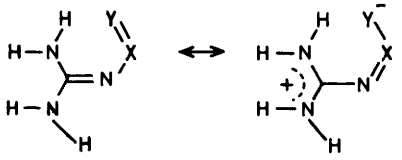
The Validity of Charton's Relation.—Charton's relation is based on eight points of which two must be discarded for special reasons. It is now accepted[1,5,7] that poles and dipoles must not be mixed in Hammett-style correlations, a point recently re-emphasised by theoretical calculation,[12] so the anion (**17**) has to be dropped from the analysis. The other point to be rejected is that for cyanoguanidine (**20**). Perhaps uniquely in this context, the second π-orbital of the cyano group must conjugate with the lone-pair electrons of the imino-nitrogen to be protonated; the resulting drop in pK_a value from that otherwise expected is >2 units (see Table 1). Removal of this point is in fact responsible for most of the difference between equation (2) and the final regression equation (3); *cf.* the parameters of Table 3.

This leaves six points of which five span less than half the pK_a range, a dangerously lopsided situation. Furthermore, while Charton's σ_I correlation was based on the reasonable evidence[9] that σ_I, σ_m, and σ_p account for the variation in pK_a successively

Table 1. pK_a and parameter values for substituted guanidines (3)[a]

	R	pK_a Ref. 9	pK_a *b*	pK_a This work	σ_I	σ_R	*H*	pK_β	ΔpK_a
(4)	Me		**14.1**		−0.01	−0.16	0	0	−0.31
(5)	H	14.46	**14.38**		0.00	0.00	0	0	+0.20
(6)	Ph	**10.77**			0.12	−0.11	0	0	−0.70
(7)	NH_2		**11.04**		0.17	−0.80	0	0	+0.70
(8)	$4\text{-}NO_2C_6H_4$		9.28	**9.13** ± 0.02	0.23	0.03	0	0	+0.14
(9)	OH			**7.96** ± 0.04	0.24	−0.62	0	0	−0.80
(10)	NHCOPh			**7.94** ± 0.06	0.28	−0.47	0	0	+0.08
(11)	$CONH_2$	7.85		**8.11** ± 0.05	0.28	0.08	1	2.76	+0.25
(12)	COMe	8.33		**8.20** ± 0.05	0.30	0.20	1	1.50	+0.79
(13)	COPh			**6.98** ± 0.05	0.30[c]	0.11	1	1.41	−0.43
(14)	CO_2Et			**7.03** ± 0.05	0.30	0.11	1	1.39	−0.38
(15)	NHPh			**8.26** ± 0.06	0.30	−0.86	0	0	+0.85
(16)	OMe	**7.46**			0.30	−0.58	0	0	+0.05
(17)	$OCH_2CO_2^-$	7.51			0.32				(+0.56)[h]
(18)	$CSNH_2$			**5.56** ± 0.02	0.38[d]	0.11[e]	1	1.60	−0.04
(19)	SO_2NH_2			**1.83** ± 0.02	0.53[c,f]	0.05[g]	1	1.10	−0.28
(20)	CN	−0.4		−0.85 ± 0.05	0.57				(−2.16)[h]
(21)	NO_2	−0.93		**−0.98** ± 0.02	0.67	0.10	1	1.20	−0.03

[a] pK_a Values in bold type are those used in the correlations; σ_I and σ_R values are from ref. 5 unless otherwise stated; the meaning of *H* and pK_β is defined in the text; ΔpK_a is the deviation of the observed from the calculated pK_a value according to equation (3). [b] D. D. Perrin, 'Dissociation Constants of Bases in Aqueous Solution,' Butterworths, London, 1972, with statistical corrections × 5 applied to (4) and × 6 to (5). [c] From equation (1). [d] M. Charton, personal communication. [e] R. T. C. Brownlee and M. Sadek, *Aust. J. Chem.*, 1981, **34**, 1593. [f] O. Exner, *Collect. Czech. Chem. Commun.*, 1966, **31**, 65. [g] Ref. 3. [h] Not included in the correlation.



(22a) (22b)

less well, the fully fledged DSP treatment did not exist at that time and many more values of σ_I and σ_R have become available in the intervening years. We therefore set out to measure the pK_a for as many guanidines as possible of established substituent σ_I and σ_R value. The results of this re-investigation are assembled in Table 1.

In assembling these compounds we were concerned both to cover the pK_a range as evenly as we could and to watch out for possible systematic deviations independent of the σ_I–σ_R dichotomy. On the first objective we have to report indifferent success. Compound unavailability still restricts the number of compounds of low pK_a: such desirables as (3; R = OPh) and (3; R = $COCF_3$) were unobtainable, whereas for (3; R = SO_2Me) and (3; R = SO_2Ph) pK_a measurement proved impossible since the u.v. change on protonation is too small. Nevertheless, there have been certain compensations: despite the usual annoying clustering of substituents at σ_I *ca.* 0.3, their spread in σ_R has proved of great value in establishing an unequivocal result.

In guarding against systematic deviations we had two main sources in mind. The first is the possibility, for which there is precedent,[13] of a unidirectional resonance effect. It is easy to see how resonance acceptors may stabilise the free base, as in the generalised structure (**22b**), and indeed we possess spectroscopic evidence[14] for strong resonance interactions in acylguanidines; it is much more difficult to see how resonance donors can stabilise the cation (except in the form of σ resonance: see later). We have attempted to test for this by using σ_R in two forms: as written in Table 1, or alternatively with the values for all donors set equal to zero. This latter set is designated σ_{R^-} in Tables 2 and 3. The other possibility is that of differential effects due to intramolecular hydrogen bonding, which is possible, and indeed

Table 2. Correlation matrix for the variables of Table 1[a]

	pK_a	σ_I	σ_R	σ_{R^-}	*H*	pK_β
pK_a	1.00					
σ_I	−0.99	1.00				
σ_R	−0.25	0.22	1.00			
σ_{R^-}	−0.45	0.47	0.70	1.00		
H	−0.65	0.64	0.72	0.87	1.00	
pK_β	−0.46	0.47	0.65	0.79	0.92	1.00

[a] For meaning of σ_{R^-} see text.

virtually obligatory, for seven of these 16 compounds.* This again is illustrated for structure (**22**): while such a bond will exist in both free base and cation, it should be stronger in the latter and so base-strengthening. We have chosen two ways of testing for this. The first is by use of the indicator variable *H*, which is given a value of unity where bonding is present and zero where not. Alternatively, we have attempted to quantify the strength of this bond by means of our recent pK_β scale of proton-acceptor ability;[15] this potentially introduces scaling difficulties but avoids the assumption of an all-or-nothing response.

The most relevant regression equations are listed in Table 3. Equation (3), a simple update of Charton's equation (2), emerges as easily the best; it is depicted graphically in the Figure. No other single-variable equation exceeds chance expectation, so none is listed. No other variable in combination with σ_I possesses greater than chance expectation either, so equations (4)—(7) are invalid also. The nearest approach to significance among the remaining possibilities is the three-variable equation (8), in which σ_{R^-} and *H* appear. However, the probable reason for this is the high mutual correlation (r 0.87) of these variables (Table 2); here a tell-tale indication is that the *sign* of each is the reverse of expectation. We also list an equation (9) which differs from (3) only in that methylguanidine

* It is possible that compound (**10**) could form an intramolecular bond in the free base, but unlikely that this would survive in the cation, so it has been counted as non-bonding.

Table 3. Correlation equations for pK_a based on the data of Table 1 [a]

	n	R^2%	s	F	p
(3) 14.18 (±0.25) − 22.58 (±0.78) σ_I	16	98.4	0.51	835	
(4) 14.07 (±0.28) − 22.42 (±0.80) σ_I − 0.36 (±0.38) σ_R	16	98.5	0.51	415	0.358
(5) 14.18 (±0.26) − 22.84 (±0.91) σ_I + 1.53 (±2.45) σ_R·	16	98.4	0.52	399	0.55
(6) 14.17 (±0.26) − 22.40 (±1.06) σ_I − 0.09 (±0.35) H	16	98.4	0.53	390	1.00
(7) 14.18 (±0.26) − 22.65 (±0.92) σ_I + 0.03 (±0.18) pK_β	16	98.4	0.53	388	1.00
(8) 14.07 (±0.25) − 21.98 (±1.03) σ_I + 7.35 (±4.43) σ_R· − 0.96 (±0.62) H	16	98.7	0.50	296	0.123 [b] 0.146 [c]
(9) 14.17 (±0.29) − 22.53 (±1.04) σ_I	15	98.3	0.53	473	

[a] Statistics: n = number of points, R = correlation coefficient for fit of points to line, s = standard error of fit, F = F statistic (significance of fit allowing for degrees of freedom), p = probability, on a scale of 0 to 1, that each variable is present due to chance alone. For significance, $p < 0.05$ is required; p is listed for variables other than σ_I, for which $p < 0.000\,01$ throughout. [b] For σ_R·. [c] For H.

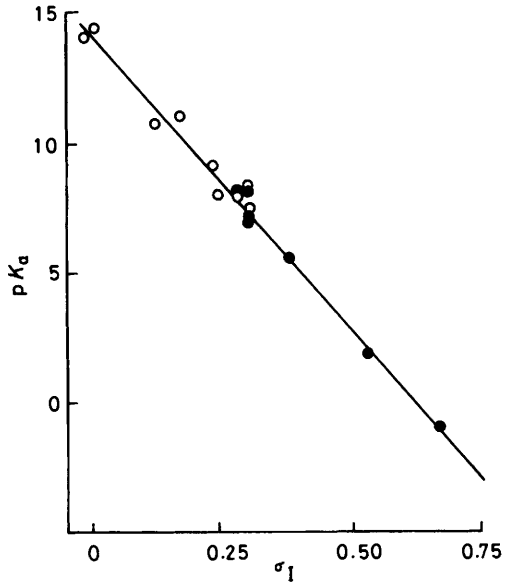
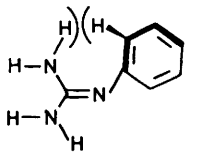
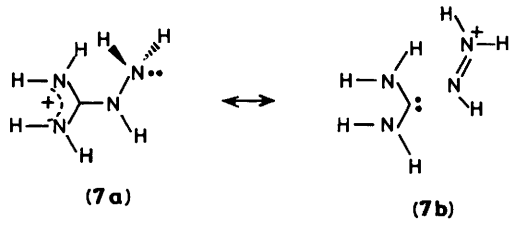


Figure. pK_a *versus* σ_I for the compounds of Table 1. Full circles denote compounds in which an intramolecular hydrogen bond is likely to be present; open circles denote the remainder. The correlation line is that of equation (3)



(6)



(7a) (7b)

(4) has been dropped. Charton [9] reasonably excluded (4) on the ground that its tautomeric form is equivocal and indeed tautomer (3c) may be preferred. However, it is probable that K_T is close to unity for this compound, in which case our statistical correction (Table 1) should cope. In no other case is tautomeric ambiguity likely to be present.[16] The improvement in statistics on going from (9) to (3) is substantial while the equations themselves are not appreciably different, so we prefer the latter.

Because of the possibility of an enhanced resonance interaction, we have also re-run these correlations with σ_R^- in place of σ_R; since every such correlation was even worse, none is listed.

The absence of a resonance component in the pK_a values of substituted guanidines is at first sight rather surprising. One reason of course is the very large value for ρ_I, the result presumably of the enormous difference in dipole gradient and direction adjacent to the substituent in free base form (3a) and cation (3b). Nevertheless any sizeable ρ_R should have been picked up by the present statistical treatment, and a more subtle factor may also be present. This is the ability, absent of course from simple amines, of cation as well as free base to act as a π-donor, since lone pairs are still present after protonation; we have recently reported the use of a guanidinium *cation* as a nucleophile.[17] The resulting stabilisation of both species by resonance acceptors could be nearly self-cancelling, in sharp contrast to the unequivocal sign expected for the field effect.

The apparent absence of any influence from an intramolecular hydrogen bond is also quite surprising. Again we suspect a cancellation of factors. The preferred bond angle at imino-nitrogen is[17] *ca.* 110°; in the cation however this will open to 120° or so. The resulting hydrogen-bond lengthening in the cation could cancel the otherwise expected greater bond strength, with nil nett effect on basicity. Other bond length and angle changes will also take place, of course, but this would appear to be the most relevant one.[17] The Figure clearly indicates, *inter alia*, the absence of systematic deviations due to hydrogen bonding.

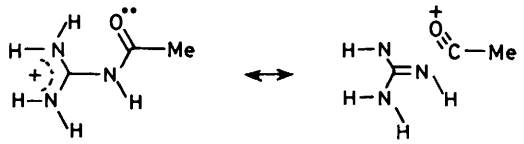
It would be tempting at this point to accept equation (3) without further ado. However, and despite the fact that points which exceed one s.d. from the regression line are the expected five out of 16, there is strong statistical evidence for *non-random* deviations. This is shown by the fact that the s.d. for the regression, at 0.51, is vastly greater than the mean s.e., <0.1, for the experimental points. Hence we need to explain at least the larger deviations, or there is a risk that the regression line is inapplicable for some special reason to the heterocycles we wish to examine. This is emphasised by the fact that here we are dealing not with XGY in Charton's formalism,[5] but with XY; *i.e.* with probe Y directly adjacent to substituent X, rather than separated from it by a buffer zone G, as is present in (1) and (2) and in the classical benzenoid context.

Reasonable explanations can in fact be offered for most of the more seriously deviant points. The benzene ring of phenylguanidine (6) is very likely twisted out of plane by the 1,6-interaction shown; the resulting partial π-overlap with the imino-nitrogen lone pair could lead to the base weakening observed. This is then similar to, but much less developed than,

Table 4. pK_a and derived σ_I values for heterocyclic guanidines[a,b]

	pK_a	σ_I
Pyridin-2-yl	10.17	0.18
6-Methylpyridin-2-yl	10.26	0.17
Quinolin-6-yl	10.23	0.17
Pyrazin-2-yl	8.53	0.25
6-Methylpyrazin-2-yl	8.76	0.24
5,6-Dimethylpyrazin-2-yl	9.07	0.23
Quinoxazin-2-yl	8.15	0.27
4,6-Dimethylpyrimidin-2-yl	9.30	0.21[c]
4-Phenylpyrimidin-2-yl	9.34	0.21[c]
4-Methylquinazolin-2-yl	9.19	0.22[c]
2-Methylpyrimidin-4-yl	8.55	0.25
2-Phenylpyrimidin-4-yl	7.90	0.28
Pyridazin-3-yl	8.28	0.26
3-Methyl-1,2,4-triazin-6-yl	6.09	0.36
3-Phenyl-1,2,4-triazin-6-yl	5.03	0.40
4-Methyloxazol-2-yl	5.85	0.37
4,5-Dimethyloxazol-2-yl	6.30	0.35
Benzoxazol-2-yl	4.95	0.41
Thiazol-2-yl	6.57	0.34
4-Methylthiazol-2-yl	7.05[d]	0.32
4,5-Dimethylthiazol-2-yl	7.25	0.31
Benzothiazol-2-yl	5.73	0.37
4-Methylimidazol-2-yl	8.39	0.26
1-Methylimidazol-2-yl	8.34	0.26
Benzimidazol-2-yl	6.97	0.32
1-Phenylpyrazol-3-yl	9.50	0.21
3-Methyl-1,2,4-oxadiazol-5-yl	3.36	0.48
3-Methyl-1,2,4-thiadiazol-5-yl	5.17	0.40
2*H*-1,2,3-triazol-2-yl	4.85	0.41
Tetrazol-5-yl	3.16	0.49
3,4-Dihydro-4-oxopyrimidin-2-yl	4.95	0.41
3,4-Dihydro-4-oxoquinazolin-2-yl	5.09	0.40
1,6-Dihydro-6-oxopyridazin-3-yl	8.00	0.27
1,6-Dihydro-1-methyl-6-oxopyridazin-3-yl	7.81	0.28
4,5-Dihydro-4-oxothiazol-2-yl	3.88	0.46

[a] Substituent as R in (**3a**); σ_I deduced from pK_a by means of equation (3). [b] Mean s.d. ±0.05 in pK_a. [c] Minimum value: see text. [d] Ref. 16.



(**12a**) (**12b**)

the effect found for cyanoguanidine (**20**). Its absence in the nitro derivative (**8**) may be due to an enhanced resonance interaction across the benzene ring that tends to force planarity, or just to the increased length of the dipole. Both simple amino derivatives, (**7**) and (**15**), by contrast possess enhanced basicity; this may be due to σ resonance in the cation of the type illustrated for (**7**) as form (**7b**). Particularly in view of this, the anomalously low basicity of hydroxyguanidine (**9**) is a surprise. We offer no explanation,* but would note that lyate anions and related species are notorious for bad behaviour; OH as substituent is singled out by Charton[5] in this respect. The related methoxy derivative (**16**) behaves normally.

The enhanced basicity of acetylguanidine (**12**) is particularly disturbing in view of its selection as a putative model for the guanidinoheterocycles.[14,16,17] We believe this anomalous rise to be due to σ resonance in the cation of the type portrayed in structure (**12b**). This should be much less important for the guanidinoheterocycles (**23**) since the ring nitrogen lone pair is poorly placed to act as donor, and is also likely to be largely suppressed when Me in (**12b**) is replaced by a σ acceptor such as aryl or a heteroatom. Consistently, the appropriate derivatives (**11**), (**13**), and (**14**) show a mean deviation of −0.19 (±0.38) from the regression line. If these are reasonable models for the guanidinoheterocycles, there is no reason why the regression line of equation (3) should not be valid.

We therefore conclude that, with minor exceptions to be discussed, equation (3) does in fact represent a valid basis for estimating σ_I values for heterocycles as substituents (X in XGY[5]). Its very large ρ_I value means that an error of *ca.* ±0.25 pK_a units is required to introduce a corresponding error of ±0.01 in σ_I. Put another way, the standard error of the regression imposes an imprecision in σ_I of *ca.* ±0.02 at the centre of the observed scale and rather more than this towards its extremes. While this is somewhat below the standards of the best data assessed by Charton,[5] it may be considered acceptable in present circumstances.

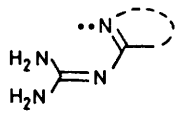
The Heteroaryl σ_I *Values.*—Table 4 lists the guanidinoheterocycle pK_a values we have measured and the σ_I values derived therefrom. Where the parent heterocycle pK_a (and therefore σ_I) value is missing, some of these data allow their extrapolation; in particular, it is possible to deduce a mean decrease in σ_I of 0.012 (±0.007) per added methyl group. In Table 5 we assemble these parent heterocycle σ_I values, some of

Table 5. σ_I and σ_R values for parent heteroaryl and some reference substituents[a]

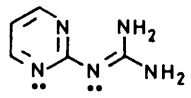
	σ_I	σ_R	σ_{R+}	σ_{R-}
Phenyl	0.12[b]	−0.11[b]	−0.17[b]	−0.11[b]
Acetyl	0.30[b]	0.20[b]	0.06[b]	0.41[b]
Amino	0.17[b]	−0.80[b]	−1.10[b]	−0.55[b]
Pyridin-2-yl	0.18			(0.57)[c]
	0.20[b]			
Quinolin-6-yl	0.17			
Pyrazin-2-yl	0.25			
Quinoxazin-2-yl	0.27			
Pyrimidin-2-yl	0.23[d]			
Quinazolin-2-yl	0.23[d]			
Pyrimidin-4-yl	0.26			
Pyridazin-3-yl	0.26			
1,2,4-Triazin-6-yl	0.37			
2-Furyl	0.17[b]	−0.19[b]	(−0.52)[c]	(0.04)[c]
3-Furyl	0.10[c]		(−0.55)[c]	
2-Thienyl	0.19[b]	−0.19[b]	(−0.52)[c]	(0.00)[c]
3-Thienyl	0.10[c]	(−0.12)[c]	(−0.48)[c]	(0.03)[c]
Pyrrol-2-yl	0.17[b]			
Indol-3-yl	0.01[b,c]			
1-Phenylpyrazol-3-yl	0.21			
Oxazol-2-yl	0.38			
Benzoxazol-2-yl	0.41	(−0.07)[c]		(0.27)[c]
Thiazol-2-yl	0.34			
Benzothiazol-2-yl	0.37	(−0.03)[c]		(0.28)[c]
Imidazol-2-yl	0.27			
1-Methylimidazol-2-yl	0.26			
Benzimidazol-2-yl	0.32			(0.16)[c]
Imidazol-4(5)-yl	0.08[b]			
1,2,4-Oxadiazol-5-yl	0.49			
1,2,4-Thiadiazol-5-yl	0.41			
2*H*-1,2,3-Triazol-2-yl	0.41	(−0.05)[c]		
Tetrazol-5-yl	0.49			

[a] σ_I Values from this work except where otherwise indicated. Some have been corrected for methyl substitution (see text). Values in parentheses are deductions made from data in the reference indicated using the σ_I value given and the appropriate form of equation (1) (see text). [b] Ref. 5. [c] Ref. 6. [d] Minimum value: see text.

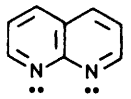
* The amine oxide tautomer is one marginal possibility[18] which, if present, would possess a base-weakening effect.



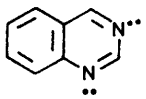
(23)



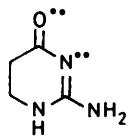
(24)



(25)



(26)



(27)

them deduced as above, alongside Charton's slim but definitive list[5] plus a handful of others which can be deduced from published data using equally rigorous criteria. Two of these criteria[5] are of special importance. First, as noted above, σ_I may be deduced directly only from reaction series known to be free of perturbing influences. Secondly, when σ_I (and σ_R) are deduced indirectly from some form of equation (1) and the equivalent relation involving σ_m, it must be known with certainty that the electrical composition of σ_m and σ_p is not appreciably different from the defining reaction of benzoic acid ionisation in water. This criterion rules out virtually all work carried out in non-protonic media[5,19] and casts doubt even on otherwise acceptable pK_a series in some mixed aqueous solvents. Hence the values of Table 5 are much more restrictive than in former compilations;[6] they are confined to data which appear to satisfy these criteria.

In addition, our values for pyrimidin-2-yl and quinazolin-2-yl have to be questioned. This is because of lone pair repulsion in the free base, illustrated as (**24**), which must increase the tendency to protonate and therefore result in a spuriously low σ_I value. It is difficult to find any good model for estimating the magnitude of this effect. 1,8-Naphthyridine (**25**) is a stronger base by *ca.* 1.4 pK_a units than quinazoline (**26**), but not all that effect is certainly due to this cause. In the model acylguanidines (**12**) and (**27**) the difference in pK_a of 1.15 units[17] lies in the wrong direction; this is probably dictated by the conformation of the acyl group,[17,19] which swamps out any opposing effect that may exist due to lone pair repulsion in (**27**). One might anticipate a value between that for pyrazin-2-yl (0.25) where the second nitrogen atom is more remote, and 1,2,4-triazin-6-yl (0.37) where a third has been added, but beyond that it is not possible to go.

Inside either series of six- or five-membered-ring heterocycles, trends are as expected. Among the azines, increasing nitrogen substitution increases σ_I, and among the diazines (except possibly when the substituent position is flanked by two nitrogens) it does not much matter where the second nitrogen atom is. It is interesting that two nitrogen atoms are required to approach the field effect due to the single oxygen atom of the acetyl group. While no σ_I value is known for any simple imino-substituent, this difference is consistent with the difference in electronegativity between N and O, and therefore with the expected dipole gradient. Similar trends are found for five-membered rings, where σ_I rises with the number of heteroatoms and, when only the π-donor atom is changed, in the order NR < S < O. This order cannot be reproduced by any simple blend of inductive pull and resonance push and may help to suggest that, at these very short ranges, electronegativity effects[20] are operative. A similar blend appears to be needed to account for non-additivity in the partition coefficients for substituted azoles.[21] All in all a considerable range in σ_I is covered; the top end of it approaches that of sulphonamide (see Table 1).

An unexpected feature of these results is the tendency of five-membered rings to possess higher σ_I values than six-membered-ring heterocycles at any given degree of heteroatom substitution. At first sight this contrasts with the classical view[22] of six- and five-membered-ring heterocycles as 'π-deficient' and 'π-excessive' respectively. However, there is no real clash: the above distinction derives from reactivity indices for such reactions as nitration and protodetritiation[23] which depend on a high degree of resonance involvement in the transition state. Two contributing factors to the observed trend in σ_I may be suggested.* The first is a compression effect related to the smaller size of five-membered rings, which will show itself here as increased intramolecular electrostatic repulsion leading to a steeper dipole gradient. The second is conformation: the smaller internal bond angle of five-membered rings will lead to a dipole more nearly aligned along the axis of the guanidine unit. The importance of conformation on σ_I has recently been demonstrated theoretically;[19] while as the authors observe it is difficult to identify experimental evidence for this, it seems possible that some of the residual standard error in equation (3) derives from conformational differences for some substituents relative to the XGY defining situation. Nevertheless the rigidity of heteroaryl substituents makes it probable that this effect of ring size on σ_I will remain valid in other contexts.

The information on σ_R summarised in Table 5 is so scanty that little can be said, but there are suggestive trends nevertheless. Most obviously, the π-deficient–π-excessive contrast *e.g.* between pyridine and furan shows up in σ_R^+ and σ_R^-, not in σ_I. Another interesting point concerns substituent positions sandwiched between a π-donor and a π-acceptor heteroatom, as *e.g.* in benzoxazole. Here there are fragmentary indications that the *sign* of σ_R may change with context; that is, the ring may act as a π-donor towards π-acceptors, and *vice versa*. We have suggestive evidence for this in the context of partition coefficients.[21] However, definitive information on this and other trends must await the further work that we hope these results will stimulate.

Experimental

Compounds were either commercial samples or obtained from the I.C.I. compound collection; the latter were authenticated by n.m.r. spectroscopy before use. pK_a Measurement was carried out at 25 °C in water, mostly by u.v. spectrophotometry but sometimes where appropriate by potentiometric titration, and by the methods described previously.[16,17] Standard errors are given in the Tables.

* As pointed out by a referee, the high 2,3-bond order may also contribute.

Acknowledgements

We thank Messrs. R. G. Button, G. A. Cockayne, S. Nicholson, and N. J. Keene for experimental assistance. We also thank Dr. J. Tomenson for advice on statistical matters and Dr. P. N. Edwards for wide-ranging discussions.

References

1 J. Shorter, 'Correlation Analysis of Organic Reactivity,' Wiley, Chichester, 1982.
2 N. B. Chapman and J. Shorter, eds., 'Correlation Analysis in Chemistry: Recent Advances,' Plenum Press, New York, 1978.
3 S. Ehrenson, R. T. C. Brownlee, and R. W. Taft, *Prog. Phys. Org. Chem.*, 1973, **10**, 1.
4 D. J. Craik and R. T. C. Brownlee, *Prog. Phys. Org. Chem.*, 1983, **14**, 1.
5 M. Charton, *Prog. Phys. Org. Chem.*, 1981, **13**, 119.
6 See, *e.g.*, O. Exner in ref. 2, p. 439; M. Charton, *ibid.*, p. 175.
7 W. F. Reynolds, *Prog. Phys. Org. Chem.*, 1983, **14**, 165.
8 M. K. A. Khan and K. J. Morgan, *Tetrahedron*, 1965, **21**, 2197.
9 M. Charton, *J. Org. Chem.*, 1965, **30**, 969.
10 D. J. Gilman, D. F. Jones, K. Oldham, J. M. Wardleworth, and T. O. Yellin, Royal Society of Chemistry Special Publication No. 42, ed. A. M. Creighton and S. Turner, 1982, p. 58.
11 See, *e.g.*, G.B. p. 2 001 624A/1979; EP 30 092/1981; EP 60 094/1982; and EP 60 697/1982, for various classes of heterocycle.
12 S. Marriott, W. F. Reynolds, and R. D. Topsom, *J. Org. Chem.*, 1985, **50**, 741.
13 M. Taagepera, K. D. Summerhays, W. J. Hehre, R. D. Topsom, A. Pross, L. Radom, and R. W. Taft, *J. Org. Chem.*, 1981, **46**, 891.
14 J. V. Greenhill, M. J. Ismail, G. R. Bedford, P. N. Edwards, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1265.
15 J. J. Morris and P. J. Taylor, manuscript in preparation.
16 R. G. Button, J. P. Cairns, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1555.
17 J. V. Greenhill, M. J. Ismail, P. N. Edwards, and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1255.
18 P. A. S. Smith, 'Open-chain Nitrogen Compounds,' Benjamin, New York, 1965, Vol. I, Ch. 1.
19 S. Marriott and R. D. Topsom, *J. Am. Chem. Soc.*, 1984, **106**, 7.
20 S. Marriott, W. F. Reynolds, R. W. Taft, and R. D. Topsom, *J. Org. Chem.*, 1984, **49**, 959.
21 P. J. Taylor, unpublished observations.
22 A. Albert, 'Heterocyclic Chemistry: an Introduction,' 2nd edn., Athlone Press, London, 1968.
23 See, *e.g.*, S. Clementi, P. P. Forsythe, C. D. Johnson, A. R. Katritzky, and B. Terem, *J. Chem. Soc., Perkin Trans. 2*, 1974, 399; A. R. Katritzky, S. Clementi, G. Milletti, and G. V. Sebastiani, *ibid.*, 1978, 613.

Received 24*th February* 1986: *Paper* 6/387